

Substrate specificity and regulation of the human SMG1-8-9 kinase complex

Dissertation von Lukas Michael Langer

München 2022

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

Substrate specificity and regulation of the human SMG1-8-9 kinase complex

Lukas Michael Langer

aus

München, Deutschland

2022

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Prof. Dr. Elena Conti betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, August 3, 2022

Lukas Langer

Dissertation eingereicht am: **08.08.2022**
1. Gutachterin/Gutachter: Prof. Dr. Elena Conti
2. Gutachterin/Gutachter: Prof. Dr. Karl-Peter Hopfner
Mündliche Prüfung am: **28.09.2022**

Contents

Summary	ix
1 Preface	1
2 Introduction	3
2.1 Metazoan co-translational mRNA quality control	4
2.1.1 Non-stop and no-go decay	5
2.2 Nonsense-mediated mRNA decay in quality control and gene expression regulation	6
2.2.1 NMD-specific protein factors	7
2.2.2 EJC-dependent NMD	11
2.2.3 EJC-independent NMD	15
2.2.4 Other NMD-related mRNA surveillance pathways	16
2.2.5 NMD in human disease	17
2.3 The PIKK family	18
2.3.1 SMG1-8-9, a PIKK complex at the heart of metazoan NMD	20
2.3.2 Structural insights into the SMG1-8-9 complex	20
2.4 Objective of this thesis	25
3 Materials and Methods	27
3.1 Materials	27
3.1.1 Antibodies used in this study	27
3.2 Methods	28
3.2.1 Expression and purification of eukaryotic release factors 1 and 3a	28
3.2.2 <i>In vitro</i> transcription	29
3.2.3 <i>In vitro</i> translation	29
3.2.4 Purification of stalled ribosome - nascent chain complexes	30
3.2.5 Western blotting	31
4 Results	33
4.1 Structure of substrate-bound SMG1-8-9 kinase complex reveals the molecular basis for phosphorylation specificity	33

4.2	Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8	62
4.3	Reconstitution of stalled ribosome translation termination complexes	102
5	Discussion	107
5.1	SMG1-8-9 in NMD	107
5.2	SMG1-8-9 and other PIKKs	112
6	Outlook	117
	Acknowledgements	134

List of Figures

2.1	Transcripts targeted by co-translational mRNA quality control.	4
2.2	NMD-targeted transcripts and NMD mechanisms in metazoans.	7
2.3	Domain organization of important mammalian NMD factors.	10
2.4	Major steps of eukaryotic translation termination.	12
2.5	EJC-dependent NMD.	14
2.6	Domain organization and specificity of the PIKK family.	19
2.7	Structure of human SMG1 and the SMG1-8-9 complex.	24
4.1	Stalling and purification of defined ribosome - nascent chain translation termination complexes using release factor mutants.	104
4.2	Size-selective double pull down to enrich ribosome - nascent chain translation termination complexes.	106
5.1	Hypothetical model for SMG1 activity regulation.	109
5.2	Structures of human SMG1, SMG1-9 and SMG1-8-9.	111
5.3	Superposition of PIKK active site-bound inhibitors with SMG1i.	113
5.4	Superposition of UPF1-LSQ peptide and AMPPNP with human ATM and DNA-PK active sites.	115

List of Tables

3.1	Primary antibodies used in this study.	27
3.2	Secondary antibodies used in this study.	27
3.3	Transcription mix used in this study.	29
3.4	10 x transcription buffer used in this study.	29
3.5	Buffers used in ribosome - nascent chain complex purifications.	30
3.6	PVDF transfer buffer used in this study.	31

Summary

Quality control pathways are central to maintaining cellular homeostasis. On the mRNA level, nonsense-mediated mRNA decay (NMD) co-translationally detects and degrades transcripts with premature termination codons to prevent synthesis of truncated proteins, and regulates levels of physiological mRNAs. The phosphorylation of the RNA helicase UPF1 by the SMG1 kinase is considered a key step in this surveillance pathway: upon recognition of a premature translation termination event, it commits the targeted mRNA to degradation. SMG1 is known to recognize specific motifs within UPF1 for phosphorylation and to be regulated by both intra- as well as intermolecular interactions, but the details of both substrate recognition and activity regulation remain elusive. How the NMD factors SMG1 and UPF1 interface with the translation machinery is not understood.

To address these questions, recombinant human SMG1 in complex with its binding partners SMG8 and SMG9 was purified from modified HEK293T suspension cells. In the first part of this study, the ~600 kDa kinase complex was reconstituted with a non-hydrolyzable ATP-analogue and a UPF1-derived peptide and subjected to single-particle cryo-electron microscopy (cryo-EM) analysis. The resulting 2.9 Å reconstruction allowed to model a substrate peptide and a nucleotide analogue in the kinase active site. To further dissect the contribution of individual residues, modified SMG1 substrate peptides were assayed using mass spectrometry-based phosphorylation experiments. These results elucidated the molecular basis of specific phosphorylation site selection by SMG1 and provided insights into similarities and differences to related kinases, such as ATM and mTOR.

Next, SMG1-8-9 was combined with a small molecule inhibitor specific for SMG1. The obtained structure provided insights into determinants of inhibitor specificity. Furthermore, the reconstruction revealed density attributed to a regulatory domain - the SMG1 insertion - blocking the substrate binding path within the kinase active site. Importantly, the reconstruction of a SMG1-9 complex calculated from the same data set did not show ordered SMG1 insertion density. By integration of the obtained structural data with crosslinking mass spectrometry, AI-based protein structure prediction and pull-down assays, this work

found that SMG1 autoinhibition by its insertion domain is stabilized by the globular SMG8 C-terminal domain. These results shed light on the intricate regulatory network finetuning activity of this kinase complex.

Additionally, leveraging *in vitro* translation methods, the reconstitution of four distinct eukaryotic translation termination complexes was established.

Taken together, the work presented here describes the first structural model of a substrate-bound SMG1 kinase active site, revealing the molecular details of a central step of the NMD surveillance pathway. It provides insights into the regulation of the SMG1-8-9 kinase complex within NMD and offers a basis for rational design of small molecule compounds targeting SMG1. Finally, it paves the way to understanding NMD initiation by investigating the interaction between NMD factors and translation termination complexes.

Chapter 1

Preface

Recognition and destruction of error-containing mRNAs in metazoans involve various surveillance pathways harnessing diverse enzymatic activities. Nonsense-mediated mRNA decay (NMD) contributes to homeostasis by co-translationally recognizing and degrading mRNAs with premature termination codons, and regulating levels of many other mRNAs. Dysfunction of this mechanism has been linked to several human diseases. A key catalytic activity within the NMD quality control pathway is the phosphorylation of the RNA helicase UPF1 by the SMG1-8-9 kinase complex. The work presented here describes biochemical and structural insights into the specific phosphorylation of target sites within UPF1 by SMG1, regulation of its catalytic activity and a first step towards exploiting this kinase for therapeutic intervention by using small molecules. Finally, the reconstitution of four different translation termination complexes is described, facilitating future studies aiming at understanding the interplay between NMD factors and the translation machinery.

This dissertation is written in a cumulative way, detailing scientific results obtained under the supervision of Prof. Elena Conti, PhD, in the Department of Structural Cell Biology at the Max Planck Institute of Biochemistry, Martinsried, since January 2018. Chapter 2 of this thesis provides an in-depth introduction to metazoan NMD and the SMG1-8-9 kinase complex and defines the objective of this thesis. Chapter 3 details materials and methods important for reproducing unpublished results discussed in this work. Chapter 4 presents the two first-authored, original scientific manuscripts that form the basis of this thesis. In addition, it reports unpublished work describing the reconstitution of different translation termination complexes. Chapter 5 contains an extended discussion, building on the results presented before and supplementing the discussions included in the published manuscripts. Finally, Chapter 6 offers an outlook into the future perspectives of the field.

Chapter 2

Introduction

Gene expression is at the heart of cellular health: the tight regulation of transcription and translation and their continuous surveillance for aberrant events are crucial to life across pro- and eukaryotes. In eukaryotes, several specialized mRNA quality control pathways have evolved, allowing them to survey the integrity of a transcript undergoing translation (Isken & Maquat, 2007; Shoemaker & Green, 2012; Karousis & Mühlemann, 2019). These mechanisms employ *cis*- and *trans*-acting factors that co-translationally integrate diverse information provided by the messenger ribonucleoprotein (mRNP) environment, such as ribosomal collisions, 3' untranslated regions (UTRs) and translation termination events. Upon detection of aberrant events, sophisticated cascades of interdependent enzymatic activities cause the degradation of the erroneous transcript, thus maintaining cellular health. Eukaryotic co-translational mRNA quality control has been historically classified into three separate pathways, each conserved from yeast to human and named after the mRNA features they detect: Non-stop, no-go and nonsense-mediated mRNA decay (Shoemaker & Green, 2012). For all of these surveillance pathways, mechanisms which specifically destabilize the resulting faulty nascent polypeptide chain have been described, functionally linking mRNA and protein quality control (Brandman & Hegde, 2016; Chu *et al.*, 2021).

The following pages will provide a general introduction to co-translational mRNA quality control pathways, with a particular focus on nonsense-mediated mRNA decay. NMD will be described mechanistically, highlighting both its quality control as well as its gene regulatory function. Detailing the protein factors involved in carrying out NMD, special attention will be given to the SMG1-8-9 kinase complex and its catalytic activity towards the RNA helicase UPF1. Recent insights into the architecture of the human SMG1-8-9 kinase complex will be described. The phosphatidylinositol 3-kinase-related kinase (PIKK) SMG1 will also be discussed in the context of closely related kinases, which are central to

other cellular surveillance mechanisms.

