

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Synthese und Untersuchung Kohlenstoffsubstituierter DNA-Basen

Charlotte Ebert
aus
Frankfurt am Main

2017

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Thomas Carell betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde selbstständig, ohne unerlaubte Hilfe erarbeitet.

München, den 05.10.2017

Charlotte Ebert

Dissertation eingereicht am: 31.07.2017

- | | |
|---------------|---------------------------|
| 1. Gutachter: | Prof. Dr. T. Carell |
| 2. Gutachter: | Prof. Dr. K. Karaghiosoff |

Mündliche Prüfung am: 29.09.2017

Teile dieser Arbeit wurden in folgenden Journals oder auf Fachkonferenzen publiziert.

Publikationen

C. Ebert*, N. Simon*, S. Schneider, T. Carell; Structural insights into the recognition of *N*²-aryl- and C8-aryl DNA lesions by the repair protein XPA/Rad14, *ChemBioChem* **2017**, DOI: 10.1002/cbic.201700169.

S. Schiffers, C. Ebert, R. Rahimoff, O. Kosmatchev, J. Steinbacher, A. Bohne, F. Spada, S. Michalakis, J. Nickelsen, M. Müller, T. Carell; Quantitative LC-MS provides no evidence for m⁶dA or m⁴dC in the genome of mouse ESCs and tissues, *Angew. Chem. Int. Ed.* **2017**, DOI: 10.1002/anie.201700424.

N. Simon*, C. Ebert, S. Schneider; Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14, *Chem. Eur. J.* **2016**, 22, 10782-10785.

S. Koch, N. Simon, C. Ebert, T. Carell; Molecular mechanisms of xeroderma pigmentosum (XP) proteins, *Q. Rev. Biophys.* **2016**, 49, e5.

* Beitrag der Autoren zu gleichen Teilen.

Konferenzbeiträge

8th Harvard-LMU Young Scientists' Forum, Cambridge, MA, USA (2016) – Posterpräsentation: “Crystallization of deoxyguanosine adducts with Rad14 for investigation of DNA damage recognition in NER”.

JCF Frühjahrssymposium, Kiel (2016) – Posterpräsentation “Synthesis of deoxyadenosine adducts for investigation of DNA damage recognition by XPA/Rad14 in NER”.

SFB 749 Meeting, München (2015) – Posterpräsentation “DNA damage recognition by XPA/Rad14 in Nucleotide Excision repair”.

Danksagung

An dieser Stelle möchte ich meinen besonderen Dank nachstehenden Personen entgegenbringen, ohne deren Mithilfe die Anfertigung dieser Promotionsarbeit niemals möglich gewesen wäre:

Meinem Doktorvater *Prof. Dr. Thomas Carell* danke ich für die Aufnahme in diesem Arbeitskreis, für die hochinteressanten und herausfordernden Themen, die ich bearbeiten durfte, und für die hervorragenden Forschungsbedingungen.

Prof. Dr. Konstantin Karaghiosoff danke ich für die Übernahme des Zweitgutachtens.

Ich danke *Dr. Markus Müller* für sein offenes Ohr bei allen Fragen und Problemen in den letzten Jahren.

Bei *Frau Slava Gärtner* möchte ich mich für die reibungslose Abwicklung aller bürokratischen Belange bedanken.

Kristof Hufnagel und *Kerstin Kurz* gebührt mein Dank für die unermüdliche Unterstützung im Laboralltag.

Nina Simon und *Sarah Schiffers* danke ich für die erfolgreiche Zusammenarbeit auf unseren Projekten.

Dr. Sabine Schneider danke ich dafür, dass sie den weiten Weg in die Schweiz wieder und wieder auf sich genommen hat um unsere Kristalle zu messen, sowie für die gewissenhafte Auswertung der Daten im Anschluss.

Meinen Forschungspraktikanten und Bacheloranden *Jennifer Teichert*, *Jonas Feldmann*, *Alexandra Leutbecher* und *Lina Beisswenger* danke ich für ihren Einsatz im Labor. Ich wünsche euch viel Erfolg auf eurem weiteren Weg.

Michael Stadlmeier, *Katharina Iwan*, *Sarah Schiffers*, *Nina Simon*, *Miriam Wallner*, *Angie Kirchner*, *René Rahimoff* und *Dr. Markus Müller* danke ich für das schnelle und gewissenhafte Korrekturlesen großer Teile meiner Arbeit.

Dem Labor F 4.012 mit seinen aktuellen und ehemaligen Mitgliedern *Michael Stadlmeier*, *Bastien Viverge*, *Kristof Hufnagel*, *Meng Su*, *Florian Steinmann*, *Hidenori Okamura* und *Clemens Dialer* danke ich für die tolle Atmosphäre und die gute Musik, die mich in den letzten Jahren begleitet haben.

Ein großer Dank an den gesamten AK Carell für die angenehme Arbeitsatmosphäre, die hilfreichen Diskussionen und die vielen unvergesslichen Feiern und Ausflüge.

Meiner Familie und insbesondere meinen Eltern *Ute* und *Thomas* möchte ich dafür danken, dass sie immer für mich da waren, dass sie mir diese Ausbildung ermöglicht und mich in allen meinen Entscheidungen und Projekten unterstützt haben.

Meinem Freund *Andreas* danke ich für seine Unterstützung im letzten Jahr und dafür, dass er alle Höhen und Tiefen mit mir durchgestanden hat.

Inhaltsverzeichnis

Zusammenfassung	I
Summary.....	III
1 Einleitung	1
1.1 DNA-Schäden durch bulky DNA-Addukte.....	2
1.1.1 Aminoflurenlyl (AF) und Acetylaminofluorenlyl (AAF) Schäden	3
1.1.2 Aminonaphthalin (AN) Schäden	9
1.1.3 Aminopyren (AP) Schäden	9
1.2 Methylierte DNA-Basen.....	10
1.2.1 6-Methyl-2'-desoxyadenosin (m^6 dA).....	10
1.2.2 5-Methyl-2'-desoxycytidin (m^5 dC)	13
2 Aufgabenstellung.....	19
3 Ergebnisse und Diskussion	21
3.1 Arylmodifizierte DNA-Basen	21
3.1.1 Nukleotidexzisionsreparatur (NER) von C8- und N ² -Addukten	21
3.1.2 Kristallisation von C8-AAF-dG geschädigter DNA mit Rad14	25
3.1.3 Synthese von bulky Addukten	26
3.1.4 Schmelzpunktbestimmung	44
3.1.5 Bindungsstudien von geschädigter DNA und XPA/Rad14	47
3.2 Synthese von isotopenmarkierten Standards methylierter DNA-Basen	94
3.2.1 Quantifizierung von DNA-Basen mittels UHPLC-MS	94
3.2.2 Quantifizierung von 4-Methyl-2'-desoxycytidin (m^4 dC) und 6-Methyl-2'-desoxyadenosin (m^6 dA)	95
3.2.3 Synthese von 5-Methyl-2'-desoxycytidin (m^5 dC) und 5-[D ₃]-Methyl-2'-desoxy-[6-D, 1,3- ¹⁵ N ₂]-cytidin.....	120
4 Zusammenfassung und Ausblick	126
5 Experimenteller Teil.....	129
5.1 Chemische Materialien und Methoden	129
5.2 Methoden zur DNA-Synthese, Entschützung und Aufreinigung	131
5.3 Biochemische Materialien und Methoden	134
5.4 Synthese der bulky Addukte.....	136
5.4.1 Synthese des N ² -AAB-dG Phosphoramidits	136
5.4.2 Synthese des N ² -AAF-dG Phosphoramidits	141
5.4.3 Synthese des N ² -AAN-dG Phosphoramidits	150
5.4.4 Synthese des C8-AAF-dA Phosphoramidits	154
5.4.5 Synthese von Di- <i>iso</i> -Propylammoniumtetrazolat (DIAT)	161

5.5	Synthese von 5-Methyl-2'-desoxycytidin (m^5dC).....	161
5.5.1	Synthese von Hoffer's Chlorozucker	161
5.5.2	Synthese von 3',5'-bis-O- <i>p</i> -Toluoyl-2'-desoxyuridin.....	163
5.5.3	Synthese von 3',5'-bis-O- <i>p</i> -Toluoyl-5-iod-2'-desoxyuridin.....	164
5.5.4	Synthese von 3',5'-Bis-O- <i>p</i> -Toluoyl-5-iod-2'-desoxycytidin.....	165
5.5.5	Synthese von 5-Methyl-2'-desoxycytidin (m^5dC) und 2'-Desoxycytidin (dC) ..	167
5.6	Synthese von 5-[D ₃]-Methyl-2'-desoxy-[6-D, 1,3- ¹⁵ N ₂]-cytidin	169
5.6.1	[D ₂]-Propiolsäure	169
5.6.2	[D ₂ , ¹⁵ N ₂]-Uracil (29) und [D, ¹⁵ N ₂]-Uracil	169
5.6.3	[D ₂ , ¹⁵ N ₂]-3',5'-Bis-O-(<i>p</i> -Toluoyl)-2'-desoxyuridin (30)	170
5.6.4	β -[D, ¹⁵ N ₂]-3',5'-Bis-O-(<i>p</i> -Toluoyl)-5-iod-2'-desoxyuridin (33)	171
5.6.5	β -[D, ¹⁵ N ₂]-3',5'-Bis-O-(<i>p</i> -Toluoyl)-5-iod-2'-desoxycytidin (34)	172
5.6.6	5-[D ₃]-Methyl-2'-desoxy-[6-D, 1,3- ¹⁵ N ₂]-cytidin und 2'-Desoxy-[6-D, 1,3- ¹⁵ N ₂]-cytidin.....	173
6	Abkürzungsverzeichnis.....	175
7	Literaturverzeichnis	179
	Anhang.....	196

Zusammenfassung

Diese Promotionsarbeit befasst sich mit aryl- und alkylmodifizierten DNA-Basen. Zum einen wurden die Phosphoramidite arylsubstituierter DNA-Basen für die Untersuchung des Mechanismus der Nukleotidexzisionsreparatur (NER) synthetisiert und in DNA-Stränge eingebaut. Zum anderen wurden isotopenmarkierte Standards methylierter DNA-Basen für deren Massenspektrometrie-basierte Quantifizierung in unterschiedlichen Zellen und Geweben hergestellt.

Die Aktivierung bzw. Stilllegung bestimmter Gene zu bestimmten Zeiten, wird in Säugetieren unter anderem durch die Methylierung von 2'-Desoxycytidin (dC) zu 5-Methyl-2'-desoxycytidin (m^5 dC) in der Promotorenregion des jeweiligen Gens durch DNA-Methyltransferasen mit S-Adenosylmethionin (SAM) als Cofaktor bewerkstelligt. Im Arbeitskreis Carell werden Fütterungsexperimente durchgeführt, bei denen dem Medium embryonaler Stammzellen der Maus (mES-Zellen) [$^{13}\text{CD}_3$]-Methionin zugefügt wird. Dieses wird zu SAM verstoffwechselt, sodass die deuterierte, ^{13}C -markierte Methylgruppe letztendlich [$^{13}\text{CD}_3$]- m^5 dC generiert. Dank der Isotopenmarkierung lässt sich das Schicksal der Methylgruppe verfolgen. Im Zuge dieser Arbeit wurde 5-[D₃]-Methyl-2'-desoxy-[6-D, 1,3- $^{15}\text{N}_2$]-cytidin ([D₄, $^{15}\text{N}_2$]- m^5 dC) als interner m^5 dC[+6] Standard für die Quantifizierung des schweren m^5 dC synthetisiert.

Des Weiteren wurden 4-Methyl-2'-desoxy-[1,3- $^{15}\text{N}_2$]-cytidin ([$^{15}\text{N}_2$]- m^4 dC) und 6-[D₃]-Methyl-2'-desoxyadenosin ([D₃]- m^6 dA) als Standards für die Quantifizierung von m^4 dC und m^6 dA in unterschiedlichen Organismen, sowie unmarkiertes m^6 dA für die Erstellung einer Eichgeraden synthetisiert. Diese beiden Modifikationen sind seit langem in Bakterien bekannt und 2015 konnte m^6 dA zusätzlich in verschiedenen eukaryotischen Organismen nachgewiesen werden. Durch die hier synthetisierten Standards konnten wir eine ultrasensitive UHPLC-MS-Methode (*Ultra High Performance Liquid Chromatography* gekoppelt mit Massenspektrometrie) entwickeln, mit der auch der Nachweis kleinsten Mengen dieser methylierten Basen in der DNA möglich war. m^4 dC konnte einzig in Bakterien nachgewiesen werden, m^6 dA zusätzlich noch in der Grünalge *Chlamydomonas reinhardtii*, welche als Positivkontrolle verwendet wurde. Anders als in der Literatur beschrieben, konnten wir jedoch kein m^6 dA in mES-Zellen oder im Gewebe erwachsener Mäuse detektieren. Dies relativiert die zahlreichen Berichte über den Nachweis dieser Modifikation in den unterschiedlichsten Wirbeltieren.

Der andere Teil dieser Promotionsarbeit befasst sich mit arylsubstituierten DNA-Basen. Die sperrigen DNA-Addukte zwischen 2'-Desoxyguanosin (dG) und aromatischen Amiden bzw. Aminen werden durch die NER erkannt und repariert. Die Proteine der *Xeroderma Pigmentosum* (XP) Familie und insbesondere das XPA-Protein sind maßgeblich an der

Erkennung und an der Reparatur dieser Schäden beteiligt. 2015 konnte im Arbeitskreis Carell die Kristallstruktur mit Acetylaminofluoren (AAF) in C8-Position von dG geschädigter DNA und Rad14, dem Hefehomolog von XPA, erhalten werden. Diese Struktur weist auf eine Rolle von XPA/Rad14 in der Schadenserkennung hin, wobei die DNA auf ihre Biegsamkeit überprüft wird. Im Zuge dieser Arbeit wurden DNA-Stränge mit weiteren C8-dG-Addukten hergestellt, welche sich in der Größe des aromatischen Rests unterschieden. Zusätzlich wurden die Phosphoramidite verschiedener N^2 -dG-Addukte synthetisiert, welche ebenfalls unterschiedlich große aromatische Reste besaßen, ebenso wie das Phosphoramidit des C8-AAF-Addukts mit 2'-Desoxyadenosin (dA). Auch die Synthese von DNA-Strängen, welche diese Addukte enthielten, war erfolgreich. EMSA-Studien (*Electrophoretic Mobility Shift Assay*) konnten eine höhere Affinität von XPA für sämtliche synthetisierte geschädigte doppelsträngige DNA im Vergleich zu ungeschädigter DNA bestätigen.

Für alle Doppelstränge, die C8-Addukte enthielten, sowie für den Doppelstrang mit N^2 -Acetylaminonaphthalin(AAN)-dG als Schaden, konnten Kristallstrukturen mit Rad14 erhalten werden. Analog zur C8-AAF-dG Struktur binden in all diesen Strukturen immer zwei Rad14-Moleküle an die DNA, wobei der Kontakt hauptsächlich über die β -Schleife mit dem Phosphatrückgrat stattfindet, also sequenzunabhängig ist. Hierbei kommt es zu keinem direkten Kontakt mit dem Schaden, was mit der Tatsache im Einklang ist, dass die NER eine große Anzahl an strukturell sehr unterschiedlichen Substraten erkennen und reparieren kann. Die Enden der DNA werden aufgeschmolzen und nur ein 13mer Erkennungsmotiv mit dem Schaden in zentraler Position, durch den eine pseudo C2-Symmetriearchse läuft, bleibt intakt. Den erhaltenen Kristallstrukturen ist gemein, dass die DNA geknickt vorliegt (ca. 70° - 78°). Aus der C8-AAF-dG Struktur wird ersichtlich, dass dG mit AAF eine Ebene bildet, welche eine große Oberfläche für π -stacking Interaktionen mit den angrenzenden Basenpaaren bietet, wobei deren Watson-Crick Basenpaarung intakt bleibt. Die Hypothese, dass diese Interaktionen zur Stabilisierung des DNA-Knicks beitragen, wurde durch die neuen Strukturen gefestigt.

Summary

This PhD-thesis deals with aryl- and alkylmodified DNA-bases. The synthesis of phosphoramidites of bulky DNA lesions and their incorporation into oligonucleotide strands for the investigation of the nucleotide excision repair (NER) mechanism was carried out as well as the synthesis of isotopically labelled standards of methylated DNA bases for their quantification in different cells and tissues.

The activation or repression of certain genes at specific times in mammals, is often accomplished by methylation of 2'-deoxycytidine (dC) to 5-methyl-2'-deoxycytidine (m^5 dC) in promoter regions using S-adenosylmethionine (SAM) as a cofactor for DNA-methyltransferases. In the Carell group feeding experiments are carried out, in which [$^{13}CD_3$]-methionine is added to the medium of murine embryonic stem cells (mESC). It is metabolized to SAM, leading eventually to the formation of [$^{13}CD_3$]- m^5 dC. Because of the isotopic label it is possible to track the fate of the methyl group. In the course of this work 5-[D₃]-methyl-2'-deoxy-[6-D, 1,3- $^{15}N_2$]-cytidine ([D₄, $^{15}N_2$]- m^5 dC) was synthesized as an internal m^5 dC[+6] standard for the quantification of isotope-labelled m^5 dC.

Furthermore, 4-methyl-2'-deoxy-[1,3- $^{15}N_2$]-cytidine ([$^{15}N_2$]- m^4 dC) and 6-[D₃]-methyl-2'-deoxyadenosine ([D₃]- m^6 dA) were synthesized as standards for the quantification of m^4 dC and m^6 dA in different organisms, as well as non-labelled m^6 dA for the construction of an internal calibration curve. These two modifications have been known in bacteria for a long time and in addition, in 2015, m^6 dA was detected in different eukaryotic organisms. With the standards synthesized here, we were able to develop an ultrasensitive UHPLC-MS method (Ultra High-Performance Liquid Chromatography coupled with mass spectrometry), which allows detection of very small amounts of these methylated bases in DNA. m^4 dC was detected only in bacteria while m^6 dA was found in the green algae *Chlamydomonas reinhardtii* as well, which was used as a positive control. In contrast to the literature, we could not detect m^6 dA neither in mESC nor in tissues of adult mice. These findings put the numerous publications on the detection of this modification in different vertebrates into perspective.

The other part of this thesis focuses on aryl DNA-bases. The bulky DNA adducts between 2'-deoxyguanosine (dG) and aromatic amides or amines are recognized and repaired by the NER machinery. The proteins of the *Xeroderma Pigmentosum* (XP) family and especially XPA are the main proteins involved in the recognition and repair of these lesions. In 2015, in the Carell group, the crystal structure of double stranded DNA (dsDNA) – damaged with acetylaminofluorene (AAF) in the C8-position of dG – in complex with Rad14, the yeast homologue of XPA, was obtained. This structure indicates a potential role of XPA/Rad14 during damage recognition, in which the bendability of the DNA strand is tested. During this

work, DNA strands with further C8-dG adducts were synthesized, which differed in the size of the aromatic ring system. In addition, the phosphoramidites of N^2 -dG-adducts were synthesized which differed regarding the size of the aromatic residue, as well as the phosphoramidite of the C8-AAF adduct with 2'-deoxyadenosine (dA). The DNA strands containing these adducts were successfully synthesized. Electrophoretic Mobility Shift Assays (EMSA) confirmed a higher affinity of XPA to all the synthesized damaged DNA strands compared to undamaged dsDNA.

For the dsDNA that contained C8-adducts as well as for the double strand containing the N^2 -acetylaminonaphthalene(AAN)-dG adduct, crystal structures with Rad14 were obtained. Analogously to the C8-AAF-dG structure, in all these structures two Rad14 molecules bind the DNA, contacting mainly the phosphate backbone in a sequence independent way via a β -hairpin. No contact with the damaged base itself is observed which is consistent with the fact that the NER machinery can recognize and repair a large number of structurally very different substrates. The ends of the DNA doublestrand are opened leaving only a 13mer recognition motif around the centre intact. A pseudo C2 symmetry axis runs through the damaged site. In all the obtained crystal structures, the DNA is kinked by about 70° to 78°. In the C8-AAF-dG structure, the dG is in plane with the AAF residue providing a large surface for π -stacking interactions with the adjacent base pairs whose Watson-Crick base pairing remains intact. The hypothesis that these interactions stabilize the DNA kink is supported by the new structures.

1 Einleitung

Die Desoxyribonukleinsäure (DNA) ist das zentrale Molekül des Lebens, da es als Träger der Erbinformation fungiert. Aufgebaut ist die DNA aus den vier kanonischen Nukleosiden 2'-Desoxyadenosin (dA), 2'-Desoxythymidin (dT), 2'-Desoxycytidin (dC) und 2'-Desoxyguanosin (dG), welche über Phosphorsäurediester miteinander verknüpft sind. Hierbei geht jeweils eine Purinbase Wasserstoffbrückenbindungen mit einer Pyrimidinbase des DNA-Gegenstrangs ein, was zu spezifischen Bindungen zwischen dA und dT, sowie zwischen dC und dG, führt (Abbildung 1).^[1-2]

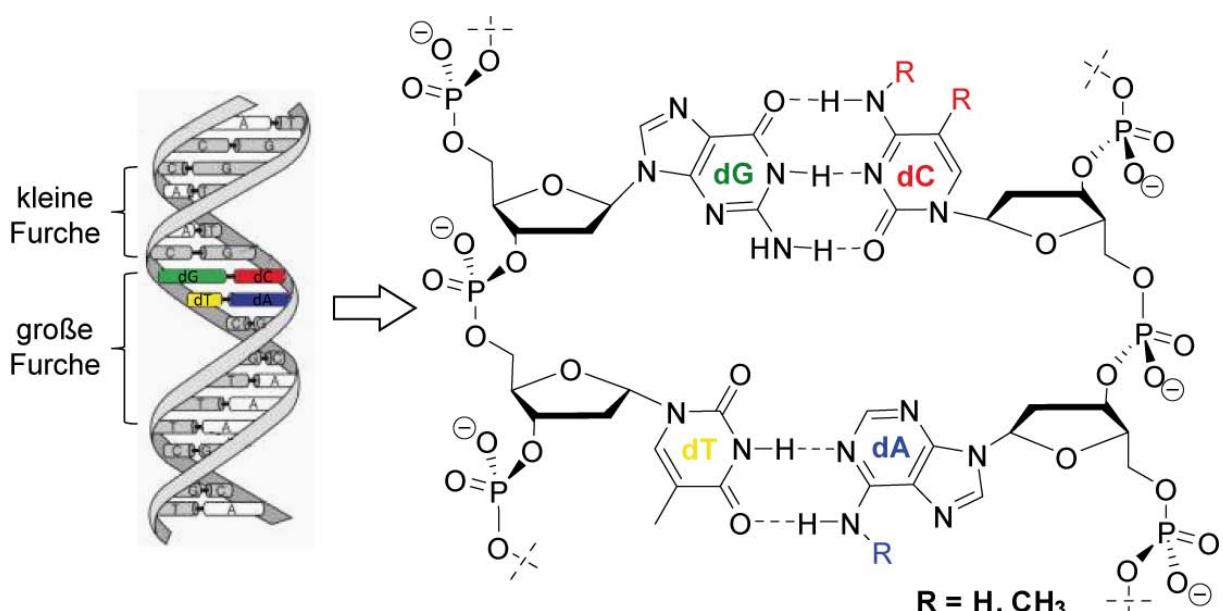


Abbildung 1: Aufbau der DNA-Doppelhelix. Basenpaarung zwischen dG und dC über drei und zwischen dT und dA über zwei Wasserstoffbrückenbindungen. Methylierung von dC oder dA führt nicht zur Beeinträchtigung der Watson-Crick-Basenpaarung.

Es kommt zur Ausbildung einer DNA-Doppelhelix, wobei die natürlich vorliegende B-Form über eine kleine und eine große Furche verfügt. Über diese Furchen können die Nukleobasen mit den Proteinen und Solvensmolekülen interagieren. Da die kleine Furche nur 1,2 nm breit ist, können hier nur kleinere Moleküle binden, während die große Furche 2,2 nm misst, sodass hier sequenzspezifische Interaktionen mit Proteinen möglich sind.^[3-6] Die Nukleoside, aus denen Ribonukleinsäure (RNA) und DNA aufgebaut sind, unterscheiden sich hauptsächlich durch eine OH-Gruppe in 2'-Position der Ribose. Während in der RNA eine Vielzahl verschiedener modifizierter Basen bekannt sind, konnten in DNA bisher jedoch nur wenige Modifikationen zusätzlich zu den kanonischen Nukleobasen nachgewiesen werden. Zu diesen gehören die alkylmodifizierten Basen. Die Methylgruppe der alkylierten DNA-Basen 5-Methyl-2'-desoxycytidin (m^5 dC), 6-Methyl-2'-desoxyadenosin (m^6 dA) und

4-Methyl-2'-desoxycytidin (m^4dC) führt zu keinerlei Beeinträchtigung der Watson-Crick-Basenpaarung und zeigt jeweils in die große Furche der DNA (Abbildung 1). Diese Modifikationen werden in bestimmten Organismen *via* DNA-Methyltransferasen gezielt in die DNA eingebaut und können dort verschiedene Rollen übernehmen, unter anderem bei der Genexpression.^[7-9] Hierbei wird zunächst die DNA zu messenger RNA (mRNA) transkribiert, welche während der Translation in die Aminosäuresequenz des zu synthetisierenden Proteins übersetzt wird.^[10] Ein Basentriplett kodiert für eine Aminosäure.^[6, 10-11] In anderen Organismen hingegen ist noch nicht geklärt, ob diese Modifikationen überhaupt vorkommen und eine Funktion ausüben, oder ob sie eher als DNA-Schäden einzuordnen sind.

Es gibt eine Vielzahl von bekannten DNA-Schäden, der erste Teil dieser Arbeit konzentriert sich jedoch auf die arylsubstituierten Basen, welche über die Nukleotidexzisionsreparatur (NER) repariert werden. Die Wichtigkeit und Aktualität dieses Forschungsgebiets wurde durch die Verleihung des Chemie-Nobelpreises 2015 unterstrichen. Aziz Sancar wurde für seine Forschung an diesem Reparaturmechanismus ausgezeichnet, zusammen mit Tomas Lindahl und Paul Modrich, die herausragende Erkenntnisse über die Basenexzisionsreparatur (BER) bzw. die *Mismatch*-Reparatur erzielen konnten.

1.1 DNA-Schäden durch *bulky* DNA-Addukte

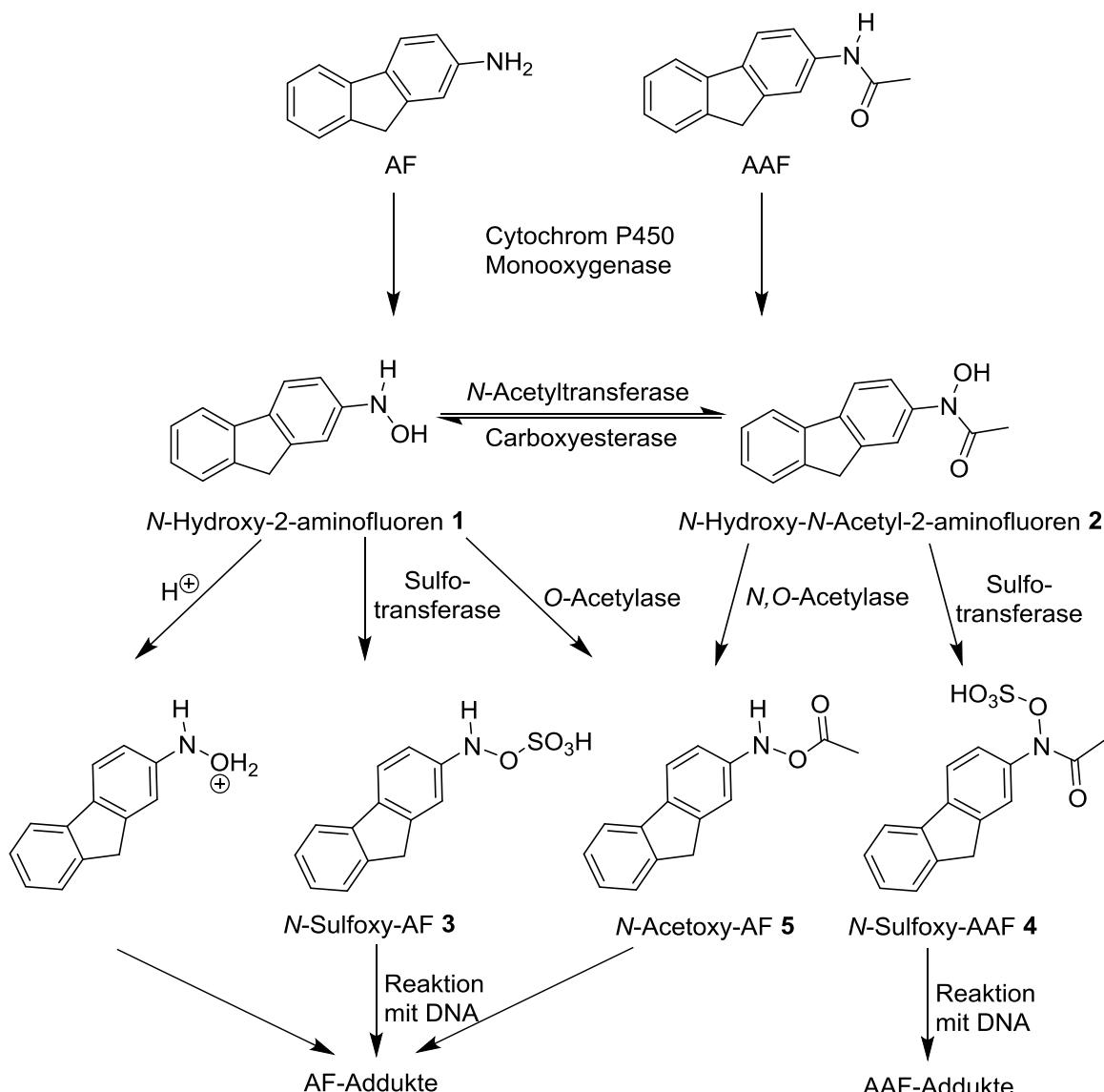
Täglich kommt es in jeder menschlichen Zelle zur Bildung von ca. 10^4 bis 10^6 DNA-Schäden.^[12-13] Schäden auf kodierenden Sequenzen können schwerwiegende Folgen haben, wenn sie zu Mutationen führen. Eine Änderung der Basenabfolge kann in einer Veränderung des gesamten genetischen Codes resultieren, wodurch eine korrekte Synthese von Proteinen nicht mehr gewährleistet ist. Modifizierte Proteinstrukturen können die Funktionalität des Proteins beeinträchtigen oder selbiges sogar ausschalten.^[6, 14] Die Natur hat deshalb verschiedene DNA-Reparaturmechanismen entwickelt um diese unvermeidlichen Schäden kontinuierlich zu beheben.

Sperrige DNA-Addukte entstehen, wenn bestimmte chemische Verbindungen unter Ausbildung kovalenter Bindungen mit DNA-Basen reagieren. Je nach Größe und Position können diese Addukte die Replikation beeinträchtigen, da beispielsweise eine korrekte Basenpaarung nicht mehr möglich ist.^[15] Zu diesen Verbindungen gehören Basenaddukte polycyclischer aromatischer Kohlenwasserstoffe (PAK), heterocyclischer aromatischer Amine oder von aromatischen Aminen und Amidien. Solche Substanzen, die Addukte bilden, finden sich in Tabakrauch, verbranntem Fleisch oder Fisch und in Fahrzeugabgasen, also überall, wo es zu einer unvollständigen Verbrennung organischer Materialien kommt.^[16-18] Es handelt sich hierbei um indirekte Mutagene, die erst gefährlich für unsere Gesundheit werden, indem sie zu reaktiven Elektrophilen metabolisiert werden, welche mit nukleophilen Zentren der

DNA-Basen reagieren können. Deshalb bilden sich solche Addukte bevorzugt an der C8- bzw. N^2 -Position von Guanosin oder an der exozyklischen N^6 -Aminogruppe von Adenosin.^[19] Diese Arbeit befasst sich mit Addukt-Schäden, welche durch aromatische Amine und Amide, wie 2-Aminofluoren (AF) und 2-N-Acetyl-Aminofluoren (AAF) sowie auch 2-Aminonaphthalin (AN), 4-Aminobiphenyl und 2-Aminotoluol und ihre acetylierten Derivate, entstehen. All diese Verbindungen gelten als starke Krebsreger, deren Mutagenität sowohl *in vitro* als auch *in vivo* bewiesen wurde.^[20-27]

1.1.1 Aminoflurenlyl (AF) und Acetylaminofluorenlyl (AAF) Schäden

Ursprünglich wurden AF sowie auch sein acetyliertes Derivat als Insektizide entwickelt. Auf Grund ihrer Mutagenität werden sie heute jedoch nur noch in der Krebsforschung für die Untersuchung von Wirkungsmechanismen der *bulky* DNA-Addukte verwendet.^[28] 2-Nitrofluoren (NF), ein Nebenprodukt der Diesel- und Kerosinverbrennung,^[29-30] wird ebenso wie AF und AAF zu instabilen *N*-Hydroxyderivaten metabolisiert, deren Reaktion mit der DNA letztendlich zur Entstehung der *bulky* Addukte führt.^[31] Die Metabolisierung dieser aromatischen Verbindungen wurde bereits in den 90er Jahren erforscht und ist in Schema 1 dargestellt.



Schema 1: Metabolisierung von AF und AAF.^[19, 32-35]

Die Aktivierung von AF bzw. AAF erfolgt in der Leber, indem das Stickstoffatom durch die Cytochrom P450 Monooxygenase zu den entsprechenden *N*-Hydroxyderivaten **1** und **2** oxidiert wird.^[25-26, 36] *N*-Hydroxyderivat **1** kann mittels einer *N*-Acetyltransferase in Verbindung **2** umgewandelt werden. Die Rückreaktion wird von einer Carboxyesterase katalysiert. Durch säurekatalysierte Solvolyse kann Verbindung **1** direkt mit der DNA zu AF-Addukten reagieren. Alternativ dazu kann die Hydroxygruppe mit Hilfe einer Sulfotransferase sulfoniert werden, sodass die reaktiven Ester **3** und **4** entstehen. Durch Acetylierung von Derivat **1** mittels einer O-Acetylase oder Umacetylierung von Derivat **2** mit Hilfe einer *N,O*-Acetylase, wird der aktivierte Ester **5** gebildet. Dieser kann nun analog zu den Verbindungen **3** und **4** über ein Nitreniumintermediat mit Adenosin bzw. Guanosin reagieren und so *bulky* Addukte in der DNA bilden.^[19, 32-35]

AF und AAF bilden bevorzugt Addukte an der C8-Position von dG. Da die Metabolisierung dieser Substanzen in der Leber stattfindet, wurden diese Zellen von Ratten mit AAF oder mit N-Hydroxyderivat **2** behandelt und sowohl *in vitro* als auch *in vivo* untersucht. *In vitro* entsteht hauptsächlich (ca. 60 %) das acetylierte C8-Acetylaminofluorenyl-2'-desoxyguanosin (C8-AAF-dG), während dieses Addukt *in vivo* nur zu ca. 15 % entsteht.^[24, 37-39] Hier kommt es primär (ca. 80 %) zur Bildung des nicht acetylierten C8-Aminofluorenyl-2'-desoxyguanosin (C8-AF-dG).^[24, 37, 40] Das acetylierte *N*²-Acetylaminofluorenyl-2'-desoxyguanosin (*N*²-AAF-dG) entsteht *in vivo* nur zu ca. 5 %.^[37-38] Allerdings ist die Lebensdauer von *N*²-AAF-dG im Organismus weit länger als die des entsprechenden C8-Addukts. C8-AAF-dG ist nur wenige Tage nach der Zuführung nachweisbar, während *N*²-AAF-dG über längere Zeit hinweg persistent ist.^[24, 37-38] *Beland et al.* haben den Tieren über einen längeren Zeitraum hinweg Hydroxyderivat **2** zugeführt, um die Folgen einer chronischen Aussetzung zu untersuchen. Das in *N*²-Position substituierte Addukt konnte im Gewebe von mit AAF behandelten Tieren noch nach zwei Wochen nachgewiesen werden, während das C8-AAF-Addukt nur bis einen Tag nach der Zuführung nachweisbar war. Die anhaltende Zuführung von N-Hydroxy-AAF resultierte in einer additiven Akkumulierung von *N*²-AAF-dG sowie von C8-AF-dG, was auch hier als Hauptprodukt entstand.^[41] *Poirier et al.* stellten fest, dass eine kontinuierliche Fütterung von AAF nach einiger Zeit zu einer konstanten Adduktkonzentration führt. Zunächst bilden sich auch acetylierte Addukte, im Gleichgewichtszustand liegen jedoch hauptsächlich nicht acetylierte Addukte vor. Die Abbaurate dieser Addukte wurde durch fortlaufende AAF-Fütterung nicht beeinflusst.^[42-44] In den Nieren und Milchdrüsen von Ratten konnte ausschließlich C8-AF-dG nachgewiesen werden, wenn auch in weitaus geringeren Mengen als in der Leber.^[41, 44-45] Fütterung von NF führte unter anderem zur Bildung von C8-AF-dG und *N*²-AAF-dG in Leber, Niere und Milz. Über elf Monate nahm die Konzentration an *N*²-AAF-dG stark zu, während die Menge an C8-AF-dG nahezu konstant blieb.^[46] *N*²-Addukte könnten dementsprechend weitaus mutagener als die C8-Addukte sein.

Die beiden C8-Addukte unterscheiden sich ausschließlich durch die Acetylgruppe am Stickstoff, was jedoch vor allem in Bakterien zu gravierenden Unterschieden in ihrer Mutagenität führt. Während C8-AF-dG dG•dC → dT•dA Punktmutationen auslöst, führt C8-AAF-dG hauptsächlich zum Verlust von Basenpaaren, wodurch es zu -1 und -2 *frameshift*-Mutationen kommt.^[33, 47-48] In Säugetieren verursachen hingegen beide Schäden, wenn sie nicht rechtzeitig erkannt und repariert werden, dG•dC → dT•dA Punktmutationen.^[47] Diese unterschiedlichen Eigenschaften könnten auf die verschiedenen Konformationen, in denen die geschädigte DNA vorliegt, zurückzuführen sein. Für C8-AF-dG sind in B-DNA mit normaler Watson-Crick-Basenpaarung aus NMR-Strukturanalysen zwei Konformationen bekannt.

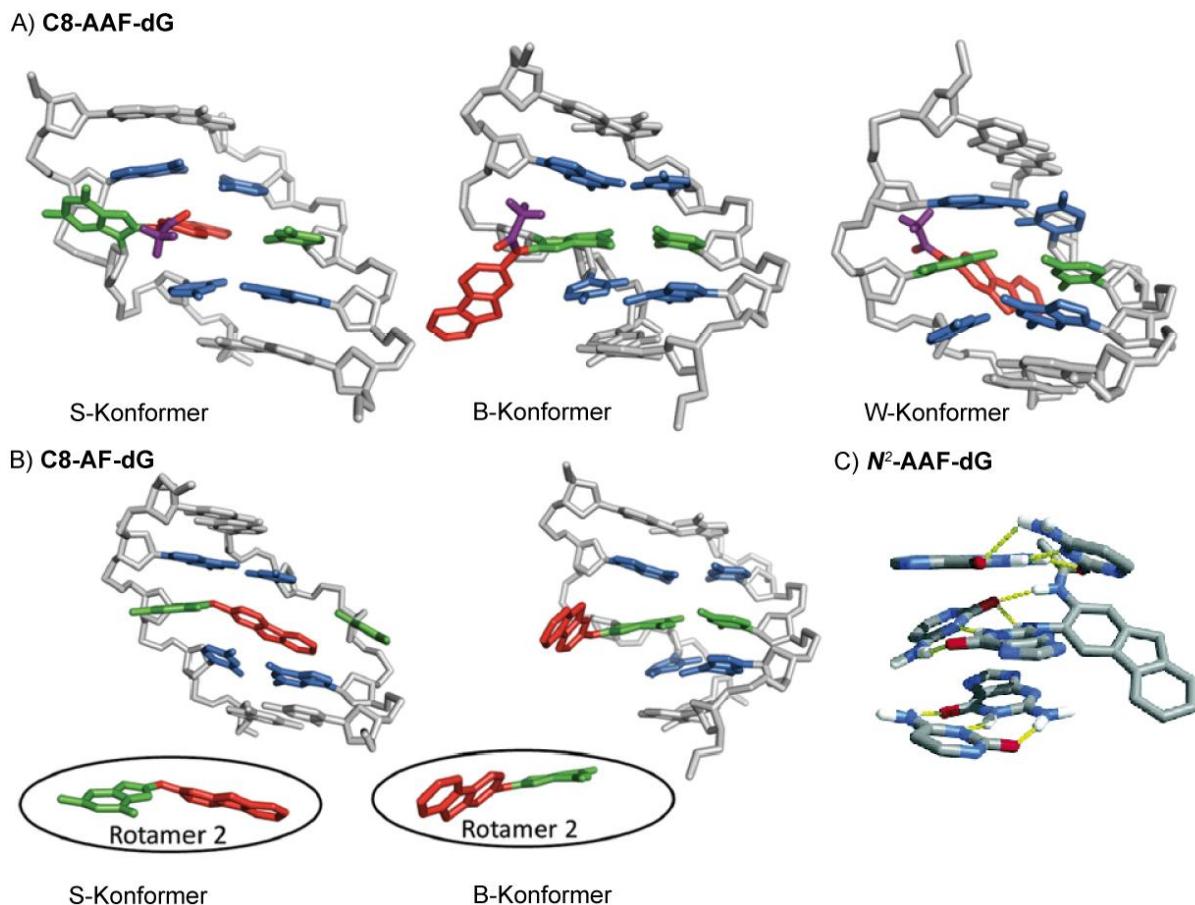
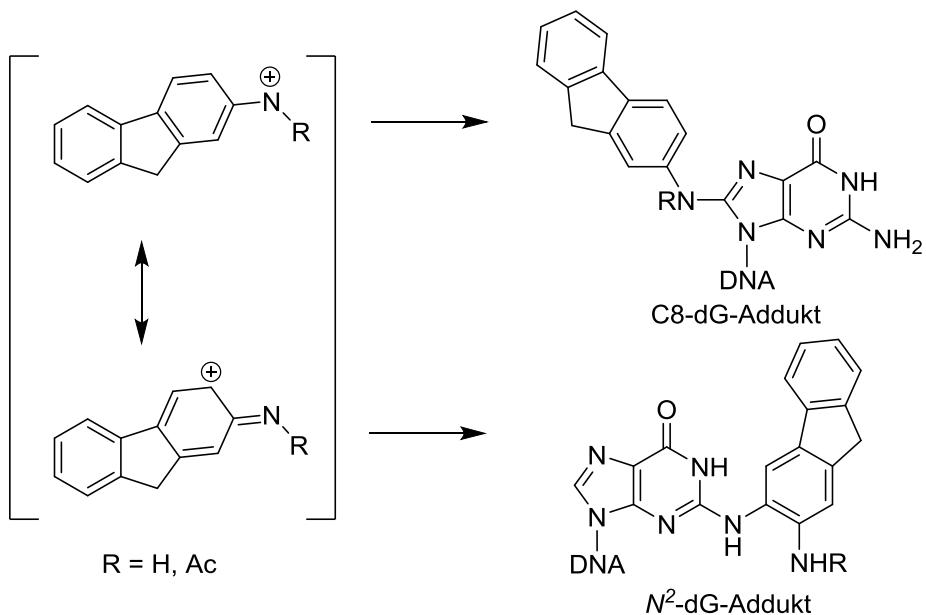


Abbildung 2: Strukturen von AF bzw. AAF geschädigter dsDNA. A) C8-AAF-dG geschädigte DNA kann als S-, B- oder W-Konformer vorliegen. Im S-Konformer interkaliert AAF in den DNA-Duplex, wodurch dG eine *syn*-Konformation einnimmt und in die große Furche verdrängt wird. Im B-Konformer liegt dG in der normalen *anti*-Konformation vor, wobei AAF in die große Furche zeigt. Im W-Konformer zeigt AAF in die kleine Furche. B) C8-AF-dG geschädigte DNA kann als S- oder B-Konformer vorliegen. Der Fluorenylrest kann jeweils in zwei Rotameren vorliegen. Der Schaden ist jeweils in rot dargestellt, das geschädigte Basenpaar in grün, die benachbarten Basenpaare in blau und die Acetylgruppe in violett. C) In N²-AAF-dG geschädigter DNA zeigt AAF in die kleine Furche, während die Basenpaarung intakt bleibt. Der Schaden ist in grau dargestellt, Wasserstoffbrückenbindungen in gelb. Die Abbildung wurde nach *Mu et al.* und *Zaliznyak et al.* adaptiert. Verwendung der Abbildung mit Erlaubnis der Verlage; Copyright 2006 ACS Publications und 2012 Oxford University Press.^[49-50]

Zum einen kann Guanosin die normale *anti*-Konformation (B-Konformer, Abbildung 2B) um die glykosidische Bindung beibehalten, wodurch auch die Wasserstoffbrücken mit dem Gegenstrang intakt bleiben. Der Fluorenylrest zeigt in die große Furche und kann zwischen zwei Rotameren (180°) um die Guanosin(N⁸)-Fluorenyl(C2)-Bindung hin und her wechseln. Diese Konformationen führen nur zu einer minimalen Deformation der DNA-Helix.^[51-54] Zum anderen kann der Fluorenylrest in die Helix interkalieren, wodurch Guanosin eine *syn*-Konformation (S-Konformer, Abbildung 2B) um die glykosidische Bindung einnimmt und in die große Furche zeigt. Die benachbarten Basenpaare bleiben hierbei intakt, während eine

Watson-Crick-Basenpaarung zwischen dG und dC nicht mehr möglich ist. Auch in diesem Fall liegen zwei Rotamere um die N-C-Bindung vor.^[54-55] Die hier beschriebenen NMR-Strukturen wurden zusätzlich über UV-Schmelzpunktmessungen, Circulardichroismus- (CD) und ¹⁹F-NMR-Spektroskopie charakterisiert.^[56-57] 1993 wurde über eine erste mittels NMR-Spektroskopie ermittelte Konformation für das C8-AAF-dG Addukt in dsDNA berichtet. Guanosin liegt in *syn*-Konformation (S-Konformer, Abbildung 2A) vor, während der Fluorenylrest in die Helix interkaliert und mit einem benachbarten Basenpaar π -stacking Interaktionen eingehen kann. Analog zur interkalierten Konformation des C8-AF-dG ist auch hier keine Ausbildung von H-Brücken zur Gegenbase möglich. Der äußere Ring des Fluorenylrests zeigt in die kleine Furche.^[58] Das Spektrum zeigte, dass auch noch ein geringer Prozentsatz an anderen Konformationen vorliegen muss, deren genaue Konformationen konnten aber erst einige Jahre später von der Gruppe um Cho mittels ¹⁹F-NMR-Spektroskopie aufgeklärt werden. Ähnlich wie das C8-AF-dG kann auch dieses Addukt in geringen Mengen als B-Konformer (Abbildung 2A) vorliegen, in der die Basenpaarung erhalten bleibt und der Fluorenylrest in die große Furche zeigt. Weiterhin ist es möglich, dass Guanosin eine *syn*-Konformation einnimmt und AAF in die kleine Furche zeigt (W-Konformer, Abbildung 2A).^[59-60] Welche Konformation vorliegt ist stark sequenzabhängig und auch in der gleichen Sequenz können unterschiedliche Konformationen vorliegen, wobei C8-AAF-dG meist in *syn*-Konformation vorliegt, während C8-AF-dG die *anti*-Konformation bevorzugt.^[53, 58-61] Auf Grund der verschiedenen Konformationen sind C8-AF-dG und sein acetyliertes Derivat unterschiedlich gute Substrate für die NER (Kapitel 3.1.1).

C8-Addukte werden durch Reaktion eines reaktiven Nitreniumions mit Guanosin gebildet. Auf Grund der höheren Nukleophilie des N7-Stickstoffatoms, wurde zunächst ein Angriff an dieser Position mit Umlagerung zum C8-Addukt diskutiert.^[62] Inzwischen geht man jedoch von einem direkten Angriff an der C8-Position aus.^[63] Alternativ kann das Carbokation in *ortho*, welches durch das Arylnitreniumion resonanzstabilisiert ist, die N^2 -Position von Desoxyguanosin angreifen.^[64] Die beiden mesomeren Grenzstrukturen, die zur Ausbildung von C8- bzw. N^2 -Addukten führen, sind in Schema 2 abgebildet.



Schema 2: Mesomere Grenzstrukturen von AF und AAF und daraus resultierende Adduktbildung mit 2'-Desoxyguanosin.^[64]

Es ist bekannt, dass N^2 -AAF-dG, ebenso wie die C8-dG Adduktschäden, zu dG•dC → dT•dT Punktmutationen führen.^[65] 1997 haben *Grad et al.* mittels computerbasierter Berechnungen eine erste Struktur für N^2 -AAF-dG geschädigte DNA vorgeschlagen,^[66] neun Jahre später wurde eine weitestgehend damit übereinstimmende NMR-Struktur von *Zaliznyak et al.* präsentiert (Abbildung 2C).^[49] In dieser Struktur kommt es nur zu einer minimalen Perturbation der DNA-Helix. Sämtliche Nukleotide liegen in *anti*-Konformation vor und die *Watson-Crick*-Basenpaarung ist nicht beeinträchtigt. Auch das *stacking* der Basen bleibt erhalten, ist jedoch am Schaden nur eingeschränkt möglich. AAF und dG formen einen Winkel von 41°. Der Fluorenylrest befindet sich tief in der kleinen Furche in Richtung des 5'-Endes des geschädigten Strangs. Die Furche ist an der geschädigten Stelle sichtbar verengt und vertieft. Dadurch wird das hydrophobe AAF weitest möglich von den Wassermolekülen der Lösung abgeschirmt. Nur 10 % seiner Oberfläche kommen mit den Lösungsmittelmolekülen in Kontakt.^[49]

Die Bildung von Adenosinaddukten mit AAF wurde bisher nur bei *in vitro* Studien mit poly(dA•dT)-Strängen beobachtet, in denen kein dG, welches stattdessen reagieren könnte, vorhanden ist.^[67] Die Bildung von *bulky* Addukten an der C8-Position von dA mit Arylaminen oder -amiden ist selten, aber wurde schon beobachtet.^[68] Die Bildung von N^6 -Addukten ist hingegen verbreiteter.^[69-70]

1.1.2 Aminonaphthalin (AN) Schäden

AN-Addukte sind weitaus weniger gut erforscht als die AF- und AAF-Addukte. Bereits 1895 kam der Verdacht auf, dass die aromatischen Amine Blasenkrebs auslösen könnten,^[71] aber erst 1938 rückte 2-AN in den Fokus der Forschung, als Tumore bei mit 2-AN gefütterten Hunden nachgewiesen werden konnten.^[72] Später wurde der Zusammenhang zwischen einem erhöhten Kontakt mit 2-AN und Blasentumoren bei Arbeitern in der Farbindustrie hergestellt.^[73] Der Nachweis von AN-Addukten in Leber- und Blasenzellen wurde 1977 erbracht.^[74] *In vitro* wurden drei verschiedene Addukte identifiziert: Das Hauptprodukt (50 %) war ein Derivat von C8-AN-dG mit Ringöffnung (1-[5-(2,6-Diamino-4-oxopyrimidinyl-N⁶-deoxyriboside)]-3-(2-naphthyl)harnstoff gefolgt von N²-Aminonaphthyl-2'-desoxyguanosin (N²-AN-dG, 30 %) und N⁶-Aminonaphthyl-2'-desoxyadenosin (N⁶-AN-dA, 15 %).^[69] Analog zu AF und AAF, wird auch AN zunächst zu *N*-Hydroxy-2-aminonaphthalin (*N*-HO-AN) metabolisiert.^[19] Die darauffolgende Bindung an Glucuronsäure führt dazu, dass der Metabolit wasserlöslich wird. Diese *N*-Glucoronidierung ermöglicht die Ausscheidung des unpolaren Aromaten über Leber und Niere.^[74] Da der pH-Wert von Urin ca. 5-6 beträgt und dieses folglich leicht sauer ist, kann es in der Blase jedoch zur Hydrolyse des Glucuronsäureaddukts kommen. Das Mutagen *N*-HO-AN wird zurückgebildet und kann direkt mit der DNA reagieren.^[75-77] Auch Zuführung des acetylierten Metaboliten *N*-Hydroxy-*N*-acetyl-2-aminonaphthalin (*N*-HO-AAN) in Ratten, Hasen und Meerschweinchen führte zur Bildung des entsprechenden Glucuronsäureaddukts.^[78]

1.1.3 Aminopyren (AP) Schäden

Zusätzlich zu den aromatischen Fluorenyl- und Naphthyladdukten mit Desoxyguanosin konnte auch die Entstehung eines Pyrenyl-dG-Addukts nachgewiesen werden. Howard *et al.* konnten C8-Aminopyrenyl-2'-desoxyguanosin (C8-AP-dG) nach Inkubation des Bakteriums *Salmonella typhimurium* mit 1-Nitropyren (1-NP) charakterisieren.^[79] 1-NP wird wahrscheinlich über 1-Nitrosopyren zu *N*-Hydroxy-1-aminopyren metabolisiert, welches dann mit der DNA reagieren kann.^[80] Auch in Ratten, die mit 1-NP gefüttert wurden, konnte die gleiche Mutation nachgewiesen werden.^[81] Mao *et al.* konnten 1996 eine NMR-Struktur dieses Addukts erhalten. Der Pyrenylrest interkaliert in den DNA-Duplex, sodass das geschädigte dG und die dC-Gegenbase in die große Furche verdrängt werden. dG nimmt eine *syn*-Konformation um die glykosidische Bindung ein und es kommt zum *stacking* mit dem dC des gleichen Stranges in 5'-Richtung. Die Wasserstoffbrückenbindungen der benachbarten Basenpaare bleiben intakt.^[82] C8-AP-dG führt in *E. coli* zu *frameshift*-Mutationen und zum Stopp der RNA Polymerase *in vitro*.^[83-84] Herreno-Saenz *et al.* konnten in der DNA verschiedener Säugetiere noch zwei weitere Addukte an der N²-Position von

Desoxyguanosin identifizieren: N^2 -(6-Aminopyrenyl)-2'-desoxyguanosin und N^2 -(8-Aminopyrenyl)-2'-desoxyguanosin.^[85]

1.2 Methylierte DNA-Basen

Weitere in der DNA bekannte Modifikationen sind die drei methylierten Basen m^5dC , m^6dA und m^4dC . In Bakterien und Archaeen-DNA konnten alle drei Modifikationen identifiziert werden – m^6dA ^[86] und m^5dC ^[87] bereits in den fünfziger bzw. sechziger Jahren, m^4dC erst ca. 20 Jahre später.^[88] Letzteres konnte bisher nur in Prokaryoten identifiziert werden. Diese schützen ihre eigene DNA durch das Restriktions-Modifikations-System (R/M-System). Methyltransferasen methylieren hierbei die zelleigene DNA zu m^6dA , m^5dC und m^4dC , sodass diese von einem Restriktionsenzym nicht abgebaut werden kann und somit nur zellfremde DNA geschnitten wird.^[89] Des Weiteren spielt m^6dA in Bakterien eine Rolle bei der Replikation und Reparatur, der Transposition und der Genregulation.^[7, 90] Die Kapitel 1.2.1 und 1.2.2 beschäftigen sich mit dem Vorkommen von m^6dA und m^5dC in Eukaryoten.

1.2.1 6-Methyl-2'-desoxyadenosin (m^6dA)

1.2.1.1 Vorkommen in Eukaryoten

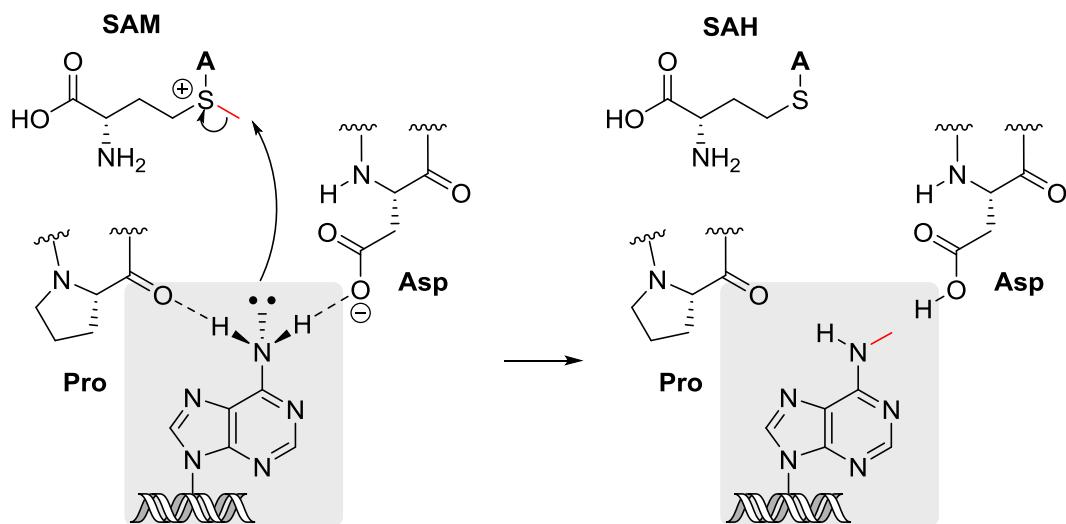
Auch in Eukaryoten ist m^6dA eine seit langem bekannte Modifikation.^[91-92] In den letzten Jahren ist m^6dA wieder in den Fokus der Forschung gerückt und konnte 2015 in der Grünalge *Chlamydomonas reinhardtii* (0,4 % m^6dA/dA),^[93] sowie auch in den multizellulären Eukaryoten *Caenorhabditis elegans* (0,01-0,4 % m^6dA/dA)^[94] und in *Drosophila melanogaster* (0,001-0,07 % m^6dA/dA)^[95] nachgewiesen werden. Ein R/M-System, wie in den Bakterien, ist bis dato in Eukaryoten nicht bekannt, weshalb m^6dA in diesen Organismen andere Funktionen übernehmen muss, welche bisher jedoch noch nicht zweifelsfrei feststehen. Beobachtungen deuten darauf hin, dass die Adenosin-Methylierung in verschiedenen Mechanismen unterschiedliche biologische Prozesse regulieren könnte.^[96-101] In *C. reinhardtii* ist m^6dA hauptsächlich an den Startstellen der Transkription (*Transcription Start Sites*, TSS) aktiver Gene angereichert, während an den TSS stillgelegter Gene ein geringerer Gehalt an m^6dA beobachtet wurde.^[93] In diesem Organismus könnte m^6dA also in der Genregulation eine Rolle spielen. In *C. elegans* hingegen ist diese Modifikation gleichmäßig über das gesamte Genom verteilt.^[94] Allerdings wurde sie verstärkt in Mutanten gefunden, die einen erhöhten Gehalt an Lysin 4-dimethylierten Histonen H3 (H3K4me2) aufweisen. Diese Mutation wird mit aktiver Transkription assoziiert, was darauf hindeuten könnte, dass m^6dA auch in *C. elegans* aktive Gene kennzeichnet. In diesem Fadenwurm konnte eine m^6dA Demethylase (N6-Methyladenin Demethylase 1, NMAD-1) aus der ALKB

Familie identifiziert werden.^[94] Auch eine mögliche Methyltransferase (DNA N6-Adenin Methyltransferase 1, DAMT-1) aus der MTA-70 Familie wurde in diesem Organismus gefunden.^[94] In *D. melanogaster* ist m⁶dA an Transposons angereichert.^[95] Der Verlust der mutmaßlichen DNA m⁶dA Demethylase (DMAD) führte zu einer erhöhten Expression dieser Transposons.^[95] In allen drei eukaryotischen Organismen ist das Auftreten von m⁶dA also mit erhöhter Genexpression verknüpft. Zu diesem Zeitpunkt lässt sich allerdings noch nicht sagen, ob diese Beobachtungen allgemein gelten und ob m⁶dA eine Rolle in der Genregulation spielt. Erste m⁶dA Methyltransferasen und Demethylasen, die diesen Mechanismus steuern könnten, wurden zwar bereits vorgeschlagen, deren eindeutige Identifizierung steht aber noch aus.^[94-95]

Diese Entdeckungen steigerten das Interesse an m⁶dA, sodass 2016 mehrere Gruppen diese Modifikation in Vertebraten nachgewiesen haben.^[102-104] Koziol et al. haben m⁶dA im Krallenfrosch, der Maus, dem Zebrafisch und in menschlichen Zellen nachgewiesen, allerdings in weitaus geringeren Mengen (ca. 0,00009 % m⁶dA/dA). Das methylierte dA wurde zunächst via Dot Blot nachgewiesen und mittels Ultra-high-performance Flüssigchromatografie gekoppelt mit Tandem-Massenspektrometrie (UHPLC-MS/MS) unter Zuhilfenahme einer externen Eichgeraden quantifiziert. In *Xenopus laevis* und *Mus musculus* Geweben soll die methylierte Base hauptsächlich in Introns, also in nicht codierenden Sequenzen der DNA präsent sein.^[102] Wu et al. können diese Ergebnisse nicht bestätigen. Sie konnten m⁶dA lediglich in embryonalen Mäusestammzellen (mES-Zellen, 6-7 ppm m⁶dA/dA), nicht jedoch im Gewebe von erwachsenen Mäusen, nachweisen. Für die Quantifizierung wurde ein [¹⁵N₅]-m⁶dA Standard während der LC-MS/MS verwendet. Als Demethylase wurde *Alkbh1* identifiziert. Die Modifikation soll vor allem an jungen Vollängen LINE-1 Transposons (L1-Elemente) der X-Chromosome angereichert sein (25-30 ppm m⁶dA/dA) und im Gegenteil zu den einfachen Eukaryoten zur Stilllegung bestimmter Gene führen.^[103] Wenig später haben Liu et al. gezielt nach biologischen Prozessen in Schweinen und Zebrafischen gesucht, während denen es zu einer Erhöhung des m⁶dA Gehalts kommt. Während der Embryonalentwicklung konnten sowohl in genomischer Spermien-DNA (Zebrafisch: ca. 0,003 % m⁶dA/dA, Schwein: ca. 0,015 % m⁶dA/dA) als auch in der Eizelle (Zebrafisch: ca. 0,015 % m⁶dA/dA, Schwein: ca. 0,09 % m⁶dA/dA) erhöhte m⁶dA Gehalte gemessen werden. Nach der Befruchtung steigt der m⁶dA Gehalt weiter (Zebrafisch: ca. 0,1 % m⁶dA/dA, Schwein: ca. 0,17 % m⁶dA/dA) um dann auf Hintergrundniveau abzufallen. Für die Quantifizierung mittels UHPLC-MS/MS wurde eine externe Eichgerade verwendet.^[104] m⁶dA ist an den TSS angereichert, aber im Gegensatz zu den Beobachtungen von Koziol et al. im *X. laevis* Genom tritt es hauptsächlich in Exons auf.^[102, 104]

1.2.1.2 N-Methylierung von DNA

Die Methylierung exozyklischer Aminogruppen wie im Fall von m^6 dA und m^4 dC verläuft als nukleophile Substitution 2. Ordnung ohne Bildung eines kovalenten Intermediats.^[105] Auf Grund der Konjugation der freien Elektronenpaare des exozyklischen Stickstoffs mit dem aromatischen Ring, sind weder dA noch dC an dieser Position besonders nukleophil. Da sowohl m^6 dA DNA-Methyltransferasen dC als Substrat akzeptieren als auch andersherum, sollte der Mechanismus der Methylierung identisch sein.^[106-107] In diesen N-Methyltransferasen ist die Aminosäuresequenz (Asp/Asn/Ser)-Pro-Pro-(Tyr/Phe) stark konserviert. Die Hauptfunktion dieses Tetrapeptids ist die Ausbildung von Wasserstoffbrückenbindungen der (Asp/Asn/Ser)-Seitenkette und der Hauptketten-Carbonylgruppe des ersten Prolins zur exozyklischen Aminogruppe. Auf Grund der tetraedrischen Geometrie dieser Akzeptorgruppen wird der Stickstoff wahrscheinlich von sp^2 zu sp^3 umhybridisiert, sodass das freie Elektronenpaar an der N^6 -Position lokalisiert wird und somit für den nukleophilen Angriff zur Verfügung steht (Schema 3).^[108-109]



Schema 3: Katalytischer Mechanismus von DNA-N-Methyltransferasen.^[108-109]

Die Methylgruppe wird von S-Adenosylmethionin (SAM) bereitgestellt. SAM entsteht als ein wichtiges Stoffwechselprodukt aus der Aminosäure Methionin. Katalysiert durch die Methionin-Adenosyl-Transferase reagiert Methionin mit Adenosintriphosphat (ATP) in Anwesenheit von zweiwertigen Kationen zu SAM. In einer Transmethylierungsreaktion kann dieses nun seine Methylgruppe auf DNA-Basen übertragen, wodurch S-Adenosylhomocystein (SAH) als Abbauprodukt entsteht, welches wiederum zu Homocystein und Adenosin hydrolysiert wird. Der (Asp/Asn/Ser)-Rest fungiert wahrscheinlich als Protonenakzeptor, wenn die $N\text{-CH}_3$ -Bindung schon fast vollständig ausgebildet ist, und

gibt das Proton dann schnell über andere Reste an Wasser ab. Der dadurch entstehende kationische Übergangszustand wird vermutlich auch durch π -Wechselwirkungen mit den Resten aromatischer Aminosäuren stabilisiert, die in allen Bindungstaschen von N-Methyltransferasen mit den DNA-Basen interagieren.^[110-113]

1.2.2 5-Methyl-2'-desoxycytidin (m^5dC)

Bereits 1925 wurde von Johnson et al. über die Entdeckung von m^5dC im Tuberkulosebazillus berichtet.^[114] In Ermangelung an zuverlässigen analytischen Methoden zu jener Zeit, konnte dessen Existenz jedoch erst 1950 zweifelsfrei in Eukaryoten nachgewiesen werden.^[115] Heutzutage ist m^5dC die am besten erforschte DNA-Modifikation, die vor allem in der eukaryotischen Epigenetik eine große Rolle spielt.^[116-120]

1.2.2.1 Epigenetik

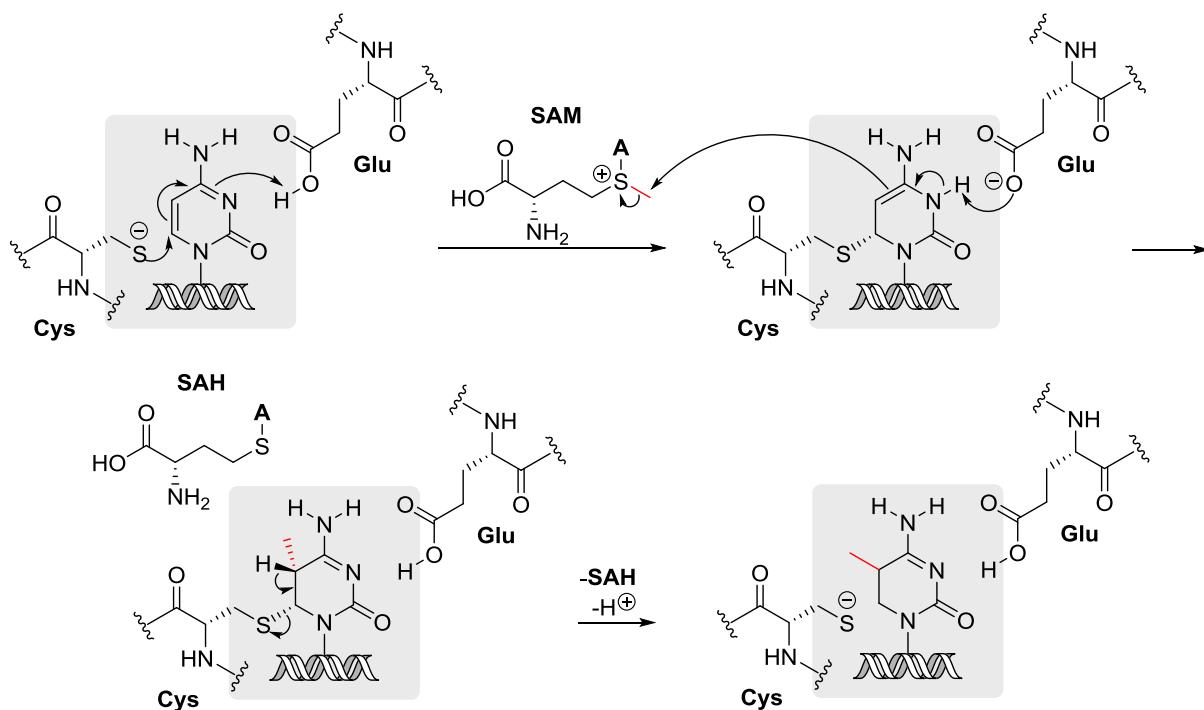
In einem multizellulären Organismus entspringen alle Zellen aus derselben Zygote und verfügen somit alle über das gleiche genetische Material. Nach der Zelldifferenzierung führen alle Zellen jedoch verschiedene Funktionen aus. Bestimmte Gene müssen also im Laufe der Zellentwicklung ein- oder ausgeschaltet werden, um den unterschiedlichen Aufgaben der Zellen gerecht zu werden. Man spricht bei der Steuerung dieser Prozesse von Genregulation. Die Epigenetik (von griechisch: epi = über) befasst sich mit allen meiotischen und mitotischen Veränderungen, welche an die Tochterzelle vererbt werden und nicht in der DNA-Sequenz, welche unverändert bleibt, festgelegt sind.^[121] Stattdessen kommt es zu kovalenten DNA-Modifikationen oder zur Modifikation von Histonen. DNA-Methylierung führt zu einer meist permanenten Repression von Genen, während es sich bei Modifikationen von Histonen um eher reversible Prozesse handelt. Inwiefern sich diese beiden regulatorischen Prozesse gegenseitig bedingen, ist derzeit Gegenstand der Forschung. Fehlerhafte Modifikationsmuster können zur Expression oder Repression der falschen Gene und letztendlich zur Entstehung von Krebs-, Autoimmun-, neurodegenerativen und Stoffwechselerkrankungen führen.^[116-118, 122-123]

1.2.2.2 C-Methylierung von DNA

In Säugetierzellen sind 4-6 % aller Desoxycytidine methyliert,^[124-125] wobei m^5dC hauptsächlich in CpG-Dinukleotiden auftritt, in denen zelltypspezifisch 70-80 % der Cytidine methyliert sind.^[118, 126] Stellen im Genom, in denen eine besonders hohe Dichte an CpG-Dinukleotiden vorherrscht, nennt man CpG-Inseln. Diese befinden sich häufig in Promotorenregionen und machen ca. 10 % aller CpG-Dinukleotide aus. Meist sind diese

CpG-Inseln hypomethyliert, also vor Methylierung geschützt.^[118, 127] Der Methylierung von dC an den TSS wurde bereits 1975 von zwei Gruppen unabhängig voneinander eine regulatorische Funktion zugesprochen.^[8-9] m⁵dC befindet sich in der großen Furche der DNA, wodurch es die Wechselwirkung von Transkriptionsfaktoren und Chromatin remodulierenden Proteine mit der DNA beeinflussen kann.^[119] Die DNA-Methylierung kann also zu einer indirekten Hemmung der Transkription, d.h. zur Stilllegung von Genen, führen.^[128-129] Dadurch spielt sie eine wichtige Rolle bei der X-Chromosom-Inaktivierung, der genomischen Prägung und der dauerhaften Stilllegung der Pluripotenzgene während der zellulären Entwicklung.^[116-117]

Zur Einführung der Methylgruppe an die 5-Position von dC kommt es postreplikativ durch DNMT mit Hilfe von SAM als Cofaktor (Schema 4).



Schema 4: Katalytischer Mechanismus von DNA-C-Methyltransferasen.^[130]

Zunächst bildet die jeweilige DNMT einen ternären Komplex mit der DNA und mit SAM.^[131] Im Zuge eines *flipping*-Mechanismus wird Cytosin aus der DNA-Doppelhelix herausgedreht. Aufgrund der geringen Elektronendichte von Cytosin kann die C5-Position die Methylgruppe von SAM nicht ohne weiteres nukleophil angreifen. Deshalb geht zunächst ein Cysteinrest der Bindungstasche des Enzyms - in einer *Michaeli*-ähnlichen 1,4-Addition - mit dem Kohlenstoff in Position 6 eine kovalente Bindung ein. Die transiente Protonierung des endozyklischen N3-Stickstoffs führt zu einer ausreichenden Aktivierung der C5-Position um die Methylgruppe von SAM nukleophil angreifen zu können. Das überschüssige Proton in C5-Position wird basisch abstrahiert und es kommt zur Rearomatisierung der

Nukleobase.^[108, 130] Katalysiert wird die Methylierung von dC also durch DNMTs. In Eukaryoten sind drei DNMTs bekannt. Die *maintenance* Methyltransferase DNMT1 sorgt zusammen mit UHRF1 (*Ubiquitin-like plant Homeodomain and RING Finger domain 1*) für die Erhaltung bestehender Methylierungen. Dadurch wird m⁵dC zu einer vererbaren Modifikation.^[132-133] Die Methyltransferasen DNMT3a und DNMT3b sind hingegen für die *de novo*-Methylierung zuständig, welche nach der globalen Demethylierung der befruchteten Eizelle stattfindet (siehe Kapitel 1.2.2.3).^[108, 133] DNA-Methylierung und somit auch die Methyltransferasen sind für Säugetiere unentbehrlich, wie anhand von *Knockout*-Mäusen, die im Verlauf ihrer frühen Entwicklung starben, demonstriert werden konnte.^[133-136] Zusätzlich gibt es die katalytisch inaktive DNMT3I, die die DNA-Methylierung durch DNMT3a im Menschen anregt.^[137]

1.2.2.3 Aktive und passive Demethylierung

Alle Säugetierzellen stammen von einer Eizelle, der Zygote, ab. Auf ihrem Weg entlang des Eileiters zur Gebärmutter teilt sich diese über die 2- 8-Zellstadien, zur Morula und zur Blastozyste weiter. Die ES-Zellen der Blastozyste sind pluripotent, d.h. sie können sich in alle drei Keimblätter (Ento-, Meso-, Ektoderm) des Embryos weiterentwickeln, und so zu jeder Gewebezelle im Organismus werden. Diese ES-Zellen können mit Hilfe von geeigneten Medien so kultiviert werden, dass ihr pluripotenter Status erhalten bleibt und sie in unterschiedliche Entwicklungsstadien versetzt bzw. in viele Zelltypen ausdifferenziert werden können, wodurch man den natürlichen Gegebenheiten in der Blastozyste sehr nahekommt.^[138-140] Um diesen pluripotenten Status zu etablieren, kommt es während der Differenzierung von der Zygote zur Blastozyste, zu einer nahezu vollständigen globalen Demethylierung sowohl des männlichen als auch des weiblichen Genoms.^[141-143]

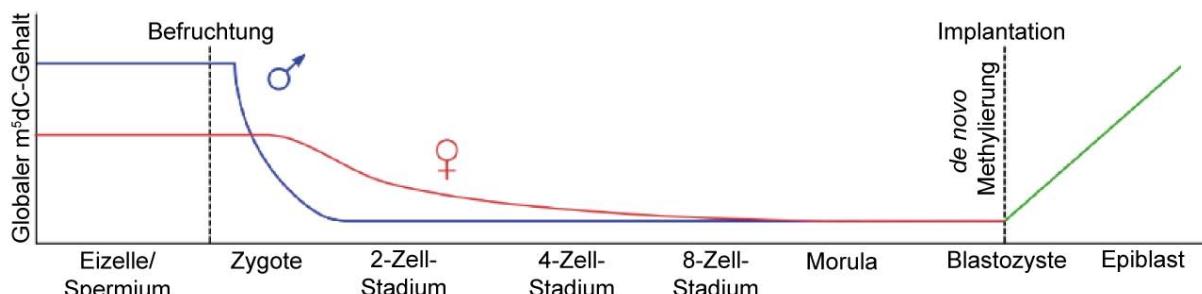
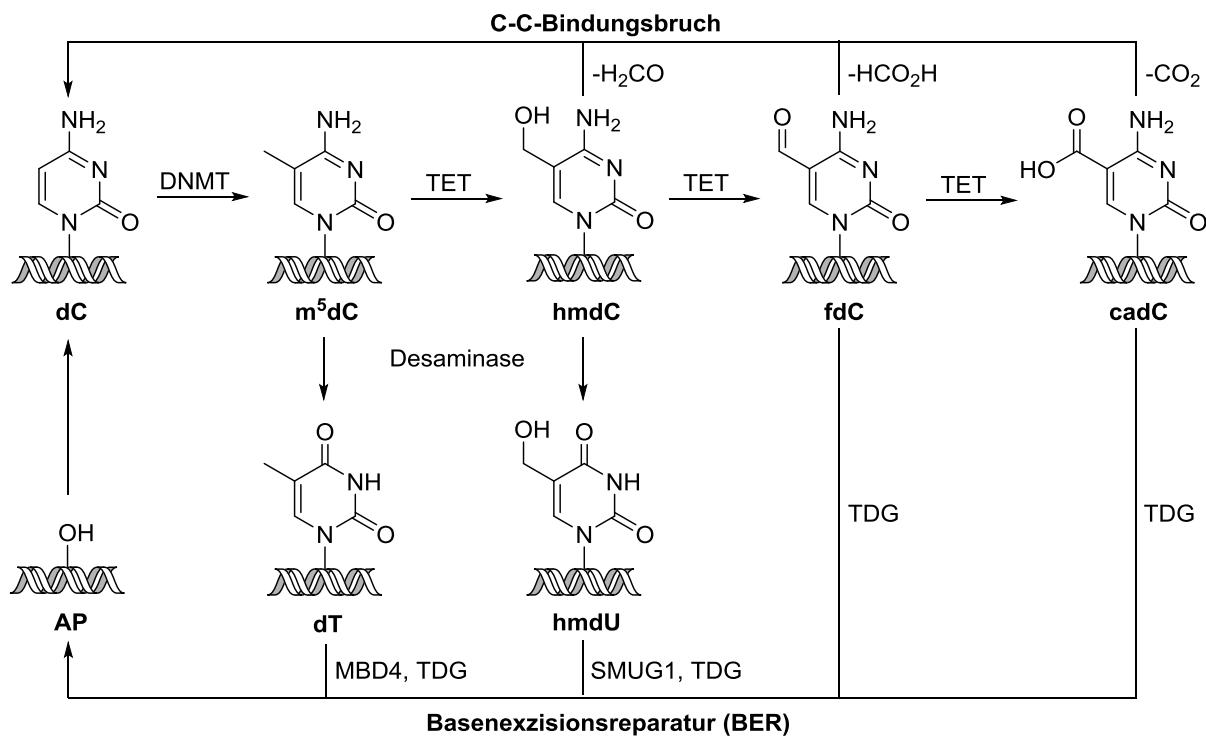


Abbildung 3: Veränderung des globalen m⁵dC-Gehalts während der Embryogenese. Nach der Befruchtung der Eizelle auf deren Wanderung entlang des Eileiters zur Gebärmutter, wird das paternale Genom weitgehend aktiv demethyliert (blaue Linie). Im maternalen Genom sinkt der m⁵dC-Gehalt hauptsächlich durch passive Demethylierung (rote Linie). Nach der Implantation kommt es zur *de novo* Methylierung des Genoms (grüne Linie). Die Abbildung wurde von Wu et al. adaptiert.^[144]

Die Demethylierung des weiblichen Genoms verläuft dabei langsamer ab, als die des männlichen Genoms, weshalb man lange nur von einem rein passiven Demethylierungsmechanismus ausging, d.h. dem Verlust der Methylgruppe durch Verdünnung bei der DNA-Replikation.^[145-146] 2014 haben Guo *et al.* jedoch eine neue Studie veröffentlicht, die auch im Fall des maternalen Genoms auf eine teilweise aktive, d.h. Enzym-vermittelte Demethylierung hinweist.^[147] Die Demethylierung der elterlichen Genome kommt einer epigenetischen Reprogrammierung gleich und ist wichtig, um die Entwicklung des Embryos aus nur einer Zelle nicht zu gefährden. Nach dem Einnisten der Eizelle in die Gebärmutterschleimhaut, der sogenannten Implantation, kommt es zu einer gezielten *de novo* Methylierung durch DNMT3a und b bestimmter Bereiche beider Genome.^[141, 144, 148] In ausdifferenzierten Zellen ist eine aktive Demethylierung bis dato nur vereinzelt und lokusspezifisch in Neuronen und Immunzellen bekannt.^[149-150]

Wie diese aktive Entfernung von Methylierungsmustern genau abläuft, ist bis heute nicht zweifelsfrei geklärt. In Pflanzen existieren spezielle Glykosylasen, die m⁵dC direkt ausschneiden, sodass die dadurch entstehende abasische Seite mittels Basenexzisionsreparatur (BER) repariert werden kann.^[120, 151-154] In Säugetieren konnte bisher keine solche Glycosylase identifiziert werden und die direkte Demethylierung von m⁵dC ist auf Grund der sehr stabilen C-C-Doppelbindung an Position 5 unwahrscheinlich. Deshalb müssen ein oder mehrere alternative Wege existieren, bei denen dieser Bindungsbruch umgangen wird. In Schema 5 sind mögliche Wege der aktiven Demethylierung von m⁵dC zu dC dargestellt.



Schema 5: Schematische Darstellung möglicher aktiver Demethylierungswege von m^5dC zu dC . AP = abasische Stelle.

Zunächst wurde eine gezielte Desaminierung von m^5dC zu dT untersucht (Schema 5, unterer Weg). Das entstehende $dG:dT$ *Mismatch* würde von der Thymin DNA Glykosylase (TDG) und dem Methyl-CpG-Bindungsprotein 4 (MBD4) erkannt, und dT durch dC ersetzt.^[155-158] *In vitro* konnte gezeigt werden, dass die Desaminasen AID (*Activation-Induced Deaminase*) und APOBEC1 (*Apolipoprotein B pre-mRNA Editing Enzyme, Catalytic Polypeptide 1*) m^5dC zu dT desaminieren können.^[159-160] AID/MBD4 sind in Zebrafischen für die DNA-Demethylierung verantwortlich,^[161] in der Maus hingegen wird nur ein kleiner Teil des m^5dC so demethyliert und ein *Knockout* dieser Enzyme führt zu keinerlei Entwicklungsstörungen.^[162-163] Auch eine Desaminierung durch DNMTs wurde diskutiert, sollte aber biologisch nicht relevant sein.^[120, 164-165] 2009 gelang ein Durchbruch in der Untersuchung der aktiven Demethylierungsmechanismen, indem nachgewiesen werden konnte, dass TET-Enzyme m^5dC zu 5-Hydroxymethyl-2'-desoxycytidin (hmdC) oxidieren können.^[166] Es wurde eine AID/APOBEC1-vermittelte Desaminierung mit anschließender BER durch die Glykosylasen SMUG1 (*Single-strand selective Monofunctional Uracil DNA Glycosylase*) und TDG beschrieben (Schema 5, unterer Weg),^[167-170] die jedoch in mES-Zellen vernachlässigbar ist.^[171] All die beschriebenen Mechanismen beruhen letztendlich auf der Metabolisierung von m^5dC zu einer Base, die in Kombination mit dG als *Mismatch* erkannt wird und daraufhin durch eine geeignete Glykosylase aus der DNA geschnitten wird. Dies führt jedoch zur Bildung von abasischen Stellen, was bei einem rapiden Demethylierungsprozess, wie dem des paternalen Genoms nach Befruchtung der Eizelle, in

einer hohen genomischen Instabilität münden könnte. Vor allem in m^5dC -reichen Promoterregionen könnten als Folge eine große Anzahl an DNA-Doppelstrangbrüchen stattfinden.^[172] Eine interessante Alternative bietet die stufenweise Oxidation von $hmCdC$ über die 2011 entdeckten Intermediate 5-Formyl- (fdC) und 5-Carboxy-2'-desoxycytidin ($cadC$).^[173-174] Zwar können auch diese direkt mittels TDG aus der DNA geschnitten werden,^[173, 175-178] aber auch ein C-C-Bindungsbruch an Position 5 dieser Modifikationen wäre eine denkbare Alternative ohne Bildung abasischer Stellen (Schema 5, oberer Weg). Guo *et al.* haben zudem beobachtet, dass es in Zygoten ohne TDG-Aktivität auch zur Demethylierung von m^5dC zu dC kommt.^[147] *In vitro* konnte sowohl die Abspaltung von Formaldehyd von $hmCdC$,^[179] als auch die Deformylierung von fdC mit Mercaptoethanol beobachtet werden.^[180] Eine Decarboxylierung von $cadC$ konnte ebenso in Stammzellextrakt nachgewiesen werden,^[181] wie auch *in vitro* mittels eines Komplexes aus den Methyltransferasen DNMT3a/DNMT3l.^[182] Hinweise, dass auch *in vivo* ein C-C-Bindungsbruch möglich ist, konnten im Arbeitskreis Carell zunächst durch *Isotope-Tracing*-Studien in TDG-Knockout mES-Zellen erhalten werden.^[183] Untermauert wurden diese Ergebnisse mit Hilfe von 2'-Fluor-substituierter Analoga der modifizierten Basen, die zu einer Hemmung der TDG-Aktivität führten, aber ansonsten keinerlei Beeinträchtigung epigenetisch relevanter Enzyme auf Grund der F-Substitution zeigten.^[183-184] Die Klärung, welche dieser Mechanismen *in vivo* tatsächlich stattfinden, ist von höchstem wissenschaftlichen Interesse, da dies zum Verständnis der Steuerung der Zellentwicklung beitrüge. Zum einen könnte so aufgeklärt werden, in welcher Korrelation bestimmte Krankheitsbilder mit einer fehlerbehafteten Genexpression stehen, zum anderen wäre eine gezielte Reprogrammierung von bereits ausdifferenzierten Zellen ein Durchbruch in der Medizin.^[185]

2 Aufgabenstellung

Diese Arbeit befasst sich mit aryl- bzw. alkylsubstituierten DNA-Basen. Zum einen sollen verschiedene arylgeschädigte DNA-Stränge mit der DNA-Bindungsdomäne von Rad14, dem Hefehomolog von XPA, kristallisiert werden. Zum anderen sollen isotopenmarkierte methylierte DNA-Basen für die Quantifizierung der natürlich vorkommenden Basen in verschiedenen Organismen und Zelltypen hergestellt werden um deren biologische Relevanz zu untersuchen.

2015 konnte im Arbeitskreis Carell die Kristallstruktur eines 15mer DNA-Strangs, welcher C8-AAF-dG als Schaden enthielt, im Komplex mit zwei Rad14 Molekülen erhalten werden.^[186] In dieser Struktur liegt die DNA geknickt vor. Der Schaden und Desoxyguanosin liegen in einer Ebene und bieten somit eine große Oberfläche für π -stacking Interaktionen, welche den Knick stabilisieren und somit dazu beitragen könnten, dass dieser Schaden von XPA/Rad14 erkannt wird. Es soll untersucht werden inwieweit diese π -stacking Interaktionen zur Stabilisierung des Knicks beitragen. Aus diesem Grund sollen C8-Desoxyguanosinaddukte mit verschiedenen großen aromatischen Resten mittels DNA-Festphasensynthese in DNA-Stränge eingebaut und mit Rad14 co-kristallisiert werden. Zum Vergleich soll das C8-AAF-Phosphoramidit von 2'-Desoxyadenosin **6** sowie die N^2 -dG-Addukte **7** - **9**, welche sich analog zu den C8-dG-Addukten in der Größe des aromatischen Rests unterscheiden, synthetisiert (Abbildung 4), in die DNA eingebaut und ebenfalls mit Rad14 co-kristallisiert werden.

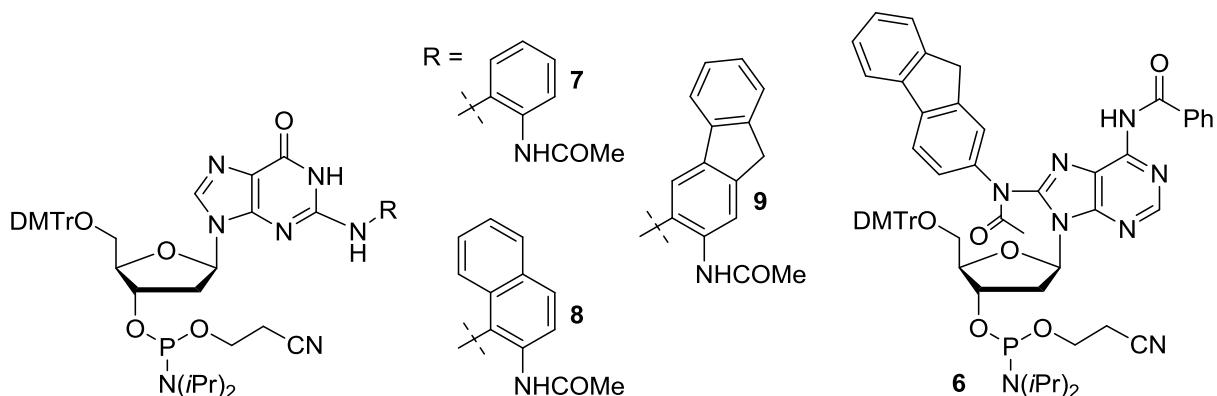


Abbildung 4: Phosphoramidite der N^2 -dG Addukte und des C8-AAF-dA Addukts.

Im zweiten Teil der Arbeit sollen isotopenmarkierte Nukleoside der in der Natur vorkommenden methylierten DNA-Basen m⁵dC, m⁴dC und m⁶dA synthetisiert werden (Abbildung 5).

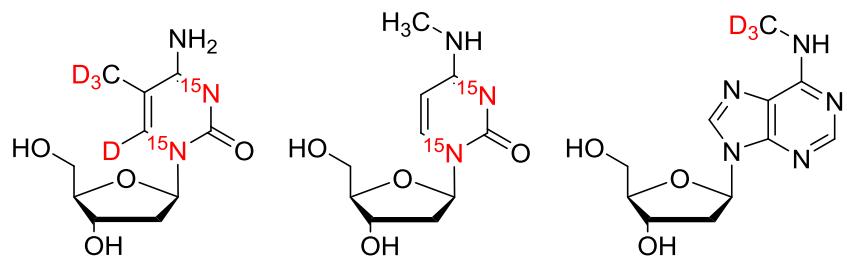


Abbildung 5: Isotopenmarkierte Nukleoside [$D_4, ^{15}N_2$]- m^5dC , $^{15}N_2$ - m^4dC und D_3 - m^6dA .

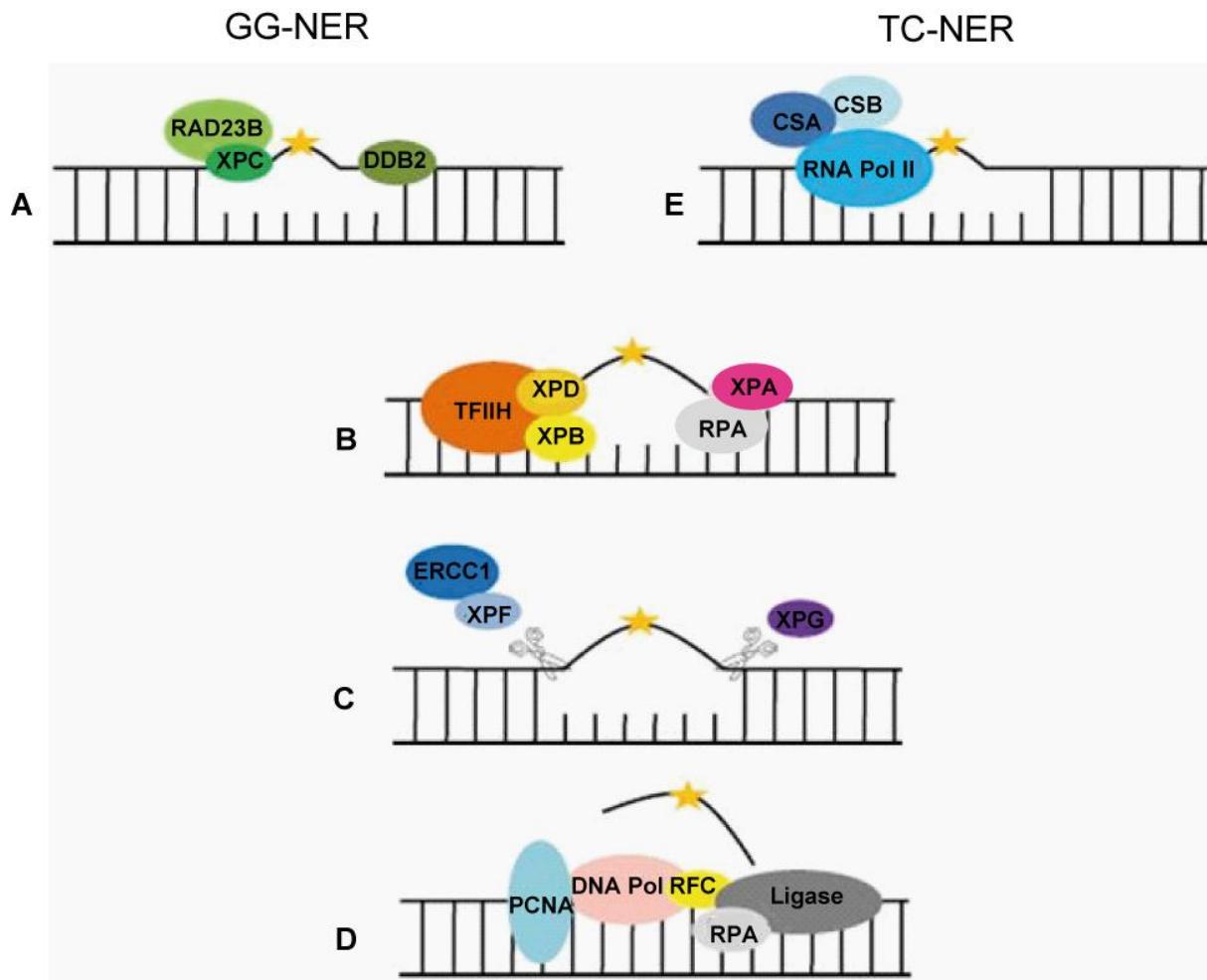
[$^{15}N_2$]- m^4dC und [D_3]- m^6dA sollen für die massenspektrometrische Quantifizierung der natürlichen Nukleoside m^4dC und m^6dA in Bakterien und in Geweben höherer Lebewesen verwendet werden. [$D_4, ^{15}N_2$]- m^5dC wird für die Quantifizierung von [$D_3, ^{13}C$]- m^5dC aus Fütterungsexperimenten mit schwerem Methionin benötigt.

3 Ergebnisse und Diskussion

3.1 Arylmodifizierte DNA-Basen

3.1.1 Nukleotidexzisionsreparatur (NER) von C8- und N²-Addukten

Sperrige Adduktschäden, die beispielsweise in C8- und N²-Position von dG entstehen können, werden von der Nukleotidexzisionreparatur (NER) erkannt und entfernt. Eine defekte NER führt zu schweren Krankheiten wie *Xeroderma Pigmentosum* (XP) und dem Cockayne Syndrom (CS). Es sind acht XP Komplementationsgruppen bekannt, von denen sieben (XPA-XPG) durch Mutationen in Genen, die im NER Prozess involviert sind, verursacht werden. Die NER erkennt eine Vielzahl strukturell unterschiedlicher Substrate und verläuft nach einem „cut-and-patch“-Mechanismus, wobei ein 24-32 Nukleotide-langes geschädigtes Stück aus der DNA herausgeschnitten und der intakte Gegenstrang als Vorlage für die Synthese des neuen Stranges verwendet wird. Unterschieden wird zwischen der Globalen Genom NER (GG-NER), die Schäden im gesamten Genom erkennen und reparieren kann, und der Transkriptions gekoppelten NER (TC-NER), die Schäden im Zuge der Transkription aktiver Gene erkennt und repariert.^[187-188] Die Entfernung und Reparatur des Schadens verläuft hierbei durch die gleichen Enzyme, die Schadenserkennung hingegen erfolgt hingegen auf unterschiedliche Art und Weise.^[189] Man nimmt an, dass die große Substratvielfalt der GG-NER möglich ist, da diese sowohl fehlerhafte Basenpaarungen als auch diverse DNA-Verformungen detektieren kann, also kurz alle Parameter, welche die thermodynamische Stabilität beeinflussen.^[190-191] Die Reparaturrate ist hierbei proportional zur Verringerung der thermodynamischen Stabilität, welche durch einen bestimmten Schaden hervorgerufen wird.^[192] In Schema 6 ist der Verlauf der NER in Eukaryoten dargestellt.



Schema 6: Mechanismus der NER. GG-NER und TC-NER unterscheiden sich in der Schadenserkennung, während dessen Entfernung durch die gleichen Kernenzyme erfolgt. PCNA = *Proliferating-Cell-Nuclear-Antigen*, RFC = Replikationsfaktor C.

Es wird angenommen, dass XPC im Komplex mit RAD23B die Schadenserkennung in der GG-NER übernimmt.^[193] Im Fall von UV-Schäden hingegen wird XPC vom primären Schadenserkenner UV-DDB (UV-DNA Schadenbindungsprotein) zum Schaden rekrutiert (Schema 6A).^[194-196] Erkennt XPC eine Strukturabweichung so rekrutiert es seinerseits den Transkriptionsfaktor TFIID zur verdächtigen Stelle. Dessen Untereinheiten XPD und XPB entwinden ca. 25-30 Nukleotide des DNA-Doppelstrangs und ersteres verifiziert den Schaden.^[197-199] Im Anschluss werden XPA und das Replikationsprotein RPA zur Bildung des Präinzisionskomplexes beordert (Schema 6B). Auch XPA könnte in der Schadensverifizierung involviert sein.^[200-201] Die Endonuklease XPG stabilisiert den Schaden und sorgt für die Abspaltung von XPC-RAD23B.^[202-203] XPA spielt eine Schlüsselrolle bei der NER, da es die beteiligten Faktoren vor der Inzision richtig gruppiert und ausrichtet.^[204-212] Seine hohe konformationelle Flexibilität könnte ein Grund für die Tatsache sein, dass XPA mit einer Vielzahl strukturell unterschiedlicher Verbindungen interagieren kann.^[213-215] Auch RPA interagiert mit der DNA und den beteiligten Enzymen und dient so zusammen mit XPA

der Stabilisierung während der NER.^[202, 216-217] XPA sitzt hierbei am 5'-Ende des Schadens^[211], während sich RPA auf dem intakten Stang befindet.^[218-219] XPA und RPA positionieren nun die Endonukleasen XPG (für die 3'-Inzision) und XPF-ERCC1 (für die 5'-Inzision).^[220-221] Die Endonukleasen XPF/ERCC1 und XPG schneiden die beschädigte DNA heraus, damit diese entfernt werden kann (Schema 6C).^[222-224] Zum Abschluss wird der entfernte Strang ersetzt. Diese Aufgabe wird von den normalen Replikationsenzymen übernommen. Hierbei wird der Gegenstrang als Vorlage genutzt. Eine Ligase „kittet“ den Einschnitt durch Verknüpfung des neusynthetisierten Strangs mit dem alten Strang (Schema 6D). Eine detaillierte Beschreibung des Mechanismus und der an der NER beteiligten Enzyme findet sich im Review „*Molecular mechanisms of xeroderma pigmentosum (XP) proteins*“ im Anhang.

Bei der TC-NER hingegen übernimmt die RNA Polymerase II die Schadenserkennung und stoppt, wenn sie während der Transkription auf einen Schaden stößt. Da ein Aussetzen der Transkription zum Erliegen der zellulären Funktionen führen kann, werden bei der TC-NER unterschiedliche Arten von Schäden mit ähnlicher Effizienz entfernt.^[225-227] CSB wird zur Schadensverifizierung rekrutiert.^[228-229] Kleinerer Schäden können von der Polymerase überlesen werden,^[230-231] während größere Schäden (UV-induzierte, Cisplatin oder *Bulky*-Addukt-Schäden) entfernt werden müssen. CSB beordert hierzu die Kernenzyme der NER und CSA in die Nähe des Schadens (Schema 6E).^[232-234] Die Beseitigung des Schadens erfolgt analog der GG-NER.

Da alle Desoxyguanosinaddukte den DNA-Doppelstrang unterschiedlich stark verformen, verläuft auch die NER dieser Addukte unterschiedlich gut und nicht jedes Addukt kann auf dem gleichen Weg erkannt und repariert werden. Schmelzpunktmessungen verschiedener geschädigter dsDNA-Stränge zeigen, dass C8-AAF-dG Addukte die thermodynamische Stabilität des Strangs am meisten absenken, C8-AF-dG Addukte zu einer weitaus geringeren Destabilisierung führen und *N*²-AAF-dG Addukte sogar für eine thermodynamische Stabilisierung des Doppelstrangs sorgen.^[49-50, 56, 235] Mu et al. haben C8-AAF-dG und C8-AF-dG an drei verschiedene Stellen der *NarI*-Sequenz (G₁G₂CG₃CC) eingebaut, die als Mutagenese *hot-spot* bekannt ist, und die NER Effizienz, mit der diese Addukte repariert werden, verglichen.^[236-239] Am besten wird ein C8-AAF-dG Schaden an Position G₂ repariert, während ein C8-AF-dG Schaden an gleicher Stelle am schlechtesten von der NER-Maschinerie erkannt wird. Außerdem werden C8-AAF-dG Addukte in jedem Fall besser repariert als die C8-AF-dG Addukte.^[50] Die Populationen der einzelnen Konformationen an den verschiedenen Positionen der *NarI*-Sequenz wurden mittels ¹⁹F-NMR-Spektroskopie ermittelt. Da dies aber eine stark temperaturabhängige Methode ist und sich die Konformation in Anwesenheit der Enzyme der NER ändern kann, ist nicht davon auszugehen, dass die Verteilung immer identisch ist.^[50, 59] Es ist aus Studien mit

verschiedenen sperrigen Addukten bekannt, dass sie in der S-Konformation (Abbildung 2) am besten durch die GG-NER erkannt und repariert werden können, da diese zur größten Verformung der Doppelhelix führt.^[190, 240-245] Vergleicht man die S-Konformation von C8-AF-dG und C8-AAF-dG, so werden einige Unterschiede sichtbar, die zeigen, dass erstere zu einer weitaus geringeren thermodynamischen Destabilisierung der DNA und damit zu einer schlechteren Erkennung durch NER-Enzyme führt. Interkaliert C8-AAF in die Doppelhelix, so steht es auf Grund der sterischen Hinderung der Acetylgruppe fast senkrecht zum geschädigten dG. Dadurch wird das *stacking* zwischen dem Fluorenylrest und den Basen des Gegenstrangs zwar optimiert, dG ist aber nicht mehr in der Lage π -Wechselwirkungen mit seinen Nachbarbasen einzugehen. Im Fall von C8-AF-dG sind die geschädigte Base und der Fluorenylrest beinahe komplanar, sodass beide π -Wechselwirkungen eingehen können, die zu einer Stabilisierung der Doppelhelix führen.^[50] Um die Sequenzabhängigkeit der NER zu erklären, haben *Mu et al.* Molekulardynamik Simulationen des C8-AAF-dG Addukts an den drei Positionen der *NarI*-Sequenz durchgeführt. Dabei sollte das *stacking* zwischen AAF und den Basen des Gegenstrangs optimiert werden, da dieses maßgeblich zur Stabilisierung der dsDNA beiträgt. Dies führt an Position G₂ zur größten Entwindung der DNA und Vergrößerung der kleinen Furche – beides Faktoren die letztendlich zu einer thermodynamischen Destabilisierung führen.^[50]

*N*²-AAF-dG führt hingegen zu einer thermodynamischen Stabilisierung der DNA, da es sich in der kleinen Furche befindet.^[49] Daher sollte es kein gutes Substrat für die GG-NER sein.^[191] Die Proteine CSA und CSB werden beide für die TC-NER benötigt.^[187] *Kitsera et al.* führten Expressionsvektoren, die auf bestimmten Positionen des Reportergens *Enhanced Green Fluorescent Protein* (EGFP) eine C8-AAF-dG, eine TT-Dimer oder eine *N*²-AAF-dG Modifikation trugen, in CSB oder CSA defiziente Zellen ein. Dadurch konnten sie bestätigen, dass die Reparatur von C8-AAF-dG wie erwartet nicht CSA bzw. CSB abhängig ist, sondern dass dieser Schaden durch GG-NER repariert wird. Erstaunlicherweise wurden TT-Dimere, die allgemein als Modelschäden für die TC-NER gelten, zu größeren Teilen als der *N*²-AAF-dG Schaden repariert, der zu einem vollkommenen Stopp der RNA Polymerase II (RNAPII) führte. Wurde der Schaden jedoch in den kodierenden Strang oder in eine Genregion, die nicht transkribiert wird, eingebaut, so kam es zu keinerlei Behinderung der Transkription. Diese Beobachtung lässt vermuten, dass eine Interaktion von *N*²-AAF-dG mit dem Transkriptionskomplex vornötig ist um die EGFP-Expression zu inhibieren. Der Vergleich verschiedener Zelllinien, die entweder CSA, CSB, XPC oder XPA-defizient waren, zeigte, dass *N*²-AAF-dG nur beim Fehlen der TC-NER relevanten Enzyme CSA, CSB und XPA zum Transkriptionsstopp führte. Das Fehlen von XPC, dem Erkennungsprotein der GG-NER, hatte keinerlei Auswirkung auf die Genexpression.^[246]

Über welchen der beiden Wege C8-AF-dG repariert wird ist derzeit noch nicht vollständig geklärt. Bei Fütterung von menschlichen Fibroblasten mit *N*-Acetoxy-2-AAF entstand C8-AF-dG als Hauptprodukt. Beim Vergleich von normalen und Zellen von CS-Patienten konnte weder bei der Induktion noch bei der Reparatur des Schadens ein Unterschied festgestellt werden, sodass davon ausgegangen werden kann, dass TC-NER bei der Reparatur keine signifikante Rolle spielt.^[247]

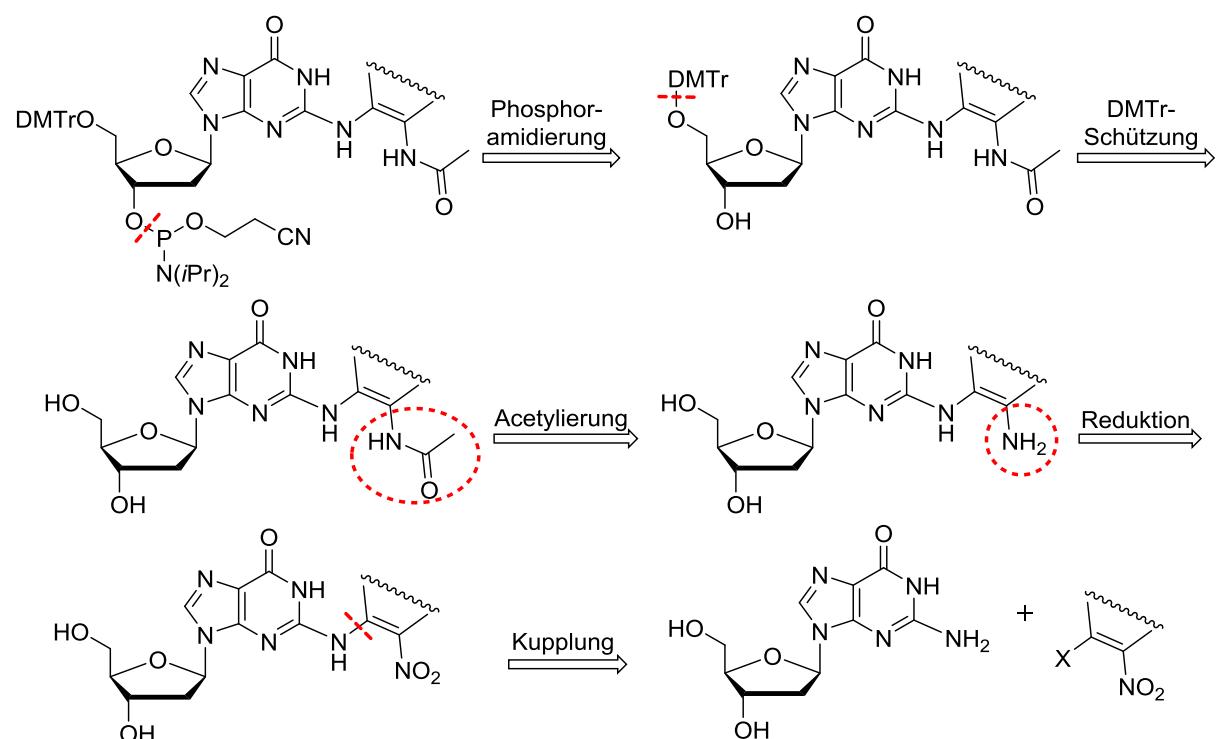
3.1.2 Kristallisation von C8-AAF-dG geschädigter DNA mit Rad14

XPA ist als Schlüsselprotein sowohl für die GG-NER als auch für die TC-NER unentbehrlich. Da dieses Protein eine erhöhte Affinität für geschädigte DNA besitzt, spielt es möglicherweise zusätzlich auch eine Rolle in der Schadenserkennung.^[200-201, 248-249] In unserer Gruppe konnte 2015 eine Kristallstruktur von geschädigter C8-AAF-dG dsDNA im Komplex mit zwei Rad14-Molekülen erhalten werden, welche an dieser Stelle kurz beschrieben werden soll.^[186] Im Einklang mit EMSA-Studien bindet bei dieser Struktur das Reparaturprotein als Homodimer an die DNA. Die α/β -Faltung von Rad14 ist hierbei der Faltung des menschlichen Homologs XPA in der NMR-Struktur sehr ähnlich.^[250-251] Der Kontakt zwischen den beiden Proteinen und dem DNA-Duplex findet nicht durch die Zinkfingerdomänen, sondern hauptsächlich über die β -Schleife statt. Dabei werden die Enden der DNA aufgeschmolzen und nur ein 13mer Erkennungsmotiv mit dem Schaden in zentraler Position bleibt intakt. Mit Ausnahme von Asn256 werden nur Wasserstoffbrückenbindungen zum Phosphatrückgrat ausgebildet. Die Schadenserkennung ist demnach sequenzunabhängig. Es kommt zu π -stacking Interaktionen zwischen den letzten intakten Basenpaaren und Phe262 und His258. Die Struktur besitzt eine pseudo C2-Symmetriearchse, welche durch den Schaden führt. Der DNA-Doppelstrang liegt in zwei unterschiedlichen Orientierungen mit einer Verteilung von je ca. 45-55 % vor. Er weist einen Knick von ca. 70° an der geschädigten Position auf, was darauf hindeutet, dass Rad14 nicht den Schaden an sich erkennt, sondern die Flexibilität des Strangs testet, welche durch den Schaden erhöht wird. Der C8-AAF-dG Schaden liegt in der Kristallstruktur in syn-Konformation vor, wobei dG und AAF beinahe in einer Ebene liegen und somit eine große Oberfläche für π -stacking Interaktionen bieten. Diese finden nicht nur mit dem oben angrenzenden dT:dA Basenpaar, sondern auch mit dem unteren dC:dG Basenpaar statt. Sie stabilisieren den DNA-Knick, was dazu beitragen könnte, dass dieser Schaden sehr gut von Rad14/XPA erkannt wird.