2.1 Metazoan co-translational mRNA quality control

The three major co-translational mRNA surveillance pathways described in eukaryotes recognize and respond to specific mRNA defects (Shoemaker & Green, 2012). Non-stop decay (NSD) recognizes mRNAs without an in-frame stop codon before the poly(A) tail (Frischmeyer *et al.*, 2002), while no-go decay (NGD) targets transcripts with features blocking the ribosome before reaching a termination codon, comprising also situations causing empty ribosomal A-sites (Doma & Parker, 2006) (see Figure 2.1). Finally, nonsense-mediated mRNA decay (NMD) was defined as a mechanism detecting transcripts with a premature termination codon, hence coding for C-terminally truncated proteins (Maquat *et al.*, 1981; Peltz *et al.*, 1993). While non-stop and no-go decay are believed to share factors important for initiating steps of the pathway as well as for degradation of the resulting polypeptide, a separate set of factors is responsible for NMD (Brandman & Hegde, 2016; Karousis *et al.*, 2016). Despite these pathways having been discovered decades ago, many important aspects remain unclear and research over recent years has provided substantial new insights into their complex molecular playbook and mechanistic diversity.

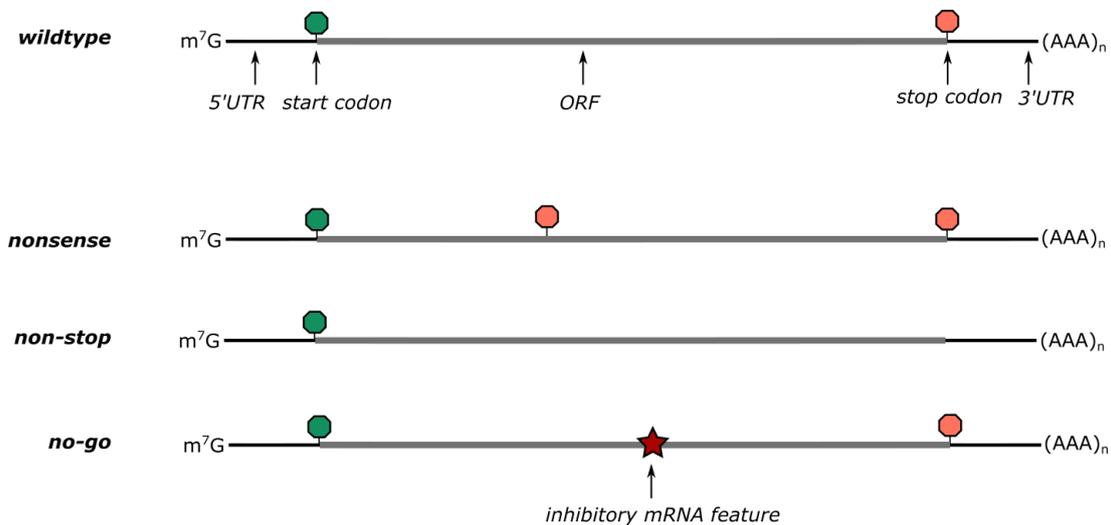


Figure 2.1: Transcripts targeted by co-translational mRNA quality control. Simplified schematic illustrating exemplary features or alterations within mRNA sequences known to trigger co-translational mRNA quality control pathways.

2.1.1 Non-stop and no-go decay

While the inducing features of the mRNA are different, being either a transcript without stop codon before the poly(A) tail for NSD or a physical barrier such as a stably folded stem loop or empty A-sites as in the case of NGD, both pathways are known to be interconnected (Shoemaker & Green, 2012). On an NSD substrate the leading ribosome will stall when translating the poly(A) tail, due to the coincident detection of poly(A) RNA in its decoding center and the poly-lysine chain within its peptide exit tunnel (Chandrasekaran *et al.*, 2019). Stalling of the leading ribosome can result in a collision with the following ribosome. These events result in stable structures termed 'collided disomes' which display a unique interface made up by both ribosomes that is crucial for recognition by some surveillance factors, such as the E3 ligase ZNF598 ubiquitinating ribosomal proteins and the endonuclease NONU-1 (Simms *et al.*, 2017; Juszkiwicz & Hegde, 2017; Juszkiwicz *et al.*, 2018; Glover *et al.*, 2020; Pochopien *et al.*, 2021). Cleavage of the targeted transcript by an endonuclease would ultimately result in the production of NGD substrates with ribosomes translating to the end of the message and consequently displaying empty A-sites. Such cases are recognized by the central NGD factors Pelota-Hbs1, which are structurally related to the canonical eukaryotic release factors eRF1-eRF3a (Shoemaker & Green, 2011, 2012). As an important difference to eRF1, Pelota lacks both the GGQ motif necessary to release the nascent chain as well as the NIKS motif important for stop codon recognition (for details, see section 2.2.2). Since initial recognition of an empty A-site by Pelota-Hbs1 leads to the recruitment of a ribosomal subunit splitting factor, they are crucial for recycling of ribosomes stalled on faulty transcripts. Further recycling or degradation of tRNA and nascent chain still bound to the resulting 60S subunit is mediated by other pathways with separate dedicated factors (Yip & Shao, 2021). The mRNA targeted in both pathways can be degraded by the exonuclease XRN1 in the 5' to 3' direction, and the cytoplasmic exosome in the 3' to 5' direction (Anderson & Parker, 1998; Van Hoof *et al.*, 2002). The latter works closely together with the SKI complex, which has recently been shown to use an intrinsic gatekeeping mechanism to direct mRNAs from the ribosome exit tunnel to the cytoplasmic exosome (Kögel *et al.*, 2022). The SKI complex is tethered to the cytoplasmic exosome by HBS1L3, which interacts with the exosome in a way similar to the yeast Ski7 protein (Kowalinski *et al.*, 2016).

2.2 Nonsense-mediated mRNA decay in quality control and gene expression regulation

The detrimental effect of nonsense mutations on mRNA half life in yeast and humans has been initially described over four decades ago (Chang & Kan, 1979; Losson & Lacroute, 1979; Maquat *et al.*, 1981). It soon became clear that a specific cellular pathway is degrading prematurely terminating transcripts in a manner dependent on active translation: nonsense-mediated mRNA decay (Belgrader *et al.*, 1993; Carter *et al.*, 1995; Lykke-Andersen *et al.*, 2000; Karousis & Mühlemann, 2019). Genetic screens using nematodes and budding yeast revealed a core set of protein factors crucial for its functionality: Upf1 to 3 (up-frameshift) in *Saccharomyces cerevisiae* and SMG-2 to SMG-7 (suppressor of morphogenetic effects on genitalia) in *Caenorhabditis elegans* (wherein SMG-2 to SMG-4 are orthologs to Upf1 to 3) (Leeds *et al.*, 1992; Hodgkin *et al.*, 1989). These factors were found to integrate information provided by the translated mRNP to target transcripts with premature termination codons (PTCs), rather than recognizing the PTC directly. In yeast, an important NMD signal was found to be long 3'UTRs (Hogg & Goff, 2010), while in metazoans the signal for degradation by NMD was initially discovered to be an exon-junction complex (EJC). This protein complex is deposited onto the mRNA 20-24 nucleotides upstream of an exon-exon junction by the spliceosome (Le Hir *et al.*, 2000, 2001). EJCs located within open reading frames will usually be removed by translating ribosomes, highlighting the presence of an EJC downstream on a termination codon as a clear indicator of an aberrant transcript (Gehring *et al.*, 2009; Saulière *et al.*, 2012). Importantly, if a PTC was located 50-55 nucleotides upstream of an exon-exon junction, the message in question was efficiently degraded by NMD factors, hence constituting splicing-dependent NMD (Nagy & Maquat, 1998).

More recently, it was discovered that other features than an EJC can trigger NMD factors to degrade transcripts in metazoans, such as upstream open reading frames (uORFs) and long 3'UTRs (Gaba *et al.*, 2005; Bühler *et al.*, 2006; Eberle *et al.*, 2008; Nyikó *et al.*, 2009). Interestingly, it has been established that NMD not only targets aberrant transcripts, but also regulates steady-state levels of physiological transcripts (see Figure 2.2). How some physiological transcripts are selected by the NMD machinery remains poorly understood, as exemplified by the undefined contribution of out-of-frame translation to creating NMD targets (Lykke-Andersen & Jensen, 2015; Boersma *et al.*, 2019). Overall, NMD carries out diverse quality control and regulatory functions, thereby affecting around ~10% of cellular

mRNAs (Kishor *et al.*, 2019; Schweingruber *et al.*, 2013).

Taken together, the understanding of metazoan NMD has evolved from a purely quality control-focused pathway to a multipronged cellular tool to control and regulate expression of diverse transcripts (see Figure 2.2). In the following, the above mentioned facets of NMD and the protein factors involved will be discussed in more detail.

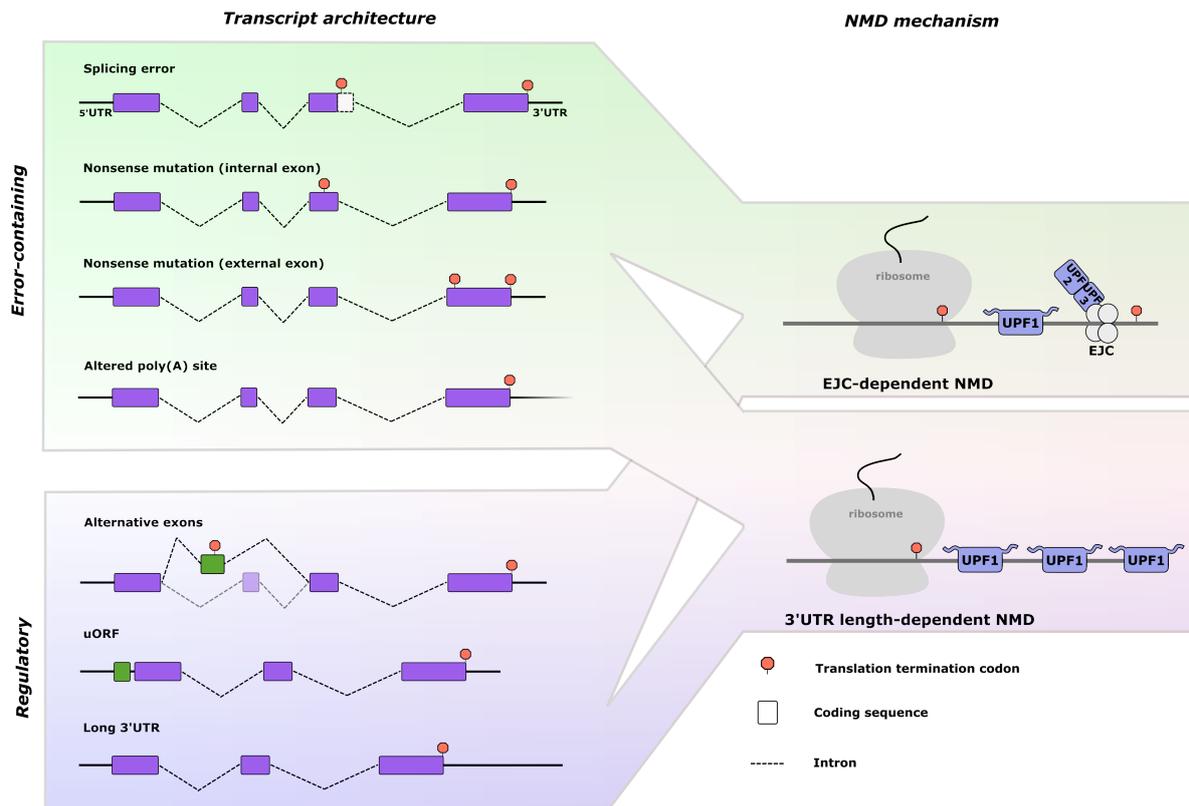


Figure 2.2: NMD-targeted transcripts and NMD mechanisms in metazoans. Different transcripts displaying NMD-inducing features are shown. Depending on their features, these transcript may be subject to the quality control or regulatory branch of NMD. The composition of the mRNP 3'UTRs will determine their degradation via EJC- or 3'UTR length- dependent NMD. The figure was based on Kishor *et al.*, 2019.

2.2.1 NMD-specific protein factors

Early genetic screens in yeast and nematodes identified a set of core NMD factors that remain at the center of all current NMD models: the up-frameshift proteins (Upf1-3) discovered in yeast, concomitantly described in nematodes as SMG-2 to SMG-4 (see Figure 2.3) (Hodgkin *et al.*, 1989; Leeds *et al.*, 1992).

UPF1 is a 120 kDa SF1 ATP-dependent RNA helicase whose ATPase activity is crucial for NMD in yeast and metazoans, while its helicase activity is dispensable (Leeds *et al.*, 1992; Jankowsky *et al.*, 2010; Franks *et al.*, 2010; Chapman *et al.*, 2022). UPF1 is believed to take a central role in NMD target recognition (Kashima *et al.*, 2006; Hogg & Goff, 2010). Its helicase domain is comprised of two RecA domains which form a catalytic cleft that can accommodate ATP, as well as additional regulatory domains (Cheng *et al.*, 2007). This helicase domain is flanked at its N-terminus by a zinc-knuckle domain called cysteine-histidine-rich CH domain. At its C-terminus, the helicase domain is neighbored by a largely unstructured serine/threonine-glutamine-rich (SQ) domain. The SQ motifs present in this domain are subject to phosphorylation by the SMG1 kinase upon recognition of an NMD target. Additionally, some SQ motifs are found in the very N-terminal part of UPF1. Phosphorylation of SQ motifs allows UPF1 to recruit downstream NMD effectors, such as SMG6 and the SMG5-7 heterodimer (see below) (Ohnishi *et al.*, 2003; Okada-Katsuhata *et al.*, 2012; Chakrabarti *et al.*, 2014). Both CH and SQ domains play crucial roles in the regulation of UPF1 activity. First, the SQ domain has been shown to inhibit UPF1 ATPase and helicase activity in its unphosphorylated state by interacting with the helicase domain (Fiorini *et al.*, 2013). The effect of SQ motif phosphorylation on this inhibitory function remains unclear. Second, the CH domain interacts with UPF2 to activate UPF1 ATPase and helicase activity (Kadlec *et al.*, 2006; Chamieh *et al.*, 2008; Chakrabarti *et al.*, 2011).

By binding to a portion of the CH domain that can also engage in an intra-molecular interaction with one of the UPF1 RecA domains, UPF2 binding causes a dramatic structural rearrangement of the CH domain, thereby switching UPF1 from an RNA clamping into an RNA unwinding mode (Chakrabarti *et al.*, 2011). Located at the UPF2 C-terminus, a UPF1-binding domain (UBD) mediates this crucial interaction (Clerici *et al.*, 2009). Furthermore, the 140 kDa large UPF2 is acting as a scaffold, thus physically linking UPF1 and UPF3. Next to the C-terminal UBD, UPF2 is characterized by three middle-portion-of-eIF4G (MIF4G) domains, one of which interacts with UPF3 (Chamieh *et al.*, 2008).

UPF3 is known to be present in two isoforms in metazoans, UPF3A and UPF3B. The latter has been found to be the isoform important for NMD, while recently UPF3A has been described to be able to largely compensate for the absence of UPF3B (Buchwald *et al.*, 2010; Wallmeroth *et al.*, 2022; Yi *et al.*, 2022). UPF3B is a ~50 kDa protein comprising an RRM-like domain that interacts with the described UPF2 MIF4G domain, as well as a C-terminal EJC binding motif (EBM). The latter has been shown to tether UPF3 to

the EJC, thereby providing a bridge to UPF2 and UPF1 (Chamieh *et al.*, 2008; Buchwald *et al.*, 2010; Melero *et al.*, 2012). Furthermore, UPF3B was reported to delay translation termination by directly interacting with eRF3a and the ribosome (Neu-Yilik *et al.*, 2017). It has also been suggested to be able to directly interact with UPF1.

Studies disentangling the complex interplay of the aforementioned NMD core factors have provided important insights into the NMD mechanism and its conservation, but also led to new open questions. For example, binding of a nucleotide as well as of UPF2 has been shown to drastically reduce the RNA binding affinity of UPF1, despite being crucial for UPF1 helicase activation (Chakrabarti *et al.*, 2011; Xue *et al.*, 2022). The SQ domain of UPF1 has been implicated in intramolecular regulation of UPF1, but its role within a UPF1-2-3 complex and the influence of its phosphorylation status therein remains unclear. In addition to the three core factors, several other metazoan NMD factors have been characterized (see Figure 2.3). Early work in nematodes identified SMG-1 and SMG-5 to 7 (Leeds *et al.*, 1992). While no homolog to SMG-1 is known in yeast (consistent with the absence of SQ motifs in yeast Upf1), factors potentially carrying out the roles of SMG-5 to SMG-7 have recently been identified (Dehecq *et al.*, 2018). Interaction studies in human cells have revealed SMG8 and SMG9 as stable interactors of SMG1 (Yamashita *et al.*, 2009). Further studies have also implicated the SF2 RNA helicase DHX34 and NBAS in NMD (Longman *et al.*, 2007).

SMG1 is a 410kDa Ser/Thr protein kinase that phosphorylates UPF1 on its SQ motifs (Denning *et al.*, 2001; Yamashita *et al.*, 2001). These phosphorylation events are considered a key step in committing a transcript to degradation by NMD (see above). Furthermore, SMG1 has been shown to tightly interact and to be regulated by the G-domain proteins SMG8 and SMG9, together forming the ~600kDa SMG1-8-9 kinase complex (Yamashita *et al.*, 2009; Arias-Palomo *et al.*, 2011; Melero *et al.*, 2014; Li *et al.*, 2017). A more detailed introduction to SMG1, SMG8 and SMG9 is provided in subsection 2.3.1.

SMG5 and SMG7 are proteins harboring phospho-serine binding 14-3-3-like domains and form a heterodimer (Fukuhara *et al.*, 2005; Jonas *et al.*, 2013). They interact with phosphorylated UPF1 and can recruit the CCR4-NOT deadenylase complex (Ohnishi *et al.*, 2003; Chakrabarti *et al.*, 2014; Loh *et al.*, 2013). By interaction with protein phosphatase 2A (PP2A) they mediate dephosphorylation of UPF1 (Ohnishi *et al.*, 2003). SMG5 carries a C-terminal PIN-like domain that lacks catalytic residues and therefore nucleolytic activity (Glavan *et al.*, 2006).

SMG6 is a 160kDa endonuclease. As described for SMG5 and SMG7 it harbors a 14-3-3-

like domain, which is part of a central TPR domain known to mediate phospho-dependent and independent interactions with UPF1 (Chakrabarti *et al.*, 2014; Nicholson *et al.*, 2014). The TPR domain is flanked on its N-terminus by a region carrying two EBMs for direct interaction with the EJC (Kashima *et al.*, 2010). Whether and how these EBMs compete with the EBM present in UPF3 (see above) for EJC binding remains to be clarified. C-terminally, the TPR domain is followed by a PIN domain, which endonucleolytically cleaves mRNAs targeted by NMD (Glavan *et al.*, 2006; Huntzinger *et al.*, 2008). While specific interactions between NMD factors have been highlighted above, the following sections will present the mechanistic context of their interplay.