3.1.3 Synthese von *bulky* Addukten

Im Folgenden ist die Synthese der Phosphoramidite von N^2 -Acetylaminophenyl- (N^2 -AAB-dG, **7**), N^2 -Acetylaminonaphthyl- (N^2 -AAN-dG, **8**) und N^2 -AAF-dG (**9**), sowie des Phosphoramidits von C8-AAF-dA (**6**) dargestellt, ebenso wie deren Einbau in DNA-Stränge. Auch verschiedene große C8-dG-Addukte wurden in DNA-Stränge eingebaut. Die erhaltenen Kristallstrukturen von dsDNA und Rad14 sollen weitere Einblicke in die Bedeutung der oben beschriebenen π -stacking Interaktionen geben.

Die Phosphoramidite der *bulky* N^2 -dG-Addukte (Abbildung 4) sollten retrosynthetisch nach Schema 7 hergestellt werden.

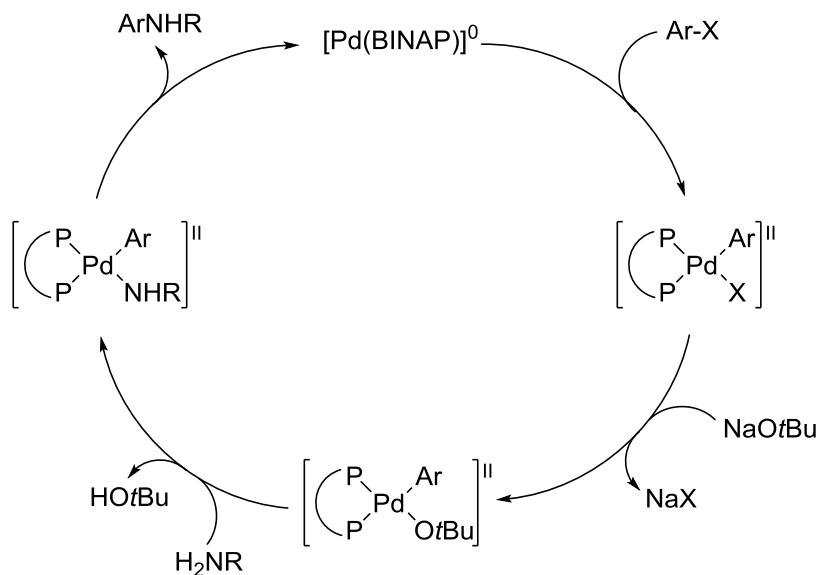


Schema 7: Retrosynthese von *bulky* N^2 -dG-Addukten. Zunächst sollte das Nitroderivat des jeweiligen aromatischen Rests an die N^2 -Position von dG gekuppelt werden und im Anschluss reduziert und acetyliert werden. Zum Abschluss sollte der $O^{5'}$ -Sauerstoff der Ribose DMTr-geschützt werden und das Phosphoramidit am O^3' -Sauerstoff installiert werden. X = Halogenid, Triflat.

Zunächst soll das Nitroderivat des jeweiligen aromatischen Rests nach *Buchwald-Hartwig* an die N^2 -Position von dG gekuppelt werden. 2'-Desoxyguanosin fungiert bei dieser Reaktion als Nukleophil, während der Nitroaromat als Elektrophil wirkt. Ziel der Synthese ist die Herstellung von Phosphoramiditen von N^2 -Desoxyguanosinderivaten mit aromatischen Schäden, welche eine Acetylaminogruppe besitzen. Da es sich bei Acetylaminoaromaten um elektronenreiche Verbindungen handelt, kann die Kupplung nicht direkt mit diesen durchgeführt werden. Stattdessen sollen Nitroaromaten verwendet werden, welche dann

zum Amin reduziert und im Anschluss acetyliert werden sollen, um zur gewünschten Acetylaminofunktion zu gelangen. Da sich elektronenarme aromatische Halogenide und Triflate aufgrund der hohen Elektrophilie gut für *Buchwald-Hartwig-Kupplungen* eignen, sind Aromaten mit einer elektronenziehenden Nitrogruppe gute Substrate für diese Reaktion. Zum Abschluss soll die Ribose für die DNA-Synthese modifiziert werden. Der $O^{5'}$ -Sauerstoff soll DMTr-geschützt werden und das Phosphoramidit am $O^{3'}$ -Sauerstoff installiert werden.

Schema 8 zeigt den Katalysezyklus der *Buchwald-Hartwig-Kupplung*, welche den Schlüsselschritt der Synthesen darstellt.^[252-255] Hierbei bilden Palladium und der Ligand zunächst einen katalytischen Komplex, in welchem das Palladium eine Oxidationszahl von 0 besitzt. Nach Zugabe der Kupplungskomponenten und der benötigten Base erfolgt eine oxidative Addition des aromatischen Halogenids oder Triflates an den Palladiumkomplex. Die Oxidationszahl des Palladiums steigt hierbei von 0 auf +II. Durch die Base wird eine Abgangsgruppe aus diesem Komplex verdrängt und das Desoxyguanosinderivat kann mit der primären Aminogruppe daran koordinieren. Es kommt zur reduktiven Eliminierung, wobei das N-C-gekoppelte Addukt abgespalten wird. Der Palladiumkomplex kann nun einen neuen katalytischen Zyklus eingehen.



Schema 8: Katalysezyklus der *Buchwald-Hartwig-Kupplung*.

Die Ausbeute und die Geschwindigkeit von *Buchwald-Hartwig-Kupplungen* hängen von vielen unterschiedlichen Faktoren ab. Entscheidend ist zunächst die Nukleophilie bzw. Elektrophilie der Kupplungspartner. Aber auch die verwendete Palladiumspezies sowie der Ligand, das Lösungsmittel und die gewählte Base spielen eine große Rolle.^[256]

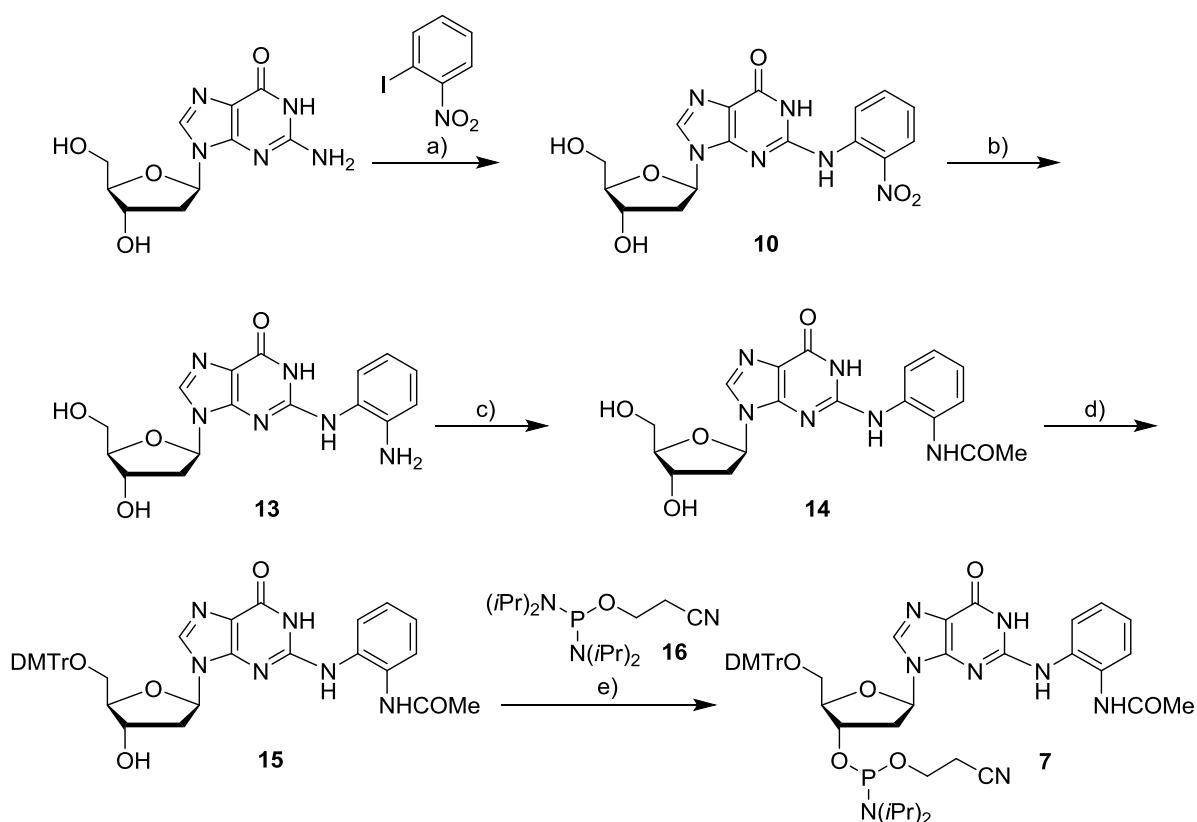
Die *Buchwald-Hartwig-Kupplung* sollte am ungeschützten dG durchgeführt werden, um möglichst wenige Schützungs- und Entschützungsschritte zu haben, welche die

Gesamtausbeute verringern. Nachteile bei der Verwendung von ungeschützten Nukleosiden sind deren geringe Löslichkeit in organischen Lösungsmitteln und die freien funktionellen Gruppen, welche zu Nebenreaktionen führen können. Klassischerweise werden die OH-Gruppen der Ribose mit *tert*-Butyldimethylsilylchlorid (TBS) geschützt um eine gute Löslichkeit zu garantieren, sowie die O⁶-Position des Guanosins in Form eines Enolats, um das saure Proton am N1-Amid zu entfernen. *Takamura-Enya et al.* konnten 2006 zeigen, dass es möglich ist auch ungeschütztes dG zu kuppeln, insofern man reaktivere Iodoaromaten verwendet. Als Lösungsmittel wird DMSO verwendet, den katalytischen Komplex bilden Tris(dibenzylidenaceton)dipalladium ($\text{Pd}_2(\text{dba})_3$) und 4,5-Bis(diphenylphosphino)-9,9-dimethylxanthen (Xantphos), während Tetraethylammoniumfluorid (TEAF) als Base fungiert.^[257] Das N²-Nitrophenyladdukt **10** konnte auf diesem Weg aus kommerziell erhältlichem dG und 1-Bromo-2-nitrobenzol in 99 % Ausbeute erhalten werden.

Für die N²-Addukte mit Naphthyl- und Fluorenylresten war diese Strategie nicht von Vorteil, da zunächst die Iodonitroaromaten synthetisiert werden mussten, welche nur in geringen Ausbeuten erhalten werden konnten und zwei zusätzliche Synthesestufen mit sich brachten. Deshalb sollten diese beiden Addukte ausgehend von O^{3'}- und O^{5'}-TBS geschützten dG synthetisiert werden. Als katalytischer Komplex fungierten ebenfalls $\text{Pd}_2(\text{dba})_3$ und Xantphos, während Cs₂CO₃ als Base verwendet wurde. Das N²-Nithronaphthyladdukt **11** wurde in Dioxan mit einer Ausbeute von 97 % erhalten, während das N²-Nitrofluorenyladdukt **12** in Toluol mit einer Ausbeute von 57 % hergestellt wurde. Es wurden 0,15 Äquivalente $\text{Pd}_2(\text{dba})_3$ und 0,3 Äquivalente Xantphos verwendet, damit dieser leichte Ligandenüberschuss den Katalysator auch bei den langen Reaktionszeiten von 14 bis 18 Stunden stabilisiert.

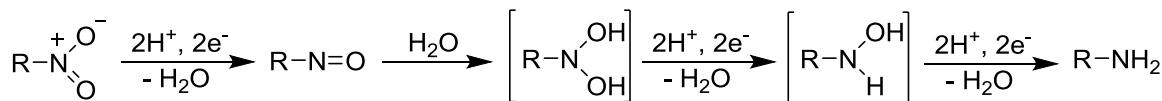
3.1.3.1 Synthese des N²-AAB-dG Phosphoramidits

Phosphoramidit **7** wurde über eine *Buchwald-Hartwig-Kupplung* ausgehend von dG synthetisiert (Schema 9).



Schema 9: Synthese des Phosphoramidits von *N*²-AAB-dG. a) Pd₂(dba)₃, Xantphos, TEAF, DMSO, 16 h, 100 °C, quant. b) Na₂S₂O₄, THF/H₂O, 4 h, RT, 80 % c) CH₃COSH, Cu₂SO₄, MeOH, 16 h, RT, 75 % d) DMTr-Cl, NEt₃, Pyridin, 4 h, RT, 39 % e) DIAT, DCM, 1,5 h, RT, quant.

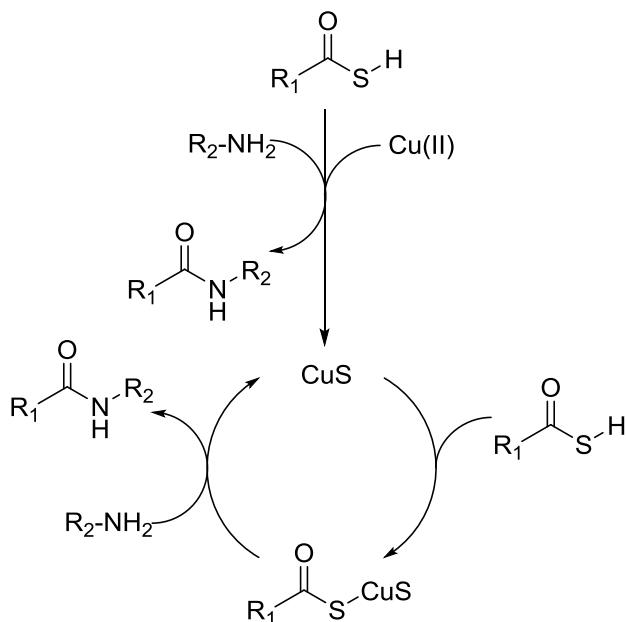
Die Nitrogruppe des gekuppelten Addukts **10** sollte im nächsten Schritt zum Amin **13** reduziert werden um dann acetyliert zu werden. Die Reduktion sollte analog zu *Takamura-Enya et al.* mit Zink in einem 500 mM Ammoniumacetat-Puffersystem bei pH = 4,5 durchgeführt werden.^[257] Allerdings konnte diese Reaktion nicht reproduziert werden. Auch eine Umsetzung mit Palladium auf Kohle führte zu keiner Reaktion. Die Umsetzung mit 5 Äquivalenten Natriumdithionit in THF/H₂O führte schließlich zum gewünschten Produkt **13**. Bei der Reaktion kam es zu einem Farbumschlag der Lösung von gelb zu farblos, welcher auf die Umwandlung des stark elektronenziehenden Substituenten NO₂ am aromatischen Ring in einen elektronenschiebenden Substituenten NH₂ zurückzuführen ist. Die Reduktion der Nitrogruppe zur Aminogruppe verläuft über die Bildung der Nitrosoverbindung, welche weiter zum Hydroxylamin und schließlich zum Amin reduziert wird (Schema 10).



Schema 10: Reduktion der Nitrogruppe zum Amin über Bildung der Nitrosoverbindung und des Hydroxylamins.

In Wasser reagiert Natriumdithionit zu zwei Schwefeldioxid-Molekülen, wobei zwei Elektronen für die Reduktion der Nitrogruppe freigesetzt werden. Der Schwefel wird hierbei von +III zu +IV oxidiert. Verbindung **13** konnte auf diese Weise mit 80 % Ausbeute erhalten werden. Das breite Singulett bei $\delta = 4,90$ ppm im NMR-Spektrum konnte der NH₂-Gruppe zugeordnet werden.

Verbindung **13** verfügt über ein primäres Amin am aromatischen Rest, ein sekundäres Amin in N^{\sharp} -Position und über die zwei ungeschützten OH-Gruppen der Ribose. Die Acetylierung sollte selektiv am primären Amin stattfinden, welches die nukleophilste Gruppe sein sollte. Die Acetylierung mit Essigsäureanhydrid in DMF/Pyridin führte auch mit einem nur minimalen Überschuss von 1,1 Äquivalenten zu mono- und diacetylierten Spezies. Außerdem war weiterhin Edukt **13** im Reaktionsgemisch enthalten. Deshalb wurden weitere 4,0 Äquivalente Essigsäureanhydrid zugegeben, was zu ausschließlich dreifach acetyliertem Produkt führte. Auf Grund der Polarität der Verbindung war eine wässrige Aufarbeitung mit Extraktion in ein organisches Lösungsmittel nicht möglich. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand in konzentriertem methanolischem Ammoniak suspendiert. Die monoacetylierte Verbindung **14** konnte auf diese Weise lediglich mit einer Ausbeute von 38 % erhalten werden. Deshalb wurden milder Acetylierungsbedingungen getestet. Nach Mali *et al.* ist bei Einsatz von Thioessigsäure eine selektive Acylierung von primären Aminen möglich, auch wenn im Molekül andere funktionelle Gruppen vorhanden sind.^[258] Diese Reaktion verläuft kupferkatalysiert ab. Als katalytischer Komplex wurde (CuSO₄ • 5H₂O) verwendet. Es wird vermutet, dass das unlösliche *in situ* generierte Nebenprodukt CuS für die katalytische Wirkung verantwortlich ist (Schema 11).^[259]

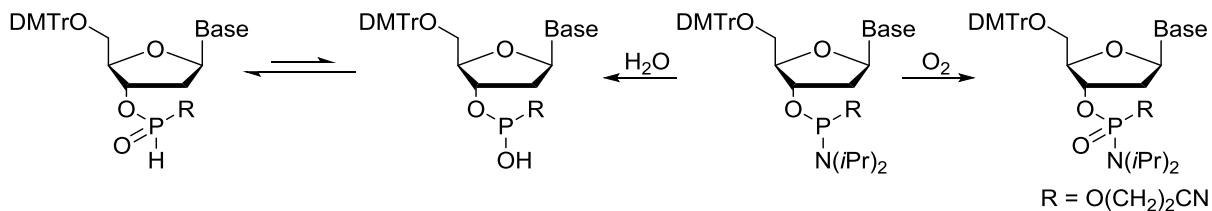


Schema 11: Vorgeschlagener Mechanismus für die von CuSO₄ • 5H₂O hervorgerufene und von CuS katalysierte Amidbindungsbildung. Abbildung adaptiert von Mali *et al.*.^[259]

Es wird angenommen, dass die Reaktion durch $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ initiiert wird. Dieses bildet mit Thioessigsäure bereits in geringen Mengen das gewünschte Produkt und Kupfer(II)-sulfid, welches als Katalysator für die Produktsynthese fungiert. Mit dieser Methode konnte das monoacetylierte N^2 -Phenyladdukt **14** in 75 %iger Ausbeute hergestellt werden. Im NMR-Spektrum konnte das Singulett bei $\delta = 2.05$ ppm mit einem Integral von 3 eindeutig der Acetylgruppe zugeordnet werden.

In den letzten beiden Syntheseschritten sollte die 2'-Desoxyriboseeinheit so modifiziert werden, dass der Baustein für die DNA-Synthese verwendet werden kann. Für diese muss das 5'-OH mit einer 4,4'-Dimethoxytritylgruppe (DMTr) geschützt und am 3'-OH ein Phosphoramidit eingeführt werden. Während der DNA-Synthese wird die DMTr-Schutzgruppe selektiv abgespalten, wobei die Menge an abgespaltenem DMTr als Maß für die Kupplungseffizienz dient. Das Phosphoramidit dient als Grundlage für die Phosphordiesterbrückenbildung und somit der Vernetzung der DNA-Basen. In 5'-Position befindet sich ein primärer Alkohol, welcher reaktiver als der sekundäre Alkohol in 3'-Position ist. Aus diesem Grund wird zur Synthese der benötigten Bausteine zunächst der 5'-Alkohol mit einer DMTr-Schutzgruppe geschützt, während die Phosphoramiditherstellung im letzten Schritt erfolgt. Der Angriff des sterisch anspruchsvollen DMTr-Chlorids erfolgt zu ca. 99 % am 5'-OH. Die Reaktion wurde mit einem leichten Überschuss (1,1 Äq.) an DMTr-Chlorid durchgeführt um die Entstehung von doppeltgeschütztem Produkt zu vermeiden. Als Lösungsmittel wird standardmäßig Pyridin verwendet, da eine Base zum Auffangen der entstehenden HCl benötigt wird und zusätzlich wurde Triethylamin zugegeben. Da DMTr-Chlorid hydrolyseempfindlich ist, wurde unter Schutzgas mit trockenen Lösungsmitteln und Reagenzien gearbeitet. Verbindung **15** konnte zwar nur mit einer Ausbeute von 39 % erhalten werden, aber nicht umgesetztes Edukt **14** konnte nach und nach zu der gewünschten Verbindung umgesetzt werden.

Im letzten Schritt wurde das Phosphoramidit des N^2 -Phenyladduktes **7** hergestellt. Phosphoramidite sind extrem empfindlich, da der Phosphor(III) durch den Luftsauerstoff leicht zu Phosphor(V) und damit zu einem Phosphoramidat oxidiert werden kann. Außerdem reagieren Phosphoramidite leicht mit Nukleophilen. Das ist einerseits wichtig für die DNA-Synthese, andererseits bedeutet es, dass kein Wasser im Lösungsmittel enthalten sein darf, da es sonst zu einer unerwünschten Nebenreaktion kommt. Diese führt schlussendlich auch zu einer oxidierten Phosphorspezies (Schema 12).

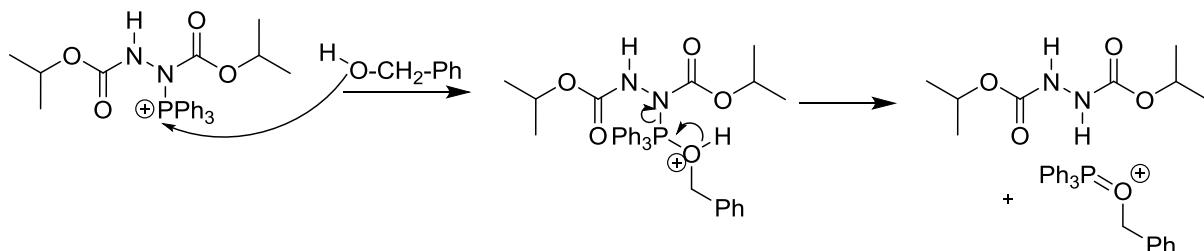


Schema 12: Reaktionen von Phosphoramiditen mit Wasser und Luftsauerstoff.

Die oxidierten Spezies können nicht mehr für die DNA-Synthese verwendet werden, weshalb bei der Synthese und danach darauf geachtet werden muss sauerstoff- und wasserfrei zu arbeiten. Für die Reaktion wurde Phosphoramidierungsreagenz **16** mit Di-*iso*-propylammoniumtetrazolat (DIAT) als Katalysator verwendet. Die Synthese von Phosphoramidit **7** verlief quantitativ. DIAT wurde aus Tetrazol und Di-*iso*-propylamin in Acetonitril hergestellt. Hierzu wurde das Lösungsmittel *in vacuo* entfernt, der Rückstand im Hochvakuum getrocknet und das Salz konnte ohne weitere Aufarbeitung verwendet werden.

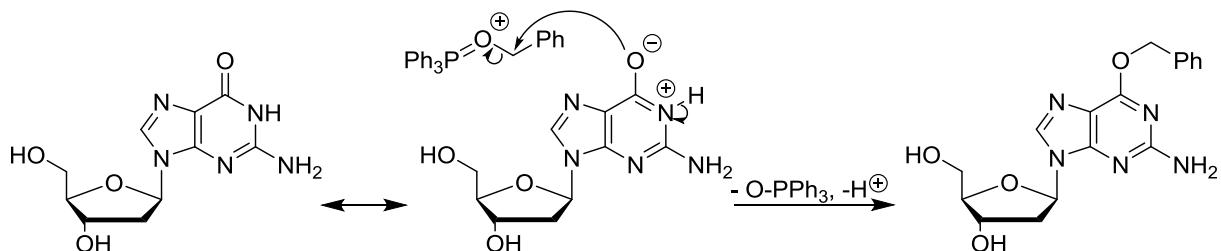
3.1.3.2 Synthese des N²-AAF-dG Phosphoramidits

Die Synthese des Phosphoramidits von *N*²-AAF-dG **9** wurde bereits 2006 von Bonala *et al.* veröffentlicht.^[64] Der Schlüsselschritt – die Buchwald-Hartwig-Kupplung – wird an dG, welches in O^{3'}- und O^{5'}-Position TBS- und in O⁶-Position benzylgeschützt ist, durchgeführt. Es sollte getestet werden, ob eine direkte Kupplung am ungeschützten 2'-Desoxyguanosin möglich ist. Dies würde die Synthese deutlich vereinfachen, da weniger Schützungs- und Entschützungsschritte notwendig wären. Vor allem die Benzylschützung ist kompliziert und könnte durch die alternative Strategie umgangen werden. Bei der Reaktion am O⁶-Sauerstoff handelt es sich um eine *Mitsunobu*-Reaktion, welche die Synthese von aromatischen Ethern aus primären Alkoholen ermöglicht. Es kommt zuerst zur Adduktbildung zwischen Di-*iso*-propylazodicarboxylat (DIAD) und Triphenylphosphin. Das Sauerstoffatom des Benzylalkohols greift nun nukleophil das positiv geladene Phosphorzentrum an. Unter Abspaltung eines Hydrazinderivats wird der Phosphor(III) zu Phosphor(V) oxidiert (Schema 13).



Schema 13: Angriff des Alkohols am Phosphorzentrum im Zuge der *Mitsunobu*-Reaktion.

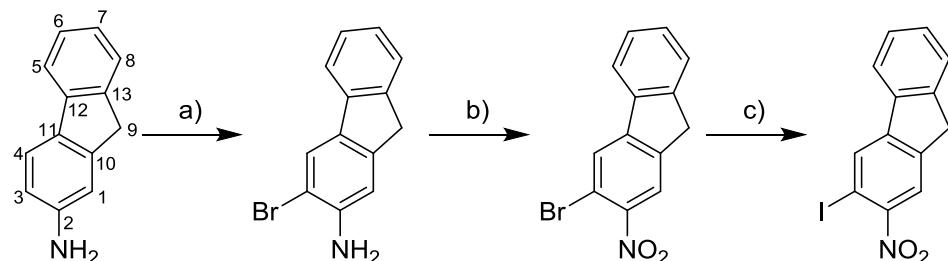
Das Addukt aus Phosphorspezies und Benzylalkohol wird daraufhin am benzyischen C-Atom gemäß einer S_N2-Reaktion vom Amidsauerstoff des Desoxyguanosins angegriffen. Triphenylphosphinoxid wird dabei als gute Abgangsgruppe abgespalten (Schema 14).



Schema 14: Einführung der Benzylschutzgruppe unter Abspaltung von Triphenylphosphinoxid.

Bei dieser Reaktion entstehen viele Nebenprodukte, die auch durch Säulenchromatografie schwer bzw. teilweise nicht abtrennbar sind.

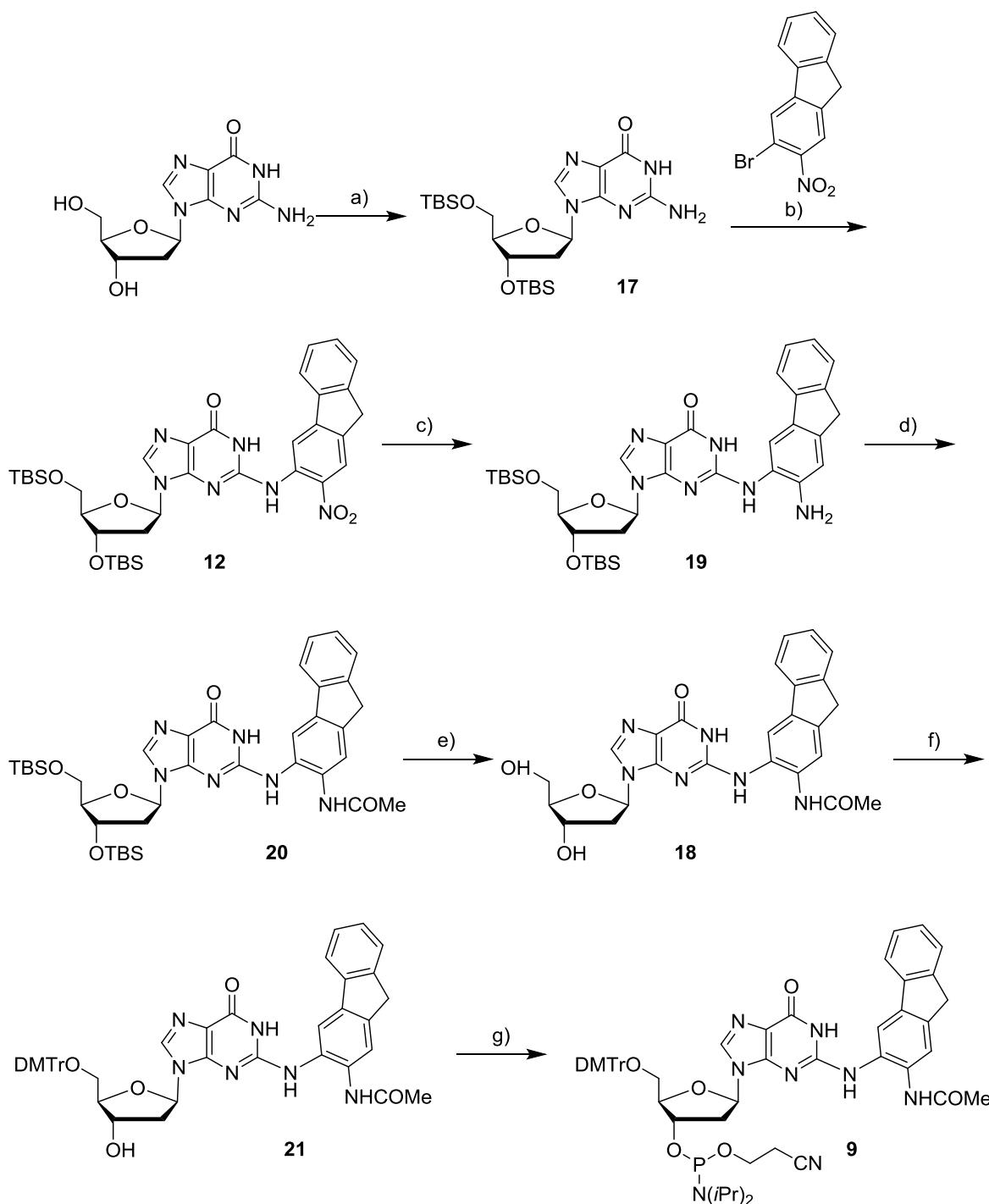
Analog zum N²-Phenyladdukt **7** sollte das N²-Fluorenyladdukt über eine Buchwald-Hartwig-Kupplung zwischen 2'-Desoxyguanosin und 3-Iod-2-nitrofluoren nach *Takamura-Enya et al.* hergestellt werden.^[257] Die aromatische Kupplungskomponente 3-Iod-2-nitrofluoren konnte über eine dreistufige Synthese hergestellt werden (Schema 15).



Schema 15: Synthese von 3-Iod-2-nitrofluoren. a) NBS, DCM, 1 min, 0 °C, 88 % b) *m*CPBA, DCM, 72 h, RT, 86 % c) 1,3-Diaminopropan, Cul, NaI, *n*-butanol, 24 h, 130 °C, 29 %.

Aufgrund der geringen Elektrophilie des Iods ist eine direkte Iodinierung von Aromaten meist problematisch.^[260] Eine gebräuchliche Methode ist hingegen die Nickel- bzw. Kupferkatalysierte Substitution der leichter zugänglichen Bromarene, also eine *Finkelstein*-Reaktion am Aromaten. Dabei zeigte sich, dass die Verwendung von Diaminliganden eine Erhöhung der Ausbeute bewirken und milder Reaktionsbedingungen ermöglichen kann.^[261] Unter Anwendung dieser Methode sollte 3-Iod-2-nitrofluoren ausgehend von 2-Aminofluoren synthetisiert werden. Zunächst wurde 2-Aminofluoren mit *N*-Bromsuccinimid am C3 bromiert. Die elektronenschiebende Aminogruppe ist *ortho*- und *para*-dirigierend. Da die *para*-Position im 2-Aminofluoren nicht zugänglich ist, verläuft die Substitution selektiv in *ortho*. Eine Reaktion am C3-Kohlenstoff sollte hierbei dem C1-Kohlenstoff gegenüber bevorzugt sein –

einerseits durch sterische Effekte und andererseits durch elektronische Effekte des Alkylsubstituenten in *para*. Mittels NMR-Spektroskopie konnte nachgewiesen werden, dass die Substitution am C3 stattgefunden hat. Im Bereich der aromatischen Protonen zwischen 6,50 und 8,00 ppm sind sechs Signale sichtbar. Die Signale bei $\delta = 7,66, 7,42, 7,24$ und $7,13$ weisen das typische Kopplungsmuster des nicht substituierten aromatischen Rings auf. Das Signal bei $\delta = 7,85$ ist ein Singulett, welches dem Proton am C4 zugeordnet wurde. Das verbleibende Signal bei $\delta = 6,97$ ist ein Triplet. Es wurde dem Proton an Position 1 zugeordnet, welches mit den beiden Protonen der CH₂-Gruppe in Position 9 koppeln kann. Dies wurde durch das COSY-Spektrum bestätigt. 2-Amino-3-bromfluoren konnte mit einer Ausbeute von 88 % erhalten werden. Das Produkt formte sich direkt nach dem Zutropfen der NBS-Lösung in DCM, sodass die Reaktion sofort aufgearbeitet werden konnte. Alternativ wurde die Bromierung mit HBr getestet, allerdings waren die Ausbeuten schlechter (ca. 50 %) und die Reaktionszeit deutlich länger. Das so erhaltene, bromierte AF wurde im zweiten Schritt mit *meta*-Chlorperbenzoësäure zum 3-Brom-2-nitrofluoren oxidiert. Die Reaktion musste über mehrere Tage röhren, um eine vollständige Oxidation zur Nitrogruppe über das Hydroxylamin und die Nitrosoverbindung zu gewährleisten. Die Ausbeute betrug 86 %. Zuletzt wurde der Halogenaustausch am C3 durchgeführt. Als Iodquelle wurde Natriumiodid verwendet. 3-Iod-2-nitrofluoren konnte mit einer Ausbeute von 29 % erhalten werden und wurde für die Buchwald-Hartwig-Reaktion mit dG nach *Takamura-Enya et al.* verwendet.^[257] Allerdings war die direkte Kupplung am ungeschützten 2'-Desoxyguanosin auf diesem Weg nicht möglich. Deshalb wurde dieses am O^{3'}- und O^{5'}-Sauerstoff TBS-geschützt (Schema 16). Die Reaktion mit TBS-Chlorid in Anwesenheit von Imidazol verläuft quantitativ. Die Aufarbeitung ist wenig zeitintensiv, da das Produkt in DMF bereits schlecht löslich ist und aus kaltem Ethanol gefällt werden kann. Ausgehend von der TBS-geschützten Verbindung **17** konnte Phosphoramidit **9** nach Schema 16 synthetisiert werden.



Schema 16: Synthese von N^2 -AAF-dG. a) TBS-Cl, Imidazol, DMF, 7 h, RT, quant. b) Pd₂(dba)₃, Xantphos, Cs₂CO₃, Toluol, 16 h, 100 °C, 57 % c) Pd/C, H₂, EtOAc/MeOH 1:1, 16 h, RT, 49 % d) 1) Ac₂O, DCM, 16 h, RT 2) NH₃ in MeOH, 73 % e) TBAF, THF, 6 h, RT, 50 % f) DMTr-Cl, NEt₃, Pyridin, 16 h, RT g) DIAT, 16, DCM, 1,5 h, RT, quant.

Für die Buchwald-Hartwig-Kupplung wurde 3-Brom-2-nitrofluoren verwendet, welches in besseren Ausbeuten als 3-Iod-2-nitrofluoren zugänglich ist. Kupplungsaddukt **12** konnte in 57 % Ausbeute erhalten werden.

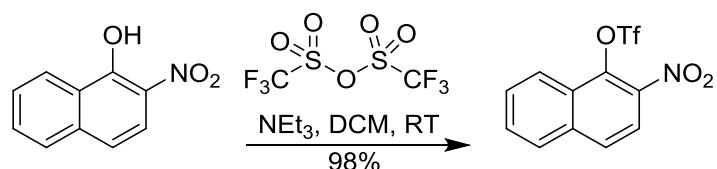
Bonala et al. behaupten, dass der klassische Syntheseweg – also zunächst die Modifizierung des aromatischen Schadens und dann die der Riboseeinheit – bei diesem Addukt nicht möglich wäre.^[64] Diese Annahme basiert auf der Tatsache, dass die TBS-entschützte Verbindung **18** sehr unlöslich in organischen Lösungsmitteln ist, sodass die Schützung mit DMTr-Chlorid nicht durchführbar wäre. Bereits die DMTr-Schützung zur Verbindung **15** als Zwischenstufe der Synthese des *N*²-Phenyladdukts wurde ausgehend von der in Pyridin unlöslichen Verbindung **14** in Suspension durchgeführt. Deshalb sollte die Behauptung von *Bonala et al.* überprüft werden. Die DMTr-Schutzgruppe gehört zu den labileren Schutzgruppen, sodass ihre Einführung vorzugsweise in einem späteren Syntheseschritt stattfinden sollte.

Die Synthese sollte nach dem Schema des *N*²-Phenyladdukts verlaufen. Die Reduktion der Nitrogruppe zum Amin wurde mit Palladium auf Kohle unter H₂-Atmosphäre, mit Eisen, mit Zink und mit Natriumdithionit getestet. Lediglich die Reduktionen mit Pd/C und mit Zink führten zur gewünschten Verbindung **19**, wobei mit Zink nur Spuren des Produkts erhalten werden konnten. Durch Reduktion mit Palladium auf Kohle konnte Amin **19** mit einer Ausbeute von 49 % erhalten werden. Im nächsten Schritt sollte Amin **19** zu Acetylamin **20** umgewandelt werden. Die Kupfer-katalysierte Acetylierung mit Thioessigsäure, die für das *N*²-Phenyladdukt angewendet wurde, führte in diesem Fall zu einer Zersetzung und nicht zum gewünschten Produkt. Die Acetylierung mit Essigsäureanhydrid war hingegen erfolgreich, wobei es auch in diesem Fall zur Acetylierung beider freier Aminogruppen kam. Das sekundäre *N*²-Amin konnte mit methanolischem Ammoniak wieder entschützt werden. Acetylamin **20** wurde mit einer Ausbeute von 73 % erhalten. Die TBS-Schutzgruppen der Alkoholfunktionen der Riboseeinheit konnten mit 1 M TBAF in THF entfernt werden und Nukleosid **18** wurde mit 50 % Ausbeute erhalten.

Im nächsten Schritt sollte der O⁵-Sauerstoff mit DMTr-Chlorid für die DNA-Synthese geschützt werden. Nukleosid **18** konnte, wie von *Bonala et al.* beschrieben, nicht in Pyridin gelöst werden und wurde demnach in trockenem Pyridin mit Molsieb (3 Å) suspendiert und zwei Stunden gerührt, um Restfeuchtigkeit zu entfernen. Trockenes Triethylamin und 1,2 Äquivalente DMTr-Chlorid wurden zugegeben und das Reaktionsgemisch 16 Stunden gerührt. Mit dieser Methode war es zwar nicht möglich das Edukt komplett zur Verbindung **21** umzusetzen, aber nicht reagiertes Edukt **18** konnte mittels Säulenchromatografie zurückgewonnen werden. Verbindung **21** wurde auf Grund seiner Instabilität direkt zu Phosphoramidit **9** umgesetzt. Es konnte gezeigt werden, dass der klassische Syntheseweg auch im Fall des *N*²-Fluorenyladdukts möglich ist.

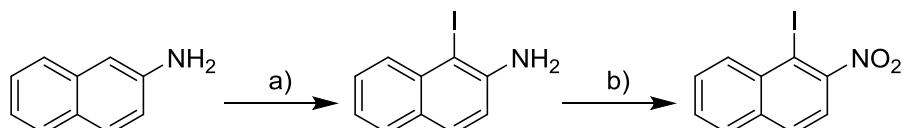
3.1.3.3 Synthese des N²-AAN-dG Phosphoramidits

Um die Anzahl an Schützungs- und Entschützungsschritten zu reduzieren, sollte auch Phosphoramidit **8** nach *Takamura-Enya et al.* ausgehend von ungeschütztem dG synthetisiert werden.^[257] Zunächst wurde die Kupplung zwischen 2'-Desoxyguanosin und 2-Nitronaphthalin-1-triflat getestet. Die aromatische Kupplungskomponente ist in einem Schritt aus kommerziell erhältlichem 2-Nitronaphthol mit hoher Ausbeute herstellbar und benötigt keine säulenchromatografische Aufreinigung (Schema 17).^[262]



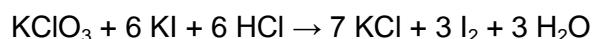
Schema 17: Synthese von 2-Nitronaphthalin-1-triflat.

Allerdings konnte das gewünschte Kupplungsprodukt auf diesem Wege nicht erhalten werden, da die Triflatverbindung nicht reaktiv genug war. Deshalb wurde die entsprechende reaktivere Iodverbindung hergestellt (Schema 18).



Schema 18: Synthese von 1-Iodo-2-nitronaphthalin. a) KI , KClO_3 , HCl , MeCN , $80\text{ }^\circ\text{C}$, 3 Tage, 89 % b) $m\text{CPBA}$, DCM , RT , 18 %.

Diese konnte ausgehend vom kommerziell erhältlichem 2-Aminonaphthalin in zwei Stufen erhalten werden. Im ersten Schritt wurde Iod am C1 durch eine säurekatalysierte Reaktion mit Kaliumchlorat als Oxidationsmittel und Kaliumiodid als Iodquelle eingeführt. Das Iodid mit Oxidationsstufe -1 wird hierbei zu elementarem Iod(0) oxidiert, während das Chlor(V) des Chlorats zum Chlorid(-I) reduziert wird:



Auf Grund des elektronenschiebenden Effekts der Aminogruppe finden Zweitsubstitutionen am aromatischen Ring in *ortho*- und *para*-Position statt. Im 2-AN ist die *para*-Position nicht zugänglich, sodass die Substitution in *ortho* stattfindet. Hierbei ist die Position 1 reaktiver als die Position 3. Über NMR-Spektroskopie konnte nachgewiesen werden, dass die Substitution am C1-Kohlenstoff stattgefunden hat. Die Signale bei $\delta = 7,73$, $7,62$, $7,40$ und

7,15 ppm wurden dem unsubstituierten Naphthyrling zugeordnet. Das Duplett bei $\delta = 7,62$ ppm koppelt mit einer Kopplungskonstante von $J = 8,6$ Hz mit dem Duplett bei $\delta = 7,07$ ppm. Die beiden Signale wurden respektive dem Proton am C4 und am C3 zugeordnet. Hätte die Substitution am C3 stattgefunden, so wären im NMR-Spektrum zwei Singuletsignale sichtbar gewesen. Das COSY-Spektrum konnte diese Zuordnung bestätigen. 1-Amino-2-iodnaphthalin konnte mit einer Ausbeute von 89 % hergestellt werden. Im nächsten Schritt wurde die Aminogruppe mit *m*CPBA zur Nitrogruppe oxidiert. Beide Stufen mussten mittels Säulenchromatografie aufgereinigt werden und die Ausbeute der Oxidation zur Nitrogruppe betrug nur 18 %. Die TBS-Schützung von 2'-Desoxyguanosin hingegen benötigt keine chromatografische Aufreinigung und verläuft quantitativ. Deswegen wurde dieser Weg nicht weiterverfolgt. Die Kupplung von TBS-geschütztem 2'-Desoxyguanosin **17** mit 2-Nitronaphthalin-1-triflat zu Verbindung **11** konnte erfolgreich mit einer Ausbeute von 97 % durchgeführt werden.

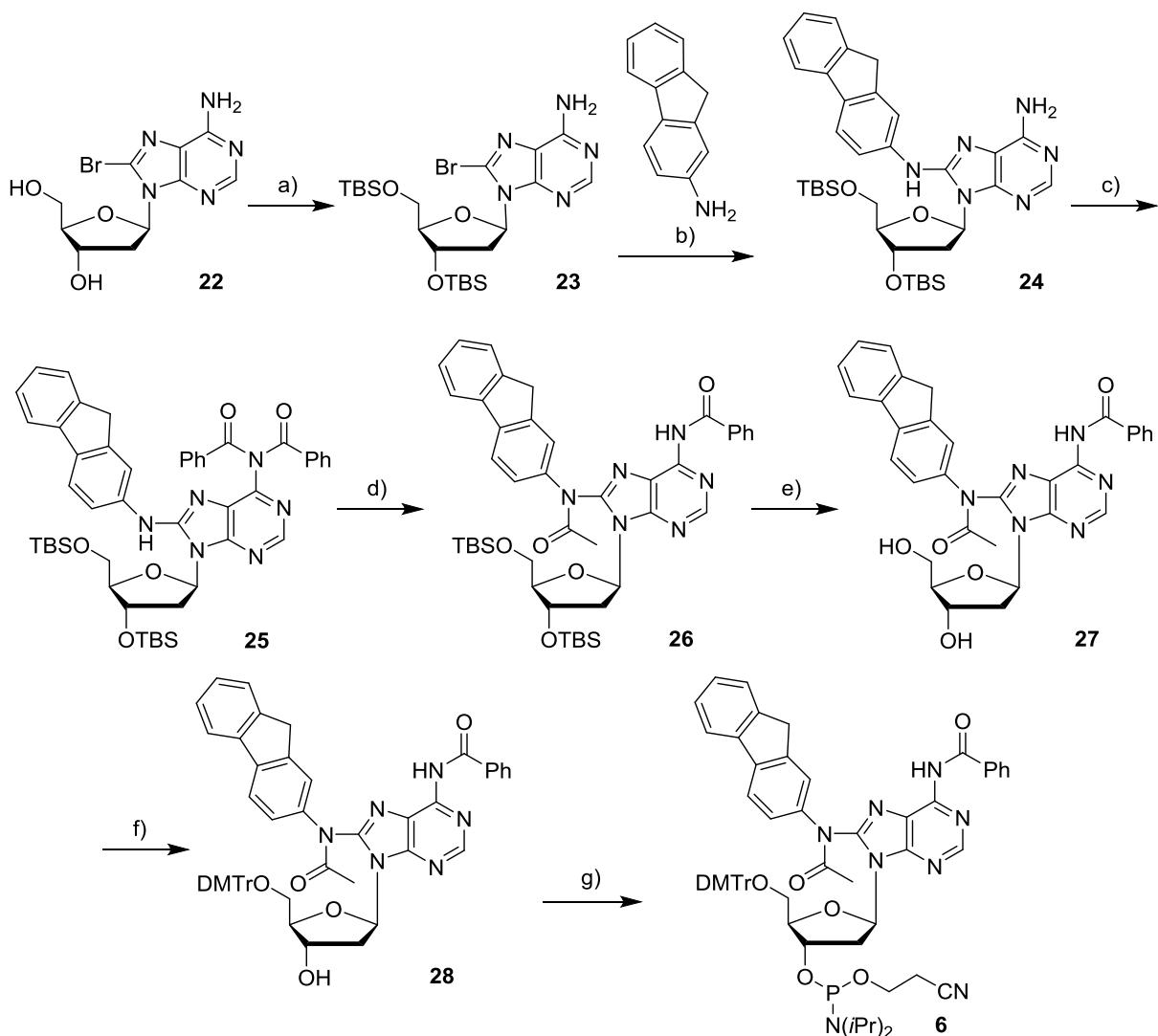
Analog zu den anderen N^2 -Addukten sollte zunächst der aromatische Rest modifiziert werden und anschließend die Ribose. Als erstes sollte die Nitrogruppe am Naphthalinderivat in N^2 -Position zum Amin reduziert werden. Es wurden verschiedene Reduktionsmittel getestet: $\text{Na}_2\text{S}_2\text{O}_4$, Pd/C unter H_2 -Atmosphäre, Zink, Eisen und Zinnchlorid. Mit keinem dieser Reaktionsmittel konnte das Amin erhalten werden. Aus diesem Grund wurde die Synthese von Phosphoramidit **8** nach *Bonala et al.* durchgeführt.^[64]

3.1.3.4 Synthese des C8-AAF-dA Phosphoramidits

Bei den Addukten von Desoxyadenosin und Desoxyguanosin, welche am C8 modifiziert sind, ist der Schlüsselschritt der Synthese ebenfalls eine *Buchwald-Hartwig*-Kupplung. In diesem Fall muss jedoch nicht der Aromat, sondern das Nukleosid mit einem Halogenid funktionalisiert werden. Als Kupplungspartner können für die Synthese dieser Addukte also aromatische Amine verwendet werden, welche in der Regel kommerziell erhältlich sind.

Das C8-AAF-dA Phosphoramidit sollte nach *Szombati et al.* synthetisiert werden.^[263] Die Autoren berichten von der Synthese von Verbindung **6**, konnten aber auch auf Nachfrage keinerlei analytische Charakterisierung des Moleküls zur Verfügung stellen. Phosphoramidit **6** und alle Zwischenprodukte der Synthese wurden aus diesem Grund in dieser Arbeit massenspektrometrisch und NMR-spektroskopisch charakterisiert.

Zunächst wurde das in C8-Position bromierte 2'-Desoxyadenosin **22** mit TBS-Chlorid und Imidazol am $O^{3'}$ - und $O^{5'}$ -Sauerstoff TBS-geschützt (Schema 19). Verbindung **23** konnte mit 82 % Ausbeute erhalten werden.



Schema 19: Synthese von C8-AAF-dA. a) TBS-Cl, Imidazol, Pyridin, 16 h, RT, 82 % b) Pd(OAc)₂, BINAP, Cs₂CO₃, Toluol, 24 h, RT, 52 % c) Benzoylchlorid, Pyridin, 16 h, RT, 93 % d) 4-DMAP, Ac₂O, Pyridin, 16 h, RT, 51 % e) TBAF, AcOH, THF, 48 h, RT, 53 % f) DMT-Cl, NEt₃, Pyridin, 18 h, RT, 62 % g) DIAT, **16**, DCM, 2,5 h, RT, 99 %.

Die *Buchwald-Hartwig-Kupplung* zwischen Verbindung **23** und 2-Aminofluoren wurde mit Palladiumacetat und 2,2'-Bis(diphenylphosphino)-1,1'-binaphthyl (BINAP) als katalytischem Komplex durchgeführt. Als Base wurde Cs₂CO₃ verwendet und die Reaktion fand in Toluol statt. Addukt **24** wurde mit 52 % Ausbeute erhalten. Für die DNA-Synthese musste das primäre *N*⁶-Amin benzoylgeschützt werden. Diese Schutzgruppe wird auch für das Standard dA-Phosphoramidit verwendet. Als Benzoylquelle wurde Benzoylchlorid im Überschuss verwendet. Dies führte zur Schützung mit zwei Benzoylschutzgruppen am *N*⁶-Stickstoff. Verbindung **25** wurde mit einer Ausbeute von 93 % erhalten. Im nächsten Schritt wurde der aromatische Rest modifiziert, wobei auch die überflüssige Benzoylschutzgruppe am *N*⁶-Stickstoff wieder entfernt wurde. Acetylamin **26** konnte mit Essigsäureanhydrid als Acetylierungsreagenz in 51 % Ausbeute erhalten werden. Zuletzt wurde die Zuckereinheit

analog zu den N^2 -Addukten modifiziert. Die TBS-Schutzgruppen wurden mit TBAF als Fluoridquelle abgespalten (53 %, Verbindung **27**) und der $O^{5'}$ -Sauerstoff wurde mit DMTr-Chlorid geschützt. Verbindung **28** wurde mit 62 % Ausbeute erhalten und nicht-umgesetztes Edukt **27** konnte *via* Säulenchromatografie zurückgewonnen werden. Im letzten Schritt wurde das Phosphoramidit am $O^{3'}$ -Sauerstoff mit 99 % Ausbeute eingeführt (Verbindung **6**).

3.1.3.5 DNA-Synthese

Für die Kristallisationsexperimente sollten die modifizierten Phosphoramidite in 15mer Oligonukleotide eingebaut werden.

Zusätzlich wurden verschiedene C8-dG *bulky* Addukte synthetisiert (Abbildung 6). Die Synthese der DMTr-Stufen dieser Addukte wurde von *Nina Simon* durchgeführt. Die Phosphoramiditsynthese und die DNA-Synthese, sowie alle weiteren Versuche mit der fertigen DNA wurden gemeinsam durchgeführt.

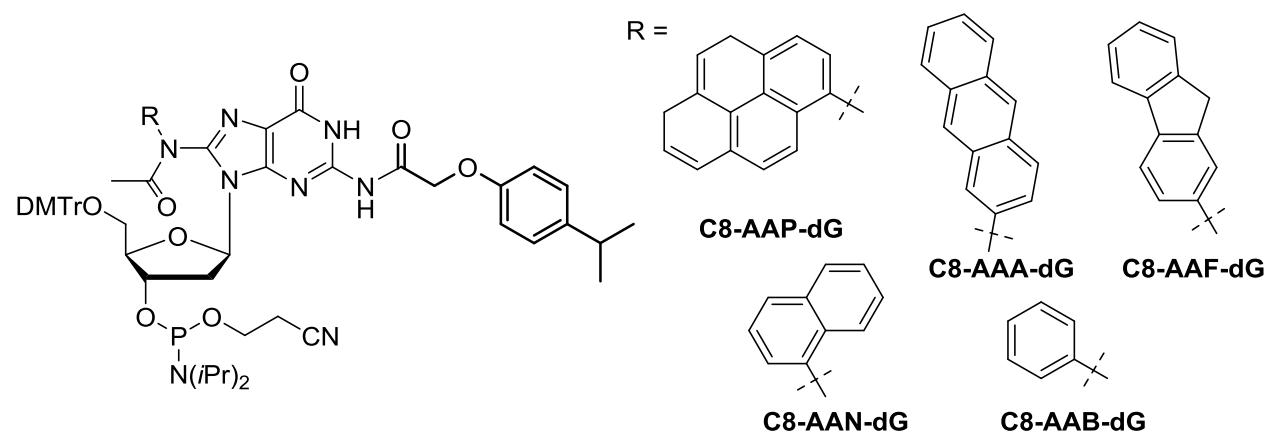
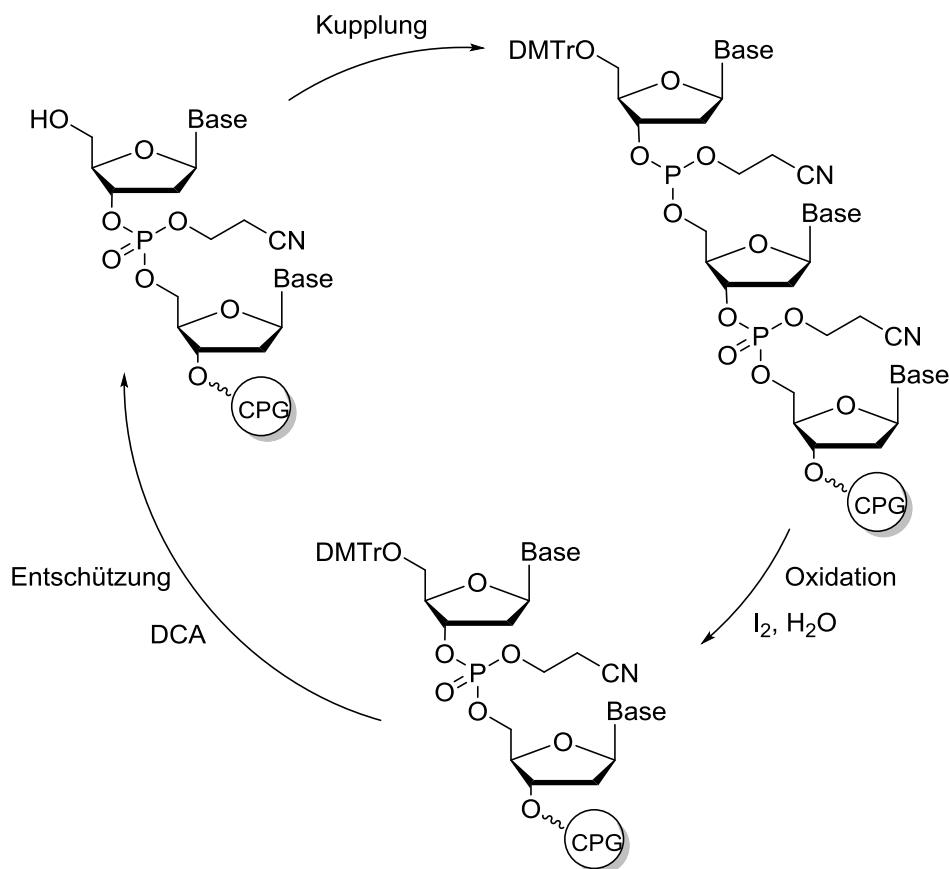


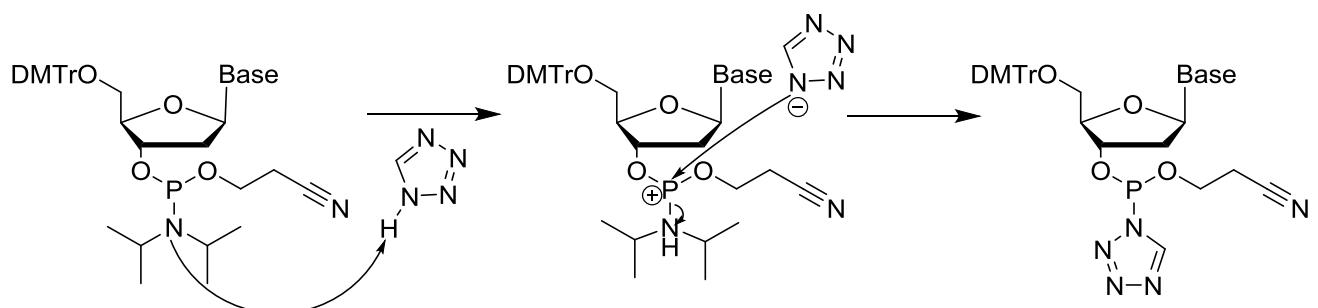
Abbildung 6: Phosphoramidite der C8-dG-Addukte.

Mit der automatisierten DNA-Festphasensynthese lassen sich Oligodesoxynukleotide mit einer spezifischen Basensequenz herstellen. Die Synthese erfolgt während die wachsende Oligonukleotidkette kovalent an einen festen polymeren Träger aus *Controlled Pore Glass* (CPG) oder Polystyrol (*Polystyrene Beads*) gebunden ist. Hierbei ist es möglich modifizierte Nukleotide an einer bestimmten Position des Stranges einzusetzen. Die Synthese erfolgt im Gegensatz zur natürlichen Syntheserichtung von 3' nach 5'. Die Phosphoramidit-Methode an der Festphase wurde von *Caruthers* entwickelt (Schema 20).^[264]



Schema 20: Reaktionszyklus der DNA-Festphasensynthese nach der Phosphoramidit-Methode.

Das erste Nukleosid der Sequenz ist über die 3'-Hydroxygruppe mit einem basenlabilen Ester als Linker an die Festphase gebunden und trägt am 5'-Ende eine DMTr-Schutzgruppe. Diese ist säurelabil und kann im richtigen Moment gezielt mit 3 %-iger Dichloressigsäure abgespalten werden, sodass ein weiteres Nukleosid selektiv an den Strang gekuppelt werden kann. Die freie, reaktive 5'-OH-Funktion kann nun nukleophil das Phosphor(III)-Atom am nächsten Nukleosids angreifen. Dieses wird vorher mit Tetrazol oder speziellen Derivaten aktiviert (Schema 21).



Schema 21: Aktivierung eines 5'-DMTr geschützten Nukleosid-3'-phosphoramidits mit Tetrazol.

Zunächst wird die Di-*iso*-Propylaminogruppe des Phosphoramidits protoniert. Das Tetrazolidion greift unter Abspaltung von Di-*iso*-Propylamin nukleophil das Phosphoratom an. Das aktivierte Phosphorzentrum kann nun mit dem nächsten Nukleotid reagieren.

Da die Synthese nicht mit einer 100 %-igen Ausbeute abläuft, müssen nach der Kupplung alle freien OH-Gruppen, welche nicht reagiert haben, deaktiviert werden. Dies geschieht durch Acylierung mit Essigsäureanhydrid (Standard DNA-Synthese) oder Phenoxyessigsäureanhydrid (Pac_2O , *ultramild*® DNA-Synthese) und wird *Capping* genannt. Ohne dieses Verfahren könnten diese OH-Gruppen im nächsten Synthesezyklus wieder reagieren, was zu verschiedenen Oligonukleotidketten unterschiedlicher Sequenz und ähnlicher Länge führen und die Aufreinigung erheblich erschweren würde. In den C8-dG-Addukten ist die N^2 -Position mit einer *iPrPac*-Schutzgruppe versehen. Essigsäureanhydrid würde diese Schutzgruppe durch eine Acetylgruppe ersetzen, die unter den milden Abspaltungsbedingungen nach erfolgreicher DNA-Synthese nicht mehr entfernt werden könnte, weshalb stattdessen Pac_2O in Lutidin für das *Capping* verwendet wird.

Nach dem *Capping* liegt ein Phosphit vor, welches nun zum Phosphat oxidiert werden muss. Zur Ausbildung der Phosphordiesterbrückenbildung wird eine wässrige Iodlösung verwendet. Zuletzt wird die DMTr-Schutzgruppe am 5'-OH-Ende abgespalten, damit ein neuer Synthesezyklus gestartet werden kann. Dieser Vorgang wird wiederholt, bis das gewünschte Oligodesoxynukleotid fertiggestellt ist. Die Menge an abgespaltenen Dimethoxytrityl-Kationen wird photometrisch oder über Messung der Leitfähigkeit bestimmt und gilt als Maß für die Kupplungseffizienz, also die Ausbeute eines Reaktionszyklus.

Für die DNA-Synthese müssen die exozyklischen Aminofunktionen als Amide geschützt werden, um eine ausreichende Stabilität bei den für die Oligonukleotidsynthese üblichen pH-Werten zu garantieren. Bei der Standard-DNA-Synthese ist das dC-Phosphoramidit acetylgeschützt, für dG wird eine *iso*-Butyrylschutzgruppe verwendet, während dA mit einer Benzoylschutzgruppe versehen ist. Das Phosphoramidit von dT besitzt keine exozyklischen Aminofunktionen und kann somit ohne Schutzgruppe verwendet werden. Für die *ultramild*® DNA-Synthese wird die primäre Aminfunktion von dA und dG jeweils mit einer *iso*-Propyl-Phenoxyacetylgruppe geschützt. Diese Schutzgruppen können nach erfolgreicher DNA-Synthese unter basischen Bedingungen wieder abgespalten werden. Die Schutzgruppen für die *ultramild*® DNA-Synthese können unter milderen Bedingungen abgespalten werden als die Schutzgruppen der Standard-DNA-Synthese. Welche der beiden Synthesen angewandt wird, hängt von den Sonderbasen ab, welche in die DNA eingebaut werden. Besitzen diese bestimmte empfindliche funktionelle Gruppen, die im fertigen DNA-Strang erhalten bleiben sollen, so werden *ultramild*® Bedingungen angewandt um zu verhindern, dass diese Gruppe bei der Entschützung abgespalten wird.

Die Oligonukleotide wurden auf einem *ABI 394 DNA/RNA Synthesizer (Applied Biosystems)* unter Verwendung der Standard Phosphoramidit Protokolle für die Festphasen DNA Synthese hergestellt. Die modifizierten Phosphoramidite wurden je nach Anteil an oxidiertem Material (ermittelt über ^{31}P -NMR-Spektroskopie) 0,1 M – 0,2 M in wasserfreiem Acetonitril angelöst. Auf Grund der schlechten Löslichkeit wurde das N^2 -AAF-dG Phosphoramidit in einem 1:1 Gemisch aus Acetonitril und DCM gelöst. Das N^2 -AAN-dG Phosphoramidit wurde in reinem DCM suspendiert. Auch bei einer Konzentration von 0,02 M, nach Zentrifugation und Filtern war die Lösung noch trüb und wurde als solche verwendet. Um der niedrigen Konzentration entgegenzuwirken wurde das Phosphoramidit vor dem nächsten Schritt im Synthesezyklus wiederholt auf die Festphase aufgezogen, sodass gute Kupplungseffizienzen erreicht werden konnten. Für alle modifizierten Phosphoramidite wurden Kupplungszeiten von 30 Minuten verwendet.

Unter Standardbedingungen wird der DNA-Strang mit einer Lösung aus Ammoniak und Methylamin im Verhältnis 1:1 von der Festphase abgespalten und entschützt. Die synthetisierten C8- und N^2 -Addukte enthalten jeweils eine Acetylgruppe an einer Aminofunktion, welche unter diesen starkbasischen Bedingungen abgespalten werden könnte. Oligonukleotide, welche diese Addukte enthalten, wurden deshalb unter *ultramild®* Bedingungen synthetisiert und mit 10 % Di-*iso*-propylamin und 0,25 M β -Mercaptoethanol in Methanol entschützt, wobei β -Mercaptoethanol als Schutz vor Oxidation der *bulky* Addukte während der langen Inkubationszeiten dient. Einzig Oligonukleotide, die das C8-AAF-dA Addukt enthielten sollte auf Grund der besonderen Instabilität der Acetylgruppe nach Szombati *et al.* mit 5 % Di-*iso*-propylamin und 0,25 M β -Mercaptoethanol in Methanol entschützt werden. Die Autoren konnten 12mer Oligonukleotide, welche verschiedene acetylierte C8-dA-Schäden enthielten, auf die Weise entschützen. Zum Teil wurden die Schäden dabei auch deacetyliert, aber der gewünschte Strang konnte *via* HPL-Chromatografie vom deacetylierten DNA-Strang abgetrennt werden. Das C8-AAF-dA Addukt wurde jedoch in keinen 12mer DNA-Strang eingebaut, sondern nur in ein 20mer Oligonukleotid. Die Autoren konnten bei der Entschützung der 20mer Oligonukleotide ausschließlich die deacetylierten Stränge erhalten.^[263] Die Ergebnisse von Szombati *et al.* lassen vermuten, dass die Länge oder die Basensequenz des Oligonukleotids entscheidend für die Deacetylierung sind. Für die Bindungs- und Kristallisationsstudien in dieser Arbeit wurden 15mer Oligonukleotide benötigt (Tabelle 2). Unter den oben genannten Entschützungsbedingungen konnte ausschließlich das C8-AF-dA Addukt isoliert werden. Stränge mit dem nicht acetylierten C8-AF-dG Addukt wurden durch Deacetylierung der entsprechenden C8-AAF-dG Stränge erhalten. Hierfür wurde zunächst mit einer 1:1 Lösung aus Ammoniumhydroxid (30 % in Wasser) und Methylamin über 5 Minuten bei 65 °C entschützt. Zur vollständigen Entschützung wurde 1 M NaOH mit 0,25 mM β -Mercaptoethanol

verwendet. Für die Schmelzpunktbestimmung wurden zum Vergleich zusätzlich geschädigte 34mer Oligonukleotide synthetisiert (Tabelle 2). Die ungeschädigten Gegenstränge wurden von *Metabion* bezogen.

Die so erhaltenen Oligonukleotide wurden mittels HPL-Chromatografie gereinigt. Für die 15mer Oligonukleotide, welche C8- oder N^2 -Addukte enthielten, wurde ein Gradient von 0 bis 25 % oder von 0 bis 30 % Puffer B in Puffer A (Kapitel 5.2) in 45 Minuten verwendet. In Abbildung 7 sind stellvertretend für alle synthetisierten Oligonukleotide die HPL-Chromatogramme des 15mer Oligonukleotids mit N^2 -AAN-dG als Schaden vor und nach der Aufreinigung gezeigt.

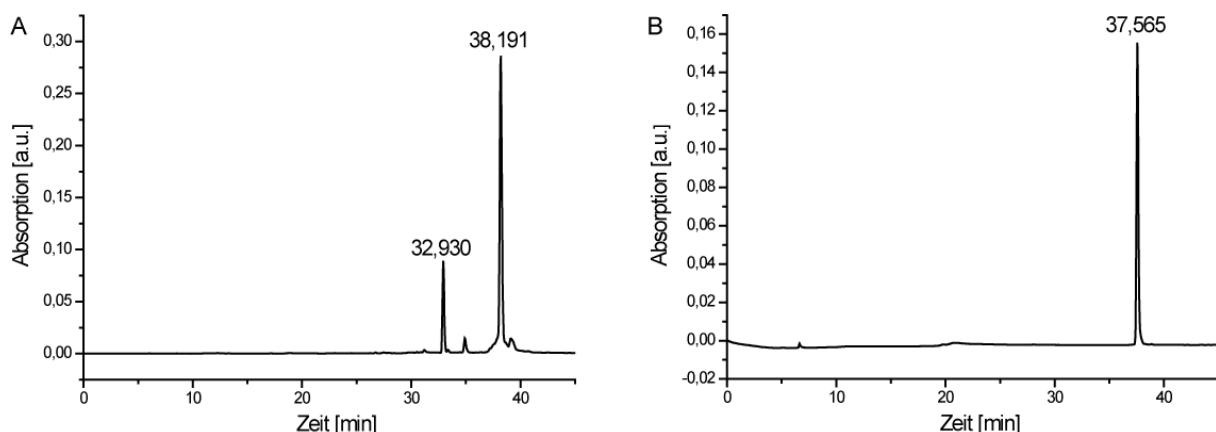


Abbildung 7: HPL-Chromatogramme des 15mer Oligonukleotids mit N^2 -AAN-dG als Schaden (A) vor und (B) nach der Aufreinigung via HPLC.

Für die 34mer Oligonukleotide, welche C8- oder N^2 -Addukte enthielten, wurde ein Gradient von 0 bis 30 % oder von 0 bis 35 % Puffer B in Puffer A (Kapitel 5.2) in 45 Minuten verwendet. Die Synthese der Oligonukleotide wurde mittels MALDI-TOF Massenspektrometrie überprüft. Im Anschluss wurden die Oligonukleotide mittels Sep-Pak® Kartuschen entsalzt.

3.1.4 Schmelzpunktbestimmung

Es ist bekannt, dass *bulky* Addukte in der DNA Auswirkungen auf die Stabilität des Doppelstrangs haben. Ob sie diesen stabilisieren oder destabilisieren, beeinflusst über welchem NER-Reparaturmechanismus diese erkannt und repariert werden können.

Die Strukturen mit Aminofluoren oder Acetylaminofluoren-geschädigter DNA sind bereits gut untersucht (Kapitel 1.1.1),^[50, 61, 265-266] allerdings sind die Konformationen, die diese Schäden annehmen, stark sequenzabhängig, wodurch auch die Reparatureffizienz sehr variiert.^[50, 59, 61, 267] In C8-AAF-dG DNA liegt die geschädigte Base meist in *syn*-Konformation vor. Acetylaminofluoren interkaliert in das dG:dC Basenpaar, sodass Desoxyguanosin in die

große Furche verdrängt wird. Die Helix wird dadurch verformt und es kommt zu einer thermodynamischen Destabilisierung.^[50, 265] Dementsprechend wird dieses Addukt leicht von GG-NER-Enzymen erkannt und entfernt.^[268] C8-AF-dG liegt hauptsächlich in *anti*-Konformation vor, wobei das Basenpaar intakt bleibt, während Aminofluoren sich in der großen Furche befindet. Letzterer Schaden führt zu einer geringeren Destabilisierung als ersterer.

Die Präsenz von *N*²-AAF-dG in einem DNA-Doppelstrang führt hingegen zu dessen Stabilisierung. Acetylaminofluoren liegt hierbei tief in der kleinen Furche und formt einen 40° Winkel mit Desoxyguanosin. Der Duplexstrang passt sich leicht an, um den hydrophoben Aromaten von der wässrigen Umgebung abzuschirmen, wodurch die kleine Furche in unmittelbarer Nähe zum Schaden noch enger wird.^[49] Als Konsequenz ist es kein gutes Substrat für die GG-NER^[191], sondern wird über den TC-NER-Mechanismus repariert.^[246]

Die Schmelzpunkte aller Duplexstränge aus den synthetisierten Oligonukleotiden und ihren unmarkierten Gegensträngen sollten bestimmt werden, um eine mögliche Veränderung im Grad der Stabilisierung oder Destabilisierung innerhalb einer Adduktgruppe nachzuweisen. Zum Vergleich wurden zusätzlich längere 34mer Nukleotide synthetisiert, welche ebenfalls verschiedene *bulky* Addukt Schäden enthielten. Sowohl die 15mer als auch die 34mer Oligonukleotide wurden in äquimolarem Verhältnis mit ihren unmarkierten Gegensträngen gemischt, mit einer Konzentration von 4 µM in Schmelzpunktspuffer (Tabelle 3) gelöst und durch Erhitzen und langsames Abkühlen (1 °C/min) hybridisiert. Für die Bestimmung der Schmelzpunkte wurde die UV Absorption bei 260 nm gemessen, während die 15mer Oligonukleotide mit einem Temperaturgradienten von 1 °C/min von 15 auf 85 °C erhitzt oder von 85 auf 15 °C abgekühlt wurden. Die 34mer Oligonukleotide wurden von 20 auf 95 °C erhitzt und von 95 auf 20 °C abgekühlt. Wird die Temperatur erhöht, so führt das Aufschmelzen der DNA, das heißt der Übergang von einer doppelsträngigen in eine einzelsträngige Struktur, zu einer Erhöhung der Absorption. Die Basenstapelung im Doppelstrang führt zu einem geringeren molaren Absorptionskoeffizienten als die Summe der molaren Absorptionskoeffizienten der einzelnen Nukleobasen. Die so ermittelten Schmelzpunkte sind in Tabelle 1 aufgeführt.

DNA-Schaden	TM [°C] 15mer	TM [°C] 34mer
dG:dC	62	84
C8-AAB-dG:dC	42	77
C8-AAN-dG:dC	46	77
C8-AAA-dG:dC	49	80
C8-AAP-dG:dC	47	80
C8-AAF-dG:dC	48	82
C8-AAF-dG:dG	39	
C8-AF-dG:dC	55	82
<i>N</i> ² -AAF-dG:dC	67	87
<i>N</i> ² -AAN-dG:dC	62	
<i>N</i> ² -AAB-dG:dC	65	
dA:dT	58	
C8-AF-dA:dC	51	

Tabelle 1: Schmelzpunkte der synthetisierten 15mer und 34mer Oligonukleotide.

Bei den acetylierten C8-dG-Addukten im 15mer DNA-Strang ist eine Zunahme der Schmelztemperatur bei Zunahme der π -stacking Oberfläche zu erkennen. Einzig der Strang, welcher das C8-AAP-dG Addukt enthält, weist im Vergleich zu den nächstkleineren π -Systemen C8-AAF-dG (48 °C) und C8-AAA-dG (49 °C) einen etwas geringeren Schmelzpunkt (47 °C) auf. Der Strang mit dem kleinen C8-AAB-dG Addukt ist deutlich destabilisiert (42 °C), während die Schmelzpunkte der anderen Stränge nah beieinander (46 – 49 °C) liegen.

Die 15mer DNA-Doppelstränge, welche die nicht acetylierten C8-Addukte C8-AF-dG und C8-AF-dA enthalten, weisen ebenfalls eine Destabilisierung auf. Diese ist weniger stark ausgeprägt als bei den acetylierten Addukten und der Strang mit dem Desoxyguanosinaddukt schmilzt bei einer höheren Temperatur (55 °C) als das der Strang mit dem dazu analogen Desoxyadenosinaddukt (51 °C). DNA-Stränge mit einem höheren GC-Gehalt haben höhere Schmelzpunkte, da zwischen diesen Basen drei Wasserstoffbrückenbindungen ausgebildet werden und nicht nur zwei wie zwischen dA- und dT-Basen. Auf Grund dessen hat der ungeschädigte Vergleichsstrang, welcher dA enthält, eine niedrigere Schmelztemperatur (58 °C) als der analoge Strang mit dG (62 °C), wodurch sich die unterschiedlichen Schmelztemperaturen der beiden geschädigten DNA-Duplexstränge relativieren und davon ausgegangen werden kann, dass die Destabilisierung in beiden Fällen ähnlich ist.

Anders sieht es bei den N^2 -Addukten aus. Diese führen nur zu einer sehr geringen Änderung der Schmelztemperatur im Vergleich zum ungeschädigten DNA-Doppelstrang. Es ist sogar eine leichte thermodynamische Stabilisierung bei N^2 -AAB-dG (65 °C) und N^2 -AAF-dG (67 °C) zu erkennen.

Bei den 34mer Oligonukleotiden sind die Unterschiede zwischen den Schmelzpunkten der geschädigten DNA-Stränge weniger stark ausgeprägt. Dennoch ist bei allen C8-Schäden eine Destabilisierung im Vergleich zum ungeschädigten Strang zu erkennen. Die Oligonukleotide, welche C8-AAF-dG oder C8-AF-dG Schäden enthalten führen zur geringsten Abweichung des Schmelzpunktes (82 °C), während beim Phenyl- und Naphthylschaden eine deutliche Änderung (77 °C) zu erkennen ist. Bei den N^2 -Addukten wurde nur ein 34mer Strang mit N^2 -AAF-dG synthetisiert. Man sieht eine deutliche Stabilisierung (87 °C) im Vergleich zur ungeschädigten DNA. Generell sind die Stabilisierungs- und Destabilisierungseffekte bei längeren DNA-Strängen weniger stark ausgeprägt als bei kürzeren Strängen, aber trotzdem noch deutlich zu beobachten.

Die gemessenen Schmelzpunkte stimmen gut mit den zu erwartenden Werten überein. Es wird deutlich, dass die Präsenz von C8-Addukten im DNA-Strang eine Erniedrigung der Schmelztemperatur zur Folge haben, was durch eine Destabilisierung des DNA-Doppelstrangs hervorgerufen wird, während die Präsenz von N^2 -Addukten zu einer Erhöhung der Schmelztemperatur in Folge einer Stabilisierung des Duplexstrangs führt. Acetylierte C8-Addukte haben eine größere Auswirkung auf die Stabilität des Doppelstrangs als nicht acetylierte Addukte. Dieses Verhalten ist im Einklang mit der Tatsache, dass C8-Addukte gute Substrate für die GG-NER sind, während N^2 -Addukte nur über TC-NER repariert werden können.

3.1.5 Bindungsstudien von geschädigter DNA und XPA/Rad14

Mit den in C8- bzw. N^2 -Position von dG oder dA geschädigten 15mer DNA-Doppelsträngen und Volllängen XPA-Protein (XPAfl) wurden zunächst *Electrophoretic Mobility Shift Assays* (EMSA) durchgeführt, um sicherzustellen, dass XPA für diese eine höhere Bindungsaffinität als für ungeschädigte DNA besitzt. Dies konnte für alle in Tabelle 2 aufgelisteten 15mer Oligonukleotide gezeigt werden. Im Anschluss wurden Kristallisationsexperimente mit der 15mer dsDNA und Rad14₁₈₈₋₃₀₂ durchgeführt. Hier konnten jedoch nur die Kristallstrukturen von Rad14 mit dsDNA, welche eines der C8-dG Addukte oder das N^2 -AAN-dG Addukt enthielt, erhalten werden. Alle anderen geschädigten Stränge konnten bis zur Abgabe dieser Arbeit nicht mit Rad14 co-kristallisiert werden. Die folgenden zwei Publikationen befassen sich mit den EMSA-Studien, sowie mit den erhaltenen Kristallstrukturen.

3.1.5.1 Zusammenfassung der Publikation zu acetylierten C8-dG Addukten

Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14.

N. Simon*, C. Ebert*, S. Schneider, *Chem. Eur. J.* **2016**, 22, 10782-10785. (*Beitrag zu gleichen Teilen)

Kopie der Publikation mit Erlaubnis des Verlags; Copyright 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Heterozyklische aromatische Amine reagieren mit Purinbasen, was zur Bildung von *bulky* DNA-Addukten führt, welche wiederum Mutationen verursachen können. Diese strukturell unterschiedlichen Schäden sind alle Substrate für die NER, welche verformte DNA Konformationen erkennen und verifizieren kann. XPA und XPC, welche als Gerüstproteine zwischen den DNA-Substraten und verschiedenen anderen NER-Proteinen agieren, spielen dabei eine wichtige Rolle. In der Kristallstruktur von Rad14, dem Hefehomolog von XPA, mit C8-AAF-dG geschädigter dsDNA liegt diese 70° geknickt vor. Diese Struktur deutete darauf hin, dass π -stacking Interaktionen eine wichtige Rolle bei der Stabilisierung des DNA-Knicks spielen. Deshalb werden hier sowohl die Synthese von dsDNA-Strängen, die die unterschiedlich großen polycyclischen, aromatischen Amidschäden Acetylaminophenyl, Acetylaminonaphthyl, Acetylaminoanthryl und Acethylaminopyrenyl in C8-Position von dG enthalten, als auch ihre Kristallstrukturen im Komplex mit Rad14 gezeigt. Den Strukturen ist gemein, dass je zwei Rad14-Moleküle an einen DNA-Doppelstrang binden und dabei einen Knick von ca. 70° induzieren. Die Struktur, welche C8-AAP-dG als Schaden enthält, ist am besten definiert, da der große Schaden die geringste Flexibilität aufweist. Diese Arbeit untermauert den indirekten Schadenserkennungsmechanismus des NER-Systems, welches destabilisierte, verformbare DNA-Strukturen erkennt.

Beitragserklärung: In Zusammenarbeit mit *Nina Simon*, die die DMTr-Stufen der oben genannten Verbindungen synthetisierte, wurden deren Phosphoramidite hergestellt. Es wurden 15mer DNA-Stränge, die in der Mitte den jeweiligen Schaden enthielten, synthetisiert und mit ihrem ungeschädigten Gegenstrang hybridisiert. Mit diesen Strängen wurden sowohl die EMSA-Bindungsstudien mit XPAfl als auch die Co-Kristallisationsexperimente mit Rad14 durchgeführt. Die Auswertung der Kristallstrukturen wurde von *Sabine Schneider* übernommen

Protein Crystal Structures

Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14

Nina Simon⁺,^[b] Charlotte Ebert⁺,^[b] and Sabine Schneider*^[a]

Abstract: Heterocyclic aromatic amines react with purine bases and result in bulky DNA adducts that cause mutations. Such structurally diverse lesions are substrates for the nucleotide excision repair (NER). It is thought that the NER machinery recognises and verifies distorted DNA conformations, also involving the xeroderma pigmentosum group A and C proteins (XPA, XPC) that act as a scaffold between the DNA substrate and several other NER proteins. Here we present the synthesis of DNA molecules containing the polycyclic, aromatic amine C8-guanine lesions acetylaminophenyl, acetylaminonaphthyl, acetylaminonanthryl, and acetylaminopyrenyl, as well as their crystal structures in complex with the yeast XPA homologue Rad14. This work further substantiates the indirect lesion-detection mechanism employed by the NER system that recognises destabilised and deformable DNA structures.

Defects in the nucleotide excision repair are known to be associated with several human, autosomal, recessive hereditary disorders, such as xeroderma pigmentosum (XP) and Cockayne Syndrome.^[1] Moreover, single-nucleotide polymorphisms (SNPs) in nucleotide excision repair (NER) genes have been associated with the susceptibility of cancers for the treatment with chemotherapeutic agents that exert their anti-tumour properties by inflicting DNA lesions.^[2] More than 30 proteins are involved in NER, which operates by a sequential assembly process. This comprises the recognition of damaged sites, excision of the oligonucleotide containing the lesion, gap filling by templated DNA synthesis and final sealing of the nick. The NER system is extremely versatile regarding the chemical structures of DNA lesions by either scanning the genome (global genome (GG)-NER) or acting on them during transcription (transcription-coupled NER).^[3] Currently it is believed that lesions are first recognised by the XPC (xeroderma pigmentosum

group C) protein^[4] followed by lesion verification by XPD (xeroderma pigmentosum group D protein), TFIIH (transcription factor IIH/XPB),^[5] XPA (xeroderma pigmentosum group A protein), replication protein A (RPA)^[6] and XPG (xeroderma pigmentosum group G protein),^[3e] resulting in the fully assembled pre-incision complex.^[3c,e,7] XPC, XPA and RPA were shown to ensure the proper 3D arrangement of the DNA intermediate for incision of the damaged single strand during the GG-NER.^[6] The interaction between RPA and XPA are essential for the NER.^[8] In addition, it was shown that XPA is a crucial component for both GG-NER as well as TC-NER,^[9] and mutations in XPA are amongst the strongest NER-phenotypes.^[10] Thus, XPA most likely acts as a scaffold during the NER.^[3e] Moreover, it was previously reported that XPA binds to a variety of DNA structures, such as kinks,^[11] DNA duplexes containing bulky adducts^[12] and cisplatin lesions.^[13]

To investigate the structural basis for the binding of the XPA homologue Rad14 from *Saccharomyces cerevisiae* to different C8-acetylated deoxyguanosine (dG) bulky-adduct lesions, we synthesised the acetylaminophenyl (dG-AAB), acetylaminonaphthyl (dG-AAN), acetylaminonanthryl (dG-AAA) and acetylaminopyrenyl (dG-AAP) phosphoramidites and incorporated the lesions into 15-mer and 34-mer DNA oligonucleotides by ultramild solid-phase DNA synthesis (Figure 1a, for details see the Supporting Information). The synthesis of the phosphoramidites dG-AAB, dG-AAA and dG-AAP was optimised by using a 2-(trimethylsilyl)ethanol (TMS) protecting group at the O⁶-position. TMS moieties can be removed in one step together with the *tert*-butyldimethylsilyl-ether (TBS) protecting groups.

The DNA duplexes containing the dG-AAB, dG-AAN, dG-AAA, or dG-AAP lesions were co-crystallised with the Rad14 minimal DNA-binding domain (DBD, residues 188–306) and the crystal structures determined by X-ray diffraction. (For data collection and structure refinement statistics see Table S3 in the Supporting Information).

As previously observed for the Rad14 binding to DNA with C8-dG-acetylaminofluorenyl (dG-AAF) or cisplatin (1,2 GG) lesions,^[14] two Rad14 protein molecules bind one DNA double strand with the DNA bent by about 75° into the major groove. The two proteins only interact with the DNA phosphodiester backbone and bury a solvent-accessible surface area of about 700 Å² per Rad14 molecule, but do not form any DNA-sequence- or DNA-lesion-specific contacts (Figure 1). In all complexes of Rad14 with damaged DNA, the bending of the DNA is achieved through the “fingers” with Gln266, Lys229 and Thr230, as well as His258 and Phe262 located on the β-hairpin, providing the anchor point by packing on the 13-mer recogni-

[a] Dr. S. Schneider

Center for Integrated Protein Science Munich CIPSM
Department of Chemistry, Technische Universität München
Lichtenbergstrasse 4, 85748 Garching (Germany)
E-mail: sabine.schneider@mytum.de

[b] N. Simon,⁺ C. Ebert⁺

Center for Integrated Protein Science Munich CIPSM
Department of Chemistry, Ludwig-Maximilians Universität München
Butenandtstrasse 13, 81377 München (Germany)

⁺ These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/chem.201602438>.

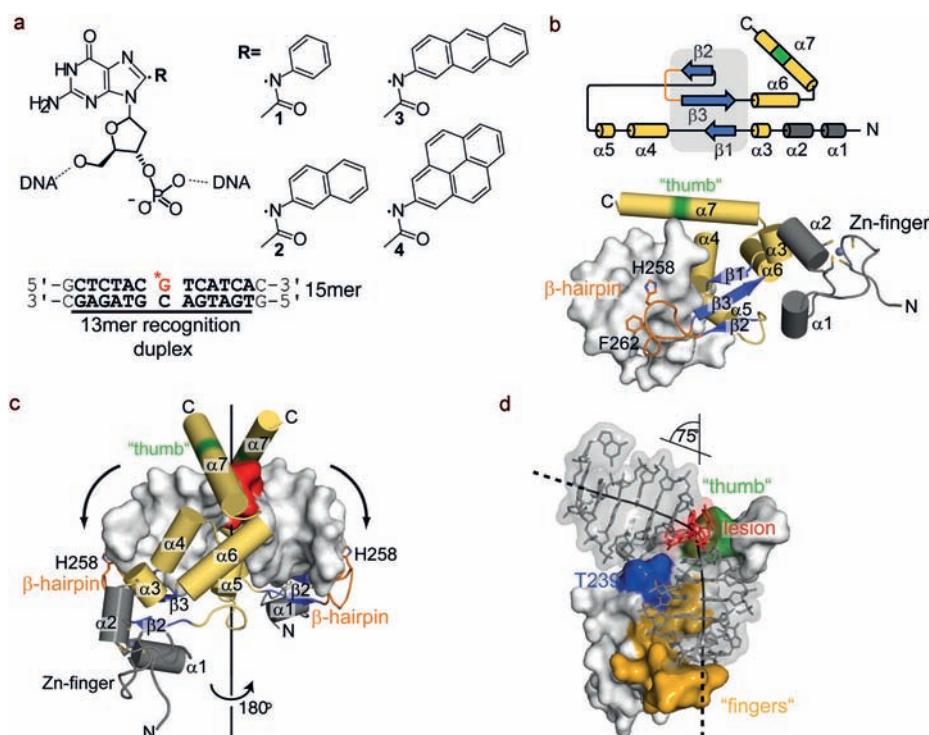


Figure 1. Structure of the Rad14 DBD in complex with DNA-containing bulky adducts. a) C8-dG-adduct lesions and DNA sequences used in this study. R = aceylaminophenyl (AAB, 1), acetylaminonaphthyl (AAN, 2), acetylamoanthryl (AAA, 3), acetylaminopyrenyl (AAP, 4). b) Overall folding topology of Rad14 and interaction with the DNA (grey surface). For clarity, only one Rad14 protein is shown. c) The DNA is bound by two Rad14 proteins and kinked at the position of the DNA lesion (red) towards the major groove (arrows). The protein is shown as cartoon and the DNA represented as grey surface. The β -hairpin with His258 and Phe262 (orange) pack on the 13-mer recognition duplex (the arginine "thumb" at the α_7 helix is highlighted in green). His258 and Phe262, as well as the Cys residues of the Zn-finger (dark grey), are shown as sticks and the Zn^{2+} as grey sphere. The black line indicates the pseudo C_2 symmetry axis. d) Rad14-DNA interface and bending mechanism; the "fingers" (gold) interact with the DNA backbone 3–6 base pairs away from the DNA-damage site. The "thumb" (green) bends the DNA at the lesion site by about 75° and Thr239 (blue) stabilises the kink. For clarity, only one Rad14 protein is depicted.

tion duplex. The arginine "thumb" (Arg293; Arg294) in the α_7 helix bends the DNA duplex with Thr239 stabilising the kinked structure (Figure 1). For the human XPA, it could recently be shown that the DNA-binding affinity was increased when the corresponding DBD (residues 98–219) was extended at the C terminus by 20 amino acids.^[15] This addition would lead to an extension at the α_7 helix resulting in a larger interaction area with the DNA duplex, which explains the higher binding affinity. All lesions are located at the central kink and the six base pairs on both sides adopt a B-form DNA topology that forms the 13-mer duplex-recognition element, in which the following bases of the DNA strand are unpaired (Figure 1). All structures exhibit a pseudo C_2 symmetry, in which the DNA is bound in two orientations, which is reflected by the mixed electron density observed for the purine–pyrimidine bases (Figure S2a in the Supporting Information). Thus, both orientations of the DNA duplexes were modelled with an occupancy of 50 %. The dC (deoxycytidine) opposite of the damaged base is flipped out and disordered, while the lesions slide into the duplex and stabilise the kinked structure. The damaged guanine base is in *syn* conformation and displaced into the narrowed major groove. In the structure of Rad14 in complex with dG-AAP, the planes of the pyrene moieties in the two duplex orientations are almost perpendicular to each other, stacking planar with its respect 5'-flanking basepair and verti-

cal on the 3'-flanking base pair (Figure 2 and Figure S2 in the Supporting Information). The phenyl, naphthyl and anthryl moieties in the other complex structures are also fully intercalated into the duplex, albeit their positions are less well defined in the electron-density maps (Figure S1 in the Supporting Information), which is most likely due to the following inherent structural properties: 1) the two alternative DNA orientations result in only 50 % occupancy for each lesion-containing strand and 2) the C_2 pseudo symmetry axis runs directly through the centre of the dsDNA at the position of the kink (Figure 1c) and causes noise in the electron-density maps, which is commonly observed along symmetry axis in protein crystals. Moreover,

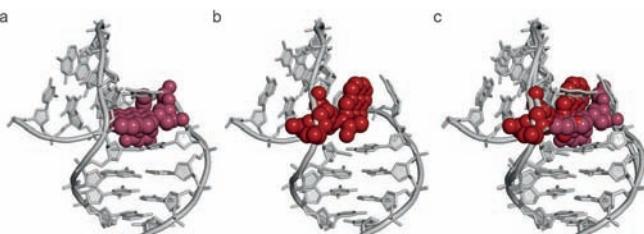


Figure 2. a,b) Position and stacking interactions of the dG-AAP lesions in the two orientations of the DNA duplex. c) Superposition of both orientations each with 50 % occupancy. The dG-AAP lesions are shown as space-fill model shaded in grey.

the smaller aromatic moieties allow a higher degree of flexibility compared to the pyrene ring system, which will provide a better stabilisation of the DNA kink due to its larger stacking area. This is also reflected by the higher temperature factors of the lesions and their neighbouring base pairs (1.2–1.5 times higher compared to the average B-factors of the DNA strands).

Previously, it was shown that DNA lesions, which do not disturb the DNA-duplex structure or neighbouring Watson–Crick base pairs and/or increase the DNA stability (such as bulky *cis*- and *trans*-benzo[*a*]pyrene-*N*²-dG adducts placed opposite an abasic site), evade the NER mechanism and, in particular, the recognition by XPC.^[16] In general, good substrates for the NER system are DNA strands with disturbed Watson–Crick base pairs, kinked/bend and destabilised DNA-duplex structures,^[17] some of which have been shown to be bound by XPC and XPA.^[11,12b,13]

Therefore, the mechanism underlying the structural variety recognised by the NER can be explained if the system employs a sequence-independent shape-readout mechanism in its lesion recognition and verification step. An overall shape-readout mechanism is also used by a number of other DNA-binding proteins, such as Holliday junction-resolving enzymes, Z-DNA-recognising and bacterial chromosomal proteins. Any complex that forms favourable enthalpy changes needs to drive unfavourable entropy changes and vice versa.^[18] The lesions used in this study destabilise DNA and reduce the melting temperature compared to the undamaged control strand, albeit the overall destabilising effect is compensated in longer DNA duplexes (66–76% 15-mer; 91–98% 34-mer; Table S2). Due to the destabilisation effect of the DNA lesions, the enthalpic cost for the DNA bending is reduced, hence, a productive recognition complex can be formed. Analogously, in the mismatch repair, the initial mispair-recognition steps are binding and bending of the linear mispair-containing DNA by roughly 50–60°.^[19] In case of *N*² adducts placed opposite an abasic site that lack pre-deformation of the DNA and possess an increased stability due to full intercalation of the aromatic moiety, the enthalpic cost that is associated with binding by NER proteins is likely to be compensated by favourable enthalpic and entropic changes. Hence, these lesions are resistant to the NER. Such an indirect shape-readout mechanism is also supported by the observations made by Reardon and Sancar; RPA, XPA and XPC recognise psoralen damage-containing DNA, but only RPA can be crosslinked to the lesion.^[20] The requirement for this chemical crosslink is the close proximity (ca. 8 Å) of the photoreactive amino acids to the pyrone side ring.^[21] Employing a nonspecific shape-readout mechanism, the lesion recognition and verification is likely to enable the NER repair machinery to deal with this vast variety of different chemical structures. Our structural data on the Rad14 DBD in complex with these bulky aromatic dG-adduct lesions further circumstantiates the hypothesis that Rad14, and most likely its human homologue XPA (sharing a 25% sequence identity and a 45% sequence similarity), is able to recognise numerous different lesions in the lesion-verification step of the NER by utilising a general mechanism of binding bendable DNA or DNA adopting already bend conformations.

Experimental Section

Phosphoramidite synthesis and DNA preparation

The dG-AAN phosphoramidite containing an isopropylphenoxyacetyl group at the *N*²-position, for the solid-phase DNA synthesis, was synthesised as previously published.^[22] For the synthesis of the dG-AAB-dG, dG-AAA and dG-AAP phosphoramidites, instead of the benzyl ether a 2-(trimethylsilyl)-ethanol (TMS) moiety was used at the *O*⁶-position, which could be removed in one step with the *tert*-butyldimethylsilyl groups, enabling a faster synthesis. The C8-lesion phosphoramidites were incorporated into DNA by solid-phase DNA synthesis under ultra-mild conditions.^[23] For synthetic and purification details, and a list of oligodesoxynucleotides (ODN) used in this study, see the Supporting Information.

Protein expression, purification, crystallisation and structure determination

The *Saccharomyces cerevisiae* Rad14 DNA-binding domain (residues 188–306) was expressed and purified as published elsewhere.^[14] For details on the complex crystallisation and structure determination see the Supporting Information. Atomic coordinates and X-ray structure data were submitted to the Protein Data Bank (<http://www.ebi.ac.uk/pdbe/>). PDB codes: 5G32, 5G33, 5G34 and 5G35.

Acknowledgements

The authors would like to thank the crystallisation facility at the Max Planck Institute of Biochemistry for setting up the crystallisation plates. We also want to thank the Swiss Light Source (SLS) and European Synchrotron radiation Facility (ESRF) for beamtime and the staff of beamlines PXI (SLS) and ID23-2, ID29, and ID30B (ESRF) for assistance with the crystal testing and data collection. We would also like to thank Thomas Carell for helpful discussions and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft, the excellence cluster CIPSM and the Fonds der Chemischen Industrie.

Keywords: bulky adducts · crystal structures · DNA damage · nucleotide excision repair · Rad14

- [1] a) M. Berneburg, J. Krutmann, *Hautarzt* **2003**, *54*, 33–40; b) J. E. Cleaver, *J. Dermatol. Sci.* **2000**, *23*, 1–11; c) C. F. Menck, V. Munford, *Genet. Mol. Biol.* **2014**, *37*, 220–233.
- [2] M. Monzo, I. Moreno, A. Navarro, R. Ibeas, R. Artells, B. Gel, F. Martinez, J. Moreno, R. Hernandez, M. Navarro-Vigo, *Oncology* **2008**, *72*, 364–370.
- [3] a) S. Alekseev, F. Coin, *Cell. Mol. Life Sci.* **2015**, *72*, 2177–2186; b) E. C. Friedberg, *Nat. Rev. Cancer* **2001**, *1*, 22–33; c) J. A. Marteijn, H. Lans, W. Vermeulen, J. H. Hoeijmakers, *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 465–481; d) I. Mellon, G. Spivak, P. C. Hanawalt, *Cell* **1987**, *51*, 241–249; e) H. Naegeli, K. Sugasawa, *DNA Repair* **2011**, *10*, 673–683; f) A. Sancar, *Science* **1994**, *266*, 1954–1956; g) S. Tornaletti, P. C. Hanawalt, *Biochimie* **1999**, *81*, 139–146.
- [4] a) M. S. Luijsterburg, G. von Bornstaedt, A. M. Gourdin, A. Z. Politi, M. J. Mone, D. O. Warmerdam, J. Goedhart, W. Vermeulen, R. van Driel, T. Hofer, *J. Cell Biol.* **2010**, *189*, 445–463; b) J. H. Min, N. P. Pavletich, *Nature* **2007**, *449*, 570–575; c) T. Riedl, F. Hanaoka, J. M. Egly, *EMBO J.* **2003**, *22*, 5293–5303; d) M. Volker, M. J. Mone, P. Karmakar, A. van Hofven, W. Schul, W. Vermeulen, J. H. Hoeijmakers, R. van Driel, A. A. van Zeeland, L. H. Mullenders, *Mol. Cell* **2001**, *8*, 213–224.

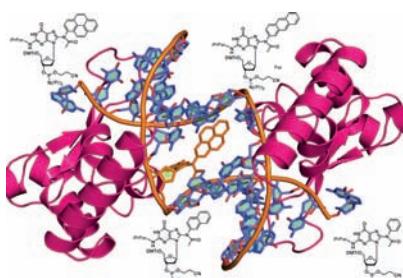
- [5] a) J. Kuper, S. C. Wolski, G. Michels, C. Kisker, *EMBO J.* **2012**, *31*, 494–502; b) N. Mathieu, N. Kaczmarek, H. Naegeli, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17545–17550; c) N. Mathieu, N. Kaczmarek, P. Ruthemann, A. Luch, H. Naegeli, *Curr. Biol.* **2013**, *23*, 204–212; d) R. A. Pugh, C. G. Wu, M. Spies, *EMBO J.* **2012**, *31*, 503–514.
- [6] Y. S. Krasikova, N. I. Rechkunova, E. A. Maltseva, I. O. Petrusheva, O. I. Lavrik, *Nucleic Acids Res.* **2010**, *38*, 8083–8094.
- [7] a) K. Sugasawa, J. Akagi, R. Nishi, S. Iwai, F. Hanaoka, *Mol. Cell* **2009**, *36*, 642–653; b) A. Tapias, J. Auriol, D. Forget, J. H. Enzlin, O. D. Scharer, F. Coin, B. Coulombe, J. M. Egly, *J. Biol. Chem.* **2004**, *279*, 19074–19083.
- [8] L. Li, X. Lu, C. A. Peterson, R. J. Legerski, *Mol. Cell. Biol.* **1995**, *15*, 5396–5402.
- [9] T. Kobayashi, S. Takeuchi, M. Saito, Y. Nakatsu, H. Morioka, E. Otsuka, M. Wakasugi, O. Nikaido, K. Tanaka, *Nucleic Acids Res.* **1998**, *26*, 4662–4668.
- [10] I. Satokata, K. Tanaka, S. Yuba, Y. Okada, *Mutat. Res.* **1992**, *273*, 203–212.
- [11] U. Camenisch, R. Dip, S. B. Schumacher, B. Schuler, H. Naegeli, *Nat. Struct. Mol. Biol.* **2006**, *13*, 278–284.
- [12] a) N. E. Geacintov, S. Broyde, T. Buterin, H. Naegeli, M. Wu, S. Yan, D. J. Patel, *Biopolymers* **2002**, *65*, 202–210; b) Y. Liu, Z. Yang, C. Utzat, G. Wang, A. K. Basu, Y. Zou, *Biochemistry* **2005**, *44*, 7361–7368; c) Z. G. Yang, Y. Liu, L. Y. Mao, J. T. Zhang, Y. Zou, *Biochemistry* **2002**, *41*, 13012–13020.
- [13] a) E. Evans, J. Fellows, A. Coffer, R. D. Wood, *EMBO J.* **1997**, *16*, 625–638; b) C. J. Jones, R. D. Wood, *Biochemistry* **1993**, *32*, 12096–12104; c) M. Missura, T. Buterin, R. Hindges, U. Hubscher, J. Kasparkova, V. Brabec, H. Naegeli, *EMBO J.* **2001**, *20*, 3554–3564.
- [14] S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8272–8277.
- [15] N. Sugitani, S. M. Shell, S. E. Soss, W. J. Chazin, *J. Am. Chem. Soc.* **2014**, *136*, 10830–10833.
- [16] a) Z. Liu, S. Ding, K. Kropachev, L. Jia, S. Amin, S. Broyde, N. E. Geacintov, *PLOS One* **2015**, *10*, e0137124; b) D. A. Reeves, H. Mu, K. Kropachev, Y. Cai, S. Ding, A. Kolbanovskiy, M. Kolbanovskiy, Y. Chen, J. Krzeminski, S. Amin, *Nucleic Acids Res.* **2011**, *39*, 8752–8764.
- [17] a) Y. Cai, D. J. Patel, N. E. Geacintov, S. Broyde, *J. Mol. Biol.* **2007**, *374*, 292–305; b) K. Kropachev, M. Kolbanovskii, Y. Cai, F. Rodriguez, A. Kolbanovskii, Y. Liu, L. Zhang, S. Amin, D. Patel, S. Broyde, *J. Mol. Biol.* **2009**, *386*, 1193–1203.
- [18] a) S. Harteis, S. Schneider, *Int. J. Mol. Sci.* **2014**, *15*, 12335–12363; b) L. Jen-Jacobson, L. E. Engler, L. A. Jacobson, *Structure* **2000**, *8*, 1015–1023; c) R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, R. S. Mann, *Annu. Rev. Biochem.* **2010**, *79*, 233–269; d) A. A. Travers, C. Vaillant, A. Arneodo, G. Muskhelishvili, *Biochem. Soc. Trans.* **2012**, *40*, 335–340.
- [19] G. L. Hura, C. L. Tsai, S. A. Claridge, M. L. Mendillo, J. M. Smith, G. J. Williams, A. J. Mastroianni, A. P. Alivisatos, C. D. Putnam, R. D. Kolodner, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17308–17313.
- [20] J. T. Reardon, A. Sancar, *Mol. Cell. Biol.* **2002**, *22*, 5938–5945.
- [21] S. S. Sastry, *Photochem. Photobiol.* **1997**, *65*, 937–944.
- [22] a) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, *3*, 565–568; b) L. C. Gillet, O. D. Scharer, *Org. Lett.* **2002**, *4*, 4205–4208.
- [23] L. C. Gillet, J. Alzeer, O. D. Scharer, *Nucleic Acids Res.* **2005**, *33*, 1961–1969.

Received: May 23, 2016

Published online on ■■■, 0000

COMMUNICATION

DNA unravelled: Crystal structures of the DNA-repair factor Rad14 in complex with damaged DNA substantiate the indirect lesion-detection mechanism of the recognition of destabilised and deformable DNA structures.