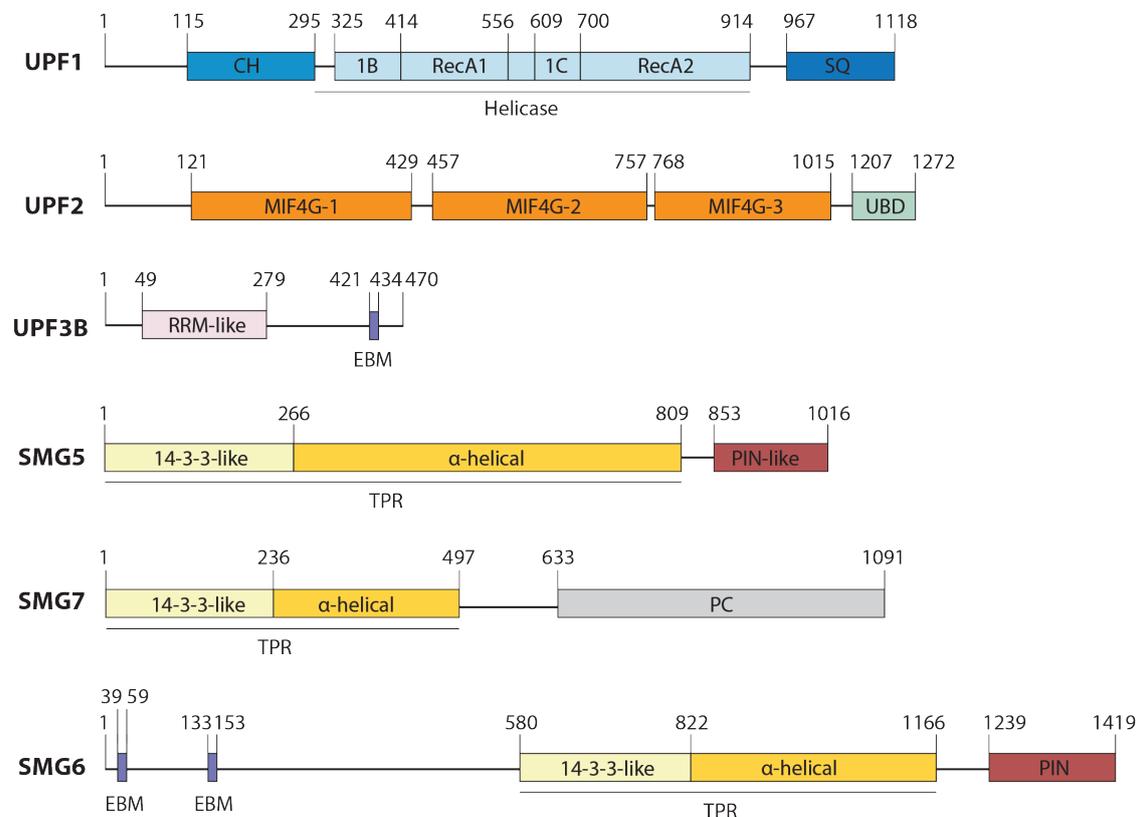


Figure 2.3: Domain organization of important mammalian NMD factors. Information for SMG1, SMG8 and SMG9 can be found in Figure 2.7. All schematics are drawn to scale and based on Karousis *et al.*, 2016. CH: cysteine-histidine rich domain; SQ: serine/theronine-glutamine-rich domain; MIF4G: middle of 4G-like domains; UBD: UPF1-binding domain; RRM: RNA recognition motif; EBM: exon-junction binding motif; PIN: PilT N-terminus; PC: C-terminal proline-rich region; TPR: tetratricopeptide repeat.

2.2.2 EJC-dependent NMD

Building upon initial work in yeast and nematodes, a model for canonical NMD in metazoans has been described as EJC-dependent NMD (see section 2.2). While certain aspects of this model are under debate, and the exon-junction complex (EJC) is now viewed as an NMD-enhancing feature more than a strict requirement, it is commonly used as a framework for further investigation in metazoans (Karousis *et al.*, 2016; Kishor *et al.*, 2019; Kurosaki *et al.*, 2019; Karousis & Mühlemann, 2022). mRNAs targeted by this pathway are characterized by a PTC located at least 50-55 nucleotides upstream of an exon-exon junction, or about 20 nucleotides upstream of an EJC (Nagy & Maquat, 1998) (see section 2.2). The reason for this distance requirement remains unknown, but might involve sterical requirements between NMD *cis*- and *trans*-acting factors and the terminating ribosome. The cytoplasmic EJC core is a heterotetramer, within which the DEAD-box helicase eIF4AIII is locked in an ATP-bound, inactive conformation onto the RNA backbone by the MAGOH-Y14 heterodimer and MLN51 (Bono *et al.*, 2006; Andersen *et al.*, 2006). The latter additionally complements the interaction with RNA. While EJCs play a crucial role in mRNP biogenesis, for example during packaging in the nucleus and export, they are therefore also central to co-translational mRNA quality control in the cytoplasm (Le Hir & Séraphin, 2008). EJCs are thought to be displaced from the mRNA by ribosomes during the first round of translation. Consequently, prematurely terminating ribosomes can result in EJCs remaining part of the translating mRNP (Karousis *et al.*, 2016; Kishor *et al.*, 2019). The causative premature termination codons may result from nonsense or frameshift mutations in the genome, errors during transcription, splicing, or other events (see Figure 2.2) (Kishor *et al.*, 2019). When ribosomes terminate on such stop codons with a correctly positioned downstream EJC, NMD factors can recognize these translation termination events as premature and trigger a complex mRNA degradation pathway.

Eukaryotic translation termination

Translation of an mRNA by a ribosome can be divided into three major phases: initiation, elongation and termination. NMD was characterized as being dependent on active translation in general, and translation termination on a premature stop codon in particular is thought to play a crucial role in eliciting NMD (Kashima *et al.*, 2006; Naeger *et al.*, 1992; Karousis & Mühlemann, 2019). Translation termination is initiated by the binding of a ternary eRF1-eRF3a-GTP complex to the ribosomal intersubunit space. In particular, the presence of a stop codon in the ribosomal A-site is recognized by the NIKS motif, an

Asn-Ile-Lys-Ser tetrapeptide within eRF1 (Frolova *et al.*, 1994, 2002; Brown *et al.*, 2015). eRF3a facilitates fast ribosomal binding by eRF1, likely by stabilizing it in a compatible conformation. GTP hydrolysis by eRF3a leads to its dissociation from the ribosomal intersubunit space and allows eRF1 to rearrange and accommodate the GGQ motif (Gly-Gly-Gln) in the peptidyl transferase center. Here, the GGQ motif coordinates water in order to hydrolyze the ester bond between the nascent polypeptide and the P-site tRNA, to release the newly synthesized protein (Frolova *et al.*, 1999; Alkalaeva *et al.*, 2006; Shao *et al.*, 2016; Lawson *et al.*, 2021). eRF1 can then interact with the ATP-dependent ribosome splitting and recycling factor ABCE1, thus directly coupling translation termination and ribosome recycling (see Figure 2.4) (Shoemaker & Green, 2011; Brown *et al.*, 2015). Co-immunoprecipitation experiments have suggested an eRF1-eRF3-ribosome complex as the unit being recognized by NMD factors (Kashima *et al.*, 2006). Whether this is indeed the translation termination complex recognized during NMD initiation or whether other ribosomal complexes play a role awaits further confirmation (Karousis & Mühlemann, 2019).

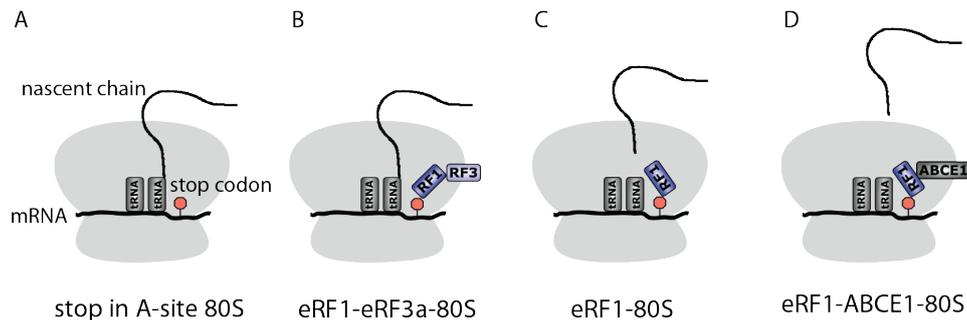


Figure 2.4: Major steps of eukaryotic translation termination. A. A translating 80S ribosome encounters a stop codon in the A-site. B. This stop codon is recognized by eRF1, binding in complex with eRF3-GTP. C. GTP hydrolysis by eRF3 leads to its dissociation from the ribosome and allows structural rearrangement of eRF1. eRF1 now hydrolyzes the P-site tRNA - nascent chain ester bond, releasing the newly synthesized polypeptide. D. eRF1 interacts with the recycling factor ABCE1 which promotes subunit splitting.

Recognition and degradation of the target mRNA

The recognition of translation termination events as premature is thought to depend on general translation termination factors as well as NMD specific proteins (see Figure 2.5). Co-immunoprecipitation experiments identified SMG1, UPF1, eRF1 and eRF3 as components of the so-called SURF complex, which is believed to represent an early step in NMD

target recognition (Kashima *et al.*, 2006). Later experiments suggested the SURF complex also includes the SMG8-SMG9 heterodimer bound to SMG1 for its recruitment and regulation (Yamashita *et al.*, 2009), as well as the RNA helicase DHX34 (Hug & Cáceres, 2014). The NMD mRNP is thought to be remodelled when SURF-bound UPF1 interacts with UPF2-UPF3B in the context of a downstream EJC, leading to phosphorylation of UPF1 by SMG1 as part of a decay-inducing complex (DECID) (Kashima *et al.*, 2006). Phosphorylated UPF1 then recruits SMG6 for endonucleolytic cleavage of the transcript as well as the SMG5-SMG7 heterodimer. The latter results in localization of the message to P-bodies for decapping, recruits the CCR4-NOT deadenylation complex and PP2A, which can dephosphorylate UPF1 (see Figure 2.5). Furthermore PNRC2 can be recruited, providing a link for recruitment of mRNA decapping factors (Kurosaki *et al.*, 2019; Karousis & Mühlemann, 2019, 2022). Hyperphosphorylation of UPF1 is believed to serve as a tool to more efficiently recruit NMD downstream effectors with increasing residence time of UPF1 on a target mRNA (Durand *et al.*, 2016). While the SMG6 and SMG5-SMG7 triggered pathways for degradation were initially believed to represent independent routes to feed NMD-targeted mRNAs to the canonical mRNA decay machinery, recent evidence suggests that they are indeed interdependent, representing a further regulatory step in NMD (Boehm *et al.*, 2021).

The SURF and DECID complex model remains to be further refined and validated, and several aspects of EJC-dependent NMD await more detailed investigation. For example, the presence of an actual eRF1-eRF3a complex on a terminating ribosome is known to be a transient state within the translation termination pathway (Lawson *et al.*, 2021). In contrast to earlier results (Amrani *et al.*, 2004), recent data suggested that NMD ensues without ribosome stalling which is known to be necessary for other co-translational quality control pathways (Karousis *et al.*, 2020). Furthermore, it is not clear how the SURF complex would assemble as a consequence of a downstream EJC, since neither SMG1-8-9 nor UPF1 are known to directly interact with the EJC. At what point and through which mechanism the SMG1 kinase is activated is not understood. How it recognizes specific motifs within UPF1 for the crucial phosphorylation step remains equally unknown. Following the initial recognition of an NMD transcript, it is not clear when and how the EJC-binding motifs in SMG6 and UPF3B compete with binding to the EJC. Additionally, the interplay of the phospho-dependent and -independent interactions between SMG6 and UPF1 within the pathway remain to be disentangled. Thus, while a general framework for canonical EJC-dependent NMD exists, many important steps, including the spatial and

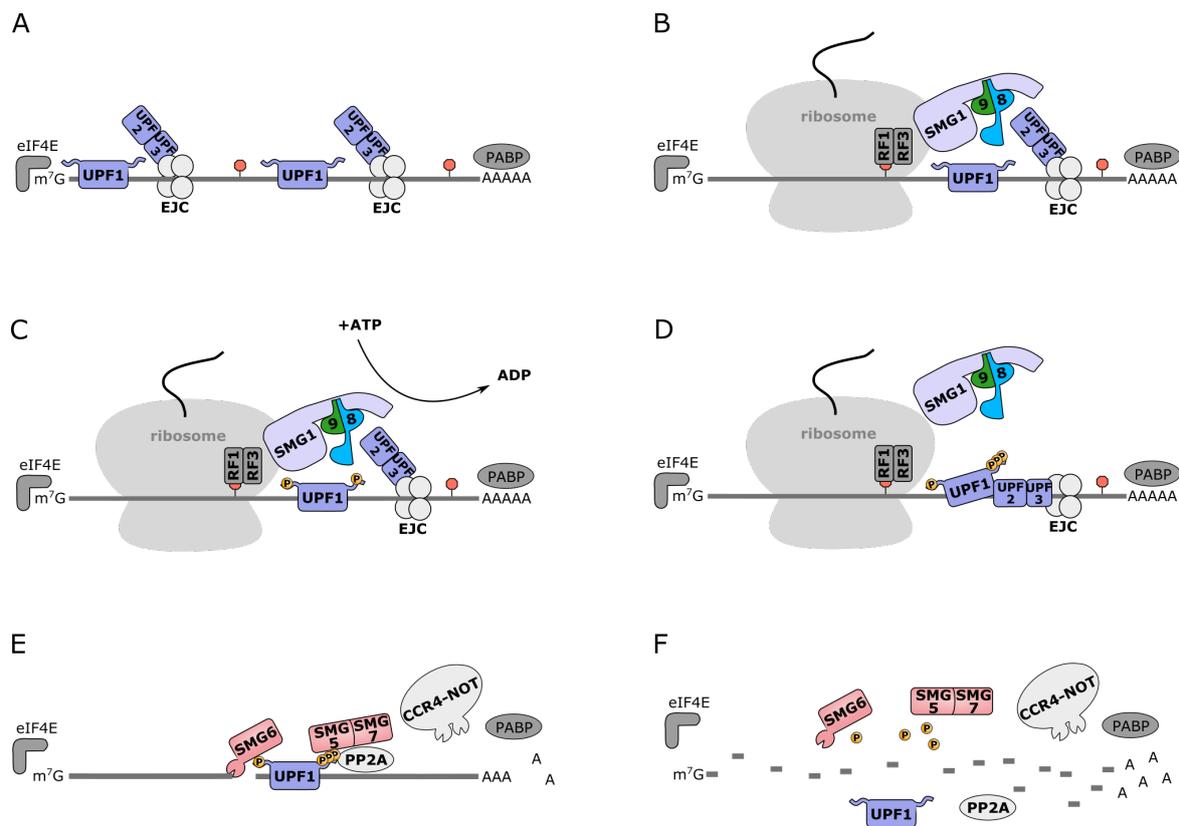


Figure 2.5: Schematic of EJC-dependent NMD. A. NMD-targeted mRNP with a PTC upstream on an EJC. B. and C. Premature translation termination triggers the formation of the SURF complex and ultimately, phosphorylation of UPF1 by SMG1. D. and E. Phosphorylation of UPF1 leads to remodelling of the mRNP and recruitment of downstream effectors, forming the DECID complex. F. SMG6 endonucleolytically cleaves the mRNA while SMG5-7 can recruit the CCR4-NOT complex to trigger 3' to 5' decay. UPF1 can be dephosphorylated by PP2A.

temporal execution, are insufficiently understood.

Fate of the truncated nascent protein chain in EJC-dependent NMD

Early models suggested NMD would exclusively occur during the so-called pioneer-round of translation on cap-binding complex (CBC)- bound messages, that is with the very first ribosome translating a transcript after its maturation and export from the nucleus to the cytoplasm (Maquat *et al.*, 2010). Contrary evidence was presented in different studies, showing that NMD does also occur at later rounds of translation on mRNAs with caps not bound by CBC, but the eIF4F complex (Durand & Lykke-Andersen, 2013; Rufener & Mühlemann, 2013). In fact, single-molecule measurements recently showed that at

least for certain NMD reporters, NMD-mediated degradation is statistically only elicited by every eighth ribosome, with each ribosome displaying the same probability to trigger NMD (Hoek *et al.*, 2019). This observation put a spotlight on a question previously only peripherally addressed: how are the truncated polypeptide chains - which are produced from an NMD mRNA before it is degraded - removed from the cells before they can potentially cause toxic effects? It has been known that the NSD and NGD pathways are tightly coupled to quality control mechanisms triggering the degradation of the resulting nascent chains, but whether truncated nascent chains resulting from NMD transcripts are specifically degraded was not clear. Existing evidence suggested a destabilization of NMD transcript-derived truncated nascent chains (Kuroha *et al.*, 2009; Udy & Bradley, 2021). Recent data now showed that in fact such NMD nascent chains are specifically fed into the Ubiquitin-proteasome system (Chu *et al.*, 2021). Furthermore, this specific polypeptide degradation is independent of the factors involved in clearing NSD and NGD nascent chains. While the detailed mechanism remains unclear, knock-downs of either UPF1 or SMG1 have been shown to impair specific degradation of NMD nascent chains and the corresponding mRNAs. Interestingly, inhibiting SMG1 kinase activity impaired NMD mRNA degradation, but not NMD nascent chain clearance, suggesting that SMG1 may take a scaffolding role in this part of the NMD mechanism (Chu *et al.*, 2021).

2.2.3 EJC-independent NMD

Studies in yeast, plants and animals have shown that mRNAs carrying long 3'UTRs can become susceptible to NMD, even in the absence of a correctly positioned EJC (Amrani *et al.*, 2004; Bühler *et al.*, 2006; Eberle *et al.*, 2008; Kertesz *et al.*, 2006; Kebaara & Atkin, 2009). Placing such 3'UTRs after ORFs initially insensitive to NMD turned these reporters into NMD targets. A possible explanation is thought to be the distance between the termination codon and the poly(A) tail (Behm-Ansmant *et al.*, 2007). The cytoplasmic poly(A) tail-binding protein (PABPC1) can stimulate translation termination, likely through interaction with the N-terminus of eRF3a, possibly competing with UPF1 and thereby suppressing NMD (Singh *et al.*, 2008; Kononenko *et al.*, 2010). In the case of long 3'UTRs this interaction could be reduced due to the physical distance between eRF3a and PABPC1, possibly resulting in slower translation termination and hence favoring NMD (Eberle *et al.*, 2008). In agreement, longer 3'UTRs increase the amount of bound UPF1, enhancing mRNA decay (Hogg & Goff, 2010). Interestingly, certain proteins like PTBP1 have been shown to protect specific mRNAs with long 3'UTRs from NMD by preventing

efficient UPF1 binding to the 3'UTR, thus adding an additional layer of regulatory complexity (Ge *et al.*, 2016). Next to other missing pieces, the absence of information regarding the three-dimensional architecture of these mRNPs hinders a more detailed understanding of the molecular mechanism underlying initiation of NMD triggered by long 3'UTRs.

2.2.4 Other NMD-related mRNA surveillance pathways

While originally discovered as a mechanism degrading mRNAs harboring PTCs, it has become clear that NMD-related pathways can be elicited by many other features (Kishor *et al.*, 2019; Lavysh & Neu-Yilik, 2020). Next to the long 3'UTR model discussed above a variety of such mechanisms has been described, some of which are dependent on specialized factors not present in canonical NMD. Still, they are all centered around the key NMD factor UPF1. Hence, they are commonly referred to as UPF1-mediated mRNA decay (UMD) mechanisms. While canonical EJC-mediated NMD is the focus of this thesis, this section aims to provide a short overview of the diverse landscape of UMD pathways (Kim & Maquat, 2019; Lavysh & Neu-Yilik, 2020).

Staufen-mediated mRNA decay (SMD) relies on a STAU-binding site, a double-stranded RNA structure formed by intra-molecular base pairing. This stem loop can be recognized by the STAU protein dimer, which in turn interacts with UPF2. UPF1 is then recruited and activated via UPF2. SMD also requires translation termination and UPF1 phosphorylation by SMG1 (Gowravaram *et al.*, 2019).

Histone mRNA decay (HMD) is based on stem-loop binding protein (SLBP) recognizing the stem loop substituting a poly(A) tail in histone mRNA. SLBP can interact with UPF1 and successful HMD requires its phosphorylation (Marzluff & Koreski, 2017).

Regnase-1 mediated mRNA decay (RMD) is important to regulate inflammation-related mRNAs. Regnase-1 recognizes a specific stem loop and contains a PIN domain that becomes endonucleolytically active once UPF1 unwinds the stem loop. Regnase-1 is thought to activate UPF1 in a manner analogous to activation by UPF2, and UPF1 N-terminal phosphorylation by SMG1 is thought to be critical for stabilization of the UPF1-Regnase complex (Mino *et al.*, 2019).

In addition to the selected UMDs discussed here, various other UMD pathways have been described and await detailed characterisation. In general, the plethora of UMD mechanisms known to date highlights the adaptability and wide range of NMD-derived mRNA quality control in metazoans (Lavysh & Neu-Yilik, 2020).

2.2.5 NMD in human disease

Estimations suggest about one-third of human inherited disorders are caused by mutations introducing PTCs. Being involved in many human diseases, NMD can have both ameliorating or aggravating effects (Lindeboom *et al.*, 2019; Bhuvanagiri *et al.*, 2010).