■ Protein Crystal Structures

N. Simon, C. Ebert, S. Schneider*



Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14



CHEMISTRY

A European Journal

Supporting Information

Structural Basis for Bulky-Adduct DNA-Lesion Recognition by the Nucleotide Excision Repair Protein Rad14

Nina Simon^{+[b]} Charlotte Ebert^{+[b]} and Sabine Schneider*^[a]

chem_201602438_sm_miscellaneous_information.pdf

Supporting Information

Structural basis for bulky adduct DNA lesion recognition by the nucleotide excision repair protein Rad14

Nina Simon, Charlotte Ebert and Sabine Schneider

General Methods

Chemicals were purchased from Sigma-Aldrich, Fluka, Alfa Aesar, ChemGenes or Acros and used without further purification. Solutions were concentrated in vacuo on a Heidolph rotary evaporator. The solvents were of reagent grade and purified by distillation. Dry solvents were bought from Sigma-Aldrich. Acetonitrile for HPLC-purification was purchased from VWR. Water was purified by a Milli-Q Plus system from Merck Millipore. Chromatographic purification of products was accomplished using flash column chromatography on Merck Geduran Si 60 (40-63 µM) silica gel (normal phase). Thin layer chromatography (TLC) was performed on Merck 60 (silica gel F₂₅₄) plates. ¹H, ¹³C and ³¹P NMR spectra were recorded in deuterated solvents on *Bruker ARX 300*, *Varian VXR400S*, *Varian Inova 400*, *Bruker AMX 600* and *Bruker Avance 800* spectrometers and calibrated to the residual solvent peak. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, hept = septet, m = multiplet, br. = broad. For assignment of the structures, additional 2D NMR spectra (COSY, HSQC, HMBC) were measured. Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectra were recorded on a Bruker Autoflex II. For MALDI-TOF measurements the samples were dialysed on a 0.025 µm VSWP filter (Millipore) against ddH₂O for 1 h and then cocrystallised in a 3-hydroxypicolinic acid matrix (HPA: 25 mg 3-hydroxypicolinic acid, 5 mg ammonium citrate, 5 µL 15-crown-5 in 0.5 mL H₂O). High resolution electrospray ionization mass spectra (HRMS-ESI) were recorded on a Thermo Finnigan LTQ-FT (ESI-FTICR). The pH-values of buffers were adjusted using a MP 220 pH-meter (Metter Toledo). UV spectra and melting profiles were measured on a Jasco V-650 spectrometer using quartz glass cuvettes with 1 cm path length. The concentration of the purified DNA was determined with a NanoDrop ND-1000 spectrophotometer (Peqlab) from Thermo Scientific (1.5 µL sample volume) using calculated extinction coefficients from Oligo Calc (online: <http://www.basic.northwestern.edu/biotools/OligoCalc.html>) and applying Lambert-Beer's law. For strands containing artificial bases the extinction coefficient of their corresponding control strand was employed without corrections.

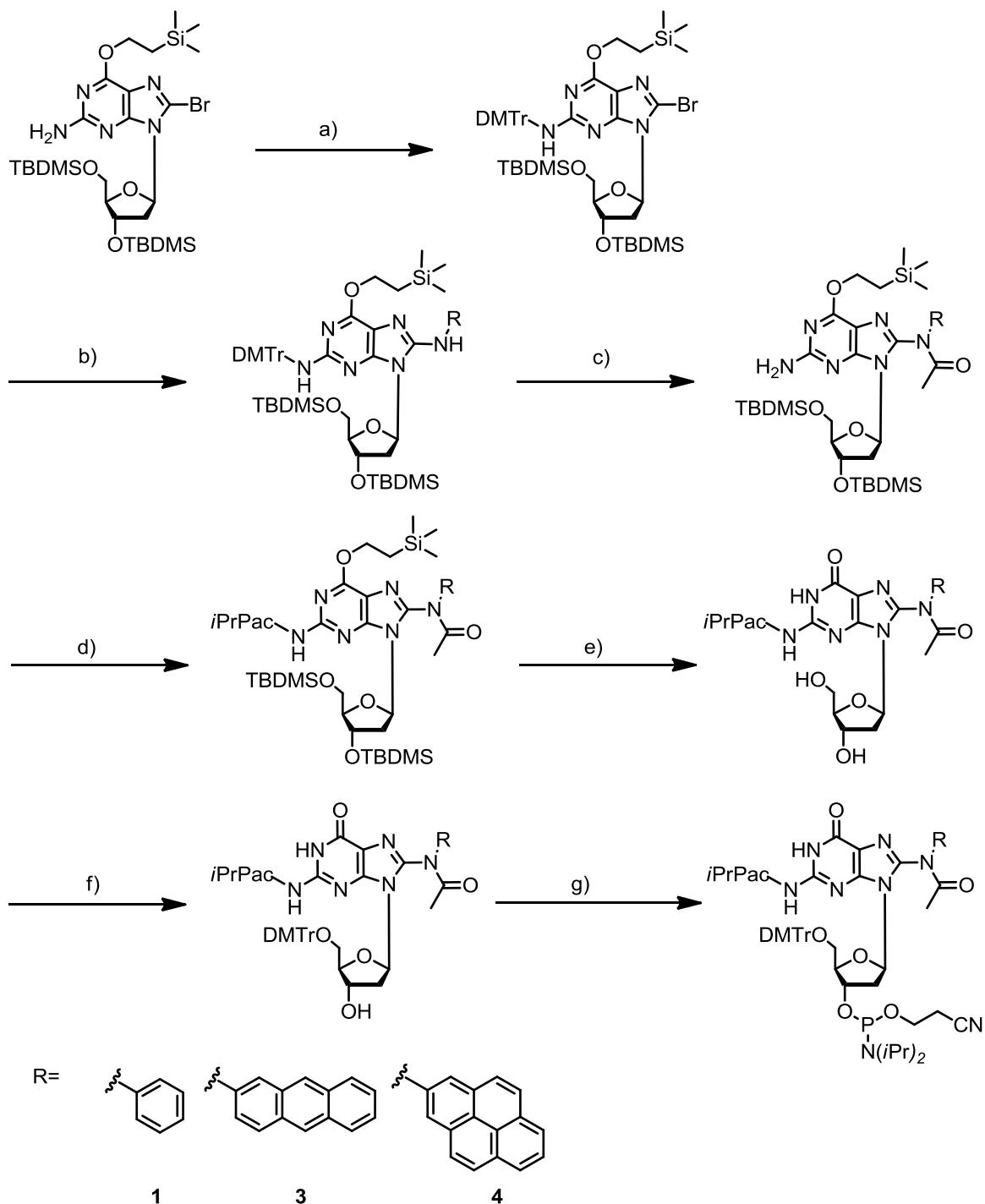
Phosphoramidite and DNA synthesis Phosphoramidites for dA, dC, dG, dT and CPG carriers were obtained from Glen Research (Sterling, VA) or Link Technologies (Bellshill, Scotland). Incorporation of the C8-lesion phosphoramidites into DNA was done by solid phase DNA synthesis with ultra-mild conditions,^[1] using an ABI 394 Nucleic Acid Synthesis System (Life Technologies, Darmstadt, Germany). The coupling time for the modified phosphoramidites was extended to 3 x 15 min. The oligonucleotides were purified by reverse phase HPL chromatography (Macherey-Nagel, Nucleosil 100-7 C18, 10 x 250 mm, 0.1 M triethylammonium acetate, linear gradient from 0% to 40% in 45 min for the 15mers or 0% to 20% acetonitrile in 65 min for the longer oligonucleotides). The purified oligonucleotides were desalted using C18 cartridges (Waters, Sep-Pak C18 Classic Cartridge) and their identity confirmed by matrix-assisted laser-desorption time-of-flight mass spectrometry.

Determination of melting temperatures The DNA duplexes (2 µM) were dissolved in 250 µL of 10 mM Tris-HCl, pH 7.4, 150 mM NaCl. The ODN 1 and ODN 3 samples were heated from 15 to 90 °C or cooled from 90 to 15 °C with a rate of 1 °C/min. The ODN 2 and ODN 4 samples were heated from 20 to 95 °C or cooled from 95 to 20 °C with a rate of 1 °C/min. Three melting profiles per sample were performed.

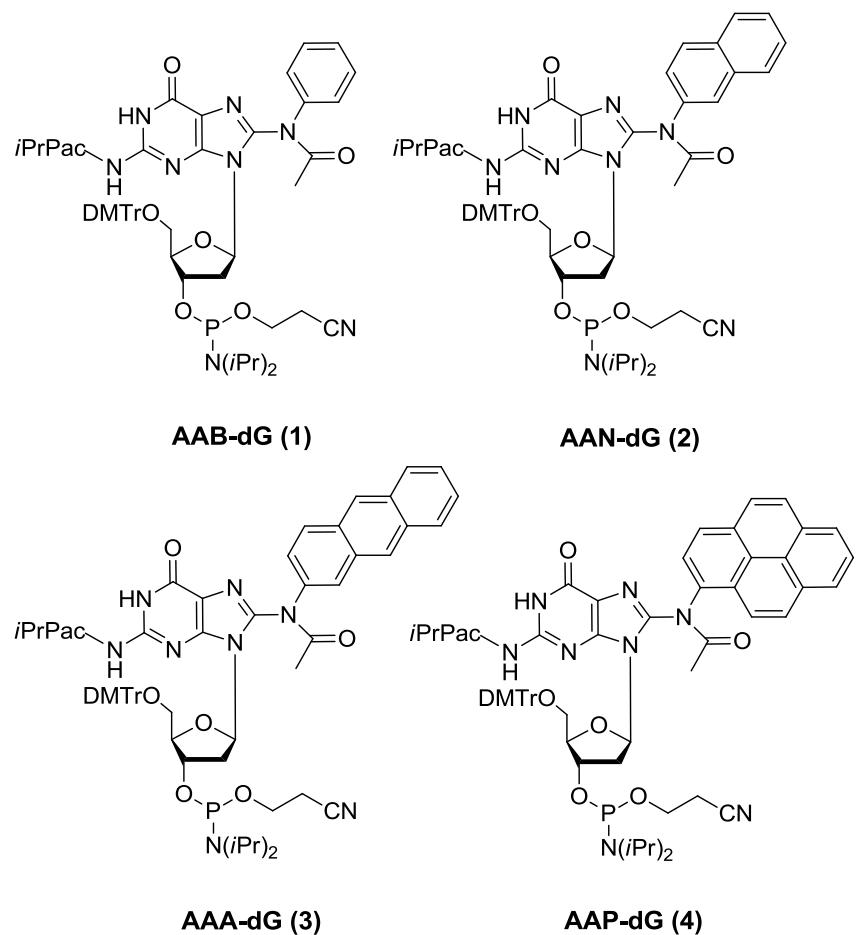
Data processing and structure determination

The *Saccharomyces cerevisiae* Rad14 DNA binding domain (residues 188-306) was expressed and purified as published in.^[2] The C8-lesion-dG containing 15mer DNA (ODN1, see Supplementary Table 1) was annealed in crystallisation buffer to its counter strand (ODN 2, Metabion). Prior to crystallisation protein and DNA were mixed in a 2:1 (protein:DNA) molar ratio and incubated for 30 min at 4 °C. Crystals were grown in 0.2-0.22 M ammonium nitrate and 38 - 44% 2-methyl-1,3,-propanediol (MPD) at 4 °C, cryo-cooled in liquid nitrogen directly from the crystallisation drop and stored in liquid nitrogen until data collection. Due to the small crystal size of 10-20 µm, diffraction data were collected with a beam size of 10 x 10 µm (PX I beamline, Swiss Light Source (SLS); ID30B beamline, European Synchrotron Radiation Facility (ESRF)). The data were processed with XDS^[3] and the resolution cut-offs were chosen according to the correlation coefficient of random half-data sets (1/2 CC) at about 50%.^[4] The structures were solved using the coordinates of the Rad14-DNA complex (PDB code 5A3D) by rigid-body refinement in REFMAC5^[5] after the atoms for the DNA were removed from the model. The models were completed by iterative cycles of manual model building in COOT^[6] and restraint and TLS refinement^[7] in REFMAC5. The pseudo C2 two fold symmetry axis that runs through the centre of the DNA double strand can be turned into a crystallographic two-fold axis by processing the data with comparable statistics in the space group P4₁2₁2, instead of P4₁. However this results in a composition of the crystals asymmetric unit (asu) of one protein with one dsDNA strand at 50% occupancy (P4₁2₁2)

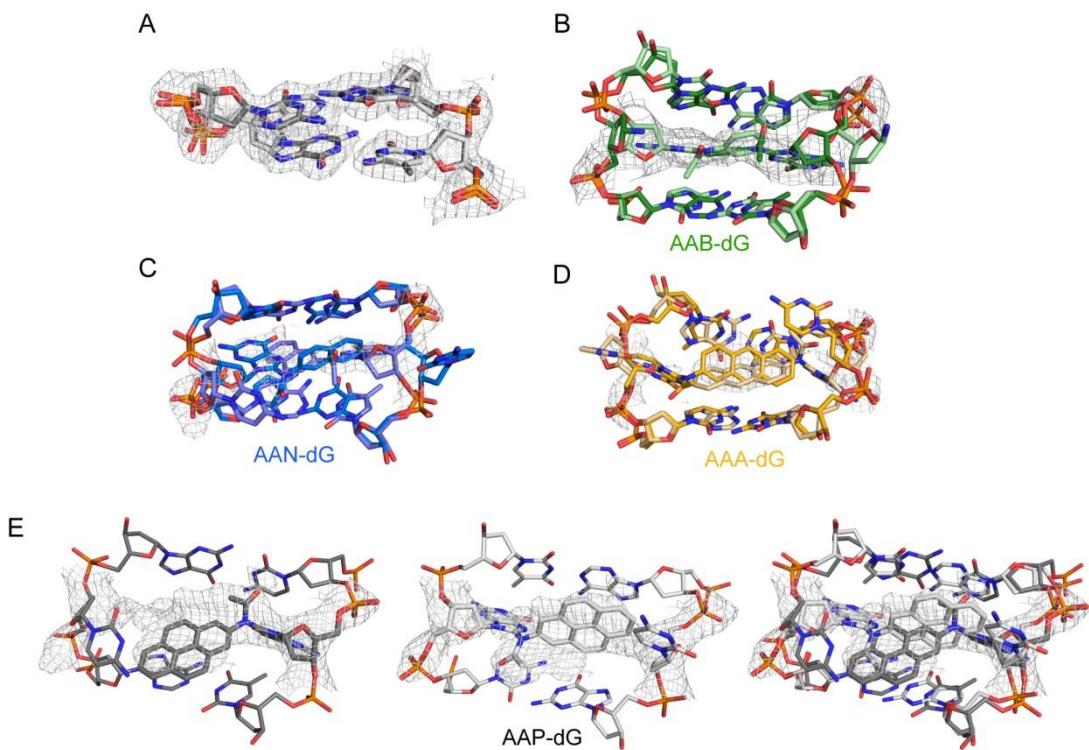
compared to two proteins with two dsDNA each with 50% occupancy (P4₁). For both space groups comparable model refinement statistics were obtained. Thus the crystal system P4₁ was chosen, which contains the biological relevant complex of two proteins recognising one dsDNA in two orientations in the asu. Simulated-annealing omit electron density map coefficients were calculated using PHENIX.^[8] Structure validation and analysis was carried out with MOLPROBITY^[9], PISA^[10] and CURVES.^[11] For data processing and structure refinement details and statistics see the Supporting Information. Structural superposition were done with SSM.^[12] All structural figures were prepared with PyMol (Delano Scientific).



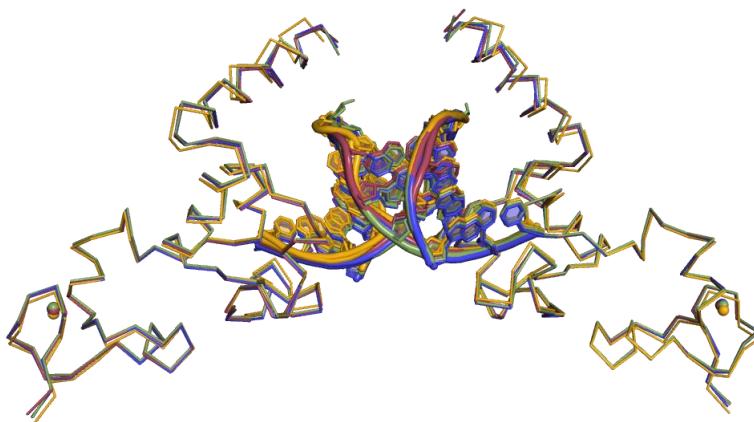
Supporting Scheme 1. Synthesis of the phosphoramidites **1**, **3** and **4**. a) DMTr-Cl, pyridine, RT, b) Pd_{2}db_3 , BINAP, NaOtBu , toluol, aryl- NH_2 (1,2 or 3), $110\ ^\circ\text{C}$, c) 1. 4-DMAP, NEt_3 , Ac_2O , pyridine, RT, 2. HCl in MeOH (0.01 M), RT, d) $i\text{PrPac-Cl}$, pyridine, RT, e) TBAF, AcOH , THF, RT, f) DMTr-Cl, pyridine, RT, g) $(i\text{Pr}_2\text{N})_2\text{PO}(\text{CH}_2)_2\text{CN}$, DIAT, DCM, RT.



Supporting Figure 1: Synthesized C8-dG lesion phosphoramidites.



Supporting Figure 2. Illustration of the electron density maps. A) Observed $2\text{Fo}-\text{DFc}$ electron density contoured at 1σ , indicating the two DNA orientations. Simulated annealing $\text{Fo}-\text{DFc}$ electron density maps, contoured at 2.5σ , calculated excluding the atoms for dG-AAB (B), AAN (C), dG-AAA (D) and dG-AAP (E).



Supporting Figure 3. Structural superposition of all Rad14-DNA structures. The structures can be superimposed with an r.m.s.d of 0.3-0.8 Å depending on the chosen reference frame. (only main chain atoms of the protein vs all atoms) The dG-AAB complex is shown in green, the dG-AAN complex in blue, the dG-AAA complex in golden and the dG-AAP complex in red. The protein is depicted as ribbon, with the DNA shown as cartoon and the Zn^{2+} as sphere.

Supporting Table 1. Oligonucleotide sequences used in this study. The damaged guanine base is marked with a star and highlighted in bold.

ODN	Oligonucleotide sequence
1	5'-GCT CTA CG* T CAT CAC-3'
2	5'-GTA GCT AGG TCG AT G* CAC AGA CAT CGC ACA CGT C-3'
3	5'-GTG ATG ACG TAG AGC-3'
4	5'-GACG TGT GCG ATG TCT GTG CAT CGA CCT AGC TAC-3'

Supporting Table 2. Destabilising effect of the different lesions on the thermal stability of duplex DNA shown by the respective melting points.

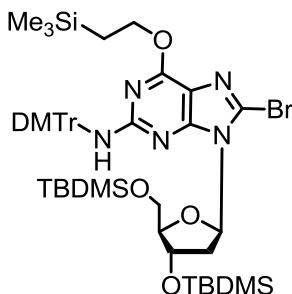
DNA lesion	T _M [°C] 15mer	T _M [°C] 34mer
dG:dC	64	84
C8-AAB-dG:dC	42	77
C8-AAN-dG:dC	46	77
C8-AAA-dG:dC	49	80
C8-AAP-dG:dC	47	80
C8-AAF-dG:dC	48	82

Supporting Table 3. Data processing and structure refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

	Rad14-dG-AAB	Rad14-dG-AAN	Rad14-dG-AAA	Rad14-dG-AAP
Wavelength	0.979	0.979	0.979	0.979
Resolution range	49.1 - 2.2 (2.3 - 2.2)	41.4 - 2.4 (2.5 - 2.4)	49.2 - 1.9 (2.0 - 1.9)	37.6 - 2.0 (2.1 - 2.0)
Space group	P 4 ₁	P 4 ₁	P 4 ₁	P 4 ₁
Unit cell	53.0 53.0 130.9	53.4 53.4 131.3	53.1 53.1 130.3	53.2 53.2 131.3
Total reflections	123,844 (11,546)	50,634 (4,920)	125,416 (8,792)	82,958 (8,405)
Unique reflections	18,268 (1,818)	14,071 (1,392)	28,204 (2,749)	24,353 (2,436)
Multiplicity	6.8 (6.4)	3.6 (3.5)	4.4 (3.2)	3.4 (3.5)
Completeness (%)	1.00 (1.00)	0.98 (0.98)	1.00 (0.97)	0.99 (0.99)
Mean I/sigma(I)	6.62 (0.97)	5.8 (1.1)	9.8 (1.4)	7.2 (1.0)
Wilson B-factor	41.30	46.9	33.0	36.0
R-merge	0.163 (1.111)	0.189 (0.994)	0.071 (0.681)	0.097 (1.248)
R-meas	0.177 (1.21)	0.223 (1.17)	0.081 (0.809)	0.115 (1.476)
CC1/2	0.995 (0.568)	0.976 (0.531)	0.998 (0.335)	0.997 (0.456)
CC*	0.999 (0.851)	0.994 (0.833)	1 (0.709)	0.999 (0.791)
Reflections used in refinement	18,255 (1,816)	14,056 (1,390)	28,199 (2,749)	24,338 (2,435)
Reflections used for R-free	895 (96)	720 (89)	1,361 (129)	1,255 (121)
R-work	0.215 (0.323)	0.255 (0.374)	0.214 (0.328)	0.214 (0.355)
R-free	0.253 (0.296)	0.307 (0.402)	0.245 (0.342)	0.2587 (0.392)
CC(work)	0.939 (0.700)	0.937 (0.739)	0.934 (0.613)	0.955 (0.608)
CC(free)	0.940 (0.793)	0.919 (0.523)	0.933 (0.464)	0.945 (0.678)
Number of non-hydrogen atoms	3,053	3,069	3,121	3,134
macromolecules	3,015	3,061	3,062	3,079
ligands	2	2	2	2
Protein residues	229	229	228	229
RMS(bonds)	0.042	0.031	0.052	0.034
RMS(angles)	1.80	1.72	1.87	2.12
Ramachandran favored (%)	96	95	98	98
Ramachandran allowed (%)	3.6	5.3	1.8	2.2
Ramachandran outliers (%)	0	0	0	0
Rotamer outliers (%)	2	3.4	0.98	0.97
Clashscore	2.5	2.1	1.1	3.5
Average B-factor	54.5	57.8	51.2	47.9
macromolecules	54.6	57.9	51.4	48.1
ligands	35.5	50.8	30.4	33.6
solvent	38.8	32.6	39.0	39.9
Number of TLS groups	20	20	20	28

Synthesis of phosphoramidites

C8-Br-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-N²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



C8-Br- O^6 -(2-(trimethylsilyl)-ethyl)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-deoxyguanosine^[13] (1.50 g, 2.23 mmol, 1.0 eq.) was dissolved in dry pyridine, molecular sieves (4 Å) were added and the solution was stirred at room temperature for 3 h. DMT-chloride (1.04 g, 2.68 mmol, 1.2 eq.) and 0.37 mL dry trimethylamine (2.7 mmol, 1.2 eq.) were added and the mixture was stirred at room temperature for 72 hours. The residue was purified via column chromatography (isohexane 30:1 EtOAc + 1% pyridine; SiO₂) and C8-Br-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-N²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (2.03 g, 2.07 mmol, 93%) was obtained as a white foam.

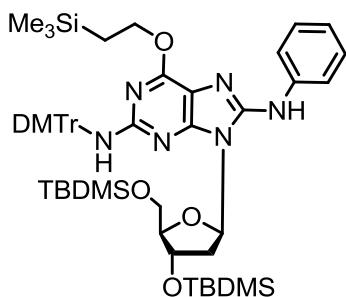
¹H-NMR (400 MHz, Acetone-*d*₆): δ = 7.36 (d, 3H, DMTr), 7.25 (d, 6H, DMTr), 6.84 (d, 4H, DMTr), 6.40 (s, *br*, 1H, NH), 6.23 (t, 1H, C1'-H), 4.84 (s, 1H, C3'-H), 3.98-3.81 (m, 4H, O⁶-CH₂, C5'-H, C4'-H), 3.77 (s, 6H, OCH₃), 3.68 (dd, 1H, C5'-H), 3.61-3.57 (m, 1H, C2'-H), 2.23-2.12 (m, 1H, C2'-H), 0.93 (s, 9H, C(CH₃)₃), 0.90–0.87 (m, 2H, O⁶-CH₂-CH₂), 0.84 (s, 9H, C(CH₃)₃), 0.16 (s, 3H, SiCH₃), 0.15 (s, 3H, SiCH₃), 0.00 (s, 3H, SiCH₃), -0.04 (s, 9H, Si(CH₃)₃), -0.05 (s, 3H, SiCH₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 159.3, 158.5, 155.0, 147.2, 139.1, 131.0, 129.8, 129.0, 128.4, 127.4, 126.5, 116.9, 113.8, 113.7, 88.1, 86.7, 65.5, 63.5, 55.5, 36.6, 26.3, 26.2, 18.9, 18.7, 17.7, -1.4, -4.3, -4.4, -5.1, -5.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₈H₇₁BrN₅O₆Si₃]⁺ ([M+H]⁺): 976.3890, found: 976.3895.

The phosphoramidites of the pyrenyl and the anthryl adducts were prepared analogous to the phenyl adduct phosphoramidite.

8-(2-aminobenzene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-N²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



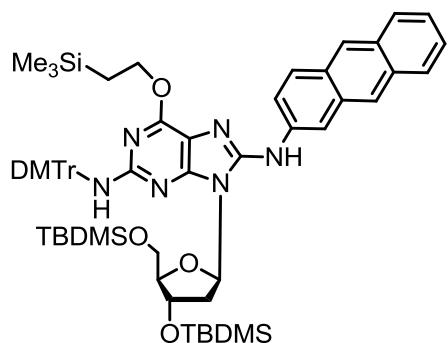
C8-Br-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-N²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (2.00 g, 2.05 mmol, 1.0 eq.) was dissolved in 35 mL dry toluene under argon atmosphere. Aniline (0.37 mL, 4.1 mmol, 2.0 eq.), Tris(dibenzylidenacetone)dipalladium(0) (0.19 g, 0.21 mmol, 0.1 eq.) and (2,2'-bis(diphenylphosphino)-1,1'-binaphthyl) (0.39 g, 0.62 mmol, 0.3 eq.) were added and the reaction mixture was heated to 110 °C. Sodium-*tert*-butanolate (0.39 g, 4.1 mmol, 2.0 eq.) dissolved in 35 mL toluene was added and the reaction mixture was stirred at 110 °C for 1.5 h. After cooling down to room temperature the mixture was filtered over SiO₂ deactivated with pyridine and the filter cake was washed with 1% pyridine in EtOAc (150 mL). The filtrate was concentrated under reduced pressure and purified via column chromatography (isohexane/EtOAc 1:0 to 7:1 + 1% pyridine; SiO₂). 8-(2-aminobenzene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-N²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.62 g, 1.64 mmol, 80%) was obtained as a red brown foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 7.96 (s, br, 1H, C8-NH), 7.73–7.69 (m, 2H, C₆H₅), 7.41–7.37 (m, 3H, DMTr), 7.28 (d, 6H, DMTr), 7.25 (d, 2H, C₆H₅), 6.93 (tt, 1H, C₆H₅), 6.83 (d, 4H, DMTr), 6.23 (s, br, 1H, N²H), 6.16 (t, 1H, C1'-H), 4.70 (s, 1H, C3'-H), 4.07–3.98 (m, 2H, O⁶-CH₂), 3.97–3.92 (m, 2H, C5'-H, C4'-H), 3.82 (dd, 1H, C5'-H), 3.76 (s, 6H, OCH₃), 3.26–3.11 (m, 1H, C2'-H), 2.15 (ddd, 1H, C2'-H), 0.99–0.94 (m, 2H, O⁶-CH₂CH₂), 0.93 (s, 9H, C(CH₃)₃), 0.85 (s, 9H, C(CH₃)₃), 0.15 (s, 3H, SiCH₃), 0.14 (s, 3H, SiCH₃), 0.04 (s, 3H, SiCH₃), 0.01 (s, 3H, SiCH₃), 0.00 (s, 9H, Si(CH₃)₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 159.1, 158.4, 157.0, 147.4, 141.6, 139.5, 131.1, 129.9, 129.7, 129.4, 128.3, 127.2, 124.6, 122.2, 119.0, 113.5, 113.3, 88.4, 85.0, 73.3, 70.9, 64.6, 63.8, 55.4, 38.7, 26.4, 26.3, 19.0, 18.7, 18.1, -1.3, -4.3, -4.4, -5.1, -5.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₅₄H₇₇N₆O₆Si₃]⁺ ([M+H]⁺): 989.5207, found: 989.5211.

8-(2-aminoanthracene)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-N²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)ethyl)-2'-deoxyguanosine:



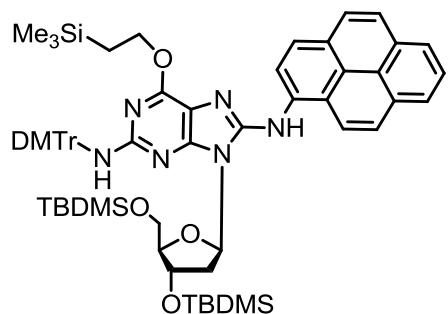
Yield: 64%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 8.42 (s, 1H, -C₁₄H₉), 8.32 (s, 1H, -C₁₄H₉), 8.28 (s, br, 1H, C8-NH), 7.99–7.96 (m, 3H, -C₁₄H₉), 7.79–7.69 (m, 1H, -C₁₄H₉), 7.63 (dd, 1H, -C₁₄H₉), 7.45–7.37 (m, 2H, -C₁₄H₉, 3H, DMTr), 7.30 (d, 6H, DMTr), 6.84 (d, 4H, DMTr), 6.28 (s, 1H, N²H), 6.23 (t, 1H, C1'-H), 4.77–4.71 (m, 1H, C3'-H), 4.08–4.05 (m, 2H, O⁶-CH₂), 4.02–3.96 (m, 2H, C5'-H, C4'-H), 3.86 (dd, 1H, C5'-H), 3.77 (s, 6H, OCH₃), 3.24 (dd, 1H, C2'-H), 2.21 (dt, 1H, C2'-H), 1.00 (dd, 2H, O⁶-CH₂CH₂), 0.93 (s, 9H, C(CH₃)₃), 0.85 (s, 9H, C(CH₃)₃), 0.17 (s, 3H, SiCH₃), 0.15 (s, 3H, SiCH₃), 0.06 (s, 9H, Si(CH₃)₃), 0.06 (s, 3H, SiCH₃), 0.03 (s, 3H, SiCH₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 159.2, 158.6, 157.2, 153.7, 150.7, 147.6, 147.2, 139.5, 138.4, 133.7, 133.3, 131.5, 131.1, 129.9, 129.5, 129.5, 129.0, 128.5, 128.3, 127.2, 126.8, 126.3, 125.3, 125.2, 124.6, 122.2, 113.6, 112.2, 88.5, 85.3, 73.4, 71.0, 64.7, 63.8, 55.5, 26.4, 26.3, 19.0, 18.7, 18.0, -1.1, -4.3, -4.4, -5.0, -5.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₆₂H₈₁N₆O₆Si₃]⁺ ([M+H]⁺): 1089.5520, found: 1089.5520.

8-(2-aminopyrene)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-N²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)ethyl)-2'-deoxyguanosine



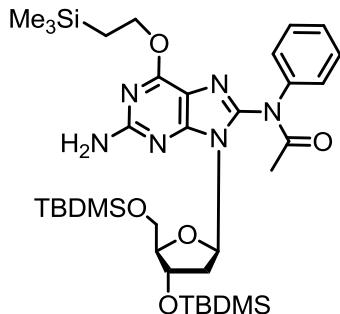
Yield: 58%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 8.67–8.59 (m, 2H, -C₁₆H₉), 8.25–7.96 (m, 8H, -C₁₆H₉, C8-NH), 7.45–7.38 (m, 2H, DMTr), 7.33–7.28 (m, 6H, DMTr), 7.26–7.18 (m, 1H, DMTr), 6.86–6.81 (m, 4H, DMTr), 6.39–6.30 (m, 1H, C1'-H), 6.29 (s, 1H, N²H), 4.78–4.69 (m, 1H, C3'-H), 4.16–4.11 (m, 1H, C4'-H), 4.04 (dd, 1H, C5'-H), 4.01–3.91 (m, 2H, O⁶-CH₂), 3.89 (dd, 1H, C5'-H), 3.77 (s, 6H, 2 x OCH₃), 3.40–3.18 (m, 1H, C2'-H), 2.33–2.21 (m, 1H, C2'-H), 0.92 (s, 9H, C(CH₃)₃), 0.90–0.83 (m, 2H, O⁶-CH₂-CH₂), 0.65 (s, 9H, C(CH₃)₃), 0.17 (s, 3H, SiCH₃), 0.17 (s, 3H, SiCH₃), -0.06 (s, 9H, Si(CH₃)₃), -0.13 (s, 3H, Si(CH₃)₃), -0.22 (s, 3H, Si(CH₃)₃) ppm.

¹³C-NMR (100 MHz, Acetone-*d*₆): δ = 159.1, 157.0, 154.0, 149.1, 147.6, 139.5, 132.6, 132.1, 131.1, 129.9, 128.3, 128.3, 127.9, 127.3, 127.2, 127.0, 126.5, 126.4, 126.2, 125.7, 125.3, 125.2, 122.0, 113.5, 88.8, 85.9, 73.5, 70.9, 64.6, 63.8, 55.4, 39.3, 26.3, 26.2, 24.4, 18.8, 18.6, 18.0, 1.3, -4.3, -4.4, -5.3, -5.4 ppm.

HR-MS (ESI⁺): m/z calculated for [C₆₄H₈₁N₆O₆Si₃]⁺ ([M+H]⁺): 1113.5520, found: 1113.5525.

8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



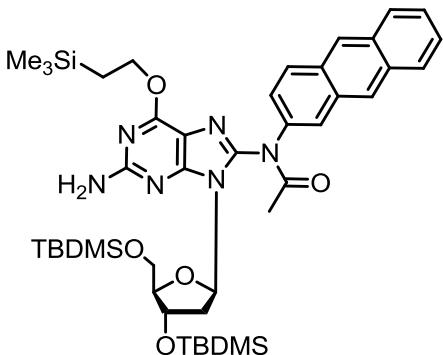
8-(2-aminobenzene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (4.37 g, 4.42 mmol, 1.0 eq.) was dissolved in 50 mL dry pyridine. 4-Dimethylaminopyridine (0.27 g, 2.2 mmol, 0.5 eq.), trimethylamine (3.08 mL, 22.1 mmol, 5.0 eq.) and acetic anhydride (2.09 mL, 22.1 mmol, 5.0 eq.) were added and the reaction mixture was stirred at room temperature for 48 h under argon atmosphere. After evaporation to dryness the residue was dissolved in EtOAc (200 mL) and washed with water (2 x 60 mL). The organic phase was dried with MgSO₄, filtered, concentrated under reduced pressure and dried for 3 h under high vacuum. Methanolic hydrochloric acid (30 mL, 0.01 M) was added and the mixture was stirred at room temperature for 20 h. After evaporation of the solvent, the residue was purified via column chromatography (*isohexane/EtOAc* 1:0 to 4:1; SiO₂). 8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (2.23 g, 3.04 mmol, 69%) was obtained as a white yellowish foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 7.56–7.15 (m, 5H, -C₆H₅), 6.13 (t, 1H, C1'-H), 5.72 (s, br, 2H, NH₂), 4.80 (s, 1H, C3'-H), 4.56 (t, 2H, O⁶-CH₂), 3.94–3.84 (m, 2H, C5'-H, C4'-H), 3.77–3.72 (m, 1H, C5'-H), 3.44 (dt, 1H, C2'-H), 2.11–2.06 (m, 4H, C2'-H, COCH₃), 1.20 (t, 2H, O⁶-CH₂-CH₂), 0.93 (s, 9H, C(CH₃)₃), 0.84 (s, 9H, C(CH₃)₃), 0.16 (s, 3H, SiCH₃), 0.13 (s, 3H, SiCH₃), 0.10 (s, 9H, Si(CH₃)₃), 0.01 (s, 3H, SiCH₃), -0.01 (s, 3H, SiCH₃) ppm.

¹³C-NMR (201 MHz, Acetone-*d*₆): δ = 162.0, 159.3, 154.3, 144.2, 141.2, 130.0, 128.7, 128.3, 127.4, 114.2, 88.4, 84.8, 73.4, 64.8, 63.6, 55.4, 37.1, 26.2, 26.1, 18.8, 18.5, 17.9, -1.3, -4.5, -4.5, -5.2, -5.3 ppm.

HR-MS (ESI⁺): calculated for [C₃₅H₆₁N₆O₅Si₃]⁺ ([M+H]⁺): 729.4006, found.: 729.4012.

8-(*N*-Acetyl-2-aminoanthracene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



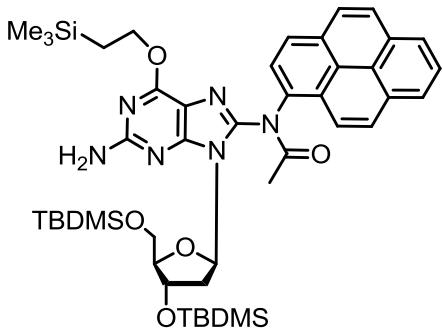
Yield: 74%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 8.56 (s, 1H, -C₁₄H₉), 8.49 (s, 1H, -C₁₄H₉), 8.12 (s, 1H, -C₁₄H₉), 8.06 (ddd, 3H, -C₁₄H₉), 7.64 (dd, 1H, -C₁₄H₉), 7.51 (dd, 2H, -C₁₄H₉), 6.21 (s, 1H, C1'-H), 5.74 (s, br, 2H, NH₂), 4.82–4.75 (m, 1H, C3'-H), 4.60 (t, 2H, O⁶-CH₂), 3.90–3.85 (m, 2H, C5'-H, C4'-H), 3.74 (d, 1H, C5'-H), 3.45 (ddd, 1H, C2'-H), 2.22–2.13 (m, 4H, C2'-H, COCH₃), 1.23 (t, 2H, O⁶-CH₂CH₂), 0.89 (s, 9H, C(CH₃)₃), 0.81 (s, 9H, C(CH₃)₃), 0.12 (s, 12H, 2× Si(CH₃)₂), -0.04 (s, 9H, Si(CH₃)₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 162.1, 160.8, 159.3, 154.5, 144.3, 138.2, 133.0, 133.0, 132.3, 130.1, 129.8, 129.1, 128.9, 128.4, 127.3, 127.1, 126.7, 126.6, 114.4, 113.70, 88.5, 85.0, 73.5, 64.9, 63.9, 37.1, 26.2, 26.2, 23.3, 18.8, 18.5, 18.1, -1.3, -4.4, -4.5, -5.2, -5.3 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₃H₆₅N₆O₅Si₃]⁺ ([M+H]⁺): 829.4319, found: 829.4319.

8-(*N*-Acetyl-2-aminopyrene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



Yield: 70%

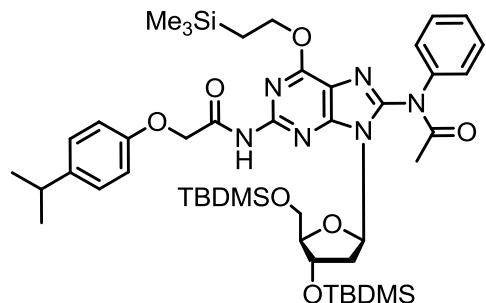
¹H-NMR (800 MHz, Acetone-*d*₆): δ = 8.71–8.62 (m, 1H, -C₁₆H₉), 8.47–7.92 (m, 8H, -C₁₆H₉), 6.64 (s, br, 2H, N²H), 6.54–6.40 (m, 1H, C1'-H), 4.81–4.72 (m, 1H, C3'-H), 4.65–4.53 (m, 1H, O⁶-CH₂), 4.49–4.29 (m, 1H, C5'-H), 4.12–3.81 (m, 3H, C4'-H, C5'-H), 3.67–3.54 (m, 1H, C2'-H), 2.40–2.34 (m, 1H, C2'-H), 1.99–1.87 (m, 3H, COCH₃), 1.27–1.13 (m, 2H, O⁶-CH₂-CH₂), 1.03–0.95 (m, 9H, C(CH₃)₃), 0.91–0.78 (m, 9H, C(CH₃)₃), 0.21 (s, 6H, Si(CH₃)₂), 0.15–0.05 (m, 9H, Si(CH₃)₃), 0.00–0.03 (m, 6H, Si(CH₃)₂) ppm.

¹³C-NMR (200 MHz, Acetone-*d*₆): δ = 159.1, 157.0, 154.0, 149.1, 147.6, 139.5, 132.6, 132.1, 131.1, 129.9, 128.3, 128.3, 127.9, 127.3, 127.2, 127.0, 126.5, 126.4, 126.2, 125.7, 125.3, 125.2, 122.0,

113.5, 88.8, 85.9, 73.5, 70.9, 64.6, 63.8, 55.4, 39.3, 26.3, 26.2, 24.4, 18.8, 18.6, 18.0, -1.3, -4.3, -4.4, -5.3, -5.4 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₅H₆₅N₆O₅Si₃]⁺ ([M+H]⁺): 853.4319, found: 853.4320.

8-(N-Acetyl-2-aminobenzene)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-N²-isopropylphenoxyacetyl-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



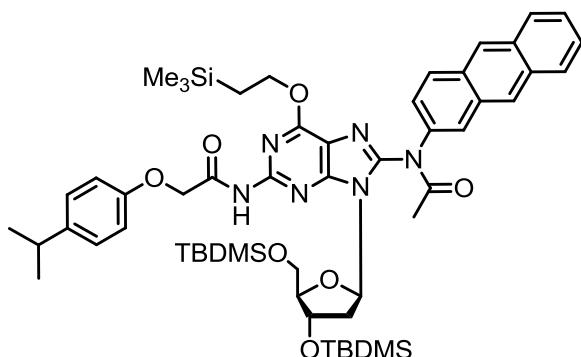
8-(N-Acetyl-2-aminobenzene)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.58 g, 2.17 mmol, 1.0 eq.) was dissolved in 50 mL dry pyridine, 4-isopropylphenoxyacetylchloride (1.50 mL, 8.67 mmol, 4.0 eq.) was added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (isohexane/EtOAc 1:0 to 7:1). 8-(N-Acetyl-2-aminobenzene)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-N²-isopropylphenoxyacetyl-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.70 g, 1.88 mmol, 86%) was obtained as a white yellowish foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.29 (s, br, 1H, N²H), 7.55–7.33 (m, 5H, -C₆H₅), 7.21 (d, 2H, iPrPac), 6.98 (d, 2H, iPrPac), 6.27 (t, 1H, C1'-H), 4.98 (s, 1H, C3'-H), 4.83 (s, 2H, iPrPac-CH₂), 4.68 (t, 2H, O⁶-CH₂), 3.99 (dd, 1H, C5'-H), 3.92–3.87 (m, 1H, C4'-H), 3.82 (dd, 1H, C5'-H), 3.44 (dt, 1H, C2'-H), 2.88 (hept, 1H, CH), 2.12 (s, 3H, COCH₃), 2.09–2.05 (m, 1H, C2'-H), 1.31–1.24 (m, 2H, O⁶-CH₂CH₂), 1.21 (d, 6H, CH(CH₃)₂), 0.92 (s, 9H, SiC(CH₃)₃), 0.81 (s, 9H, SiC(CH₃)₃), 0.16 (s, 3H, SiCH₃), 0.13 (s, 3H, SiCH₃), 0.11 (s, 9H, Si(CH₃)₃), -0.05 (s, 3H, SiCH₃), -0.05 (s, 3H, SiCH₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 166.7, 161.8, 156.9, 153.1, 152.33, 147.2, 142.9, 130.2, 130.1, 128.8, 128.3, 128.2, 127.4, 117.8, 115.6, 113.6, 89.2, 73.4, 68.9, 66.2, 64.5, 55.5, 37.9, 34.1, 26.3, 26.2, 24.5, 18.8, 18.5, 18.0, -1.3, -4.5, -4.5, -5.1, -5.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₆H₇₃N₆O₇Si₃]⁺ ([M+H]⁺): 905.4843, found: 905.4851.

8-(N-Acetyl-2-aminoanthracene)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-N²-isopropylphenoxyacetyl-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



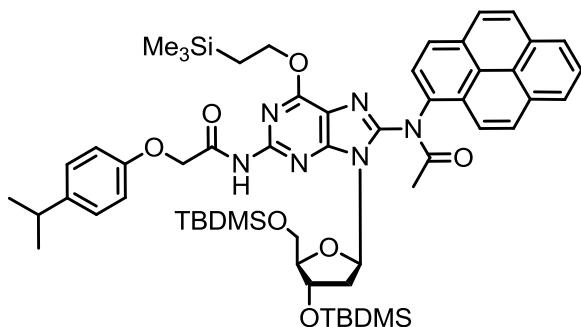
Yield: 70%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.30 (s, br, 1H, *N*²H), 8.58 (s, 1H, -C₁₄H₉), 8.51 (s, 1H, -C₁₄H₉), 8.15 (d, 1H, -C₁₄H₉), 8.07 (ddd, 3H, -C₁₄H₉), 7.67 (dd, 1H, -C₁₄H₉), 7.52 (dq, 2H, -C₁₄H₉), 7.20 (d, 2H, iPrPac), 6.97 (d, 2H, iPrPac), 6.36 (t, 1H, C1'-H), 4.97 (s, 1H, C3'-H), 4.83 (s, 2H, iPrPac-CH₂), 4.70 (t, 2H, O⁶-CH₂), 3.97 (dd, 1H, C5'-H), 3.92-3.88 (m, 1H, C4'-H), 3.81 (d, 1H, C5'-H), 3.46 (dt, 1H, C2'-H), 2.87 (hept, 1H, CH), 2.22 (s, 3H, OCH₃), 2.09-2.06 (m, 1H, C2'-H), 1.27 (t, 2H, O⁶-CH₂CH₂), 1.20 (d, 6H, CH(CH₃)₂), 0.89 (s, 9H, C(CH₃)₃), 0.78 (s, 9H, C(CH₃)₃), 0.13 (s, 3H, SiCH₃), 0.12 (s, 9H, Si(CH₃)₃), 0.07 (s, 3H, SiCH₃), -0.08 (s, 3H, SiCH₃), -0.08 (s, 3H, SiCH₃) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 166.7, 161.8, 159.3, 156.9, 153.2, 152.4, 147.2, 142.9, 138.1, 133.1, 132.2, 131.2, 130.4, 129.8, 129.1, 128.9, 128.4, 128.2, 127.5, 127.2, 126.8, 126.7, 117.9, 115.6, 113.7, 89.2, 85.3, 73.3, 68.9, 66.3, 64.4, 37.9, 34.1, 26.2, 26.2, 24.5, 23.3, 18.8, 18.5, 18.1, -1.3, -4.5, -4.5, -5.1, -5.2 ppm.

HR-MS (ESI⁺): m/z calculated for [C₅₄H₇₇N₆O₇Si₃]⁺ ([M+H]⁺): 1005.5156, found: 1005.5158.

8-(N-Acetyl-2-aminopyrene)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-N²-isopropylphenoxyacetyl-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine



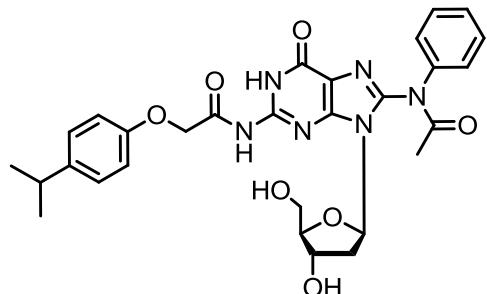
Yield: 92%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.26 (s, 1H, *N*²H), 8.46-8.02 (m, 9H, -C₁₆H₉), 7.24-7.15 (d, 2H, iPrPac), 6.97 (d, 2H, iPrPac), 6.68-6.52 (m, 1H, C1'-H), 5.25-4.92 (m, 1H, C3'-H), 4.82 (s, 2H, iPrPac-CH₂), 4.72-4.39 (m, 2H, O⁶-CH₂), 4.13-3.97 (m, 2H, C4'-H, C5'-H), 3.97-3.82 (m, 1H, C5'-H), 3.65-3.46 (m, 1H, C2'-H), 2.86 (hept, 1H, CH), 2.81 (s, 3H, COCH₃), 2.79-2.76 (m, 1H, C2'-H), 1.21 (s, 3H, CH(CH₃)₂), 1.19 (s, 3H, CHCH₃)₂), 1.15-1.11 (m, 2H, O⁶-CH₂CH₂), 0.94 (s, 9H, C(CH₃)₃), 0.82 (s, 9H, C(CH₃)₃), 0.16 (m, 6H, Si(CH₃)₂), 0.04 (s, 9H, Si(CH₃)₃), -0.04 (s, 6H, Si(CH₃)₂) ppm.

¹³C-NMR (100 MHz, Acetone-*d*₆): δ = 166.7, 156.9, 153.2, 142.9, 132.1, 128.2, 128.0, 127.6, 127.0, 126.2, 115.7, 115.6, 89.4, 85.6, 73.5, 68.9, 66.1, 34.0, 26.3, 26.2, 24.5, 21.4, 18.9, 18.5, 18.0, -1.3, -4.4, -5.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₅₆H₇₇N₆O₇Si₃]⁺ [M+H]⁺: 1029.5156, found: 1029.5156.

8-(*N*-Acetyl-2-aminobenzene)-N²-isopropylphenoxyacetyl-2'-deoxyguanosine



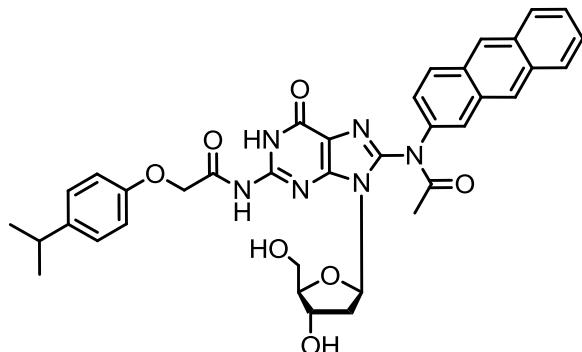
8-(*N*-Acetyl-2-aminobenzene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-N²-isopropylphenoxyacetyl-O⁶-(2-(trimethylsilyl)-ethyl)-2'-deoxyguanosine (1.10 g, 1.21 mmol, 1.0 eq.) was dissolved in dry THF (60 mL), 1 M TBAF in THF (4.86 mL, 4.0 eq.) and acetic acid (0.42 mL, 7.3 mmol, 6.0 eq.) were added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 40:1). 8-(*N*-Acetyl-2-aminobenzene)-N²-isopropylphenoxyacetyl-2'-deoxyguanosine (0.55 g, 0.95 mmol, 78%) was obtained as a yellowish foam.

¹H-NMR (400 MHz, DMSO-*d*₆): δ = 11.79 (s, br, 1H, N¹H), 11.38 (s, br, 1H, N²H), 7.64 (dd, 2H, C₆H₅), 7.29 (dd, 2H, -C₆H₅), 7.18 (d, 2H, iPrPac), 6.94 (t, 1H, -C₆H₅), 6.91 (d, 2H, iPrPac), 6.38 (t, 1H, C1'-H), 5.39 (d, br, 1H, C3'-OH), 4.84 (s, 2H, iPrPac-CH₂), 4.60 (dd, 1H, C3'-H), 4.38 (dd, 1H, C5'-H), 4.20 (dd, 1H, C5'-H), 3.97 (dt, 1H, C4'-H), 3.12 (dt, 1H, C2'-H), 2.84 (hept, 1H, CH), 2.23 (ddd, 1H, C2'-H), 1.97 (s, 3 H, COCH₃), 1.17 (d, 6H, CH(CH₃)₂) ppm.

¹³C-NMR (101 MHz, DMSO-*d*₆): δ = 170.2, 155.7, 153.8, 145.0, 141.3, 140.7, 127.2, 121.1, 117.6, 114.4, 83.9, 82.8, 70.6, 69.6, 66.5, 64.3, 37.1, 32.6, 24.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₂₉H₃₃N₆O₇]⁺ ([M+H]⁺): 577.2405, found: 577.2413.

8-(*N*-Acetyl-2-aminoanthracene)-N²-isopropylphenoxyacetyl-2'-deoxyguanosine



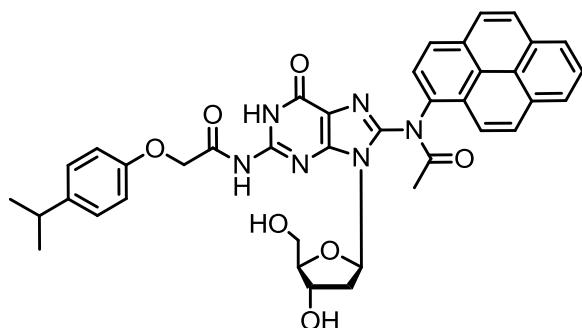
Yield: 76%

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 11.83 (s, *br*, 1H, *N*¹H), 10.07 (s, *br*, 1H, *N*²H), 8.58-8.52 (m, 2H, -C₁₄H₉), 8.15 (d, 1H, -C₁₄H₉), 8.10-8.04 (m, 3H, -C₁₄H₉), 7.67 (dd, 1H, -C₁₄H₉), 7.52 (dd, 2H, -C₁₄H₉), 7.21 (d, 2H, *iPrPac*), 7.01 (d, 2H, *iPrPac*), 6.39 (t, 1H, C1'-H), 5.97 (s, *br*, 1H, C3'-OH), 4.86 (s, 2H, *iPrPac*-CH₂), 4.61 (s, *br*, 1H, C5'-OH), 4.43-4.35 (m, 1H, C3'-H), 4.08-4.01 (m, 1H, C4'-H), 3.79 (d, 1H, C5'-H), 3.74 (s, 1H, C5'-H), 3.28-3.18 (m, 1H, C2'-H), 2.91-2.84 (m, 1H, CH), 2.20 (s, 3H, OCH₃), 2.10-2.05 (m, 1H, C2'-H), 1.20 (d, 6H, CH(CH₃)₂) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 172.0, 159.5, 156.4, 155.1, 143.4, 138.3, 133.1, 132.2, 130.5, 129.5, 129.1, 128.9, 128.8, 128.3, 128.1, 127.5, 127.2, 126.8, 126.8, 121.1, 115.8, 89.7, 86.5, 72.8, 71.1, 67.9, 65.8, 63.6, 39.1, 34.1, 24.5, 23.2 ppm.

HR-MS (ESI⁺): m/z calculated for [C₃₇H₃₇N₆O₇]⁺ ([M+H]⁺): 677.2718, found: 677.2726.

8-(*N*-Acetyl-2-aminopyrene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



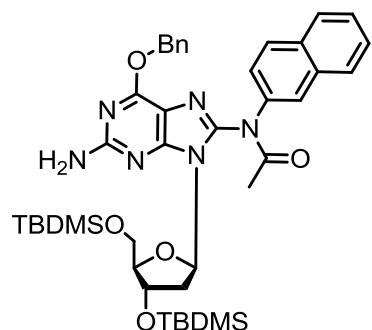
Yield 90%

¹H-NMR (400 MHz, DMSO-*d*₆): δ = 9.75 (s, *br*, 1H, *N*¹H), 9.08 (*br s*, 1H, *N*²H), 8.95-8.50 (m, 9H, -C₁₆H₉), 7.63 (d, 2H, *iPrPac*), 7.43 (d, 2H, *iPrPac*), 7.20-7.00 (m, 1H, C1'-H), 5.60-5.44 (m, 1H, C3'-H), 5.33 (s, 2H, *iPrPac*-CH₂), 5.09-4.98 (m, 1H, C3'-OH), 4.69-3.96 (m, 3H, C4'-H, C5'-H), 3.75-3.55 (m, 1H, C2'-H), 2.99-2.83 (m, 1H, CH), 2.41 (s, 3H, COCH₃), 2.30-2.16 (m, 1H, C2'-H), 1.64 (s, 3H, CH(CH₃)₂), 1.63 ppm (s, 3H, CH(CH₃)₂) ppm.

¹³C NMR (100 MHz, DMSO-*d*₆): δ = 166.7, 152.4, 141.7, 137.8, 137.8, 137.7, 136.0, 134.7, 125.3, 125.3, 78.4, 77.5, 76.1, 69.1, 43.7, 40.1, 39.9, 39.7, 39.5, 39.3, 39.2, 38.9, 34.2, 34.1, 34.1, 30.2, 30.1, 27.7, 23.5, 8.3 ppm.

HR-MS (ESI⁻): m/z calculated for [C₃₉H₃₅N₆O₇]⁻ ([M-H]⁻): 699.2573, found: 699.2579.

8-(*N*-Acetyl-2-aminonaphthalene)-3',5'-bis-*O*-(*tert*-butyl-dimethylsilyl)-*O*⁶-benzyl-2'-deoxyguanosine



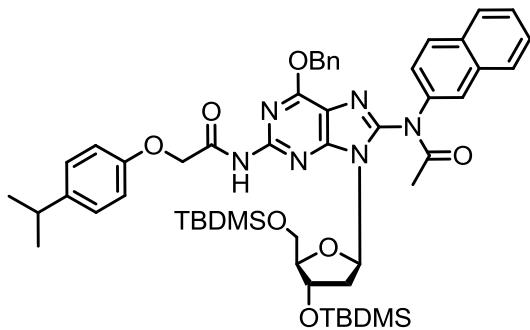
8-(2-aminonaphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-(4,4'-dimethoxytrityl)-O⁶-benzyl-ethyl)-2'-deoxyguanosine^[14] (6.47 g, 6.28 mmol, 1.0 eq.) was dissolved in 120 mL dry pyridine. 4-Dimethylaminopyridine (0.38 g, 3.14 mmol, 0.5 eq.), trimethylamine (4.38 mL, 31.4 mmol, 5.0 eq.) and acetic anhydride (2.97 mL, 31.4 mmol, 5.0 eq.) were added and the reaction mixture was stirred at room temperature for 48 h under argon atmosphere. After evaporation to dryness the residue was dissolved in EtOAc (200 mL) and washed with water (2 x 60 mL). The organic phase was dried with MgSO₄, filtered, concentrated under reduced pressure and dried for 3 h under high vacuum. Methanolic hydrochloric acid (30 mL, 0.01 M) was added and the mixture was stirred at r.t. for 20 h. After evaporation of the solvent, the residue was purified via column chromatography (*i*sohexane/EtOAc 1:0 to 7:1; SiO₂). 8-(*N*-Acetyl-2-aminonaphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-O⁶-benzyl-2'-deoxyguanosine (3.99 g, 5.19 mmol, 83%) was obtained as a white brown-yellowish foam.

¹H-NMR (400 MHz, Acetone-*d*₆): δ = 7.95-7.84 (m, 4H, -C₁₀H₇), 7.65-7.47 (m, 5H, 3x -C₁₀H₇, 2x Bn), 7.42-7.28 (m, 3H, Bn), 6.21 (s, 1H, C1'-H), 5.84 (s, *br*, 1H, N²H), 5.53 (s, *br*, 2H, Bn-CH₂), 4.81 (s, 1H, C3'-H), 3.94-3.82 (m, 2H, C4'-H, C5'-H), 3.80-3.70 (m, 1H, C5'-H), 3.46 (dd, 2H, C2'-H), 2.14 (s, 3H, COCH₃), 0.91 (s, 9H, C(CH₃)₃), 0.82 (s, 9H, C(CH₃)₃), 0.17-0.08 (m, 6H, Si(CH₃)₂), -0.01(-0.04) (m, 6H, Si(CH₃)₂) ppm.

¹³C-NMR (101 MHz, Acetone-*d*₆): δ = 161.71, 160.64, 154.74, 137.72, 134.40, 129.44, 129.24, 128.89, 128.83, 128.49, 127.42, 114.23, 88.50, 84.96, 73.43, 68.29, 63.84, 41.83, 37.08, 26.23, 26.18, 20.83, 18.79, 18.51, -4.45, -5.21 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₁H₅₇N₆O₅Si₂]⁺ ([M+H]⁺): 769.3851, found.: 769.39207.

8-(*N*-Acetyl-2-naphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-O⁶-benzyl-2'-deoxyguanosine



8-(*N*-Acetyl-2-aminonaphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-O⁶-benzyl-2'-deoxyguanosine (3.99 g, 5.19 mmol, 1.0 eq.) was dissolved in 120 mL dry pyridine, 4-isopropylphenoxyacetylchloride (3.6 mL, 20.7 mmol, 4.0 eq.) was added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (*i*sohexane/EtOAc 1:0 to 6:1). 8-(*N*-Acetyl-2-naphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-*N*²-isopropylphenoxyacetyl-O⁶-benzyl-2'-deoxyguanosine (4.54 g, 4.80 mmol, 92%) was obtained as a white yellowish foam.

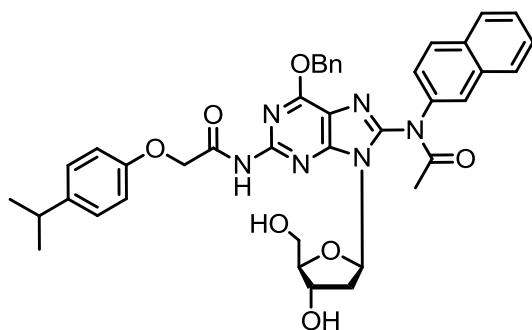
¹H-NMR (400 MHz, Acetone-*d*₆): δ = 9.44 (s, 1H, N²H), 8.00-7.85 (m, 4H, -C₁₀H₇), 7.68-7.50 (m, 5H, 3 x -C₁₀H₇, 2 x Bn), 7.41-7.29 (m, 3H, Bn), 7.23-7.18 (m, 2H, iPrPac), 7.01-6.95 (m, 2H, iPrPac), 6.40-6.31 (m, *br*, 1H, C1'-H), 5.68-5.55 (m, 2H, Bn-CH₂), 5.03-4.94 (m, 1H, C3'-H), 4.84 (s, 2H, iPrPac-CH₂), 4.00-3.78 (m, *br*, 3H, 2x C5'-H, C4'-H), 3.43-3.31 (m, 2H, C2'-H₂), 2.93-2.76 (m, 1H, CH), 2.17

(s, 3H, COCH₃), 1.21 (d, 3H, CH(CH₃)₂), 1.16 (d, 3H, CH(CH₃)₂), 0.91 (s, 9H, C(CH₃)₃), 0.78 (s, 9H, C(CH₃)₃), 0.19-0.05 (m, 6H, Si(CH₃)₂), -0.07 (s, 6H, Si(CH₃)₂) ppm.

¹³C-NMR (101 MHz, Acetone-d₆) δ = 170.46, 166.74, 161.40, 156.95, 153.34, 152.22, 147.50, 142.91, 137.21, 134.44, 129.61, 129.33, 129.11, 128.90, 128.56, 128.20, 117.73, 115.64, 89.21, 85.31, 73.27, 68.91, 64.41, 58.97, 37.96, 34.05, 26.23, 26.20, 24.50, 21.10, 18.80, 18.45, -4.47, -5.15 ppm.

HR-MS (ESI⁺): m/z calculated for [C₅₂H₆₇N₆O₇Si₂]⁺ ([M-H]⁺): 943.4688, found: 943.46036.

8-(N-Acetyl-2-aminonaphthalene)-N²-isopropylphenoxyacetyl-O⁶-benzyl-2'-deoxyguanosine



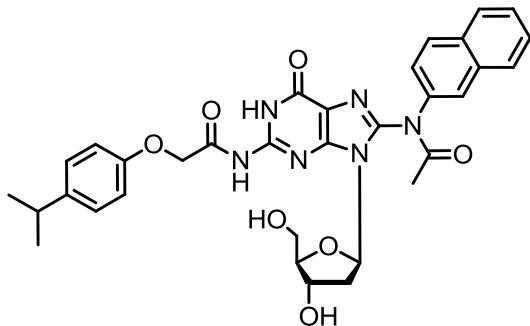
8-(N-Acetyl-2-naphthalene)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-N²-isopropylphenoxyacetyl-O⁶-benzyl-2'-deoxyguanosine (4.50 g, 4.76 mmol, 1.0 eq.) was dissolved in dry THF (120 mL), 1 M TBAF in THF (19.0 mL, 4.0 eq.) and acetic acid (1.63 mL, 28.6 mmol, 6.0 eq.) were added and the reaction mixture was stirred at room temperature for 24 h under argon atmosphere. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 40:1). 8-(N-Acetyl-2-aminonaphthalene)-N²-isopropylphenoxyacetyl-O⁶-benzyl-2'-deoxyguanosine (3.10 g, 4.32 mmol, 91%) was obtained as a yellowish foam.

¹H-NMR (400 MHz, Aceton-d₆): δ = 9.44 (s, 1H, N²-H), 8.03-7.87 (m, 4H, -C₁₀H₇), 7.71-7.50 (m, 5H, 3x -C₁₀H₇, 2x Bn), 7.41-7.30 (m, 3H, Bn), 7.22-7.14 (m, 2H, iPrPac), 7.02-6.96 (m, 2H, iPrPac), 6.43 (t, 1H, C1'-H), 5.69-5.57 (s, 2H, Bn-CH₂), 4.88 (s, 2H, iPrPac-CH₂), 4.80 (s, br, 1H, C3'-H) 4.04 (s, br, 1H, C4'-H), 3.90-3.71 (m, 2H, C5'-H), 3.30-2.90 (m, 2H, C2'-H) 2.86 (hept, 1H, CH), 1.96 (s, 3H, COCH₃), 1.20 (d, 6H, CH(CH₃)₂) ppm.

¹³C-NMR (101 MHz, Acetone-d₆): δ = 172.01, 167.33, 161.47, 156.89, 153.05, 152.11, 147.28, 142.80, 137.10, 134.41, 129.77, 129.30, 129.13, 128.92, 128.55, 128.15, 127.63, 117.75, 115.63, 89.85, 86.39, 72.60, 69.50, 68.71, 63.54, 39.13, 34.02, 24.49, 20.50 ppm.

HR-MS (ESI⁺): m/z calculated for [C₄₀H₄₁N₆O₆]⁺ ([M+H]⁺): 717.2958, found: 717.30365.

8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-*O*⁶-benzyl-2'-deoxyguanosine (0.50 g, 0.69 mmol, 1.0 eq.) was dissolved in 25 mL MeOH/THF 1:1, Pd/C (0.05g) was added and argon was exchanged with hydrogen atmosphere. The reaction mixture was stirred at room temperature for 20 h. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 20:1). 8-(*N*-Acetyl-2-aminonaphthalene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (0.40 g, 0.63 mmol, 90%) was obtained as a white solid.

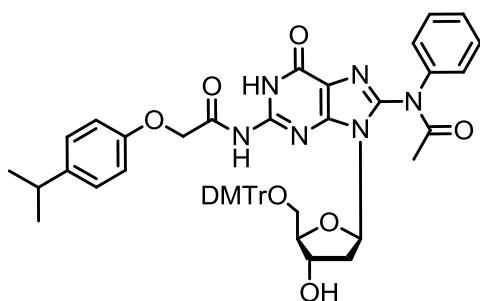
¹H-NMR (400 MHz, DMSO-*d*₆): δ = 11.92 (s, br, 2H, *N*¹-H, *N*²-H), 8.00 (d, 1H, -C₆H₅), 7.97-7.91 (m, 3H, -C₆H₅), 7.63-7.50 (m, 2H, -C₆H₅), 7.45-7.33 (m, 1H, -C₆H₅), 7.16 (d, 2H, iPrPac), 6.88 (d, 2H, iPrPac), 6.31-6.13 (m, 1H, C1'-H), 5.24 (s, br, 1H, C3'-OH), 4.93 (s, br, 1H, C5'-OH), 4.81 (s, 2H, iPrPac-CH₂), 4.51-4.40 (m, 1H, C3'-H), 3.90-3.80 (m, 1H, C4'-H), 3.68-3.61 (m, 1H, C5'-H), 3.56-3.47 (m, 1H, C5'-H), 3.36-3.29 (m, 2H, C2'-H), 2.80 (sept, 1H, CH), 2.10 (s, 3H, COCH₃), 1.17 (d, 3H, CH(CH₃)₂), 1.15 (d, 3H, CH(CH₃)₂) ppm.

¹³C-NMR (100 MHz, DMSO-*d*₆): δ = 169.3, 155.9, 149.6, 141.1, 136.1, 132.9, 129.0, 127.9, 127.6, 127.1, 127.1, 126.8, 124.9, 123.9, 118.8, 114.4, 114.2, 88.1, 84.0, 70.9, 67.0, 64.6, 62.0, 51.7, 36.9, 34.4, 32.6, 30.4, 24.1, 22.7 ppm.

HR-MS (ESI⁻): m/z calculated for [C₃₃H₃₂N₆O₇]⁻ ([M-H]⁻): 626.6591, found: 625.2417.

The DMT protection and the preparation of the phosphoramidites of the phenyl, naphthyl, pyrenyl and the anthryl adducts were carried out analogous.

8-(*N*-Acetyl-2-aminobenzene)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



8-(*N*-Acetyl-2-aminobenzene)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (200 mg, 0.347 mmol, 1.0 eq.) was evaporated two times from dry pyridine and then dissolved in 2 mL dry pyridine. Molecular sieves (4 Å) were added and the solution was stirred at room temperature for 2 h under argon atmosphere. 4,4'-dimethoxytritylchloride (153 mg,

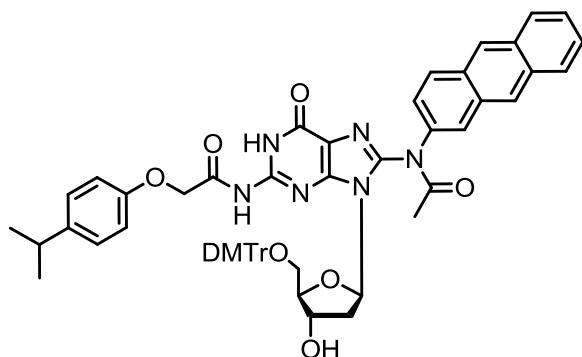
0.451 mmol, 1.3 eq.) was added and the reaction mixture was stirred at room temperature for 3 h. After evaporation to dryness the residue was purified via column chromatography on silica gel (DCM/MeOH 1:0 to 10:1 + 1% pyridine). 8-(N-Acetyl-2-aminobenzene)-5'-O-(4,4'-dimethoxytrityl)-N²-isopropylphenoxyacetyl-2'-deoxyguanosine (274 mg, 0.312 mmol, 90%) was obtained as a yellow foam.

¹H-NMR (599 MHz, DMSO-d₆): δ = 11.74 (s, 1H, N¹-H), 11.31 (s, 1H, N²-H), 7.44-7.36 (m, 4H, -C₆H₅), 7.33-7.29 (m, 1H, -C₆H₅), 7.27-7.22 (m, 2H, DMTr), 7.16-7.14 (m, 2H, iPrPac), 7.14-7.02 (m, 7H, DMTr), 6.86 (d, 2H, iPrPac), 6.72-6.65 (m, 2H, DMTr), 6.66-6.54 (m, 2H, DMTr), 6.24 (s, 1H, C1'-H), 5.17 (d, 1H, C3'-OH), 4.82 (d, 1H, CH₂), 4.70 (d, 1H, iPrPacCH₂), 4.59 (m, 1H, C3'-H), 4.00 (s, 1H, C4'-H), 3.65 (s, 3H, OCH₃), 3.62 (s, 3H, OCH₃), 3.37 (s, br, 1H, C5'-H), 3.10 (d, 1H, C5'-H), 2.81 (hept, 1H, CH), 2.04 (m, 4H, C2'-H, COCH₃), 1.14 (d, 6H, CH(CH₃)₂) ppm.

¹³C-NMR (151 MHz, DMSO-d₆): δ = 171.0, 170.5, 158.6, 158.3, 158.2, 156.1, 154.8, 147.8, 147.0, 145.3, 143.4, 141.9, 136.2, 136.0, 130.2, 130.1, 129.8, 128.2, 127.8, 127.7, 126.8, 119.6, 114.9, 113.2, 113.1, 87.2, 85.1, 84.2, 71.3, 67.0, 65.4, 55.3, 33.0, 24.5, 23.1 ppm.

HR-MS (ESI⁺): m/z calculated for [C₅₀H₅₁N₆O₉]⁺ ([M+H]⁺): 879.3712, found: 879.3723.

8-(N-Acetyl-2-aminoanthracene)-5'-O-(4,4'-dimethoxytrityl)-N²-isopropylphenoxyacetyl-2'-deoxyguanosine



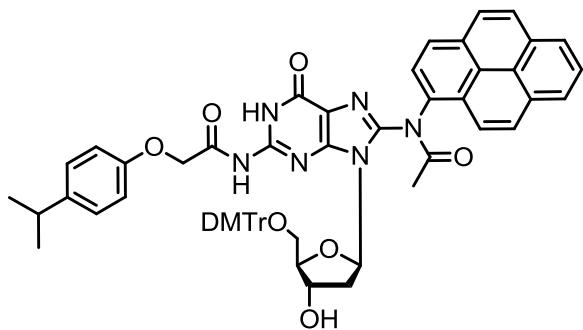
Yield: 64%

¹H-NMR (599 MHz, DMSO-d₆): δ = 11.72 (s, 1H, N¹H), 11.30 (s, 1H, N²H), 8.62-8.46 (m, 2H, -C₁₄H₉), 8.07 (dd, 2H, -C₁₄H₉), 7.78 (dd, 2H, -C₁₄H₉), 7.56 (s, 1H, -C₁₄H₉), 7.53-7.42 (m, 1H, -C₁₄H₉), 7.35 (m, 1H, -C₁₄H₉), 7.33-7.23 (m, 3H, DMTr), 7.18-7.17 (m, 2H, DMTr), 7.16-7.13 (m, 2H, iPrPac), 7.13-7.01 (m, 6H, DMTr), 6.85 (d, 2H, iPrPac), 6.74-6.47 (m, 2H, DMTr), 6.24 (s, 1H, C1'-H), 5.17 (s, 1H, C3'-OH), 4.81 (dd, 1H, CH₂), 4.69 (d, 1H, iPrPacCH₂), 4.55 (s, 1H, C3'-H), 4.10-4.00 (m, 2H, C5'-H, C4'-H), 3.87 (s, 1H, C5'-H), 3.64 (s, 6H, 2 x OCH₃), 3.18-3.08 (td, 1H, C2'-H), 2.80 (hept, 1H, CH), 2.20-2.00 (m, 4H, C2'-H, COCH₃), 1.14 (d, 6H, CH(CH₃)₂) ppm.

¹³C-NMR (151 MHz, DMSO-d₆): δ = 171.0, 158.2, 156.1, 154.8, 145, 141.9, 136.0, 131.9, 131.1, 130.1, 128.5, 128.2, 127.8, 127.7, 126.8, 126.4, 119.6, 114.9, 113.2, 113.0, 87.3, 85.6, 84.2, 71.3, 67.0, 65.5, 55.3, 55.2, 33.0, 24.5, 23.1 ppm.

HR-MS (ESI⁻): m/z calculated for [C₅₈H₅₃N₆O₉]⁻ ([M-H]⁻): 977.3879, found: 977.3882.

8-(*N*-Acetyl-2-aminopyrene)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine

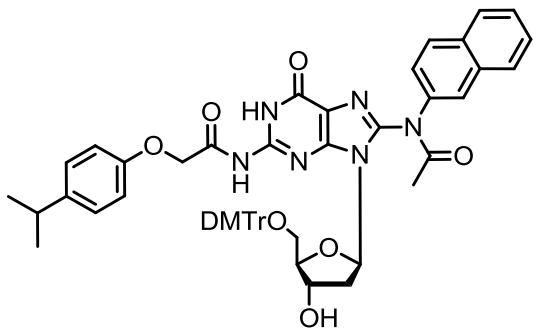


Yield: 84%

¹H-NMR (800 MHz, Acetone-*d*₆): δ = 11.65 (s, 1H, *N*¹H), 10.24 (s, 1H, *N*²H), 8.82-8.03 (m, 9H, -C₁₆H₉), 7.37-7.32 (m, 2H, DMTr), 7.33-7.23 (m, 5H, DMTr), 7.22-7.18 (m, 2H, iPrPac), 7.15-7.04 (m, 4H, DMTr), 6.99-6.90 (m, 2H, iPrPac), 6.76-6.65 (m, 2H, DMTr), 6.46-6.32 (m, 1H, C1'-H), 5.19 (s, 1H, C3'-OH), 4.86 (s, 1H, C3'-H), 4.82-4.72 (m, 2H, iPrPacCH₂), 4.68-4.26 (m, 3H, C4'-H, C5'-H), 3.69-3.60 (m, 6H, OCH₃), 3.29-3.22 (m, 1H, C2'-H), 2.83-2.76 (hept, 1H, CH), 2.00-1.95 (m, 1H, C2'-H), 1.29 (s, 3H, COCH₃), 1.20 ppm (s, 6H, 2x CH(CH₃)₂) ppm.

HR-MS (ESI⁻): m/z calculated for [C₆₀H₅₃N₆O₉]⁻ ([M-H]⁻): 1001.3879, found: 1001.3887.

8-(*N*-Acetyl-2-aminonaphthalene)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine



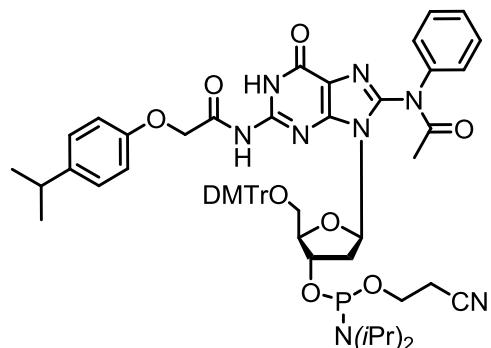
Yield: 65%

¹H NMR (599 MHz, DMSO-*d*₆) δ 11.73 (s, 1H, *N*¹H), 11.30 (s, 1H, *N*²H), 7.93 (d, 1H, -C₁₀H₇), 7.90 (d, 1H, -C₁₀H₇), 7.81 (s, 3H, -C₁₀H₇), 7.50 (t, 1H, -C₁₀H₇), 7.45 (s, 1H, -C₁₀H₇), 7.24 (d, 2H, DMTr), 7.15 (d, 2H, iPrPac), 7.13-6.92 (m, 7H, DMTr), 6.85 (d, 2H, iPrPac), 6.62 (d, 4H, DMTr), 6.31 (s, 1H, C1'-H), 5.15 (s, 1H, C3'-OH), 4.81 (d, 1H, CH₂), 4.69 (d, 1H, iPrPacCH₂), 4.54 (s, 1H, C3'-H), 4.02 (s, 1H, C4'-H), 3.63 (s, 3H, OCH₃), 3.61 (s, 3H, OCH₃), 3.37 (s, 1H, C5'-H), 3.12 (m, 1H, C5'-H), 3.10-3.00 (m, 1H, C2'-H), 2.93-2.75 (hept, 1H, CH) 2.20-2.06 (s, 4H, COCH₃, C2'-H), 1.14 (d, 6H, CH(CH₃)₂) ppm.

¹³C NMR (151 MHz, DMSO-*d*₆) δ 171.0, 158.2, 156.1, 154.8, 145.3, 141.9, 136.0, 133.4, 132.0, 130.2, 130.1, 128.2, 128.0, 127.8, 127.7, 127.2, 126.8, 119.8, 114.8, 113.2, 113.0, 87.3, 85.6, 84.3, 71.3, 67.0, 65.7, 55.3, 33.0, 24.5, 23.9 ppm.

HR-MS (ESI⁻): m/z calculated for [C₅₄H₅₁N₆O₉]⁻ ([M-H]⁻): 927.3723, found: 927.3724.

8-(*N*-Acetyl-2-aminobenzene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (1)

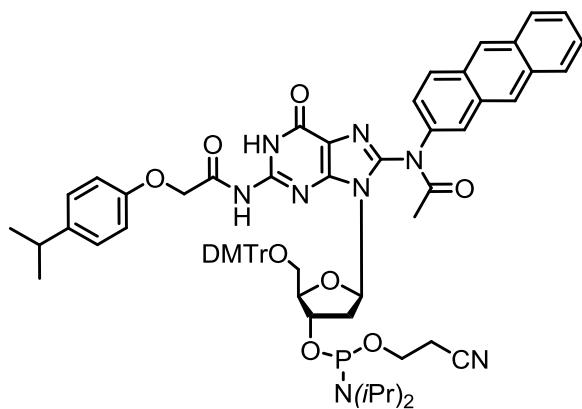


8-(*N*-Acetyl-2-aminobenzene)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (200 mg, 0.227 mmol, 1.0 eq.) and Diisopropylammoniumtetrazolate (144 mg, 0.83 mmol, 3.6 eq.) were dissolved in dry DCM (2 mL) and degassed three times. 3-((bis(diisopropylamino)phosphino)oxy)propanenitrile (146 μ L, 0.46 mmol, 2.0 eq.) was added and the mixture was stirred at room temperature for 2 h. 8-(*N*-Acetyl-2-aminobenzene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine **1** (245 mg, 0.226 mmol, 99%) was obtained as a yellow solid.

³¹P-NMR (162 MHz, CDCl₃): δ = 148.9 ppm.

HR-MS (ESI⁻): m/z calculated for [C₅₉H₆₆N₈O₁₀P]⁻ ([M-H]⁻): 1077.4645, found: 1077.4647.

8-(*N*-Acetyl-2-aminoanthracene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (3)

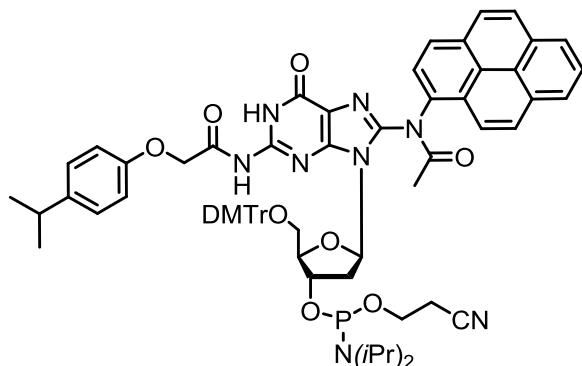


Yield: 99%

³¹P-NMR (162 MHz, CDCl₃): δ = 149.3 ppm.

HR-MS (ESI⁻): m/z calculated for [C₆₇H₇₀N₈O₁₀P]⁻ ([M-H]⁻): 1177.4958, found: 1177.4960.

8-(*N*-Acetyl-2-aminopyrene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (4)

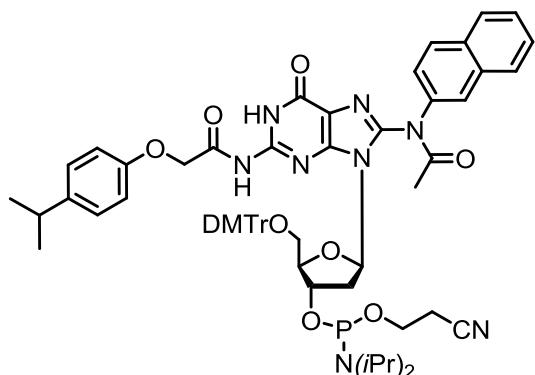


Yield: 99%

³¹P-NMR (162 MHz, CDCl₃): δ = 149.4, 148.8 ppm.

HR-MS (ESI⁻): m/z calculated for [C₆₉H₇₀N₈O₁₀P]⁻ ([M-H]⁻): 1201.4958, found: 1201.4951.

8-(*N*-Acetyl-2-aminonaphthalene)-3'-O-(2'-cyanoethoxydiisopropylamino-phosphino)-5'-O-(4,4'-dimethoxytrityl)-*N*²-isopropylphenoxyacetyl-2'-deoxyguanosine (2)



Yield: 99%

³¹P-NMR (162 MHz, CDCl₃): δ = 149.4, 149.0 ppm.

HR-MS (ESI⁻): m/z calculated for [C₆₃H₆₈N₈O₁₀P]⁻ ([M-H]⁻): 1127.4801, found: 1127.4803.

References

- [1] L. C. Gillet, J. Alzeer, O. D. Scharer, *Nucleic Acids Res.* **2005**, 33, 1961-1969.
- [2] S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 112, 8272-8277.
- [3] W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, 66, 133-144.
- [4] a) K. Diederichs, P. A. Karplus, *Acta Crystallogr. D Biol. Crystallogr.* **2013**, 69, 1215-1222; b) P. Evans, *Science* **2012**, 336, 986-987; c) P. A. Karplus, K. Diederichs, *Science* **2012**, 336, 1030-1033.

- [5] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, 67, 355-367.
- [6] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, 66, 486-501.
- [7] a) J. Painter, E. A. Merritt, *J. Appl. Crystallogr.* **2006**, 39, 109-111; b) M. D. Winn, G. N. Murshudov, M. Z. Papiz, *Methods Enzymol.* **2003**, 374, 300-321.
- [8] a) P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta Crystallogr. D Biol. Crystallogr.* **2010**, 66, 213-221; b) P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Crystallogr. D Biol. Crystallogr.* **2012**, 68, 352-367.
- [9] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, 66, 12-21.
- [10] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, 372, 774-797.
- [11] R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, K. Zakrzewska, *Nucleic Acids Res.* **2009**, 37, 5917-5929.
- [12] E. Krissinel, K. Henrick, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, 60, 2256-2268.
- [13] M. Munzel, C. Szeibert, A. F. Glas, D. Globisch, T. Carell, *J. Am. Chem. Soc.* **2011**, 133, 5186-5189.
- [14] a) L. C. Gillet, O. D. Scharer, *Org. Lett.* **2002**, 4, 4205-4208; b) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, 3, 565-568.

3.1.5.2 Zusammenfassung der Publikation zu C8-AF-dG und N²-AAN-dG

Structural insights into the recognition of N²-aryl- and C8-aryl DNA lesions by the repair protein XPA/Rad14.

C. Ebert*, N. Simon*, S. Schneider, T. Carell, *ChemBioChem* 2017, DOI: 10.1002/cbic.201700169. (*Beitrag zu gleichen Teilen)

Kopie der Publikation mit Erlaubnis des Verlags; Copyright 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Aromatische Amine sind stark karzinogene Verbindungen. Sie werden in der Leber zu reaktiven Nitreniumionen metabolisiert, die mit den Nukleobasen im DNA-Duplex reagieren können. Die Reaktion findet hauptsächlich an der C8-Position von dG statt, wodurch C8-Acetylaryl- oder C8-Aryl-dG Addukte in einer elektrophilen aromatischen Substitution entstehen. Außerdem wird die Reaktion mit der exozyklischen 2-NH₂-Gruppe beobachtet. Während C8-Addukte die Möglichkeit der *Watson-Crick*-Basenpaarung behalten, ist diese bei N²-Addukten stark beeinträchtigt. Hier zeigen wir die Kristallstrukturen der zwei DNA-Schäden N²-AAN-dG und C8-AF-dG in einem DNA-Duplex, welche durch das Reparaturenzym Rad14 erkannt werden. Die Strukturen bestätigen, dass zwei Rad14-Moleküle den Schaden durch Ausbildung eines 72° bzw. 78° Knicks an der geschädigten Stelle erkennen. Die strukturell unterschiedlichen Schäden führen zwar zu einer sehr ähnlichen Gesamtstruktur, im Bereich des Schadens kommt es jedoch zu verschiedenen π-stacking Interaktionen. Diese Beobachtungen lassen vermuten, dass das Reparaturenzym XPA/Rad14 ein Sensor ist, der die Flexibilität des DNA-Duplex testet und erkennt. Das Protein hat also die Fähigkeit strukturell voneinander abweichende Schäden in ein einheitliches stark geknicktes Erkennungsmotiv umzuwandeln.

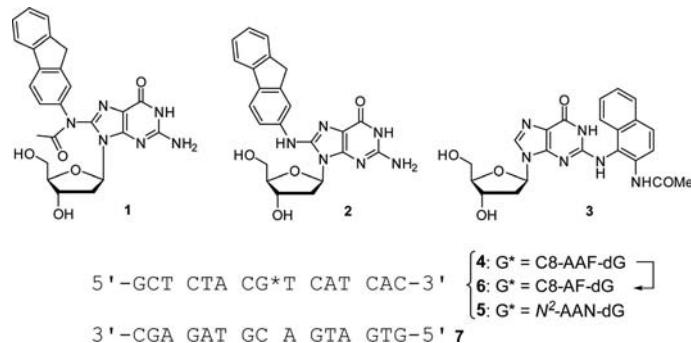
Beitragserklärung: Es wurde das Phosphoramidit von N²-AAN-dG nach *Bonala et al.* synthetisiert. In Zusammenarbeit mit *Nina Simon* wurden 15mer DNA-Stränge, die in der Mitte einen N²-AAN-dG oder einen C8-AAF-dG Schaden enthielten, synthetisiert. Die C8-AAF-dG Stränge wurden deacetyliert und alle Stränge wurden mit ihrem ungeschädigten Gegenstrang hybridisiert. Mit diesen Strängen wurden sowohl die EMSA-Bindungsstudien mit XPAfl als auch die Co-Kristallisationsexperimente mit Rad14 durchgeführt. Die Auswertung der Kristallstrukturen wurde von *Sabine Schneider* übernommen.

Structural Insights into the Recognition of N^2 -Aryl- and C8-Aryl DNA Lesions by the Repair Protein XPA/Rad14

Charlotte Ebert⁺^[a] Nina Simon⁺^[a] Sabine Schneider,^{*[b]} and Thomas Carell^{*[a]}

Aromatic amines are strongly carcinogenic. They are activated in the liver to give reactive nitrenium ions that react with nucleobases within the DNA duplex. The reaction occurs predominantly at the C8 position of the dG base, thereby giving C8-acetyl-aryl- or C8-aryl-dG adducts in an electrophilic aromatic substitution reaction. Alternatively, reaction with the exocyclic 2-NH₂ group is observed. Although the C8 adducts retain base-pairing properties, base pairing is strongly compromised in the case of the N^2 adducts. Here we show crystal structures of two DNA lesions, N^2 -acetylnaphthyl-dG and C8-fluorenyl-dG, within a DNA duplex recognized by the repair protein Rad14. The structures confirm that two molecules of the repair protein recognize the lesion and induce a 72 or 78° kink at the site of the damage. Importantly, the same overall kinked structure is induced by binding of the repair proteins, although the structurally different lesions result in distinct stacking interactions of the lesions within the duplex. The results suggest that the repair protein XPA/Rad14 is a sensor that recognizes flexibility. The protein converts the information that structurally different lesions are present in the duplex into a unifying sharply kinked recognition motif.

Aromatic amines are known to be strongly carcinogenic.^[1] The mode of action involves oxidation of the amines in the liver followed by their conversion into mono- or *N,N*-bis-acetyl compounds. After elimination of one acetyl group a reactive nitrenium ion compound is formed; this reacts with purine bases, mainly deoxyguanosine (dG), in DNA to give either C8 or N^2 bulky adducts (Scheme 1).^[1e,2] These bulky adducts are repaired by the nucleotide excision repair (NER) system.^[3] Lesion recognition is generally believed to proceed through binding of the NER protein XPC.^[4] Lesion binding was also reported for XPA, although it is assumed that XPA acts as a scaffold protein that is important for the assembly of the repair factors at the damaged site.^[3i,5] Patients with mutated XPA exhibit the strongest



Scheme 1. C8-AF-dG, N^2 -AAN-dG, and C8-AAF-dG lesions studied in complexes with Rad14; duplex used for crystallization (G^* = lesion).

xeroderma pigmentosum phenotype, thus showing that the platform protein XPA is of paramount importance for efficient repair of bulky adduct lesions.^[6]

We recently crystallized yeast Rad14 (homologue of XPA) in complex with DNA containing C8-acetylaminaryl bulky lesions (1).^[7] Because the 3D structures of XPA and Rad14 are very similar, structural insights gained with Rad14 can be directly related to lesion recognition in the human NER system.^[5d,e] The Rad14 structures clarified that Rad14 binds (2:1) to the lesion-containing DNA and that upon binding a sharp DNA kink of about 70° is induced. Because the NER process requires the assembly of a large number of proteins at the damage site to excise a stretch of 20–24 nucleobases around the lesion,^[3i,j,8] we speculated that the kink is a major recognition signal that enables recruitment of the proteins involved in the repair reaction.

Although structures of various acetylated C8-acetylaryl-dG-lesions such as 1 have been solved,^[7] the question remains as to how XPA/Rad14 recognizes other nonacetylated lesions, like 2 and particularly N^2 bulky adducts such as 3 (Scheme 1). Here we obtained crystal structures of the DNA-binding domain of Rad14 in complex with DNA duplexes containing the non-acetylated C8-aminofluorene lesion (2, C8-AF-dG) and the N^2 -acetylnaphthyl adduct (3, N^2 -AAN-dG). We show that the ultimate recognition mode (2:1 binding) and the induction of a large kink at the lesioned site is preserved, despite distinct stacking interactions of the lesions in the duplex within the (Rad14)₂:DNA complex.

First we synthesized phosphoramidites of the C8-AAF-dG and of the N^2 -AAN-dG adducts according to published procedures.^[7b,g] We used phosphoramidite-based solid phase DNA synthesis to prepare oligonucleotides 4 and 5 (Scheme 1).^[10] Oligonucleotide 6 was obtained by deacetylation of 4.^[11] 5 and 6 were subsequently annealed with the complementary strand

[a] C. Ebert,⁺ N. Simon,⁺ Prof. Dr. T. Carell

Center for Integrated Protein Science at the Department of Chemistry
Ludwig-Maximilians-Universität München
Butenandtstrasse 5–13, 81377 München (Germany)
E-mail: thomas.carell@lmu.de

[b] Dr. S. Schneider

Center for Integrated Protein Science at the Department of Chemistry
Technische Universität München
Lichtenbergstrasse 4, 85748 Garching (Germany)
E-mail: sabine.schneider@tum.de

[+] These authors contributed equally to this work.

Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201700169>.

7 to give duplexes. These were crystallized with a truncated Rad14 (Rad14_{188–302}, the minimal DNA-binding domain).^[7a] In 2014, Sugitani and others reported that DNA-binding strength is increased if the binding domain is extended at the C terminus by 20 amino acids (for XPA).^[12] Although this extended version gives better DNA binding, our shorter constructs crystallized well and afforded insight into the binding mode. We noticed, however, that in comparison to crystals with oligonucleotides containing lesion 1,^[7a] the crystals for this study grew significantly more slowly. In the case of Rad14_{188–302} and duplex 5:7, we had to wait three months before crystals were obtained, and two months for duplex 6:7, thus indicating that the system needs time to find the thermodynamically most favorable situation. We noted no protein or DNA degradation during this time (Supporting Information).

We first prepared 15-mer duplexes with C8-AF-dG and N²-AAN-dG (Table S1 in the Supporting Information) and confirmed binding with an electrophoretic mobility shift assay (EMSA). Lesion-containing duplexes were nicely recognized by XPA as expected (Figure 1 A and B). Next we crystallized duplexes 5:7 and 6:7 with Rad14_{188–302} and obtained crystals that diffracted to 2.2 Å (5:7, C8-AF-dG lesion; in 0.22 M ammonium nitrate and 38% MPD) and to 1.9 Å (6:7, N²-AAN-dG lesion; in 0.02 M CaCl₂, 0.1 M NaOAc pH 4.6 and 30% MPD). As observed in all Rad14–DNA complex structures determined so far, the DNA bound (1:1) in two orientations, with a pseudo-C₂-symmetry axis running directly through the kink (Figure 1 C and D). Thus, the lesion-containing strand is on one side of the

duplex in 50% of cases, and on the other side in the other 50%. For clarity, we show only one DNA duplex with the lesion located on one side.

Figure 1 shows the complexes between Rad14 and DNA containing C8-AF-dG (Figure 1C) or N²-AAN-dG (Figure 1D). They are, as expected, very similar. In both cases two Rad14 molecules binding to the lesion (one above, one below), and this induces a sharp kink of the duplex at the lesion site (> 70°). The DNA is kinked into the major groove, and the protein makes no direct interaction with the lesion. Despite the different lesions, the sharply kinked DNA motif induced by binding of two Rad14 molecules is similar, and we postulate that this motif induces the assembly of the full set of repair enzymes at the damaged site. This could explain why XPA/Rad14 is able to function as a scaffold protein in NER and how lesions with quite different structures induce the same NER process.

The fact that both damaged DNAs lead to the same overall structure is quite interesting if we compare the chemical structures of the C8-AF and N²-AAN lesions. In principle, C8-AF can still base pair like dG; this is however impossible for the N²-AAN-dG lesion, in which the aromatic unit fully blocks the Watson–Crick base pairing side.

In the case of the C8-AF-dG structure (Figures 2 and S1), the lesion adopts a *syn* conformation around the nucleosidic bond. The Watson–Crick face of the dG base points into the major groove, which is strongly compressed due to the 78° kink into this groove. The dG part of the lesion stacks effectively with the dT base on the 3' side (3.2 Å), whereas stacking with the 5'

dC is completely abolished. The fluorene ring system is not in plane with the dG part: the two planes are oriented at 50°. The fluorene ring points out into the minor groove, which is strongly widened due to the sharp kink. Stacking interactions are established with the dA base of the complementary strand (3.9 Å). There is no stacking interaction with the dC:dG base pair below. In fact the fluorene unit stands basically in an edge-to-face orientation on this base pair (Figure 2B). This dC:dG base pair is consequently not flat but buckled (V-type orientation, an angle of 35°). But, the dC:dG base pair is clearly formed with three H-bonds (Figure 2A); this is not the case for the dT:dA base pair above (Figure 2B). Although the dA base is in place with the Watson–Crick side pointing towards the helical axis, the dT base 3' to the lesion is shifted out of the helix, almost flipped out into the compressed major groove. This maximizes the stacking interactions with the dG part of the lesion. Again, in the duplex, optimizing stacking interactions is energetically more important than establishing H-bonding to the complementary base. The location of the dC opposite the lesion is noteworthy. Instead of pairing, the lesion adopts the described *syn* conformation around the nucleosidic bond. This places the fluorene moiety towards the dC complementary base, which is consequently flipped out of the duplex. Optimizing stacking of the large π surface of the fluo-

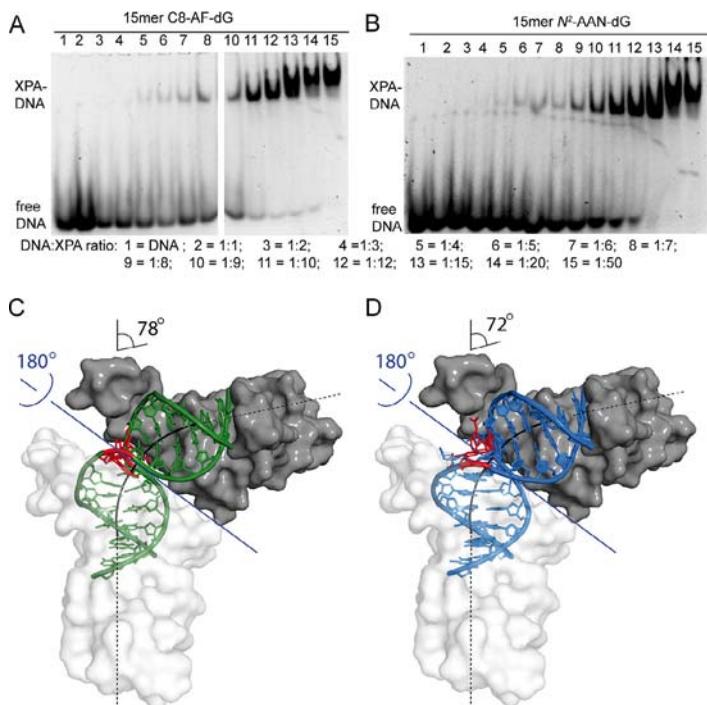


Figure 1. EMSA of XPA with DNA containing A) C8-AF-dG and B) N²-AAN-dG. Structures of the complexes between two Rad14 repair proteins (DNA-binding domains) bound to DNA containing C) C8-AF-dG or D) N²-AAN-dG. In both structures the DNA is kinked (78° and 72°, respectively). The pseudo-C₂-symmetry axis is shown in blue; the lesion (red) is positioned at the outer edge of the kink.

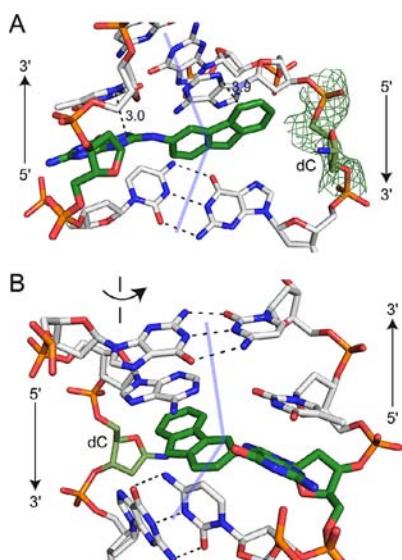


Figure 2. Detail of the C8-AF-dG crystal structure, with DNA as a stick model and the lesion highlighted in green. A) The C8-AF-dG lesion adopts a *syn* conformation around the nucleosidic bond and stacks with the dA of the complementary strand. The dG base stacks with the 3' dT base (the 3' dT:dA base pair is not formed). For the dC complementary base only the phosphodiester backbone and ribose moiety are defined in the $F_o - DF_c$ electron density (green mesh, contoured at 2.5 σ). B) The 5' dC:dG base pair is formed despite buckled distortion. The direction of the kink is indicated in light blue.

rene ring system is clearly more important than stacking of the small pyrimidine dC. We were unable to detect any electron density for the cytosine heterocycle of the dC base, thus suggesting that the heterocycle does not have a fixed conformation (Figure 2 A).

The structure of the N^2 -AAN-dG differs at the lesion site. Although the lesion looks much more helix-disrupting due to the blocked Watson–Crick face, the crystal structure shows that the disrupting effect is much less than for the C8-AF-dG lesion. First, and most importantly, both the surrounding dC:dG and the dT:dA base pairs are intact (Figure 3 A). These two base pairs are directly at the kink site, so they have a buckled rather than parallel orientation. The lesion slides into the V (formed by the buckling) to form stacking interactions with both base pairs thereby filling the void volume. This is possible because the dG part of the lesion and the naphthyl part are almost in a plane, so one large π surface is created (the angle between the two planes is only 10°). Stacking of the lesion is optimal with the dT:dA base pair on top, the dG packing below the dT base, and the naphthyl part stacking below the dA (Figure 3 B). In this situation the naphthyl part occupies the space that is reserved usually for the dC complementary base, which is flipped out of the duplex into the much widened minor groove (Figure 3 A). Due to the extensive stacking interactions of the N^2 adduct, the melting temperature of the DNA duplex is only marginally reduced compared to the undamaged control strand and the C8 adducts (Table S3).

The N -acetyl subunit of the lesion points in the same direction out of the minor groove. Again, electron density for the dC-heterocycle was not present in the structure, thus showing that the heterocycle is flexible. In addition to the flipped-out

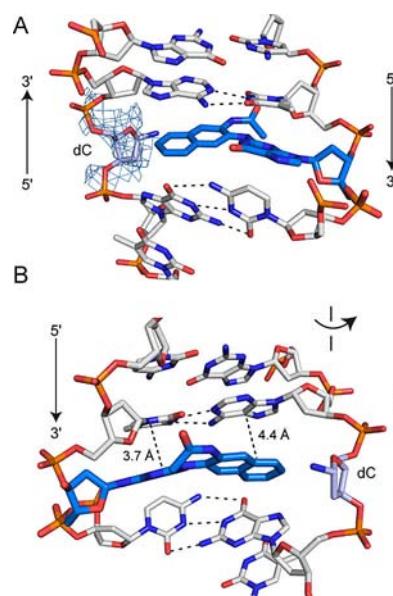


Figure 3. Detail of the N^2 -AAN-dG crystal structure, with DNA as a stick model, and the lesion highlighted in blue. A) The N^2 -AAN-dG lesion adopts an *anti* conformation around the nucleosidic bond. The base pairs above and below the lesion are intact. The dG base and the N^2 -AAN lesion are almost planar. For the dC complementary base only the phosphodiester backbone and ribose moiety are defined in the $F_o - DF_c$ electron density (blue mesh, contoured at 2.5 σ). B) The lesion stacks with the dA of the complementary strand; the dG part stacks with the dT base above.

complementary dC base, the kink at the lesion site is another shared feature (this was already observed for the C8-AAF-dG base), and hence it is possible that this dC is another structural element that induces assembly of the other repair factors at bulky adduct damaged sites. It is possible that this flipped out base functions as the recognition motif that allows the NER endonucleases to cleave specifically the damaged strand. Currently, we do not understand how the NER system distinguishes the damaged strand for incision from the undamaged strand. The Rad14:DNA structures are pseudo- C_2 symmetric, with the DNA bound 50:50 in both directions (Figure 1 C and D). It is therefore likely that the flipped-out dC allows downstream factors to differentiate the damage-containing strand from the intact strand. This hypothesis needs investigation.

A comparison between the structures obtained here (C8-AF-dG lesion) and our previous work (C8-AAF-dG) offers an interesting perspective (Figure 4). The C8-AAF-dG lesion also exists in a *syn* conformation, but, whereas the two ring systems of the C8-AF-dG deviate from planarity by 50°, the two rings of the C8-AAF-dG are almost in a plane (perfect π surface). This allows the C8-AAF-dG lesion to form much better stacking interactions, not only with the upper dT:dA base pair but also with the lower dC:dG. C8-AAF-dG therefore improves the stability of the DNA kink significantly, and this might explain why this lesion is much better recognized by XPA/Rad14. In fact, for our XPA construct we noticed that when we performed the EMSA binding studies, the difference between binding to DNA containing lesions 2 and 3 and undamaged DNA is small (Figure S2), whereas for the C8-AAF-dG lesion, duplexes containing this lesion are more strongly bound.^[7a] Our crystal structures

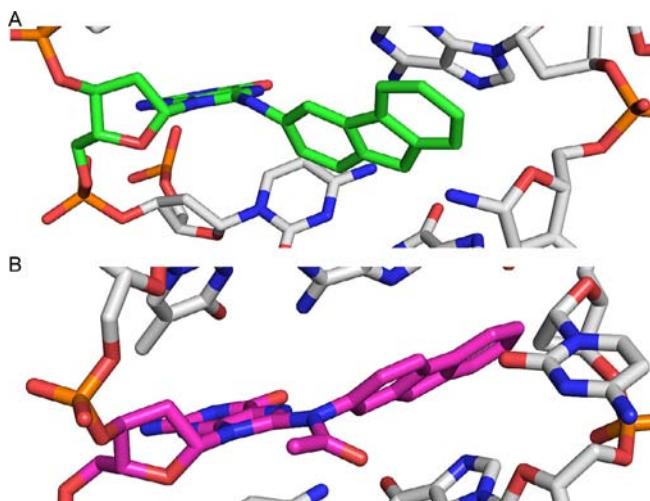


Figure 4. Detail of the A) C8-AF-dG (present study) and B) C8-AAF-dG (PDB ID: 5A3D) crystal structures. The lesions are highlighted in green and magenta, respectively. Both lesions adopt a *syn* confirmation around the nucleosidic bond. The dG base and the C8-AAF lesion are almost planar; the C8-AF lesion and dG are in planes oriented at 50°.

therefore confirm that the flexibility associated with the presence of a damaged base, as needed to establish a sharp 70–78° kink, is a major factor for efficient recognition by XPA/Rad14. XPA/Rad14 is, in this sense, a duplex flexibility sensor.

The crystal structures have been deposited in the PDB under IDs: 5LCL and 5LCM.

Acknowledgements

The authors would like to thank the crystallization facility at the Max Planck Institute of Biochemistry for setting up the crystallization plates. We also want to thank the Swiss Light Source (SLS) and European Synchrotron Radiation Facility (ESRF) for beam time and the staff of beamlines PXI (SLS), MASSIF-3 and ID30B (ESRF) for assistance with the crystal testing and data collection. We thank the Deutsche Forschungsgemeinschaft, SFB1032 (TP-A5), SFB749 (TP-A4), SPP1784 and the Excellence Cluster CiPS^M for financial support and Dr. Markus Müller for critical reading of the manuscript and helpful advice.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: bulky adducts • DNA damage • nucleosides • nucleotide excision repair • Rad14 • XPA

- [1] a) J. A. Miller, *Cancer Res.* **1970**, *30*, 559–576; b) J. A. Miller, E. C. Miller, *Prog. Exp. Tumor Res.* **1969**, *11*, 273–301; c) R. H. Wilson, F. DeEds, A. J. Cox, Jr, *Cancer Res.* **1941**, *1*, 595–608; d) P. C. Howard, D. A. Casciano, F. A. Beland, J. G. Shaddock, Jr, *Carcinogenesis* **1981**, *2*, 97–102; e) R. H.

Heflich, R. E. Neft, *Mutat. Res.* **1994**, *318*, 73–114; f) J. L. Radomski, E. Brill, *Science* **1970**, *167*, 992–993; g) J. H. Weisburger, E. K. Weisburger, *Pharmacol. Rev.* **1973**, *25*, 1–66.