Human disease connected to NMD

Historically, NMD was identified as the mechanism destabilizing β -globin mRNAs in β thalassemia patients (Maquat *et al.*, 1981). PTCs located closer to the 5' end of the β -globin mRNA are detected by NMD, triggering mRNA degradation and resulting in the recessive form of the disease. On the other hand, if PTCs are located close to the 3' end they can evade NMD, leading to the dominant form. Thus, NMD is suggested to exert its protective function by reducing the levels of toxic truncated proteins produced.

In contrast, the effect of PTC location on disease outcome is reversed in Duchenne muscular dystrophy: a severe phenotype is caused by PTCs within the dystrophin gene that are eliciting NMD, because they prevent production of a truncated protein that retains partial activity (Lindeboom *et al.*, 2019). In cases where NMD is not recognizing the PTC-containing mRNA, a truncated dystrophin protein is still produced causing a milder disease phenotype termed Becker muscular dystrophy (Bhuvanagiri *et al.*, 2010).

Another example is cystic fibrosis, where the location of the PTC within the CFTR mRNA has been linked to severity of the disease. Furthermore, the efficiency of gene editing by CRISPR-Cas9 is largely influenced by NMD. Finally, NMD has been also connected to the development of different types of cancers (Lindeboom *et al.*, 2019).

Therapeutic targeting of NMD

Its involvement in a diverse set of human diseases has spurred numerous efforts to develop effective treatments by targeting NMD. For example, the functionality of anti-sense oligonucleotides designed to prevent EJC deposition and therefore to stabilize NMD targets has been established (Kim *et al.*, 2022). Efforts have also been made to enhance translational read-through at premature termination codons to remove NMD-eliciting downstream EJCs (Dabrowski *et al.*, 2018). To this end, chemicals targeting translation have been employed, for example aminoglycosides to decrease the accuracy of codon-anticodon pairing (Leubitiz *et al.*, 2019). The read-through promoting oxadiazole-derivative Ataluren has been clinically approved for treating Duchenne muscular dystrophy caused by nonsense muta-

tions (Welch *et al.*, 2007; Campofelice *et al.*, 2019). Towards the same goal, suppressor tRNAs have been successfully used as therapeutics in an animal model by promoting stop codon read-through and inhibiting NMD (Wang *et al.*, 2022). Another approach aims at developing small molecule compounds specifically targeting NMD factors. Goals include disrupting the interaction between SMG7 and UPF1 to stall NMD as well as targeting the SMG1 kinase active site with inhibitors, thereby preventing the crucial step of UPF1 phosphorylation (Durand *et al.*, 2007; Gopalsamy *et al.*, 2012).

2.3 The PIKK family

The SMG1 kinase - an enzyme of central importance to NMD - belongs to the family of phosphatidylinositol 3-kinase-related kinases (PIKKs). Originally identified due to their apparent difference in sequence when compared to known kinases (Keith & Schreiber, 1995), PIKKs are prominent Serine/Threonine (Ser/Thr) protein kinases: ATM (ataxia telangiectasia-mutated), ATR (ataxia- and Rad3-related), DNA-PKcs (DNA protein kinase catalytic subunit), mTOR (mammalian target of rapamycin), TRRAP (transformation/transcription domain-associated protein) and SMG1 (suppressor with morphogenetic effect on genitalia 1). While these large enzymes (about 2400 to 4000 residues) share sequence similarity with phosphatidylinositol 3-kinases (PI3Ks), PIKKs lack their lipid kinase activity. With the exception of the catalytically inactive TRRAP involved in epigenetic regulation (McMahon *et al.*, 1998), all PIKKs take central roles in crucial surveillance pathways throughout eukaryotic cells (Lovejoy & Cortez, 2009). In the nucleus, ATM and DNA-PKcs initiate DNA damage repair upon detection of DNA double-strand breaks, while ATR monitors the genome for single-strand DNA (Shiloh & Ziv, 2013; Zou & Elledge, 2003). In the cytoplasm, mTOR regulates translational activity dependent on nutrient availability and SMG1 triggers nonsense-mediated mRNA decay to degrade prematurely terminating transcripts and regulate overall mRNA levels (Saxton & Sabatini, 2017; Denning *et al.*, 2001; Yamashita *et al.*, 2001). All members of the PIKK family share an overall similar domain architecture (see Figure 2.6) (Yang *et al.*, 2013; Baretic & Williams, 2014; Imseng *et al.*, 2018; Jansma & Hopfner, 2021). They all contain N-terminal α -solenoids which differ in length but constitute binding platforms for numerous interacting and regulatory factors. These HEAT repeats (Huntingtin, elongation factor 3, protein phosphatase 2A, and the yeast kinase Tor1) are followed by a FAT domain (focal adhesion targeting). The neighbouring kinase domains include, with the exception of ATM

and ATR, an N-terminal FRB domain (mTOR FKBP-rapamycin binding). Towards its C-terminus the kinase domain harbors a PRD (PIKK-regulatory domain), which can vary in size from just a few to hundreds of residues as in the case of SMG1 (also called: SMG1 insertion domain). Finally, the two lobes typical for serine/threonine protein kinase domains are neighbored by a FAT C-terminal domain (FATC) (see Figure 2.6). SMG1, ATM, ATR and DNA-PKcs are known to share phosphorylation site specificity for Ser/Thr-Gln (SQ) motifs within their substrates (O'Neill *et al.*, 2000; Kim *et al.*, 1999; Bannister *et al.*, 1993; Yamashita *et al.*, 2001). Interestingly, with the exception of SMG1, all catalytically active PIKKs have been found to be able to dimerize. Except for DNA-PKcs and SMG1, yeast orthologues have been identified for all PIKKs (Baretić & Williams, 2014).

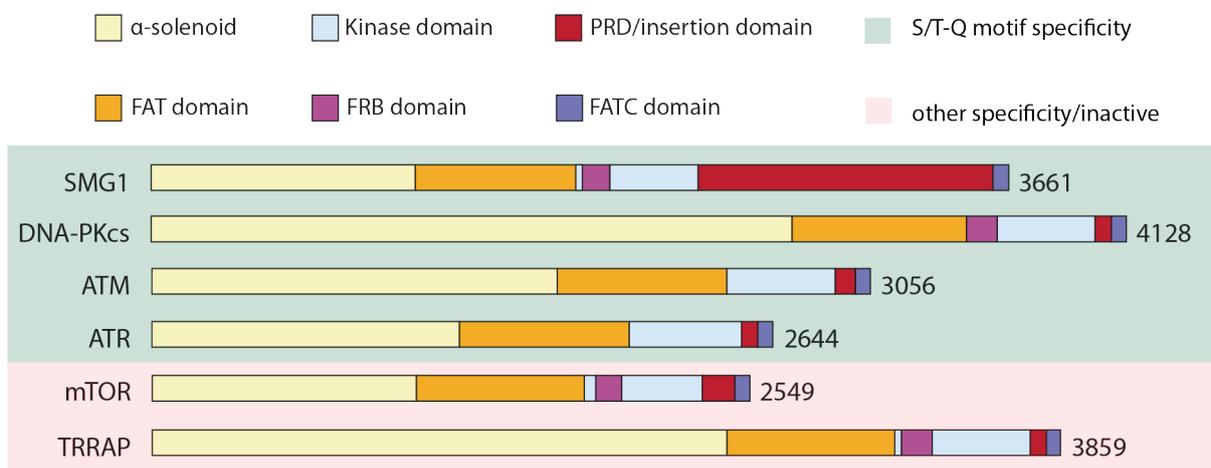


Figure 2.6: Domain organization and phosphorylation motif specificity of the PIKK family. The color scheme used for PIKK domains is shown on top. The background color indicates phosphorylation motif specificity of the respective kinases.

Initial insights into PIKK architecture were provided by a crystal structure of a truncated mTOR-LST8 complex (Yang *et al.*, 2013), a medium-resolution crystal structure of DNA-PKcs (Sibanda *et al.*, 2010) as well as several low-resolution electron microscopy reconstructions (Melero *et al.*, 2014; Yip *et al.*, 2010; Adami *et al.*, 2007; Llorca *et al.*, 2003). Advances in single-particle cryo-electron microscopy (cryo-EM) and increasingly sophisticated protein expression systems have recently allowed to determine structures of all human PIKK family members at resolutions sufficient to allow atomic model building (Yang *et al.*, 2017; Gat *et al.*, 2019; Zhu *et al.*, 2019; Stakyte *et al.*, 2021; Rao *et al.*, 2018; Liang *et al.*, 2022).

2.3.1 SMG1-8-9, a PIKK complex at the heart of metazoan NMD

The gene encoding SMG1 kinase was initially discovered by genetic screening in *C.elegans* and its protein products first characterized as a kinase important for functional NMD by studies focusing on the human orthologue (Hodgkin *et al.*, 1989; Denning *et al.*, 2001; Yamashita *et al.*, 2001). The phosphorylation of UPF1 by SMG1 upon recognition of a premature translation termination event is considered the key event in metazoan NMD committing the targeted transcript to degradation. SMG1 phosphorylates UPF1 in its unstructured N- and C-terminal regions at SQ motifs, enabling UPF1 to recruit downstream NMD effectors (see subsection 2.2.1 and section 2.2.2). Inhibition of SMG1 activity by mutations of residues important for catalysis or small molecules has been shown to stabilize levels of mRNAs otherwise subject to NMD (Yamashita *et al.*, 2001). Of note, SMG1 has also been implicated in genotoxic stress response pathways (Brumbaugh *et al.*, 2004; Gehen *et al.*, 2008; Gewandter *et al.*, 2011).

SMG8 and SMG9 have been identified as tight interactors of the SMG1 kinase (Yamashita *et al.*, 2009). As for SMG1, depletion of SMG8 and SMG9 stabilized NMD mRNA reporters, highlighting their importance for this pathway (Yamashita *et al.*, 2009). Mutations in genes coding for either protein in human patients have been linked to neurodevelopmental disorders (Shaheen *et al.*, 2016; Alzahrani *et al.*, 2020). Both SMG8 and SMG9 enter the NMD pathway early on and hence potentially have important functions in NMD target recognition (Yamashita *et al.*, 2009). Interestingly, depletion of either or both of them from the complex increased SMG1 kinase activity *in vitro*, suggesting that SMG8 and SMG9 can inhibit SMG1 (Yamashita *et al.*, 2009; Arias-Palomo *et al.*, 2011; Deniaud *et al.*, 2015; Zhu *et al.*, 2019). Conversely, *in vivo* data indicated that SMG8 possibly functions by recruiting SMG1 to an early NMD mRNP and knock-down of SMG8 reduced UPF1 phosphorylation by SMG1 in cells (Yamashita *et al.*, 2009). These findings suggest intricate regulatory functions for SMG8 and SMG9 with respect to SMG1 kinase activity, but the details of this mechanism remain elusive.

2.3.2 Structural insights into the SMG1-8-9 complex

Initial low-resolution structural data were provided by negative-stain and cryo-EM studies of either SMG1 in isolation or the assembled trimeric SMG1-8-9 complex (Arias-Palomo *et al.*, 2011; Melero *et al.*, 2014; Deniaud *et al.*, 2015). These studies revealed that SMG1 consists of a N-terminal extended 'arch' region connected to a globular 'head'. Comparison

of reconstructions of SMG1 to a SMG1-8-9 complex identified the extended arch region as the binding site for SMG8 and SMG9. Concomitantly, these reconstructions suggested a compaction of the extended arch upon SMG8-9 binding. Further low-resolution EM studies of crosslinked samples indicated the SMG1 head region as the binding site for its UPF1 substrate, as well as UPF2 (Melero *et al.*, 2014). Using the same methods, this region of SMG1 was furthermore shown to interact with DHX34, an RNA helicase suggested to be involved in recruitment of unphosphorylated UPF1 to SMG1 (Melero *et al.*, 2016). While these studies provided the first data allowing to attempt analysis of structure-function relationships and showed that these interactors can bind to a SMG1-8-9 complex, more detailed insights were hindered by the limited resolution.

The first piece of high-resolution data was provided by a crystal structure of a *C. elegans* SMG8-SMG9 (SMG-8 and SMG-9) complex (Li *et al.*, 2017). This study identified truncated constructs of SMG-8 and SMG-9 sufficient to form a stable complex and revealed a G-domain heterodimer. In contrast to SMG-9, SMG-8 displays a helical bundle made up by three α -helices. C-terminally of the G-domain like globular region, this stalk protrudes into solution connecting to a so far uncharacterized C-terminal domain which was absent in the used construct. The SMG8 stalk was found to be similar to the stalk domain found in GTPases of the dynamin family (Li *et al.*, 2017). Interestingly, only SMG-9 displayed the ability to bind nucleotides.

Overexpression of full-length human SMG1, SMG8 and SMG9 using systems based on HEK293 suspension cells and the advent of high-resolution single-particle cryo-EM led to the reporting of reconstructions of SMG1 and the SMG1-8-9 complex at resolutions better than 4 Å, allowing to build atomic models (Zhu *et al.*, 2019; Gat *et al.*, 2019). These results allowed for the first time to visualize the structure of the 410 kDa SMG1 kinase as well as the roughly 600 kDa SMG1-8-9 complex (see Figure 2.7).

In more detail, they revealed that SMG1 is structurally most similar to mTOR amongst PIKKs (Gat *et al.*, 2019). The high-resolution data confirmed the notion from negative-stain reconstructions that SMG1 consists of an extended N-terminal 'arch' region connecting to a globular C-terminal 'head' (Arias-Palomo *et al.*, 2011; Melero *et al.*, 2014). The 'arch' region is a curved solenoid made up of 14 HEAT repeats. While this region does not mediate homodimerization as observed for other PIKKs, it provides binding sites for SMG8 and SMG9. The 'arch' region is followed by the FAT domain, which is split into a N-terminal 'ring' region made up by α -helical repeats, and a C-terminal curved 'spine' that is made up of eight irregular tetratricopeptide repeats (TPRs). The latter, consistent

with structural observations for other PIKKs, has similarities to the 'spine' found in the lipid kinases PI(3)K, the namesake for the PIKK family (Keith & Schreiber, 1995; Gat *et al.*, 2019). Structurally, the FAT domain packs against the kinase domain of SMG1, which is build up by the N- and C-terminal lobes characteristic for serine/threonine protein kinases, with all residues critical for catalysis conserved amongst enzymatically active PIKKs. In between the FAT and kinase domains, a four-helix bundle with similarity to the mTOR FRB domain was observed. As an important difference to other PIKKs, the PRD of SMG1, termed 'insertion', spans more than 1,100 residues and remained largely unresolved. The very C-terminus of SMG1 shows a well ordered FATC domain, structurally complementing the kinase domain (see Figure 2.7 A and B). Interestingly, the FAT domain, which displays similarities to PI(3)Ks (see above), was found to form a cavity lined with positively charged residues pointing towards a spherical part of the Coulomb potential map not accounted for by the protein. Reversed-phase ion-pair high-performance liquid chromatography coupled to mass spectrometry (RPIP-HPLC) identified this ligand as inositol hexakisphosphate (InsP6, see Figure 2.7 D). Further biochemical experiments and re-inspection of published experimental maps revealed that InsP6 is also a ligand of the only other cytoplasmic PIKK, mTOR (Gat *et al.*, 2019). Interestingly, this ligand was only characterized in one of the published SMG1-8-9 reconstructions. While the same density is clearly present in a SMG1-only reconstruction but remained undescribed, it is absent in the cross-linked SMG1-8-9 reconstruction using the same SMG1 preparation, where it is possibly lost due to the destruction of its binding site by crosslinking primary amines (Zhu *et al.*, 2019).

Within the SMG1-8-9 complex, SMG9 was observed to be wedged in between the N-terminal HEAT repeats and the FAT domain of SMG1. Its globular G-domain core forms extensive interactions with both parts of SMG1, featuring prominently its N-terminus extending over the side of the 'ring' part of the FAT domain. A more C-terminally located extension is closely interacting with two helices of the N-terminal SMG1 solenoid (see Figure 2.7 E). Importantly, the SMG9 reconstruction displayed strong density for a bound ligand. RPIP-HPLC identified ATP as the co-purified nucleotide (Gat *et al.*, 2019). Sequence alignments and structural comparisons revealed that residues within the G-motifs of SMG9 involved in base recognition and GTP hydrolysis in bona fide GTPases have been substituted with amino acids allowing it to bind ATP instead of GTP and possibly rendering it catalytically inactive (Gat *et al.*, 2019; Langer *et al.*, 2020).

Interacting with both SMG9 and the more N-terminal HEAT repeats of the SMG1 solenoid,

SMG8 does not have a nucleotide binding site. An α -helix contacts SMG1 HEATs 8 and 10 via hydrophobic interactions. The stalk is resolved with similar boundaries as used in the *C.elegans* crystal structure, with only very weak density for the approximately C-terminal 40 % of the molecule (see Figure 2.7 C and E). This SMG8 C-terminal domain was observed to protrude towards the SMG1 kinase active site suggesting a regulatory function. Indeed, its removal resulted in kinase hyperactivation, as observed for deletion of the SMG1 insertion/PRD (Zhu *et al.*, 2019; Deniaud *et al.*, 2015). The structures of SMG1 insertion/PRD and SMG8 C-terminus as well as the mechanistic basis of the observed kinase regulatory effects remain unclear.