- [2] a) D. M. Grant, N. C. Hughes, S. A. Janezic, G. H. Goodfellow, H. J. Chen, A. Gaedigk, V. L. Yu, R. Grewal, *Mutat. Res. Fund. Mol. M.* **1997**, *376*, 61–70; b) G. R. Hoffmann, R. P. P. Fuchs, *Chem. Res. Toxicol.* **1997**, *10*, 347–359; c) J. A. Miller, E. C. Miller, *Environ. Health Perspect.* **1983**, *49*, 3–12; d) L. Möller, *Environ. Health Perspect.* **1994**, *102*, 139–146; e) D. Kim, F. P. Guengerich, *Annu. Rev. Pharmacol.* **2005**, *45*, 27–49; f) R. A. McClelland, A. Ahmad, A. P. Dicks, V. E. Licence, *J. Am. Chem. Soc.* **1999**, *121*, 3303–3310; g) F. A. Beland, W. T. Allaben, F. E. Evans, *Cancer Res.* **1980**, *40*, 834–840; h) W. G. Humphreys, F. F. Kadlubar, F. P. Guengerich, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8278–8282; i) A. Dipple, *Carcinogenesis* **1995**, *16*, 437–441; j) S. A. Kennedy, M. Novak, B. A. Kolb, *J. Am. Chem. Soc.* **1997**, *119*, 7654–7664.
- [3] a) T. Lindahl, P. Karran, R. D. Wood, *Curr. Opin. Genet. Dev.* **1997**, *7*, 158–169; b) J. H. J. Hoeijmakers, *Nature* **2001**, *411*, 366–374; c) O. D. Schärer, *Angew. Chem. Int. Ed.* **2003**, *42*, 2946–2974; *Angew. Chem.* **2003**, *115*, 3052–3082; d) N. E. Geacintov, S. Broyde, T. Buterin, H. Naegeli, M. Wu, S. Yan, D. J. Patel, *Biopolymers* **2002**, *65*, 202–210; e) D. Gunz, M. T. Hess, H. Naegeli, *J. Biol. Chem.* **1996**, *271*, 25089–25098; f) J.-C. Huang, D. S. Hsu, A. Kazantsev, A. Sancar, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12213–12217; g) L. C. J. Gillet, O. D. Schärer, *Chem. Rev.* **2006**, *106*, 253–276; h) S. Lagerwerf, M. G. Vrouwe, R. M. Overmeer, M. I. Fousteri, L. H. F. Mullenders, *DNA Repair* **2011**, *10*, 743–750; i) O. D. Schärer, *Cold Spring Harbor Perspect. Biol.* **2013**, *5*, a012609; j) S. C. Koch, N. Simon, C. Ebert, T. Carell, *Q. Rev. Biophys.* **2016**, *49*, e5.
- [4] a) J.-H. Min, N. P. Pavletich, *Nature* **2007**, *449*, 570–575; b) T. Riedl, F. Han-naoka, J.-M. Egly, *EMBO J.* **2003**, *22*, 5293–5303; c) M. Volker, M. J. Moné, P. Karmakar, A. van Hoffen, W. Schul, W. Vermeulen, J. H. J. Hoeijmakers, R. van Driel, A. A. van Zeeland, L. H. F. Mullenders, *Mol. Cell* **2001**, *8*, 213–224; d) M. S. Luijsterburg, G. von Bornstaedt, A. M. Gourdin, A. Z. Politi, M. J. Moné, D. O. Warmerdam, J. Goedhart, W. Vermeulen, R. van Driel, T. Höfer, *J. Cell Biol.* **2010**, *189*, 445–463; e) K. Sugasawa, J. M. Y. Ng, C. Masutani, S. Iwai, P. J. van der Spek, A. P. M. Eker, F. Hanaoka, D. Bootsma, J. H. J. Hoeijmakers, *Mol. Cell* **1998**, *2*, 223–232.
- [5] a) W. Yang, *Cell Res.* **2008**, *18*, 184–197; b) B. C. Feltes, D. Bonatto, *Mutat. Res. Rev. Mutat. Res.* **2015**, *763*, 306–320; c) C.-L. Li, F. M. Golebiowski, Y. Onishi, N. L. Samara, K. Sugasawa, W. Yang, *Mol. Cell* **2015**, *59*, 1025–1034; d) T. Ikegami, I. Kuraoka, M. Saijo, N. Kodo, Y. Kyogoku, K. Morikawa, K. Tanaka, M. Shirakawa, *Nat. Struct. Mol. Biol.* **1998**, *5*, 701–706; e) G. W. Buchko, S. Ni, B. D. Thrall, M. A. Kennedy, *Nucleic Acids Res.* **1998**, *26*, 2779–2788.
- [6] a) M. Kaposi, *Ann. Dermatol. Venereol.* **1883**, *4*, 29–38; b) C. F. M. Menck, V. Munford, *Genet. Mol. Biol.* **2014**, *37*, 220–233.
- [7] a) S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8272–8277; b) N. Simon, C. Ebert, S. Schneider, *Chem. Eur. J.* **2016**, *22*, 10782–10785.
- [8] M. G. Kemp, J. T. Reardon, L. A. Lindsey-Boltz, A. Sancar, *J. Biol. Chem.* **2012**, *287*, 22889–22899.
- [9] a) R. R. Bonala, M. C. Torres, S. Attaluri, C. R. Iden, F. Johnson, *Chem. Res. Toxicol.* **2005**, *18*, 457–465; b) L. C. J. Gillet, O. D. Schärer, *Org. Lett.* **2002**, *4*, 4205–4208; c) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, *3*, 565–568.
- [10] L. C. J. Gillet, J. Alzeer, O. D. Schärer, *Nucleic Acids Res.* **2005**, *33*, 1961–1969.
- [11] S. Shibusaki, R. Gentles, F. Johnson, A. P. Grollman, *Carcinogenesis* **1991**, *12*, 813–818.
- [12] a) N. Sugitani, S. M. Shell, S. E. Soss, W. J. Chazin, *J. Am. Chem. Soc.* **2014**, *136*, 10830–10833; b) B. Hilton, N. Shkriabai, P. R. Musich, M. Kvaratskhelia, S. Shell, Y. Zou, *Biosci. Rep.* **2014**, *34*, e00162.

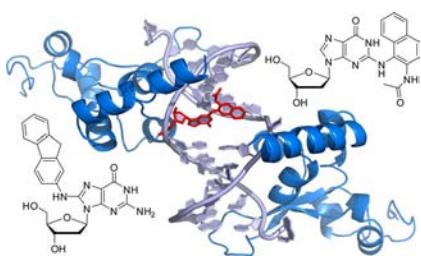
Manuscript received: March 24, 2017

Accepted manuscript online: April 26, 2017

Version of record online: ■■■, 0000

COMMUNICATIONS

Crystal structures of nucleotide excision repair factor XPA/Rad14 in complex with C8-aryl and N^2 -acetyl-aryl-dG DNA gave insights into the interactions. Binding of two Rad14 molecules to the lesion site induces a kink in the DNA and recruits the repair mechanism. XPA/Rad14 thus acts as a duplex flexibility sensor.



C. Ebert, N. Simon, S. Schneider,*
T. Carell*



Structural Insights into the
Recognition of N^2 -Aryl- and C8-Aryl
DNA Lesions by the Repair Protein
XPA/Rad14



CHEMBIO**CHEM**

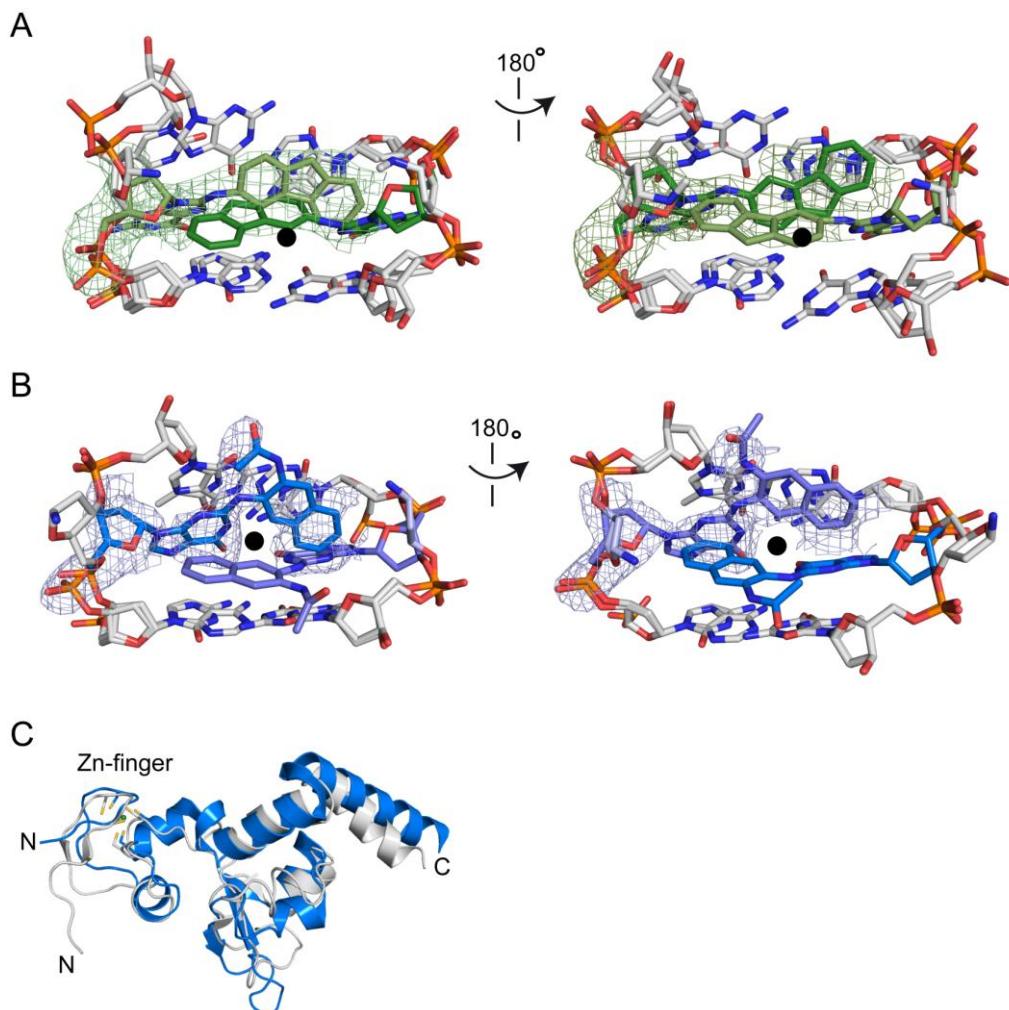
Supporting Information

Structural Insights into the Recognition of *N*²-Aryl- and C8-Aryl DNA Lesions by the Repair Protein XPA/Rad14

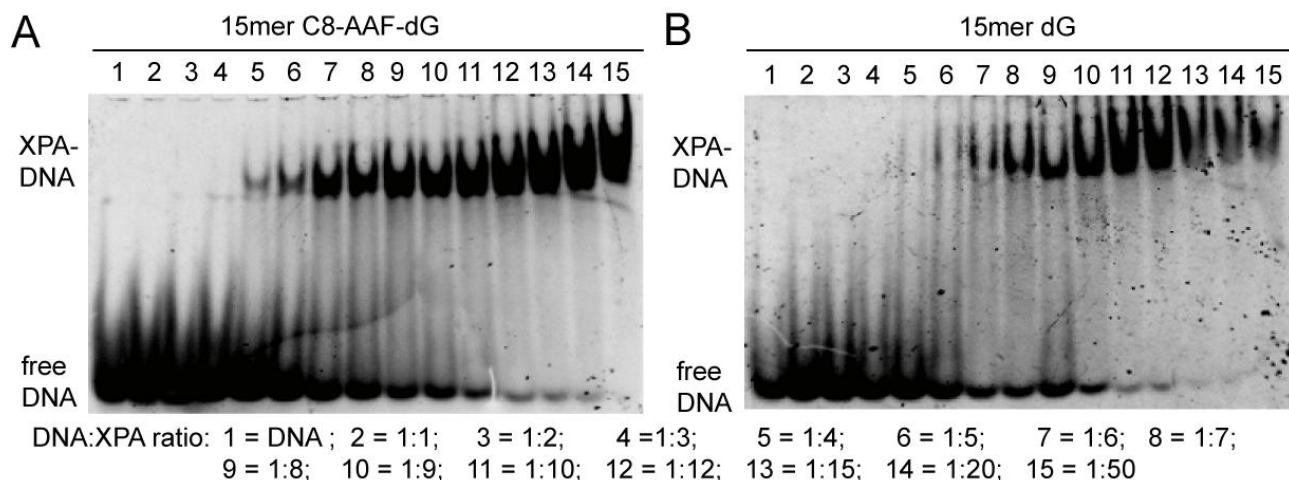
Charlotte Ebert^{+[a]} Nina Simon^{+[a]} Sabine Schneider,^{*[b]} and Thomas Carell^{*[a]}

cbic_201700169_sm_misellaneous_information.pdf

Supporting Information



Supporting Figure S1. Simulated annealing-omit-difference Fo-DFc electron density map of the C8-AF-dG (A) und N^2 -AAN-dG (B) lesions contoured at 2.5σ . The map was calculated after omitting the atoms of the lesions (both orientations). The two orientations of the DNA are overlayed in the figure and the black circle marks the C2-symmetry axis. C) Superposition of the DNA-binding domain of Rad14 (blue) and XPA (grey; NMR structure: PDB code 1XPA); r.m.s.d = 2.6 Å.



Supporting Figure S2. Electrophoretic mobility shift assay (EMSA) of XPA with (A) C8-AAF-dG and (B) dG containing DNA.

Supporting Table S1. Oligonucleotide sequences used in this study. The damaged guanine base is marked with a star and highlighted in bold.

ODN	Oligonucleotide sequence	G*
4	5'-GCT CTA CG*T CAT CAC-3'	C8-AAF-dG
5		N ² -AAN-dG
6		C8-AF-dG
7	5'-GTG ATG ACG TAG AGC-3'	-

Supporting Table S2. Data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses. * Indicates values calculated by the Molprobity server^[1](<http://molprobity.biochem.duke.edu>)

	C8-AF-dG	N ² -AAN-dG
Wavelength (Å)	0.968	0.976
Resolution range	49.3-2.2 (2.3 - 2.2)	49.2-1.9 (1.97- 1.9)
Space group	P 4 ₁	P 4 ₁
Unit cell	53.2 53.2 131.0	53.0 53.0 131.3
Total reflections	130,646 (9,621)	127,851 (12,486)
Unique reflections	18,314 (1,714)	28,333 (2,850)
Multiplicity	7.1 (5.6)	4.5 (4.4)
Completeness (%)	0.99 (0.93)	0.99 (0.99)
Mean I/sigma(I)	8.6 (1.6)	7.3 (1.4)
Wilson B-factor	31.4	28.4
R-merge	0.168 (0.901)	0.125 (0.704)
R-meas	0.181 (0.993)	0.142 (0.799)
R-pim	0.07 (0.408)	0.07 (0.373)
CC1/2	0.995 (0.566)	0.995 (0.604)
CC*	0.999 (0.85)	0.999 (0.868)
Reflections used in refinement	18,308 (1716)	28,349 (2,850)
Reflections used for R-free	895 (87)	1,370 (144)
R-work	0.216 (0.292)	0.231 (0.328)
R-free	0.257 (0.317)	0.2687 (0.331)
CC(work)	0.927 (0.72)	0.938 (0.719)
CC(free)	0.94 (0.68)	0.95 (0.65)
Number of non-hydrogen atoms		
Protein	1,877	1,632
DNA	1,157	1,156
Zinc-ions	2	2
Waters	49	112
Protein residues	227	229
RMS(bonds)	0.035	0.021
Bad bonds (protein)*	0/1,927 (0%)	0/1,967 (0%)
Bad bonds (DNA)*	3/1,217 (0.25%)	14/1,216 (1.15%)
RMS(angles)	2.1	2.5
Bad angles (protein)*	0/2598 (0%)	1/2,651 (0.04%)
Bad angles (DNA)*	33/1,861 (1.77%)	85/1,860 (4.57%)
Ramachandran favored (%)*	96	97
Ramachandran allowed (%)*	3.6	2.6
Ramachandran outliers (%)*	0	0.4
Rotamer outliers (%)*	2	1.4
Clashscore*	13.2	12.2
Average B-factor		
Protein	36.9	33.9
DNA	46.7	43.4
Zinc ions	24.4	23.0
Waters	30.2	33.2
Number of TLS groups	20	22

Supporting Table S3. Melting points.

DNA lesion	T _M [°C] 15mer
dG:dC	64
C8-AF-dG:dC	55
N ² -AAN-dG:dC	60
C8-AAF-dG:dC	48

Supporting Materials and Methods

Chemicals were purchased from Sigma-Aldrich, Fluka, Alfa Aesar, ChemGenes or Acros and used without further purification. Solutions were concentrated *in vacuo* on a Heidolph rotary evaporator. The solvents were of reagent grade and purified by distillation. Dry solvents were bought from Sigma-Aldrich. Acetonitrile for HPLC-purification was purchased from VWR. Water was purified by a Milli-Q Plus system from Merck Millipore. Chromatographic purification of products was accomplished using flash column chromatography on Merck Geduran Si 60 (40-63 µM) silica gel (normal phase). Thin layer chromatography (TLC) was performed on Merck 60 (silica gel F₂₅₄) plates. ¹H, ¹³C and ³¹P NMR spectra were recorded in deuterated solvents on *Bruker ARX 300*, *Varian VXR400S*, *Varian Inova 400*, *Bruker AMX 600* and *Bruker Avance 800* spectrometers and calibrated to the residual solvent peak. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, hept = septet, m = multiplet, br. = broad. For assignment of the structures, additional 2D NMR spectra (COSY, HSQC, HMBC) were measured. Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectra were recorded on a Bruker Autoflex II. For MALDI-TOF measurements the samples were dialysed on a 0.025 µm VSWP filter (Millipore) against ddH₂O for 1 h and then cocrystallised in a 3-hydroxypicolinic acid matrix (HPA: 25 mg 3-hydroxypicolinic acid, 5 mg ammonium citrate in 0.5 mL H₂O). High resolution electrospray ionization mass spectra (HRMS-ESI) were recorded on a Thermo Finnigan LTQ-FT (ESI-FTICR). The pH-values of buffers were adjusted using a MP 220 pH-meter (Metter Toledo). UV spectra and melting profiles were measured on a Jasco V-650 spectrometer using quartz glass cuvettes with 1 cm path length. The concentration of the purified DNA was determined with a NanoDrop ND-1000 spectrophotometer (Peqlab) from Thermo Scientific (1.5 µL sample volume) using calculated extinction coefficients from Oligo Calc (online: <http://www.basic.northwestern.edu/biotools/OligoCalc.html>) and applying Lambert-Beer's law. For strands containing artificial bases the extinction coefficient of their corresponding control strand was employed without corrections.

Phosphoramidite synthesis

The C8-AAF-dG phosphoramidite containing an *isopropylphenoxyacetyl* group at the N^2 -position for solid phase DNA synthesis and the N^2 -AAN-dG phosphoramidite were prepared as previously published.^[2]

DNA Preparation

The C8-lesion phosphoramidite containing an *isopropylphenoxyacetyl* group at the N^2 -position for solid phase DNA synthesis was incorporated into DNA using ultra mild conditions for DNA synthesis.^[3] AF-C8-dG lesion was prepared as previously published.^[4] The N^2 -AAN-dG phosphoramidite was incorporated into DNA using standard conditions.^[2c] Oligonucleotide synthesis was performed on an ABI 394 Nucleic Acid Synthesis System (Life Technologies, Darmstadt, Germany). Phosphoramidites for dA, dC, dG, dT and CPG carriers were obtained from Glen Research (Sterling, VA) or Link Technologies (Bellshill, Scotland). The coupling time for the modified phosphoramidites was extended to 3 x 15 min. The oligonucleotides were purified by reverse phase HPLC chromatography (Macherey-Nagel, Nucleosil 100-7 C18, 10 x 250 mm, 0.1 M triethylammonium acetate, linear gradient from 0% to 40% in 45 min (C8-AAF-dG and C8-AF-dG) or 0% to 25% in 45 min (N^2 -AAN-dG). The purified strands were desalted using C18 cartridges (Waters, Sep-Pak C18 Classic Cartridge). The identity of the oligonucleotides were confirmed by matrix-assisted laser-desorption time-of-flight mass spectrometry and the purity analytical HPLC.

Electromobility shift assays

The damaged oligonucleotides (ODN **4-6**) were mixed with their complementary strands (ODN **7**) in equal molar amounts and annealed to form duplex DNA. The formation of the duplex DNA was checked by native TBE polyacrylamide gel electrophoresis. A standard DNA binding reaction mixture (10 µL) contained 5 pmol of the C8- or N^2 -lesion duplex DNA and varying concentrations of purified XPAfl in binding buffer (50 mM HEPES-KOH pH 8.0, 100 mM NaCl, 5 mM MgCl₂, 1 mM DTT and 5% (v/v) glycerol). After incubation at 37 °C for 30 min, 2 µL loading dye (250 mM Tris-HCl pH 7.5, 40% glycerol, 0.2% bromphenolblue) were added and electrophoresis was performed on a 4% non-denaturing polyacrylamide TBE gel at 4 °C, which was pre-run for 30 min. Band shifts were visualized by incubating the gel in 1x TBE with SYBR Green I (1:10000, Sigma Aldrich) for 30 min and then detected via ImageReader.

Protein expression and purification

The *Saccharomyces cerevisiae* Rad14 DNA binding domain (residues 188-302) and the human XPA full length protein were expressed and purified as published in^[5]. The stability of the protein was tested via limited proteolysis^[5] and also proofed via gelfiltration.

Crystallization, data collection, structure determination and structure refinement

For co-crystallization experiments C8-AF and N²-AAN-dG containing 15mer DNA (ODN **4-6**, Supporting Table S1) was annealed in crystallization buffer to its counter strand ODN **7** (Metabion). Prior to crystallization protein and DNA were mixed in a molar ratio of 2:1 protein:DNA and incubated for 30 min at 4 °C. Crystals were grown in 0.22 M ammonium nitrate and 38 % 2-methyl-1,3-propanediol (MPD) (C8-AF-dG) or 0.02 M CaCl₂, 0.1 M NaOAc pH 4.6 and 30% MPD (N²-AAN-dG) at 4 °C, cryo-cooled in liquid nitrogen directly from the crystallisation drop and stored in liquid nitrogen until data collection.

Due to the small crystal size of 10-20 μm, diffraction data were collected with a micro beam at MASSIF-3 and ID30B beamlines, European Synchrotron Radiation Facility (ESRF)). Data processing was carried out with XDS^[6] and the resolution cut-offs were chosen according to the correlation coefficient of random half-data sets (1/2 CC) at about 50%.^[7] The coordinates of the Rad14-DNA complex (PDB code 5G34) were used to solve the structure by rigid-body refinement in REFMAC5^[8] after the atoms for the DNA were removed from the model. Completion of the models was done by iterative cycles of manual model building in COOT^[9] and restraint and TLS refinement^[10] in REFMAC5. The optimal number of TLS-groups was determined using the TLSMD-Server.^[10] Simulated-annealing omit electron density map coefficients were calculated using PHENIX.^[11] Structure validation and analysis was carried out with MOLPROBITY^[1], PISA^[12] and CURVES.^[13] For data processing and structure refinement statistics see the Supporting Table 2. Structural superposition were done with SSM^[14] and all structural figures prepared with PyMol (Delano Scientific).

Atomic coordinates and X-ray structure data were submitted to the Protein Data Bank (<http://www.ebi.ac.uk/pdbe/>). PDB codes: 5LCL (C8-AF-dG), 5LCM (N²-AAN-dG).

Melting Temperatures

The duplexes of ODN **4-6** and ODN **7** (Supporting Table S1, at concentration 2 μM) were dissolved in 250 μL of 10 mM Tris-HCl, pH 7.4, containing 150 mM NaCl. The samples were heated from 15 to 90 °C or cooled from 90 to 15 °C with a rate of 1 °C/min. Three melting profiles per sample were performed (Supporting Table S3).

References

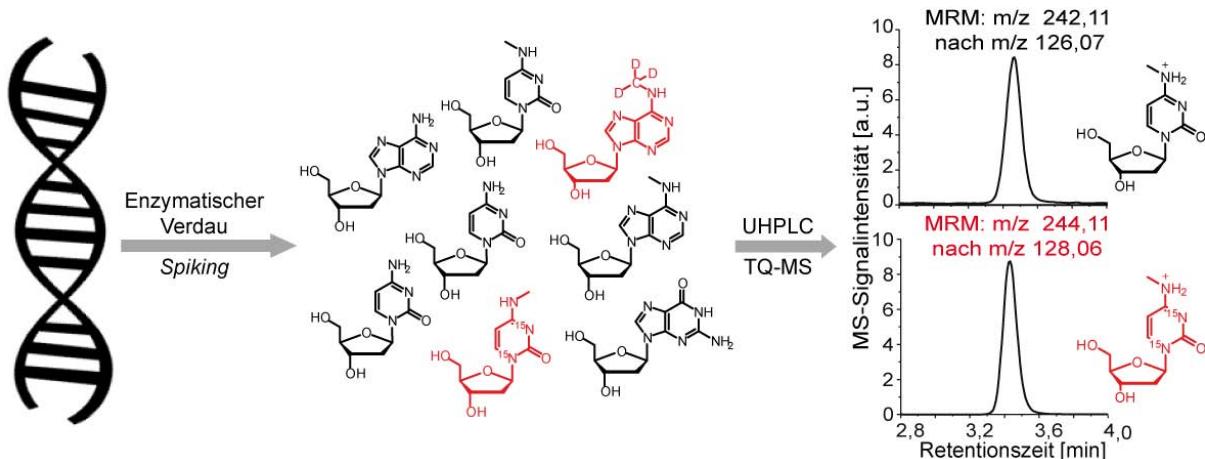
- [1] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Cryst. D* **2010**, *66*, 12-21.
- [2] a) Z. Wang, C. J. Rizzo, *Org. Lett.* **2001**, *3*, 565-568; b) L. C. Gillet, O. D. Scharer, *Org. Lett.* **2002**, *4*, 4205-4208; c) R. R. Bonala, M. C. Torres, S. Attaluri, C. R. Iden, F. Johnson, *Chem. Res. Toxicol.* **2005**, *18*, 457-465; d) N. Simon, C. Ebert, S. Schneider, *Chem. Eur. J.* **2016**, *22*, 10782-10785.
- [3] L. C. Gillet, J. Alzeer, O. D. Scharer, *Nucleic Acids Res.* **2005**, *33*, 1961-1969.
- [4] S. Shibusaki, R. Gentles, F. Johnson, A. P. Grollman, *Carcinogenesis* **1991**, *12*, 813-818.
- [5] S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *PNAS* **2015**, *112*, 8272-8277.
- [6] W. Kabsch, *Acta Cryst. D* **2010**, *66*, 133-144.
- [7] a) K. Diederichs, P. A. Karplus, *Acta Cryst. D* **2013**, *69*, 1215-1222; b) P. Evans, *Science* **2012**, *336*, 986-987; c) P. A. Karplus, K. Diederichs, *Science* **2012**, *336*, 1030-1033.
- [8] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Cryst. D* **2011**, *67*, 355-367.
- [9] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Cryst. D* **2010**, *66*, 486-501.
- [10] a) J. Painter, E. A. Merritt, *J. Appl. Crystallogr.* **2006**, *39*, 109-111; b) M. D. Winn, G. N. Murshudov, M. Z. Papiz, *Methods Enzymol.* **2003**, *374*, 300-321.
- [11] a) P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Cryst. D* **2010**, *66*, 213-221; b) P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Cryst. D* **2012**, *68*, 352-367.
- [12] E. Krissinel, K. Henrick, *J Mol Biol* **2007**, *372*, 774-797.
- [13] R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, K. Zakrzewska, *Nucleic Acids Res.* **2009**, *37*, 5917-5929.
- [14] E. Krissinel, K. Henrick, *Acta Cryst. D* **2004**, *60*, 2256-2268.

3.2 Synthese von isotopenmarkierten Standards methylierter DNA-Basen

Auch methylierte DNA-Basen können DNA-Schäden darstellen, aber anders als die arylmodifizierten Basen können diese in bestimmten Organismen auch verschiedene Funktionen ausüben (Kapitel 1.2). Die Quantifizierung dieser methylierten DNA-Basen in unterschiedlichen Organismen und Geweben und zu unterschiedlichen Zeitpunkten der Entwicklung ist von zentraler Bedeutung um zu beurteilen, ob diese eine Funktion besitzen, oder – wenn vorhanden – eher als Schaden einzustufen sind.

3.2.1 Quantifizierung von DNA-Basen mittels UHPLC-MS

In Schema 22 ist der allgemeine Arbeitsablauf für die Quantifizierung von DNA-Basen dargestellt.



Schema 22: Arbeitsablauf für die Quantifizierung von DNA-Basen. Die DNA wird enzymatisch verdaut, sodass man die einzelnen Nukleoside erhält. Isotopenmarkierte Standards werden in bekannter Menge dazugegeben und die Mischung wird via UHPLC-MS analysiert. Durch Vergleich der Integrale des Analyten und des Standards im Massenspektrum kann die Menge an Analyt quantifiziert werden. MRM = *Multiple Reaction Monitoring*.

Um DNA-Basen in einem spezifischen Gewebe quantifizieren zu können muss zunächst die DNA aus dem Gewebe extrahiert werden.^[269] Im nächsten Schritt wird diese dann enzymatisch verdaut, sodass die einzelnen Nukleoside vorliegen. Das Gemisch dieser Nukleoside kann nun über eine HPLC (*High Performance Liquid Chromatography*) aufgetrennt werden.^[269-270] Standardmäßig wird für die Detektion von DNA-Basen auf der HPLC ein UV-Detektor bei 260 nm verwendet. Koppelt man die HPLC (oder UHPLC, *Ultra High Performance Liquid Chromatography*) jedoch stattdessen an ein Massenspektrometer, so verbessert sich das Detektionslimit um ein Vielfaches und auch kleinste Spuren modifizierter Basen in der DNA-Probe können detektiert werden. Der Vorteil dieser Methode ist, dass sich die Analyten so auf Grund von zwei verschiedenen Eigenschaften identifizieren

lassen: Der Retentionszeit und dem Massenübergang. Ein bestimmtes Signal kann somit mit hoher Sicherheit einer bestimmten Base zugeordnet werden. Die Höhe eines Massesignals an sich erlaubt allerdings keine quantitative Aussage über die Menge dieses spezifischen Nukleosids im ursprünglichen Gemisch. Man hat hier zwei Möglichkeiten zur Quantifizierung der Nukleoside über eine externe Eichgerade des natürlichen Nukleosids oder durch Zusatz eines isotopenmarkierten internen Standards. Bei diesen Standards handelt es sich um die isotopenmarkierten Derivate der DNA-Basen, welche quantifiziert werden sollen. Der Einsatz von stabilen isotopenmarkierten Standards macht eine exaktere Quantifizierung möglich, die nicht von äußereren Bedingungen wie Puffer etc. abhängt. Zudem kann die Probenmenge reduziert werden, da nicht jeder Messung eine externe Eichgerade im Triplikat vorangestellt werden muss. Die Messungen werden an einem Dreifach Quadrupol Massenspektrometer (*Triple Quadrupole Mass Spectrometer TQ-MS*) vorgenommen. Hier werden die Analyten eines bestimmten Massebereichs selektiert. Die Standards müssen mindestens zwei Masseeinheiten leichter oder schwerer als das zu quantifizierende Nukleosid sein, damit Moleküle, welche natürlich vorkommende Isotopen wie ^{13}C enthalten, nicht fälschlicherweise als Standard analysiert werden. Im Fall von m^5dC reicht ein um zwei Masseeinheiten schwererer Standard jedoch nicht aus, da diese Modifikation in so großen Mengen vorkommt. Das führt dazu, dass am sensitiven TQ-MS das die Größe des Signals des +2 Massenübergangs sowohl durch den $\text{m}^5\text{dC}[+2]$ -Standard als auch durch natürliches $\text{m}^5\text{dC}[+2]$ repräsentiert würde. Der Grund hierfür ist, dass ^{13}C natürlich schon zu ca. 1,1 % vorkommt. Um eine exakte Quantifizierung zu ermöglichen wird deshalb in diesem Fall ein $\text{m}^5\text{dC}[+3]$ -Standard verwendet. Gibt man nun vor der HPLC-MS Analyse eine bekannte Menge an Standard zur Probe (*Spiking*), so kann man im Massenspektrum das Integral des Analyten (leicht) mit dem Integral des Standards (schwer) vergleichen und dadurch die Menge an Analyt in der Probe ermitteln. Diese Methode erlaubt die zeitgleiche Quantifizierung aller modifizierten DNA-Basen in einer Probe, solange sie sich in ihrer Retentionszeit oder in ihrer Masse unterscheiden.^[269-271]

3.2.2 Quantifizierung von 4-Methyl-2'-desoxycytidin (m^4dC) und 6-Methyl-2'-desoxyadenosin (m^6dA)

Die folgende Publikation befasst sich mit der Synthese von m^4dC und m^6dA sowie ihrer um zwei Masseeinheiten schwererer Isotopologe für deren Quantifizierung in verschiedenen Organismen und Geweben.

3.2.2.1 Zusammenfassung der Publikation zur Quantifizierung von m^4dC und m^6dA

Quantitative LC-MS provides no evidence for m^6dA or m^4dC in the genome of mouse ESCs and tissues.

S. Schiffers, C. Ebert, R. Rahimoff, O. Kosmatchev, J. Steinbacher, A. Bohne, F. Spada, S. Michalakis, J. Nickelsen, M. Müller, T. Carell, *Angew. Chem. Int. Ed.* **2017**, DOI: 10.1002/anie.201700424.

Kopie der Publikation mit Erlaubnis des Verlags; Copyright 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Bis vor kurzem wurde angenommen, dass Genome höherer Organismen zusätzlich zu den vier kanonischen DNA-Basen nur m^5dC als modifizierte Base zur Kontrolle epigenetischer Prozesse enthalten. In den letzten Jahren hat sich diese Ansicht auf Grund der Entdeckung von hmdC, fdC und cadC in Stammzellen und im Gehirn drastisch geändert. Seit neuestem häufen sich die Berichte über das Vorkommen von m^6dA in verschiedenen eukaryotischen Organismen. Diese Base war wie m^4dC zunächst nur als Komponente des Bakteriengenoms bekannt. In dieser Arbeit wurde der Gehalt dieser potenziell epigenetisch relevanten DNA-Basen mittels einer ultrasensitiven UHPLC-MS-Methode untersucht. Die Präsenz von m^6dA und m^4dC im Cyanobakterium *Synechocystis* und von ersterem auch in der Grünalge *C. reinhardtii* konnte zwar bestätigt werden, in mES-Zellen und im Gewebe adolescenter Mäuse konnten diese Basen jedoch nicht detektiert werden. Dieses Ergebnis stellt ihre epigenetische Relevanz in Vertebraten in Frage. Für die modifizierten Basen m^5dC , hmdC, fdC und cadC konnten hingegen quantitative Daten erhalten werden.

Beitragserklärung: Es wurden die methylierten DNA-Basen m^6dA , 6-[D₃]-Methyl-2'-desoxycytidin und 4-Methyl-2'-desoxy-[1,3-¹⁵N₂]-cytidin synthetisiert, via HPLC aufgereinigt und charakterisiert.

Quantitative LC–MS Provides No Evidence for m⁶dA or m⁴dC in the Genome of Mouse Embryonic Stem Cells and Tissues

Sarah Schiffers, Charlotte Ebert, René Rahimoff, Olesea Kosmatchev, Jessica Steinbacher, Alexandra-Viola Bohne, Fabio Spada, Stylianos Michalakis, Jörg Nickelsen, Markus Müller, and Thomas Carell*

Abstract: Until recently, it was believed that the genomes of higher organisms contain, in addition to the four canonical DNA bases, only 5-methyl-dC (m^5dC) as a modified base to control epigenetic processes. In recent years, this view has changed dramatically with the discovery of 5-hydroxymethyl-dC ($hmdC$), 5-formyl-dC (fdC), and 5-carboxy-dC ($cadC$) in DNA from stem cells and brain tissue. N⁶-methyldeoxyadenosine (m^6dA) is the most recent base reported to be present in the genome of various eukaryotic organisms. This base, together with N⁴-methyldeoxycytidine (m^4dC), was first reported to be a component of bacterial genomes. In this work, we investigated the levels and distribution of these potentially epigenetically relevant DNA bases by using a novel ultrasensitive UHPLC–MS method. We further report quantitative data for m^5dC , $hmdC$, fdC , and $cadC$, but we were unable to detect either m^4dC or m^6dA in DNA isolated from mouse embryonic stem cells or brain and liver tissue, which calls into question their epigenetic relevance.

The genetic material of living organisms is constructed from the four canonical nucleobases dA, dC, dG, and dT, which establish the sequence information that, in multicellular organisms, is stored in the nucleus of every cell (Figure 1). In addition to the canonical bases, the methylated dC base 5-methyldeoxycytidine (m^5dC) is frequently found.^[1] The presence or absence of this base in specific promoter segments determines whether the gene is actively transcribed or silenced.^[1] The cell-type-specific distribution of m^5dC thus determines the identity of a given cell. Recently, 5-hydroxymethyldeoxycytidine ($hmdC$) was found as a sixth base of

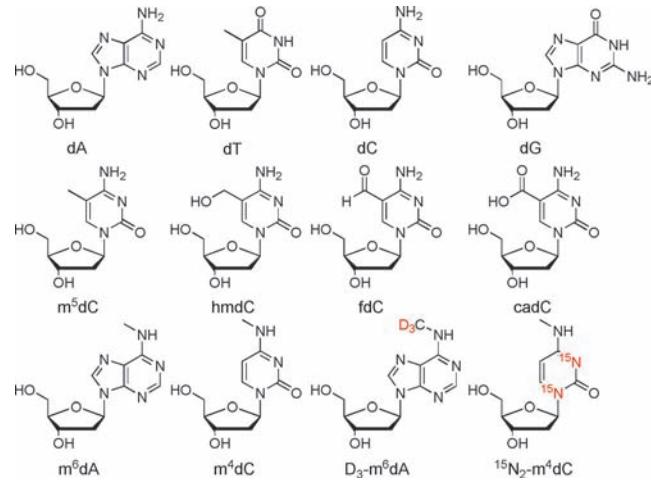


Figure 1. Depiction of the four canonical DNA bases and the epigenetic DNA marks m^5dC , $hmdC$, fdC , and $cadC$, as well as the bases m^6dA and m^4dC together with the synthesized isotopologues.

the genetic system^[2,3] and in 2011, 5-formyldeoxycytidine (fdC)^[4,5] and 5-carboxydeoxycytidine ($cadC$)^[5,6] were also discovered, particularly in DNA isolated from stem cells, but also in brain DNA. It is currently believed that fdC and $cadC$ are intermediates in an active demethylation process that allows cells to change the methylation pattern and hence the activity state of specific genes.^[7,8] For fdC , separate epigenetic functions are also envisaged.^[9]

While the genomes of bacteria are known to also contain N⁴-methyldeoxycytidine (m^4dC)^[10] and N⁶-methyldeoxyadenosine (m^6dA),^[11] attempts to detect these bases in the DNA of higher organisms have failed until recently.^[12–15] m^6dA has now been found in algae (0.4 mol % m^6dA/A),^[12] fruit flies (0.001%–0.07% m^6dA/A),^[14] and *C. elegans* (0.01%–0.4% m^6dA/A),^[13] and its presence has even been reported in the DNA of vertebrates (0.00009% in *X. laevis*^[16] and 0.00019–0.003% of dA in murine cells and tissue^[17]). These discoveries, especially concerning the DNA of vertebrates, have spurred a worldwide research interest in unraveling the function of these new bases in human genomic DNA.^[18–20]

In this study, we developed an ultrasensitive triple quadrupole mass spectrometry (QQQ-MS) method, which in combination with ultra-high-pressure chromatography (UHPLC) enables m^4dC and m^6dA to be searched for and quantified in parallel to the more established new epigenetic DNA marks m^5dC , $hmdC$, fdC and $cadC$.

[*] S. Schiffers, C. Ebert, R. Rahimoff, O. Kosmatchev, J. Steinbacher, Dr. F. Spada, Dr. M. Müller, Prof. Dr. T. Carell
Center for Integrated Protein Science (CiPSM) at the Department of Chemistry, LMU München
Butenandtstr. 5–13, 81377 München (Germany)
E-mail: Thomas.Carell@lmu.de
Homepage: <http://www.carellgroup.de>
Dr. A.-V. Bohne, Prof. Dr. J. Nickelsen
Biocenter of the LMU
Dept. Biologie 1 – Botanik, Molekulare Pflanzenwissenschaften
Grosshaderner Strasse 2–4, 82152 Planegg-Martinsried (Germany)
Dr. S. Michalakis
CiPSM, Department of Pharmacy, LMU München
Pharmacology for Life Sciences
Butenandtstr. 7, 81377 München (Germany)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<http://dx.doi.org/10.1002/anie.201700424>.

For the quantitative measurements, we first chemically synthesized the two isotopologues of m⁶dA and m⁴dC shown in Figure 1 as internal standards for the analytical method. The prepared compounds D₃-m⁶dA and ¹⁵N₂-m⁴dC are three and two mass units heavier, respectively, than the natural bases. Despite these molecular-weight differences, they have identical properties during the UHPL chromatography step so that they strictly coelute with their natural counterparts, thus allowing them to enter the mass spectrometer at exactly the same time as the internal standards. The availability of these isotopologues makes the method highly reliable and strictly quantitative. The syntheses of the two compounds, together with all analytical data, are given in the Supporting Information.

We first benchmarked our study with an investigation of genomic DNA isolated from the unicellular green algae *Chlamydomonas reinhardtii* and the cyanobacterium *Synechocystis*. In both cases, DNA was isolated after cell lysis using a standard method (see the Supporting Information). The isolated DNA was subsequently digested with a mixture of three commercially available digestion enzymes (Nuclease S1, Antarctic Phosphatase, and Snake Venom Phosphodiesterase; see the Supporting Information). We next added the isotope-labelled standards D₃-m⁶dA and ¹⁵N₂-m⁴dC to the obtained digestion mixture and performed UHPLC-QQQ analysis. For the mass spectrometry detection, we selected fragmentation of the glycosidic bond as the indicative and hence recorded mass transition. This is $m/z = 266.12 \rightarrow 150.08$ for m⁶dA and $m/z = 269.14 \rightarrow 153.10$ for its isotopologue D₃-m⁶dA. For m⁴dC, we also used fragmentation of the glycosidic bond, which gives a mass transition of $m/z = 242.11 \rightarrow 126.07$ for the natural compound m⁴dC and $m/z = 244.11 \rightarrow 128.07$ for its isotopologue ¹⁵N₂-m⁴dC (Figure 2 A).

We next modified the reported UHPLC-QQQ method^[10] for the simultaneous quantification of m⁴dC and m⁶dA, together with the other epigenetically relevant bases m⁵dC, hmdC, fdC, and cadC. To this end, the UHPLC gradient was fine-tuned to enable full separation of all six compounds. Finally, we measured precise calibration curves for all of the compounds (see Figures S1 and S2 in the Supporting Information). This subsequently allowed exact quantification of all of the discussed epigenetic DNA marks in a given sample (Figure 2 B–D).

Since m⁴dC and m⁶dA are well known in bacteria, we first analysed the cyanobacterium *Synechocystis* (PCC6803), and we indeed found both bases (Figure 2 B). The base m⁶dA was detected at a level of 8.4×10^{-3} per dN and for m⁴dC we measured a value of 5.9×10^{-3} per dN. The constitutional isomer m⁵dC and all other dC-derived epigenetic DNA marks were detectable, but were not quantified in this experiment.

Next, we analyzed two different strains of *Chlamydomonas reinhardtii* (CC-3491 and wt 7d+), in which m⁶dA has just recently been discovered,^[12] and the levels of m⁶dA were determined to be 8.4×10^{-4} per dN for CC-3491 and 6.9×10^{-4} per dN for wt 7d+ (Figure 2 C). This corresponds to about 3000 m⁶dA bases per *Chlamydomonas* genome (genome size 1.2×10^8), which at 0.7% of the dA is a relatively high number. In both strains, m⁴dC was not detected, thus showing that this base is unlikely to be a component of the genetic

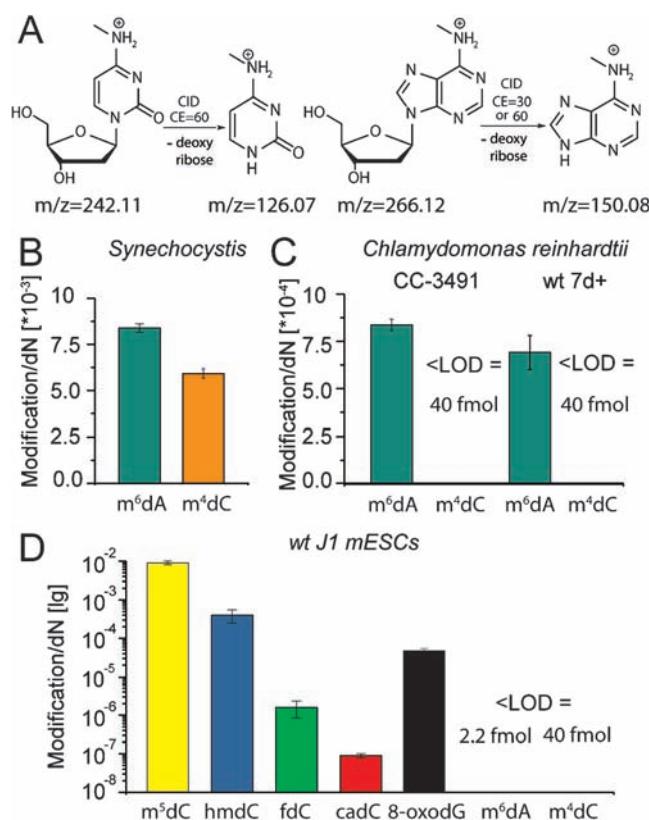


Figure 2. A) Fragmentation patterns of m⁴dC and m⁶dA. B–D) Quantitative data of the bases m⁴dC and m⁶dA in *Synechocystis* (B), *Chlamydomonas* (C) and of these bases and the other epigenetic DNA marks hmdC, fdC, cadC, and m⁵dC in mouse embryonic stem cells (D).

material of *Chlamydomonas*. This is interesting because *Synechocystis* is considered a relative of the chloroplasts present in *Chlamydomonas*.

With these positive results in hand, we extended our study to mouse embryonic stem cells (ESCs; wt J1, Figure 2 D). m⁶dA in particular was recently reported to occur as an epigenetically relevant DNA mark in mouse ESCs (mESC cell line wt TT2).^[17] When performing the measurements, we turned the mass spectrometer to maximum sensitivity. But even in this mode, we were unable to detect a signal for m⁶dA within the detection limits of our system Table S2. In contrast, the other epigenetically relevant bases hmdC, fdC, cadC, and even the oxidative lesion 8-oxodG, which we also quantified in parallel, were clearly detectable. The 8-oxodG level was 4.8×10^{-5} per dN. The rare and difficult to detect cadC was clearly seen even at levels of only 9.0×10^{-8} per dN. For m⁶dA, in contrast, a signal did not appear. We also re-measured wt TT2 cells as described and still did not detect m⁶dA over background levels (see Figure S4).^[17] To obtain unequivocal proof that m⁶dA is not present in stem cells, we added ¹³CD₃-methionine to the mESC culture. Methionine provides the methyl group for the biosynthesis of m⁶dA. With ¹³CD₃-methionine, this would lead to an *m/z*-shift of +4. We tuned the mass spectrometer to the new *m/z*-transition and again were unable to see any signal for ¹³CD₃-m⁶dA (Figure S8).

We subsequently turned our attention to adult mouse tissue and analysed DNA isolated from liver and whole brain

(Figure S3) using our UHPLC-QQQ method. Figure 3 shows the data obtained from mouse liver. The middle column shows the data we obtained for hmdC. The already reported $D_2\text{-}^{15}\text{N}_2\text{-hmdC}$ standard elutes at a retention time of 2.25 min

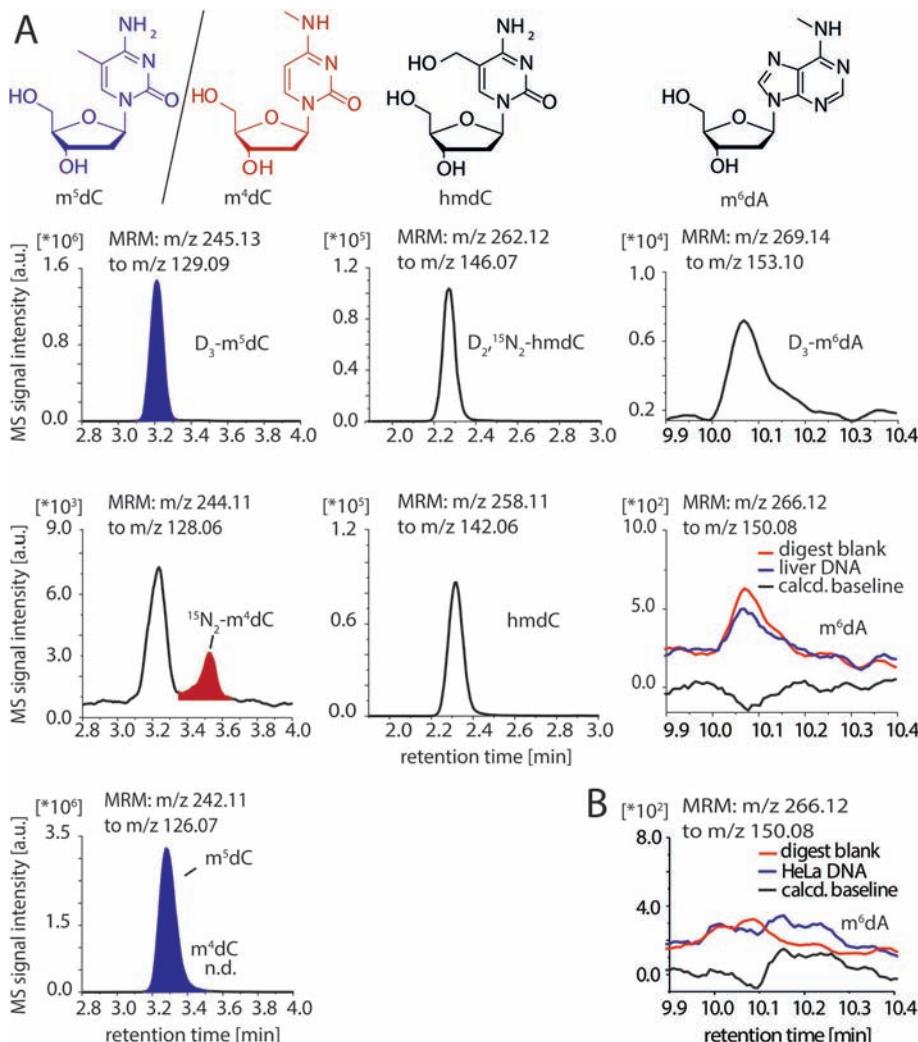


Figure 3. A) Chromatograms of the mass signal of mouse liver DNA. UHPLC-QQQ data obtained for $m^5\text{dC}$ and $m^4\text{dC}$ (left), hmdC (middle), and $m^6\text{dA}$ (right, blue line) and their corresponding isotopic standards are shown. Additionally, for $m^6\text{dA}$, the chromatogram for the digest blank is shown (red line) and a computed baseline (black line), which was determined by subtracting the digest blank from the sample. B) A chromatogram of the mass signal from UHPLC-QQQ data obtained for $m^6\text{dA}$ in DNA from HeLa cells.

and shows the expected fragmentation of the glycosidic bond, providing the fragmentation signature $m/z = 262.12 \rightarrow 146.07$, which allowed assignment of the signal. The naturally occurring hmdC is detected at exactly the same retention time with a mass transition of $m/z = 258.11 \rightarrow 142.06$, thus unequivocally demonstrating the presence of hmdC in murine liver DNA. Regarding the different monomethylated dC compounds $m^5\text{dC}$ and $m^4\text{dC}$ (left column), the epigenetic DNA mark $m^5\text{dC}$ is clearly detected at a retention time of 3.2 min, but for $m^4\text{dC}$ with a retention time of 3.5 min, there is obviously no signal present.

The $m^6\text{dA}$ data are highly interesting (Figure 3 right column). While the $D_3\text{-}m^6\text{dA}$ standard was clearly detectable at a retention time of 10.1 min, the unlabelled $m^6\text{dA}$ provided a very weak signal. We then performed a control experiment

to determine the limit of detection and investigated the digestion solution alone, which contains all of the commercial enzymes but no isolated DNA (red chromatogram). A weak signal for $m^6\text{dA}$ was again detectable at a retention time of 10.1 min. After subtracting this background signal (red) from the measured chromatogram (blue), we obtained the black line showing that the original signal at 10.1 min is purely caused by background derived from the enzyme mixture. Here it is important to note that most of these proteins are recombinant proteins obtained from bacterial expression systems and bacterial DNA contains plenty of $m^6\text{dA}$. To support the evidence that $m^6\text{dA}$ is not present in vertebrate DNA, we performed the same experiment with HeLa cells and also observed no signal for $m^6\text{dA}$ (Figure 3B).

We then determined our limit of detection for $m^6\text{dA}$ to be 3.5×10^{-7} per dN, which corresponds to 170 $m^6\text{dA}$ bases per murine genome. This is a very small number and demonstrates the excellent sensitivity of our method. It shows that the maximum number of $m^6\text{dA}$ that could be theoretically present and would not be detectable by our method is less than 170 $m^6\text{dA}$ bases per genome, which led us to conclude that $m^6\text{dA}$ is likely not epigenetically relevant but rather formed as a DNA lesion, perhaps by misguided methyltransferases. Spiking tests with synthetic nucleoside and

DNA from *Chlamydomonas* nevertheless confirmed the sensitivity of the method, since the input amount equalled the found amount (see Figure S6).

To find a potential source for $m^6\text{dA}$ in mESC DNA that could explain previous sequencing data,^[17] we thought that $m^6\text{dA}$ -containing bacterial DNA that gets degraded could provide the $m^6\text{dA}$ nucleoside, which then might get incorporated into mESC DNA. This is indeed a possibility. When we added the $m^6\text{dA}$ nucleoside to a mESCs culture, we indeed saw incorporation of some $m^6\text{dA}$ into the genome (Figure S7).

We also cannot fully exclude the possibility that the presence of few m⁶dA bases, at levels below our detection limit, could have a biological function. In addition, it is possible that at certain stages of organismal development, certain methyltransferases are activated that may induce high m⁶dA levels at specific time points that may have escaped our detection.^[18] Our data, however, show clearly that the maximum possible levels of m⁶dA in the analysed organisms and mESCs under normal conditions are far lower than so far believed.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for financial support via SFB1032 (TP-A5), SFB749 (TP-A4), SPP1784, CA275 and the Excellence Cluster CiPS^M.

Conflict of interest

The authors declare no conflict of interest.

Keywords: DNA · epigenetics · methyldeoxyadenosine · methyldeoxycytidine · mass spectrometry

- [1] J. A. Law, S. E. Jacobsen, *Nat. Rev. Genet.* **2010**, *11*, 204–220.
- [2] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, *Science* **2009**, *324*, 930–935.
- [3] S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929–930.
- [4] T. Pfaffeneder, B. Hackner, M. Truss, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 7008–7012; *Angew. Chem.* **2011**, *123*, 7146–7150.
- [5] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, *Science* **2011**, *333*, 1300–1303.
- [6] Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, G. L. Xu, *Science* **2011**, *333*, 1303–1307.
- [7] R. M. Kohli, Y. Zhang, *Nature* **2013**, *502*, 472–479.
- [8] S. Schiesser, T. Pfaffeneder, K. Sadeghian, B. Hackner, B. Steigenberger, A. S. Schröder, J. Steinbacher, G. Kashiwazaki, G. Höfner, K. T. Wanner, C. Ochsenfeld, T. Carell, *J. Am. Chem. Soc.* **2013**, *135*, 14593–14599.
- [9] M. Su, A. Kirchner, S. Stazzoni, M. Müller, M. Wagner, A. Schröder, T. Carell, *Angew. Chem. Int. Ed.* **2016**, *55*, 11797–11800; *Angew. Chem.* **2016**, *128*, 11974–11978.
- [10] M. Yu, L. Ji, D. A. Neumann, D. H. Chung, J. Groom, J. Westpheling, C. He, R. J. Schmitz, *Nucleic Acids Res.* **2015**, *43*, e148.
- [11] M. Ehrlich, M. A. Gama-Sosa, L. H. Carreira, L. G. Ljungdahl, K. C. Kuo, C. W. Gehrke, *Nucleic Acids Res.* **1985**, *13*, 1399–1412.
- [12] Y. Fu, G. Z. Luo, K. Chen, X. Deng, M. Yu, D. Han, Z. Hao, J. Liu, X. Lu, L. C. Dore, X. Weng, Q. Ji, L. Mets, C. He, *Cell* **2015**, *161*, 879–892.
- [13] E. L. Greer, M. A. Blanco, L. Gu, E. Sendinc, J. Liu, D. Aristizabal-Corralles, C. H. Hsu, L. Aravind, C. He, Y. Shi, *Cell* **2015**, *161*, 868–878.
- [14] G. Zhang, H. Huang, D. Liu, Y. Cheng, X. Liu, W. Zhang, R. Yin, D. Zhang, P. Zhang, J. Liu, C. Li, B. Liu, Y. Luo, Y. Zhu, N. Zhang, S. He, C. He, H. Wang, D. Chen, *Cell* **2015**, *161*, 893–906.
- [15] S. Hattman, C. Kenny, L. Berger, K. Pratt, *J. Bacteriol.* **1978**, *135*, 1156–1157.
- [16] M. J. Koziol, C. R. Bradshaw, G. E. Allen, A. S. Costa, C. Frezza, J. B. Gurdon, *Nat. Struct. Mol. Biol.* **2016**, *23*, 24–30.
- [17] T. P. Wu, T. Wang, M. G. Seetin, Y. Lai, S. Zhu, K. Lin, Y. Liu, S. D. Byrum, S. G. Mackintosh, M. Zhong, A. Tackett, G. Wang, L. S. Hon, G. Fang, J. A. Swenberg, A. Z. Xiao, *Nature* **2016**, *532*, 329–333.
- [18] J. Liu, Y. Zhu, G. Z. Luo, X. Wang, Y. Yue, X. Wang, X. Zong, K. Chen, H. Yin, Y. Fu, D. Han, Y. Wang, D. Chen, C. He, *Nat. Commun.* **2016**, *7*, 13052.
- [19] W. Huang, J. Xiong, Y. Yang, S.-M. Liu, B.-F. Yuan, Y.-Q. Feng, *RSC Adv.* **2015**, *5*, 64046–64054.
- [20] G. Z. Luo, F. Wang, X. Weng, K. Chen, Z. Hao, M. Yu, X. Deng, J. Liu, C. He, *Nat. Commun.* **2016**, *7*, 11301.

Manuscript received: January 13, 2017

Revised: February 15, 2017

Final Article published: ■■■■■, ■■■■■

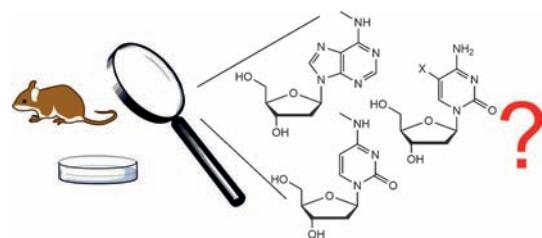
Communications



Epigenetics

S. Schiffers, C. Ebert, R. Rahimoff,
 O. Kosmatchev, J. Steinbacher,
 A.-V. Bohne, F. Spada, S. Michalakis,
 J. Nickelsen, M. Müller,
 T. Carell*

Quantitative LC-MS Provides No Evidence for m⁶dA or m⁴dC in the Genome of Mouse Embryonic Stem Cells and Tissues



Previous results challenged: Highly sensitive mass spectrometry reveals that m⁶dA is most likely not an epigenetic base in the mouse genome. It appears

that in vertebrates, the levels of this modified base, and that of N⁴-methyl-deoxycytidine, are much lower than previously thought.



Supporting Information

Quantitative LC–MS Provides No Evidence for m⁶dA or m⁴dC in the Genome of Mouse Embryonic Stem Cells and Tissues

*Sarah Schiffers, Charlotte Ebert, René Rahimoff, Olesea Kosmatchev, Jessica Steinbacher, Alexandra-Viola Bohne, Fabio Spada, Stylianos Michalakis, Jörg Nickelsen, Markus Müller, and Thomas Carell**

anie_201700424_sm_miscellaneous_information.pdf

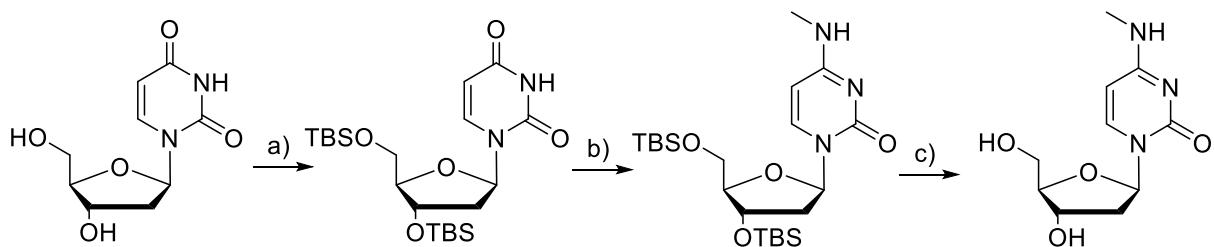
Supporting information

General Methods

Chemicals were purchased from Sigma-Aldrich, Fluka, Alfa Aesar, ChemGenes or Acros and used without further purification. Solutions were concentrated *in vacuo* on a Heidolph rotary evaporator. The solvents were of reagent grade and purified by distillation. Dry solvents were bought from Sigma-Aldrich. Acetonitrile for HPLC-purification was purchased from VWR. Water was purified by a Milli-Q Plus system from Merck Millipore. Chromatographic purification of products was accomplished using flash column chromatography on Merck Geduran Si 60 (40-63 µM) silica gel (normal phase). Thin layer chromatography (TLC) was performed on Merck 60 (silica gel F254) plates. ^1H and ^{13}C -NMR spectra were recorded in deuterated solvents on *Bruker ARX 300*, *Varian VXR400S*, *Varian Inova 400*, *Bruker AMX 600* and *Bruker Avance 800* spectrometers and calibrated to the residual solvent peak. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, hept = septet, m = multiplet, br. = broad. For assignment of the structures, additional 2D NMR spectra (COSY, HSQC, HMBC) were measured. High resolution electrospray ionization mass spectra (HRMS-ESI) were recorded on a Thermo Finnigan LTQ-FT (ESI-FTICR).

Synthesis of 2'-deoxy-4-methyl-cytidine

The final compounds were purified by reverse phase HPL chromatography (Macherey-Nagel, Nucleosil 100-7 C18, 10 × 250 mm, linear gradient from 0% to 10% acetonitrile in water in 45 min for 2'-deoxy-4-methyl-cytidine or 0% to 15% acetonitrile in water in 45 min for 2'-deoxy-6-methyladenosine).



Scheme S1: Synthesis of 2'-deoxy-4-methyl-cytidine. TBS-Cl, imidazole, pyridine, 0 °C to r.t., o/n, quant b) 1. NaH, THF, r.t., 45 min, 2. TPS-Cl, r.t., 15 h, 3. MeNH₂, r.t., 5 min, 86% c) HF in pyridine, EtOAc, r.t., o/n, quant.

Synthesis of 2'-deoxy-4-methyl-cytidine

O^{3'},O^{5'}-TBS-2'-deoxyuridine

2'-Deoxyuridine (3.00 g, 13.1 mmol, 1.0 eq) was dissolved in pyridine (50 mL) and cooled to 0 °C. Imidazole (4.00 g, 59.0 mmol, 4.5 eq) and TBS-Cl (5.90 g, 39.3 mmol, 3.0 eq) were then added. The mixture was allowed to warm to rt overnight and volatiles were subsequently removed *in vacuo*. The gummy residue was taken up in DCM (100 mL) and washed with sat. NH₄Cl (3x 100 mL). The organic phase was dried over Na₂SO₄, the organic solvent removed under reduced pressure and the resulting oil was finally coevaporated with toluene (3x) to yield a white powder in quantitative yield (5.98 g, 13.1 mmol) that was used without further purification.

O^{3'},O^{5'}-TBS-4-methyl-2'-deoxycytidine

O^{3'},O^{5'}-TBS-2'-deoxyuridine (1.00 g, 2.19 mmol, 1.0) was dissolved in THF (50 mL) and cooled to 0 °C. NaH (60% in mineral oil, 434 mg, 10.95 mmol, 5.0 eq) was added in one portion and the grey suspension was kept at 0 °C for 45 minutes. Subsequently, TPS-Cl (1.33 g, 4.38 mmol, 2.0 eq) was added and the mixture was warmed to rt and stirred for 15 hours. After all the starting material was consumed (as judged by TLC), ice cold sat. NH₄Cl (200 mL) was added cautiously and the aqueous layer was extracted with EtOAc (2x 100 mL). After drying the organic phase over Na₂SO₄, volatiles were removed *in vacuo* and the residue was charged with MeNH₂ (50 mL, 33% in EtOH). After stirring at rt for 5 minutes, the reaction was aborted through addition of NH₄Cl (200 mL) and the aqueous phase was again extracted with EtOAc (2x 50 mL). After drying over Na₂SO₄ and removal of EtOAc *in vacuo*, the compound was purified by column chromatography (DCM --> 2.5% MeOH --> 4% MeOH) and the product was yielded as a colourless foam (887 mg, 1.80 mmol, 86%).

¹H-NMR (599 MHz, CDCl₃, ppm): δ = 7.87 (d, *J* = 7.8 Hz, 1H, 6-H), 6.29 (dd, *J* = 7.8 Hz, 6.3 Hz, 1H, 1'-H), 5.55 (d, *J* = 7.8 Hz, 1H, 5-H), 4.35 (q, *J* = 5.5 Hz, 1H, 3'-H), 3.90 (dd, *J* = 11.4 Hz, 2.6 Hz, 1H, 5'-H), 3.86 (m, 1H, 4'-H), 3.75 (dd, *J* = 11.1 Hz, 2.3 Hz, 1H, 5'-H), 3.0 (ds, *J* = 4.9 Hz, 3H, CH₃-NH), 2.44 - 2.36 (m, 1H, 2'-H), 2.07 (ddd, *J* = 13.5, 6.4, 4.9 Hz, 1H, 2'-H), 0.96 - 0.82 (m, 18H, TBS-C(CH₃)₃), 0.14 - 0.01 (m, 12H, TBS-CH₃).

¹³C-NMR (151 MHz, CDCl₃, ppm): δ = 164.2 (C4), 156.2 (C2), 139.5 (C6), 94.4 (C5), 87.1 (C4'), 85.7 (C3'), 70.1 (C1'), 61.9 (C5'), 42.1 (C2'), 28.0 (CH₃), 25.9 - 25.7 (TBS-C(CH₃)₃), 18.3 - 18.0 (TBS-C(CH₃)₃), -4.6 - -5.6 (TBS-CH₃).

HR-MS (ESI+): m/z calculated for [C₂₂H₄₄N₃O₄Si₂⁺] ([M+H]⁺): 470.2865, found: 470.2864.

2'-deoxy-4-methyl-cytidine

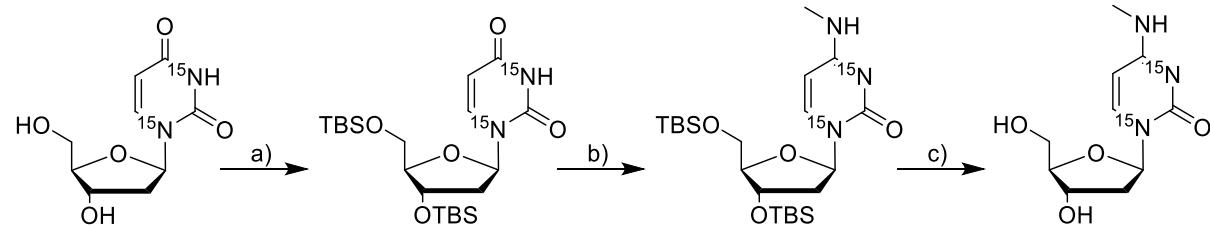
In a falcon tube, the TBS-protected nucleoside (300 mg, 0.64 mmol, 1.0 eq) was dissolved in EtOAc (5 mL). The mixture was then charged with HF•pyridine (70%, 374 µL, 14.4 mmol, 22.5 eq) and stirred at rt overnight. Then TMS•OMe (3.46 mL, 25.6 mmol, 40 eq) was added whereas a colorless precipitate was formed. The white solid was isolated through repeated centrifugation of the mixture (2x 10 minutes, 6000 rpm) and subsequent washing with EtOAc (2x 5 mL). The supernatant was discarded and the compound (154 mg, 0.64 mmol, quant.) was yielded as a white powder, that was further purified by HPLC.

¹H NMR (400 MHz, D₂O) δ 7.68 (d, *J* = 7.6 Hz, 1H, C6-H), 6.29 (t, *J* = 6.7 Hz, 1H, C1'-H), 5.98 (d, *J* = 7.6 Hz, 1H, C5-H), 4.45 (dt, *J* = 6.5, 3.9 Hz, 1H, C3'-H), 4.05 (dt, *J* = 5.1, 3.7 Hz, 1H, C4'-H), 3.84 (dd, *J* = 12.4, 3.7 Hz, 1H, C5'-H), 3.76 (dd, *J* = 12.4, 5.2 Hz, 1H, C5'-H), 2.89 (s, 3H, CH₃), 2.41 (ddd, *J* = 14.1, 6.4, 4.0 Hz, 1H, C2'-H), 2.29 (dt, *J* = 13.9, 6.8 Hz, 1H, C2'-H) ppm.

¹³C NMR (101 MHz, D₂O) δ 164.5 (C4), 157.8 (C2), 139.3 (C6), 97.3 (C5), 86.4 (C4'), 85.8 (C1'), 70.6 (C3'), 61.3 (C5'), 39.1 (C2'), 27.1 (CH₃) ppm.

HR-MS (ESI+): m/z calculated for [C₁₀H₁₆N₃O₄]⁺ ([M+H]⁺): 242.1136, found: 242.1137.

Synthesis of 2'-deoxy-4-methyl-[N¹,N³-¹⁵N₂]-cytidine



Scheme S2: Synthesis of 2'-deoxy-4-methyl-[N¹,N³-¹⁵N₂]-cytidine. a) TBS-Cl, imidazole, pyridine, r.t., 72 h, 64% b) 1. NaH, THF, r.t., 45 min, 2. TPS-Cl, r.t., 17 h, 3. MeNH₂, r.t., 10 min, 49% c) HF in pyridine, EtOAc, r.t., 17 h, 64%.

O³',O⁵'-TBS-2'-deoxy-[N¹,N³-¹⁵N₂]-uridine

2'-Deoxy-[¹⁵N,¹⁵N]-uridine (51 mg, 0.22 mmol, 1.0 eq), TBS-chloride (109 mg, 0.726 mmol, 3.3 eq) and imidazole (99 mg, 1.4 mmol, 6.6 eq) were dissolved in 1 mL pyridine and stirred at

r.t. for 72 h. The solvent was removed *in vacuo*, the residue was dissolved in DCM and washed with sat. NaHCO₃ (aq). *O*^{3'},*O*^{5'}-TBS-2'-deoxy-[*N*¹,*N*³-¹⁵N₂]-uridine (59 mg, 0.13 mmol, 64%) was obtained as a white solid.

HR-MS (ESI+): m/z calculated for [C₂₁H₄₁¹⁵N₂O₅Si₂]⁺ ([M+H]⁺): 459.2489, found: 459.2493.

***O*^{3'},*O*^{5'}-TBS-2'-deoxy-[*N*¹,*N*³-¹⁵N₂]-cytidine**

O^{3'},*O*^{5'}-TBS-2'-deoxy-[*N*¹,*N*³-¹⁵N₂]-uridine (59 mg, 0.13 mmol, 1.0 eq) was dissolved in THF (2 mL) and cooled to 0 °C. Sodium hydride (16 mg, 0.65 mmol, 5.0 eq) was added and the mixture was stirred at r.t. for 45 minutes. Then TPS-chloride (79 mg, 0.26 mmol, 2.0 eq) was added and the mixture was stirred at r.t. for another 17 h. The reaction was quenched with saturated NH₄Cl_(aq), extracted with EtOAc and washed three times with saturated NH₄Cl_(aq). The organic phase was dried with Na₂SO₄ and the solvent was removed *in vacuo*. The residue was dissolved in MeNH₂ (5 mL, 33% in EtOH) and stirred at r.t. for 10 minutes. The solvent was removed *in vacuo* and the residue was purified via column chromatography (DCM/MeOH 0.5% to 1%, SiO₂). *O*^{3'},*O*^{5'}-TBS-2'-deoxy-4-methyl-[*N*¹,*N*³-¹⁵N₂]-cytidine (30 mg, 0.064 mmol, 49%) was obtained as a white solid. Unreacted *O*^{3'},*O*^{5'}-TBS-2'-deoxy-[*N*¹,*N*³-¹⁵N₂]-uridine could be recovered via column chromatography.

HR-MS (ESI+): m/z calculated for [C₂₂H₄₄N¹⁵N₂O₄Si₂]⁺ ([M+H]⁺): 472.2806, found: 472.2804.

2'-deoxy-4-methyl-[*N*¹,*N*³-¹⁵N₂]-cytidine

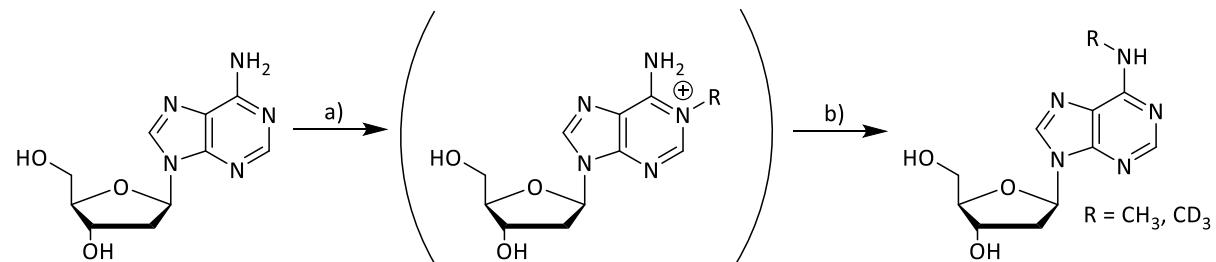
O^{3'},*O*^{5'}-TBS-2'-deoxy-4-methyl-[*N*¹,*N*³-¹⁵N₂]-cytidine (30 mg, 0.064 mmol, 1.0 eq) was dissolved in EtOAc (3 mL), HF in pyridine (40 µL, 1.4 mmol, 22.5 eq) was added and the reaction mixture was stirred at r.t. for 17 h. The mixture was cooled to 0 °C, quenched with TMS-OMe and stirred at r.t. for 4 h. The precipitation was centrifuged and the supernatant was removed and the solid was washed three times with EtOAc. The solvent was removed *in vacuo* and 2'-deoxy-4-methyl-[*N*¹,*N*³-¹⁵N₂]-cytidine (10 mg, 0.041 mmol, 64%) was obtained as a white solid.

¹H NMR (599 MHz, D₂O) δ 7.68 (d, *J* = 7.4 Hz, 1H, C6-H), 6.29 (d, *J* = 7.7 Hz, 1H, C1'-H), 5.99 (dd, *J* = 7.8, 3.9 Hz, 1H, C5-H), 4.45 (h, *J* = 4.1, 3.5 Hz, 1H, C3'-H), 4.05 (m, 1H, C4'-H), 3.84 (m, 1H, C5'-H), 3.77 (m, 1H, C5'-H), 2.90 (d, *J* = 3.3 Hz, 3H, CH₃), 2.41 (m, 1H, C2'-H), 2.31 (q, *J* = 7.3 Hz, 1H, C2'-H) ppm.

¹³C NMR (101 MHz, D₂O) δ 164.4 (d, *J* = 9.7 Hz, C4), 157.8 (C2), 139.2 (d, *J* = 13.2 Hz, C6), 97.3 (C5), 86.4 (C4'), 85.7 (d, *J* = 11.7 Hz, C1'), 70.6 (C3') , 61.3 (C5') , 39.0 (C2') , 27.0 (d, *J* = 2.0 Hz, CH₃) ppm.

HR-MS (ESI+): m/z calculated for [C₁₀H₁₆N¹⁵N₂O₄]⁺ ([M+H]⁺): 244.1076, found: 244.1078.

Synthesis of 2'-deoxy-6-methyladenosine



Scheme S3: Synthesis of 2'-deoxy-6-methyladenosine and 2'-deoxy-6-[D₃]-methyladenosine.

a) Methyl-p-toluenesulfonate, DMF, r.t., 20 h (R=CH₃); CD₃I, *N,N*-dimethylacetamide 28 °C, 20 h (R=CD₃), b) 2 M NaOH, 28 °C, 1 h, 13% (R=CH₃), 14% (R=CD₃).

2'-Deoxy-6-methyladenosine was synthesized as published in ^[1].

Methyl-*p*-toluenesulfonate (600 mg, 3.22 mmol, 4.0 eq) was dissolved in 1 mL *N,N*-dimethylformamide and 2'-Deoxyadenosine (200 mg, 0.796 mmol, 1.0 eq) was added. The solution was stirred at r.t. for 20 h. Celite was added and the solution was filtered. Acetone (30 mL) was added to the filtrate and the mixture was stirred for 1 h. The precipitation was filtered, washed with acetone and dried on high vacuum. The white solid was dissolved in 5 mL 2 M NaOH and stirred at 28 °C for 1 h. The solution was neutralized with 10% aqueous *p*-toluenesulfonic acid and the solvent was removed *in vacuo*. The residue was purified via column chromatography (DCM/MeOH 10:1 → 9:1, SiO₂) and 2'-deoxy-6-methyladenosine (27 mg, 0.10 mmol, 13%) was obtained as a white solid.

¹H NMR (800 MHz, DMSO-*d*₆) δ 8.32 (s, 1H, C8-H), 8.22 (s, 1H, C2-H), 7.77 (s, 1H, NH), 6.35 (dd, *J* = 7.9, 6.2 Hz, 1H, C1'-H), 5.30 (d, *J* = 4.0 Hz, 1H, C3'-OH), 5.24 (t, *J* = 5.9 Hz, 1H, C5'-OH), 4.41 (dq, *J* = 6.3, 3.1 Hz, 1H, C3'-H), 3.88 (q, *J* = 3.8 Hz, 1H, C4'-H), 3.62 (dq, *J* = 13.2, 4.4 Hz, 1H, C5'-H), 3.52 (ddt, *J* = 10.8, 7.7, 4.0 Hz, 1H, C5'-H), 2.95 (s, 3H, CH₃), 2.72 (ddd, *J* = 13.4, 7.7, 5.6 Hz, 1H, C2'-H), 2.26 (ddd, *J* = 13.2, 6.2, 2.8 Hz, 1H, C2'-H) ppm.

¹³C NMR (201 MHz, DMSO-*d*₆) δ 155.0 (C6), 152.4 (C2), 147.9 (C4), 139.3 (C8), 119.8 (C5), 88.0 (C4'), 84.0 (C1'), 71.0 (C3'), 61.9 (C5'), 39.4 (C2'), 27.0 (CH₃) ppm.

HR-MS (ESI+): m/z calculated for [C₁₁H₁₆N₅O₃]⁺ ([M+H]⁺): 266.1248, found: 266.1248.

Synthesis of 2'-deoxy-6-[D₃]-methyladenosine

2'-Deoxyadenosine (200 mg, 0.796 mmol, 1.0 eq) was dissolved in 1 mL *N,N*-dimethylacetamide and CD₃I (0.15 mL, 2.4 mmol, 3.0 eq) was added. The yellow solution was stirred at 28 °C for 20 h. Acetone was added and the precipitation was filtered, washed with acetone and dried on high vacuum. The white solid was dissolved in 5 mL 2 M NaOH and stirred at 28 °C for 1 h. The solution was neutralized with 10% aqueous *p*-toluenesulfonic acid and the solvent was removed *in vacuo*. The residue was purified via column chromatography (DCM/MeOH 10:1 → 9:1, SiO₂) and 2'-deoxy-6-[D₃]-methyladenosine (29 mg, 0.11 mmol, 14%) was obtained as a white solid.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.32 (s, 1H, C8-H), 8.22 (s, 1H, C2-H), 7.74 (s, 1H, NH), 6.35 (dd, *J* = 7.9, 6.1 Hz, 1H, C1'-H), 5.30 (d, *J* = 4.0 Hz, 1H, C3'-OH), 5.24 (s, 1H, C5'-OH), 4.41 (dq, *J* = 6.0, 2.9 Hz, 1H, C3'-H), 3.88 (td, *J* = 4.2, 2.5 Hz, 1H, C4'-H), 3.62 (m, 1H, C5'-H), 3.52 (d, *J* = 11.6 Hz, 1H, C5'-H), 2.72 (ddd, *J* = 13.4, 8.0, 5.7 Hz, 1H, C2'-H), 2.26 (ddd, *J* = 13.1, 6.1, 2.9 Hz, 1H, C2'-H) ppm.

¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.0 (C6), 152.3 (C2), 147.8 (C4), 139.2 (C8), 119.7 (C5), 88.0 (C4'), 83.9 (C1'), 70.9 (C3'), 61.9 (C5'), 39.4 (C2'), 26.4 (CD₃) ppm.

HR-MS (ESI+): m/z calculated for [C₁₁H₁₃D₃N₅O₃]⁺ ([M+H]⁺): 269.1436, found: 269.1437.

Cyanobacterial and Algal Strains, Culture Conditions and DNA Extraction

For DNA extractions, 150 mL of wild-type *Synechocystis* sp. PCC6803 cultures were grown at 30 °C in BG11 medium containing 5 mM glucose.^[2] For *Chlamydomonas reinhardtii* the cell-wall-deficient strain 3491 (*cw15 mt-*), obtained from the *Chlamydomonas* Center (<http://www.chlamy.org/>), and the cell-walled laboratory strain 7d+ were cultured in 150-300 mL Tris-acetate-phosphate medium containing 1% sorbitol^[3] at 25°C. All cultures were grown under orbital shaking and continuous irradiation with 30 μmol of photons m⁻² s⁻¹.

For determination of *N*⁶-methyldeoxyadenosine and *N*⁴-methyldeoxycytidine, DNA was isolated by harvesting cells at 3000 g for 20 min, followed by resuspension of the pellets in 4-8 mL DNA extraction buffer ((2% (w/v) cetyltrimethylammonium bromide (CTAB), 100 mM Tris-HCl, pH 8; 1.4 M NaCl; 20 mM EDTA pH 8; 2% (v/v) β-mercaptoethanol)). Cells were lysed upon addition of ~300 μL glass beads for 1 h at 65 °C and vortexed every 15 min. Proteins were removed by extraction with phenol:chloroform:isoamylalcohol 25:24:1 solution. RNA in the water phase was digested at 37 °C for 1 h using 0.1 mg/mL RNase A. RNase A was removed

by extraction with phenol:chloroform:isoamylalcohol 25:24:1 solution, followed by a chloroform/isoamylalcohol 24:1 extraction and isopropanol precipitation of the DNA. DNAs were dissolved in appropriate volumes of water at 4 °C overnight.

Preparation of mouse tissues

All mice used were on the wild-type C57-BL6 genetic background. All procedures concerning animals were performed with permission of the local authority (*Regierung von Oberbayern*). Brain und liver tissues of male wt mice (C57-BL6/N) were dissected at postnatal day 51 and immediately placed in 2 mL *Eppendorf*-tubes, snap frozen in liquid nitrogen and stored at -80 °C until use.

Genomic DNA isolation

Genomic DNA from murine tissues was extracted using the Qiagen Blood and Cell Culture DNA Midi Kit. Extraction was performed following the manufacturer's instructions for genomic DNA isolation from tissue samples. All buffers until loading of the sample on Genomic-tip 100/G were additionally supplemented with antioxidants 3,5-di-*tert*-butyl-4-hydroxytoluene (BHT, 200 µM) and desferoxamine mesylate salt (desferal, 200 µM). G2-buffer was additionally supplemented with the deaminase inhibitor tetrahydouridine (THU, 200 µM), according to published methods, to reduce background oxidation or deamination.^[4] Elution buffer QF was supplemented with 200 µM BHT.

mESC and HeLa samples differentiated in the presence of growth factors were lysed directly in the plates with RLT buffer (Qiagen) supplemented with BHT and desferal as described above. DNA was isolated using the Zymo Quick gDNA Midi Kit according to the manufacturer's instruction, except that elution was performed with 100 µl of bidest. water supplemented with BHT (0.2 µM).

Cell culture

Feeder independent wt J1 cells (strain 129S4/SvJae)^[5] were cultured in the presence of serum and LIF as previously described^[6] and routinely maintained on gelatinized plates in DMEM (PAA or Sigma) supplemented with 10% FBS, 1×MEM-nonessential amino acids (NEAA), 0.2 mM *L*-alanyl-*L*-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin (all from PAA), 0.1 mM β-mercaptoethanol, 20 ng/ml ($\geq 1 \times 10^3$ U/ml) mouse recombinant LIF (ORF Genetics), 1 µM PD 0325901 and 3 µM CHIR 99021 (2i; both from Axon Medchem). Before DNA isolation, 2i cultures were passaged twice (over 4 d and 5 d, respectively) in DMEM

supplemented with FBS and LIF as above but lacking 2i. With this strategy, primed mESC cultures were obtained with no sign of overt differentiation.

As an additional experiment for determination of m⁶dA as a marker of early embryogenesis, we cultured wt J1 cells in serum-free 2i/LIF on GelTrex-coated plates for more than 5 passages and harvested them with RLT buffer (Qiagen) supplemented with BHT and desferal.

Feeder independent wt TT2 cells were cultured like wt J1 cells (as described above) and primed for 4 d, 10 d, 12 d and 14 d respectively. Feeding experiments with dimethylsulfate (DMS, <0.01% DMSO) were performed on day four of the priming process and the cells were incubated with DMS for 24 h.

For isotope tracing with heavy methionine in serum-primed mESCs, 2i cultures of wt J1 cells were passaged twice (over 4 d or 5 d) without 2i in L-Metfree DMEM (Sigma) supplemented as above and with 0.2 mM of either [*methyl*-¹³C,D₃]L-Met or natural L-Met. The medium was changed after 24 h. The experiment was performed in a biological duplicate.

Feeding of wt J1 cells with m⁶dA was performed once by adding m⁶dA nucleoside (1 μM, 0.1% DMSO) to the medium on day 0 of culturing in FBS/LIF and increasing the concentration on day 2 for another two days (1 mM, 1% DMSO). DMSO controls were only treated with the respective concentration of DMSO.

HeLa cells were cultivated at 37 °C in water saturated, CO₂-enriched (5%) atmosphere. DMEM (10% FBS) was used as growing medium. When reaching a confluence of 70% to 80% the cells were passaged or harvested with RLT buffer (Qiagen) supplemented with BHT and desferal, respectively.

Digest of genomic DNA

1 μg (10 μg) of genomic DNA in 35 μl H₂O were digested as follows: An aqueous solution (7.5 μl) of 480 μM ZnSO₄, containing 42 U (63 U) nuclease S1 (*Aspergillus oryzae*, Sigma-Aldrich), 5 U (7.5 U) Antarctic phosphatase (New England BioLabs) and specific amounts of labeled internal standards were added, and the mixture was incubated at 37 °C for 3 h. After addition of 7.5 μl of a 520 μM [Na]₂-EDTA solution, containing 0.15 U (0.2 U) snake venom phosphodiesterase I (*Crotalus adamanteus*, USB corporation), the sample was incubated for another 3 h (overnight) at 37 °C. Prior to LC-MS/MS analysis, samples were filtered by using an AcroPrep Advance 96 filter plate 0.2 μm Supor (Pall Life Sciences).

For dilution experiments, m⁶dA and m⁴dC were not spiked during the digest step, but after a dilution of the digest of 1:20, or 1:200 respectively.

Furthermore, a recovering experiment was performed as follows: a digest of 10 µg of murine DNA was spiked with a small amount of *Chlamydomonas* (CC-3491) DNA or synthetic m⁶dA-nucleoside, which corresponds to a level of m⁶dA as described earlier in mouse tissue.^[7]

LC/MS analysis of DNA samples

Quantitative LC/UV-ESI-MS/MS analysis of digested DNA samples was performed using an Agilent 1290 UHPLC system equipped with a UV detector and an Agilent 6490 triple quadrupole mass spectrometer coupled with the stable isotope dilution technique. An improved method, based on earlier published work,^[8] was developed, which allowed the concurrent analysis of all nucleosides in one single analytical run. The source-dependent parameters were as follows: gas temperature 80 °C, gas flow 15 l/min (N₂), nebulizer 30 psi, sheath gas heater 275 °C, sheath gas flow 11 l/min (N₂), capillary voltage 2,500 V in the positive ion mode, capillary voltage -2,250 V in the negative ion mode and nozzle voltage 500 V. The fragmentor voltage was 380 V. Delta EMV was set to 500 (positive mode) and 800 (negative mode). Compound-dependent parameters are summarized in Table S1.

Exact quantification of the nucleosides m⁵dC, hmdC, fdC, cadC and 8oxodG was performed with a previously published method.^[6]

Table S1: Compound-dependent LC-MS/MS-parameters used for the analysis of genomic DNA. CE: collision energy; CAV: collision cell accelerator voltage; EMV: electron multiplier voltage. The nucleosides were analyzed in the positive ($[M+H]^+$ -species) as well as in the negative ($[M-H]^-$ -species) ion selected reaction monitoring mode (SRM).

Compound	Precursor ion (m/z)	MS1 resolution	Product ion (m/z)	MS2 resolution	Dwell time [ms]	CE (V)	CAV (V)	Polarity
Time segment 1.5–6 min								
$[^{15}\text{N}_2]\text{-cadC}$	274.08	Wide	158.03	Wide	50	5	5	Positive
cadC	272.09	Wide	156.04	Wide	50	5	5	Positive
$[\text{D}_2\text{-}^{15}\text{N}_2]\text{-hmdC}$	262.12	Wide	146.07	Wide	50	27	1	Positive
hmdC	258.11	Wide	142.06	Wide	50	27	1	Positive
$[\text{D}_3]\text{-m}^5\text{dC}$	245.13	Wide	129.09	Wide	55	60	1	Positive
$[^{15}\text{N}_2]\text{-m}^4\text{dC}$	244.11	Wide	128.06	Wide	55	60	1	Positive
$\text{m}^5\text{dC/m}^4\text{dC}$	242.11	Wide	126.07	Wide	55	60	1	Positive
dC	228.1	Wide	112.05	Wide	1	1	0	Positive
$[^{15}\text{N}_2]\text{-cadC_2}$	158.03	Wide	140.09	Wide	50	13	7	Positive
cadC_2	156.04	Wide	138.03	Wide	50	13	7	Positive
Time segment 6–8 min								
$[\text{D}_2]\text{-hmdU}$	259.09	Wide	216.08	Wide	48	7	5	Negative
$[\text{D}_2]\text{-hmdU_2}$	259.09	Wide	126.05	Wide	48	7	5	Negative
hmdU	257.08	Wide	214.07	Wide	48	7	5	Negative
hmdU_2	257.08	Wide	124.04	Wide	48	7	5	Negative
$[^{15}\text{N}_2]\text{-fdU}$	257.06	Wide	213.05	Wide	48	6	5	Negative
$[^{15}\text{N}_2]\text{-fdU_2}$	257.06	Wide	141	Wide	48	6	5	Negative
fdU	255.06	Wide	212.06	Wide	48	6	5	Negative
fdU_2	255.06	Wide	140	Wide	48	6	5	Negative
$[^{15}\text{N}_2]\text{-dU}$	229.06	Wide	185.06	Wide	48	5	5	Negative
dU	229.06	Wide	185.06	Wide	48	5	5	Negative
Time segment 8–12 min								
$[^{15}\text{N}_5]\text{-8oxodG}$	289.09	Wide	173.04	Wide	60	9	7	Positive
8oxodG	284.1	Wide	168.05	Wide	60	9	7	Positive
$[\text{D}_3]\text{-m}^6\text{dA}$ sensitive	269.14	Wide	153.1	Wide	55	30	1	Positive
m ⁶ dA sensitive	266.12	Wide	150.08	Wide	55	30	1	Positive
$[\text{^{13}CD}_3]\text{-m}^6\text{dA}$	270.15	Wide	154.14	Wide	55	30	1	Positive
$[\text{D}_3]\text{-m}^6\text{dA}$ insensitive	269.14	Wide	153.1	Wide	55	60	5	Positive
m ⁶ dA insensitive	266.12	Wide	150.08	Wide	55	60	5	Positive
$[^{15}\text{N}_2]\text{-fdC}$	258.09	Wide	142.04	Wide	60	5	5	Positive
fdC	256.09	Wide	140.05	Wide	60	5	5	Positive
$[^{15}\text{N}_2]\text{-fdC_2}$	142.04	Wide	98.04	Wide	60	13	7	Positive
fdC_2	140.05	Wide	97.04	Wide	60	13	7	Positive

Chromatography was performed by a Poroshell 120 SB-C18 column (Agilent, 2.7 µm, 2.1 mm × 150 mm) at 30 °C using a previously published gradient^[8] of water and MeOH, each containing 0.0075% (v/v) formic acid, at a flow rate of 0.35 ml/min. The effluent up to 1.5 min and after 12 min was diverted to waste by a Valco valve. The autosampler was cooled to 4 °C. The injection volume was amounted to 39 µl. Calibration curves are shown in Fig S4.

Method validation and data processing were performed as described in earlier published work. Calibration curves were obtained by analyzing each standard (5–8 standard concentrations) as technical triplicate and applying linear regression Origin® 6.0 (Microcal™). Therefore, the ratio of the area under the curve (A/A^*) of the natural nucleoside to the internal standard was plotted against the ratio of the amount of substance (n/n^*) of the natural nucleoside to the internal standard (see Fig. S4). Additionally, acceptable accuracy (80–120%) as well as precision (<20% RSD) was required. Accuracy was proven by calculating the amount of substance n from the obtained A/A^* ratios of the calibration standards using the respective calibration function. Here, accuracy was defined as the ratio of the used amount of substance to the calculated amount of substance in percent and had to be between 80–120% for each standard concentration. Precision was determined as follows: technical replicates of A/A^* ratios for each calibration standard had to have relative standard deviations (RSD) smaller than 20%. The lower limit of quantification (LLOQ) was defined as the lowest concentration fulfilling the requirements of accuracy and precision and achieving a response of at least three times the response compared with the blank response. A compilation of absolute and relative LLOQs is shown in Table S2.

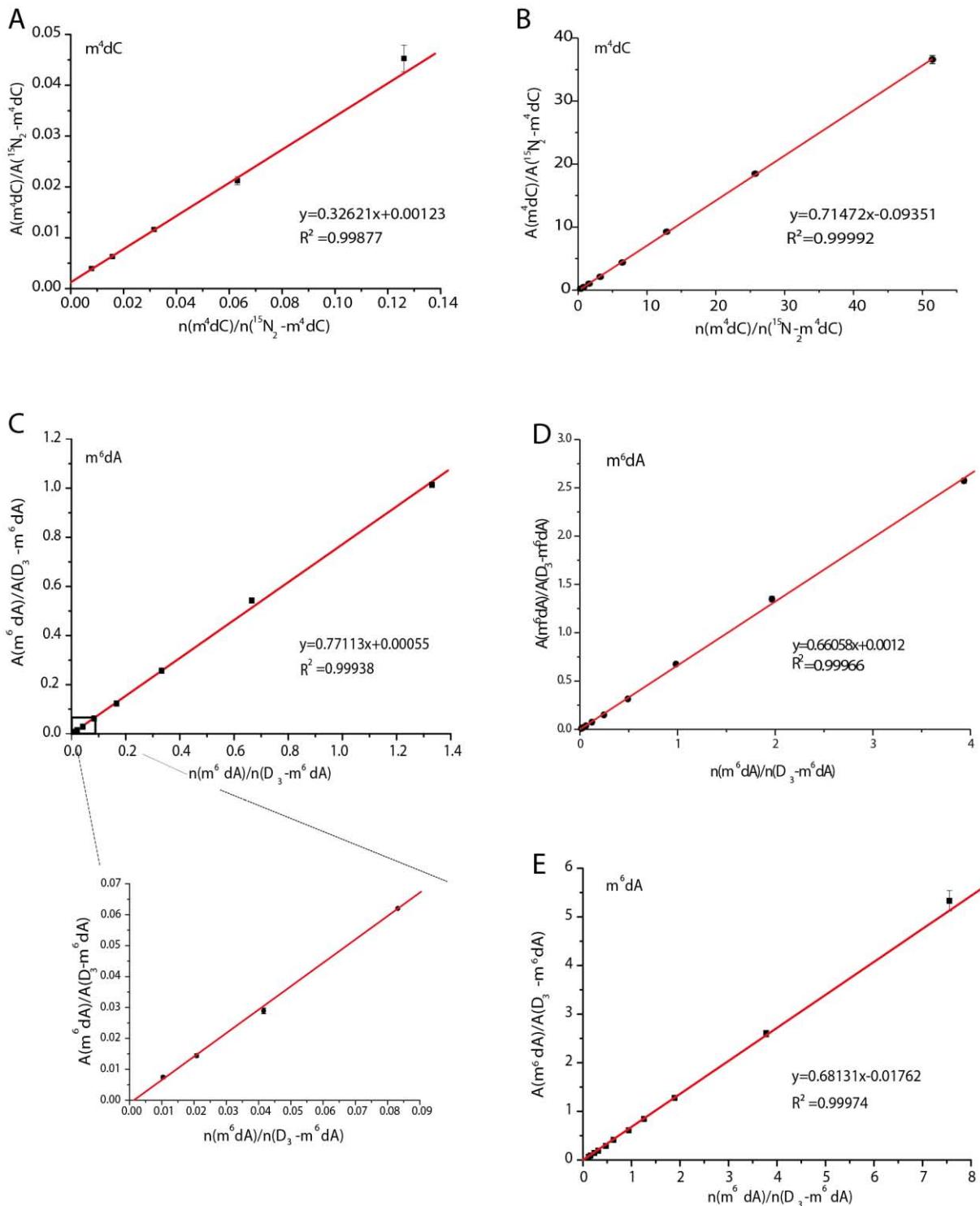


Figure S1: Calibration curves for m^4dC (upper row; A for murine, cell culture and *Chlamydomonas* DNA, B for *Synechocystis* DNA) and m^6dA (sensitive signals: C for mouse and cell culture DNA; insensitive signals: D for *Chlamydomonas* DNA and E for *Synechocystis* DNA); the ratio of the signal integral of unlabeled nucleoside towards labelled nucleoside is plotted against the ratio of the amount of natural nucleoside over the labelled nucleoside.

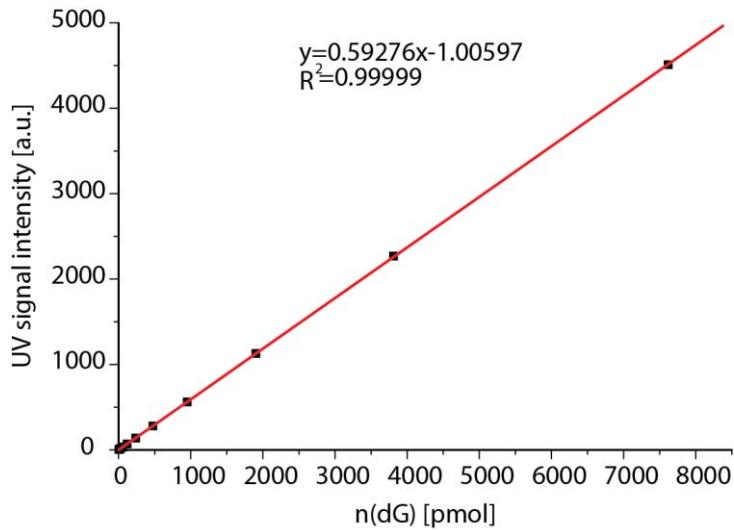


Figure S2: UV Calibration curve for dG with the UV signal intensity plotted against the amount of nucleoside [pmol].

Table S2: Compilation of absolute lower limits of quantification [fmol] (LLOQ; see Fig. S1) and relative LLOQs [per dN]. The relative LLOQs were computed by generating ratios of the absolute LLOQ [pmol] to the total amount of nucleosides (dN; [pmol]) in 10 µg of DNA (for calibration curve A and C) or 1 µg of DNA (for calibration curve B and D) or 50 ng (for calibration curve E). The total amount of nucleosides was obtained by calculating the content of dG using the respective calibration curve (see Fig. S2) and taking into account that murine DNA contains 21% dG, DNA of *Chlamydomonas* contains 32% dG and DNA of *Synechocystis* contains 23.9% dG.

	Absolute ULOQ [fmol]	Absolute LLOQ [fmol]	Relative ULOQ [per dN]	Relative LLOQ [per dN]	Relative LOD [per dN]	Absolute LOD [fmol]
m ⁴ dC (small, A)	302.1	18.8	1.1*10 ⁻⁵	6.6*10 ⁻⁷	1.4*10 ⁻⁶ (mouse/cell culture) 4.8*10 ⁻⁵ (<i>Chlamydomonas</i>)	40
m ⁴ dC (big, B)	48100	376	5.1*10 ⁻²	3.9*10 ⁻⁴	4.2*10 ⁻⁵	40
m ⁶ dA (sensitive, C)	203.8	1.6	7.4*10 ⁻⁶	5.8*10 ⁻⁸	8.0*10 ⁻⁸	2.2
m ⁶ dA (insensitive, D)	1162	18.2	1.4*10 ⁻³	2.2*10 ⁻⁵	2.0*10 ⁻⁵	17
m ⁶ dA (insensitive, E)	15500	60	2.6*10 ⁻¹	1.0*10 ⁻³	4.3*10 ⁻⁴	26

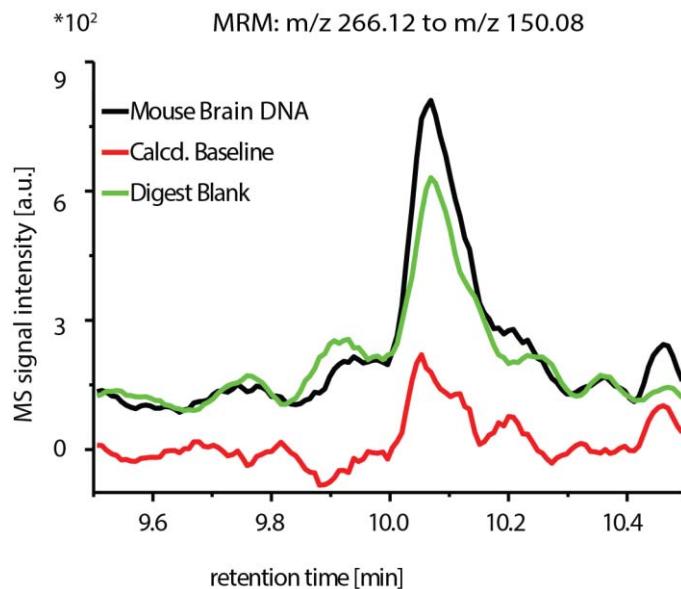


Figure S3: chromatogram of DNA from murine brain green: digest blank, in which everything was performed like in a real digest, but no DNA was added; red: calculated baseline, for which the chromatogram of the digest blank was subtracted from the chromatogram for DNA from murine brain.

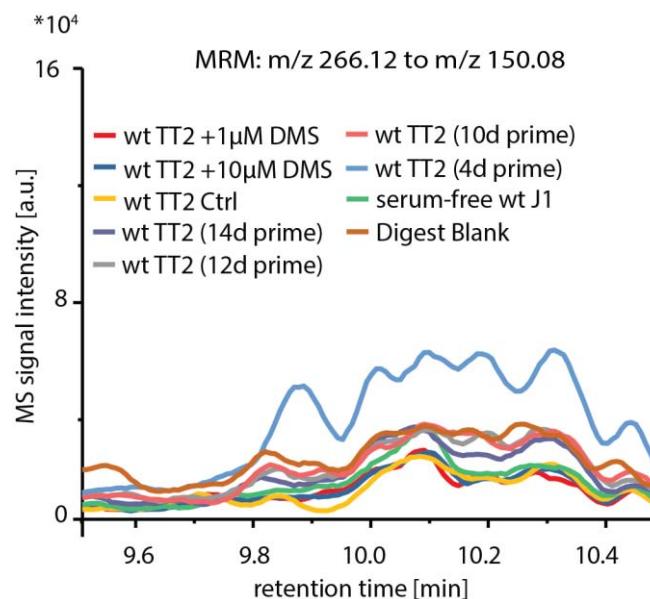


Figure S4: chromatogram of various cell culture samples; red: DNA from wt TT2 cells fed with 1 μ M DMS, blue: DNA from wt TT2 cells fed with 10 μ M DMS, yellow: DNA from a wt TT2 control, purple: DNA from wt TT2 cells after 14 d of priming, grey: DNA from wt TT2 cells after 12 d of priming, orange: DNA from wt TT2 cells after 10 d of priming, light blue: DNA from wt TT2 after 4 d of priming, green: DNA from wt J1 cells cultured in serum-free 2i/LIF, brown: digest blank, in which everything was performed like in a real digest, but no DNA was added

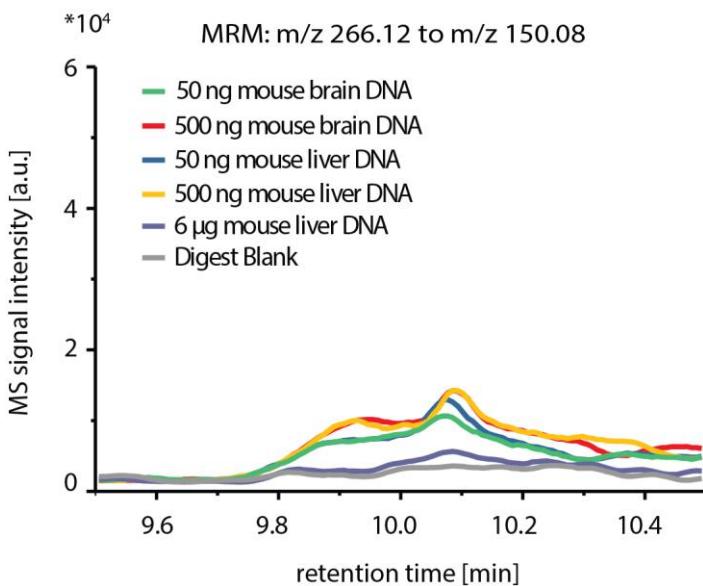


Figure S5: chromatogram of murine DNA from different tissues in different dilutions.; green: 50 ng of DNA from murine brain, red: 500 ng of DNA from murine brain, blue: 50 ng of DNA from murine liver, yellow: 500 ng of DNA from murine liver, 6 μ g of DNA from murine liver, grey: digest blank, in which everything was performed like in a real digest, but no DNA was added.

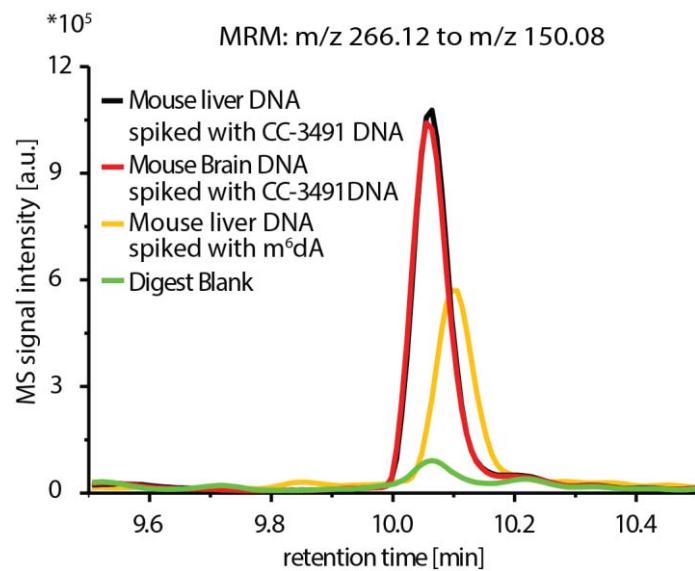


Figure S6: chromatogram of murine DNA from different tissues spiked with DNA from *Chlamydomonas reinhardtii* (CC-3491) and m⁶dA nucleoside, respectively.; black: DNA from murine liver spiked with CC-3491; green: digest blank, in which everything was performed like in a real digest, but no DNA was added, red: DNA from murine brain with CC-3491, yellow: DNA from murine liver spiked with m⁶dA nucleoside. The shown peak seems shifted in contrast to the other two peaks, but this is due to a measurement at a different time.

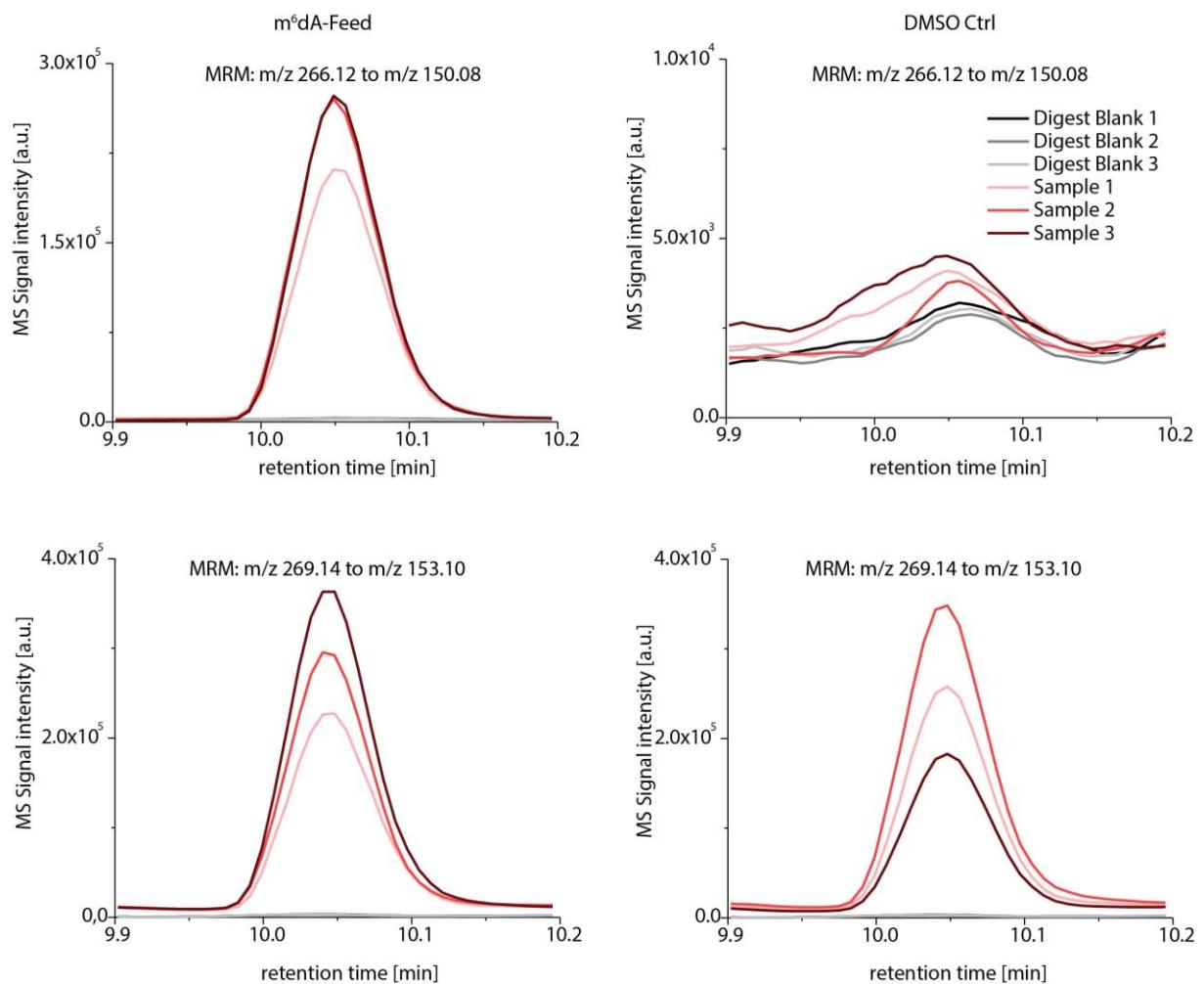


Figure S7: chromatograms for $m^6\text{dA}$ -fed cells (left) and the respective DMSO control (right) with the transition for the natural nucleoside (m/z 266.12 to m/z 150.08) displayed in the upper row and the transition for the spiked heavy labeled nucleoside (m/z 269.14 to m/z 153.10) displayed in the lower row. The red lines indicate the chromatogram for the measured samples, the black line indicates the respective digest blank, in which everything was performed like in a real digest, but without addition of DNA or $\text{D}_3\text{-m}^6\text{dA}$).

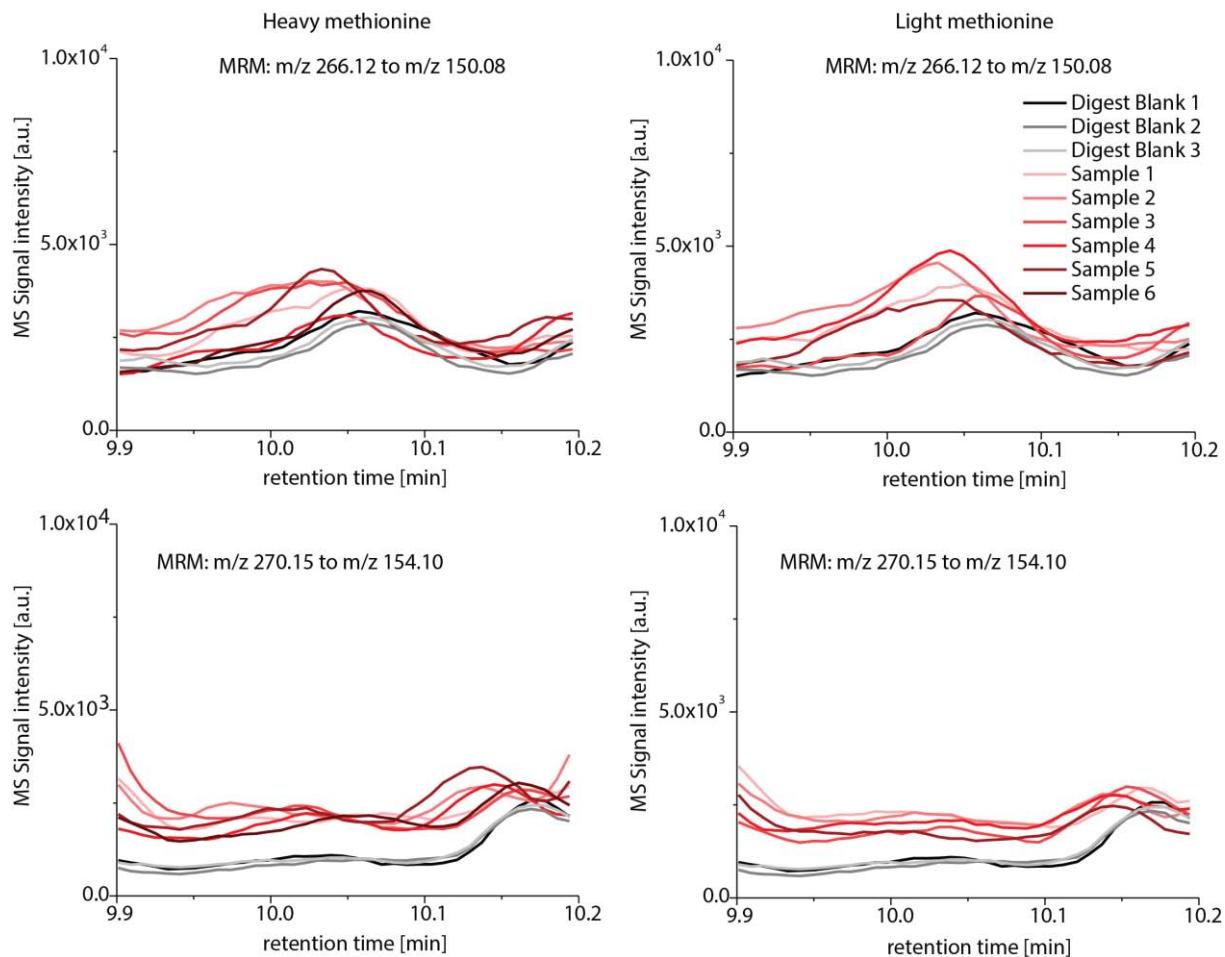
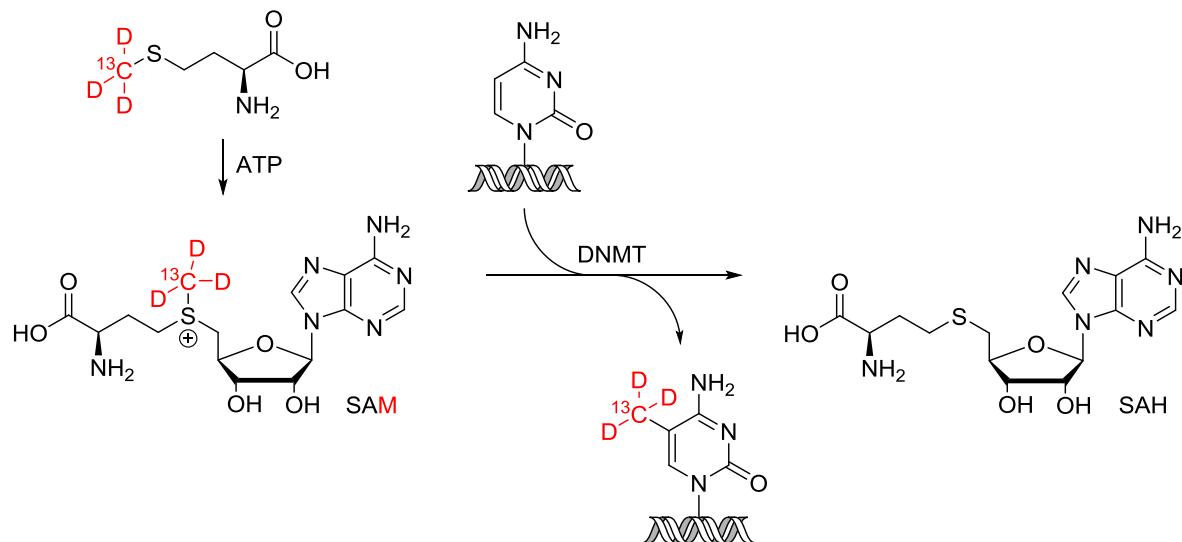


Figure S8: chromatograms for heavy methionine-fed cells (left) and the respective light methionine control (right) with the transition for the natural nucleoside (m/z 266.12 to m/z 150.08) displayed in the upper row and the transition for the heavy labeled nucleoside (m/z 270.15 to m/z 154.10; not formed) displayed in the lower row. The red lines indicate the chromatogram for the measured samples, the black line indicates the respective digest blank, in which everything was performed like in a real digest, but without addition of DNA.

- [1] J. W. Jones, R. K. Robins, *J. Am. Chem. Soc.* **1963**, 85, 193-201.
- [2] J. D. Rosmarie Rippka, John B. Waterbury, Michael Herdman, Roger Y. Stanier, *J. Gen. Microbiol.* **1979**, 111, 1-61.
- [3] D. S. Gorman, R. P. Levine, *Proc. Natl. Acad. Sci. U S A* **1965**, 54, 1665-1669.
- [4] K. Taghizadeh, J. L. McFaline, B. Pang, M. Sullivan, M. Dong, E. Plummer, P. C. Dedon, *Nat. Protoc.* **2008**, 3, 1287-1298.
- [5] E. Li, T. H. Bestor, R. Jaenisch, *Cell* **1992**, 69, 915-926.
- [6] T. Pfaffeneder, B. Hackner, M. Truss, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem. Int. Ed.* **2011**, 50, 7008-7012.
- [7] M. J. Koziol, C. R. Bradshaw, G. E. Allen, A. S. Costa, C. Frezza, J. B. Gurdon, *Nat. Struct. Mol. Biol.* **2016**, 23, 24-30.
- [8] M. Yu, L. Ji, D. A. Neumann, D. H. Chung, J. Groom, J. Westpheling, C. He, R. J. Schmitz, *Nucleic Acids Res.* **2015**, 43, e148.

3.2.3 Synthese von 5-Methyl-2'-desoxycytidin (m^5dC) und 5-[D₃]-Methyl-2'-desoxy-[6-D, 1,3-¹⁵N₂]-cytidin

In unserer Gruppe werden Fütterungsexperimente mit mES-Zellen durchgeführt. Dem Medium wird hierbei isotopenmarkiertes Methionin ([¹³C, D₃]-Methionin) zugefügt, wodurch sich der Metabolismus der Methylgruppe nachvollziehen lässt (Schema 23).

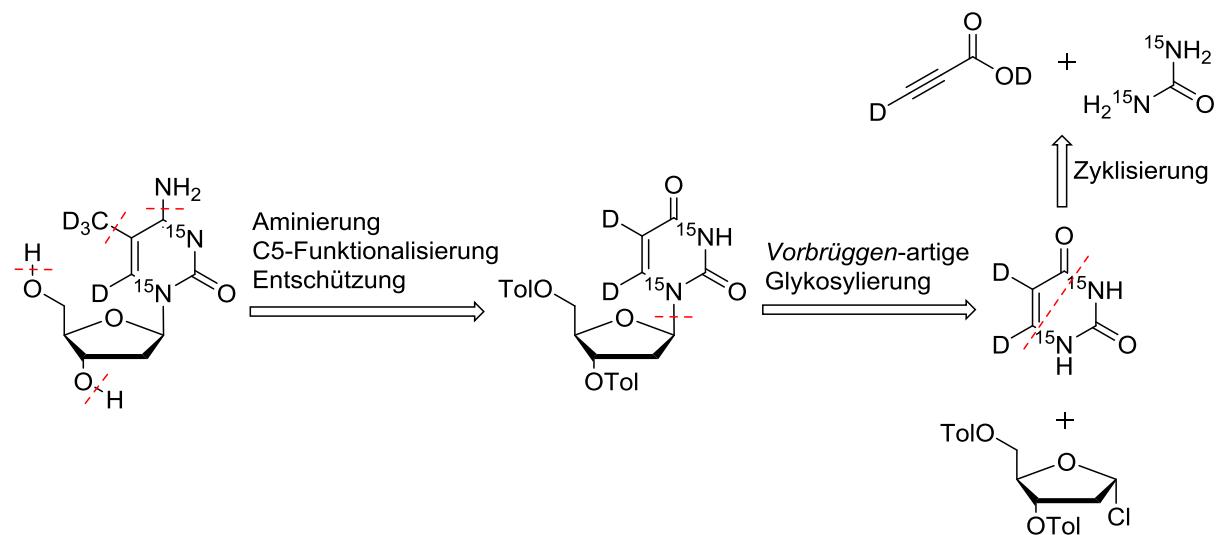


Schema 23: Metabolisierung der deuterierten, ¹³C-markierten Methylgruppe.

Die deuterierte und ¹³C-markierte Methylgruppe wird von Methionin bereitgestellt, welches katalysiert durch die Methionin-Adenosyl-Transferase mit ATP den Kofaktor SAM bildet. SAM überträgt die Methylgruppe nun in einer Transmethylierungsreaktion auf dC, und wird zu SAH abgebaut. In sogenannten *Pulse-Chase*-Experimenten wird dem Medium der mES-Zellen für eine bestimmte Anzahl von Tagen ein kommerziell erhältlicher [¹³C₁₀, ¹⁵N₂]-dT Standard zugegeben. Dieser wird von den Zellen inkorporiert und metabolisiert. Danach wird das Medium ausgetauscht und so den Zellen keine isotopenmarkierte Verbindung mehr zugefügt. Durch DNA-Replikation und Reparatur kommt es nach und nach zum Verlust der Isotopenmarkierungen. Der Verlust von dT[+12] kann beispielsweise durch einen [¹³C, ¹⁵N₂]-dT Standard zu verschiedenen Zeitpunkten quantifiziert werden und dient als Maß für die natürliche Verdünnung. Im Vergleich dazu wird der Gehalt an [¹³C, D₃]-m⁵dC bestimmt. Bis dato stand für die exakte Quantifizierung von [¹³C, D₃]-m⁵dC jedoch kein Standard zur Verfügung. Ein m⁵dC[+2]-Standard wiese zwar die nötige Massendifferenz auf, würde aber mit dem Signal des natürlich vorkommenden m⁵dC – hervorgerufen durch natürlich vorkommende Isotopen – überlagern. In Ermangelung eines Standards muss derzeit mit Parallelkulturen gearbeitet werden. Das Medium der einen Kultur enthält unmarkiertes Methionin, das der anderen Kultur [¹³C, D₃]-Methionin. Der Gehalt an

unmarkiertem m^5dC kann mit einem $m^5dC[+3]$ Standard quantifiziert und die beiden Zellkulturen so verglichen werden. Bei dieser indirekten Quantifizierung von schwerem m^5dC konnte beobachtet werden, dass der Verlust der Methylgruppe nicht allein durch replikative Verdünnung erklärt werden kann. Zusätzlich zu dieser passiven Demethylierung könnte also auch aktive Demethylierung stattfinden. Allerdings ist diese Quantifizierungsmethode keinesfalls exakt, weshalb im Rahmen dieser Arbeit ein Standard synthetisiert werden sollte, der für eine direkte Quantifizierung von $[^{13}C, D_3]-m^5dC$ und simultan des globalen m^5dC -Gehalts geeignet ist. Ein solcher Standard würde eine zusätzliche Kultur ersparen, die Anzahl der zu analysierenden Proben minimieren und daraus resultierende Fehlerquellen beheben. Auf Grund der Sensitivität des UHPLC-MS/MS muss ein solcher Standard sechs Isotopenmarkierungen enthalten, sodass ein Massenunterschied von zwei Masseneinheiten zu der zu quantifizierenden Spezies gewährleistet ist.

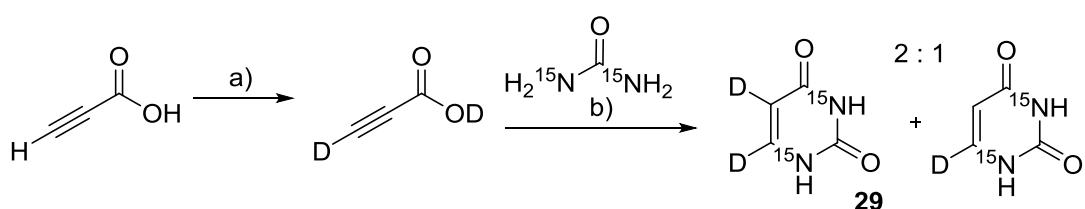
Zunächst sollte das isotopenmarkierte Uracilderivat **29** synthetisiert werden. Dies lässt sich durch Zyklisierung von $^{15}N_2$ -Harnstoff mit deuterierter Propiolsäure erreichen (Schema 24). Im nächsten Schritt sollte dieses durch eine Vorbrüggen-artige Glykosylierung mit Hoffer's Chlorozucker zum geschützten Uridin **30** umgesetzt werden. Aus diesem Uridinderivat sollte $[D_4, ^{15}N_2]-m^5dC$ nach Aminierung, Funktionalisierung der C5-Position und Entschützung der 2'-Desoxyribose erhalten werden.



Schema 24: Retrosynthese von isotopenmarkiertem $[D_4, ^{15}N_2]-m^5dC$. Abbildung adaptiert von A. Schröder.^[183]

Schlüsselschritt der Synthese ist der Aufbau des Uracilderivats **29** aus Harnstoff und Propiolsäure nach Hill et al.^[272] $^{15}N_2$ -markierter Harnstoff ist kommerziell erhältlich, sodass sich auf diese Weise bereits zwei Isotopen einführen lassen. Mit Na_2CO_3 als Base zur Neutralisation kann im Voraus der Alkinwasserstoff der Propiolsäure mit deuteriertem Wasser ausgetauscht werden. Durch saure Aufarbeitung mit Deuteroschwefelsäure kommt

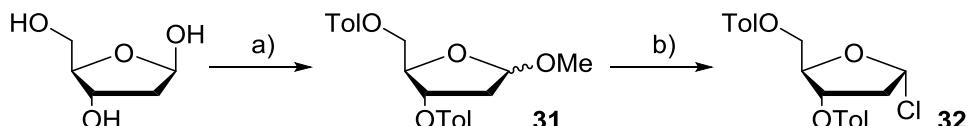
es zum Austausch des Säureprotons, sodass D₂-markierte Propiolsäure entsteht.^[273] [D₂]-Propiolsäure konnte nach Destillation mit einer Ausbeute von 15 % erhalten werden. Nach Erhitzen beider Komponenten in D₂SO₄ unter Rückfluss wurde ein Gemisch aus ca. zwei Dritteln [5,6-D₂, 1,3-¹⁵N₂]-Uracil (Ausbeute 36 %) und ca. einem Drittel [6-D, 1,3-¹⁵N₂]-Uracil (Ausbeute 18 %) erhalten (Schema 25). Im ¹H-NMR-Spektrum sind die NH-Signale deutlich aufgespalten mit Kopplungskonstanten von J = 77,4 Hz bzw. J = 82,6 Hz. Im ¹³C-NMR-Spektrum kommt es auf Grund der Isotopenmarkierungen mit D und ¹⁵N zur Aufspaltung sämtlicher Signale.



Schema 25: Synthese von [5,6-D₂, 1,3-¹⁵N₂]-Uracil und [6-D₂, 1,3-¹⁵N₂]-Uracil. a) Na₂CO₃, D₂O, RT, 2 Tage, 15 %
b) D₂SO₄, 105 °C, 18 h, 54 %.

Ein Stickstoff des Harnstoffs greift bei dieser Reaktion bevorzugt am Alkin der Propiolsäure an. Die so gebildete Zwischenstufe reagiert so schnell unter Wasserabspaltung zum Uracil weiter, dass sie nicht isoliert werden kann.^[274] Die vollständige Deuterierung am C5-Kohlenstoff ist durch erneutes Erhitzen mit Deuterosalzsäure möglich, aber nicht nötig, da diese Position in den nächsten Syntheseschritten funktionalisiert werden soll. Das Gemisch wurde als solches umgesetzt.

Für die Vorbrüggen-artige Glykosylierung musste zuerst Hoffer's Chlorozucker synthetisiert werden (Schema 26).^[275]

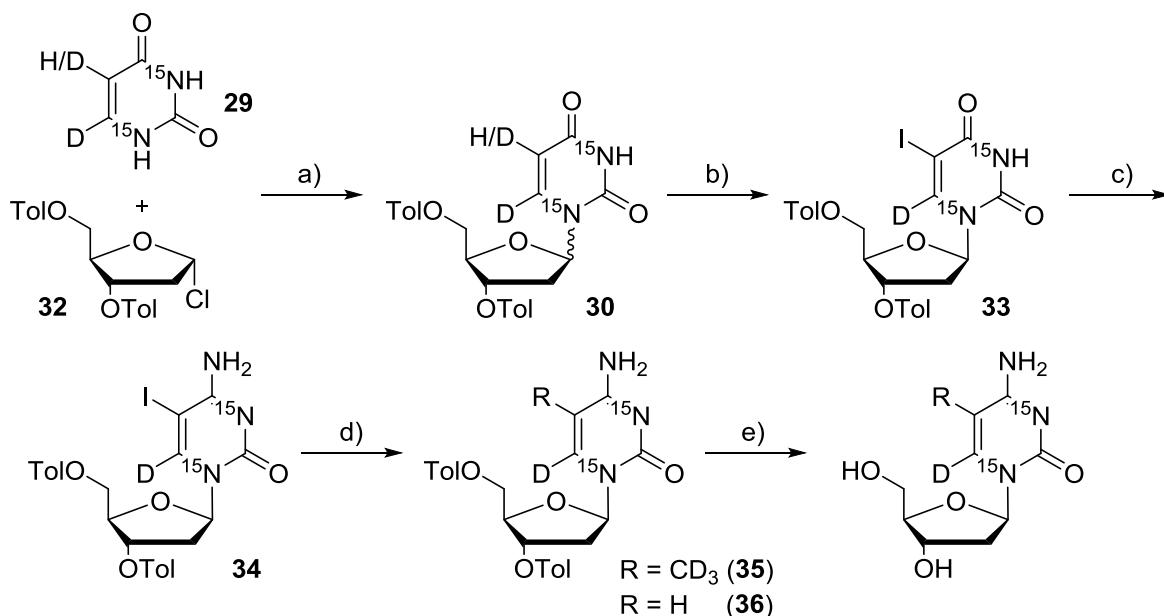


Schema 26: Synthese von Hoffer's Chlorozucker. a) 1. Acetylchlorid, MeOH, RT, 30 min 2. Ag₂CO₃ 3. Pyridin, Toluoylchlorid, 45 °C, 20 h, 92 % b) AcOH, H₂O, Acetylchlorid, 0 °C, 15 min, 69 %.

Im ersten Schritt wurde mit Methanol in Anwesenheit von Acetylchlorid säurekatalysiert am C1-Kohlenstoff von 2-Desoxyribose Methanolat als gute Abgangsgruppe installiert. Überschüssiges HCl wurde mit Silbercarbonat abgepuffert. Das Rohprodukt wurde ohne Aufarbeitung weiter umgesetzt. Mit Toluoylchlorid in Pyridin wurden die C3- und die C5-Position geschützt. 3,5-Di-*p*-toluoyl-methyl-2-desoxyribofuranosid **31** wurde mit einer Ausbeute von 92 % erhalten. Zuletzt wurde die C1-Position mit Acetylchlorid in Essigsäure

chloriert, wobei auf Grund des anomeren Effekts nur das α -Anomer **32** mit einer Ausbeute von 69 % erhalten wurde.

Die weitere Syntheseroute wurde zunächst mit unmarkiertem Uracil getestet, um die optimalen Bedingen für die Synthese der isotopenmarkierten Verbindung festzulegen. Bei der Vorbrüggen-artigen Glykosylierung wurden zu Beginn die Amidgruppen des Uracils mit Bis(trimethylsilyl)acetamid in MeCN geschützt. Nach dreimaliger Coevaporation mit Chloroform zur restlosen Entfernung von MeCN, wurde der Chlorozucker **32** zugegeben, der nun nukleophil am anomeren C1-Kohlenstoff vom N1-Stickstoff des geschützten Uracils angegriffen werden kann. Die Knüpfung der glykosidischen Bindung konnte mit 90 % Ausbeute erreicht werden, wobei jedoch ein 0,7:1 Gemisch aus α - und β -Anomer vorlag, welches säulenchromatografisch nicht getrennt werden konnte. Die gewünschte β -Verbindung wurde somit nur mit einer Ausbeute von 53 % erhalten. Nach Aoyama sollte die Zugabe einer Brönsted-Säure zu einer nahezu quantitativen Entstehung des β -Produkts führen. Der Chlorozucker war in diesem Fall mit *p*-Chlorobenzoyl geschützt und als Brönsted-Säure wurde *p*-Nitrophenol verwendet.^[276] Die analoge Reaktion mit Hoffer's Chlorozucker **32** führte jedoch nicht zu einer Verbesserung des Isomerenverhältnisses. Deshalb wurde die Reaktion ohne die Zugabe von *p*-Nitrophenol durchgeführt. Bei der isotopenmarkierten Verbindung konnte allerdings nur ein Isomerenverhältnis von α/β 1,3:1 erhalten werden mit einer daraus resultierenden Ausbeute von 12 % für das β -Anomer von **30** (Schema 27).



Schema 27: Synthese von 5-[D₃]-methyl-[6-D, 1,3-¹⁵N₂]-cytidin. a) 1. Bis(trimethylsilyl)acetamid, MeCN, 80 °C, 2 h, 2. CHCl₃, RT, 17 h, 12 % b) LiI, CAN, MeCN, 80 °C, 1 h, 52 % c) 1. 1,2,4-Triazol, MeCN, POCl₃, RT, 10 min, 2. NEt₃, 0 °C, 20 min, 3. **33**, 30 °C, 17 h, 4. NH₃, Dioxan, 30 °C, 15 min, 43 % d) 1. Ni(dppp)Cl₂, CH₃-MgI (3 M in THF), THF, RT, 3 h, 2. K₂CO₃, MeOH, RT, 17 h, 12 %. Tol = Toluoyl.

Die Funktionalisierung der C5-Position wurde über die Iodierung mit Lithiumiodid eingeleitet. Das Iodid wird hierbei von Cer(IV)ammoniumnitrat (CAN) zu einer I^+ -Spezies oxidiert, welche dann elektrophil am Uracil angreifen kann. Ein Angriff am C5-Kohlenstoff ist bevorzugt, da das freie Elektronenpaar am N1-Stickstoff das Intermediat stabilisiert. In diesem Schritt war die Trennung des α - und β -Anomers mittels Säulenchromatografie (*i*Hex/EtOAc 5:1) möglich und das iodierte β -Uridinderivat **33** konnte mit einer Ausbeute von 52 % als gelblicher Feststoff erhalten werden.

Die Aminierung des C4-Kohlenstoffs verläuft über ein Triazolid als stabile Zwischenstufe. Phosphoroxychlorid und Triazol bilden Phosphoryl-tris-triazol, welches in Anwesenheit von Triethylamin Triazol als gute Abgangsgruppe in C4-Position installiert. In konzentrierter Ammoniaklösung wird dieses durch $-\text{NH}_2$ substituiert. Cytidinderivat **34** konnte mit einer Ausbeute von 43 % erhalten werden.

Die Einführung der Methylgruppe in C5-Position wurde mit Methyl-*Grignard*-Reagenz durchgeführt, wobei $\text{Ni}(\text{dppp})\text{Cl}_2$ als Katalysator fungierte. Bei dieser Reaktion entstehen sowohl das gewünschte Produkt **35**, als auch unmodifiziertes toluoylgeschütztes 2'-Desoxy-[6-D, 1,3- $^{15}\text{N}_2$]-cytidin **36**. Methylmagnesiumiodid übt zwei Funktionen aus: Zum einen dient es als Methylanionondonor in der *Kumada*-Kupplung, zum anderen als Base, die das exozyklische Amin deprotoniert. Diese Deprotonierung ist wichtig, damit das Amin am C4-Kohlenstoff nicht mit der Kreuzkupplung interferiert und es nicht zu einer 1,3-Protonenwanderung kommt. Allerdings ist die C5-Position in Cytidinderivat **34** so reaktiv, dass es nicht nur zur gewünschten Deprotonierung sondern auch zu einem Iod/Magnesium-Austausch kommt, welcher zu ebendieser energetisch günstigen Protonenwanderung führt.^[277-278] Arne Schröder hat dieses Problem in seiner Dissertation mit der Verwendung von Trimethylaluminium unter *Kumada*-Bedingungen gelöst,^[183] allerdings ist eine deuterierte Version dieses Reagenzes nicht kommerziell erhältlich, weshalb diese Bedingungen in dieser Arbeit nicht anwendbar waren. Bei der *Kumada*-Kupplung wird zunächst *in situ* aus $\text{Ni}(\text{dppp})\text{Cl}_2$ mit dem Methyl-*Grignard*-Reagenz ein Nickel(0)-Komplex gebildet, wodurch es zur Bildung von Ethan als homogekuppeltes Nebenprodukt kommt. Im katalytischen Zyklus wird die Iodoverbindung oxidativ an den Nickel(0)-Komplex addiert. Mit dem Methyl-*Grignard*-Reagenz kommt es zur Transmetallierung, sodass beide Kupplungspartner am Komplex koordiniert sind und das Produkt im nächsten Schritt reduktiv vom Komplex abgespalten werden kann, welcher nun für einen neuen Synthesezyklus zur Verfügung steht.^[279] Das geschützte Methylcytidin **35** und das geschützte Cytidin **36**, welches als Nebenprodukt entstanden ist, konnten auf Grund der identischen Retentionszeit nicht *via* Säulenchromatografie voneinander getrennt werden. Das Gemisch der beiden Verbindungen wurde als solches in Methanol gelöst und mit Kaliumcarbonat versetzt, sodass die Toluoylschutzgruppen als *p*-Methylbenzoësäuremethylester abgespalten werden konnten.

Mittels präparativer Umkehrphasen HPLC war es möglich das isotopenmarkierte m^5dC (12 %) und $[D, {}^{15}N_2]-dC$ (17 %) zu trennen. Das Nebenprodukt $dC[+3]$ kann für die Quantifizierung von dC verwendet werden.

Für die Aufreinigung und Trennung wurde ein Gradient von 0 bis 15 % Acetonitril in Wasser in 45 Minuten verwendet. Abbildung 8A zeigt das HPL-Chromatogramm von $[D_4, {}^{15}N_2]-m^5dC$, Abbildung 8B das von $[D, {}^{15}N_2]-dC$ nach der Aufreinigung. Die Retentionszeiten der markierten Verbindungen stimmen mit den Retentionszeiten der unmarkierten Verbindungen überein. Die Retentionszeiten von m^5dC und dC unterscheiden sich auf der analytischen HPLC um ca. 6 Minuten, wodurch die Trennung der beiden Verbindungen möglich ist.

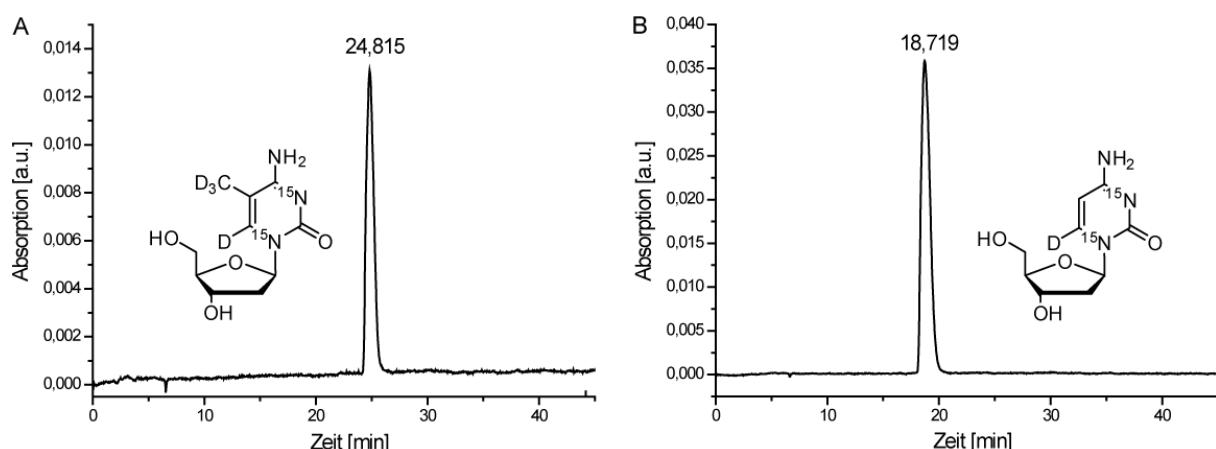


Abbildung 8: HPL-Chromatogramme von (A) $[D_4, {}^{15}N_2]-m^5dC$ und (B) $[D, {}^{15}N_2]-dC$ nach der Aufreinigung.

Da im 1H -NMR-Spektrum in D_2O die Isotopenmarkierung von $[D_4, {}^{15}N_2]-m^5dC$ nicht nachgewiesen werden kann und im ${}^{13}C$ -NMR-Spektrum auf Grund der geringen Subzanzmenge nur die Protonen der Zuckereinheit sichtbar waren, wurden ${}^{15}N$ - und D-Spektren aufgenommen. Allerdings konnten auch in diesen Spektren keine Signale detektiert werden. Letztendlich konnte im ${}^{15}N$ -HMBC-Spektrum die Kopplung des ${}^{15}N$ -Stickstoffs in Position 1 mit den Protonen der Ribose in Position 1 und 2 nachgewiesen werden (Abbildung 9). Auf Grund der Signalverstärkung durch die Kopplung mit dem Protonenspektrum, war im internen ${}^{15}N$ -Spektrum ein Signal für den N1-Stickstoff sichtbar.

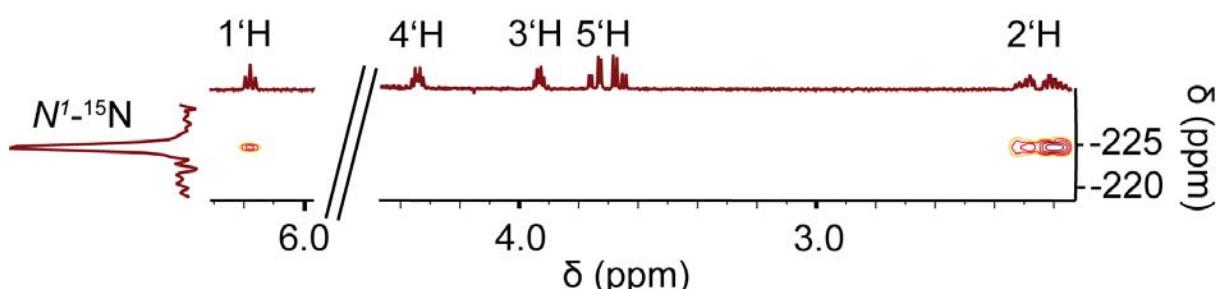


Abbildung 9: ${}^{15}N$ -HMBC-Spektrum von m^5dC . Der ${}^{15}N$ -Stickstoff in Position 1 koppelt mit den Protonen der Ribose in Position 1 und 2.

4 Zusammenfassung und Ausblick

In dieser Arbeit wurden zwei verschiedene Projekte bearbeitet, welche sich beide mit DNA-Modifikationen beschäftigen.

Die Phosphoramidite der drei N^2 -arylsesubstituierten Desoxyguanosinaddukte von Acetylaminophenyl, -naphthyl und -fluorenyl konnten erfolgreich synthetisiert werden, ebenso wie das Phosphoramidit von C8-AAF-dA. Die N^2 -Addukte wurden ebenso wie die entsprechenden C8-Addukte in verschiedene DNA-Stränge eingebaut. Das C8-Desoxyadenosin Addukt konnte zwar in DNA-Stränge eingebaut werden, allerdings kam es bei der Entschützung zur Deacetylierung zu C8-AF-dA. Nach Hybridisierung mit den komplementären ungeschädigten Gegensträngen wurden die Schmelzpunkte aller Addukt-DNA-Doppelstränge bestimmt. Es konnte bestätigt werden, dass C8-Addukte zu einer Erniedrigung des Schmelzpunktes führen, wobei die acetylierten Addukte zu einer größeren Destabilisierung der DNA führen als die nicht acetylierten Addukte. Die Präsenz eines N^2 -Addukts führt hingegen zu einer Stabilisierung und somit zu einem höheren Schmelzpunkt. Mittels EMSA konnte die Erkennung aller geschädigten DNA-Stränge durch XPA gezeigt werden. Da lange, flexible Bereiche die Kristallisation erschweren, wurde die DNA-Bindungsdomäne von Rad14, dem Hefehomolog von XPA, für die Kristallisationsexperimente verwendet. Alle C8-Desoxyguanosinaddukte in 15mer dsDNA konnten erfolgreich mit Rad14 co-kristallisiert werden. Außerdem ist es gelungen die erste Kristallstruktur mit einem N^2 -acetylarylamino-substituierten Desoxyguanosinaddukt (N^2 -AAN-dG) zu erhalten. Die Kristallisation der anderen beiden N^2 -Addukte und des C8-Adenosinaddukts war bei Fertigstellung dieser Arbeit noch nicht erfolgreich. Den erhaltenen Kristallstrukturen ist gemein, dass die DNA geknickt vorliegt (ca. 70 - 78°). Sie legen nahe, dass die Fähigkeit der geschädigten Base, diesen Knick zu stabilisieren, eine wichtige Rolle bei der Schadenserkennung durch XPA/Rad14 spielt, da diese die DNA auf ihre Flexibilität testen. Weiterhin ist in allen Kristallstrukturen die dC-Gegenbase aus der Helix herausgedreht, was ein weiteres Erkennungsmerkmal sein könnte, um geschädigte von ungeschädigter DNA zu unterscheiden. Neue EMSA-Bindungsstudien mit XPAfl und eine Kristallstruktur mit Rad14 von C8-AAF-dG geschädigter DNA, in welcher sich der Schaden im Zentrum einer *Mismatch*-bedingten drei Basen großen Blase (*bubble*) befindet, deuten jedoch darauf hin, dass dem π -stacking doch eine geringere Rolle als angenommen zukommt. Dieses Motiv wird ebenfalls gut durch XPA/Rad14 erkannt, aber π -stacking Interaktionen sollten in einer solchen Struktur keine große Bedeutung haben. Allerdings ist die Struktur rund um den Schaden auf Grund der hohen Flexibilität nicht gut aufgelöst, sodass über die genaue Struktur in diesem Bereich keine Aussage getroffen werden kann.

Um ein vollständiges Bild von der Erkennung und Reparatur von *bulky* Addukten durch NER-Enzyme zu bekommen, reicht es allerdings nicht, sich die Interaktion eines einzelnen Proteins mit der DNA alleine anzuschauen. An der NER ist eine Vielzahl von Enzymen beteiligt, wobei XPA als Gerüst für alle anderen Proteine fungiert. Da die Röntgenbeugung bei solch großen Proteinkomplexen an ihre Grenzen stößt, soll im nächsten Schritt die Komplexbildung zwischen dsDNA, Volllängen-XPA und RPA mittels Kryo-Elektronenmikroskopie (Kryo-EM) untersucht werden. Der Vorteil der Kryo-EM ist, dass diese Komplexe in ihrer natürlichen Umgebung untersucht werden könnten, ohne vorher fixiert werden zu müssen.

Im zweiten Teil der Arbeit wurden die methylierten Nukleoside m^5dC und m^6dA sowie ihre isotopenmarkierten Analoga $[D_4, ^{15}N_2]-m^5dC$, $[D_3]-m^6dA$ und $[^{15}N_2]-m^4dC$ für die Quantifizierung dieser DNA-Modifikationen in verschiedenen Zelltypen hergestellt.

Es konnte bestätigt werden, dass m^4dC und m^6dA im Cyanobakterium *Synechocystis* vorkommen und m^6dA konnte außerdem in der Grünalge *C. reinhardtii* gefunden werden. Die Frage, ob diese Modifikationen auch in Eukaryoten eine Rolle spielen, ist derzeit Gegenstand der Forschung. Sowohl in mES-Zellen als auch in ausdifferenzierten Zellen lag der Anteil der beiden Modifikationen unter dem Detektionslimit unserer Methode. Dies weist darauf hin, dass diese Basen in diesen Zellen keine relevante epigenetische Funktion ausüben, sondern falls vorhanden eher als DNA-Schäden, resultierend aus fehlgeleiteten Methyltransferasen, einzustufen sind. In früheren Stufen der embryonalen Entwicklung des Zebrafischs und des Schweins soll m^6dA allerdings in relevanten Mengen nachgewiesen worden sein. Mit unserer ultrasensitiven Methode, und der im Rahmen dieser Arbeit synthetisierten Standards, ließe sich dies überprüfen. Die konkrete Suche nach biologischen Prozessen zu bestimmten Zeitpunkten der Entwicklung, in denen eine dieser Basen involviert ist, ist in jedem Fall ein interessanter Ansatz für die weitere Forschung. Die beiden synthetisierten Standards eignen sich generell für die Quantifizierung des Gehalts an m^4dC und m^6dA in allen erdenklichen Zelltypen und Geweben zu jedem beliebigen Zeitpunkt.

Der Standard für die Quantifizierung von schwerem m^5dC aus Fütterungsexperimenten von mES-Zellen mit $[^{13}C, D_3]$ -Methionin wurde bis zur Fertigstellung dieser Arbeit noch nicht für massenspektroskopische Experimente verwendet. In unserer Gruppe wurde bei Pulse-Chase-Experimenten beobachtet, dass der Verlust der Methylgruppe an dC vor der Implantation der Eizelle nicht allein durch passive Demethylierung zu erklären ist. $[D_4, ^{15}N_2]-m^5dC$ soll für die exakte Quantifizierung des globalen $[^{13}C, D_3]-m^5dC$ Gehalts zu verschiedenen Zeitpunkten der embryonalen Entwicklung verwendet werden. Somit könnte der Anteil an m^5dC , das aktiv demethyliert wird, bestimmt werden. Derzeit werden diese Experimente mit Parallelzellkulturen durchgeführt, wobei dem Medium der ersten Kultur Methionin beigefügt wird, dem Medium der zweiten Kultur hingegen $[^{13}C, D_3]$ -Methionin. Der

Gehalt an unmarkiertem m^5dC kann mit einem $m^5dC[+3]$ Standard quantifiziert und die beiden Zellkulturen verglichen werden. Für die Quantifizierung von $[^{13}C, D_3]-m^5dC$ ist dieser Standard jedoch nicht geeignet, da der Massenunterschied nur 1 Da beträgt. Der im Zuge dieser Arbeit synthetisierte $[D_4, ^{15}N_2]-m^5dC$ Standard ermöglicht erstmals die simultane Quantifizierung von m^5dC und $[D_4, ^{15}N_2]-m^5dC$ aus einer Zellkultur. Die zweite Zellkultur wird dadurch obsolet, was eine enorme Arbeitsersparnis bedeutet. Außerdem ist die bis dato durchgeführte indirekte Quantifizierung keineswegs exakt, die Quantifizierung mit dem neuen Standard hingegen schon. Als Nebenprodukt entstand bei der Synthese $[D, ^{15}N_2]-dC$. Dieser $dC[+3]$ Standard kann für die Quantifizierung von dC verwendet werden.

5 Experimenteller Teil

5.1 Chemische Materialien und Methoden

Alle Lösungsmittel und Reagenzien wurden in den kommerziell erhältlichen Reinheiten (*puriss.*, *p.a.* oder *purum*) käuflich erworben und auch so verwendet. Trockene Lösungsmittel wurden von *Fluka* (Sigma-Aldrich-Chemie GmbH), *Acros* (KMF Laborchemie Handels GmbH), *Aldrich*, *ABCR*, *Alfa Aesar*, *Sigma*, *TCI* oder *Merck* gekauft oder mittels 3 bzw. 4 Å Molekularsieb selbst getrocknet. Das hierfür verwendete Molsieb wurde vor der eigentlichen Nutzung zuerst mit Wasser gewaschen um Staub zu entfernen und anschließend in einem Heizschrank für mindestens 24 h bei 120 °C vorgetrocknet und abschließend für 12 Stunden in einem Sandbad (bei ca. 340 °C) im Hochvakuum weiter ausgeheizt. Für Säulenchromatographie und Extraktionen wurden Lösemittel von technischer Qualität verwendet die vorher an einem Rotationsverdampfer (Firma *Heidolph*) mittels eines Membranpumpenvakuums (*Vario PC3001-Membranpumpe* – Firma *Vacuubrand*) destilliert wurden. Zum Entgasen von Lösemittel wurde dieses dreimal in flüssigem Stickstoff eingefroren und jeweils mittels Vakuum und Argon ein Gasaustausch vorgenommen. Zum Trocknen von Substanzen im Hochvakuum verwendete man eine RZ-4 Öldrehschieberpumpe (Firma *Vacuubrand*) mit einem Vakuum zwischen 10⁻² und 10⁻³ mbar. Zum Entfernen von Wasser bzw. wässriger Puffer wurde eine SpeedVac *RVC2-25* bzw. *RVC2-33 IR* (Firma *Christ*) oder eine SpeedVac *Plus SC110A* bzw. SpeedVac *SPD111V* (Firma *Savant*) verwendet. Alle Geräte waren mit den Kühlfallen *CT62-50-cooling trap* (Firma *Christ*) ausgestattet. Für die Gefriertrocknung wässriger Lösungen wurde eine *Alpha 2-4 plus Lyophille* (Firma *Christ*) verwendet. Alle feuchtigkeits- und sauerstoffempfindlichen Reaktionen wurden in ausgeheizten, evakuierten und mit Argon gefüllten Schlenkgefäßen durchgeführt. Iodhaltige Verbindungen wurden unter Lichtausschluss gehandhabt.

Dünnschichtchromatographie (DC) wurde auf vorgefertigten Aluminiumplatten (Kieselgel 60 F₂₅₄, Schichtdicke 0,25 mm, Firma *Merck*) ausgeführt. Die Substanzen wurden durch UV-Licht ($\lambda = 254$ nm oder 366 nm), Ninhydrin-Lösung (1,5 g Ninhydrin, 100 mL *n*Butanol, 3 mL Eisessig) oder Anisaldehyd-Lösung (0,5 mL Anisaldehyd, 50 mL Eisessig, 1 mL H₂SO₄) sichtbar gemacht.

Säulenchromatographie zur Aufreinigung von Substanzgemischen wurde unter der Verwendung von Stickstoffüberdruck nach der Flash-Methode durchgeführt. Hierbei wurde Silicagel (Korngröße 40 - 63 µm Kieselgel 60 Firma *Merck*) verwendet.

NMR-Spektren wurden an folgenden Geräten aufgenommen: *ARX-300* und *AMX-600* (Firma Bruker) sowie *Mercury 200VX*, *INOVAN XL400* (400 MHz), *Varian Inova 400* (400 MHz) (Firma Varian). Die Messungen wurden bei Raumtemperatur durchgeführt. Die Messfrequenz und das verwendete deuterierte Lösungsmittel sind in Klammern angegeben. Die chemischen Verschiebungen δ für die ^1H und ^{13}C Spektren sind in ppm relativ zu Tetramethylsilan ($\delta = 0$) und für ^{31}P Spektren relativ zu Phosphorsäure (85 % in Wasser, $\delta = 0$) angegeben. Die jeweiligen Kopplungskonstanten J sind in Hertz (Hz) angegeben. Die Multiplizitäten wurden wie folgt abgekürzt: s = Singulett, d = Dublett, t = Triplet, q = Quartett, m = Multiplett. Die chemischen Verschiebungen wurden relativ zu CHCl_3 ($\delta = 7.26$) in CDCl_3 , DMSO ($\delta = 2.50$) in $\text{DMSO-d}6$, MeOH ($\delta = 3.31$) in CD_3OD , H_2O ($\delta = 4.80$) in D_2O , CH_3CN ($\delta = 1.94$) in CD_3CN oder Aceton ($\delta = 2.04$) in Aceton- $d6$.

Massenspektrometrische Messungen wurden an einem *Varian MAT CH 7A* (Firma Varian) für Elektronenstoßionisation (EI) sowie an einem *Thermo Finnigan LTQ-FT-ICR* (Firma Finnigan) für Elektrosprayionisationen (ESI) durchgeführt. Für Reaktionskontrollen wurde eine *Dionex (Thermo) Ultimate 3000* gekoppelt an ein *Thermo MSQ*-Massenspektrometer verwendet. Für MALDI-TOF Messungen benutzte man das *Autoflex II* (Firma Bruker) mit einem MTP AnchorChip var/348 Träger. Als Matrix wurde eine HPA-Matrix (50 mg 3-Hydroxypicolinsäure, 10 mg Ammoniumhydrogencitrat, 1 mL ddH₂O/CH₃CN 1:1) verwendet. Salzhaltige Proben wurden mittels eines *MF-Millipore™* Membranfilters (0,025 µm Porengröße) entsalzt.

IR-Spektren wurden an einem *Spectrum BX* FT-IR System von *Perkin Elmer* mit einem *DuraSamplIR II* Messkopf von *Smiths* und einer Diamant-ATR (Attenuated Total Reflection) Einheit ohne weitere Präparation der Proben aufgenommen. Die Wellenzahlen wurden in cm^{-1} angegeben. Die Intensität der Signale wurden mit s (stark), m (medium), w (schwach) konkretisiert. Die Bezeichnung br steht für sehr breite Signale.

Schmelzpunkte wurden in offenen Glaskapillaren an einem *Melting Point B-450* (Firma Büchi) bestimmt und sind nicht korrigiert.

Doppelt destilliertes Wasser (ddH₂O): Für biochemische Untersuchungen und zur Herstellung flüssiger Phasen für die HPLC wurde doppelt destilliertes Wasser (Bidestwasser) mit Hilfe einer Milli-Q Plus-Anlage mit einer QPAK® 2 Kartusche erzeugt.

5.2 Methoden zur DNA-Synthese, Entschützung und Aufreinigung

DNA-Synthesen wurden an einem *ABI 394 DNA/RNA Synthesizer (Applied Biosystems)* im 1 µmol oder 200 nmol Maßstab durchgeführt (es dienten die Standardprotokolle des Geräteherstellers als Grundlage). Zum Einbau der einzelnen Modifikationen (C8-AAF-dA, C8-AAF-dG, C8-AAN-dG, C8-AAB-dG, C8-AAA-dG, C8-AAP-dG, N^2 -AAF-dG, N^2 -AAN-dG, N^2 -AAB-dG) wurden die Kupplungszeiten auf 30 min verlängert. Die Phosphoramidite der Standardbasen Bz-dA, Bz-dC, *i*Bu-dG und dT sowie die *ultramild®* Phosphoramidite Pac-dA, Ac-dC und *i*Pr-Pac-dG und die mit der ersten Base beladenen CPG- oder *ultramild Q*-CPG-Träger wurden von den Firmen *Glen Research* oder *Link Technologies* bezogen. Als Aktivator wurde *Aktivator 42® Solution* der Firma *Sigma-Aldrich* verwendet. Alle Reagenzien (Amidite, Oxidationslösung, Aktivator, Waschlösung und *capping*) wurden in Acetonitril (Firma *Roth* - Wassergehalt < 10 ppm) gelöst. Das N^2 -AAF-dG Phosphoramidit wurde in einem 1:1 Gemisch aus Acetonitril und DCM gelöst, das N^2 -AAN-dG Phosphoramidit wurde in DCM gelöst. Für alle anderen Lösungen wurde Acetonitril (Firma *Roth* - Wassergehalt < 30 ppm) eingesetzt. Die Konzentration der Phosphoramidite betrug 0,1 M in Acetonitril (< 10 ppm H₂O).

Capping der Standardbasen:

Cap A: 20 % Essigsäureanhydrid (Ac₂O) in 50 % CH₃CN und 30 % 2,6-Lutidin

Cap B: 16 % *N*-Methylimidazol in CH₃CN

Capping während der *ultramild* Synthesen:

Cap A: 0,5 M Pac₂O in CH₃CN/2,6-Lutidin 9:1

Cap B: 16 % *N*-Methylimidazol in CH₃CN

Detritylierung:

3 % Dichloressigsäure (DCA) in DCM

Oxidation:

2,5 mM I₂ in 2,6-Lutidin in CH₃CN/H₂O 6:65:30 v/v

DNA-Entschützung: Die terminale 5'-DMTr Gruppe wurde am Ende der Synthesen abgespalten (*DMT-off* Synthese). Oligonukleotide, welche das Formylcytidin Phosphoramidit enthielten, wurden zunächst basisch (Entfernung der Aminschutzgruppe, Abspaltung vom Trägermaterial) mit konz. NH₄OH (33 % NH₃ in Wasser) bei 25 °C über 17 h entschützt, an einer SpeedVac oder Lyophille eingeengt und mit 0,225 M NaOAc in EtOH/Wasser 8:2 gefällt. Die Carbonylfunktion wurde sauer mit 80 % Essigsäure in Wasser bei 20 °C für 6 h entschützt an einer SpeedVac oder Lyophille eingeengt und mit 0,15 M NaOAc in EtOH/Wasser 8:2 gefällt. Für die Entschützung und Abspaltung unter *ultramild*-Bedingungen (N^2 - und C8-Addukte) wurde eine Lösung aus 10 % Di-*iso*-propylamin und 0,25 M β-Mercaptoethanol in Methanol bei 55 °C über 18 h in einem Thermoschüttler (mind. 650 rpm)

verwendet. Für das C8-AAF-dA Addukt wurden nur 5 % Di-*iso*-propylamin und 0,25 M β-Mercaptoethanol in Methanol bei 55 °C über 18 h verwendet, allerdings konnte nur das C8-AF-dA Addukt isoliert werden. Die Lösungen wurden anschließend an einer *SpeedVac* oder Lyophile eingeeengt und das erhaltene Pellet in ddH₂O aufgenommen. Um die Festphase für die anschließende HPLC-Reinigung zu entfernen wurde die Lösung durch einen 0,2 µm Spritzenfilter filtriert.

Sofern nicht selbst hergestellt, wurden die verwendeten Oligonukleotide von der *Metabion GmbH* (in HPLC-gereinigter Qualität) bezogen. Eine vollständige Liste aller synthetisierten geschädigten Oligonukleotide und der kommerziell erworbenen Gegenstränge und ihrer Massen ist in Tabelle 2 dargestellt.

ODN	Oligonukleotidsequenz	G*	Masse [g/mol]
1	5'-GCT CTA CG*T CAT CAC-3'	N ² -AAB-dG	4621
2		N ² -AAF-dG	4709
3		N ² -AAN-dG	4671
4		C8-AAB-dG	4621
5		C8-AAF-dG	4709
6		C8-AAN-dG	4671
7		C8-AAA-dG	4721
8		C8-AAP-dG	4745
9		C8-AF-dG	4667
10		C8-AF-dA	4654
11	5'-GTG ATG ACG TAG AGC-3'		4657
12	5'-GTG ATG ATG TAG AGC-3'		4672
13	5'-GTA GCT AGG TCG ATG* CAC AGA CAT CGC ACA CGT C-3'	N ² -AAF-dG	10658
14		C8-AAB-dG	10570
15		C8-AAN-dG	10620
16		C8-AAA-dG	10670
17		C8-AAP-dG	10694
18		C8-AAF-dG	10658
19		C8-AF-dG	10617
20	5'-GAC GTG TGC GAT GTC TGT GCA TCG ACC TAG CTA C-3'		10450

Tabelle 2: Liste aller synthetisierter Oligonukleotide und deren Massen.

HPL-Chromatographie der synthetisierten Oligonukleotide: Die analytische als auch präparative *reversed phase (RP)* HPL-Chromatographie wurde auf den Anlagen *L-7400 UV*- und *L-7420 UV/VIS-Detektor*, *L-7100-* und *L-7150 Pumpe*, *D-7200 Autosampler*, *L-7000, Interfacemodul* und *L-7612 bzw. ERC-3415* (Firma *Merck-Hitachi*) oder *2487 Dual λ Absorbance Detector*, *1525 Binary HPLC Pump*, *2695 Separations Module* und *2996*, bzw. *996 Photodiode Array Detector* (Firma *Waters*) durchgeführt. Bei den Entgasern handelte es sich um die jeweiligen Modelle der Hersteller. Die präparativen Trennungen wurden auf folgenden *RP*-Säulen durchgeführt: *CC 250/10 Nucleosil 100-7 C18* oder *CC 250/10 Nucleodur 100-7 C18* (Firma *Macherey-Nagel*). Für die analytischen HPLC Trennungen wurden die Säulen *CC 250/4 Nucleosil 4 120-3 C18* und *CC 250/4 Nucleodur 4 120-3 C18* (Firma *Macherey-Nagel*) eingesetzt. Die präparative *Waters* HPLC erlaubte eine Beobachtung bei zwei Wellenlängen, was eine zusätzliche Detektion der aromatischen Acetylamino-Einheiten bei 305 nm ermöglichte.

Für die Auftrennung von Oligonukleotiden wurden folgende Puffersysteme verwendet:

Puffersystem A: 0,1 M Triethylammoniumacetat (TEAA) in ddH₂O, pH 7,0.

Puffersystem B: 0,1 M TEAA in Acetonitril/ddH₂O 8:2, pH 7,0.

Für die Aufreinigung von Oligonukleotiden mit 15 Basen wurde ein Gradient von 0–25 % oder 0–30 % Puffer B in 45 min (analytische HPLC: Flussrate: 0,5 mL/min, präparative HPLC: Flussrate: 5 mL/min). Je nach beobachteter Retentionszeit wurde der oben angeführte Standardgradient modifiziert. Die gereinigten Oligonukleotide wurden mittels *Sep-Pak®* Kartuschen (Firma *Waters*) anhand des vom Hersteller empfohlenen Protokolls entsalzt.

Extinktionskoeffizienten von Oligonukleotiden bei 260 nm wurden aus der Summe der einzelnen Extinktionskoeffizienten der Einzelnukelotide des jeweiligen DNA-Stranges berechnet. Diese betrugen für dA 15,2 L mmol⁻¹ cm⁻¹, dC 7,1 L mmol⁻¹ cm⁻¹, dG 12,0 L mmol⁻¹ cm⁻¹ und für dT 8,4 L mmol⁻¹ cm⁻¹. Für die aromatisch modifizierten Monomere nahm man den Extinktionskoeffizienten des dGs bzw. dAs an. Die Konzentration von DNA-Strängen wurde mit einem *NanoDrop® Spektrophotometer ND-100* von *Peqlab Biotechnologie GmbH* bestimmt. Die Absorption wurde bei 260 nm gemessen.

Hybridisierungen wurden in einem *Mastercycler personal* Thermocycler (Firma *Eppendorf*) durchgeführt. Zur Hybridisierung wurden die beiden komplementären DNA Stränge in 5 min auf 105 °C erhitzt und anschließend mit 1 °C min⁻¹ bis auf 4 °C abgekühlt.

5.3 Biochemische Materialien und Methoden

Schmelztemperaturen. Schmelzkurven wurden mit einem *Cary 100 Bio* Spektralphotometer mit *Cary Temperature Controller*, *Sample Transport Accessory* und *Multi Cell Block* der Firma *Varian* ermittelt. Die Duplexe (Konzentration 4 µM) wurden in einem Volumen von 250 µL in Schmelzpunktpuffer (Tabelle 3) gelöst und in den Küvetten mit Dimethylpolysiloxan überschichtet. Die 15mer Oligonukleotide wurden mit einem Temperaturgradienten von 1 °C/min von 15 auf 85 °C erhitzt oder von 85 auf 15 °C abgekühlt, während die 34mer Oligonukleotide von 20 auf 95 °C erhitzt und von 95 auf 20 °C abgekühlt wurden. Pro Duplex wurden jeweils drei Schmelzprofile aufgenommen. Schmelzpunkte (T_m) von Oligonukleotiden ergeben sich aus dem Nulldurchgang der geglätteten zweiten Ableitung der Schmelzkurve eines Oligonukleotiddoppelstrangs.

Schmelzpunktpuffer	Bindungspuffer	Auftragspuffer
10 mM Tris Puffer, pH 7,4	50 mM HEPES pH 8,0	250 mM Tris-HCl pH 7,5
150 mM NaCl	50 mM MgCl ₂	40 % Glycerin
	5 % Glycerol	0,2 % Bromphenolblau
	5 mM DTT	
	10 µM ZnCl ₂	

Tabelle 3: Zusammensetzung der Puffer für biochemische Experimente.

Native Polyacrylamidgelektrophorese. Für EMSA-Experimente wurden native, nicht denaturierende Gele verwendet. In Tabelle 4 ist die Zusammensetzung für sechs native Polyacrylamidgele angegeben.

Reagenzien	V [mL]
ddH ₂ O	31
30 % Rotihoresce	6
10 x TBE	2
Glycerin	1
10 %ige APS-Lösung	0,3
TEMED	0,02

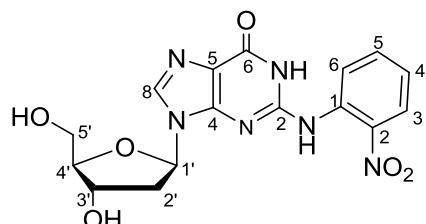
Tabelle 4: Zusammensetzung von sechs 4 % nicht-denaturierenden Gelen.

Electrophoretic Mobility Shift Assays (EMSA). Ungeschädigte oder geschädigte Oligonukleotide wurden mit ihrem komplementären Strang in einem äquimolaren Verhältnis vermischt und hybridisiert. Ein Standard Reaktionsgemisch bestand aus 5 pmol der Duplex DNA mit variierenden Konzentrationen an aufgereinigtem Volllängen-XPA (XPAfl) mit entweder nur einem *His-Tag* am N-Terminus oder auch einem *Strep-Tag* am C-Terminus in Bindungspuffer (Tabelle 3). Nach einer Inkubationszeit von 30 min bei 37 °C wurden 2 µL Auftragspuffer (Tabelle 3) zugegeben und die Elektrophorese auf einem 4 % nicht-denaturierendem Gel durchgeführt (100 V konstant, 60 min, 4 °C). Das Gel wurde zuvor 30 min bei 100 V laufen gelassen und als Laufpuffer wurde 1× TBE verwendet. Die dsDNA wurde anschließend mit *SYBR Green I* über 20 – 30 Minuten angefärbt und das Gel wurde mit Hilfe eines *Image Reader LAS 3000* detektiert.

5.4 Synthese der *bulky* Addukte

5.4.1 Synthese des *N*²-AAB-dG Phosphoramidits

5.4.1.1 Synthese von *N*²-(2-Nitrophenyl)-2'-desoxyguanosin (10)



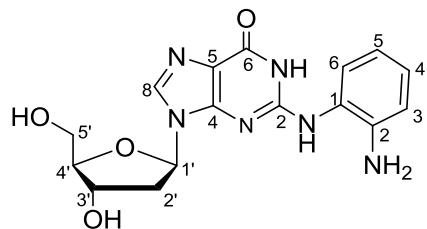
2'-Desoxyguanosin Monohydrat (2,000 g, 7,011 mmol, 1,0 Äq.), 1-Bromo-2-nitrobenzol (2,270 g, 9,114 mmol, 1,3 Äq.), Pd₂(dba)₃ (0,642 g, 0,701 mmol, 0,1 Äq.) und Xantphos (1,217 g, 2,103 mmol, 0,3 Äq.) wurden drei Mal entgast, in 21 mL trockenem DMSO gelöst und 10 Minuten bei RT gerührt. TEAF Hydrat (2,093 g, 14,02 mmol, 2,0 Äq.) wurde hinzugegeben und das Reaktionsgemisch wurde 14 h bei 100 °C gerührt. Die Suspension wurde mit 150 mL DCM/MeOH 100:1 verdünnt und *via* Säulenchromatografie (DCM/MeOH 10 → 6:1) gereinigt. *N*²-(2-Nitrophenyl)-2'-desoxyguanosin **10** (2,695 g, 6,941 mmol, 99 %) wurde als gelbes Pulver erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 11.82 (s, 1H, NH), 9.61 (s, 1H, NH), 8.22 (dd, *J* = 8.4, 1.1, 1H, C₆H₄NO₂-H), 8.07 (dd, *J* = 8.4, 1.5, 1H, C₆H₄NO₂-H), 8.04 (s, 1H, C8-H), 7.74 – 7.68 (m, 1H, C₆H₄NO₂-H), 7.26 (ddd, *J* = 8.5, 7.3, 1.3, 1H, C₆H₄NO₂-H), 6.06 (t, *J* = 6.8, 1H, C1'-H), 5.22 (d, *J* = 4.1, 1H, OH), 4.83 (t, *J* = 5.4, 1H, OH), 4.26 (dd, *J* = 6.0, 3.3, 1H, C3'-H), 3.75 (dd, *J* = 7.8, 4.6, 1H, C4'-H), 3.54 – 3.37 (m, 2H, C5'-H), 2.50 (dt, *J* = 6.6, 4.2, 1H, C2'-H), 2.19 (ddd, *J* = 13.3, 6.4, 3.5, 1H, C2'-H) ppm.

¹³C-NMR (101 MHz, DMSO-*d*₆) δ 157.1 (C6), 149.1 (C2), 149.0 (C4), 139.7 (C₆H₄NO₂), 137.4 (C8), 135.1 (C₆H₄NO₂), 133.8 (C₆H₄NO₂), 125.9 (C₆H₄NO₂), 124.5 (C₆H₄NO₂), 123.9 (C₆H₄NO₂), 119.8 (C5), 88.1 (C4'), 83.5 (C1'), 71.0 (C3'), 62.0 (C5'), 39.9 (C2') ppm.

HR-MS (ESI+): m/z berechnet für [C₁₆H₁₇N₆O₆]⁺ ([M+H]⁺): 389,1204, gefunden: 389,1205.

5.4.1.2 Synthese von *N*²-(2-Aminophenyl)-2'-desoxyguanosin (13)



*N*²-(2-Nitrophenyl)-2'-desoxyguanosin **10** (2,717 g, 7,011 mmol, 1,0 Äq.) und Na₂S₂O₄ (6,067 g, 35,06 mmol, 5,0 Äq.) wurden in 75 mL THF und 40 mL H₂O gelöst und das Reaktionsgemisch wurde 15 h bei RT gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand *via* Säulenchromatografie (DCM/MeOH 10 → 3:1) gereinigt. *N*²-(2-Aminophenyl)-2'-desoxyguanosin **13** (2,010 g, 5,609 mmol, 80 %) wurde als weißer Feststoff erhalten

¹H-NMR (400 MHz, DMSO-*d*₆) δ 10.60 (s, 1H, NH), 8.00 (s, 1H, NH), 7.92 (s, 1H, C8-H), 7.36 (dd, *J* = 7.9, 1.5 Hz, 1H, -C₆H₄), 6.87 (ddd, *J* = 8.0, 7.2, 1.5 Hz, 1H, -C₆H₄), 6.73 (dd, *J* = 8.0, 1.5 Hz, 1H, -C₆H₄), 6.56 (ddd, *J* = 7.9, 7.2, 1.5 Hz, 1H, -C₆H₄), 6.00 (dd, *J* = 7.6, 6.3 Hz, 1H, C1'-H), 5.17 (d, *J* = 3.8 Hz, 1H, C3'-OH), 4.90 (bs, 2H, NH₂), 4.77 (bs, 1H, C5'-OH), 4.23 (d, *J* = 6.9 Hz, 1H, C3'-H), 3.73 (q, *J* = 4.3, 3.8 Hz, 1H, C4'-H), 3.46 – 3.35 (m, 2H, C5'-H), 2.51 (ddd, *J* = 11.8, 6.9, 5.1 Hz, 1H, C2'-H), 2.13 (ddd, *J* = 13.2, 6.3, 3.1 Hz, 1H, C2'-H) ppm.

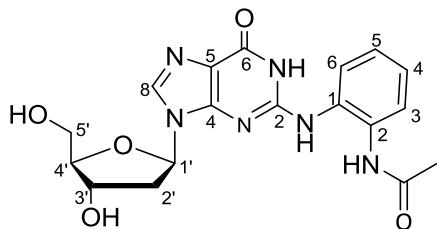
¹³C-NMR (101 MHz, DMSO-*d*₆) δ 157.1 (C6), 151.2 (C2), 150.5 (C4), 142.6 (-C₆H₄), 136.5 (C8), 126.0 (-C₆H₄), 125.5 (-C₆H₄), 123.6 (-C₆H₄), 118.1 (C5), 116.9 (-C₆H₄), 116.3 (-C₆H₄), 88.1 (C4'), 83.1 (C1'), 71.2 (C3'), 62.2 (C5'), 40.2 (C2') ppm.

HR-MS (ESI+): m/z berechnet für [C₁₆H₁₉N₆O₄]⁺ ([M+H]⁺): 359,1463, gefunden: 359,1461.

IR: $\tilde{\nu}$ = 3451 (w), 3348 (w), 2928 (w), 2360 (w), 1674 (m), 1601 (w), 1576 (m), 1504 (w), 1447 (w), 1308 (w), 1254 (m), 1216 (s), 1168 (s), 1041 (w), 1025 (m), 962 (s), 940 (s), 890 (w), 753 (m), 662 (m) cm⁻¹.

Schmelzpunkt: 223 °C.

5.4.1.3 Synthese von N^2 -(2-Acetylaminophenyl)-2'-desoxyguanosin (14)



Acetylierung mit Essigsäureanhydrid: N^2 -(2-Aminophenyl)-2'-desoxyguanosin **13** (100 mg, 0,279 mmol, 1,0 Äq.) wurde in 4 mL DMF/Pyridin 2:1 suspendiert und 0,03 mL Essigsäureanhydrid (0,3 mmol, 1,1 Äq.) wurden zugetropft. Die Suspension wurde 20 h bei RT gerührt und weitere 4,0 Äq. Essigsäureanhydrid (0,12 mL, 1,1 mmol) wurden zugetropft. Nach 20 h Röhren bei RT wurde das Lösungsmittel *in vacuo* entfernt und der Rückstand in 2 mL konz. methanolischen Ammoniak suspendiert. Die Suspension wurde 20 h bei RT gerührt, das Lösungsmittel *in vacuo* entfernt und der Rückstand *via* Säulenchromatografie (DCM/MeOH 10 → 5:1 + 1 % NH₄OH) gereinigt. N^2 -(2-Acetylaminophenyl)-2'-desoxyguanosin **14** (43 mg, 0,11 mmol, 38 %) wurde als weißer Feststoff erhalten.

Acetylierung mit Thioessigsäure: N^2 -(2-Aminophenyl)-2'-desoxyguanosin **13** (50 mg, 0,14 mmol, 1,0 Äq.), CuSO₄ Pentahydrat (10 mg, 0,040 mmol, 0,3 Äq.) und Thioessigsäure (0,010 mL, 0,14 mmol, 1,0 Äq.) wurden in 0,5 mL MeOH gelöst und 15 h bei RT gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand *via* Säulenchromatografie (DCM/MeOH 10 → 3:1 + 1 % NH₄OH) gereinigt. N^2 -(2-Acetylaminophenyl)-2'-desoxyguanosin **14** (42 mg, 0,11 mmol, 75 %) wurde als weißer Feststoff erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 11.24 (s, 1H, NH), 9,57 (s, 1H, NH), 7,97 (m, 3H, NH, -C₆H₄, C8-H), 7,31 (d, *J* = 7.8, 1H, -C₆H₄), 7.21 (t, *J* = 7.2, 1H, -C₆H₄), 7.07 (td, *J* = 7.7, 1.4, 1H, -C₆H₄), 6.09 (t, *J* = 6.8, 1H, C1'-H), 5.23 (d, *J* = 3.9, 1H, C3'-OH), 4.82 (t, *J* = 5.3, 1H, C5'-OH), 4.29 (d, *J* = 16.9, 1H, C3'-H), 3.76 (dd, *J* = 7.7, 4.7, 1H, C4'-H), 3.54 – 3.38 (m, 2H, C5'-H), 2.64 – 2.51 (m, 1H, C2'-H), 2.19 (ddd, *J* = 13.2, 6.2, 3.3, 1H, C2'-H), 2.05 (d, *J* = 3.2, 3H, CH₃CO) ppm.

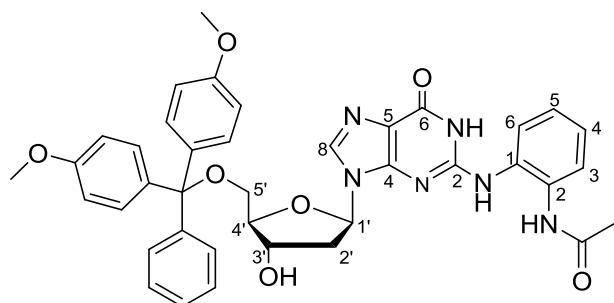
¹³C-NMR (101 MHz, DMSO-*d*₆) δ 169.6 (CH₃CO), 157.2 (C6), 150.3 (C2), 150.0 (C4), 136,9 (C8), 132,9 (-C₆H₄), 129,8 (-C₆H₄), 126,5 (-C₆H₄), 126,2 (-C₆H₄), 124,2 (-C₆H₄), 123,6 (-C₆H₄), 118,7 (C5), 88,1 (C4'), 83,3 (C1'), 71,1 (C3'), 62,1 (C5'), 40,1 (C2'), 23,8 (CH₃CO) ppm.

HR-MS (ESI+): m/z berechnet für $[C_{18}H_{20}N_6O_5Na]^+$ ($[M+Na]^+$): 423,1387, gefunden: 423,1387.

IR: $\tilde{\nu}$ = 3328 (m), 3126 (w), 2935 (w), 2360 (w), 1681 (m), 1590 (m), 1566 (s), 1451 (m), 1447 (s), 1171 (w), 1097 (m), 1054 (s), 930 (w), 780 (m), 757 (s), 658 (s) cm^{-1} .

Schmelzpunkt: 93 °C.

5.4.1.4 Synthese von N^2 -(2-Acetylaminophenyl)-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin (15)



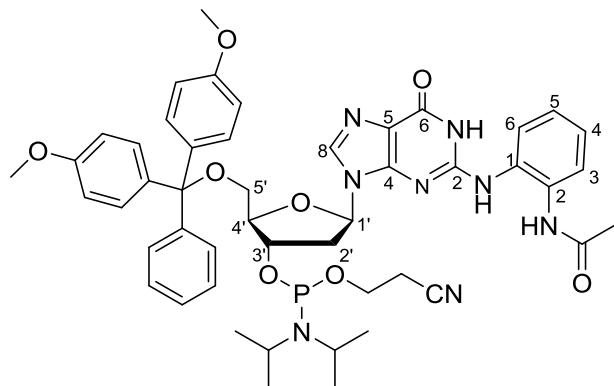
In einem Kolben mit Molsieb (3 Å) wurde N^2 -(2-Acetylaminophenyl)-2'-desoxyguanosin **14** (0,330 g, 0,824 mmol, 1,0 Äq.) in 10 mL trockenem Pyridin gelöst und 1,5 h bei Raumtemperatur gerührt. Triethylamin (0,12 mL, 0,87 mmol, 1,1 Äq.) und DMTr-Cl (0,352 g, 0,907 mmol, 1,1 Äq.) wurden hinzugegeben und das Reaktionsgemisch 3 h bei RT gerührt. Es wurden 4 mL MeOH zugegeben und 30 Minuten gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand *via* Säulenchromatografie (DCM/MeOH 30 → 10:1 + 1 % Pyridin) gereinigt. N^2 -(2-Acetylaminophenyl)-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin **15** (0,223 g, 0,317 mmol, 39 %) wurde als rotbrauner Schaum erhalten.

1H -NMR (400 MHz, DMSO-*d*₆) δ 11.25 (s, 1H, NH), 9.52 (s, 1H, NH), 7.99 (s, 1H, NH), 7.91 (dd, *J* = 7.9, 1.6 Hz, 1H, -C₆H₄), 7.86 (s, 1H, C8-H), 7.32 – 7.24 (m, 3H, -C₆H₄, DMT), 7.23 – 7.12 (m, 6H, DMT), 7.06 (dtd, *J* = 15.0, 7.7, 1.7 Hz, 2H, -C₆H₄), 6.85 – 6.65 (m, 4H, DMT), 6.13 (t, *J* = 6.5 Hz, 1H, C1'-H), 5.26 (d, *J* = 4.6 Hz, 1H, C3'-OH), 4.24 (t, *J* = 5.7 Hz, 1H, C3'-H), 3.87 (dt, *J* = 6.9, 4.1 Hz, 1H, C4'-H), 3.67 (s, 6H, OCH₃), 3.07 (qd, *J* = 10.2, 5.0 Hz, 2H, C5'-H), 2.66 (dt, *J* = 13.3, 6.6 Hz, 1H, C2'-H), 2.35 – 2.16 (m, 1H, C2'-H), 2.05 (s, 3H, CH₃CO) ppm.

¹³C-NMR (101 MHz, DMSO-*d*₆) δ 169.6 (CH_3CO), 158.4 (DMT), 157.1 (C6), 150.1 (C2), 150.0 (C4), 145.3 (DMT), 136.9 (C8), 135.9 (DMT), 132.9 (-C₆H₄), 130.1 (DMT), 130.0 (-C₆H₄), 128.1 (DMT), 127.0 (DMT), 126.5 (-C₆H₄), 126.1 (-C₆H₄), 124.1 (-C₆H₄), 123.3 (-C₆H₄), 119.0 (C5), 113.5 (DMT), 86.1 (C4'), 85.9 (CAr₃), 83.3 (C1'), 71.0 (C3'), 64.7 (C5'), 55.4 (OCH₃), 39.4 (C2'), 23.8 (CH_3CO) ppm.

HR-MS (ESI+): m/z berechnet für [C₃₉H₃₈N₆O₇Na]⁺ ([M+Na]⁺): 725,2694, gefunden: 725,2716.

5.4.1.5 Synthese von N²-(2-Acetylaminophenyl)-3'-O-[N,N-diisopropylamino-(2-cyanoethoxy)-phosphinyl]-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin (7)



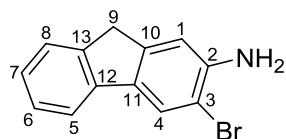
*N*²-(2-Acetylaminophenyl)-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin **15** (114 mg, 0,162 mmol, 1,0 Äq.) wurde in 3 mL trockenem DCM gelöst und DIAT (33 mg, 0,20 mmol, 1,2 Äq.) wurde zugegeben. Das Reaktionsgemisch wurde drei Mal entgast, (iPr₂N)₂PO(CH₂)₂CN (59 mg, 0,20 mmol, 1,2 Äq.) wurde zugegeben und es wurde 1,5 h gerührt. Der Rückstand wurde säulenchromatografisch (DCM/MeOH 15:1 + 1 % Pyridin) aufgereinigt und *N*²-(2-Acetylaminophenyl)-3'-O-[N,N-diisopropylamino-(2-cyanoethoxy)-phosphinyl]-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin (146 mg, 0,162 mmol, quant.) wurde als rotbrauner Schaum erhalten.

³¹P-NMR (162 MHz, CDCl₃) δ 149.03, 148.67 ppm.

HR-MS (ESI+): m/z berechnet für [C₄₈H₅₆N₈O₈P]⁺ ([M+H]⁺): 903,3880, gefunden: 903,3953.

5.4.2 Synthese des *N*²-AAF-dG Phosphoramidits

5.4.2.1 Synthese von 2-Amino-3-bromfluoren



Eine Lösung von *N*-Bromsuccinimid (5,17 g, 29,0 mmol, 1,2 Äq.) in DCM (200 mL) wurde bei 0 °C tropfenweise unter Rühren zu einer Lösung von 2-Aminofluoren (4,39 g, 24,2 mmol, 1,0 Äq.) in DCM (45 mL) gegeben. Nach dem Entfernen des Lösungsmittels *in vacuo* wurde der Rückstand mittels Säulenchromatografie (*i*Hex/EtOAc 30:1) gereinigt. 2-Amino-3-bromfluoren (5,57 g, 21,4 mmol, 88 %) wurde als beiger Feststoff erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 7,85 (s, 1H, C4-H), 7,66 (ddd, 1H, ³J=7,50, ⁴J=0,88, ⁵J=0,88 Hz, C5-H), 7,42 (ddd, 1H, ³J=7,37, J=0,93, J=0,93 Hz, C8-H), 7,24 (ddd, 1H, ³J=7,56, ³J=7,56, ⁴J=1,07 Hz, C6-H), 7,13 (ddd, 1H, ³J=7,46, ³J=7,46, ⁴J=1,20 Hz, C7-H), 6,97 (t, 1H, ⁴J=0,90 Hz, C1-H), 5,35 (bs, 2H, NH₂), 3,72 (s, 2H, C9-H) ppm.

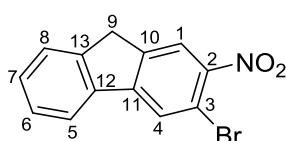
¹³C-NMR (101 MHz, DMSO-*d*₆) δ 145,2 (C2), 144,3 (C10), 142,3 (C13), 141,4 (C12), 132,0 (C11), 127,1 (C6), 125,6 (C7), 125,2 (C8), 124,1 (C4), 119,1 (C5), 112,1 (C1), 107,0 (C3), 36,4 (C9) ppm.

HR-MS (ESI+): m/z berechnet für [C₁₃H₁₁BrN]⁺ ([M+H]⁺): 260,0070, gefunden: 260,0069.

IR: ν = 3421,8 (m), 3336,9 (m), 3039,0 (w), 1612,2 (m), 1411,8 (m), 870,5 (m), 749,2 (s), 730,3 (s) cm⁻¹.

Schmelzbereich: 143 – 144 °C.

5.4.2.2 Synthese von 3-Brom-2-nitrofluoren



2-Amino-3-bromfluoren (1,63 g, 6,27 mmol, 1,0 Äq.) wurde in 100 mL DCM bei RT gelöst. Unter Rühren wurden 5,77 g *m*CPBA 75 % (25,1 mmol, 4,0 Äq.) zugegeben. Nach 72 Stunden wurde die Suspension mit gesättigter NaHSO₄-Lösung neutralisiert und mit DCM extrahiert. Die organische Phase wurde mit Wasser gewaschen, mit MgSO₄ getrocknet und das Lösungsmittel wurde evaporiert. Der Rückstand wurde mittels Säulenchromatografie (*i*Hex/EtOAc 20:1) gereinigt. 3-Brom-2-nitrofluoren (1,57 g, 5,40 mmol, 86 %) wurde als gelbes Pulver erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 8,44 (s, 1H, C4-H), 8,21 (s, 1H, C1-H), 8,07 (m, 1H, C5-H), 7,63 (m, 1H, C8-H), 7,43 (m, 2H, C6-H und C7-H), 3,97 (s, 2H, CH₂) ppm.

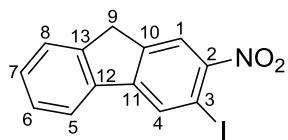
¹³C-NMR (101 MHz, DMSO-*d*₆) δ 148,0 (C2), 147,0 (C11), 145,3 (C13), 143,9 (C10), 138,5 (C12), 129,4 (C7), 127,7 (C6), 126,1 (C4), 125,9 (C8), 122,9 (C1), 122,4 (C5), 112,6 (C3), 36,9 (C9) ppm.

HR-MS (ESI+): m/z berechnet für [C₁₃H₇O₂NBr]⁺ ([M-H]⁺): 287,9665, gefunden: 287,9665.

IR: ν = 1514,6 (m), 1327,0 (m), 870,6 (m), 749,0 (s), 765,1 (s), 701,3 (s) cm⁻¹.

Schmelzbereich: 126 – 130 °C.

5.4.2.3 Synthese von 3-Iod-2-nitrofluoren



3-Brom-2-nitrofluoren (100 mg, 0,345 mmol, 1,0 Äq.), Cul (3 mg, 0,02 mmol, 6 mol%) und NaI (103 mg, 0,689 mmol, 2,0 Äq.) wurden unter Argonatmosphäre vorgelegt und entgast. Nach der Zugabe von 1,3-Propandiamin (3 mg, 0,04 mmol, 0,1 Äq.) und *n*-Butanol (0,35 mL) wurde die Suspension 24 h bei 130 °C gerührt. Das Reaktionsgemisch wurde auf RT gekühlt, mit konz. NH₃ (1 mL) verdünnt und in Wasser (20 mL) gegeben. Nach der Extraktion mit DCM wurde die organische Phase über MgSO₄ getrocknet. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand mittels Säulenchromatografie (*i*Hex/EtOAc 40:1) gereinigt. 3-Iod-2-nitrofluoren (35 mg, 0,10 mmol, 29 %) wurde als gelber Feststoff erhalten.

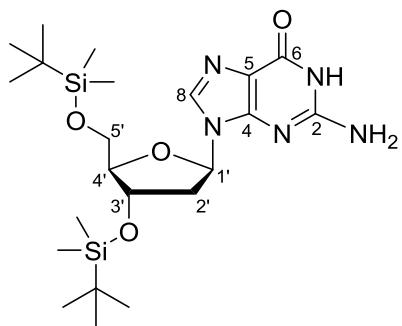
¹H-NMR (400 MHz, DMSO-*d*₆): δ 8,53 (s, 1H, C4-H), 8,07 (s, 1H, C1-H), 7,97 (m, 1H, C5-H), 7,55 (m, 1H, C8-H), 7,35 (m, 2H, C6-H und C7-H), 3,87 (s, 2H, CH₂) ppm.

¹³C-NMR (101 MHz, DMSO-*d*₆): δ 151,6 (C2), 147,0 (C11), 145,2 (C13), 144,5 (C10), 138,4 (C12), 132,6 (C4), 129,3 (C7), 127,7 (C6), 125,9 (C8), 122,5 (C1), 122,3 (C5), 86,8 (C3), 36,9 (C9) ppm.

HR-MS (ESI-): m/z berechnet für [C₁₃H₇INO₂]⁺ ([M-H]⁻): 335,9527, gefunden: 335,9529.

IR: $\tilde{\nu}$ = 1600,2 (w), 1447,6 (m), 1513,8 (m), 1439,8 (w), 1396,4 (m), 1321,2 (s), 1100,0 (m), 953,4 (w), 924,5 (m), 894,3 (m), 804,2 (m), 767,9 (s), 749,8 (s), 723,9 (s) cm⁻¹.

5.4.2.4 Synthese von 3',5'-Bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin (17)



2'-Desoxyguanosin Monohydrat (7,00 g, 24,5 mmol, 1,0 Äq.), TBS-Chlorid (12,2 g, 81,1 mmol, 3,3 Äq.) und Imidazol (11,0 g, 161 mmol, 6,6 Äq.) wurden in 25 mL DMF suspendiert und 7 Stunden bei RT gerührt. Nach Zugabe von 50 mL Ethanol wurde die Suspension 16 Stunden bei -18 °C aufbewahrt. Die Suspension wurde filtriert und der Filterkuchen mit 25 mL Ethanol gewaschen. 3',5'-Bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin **17** (12,0 g, 24,5 mmol, quant.) wurde als weißer Feststoff erhalten.

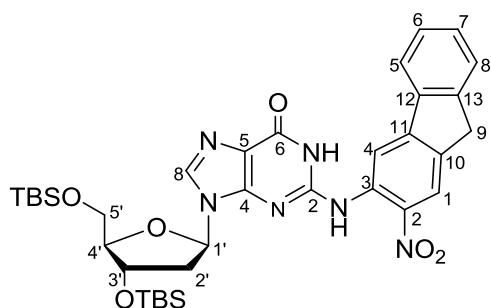
¹H-NMR (400 MHz, DMSO-*d*₆) δ 10,56 (s, 1H, NH), 7,82 (s, 1H, C8-H), 6,42 (s, 2H, NH₂), 6,07 (dd, ³J=7,6, ³J=6,0 Hz, 1H, C1'-H), 4,45 (dt, ³J=5,6, ³J=3,0 Hz, 1H, C3'-H), 3,77 (m, 1H, C4'-H), 3,63 (ddd, J=15,6, J=11,0, J=5,2 Hz, 2H, C5'-H), 2,60 (ddd, ²J=13,2, ³J=7,7, ³J=5,6 Hz, 1H, C2'-H), 2,20 (ddd, ²J=13,2, ³J=6,0, ³J=3,2 Hz, 1H), 0,85 (s, 9H, (CH₃)₃C), 0,84 (s, 9H, (CH₃)₃C), 0,07 (s, 6H, CH₃Si), 0,01 (s, 3H, CH₃Si), 0,00 (s, 3H, CH₃Si) ppm.

¹³C-NMR (101 MHz, DMSO-*d*₆) δ 157,1 (C6), 154,1 (C2), 151,4 (C4), 135,3 (C8), 117,1 (C5), 87,4 (C4'), 82,6 (C1'), 72,6 (C3'), 63,2 (C5'), 39,6 (C2'), 26,2 ((CH₃)₃C), 26,1 ((CH₃)₃C), 18,4 ((CH₃)₃C), 18,1 ((CH₃)₃C), -4,3 (CH₃Si), -4,5 (CH₃Si), -5,0 (CH₃Si), -5,1 (CH₃Si) ppm.

HR-MS (ESI+): m/z berechnet für $[C_{22}H_{42}N_5O_4Si_2]^+$ ($[M+H]^+$): 496,2770, gefunden: 496,2770.

IR: $\tilde{\nu}$ = 3314,9 (m), 3163,0 (m), 2952,0 (w), 2926,0 (m), 2855,0 (m), 2737,0 (w), 1689,6 (s), 1602,0 (m), 1380,8 (m), 1152,9 (m), 1128,9 (m), 820,0 (s), 774,0 (s) cm^{-1} .

5.4.2.5 Synthese von *N*²-(2-Nitrofluoren-3-yl)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyguanosin (12)



$\text{Pd}_2(\text{dba})_3$ (55 mg, 0,060 mmol, 0,15 Äq.) und Xantphos (70 mg, 0,12 mmol, 0,30 Äq.) wurden in 1 mL trockenem Toluol gelöst und 10 min bei RT gerührt. 3',5'-Bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyguanosin **17** (200 mg, 0,403 mmol, 1,00 Äq.), 3-Bromo-2-nitrofluoren (140 mg, 0,484 mmol, 1,20 Äq.) und Cs_2CO_3 (263 mg, 0,806 mmol, 2,00 Äq.) und 1 mL trockenes Toluol wurden hinzugegeben und das grünbraune Reaktionsgemisch wurde 16 h bei 70 °C gerührt, sodass eine rote Suspension entstand. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand wurde *via* Säulenchromatografie (*dryload*, DCM/MeOH 80 → 40:1) gereinigt. *N*²-(2-Nitrofluorenyl)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyguanosin **12** (161 mg, 0,229 mmol, 57 %) wurde als oranger Feststoff erhalten.

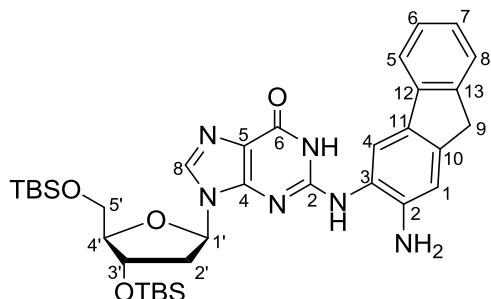
¹H-NMR (599 MHz, CDCl_3) δ 12.86 (s, 1H, NH), 10.25 (s, 1H, NH), 8.75 (s, 1H, $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-H}$), 8.13 (s, 1H, $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-H4}$), 7.94 (s, 1H, C8-H), 7.79 (d, J = 7.4 Hz, 1H, $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-H5}$), 7.48 (d, J = 7.3 Hz, 1H, $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-H8}$), 7.32 – 7.19 (m, 2H, $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-H6}$ und $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-H7}$), 6.19 (t, J = 6.7 Hz, 1H, C1'-H), 4.46 (q, J = 3.7 Hz, 1H, C3'-H), 4.03 – 3.84 (m, 3H, CH₂ und C4'-H), 3.70 (dd, J = 10.7, 3.6 Hz, 1H, C5'-H), 3.64 (dd, J = 11.0, 5.4 Hz, 1H, C5'-H), 2.66 – 2.51 (m, 1H, C2'-H), 2.43 (d, J = 12.3 Hz, 1H, C2'-H), 0.83 (s, 18H, $(\text{CH}_3)_3\text{C}$), 0.03 (s, 3H, SiCH₃), 0.03 (s, 3H, SiCH₃), -0.01 (s, 3H, SiCH₃), -0.04 (s, 3H, SiCH₃) ppm.

¹³C-NMR (151 MHz, CDCl_3) δ 158.7 (C6), 149.2 (C4), 148.5 (C2), 148.0 ($-\text{C}_{13}\text{H}_8\text{NO}_2$), 145.5 ($-\text{C}_{13}\text{H}_8\text{NO}_2$), 139.5 ($-\text{C}_{13}\text{H}_8\text{NO}_2$), 137.7 ($-\text{C}_{13}\text{H}_8\text{NO}_2$), 137.0 (C8), 133.8 ($-\text{C}_{13}\text{H}_8\text{NO}_2$), 128.8 ($-\text{C}_{13}\text{H}_8\text{NO}_2\text{-C6}$ oder $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-C7}$), 127.3 ($-\text{C}_{13}\text{H}_8\text{NO}_2\text{-C6}$ oder $-\text{C}_{13}\text{H}_8\text{NO}_2\text{-C7}$), 125.2

(-C₁₃H₈NO₂-C8), 121.8 (-C₁₃H₈NO₂-C4), 121.1 (-C₁₃H₈NO₂-C5), 119.3 (C5), 114.6 (-C₁₃H₈NO₂-C1), 87.9 (C4'), 84.3 (C1'), 72.3 (C3'), 63.0 (C5'), 40.8 (C2'), 36.4 (CH₂), 25.9 ((CH₃)₃C), 25.7 ((CH₃)₃C), 18.3 ((CH₃)₃C), 17.9 ((CH₃)₃C), -4.73 (SiCH₃), -4.9 (SiCH₃), -5.52 (SiCH₃), -5.6 (SiCH₃) ppm.

HR-MS (ESI-): m/z berechnet für [C₃₅H₄₇N₆O₆Si₂]⁻ ([M-H]⁻): 703,3101, gefunden: 703,3112.

5.4.2.6 Synthese von N²-(2-Aminofluoren-3-yl)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin (19)



N²-(2-Nitrofluoren-3-yl)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin **12** (2,32 g, 3,29 mmol, 1,0 Äq.) wurde in einem Gemisch aus EtOAc (55 mL) und MeOH (35 mL) gelöst und Pd/C (564 mg, 0,329 mmol, 0,1 Äq.) wurde zugegeben. Anschließend wurde die Ar-Atmosphäre durch H₂-Atmosphäre ausgetauscht. Nach Röhren bei RT über Nacht und säulenchromatographischer Reinigung (DCM/MeOH 30 → 15:1) wurde N²-(2-Aminofluoren-3-yl)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin **19** (1,09 g, 1,61 mmol, 49 %) als brauner Feststoff erhalten.

¹H-NMR (599 MHz, CDCl₃) δ 9.05 (s, 1H, NH), 7.74 (s, 1H, C₁₃H₈NH₂-H4), 7.64 (s, 1H, C8-H), 7.50 (d, J = 7.5 Hz, 1H, -C₁₃H₈NH₂-H5), 7.39 (d, J = 7.6 Hz, 1H, -C₁₃H₈NH₂-H7), 7.21 (d, J = 7.5 Hz, 1H, -C₁₃H₈NH₂-H6), 7.12 (t, J = 7.5 Hz, 1H, -C₁₃H₈NH₂-H8), 6.82 (s, 1H, -C₁₃H₈NH₂-H1), 5.96 (s, 1H, C1'-H), 4.23 (s, 1H, C3'-H), 3.77 (s, 1H, C4'-H), 3.68 (s, 2H, CH₂), 3.55 (d, J = 10.7 Hz, 1H, C5'-H), 3.52 – 3.38 (m, 1H, C5'-H), 2.47 (s, 1H, C2'-H), 2.10 (s, 1H, C2'-H), 0.78 (s, 9H, C(CH₃)₃), 0.76 (s, 9H, C(CH₃)₃), -0.08 (s, 6H, SiCH₃), -0.10 (s, 6H, SiCH₃) ppm.

¹³C-NMR (151 MHz, CDCl₃) δ 159.1 (C6), 151.3 (C2), 150.6 (C4), 142.5 (-C₁₃H₈NH₂-C12), 142.2 (-C₁₃H₈NH₂-C13), 141.5 (-C₁₃H₈NH₂), 136.3 (C8), 132.7 (-C₁₃H₈NH₂), 126.6 (-C₁₃H₈NH₂-C6), 125.0 (-C₁₃H₈NH₂-C8), 124.6 (-C₁₃H₈NH₂-C7), 123.1 (-C₁₃H₈NH₂), 118.5 (-C₁₃H₈NH₂-C5), 118.1 (-C₁₃H₈NH₂-C4), 113.0 (-C₁₃H₈NH₂-C1), 87.8 (C4'), 84.3 (C1'), 72.1

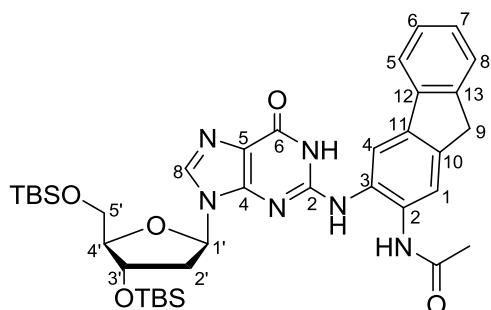
(C3'), 62.9 (C5'), 40.3 (C2'), 36.5 CH₂), 25.8 (C(CH₃)₃), 25.6 (C(CH₃)₃), 18.2 (C(CH₃)₃), 17.8 (C(CH₃)₃), -4.8 (SiCH₃), -5.0 (SiCH₃), -5.5 (SiCH₃), -5.6 (SiCH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₃₅H₅₁N₆O₄Si₂]⁺ ([M+H]⁺): 675,3505, gefunden: 675,3516.

IR: $\tilde{\nu}$ = 2952,2 (w); 2927,7 (w); 2855,6 (w); 1689,1 (m); 1573,3 (m); 1456,0 (m); 1360,5 (m); 1251,3 (m); 1074,4 (m); 1027,2 (m); 8125,5 (m); 776,7 (s); 761,8 (s); 724,3 (s); 666,9 (m) cm⁻¹.

Schmelzbereich: 194 – 196 °C.

5.4.2.7 Synthese von N²-(2-Acetylaminofluoren-3-yl)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin (20)



Zu einer Lösung aus N²-(2-Aminofluoren-3-yl)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin **19** (976 mg, 1,45 mmol, 1,0 Äq.) in Pyridin (7,8 mL) und DCM (23 mL) wurde Ac₂O (738 mg, 7,23 mmol, 5,0 Äq.) gegeben und 20 h bei RT gerührt. Nach Zugabe von einer gesättigten NaHCO₃-Lösung wurde 10 min gerührt und mit DCM extrahiert. Die organische Phase wurde über MgSO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde in 10 mL MeOH und 7,5 mL NH₃ (7 N in MeOH) gelöst und die Lösung wurde 20 h bei RT gerührt. Nach säulenchromatographischer Aufreinigung (DCM/MeOH 50 → 20:1) wurde N²-(2-Acetylaminofluoren-3-yl)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin **20** (759 mg, 1,06 mmol, 73 %) als gelblicher Feststoff erhalten.

¹H-NMR (599 MHz, CDCl₃) δ 11.56 (s, 1H, NH), 9.56 (s, 1H, NH), 8.61 (s, 1H, -C₁₃H₈NHAc), 8.20 (s, 1H, C8-H), 7.69 (d, *J* = 7.6 Hz, 1H, -C₁₃H₈NHAc-H5), 7.54 (s, 1H, -C₁₃H₈NHAc), 7.52 (d, *J* = 7.6 Hz, 1H, -C₁₃H₈NHAc-H8), 7.34 (t, *J* = 7.4 Hz, 1H, -C₁₃H₈NHAc-H6), 7.28 (t, *J* = 7.4 Hz, 1H, -C₁₃H₈NHAc-H7), 6.30 (s, 1H, C1'-H), 4.52 (s, 1H, C3'-H), 4.03 – 3.93 (m, 1H, C4'-H), 3.88 (s, 3H, C5'-H und CH₂), 3.81 – 3.72 (m, 1H, C5'-H), 2.46 (d, *J* = 5.3 Hz, 1H,

C2'-H), 2.13 (s, 3H, COCH₃), 0.91 (s, 9H, C(CH₃)₃), 0.80 (s, 9H, C(CH₃)₃), 0.08 (s, 6H, SiCH₃), -0.01 (s, 3H, SiCH₃), -0.02 (s, 3H, SiCH₃) ppm.

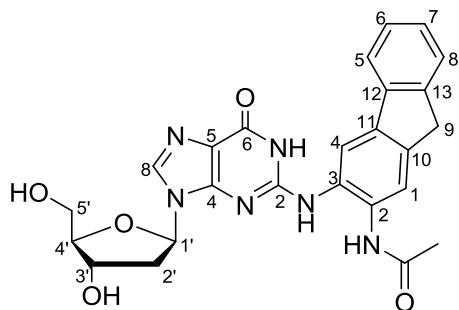
¹³C-NMR (151 MHz, CDCl₃) δ 171.5 (COCH₃), 158.3 (C6), 150.6 (C2), 150.1 (C4), 143.9 (-C₁₃H₈NHAc), 141.5 (-C₁₃H₈NHAc), 140.7 (-C₁₃H₈NHAc-C10), 138.6 (-C₁₃H₈NHAc), 136.7 (C8), 127.3 (-C₁₃H₈NHAc), 126.8 (-C₁₃H₈NHAc-C6), 126.6 (-C₁₃H₈NHAc-C7), 125.0 (-C₁₃H₈NHAc-C8), 119.5 (-C₁₃H₈NHAc-C5), 117.3 (C5), 113.5 (-C₁₃H₈NHAc), 87.6 (C4'), 84.5 (C1'), 70.3 (C3'), 62.1 (C5'), 42.2 (C2'), 36.6 (CH₂), 26.0 (C(CH₃)₃), 25.6 (C(CH₃)₃), 23.3 (COCH₃), 18.4 (C(CH₃)₃), 17.9 (C(CH₃)₃), -4.7 (SiCH₃), -4.9 (SiCH₃), -5.5 (SiCH₃) ppm.

HR-MS (ESI-): m/z berechnet für [C₃₇H₅₁N₆O₅Si₂]⁻ ([M+H]⁻): 715,3465, gefunden: 715,3460.

IR: $\tilde{\nu}$ = 3203,0 (w); 2952,6 (w); 2928,2 (w); 2855,9 (w); 1682,4 (m); 1591,0 (s); 1456,2 (m); 1427,7 (m); 1252,9 (m); 1093,6 (m); 812,4 (m); 777,4 (s); 764,7 (s); 725,8 (m); 668,5 (m) cm⁻¹.

Schmelzbereich: 195 – 197 °C.

5.4.2.8 Synthese von N²-(2-Acetylaminofluoren-1-yl)-2'-desoxyguanosin (18)



Zu einer Lösung aus N²-(2-Acetylaminofluoren-3-yl)-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyguanosin **20** (759 mg, 1,06 mol, 1,0 Äq.) in THF (16 mL) wurden AcOH (0,363 mL, 6,35 mol, 6,0 Äq.) und TBAF (3,18 mL, 3,17 mol, 3,0 Äq.) gegeben. Nach 20 h Rühren bei RT wurden zu der braunen Lösung erneut AcOH (0,363 mL, 6,35 mol, 6,0 Äq.) und TBAF (3,18 mL, 3,17 mol, 3,0 Äq.) gegeben und 24 h bei RT gerührt. Nach zweimaliger säulenchromatographischer Reinigung (DCM/MeOH 30 → 3:1 + 1 % NH₄OH; DCM/MeOH 7 → 5:1 + NH₄OH) wurde N²-(2-Acetylaminofluoren-1-yl)-2'-desoxyguanosin **18** (258 mg, 0,528 mmol, 50 %) als gelblicher Feststoff erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 11.36 (s, 1H, NH), 9.64 (s, 1H, NH), 8.61 (s, 1H, -C₁₃H₈NHAc-H1), 8.35 (bs, 1H, NH), 8.05 (s, 1H, C8-H), 7.83 (d, *J* = 7.5 Hz, 1H, -C₁₃H₈NHAc-H5), 7.61 (s, 1H, -C₁₃H₈NHAc-H4), 7.58 (d, *J* = 7.4 Hz, 1H, -C₁₃H₈NHAc-H8), 7.37 (dd, *J* = 7.4, 1.3 Hz, 1H, -C₁₃H₈NHAc-H6), 7.31 (td, *J* = 7.4, 1.2 Hz, 1H, -C₁₃H₈NHAc-H7), 6.20 (dd, *J* = 7.9, 5.9 Hz, 1H, C1'-H), 5.29 (bs, 1H, C3'-OH), 4.87 (s, 1H, C5'-OH), 4.27 (dd, *J* = 5.6, 2.9 Hz, 1H, C3'-H), 3.91 (s, 2H, CH₂), 3.87 – 3.78 (m, 1H, C4'-H), 3.45 (m, 2H, C5'-H), 2.62 (ddd, *J* = 13.3, 8.0, 5.6 Hz, 1H, C2'-H), 2.32 – 2.22 (m, 1H, C2'-H), 2.13 (s, 3H, CH₃) ppm.

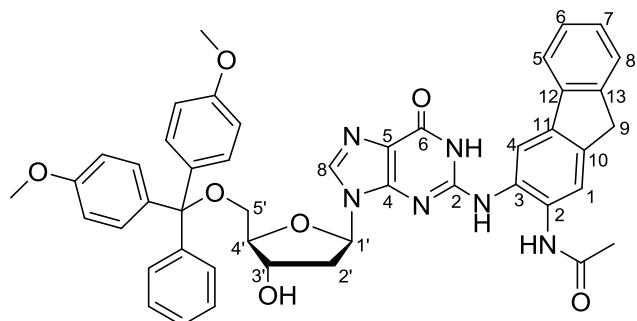
¹³C-NMR (101 MHz, DMSO-*d*₆) δ 169.8 (COCH₃), 157.2 (C6), 150.5 (C2), 150.1 (C4), 144.0 (-C₁₃H₈NHAc), 141.3 (-C₁₃H₈NHAc), 139.1 (-C₁₃H₈NHAc), 138.7 (-C₁₃H₈NHAc), 136.6 (C8), 131.9 (-C₁₃H₈NHAc), 128.9 (-C₁₃H₈NHAc), 127.4 (-C₁₃H₈NHAc-C6), 127.1 (-C₁₃H₈NHAc-C7), 125.6 (-C₁₃H₈NHAc-C8), 123.0 (-C₁₃H₈NHAc-C4), 120.2 (-C₁₃H₈NHAc-C5), 118.6 (C5), 114.7 (-C₁₃H₈NHAc-C1), 88.3 (C4'), 83.6 (C1'), 71.2 (C3'), 62.1 (C5'), 39.4 (C2'), 36.5 (CH₂), 24.0 (CH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₂₅H₂₅N₆O₅]⁺ ([M+H]⁺): 489,1881, gefunden: 489,1886.

IR: $\tilde{\nu}$ = 3122,5 (w); 2960,4 (m); 2874,8 (w); 2250,3 (w); 2124,8 (w); 1686,4 (s); 1574,9 (s); 1506,9 (m); 1455,8 (s); 1427,4 (s); 1312,6 (m); 1153,9 m); 1092,7 (m); 1048,9 (s); 1023,5 (s); 1003,8 (s); 951,1 (m); 827,3 (m); 779,1 (s); 763,9 (s); 713,3 (s) cm⁻¹.

Zersetzungsbereich: 156 – 160 °C.

5.4.2.9 Synthese von N²-(2-Acetylaminofluoren-1-yl)-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin (21)



Zu einer Lösung aus N²-(2-Acetylaminofluoren-1-yl)-2'-desoxyguanosin **18** (258 mg, 0,528 mmol, 1,0 Äq.) in Pyridin (11 mL) wurde trockenes NEt₃ (0,110 mL, 0,792 mmol,

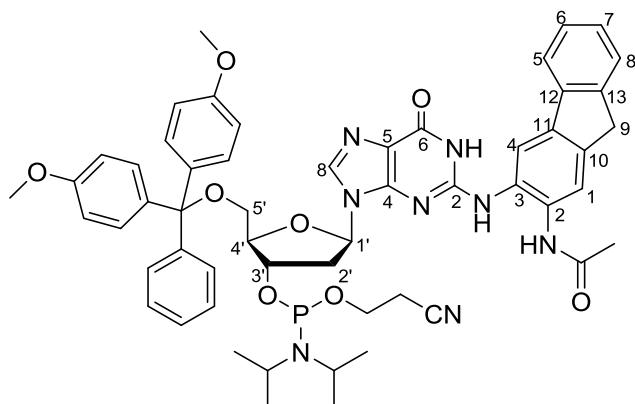
1,5 Äq.) und DMTr-Cl (308 mg, 0,792 mmol, 1,5 Äq.) gegeben und 2 h bei RT gerührt. Nach weiterer Zugabe von DMTr-Cl (103 mg, 0,264 mmol, 0,5 Äq.) und 17 h Röhren bei RT wurde nach säulenchromatographischer Aufreinigung (DCM/MeOH 20 → 15:1 + 1 % Pyridin) *N*²-(2-Acetylaminofluoren-1-yl)-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin **21** als gelber Feststoff erhalten, welcher auf Grund seiner Instabilität direkt weiter umgesetzt wurde.

¹H-NMR (800 MHz, DMSO-*d*₆) δ 11.34 (s, 1H, NH), 9.64 (s, 1H, NH), 8.71 (s, 1H, -C₁₃H₈NHAc), 8.21 (s, 1H, NH), 7.94 (s, 1H, C8-H), 7.85 (d, *J* = 7.4 Hz, 1H, -C₁₃H₈NHAc), 7.61 (m, 2H, -C₁₃H₈NHAc), 7.39 (m, 1H, -C₁₃H₈NHAc), 7.34 (td, *J* = 7.3, 1.2 Hz, 1H, -C₁₃H₈NHAc), 7.28 – 7.22 (m, 2H, DMTr), 7.22 – 7.15 (m, 3H, DMTr), 7.15 – 7.10 (m, 4H, DMTr), 6.81 – 6.74 (m, 4H, DMTr), 6.22 (t, *J* = 6.4 Hz, 1H, C1'-H), 5.32 (d, *J* = 4.4 Hz, 1H, C3'-OH), 4.30 (dq, *J* = 10.3, 5.3, 4.7 Hz, 1H, C3'-H), 4.12 (q, *J* = 5.2 Hz, 1H, C5'-OH), 3.92 (d, *J* = 21.8 Hz, 1H, CH₂), 3.88 (d, *J* = 21.9 Hz, 1H, CH₂), 3.78 (dt, *J* = 6.0, 4.1 Hz, 1H, C4'-H), 3.71 (s, 3H, OCH₃), 3.70 (s, 3H, OCH₃), 3.12 (dd, *J* = 10.5, 4.4 Hz, 1H, C5'-H), 3.09 – 3.02 (m, 1H, C5'-H), 2.85 (dt, *J* = 13.0, 6.3 Hz, 1H, C2'-H), 2.35 (ddd, *J* = 13.2, 6.2, 4.2 Hz, 1H, C2'-H), 2.14 (s, 3H, COCH₃) ppm.

¹³C-NMR (201 MHz, DMSO-*d*₆) δ 169.8 (CH₃CO), 158.4 (DMTr), 157.2 (C6), 150.2 (C4), 145.4 (DMTr), 144.0 (-C₁₃H₈NHAc), 141.4 (-C₁₃H₈NHAc), 139.1 (-C₁₃H₈NHAc), 138.4 (-C₁₃H₈NHAc), 136.0 (DMTr), 136.0 (C8), 132.0 (-C₁₃H₈NHAc), 130.1 (DMTr), 130.0 (DMTr), 128.4 (-C₁₃H₈NHAc), 128.2 (DMTr), 128.0 (DMTr), 127.4 (-C₁₃H₈NHAc), 127.1 (-C₁₃H₈NHAc), 127.0 (-C₁₃H₈NHAc), 125.6 (-C₁₃H₈NHAc), 123.2 (-C₁₃H₈NHAc), 120.2 (-C₁₃H₈NHAc), 119.0 (C5), 114.2 (-C₁₃H₈NHAc), 113.5 (DMTr), 113.5 (DMTr), 86.2 (DMTr), 85.9 (C4'), 83.7 (C1'), 71.1 (C3'), 64.5 (C5'), 55.4 (OCH₃), 39.1 (C2'), 36.6 (CH₂) 23.9 (CH₃CO) ppm.

HR-MS (ESI+): m/z berechnet für [C₄₆H₄₃N₆O₇]⁺ ([M+H]⁺): 791,3187, gefunden: 791,3194.

5.4.2.10 Synthese von *N*²-(2-Acetylaminofluoren-1-yl)-3'-O-[*N,N*-diisopropylamino-(2-cyanoethoxy)-phosphinyl]-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin (9)



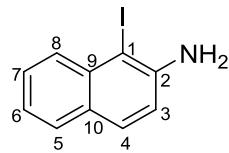
*N*²-(2-Acetylaminofluoren-1-yl)-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin **21** (150 mg, 0,190 mmol) wurde in 2 mL trockenem DCM gelöst und DIAT (39 mg, 0,23 mmol, 1,2 Äq.) wurde zugegeben. Zu der Reaktionslösung wurde (*i*Pr₂N)₂PO(CH₂)₂CN (69 mg, 0,23 mmol, 1,2 Äq.) gegeben und 3 h bei RT gerührt. Nach säulenchromatographischer Aufreinigung (DCM/MeOH 20:1 + 1 %Pyridin) wurde *N*²-(2-Acetylaminofluoren-1-yl)-3'-O-[*N,N*-diisopropylamino-(2-cyanoethoxy)-phosphinyl]-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyguanosin **9** (188 mg, 0,190 mmol, 100 %) als gelber Feststoff erhalten.

³¹P-NMR (162 MHz, CDCl₃) δ 148.99, 148.46 ppm.

HR-MS (ESI-): m/z berechnet für [C₅₅H₅₈N₈O₈P]⁺ ([M+H]⁺): 989,4121, gefunden: 989,4131.

5.4.3 Synthese des *N*²-AAN-dG Phosphoramidits

5.4.3.1 Synthese von 1-Iodo-2-Naphthylamin



2-Aminonaphthalin (500 mg, 3,49 mmol, 1,00 Äq.), Kaliumchlorat (638 mg, 5,20 mmol, 0,33 Äq.) und Kaliumiodid (638 mg, 3,84 mmol, 1,00 Äq.) wurden in 9 mL Methanol und 0,5 mL Wasser gelöst. Unter starkem Rühren wurden 0,32 mL Salzsäure hinzugegeben. Das Reaktionsgemisch wurde unter Lichtausschluss 24 h bei 80 °C gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand wurde mittels Säulenchromatografie (*dryload*,

*i*Hex/EtOAc 50 → 35:1) gereinigt. 1-Iodo-2-Naphthylamin (600 mg, 2,23 mmol 64 %) wurde als schwarzer Feststoff erhalten, der unter Lichtausschluss aufbewahrt wurde.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 7.73 (dt, *J* = 8.5, 0.9 Hz, 1H, C8-H), 7.62 (d, *J* = 7.2 Hz, 1H, C5-H), 7.62 (d, *J* = 8.6 Hz, 1H, C4-H), 7.40 (ddd, *J* = 8.4, 6.8, 1.4 Hz, 1H, C7-H), 7.15 (ddt, *J* = 7.7, 6.8, 0.9 Hz, 1H, C6-H), 7.07 (dd, *J* = 8.8, 1.2 Hz, 1H, C3-H), 5.66 (s, 2H, NH₂) ppm.

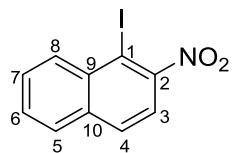
¹³C-NMR (101 MHz, DMSO-*d*₆) δ 148.0 (C2), 135.7 (C9), 129.8 (C4 oder C5), 129.1 (C8), 128.6 (C4 oder C5), 128.3 (C7), 127.8 (C10), 122.3 (C6), 118.0 (C3), 80.7 (C1) ppm.

MS (EI): m/z berechnet für [C₁₀H₉NI] ([M]): 268,9701, gefunden: 268,9703.

IR: $\tilde{\nu}$ = 3396 (w), 3298 (w), 3189 (w), 2964 (w), 2344 (w), 1619 (w), 1500 (w), 1259 (m), 1015 (s), 915 (w), 799 (s), 760 (m), 691 (w) cm⁻¹.

Schmelzpunkt: 63 °C.

5.4.3.2 Synthese von 1-Iodo-2-nitronaphthalin



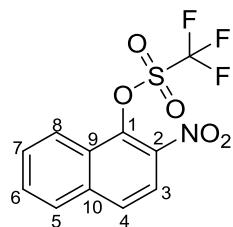
1-Iodo-2-aminoaphthalin (100 mg, 0,373 mmol, 1,0 Äq.) wurde unter Lichtausschluss in 3 mL DCM gelöst und *m*CPBA (343 mg, 1,49 mmol, 4,0 Äq.) wurde hinzugegeben. Das Reaktionsgemisch wurde 18 h bei RT gerührt und mit NaHCO₃-Lösung und H₂O gewaschen. Die organischen Phasen wurden über MgSO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde säulenchromatografisch (*i*Hex/EtOAc 50:1) gereinigt. 1-Iodo-2-nitronaphthalin (20 mg, 0,069 mmol, 18 %) wurde als brauner Feststoff erhalten, der unter Lichtausschluss aufbewahrt wurde.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 8.25 (dd, *J* = 8.4, 0.7 Hz, 1H, C8-H), 8.14 (d, *J* = 8.5, 1H, C4-H), 8.03 (d, *J* = 8.2, 1H, C5-H), 7.82 (dt, *J* = 8.7, 0.7 Hz, 1H, C3-H), 7.77 (m, 1H, C7-H), 7.72 (m, 1H, C6-H) ppm.

¹³C-NMR (101 MHz, DMSO-*d*₆) δ 154.3 (C2), 134.3 (C10), 134.0 (C8), 134.0 (C9), 131.2 (C4), 130.5 (C7), 129.5 (C6), 129.4 (C5), 120.7 (C3), 94.0 (C1) ppm.

MS (EI): m/z berechnet für [C₁₀H₇O₂NI] ([M]): 298,9443, gefunden: 298,9437.

5.4.3.3 Synthese von 2-Nitronaphthalin-1-triflat



2-Nitronaphthol (3,00 g, 15,9 mmol, 1,0 Äq.) wurde in 65 mL DCM gelöst und 6,59 mL NEt₃ (47,6 mmol, 3,0 Äq.) wurden unter Rühren bei RT zugegeben. Die Lösung wurde auf -10°C gekühlt und Trifluormethansulfonsäureanhydrid (4,00 mL, 23,8 mmol, 1,5 Äq.) wurde langsam zugetropft. Nach 30 Minuten wurde die Lösung auf RT gebracht mit DCM verdünnt, zweimal mit NH₄Cl-Lösung, zweimal mit Wasser und einmal mit gesättigter NaCl-Lösung gewaschen. Die organische Phase wurde mit MgSO₄ getrocknet und das Lösungsmittel wurde evaporiert. 2-Nitronaphthalin-1-triflat (5,01 g, 15,6 mmol, 98 %) wurde als brauner Feststoff erhalten.

¹H-NMR (400 MHz, CDCl₃) δ 8,28 (m, 1H, C8-H), 8,08 (d, 1H, ³J=9,01 Hz, C3-H), 8,00 (d, 1H, ³J=9,01 Hz, C4-H), 7,99 (m, 1H, C5-H), 7,79 (m, 2H, C6-H und C7-H) ppm.

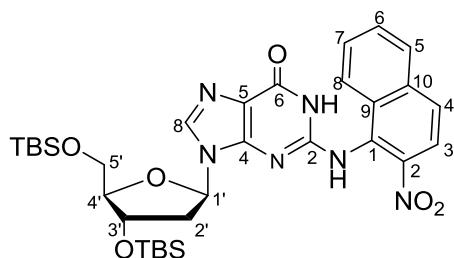
¹³C-NMR (101 MHz, CDCl₃) δ 139,1 (C1), 138,0 (C9), 136,4 (C2), 130,4 (C6), 129,5 (C7), 129,2 (C5), 128,3 (C4), 126,6 (C9), 123,1 (C8), 120,7 (C3), 116,9 (CF₃) ppm.

MS (EI): m/z berechnet für [C₁₁H₆O₅NF₃S] ([M]): 320,9919, gefunden: 320,9141.

IR: $\tilde{\nu}$ = 1529,3 (m), 1424,4 (m), 1344,1 (m), 1204,8 (s), 1134,0 (s), 1052,8 (m), 895,6 (m), 808,0 (s), 798,6 (s), 751,2 (s) cm⁻¹.

Schmelzbereich: 65 – 69 °C.

5.4.3.4 Synthese von *N*²-(2-Nitronaphthalin-1-yl)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyguanosin (**11**)



3',5'-Bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyguanosin **17** (247 mg, 0,500 mmol, 1,0 Äq.), 2-Nitronaphthalin-1-triflat (241 mg, 0,750 mmol, 1,5 Äq.), Pd₂(dba)₃ (46 mg, 0,050 mmol, 0,1 Äq.) und Xantphos (87 mg, 0,15 mmol, 0,3 Äq.) wurden unter Argonatmosphäre vorgelegt, dreimal entgast und mit 1,5 mL trockenem Dioxan versetzt. Das Reaktionsgemisch wurde 10 min bei RT gerührt und Cs₂CO₃ (326 mg, 1,00 mmol, 2,0 Äq.) wurde hinzugegeben. Die Suspension wurde erneut dreimal entgast und 18 h bei 100 °C gerührt. Das Lösemittel wurde *in vacuo* entfernt und der Rückstand wurde säulenchromatografisch (DCM/MeOH 100:1) gereinigt. *N*²-(2-Nitronaphthalin-1-yl)-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyguanosin **11** (352 mg, 0,487 mmol, 97 %) wurde als gelbbrauner Feststoff erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 11.27 (s, 1H, NH), 9.41 (s, 1H, NH), 8.31 (d, J= 8.1 Hz, 1H, -C₁₀H₆NO₂), 8.17 (d, 1H, J= 7.8 Hz, -C₁₀H₆NO₂), 8.10 (d, J= 9.0 Hz, 1H, -C₁₀H₆NO₂), 8.03 (d, J= 9.0 Hz, 1H, -C₁₀H₆NO₂), 7.96 (d, J= 6.4 Hz, 1H, C8 H), 7.85 – 7.75 (m, 2H, -C₁₀H₆NO₂), 5.84 (dt, J= 19.0 Hz, 3J= 9.5 Hz, 1H, C1'-H), 4.18 (dd, J= 11.4 Hz, J= 9.0 Hz, 1H, C3'-H), 3.77 – 3.62 (m, 1H, C4'-H), 3.41 (s, 2H, C5'-H) 2.40 (ddd, J= 10.1 Hz, J= 5.8 Hz, J= 3.7 Hz, 1H, C2'-H), 2.07 – 1.90 (m, 1H, C2'-H), 0.88 (s, 9H, C(CH₃)₃), 0.86 (s, 9H, C(CH₃)₃), 0.07 (s, 6H, SiCH³), 0.00 (s, 6H, SiCH₃) ppm.

¹³C-NMR (101 MHz, DMSO-*d*₆) δ 156.9 (C6), 150.6 (C2), 149.4 (C4), 142.3 (-C₁₀H₆NO₂), 135.7 (C8), 129.9 (-C₁₀H₆NO₂), 129.6 (-C₁₀H₆NO₂), 128.9 (-C₁₀H₆NO₂), 128.6 (-C₁₀H₆NO₂), 127.4 (-C₁₀H₆NO₂), 125.3 (-C₁₀H₆NO₂), 121.1 (C5), 87.7 (C4'), 83.4 (C1'), 72.8 (C3'), 63.1 (C5'), 39.4 (C2'), 26.2 (C(CH₃)₃), 26.1 (C(CH₃)₃), 18.3 (C(CH₃)₃), 18.1 (C(CH₃)₃), -4.4 (SiCH₃), -4.5 (SiCH₃), -5.1 (SiCH₃) ppm.

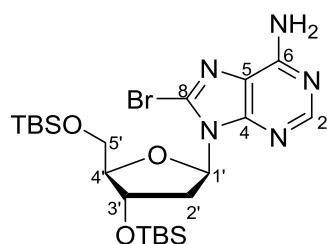
HR-MS (ESI+): m/z berechnet für [C₃₂H₄₆N₆O₆Si₂Na]⁺ ([M+Na]⁺): 689,2919; gefunden: 689,2909.

IR: $\tilde{\nu}$ = 2951 (w), 2927 (w), 2856 (w), 2365 (w), 1702 (s), 1596 (m), 1578 (m), 1469 (m), 1403 (m), 1360 (m), 1253 (m), 1138 (m), 1093 (m), 1027 (m), 937 (m), 832 (s), 817 (s), 778 (s), 665 (w) cm^{-1} .

Die Synthese des N^2 -AAN-dG Phosphoramidits wurde letztendlich nach Bonala et al. durchgeführt.^[64]

5.4.4 Synthese des C8-AAF-dA Phosphoramidits

5.4.4.1 Synthese von C8-Brom-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin (23)



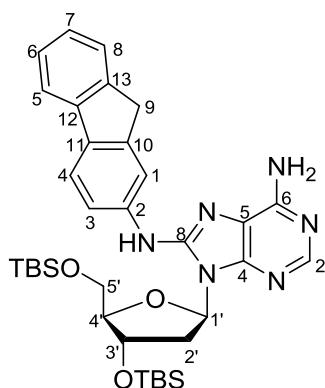
C8-Brom-2'-desoxyadenosin **22** (1,05 g, 3,18 mmol, 1,0 Äq.) wurde in 10 mL Pyridin suspendiert. Es wurden TBS-Chlorid (1,40 g, 9,30 mmol, 3,0 Äq) und Imidazol (0,64 g, 9,4 mmol, 3,0 Äq.) zugegeben und die Reaktionsmischung über 6 Stunden bei RT gerührt. Der Rückstand wurde säulenchromatografisch gereinigt (DCM/MeOH 200:1) und C8-Brom-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin **23** (1,45 g, 2,61 mmol, 82 %) wurde als weißer Feststoff erhalten.

¹H-NMR (599 MHz, CDCl₃) δ 8.24 (s, 1H, C2-H), 6.34 (t, J = 6.94, 1H, C1'-H), 5.80 (s, 2H, NH₂), 4.92 – 4.81 (m, 1H, C3'-H), 3.95 (dd, J = 4.7, 3.6 Hz, 1H, C4'-H), 3.91 (dd, J = 10.8, 6.5 Hz, 1H, C5'-H), 3.67 (m, 2H, C2'-H und C5'-H), 2.23 (ddd, J = 13.1, 7.0, 4.2 Hz, 1H, C2'-H), 0.93 (s, 9H, C(CH₃)₃), 0.82 (s, 9H, C(CH₃)₃), 0.14 (s, 6H, SiCH₃), -0.01 (s, 3H, SiCH₃), -0.05 (s, 3H, SiCH₃) ppm.

¹³C-NMR (151 MHz, CDCl₃) δ 154.2 (C6), 152.5 (C2), 150.9 (C4), 128.2 (C8), 120.5 (C5), 87.7 (C4'), 86.3 (C1'), 72.3 (C3'), 62.6 (C5'), 36.7 (C2'), 25.8 (C(CH₃)₃), 18.3 (C(CH₃)₃), 18.2 (C(CH₃)₃), -4.7 (SiCH₃), -4.7 (SiCH₃), -5.4 (SiCH₃), -5.5 (SiCH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₂₂H₄₁BrN₅O₃Si₂]⁺ ([M+H]⁺): 558,1926, gefunden: 558,1939.

5.4.4.2 Synthese von C8-Aminofluorenyl-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin (24)



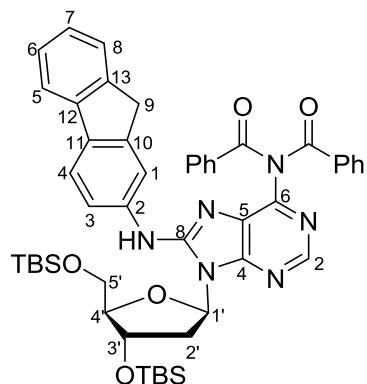
Pd(OAc)₂ (123 mg, 0,255 mmol, 0,15 Äq.) und BINAP (317 mg, 0,510 mmol, 0,30 Äq.) wurden in 16 mL trockenem Toluol gelöst und 30 Minuten bei RT gerührt. Es wurden C8-Bromo-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin **23** (1,000 g, 1,699 mmol, 1,00 Äq.), 2-Aminofluoren (462 mg, 2,54 mmol, 1,5 Äq.) und Cs₂CO₃ (1,384 g, 4,248 mmol, 2,5 mmol) zugegeben und über 16 h bei 100 °C gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand *via* Säulenchromatografie (DCM/MeOH 200 → 100:1) gereinigt. C8-Aminofluorenyl-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin **24** (578 mg, 0,878 mmol, 52 %) wurde als gelber Schaum erhalten.

¹H-NMR (599 MHz, CDCl₃) δ 8.21 (s, 1H, C2-H), 8.09 (s, 1H, NH), 7.87 (d, *J* = 2.1 Hz, 1H, -C₁₃H₉), 7.79 (dd, *J* = 8.0, 2.2 Hz, 1H, -C₁₃H₉), 7.63 (d, *J* = 7.3 Hz, 1H, -C₁₃H₉), 7.46 (m, 3H, -C₁₃H₉), 7.23 (dt, *J* = 7.4, 4.2 Hz, 1H, -C₁₃H₉), 6.45 (m, 1H, C1'-H), 5.32 (s, 2H, CH₂), 4.58 (tt, *J* = 6.0, 3.0 Hz, 1H, C3'-H), 4.08 (q, *J* = 3.2 Hz, 1H, C4'-H), 3.99 (m, 1H, C5'-H), 3.87 (dt, *J* = 11.5, 3.7 Hz, 1H, C5'-H), 2.92 (m, 1H, C2'-H), 2.37 (m, 1H, C2'-H), 0.93 (s, 9H, C(CH₃)₃), 0.83 (s, 9H, C(CH₃)₃), 0.13 (s, 3H, SiCH₃), 0.12 (s, 3H, SiCH₃), 0.07 (s, 3H, SiCH₃), 0.04 (s, 3H, SiCH₃) ppm.

¹³C-NMR (151 MHz, CDCl₃) δ 152.4 (C6), 150.4 (C2), 149.3 (C4), 147.8 (C8), 144.8 (-C₁₃H₉), 140.4 (-C₁₃H₉), 138.4 (-C₁₃H₉), 135.2 (-C₁₃H₉), 134.8 (-C₁₃H₉), 134.3 (-C₁₃H₉), 128.2 (-C₁₃H₉), 124.9 (-C₁₃H₉), 124.4 (-C₁₃H₉), 120.8 (-C₁₃H₉), 119.8 (-C₁₃H₉), 117.3 (C5), 115.7 (-C₁₃H₉), 87.9 (C4'), 84.9 (C1'), 69.1(C3'), 62.8 (C5'), 39.9 (C2'), 26.0 (C(CH₃)₃), 25.8 (C(CH₃)₃), 18.4 (C(CH₃)₃), 18.0 (C(CH₃)₃), -4.6 (SiCH₃), -4.7 (SiCH₃), -5.2 (SiCH₃), -5.3 (SiCH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₃₅H₅₁N₆O₃Si₂]⁺ ([M+H]⁺): 659,3556, gefunden: 659,3560.

5.4.4.3 Synthese von C8-Aminofluorenyl-N⁶-dibenzoyl-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin (25)



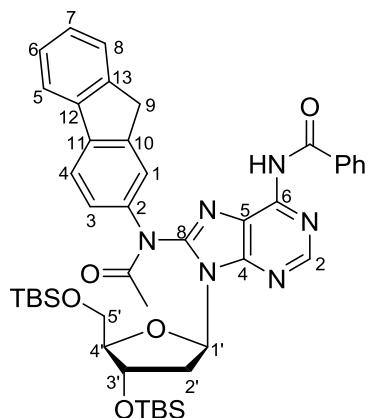
C8-Aminofluorenyl-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin **24** (100 mg, 0,150 mmol, 1,0 Äq.) wurde in 1 mL Pyridin gelöst und Benzoylchlorid (0,09 mL, 0,8 mmol) wurde zugegeben. Das Reaktionsgemisch wurde 18 h gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand wurde säulenchromatografisch (*i*Hex/EtOAc 8 → 5:1) gereinigt. C8-Aminofluorenyl-N⁶-dibenzoyl-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin **25** (123 mg, 0,142 mmol, 93 %) wurde als gelber Schaum erhalten.

¹H-NMR (599 MHz, CDCl₃) δ 8.63 (s, 1H, C2-H), 7.77 (dt, *J* = 7.6, 0.9 Hz, 1H, -C₁₃H₉), 7.73 (d, *J* = 8.2 Hz, 1H, -C₁₃H₉), 7.67 (d, *J* = 7.7 Hz, 4H, -C₆H₅), 7.60 – 7.53 (m, 2H, -C₁₃H₉), 7.45 – 7.40 (m, 2H, -C₆H₅), 7.40 – 7.37 (m, 1H, -C₁₃H₉), 7.33 (td, *J* = 7.5, 1.2 Hz, 1H, -C₁₃H₉), 7.31 – 7.19 (m, 5H, -C₁₃H₉ + -C₆H₅), 6.25 (dd, *J* = 7.9, 6.2 Hz, 1H, C1'-H), 4.66 (s, 1H, C3'-H), 4.01 – 3.94 (m, 2H, C4'-H + C5'-H), 3.85 (m, 2H, CH₂), 3.78 (dd, *J* = 9.4, 4.0 Hz, 1H, C5'-H), 3.18 (s, 2H, C2'-H), 0.90 (s, 9H, C(CH₃)₃), 0.83 (s, 9H, C(CH₃)₃), 0.07 (s, 3H, SiCH₃), 0.05 (d, *J* = 7.9 Hz, 3H, SiCH₃), 0.02 (s, 3H, SiCH₃), 0.01 (s, 1H, SiCH₃) ppm.

¹³C-NMR (151 MHz, CDCl₃) δ 172.1 (CO), 170.7 (CO), 152.8 (C6), 151.9 (C2), 150.9 (C4), 150.3 (C8), 144.5 (-C₁₃H₉), 143.5 (-C₁₃H₉), 141.3 (-C₁₃H₉), 140.6 (-C₁₃H₉), 138.7, 134.2 (-C₆H₅), 132.6 (-C₆H₅), 131.4 (-C₆H₅), 129.3 (-C₆H₅), 128.8 (-C₁₃H₉), 128.5 (-C₆H₅), 128.2 (-C₆H₅), 127.2 (-C₁₃H₉), 127.0 (-C₁₃H₉), 126.8, 125.5, 125.1 (-C₁₃H₉), 123.3, 120.6 (-C₁₃H₉), 120.1 (-C₁₃H₉), 88.2 (C4'), 85.2 (C1'), 72.5 (C3'), 63.1 (C5'), 36.9 (CH₂), 36.6 (C2'), 25.9 (C(CH₃)₃), 25.8 (C(CH₃)₃), 18.4 (C(CH₃)₃), 17.9 (C(CH₃)₃), -4.8 (SiCH₃), -4.9 (SiCH₃), -5.3 (SiCH₃), -5.3 (SiCH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₄₉H₅₉N₆O₅Si₂]⁺ ([M+H]⁺): 867,4080, gefunden: 867,4099.

5.4.4.4 Synthese von C8-Acetylaminofluorenyl-N⁶-benzoyl-3',5'-bis-O-(tert-butyl-dimethylsilyl)-2'-desoxyadenosin (26)



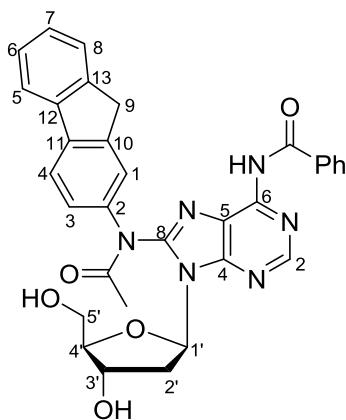
C8-Aminofluorenyl-N⁶-dibenzoyl-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyadenosin **25** (123 mg, 0,142 mmol, 1,0 Äq.) und 4-DMAP (52 mg, 0,43 mmol, 3,0 Äq.) wurden in 1 mL trockenem Pyridin gelöst und 0,04 mL Ac₂O (0,4 mmol) wurden zugetropft. Das Reaktionsgemisch wurde 72 h bei RT gerührt. Der Rückstand wurde via Säulenchromatografie (iHex/EtOAc 3:1) gereinigt und C8-Acetylaminofluorenyl-N⁶-benzoyl-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyadenosin **26** (58 mg, 0,072 mmol, 51 %) wurde als gelber Schaum erhalten.

¹H-NMR (400 MHz, CDCl₃) δ 8.93 (s, 1H, NH), 8.61 (s, 1H, C2-H), 8.04 (m, 1H, -C₁₃H₉), 7.66 (m, 2H, -C₁₃H₉), 7.58 (d, J = 7.7 Hz, 2H, -C₆H₅), 7.46 (d, J = 7.4 Hz, 1H, -C₁₃H₉), 7.39 (bs, 1H, -C₁₃H₉), 7.32 (m, 2H, -C₁₃H₉), 7.23 (m, 4H, -C₁₃H₉, -C₆H₅), 6.26 (t, J = 6.9 Hz, 1H, C1'-H), 4.69 (s, 1H, C3'-H), 4.05 – 3.90 (m, 2H, C4'-H, C5'-H), 3.87 – 3.66 (m, 3H, CH₂, C5'-H), 3.48 – 3.19 (m, 1H, C2'-H), 2.43 (s, 3H, CH₃CO), 0.82 (s, 9H, C(CH₃)₃), 0.78 (s, 9H, C(CH₃)₃), 0.04 (s, 3H, SiCH₃), 0.00 (s, 3H, SiCH₃), -0.02 (s, 6H, SiCH₃) ppm.

¹³C-NMR (151 MHz, CDCl₃) δ 170.8 (COPh), 170.1 (COCH₃), 152.1 (C2), 150.8 (C6), 148.7 (C4), 148.67, 144.6 (-C₁₃H₉), 143.4 (-C₁₃H₉), 141.4 (-C₁₃H₉), 140.5 (-C₁₃H₉), 131.8 (-C₁₃H₉), 130.0 (-C₁₃H₉), 129.83, 129.1 (-C₆H₅), 128.4 (-C₆H₅), 128.3 (-C₁₃H₉), 127.2 (-C₁₃H₉), 126.9 (-C₁₃H₉), 125.0 (-C₁₃H₉), 120.6 (-C₁₃H₉), 120.3 (C5), 120.1 (-C₁₃H₉), 88.2 (C4'), 85.3 (C1'), 72.6 (C3'), 63.0 (C5'), 36.9 (CH₂), 36.4 (C2'), 25.9 (C(CH₃)₃), 25.7 (C(CH₃)₃), 25.5 (CH₃CO) 18.3 (C(CH₃)₃), 17.9 (C(CH₃)₃), -4.8 (SiCH₃), -4.8 (SiCH₃) -5.3 (SiCH₃), -5.4 (SiCH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₄₄H₅₇N₆O₅Si₂]⁺ ([M+H]⁺): 805,3924, gefunden: 805,3926.

5.4.4.5 Synthese von C8-Acetylaminofluorenyl-N⁶-benzoyl-2'-desoxyadenosin (27)



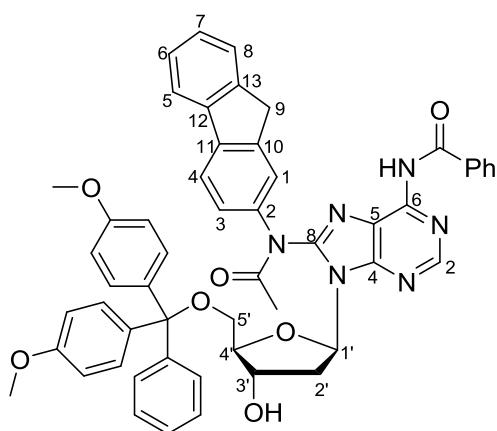
C8-Acetylaminofluorenyl-*N*⁶-benzoyl-3',5'-bis-O-(*tert*-butyl-dimethylsilyl)-2'-desoxyadenosin **26** (58 mg, 0,072 mmol, 1,0 Äq.) wurde in 1 mL THF gelöst und 0,22 mL TBAF (1 M in THF, 3,0 Äq.) und 0,03 mL Essigsäure (0,4 mmol, 6,0 Äq.) wurden zugetropft. Das Reaktionsgemisch wurde über 48 h bei RT gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand wurde säulenchromatografisch (DCM/MeOH 80 → 30:1) aufgereinigt. C8-Acetylaminofluorenyl-*N*⁶-benzoyl-2'-desoxyadenosin **27** (22 mg, 0,038 mmol, 53 %) wurde als gelber Schaum erhalten.

¹H-NMR (800 MHz, DMSO-*d*₆) δ 10.70 (s, 1H, NH), 8.67 (s, 1H, C2-H), 7.89 (d, *J* = 8.1 Hz, 2H, -C₁₃H₉), 7.69 – 7.49 (m, 4H, -C₆H₅ und -C₁₃H₉), 7.49 – 7.44 (m, 1H, -C₁₃H₉), 7.43 -7.25 (m, 5H, -C₆H₅ und -C₁₃H₉), 6.38 (s, 1H, C1'-H), 5.35 (s, 2H, OH), 4.54 (m, 1H, C3'-H), 4.01 – 3.84 (m, 3H, C4'-H und CH₂), 3.75 (d, *J* = 11.5 Hz, 1H, C5'-H), 3.59 (m, 1H, C5'-H), 2.53 (m, 1H, C2'-H), 2.48 (m, 1H, C2'-H), 2.18 (s, 3H, CH₃CO) ppm.

¹³C-NMR (201 MHz, DMSO-*d*₆) δ 170.6 (COPh), 169.3 (COCH₃), 152.2 (C2), 151.4 (C4), 149.9 (C6), 143.9 (-C₁₃H₉), 140.8 (-C₁₃H₉), 132.2 (-C₁₃H₉), 129.2 (-C₁₃H₉), 128.9 (-C₆H₅), 127.7 (-C₆H₅), 127.4 (-C₁₃H₉), 125.7 (-C₁₃H₉), 122.4 (C5), 121.2 (-C₁₃H₉), 120.8 (-C₁₃H₉), 89.0 (C4'), 85.8 (C1'), 71.4 (C3'), 62.5 (C5'), 40.2 (C2'), 36.9 (CH₂), 24.8 (CH₃CO) ppm.

HR-MS (ESI+): m/z berechnet für [C₃₂H₂₉N₆O₅]⁺ ([M+H]⁺): 577,2199, gefunden: 577,2194.

5.4.4.6 Synthese von C8-Acetylaminofluorenyl-N⁶-benzoyl-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyadenosin (28)

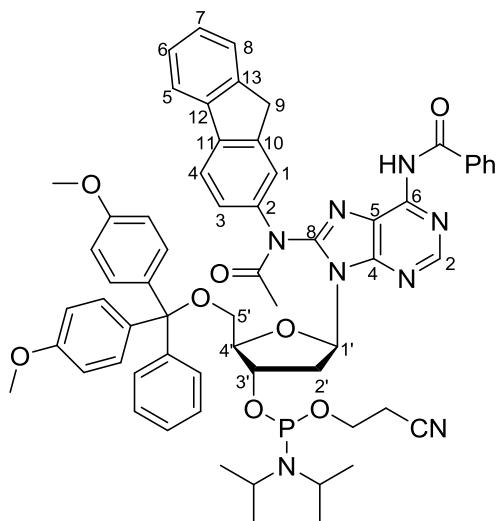


C8-Acetylaminofluorenyl-N⁶-benzoyl-2'-desoxyadenosin **27** (100 mg, 0,170 mmol, 1,0 Äq.) wurde in 1 mL trockenem Pyridin (Molsieb) gelöst und über 1,5 h bei RT gerührt. Es wurden DMTr-Chlorid (81 mg, 0,21 mmol, 1,2 Äq.) und trockenes Triethylamin (0,03 mL, 0,2 mmol, 1,2 Äq.) zugegeben und das Reaktionsgemisch wurde 18 h bei RT gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand wurde *via* Säulenchromatografie gereinigt. C8-Acetylaminofluorenyl-N⁶-benzoyl-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyadenosin **28** (92 mg, 0,11 mmol, 62 %) wurde als gelber Schaum erhalten. Unreagiertes Edukt **27** (19 mg, 0,033 mmol, 19 %) konnte zurückgewonnen werden.

¹H-NMR (400 MHz, CDCl₃) δ 8.30 (s, 1H, NH), 8.26 (s, 1H, C2-H), 7.65 – 7.05 (m, 21H, DMT, -C₆H₅, -C₁₃H₉), 6.70 – 6.55 (m, 4H, DMT), 6.29 (t, J = 6.7 Hz, 1H, C1'-H), 4.81 (s, 1H, C3'-H), 4.17 (td, J = 6.1, 3.6 Hz, 1H, C4'-H), 3.78 (bs, 1H, C5'-H), 3.72 (bs, 1H, C5'-H), 3.67 (s, 3H, OCH₃), 3.40 (m, 2H, C2'-H) 3.65 (s, 3H, OCH₃), 2.45 (s, 3H, CH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₅₃H₄₇N₆O₇]⁺ ([M+H]⁺): 879,3501, gefunden: 879,3503.

5.4.4.7 Synthese von C8-Acetylaminofluorenyl-3'-O-[N,N-diisopropylamino-(2-cyanoethoxy)-phosphinyl]-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyadenosin (6)

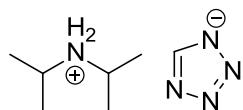


C8-Acetylaminofluorenyl-*N*⁶-benzoyl-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyadenosin **28** (50 mg, 0,057 mmol, 1,0 Äq.) wurde in 2,5 mL trockenem DCM gelöst und DIAT (21 mg, 0,068 mmol, 1,2 Äq.) wurde zugegeben. Das Reaktionsgemisch wurde drei Mal entgast, (*i*Pr₂N)₂PO(CH₂)₂CN (0,02 mL, 0,07 mmol, 1,4 Äq.) wurde zugegeben und es wurde 2,5 h gerührt. Der Rückstand wurde säulenchromatografisch (DCM/MeOH 40:1 + 1 % Pyridin) aufgereinigt und C8-Acetylaminofluorenyl-3'-O-[N,N-diisopropylamino-(2-cyanoethoxy)-phosphinyl]-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyadenosin **6** (61 mg, 0,056 mmol, 99 %) wurde als gelber Schaum erhalten.

³¹P-NMR (162 MHz, CDCl₃) δ 149.34 ppm.

HR-MS (ESI+): m/z berechnet für [C₆₂H₆₄N₈O₈P]⁺ ([M+H]⁺): 1079,4579, gefunden: 1079,4581.

5.4.5 Synthese von Di-*iso*-Propylammoniumtetrazolat (DIAT)



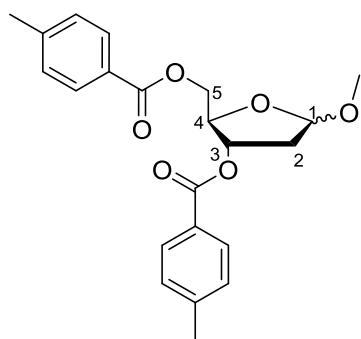
Tetrazol (0,45 M in Acetonitril, 44,4 mL, 20,0 mmol, 1,0 Äq.) und Di-*iso*-propylamin (4,5 mL, 32 mmol, 1,6 Äq.) wurden 20 min bei RT gerührt. Das Lösungsmittel wurde *in vacuo* entfernt. DIAT (3,42 g, 20,0 mmol, 100 %) wurde als weißer Feststoff erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 8.19 (s, 1H, ar. H), 3.40 (hept, 2H, *J* = 6.3 Hz, -CH(CH₃)₂), 1.25 (d, 12H, *J* = 6.3 Hz, -CH(CH₃)₂) ppm.

5.5 Synthese von 5-Methyl-2'-desoxycttidin (**m⁵dC**)

5.5.1 Synthese von Hoffer's Chlorozucker

5.5.1.1 *Synthese von 3,5-Di-p-toluoyl-methyl-2-desoxy-ribofuranosid (31)*

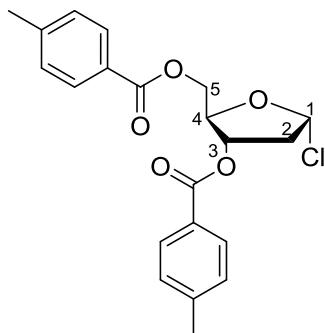


2-Desoxyribose (10,70 g, 79,77 mmol, 1,00 Äq.) wurde in 200 mL trockenem Methanol gelöst und es wurden 0,36 mL Acetylchlorid (5,0 mmol, 0,06 Äq.) zugegeben. Die Suspension wurde 30 min bei RT gerührt, und 3,57 g Ag₂CO₃ wurden zugegeben um überschüssige Salzsäure abzupuffern. Ag₂CO₃ wurde abfiltriert und das Filtrat wurde *in vacuo* eingeengt. Das braune Öl wurde in 65 mL Pyridin gelöst und 21,10 mL Toluoylchlorid (159,5 mmol, 2,00 Äq.) wurde bei 0 °C zugegeben. Die Lösung wurde 20 h bei 45 °C gerührt. Es wurde Wasser zugegeben und mit EtOAc extrahiert. Die organische Phase wurde mit 2 M HCl und gesättigter wässriger NaHCO₃-Lösung gewaschen und über Mg₂SO₄ getrocknet. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand wurde via

Säulenchromatografie gereinigt (*i*Hex/EtOAc 8:1). 3,5-Di-*p*-toluoyl-methyl-2-desoxy-ribofuranosid **31** (28,22 g, 73,41 mmol, 92 %) wurde als farbloses Öl erhalten.

HR-MS (ESI+): m/z berechnet für $[C_{22}H_{28}O_6]^+$ ($[M+NH_4]^+$): 402.1911, gefunden: 402.1917.

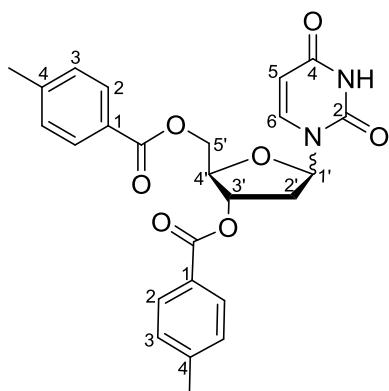
5.5.1.2 Synthese von α -3,5-Di-*p*-toluoyl-methyl-2-desoxy-ribofuranosylchlorid (32)



3,5-Di-*p*-toluoyl-methyl-2-desoxy-ribofuranosid **31** (28,22 g, 73,41 mmol, 1,0 Äq.) wurde in 50 mL Essigsäure gelöst. Die Lösung wurde auf 0 °C gekühlt und 3,5 mL Wasser (0,20 mol, 2,7 Äq.) und 13 mL Acetylchlorid (0,18 mol, 2,5 Äq.) wurden zugefügt. Es wurde 15 min gerührt. Der weiße Niederschlag wurde abfiltriert und mit Diethylether gewaschen. α -3,5-Di-*p*-toluoyl-methyl-2-desoxy-ribofuranosylchlorid **32** (19,61 g, 50,44 mmol, 69 %) wurde als weißer Feststoff erhalten.

¹H-NMR (400 MHz, CDCl₃) δ 7.97 – 7.89 (m, 2H, Tol), 7.85 – 7.78 (m, 2H, Tol), 7.22 – 7.18 (m, 2H, Tol), 7.16 (m, 2H, Tol), 6.40 (d, ³J = 5.1 Hz, 1H, C1'-H), 5.49 (ddd, ³J = 7.4 Hz, ³J = 2.9 Hz, ³J = 1.1 Hz, 1H, C3'-H), 4.78 (dt, ³J = 4.1, ³J = 3.0 Hz, 1H, C4'-H), 4.61 (dd, ²J = 12.1, ³J = 3.2 Hz, 1H), 4.52 (dd, ²J = 12.1, ³J = 4.3 Hz, 1H), 2.80 (ddd, ²J = 15.1 Hz, ³J = 7.4 Hz, ³J = 5.2 Hz, 1H, C2'-H), 2.67 (dd, ²J = 15.1 Hz, ³J = 1.1 Hz, 1H, C2'-H), 2.35 (s, 3H, CH₃), 2.34 (s, 3H, CH₃) ppm.

5.5.2 Synthese von 3',5'-bis-O-p-Toluoyl-2'-desoxyuridin



Zu einer Lösung aus Uracil (200 mg, 1,78 mmol, 1,0 Äq.) in trockenem MeCN (12 mL) wurde Bis(trimethylsilyl)acetamid (1,3 mL, 5,4 mmol, 3,0 Äq.) gegeben und die farblose Lösung wurde 2 h bei 80 °C refluxiert. Nach Entfernung des Lösungsmittels *in vacuo* (50 °C, 100 mbar) wurde der Rückstand drei Mal mit trockenem CHCl₃ coevaporiert. Es wurden trockenes CHCl₃ (6 mL) und Hoffer's Chlorozucker (763 mg, 1,96 mmol, 1,1 Äq.) zugegeben. Die braune Lösung wurde 18 h bei RT gerührt. Nach säulenchromatographischer Reinigung (iHex/EtOAc 5 → 1: 1) wurde ein Gemisch aus α- und β-3',5'-Bis-O-p-Toluoyl-2'-desoxyuridin (742 mg, 1,60 mmol, 90 %, α/β 0,7: 1) als weißer Feststoff erhalten.

β-3',5'-Bis-O-p-Toluoyl-2'-desoxyuridin:

¹H-NMR (800 MHz, CDCl₃) δ 8.24 (s, 1H, NH), 7.97 (d, *J* = 1.8 Hz, 2H, Tol-H2), 7.94 – 7.91 (m, 2H, Tol-H2), 7.55 (d, *J* = 8.2 Hz, 1H, C6-H), 7.31 – 7.29 (m, 4H, Tol-H3), 6.42 (dd, *J* = 8.4, 5.6 Hz, 1H, C1'-H), 5.64 – 5.63 (m, 1H, C3'-H), 5.61 (dd, *J* = 8.2, 2.2 Hz, 1H, C5-H), 4.75 (dd, *J* = 12.2, 3.1 Hz, 1H, C5'-H), 4.70 (dd, *J* = 12.2, 3.5 Hz, 1H, C5'-H), 4.57 (d, *J* = 3.0 Hz, 1H, C4'-H), 2.77 (ddd, *J* = 14.4, 5.6, 2.0 Hz, 1H, C2'-H), 2.46 (s, 3H, CH₃), 2.45 (s, 3H, CH₃), 2.32 (ddd, *J* = 14.7, 8.4, 6.6 Hz, 1H, C2'-H) ppm.

¹³C-NMR (201 MHz, CDCl₃) δ 165.0 (CO(Tol)), 165.0 (CO(Tol)), 161.3 (C4), 148.9 (C2), 143.6 (Tol-C4), 143.6 (Tol-C4), 137.8 (C6), 128.8 (Tol-C2), 128.5 (Tol-C2), 128.4 (Tol-C3), 128.3 (Tol-C3), 125.5 (Tol-C1), 125.2 (Tol-C1), 101.9 (C5), 84.4 (C1'), 82.0 (C4'), 73.6 (C3'), 63.0 (C5'), 37.3 (C2'), 20.7 (CH₃) ppm.

α-3',5'-Bis-O-p-Toluoyl-2'-desoxyuridin:

¹H-NMR (800 MHz, CDCl₃) δ 8.32 (s, 1H, NH), 7.97 – 7.94 (m, 2H, Tol-H2), 7.83 – 7.80 (m, 2H, Tol-H2), 7.61 (d, *J* = 8.2 Hz, 1H, C6-H), 7.32 – 7.28 (m, 2H, Tol-H3), 7.28 – 7.25 (m, 2H,

Tol-H3), 6.33 (dd, $J = 7.0, 1.6$ Hz, 1H, C1'-H), 5.74 (dd, $J = 8.1, 2.1$ Hz, 1H, C5-H), 5.63 (d, $J = 6.4$ Hz, 1H, C3'-H), 4.90 – 4.88 (m, 1H, C4'-H), 4.58 (dd, $J = 12.0, 4.2$ Hz, 1H, C5'-H), 4.55 (dd, $J = 12.0, 4.2$ Hz, 1H, C5'-H), 2.98 (dt, $J = 15.6, 6.7$ Hz, 1H, C2'-H), 2.58 – 2.55 (m, 1H, C2'-H), 2.46 (s, 3H, CH₃), 2.44 (s, 3H, CH₃) ppm.

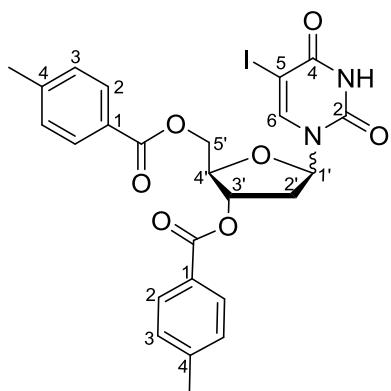
¹³C-NMR (201 MHz, CDCl₃) δ 165.1 (CO(Tol)), 164.6 (CO(Tol)), 161.8 (C4), 148.9 (C2), 143.9 (Tol-C4), 143.4 (Tol-C4), 138.4 (C6), 128.7 (Tol-C2), 128.6 (Tol-C2), 128.4 (Tol-C3), 128.4 (Tol-C3), 125.5 (Tol-C1), 124.9 (Tol-C1), 100.6 (C5), 86.93 (C1'), 84.6 (C4'), 73.7 (C3'), 63.0 (C5'), 38.0 (C2'), 20.7 (CH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₂₅H₂₃N₂O₇]⁺ ([M-H]⁺): 463,1511, gefunden: 463,1504.

IR: $\tilde{\nu} = 3225,7$ (m); 2961,7 (m); 2874,3 (w); 1590,9 (s); 1507,1 (m); 1455,1 (m); 1428,5 (s); 1408,5 (m); 1300,4 (m); 1257,3 (s); 1176,5 (m); 1093,3 (s); 1047,6 (s); 1019,4 (s); 1005,4 (s); 949,8 (m); 873,5 (m); 779,6 (s); 752,1 (s); 713,7 (s); 690,0 (s); 671,3 (m) cm⁻¹.

Schmelzbereich: 189 – 190 °C.

5.5.3 Synthese von 3',5'-bis-O-p-Toluoyl-5-iod-2'-desoxyuridin



Zu einer Lösung aus α - und β -3',5'-Bis-O-p-Toluoyl-2'-desoxyuridin (50 mg, 0,11 mmol, 1,0 Äq.) und LiI (17 mg, 0,13 mmol, 1,2 Äq.) in trockenem MeCN (1 mL) wurde CAN (118 mg, 0,215 mmol, 2,0 Äq.) gegeben und 30 min bei 80 °C refluxiert. Anschließend wurde die Reaktionslösung mit Brine (2 × 5 mL) gewaschen, mit DCM (4 × 3 mL) extrahiert und das Lösungsmittel *in vacuo* entfernt. Nach säulenchromatographischer Reinigung (*t*Hex/EtOAc 4 → 2:1) wurde ein Gemisch aus α - und β -3',5'-Bis-O-p-Toluoyl-5-iod-2'-desoxyuridin (57 mg, 97 µmol, 90 %, α/β 2:1) als weißer Feststoff erhalten.

β -3',5'-Bis-O-pToluoyl-5-iod-2'-desoxyuridin

$^1\text{H-NMR}$ (599 MHz, CDCl_3) δ 8.68 (s, 1H, NH), 7.96 (s, 1H, C6-H), 7.95 – 7.89 (m, 4H, Tol-H2), 7.30 - 7.24 (m, 4H, Tol-H3), 6.36 (dd, J = 8.6, 5.5 Hz, 1H, C1'-H), 5.60 (d, J = 6.3 Hz, 1H, C3'-H), 4.76 – 4.69 (m, 2H, C5'-H), 4.56 (q, J = 2.9 Hz, 1H, C4'-H), 2.76 (dd, J = 14.3, 5.6 Hz, 1H, C2'-H), 2.42 (s, 3H, CH_3), 2.41 (s, 3H, CH_3), 2.27 (ddd, J = 14.7, 8.7, 6.5 Hz, 1H, C2'-H) ppm.

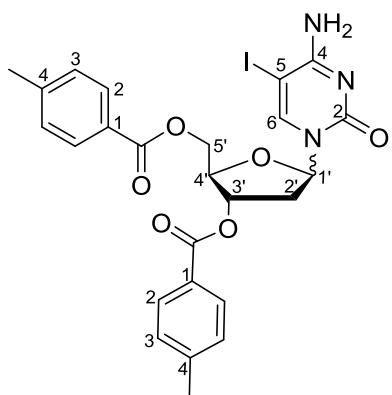
$^{13}\text{C-NMR}$ (151 MHz, CDCl_3) δ 166.1 (CO(Tol)), 166.0 (CO(Tol)), 159.5 (C4), 149.6 (C2), 144.6 (Tol-C4), 144.6 (Tol-C4), 143.6 (C6), 129.8 (Tol-C2), 129.7 (Tol-C2), 129.6 (Tol-C3), 129.3 (Tol-C3), 126.4 (Tol-C1), 126.2 (Tol-C1), 85.8 (C1'), 83.4 (C4'), 74.9 (C3'), 68.8 (C5), 64.1 (C5'), 38.7 (C2'), 21.7 (CH_3) ppm.

HR-MS (ESI+): m/z berechnet für $[\text{C}_{25}\text{H}_{24}\text{IN}_2\text{O}_7]^+$ ($[\text{M}+\text{H}]^+$): 591,0623, gefunden: 591,0644.

IR: $\tilde{\nu}$ = 3226,5 (w); 3081,1 (w); 2979,7 (w); 2808,2 (w); 2361,8 (w); 1702,1 (m); 1688,9 (s); 1608,1 (m); 1440,3 (m); 1300,3 (m); 1271,7 (s); 1258,2 (s); 1178,0 (m); 1098,7 (m); 1088,1 (s); 1079,5 (s); 1048,2 (m); 1018,0 (m); 950,3 (m); 790,9 (m); 753,4 (s) cm^{-1} .

Schmelzbereich: 196 – 198 °C

5.5.4 Synthese von 3',5'-Bis-O-p-Toluoyl-5-iod-2'-desoxycytidin



Zu einer Lösung aus 1,2,4-Triazol (158 mg, 2,29 mmol, 9,0 Äq.) in trockenem MeCN (6 mL) wurde POCl_3 (0,052 mL, 0,56 mmol, 2,2 Äq.) auf Eis zugegeben und 15 min gerührt. Trockenes NEt_3 (0,317 mL, 2,29 mmol, 9,0 Äq.) wurde tröpfchenweise zugegeben und 20 min auf Eis gerührt. Nach Zugabe von 3',5'-Bis-O-p-Toluoyl-5-iod-2'-desoxyuridin (150 mg, 0,254 mmol, 1,0 Äq., α/β 1:0,3) wurde die gelbe Lösung 18 h bei 30 °C gerührt.

Anschließend wurden NEt₃ (1,2 mL) und H₂O (0,3 mL) zugegeben und 10 min bei 30 °C gerührt. Nach Extraktion mit DCM (3 × 30 mL) und Waschen mit einer 1:1-Lösung aus eiskaltem Brine und einer gesättigten NaHCO₃-Lösung (1 × 20 mL) wurde die organische Phase über MgSO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Danach wurden 1,4-Dioxan (35 mL) und konzentrierte wässrige NH₃-Lösung (5,7 mL) zugegeben und 15 min bei 30 °C gerührt. Die Lösung wurde mit einer gesättigten NH₄Cl-Lösung (40 mL) versetzt und nach 30 min Röhren mit DCM (4 × 40 mL) extrahiert. Die organische Phase wurde über MgSO₄ getrocknet und das Lösungsmittel wurde *in vacuo* entfernt. Der Rückstand wurde säulenchromatographisch gereinigt (DCM/MeOH 100 → 50:1) und ein Gemisch aus α- und β-3',5'-Bis-O-*p*-Toluoyl-5-iod-2'-desoxycytidin (112 mg, 0,190 mmol, 75 %, α/β 4:1) wurde als weißer Feststoff erhalten.

β-3',5'-Bis-O-*p*-Toluoyl-5-iod-2'-desoxycytidin:

¹H-NMR (400 MHz, CDCl₃) δ 7.99 (s, 1H, C6-H), 7.93 (d, *J* = 8.2 Hz, 2H, Tol-H2), 7.76 (m, 2H, Tol-H2), 7.33 – 7.25 (m, 2H, Tol-H3), 7.30 – 7.21 (m, 2H, Tol-H3), 6.39 (dd, *J* = 8.4, 5.4 Hz, 1H, C1'-H), 5.61 (d, *J* = 6.4 Hz, 1H, C3'-H), 5.51 (s, 2H, NH₂), 4.79 (dd, *J* = 12.3, 3.0 Hz, 1H, C5'-H), 4.70 (dd, *J* = 12.3, 3.4 Hz, 1H, C5'-H), 4.61 (d, *J* = 2.1 Hz, 1H, C4'-H), 3.00 (d, *J* = 6.6 Hz, 1H, C2'-H), 2.44 (s, 6H, CH₃), 2.23 – 2.18 (m, 1H, C2'-H) ppm.

¹³C-NMR (101 MHz, CDCl₃) δ 166.1 (CO(Tol)), 165.5 (CO(Tol)), 163.8 (C4), 154.7 (C2), 146.1 (C6), 144.5 (Tol-C4), 144.3 (Tol-C4), 129.8 (Tol-C2), 129.7 (Tol-C2), 129.5 (Tol-C3), 129.3 (Tol-C3), 126.3 (Tol-C1), 126.1 (Tol-C1), 87.1 (C1'), 83.6 (C4'), 75.3 (C3'), 64.2 (C5'), 55.1 (C5), 39.5 (C2'), 21.8 (CH₃), 21.7 (CH₃) ppm.

α-3',5'-Bis-O-*p*-Toluoyl-5-iod-2'-desoxycytidin:

¹H-NMR (400 MHz, CDCl₃) δ 7.97 (s, 1H, C6-H), 7.96 (d, *J* = 8.3 Hz, 2H, Tol-H2), 7.76 (d, *J* = 8.2 Hz, 2H, Tol-H2), 7.30 (d, *J* = 7.7 Hz, 2H, Tol-H3), 7.28 – 7.21 (m, 2H, Tol-H3), 6.33 (dd, *J* = 6.9, 1.2 Hz, 1H, C1'-H), 5.65 (d, *J* = 5.9 Hz, 1H, C3'-H), 5.51 (s, 2H, NH₂), 4.94 (t, *J* = 4.3 Hz, 1H, C4'-H), 4.57 (dd, *J* = 12.1, 4.2 Hz, 1H, C5'-H), 4.53 (dd, *J* = 12.0, 4.5 Hz, 1H, C5'-H), 3.03 – 2.90 (m, 1H, C2'-H), 2.67 – 2.59 (m, 1H, C2'-H), 2.45 (s, 3H, CH₃), 2.43 (s, 3H, CH₃) ppm.

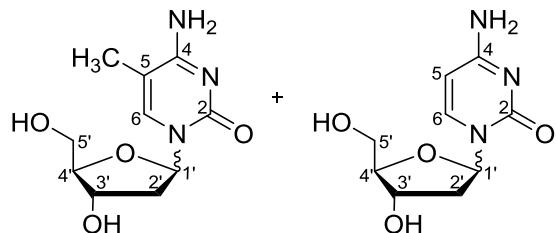
¹³C-NMR (101 MHz, CDCl₃) δ 166.1 (CO(Tol)), 165.5 (CO(Tol)), 163.8 (C4), 154.7 (C2), 146.6 (C6), 144.5 (Tol-C4), 144.3 (Tol-C4), 129.7 (Tol-C2), 129.7 (Tol-C2), 129.5 (Tol-C3), 129.4 (Tol-C3), 126.5 (Tol-C1), 126.1 (Tol-C1), 89.1 (C1'), 85.8 (C4'), 74.8 (C3'), 64.1 (C5'), 55.1 (C5), 39.2 (C2'), 21.8 (CH₃), 21.7 (CH₃) ppm.

HR-MS (ESI+): m/z berechnet für $[C_{25}H_{25}IN_3O_6]^+$ ($[M+H]^+$): 590,0783; gefunden: 590,0778.

IR: $\tilde{\nu}$ = 3035,0 (w); 1926,3 (w); 1715,5 (m); 1624,7 (s); 1610,9 (s); 1484,4 (m); 1264,8 (s); 1176,9 (m); 1088,9 (s); 1018,4 (m); 728,2 (s); 689,4 (m) cm^{-1} .

Schmelzbereich: 108 – 110 °C.

5.5.5 Synthese von 5-Methyl-2'-desoxycytidin (m^5dC) und 2'-Desoxycytidin (dC)



Ein Gemisch aus α - und β -3',5'-Bis-O-p-Toluoyl-5-iod-2'-desoxycytidin (50 mg, 90 μ mol, 1,0 Äq.) und Ni(dppp)Cl₂ (23 mg, 40 μ mol, 0,5 Äq.) wurden 30 min im Hochvakuum getrocknet, in 1,7 mL THF (50 mM) gelöst und 62 μ L einer 3 M H₃C-MgI-Lösung in Diethylether (0,19 mmol, 2,2 Äq.) wurden langsam auf Eis zugegeben. Nach 1,5 h Rühren bei RT wurden weitere 31 μ L 3 M H₃C-MgI-Lösung in Diethylether (90 μ mol, 1,1 Äq.) langsam auf Eis zu der Reaktionslösung gegeben und 1 h bei RT gerührt. Anschließend wurde mit einer gesättigten NH₄Cl-Lösung (4 mL) neutralisiert und mit EtOAc (4 \times 25 mL) extrahiert. Die organische Phase wurde über MgSO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Nach säulenchromatographischer Reinigung (DCM/ MeOH 70:1 \rightarrow 10:1) wurde eine Mischung aus α - und β -3',5'-bis-O-p-Toluoyl-5-methyl-2'-desoxycytidin und α - und β -3',5'-bis-O-p-Toluoyl-2'-desoxycytidin erhalten. Dieses Gemisch wurde in MeOH (3,6 mL) gelöst, K₂CO₃ (38 mg, 0,28 mmol, 11 Äq.) zugegeben und 96 h bei RT gerührt. Nach Aufreinigung mittels HPLC (0 \rightarrow 15 % MeCN in H₂O in 45 min) wurden ein Gemisch aus α - und β -5-Methyl-2'-desoxycytidin (0,3 mg, 1,2 μ mol, 3 %, α/β 2: 1), sowie reines α -5-Methyl-2'-desoxycytidin (1,3 mg, 5,4 μ mol, 14 %) als weißer Feststoff erhalten. Außerdem wurde α -Desoxycytidin (1,0 mg, 4,4 μ mol, 11 %) und β -Desoxycytidin (0,5 mg, 2,2 μ mol, 6 %) erhalten.

β -5-Methyl-2'-desoxycytidin:

¹H-NMR (800 MHz, D₂O) δ 7.56 (d, J = 1.3 Hz, 1H, C6-H), 6.19 (t, J = 6.7 Hz, 1H, C1'-H), 4.35 (dt, J = 6.7, 4.1 Hz, 1H, C3'-H), 3.94 (q, J = 4.2 Hz, 1H, C4'-H), 3.75 (dd, J = 12.5,

3.5 Hz, 1H, C5'-H), 3.67 (dd, $J = 12.5, 5.1$ Hz, 1H, C5'-H), 2.30 (ddd, $J = 14.1, 6.5, 4.1$ Hz, 1H, C2'-H), 2.21 (dt, $J = 13.9, 6.8$ Hz, 1H, C2'-H), 1.88 – 1.87 (m, 3H, CH₃) ppm.

¹³C-NMR (201 MHz, D₂O) δ 166.0 (C4), 157.4 (C2), 138.6 (C6), 104.7 (C5), 86.4 (C4'), 85.7 (C1'), 70.4 (C3'), 61.1 (C5'), 39.1 (C2'), 12.3 (CH₃) ppm.

α -5-Methyl-2'-desoxycytidin:

¹H-NMR (800 MHz, D₂O) δ 7.61 (d, $J = 1.3$ Hz, 1H, C6-H), 6.05 (dd, $J = 7.1, 3.0$ Hz, 1H, C1'-H), 4.34 - 4.31 (m, 2H, C4'-H und C3'-H), 3.63 (dd, $J = 12.3, 3.6$ Hz, 1H, C5'-H), 3.54 (dd, $J = 12.3, 5.3$ Hz, 1H, C5'-H), 2.68 – 2.59 (m, 1H, C2'-H), 2.05 (dt, $J = 14.7, 2.9$ Hz, 1H, C2'-H), 1.90 – 1.89 (m, 3H, CH₃) ppm.

¹³C-NMR (201 MHz, D₂O) δ 166.0 (C4), 157.3 (C2), 139.0 (C6), 103.8 (C5), 88.6 (C4'), 87.6 (C1'), 70.8 (C3'), 61.5 (C5'), 40.0 (C2'), 12.3 (CH₃) ppm.

HR- MS (ESI+): m/z berechnet für [C₁₀H₁₆N₃O₄]⁺ ([M+H]⁺): 242,1135, gefunden: 242,1134.

α -2'-Desoxycytidin:

¹H-NMR (400 MHz, D₂O) δ 7.75 (d, $J = 7.5$ Hz, 1H, C6-H), 6.01 (dd, $J = 7.1, 2.5$ Hz, 1H, C1'-H), 5.91 (d, $J = 7.5$ Hz, 1H, C5-H), 4.30 (dq, $J = 5.0, 3.1, 2.2$ Hz, 2H, C3'-H und C4'-H), 3.60 (dd, $J = 12.3, 3.9$ Hz, 1H, C5'-H), 3.51 (dd, $J = 12.3, 5.5$ Hz, 1H, C5'-H), 2.59 (dt, $J = 14.7, 6.5$ Hz, 1H, C2'-H), 2.04 (dt, $J = 14.9, 2.5$ Hz, 1H, C2'-H).

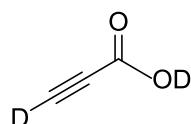
β -2'-Desoxycytidin:

¹H-NMR (400 MHz, D₂O) δ 7.71 (d, $J = 7.5$ Hz, 1H, C6-H), 6.15 (t, $J = 6.6$ Hz, 1H, C1'-H), 5.93 (d, $J = 7.5$ Hz, 1H, C5-H), 4.31 (dd, $J = 6.9, 3.7$ Hz, 1H, C3'-H), 3.93 (q, $J = 4.2$ Hz, 1H, C4'-H), 3.71 (dd, $J = 12.5, 3.6$ Hz, 1H, C5'-H), 3.63 (dd, $J = 12.5, 5.2$ Hz, 1H, C5'-H), 2.30 (ddd, $J = 14.2, 6.5, 4.1$ Hz, 1H, C2'-H), 2.25 – 2.12 (m, 1H, C2'-H).

MS (ESI+): m/z berechnet für [C₉H₁₄N₃O₄]⁺ ([M+H]⁺): 228,1, gefunden: 228,1.

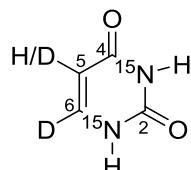
5.6 Synthese von 5-[D₃]-Methyl-2'-desoxy-[6-D, 1,3-¹⁵N₂]-cytidin

5.6.1 [D₂]-Propiolsäure



Na₂CO₃ (6,57 g, 62,0 mmol, 1,5 Äq.) wurde in D₂O (25 mL) gelöst und Propiolsäure (2,60 mL, 42,0 mmol, 1,0 Äq.) innerhalb von 3 min langsam zugetropft. Nach 18 h Rühren bei RT wurde das Lösungsmittel *in vacuo* entfernt und es wurde erneut D₂O (21 mL) hinzugefügt und 14 h bei RT gerührt. Ein pH-Wert von 1 wurde mit D₂SO₄ eingestellt. Nach der Extraktion mit DCM wurde die organische Phase über MgSO₄ getrocknet. Das Lösungsmittel wurde entfernt (600 mbar, 40 °C) und die fraktionierte Destillation (27 mbar, 39 °C) ergab [D₂]-Propiolsäure (452 mg, 6,27 mmol, 15 %) als farblose Flüssigkeit und wurde direkt weiter umgesetzt.

5.6.2 [D₂, ¹⁵N₂]-Uracil (29) und [D, ¹⁵N₂]-Uracil

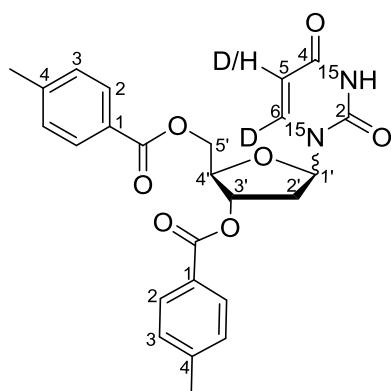


[¹⁵N₂]-Harnstoff (150 mg, 2,41 mmol, 1,0 Äq.) wurde in D₂SO₄ (3 mL) gelöst und [D₂]-Propiolsäure (0,161 mL, 2,65 mmol, 1,1 Äq.) wurde langsam zugetropft. Die Reaktionsmischung wurde auf 105 °C erhitzt und 18 h gerührt. Die braun-schwarze Suspension wurde mit Wasser (50 mL) verdünnt und mit Na₂CO₃ auf einen pH-Wert von 10 gebracht. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand mittels Säulenchromatographie (DCM/MeOH 4:1) gereinigt. Ein Gemisch aus [D₂, ¹⁵N₂]-Uracil **29** (101 mg, 0,864 mmol, 36 %) und [D, ¹⁵N₂]-Uracil (50 mg, 0,43 mmol, 18 %) wurde als farbloses Pulver erhalten.

¹H-NMR (400 MHz, DMSO-d₆) δ 11.03 (dd, ¹J = 77.2 Hz, ⁴J = 1.8 Hz, 1H, N3-H), 10.80 (dd, ¹J = 82.6 Hz, ⁴J = 1.9 Hz, 1H, N1-H), 5,46 (dd, ³J = 4.3 Hz, ⁴J = 2.5 Hz, C5-H) ppm.

¹³C-NMR (101 MHz, DMSO-d₆) δ 164.9 (t, ¹J = 9.7 Hz, C4), 151.9 (m, C2), 142.3 (m, C6), 101.4-99.5 (m, C5) ppm.

5.6.3 [D₂, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-2'-desoxyuridin (30)



Das Gemisch aus [D₂, ¹⁵N₂]-Uracil **29** und [D, ¹⁵N₂]-Uracil (37 mg, 0,32 mmol, 1,0 Äq.) wurde in MeCN (2 mL) gelöst. Bis(trimethylsilyl)acetamid (0,23 mL, 0,96 mmol, 3,0 Äq.) wurde langsam zugetropft und die Reaktionsmischung wurde 3 h bei 80 °C gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und drei Mal mit Chloroform (6 mL) coevaporiert. Der Rückstand wurde in Chloroform (2 mL) gelöst und Hoffer's Chlorozucker (149 mg, 0,383 mmol, 1,2 Äq.) wurde hinzugefügt und 17 h bei RT gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand mittels Säulenchromatographie (iHex/EtOAc 1:1) gereinigt. Es wurde ein Gemisch aus α- und β-[D₂, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-2'-desoxyuridin und α- und β-[D, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-2'-desoxyuridin **30** (41 mg, 0,088 mmol, 28 %, α/β: 1,28:1) als gelber Schaum erhalten.

α-[D₂, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-2'-desoxyuridin:

¹H-NMR (400 MHz, CDCl₃): δ = 9.06 (d, ¹J = 91.0 Hz, 1H, NH), 7.89 (d, ³J = 8.0 Hz, 2H, Tol2), 7.78 (d, ³J = 8.0 Hz, 2H, Tol2), 7.31-7.20 (m, 4H, Tol3), 6.31 (d, ³J = 6.8 Hz, 1H, C1'-H), 5.72 (dt, ³J = 4.7 Hz, ³J = 2.4 Hz, 1H, C5-H), 5.61 (m, 1H, C3'-H), 4.86 (t, ³J = 4.3 Hz, 1H, C4'-H), 4.54 (m, 2H, C5'-H), 3.01-2.89 (m, 1H, C2'-H), 2.55 (d, ³J = 15.5 Hz, 1H, C2'-H), 2.43 (s, 3H, CH₃), 2.41 (s, 3H, CH₃) ppm.

¹³C-NMR (101 MHz, CDCl₃): δ = 165.7 (COO), 166.1 (COO), 150.1 (C2), 144.9 (Tol4), 144.4 (Tol4), 129.7 (Tol2), 129.6 (Tol2), 129.5 (Tol3), 129.4 (Tol3), 126.5 (Tol1), 125.9 (Tol1), 101.4 (C5), 88.0 (d, ¹J = 10.0 Hz, C1'), 85.7 (C4'), 74.7 (C3'), 64.0 (C5'), 39.0 (C2'), 21.7 (CH₃) ppm.

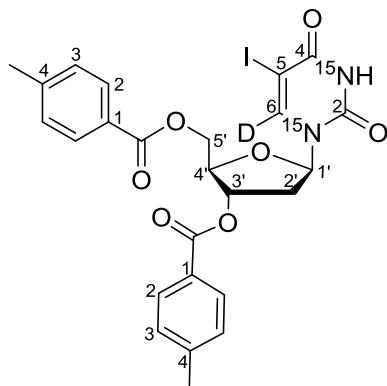
β -[D₂, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-2'-desoxyuridin (30):

¹H-NMR (400 MHz, CDCl₃): δ = 8.93 (d, ¹J = 91.0 Hz, 1H, NH), 7.94 (d, ³J = 8.0 Hz, 2H, Tol2), 7.93 (d, ³J = 8.0 Hz, 2H, Tol2) 7.31-7.20 (m, 4H, Tol3), 6.41 (dd, ³J = 8.3 Hz, ³J = 5.7 Hz, 1H, C1'-H), 5.60 (m, 2H, C3'-H und C5-H), 4.73 (dd, ³J = 12.3 Hz, ⁴J = 3.1 Hz, 1H, C5'-H), 4.67 (dd, ³J = 12.3 Hz, ⁴J = 3.5 Hz, 1H, C5'-H), 4.60-4.49 (m, 1H, C4'-H), 2.75 (dd, ³J = 14.3 Hz, ³J = 5.7 Hz, 1H, C2'-H), 2.43 (s, 6H, CH₃), 2.30 (m, 1H, C2'-H) ppm.

¹³C-NMR (101 MHz, CDCl₃): δ = 165.7 (COO), 166.1 (COO), 150.1 (C2), 144.7 (Tol4), 129.9 (Tol2), 129.5 (Tol2), 129.4 (Tol3), 129.3 (Tol3), 126.5 (Tol1), 125.9 (Tol1), 101.4 (C5), 85.4 (d, ¹J = 12.8 Hz, C1') 83.0 (C4'), 74.7 (C3'), 64.0 (C5'), 38.4 (C2'), 21.7 (CH₃) ppm.

MS (ESI+): m/z berechnet für [C₂₅H₂₃D₂O₇¹⁵N₂]⁺ ([M-H]⁺): 469,17, gefunden: 469,15.

5.6.4 β -[D, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-5-iod-2'-desoxyuridin (33)



Eine Mischung aus α - und β -[D₂, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-2'-desoxyuridin und α - und β -[D₂, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-2'-desoxyuridin (267 mg, 0,570 mmol, 1,0 Äq.) wurde in MeCN (6 mL) gelöst. LiI (92 mg, 0,68 mmol, 1,2 Äq.) und Ammoniumcer(IV)-nitrat (625 mg, 1,14 mmol, 2,0 Äq.) wurden zugefügt. Die Suspension wurde bei 80 °C 3 h geheizt und anschließend wurde eine gesättigte NaCl-Lösung (30 mL) hinzugefügt. Nach Extraktion mit DCM wurde die organische Phase über MgSO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde mittels Säulenchromatographie (*i*Hex/EtOAc 1:1) gereinigt und es wurde β -[D, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-5-iod-2'-desoxyuridin **33** (177 mg, 0,298 mmol, 52 %) als gelber Feststoff erhalten.

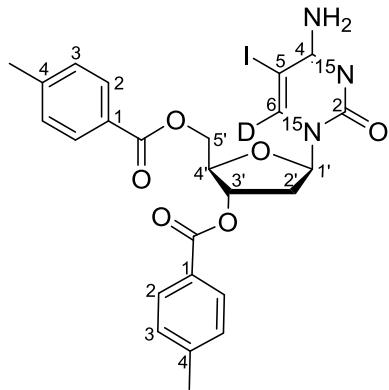
¹H-NMR (400 MHz, CDCl₃): δ = 8.99 (d, ¹J = 93.6 Hz, 1H, NH), 7.94 (d, ³J = 3.5 Hz, 2H, Tol2), 7.92 (d, ³J = 3.6 Hz, 2H, Tol2), 7.27 (d, ³J = 7.7 Hz, 4H, Tol3), 6.37 (dd, ³J = 8.7 Hz, ³J

= 5.4 Hz, 1H, C1'-H), 5.61 (d, 3J = 6.6 Hz, 1H, C3'-H), 4.73 (t, 3J = 2.7 Hz, 2H, C5'-H), 4.57 (q, 3J = 2.7 Hz, 1H, C4'-H), 2.78 (dd, 3J = 14.3, 3J = 5.5 Hz, 1H, C2'-H), 2.43 (s, 3H, CH₃), 2.42 (s, 3H, CH₃), 2.28 (m, 1H, C2'-H) ppm.

$^{13}\text{C-NMR}$ (101 MHz, CDCl₃): δ = 166.1 (COO), 166.0 (COO), 159.6 (d, 1J = 10.4 Hz, C4), 149.7 (t, 1J = 18.9 Hz, C2), 144.7 (Tol4), 144.6 (Tol4), 129.9 (Tol2), 129.7 (Tol2), 129.6 (Tol3), 129.3 (Tol3), 126.4 (Tol1), 126.2 (Tol1), 85.8 (d, 1J = 12.3 Hz, C1'), 83.5 (C4'), 74.9 (C3'), 68.8 (d, 2J = 9.0 Hz, C5), 64.2 (C5'), 38.8 (C2'), 21.8 (CH₃) ppm.

HR-MS (ESI+): m/z berechnet für [C₂₅H₂₃DO₇¹⁵N₂I]⁺ ([M+H]⁺): 592,0480, gefunden: 592,0476.

5.6.5 β -[D, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-5-iod-2'-desoxycytidin (34)



1,2,4-Triazol (42 mg, 0,60 mmol, 9,0 Äq.) wurde in MeCN (1 mL) gelöst. POCl₃ (0,028 mL, 0,29 mmol, 4,4 Äq.) wurde bei 0 °C hinzugefügt und bei 0 °C 10 min gerührt. NEt₃ (0,084 mL, 0,60 mmol, 9 Äq.) wurde hinzugefügt und 20 min bei RT gerührt. β -[D, ¹⁵N₂]-3',5'-Bis-O-(p-Toluoyl)-5-iod-2'-desoxyuridin **33** (40 mg, 0,07 mmol, 1,0 Äq.) wurde in MeCN (1,5 mL) gelöst, zur Reaktionsmischung zugegeben und 18 h bei 30 °C gerührt. Nach Zugabe von NEt₃ (0,4 mL) und H₂O (0,1 mL) wurde die Suspension 10 min bei 30 °C gerührt. Eine gesättigte NaCl-Lösung (5 mL) und gesättigte NaHCO₃-Lösung (5 mL) wurden auf 0 °C gekühlt und zur Reaktionsmischung dazugegeben. Nach Extraktion mit DCM wurde die organische Phase über MgSO₄ getrocknet und das Lösungsmittel wurde *in vacuo* entfernt. Dioxan (10 mL) und konz. NH₃-Lösung (1,5 mL) wurden zugegeben und 15 min bei 30 °C gerührt. Gesättigte NH₄Cl-Lösung (10 mL) wurde hinzugefügt und mit DCM extrahiert. Die organische Phase wurde über MgSO₄ getrocknet und das Lösungsmittel wurde *in vacuo* entfernt. Der Rückstand wurde durch Säulenchromatographie (DCM/MeOH 50:1) gereinigt

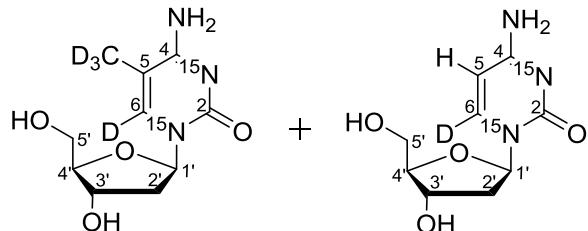
und $[\text{D}, ^{15}\text{N}_2]\text{-}3',5'\text{-Bis-O-(p-Toluoyl)-5-iod-2'-desoxycytidin}$ **34** (17 mg, 0,029 mmol, 43 %) wurde als gelber Feststoff erhalten.

$^1\text{H-NMR}$ (400 MHz, CDCl_3): δ = 7.94 (d, 3J = 8.0 Hz, 2H, Tol2), 7.90 (d, 3J = 8.1 Hz, 2H, Tol2), 7.29-7.22 (m, 4H, Tol3), 6.36 (dd, 3J = 8.4 Hz, 3J = 5.4 Hz, 1H, C1'-H), 5.58 (d, 3J = 6.4 Hz, 1H, C3'-H), 5.49 (s, 2H, NH_2), 4.76 (dd, 3J = 12.3 Hz, 3J = 3.0 Hz, 1H, C5'-H), 4.68 (dd, 3J = 12.3 Hz, 3J = 3.4 Hz, 1H, C5'-H), 4.59 (q, 3J = 2.7 Hz, 1H, C4'-H), 3.02-2.92 (m, 1H, C2'-H), 2.42 (s, 3H, CH_3), 2.41 (s, 3H, CH_3), 2.17 (m, 1H, C2'-H) ppm.

$^{13}\text{C-NMR}$ (101 MHz, CDCl_3): δ = 166.1 (COO), 163.7 (C4), 154.9 (C2), 144.5 (Tol4) 144.5 (Tol4), 129.8 (Tol2), 129.7 (Tol2), 129.5 (Tol3), 129.3 (Tol3), 126.4 (Tol1), 126.3 (Tol1), 87.1 (d, 1J = 11.2 Hz, C1'), 83.6 (C4'), 75.3 (C3'), 64.2 (C5'), 56.1 (C5), 39.5 (C2'), 21.8 (CH_3) ppm.

HR-MS (ESI+): m/z berechnet für $[\text{C}_{25}\text{H}_{24}\text{DO}_6\text{N}^{15}\text{N}_2\text{I}]^+$ ($[\text{M}+\text{H}]^+$): 593,0792, gefunden: 593.0791.

5.6.6 5-[D_3]-Methyl-2'-desoxy-[6-D, 1,3- $^{15}\text{N}_2$]-cytidin und 2'-Desoxy-[6-D, 1,3- $^{15}\text{N}_2$]-cytidin



$\text{Ni}(\text{dppp})\text{Cl}_2$ (8 mg, 0,01 mmol, 0,5 Äq.) wurde in einen Schlenkkolben überführt und 30 min *in vacuo* getrocknet. β -[D, $^{15}\text{N}_2]\text{-}3',5'\text{-Bis-O-(p-Toluoyl)-5-iod-2'-desoxycytidin}$ **34** (17 mg, 0,029 mmol, 1,0 Äq.) wurde in THF (1 mL) gelöst und zur Reaktionsmischung bei 0 °C hinzugefügt. D_3CMgI (0,072 mL, 0,072 mmol, 2,5 Äq.) wurde tropfenweise hinzugefügt und bei RT 1,5 h gerührt. D_3CMgI (0,072 mL, 0,072 mmol, 2,5 Äq.) wurde erneut tropfenweise hinzugefügt und 1 h bei RT gerührt. D_3CMgI (0,20 mL, 0,20 mmol, 6,9 Äq.) und $\text{Ni}(\text{dppp})\text{Cl}_2$ (8 mg, 0,01 mmol, 0,5 Äq.) wurde hinzugefügt und 17 h bei RT gerührt. Gesättigte NH_4Cl -Lösung (8 mL) wurde hinzugefügt. Nach Extraktion mit EtOAc wurde die organische Phase über MgSO_4 getrocknet und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde durch Säulenchromatographie (DCM/MeOH 20:1) gereinigt und eine Mischung aus β -5-[D_3]-Methyl-3',5'-Bis-O-(p-Toluoyl)-2'-desoxy-[D, $^{15}\text{N}_2$]-cytidin **35** und β -3',5'-Bis-O-(p-Toluoyl)-2'-desoxy-

[D, $^{15}\text{N}_2$]-cytidin **36** wurde als gelber Schaum erhalten. Diese Mischung wurde in MeOH (1,2 mL) gelöst, K_2CO_3 (13 mg, 0,094 mmol, 4,7 Äq.) wurde hinzugefügt und 18 h bei RT gerührt. Es wurde erneut K_2CO_3 (13 mg, 0,094 mmol, 4,7 Äq.) und MeOH (1,2 mL) hinzugefügt und 17 h an RT gerührt. Nach Entfernen des Lösungsmittels *in vacuo* wurde der Rückstand *via* HPLC (*Macherey-Nagel*, Nucleosil 100-7 C18, 10 × 250 mm, linearer Gradient von 0 % bis 10 % Acetonitril in Wasser in 45 min) gereinigt und [D₄, $^{15}\text{N}_2$]-m⁵dC (0,9 mg, 0,004 mmol, 12 %) wurde als farbloser Feststoff erhalten. Außerdem wurde 2'-Desoxy-[D, $^{15}\text{N}_2$]-cytidin (1,1 mg, 0,0048 mmol, 17 %) als farbloser Feststoff erhalten.

5-[D₃]-Methyl-2'-Desoxy-[6-D, 1,3- $^{15}\text{N}_2$]-cytidin:

¹H-NMR (800 MHz, D₂O): δ = 6.30 (t, 3J = 6.7 Hz, 1H, C1'-H), 4.46 (dt, 3J = 6.6 Hz, 3J = 4.1 Hz, 1H, C3'-H), 4.05 (q, 3J = 4.2 Hz, 1H, C4'-H), 3.86 (dd, 3J = 12.5 Hz, 3J = 3.6 Hz, 1H, C5'-H), 3.78 (dd, 3J = 12.5 Hz, 3J = 5.1 Hz, 1H, C5'-H), 2.47-2.36 (m, 1H, C2'-H), 2.38-2.25 (m, 1H, C2'-H) ppm.

¹³C-NMR (201 MHz, D₂O): δ = 89.1 (C4'), 88.4 (d, 1J = 13.0 Hz, C1'), 73.1 (C3'), 63.8 (C5'), 41.8 (C2') ppm.

HR-MS (ESI+): m/z berechnet für [C₁₀H₁₂D₄O₄N¹⁵N₂]⁺ ([M+H]⁺): 248,1327, gefunden: 248,1326.

2'-Desoxy-[6-D, 1,3- $^{15}\text{N}_2$]-cytidin:

¹H-NMR (800 MHz, D₂O): δ = 6.17 (t, 3J = 6.7 Hz, 1H, C1'-H), 5.95 (d, 3J = 3.7 Hz, 1H, C5-H), 4.35-4.33 (m, 1H, C3'-H), 3.96 (dt, 3J = 5.3, 3.8 Hz, 1H, C4'-H), 3.78-3.70 (m, 1H, C5'-H), 3.68-3.60 (m, 1H, C5'-H), 2.38-2.30 (m, 1H, C2'-H), 2.24-2.18 (m, 1H, C2'-H) ppm.

¹³C-NMR (201 MHz, D₂O): δ = 166.18 (C4), 157.58 (C2), 96.02 (C5), 86.6 (C4'), 86.0 (d, 1J = 11.8 Hz, C1'), 70.5 (C3'), 61.2 (C5'), 39.19 (C2') ppm.

HR-MS (ESI+): m/z berechnet für [C₉H₁₃DO₄N¹⁵N₂]⁺ ([M+H]⁺): 231,0982, gefunden: 231,0981.

6 Abkürzungsverzeichnis

AAA	2-Acetylaminoanthracen
AAB	2-Acetylaminobenzol
AAF	2-Acetylaminofluoren
AAN	2-Acetylaminonaphthalin
AAP	2-Acetylaminonpyren
Ac	Acetyl
AF	2-Aminofluoren
AFB1	Aflatoxin B1
AID	Activation-Induced Deaminase
AN	2-Aminonaphthalin
AP	2-Aminopyren
APOBEC1	Apolipoprotein B pre-mRNA Editing Enzyme, Catalytic Polypeptide 1
Äq.	Äquivalente
Asn	Asparagin
Asp	Asparaginsäure
ATP	Adenosintriphosphat
B[a]P	Benzo[a]pyren
B[a]A	(1S)-Benzo[a]anthracen
BER	Basenexzisionsreparatur
bidest.	bidestilliert
BINAP	2,2'-Bis(diphenylphosphino)-1,1'-binaphthyl
cadC	5-Carboxy-2'-desoxycytidin
CD	Circulardichronismus
COSY	Correlated Spectroscopy
CPG	Controlled Pore Glass
CS	Cockayne Syndrom
δ	chemische Verschiebung (NMR)
dA	2'-Desoxyadenosin
DAMT-1	DNA N6-Adenin Methyltransferase 1
dba	Dibenzylidenaceton
dC	2'-Desoxycytosin
DC	Dünnschichtchromatografie
DCM	Dichlormethan
dG	2'-Desoxyguanosin
DIAD	Di- <i>iso</i> -propylazodicarboxylat

DIAT	Di- <i>iso</i> -propylammoniumtetrazolid
DMAD	DNA N6-Methyladenin Demethylase in Drosophila
DMF	<i>N,N'</i> -Dimethylformamid
DMSO	Dimethylsulfoxid
DMTr	4,4'-Dimethoxytrityl
DNA	Desoxyribonukleinsäure
dsDNA	doppelsträngige DNA
dT	2'-Desoxythymidin
DTT	Dithiothreitol
EGFP	Enhanced Green Fluorescent Protein
EMSA	Electrophoretic Mobility Shift Assay
ESI	Electron Spray Ionization
ERCC-1	Excision repair cross-complementing group 1
ES-Zellen	Embryonale Stammzellen
Et	Ethyl
fdC	5-Formyl-2'-desoxycytidin
g	Gramm
GG-NER	Globale Genom NER
H3K4me2	Histon H3 mit Dimethylierung von Lysin 4
HMBC	Heteronuclear Multiple Bond Correlation
hmdC	5-Hydroxymethyl-2'-desoxycytidin
HMGN1	High mobility group nucleosome binding domain 1
HPLC	High Perfomance Liquid Chromatography
HRMS	High Resolution Mass Spectorscopy
HSQC	Heteronuclear Single Quantum Coherence
HV	Hochvakuum
Hz	Hertz
<i>i</i> Hex	<i>iso</i> -Hexan
IR	Infrarot
<i>i</i> Pr	<i>iso</i> -Propyl
J	Kopplungskonstante
L	Liter
L1-Elemente	junge Vollängen LINE-1 Transposons
M	Molar (mol/Liter)
m ⁴ dC	4-Methyl-2'-desoxycytidin
m ⁵ dC	5-Methyl-2'-desoxycytidin
m ⁶ dA	6-Methyl-2'-desoxyadenosin

MALDI	Matrix-unterstützte Laser-Desorption/Ionisation
MBD4	Methyl-CpG-Bindungsprotein 4
<i>m</i> CPBA	<i>meta</i> -Chlorperbenzoësäure
Me	Methyl
mES-Zellen	murine embryonale Stammzellen
mg	Milligramm
MS	Massenspektrometrie
MS/MS	Tandemmassenspekrometrie
v	Wellenzahl (IR)
NER	Nukleotidexzisionsreparatur
NF	Nitrofluoren
<i>N</i> -HO-AAN	<i>N</i> -Hydroxyacetylaminonaphthalin
<i>N</i> -HO-AN	<i>N</i> -Hydroxyaminonaphthalin
nm	Nanometer
NMR	Kernresonanzspektroskopie
NP	Nitropyren
ODN	Oligodesoxynukleotid
Pac ₂ O	Phenoxyessigsäureanhydrid
PAK	Polycyclischer Aromatischer Kohlenwasserstoff
Phe	Pphenylalanin
Pol	Polymerase
ppm	parts per million
Pro	Prolin
R/M-System	Restriktions-Modifikations-System
RNA	Ribonukleinsäure
RNAPII	RNA Polymerase II
RPA	Replikationsprotein A
ROS	Reactive oxygen species
RT	Raumtemperatur
SAH	S-Adenosylhomocysteïne
SAM	S-Adenosylmethionin
Ser	Serin
SMUG1	Single-strand selective Monofunctional Uracil DNA Glycosylase
ssDNA	einzelsträngige DNA
TBAF	<i>tert</i> -Butylammoniumfluorid
TBS	<i>tert</i> -Butyldimethylsilyl
<i>t</i> -Bu	<i>tert</i> -Butyl

TC-NER	Transkriptions gekoppelte NER
TDG	Thymin DNA Glykolase
TEAF	Tetraethylammoniumfluorid
TEMED	<i>N,N,N',N'</i> -Tetramethylethylendiamin
TFIIC	Transkriptionsfaktor II C
TFIIS	Transkriptionsfaktor II S
THF	Tetrahydrofuran
TLS	Translesion DNA Synthesis
-TOF	Massenspektrometrie mit Flugzeitanalysator (time of flight)
TPS	2,4,6-Tri- <i>iso</i> -propylbenzolsulfonyl
TQ	Triple Quadrupole
TSS	Transkriptions Startstellen (Transcription Start Sites)
TTDA	Trichothiodystrophie Gruppe A
Tyr	Tyrosin
UHPLC	Ultra-high-performance Flüssigchromatografie
UHRF1	Ubiquitin-like plant and RING Finger domain 1
UV	Ultraviolett
Xantphos	4,5-Bis(diphenylphosphino)-9,9-dimethylxanthen
XP	Xeroderma Pigmentosum

7 Literaturverzeichnis

- [1] J. D. Watson, F. H. Crick, *Nature* **1953**, 171, 737-738, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.*
- [2] E. Chargaff, S. Zamenhof, C. Green, *Nature* **1950**, 165, 756-757, *Composition of human desoxypentose nucleic acid.*
- [3] R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Nature* **1980**, 287, 755-758, *Crystal structure analysis of a complete turn of B-DNA.*
- [4] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemie* 6. Auflage, Spektrum Akademischer Verlag, München, **2002**.
- [5] C. O. Pabo, R. T. Sauer, *Annu. Rev. Biochem.* **1984**, 53, 293-321, *Protein-DNA recognition.*
- [6] J. L. T. J. M. Berg, L. Stryer, *Biochemistry*, Vol. 6, W. H. Freeman, New York, **2006**.
- [7] D. Wion, J. Casadesús, *Nat. Rev. Microbiol.* **2006**, 4, 183-192, *N(6)-methyl-adenine: an epigenetic signal for DNA-protein interactions.*
- [8] R. Holliday, J. Pugh, *Science* **1975**, 187, 226-232, *DNA modification mechanisms and gene activity during development.*
- [9] A. D. Riggs, *Cytogenet. Genome Res.* **1975**, 14, 9-25, *X inactivation, differentiation, and DNA methylation.*
- [10] F. Crick, *Nature* **1970**, 227, 561-563, *Central Dogma of Molecular Biology.*
- [11] F. H. Crick, *Symp. Soc. Exp. Biol.* **1958**, 12, 138-163, *On protein synthesis.*
- [12] T. Lindahl, *Nature* **1993**, 362, 709-715, *Instability and decay of the primary structure of DNA.*
- [13] T. Lindahl, B. Nyberg, *Biochemistry* **1972**, 11, 3610-3618, *Rate of depurination of native deoxyribonucleic acid.*
- [14] Knippers, *Molekulare Genetik* 9. Auflage, Thieme, **2006**.
- [15] J. A. Miller, *Cancer Res.* **1970**, 30, 559-576, *Carcinogenesis by Chemicals: An Overview—G. H. A. Clowes Memorial Lecture.*
- [16] J. S. Felton, M. G. Knize, N. H. Shen, B. D. Andresen, L. F. Bjedanee, F. T. Hatch, *Env. Health Perspect.* **1986**, 57, 17-24, *Identification of the mutagens in cooked beef.*
- [17] M. Maclure, R. B. Katz, M. S. Bryant, P. L. Skipper, S. R. Tannenbaum, *Am. J. Public Health* **1989**, 79, 1381-1384, *Elevated blood levels of carcinogens in passive smokers.*
- [18] R. J. Turesky, *Toxicol. Lett.* **2007**, 168, 219-227, *Formation and biochemistry of carcinogenic heterocyclic aromatic amines in cooked meats.*
- [19] F. A. Beland, F. F. Kadlubar, *Environ. Health Perspect.* **1985**, 62, 19-30, *Formation and persistence of arylamine DNA adducts in vivo.*
- [20] A. K. Basu, J. M. Essigmann, *Chem. Res. Toxicol.* **1988**, 1, 1-18, *Site-specifically modified oligodeoxynucleotides as probes for the structural and biological effects of DNA-damaging agents.*
- [21] K. Hemminki, *Arch. Toxicol.* **1983**, 52, 249-285, *Nucleic acid adducts of chemical carcinogens and mutagens.*
- [22] B. Singer, J. T. Kusmierek, *Annu. Rev. Biochem.* **1982**, 51, 655-691, *Chemical Mutagenesis.*
- [23] F. F. Kadlubar, J. F. Anson, K. L. Dooley, F. A. Beland, *Carcinogenesis* **1981**, 2, 467-470, *Formation of urothelial and hepatic DNA adducts from the carcinogen 2-naphthylamine.*

- [24] E. Kriek, *Chem. Biol. Interact.* **1971**, 3, 19-28, *On the mechanism of action of carcinogenic aromatic amines: II. Binding of N-hydroxy-N-acetyl-4-aminobiphenyl to rat-liver nucleic acids in vivo.*
- [25] E. Kriek, *Proceedings of the 11th International Cancer Congress*, Vol. 2, Excerpta Medica, Amsterdam, **1975**.
- [26] S. Hecht, K. El-Bayoumy, A. Riverson, E. Fiala, *Cancer Lett.* **1982**, 16, 103-108, *A study of chemical carcinogenesis. 43. Comparative carcinogeneity of ortho-toluidine hydrochloride and ortho-nitrosotoluene in F344 rats.*
- [27] D. Grundberger, B. Weinstein, *Chemical Carcinogens and DNA*, Vol. 2, CRC Press, Boca Raton, FL., **1979**.
- [28] R. H. Wilson, F. DeEds, A. J. Cox, *Cancer Res.* **1941**, 1, 595-608, *The Toxicity and Carcinogenic Activity of 2-Acetaminofluorene.*
- [29] R. M. Campbell, M. L. Lee, *Anal. Chem.* **1984**, 56, 1026-1030, *Capillary column gas chromatographic determination of nitro polycyclic aromatic compounds in particulate extracts.*
- [30] B. Beije, L. Möller, *Mutat. Res.-Rev. Genet.* **1988**, 196, 177-209, *2-nitrofluorene and related compounds: prevalence and biological effects.*
- [31] F. Beland, M. Marques, *IARC Sci. Publ.* **1994**, 125, 229-244, *DNA Adducts of Nitropolycyclic Aromatic Hydrocarbons.*
- [32] R. H. Helflich, R. E. Neft, *Mutat. Res.* **1994**, 318, 73-114, *Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites.*
- [33] G. R. Hoffmann, R. P. P. Fuchs, *Chem. Res. Toxicol.* **1997**, 10, 347-359, *Mechanisms of Frameshift Mutations: Insight from Aromatic Amines.*
- [34] D. M. Grant, N. C. Hughes, S. A. Janezic, G. H. Goodfellow, H. J. Chen, A. Gaedigk, V. L. Yu, R. Grewal, *Mut. Res.-Fund. Mol. M.* **1997**, 376, 61-70, *Human acetyltransferase polymorphisms.*
- [35] C. M. King, S. J. Land, R. F. Jones, M. Debiec-Rychter, M.-S. Lee, C. Y. Wang, *Mutat. Res.- Fund. Mol. Mech. Mutagen.* **1997**, 376, 123-128, *Role of acetyltransferases in the metabolism and carcinogenicity of aromatic amines.*
- [36] J. A. Miller, E. C. Miller, *Prog. Exp. Tumor. Res.* **1969**, 11, 273-301, *The metabolic activation of carcinogenic aromatic amines and amides.*
- [37] E. Kriek, *Cancer Res.* **1972**, 32, 2042-2048, *Persistent Binding of a New Reaction Product of the Carcinogen N-Hydroxy-N-2-Acetylaminofluorene with Guanine in Rat Liver DNA in Vivo.*
- [38] J. G. Westra, E. Kriek, H. Hittenhausen, *Chem. Biol. Interact.* **1976**, 15, 149-164, *Identification of the persistently bound form of the carcinogen N-acetyl-2-aminofluorene to rat liver DNA in vivo.*
- [39] P. C. Howard, D. A. Casciano, F. A. Beland, J. G. Shaddock, *Carcinogenesis* **1981**, 2, 97-102, *The binding of N-hydroxy-2-acetylaminofluorene to DNA and repair of the adducts in primary rat hepatocyte cultures.*
- [40] C. C. Irving, R. A. Veazey, *Cancer Res.* **1969**, 29, 1799-1804, *Persistent Binding of 2-Acetylaminofluorene to Rat Liver DNA in Vivo and Consideration of the Mechanism of Binding of N-Hydroxy-2-acetylaminofluorene to Rat Liver Nucleic Acids.*
- [41] F. A. Beland, K. L. Dooley, C. D. Jackson, *Cancer Res.* **1982**, 42, 1348-1354, *Persistence of DNA Adducts in Rat Liver and Kidney after Multiple Doses of the Carcinogen N-Hydroxy-2-acetylaminofluorene.*

- [42] M. C. Poirier, J. M. Hunt, B. A. True, B. A. Laishes, J. F. Young, F. A. Beland, *Carcinogenesis* **1984**, 5, 1591-1596, *DNA adduct formation, removal and persistence in rat liver during one month of feeding 2-acetylaminofluorene*.
- [43] M. C. Poirier, B. A. True, B. A. Laishes, *Environ. Health Perspect.* **1983**, 49, 93-99, *Determination of 2-acetylaminofluorene adducts by immunoassay*.
- [44] M. C. Poirier, B. A. True, B. A. Laishes, *Cancer Res.* **1982**, 42, 1317-1321, *Formation and Removal of (Guan-8-yl)-DNA-2-Acetylaminofluorene Adducts in Liver and Kidney of Male Rats Given Dietary 2-Acetylaminofluorene*.
- [45] W. T. Allaben, C. C. Weis, N. F. Fullerton, F. A. Beland, *Carcinogenesis* **1983**, 4, 1067-1070, *Formation and persistence of DNA adducts from the carcinogen N-hydroxy-2-acetylaminofluorene in rat mammary gland in vivo*.
- [46] X.-S. Cui, L. C. Eriksson, L. Möller, *Mut. Res.-Genet. Toxicol. Environ. Mutagen.* **1999**, 442, 9-18, *Formation and persistence of DNA adducts during and after a long-term administration of 2-nitrofluorene*.
- [47] X. Tan, N. Suzuki, A. P. Grollman, S. Shibutani, *Biochemistry* **2002**, 41, 14255-14262, *Mutagenic Events in Escherichia coli and Mammalian Cells Generated in Response to Acetylaminofluorene-Derived DNA Adducts Positioned in the Nar I Restriction Enzyme Site*.
- [48] S. Schorr, T. Carell, *ChemBioChem* **2010**, 11, 2534-2537, *Mechanism of Acetylaminofluorene-dG Induced Frameshifting by Polymerase η* .
- [49] T. Zaliznyak, R. Bonala, F. Johnson, C. de los Santos, *Chem. Res. Toxicol.* **2006**, 19, 745-752, *Structure and Stability of Duplex DNA Containing the 3-(Deoxyguanosin-N2-yl)-2-acetylaminofluorene (dG(N2)-AAF) Lesion: A Bulky Adduct that Persists in Cellular DNA*.
- [50] H. Mu, K. Kropachev, L. Wang, L. Zhang, A. Kolbanovskiy, M. Kolbanovskiy, N. E. Geacintov, S. Broyde, *Nucleic Acids Res.* **2012**, 40, 9675-9690, *Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies*.
- [51] B. Mao, B. E. Hingerty, S. Broyde, D. J. Patel, *Biochemistry* **1998**, 37, 95-106, *Solution Structure of the Aminofluorene [AF]-External Conformer of the anti-[AF]-C8-dG Adduct Opposite dC in a DNA Duplex*.
- [52] B. P. Cho, F. A. Beland, M. M. Marques, *Biochemistry* **1994**, 33, 1373-1384, *NMR structural studies of a 15-mer duplex from a ras protooncogene modified with the carcinogen 2-aminofluorene: conformational heterogeneity*.
- [53] B. P. Cho, *J. Environ. Sci. Health C Environ. Carcinog. Ecotoxicol. Rev.* **2004**, 22, 57-90, *Dynamic conformational heterogeneities of carcinogen-DNA adducts and their mutagenic relevance*.
- [54] L. Zhou, G. Rajabzadeh, D. D. Traficante, B. P. Cho, *JACS* **1997**, 119, *Conformational heterogeneity of arylamine-modified DNA: 19F NMR evidence*.
- [55] B. Mao, B. E. Hingerty, S. Broyde, D. J. Patel, *Biochemistry* **1998**, 37, 81-94, *Solution Structure of the Aminofluorene [AF]-Intercalated Conformer of the syn-[AF]-C8-dG Adduct Opposite dC in a DNA Duplex*.
- [56] N. Jain, Y. Li, L. Zhang, S. R. Meneni, B. P. Cho, *Biochemistry* **2007**, 46, 13310-13321, *Probing the Sequence Effects on NarI-Induced -2 Frameshift Mutagenesis by Dynamic 19F NMR, UV, and CD Spectroscopy*.
- [57] S. R. Meneni, S. M. Shell, L. Gao, P. Jurecka, W. Lee, J. Sponer, Y. Zou, M. P. Chiarelli, B. P. Cho, *Biochemistry* **2007**, 46, 11263-11278, *Spectroscopic and*

Theoretical Insights into Sequence Effects of Aminofluorene-Induced Conformational Heterogeneity and Nucleotide Excision Repair.

- [58] S. F. O'Handley, D. G. Sanford, R. Xu, C. C. Lester, B. E. Hingerty, S. Broyde, T. R. Krugh, *Biochemistry* **1993**, 32, 2481–2497, *Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization and molecular dynamics.*
- [59] V. Jain, B. Hilton, S. Patnaik, Y. Zou, M. P. Chiarelli, B. P. Cho, *Nucleic Acids Res.* **2012**, 40, 3939-3951, *Conformational and thermodynamic properties modulate the nucleotide excision repair of 2-aminofluorene and 2-acetylaminofluorene dG adducts in the Narl sequence.*
- [60] S. Patnaik, B. P. Cho, *Chem. Res. Toxicol.* **2010**, 23, 1650-1652, *Structures of 2-Acetylaminofluorene Modified DNA Revisited: Insight into Conformational Heterogeneity.*
- [61] D. J. Patel, B. Mao, Z. Gu, B. E. Hingerty, A. Gorin, A. K. Basu, S. Broyde, *Chem. Res. Toxicol.* **1998**, 11, 391-407, *Nuclear Magnetic Resonance Solution Structures of Covalent Aromatic Amine-DNA Adducts and Their Mutagenic Relevance.*
- [62] W. G. Humphreys, F. F. Kadlubar, F. P. Guengerich, *PNAS* **1992**, 89, 8278-8282, *Mechanism of C8 alkylation of guanine residues by activated arylamines: evidence for initial adduct formation at the N7 position.*
- [63] R. A. McClelland, A. Ahmad, A. P. Dicks, V. E. Licence, *JACS* **1999**, 121, 3303-3310, *Spectroscopic Characterization of the Initial C8 Intermediate in the Reaction of the 2-Fluorenylnitrenium Ion with 2'-Deoxyguanosine.*
- [64] R. R. Bonala, M. C. Torres, S. Attaluri, C. R. Iden, F. Johnson, *Chem. Res. Toxicol.* **2005**, 18, 457-465, *Incorporation of N2-Deoxyguanosine Metabolic Adducts of 2-Aminonaphthalene and 2-Aminofluorene into Oligomeric DNA.*
- [65] M. Yasui, H. Dong, R. R. Bonala, N. Suzuki, H. Ohmori, F. Hanaoka, F. Johnson, A. P. Grollman, S. Shibutani, *Biochemistry* **2004**, 43, 15005-15013, *Mutagenic Properties of 3-(Deoxyguanosin-N2-yl)-2-acetylaminofluorene, a Persistent Acetylaminofluorene-Derived DNA Adduct in Mammalian Cells†.*
- [66] R. Grad, R. Shapiro, B. E. Hingerty, S. Broyde, *Chem. Res. Toxicol.* **1997**, 10, 1123-1132, *A Molecular Mechanics and Dynamics Study of the Minor Adduct between DNA and the Carcinogen 2-(Acetylamo)fluorene (dG-N2-AAF).*
- [67] R. Saffhill, P. J. Abbott, *Chem.-Biol. Interact.* **1983**, 44, 95-110, *The formation of acetylaminofluorene adducts in poly(dC-dG) and poly(dA-dT) on reaction with N-acetoxy-2-acetylaminofluorene and the effect of such modification upon the polymers as templates for DNA polymerases.*
- [68] F. A. Beland, D. T. Beranek, K. L. Dooley, R. H. Heflich, F. F. Kadlubar, *Environ. Health Perspect.* **1983**, 49, 125-134, *Arylamine-DNA adducts in vitro and in vivo: their role in bacterial mutagenesis and urinary bladder carcinogenesis.*
- [69] F. F. Kadlubar, L. E. Unruh, F. A. Beland, K. M. Straub, F. E. Evans, *Carcinogenesis* **1980**, 1, 139-150, *In vitro reaction of the carcinogen, N-hydroxy-2-naphthylamine, with DNA at the C-8 and N2 atoms of guanine and at the N6 atom of adenine.*
- [70] D. L. Tillis, K. M. Straub, F. F. Kadlubar, *Chem.-Biol. Interact.* **1981**, 38, 15-27, *A comparison of the carcinogen-DNA adducts formed in rat liver in vivo after administration of single or multiple doses of N-methyl-4-aminoazobenzene.*
- [71] L. Rehn, *Arch. Klin. Chir.* **1895**, 50, 588-600, *Blasengeschwüste bei Fuchsarbeitern. .*

- [72] W. C. Hueper, F. Wiley, H. D. Wolfe, *J. Ind. Hyg. Toxicol.* **1938**, 20, 46, *Experimental production of bladder tumors in dogs by administration of beta-naphthylamine*.
- [73] R. A. M. Case, M. E. Hosker, D. B. McDonald, J. Pearson, *Br. J. Ind. Med.* **1954**, 11, 75-104, *Tumors of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the british chemical industry*.
- [74] F. F. Kadlubar, J. A. Miller, E. C. Miller, *Cancer Res.* **1977**, 37, 805-814, *Hepatic Microsomal N-Glucuronidation and Nucleic Acid Binding of N-Hydroxy Arylamines in Relation to Urinary Bladder Carcinogenesis*.
- [75] J. L. Radomski, E. Brill, W. B. Deichmann, E. M. Glass, *Cancer Res.* **1971**, 31, 1461-1467, *Carcinogenicity Testing of N-Hydroxy and Other Oxidation and Decomposition Products of 1- and 2-Naphthylamine*.
- [76] L. A. Poirier, F. J. de Serres, *J. Natl. Cancer Inst.* **1979**, 62, 919-926, *Initial National Cancer Institute Studies on Mutagenesis as a Prescreen for Chemical Carcinogens: An Appraisal*.
- [77] L. Robbiano, R. Carrozzino, M. Bacigalupo, C. Corbu, G. Brambilla, *Toxicology* **2002**, 179, 115-128, *Correlation between induction of DNA fragmentation in urinary bladder cells from rats and humans and tissue-specific carcinogenic activity*.
- [78] E. Boyland, D. Manson, *Biochem. J.* **1966**, 101, 84-102, *The biochemistry of aromatic amines: The metabolism of 2-naphthylamine and 2-naphthylhydroxylamine derivatives*.
- [79] P. C. Howard, R. H. Heflich, F. E. Evans, F. A. Beland, *Cancer Res.* **1983**, 43, 2052-2058, *Formation of DNA Adducts in Vitro and in Salmonella typhimurium upon Metabolic Reduction of the Environmental Mutagen 1-Nitropyrene*.
- [80] R. H. Heflich, P. C. Howard, F. A. Beland, *Mut. Res.-Fund. Mol. M.* **1985**, 149, 25-32, *1-nitrosopyrene: An intermediate in the metabolic activation of 1-nitropyrene to a mutagen in *Salmonella typhimurium* TA1538*.
- [81] C. A. Stanton, F. L. Chow, D. H. Phillips, P. L. Grover, R. C. Garner, C. N. Martin, *Carcinogenesis* **1985**, 6, 535-538, *Evidence for N-(deoxyguanosin-8-yl)-1-aminopyrene as a major DNA adduct in female rats treated with 1-nitropyrene*.
- [82] B. Mao, R. R. Vyas, B. E. Hingerty, S. Broyde, A. K. Basu, D. J. Patel, *Biochemistry* **1996**, 35, 12659-12670, *Solution conformation of the N-(deoxyguanosin-8-yl)-1-aminopyrene ([AP]dG) adduct opposite dC in a DNA duplex*.
- [83] R. R. Vyas, A. K. Basu, *Carcinogenesis* **1995**, 16, 811-816, *DNA polymerase action on an oligonucleotide containing a site-specifically located N-(deoxyguanosin-8-yl)-1-aminopyrene*.
- [84] S. A. Malia, R. R. Vyas, A. K. Basu, *Biochemistry* **1996**, 35, 4568-4577, *Site-Specific Frame-Shift Mutagenesis by the 1-Nitropyrene-DNA Adduct N-(Deoxyguanosin-8-yl)-1-aminopyrene Located in the (CG)3 Sequence: Effects of SOS, Proofreading, and Mismatch Repair*.
- [85] D. Herreno-Saenz, F. Evans, F. Beland, P. Fu, *Chem. Res. Toxicol.* **1995**, 8, 269-277, *Identification of two N2-deoxyguanosinyl DNA adducts upon nitroreduction of the environmental mutagen 1-nitropyrene*.
- [86] D. B. Dunn, J. D. Smith, *Biochem. J.* **1958**, 68, 627-636, *The occurrence of 6-methylaminopurine in deoxyribonucleic acids*.
- [87] J. Doskočil, Z. Šormová, *Biochim. Biophys. Acta - Nucleic Acids Protein Synth.* **1965**, 95, 513-515, *The occurrence of 5-methylcytosine in bacterial deoxyribonucleic acids*.

- [88] M. Ehrlich, M. A. Gama-Sosa, L. H. Carreira, L. G. Ljungdahl, K. C. Kuo, C. W. Gehrke, *Nucleic Acids Res.* **1985**, 13, 1399-1412, *DNA methylation in thermophilic bacteria: N 4 -methylcytosine, 5-methylcytosine, and N 5 methyladenine.*
- [89] T. Naito, K. Kusano, I. Kobayashi, *Science* **1995**, 267, 897-899, *Selfish behavior of restriction-modification systems.*
- [90] K. Vasu, V. Nagaraja, *Microbiol. Mol. Biol. Rev.* **2013**, 77, 53-72, *Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense.*
- [91] M. A. Gorovsky, S. Hattman, G. L. Pleger, *J. Cell Biol.* **1973**, 56, 697-701, *[(6)N]Methyl Adenine in the Nuclear DNA of a Eucaryote, Tetrahymena Pyriformis.*
- [92] D. J. Cummings, A. Tait, J. M. Goddard, *Biochim. Biophys. Acta* **1974**, 374, 1-11, *Methylated bases in DNA from Paramecium aurelia.*
- [93] Y. Fu, G.-Z. Luo, K. Chen, X. Deng, M. Yu, D. Han, Z. Hao, J. Liu, X. Lu, Louis C. Doré, X. Weng, Q. Ji, L. Mets, C. He, *Cell* **2015**, 161, 879-892, *N6-Methyldeoxyadenosine Marks Active Transcription Start Sites in Chlamydomonas.*
- [94] E. L. Greer, M. A. Blanco, L. Gu, E. Sendinc, J. Liu, D. Aristizábal-Corrales, C.-H. Hsu, L. Aravind, C. He, Y. Shi, *Cell* **2015**, 161, 868-878, *DNA methylation on N(6)-adenine in C. elegans.*
- [95] G. Zhang, H. Huang, D. Liu, Y. Cheng, X. Liu, W. Zhang, R. Yin, D. Zhang, P. Zhang, J. Liu, C. Li, B. Liu, Y. Luo, Y. Zhu, N. Zhang, S. He, C. He, H. Wang, D. Chen, *Cell* **2015**, 161, 893-906, *N6-Methyladenine DNA Modification in Drosophila.*
- [96] J. Bang, S. H. Bae, C. J. Park, J. H. Lee, B. S. Choi, *JACS* **2008**, 130, 17688-17696, *Structural and dynamics study of DNA dodecamer duplexes that contain un-, hemi-, or fully methylated GATC sites.*
- [97] F. Tronche, A. Rollier, I. Bach, M. C. Weiss, M. Yaniv, *Mol. Cell. Biol.* **1989**, 9, 4759-4766, *The rat albumin promoter: cooperation with upstream elements is required when binding of APF/HNF1 to the proximal element is partially impaired by mutation or bacterial methylation.*
- [98] K. Sugimoto, S. Takeda, H. Hirochika, *Plant J.* **2003**, 36, 550-564, *Transcriptional activation mediated by binding of a plant GATA-type zinc finger protein AGP1 to the AG-motif (AGATCCAA) of the wound-inducible Myb gene NtMyb2.*
- [99] S. Lichtsteiner, U. Schibler, *Cell* **1989**, 57, 1179-1187, *A glycosylated liver-specific transcription factor stimulates transcription of the albumin gene.*
- [100] D. Ratel, S. Boisseau, S. M. Davidson, B. Ballester, J. Mathieu, M. Morange, D. Adamski, F. Berger, A. L. Benabid, D. Wion, *Biochem. Biophys. Res. Commun.* **2001**, 285, 800-805, *The bacterial nucleoside N(6)-methyldeoxyadenosine induces the differentiation of mammalian tumor cells.*
- [101] R. van Blokland, S. Ross, G. Corrado, C. Scollan, P. Meyer, *Plant J.* **1998**, 15, 543-551, *Developmental abnormalities associated with deoxyadenosine methylation in transgenic tobacco.*
- [102] M. J. Koziol, C. R. Bradshaw, G. E. Allen, A. S. H. Costa, C. Frezza, J. B. Gurdon, *Nat. Struct. Mol. Biol.* **2016**, 23, 24-30, *Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications.*
- [103] T. P. Wu, T. Wang, M. G. Seetin, Y. Lai, S. Zhu, K. Lin, Y. Liu, S. D. Byrum, S. G. Mackintosh, M. Zhong, A. Tackett, G. Wang, L. S. Hon, G. Fang, J. A. Swenberg, A. Z. Xiao, *Nature* **2016**, 532, 329-333, *DNA methylation on N6-adenine in mammalian embryonic stem cells.*

- [104] J. Liu, Y. Zhu, G.-Z. Luo, X. Wang, Y. Yue, X. Wang, X. Zong, K. Chen, H. Yin, Y. Fu, D. Han, Y. Wang, D. Chen, C. He, *Nature Comm.* **2016**, 7, 13052, *Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig*.
- [105] A. L. Pogolotti, A. Ono, R. Subramaniam, D. V. Santi, *J. Biol. Chem.* **1988**, 263, 7461-7464, *On the mechanism of DNA-adenine methylase*.
- [106] A. Jeltsch, F. Christ, M. Fatemi, M. Roth, *J. Biol. Chem.* **1999**, 274, 19538-19544, *On the Substrate Specificity of DNA Methyltransferases: Adenine-N6 DNA Methyltransferases also Modify Cytosine Residues at Position N4*.
- [107] A. Jeltsch, *bchm* **2001**, 382, 707-710, *The Cytosine N4-Methyltransferase M.Pvull Also Modifies Adenine Residues*.
- [108] A. Jeltsch, *ChemBioChem* **2002**, 3, 274-293, *Beyond Watson and Crick: DNA Methylation and Molecular Enzymology of DNA Methyltransferases*.
- [109] K. Goedecke, M. Pignot, R. S. Goody, A. J. Scheidig, E. Weinhold, *Nat. Struct. Mol. Biol.* **2001**, 8, 121-125, *Structure of the N6-adenine DNA methyltransferase M• Taql in complex with DNA and a cofactor analog*.
- [110] T. Friedrich, M. Roth, S. Helm-Kruse, A. Jeltsch, *bchm* **1998**, 379, 475-480, *Functional Mapping of the EcoRV DNA Methyltransferase by Random Mutagenesis and Screening for Catalytically Inactive Mutants*.
- [111] A. Jeltsch, M. Roth, T. Friedrich, *J. Mol. Biol.* **1999**, 285, 1121-1130, *Mutational Analysis of Target Base Flipping by the EcoRV Adenine-N6 DNA Methyltransferase*.
- [112] B. Holz, N. Dank, J. E. Eickhoff, G. Lipps, G. Krauss, E. Weinhold, *J. Biol. Chem.* **1999**, 274, 15066-15072, *Identification of the Binding Site for the Extrahelical Target Base in N 6-Adenine DNA Methyltransferases by Photo-cross-linking with Duplex Oligodeoxyribonucleotides Containing 5-Iodouracil at the Target Position*.
- [113] G. Schluckebier, J. Labahn, J. Granzin, W. Saenger, *bchm* **1998**, 379, 389-400, *M.Taql: possible catalysis via cation-pi interactions in N-specific DNA methyltransferases*.
- [114] T. B. Johnson, R. D. Coghill, *JACS* **1925**, 47, 2838-2844, *Researches on pyrimidines. C111. The discovery of 5-methyl-cytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus1*.
- [115] G. R. Wyatt, *Nature* **1950**, 166, 237-238, *Occurrence of 5-Methyl-Cytosine in Nucleic Acids*.
- [116] R. Bonasio, S. Tu, D. Reinberg, *Science* **2010**, 330, 612-616, *Molecular Signals of Epigenetic States*.
- [117] H. Cedar, Y. Bergman, *Nat. Rev. Genet.* **2009**, 10, 295-304, *Linking DNA methylation and histone modification: patterns and paradigms*.
- [118] Z. D. Smith, A. Meissner, *Nat. Rev. Genet.* **2013**, 14, 204-220, *DNA methylation: roles in mammalian development*.
- [119] P. A. Jones, D. Takai, *Science* **2001**, 293, 1068-1070, *The Role of DNA Methylation in Mammalian Epigenetics*.
- [120] J. A. Law, S. E. Jacobsen, *Nat. Rev. Genet.* **2010**, 11, 204-220, *Establishing, maintaining and modifying DNA methylation patterns in plants and animals*.
- [121] R. Holliday, *Biol. Rev.* **1990**, 65, 431-471, *Mechanisms for the Control of Gene Activity during Development*.
- [122] A. Portela, M. Esteller, *Nat. Biotech.* **2010**, 28, 1057-1068, *Epigenetic modifications and human disease*.

- [123] T. Bartke, M. Vermeulen, B. Xhemalce, S. C. Robson, M. Mann, T. Kouzarides, *Cell* **2010**, 143, 470-484, *Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation*.
- [124] M. Ehrlich, M. A. Gama-Sosa, L. H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune, C. Gehrke, *Nucleic Acids Res.* **1982**, 10, 2709-2721, *Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells*.
- [125] R. Lister, M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q.-M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren, J. R. Ecker, *Nature* **2009**, 462, 315-322, *Human DNA methylomes at base resolution show widespread epigenomic differences*.
- [126] R. M. Kohli, Y. Zhang, *Nature* **2013**, 502, 472-479, *TET enzymes, TDG and the dynamics of DNA demethylation*.
- [127] P. A. Jones, *Nat. Rev. Genet.* **2012**, 13, 484-492, *Functions of DNA methylation: islands, start sites, gene bodies and beyond*.
- [128] A. Kuroda, T. A. Rauch, I. Todorov, H. T. Ku, I. H. Al-Abdullah, F. Kandeel, Y. Mullen, G. P. Pfeifer, K. Ferreri, *PLOS ONE* **2009**, 4, e6953, *Insulin Gene Expression Is Regulated by DNA Methylation*.
- [129] J. P. Thomson, P. J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A. R. W. Kerr, A. Deaton, R. Andrews, K. D. James, D. J. Turner, R. Illingworth, A. Bird, *Nature* **2010**, 464, 1082-1086, *CpG islands influence chromatin structure via the CpG-binding protein Cfp1*.
- [130] S. Klimasauskas, S. Kumar, R. J. Roberts, X. Cheng, *Cell* **1994**, 76, 357-369, *Hhal methyltransferase flips its target base out of the DNA helix*.
- [131] S. Bheemanaik, Yeturu V. R. Reddy, Desirazu N. Rao, *Biochem. J.* **2006**, 399, 177-190, *Structure, function and mechanism of exocyclic DNA methyltransferases*.
- [132] A. Hermann, H. Gowher, A. Jeltsch, *Cell. Mol. Life Sci.* **2004**, 61, 2571-2587, *Biochemistry and biology of mammalian DNA methyltransferases*.
- [133] R. Z. Jurkowska, T. P. Jurkowski, A. Jeltsch, *ChemBioChem* **2011**, 12, 206-222, *Structure and Function of Mammalian DNA Methyltransferases*.
- [134] M. Okano, E. Li, *J. Nutr.* **2002**, 132, 2462S-2465S, *Genetic Analyses of DNA Methyltransferase Genes in Mouse Model System*.
- [135] M. Okano, D. W. Bell, D. A. Haber, E. Li, *Cell* **1999**, 99, 247-257, *DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development*.
- [136] G. Liang, M. F. Chan, Y. Tomigahara, Y. C. Tsai, F. A. Gonzales, E. Li, P. W. Laird, P. A. Jones, *Mol. Cell. Biol.* **2002**, 22, 480-491, *Cooperativity between DNA Methyltransferases in the Maintenance Methylation of Repetitive Elements*.
- [137] F. Chédin, M. R. Lieber, C.-L. Hsieh, *PNAS* **2002**, 99, 16916-16921, *The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a*.
- [138] K. Blaschke, K. T. Ebata, M. M. Karimi, J. A. Zepeda-Martinez, P. Goyal, S. Mahapatra, A. Tam, D. J. Laird, M. Hirst, A. Rao, M. C. Lorincz, M. Ramalho-Santos, *Nature* **2013**, 500, 222-226, *Vitamin[thiols]C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells*.
- [139] G. Ficz, Timothy A. Hore, F. Santos, Heather J. Lee, W. Dean, J. Arand, F. Krueger, D. Oxley, Y.-L. Paul, J. Walter, Simon J. Cook, S. Andrews, Miguel R. Branco, W. Reik, *Cell Stem Cell* **2013**, 13, 351-359, *FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency*.

- [140] J. Nichols, A. Smith, *Cold Spring Harb. Perspect. Biol.* **2012**, 4, Pluripotency in the Embryo and in Culture.
- [141] S. Feng, S. E. Jacobsen, W. Reik, *Science* **2010**, 330, 622-627, Epigenetic Reprogramming in Plant and Animal Development.
- [142] W. Mayer, A. Niveleau, J. Walter, R. Fundele, T. Haaf, *Nature* **2000**, 403, 501-502, Embryogenesis: Demethylation of the zygotic paternal genome.
- [143] J. Oswald, S. Engemann, N. Lane, W. Mayer, A. Olek, R. Fundele, W. Dean, W. Reik, J. Walter, *Curr. Biol.* **2000**, 10, 475-478, Active demethylation of the paternal genome in the mouse zygote.
- [144] H. Wu, Y. Zhang, *Cell* **2014**, 156, 45-68, Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions.
- [145] K. Iqbal, S.-G. Jin, G. P. Pfeifer, P. E. Szabó, *PNAS* **2011**, 108, 3642-3647, Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine.
- [146] F. Santos, B. Hendrich, W. Reik, W. Dean, *Dev. Biol.* **2002**, 241, 172-182, Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo.
- [147] F. Guo, X. Li, D. Liang, T. Li, P. Zhu, H. Guo, X. Wu, L. Wen, T.-P. Gu, B. Hu, Colum P. Walsh, J. Li, F. Tang, G.-L. Xu, *Cell Stem Cell* **2014**, 15, 447-458, Active and Passive Demethylation of Male and Female Pronuclear DNA in the Mammalian Zygote.
- [148] M. A. Surani, K. Hayashi, P. Hajkova, *Cell* **2007**, 128, 747-762, Genetic and Epigenetic Regulators of Pluripotency.
- [149] K. Martinowich, D. Hattori, H. Wu, S. Fouse, F. He, Y. Hu, G. Fan, Y. E. Sun, *Science* **2003**, 302, 890-893, DNA Methylation-Related Chromatin Remodeling in Activity-Dependent *Bdnf* Gene Regulation.
- [150] D. Bruniquel, R. H. Schwartz, *Nat. Immunol.* **2003**, 4, 235-240, Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process.
- [151] M. Gehring, K. L. Bubb, S. Henikoff, *Science* **2009**, 324, 1447-1451, Extensive Demethylation of Repetitive Elements During Seed Development Underlies Gene Imprinting.
- [152] M. Gehring, J. H. Huh, T.-F. Hsieh, J. Penterman, Y. Choi, J. J. Harada, R. B. Goldberg, R. L. Fischer, *Cell* **2006**, 124, 495-506, DEMETER DNA Glycosylase Establishes MEDEA Polycomb Gene Self-Imprinting by Allele-Specific Demethylation.
- [153] J. Penterman, D. Zilberman, J. H. Huh, T. Ballinger, S. Henikoff, R. L. Fischer, *PNAS* **2007**, 104, 6752-6757, DNA demethylation in the *Arabidopsis* genome.
- [154] A. P. Ortega-Galisteo, T. Morales-Ruiz, R. R. Ariza, T. Roldán-Arjona, *Plant Mol. Biol.* **2008**, 67, 671-681, *Arabidopsis* DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks.
- [155] K. Wiebauer, J. Jiricny, *PNAS* **1990**, 87, 5842-5845, Mismatch-specific thymine DNA glycosylase and DNA polymerase beta mediate the correction of G.T mispairs in nuclear extracts from human cells.
- [156] B. Hendrich, A. Bird, *Mol. Cell. Biol.* **1998**, 18, 6538-6547, Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins.
- [157] A. Bellacosa, A. C. Drohat, *DNA Repair* **2015**, 32, 33-42, Role of base excision repair in maintaining the genetic and epigenetic integrity of CpG sites.
- [158] A. B. Sjolund, A. G. Senejani, J. B. Sweasy, *Mut. Res.-Fund. Mol. M.* **2013**, 743–744, 12-25, MBD4 and TDG: Multifaceted DNA glycosylases with ever expanding biological roles.

- [159] R. Bransteitter, P. Pham, M. D. Scharff, M. F. Goodman, *PNAS* **2003**, *100*, 4102-4107, *Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase*.
- [160] H. D. Morgan, W. Dean, H. A. Coker, W. Reik, S. K. Petersen-Mahrt, *J. Biol. Chem.* **2004**, *279*, 52353-52360, *Activation-induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues: Implications for Epigenetic Reprogramming*.
- [161] K. Rai, I. J. Huggins, S. R. James, A. R. Karpf, D. A. Jones, B. R. Cairns, *Cell* **2008**, *135*, 1201-1212, *DNA Demethylation in Zebrafish Involves the Coupling of a Deaminase, a Glycosylase, and Gadd45*.
- [162] C. Popp, W. Dean, S. Feng, S. J. Cokus, S. Andrews, M. Pellegrini, S. E. Jacobsen, W. Reik, *Nature* **2010**, *463*, 1101-1105, *Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency*.
- [163] J. R. Morrison, C. Pászty, M. E. Stevens, S. D. Hughes, T. Forte, J. Scott, E. M. Rubin, *PNAS* **1996**, *93*, 7154-7159, *Apolipoprotein B RNA editing enzyme-deficient mice are viable despite alterations in lipoprotein metabolism*.
- [164] Y.-Q. Li, P.-Z. Zhou, X.-D. Zheng, C. P. Walsh, G.-L. Xu, *Carcinogenesis* **2007**, *14*, 15, *Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair*.
- [165] R. Metivier, R. Gallais, C. Tiffoche, C. Le Peron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, V. Benes, A. Jeltsch, F. Gannon, G. Salbert, *Nature* **2008**, *452*, 45-50, *Cyclical DNA methylation of a transcriptionally active promoter*.
- [166] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, *Science* **2009**, *324*, 930-935, *Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1*.
- [167] R. J. Boorstein, A. Cummings, D. R. Marenstein, M. K. Chan, Y. Ma, T. A. Neubert, S. M. Brown, G. W. Teebor, *J. Biol. Chem.* **2001**, *276*, 41991-41997, *Definitive Identification of Mammalian 5-Hydroxymethyluracil DNAN-Glycosylase Activity as SMUG1*.
- [168] S. Cortellino, J. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, Lara K. Abramowitz, Marisa S. Bartolomei, F. Rambow, Maria R. Bassi, T. Bruno, M. Fanciulli, C. Renner, Andres J. Klein-Szanto, Y. Matsumoto, D. Kobi, I. Davidson, C. Alberti, L. Larue, A. Bellacosa, *Cell* **2011**, *146*, 67-79, *Thymine DNA Glycosylase Is Essential for Active DNA Demethylation by Linked Deamination-Base Excision Repair*.
- [169] M. C. Hollstein, P. Brooks, S. Linn, B. N. Ames, *PNAS* **1984**, *81*, 4003-4007, *Hydroxymethyluracil DNA glycosylase in mammalian cells*.
- [170] A. Masaoka, M. Matsubara, R. Hasegawa, T. Tanaka, S. Kurisu, H. Terato, Y. Ohyama, N. Karino, A. Matsuda, H. Ide, *Biochemistry* **2003**, *42*, 5003-5012, *Mammalian 5-Formyluracil-DNA Glycosylase. 2. Role of SMUG1 Uracil-DNA Glycosylase in Repair of 5-Formyluracil and Other Oxidized and Deaminated Base Lesions*.
- [171] T. Pfaffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljarova, D. Schuermann, S. Michalakis, O. Kosmatchev, S. Schiesser, B. Steigenberger, N. Raddaoui, G. Kashiwazaki, U. Müller, C. G. Spruijt, M. Vermeulen, H. Leonhardt, P. Schär, M. Müller, T. Carell, *Nat.*

Chem. Biol. **2014**, *10*, 574-581, *Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA*.

- [172] L. Schomacher, D. Han, M. U. Musheev, K. Arab, S. Kienhofer, A. von Seggern, C. Niehrs, *Nat. Struct. Mol. Biol.* **2016**, *23*, 116-124, *Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation*.
- [173] Y.-F. He, B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C.-X. Song, K. Zhang, C. He, G.-L. Xu, *Science* **2011**, *333*, 1303-1307, *Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA*.
- [174] T. Pfaffeneder, B. Hackner, M. Truß, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 7008-7012, *The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA*.
- [175] A. Maiti, A. C. Drohat, *J. Biol. Chem.* **2011**, *286*, 35334-35338, *Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: Potential Implications for Active Demethylation of CpG Sites*.
- [176] L. Zhang, X. Lu, J. Lu, H. Liang, Q. Dai, G.-L. Xu, C. Luo, H. Jiang, C. He, *Nat. Chem. Biol.* **2012**, *8*, 328-330, *Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA*.
- [177] A. Maiti, A. Z. Michelson, C. J. Armwood, J. K. Lee, A. C. Drohat, *JACS* **2013**, *135*, 15813-15822, *Divergent Mechanisms for Enzymatic Excision of 5-Formylcytosine and 5-Carboxylcytosine from DNA*.
- [178] A. C. Drohat, A. Maiti, *Org. Biomol. Chem.* **2014**, *12*, 8367-8378, *Mechanisms for enzymatic cleavage of the N-glycosidic bond in DNA*.
- [179] Z. Liutkevičiūtė, G. Lukinavicius, V. Masevicius, D. Daujotyte, S. Klimasauskas, *Nat. Chem. Biol.* **2009**, *5*, 400-402, *Cytosine-5-methyltransferases add aldehydes to DNA*.
- [180] S. Schiesser, T. Pfaffeneder, K. Sadeghian, B. Hackner, B. Steigenberger, A. S. Schröder, J. Steinbacher, G. Kashiwazaki, G. Höfner, K. T. Wanner, C. Ochsenfeld, T. Carell, *JACS* **2013**, *135*, 14593-14599, *Deamination, Oxidation, and C–C Bond Cleavage Reactivity of 5-Hydroxymethylcytosine, 5-Formylcytosine, and 5-Carboxylcytosine*.
- [181] S. Schiesser, B. Hackner, T. Pfaffeneder, M. Müller, C. Hagemeier, M. Truss, T. Carell, *Angew. Chem. Int. Ed.* **2012**, *51*, 6516-6520, *Mechanism and Stem-Cell Activity of 5-Carboxycytosine Decarboxylation Determined by Isotope Tracing*.
- [182] Z. Liutkevičiūtė, E. Kriukienė, J. Ličytė, M. Rudytė, G. Urbanavičiūtė, S. Klimašauskas, *JACS* **2014**, *136*, 5884-5887, *Direct Decarboxylation of 5-Carboxylcytosine by DNA C5- Methyltransferases*.
- [183] A. S. Schröder, *Aufklärung des Stoffwechselwegs von 5-Methyl-desoxycytidin mittels Synthese innovativ-markierter Desoxycytidin-Derivate und Entwicklung einer reversiblen Proteinmarkierung*, Dissertation, LMU (München), **2016**.
- [184] A. S. Schröder, E. Parsa, K. Iwan, F. R. Traube, M. Wallner, S. Serdjukow, T. Carell, *Chem. Comm.* **2016**, *52*, 14361-14364, *2[prime or minute]-*(R*)-Fluorinated mC, hmC, fC and caC triphosphates are substrates for DNA polymerases and TET-enzymes*.
- [185] S. C. Wu, Y. Zhang, *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 607-620, *Active DNA demethylation: many roads lead to Rome*.
- [186] S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eisen, S. Geiger, S. Schneider, C. Kisker, T. Carell, *PNAS* **2015**, *112*, 8272-8277, *Structural insights into the recognition of cisplatin and AAF-dG lesion by Rad14 (XPA)*.

- [187] J. Q. Svejstrup, *Nat. Rev. Mol. Cell Biol.* **2002**, 3, 21-29, *Mechanisms of transcription-coupled DNA repair*.
- [188] P. C. Hanawalt, G. Spivak, *Nat. Rev. Mol. Cell Biol.* **2008**, 9, 958-970, *Transcription-coupled DNA repair: two decades of progress and surprises*.
- [189] S. Lagerwerf, M. G. Vrouwe, R. M. Overmeer, M. I. Fousteri, L. H. F. Mullenders, *DNA Repair* **2011**, 10, 743-750, *DNA damage response and transcription*.
- [190] D. Gunz, M. T. Hess, H. Naegeli, *J. Biol. Chem.* **1996**, 271, 25089-25098, *Recognition of DNA Adducts by Human Nucleotide Excision Repair: Evidence for a Thermodynamic Probing Mechanism*.
- [191] M. T. Hess, U. Schwitter, M. Petretta, B. Giese, H. Naegeli, *PNAS* **1997**, 94, 6664-6669, *Bipartite substrate discrimination by human nucleotide excision repair*.
- [192] N. E. Geacintov, S. Broyde, T. Buterin, H. Naegeli, M. Wu, S. Yan, D. J. Patel, *Biopolymers* **2002**, 65, 202-210, *Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery*.
- [193] K. Sugasawa, J. M. Y. Ng, C. Masutani, S. Iwai, P. J. van der Spek, A. P. M. Eker, F. Hanaoka, D. Bootsma, J. H. J. Hoeijmakers, *Mol. Cell* **1998**, 2, 223-232, *Xeroderma Pigmentosum Group C Protein Complex Is the Initiator of Global Genome Nucleotide Excision Repair*.
- [194] K. Sugasawa, Y. Okuda, M. Saijo, R. Nishi, N. Matsuda, G. Chu, T. Mori, S. Iwai, K. Tanaka, K. Tanaka, F. Hanaoka, *Cell* **2005**, 121, 387-400, *UV-Induced Ubiquitylation of XPC Protein Mediated by UV-DDB-Ubiquitin Ligase Complex*.
- [195] M. E. Fitch, S. Nakajima, A. Yasui, J. M. Ford, *J. Biol. Chem.* **2003**, 278, 46906-46910, *In Vivo Recruitment of XPC to UV-induced Cyclobutane Pyrimidine Dimers by the DDB2 Gene Product*.
- [196] Q.-E. Wang, Q. Zhu, G. Wani, J. Chen, A. A. Wani, *Carcinogenesis* **2004**, 25, 1033-1043, *UV radiation-induced XPC translocation within chromatin is mediated by damaged-DNA binding protein, DDB2*.
- [197] O. D. Schärer, *Cold Spring Harb. Persp. Biol.* **2013**, 5, a012609, *Nucleotide Excision Repair in Eukaryotes*.
- [198] C.-L. Li, Filip M. Golebiowski, Y. Onishi, Nadine L. Samara, K. Sugasawa, W. Yang, *Mol. Cell* **2015**, 59, 1025-1034, *Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair*.
- [199] N. Mathieu, N. Kaczmarek, P. Rüthemann, A. Luch, H. Naegeli, *Curr. Biol.* **2013**, 23, 204-212, *DNA Quality Control by a Lesion Sensor Pocket of the Xeroderma Pigmentosum Group D Helicase Subunit of TFIIH*.
- [200] H. Naegeli, K. Sugasawa, *DNA Repair* **2011**, 10, 673-683, *The xeroderma pigmentosum pathway: Decision tree analysis of DNA quality*.
- [201] A. Tapias, J. Auriol, D. Forget, J. H. Enzlin, O. D. Schärer, F. Coin, B. Coulombe, J.-M. Egly, *J. Biol. Chem.* **2004**, 279, 19074-19083, *Ordered Conformational Changes in Damaged DNA Induced by Nucleotide Excision Repair Factors*.
- [202] T. Riedl, F. Hanaoka, J. M. Egly, *EMBO J.* **2003**, 22, 5293-5303, *The comings and goings of nucleotide excision repair factors on damaged DNA*.
- [203] M. Wakasugi, A. Sancar, *PNAS* **1998**, 95, 6669-6674, *Assembly, subunit composition, and footprint of human DNA repair excision nuclease*.
- [204] G. W. Buchko, G. W. Daughdrill, R. de Lorimier, N. G. Isern, J. M. Lingbeck, J.-S. Taylor, M. S. Wold, M. Gochin, L. D. Spicer, D. F. Lowry, M. A. Kennedy, *Biochemistry* **1999**, 38, 15116-15128, *Interactions of Human Nucleotide Excision*

Repair Protein XPA with DNA and RPA70ΔC327: Chemical Shift Mapping and 15N NMR Relaxation Studies.

- [205] L. Li, C. A. Peterson, X. Lu, R. J. Legerski, *Mol. Cell. Biol.* **1995**, 15, 1993-1998, *Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair.*
- [206] L. Li, X. Lu, C. A. Peterson, R. J. Legerski, *Mol. Cell. Biol.* **1995**, 15, 5396-5402, *An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair.*
- [207] C. H. Park, A. Sancar, *PNAS* **1994**, 91, 5017-5021, *Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins.*
- [208] K. Tripsianes, G. E. Folkers, C. Zheng, D. Das, J. S. Grinsteadi, R. Kaptein, R. Boelens, *Nucleic Acids Res.* **2007**, 35, 5789-5798, *Analysis of the XPA and ssDNA-binding surfaces on the central domain of human ERCC1 reveals evidence for subfunctionalization.*
- [209] O. V. Tsodikov, D. Ivanov, B. Orelli, L. Staresincic, I. Shoshani, R. Oberman, O. D. Schärer, G. Wagner, T. Ellenberger, *EMBO J.* **2007**, 26, 4768-4776, *Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA.*
- [210] J.-S. You, M. Wang, S.-H. Lee, *J. Biol. Chem.* **2003**, 278, 7476-7485, *Biochemical Analysis of the Damage Recognition Process in Nucleotide Excision Repair.*
- [211] Y. S. Krasikova, N. I. Rechkunova, E. A. Maltseva, I. O. Petruseva, O. I. Lavrik, *Nucleic Acids Res.* **2010**, 38, 8083-8094, *Localization of xeroderma pigmentosum group A protein and replication protein A on damaged DNA in nucleotide excision repair.*
- [212] M. Missura, T. Buterin, R. Hindges, U. Hübscher, J. Kaspáková, V. Brabec, H. Naegeli, *EMBO J.* **2001**, 20, 3554-3564, *Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair.*
- [213] L. M. Iakoucheva, A. L. Kimzey, C. D. Masselon, R. D. Smith, A. K. Dunker, E. J. Ackerman, *Protein Science* **2001**, 10, 1353-1362, *Aberrant mobility phenomena of the DNA repair protein XPA.*
- [214] J. E. Cleaver, J. C. States, *Biochem. J.* **1997**, 328, 1-12, *The DNA damage-recognition problem in human and other eukaryotic cells: the XPA damage binding protein.*
- [215] T. Ikegami, I. Kuraoka, M. Saijo, N. Kodo, Y. Kyogoku, K. Morikawa, K. Tanaka, M. Shirakawa, *J. Biochem.* **1999**, 125, 495-506, *Resonance Assignments, Solution Structure, and Backbone Dynamics of the DNA- and RPA-Binding Domain of Human Repair Factor XPA.*
- [216] S. Rademakers, M. Volker, D. Hoogstraten, A. L. Nigg, M. J. Moné, A. A. van Zeeland, J. H. J. Hoeijmakers, A. B. Houtsmailler, W. Vermeulen, *Mol. Cell. Biol.* **2003**, 23, 5755-5767, *Xeroderma Pigmentosum Group A Protein Loads as a Separate Factor onto DNA Lesions.*
- [217] M. Volker, M. J. Moné, P. Karmakar, A. van Hoffen, W. Schul, W. Vermeulen, J. H. J. Hoeijmakers, R. van Driel, A. A. van Zeeland, L. H. F. Mullenders, *Mol. Cell* **2001**, 8, 213-224, *Sequential Assembly of the Nucleotide Excision Repair Factors In Vivo.*
- [218] W. L. de Laat, E. Appeldoorn, K. Sugawara, E. Weterings, N. G. J. Jaspers, J. H. J. Hoeijmakers, *Genes & Development* **1998**, 12, 2598-2609, *DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair.*
- [219] I. L. Hermanson-Miller, J. J. Turchi, *Biochemistry* **2002**, 41, 2402-2408, *Strand-Specific Binding of RPA and XPA to Damaged Duplex DNA†.*

- [220] W. L. de Laat, E. Appeldoorn, N. G. J. Jaspers, J. H. J. Hoeijmakers, *J. Biol. Chem.* **1998**, 273, 7835-7842, *DNA Structural Elements Required for ERCC1-XPF Endonuclease Activity*.
- [221] T. Matsunaga, C.-H. Park, T. Bessho, D. Mu, A. Sancar, *J. Biol. Chem.* **1996**, 271, 11047-11050, *Replication Protein A Confers Structure-specific Endonuclease Activities to the XPF-ERCC1 and XPG Subunits of Human DNA Repair Excision Nuclease*.
- [222] T. Bessho, A. Sancar, L. H. Thompson, M. P. Thelen, *J. Biol. Chem.* **1997**, 272, 3833-3837, *Reconstitution of Human Excision Nuclease with Recombinant XPF-ERCC1 Complex*.
- [223] A. O'Donovan, A. A. Davies, J. G. Moggs, S. C. West, R. D. Wood, *Nature* **1994**, 371, 432-435, *XPG endonuclease makes the 3[prime] incision in human DNA nucleotide excision repair*.
- [224] A. M. Sijbers, W. L. de Laat, R. R. Ariza, M. Biggerstaff, Y.-F. Wei, J. G. Moggs, K. C. Carter, B. K. Shell, E. Evans, M. C. de Jong, S. Rademakers, J. de Rooij, N. G. J. Jaspers, J. H. J. Hoeijmakers, R. D. Wood, *Cell* **1996**, 86, 811-822, *Xeroderma Pigmentosum Group F Caused by a Defect in a Structure-Specific DNA Repair Endonuclease*.
- [225] A. v. Hoffen, J. Venema, R. Meschini, A. A. v. Zeeland, L. H. Mullenders, *Embo J.* **1995**, 14, 360-367, *Transcription-coupled repair removes both cyclobutane pyrimidine dimers and 6-4 photoproducts with equal efficiency and in a sequential way from transcribed DNA in xeroderma pigmentosum group C fibroblasts*.
- [226] P. Hanawalt, I. Mellon, *Curr. Biol.* **1993**, 3, 67-69, *Stranded in an active gene*.
- [227] E. C. Friedberg, *Annu. Rev. Biochem.* **1996**, 65, 15-42, *Relationships Between DNA Repair and Transcription*.
- [228] B. A. Donahue, S. Yin, J. S. Taylor, D. Reines, P. C. Hanawalt, *PNAS* **1994**, 91, 8502-8506, *Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template*.
- [229] S. Tornaletti, D. Reines, P. C. Hanawalt, *J. Biol. Chem.* **1999**, 274, 24124-24130, *Structural Characterization of RNA Polymerase II Complexes Arrested by a Cyclobutane Pyrimidine Dimer in the Transcribed Strand of Template DNA*.
- [230] E. Citterio, V. Van Den Boom, G. Schnitzler, R. Kanaar, E. Bonte, R. E. Kingston, J. H. J. Hoeijmakers, W. Vermeulen, *Mol. Cell. Biol.* **2000**, 20, 7643-7653, *ATP-Dependent Chromatin Remodeling by the Cockayne Syndrome B DNA Repair-Transcription-Coupling Factor*.
- [231] D. Tantin, A. Kansal, M. Carey, *Molecular and Cellular Biology* **1997**, 17, 6803-6814, *Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes*.
- [232] S. Kamiuchi, M. Saijo, E. Citterio, M. de Jager, J. H. J. Hoeijmakers, K. Tanaka, *PNAS* **2002**, 99, 201-206, *Translocation of Cockayne syndrome group A protein to the nuclear matrix: Possible relevance to transcription-coupled DNA repair*.
- [233] M. Fousteri, W. Vermeulen, A. A. van Zeeland, L. H. F. Mullenders, *Mol. Cell* **2006**, 23, 471-482, *Cockayne Syndrome A and B Proteins Differentially Regulate Recruitment of Chromatin Remodeling and Repair Factors to Stalled RNA Polymerase II In Vivo*.
- [234] M. Saijo, *Mech. Ageing Dev.* **2013**, 134, 196-201, *The role of Cockayne syndrome group A (CSA) protein in transcription-coupled nucleotide excision repair*.

- [235] P. Koehl, P. Valladier, J.-F. Lefèvre, R. P. P. Fuchs, *Nucleic Acids Res.* **1989**, 17, 9531-9541, *Strong structural effect of the position of a single acetylaminofluorene adduct within a mutation hot spot.*
- [236] R. P. P. Fuchs, N. Schwartz, M. P. Daune, *Nature* **1981**, 294, 657-659, *Hotspots of frameshift mutations induced by the ultimate carcinogen N-acetoxy-N-2-acetylaminofluorene.*
- [237] N. Koffel-Schwartz, J.-M. Verdier, M. Bichara, A.-M. Freund, M. P. Daune, R. P. P. Fuchs, *J. Mol. Biol.* **1984**, 177, 33-51, *Carcinogen-induced mutation spectrum in wild-type, uvrA and umuC strains of Escherichia coli.*
- [238] N. Koffel-Schwartz, R. P. P. Fuchs, *J. Mol. Biol.* **1995**, 252, *Sequence determinants for -2 frameshift mutagenesis at NarI-derived hot spots.*
- [239] R. Janel-Bintz, J. Wagner, L. Haracska, M. C. M. Mah-Becherel, M. Bichara, R. P. Fuchs, A. M. Cordonnier, *PLOS ONE* **2012**, 7, e36004, *Evidence for a Rad18-Independent Frameshift Mutagenesis Pathway in Human Cell-Free Extracts.*
- [240] S. Hoare, Y. Zou, V. Purohit, R. Krishnasamy, M. Skorvaga, B. Van Houten, N. E. Geacintov, A. K. Basu, *Biochemistry* **2000**, 39, 12252-12261, *Differential Incision of Bulky Carcinogen-DNA Adducts by the UvrABC Nuclease: Comparison of Incision Rates and the Interactions of Uvr Subunits with Lesions of Different Structures.*
- [241] D. A. Reeves, H. Mu, K. Kropachev, Y. Cai, S. Ding, A. Kolbanovskiy, M. Kolbanovskiy, Y. Chen, J. Krzeminski, S. Amin, D. J. Patel, S. Broyde, N. E. Geacintov, *Nucleic Acids Res.* **2011**, 39, 8752-8764, *Resistance of bulky DNA lesions to nucleotide excision repair can result from extensive aromatic lesion-base stacking interactions.*
- [242] Y. Cai, K. Kropachev, M. Kolbanovskiy, A. Kolbanovskiy, S. Broyde, D. J. Patel, N. E. Geacintov, in *The Chemical Biology of DNA Damage*, Wiley-VCH Verlag GmbH & Co. KGaA, **2010**, pp. 261-298.
- [243] Y. Cai, N. E. Geacintov, S. Broyde, *Biochemistry* **2012**, 51, 1486-1499, *Nucleotide Excision Repair Efficiencies of Bulky Carcinogen-DNA Adducts Are Governed by a Balance between Stabilizing and Destabilizing Interactions.*
- [244] M. T. Hess, D. Gunz, N. Luneva, N. E. Geacintov, H. Naegeli, *Mol. Cell. Biol.* **1997**, 17, 7069-7076, *Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes.*
- [245] D. Mu, E. Bertrand-Burggraf, J. C. Huang, R. P. Fuchs, A. Sancar, B. P. Fuchs, *Nucleic Acids Res.* **1994**, 22, 4869-4871, *Human and E.coli excinucleases are affected differently by the sequence context of acetylaminofluorene-guanine adduct.*
- [246] N. Kitsera, K. Gasteiger, B. Lühnsdorf, J. Allgayer, B. Epe, T. Carell, A. Khobta, *PLOS ONE* **2014**, 9, e94405, *Cockayne Syndrome: Varied Requirement of Transcription-Coupled Nucleotide Excision Repair for the Removal of Three Structurally Different Adducts from Transcribed DNA.*
- [247] M. F. van Oosterwijk, A. Versteeg, R. Filon, A. A. van Zeeland, L. H. Mullenders, *Mol. Cell. Biol.* **1996**, 16, 4436-4444, *The sensitivity of Cockayne's syndrome cells to DNA-damaging agents is not due to defective transcription-coupled repair of active genes.*
- [248] C. J. Jones, R. D. Wood, *Biochemistry* **1993**, 32, 12096-12104, *Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA.*
- [249] P. Robins, C. J. Jones, M. Biggerstaff, T. Lindahl, R. D. Wood, *EMBO J.* **1991**, 10, 3913-3921, *Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA.*

- [250] T. Ikegami, I. Kuraoka, M. Saijo, N. Kodo, Y. Kyogoku, K. Morikawa, K. Tanaka, M. Shirakawa, *Nat. Struct. Mol. Biol.* **1998**, 5, 701-706, *Solution structure of the DNA-and RPA-binding domain of the human repair factor XPA*.
- [251] G. W. Buchko, S. Ni, B. D. Thrall, M. A. Kennedy, *Nucleic Acids Res.* **1998**, 26, 2779-2788, *Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA*.
- [252] S. Wagaw, R. A. Rennels, S. L. Buchwald, *JACS* **1997**, 119, 8451-8458, *Palladium-Catalyzed Coupling of Optically Active Amines with Aryl Bromides*.
- [253] J. P. Wolfe, S. L. Buchwald, *J. Org. Chem.* **2000**, 65, 1144-1157, *Scope and Limitations of the Pd/BINAP-Catalyzed Amination of Aryl Bromides*.
- [254] J. P. Wolfe, S. Wagaw, S. L. Buchwald, *JACS* **1996**, 118, 7215-7216, *An Improved Catalyst System for Aromatic Carbon–Nitrogen Bond Formation: The Possible Involvement of Bis(Phosphine) Palladium Complexes as Key Intermediates*.
- [255] J. F. Hartwig, S. Richards, D. Barañano, F. Paul, *JACS* **1996**, 118, 3626-3633, *Influences on the Relative Rates for C–N Bond-Forming Reductive Elimination and β-Hydrogen Elimination of Amides. A Case Study on the Origins of Competing Reduction in the Palladium-Catalyzed Amination of Aryl Halides*.
- [256] D. S. Surry, S. L. Buchwald, *Chem. Sci.* **2011**, 2, 27-50, *Dialkylbiaryl phosphines in Pd-catalyzed amination: a user's guide*.
- [257] T. Takamura-Enya, S. Enomoto, K. Wakabayashi, *J. Org. Chem.* **2006**, 71, 5599-5606, *Palladium-Catalyzed Direct N-Arylation of Nucleosides, Nucleotides, and Oligonucleotides for Efficient Preparation of dG-N2 Adducts with Carcinogenic Amino-/Nitroarenes*.
- [258] S. M. Mali, R. D. Bhaisare, H. N. Gopi, *J. Org. Chem.* **2013**, 78, 5550-5555, *Thioacids Mediated Selective and Mild N-Acylation of Amines*.
- [259] S. M. Mali, S. V. Jadhav, H. N. Gopi, *Chem. Comm.* **2012**, 48, 7085-7087, *Copper(II) Mediated Facile and Ultra Fast PeptideSynthesis in Methanol*.
- [260] E. B. Merkushev, *Synthesis* **1988**, 1988, 923-937, *Advances in the Synthesis of Iodoaromatic Compounds*.
- [261] A. Klapars, S. L. Buchwald, *JACS* **2002**, 124, 14844-14845, *Copper-Catalyzed Halogen Exchange in Aryl Halides: An Aromatic Finkelstein Reaction*.
- [262] S. Schann, S. Mayer, L. Foucher, *Vol. EP 1777219 A1 (Ed.: E. P. Appl.)*, **2007**.
- [263] Z. Szombati, S. Baerns, A. Marx, C. Meier, *ChemBioChem* **2012**, 13, 700-712, *Synthesis of C8-Arylamine-Modified 2'-Deoxyadenosine Phosphoramidites and their Site-Specific Incorporation into Oligonucleotides*.
- [264] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, 22, 1859-1862, *Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis*.
- [265] J.-E. Yeo, A. Khoo, A. F. Fagbemi, O. D. Schärer, *Chem. Res. Toxicol.* **2012**, 25, 2462-2468, *The Efficiencies of Damage Recognition and Excision Correlate with Duplex Destabilization Induced by Acetylaminofluorene Adducts in Human Nucleotide Excision Repair*.
- [266] B. Mao, Z. Gu, B. E. Hingerty, S. Broyde, D. J. Patel, *Biochemistry* **1998**, 37, 94, *Solution structure of the aminofluorene [AF]-intercalated conformer of the syn [AF]-C8-dG adduct opposite dC in a DNA duplex*.
- [267] P. Belguise-Valladier, R. P. P. Fuchs, *J. Mol. Biol.* **1995**, 249, *N-2-aminofluorene and N-2-acetylaminofluorene adducts: the local sequence context of an adduct and its chemical structure determine its replication properties*.

- [268] S. F. O'Handley, D. G. Sanford, R. Xu, C. C. Lester, B. E. Hingerty, S. Broyde, T. R. Krugh, *Biochemistry* **1993**, 32, 2481-2497, *Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics.*
- [269] D. Globisch, M. Münz, M. Müller, S. Michalakis, M. Wagner, S. Koch, T. Brückl, M. Biel, T. Carell, *PLOS ONE* **2010**, 5, e15367, *Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates.*
- [270] M. Münz, D. Globisch, T. Brückl, M. Wagner, V. Welzmiller, S. Michalakis, M. Müller, M. Biel, T. Carell, *Angew. Chem. Int. Ed.* **2010**, 49, 5375-5377, *Quantification of the Sixth DNA Base Hydroxymethylcytosine in the Brain.*
- [271] M. Sargent, ISBN 978-0-948926-27-3, **2013**.
- [272] R. K. Hill, N. D. Ledford, L. A. Renbaum, *J. Label. Compd. Rad.* **1985**, 22, 143-148, *Convenient preparation of deuterated uracils and dihydrouracils.*
- [273] J. Thiem, H. Mohn, A. Heesing, *Synthesis* **1985**, 1985, 775-778, *Anwendung der Perdeuteroallyl Schutz-Gruppe.*
- [274] R. J. De Pasquale, *J. Org. Chem.* **1977**, 42, 2185-2187, *Preparation of uracil.*
- [275] M. Hoffer, R. Duschinsky, J. J. Fox, N. Yung, *JACS* **1959**, 81, 4112-4113, *Simple Synthesis of Pyrimidine-2'-Deoxy-Ribonucleosides.*
- [276] H. Aoyama, *Bull. Chem. Soc. Jpn.* **1987**, 60, 2073-2077, *Stereoselective Synthesis of Anomers of 5-Substituted 2'-Deoxyuridines.*
- [277] T. Brückl, I. Thoma, A. J. Wagner, P. Knochel, T. Carell, *Eur. J. Org. Chem.* **2010**, 2010, 6517-6519, *Efficient Synthesis of Deazaguanosine-Derived tRNA Nucleosides PreQ0, PreQ1, and Archaeosine Using the Turbo-Grignard Method.*
- [278] P. Knochel, W. Dohle, N. Gommermann, F. F. Kneisel, F. Kopp, T. Korn, I. Sapountzis, V. A. Vu, *Angew. Chem. Int. Ed.* **2003**, 42, 4302-4320, *Highly Functionalized Organomagnesium Reagents Prepared through Halogen-Metal Exchange.*
- [279] P. Knochel, *OC Skript 4*, **2017**.

Anhang

S. Koch, N. Simon, C. Ebert, T. Carell, Q. *Rev. Biophys.* **2016**, 49, e5, *Molecular mechanisms of xeroderma pigmentosum (XP) proteins*.

**REVIEW**

Molecular mechanisms of xeroderma pigmentosum (XP) proteins

Sandra C. Koch, Nina Simon, Charlotte Ebert and Thomas Carell*

Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians Universität München, Butenandtstr. 5-13, 81377 Munich, Germany

Quarterly Reviews of Biophysics (2016), 49, e5, pages 1 of 32 doi:10.1017/S0033583515000268

Abstract. Nucleotide excision repair (NER) is a highly versatile and efficient DNA repair process, which is responsible for the removal of a large number of structurally diverse DNA lesions. Its extreme broad substrate specificity ranges from DNA damages formed upon exposure to ultraviolet radiation to numerous bulky DNA adducts induced by mutagenic environmental chemicals and cytotoxic drugs used in chemotherapy. Defective NER leads to serious diseases, such as xeroderma pigmentosum (XP). Eight XP complementation groups are known of which seven (XPA–XPG) are caused by mutations in genes involved in the NER process. The eighth gene, XPV, codes for the DNA polymerase η , which replicates through DNA lesions in a process called translesion synthesis (TLS). Over the past decade, detailed structural information of these DNA repair proteins involved in eukaryotic NER and TLS have emerged. These structures allow us now to understand the molecular mechanism of the NER and TLS processes in quite some detail and we have begun to understand the broad substrate specificity of NER. In this review, we aim to highlight recent advances in the process of damage recognition and repair as well as damage tolerance by the XP proteins.

Key words: NER, xeroderma pigmentosum, crystal structures, DNA repair.

1. Introduction	2
1.1. DNA damages	2
1.2. NER and NER-related diseases	2
1.3. NER mechanism	3
1.4. DNA-structure/repair relationships	4
1.5. Damage recognition in NER	4
1.6. Overview of each XP protein and its role in GG-NER	5
2. XPE (UV-DDB2)	6
2.1. Crystal structure of human XPE with 6–4PP damaged dsDNA	7
2.2. Crystal structure of human XPE with CPD damaged dsDNA	7
2.3. Crystal structure of human XPE with an abasic lesion site in dsDNA	7
3. XPC	7
3.1. Crystal structure of yeast XPC (Rad4) in complex with a CPD lesion	9
4. XPB	11
4.1. Crystal structure of <i>Archaeoglobus fulgidus</i> XPB (AfXPB)	11
4.2. Crystal structure of human XPB	12
5. XPD	12
5.1. Crystal structure of <i>Sulfolobus acidocaldarius</i> XPD (SaXPD)	13

* Authors for correspondence: Thomas Carell, Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians Universität München, Butenandtstr. 5-13, 81377 Munich, Germany. Tel.: +49 (0)89 2180 77755 Email: thomas.carell@cup.uni-muenchen.de

5.2. Crystal structure of <i>T. acidophilum</i> (TaxPD) in complex with ssDNA	13
5.3. Proposed model of damage verification	14
6. XPA	15
6.1. Nuclear magnetic resonance (NMR) structure of the DBD of human XPA	15
6.2. Crystal structures of Rad14 with AAF-dG and cisplatin containing dsDNA	17
7. XPF–ERCC1	17
7.1. Crystal structure of human XPF–ERCC1	18
7.2. NMR structure of C-terminal human XPF in complex with ssDNA	18
8. XPG	19
8.1. Crystal structure of the yeast XPG (Rad2) in complex with dsDNA	19
9. XPV	21
9.1. Crystal structure of the yeast XPV (Rad30) protein	21
9.2. Crystal structure of the yeast XPV (Rad30) in complex with 1,2-GG cisplatin	21
9.3. Crystal structure of the yeast XPV (Rad30) in complex with the CPD lesion	21
9.4. Crystal structure of the human XPV in complex with the CPD lesion	22
10. Conclusion	23
Acknowledgements	24
References	24

1. Introduction

1.1 DNA damages

The DNA is attacked daily by ultraviolet (UV) radiation, chemical carcinogens, cellular metabolites as well as anti-cancer drugs such as cisplatin (Hoeijmakers, 2001) leading to more than 50 000 damages per cell per day (Rajska *et al.* 2000). To preserve the integrity of the genetic information, all organisms are equipped with a number of DNA repair or tolerance mechanisms such as DNA recombination, base excision repair (BER), mismatch repair (MMR), nucleotide excision repair (NER) and translesion synthesis (TLS) that allow cells to cope with the different lesions (Friedberg, 2005; Ohmori *et al.* 2001; Rupp & Howard-Flanders, 1968). In this review, we will focus on eukaryotic NER and TLS.

1.2 NER and NER-related diseases

One of the most prominent processes among the DNA repair pathways is the NER. NER is a central pathway safeguarding the genome and cells against induced carcinogenesis, because of its capacity to eliminate a broad range of structurally very different DNA lesions (de Laat *et al.* 1999; Geacintov *et al.* 2002; Gillet & Scharer, 2006; Gunz *et al.* 1996; Huang *et al.* 1994) including UV lesions such as cyclobutane–pyrimidine dimers (CPDs) and 6–4 pyrimidine–pyrimidone photoproducts (6-4PPs), numerous bulky chemical adducts, intrastrand crosslinks caused by chemotherapeutics such as cisplatin and oxidative damages (Fig. 1) (Brooks *et al.* 2000; Demple & Harrison, 1994). All these lesions do not share the same chemical structures, but they are bulky and they thermodynamically destabilize the DNA helix, except for CPDs.

In humans, mutations of NER components lead to three severe diseases: xeroderma pigmentosum (XP), Cockayne syndrome (CS) and trichothiodystrophy (TTD) (Cleaver, 2000; Lehmann, 2001). The NER pathway is divided in two subpathways: the global genome NER (GG-NER) and the transcription-coupled NER (TC-NER) (Gillet & Scharer, 2006; Hanawalt & Spivak, 2008) differing only in their damage recognition step. But NER in chromatin is a complex process and probably uses the proposed ‘access–repair–restore’ mechanism, which comprises many components (Scharer, 2013). A defect in TC-NER along with a mild transcription effect can cause CS and TTD (Cleaver *et al.* 2009).

The CS is characterized by developmental impairment, premature aging and sunlight sensitivity. TTD leads to developmental impairment and mental retardation without showing skin cancer predisposition (Cleaver *et al.* 2009; Lehmann, 2003). Mutations inactivating the GG-NER cause the genetic disorder XP (Cleaver *et al.* 2009). Persons affected by XP ('children of the moon') need to avoid exposure to sunlight, because they show a high predisposition to UV-induced skin cancer (Berneburg & Lehmann, 2001; Hebra & Kaposi, 1874; Lehmann, 2003). Seven complementation groups (XPA–XPG)

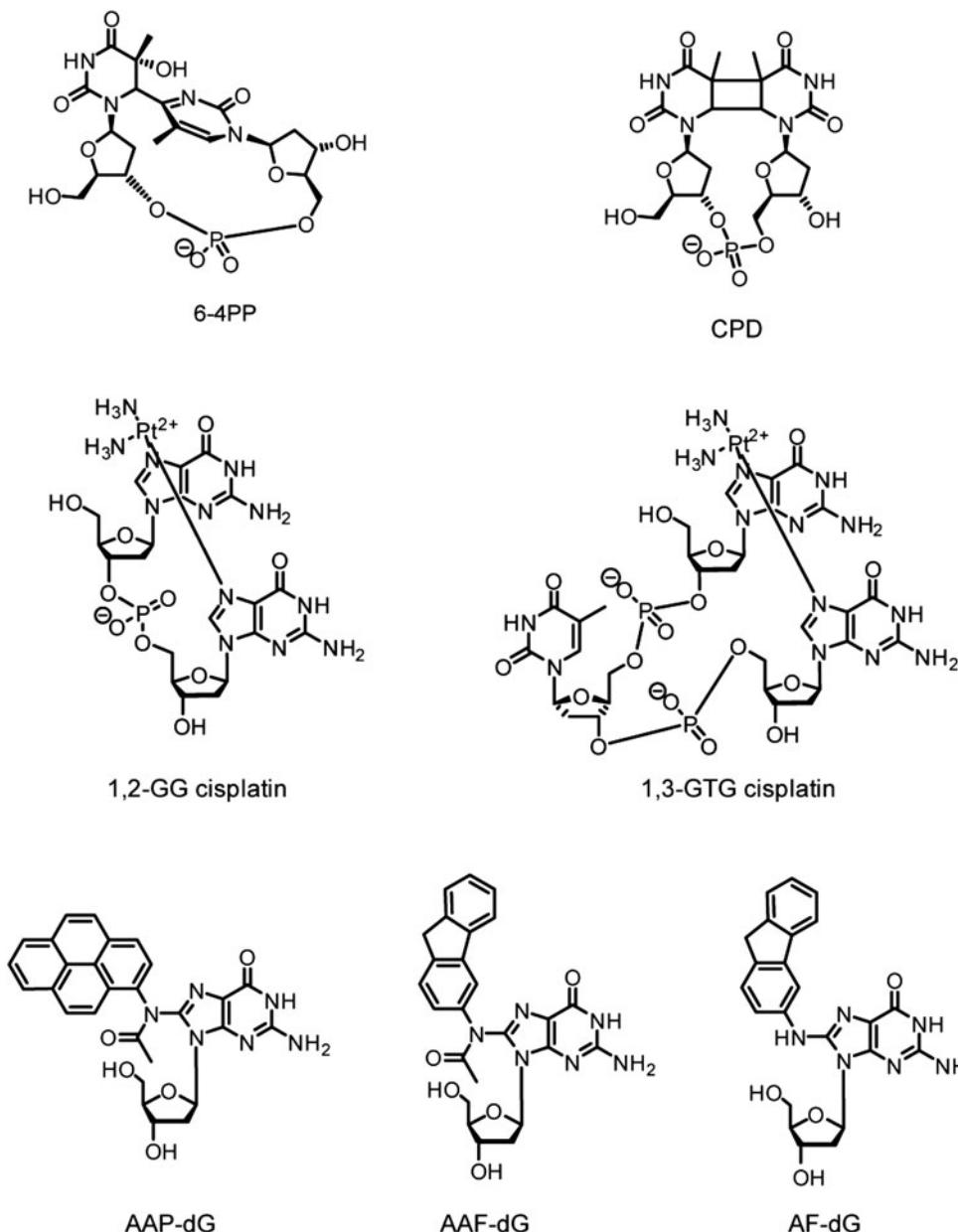


Fig. 1. Examples of DNA adducts repaired by NER. AAF-dG: *N*-(2'-desoxyguanosin-8-yl)-2-acetylaminofluorene, AF-dG: *N*-(2'-desoxyguanosin-8-yl)-2-aminofluorene AAP-dG: *N*-(2'-desoxyguanosin-8-yl)-2-acetylaminopyrene 6-4PP: 6-4PP, CPD: cyclobutane-pyrimidine dimer, 1,2-GG and 1,3-GTG: major cisplatin-induced intrastrand crosslinks lesions.

associated with different NER gene defects were discovered by cell fusion techniques (De Weerd-Kastelein *et al.* 1972). The discovery of NER and its association with genetic disorders together with the mechanism of action have already been reviewed (Friedberg, 1995). Mutations associated with XP and posttranslational modifications of the XP proteins have also been recently reported (Feltes & Bonatto, 2015). The variant form of the human syndrome xeroderma pigmentosum (XPV) (Masutani *et al.* 1999b) is caused by a defect in DNA polymerase η . In this paper, we review the eukaryotic structures of XP proteins (XPA–XPG), including the XP variant DNA polymerase η (called XPV) in complex with damaged or undamaged DNA. This review will provide insights at the atomic level of how lesions in DNA are processed by these proteins.

1.3 NER mechanism

With the exception of chromatin, the NER mechanism is well understood in *Escherichia coli* as well as in eukaryotes (Naegeli & Sugasawa, 2011; Sancar & Tang, 1993; Scharer, 2013). While the repair process in eukaryotes is accomplished by up to 30

proteins, the precise role of each protein has not been fully elucidated so far. Despite this, a general core repair mechanism has emerged (Fig. 2). In eukaryotes, the NER is divided in two distinct subpathways differing only in their damage recognition step (de Laat *et al.* 1999; Hanawalt & Spivak, 2008; Mellon *et al.* 1987). TC-NER ensures the repair of lesions which are located on the transcribed strand of active genes (Hanawalt & Spivak, 2008; Hoeijmakers, 2001; Mellon *et al.* 1987; Sweder & Hanawalt, 1992). Here the RNA-polymerase II seems to act as the initial damage sensor. During transcription the polymerase gets stalled at the lesion site (Donahue *et al.* 1994; Tornaletti & Hanawalt, 1999) and by addition of the protein Cockayne syndrome group B (CSB) a stable CSB–RNAP–DNA complex can form by ATP-hydrolysis (Citterio *et al.* 2000; Tantin *et al.* 1997). In case the lesion blocks any read through, downstream factors such as Cockayne syndrome group A protein (CSA) (Kamiuchi *et al.* 2002), TFIIH (Tantin, 1998) and XPG (Iyer *et al.* 1996) are recruited to the damage site and repair is initiated (Hanawalt & Spivak, 2008). The GG-NER in contrast detects and removes helix distorting lesions throughout the whole genome. The exact mechanism of the damage recognition steps is still under debate. Aboussekha *et al.* and Mu *et al.* were able to reconstitute the repair mechanism *in vitro* allowing deciphering the core excision reaction of NER (Aboussekha *et al.* 1995; Mu *et al.* 1995). The NER assembly seems to be a sequential and coordinated process in which damage excision and repair is mostly error-free. This repair process is characterized by five main steps (steps I–V): after damage recognition (I) a multi-protein complex ('preincision complex' (PIC)) binds to the damage site (II) leading to a double asymmetric cut of the damage-containing oligonucleotide strand. The damage-containing single-stranded DNA (ssDNA) fragment is eliminated (III). Afterwards the double strand is reconstituted by filling the gap with the DNA polymerase (IV) and ligation of the new synthesized strand takes place (V) (Shivji *et al.* 1995). In this work, we will only consider the GG-NER pathway.

1.4 DNA-structure/repair relationships

Because of its broad substrate specificity, it was speculated early on that the NER system does not recognize the lesion itself but the distortion of the DNA structure. In 1997 Naegeli and co-workers discovered that for damage recognition a chemical modification of the nucleotide in addition to distortion in the DNA helix had to be present (Hess *et al.* 1996b). They could show that neither a modified ribose moiety with intact Watson–Crick base pairing nor a DNA bubble of three base-pairs was recognized and repaired by the NER system. However, placing the modified ribose moiety into a DNA bubble, created a well-recognized NER substrate. These observations helped to formulate the '*bipartite substrate discrimination*'-model (Hess *et al.* 1996a, 1997a). This model shows that only DNA damages that cause disrupted base pairing can be recognized as NER substrates. Furthermore, a DNA helix distortion might be additionally needed for the correct recruitment of the NER machinery (Hess *et al.* 1996b). This hypothesis explains why helix-distorting damages are more efficiently processed by NER (Gunz *et al.* 1996; Sugasawa *et al.* 2001). These ideas can also explain why 6–4PPs are better substrates than CPD lesions. Not all adducts, induced by polycyclic aromatic hydrocarbons or aromatic amines, are good substrates for NER, because some adducts do not change the duplex structure, e.g. *N*-(deoxyguanosin-8-yl)-2-aminoanthracene. The aromatic ring of the lesion is easily accommodated in the minor or major grooves of the DNA duplex. Lesions causing helix disruption upon intercalation of their aromatic ring inside the DNA helix are typically good substrates for NER, e.g. *N*-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF), (Gillet *et al.* 2005; Mu *et al.* 1994) or benzo[a]pyrene (Buterin *et al.* 2000; Geacintov *et al.* 2002; Hess *et al.* 1997a; Yan *et al.* 2003). Nowadays, it is assumed that all parameters able to reduce the thermodynamic stability of the DNA play a role for the NER damage recognition step. The repair rate is apparently proportional to the decrease of thermodynamic stability of the duplex induced by the discrete DNA lesions (Aboussekha *et al.* 1995; Geacintov *et al.* 2002; Gillet & Scharer, 2006; Gunz *et al.* 1996).

1.5 Damage recognition in NER

How does discrimination between undamaged and damaged DNA take place in the context of dominating undamaged DNA present in the nucleus? The exact damage recognition process in NER is still unclear and as such it is the most challenging current question associated with the NER process.

Several experiments with fluorescein-labeled proteins, such as ERCC1 (Houtsmuller *et al.* 1999), XPB (Hoogstraten *et al.* 2002) and XPA (Rademakers *et al.* 2003), in combination with photobleaching techniques showed that the repair proteins move freely in the nucleus and accumulate successively at the damaged site. These observations argue that recognition factors must initially bind to the damage site and subsequently recruit downstream factors. At the beginning of NER research XPA was thought to be the initial damage recognizer in NER (Asahina *et al.* 1994; Jones & Wood, 1993). However, in 1998 Sugasawa *et al.* (Sugasawa *et al.* 1998) demonstrated in an *in vitro* experiment, that the complex XPC–RAD23B was the first damage binder followed by participation of XPA. While for XPC–RAD23B the involvement in the GG-NER pathway has been clearly shown, the complex is not necessary for TC-NER. XPA in contrast, is indispensable for both GG-NER and TC-NER, which perhaps supports its central function in a later step of NER (Kobayashi *et al.* 1998; Sugasawa *et al.* 1998). The role of XPA requires further investigations.

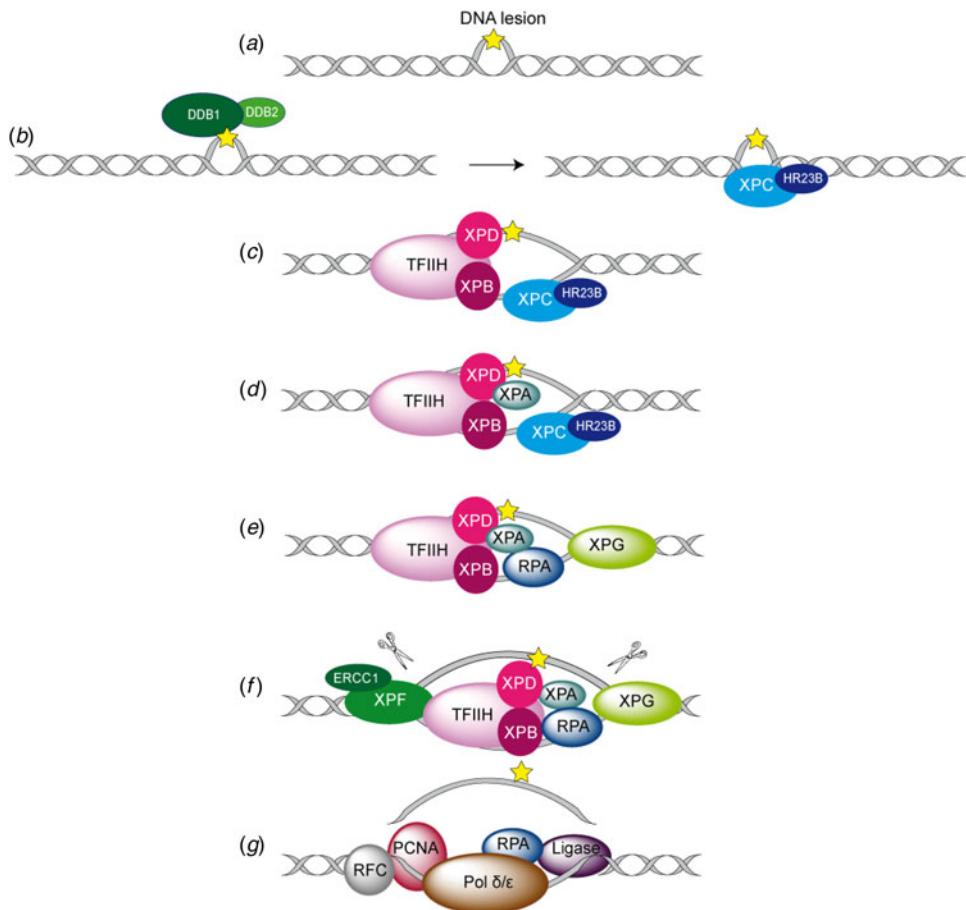


Fig. 2. Overview of the global genome-NER mechanism (Friedberg, 2001; Li *et al.* 2015).

Biochemical experiments demonstrated that the UV-damaged DNA-binding protein (UV-DDB) (Fujiwara *et al.* 1999; Keeney *et al.* 1993; Reardon *et al.* 1993), XPC–RAD23B (Batty *et al.* 2000a; Raoul *et al.* 1995; Sugasawa *et al.* 1998, 2001), XPA (Buschta-Hedayat *et al.* 1999; Jones & Wood, 1993; Li *et al.* 1995a) and replication protein A (RPA) (Burns *et al.* 1996; Mellon *et al.* 1987; Reardon & Sancar, 2002; Schweizer *et al.* 1999) show all affinity to damage-containing DNA. XPC–RAD23B binds specifically to DNA lesions such as 6–4PP (Batty *et al.* 2000a; Kusumoto *et al.* 2001; Raoul *et al.* 1995), cisplatin (Evans *et al.* 1997b; Hey *et al.* 2002; Moggs *et al.* 1996, 1997; Sugasawa *et al.* 1998), dG-AAF (Sugasawa *et al.* 1998) and cholesterol adducts (Kusumoto *et al.* 2001; Sugasawa *et al.* 2001). It recognizes CPD damages (Kusumoto *et al.* 2001) poorly although their repair is strongly dependent on the presence of XPC (Venema *et al.* 1991). The affinity of XPC–RAD23B to CPDs can be enhanced by mismatches surrounding the lesion (Sugasawa *et al.* 2001). XPA alone or in combination with RPA binds to DNA containing bulky adducts such as dG-AAF (Liu *et al.* 2005; Yang *et al.* 2002). By far the strongest DNA binding for XPA are DNA junctions which suggest that XPA interacts with an intermediate structure rather than directly with a lesion (Missura *et al.* 2001). Photoaffinity-crosslinking experiments demonstrate that XPC–RAD23B (Maltseva *et al.* 2007), as well as XPA (Wakasugi & Sancar, 1999) show a high affinity to damaged DNA without binding directly to the lesion (Reardon & Sancar, 2002). Reconstitution experiments with NER-deficient cells with all the different NER proteins in different order delivered contradictory results. Sugasawa *et al.* showed a fast repair rate when the cells were preincubated with XPC–RAD23B (Sugasawa *et al.* 1998). Preincubation of the damaged plasmids with XPA–RPA slowed down the repair velocity. Experiments from Wakasugi *et al.* however showed a more effective repair when damaged DNA was preincubated with XPA/RPA (Wakasugi & Sancar, 1999).

1.6 Overview of each XP protein and its role in GG-NER

In Fig. 2, the different steps of the GG-NER mechanism are depicted. It is believed that XPC–RAD23B initiates the GR-NER by recognition of the damage (Fig. 2a, b) (Reardon & Sancar, 2003; Sugasawa *et al.* 1998; Wakasugi & Sancar, 1998). XPC ignores the lesion but binds to the ssDNA opposite to the lesion (Min & Pavletich, 2007). In the case of UV lesions, the

Table 1. Overview of the XP proteins involved in NER and TLS. ERCCX: Excision repair cross-complementing rodent repair deficiency complementation group X. Af: *Archaeoglobus fulgidus*, Dr: *Danio rerio*, Sa: *Sulfolobus acidocaldarius*, Ape: *Aeropyrum pernix*, Ta: *Thermoplasma acidophilum*, AAF-dG: N-(2'-desoxyguanosin-8-yl)-2-acetylaminofluorene, 6-4PP: 6-4PP, CPD: cyclobutane-pyrimidine dimer. PDB codes depicted in bold refer to NMR structures.

Protein (gene)	Function	Protein size (aa)	Interaction partner	PDB code	DNA substrate
XPA (XPA)	Damage verification	31 kDa (273)	RPA, TFIIH, XPF-ERCC1, DDB2, PCNA	1XPA	–
				5A39 (Rad14)	cisplatin
				5A3D (Rad14)	AAF-dG
XPB (ERCC3)	3'-5' DNA helicase	89 kDa (782)	XPC, XPA, XPG, XPF	2FWR (Af)	–
				4ERN	–
XPC-RAD23B (XPC)	Damage recognition	XPC: 106 kDa (940) RAD23B: 43 kDa (409)	XPA, TFIIH, RAD23B, UV-DDB2,	2QSG (Rad4)	CPD
XPD (ERCC2)		87 kDa (760)	XPC, XPA, XPG, XPF	3CRV (Sa) 4A15 (Ta)	– ssDNA
XPE (UV-DDB1/2)	Damage recognition of UV lesions	DDB1: 127 kDa (1149) DDB2: 48 kDa (428)	XPC, RPA, XPA	3EI1 (Hs-Dr) 4A08	6-4PP CPD
				4E5Z	abasic lesion
XPF-ERCC1 (ERCC4)	5'-endonuclease	104 kDa (916) 33 kDa (297)	TFIIH, XPA, ERCC1	2A1J 2KN7	– ssDNA
XPG (ERCC5)	3'-endonuclease	133 kDa (1186)	XPA, TFIIH, PCNA, XPD	4Q0W (Rad2)	dsDNA with ss overhang
XPV (POLH)	Damage bypass	78 kDa (713)	PCNA	1JIH (Rad30) 2R8K (Rad30) 3MFI (Rad30) 3MR3	– cisplatin CPD CPD

recognition process by XPC-RAD23B is sustained by the UV-DNA damage-binding protein 2 (UV-DDB2) (Luijsterburg *et al.* 2010; Min & Pavletich, 2007; Riedl *et al.* 2003; Volker *et al.* 2001). After this step, the general transcription factor TFIIH is recruited to the lesion site using the energy provided by the ATPase activity of XPB (Fig. 2c). Normally, TFIIH is involved in transcription by RNA polymerases I and II, but in the presence of DNA damage it is clearly recruited 5' to the damage site (Hoogstraten *et al.* 2002). Two subunits of TFIIH, the ATP-dependent DNA helicases XPB and XPD, are needed for ‘hooking up’ TFIIH at the lesion site performing further unwinding and damage verification. It is believed that XPB opens the DNA and facilitates the binding of XPD to the DNA which verifies the chemical modification of the lesion (Scharer, 2013). In addition, XPD scans the strand with the lesion and both helicases are stalled when XPD encounters a bulky lesion (Li *et al.* 2015; Mathieu *et al.* 2013). In the next step, XPA and RPA (Fig. 2d, e) are thought to be recruited to the damaged site to form the PIC which is stabilized by the structure specific endonucleases XPG (Fig. 2e). Upon arrival of XPG to the PIC, XPC-RAD23B leaves the complex (Riedl *et al.* 2003; Wakasugi & Sancar, 1998). Besides XPD, XPA might be involved in damage verification, however, the details of this event remain elusive (Naegeli & Sugasawa, 2011; Tapias *et al.* 2004). The ssDNA-binding protein RPA now binds the undamaged strand, while TFIIH opens the duplex. XPA is then thought to recruit the structure specific endonuclease XPF-ERCC1 heterodimer (Li *et al.* 1995a; Tsodikov *et al.* 2007) and incision on both sites of the lesion takes place. The first incision 5' of the lesion is made by XPF-ERCC1 (Staresinic *et al.* 2009). Then, XPG cuts 3' to the lesion (Fagbemi *et al.* 2011) triggering the elimination of a 24–32 oligonucleotide containing the lesion (Fig. 2f) (Edenberg & Hanawalt, 1972; Evans *et al.* 1997b; Ikegami *et al.* 1998). After that, DNA replication by the polymerases $\kappa/\delta/\epsilon$ takes place sustained by other factors (PCNA, RFC and RPA) (Ogi *et al.* 2010). Finally, the new synthesized DNA strand is mostly sealed by DNA ligase III/XRCC1. DNA ligase I is only recruited to sites of UV damage in proliferating cells (Fig. 2g) (Moser *et al.* 2001, 2007; Reardon & Sancar, 2005; Wood, 1997). Table 1 gives an overview of the XP proteins, their function and size, their interaction partners as well as the PDB codes of solved structures associated to the DNA substrates.

2. XPE (UV-DDB2)

For the recognition of UV lesions such as 6-4PPs and CPD, the UV-DDB protein seems to promote binding of XPC-RAD23B to the damaged site (Scrima *et al.* 2008). The UV-DDB protein, consists of the two subunits DDB1 (also called p127 with 127 kDa, 1149 amino acids) and DDB2 (called p48 with 48 kDa, 428 amino acids) (Fig. 3a). It shows a 100- to 1000-fold higher affinity to photolesions than XPC-RAD23B (Keeney *et al.* 1993; Takao *et al.* 1993). Both subunits are required for the activity of the complex (Hwang *et al.* 1998a; Keeney *et al.* 1993; Nichols *et al.* 1996). UV-DDB is known to be involved in the DNA damage recognition process of GG-NER (Wakasugi *et al.* 2002). The DDB2 subunit binds with

preference to 6–4PPs but it also recognizes CPD lesions and it also binds abasic sites as well as DNA bubbles (Fujiwara *et al.* 1999). DDB1 is not involved in DNA binding. Mutation of the DDB2 subunit leads to a loss of function of the UV-DDB protein found in XP complementation group E (XPE) in patients (Chu & Chang, 1988). XPE-deficient cells are shown to be defective in GG-NER (Hwang *et al.* 1999). Wakasugi *et al.* showed that UV-DDB rapidly accumulates *in vivo* independently of XPA and XPC at DNA damage sites after UV-irradiation (Wakasugi *et al.* 2002). Furthermore, they showed that UV lesions are mostly recognized by UV-DDB which then recruits XPA and XPC.

2.1 Crystal structure of human XPE with 6–4PP damaged dsDNA

Scrima *et al.* (2008) solved the crystal structure of the UV-DDB (human DDB1, HsDDB1 and the zebrafish DDB2 ortholog, DrDDB2) protein in complex with double-stranded DNA (dsDNA) containing the 6–4PP photolesion (Fig. 3b). In order to achieve damage recognition the UV-DDB2 protein inserts a β -hairpin into the minor groove of the DNA. In addition, the damaged dinucleotide is flipped out of the double helix and inserted into a binding pocket of the protein leading to a DNA kink of about 40° at the lesion site. This destabilized conformation of the DNA might trigger the subsequent recruitment of XPC–RAD23B and further downstream factors. Due to the importance of DDB2, defective cells show a remarkable reduced CPD repair activity that gives rise to the XP-E phenotype (Hwang *et al.* 1998a; Tang *et al.* 2000).

2.2 Crystal structure of human XPE with CPD damaged dsDNA

Fischer *et al.* (2011) obtained insight into the molecular basis of CPD lesion recognition by solving the crystal structure of the UV-DDB protein in complex with CPD containing DNA. For obtaining the structure, residues creating crystal contacts in the apo-protein that prevented protein crystallization were mutated. Consistent with the structure of Scrima *et al.* (2008) the CPD lesion is held by the DDB2 WD40 propeller. The crystal structure shows that the lesion recognition takes place with the help of a β -hairpin ('finger') consisting of the residues Phe371, Gln372 and His373 (Fig. 3c). This β -hairpin inserts into the minor groove of the DNA duplex at the damage site and unwinds the duplex by 12.6°. Consequently the CPD lesion is extruded into an extrahelical, flipped-out state. Finally, the DNA shows also a kink angle of about 45° at the *cis*–*syn* CPD lesion site.

The UV-DDB protein is part of an ubiquitin ligase complex. The displacement of COP9 signalosome (CSN) from this complex by DDB2 substrate binding leads to ubiquitination of XPC and DDB2 (Groisman *et al.* 2003; Luijsterburg *et al.* 2007). This is thought to mediate the handover to XPC by stabilizing it, resulting in an increase of its affinity for damaged DNA, while the affinity of DDB2 to DNA containing a lesion is reduced (El-Mahdy *et al.* 2006; Sugasawa *et al.* 2005). Histones surrounding the lesion are ubiquitinated as well, which is suggested to slacken the nucleosome structure to promote recruitment of NER repair enzymes to the damage (Wang *et al.* 2006).

2.3 Crystal structure of human XPE with an abasic lesion site in dsDNA

Yeh *et al.* (2012) solved a dimeric crystal structure of UV-DDB (human DDB1 and DDB2) protein in complex dsDNA containing an abasic site (THF) (Fig. 3d). Each DDB2 subunit is bound to one double-stranded oligonucleotide. The DDB2 residues Asn360/Asn360 are located between the two DNA strands spanning two antiparallel β -strands (β -wing) and giving rise to a 2-fold symmetry axis of the dimer. The crystal structure shows that the residues on the leading β -strand and loop form contacts with the undamaged DNA strand and the residues on the loop and the retreating β -strand form contacts with the neighboring undamaged DNA strand. These contacts are supported by additional interactions of the N-terminal-helical domain of DDB2 (α -paddle) with the neighboring DNA-strand which lead to dimerization, explaining the high affinity to damaged DNA.

3. XPC

The human XPC protein is a polypeptide of 940 amino acids with a size of 106 kDa, forming a stable complex with the RAD23B protein (43 kDa, 409 amino acids). It is currently believed that XPC–RAD23B is responsible for the initial damage recognition with the exception of CPD lesions, where DDB2 is the first binding factor (Hey *et al.* 2002; Sugasawa *et al.* 2001). The complex is known to interact with several other proteins, e.g. XPA and TFIIH (Fig. 4a, b) (Araujo *et al.* 2001; Bernardes de Jesus *et al.* 2008; Cleaver *et al.* 2009). After binding, the protein recruits further downstream factors to the damaged site (Sugasawa *et al.* 1998). Biochemical assays show that the XPC–RAD23B complex binds with high affinity to different DNA lesions, e.g. 6–4PP and AAF-adducts (Batty & Wood, 2000b; Sugasawa *et al.* 2002). The breakthrough came in 2001 when the Mullenders group was able to show that XPC might be the initial damage recognizer also for UV lesions. In the experiment, nuclei of living cells were irradiated with UV light at defined positions through a porous polycarbonate filter (Katsumi *et al.* 2001; Mone *et al.* 2001; Volker *et al.* 2001). The generated UV lesions could be detected by antibodies specific against CPDs and 6–4PPs. To identify which NER factors were recruited to these sites they were also visualized

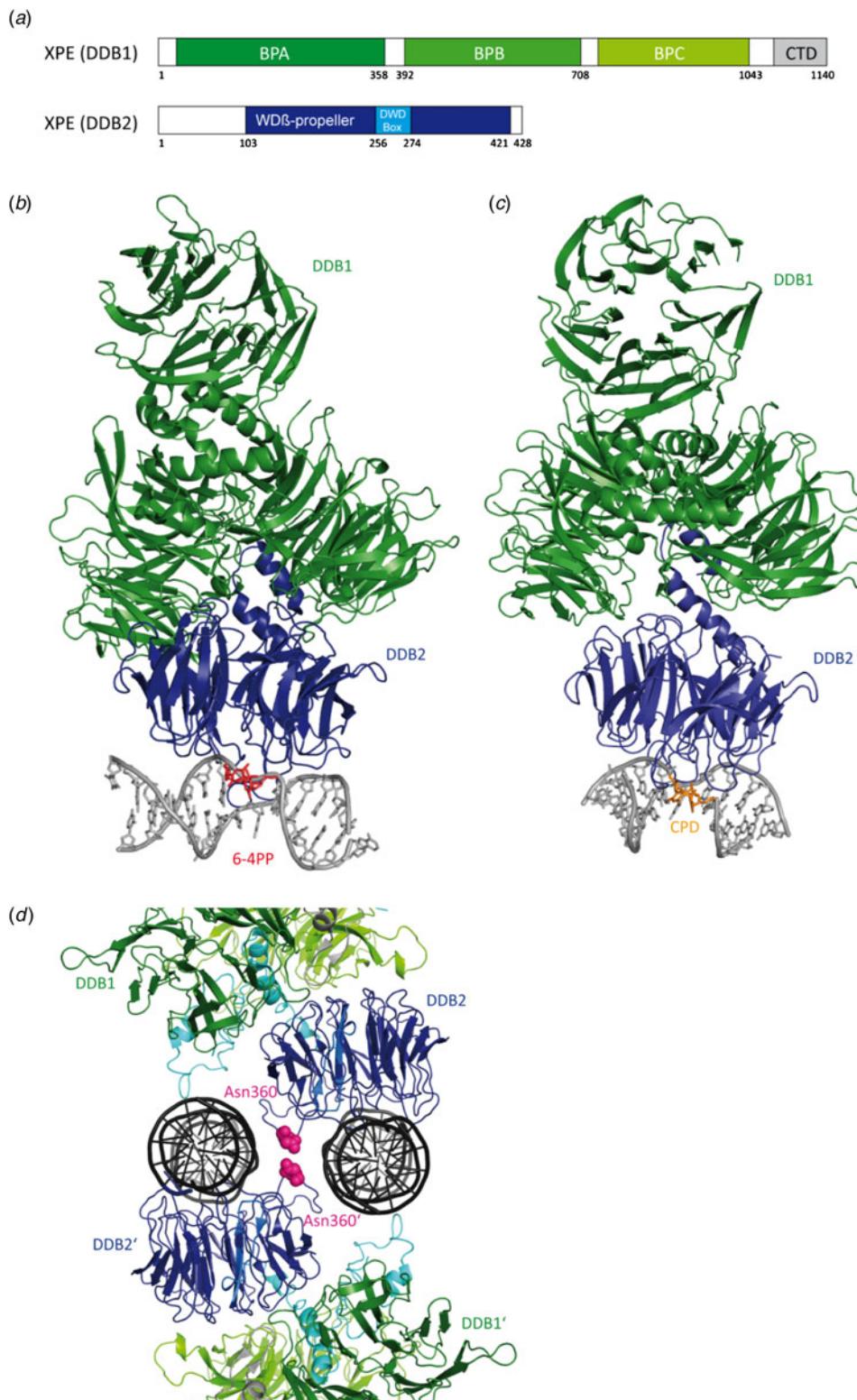


Fig. 3. Crystal structures of the human XPE protein (UV-DDB). (a) Schematic representation of the human UV-DDB protein (DDB1 and DDB2) with the different domains. (b) Crystal structure of the XPE protein in complex with 6-4PP damaged DNA (PDB code: 3EI1) (Scrima *et al.* 2008). DDB1 is depicted in green, DDB2 in blue, DNA in gray and the 6-4PP dinucleotide in red. (c) Crystal structure of XPE in complex with the CPD lesion (PDB code: 4A08) (Fischer *et al.* 2011). DDB1 is depicted in green, DDB2 in blue, DNA in gray and the CPD lesion in orange. (d) Crystal structure of the dimeric human UV-DDB in a complex with an abasic lesion (PDB code: 4e5z) (Yeh *et al.* 2012). DDB1 is depicted in green, DDB2 in blue, DNA in dark gray, DDB2 α -paddle in cyan and the amino acid Asn360 in pink.

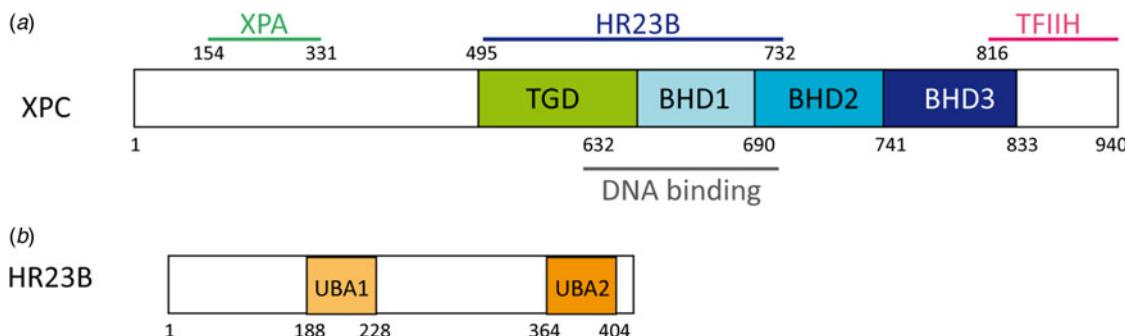


Fig. 4. Schematic representation of human XPC and RAD23B. (a) The TGD domain of XPC is depicted in green, the BHD1 domain in light blue, the BHD2 domain in blue and the BHD3 domain in dark blue. (b) The UBA1 and UBA2 domains of RAD23B are depicted in light orange and orange, respectively.

with an appropriate antibody. This experiment revealed that XPC colocalized with UV lesion sites in XPA-deficient cells, while XPA was not recruited in XPC-deficient cells (Volker *et al.* 2001). Thus, it was concluded that XPA needs the presence of XPC to be recruited to the damaged site and that XPC is the first factor being recruited to the damaged site. Further investigations pointed out that XPC is also required to unwind the DNA at the damaged site in order to recruit NER downstream factors (Katsumi *et al.* 2001; Mone *et al.* 2001). It has to be addressed that these experiments refer to the recognition of UV lesions. The general lesion recognition strategy adopted by XPC was established by a variety of experiments with different DNA lesions (Tapias *et al.* 2004). The helicase activity of XPD is not required for transcription but is indispensable for the NER process. The combined data show that XPC might be more structure specific than lesion specific. The XPC–RAD23B complex is able to recognize ssDNA (Maillard *et al.* 2007) and DNA containing destabilizing lesions. Additionally it was shown, that the complex binds the DNA lesion in an asymmetric fashion (Sugasawa *et al.* 2002; Wakasugi & Sancar, 1999). Interestingly, the complex also binds to 3–5 unpaired bases independently of the presence of damage. However, repair takes place only in the presence of DNA damage (Hess *et al.* 1997b; Sugasawa *et al.* 2001).

Hence, XPC might be a sensor for unpaired base pairs that occur in bubbles and loops and binds those in complex with RAD23B (Araki *et al.* 2001; Van der Spek *et al.* 1996) and Centrin2 (Araki *et al.* 2001). This would explain why 6–4PP, AAF-dG and intrastrand cisplatin crosslinks, which introduce a severe helical kink into the DNA structure (Lukin & de Los Santos, 2006), are good substrates. On the other hand, CPDs are not recognized, possibly because they do not disturb the helical structure (Kusumoto *et al.* 2001; Reardon & Sancar, 2003; Sugasawa *et al.* 2001). As mentioned previously the affinity of XPC–RAD23B to CPDs can be enhanced by introducing mismatches at the lesion site (Sugasawa *et al.* 2001).

3.1 Crystal structure of yeast XPC (Rad4) in complex with a CPD lesion

Min & Pavletich, (2007) obtained the crystal structure of the yeast homolog of XPC–RAD23B (Rad4-Rad23, Fig. 5a) in complex with a CPD lesion in a mismatched situation with two thymidines (Fig. 5b). A shorter version of the protein (the DNA-binding domain) showing the same affinity to damaged DNA as the full length protein was crystallized. The crystal structure shows that Rad4 inserts a β -hairpin (beta hairpin domain 3, (BHD3)) into the destabilized DNA duplex, inducing the two damaged base pairs to completely flip out of the double helix (base flipping). A kink of 42° is created which could be confirmed by atomic force microscopy (Hwang *et al.* 1999). The protein recognizes the expelled nucleotides of the undamaged strand through Van der Waals interactions with BHD1 and transglutaminase homology domain (TGD), while the two CPD-linked nucleotides are disordered in the structure and could not be resolved. The protein undergoes a conformational change during binding to damaged DNA (*induced fit*). The authors deduced from these findings that the lesions recognized by XPC must thermodynamically destabilize the Watson–Crick double helix in a manner that facilitates the flipping-out of two bases. The protein recognizes a weakened and distorted DNA helix without binding directly to the damage.

The structure, however, does not explain why the CPD damage alone, without a mismatch, is so efficiently repaired by NER. For UV lesion repair, additional damage recognition factors must exist. New experiments show that the UV-DDB protein plays a very important role and that this protein recruits the XPC–RAD23B protein to the UV-damaged sites (Fitch *et al.* 2003; Wakasugi *et al.* 2002). Also the 6–4PP lesion which is a better NER substrate than CPD (Mitchell & Nairn, 1989), requires XPC–RAD23B as well as UV-DDB for recognition. See Section 2.

Camenisch *et al.* (2009) proposed a two-stage detection process for XPC that recognizes DNA lesions without involvement of a DNA-binding domain (BHD3). BHD1/BHD2 and a β -turn subdomain act as a damage sensor that scans the double helix

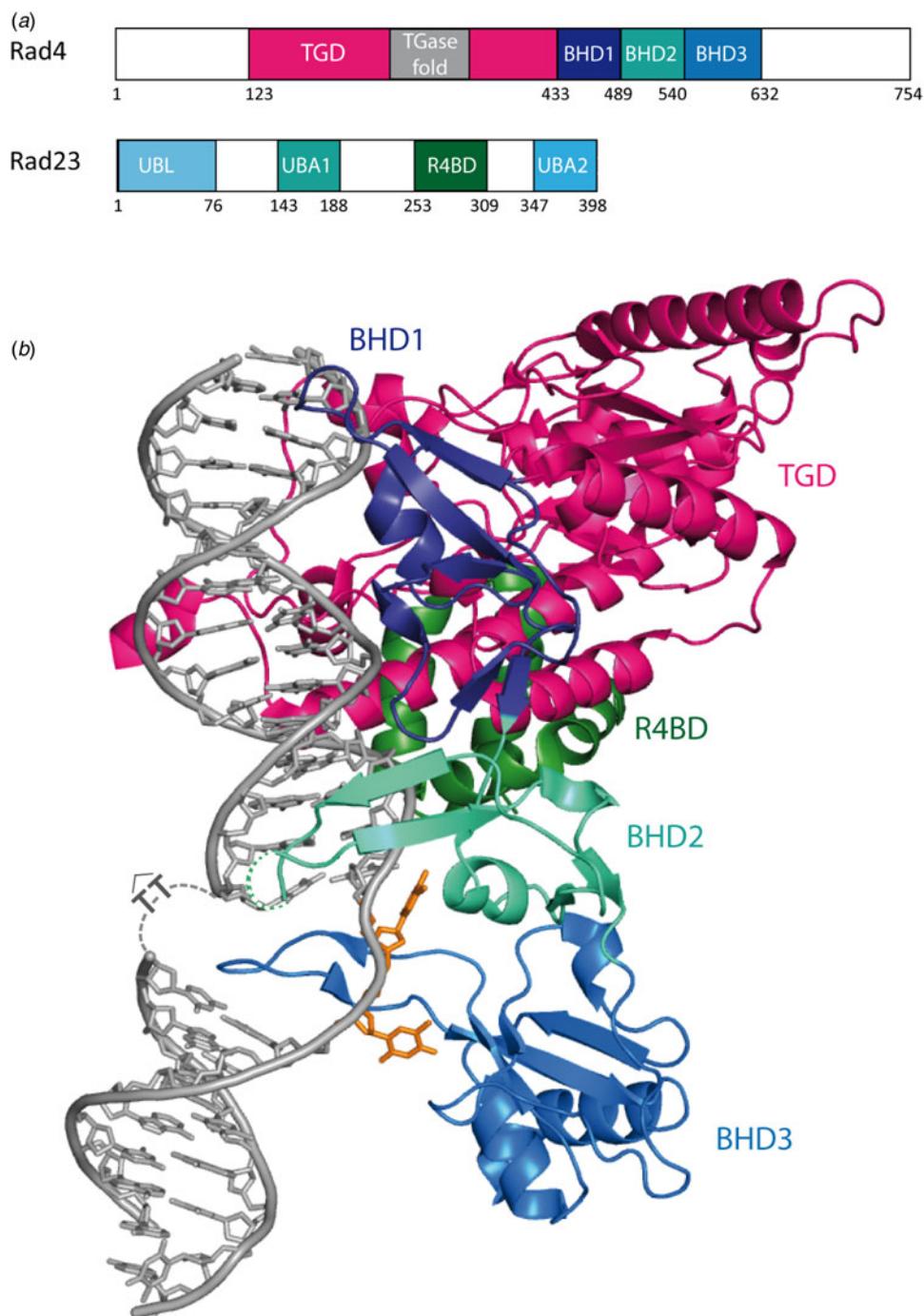


Fig. 5. (a) Schematic representation of Rad4 and Rad23. The TGD domain is depicted in pink, BHD1 in dark blue, BHD2 in cyan, BHD3 in blue and R4BD in green. (b) Crystal structure of the yeast Rad4–Rad23 in complex with a CPD damaged DNA (PDB code: 2QSG) (Min & Pavletich, 2007). Same color code as in (a). The extruded Ts in front of the CPD lesion are depicted in orange.

for unpaired bases, which forms a labile recognition complex. Subsequently, a stable recognition is formed through binding of BHD3 to the flipped out nucleotides in the undamaged strand across the lesion (Camenisch *et al.* 2009).

Chen *et al.* (2015) obtained the crystal structure of the yeast homolog of XPC–RAD23B (Rad4–Rad23) in complex with undamaged DNA using disulfide crosslinking. The crystal structure shows that Rad4 flips out normal nucleotide pairs like with damaged DNA. Therefore, structural discrimination of Rad4/XPC cannot play an important role for lesion recognition. The authors propose a ‘kinetic-gating’ mechanism for Rad4/XPC lesion recognition, which is controlled by two kinetic parameters (opening time and residence time). The weakened base stacking and hydrogen bonding within a lesion lead to a larger

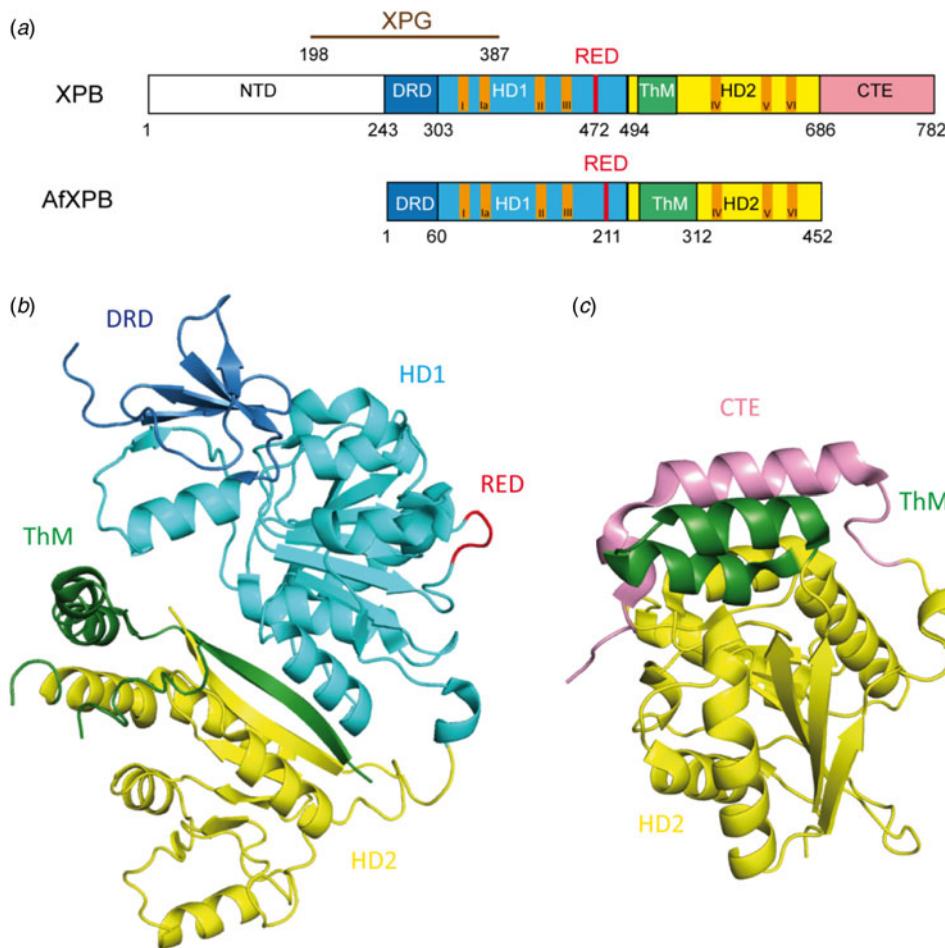


Fig. 6. (a) Schematic alignment of XPB and AfXPB showing the different domains: The DRD domain is depicted in blue, HD1 in light blue, HD2 in yellow, ThM in green and CTE (C-terminal extension) in pink. The conserved helicase motifs (Walker A motif I–VI) are depicted in orange. (b) Crystal structure of the AfXPB protein (PDB code: 2FWR). (Fan *et al.* 2006). Same color code as in (a). (c) Crystal structure of C-terminal domain of the human XPB protein (PDB code: 4ERN) (Hilario *et al.* 2013). Same color code as in (a).

opening time for damaged DNA than for undamaged DNA to form a stable recognition complex. Also, the residence time is larger for damaged DNA indicating that Rad4/XPC opens damaged DNA with a higher probability (Chen *et al.* 2015).

4. XPB

XPB (also called ERCC3) is a 3'-5' ATP-dependent DNA helicase and a subunit of the TFIIH complex necessary for transcription and DNA repair (Gillet & Scharer, 2006; Lehmann, 2003; Schaeffer *et al.* 1993). Mutations in XPB lead to a 1000-fold increase of the melanoma risk and neurological abnormalities because of a defective TFIIH function in NER (Lehmann, 2003). The human protein consists of 782 amino acids (89 kDa). NER helicases (XPB and XPD) all include conserved Rad51/RecA like ATPase domains called helicase domains 1 and 2 (HD1 and HD2), containing the ATPase and helicase motifs characteristic for superfamily 2 (SF2) helicases (Shin *et al.* 2003; Story & Steitz, 1992). Furthermore, NER helicases consist of two accessory domains, which are responsible for transmitting the helicase domain changes upon binding to DNA. The ATPase activity of XPB is necessary during NER to anchor TFIIH to the lesion site (Abdulrahman *et al.* 2013; Coin *et al.* 1998; Dubaele *et al.* 2003; Oksenyech *et al.* 2009). Its helicase activity is, however, not needed for the repair process (Coin *et al.* 2007; Oksenyech *et al.* 2009; Tirode *et al.* 1999).

4.1 Crystal structure of *Archaeoglobus fulgidus* XPB (AfXPB)

Fan *et al.* (2006) determined the crystal structure of an *A. fulgidus* XPB homolog (AfXPB) with a sequence equivalent to residues 240–686 of the human XPB (Fig. 6b). The AfXPB structure is built up by two Rad51/RecA-like helicase domains, HD1 and HD2, containing seven conserved helicase motifs (Walker motif I, Ia and II–VI) in the middle of the polypeptide (Fig. 6a)

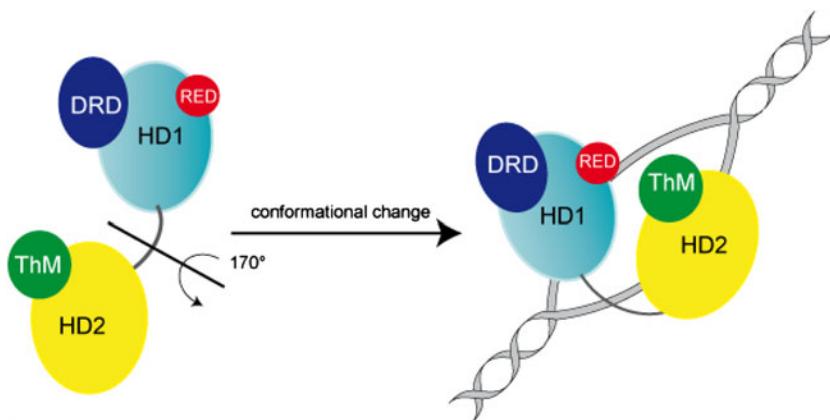


Fig. 7. Proposed model for binding of XPB to DNA (adapted from Fan *et al.* (2006)). The apo-protein is in an opened conformation. By binding to DNA, a rotation (170°) of the second helicase domain (HD2) together with the ThM domain, facilitated by HD1-mediated ATP hydrolysis, takes place and forms the closed and stable XPB–DNA complex. Same color code as in Fig. 6a.

(Fuss & Tainer, 2011; Tuteja & Tuteja, 1996). These domains use the binding and hydrolysis of ATP to induce conformational changes of the protein. Furthermore, the protein contains a DNA damage recognition domain (DRD), a unique RED motif (Arg–Glu–Asp) and a thumb domain (ThM). These domains are highly conserved and they are also found in the corresponding human protein. The structure suggests that the ATP hydrolysis provides the energy to induce a flip of about 170° in the orientation of HD2 relative to HD1 after DNA binding (Fig. 7). This conformational change brings the RED domain and the ThM domain in close vicinity and stabilizes TFIIH on the DNA by inserting a wedge (Glu473) into the double strand, gripped by the ThM domain. The motion might be important for melting the DNA duplex and anchoring TFIIH (Coin *et al.* 2007). Mutations in the RED and ThM domains as well as mutations in the ATPase domain diminish the helicase activity, suggesting that these three domains ensure the correct recruitment of TFIIH to the damaged site.

4.2 Crystal structure of human XPB

Hilario *et al.* (2013) successfully solved the crystal structure of the C-terminal fragment of human XPB (residues 494–782) containing the helicase domain 2 (HD2), the thumb-like domain (ThM) and the C-terminal extension (Fig. 6c). They investigated the structural basis for the role of the C-terminus of XPB in transcription and DNA repair. Because of lack of electron density in the structure, only the positions of the residues 502–730 could be determined. The observed structure consists of a globular domain containing a central sheet of seven parallel β -strands sandwiched by three α -helices on one side and two α -helices on the other (Fig. 6c). The C-terminal extension (residues 670–730) is built up by two α -helices linked by a loop and a β -strand, which forms the edge strand of the seven-stranded central sheet, and an unstructured tail that possibly extends to the C-terminus. The ThM domain and one α -helix of the C-terminal extension are located at the top of the central sheet. Thus, the recruitment of TFIIH to the damage site is an active process controlled by the ATPase motif of XPB. This suggests that this subunit acts as an ATP-dependent hook that stabilizes the binding of the TFIIH to damaged DNA (Hilario *et al.* 2013). In fact, a mutated ATPase activity of XPB abolishes the NER activity (Guzder *et al.* 1994; Sung *et al.* 1988). Recruiting TFIIH to the damage site by the ATPase activity of XPB may reorganize the protein–DNA complexes in order to allow new protein–protein or protein–DNA contacts to happen.

5. XPD

The XPD (also called ERCC2) protein with a size of 87 kDa (760 amino acids) is known to be involved in transcription and DNA damage repair. It is a key subunit of the TFIIH complex and acts as an ATP-dependent 5'-3' DNA helicase that belongs to the superhelicase 2 family (SF2) proteins (Araujo *et al.* 2000; Coin *et al.* 2007). The helicase activity of XPD is not required for transcription but is indispensable for the NER process (Winkler *et al.* 2000). The TFIIH complex interacts with other NER factors such as XPA, XPC, XPG and XPF–ERCC1 (Evans *et al.* 1997a; Gillet *et al.* 2005). Mutations in human XPD lead to three different inherited diseases: XP, the CS and TTD. XPD is thought to be responsible for the damage verification (Mathieu *et al.* 2010; Naegeli *et al.* 1993) before the recruitment of downstream NER factors to the lesion site occurs. This idea has been postulated based on the observation that the helicase activity of XPD becomes inhibited by the presence of DNA damage (Mathieu *et al.* 2010, 2013; Naegeli *et al.* 1992).

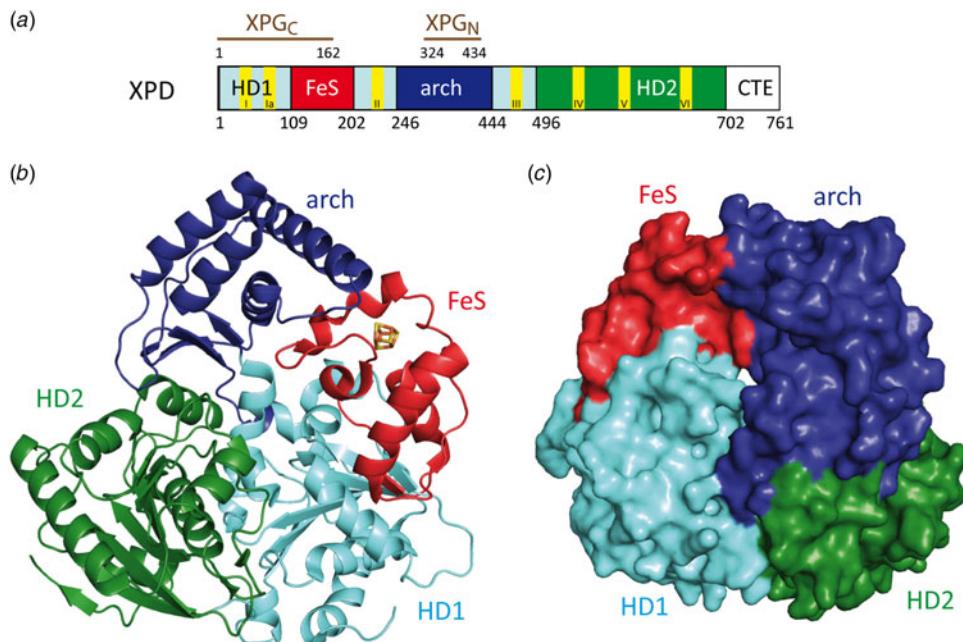


Fig. 8. (a) Schematic representation of human XPD (Fan *et al.* 2008). The protein interacts with XPG_C (C-terminal domain of XPG) and XPG_N (N-terminal domain of XPG) depicted in brown. (b) Crystal structure of SaXPD (PDB code: 3CRV) (Fan *et al.* 2008) with the HD1 domain (light blue) and the HD2 domain (green) form the ATP-binding interface. The arch domain is depicted in blue and FeS domain in red. The iron–sulfur cluster (4Fe4S) is depicted in orange and yellow. (c) Surface representation of SaXPD showing the tunnel formed by the different domains, which is large enough to pass through ssDNA.

5.1 Crystal structure of *Sulfolobus acidocaldarius* XPD (SaXPD)

Several crystal structures of the catalytic domain of XPD from different organisms (*S. acidocaldarius* (Fan *et al.* 2008) (SaXPD; Fig. 8), *Thermoplasma acidophilum* (Wolski *et al.* 2008) (TaXPD) and *Sulfolobus tokodaii* (Liu *et al.* 2008)) reveal a four domain organization: a domain harboring the 4Fe4S iron–sulfur cluster (FeS domain), an arch domain and two canonical Rad51/RecA like domains, the helicase domains HD1 and HD2 (Fig. 8a). The FeS domain is essential for the helicase activity of XPD.

The different domains form a tunnel that is large enough to harbor ssDNA (Fig. 8b, c) (Fan *et al.* 2008; Liu *et al.* 2008; Wolski *et al.* 2008). Fan *et al.* proposed a model of XPD binding to ssDNA (Fig. 9) (Fan *et al.* 2008). The iron–sulfur cluster is supposed to act as a molecular plough, which scans blocking damages by breaking hydrogen bonds (Lehmann, 2008; White, 2009). In this process, one strand of the duplex is supposed to be pulled through the internal channel with the help of the ATP-dependent motor domain, while damaged DNA will plug the channel (Fan *et al.* 2008; Liu *et al.* 2008; Mathieu *et al.* 2013; Pugh *et al.* 2012; Wolski *et al.* 2008). In this way, the protein undergoes a conformational change by closing the domain between the arch and the FeS motifs. The helicase domains HD1 and HD2 furnish the necessary flexibility for translocation (Fan *et al.* 2008).

Because the channel is not narrow enough to distinguish between damaged and undamaged bases 2008, Wolski *et al.* proposed that during the unwinding process single bases are analyzed through insertion of the DNA backbone into a neighboring binding pocket on the protein surface. Damaged bases would not fit into the pocket, which could be the signal that leads to blocking of the helicase activity (Wolski *et al.* 2008).

5.2 Crystal structure of *T. acidophilum* (TaXPD) in complex with ssDNA

Finally, in 2012 Kuper *et al.* successfully solved the crystal structure of TaXPD in complex with a ssDNA oligonucleotide (22mer), shedding light on the mechanism of polarity and translocation of the protein along the DNA (Kuper *et al.* 2012). At the same time Pugh *et al.* showed that the 5'-3' translocation polarity needs conserved residues in the HD1 and the FeS domains. They propose that the FeS domain provides a wedge-like feature enabling the duplex opening during unwinding (Pugh *et al.* 2012).

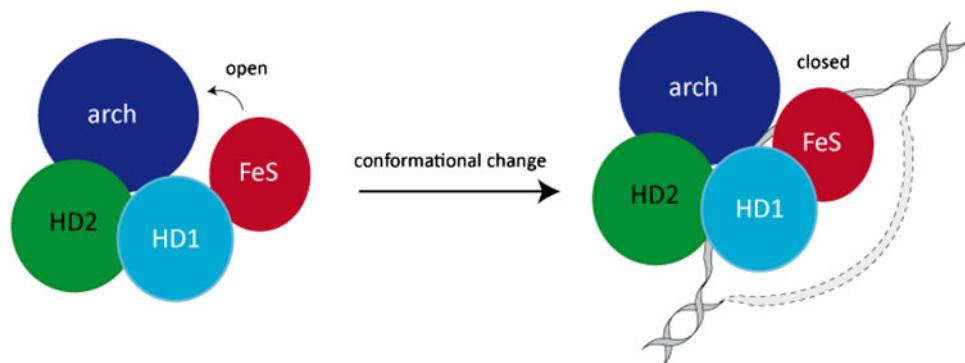


Fig. 9. DNA-binding model of SaXPD adapted from Fan *et al.* (2008). The HD1 domain is depicted in light blue, the HD2 domain in green, the arch domain is depicted in blue, the FeS domain in red and the DNA in gray. Upon binding to the DNA the protein undergoes a conformational change by closing the domain between the arch and the FeS domains.

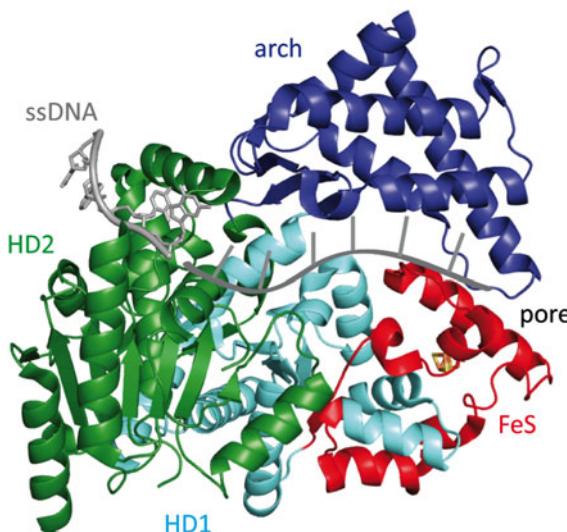


Fig. 10. Crystal structure of *Thermoplasma acidophilum* XPD (TaXPD) in complex with a 10 nt ssDNA. The arch domain is depicted in dark blue, the FeS domain in red, the HD1 domain in cyan and the HD2 domain in green. The four bases (TACG), showing electron density, are depicted as gray sticks. The modeled ssDNA in the pore is schematically represented in gray. The modeled structure is adapted from Kuper *et al.* (2012).

The structure allowed the Kisker group to define the translocation path of the DNA substrate through the protein and to determine the DNA-binding domain (DBD) of TaXPD by mutational and biochemical analysis. Only four bases (TACG) of the oligonucleotide could be identified in the electron density. These bases are located in the cleft of the HD2 domain (Fig. 10). This domain was already predicted as a potential binding cleft by Wolski *et al.* (2008). The structure shows stacking of Trp549 with the adenine of TACG. Arg584 builds hydrogen bonds with N1 and N2 of the phosphate backbone of adenine and cytosine. Phe538 does not directly contact the DNA, but it is supposed to stabilize larger bases like adenine or guanine. A fourth residue Asp582 stabilizes Arg584 by hydrogen bonding to maintain it in position for DNA binding. The 5' end points away from the HD2 domain and the 3' end points towards HD1 domain. Kuper *et al.* propose a model (Fig. 10) which shows how the ssDNA is situated and translocated in the protein.

5.3 Proposed model of damage verification

Based on these new data about XPD, Fuss and Tainer proposed an extended model of the damage verification process, the so-called ‘bind-pry-unwind’-model: (Fuss & Tainer, 2011) here XPB, XPD and CAK of the TFIIH complex are all considered to be essential for the coordination of damage recognition, verification, repair and signaling. In this model, the damage recognition proteins not only recruit the TFIIH complex and other NER factors to the damage site, but also fix the DNA at the same time. Kim *et al.* suggested that XPB facilitates the unwinding of the DNA duplex during transcription by distorting the

DNA underneath the transcription bubble, which is fixed by the transcription factors and RNAP II (Kim *et al.* 2000). Similarly XPB could unwind and open the DNA. In that way, the damage verification step and the DNA unwinding could be linked. These findings are in accord with the low helicase activity of XPB and its strong conformational change upon binding to DNA (Fan *et al.* 2006). Furthermore, TFIIH can thereby position itself directly next to XPC–RAD23B or RNAP II, which establishes a common mechanism of how XPB may act in GG-NER, TC-NER and transcription. Through its short binding canal, XPB could place TFIIH vertically to the DNA and 5' of the recognition protein. Then XPB could break the DNA strand, so that TFIIH bends it towards the DNA enabling XPD to bind about 22 base pairs away from the lesion on the damaged strand. Fuss and Tainer assume that the distance between XPB and XPD determines the size of the open DNA-bubble, corresponding to about 27 base pairs that are eliminated during NER. XPA catalyzes the release of CAK from the TFIIH complex (Coin *et al.* 2008). This step signals that the repair process can be continued and promotes the incision step. The XPD helicase unwinds the duplex until the 4Fe4S-cluster docks onto the damage and is blocked by it (Mathieu *et al.* 2010; Naegeli *et al.* 1993). Recent studies show that XPD exists in different conformations depending on the type of damage (Buechner *et al.* 2014). After the blockage, XPD and XPC–RAD23B anchor the TFIIH complex and build a platform for binding of ERCC1-XPF and XPG (Araujo *et al.* 2001).

Following this model RPA, XPA and XPG probably reach the lesion site independently and build together with TFIIH the pre-incision complex (Rademakers *et al.* 2003; Riedl *et al.* 2003; Volker *et al.* 2001; Wakasugi & Sancar, 1998). Here, all NER factors, except ERCC1-XPF, are present and stably bound to the open DNA bubble.

Recently, the Kisker group showed by mutational analysis that the FeS domain of XPD is crucial for NER (Kuper *et al.* 2014). Furthermore, the successful repair process is dependent on the enzymatic activities of XPD which involves DNA binding, ATPase activity, helicase activity and the correct interaction with other TFIIH subunits. In 2013, the Naively group showed by uncoupling the DNA unwinding activity from the damage verification activity of XPD, that residues Tyr192 and Arg196 form a sensor pocket located ahead of the narrow central protein pore responsible for damage recognition. By mutating these two residues, the helicase and ATPase activities are not affected but the damage sensing during ATP-driven scanning movement along nucleic acid lattices fails and the ability of XPD to form stable DNA damage recognition intermediates is significantly reduced (Mathieu *et al.* 2013). The activity of XPD however is not relevant for transcription and DNA binding.

6. XPA

The XPA protein (xeroderma pigmentosum group complementation group A) consists of 273 amino acids (31 kDa) and can occur as either a homodimer or in complex with other NER proteins (Liu *et al.* 2005; Yang *et al.* 2002). The protein is constituted of several domains (Camenisch *et al.* 2006; Camenisch & Nägeli, 2008): the N-terminal domain (XPA_{4–97}) includes the RPA34- and ERCC1-binding regions (He *et al.* 1995; Li *et al.* 1995a). The C-terminal domain (XPA_{226–273}) interacts with TFIIH (Camenisch & Nägeli, 2008; Ikegami *et al.* 1998; Park *et al.* 1995a). The central domain (XPA_{98–219}) is the minimal DBD (XPA-DBD) (Cleaver & States, 1997; Kuraoka *et al.* 1996). The protein contains a zinc-finger binding domain (Cys105-X₂-Cys108-X₁₇-Cys126-X₂-Cys129) (Hess *et al.* 1998). The C₄-zinc finger sustains the stability of the protein and plays a role in the interaction with RPA70 (Fig. 11a) (Ikegami *et al.* 1998; Kuraoka *et al.* 1996; Morikawa & Shirakawa, 2000; Morita *et al.* 1996; Sugitani *et al.* 2014). Mutation of one or all zinc-binding cysteines leads to a loss of function of XPA and to an almost complete loss of function of NER (Li *et al.* 1995b; Mellon *et al.* 1987).

It has been demonstrated that XPA binds ssDNA, dsDNA and denatured stretches of DNA in cooperation with RPA (Buschta-Hedayat *et al.* 1999; Missura *et al.* 2001; Wang *et al.* 2000; Yang *et al.* 2006). XPA shows binding affinity to irradiated DNA (Jones & Wood, 1993; Liu *et al.* 2005), DNA containing bulky adducts and cisplatin lesions (Jones & Wood, 1993; Liu *et al.* 2005), as well as to DNA duplexes containing loops and bulges (Camenisch *et al.* 2006). Camenisch *et al.* could show by mutational analyses that the residues Lys141 and Lys179 found on the surface of XPA next to the DNA-binding pocket are responsible for the sensing of DNA deformability as the common characteristics of damaged sites (Camenisch *et al.* 2006; Miyamoto *et al.* 1992). It was postulated that these two lysines interact with the DNA backbone, when XPA is in complex with RPA and kinked DNA (Read *et al.* 1993; Weir *et al.* 1993). Photocrosslinking experiments demonstrated that the mutated protein does interact with kinked DNA structures (Camenisch *et al.* 2006). This led to the assumption that the protein scans the DNA duplex for the presence of the kinked DNA structure with its positively charged pocket. The loop-rich domain of XPA was determined to be responsible for DNA binding. The ternary structure is stabilized by the zinc-finger domain (Miyamoto *et al.* 1992; Yan *et al.* 2002).

6.1 Nuclear magnetic resonance (NMR) structure of the DBD of human XPA

Ikegami *et al.* (1998) determined the NMR structure of the DBD of XPA (XPA_{98–219}, called XPA-DBD) and they found that the protein contains α -helices and β -sheets (Fig. 11b). The structure shows the presence of a positively charged (basic amino

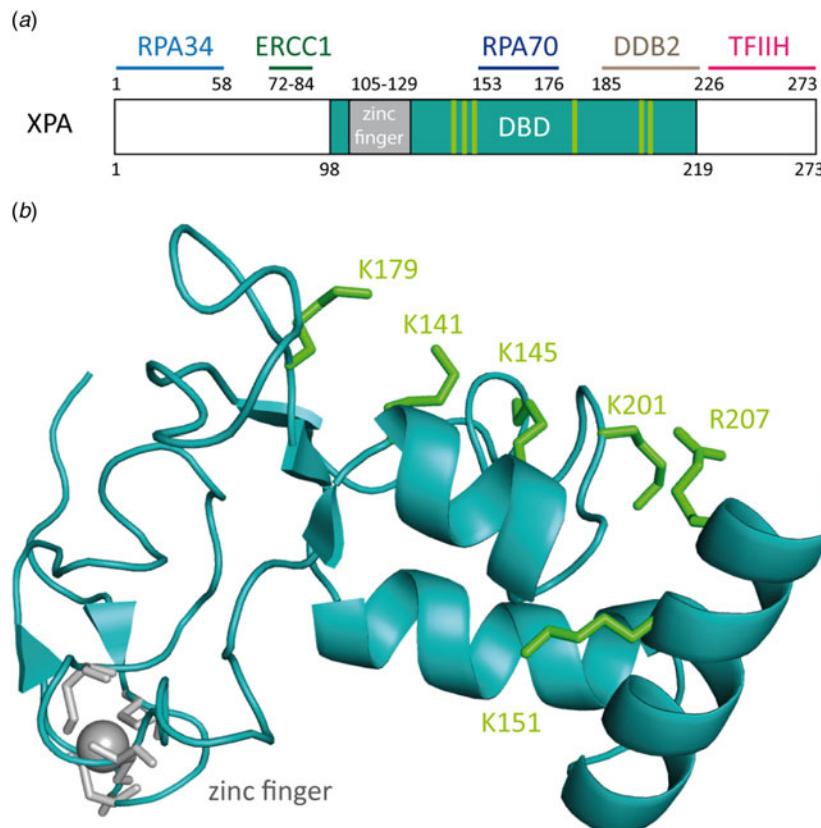


Fig. 11. (a) Schematic view of the human XPA protein with XPA-DBD (DNA-binding domain, XPA_{98–219}) in cyan the zinc-finger domain in gray. The binding domains of XPA with other NER proteins are depicted in light blue for RPA34, green for ERCC1, dark blue for RPA70 and pink for TFIIH. (b) NMR-structure of the XPA-DBD (Buchko *et al.* 2001; Camenisch *et al.* 2006; Ikegami *et al.* 1998) (PDB code: 1XPA). Positively charged amino acids that build the cleft are depicted as green sticks. The zinc atom is shown in gray.

acids) cleft with the right dimension to bind dsDNA (Buchko *et al.* 1998, 2001). This positively charged cleft was speculated to interact with the negatively charged DNA backbone.

Sugitani *et al.* (2014) redefined the DBD of XPA. By extending the originally characterized XPD-DBD at its C-terminus by 20 amino acids (XPA_{98–239}), they improved the DNA-binding properties. The new construct with the additional critical basic residues and helical elements at its C-terminus was found to bind to DNA substrates with the same affinity as the full-length protein. The group reported, based on NMR data, that XPA_{98–239} contains the same globular core as XPA_{98–219} and similarly undergoes a conformational change when binding to Y-shaped ss/ds-DNA junctions.

The rather small XPA protein interacts with a large number of other NER factors which supports its function as a scaffold protein for the assembly of the incision complex during repair (Gillet & Scherer, 2006; Naegeli & Sugasawa, 2011). Probably, XPA binds best to already prekinked DNA structures, which explains why XPA is able to bind defect Watson–Crick base pairing, created by mismatches, bubbles or three-way junctions (Bootsma & Hoeijmakers, 1993; Buschta-Hedayat *et al.* 1999; Missura *et al.* 2001). These structures can form easily kinked DNA structures. XPA shows a slight binding preference to UV-damaged DNA and bulky adducts over undamaged DNA (Jones & Wood, 1993; Robins *et al.* 1991; Thoma & Vasquez, 2003). Similarly to XPC, XPA recognizes 6–4PP lesions but not CPDs (Jones & Wood, 1993). CPD might block the bending of major grooves because of its stiff four-membered ring which points into the major groove. XPA shows a high affinity to DNA containing aromatic adducts and cisplatin crosslinks. Cisplatin crosslinks also kink the DNA structure. The damage recognition capability of XPA increases with RPA (Jones & Wood, 1993; Li *et al.* 1995a; Mellon *et al.* 1987) and ERCC1 (Nocentini *et al.* 1997; Saijo *et al.* 1996). XPA does not bind ssDNA, in agreement with the observed size of the positively charged cleft able to fit dsDNA (Buchko *et al.* 1998; Ikegami *et al.* 1998). For single-base bulky adduct lesions such as AAF-dG, it is unclear to which extent they prekink the duplex. Bulky adducts can increase the local conformational flexibility of DNA (Isaacs & Spielmann, 2004). XPA might employ a DNA flexibility scanning mechanism to create a kink, which could be stabilized by intercalation of the aromatic ring into the duplex and may be the secret behind the broad substrate

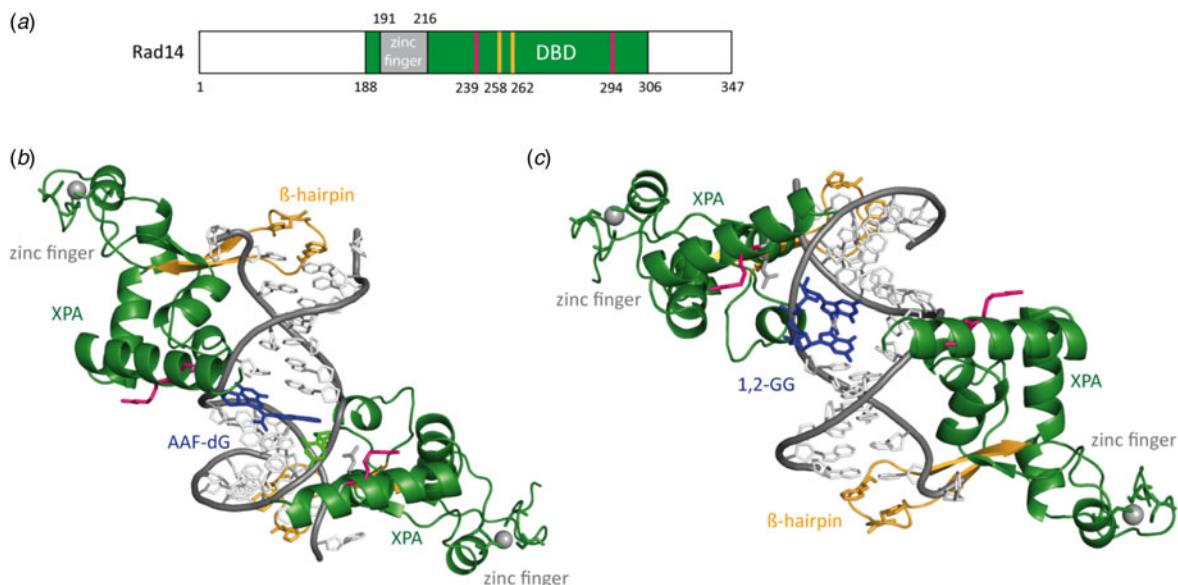


Fig. 12. (a) Schematic sequence of Rad14 with the DBD depicted in green, the zinc-finger domain in gray and the residues interacting with the DNA in orange and pink. (b) Crystal structure of Rad14 in complex with AAF-dG crosslinked DNA (PDB code: 5A3D) (Koch *et al.* 2015). Same color code as in (a). The AAF-dG residue is depicted in dark blue. (c) Crystal structure of Rad14 in complex with 1,2-GG cisplatin (PDB code: 5A39) (Koch *et al.* 2015). Same color code as in (a). The 1,2-GG cisplatin crosslink is depicted in dark blue.

promiscuity of NER. Koch *et al.* could obtain two crystal structures of Rad14 with a bulky lesion, which confirm this hypothesis (Koch *et al.* 2015).

6.2 Crystal structures of Rad14 with AAF-dG and cisplatin containing dsDNA

Recently, Koch *et al.* (2015) solved two crystal structures of the yeast homolog of XPA (Rad14) in complex with AAF-dG and 1,2-GG cisplatin containing DNA (Fig. 12b, c). The DBD of the protein (residues 188–306) was crystallized with damaged DNA (Fig. 12a). The structures allow us to uncover the molecular mechanism of XPA interacting with damaged DNA and are consistent with previous biochemical data showing the binding of XPA to kinked DNA structures (Buschta-Hedayat *et al.* 1999; Camenisch *et al.* 2006). The protein assumes an overall α/β -fold quite similar to its human homolog (Ikegami *et al.* 1998; Koch *et al.* 2015). Comparable with Rad4, Rad14 does not bind the lesion directly; instead it recognizes a weakened DNA duplex structure (Min & Pavletich, 2007). By the binding of two Rad14 proteins to the duplex, the DNA undergoes a sharp kink of 70° at the lesion site and single stranded regions are formed above and below the 13mer binding motif. Through insertion of a β -hairpin (His258 and Phe262) from each protein (acting like ‘fingers’) exactly six base pairs away from the lesion, a 13mer recognition motif is generated. Another domain composed of helix $\alpha 7$ (‘thumb’ containing Arg234 and Arg239) interacts with the DNA backbone close to the lesion, so that DNA bending takes place into the major groove. The lesion is located inside the 13mer duplex instead of being flipped out. Because of the relatively few interactions between Rad14 and the duplex, the protein might preferably bind already pre-kinked DNA structures (Camenisch *et al.* 2006). The 2:1 complex can coordinate contacts between TFIIH, DNA and RPA and recruit ERCC1–XPF (Li *et al.* 1995a; Park *et al.* 1995b; Tripsianes *et al.* 2007).

7. XPF–ERCC1

The human DNA repair endonuclease XPF (gene called ERCC4) is a protein of 916 amino acids (104 kDa) (Brookman *et al.* 1996). It forms a stable ERCC1–XPF heterodimer enzyme complex that participates in DNA repair and DNA recombination (Sijbers *et al.* 1996; Tsodikov *et al.* 2005). ERCC1 is a protein of 297 amino acids (32.5 kDa) that mediates DNA–protein and protein–protein interactions. The XPF–ERCC1 complex cuts specific structures near an ss/dsDNA junction at the 5'-end of single stranded damaged DNA strand and hence makes the first incision in the NER process. Its nuclease activity is absolutely essential in the NER process (de Laat *et al.* 1999; Sijbers *et al.* 1996). In the complex, XPF provides the endonuclease active site and is involved in DNA binding and establishing additional protein–protein interactions. XPF–ERCC1 is known to be the last

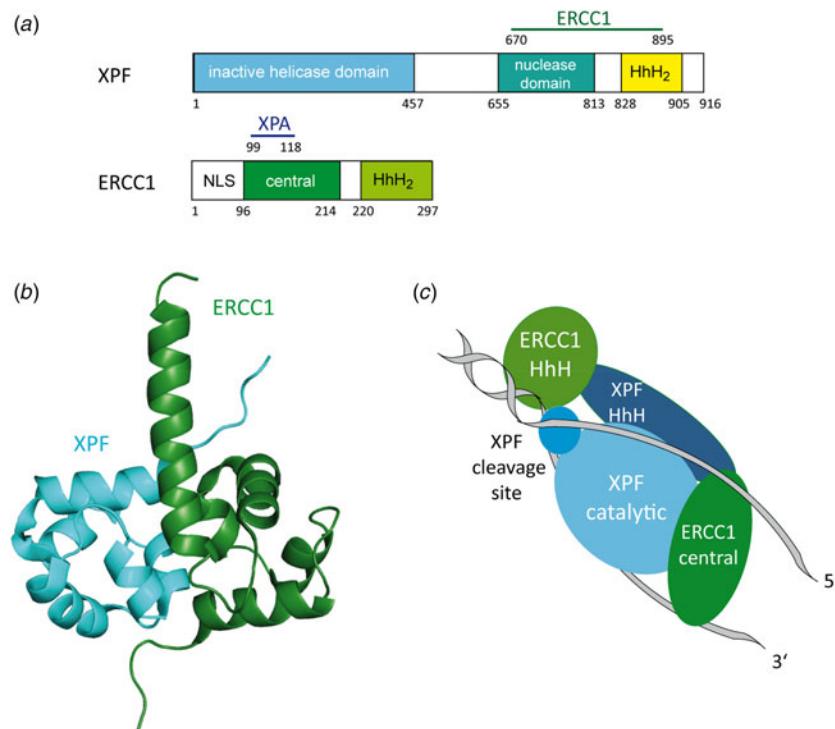


Fig. 13. (a) Structural domains of the human XPF and ERCC1 proteins. XPF interacts with ERCC1 (residues 670–895) depicted in green. The inactive helicase domain is depicted in light blue, the nuclease domain in cyan and the HhH₂ domain in yellow. ERCC1 interacts with XPA (residues 99–118) depicted in dark blue. The central domain is depicted in green and the HhH₂ domain in light green. (b) Structure of the human XPF–ERCC1 heterodimer (PDB code: 2A1J) (Tsodikov *et al.* 2005) with ERCC1 depicted in green and XPF depicted in light blue. (c) Model of the XPF–ERCC1 complex binding to a splayed DNA substrate (depicted in gray). The ERCC1 central domain and HhH domain are depicted in green. XPF, XPF cleavage site and HhH domain are depicted in light blue.

factor joining the NER complex before dual incision takes place (Riedl *et al.* 2003; Volker *et al.* 2001; Wakasugi & Sancar, 1998). Evidence exists that it is recruited by XPA (Fig. 13a) (Li *et al.* 1994, 1995a; Park & Sancar, 1994; Sajio *et al.* 1996).

7.1 Crystal structure of human XPF–ERCC1

Tsodikov *et al.* (2005) solved the crystal structure of the minimal dimerization domains of the human XPF and ERCC1 (Fig. 13b). The crystallized domain shows a similar endonuclease activity compared to that of the full-length complex. Both protein domains exhibit double helix–hairpin–helix motifs (HhH₂) related by a pseudo-2-fold symmetry axis (Tripsianes *et al.* 2005). The HhH₂ motifs are known to be ds- as well as ss-DNA-binding motifs (Newman *et al.* 2005; Singh *et al.* 2002). Both C-terminal HhH₂ domains of XPF and ERCC1 are essential and sufficient for a stable heterodimer formation (de Laat *et al.* 1998). XPF might play a role as a scaffold protein, allowing proper folding of the ERCC1 domain. All XPF family members contain a catalytic domain comprising the active site motive (GDX_nERKX₃D) and they need a divalent cation for their nuclelease activity (Nishino *et al.* 2003; Sijbers *et al.* 1996).

7.2 NMR structure of C-terminal human XPF in complex with ssDNA

Das *et al.* (2012) solved the NMR structure of the C-terminal domain of the human XPF protein (residues 823–905) in complex with a 10-nucleotide ssDNA fragment. In earlier NMR studies of the apo-protein, the group showed that the XPF homodimer complex is more stable than the XPF–ERCC1 heterodimer complex (Das *et al.* 2008). The structure shows that a positively charged region of one HhH domain interacts with the ssDNA phosphate backbone. The central guanine (G5) is extruded from the duplex and inserted into a pocket formed by residues from both HhH motifs of XPF (Das *et al.* 2008; Tripsianes *et al.* 2005). This cavity offers a large interaction surface area for the protein with the DNA. The DNA forms a right-handed structure binding to the protein, which mainly interacts with the ssDNA phosphate backbone (Fig. 14). In addition, the protein interacts with three bases: thymidine3 (His857), thymidine4 (Ser875, Ile876) and G5 (Leu877, Gly878 and Asn879). Combined with previous data (Tripsianes *et al.* 2005) the group proposes a binding model of XPF–ERCC1 to ss/dsDNA junctions (Das *et al.* 2012). They modeled the XPF–ERCC1 complex bound to a splayed-arm DNA substrate showing

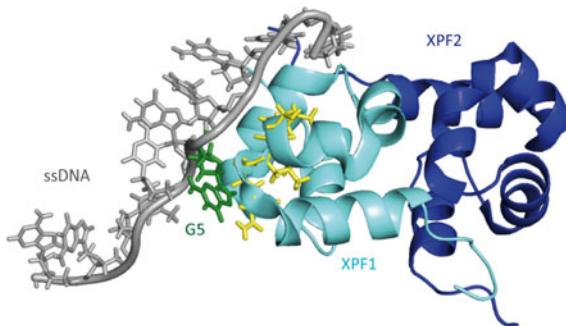


Fig. 14. NMR structure of the C-terminal domain of human XPF in complex with an ssDNA fragment (PDB code: 2KN7) (Das *et al.* 2012). The first XPF (XPF1) protein is depicted in light blue showing the interacting residues with the DNA in yellow. The second XPF (XPF2) protein is depicted in light blue. The central guanine (G5) is depicted in green.

that the conserved HhH domain of ERCC1 contacts the phosphate backbone of dsDNA in the minor groove with its two hairpins H1 and H2 while the XPF HhH domain stabilizes the ssDNA via interactions with a 5'/3' polarity (Fig. 13c). The damaged strand gets cleaved by the catalytic residues of the XPF nuclease domain and the ERCC1 central domain connects with XPA and ssDNA on the NER complex (de Laat *et al.* 1998).

8. XPG

XPG (gene called ERCC5) is a single-stranded structure-specific DNA 3'-endonuclease involved in NER. It belongs to the family of FEN-1 endonucleases (Flap endonuclease) (Lieber, 1997). The human protein consists of 1186 amino acids (133 kDa) (O'Donovan *et al.* 1994b) and it interacts with several other NER proteins (Fig. 15). Mutations in this gene cause the XP-G phenotype, characterized by hypersensitivity to UV light and increased susceptibility for skin cancer development following UV exposure. Furthermore, it can lead to the CS characterized by severe growth defects and mental retardation (Constantinou *et al.* 1999). The XPG protein cleaves DNA near junctions of dsDNA to ssDNA with a defined polarity moving 3'-5' away from the junction. It has been observed that XPG needs a duplex opening of more than five base pairs for efficient cleavage (Evans *et al.* 1997a). The protein is known to interact with XPB, XPD, RPA and PCNA (Feltes & Bonatto, 2015; Iyer *et al.* 1996). XPG is needed for the proper recruitment of XPF-ERCC1 to the preincision complex. Once the 5'-incision is made by the XPF-ERCC1 nuclease, XPG becomes fully active (Staresinicic *et al.* 2009). Despite this knowledge, the exact mechanism of substrate recognition by XPG is still unclear.

The human protein possesses a globular structure like the FEN-1 endonuclease (Hosfield *et al.* 1998; Hwang *et al.* 1998b). It contains an N-terminal domain and an I-region linked by a 600 residue long spacer region (Dunand-Sauthier *et al.* 2005; Hohl *et al.* 2007). The N-terminal and the I-region are evolutionary conserved among similar endonucleases (T4 RNaseH, Taq Polymerase). The spacer region is not involved in catalysis. It is responsible for proper DNA cleavage and for the recruitment and positioning of XPG to the lesion site (Constantinou *et al.* 1999; Dunand-Sauthier *et al.* 2005; He *et al.* 1995; Thorel *et al.* 2004). It is also known to interact with RPA and TFIIH. The protein contains a D1 box (residues 184–210) and a D2 box (residues 890–984) which are evolutionary conserved (Dunand-Sauthier *et al.* 2005; Houle & Friedberg, 1999). One cation is required for the endonuclease activity in the active site. The protein shows a higher affinity for Mn²⁺ than for Mg²⁺ and uses the more abundant one for catalysis (Hohl *et al.* 2007; Iyer *et al.* 1996). Two residues in the active site (Glu791 and Asp812) coordinate the divalent cations (Constantinou *et al.* 1999) (light blue sticks in Fig. 15). The protein incises flap DNA and bubbled DNA structures with a defined polarity (Evans *et al.* 1997b; Hohl *et al.* 2003; O'Donovan *et al.* 1994a; Tsutakawa *et al.* 2011).

8.1 Crystal structure of the yeast XPG (Rad2) in complex with dsDNA

To understand the mechanism of DNA binding by XPG Mietus *et al.* (2014) solved in 2014 the crystal structure of the catalytic core of the *Saccharomyces cerevisiae* homolog of XPG (Rad2-ΔSC, Fig. 16a) in complex with four different DNA substrates. For crystallization a XPG variant missing the mostly disordered spacer residues 111–730 between the two sequence segments N and I of the nuclease domain was used. The double-stranded oligonucleotides (12–15 bp) were designed with single-stranded overhangs constructed with one to six adenines or thymidines. In the crystal structure two Rad2 proteins bind to one DNA substrate with a single-protein molecule covering 13 bp of the double-stranded portion. Rad2 utilizes



Fig. 15. Schematic representation of human XPG with its different domains and the region interacting with other NER proteins. The D1 box is depicted in light green, the D2 box in green and the 600 aa spacer region in orange. The XPB and XPD interacting region is highlighted in purple, RPA interacting region in dark blue and the PCNA interacting region in pink. Glu791 and Asp812 in the active sites are depicted as light blue bars.

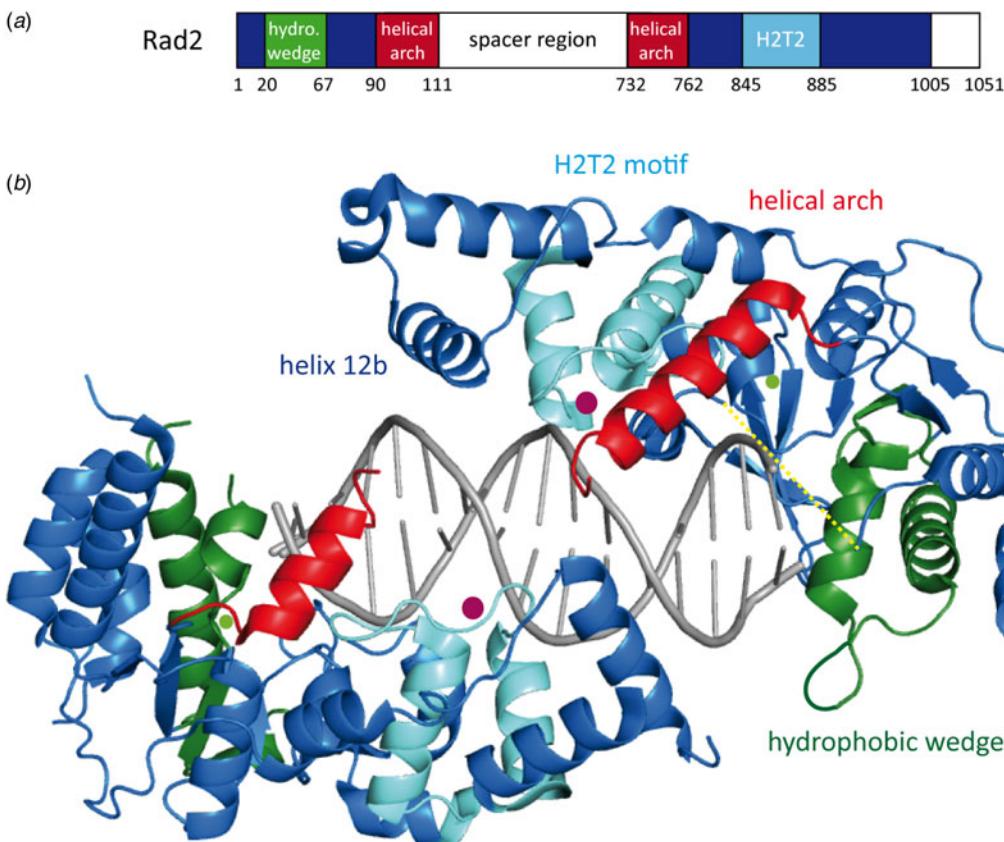


Fig. 16. Crystal structure of the Rad2–DNA complex (PDB code: 4Q0W) (Mietus *et al.* 2014). (a) Schematic representation of the protein. The hydrophobic wedge is depicted in green, the helical arch in red, the H2TH domain in light blue and the rest of the crystallized structure in dark blue. (b) Structure of Rad2 with dsDNA. Same color code as in (a). Calcium is depicted in light green and potassium in purple. The helix 12b is named as referred by Tsutakawa *et al.* (2011). The spacer region omitted in the crystal structure is depicted as a yellow dashed line.

three structural domains to recognize the double-stranded portion of the DNA but it does not specifically recognize the single-stranded portion of the duplex. The structures reveal that the protein recognizes a single 5'-nucleotide of the single-stranded portion of the DNA and a 3'-phosphate group of the ds/ssDNA junction. The helix two turn helix (H2TH) motif including a Rad2-specific α -helix 12b and a hydrophobic wedge, which binds the first base pair of the double-stranded portion of the DNA, is responsible for binding to DNA. The helical arch may assume a different structure in Rad2 because its active site is more accessible compared with related enzymes such as FEN1 and EXO1 (Mietus *et al.* 2014). This feature allows the protein to form a gap and create an exit route from the active site. As expected, the catalytic core of Rad2 is quite comparable with other flap nucleases (Orans *et al.* 2011). The protein has an oblong shape with its active site situated in the middle. A twisted central β -sheet composed of seven strands flanked by α -helices build the central core of the protein (Fig. 16b). Gln37 forms a hydrogen bond with the base of the cleaved strand and represents an important residue of substrate recognition by Rad2. The H2TH module is responsible for the key protein–DNA interactions through its Rad2-specific helix α 12b and the hydrophobic wedge contacting the ss/dsDNA junction.

9. XPV

The repair of lesions by NER is a powerful method to clean the genome from the plethora of lesions that are generated daily. NER is hence the prime pathway that allows cells to cope with DNA damage. Because lesions can occur at any time in cellular DNA, even when the NER complex cannot assemble properly, e.g. while the cell is undergoing cell division, a second system has evolved that allows cells to cope with lesions. This system does not repair lesions but is able to give cells the ability to tolerate lesions, so that repair can be shifted to later times. This lesion tolerance process (called TLS) requires the action of special lesion tolerance polymerases. Among them the polymerase η (XPV) is of prime importance, because it provides tolerance of dinucleotide lesions such as CPDs and cisplatin adducts, which are the prime repair targets for NER. XPV (Polymerase η) is a Y-family DNA polymerase responsible for TLS enabling replication through bulky DNA adducts like cisplatin, UV-induced CPDs and 6–4PP as well as acetylaminofluorene adducts and oxidative lesions (like 8-oxo-G) (Alt *et al.* 2007; Biertumpfel *et al.* 2010; Haracska *et al.* 2000; Reissner *et al.* 2010; Schorr *et al.* 2010b; Silverstein *et al.* 2010a, b). The human protein consists of 713 amino acids with a size of 78 kDa (Masutani *et al.* 1999a). Indeed, mutation in the XPV gene leads to just another form of XP called the variant type also associated with a high risk of developing skin cancer (Johnson *et al.* 2000; Lehmann *et al.* 1975; Masutani *et al.* 1999b). Lesion tolerance, however, has a major drawback. Because the polymerases are able to read through lesions they also have a strongly reduced replication fidelity, which gives rise to mutations. That is why repair is always the first and preferred defense pathway.

9.1 Crystal structure of the yeast XPV [Rad30] protein

Kondratick *et al.* (2001) solved the crystal structure of the catalytic core (1–513 amino acids) of the *S. cerevisiae* apo-enzyme (Fig. 17b). The catalytic core possesses the same bypass activity as the full-length protein (1–632 amino acids). The structure shows that the polymerase possesses a right-hand shape built up by four domains: thumb, fingers, PAD-domain (polymerase-associated domain also named ‘little fingers’ (LF)) and the palm domain (Fig. 17a). Characterized by a wide open and possibly flexible active site, the polymerase is able to replicate over bulky dinucleotide lesions, however, at the cost of misincorporation and only a low processivity in contrast to high-fidelity polymerases. Because of the helix deforming action of lesions like intrastrand crosslinks (CPD and 1,2-GG) their bypass by high-fidelity DNA polymerases is impossible. Being able to bypass cisplatin crosslinks, XPV enables cells to gain resistance against cisplatin-based chemotherapy.

9.2 Crystal structure of the yeast XPV [Rad30] in complex with 1,2-GG cisplatin

Alt *et al.* (2007) solved the crystal structure of XPV replicating over 1,2-GG containing DNA. This structure provided insights into the molecular mechanism that allows replication of XPV across strongly distorting DNA lesions. 1,2-GG and 1,3-GTG cisplatin adducts are formed in a typical cancer therapy with cisplatin. In the structure, XPV accommodates the full 1,2-GG adduct in its active site where it properly base pairs the 3' G of the 1,2-GG lesion with a dCTP (Fig. 18). For the nucleotidyl transfer, the DNA rotates into an active conformation, supported by hydrogen bonding of the temple base to the dCTP. The bypass of the 3'-dG of 1,2-GG is more efficient and accurate than the bypass of the 5'-dG. In the palm domain, Asp30 and Asp155 coordinate the divalent cations in the active site and Glu156 is responsible for the catalysis. Glu39 plays a structural role too, because it maintains the integrity of the fingers domain. Arg73 coordinates the incoming dNTP for the lesion bypass step and stacks on top of the base moiety. Two magnesium ions are coordinated by the catalytically essential residues.

9.3 Crystal structure of the yeast XPV [Rad30] in complex with the CPD lesion

Silverstein *et al.* (2010b) determined the crystal structure of yeast Rad30 in complex with a *cis-syn* T-T dimer lesion. These authors argue that crystal contacts between the palm and the PAD domain of two neighboring proteins in the crystal lattice influence the binding of the protein to the DNA. The group performed a double-mutation (K140A, S144W) to eliminate these contacts, and indeed crystals were obtained with a different space group. Deeper analysis of the structure shows that the finger domain is a little more closed but the situation in the active site is almost unchanged compared with the yeast structure (Alt *et al.* 2007). Figure 19 shows the XPV structure with its different domains, the containing T-T dimer and the incoming dATP. Again the large active site accommodates the dinucleotide lesion very similar to what was already shown by Alt *et al.* for the cisplatin lesion. The 3'T of the dimer interacts with the incoming dATP by Watson–Crick hydrogen bonding. The 5'T of the lesion builds van der Waals contacts with Met74 and Arg73 of the finger domain. The active site residues of the palm domain (Asp30, Asp155 and Glu156) catalyze the nucleotidyl transfer reaction. By mutational analysis and biochemical assays Silverstein *et al.* could show that the residues Gln55, Arg73 and Met74 interact with the T-T dimer to keep the lesion in the active site.

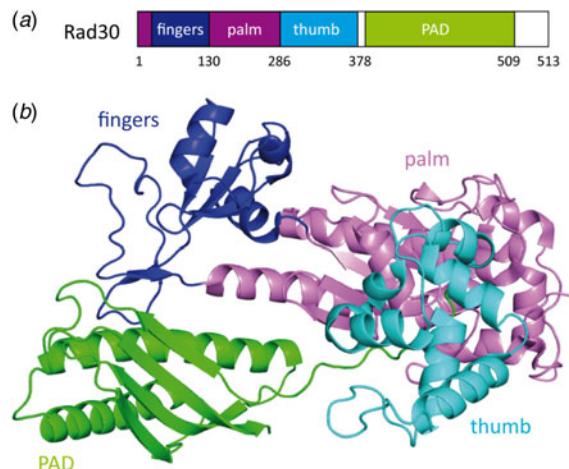


Fig. 17. (a) Structural domains of yeast XPV (Rad30) (Silverstein *et al.* 2010b). (b) Structure of the catalytic domain of *S. cerevisiae* XPV showing the four domains (PDB code: 1J1H) (Trincao *et al.* 2001). The fingers domain is depicted in dark blue, the PAD domain in green, the ThM in cyan and the palm domain in purple.

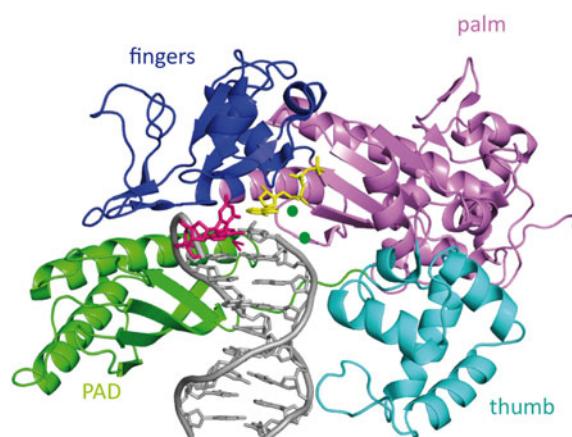


Fig. 18. Crystal structure of Rad30 in complex with 1,2-GG cisplatin adduct (PDB code: 2R8K) (Alt *et al.* 2007). Rad30 has right-hand structure containing four domains: palm (violet), thumb (cyan), finger domain (dark blue) and the additional PAD domain (green) which is unique for Y-family polymerases. The cisplatin adduct is depicted in pink, the dCTP in yellow and the two metal ions in green.

9.4 Crystal structure of the human XPV in complex with the CPD lesion

Biertumpfel *et al.* (2010) reported the crystal structure of the catalytic core (residues 1–432) of human polymerase η at four consecutive steps during DNA synthesis through *cis-syn* CPD lesion (Fig. 20b). They could show on a molecular level how human XPV efficiently bypasses the lesion in TLS. Human XPV, similar to its yeast homolog, consists of four domains: palm, PAD (little fingers), thumb and the fingers (Fig. 20a). The active site is situated in the palm domain and the DNA is bound between the thumb and the finger domains. Acting as a molecular splint the polymerase keeps the damaged DNA straight and rigid in its B-form conformation in contrast to the bent and unwound structures of CPD containing DNA with other repair enzymes or alone. The β -strand in the little finger is nearly parallel to the template strand and hydrogen bonding of the main chain amides to the phosphate supports this conformation. Again, the most prominent feature is the large active site of the protein already reported by Alt *et al.* (2007), which is able to accommodate two bases of the template strand instead of one. The 3' thymine of CPD is base paired with the incoming dATP, while the 5' thymine is turned and moved closer to the finger domain which leads to tighter interactions than the ones observed with undamaged DNA. Furthermore, the two residues Arg61 and Glu38 that are conserved among the XPV homologs create specific hydrogen bonds with the incoming dATP and the CPD lesion. Again, the human XPV shows a more closed finger domain, which contacts the replicating base pair more intensively. Based on all these structures we can therefore assume that the finger domain is quite flexible and that the more closed structure is the conformation the protein assumes in the processive mode.

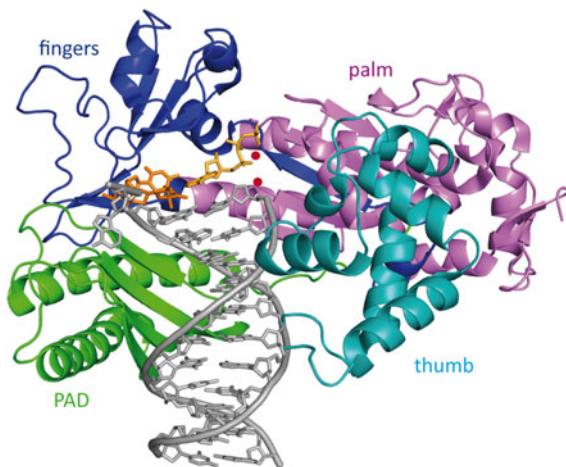


Fig. 19. Structure of Rad30 with a CPD lesion. The *cis*-syn T–T dimer is depicted in orange and the incoming dATP is shown in yellow (PDB code: 3MFI) (Silverstein *et al.* 2010b). The four domains are depicted with different colors: fingers (blue), palm (violet), PAD (green) and thumb (light blue). The two Mg^{2+} ions in the active site are colored in red.

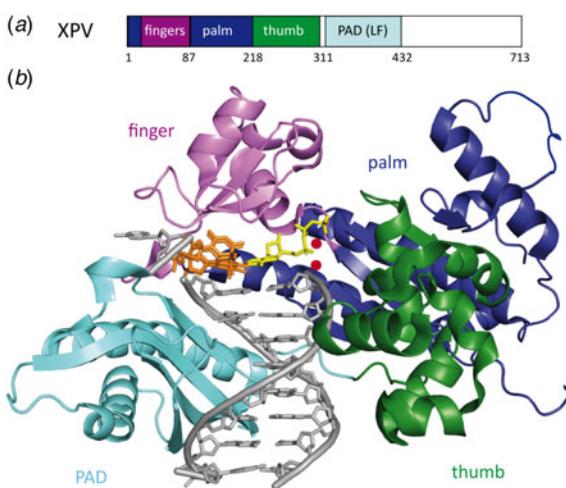


Fig. 20. (a) Schematic representation of human XPV showing the different domains. The finger domain is depicted in purple, the palm domain in dark blue, the ThM in green and PAD domain (little fingers) in cyan. (b) Crystal structure of human XPV in complex with CPD damaged DNA (PDB code: 3MR3) (Biertumpfel *et al.* 2010). Same color code as in (a). The CPD lesion is shown in orange and the dAMPNPP (2'-deoxy-adenosine-5'-[α,β]-imido]-triphosphate) is shown in yellow. The two Mg^{2+} ions of the active site are depicted in red.

Further crystal structures of human and yeast XPV could be resolved with different DNA damages such as 1,3-GTG cisplatin or 8-oxoG (Reissner *et al.* 2010; Silverstein *et al.* 2010a). Schorr & Carell (2010a) crystallized the *S. cerevisiae* XPV (Rad30) with an AAF-dG lesion and they proposed a model for lesion bypass. *In vivo* experiments show that Pol η indeed replicates through acetylated aromatic amine lesions such as AAF-dG (Bresson & Fuchs, 2002). *In vitro* studies revealed that both human and yeast XPV base pair the AAF-dG lesion correctly with dC (Bresson & Fuchs, 2002; Livneh, 2001; Masutani *et al.* 2000; Matsuda *et al.* 2000) and are able to catalyze the extension from the lesion, however, with reduced efficiency (Masutani *et al.* 2000). In addition, XPV also incorporates a correct dCTP opposite the N^2 -AAF-dG adduct (Yasui *et al.* 2004). It has also been shown that XPV is responsible for the occurrence of frameshift mutations in response to the presence of aromatic amine lesions during replication (Bresson & Fuchs, 2002; Cordonnier *et al.* 1999).

10. Conclusion

In this review, we presented the most recent eukaryotic structures of the XP proteins in complex with their DNA substrate. When no co-crystal structures with the DNA substrate were obtained we discussed the models of protein–DNA interaction

that were proposed. All these structures led to a dramatic increase of our understanding of the role of the individual XP proteins in DNA damage repair and tolerance. The structures allow to determine binding partner regions and to elucidate on a molecular level the interaction between the proteins and their DNA substrates. Now, further studies have to be done in order to better understand the correlation between genetic defects and the role of the various XP proteins in NER and TLS. We also need further human structural information and most importantly structures of the XP proteins in functional complexes with the other involved NER proteins. This would allow deciphering of the protein–protein interactions on the DNA damage site and it may shed light on the correlation to genetic disorders. Understanding the mechanism of NER and TLS in correlation with human genetic disorders unveils the biological basis for oncogenesis and aging. Last but not least, proteomics data are needed to uncover the functional relations of the proteins involved in replication, transcription and repair.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft SFB-646, SFB-749 and the excellence cluster CIPS^M for financial support. We also thank Dr. Jochen Kuper, Wolfgang Koelmel, Dr. Dorothea Matschkal, Dr. Iakovos N. Michaelides and Dr. Sabine Schneider for critical reading of the manuscript. We thank Dr. Markus Mueller for helpful advice.

References

- ABDULRAHMAN, W., ILTIS, I., RADU, L., BRAUN, C., MAGLOTT-ROTH, A., GIRAUDON, C., EGLY, J. M. & POTERSZMAN, A. (2013). ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E633–E642.
- ABOUSSEKHA, A., BIGGERSTAFF, M., SHIVJI, M. K., VILPO, J. A., MONCOLLIN, V., PODUST, V. N., PROTIC, M., HUBSCHER, U., EGLY, J. M. & WOOD, R. D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* **80**, 859–868.
- ALT, A., LAMMENS, K., CHIOCCHINI, C., LAMMENS, A., PIECK, J. C., KUCH, D., HOPFNER, K. P. & CARELLI, T. (2007). Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. *Science* **318**, 967–970.
- ARAKI, M., MASUTANI, C., TAKEMURA, M., UCHIDA, A., SUGASAWA, K., KONDOKI, J., OHKUMA, Y. & HANAOKA, F. (2001). Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. *The Journal of Biological Chemistry* **276**, 18665–18672.
- ARAUJO, S. J., NIGG, E. A. & WOOD, R. D. (2001). Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairstosome. *Molecular and Cellular Biology* **21**, 2281–2291.
- ARAUJO, S. J., TIRODE, F., COIN, F., POSPIECH, H., SYVAOJA, J. E., STUCKI, M., HUBSCHER, U., EGLY, J. M. & WOOD, R. D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes & Development* **14**, 349–359.
- ASAHIKA, H., KURAOKA, I., SHIRAKAWA, M., MORITA, E. H., MIURA, N., MIYAMOTO, I., OHTSUKA, E., OKADA, Y. & TANAKA, K. (1994). The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage. *Nature Structural & Molecular Biology* **315**, 229–237.
- BATTY, D., RAPIC'-OTRIN, V., LEVINE, A. S. & WOOD, R. D. (2000a). Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *Journal of Molecular Biology* **300**, 275–290.
- BATTY, D. P. & WOOD, R. D. (2000b). Damage recognition in nucleotide excision repair of DNA. *Gene* **241**, 193–204.
- BERNARDES DE JESUS, B. M., BJORAS, M., COIN, F. & EGLY, J. M. (2008). Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. *Molecular and Cellular Biology* **28**, 7225–7235.
- BERNEBURG, M. & LEHMANN, A. R. (2001). Xeroderma pigmentosum and related disorders: defects in DNA repair and transcription. *Advances in Genetics* **43**, 71–102.
- BIERTUMPFEL, C., ZHAO, Y., KONDO, Y., RAMON-MAIQUES, S., GREGORY, M., LEE, J. Y., MASUTANI, C., LEHMANN, A. R., HANAOKA, F. & YANG, W. (2010). Structure and mechanism of human DNA polymerase eta. *Nature* **465**, 1044–1048.
- BOOTSMA, D. & HOEIJMakers, J. H. (1993). DNA repair. Engagement with transcription. *Nature* **363**, 114–115.
- BRESSON, A. & FUCHS, R. P. (2002). Lesion bypass in yeast cells: Pol eta participates in a multi-DNA polymerase process. *The EMBO Journal* **21**, 3881–3887.
- BROOKMAN, K. W., LAMERDIN, J. E., THELEN, M. P., HWANG, M., REARDON, J. T., SANCAR, A., ZHOU, Z. Q., WALTER, C. A., PARRIS, C. N. & THOMPSON, L. H. (1996). ERCC4 (XPF) encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Molecular and Cellular Biology* **16**, 6553–6562.
- BROOKS, P. J., WISE, D. S., BERRY, D. A., KOSMOSKI, J. V., SMERDON, M. J., SOMERS, R. L., MACKIE, H., SPOONDE, A. Y., ACKERMAN, E. J., COLEMAN, K., TARONE, R. E. & ROBBINS, J. H. (2000). The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. *The Journal of Biological Chemistry* **275**, 22355–22362.
- BUCHKO, G. W., NI, S., THRALL, B. D. & KENNEDY, M. A. (1998). Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA. *Nucleic Acids Research* **26**, 2779–2788.

- BUCHKO, G. W., TUNG, C. S., MCATEER, K., ISERN, N. G., SPICER, L. D. & KENNEDY, M. A. (2001). DNA-XPA interactions: a ^{31}P NMR and molecular modeling study of dCCAATAACC association with the minimal DNA-binding domain (M98-F219) of the nucleotide excision repair protein XPA. *Nucleic Acids Research* **29**, 2635–2643.
- BUECHNER, C. N., HEIL, K., MICHELS, G., CARELL, T., KISKER, C. & TESSMER, I. (2014). Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility. *The Journal of Biological Chemistry* **289**, 3613–3624.
- BURNS, J. L., GUZDER, S. N., SUNG, P., PRAKASH, S. & PRAKASH, L. (1996). An affinity of human replication protein A for ultraviolet-damaged DNA. *The Journal of Biological Chemistry* **271**, 11607–11610.
- BUSCHTA-HEDAYAT, N., BUTERIN, T., HESS, M. T., MISSURA, M. & NAEGELI, H. (1999). Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6090–6095.
- BUTERIN, T., HESS, M. T., LUNEVA, N., GEACINTOV, N. E., AMIN, S., KROTH, H., SEIDEL, A. & NAEGELI, H. (2000). Unrepaired Fjord region polycyclic aromatic hydrocarbon-DNA adducts in ras codon 61 mutational hot spots. *Cancer Research* **60**, 1849–1856.
- CAMENISCH, U., DIP, R., SCHUMACHER, S. B., SCHULER, B. & NAEGELI, H. (2006). Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nature Structural and Molecular Biology* **13**, 278–284.
- CAMENISCH, U. & NÄGELI, H. (2008). XPA gene, its product and biological roles. *Advances in Experimental Medicine and Biology* **637**, 28–38.
- CAMENISCH, U., TRAUTLEIN, D., CLEMENT, F., FEI, J., LEITENSTORFER, A., FERRANDO-MAY, E. & NAEGELI, H. (2009). Two-stage dynamic DNA quality check by xeroderma pigmentosum group C protein. *EMBO Journal* **28**, 2387–2399.
- CHEN, X., VELMURUGU, Y., ZHENG, G., PARK, B., SHIM, Y., KIM, Y., LIU, L., VAN HOUTEN, B., HE, C., ANSARIA, A. & MIN, J. (2015). Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. *Nature Communications* **6**, 5849.
- CHU, G. & CHANG, E. (1988). Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. *Science* **242**, 564–567.
- CITTERIO, E., VAN DEN BOOM, V., SCHNITZLER, G., KANAAR, R., BONTE, E., KINGSTON, R. E., HOEIJMAKERS, J. H. & VERMEULEN, W. (2000). ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor. *Molecular and Cellular Biology* **20**, 7643–7653.
- CLEAVER, J. E. (2000). Common pathways for ultraviolet skin carcinogenesis in the repair and replication defective groups of xeroderma pigmentosum. *Journal of Dermatological Science* **23**, 1–11.
- CLEAVER, J. E., LAM, E. T. & REVET, I. (2009). Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nature Reviews Genetics* **10**, 756–768.
- CLEAVER, J. E. & STATES, J. C. (1997). The DNA damage-recognition problem in human and other eukaryotic cells: the XPA damage binding protein. *Biochemical Journal* **328**, 1–12.
- COIN, F., MARINONI, J. C., RODOLFO, C., Fribourg, S., PEDRINI, A. M. & EGGLY, J. M. (1998). Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nature Genetics* **20**, 184–188.
- COIN, F., OKSENYCH, V. & EGGLY, J. M. (2007). Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Molecular Cell* **26**, 245–256.
- COIN, F., OKSENYCH, V., MOCQUET, V., GROH, S., BLATTNER, C. & EGGLY, J. M. (2008). Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Molecular Cell* **31**, 9–20.
- CONSTANTINOU, A., GUNZ, D., EVANS, E., LALLE, P., BATES, P. A., WOOD, R. D. & CLARKSON, S. G. (1999). Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair. *The Journal of Biological Chemistry* **274**, 5637–5648.
- CORDONNIER, A. M., LEHMANN, A. R. & FUCHS, R. P. (1999). Impaired translesion synthesis in xeroderma pigmentosum variant extracts. *Molecular and Cellular Biology* **19**, 2206–2211.
- DAS, D., FOLKERS, G. E., VAN DIJK, M., JASPERS, N. G., HOEIJMAKERS, J. H., KAPTEIN, R., BOELENS, R. & FOLKERS, G. E. (2012). The structure of the XPF-ssDNA complex underscores the distinct roles of the XPF and ERCC1 helix-hairpin-helix domains in ss/ds DNA recognition. *Structure* **20**, 667–675.
- DAS, D., TRIPSIANES, K., JASPERS, N. G., HOEIJMAKERS, J. H., KAPTEIN, R., BOELENS, R. & FOLKERS, G. E. (2008). The HhH domain of the human DNA repair protein XPF forms stable homodimers. *Proteins* **70**, 1551–1563.
- DE LAAT, W. L., APPEDOORN, E., SUGAWARA, K., WETERINGS, E., JASPERS, N. G. J. & HOEIJMAKERS, J. H. J. (1998). DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes & Development* **12**, 2598–2609.
- DE LAAT, W. L., JASPERS, N. G. & HOEIJMAKERS, J. H. (1999). Molecular mechanism of nucleotide excision repair. *Genes & Development* **13**, 768–785.
- DE WEERD-KASTELEIN, E. A., KEIJZER, W. & BOOTSMA, D. (1972). Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization. *Nature: New Biology* **238**, 80–83.
- DEMPEL, B. & HARRISON, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. *Annual Review of Biochemistry* **63**, 915–948.
- DONAHUE, B. A., YIN, S., TAYLOR, J. S., REINES, D. & HANAWALT, P. C. (1994). Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 8502–8506.
- DUBAELE, S., PROIELLI DE SANTIS, L., BIENSTOCK, R. J., KERIEL, A., STEFANINI, M., VAN HOUTEN, B. & EGGLY, J. M. (2003). Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Molecular Cell* **11**, 1635–1646.
- DUNAND-SAUTHIER, I., HOHL, M., THOREL, F., JAQUIER-GUBLER, P., CLARKSON, S. G. & SCHARER, O. D. (2005). The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. *The Journal of Biological Chemistry* **280**, 7030–7037.
- EDENBERG, H. & HANAWALT, P. (1972). Size of repair patches in the DNA of ultraviolet-irradiated HeLa cells. *Biochimica et Biophysica Acta* **272**, 361–372.

- EL-MAHDY, M. A., ZHU, Q., WANG, Q. E., WANI, G., PRAETORIUS-IBBA, M., WANI, A. A. (2006). Cullin 4A-mediated proteolysis of DDB2 protein at DNA damage sites regulates *in vivo* lesion recognition by XPC. *Journal of Biological Chemistry* **281**, 13404–13411.
- EVANS, E., FELLOWS, J., COFFER, A. & WOOD, R. D. (1997a). Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *The EMBO Journal* **16**, 625–638.
- EVANS, E., MOGGS, J. G., HWANG, J. R., EGLY, J. M. & WOOD, R. D. (1997b). Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *The EMBO Journal* **16**, 6559–6573.
- FAGBEMI, A. F., ORELLI, B. & SCHÄRER, O. D. (2011). Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair* **10**, 722–729.
- FAN, L., ARVAI, A. S., COOPER, P. K., IWAI, S., HANAOKA, F. & TAINER, J. A. (2006). Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. *Molecular Cell* **22**, 27–37.
- FAN, L., FUSS, J. O., CHENG, Q. J., ARVAI, A. S., HAMMEL, M., ROBERTS, V. A., COOPER, P. K. & TAINER, J. A. (2008). XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell* **133**, 789–800.
- FELTES, B. C. & BONATTO, D. (2015). Overview of xeroderma pigmentosum proteins architecture, mutations and post-translational modifications. *Mutation Research/Reviews in Mutation Research* **763**, 306–320.
- FISCHER, E. S., SCRIMA, A., BOHM, K., MATSUMOTO, S., LINGARAJU, G. M., FAY, M., YASUDA, T., CAVADINI, S., WAKASUGI, M., HANAOKA, F., IWAI, S., GUT, H., SUGASAWA, K. & THOMA, N. H. (2011). The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024–1039.
- FITCH, M. E., NAKAJIMA, S., YASUI, A. & FORD, J. M. (2003). *In vivo* recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *The Journal of Biological Chemistry* **278**, 46906–46910.
- FRIEDBERG, E. C. (1995). *DNA Repair and Mutagenesis*. Washington, D.C.: ASM Press.
- FRIEDBERG, E. C. (2001). How nucleotide excision repair protects against cancer. *Nature Review Cancer* **1**, 22–33.
- FRIEDBERG, E. C. (2005). Suffering in silence: the tolerance of DNA damage. *Nature Reviews Molecular Cell Biology* **6**, 943–953.
- FUJIWARA, Y., MASUTANI, C., MIZUKOSHI, T., KONDO, J., HANAOKA, F. & IWAI, S. (1999). Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *The Journal of Biological Chemistry* **274**, 20027–20033.
- FUSS, J. O. & TAINER, J. A. (2011). XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. *DNA Repair* **10**, 697–713.
- GEACINTOV, N. E., BROYDE, S., BUTERIN, T., NAEGLI, H., WU, M., YAN, S. & PATEL, D. J. (2002). Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers* **65**, 202–210.
- GILLET, L. C., ALZEER, J. & SCHÄRER, O. D. (2005). Site-specific incorporation of N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) into oligonucleotides using modified ‘ultra-mild’ DNA synthesis. *Nucleic Acids Research* **33**, 1961–1969.
- GILLET, L. C. & SCHÄRER, O. D. (2006). Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chemical Reviews* **106**, 253–276.
- GROISMAN, R., POLANOWSKA, J., KURAOKA, I., SAWADA, J., SAIJO, M., DRAPKIN, R., KISSELEV, A. F., TANAKA, K., NAKATANI, Y. (2003). The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**, 357–367.
- GUNZ, D., HESS, M. T. & NAEGLI, H. (1996). Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *The Journal of Biological Chemistry* **271**, 25089–25098.
- GUZDER, S. N., SUNG, P., BAILLY, V., PRAKASH, L. & PRAKASH, S. (1994). RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. *Nature* **369**, 578–581.
- HANAWALT, P. C. & SPIVAK, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nature Reviews Molecular Cell Biology* **9**, 958–970.
- HARACSKA, L., YU, S. L., JOHNSON, R. E., PRAKASH, L. & PRAKASH, S. (2000). Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. *Nature Genetics* **25**, 458–461.
- HE, Z., HENRICKSEN, L. A., WOLD, M. S. & INGLES, C. J. (1995). RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature* **374**, 566–569.
- HEBRA, F. & KAPOSI, M. (1874). On diseases of the skin including the exanthemata. *The New Sydenham Society* **61**, 252–258.
- HESS, M. T., GUNZ, D., LUNEVA, N., GEACINTOV, N. E. & NAEGLI, H. (1997a). Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes. *Molecular and Cellular Biology* **17**, 7069–7076.
- HESS, M. T., GUNZ, D. & NAEGLI, H. (1996a). A repair competition assay to assess recognition by human nucleotide excision repair. *Nucleic Acids Research* **24**, 824–828.
- HESS, M. T., SCHWITTER, U., PETRETTA, M., GIESE, B. & NAEGLI, H. (1996b). Site-specific DNA substrates for human excision repair: comparison between deoxyribose and base adducts. *Chemical Biology* **3**, 121–128.
- HESS, M. T., SCHWITTER, U., PETRETTA, M., GIESE, B. & NAEGLI, H. (1997b). Bipartite substrate discrimination by human nucleotide excision repair. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 6664–6669.
- HESS, N. J., BUCHKO, G. W., CONRADSON, S. D., ESPINOSA, F. J., NI, S., THRALL, B. D. & KENNEDY, M. A. (1998). Human nucleotide excision repair protein XPA: extended X-ray absorption fine-structure evidence for a metal-binding domain. *Protein Science* **7**, 1970–1975.
- HEY, T., LIPPS, G., SUGASAWA, K., IWAI, S., HANAOKA, F. & KRAUSS, G. (2002). The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay. *Biochemistry* **41**, 6583–6587.
- HILARIO, E., LI, Y., NOBUMORI, Y., LIU, X. & FAN, L. (2013). Structure of the C-terminal half of human XPB helicase and the impact of the disease-causing mutation XP11BE. *Acta Crystallographica D: Biological Crystallography* **69**, 237–246.

- HOEIJMAKERS, J. H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366–374.
- HOHL, M., DUNAND-SAUTHIER, I., STARESINCIC, L., JAQUIER-GUBLER, P., THOREL, F., MODESTI, M., CLARKSON, S. G. & SCHARER, O. D. (2007). Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity. *Nucleic Acids Research* **35**, 3053–3063.
- HOHL, M., THOREL, F., CLARKSON, S. G. & SCHARER, O. D. (2003). Structural determinants for substrate binding and catalysis by the structure-specific endonuclease XPG. *The Journal of Biological Chemistry* **278**, 19500–19508.
- HOOGSTRATEN, D., NIGG, A. L., HEATH, H., MULLENDERS, L. H., VAN DRIEL, R., HOEIJMAKERS, J. H., VERMEULEN, W. & HOUTSMULLER, A. B. (2002). Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair *in vivo*. *Molecular Cell* **10**, 1163–1174.
- HOSFIELD, D. J., MOL, C. D., SHEN, B. & TAINER, J. A. (1998). Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. *Cell* **95**, 135–146.
- HOULE, J. F. & FRIEDBERG, E. C. (1999). The *Drosophila* ortholog of the human XPG gene. *Gene* **234**, 353–360.
- HOUTSMULLER, A. B., RADEMAKERS, S., NIGG, A. L., HOOGSTRATEN, D., HOEIJMAKERS, J. H. & VERMEULEN, W. (1999). Action of DNA repair endonuclease ERCC1/XPF in living cells. *Science* **284**, 958–961.
- HUANG, J. C., HSU, D. S., KAZANTSEV, A. & SANCAR, A. (1994). Substrate spectrum of human excinuclease: repair of abasic sites, methylated bases, mismatches, and bulky adducts. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 12213–12217.
- HWANG, B. J., FORD, J. M., HANAWALT, P. C. & CHU, G. (1999). Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 424–428.
- HWANG, B. J., TOERING, S., FRANCKE, U. & CHU, G. (1998a). p48 Activates a UV-damaged-DNA binding factor and is defective in xeroderma pigmentosum group E cells that lack binding activity. *Molecular and Cellular Biology* **18**, 4391–4399.
- HWANG, K. Y., BAEK, K., KIM, H. Y. & CHO, Y. (1998b). The crystal structure of flap endonuclease-1 from *Methanococcus jannaschii*. *Nature Structural & Molecular Biology* **5**, 707–713.
- IKEGAMI, T., KURAOKA, I., SAIJO, M., KODO, N., KYOGOKU, Y., MORIKAWA, K., TANAKA, K. & SHIRAKAWA, M. (1998). Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. *Nature Structural & Molecular Biology* **5**, 701–706.
- ISAACS, R. J. & SPIELMANN, H. P. (2004). A model for initial DNA lesion recognition by NER and MMR based on local conformational flexibility. *DNA Repair* **3**, 455–464.
- IYER, N., REAGAN, M. S., WU, K. J., CANAGARAJAH, B. & FRIEDBERG, E. C. (1996). Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochemistry* **35**, 2157–2167.
- JOHNSON, R. E., WASHINGTON, M. T., HARACSKA, L., PRAKASH, S. & PRAKASH, L. (2000). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* **406**, 1015–1019.
- JONES, C. J. & WOOD, R. D. (1993). Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochemistry* **32**, 12096–12104.
- KAMIUCHI, S., SAIJO, M., CITTERIO, E., DE JAGER, M., HOEIJMAKERS, J. H. & TANAKA, K. (2002). Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 201–206.
- KATSUMI, S., KOBAYASHI, N., IMOTO, K., NAKAGAWA, A., YAMASHINA, Y., MURAMATSU, T., SHIRAI, T., MIYAGAWA, S., SUGIURA, S., HANAOKA, F., MATSUNAGA, T., NIKAIDO, O. & MORI, T. (2001). *In situ* visualization of ultraviolet-light-induced DNA damage repair in locally irradiated human fibroblasts. *Journal of Investigative Dermatology* **117**, 1156–1161.
- KEENEY, S., CHANG, G. J. & LINN, S. (1993). Characterization of a human DNA damage binding protein implicated in xeroderma pigmentosum E. *The Journal of Biological Chemistry* **268**, 21293–21300.
- KIM, T. K., EBRIGHT, R. H. & REINBERG, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**, 1418–1422.
- KOBAYASHI, T., TAKEUCHI, S., SAIJO, M., NAKATSU, Y., MORIOKA, H., OTSUKA, E., WAKASUGI, M., NIKAIDO, O. & TANAKA, K. (1998). Mutational analysis of a function of xeroderma pigmentosum group A (XPA) protein in strand-specific DNA repair. *Nucleic Acids Research* **26**, 4662–4668.
- KOCH, S. C., KUPER, J., GASTEIGER, K. L., SIMON, N., STRASSER, R., EISEN, D., GEIGER, S., SCHNEIDER, S., KISKER, C. & CARELL, T. (2015). Structural insights into the recognition of cisplatin and AAF-dG lesion by Rad14 (XPA). *Proceedings of the National Academy of Sciences of the United States of America* **112**, 8272–8277.
- KONDRATICK, C. M., WASHINGTON, M. T., PRAKASH, S. & PRAKASH, L. (2001). Acidic residues critical for the activity and biological function of yeast DNA polymerase eta. *Molecular and Cellular Biology* **21**, 2018–2025.
- KUPER, J., BRAUN, C., ELIAS, A., MICHELS, G., SAUER, F., SCHMITT, D. R., POTERSZMAN, A., EGLY, J. M. & KISKER, C. (2014). In TFIIH, XPD helicase is exclusively devoted to DNA repair. *PLoS Biology* **12**, e1001954.
- KUPER, J., WOLSKI, S. C., MICHELS, G. & KISKER, C. (2012). Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *The EMBO Journal* **31**, 494–502.
- KURAOKA, I., MORITA, E. H., SAIJO, M., MATSUDA, T., MORIKAWA, K., SHIRAKAWA, M. & TANAKA, K. (1996). Identification of a damaged-DNA binding domain of the XPA protein. *Mutation Research* **362**, 87–95.
- KUSUMOTO, R., MASUTANI, C., SUGASAWA, K., IWAI, S., ARAKI, M., UCHIDA, A., MIZUKOSHI, T. & HANAOKA, F. (2001). Diversity of the damage recognition step in the global genomic nucleotide excision repair *in vitro*. *Mutation Research* **485**, 219–227.
- LEHMANN, A. R. (2001). The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes & Development* **15**, 15–23.

- LEHMANN, A. R. (2003). DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie* **85**, 1101–1111.
- LEHMANN, A. R. (2008). XPD structure reveals its secrets. *DNA Repair* **7**, 1912–1915.
- LEHMANN, A. R., KIRK-BELL, S., ARLETT, C. F., PATERSON, M. C., LOHMAN, P. H., DE WEERD-KASTELEIN, E. A. & BOOTSMA, D. (1975). Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 219–223.
- LI, L., ELLEDGE, S. J., PETERSON, C. A., BALES, E. S. & LEGERSKI, R. J. (1994). Specific association between the human DNA repair proteins XPA and ERCC1. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5012–5016.
- LI, C., GOLEBIOWSKI, F. M., ONISHI, Y., SAMARA, N. L., SUGASAWA, K. & YANG, W. (2015). Tripartite DNA lesion recognition and verification by XPC, TFIIH, and XPA in nucleotide excision repair. *Molecular Cell* **59**, 1025–1034.
- LI, L., LU, X., PETERSON, C. A. & LEGERSKI, R. J. (1995a). An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Molecular and Cellular Biology* **15**, 5396–5402.
- LI, L., PETERSON, C. A., LU, X. & LEGERSKI, R. J. (1995b). Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Molecular and Cellular Biology* **15**, 1993–1998.
- LIEBER, M. R. (1997). The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **19**, 233–240.
- LIU, H., RUDOLF, J., JOHNSON, K. A., McMAHON, S. A., OKE, M., CARTER, L., MCROBBIE, A. M., BROWN, S. E., NAISMITH, J. H. & WHITE, M. F. (2008). Structure of the DNA repair helicase XPD. *Cell* **133**, 801–812.
- LIU, Y., YANG, Z., UTZAT, C., WANG, G., BASU, A. K. & ZOU, Y. (2005). Cooperative interaction of human XPA stabilizes and enhances specific binding of XPA to DNA damage. *Biochemistry* **44**, 7361–7368.
- LIVNEH, Z. (2001). DNA damage control by novel DNA polymerases: translesion replication and mutagenesis. *The Journal of Biological Chemistry* **276**, 25639–25642.
- LUIJSTERBURG, M. S., GOEDHART, J., MOSER, J., KOOL, H., GEVERTS, B., HOUTSMULLER, A. B., MULLENDERS, L. H., VERMEULEN, W. & VAN DRIEL, R. (2007). Dynamic *in vivo* interaction of DDB2 E3 ubiquitin ligase with UV-damaged DNA is independent of damage-recognition protein XPC. *Journal of Cell Science* **120**, 2706–2716.
- LUIJSTERBURG, M. S., VON BORNSTAEDT, G., GOURDIN, A. M., POLITI, A. Z., MONE, M. J., WARMERDAM, D. O., GOEDHART, J., VERMEULEN, W., VAN DRIEL, R. & HOFER, T. (2010). Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. *The Journal of Cell Biology* **189**, 445–463.
- LUKIN, M. & DE LOS SANTOS, C. (2006). NMR structures of damaged DNA. *Chemical Reviews* **106**, 607–686.
- MAILLARD, O., SOLYOM, S. & NAEGELI, H. (2007). An aromatic sensor with aversion to damaged strands confers versatility to DNA repair. *PLoS Biology* **5**(4), e79.
- MALTSEVA, E. A., RECHKUNOVA, N. I., GILLET, L. C., PETRUSEVA, I. O., SCHARER, O. D. & LAVRIK, O. I. (2007). Crosslinking of the NER damage recognition proteins XPC-HR23B, XPA and RPA to photoreactive probes that mimic DNA damages. *Biochimica et Biophysica Acta* **1770**, 781–789.
- MASUTANI, C., ARAKI, M., YAMADA, A., KUSUMOTO, R., NOGIMORI, T., MAEKAWA, T., IWAI, S. & HANAOKA, F. (1999a). Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *The EMBO Journal* **18**, 3491–3501.
- MASUTANI, C., KUSUMOTO, R., IWAI, S. & HANAOKA, F. (2000). Mechanisms of accurate translesion synthesis by human DNA polymerase eta. *The EMBO Journal* **19**, 3100–3109.
- MASUTANI, C., KUSUMOTO, R., YAMADA, A., DOHMAE, N., YOKOI, M., YUASA, M., ARAKI, M., IWAI, S., TAKIO, K. & HANAOKA, F. (1999b). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature* **399**, 700–704.
- MATHIEU, N., KACZMAREK, N. & NAEGELI, H. (2010). Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 17545–17550.
- MATHIEU, N., KACZMAREK, N., RUTHEMANN, P., LUCH, A. & NAEGELI, H. (2013). DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH. *Current Biology* **23**, 204–212.
- MATSUDA, T., BEBENEK, K., MASUTANI, C., HANAOKA, F. & KUNKEL, T. A. (2000). Low fidelity DNA synthesis by human DNA polymerase-eta. *Nature* **404**, 1011–1013.
- MELLON, I., SPIVAK, G. & HANAWALT, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**, 241–249.
- MIETUS, M., NOWAK, E., JACIUK, M., KUSTOSZ, P., STUDNICKA, J. & NOWOTNY, M. (2014). Crystal structure of the catalytic core of Rad2: insights into the mechanism of substrate binding. *Nucleic Acids Research* **42**, 10762–10775.
- MIN, J. H. & PAVLETICH, N. P. (2007). Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* **449**, 570–575.
- MISSURA, M., BUTERIN, T., HINDGES, R., HUBSCHER, U., KASPARKOVA, J., BRABEC, V. & NAEGELI, H. (2001). Double-check probing of DNA bending and unwinding by XPA-RPA: an architectural function in DNA repair. *The EMBO Journal* **20**, 3554–3564.
- MITCHELL, D. L. & NAIRN, R. S. (1989). The biology of the (6–4) photoproduct. *Photochemistry and Photobiology* **49**, 805–819.
- MIYAMOTO, I., MIURA, N., NIWA, H., MIYAZAKI, J. & TANAKA, K. (1992). Mutational analysis of the structure and function of the xeroderma pigmentosum group A complementing protein. Identification of essential domains for nuclear localization and DNA excision repair. *The Journal of Biological Chemistry* **267**, 12182–12187.
- MOGGS, J. G., SZYMOWSKI, D. E., YAMADA, M., KARRAN, P. & WOOD, R. D. (1997). Differential human nucleotide excision repair of paired and mispaired cisplatin-DNA adducts. *Nucleic Acids Research* **25**, 480–491.

- MOGGS, J. G., YAREMA, K. J., ESSIGMANN, J. M. & WOOD, R. D. (1996). Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *The Journal of Biological Chemistry* **271**, 7177–7186.
- MONE, M. J., VOLKER, M., NIKAIDO, O., MULLENDERS, L. H., VAN ZEELAND, A. A., VERSCHURE, P. J., MANDERS, E. M. & VAN DRIEL, R. (2001). Local UV-induced DNA damage in cell nuclei results in local transcription inhibition. *EMBO Reports* **2**, 1013–1017.
- MORIKAWA, K. & SHIRAKAWA, M. (2000). Three-dimensional structural views of damaged-DNA recognition: T4 endonuclease V, *E. coli* Vsr protein, and human nucleotide excision repair factor XPA. *Mutation Research* **460**, 257–275.
- MORITA, E. H., OHKUBO, T., KURAOKA, I., SHIRAKAWA, M., TANAKA, K. & MORIKAWA, K. (1996). Implications of the zinc-finger motif found in the DNA-binding domain of the human XPA protein. *Genes to Cells* **1**, 437–442.
- MOSER, J., KOOL, H., GIAKZIDIS, J., CALDECOTT, K., MULLENDERS, L. & FOUSTERI, M. I. (2001). Sealing of Chromosomal DNA Nicks during Nucleotide Excision Repair Requires XRCC1 and DNA Ligase III α in a cell-cycle-specific manner. *Molecular Cells* **27**, 311–323.
- MOSER, J., KOOL, H., GIAKZIDIS, I., CALDECOTT, K., MULLENDERS, L. H. & FOUSTERI, M. I. (2007). Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Molecular Cell* **27**(2), 311–323.
- MU, D., BERTRAND-BURGGRAF, E., HUANG, J. C., FUCHS, R. P., SANCAR, A. & FUCHS, B. P. (1994). Human and *E. coli* excinucleases are affected differently by the sequence context of acetylaminofluorene-guanine adduct. *Nucleic Acids Research* **22**, 4869–4871.
- MU, D., PARK, C. H., MATSUNAGA, T., HSU, D. S., REARDON, J. T. & SANCAR, A. (1995). H reconstitution of human DNA repair excision nuclease in a highly defined system. *Journal of Biological Chemistry* **270**, 2415–2418.
- NAEGELI, H., BARDWELL, L. & FRIEDBERG, E. C. (1992). The DNA helicase and adenosine triphosphatase activities of yeast Rad3 protein are inhibited by DNA damage. A potential mechanism for damage-specific recognition. *The Journal of Biological Chemistry* **267**, 392–398.
- NAEGELI, H., MODRICH, P. & FRIEDBERG, E. C. (1993). The DNA helicase activities of Rad3 protein of *Saccharomyces cerevisiae* and helicase II of *Escherichia coli* are differentially inhibited by covalent and noncovalent DNA modifications. *The Journal of Biological Chemistry* **268**, 10386–10392.
- NAEGELI, H. & SUGASAWA, K. (2011). The xeroderma pigmentosum pathway: decision tree analysis of DNA quality. *DNA Repair* **10**, 673–683.
- NEWMAN, M., MURRAY-RUST, J., LALLY, J., RUDOLF, J., FADDEN, A., KNOWLES, P. P., WHITE, M. F. & McDONALD, N. Q. (2005). Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. *The EMBO Journal* **24**, 895–905.
- NICHOLS, A. F., ONG, P. & LINN, S. (1996). Mutations specific to the xeroderma pigmentosum group E Ddb-phenotype. *The Journal of Biological Chemistry* **271**, 24317–24320.
- NISHINO, T., KOMORI, K., ISHINO, Y. & MORIKAWA, K. (2003). X-ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease: similarity between its endonuclease domain and restriction enzymes. *Structure* **11**, 445–457.
- NOCENTINI, S., COIN, F., SAJJO, M., TANAKA, K. & EGLY, J. M. (1997). DNA damage recognition by XPA protein promotes efficient recruitment of transcription factor II H. *The Journal of Biological Chemistry* **272**, 22991–22994.
- O'DONOVAN, A., DAVIES, A. A., MOGGS, J. G., WEST, S. C. & WOOD, R. D. (1994a). XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371**, 432–435.
- O'DONOVAN, A., SCHERLY, D., CLARKSON, S. G. & WOOD, R. D. (1994b). Isolation of active recombinant XPG protein, a human DNA repair endonuclease. *The Journal of Biological Chemistry* **269**, 15965–15968.
- OGI, T., LIMSIRICHAIKUL, S., OVERMEER, R. M., VOLKER, M., TAKENAKA, K., CLONEY, R., NAKAZAWA, Y., NIIMI, A., MIKI, Y., JASPERS, N. G., MULLENDERS, L. H., YAMASHITA, S., FOUSTERI, M. I. & LEHMANN, A. R. (2010). Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Molecular Cell* **37**(5), 714–727.
- OHMORI, H., FRIEDBERG, E. C., FUCHS, R. P., GOODMAN, M. F., HANAOKA, F., HINKLE, D., KUNKEL, T. A., LAWRENCE, C. W., LIVNEH, Z., NOHMI, T., PRAKASH, L., PRAKASH, S., TODO, T., WALKER, G. C., WANG, Z. & WOODGATE, R. (2001). The Y-family of DNA polymerases. *Molecular Cell* **8**, 7–8.
- OKSENYCH, V., BERNARDES DE JESUS, B., ZHOVMER, A., EGLY, J. M. & COIN, F. (2009). Molecular insights into the recruitment of TFIIH to sites of DNA damage. *The EMBO Journal* **28**, 2971–2980.
- ORANS, J., MCSWEENEY, E. A., IYER, R. R., HAST, M. A., HELLINGA, H. W., MODRICH, P. & BESE, L. S. (2011). Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell* **145**, 212–223.
- PARK, C. H., BESSHIO, T., MATSUNAGA, T. & SANCAR, A. (1995a). Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *The Journal of Biological Chemistry* **270**, 22657–22660.
- PARK, C. H., MU, D., REARDON, J. T. & SANCAR, A. (1995b). The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIIE transcription factor. *The Journal of Biological Chemistry* **270**, 4896–4902.
- PARK, C. H. & SANCAR, A. (1994). Formation of a ternary complex by human XPA, ERCC1, and ERCC4 (XPF) excision repair proteins. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5017–5021.
- PUGH, R. A., WU, C. G. & SPIES, M. (2012). Regulation of translocation polarity by helicase domain 1 in SF2B helicases. *The EMBO Journal* **31**, 503–514.
- RADEMAKERS, S., VOLKER, M., HOOGSTRATEN, D., NIGG, A. L., MONE, M. J., VAN ZEELAND, A. A., HOEIJMERS, J. H., HOUTSMULLER, A. B. & VERMEULEN, W. (2003). Xeroderma pigmentosum group A protein loads as a separate factor onto DNA lesions. *Molecular and Cellular Biology* **23**, 5755–5767.
- RAJSKI, S. R., JACKSON, B. A. & BARTON, J. K. (2000). DNA repair: models for damage and mismatch recognition. *Mutation Research* **447**, 49–72.

- RAOUL, S., BARDET, M. & CADET, J. (1995). Gamma irradiation of 2'-deoxyadenosine in oxygen-free aqueous solutions: identification and conformational features of formamidopyrimidine nucleoside derivatives. *Chemical Research in Toxicology* **8**, 924–933.
- READ, C. M., CARY, P. D., CRANE-ROBINSON, C., DRISCOLL, P. C. & NORMAN, D. G. (1993). Solution structure of a DNA-binding domain from HMG1. *Nucleic Acids Research* **21**, 3427–3436.
- REARDON, J. T., NICHOLS, A. F., KEENEY, S., SMITH, C. A., TAYLOR, J. S., LINN, S. & SANCAR, A. (1993). Comparative analysis of binding of human damaged DNA-binding protein (XPE) and *Escherichia coli* damage recognition protein (UvrA) to the major ultraviolet photoproducts: T[c, s]T, T[t,s]T, T[6–4]T, and T[Dewar]T. *The Journal of Biological Chemistry* **268**, 21301–21308.
- REARDON, J. T. & SANCAR, A. (2002). Molecular anatomy of the human excision nuclease assembled at sites of DNA damage. *Molecular and Cellular Biology* **22**, 5938–5945.
- REARDON, J. T. & SANCAR, A. (2003). Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes & Development* **17**, 2539–2551.
- REARDON, J. T. & SANCAR, A. (2005). Nucleotide excision repair. *Progress in Nucleic Acid Research and Molecular Biology* **79**, 183–235.
- REISSNER, T., SCHNEIDER, S., SCHORR, S. & CARELL, T. (2010). Crystal structure of a cisplatin-(1,3-GTG) cross-link within DNA polymerase eta. *Angewandte Chemie International Edition* **49**, 3077–3080.
- RIEDL, T., HANAOKA, F. & EGLY, J. M. (2003). The comings and goings of nucleotide excision repair factors on damaged DNA. *The EMBO Journal* **22**, 5293–5303.
- ROBINS, P., JONES, C. J., BIGGERSTAFF, M., LINDAHL, T. & WOOD, R. D. (1991). Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA. *The EMBO Journal* **10**, 3913–3921.
- RUPP, W. D. & HOWARD-FLANDERS, P. (1968). Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *Journal of Molecular Biology* **31**, 291–304.
- SAIJO, M., KURAOKA, I., MASUTANI, C., HANAOKA, F. & TANAKA, K. (1996). Sequential binding of DNA repair proteins RPA and ERCC1 to XPA *in vitro*. *Nucleic Acids Research* **24**, 4719–4724.
- SANCAR, A. & TANG, M. S. (1993). Nucleotide excision repair. *Photochemistry and Photobiology* **57**, 905–921.
- SCHAFFER, L., ROY, R., HUMBERT, S., MONCOLLIN, V., VERMEULEN, W., HOEIJMAKERS, J. H., CHAMBON, P. & EGLY, J. M. (1993). DNA repair helicase: a component of BTF2 (TFIIF) basic transcription factor. *Science* **260**, 58–63.
- SCHARER, O. D. (2013). Nucleotide excision repair in eukaryotes. *Cold Spring Harbor Perspectives Biology* **5**, a012609.
- SCHORR, S. & CARELL, T. (2010a). Mechanism of acetylaminofluorene-dG induced frameshifting by polymerase eta. *Chembiochem* **11**, 2534–2537.
- SCHORR, S., SCHNEIDER, S., LAMMENS, K., HOPFNER, K. P. & CARELL, T. (2010b). Mechanism of replication blocking and bypass of Y-family polymerase η by bulky acetylaminofluorene DNA adducts. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20720–20725.
- SCHWEIZER, U., HEY, T., LIPPS, G. & KRAUSS, G. (1999). Photocrosslinking locates a binding site for the large subunit of human replication protein A to the damaged strand of cisplatin-modified DNA. *Nucleic Acids Research* **27**, 3183–3189.
- SCRIMA, A., KONICKOVA, R., CZYZEWSKI, B. K., KAWASAKI, Y., JEFFREY, P. D., GROISMAN, R., NAKATANI, Y., IWAI, S., PAVLETICH, N. P. & THOMA, N. H. (2008). Structural basis of UV DNA-damage recognition by the DDB1–DDB2 complex. *Cell* **135**, 1213–1223.
- SHIN, D. S., PELLEGRINI, L., DANIELS, D. S., YELENT, B., CRAIG, L., BATES, D., YU, D. S., SHIVJI, M. K., HITOMI, C., ARVAI, A. S., VOLKMANN, N., TSURUTA, H., BLUNDELL, T. L., VENKITARAMAN, A. R. & TAINER, J. A. (2003). Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *The EMBO Journal* **22**, 4566–4576.
- SHIVJI, M. K., PODUST, V. N., HUBSCHER, U. & WOOD, R. D. (1995). Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry* **34**, 5011–5017.
- SIJBERS, A. M., DE LAAT, W. L., ARIZA, R. R., BIGGERSTAFF, M., WEI, Y. F., MOGGS, J. G., CARTER, K. C., SHELL, B. K., EVANS, E., DE JONG, M. C., RADEMAKERS, S., DE ROOIJ, J., JASPERS, N. G., HOEIJMAKERS, J. H. & WOOD, R. D. (1996). Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**, 811–822.
- SILVERSTEIN, T. D., JAIN, R., JOHNSON, R. E., PRAKASH, L., PRAKASH, S. & AGGARWAL, A. K. (2010a). Structural basis for error-free replication of oxidatively damaged DNA by yeast DNA polymerase eta. *Structure* **18**, 1463–1470.
- SILVERSTEIN, T. D., JOHNSON, R. E., JAIN, R., PRAKASH, L., PRAKASH, S. & AGGARWAL, A. K. (2010b). Structural basis for the suppression of skin cancers by DNA polymerase eta. *Nature* **465**, 1039–1043.
- SINGH, S., FOLKERS, G. E., BONVIN, A. M., BOELENS, R., WECHSELBERGER, R., NIZTAYEV, A. & KAPTEIN, R. (2002). Solution structure and DNA-binding properties of the C-terminal domain of UvrC from *E.coli*. *The EMBO Journal* **21**, 6257–6266.
- STARESINCIC, L., FAGBEMI, A. F., ENZLIN, J. H., GOURLIN, A. M., WIJGERS, N., DUNAND-SAUTHIER, I., GIGLIA-MARI, G., CLARKSON, S. G., VERMEULEN, W. & SCHÄFER, O. D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. *The EMBO Journal* **28**, 1111–1120.
- STORY, R. M. & STEITZ, T. A. (1992). Structure of the recA protein–ADP complex. *Nature* **355**, 374–376.
- SUGASAWA, K., NG, J. M., MASUTANI, C., IWAI, S., VAN DER SPEK, P. J., EKER, A. P., HANAOKA, F., BOOTSMA, D. & HOEIJMAKERS, J. H. (1998). Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Molecular Cell* **2**, 223–232.
- SUGASAWA, K., OKAMOTO, T., SHIMIZU, Y., MASUTANI, C., IWAI, S. & HANAOKA, F. (2001). A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes & Development* **15**, 507–521.
- SUGASAWA, K., OKUDA, Y., SAIJO, M., NISHI, R., MATSUDA, N., CHU, G., MORI, T., IWAI, S., TANAKA, K., TANAKA, K., HANAOKA, F. (2005). UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* **121**, 387–400.
- SUGASAWA, K., SHIMIZU, Y., IWAI, S. & HANAOKA, F. (2002). A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA Repair* **1**, 95–107.

- SUGITANI, N., SHELL, S. M., SOSS, S. E. & CHAZIN, W. J. (2014). Redefining the DNA-binding domain of human XPA. *Journal of the American Chemical Society* **136**, 10830–10833.
- SUNG, P., HIGGINS, D., PRAKASH, L. & PRAKASH, S. (1988). Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. *The EMBO Journal* **7**, 3263–3269.
- SWEDER, K. S. & HANAWALT, P. C. (1992). Preferential repair of cyclobutane pyrimidine dimers in the transcribed strand of a gene in yeast chromosomes and plasmids is dependent on transcription. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 10696–10700.
- TAKAO, M., ABRAMIC, M., MOOS, M. Jr., OTRIN, V. R., WOOTTON, J. C., MCLENIGAN, M., LEVINE, A. S. & PROTIC, M. (1993). A 127 kDa component of a UV-damaged DNA-binding complex, which is defective in some xeroderma pigmentosum group E patients, is homologous to a slime mold protein. *Nucleic Acids Research* **21**, 4111–4118.
- TANG, J. Y., HWANG, B. J., FORD, J. M., HANAWALT, P. C. & CHU, G. (2000). Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis. *Molecular Cell* **5**, 737–744.
- TANTIN, D. (1998). RNA polymerase II elongation complexes containing the Cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IIH components xeroderma pigmentosum B and p62. *The Journal of Biological Chemistry* **273**, 27794–27799.
- TANTIN, D., KANSAL, A. & CAREY, M. (1997). Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Molecular and Cellular Biology* **17**, 6803–6814.
- TAPIAS, A., AURIOL, J., FORGET, D., ENZLIN, J. H., SCHÄRER, O. D., COIN, F., COULOMBE, B. & EGLY, J. M. (2004). Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *The Journal of Biological Chemistry* **279**, 19074–19083.
- THOMA, B. S. & VASQUEZ, K. M. (2003). Critical DNA damage recognition functions of XPC-hHR23B and XPA-RPA in nucleotide excision repair. *Molecular Carcinogenesis* **38**, 1–13.
- THOREL, F., CONSTANTINOU, A., DUNAND-SAUTHIER, I., NOUSPIKEL, T., LALLE, P., RAAMS, A., JASPERS, N. G., VERMEULEN, W., SHIVJI, M. K., WOOD, R. D. & CLARKSON, S. G. (2004). Definition of a short region of XPG necessary for TFIIH interaction and stable recruitment to sites of UV damage. *Molecular and Cellular Biology* **24**, 10670–10680.
- TIRODE, F., BUSSO, D., COIN, F. & EGLY, J. M. (1999). Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Molecular Cell* **3**, 87–95.
- TORNALETTI, S. & HANAWALT, P. C. (1999). Effect of DNA lesions on transcription elongation. *Biochimie* **81**, 139–146.
- TRINCAO, J., JOHNSON, R. E., ESCALANTE, C. R., PRAKASH, S., PRAKASH, L. & AGGARWAL, A. K. (2001). Structure of the catalytic core of *S. cerevisiae* DNA polymerase eta: implications for translesion DNA synthesis. *Molecular Cell* **8**, 417–426.
- TRIPSIANES, K., FOLKERS, G., AB, E., DAS, D., ODIJK, H., JASPERS, N. G., HOEIJMAKERS, J. H., KAPTEIN, R. & BOELENS, R. (2005). The structure of the human ERCC1/XPF interaction domains reveals a complementary role for the two proteins in nucleotide excision repair. *Structure* **13**, 1849–1858.
- TRIPSIANES, K., FOLKERS, G. E., ZHENG, C., DAS, D., GRINSTEAD, J. S., KAPTEIN, R. & BOELENS, R. (2007). Analysis of the XPA and ssDNA-binding surfaces on the central domain of human ERCC1 reveals evidence for subfunctionalization. *Nucleic Acids Research* **35**, 5789–5798.
- TSODIKOV, O. V., ENZLIN, J. H., SCHÄRER, O. D. & ELLENBERGER, T. (2005). Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF–ERCC1. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11236–11241.
- TSODIKOV, O. V., IVANOV, D., ORELLI, B., STARESINCIC, L., SHOSHANI, I., OBERMAN, R., SCHÄRER, O. D., WAGNER, G. & ELLENBERGER, T. (2007). Structural basis for the recruitment of ERCC1–XPF to nucleotide excision repair complexes by XPA. *The EMBO Journal* **26**, 4768–4776.
- TSUTAKAWA, S. E., CLASSEN, S., CHAPADOS, B. R., ARVAI, A. S., FINGER, L. D., GUENTHER, G., TOMLINSON, C. G., THOMPSON, P., SARKER, A. H., SHEN, B., COOPER, P. K., GRASBY, J. A. & TAINER, J. A. (2011). Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**, 198–211.
- TUTEJA, N. & TUTEJA, R. (1996). DNA helicases: the long unwinding road. *Nature Genetics* **13**, 11–12.
- VAN DER SPEK, P. J., EKER, A., RADEMAKERS, S., VISSER, C., SUGASAWA, K., MASUTANI, C., HANAOKA, F., BOOTSMA, D. & HOEIJMAKERS, J. H. (1996). XPC and human homologs of RAD23: intracellular localization and relationship to other nucleotide excision repair complexes. *Nucleic Acids Research* **24**, 2551–2559.
- VENEMA, J., VAN HOFFEN, A., KARCAGI, V., NATARAJAN, A. T., VAN ZEELAND, A. A. & MULLENDERS, L. H. (1991). Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. *Molecular and Cellular Biology* **11**, 4128–4134.
- VOLKER, M., MONE, M. J., KARMAKAR, P., VAN HOFFEN, A., SCHUL, W., VERMEULEN, W., HOEIJMAKERS, J. H., VAN DRIEL, R., VAN ZEELAND, A. A. & MULLENDERS, L. H. (2001). Sequential assembly of the nucleotide excision repair factors *in vivo*. *Molecular Cell* **8**, 213–224.
- WAKASUGI, M., KAWASHIMA, A., MORIOKA, H., LINN, S., SANCAR, A., MORI, T., NIKAIDO, O. & MATSUNAGA, T. (2002). DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair. *The Journal of Biological Chemistry* **277**, 1637–1640.
- WAKASUGI, M. & SANCAR, A. (1998). Assembly, subunit composition, and footprint of human DNA repair excision nuclease. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 6669–6674.
- WAKASUGI, M. & SANCAR, A. (1999). Order of assembly of human DNA repair excision nuclease. *The Journal of Biological Chemistry* **274**, 18759–18768.
- WANG, M., MAHRENHOLZ, A. & LEE, S. H. (2000). RPA stabilizes the XPA-damaged DNA complex through protein–protein interaction. *Biochemistry* **39**, 6433–6439.

- WANG, H., ZHAI, L., XU, J., JOO, H. Y., JACKSON, S., ERDJUMENT-BROMAGE, H., TEMPST, P., XIONG, Y. & ZHANG, Y. (2006). Histone H3 and H4 ubiquitylation by the CUL4-DDB-ROC1 ubiquitin ligase facilitates cellular response to DNA damage. *Mol. Cell* **22**, 383–394.
- WEIR, H. M., KRAULIS, P. J., HILL, C. S., RAINES, A. R., LAUE, E. D. & THOMAS, J. O. (1993). Structure of the HMG box motif in the B-domain of HMG1. *The EMBO Journal* **12**, 1311–1319.
- WHITE, M. F. (2009). Structure, function and evolution of the XPD family of iron-sulfur-containing 5'→3' DNA helicases. *Biochemical Society Transactions* **37**, 547–551.
- WINKLER, G. S., ARAUJO, S. J., FIEDLER, U., VERMEULEN, W., COIN, F., EGLY, J. M., HOEIJMAKERS, J. H., WOOD, R. D., TIMMERS, H. T. & WEEDA, G. (2000). TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *The Journal of Biological Chemistry* **275**, 4258–4266.
- WOLSKI, S. C., KUPER, J., HANZELMANN, P., TRUGLIO, J. J., CROTEAU, D. L., VAN HOUTEN, B. & KISKER, C. (2008). Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS Biology* **6**, e149.
- WOOD, R. D. (1997). Nucleotide excision repair in mammalian cells. *The Journal of Biological Chemistry* **272**, 23465–23468.
- YAN, H., YUAN, W., VELCULESCU, V. E., VOGELSTEIN, B. & KINZLER, K. W. (2002). Allelic variation in human gene expression. *Science* **297**, 1143.
- YAN, S., WU, M., BUTERIN, T., NAEGELE, H., GEACINTOV, N. E. & BROYDE, S. (2003). Role of base sequence context in conformational equilibria and nucleotide excision repair of benzo[a]pyrene diol epoxide–adenine adducts†. *Biochemistry* **42**, 2339–2354.
- YANG, Z., ROGINSKAYA, M., COLIS, L. C., BASU, A. K., SHELL, S. M., LIU, Y., MUSICH, P. R., HARRIS, C. M., HARRIS, T. M. & ZOU, Y. (2006). Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3'- and/or 5'-ssDNA branches. *Biochemistry* **45**, 15921–15930.
- YANG, Z. G., LIU, Y., MAO, L. Y., ZHANG, J. T. & ZOU, Y. (2002). Dimerization of human XPA and formation of XPA2-RPA protein complex. *Biochemistry* **41**, 13012–13020.
- YASUI, M., DONG, H., BONALA, R. R., SUZUKI, N., OHMORI, H., HANAOKA, F., JOHNSON, F., GROLLMAN, A. P. & SHIBUTANI, S. (2004). Mutagenic properties of 3-(deoxyguanosin-N2-yl)-2-acetylaminofluorene, a persistent acetylaminofluorene-derived DNA adduct in mammalian cells. *Biochemistry* **43**, 15005–15013.
- YEH, J. I., LEVINE, A. S., DU, S., CHINTE, U., GHODKE, H., WANG, H., SHI, H., HSIEH, C. L., CONWAY, J. F., VAN HOUTEN, B. & RAPIC-OTRIN, V. (2012). Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2737–E2746.