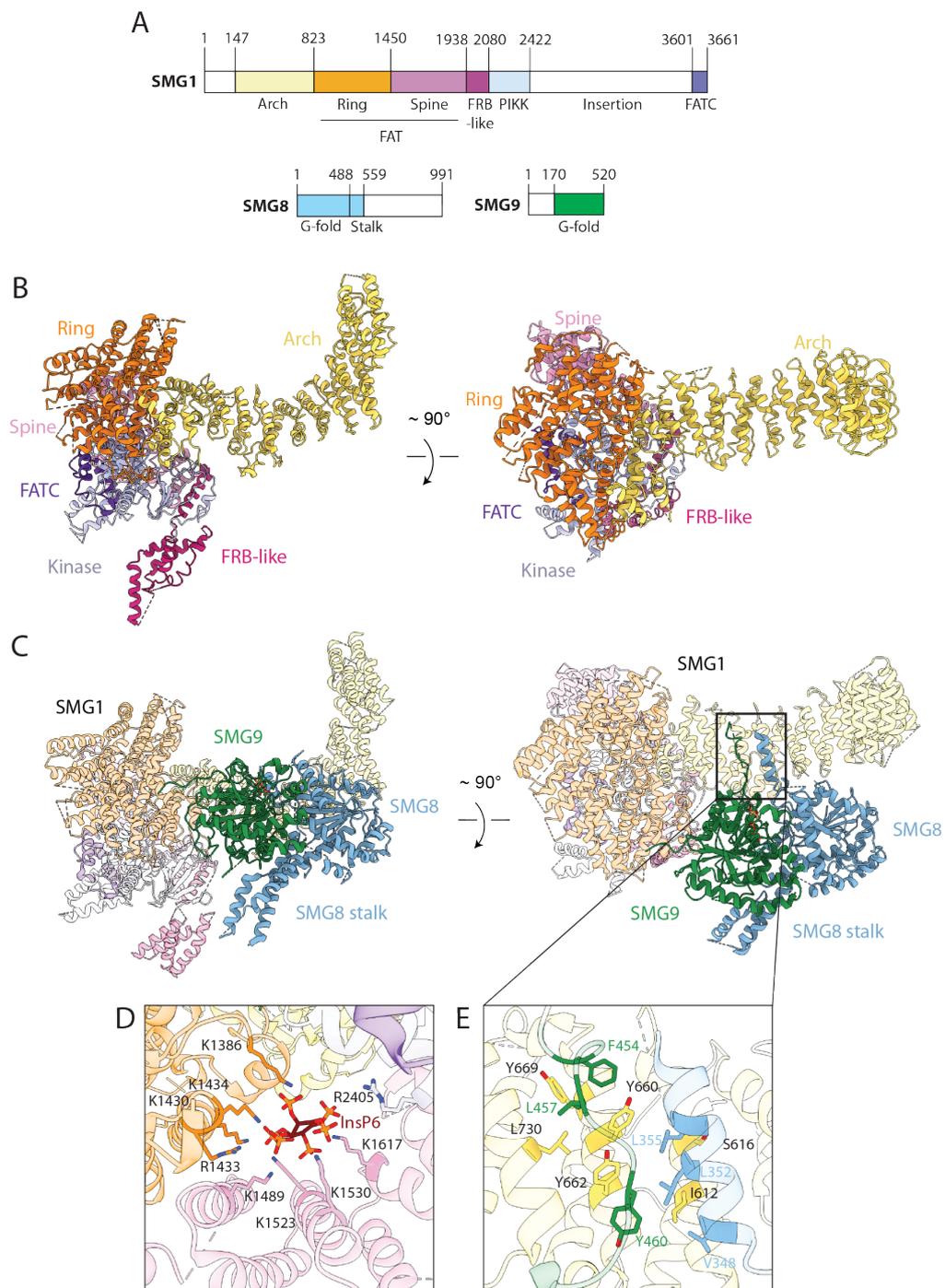


Figure 2.7: Structure of human SMG1 and the SMG1-8-9 complex. A. Domain organization of SMG1, SMG8 and SMG9. Unmodelled regions are transparent. B. Structure of the SMG1 kinase (PDB: 6L53) shown in two orientations. Domains colored according to A. C. Structure of human SMG1-8-9 complex (PDB: 6SYT) shown in two orientations. Coloring according to A. D. The InsP6 ligand sandwiched between ring, spine and kinase domains. Coordinating, positively charged SMG1 residues are shown as sticks. Coloring according to A. E. Interaction of SMG8 and SMG9 segments with the SMG1 arch domain. Hydrophobic residues involved in interactions are highlighted as sticks. Coloring according to A.

2.4 Objective of this thesis

UPF1 phosphorylation by SMG1-8-9 marks the committed step in metazoan NMD. While structural and biochemical studies have provided insights into the architecture of SMG1-8-9, how substrate specificity is achieved and how this critical kinase activity is regulated remains largely unknown.

This work builds on previous structural work that provided the first high-resolution structure of SMG1-8-9 (Gat *et al.*, 2019). While it is known that UPF1 is phosphorylated on its N- and C-terminal SQ motifs, it is not understood how SMG1 specifically selects these motifs within UPF1, and whether additional residues can contribute to phosphorylation site selection. Understanding SMG1 kinase specificity on a structural and biochemical level will shed light on a key event in the crucial NMD surveillance pathway.

Biochemical assays have shown that several parts of the SMG1-8-9 complex contribute to its regulation - in particular the SMG1 insertion/PRD and the SMG8 C-terminal domain - but the mechanism remains unclear. Insights into this process can deepen our understanding of SMG1 regulation and build the foundation for integrating SMG1-8-9 activity into the NMD pathway.

Small molecule inhibitors have proven to be an invaluable tool in targeting specific kinases to combat devastating human diseases. Several compounds have been described for different PIKKs with some already advanced to the clinic, but a limited understanding of chemical features providing specificity to certain PIKKs has slowed their development. While some structural data are available for mTOR inhibitors, no insights into small molecules specifically targeting SMG1 have been obtained. Making such data available can provide the foundation for structure-based drug design, with the potential to unlock so far untapped therapeutic potential.

Finally, investigating the mechanistic basis of crosstalk between the translation machinery and early-acting NMD factors such as SMG1-8-9 and UPF1 can provide unprecedented insight into the recognition of transcripts by the NMD machinery and therefore co-translational mRNA quality control in general.

Answering the outlined questions will not only improve the understanding of metazoan NMD and point into new directions of investigation, but also provide more general insights into PIKK biology: no structural data on kinase-substrate interactions are available for any of the PIKK family members and the dissection of the regulatory mechanisms governing their activity remains an area of active research.

Chapter 3

Materials and Methods

This chapter will detail material and methods essential to reproduce experiments reported in Chapter 4 that have not been published.

3.1 Materials

3.1.1 Antibodies used in this study

All antibodies used in this work are listed in Table 3.1 and Table 3.2. All stocks were diluted in 5% (v/v) non-fat dry milk in PBS-T buffer (1xPBS, Life Technologies and 0.1% Tween-20, Biorad) for usage.

Table 3.1: Primary antibodies used in this study. Target species for all antibodies is human.

Antigen	Ab species	Ab type	Source (Cat. No.)	Dilution
FLAG tag	mouse	monoclonal	Sigma-Aldrich (F3165)	1:5000
eRF1	rabbit	polyclonal	ThermoFisher (PA5-28777)	1:3000
eRF3a	rabbit	polyclonal	ThermoFisher (PA5-28256)	1:3000
ABCE1	rabbit	monoclonal	Abcam (ab185548)	1:2000
RPS24	rabbit	monoclonal	Abcam (ab196652)	1:1000

Table 3.2: Secondary antibodies used in this study. All HRP-conjugated.

Target species	Ab species	Ab type	Source (Cat. No.)	Dilution
rabbit	goat	polyclonal	Bio-Rad (172-1019)	1:10000
mouse	goat	polyclonal	Bio-Rad (172-1011)	1:10000

3.2 Methods

3.2.1 Expression and purification of eukaryotic release factors 1 and 3a

All human eukaryotic release factor (eRF) constructs were expressed in *Escherichia coli* (*E. coli*) BL21 Star pRARE cells. Cultures were induced with 0.5 mM IPTG at an OD of 2 and proteins were expressed overnight at 18 °C. eRF1 wildtype and GGQ-to-AAQ mutant (Frolova *et al.*, 1999) were essentially purified as described before (Song *et al.*, 2000; Brown *et al.*, 2015).

eRF3a wildtype and H299L mutant (Salas-Marco & Bedwell, 2004) were expressed with a N-terminal 6xHis-tag followed by a TEV cleavage site. Bacterial pellets were resuspended in lysis buffer (LB) containing 10 mM Tris-Cl pH 7.5, 500 mM NaCl, 1 mM DTT, 10 % (v/v) glycerol, 0.01 % (v/v) NP-40, 1 mM PMSF and cOmplete protease inhibitor (Roche) at a buffer-to-pellet ratio of 2:1. Cells were lysed by sonication and debris was pelleted for 30 min, 25,000 rpm at 4 °C using a JA25-50 rotor (Beckman Coulter). Cleared and filtered lysate was applied to a Ni²⁺-NTA affinity column. Bound protein was washed with 15 column volumes (CV) of LB and eluted using LB supplemented with 300 mM imidazole pH 8. The elution was combined with His-tagged TEV protease and dialysed against 20mM HEPES pH 7.4, 150mM NaCl, 10 % (v/v) glycerol and 1 mM DTT (Hep A) overnight at 4 °C. The dialysis product was supplemented with imidazole to a final concentration of 10 mM and passed over Ni²⁺-NTA affinity column to remove uncleaved protein and protease and directly loaded on a Heparin column. A salt gradient from 0 % to 50 % over 20 CVs using Hep A and Hep B (20mM HEPES pH 7.4, 1000 mM NaCl, 10 % (v/v) glycerol and 1 mM DTT) was used for elution. Peak fractions were collected and concentrated (Amicon concentrator MWCO 30) before sizing the protein content over an S200 HiLoad size exclusion column (GE Healthcare) equilibrated with 20mM HEPES pH 7.4, 100 mM KCl, 10 % (v/v) glycerol and 1 mM DTT. Peak fractions were concentrated as before, plunge frozen using liquid nitrogen and stored at -80 °C. All purifications steps were monitored using SDS-polyacrylamide gel electrophoresis (SDS-PAGE) followed by Coomassie-staining.

3.2.2 *In vitro* transcription

In vitro transcription was carried out with T7 RNA Polymerase (in-house) using PCR-generated, gel-purified template at 1 pmol per 10 μ l reaction and scaled up as needed. The reaction was incubated for four hours at 37°C, before 0.5 μ l RNase-free DNase I (Roche) per 10 μ l reaction was added for one hour at 37°C. Progression of the transcription reaction was judged by the appearance of pyrophosphate precipitate. The RNA was purified by LiCl₂ precipitation followed by EtOH precipitation, finally resuspended in RNase-free H₂O and stored at -20°C. The composition of a transcription reaction is listed in Table 3.3.

Table 3.3: Transcription mix used in this study.

[final]	reagent
2.5 mM each	NTP mix
1x	transcription buffer
2 μ M	DNA template
100 U/ μ l	T7 polymerase
10 mM	Spermidine (Sigma No. 05292)
\emptyset	ddH ₂ O

The composition of the 10x transcription buffer mix is listed in Table 3.4.

Table 3.4: 10x transcription buffer used in this study.

[final]	reagent	stock
400 mM	Tris-HCl pH 8.0	1 M
280 mM	MgCl ₂	1 M
0.1 % (v/v)	Triton X-100 (Fluka No. 93443)	10 % (v/v)
50 mM	DTT	2 M
10 mM	Spermidine (Sigma No. 05292)	0.1 M
\emptyset	ddH ₂ O	\emptyset

3.2.3 *In vitro* translation

In vitro translation reactions were prepared using 70% (v/v) processed rabbit reticulocyte lysate (RRL, Promega) per reaction. RNA was added to a final concentration of 0.5 pmol/ μ l. Recombinantly purified proteins were used at a final concentration of 4 μ M. The reactions were brought up to 100% using ddH₂O or a compensation buffer. Translation mixtures were incubated at 32°C for 10 min. The incubation time was extended with increasing reaction volumes. Finally, reactions were quenched by transfer to ice and either

addition of SDS sample buffer (for western blotting, no boiling) or pull down buffer (for purification) at a 1:1 (v/v) ratio.

3.2.4 Purification of stalled ribosome - nascent chain complexes

All buffers used during ribosome - nascent chain (RNC) complex purification are listed in Table 3.5 and were derived from previously published protocols (Brown *et al.*, 2015; Shao *et al.*, 2016; Feng & Shao, 2018). Directly following *in vitro* translation, reactions were diluted 1:1 with RNCW-1 buffer and input samples were set aside. The diluted reactions were then incubated with pre-equilibrated magnetic, porous Agarose anti-FLAG beads (Sigma, use 12 % of translation reaction volume of bead slurry) for 20 min at 4 °C. Next, the flow-through was collected and a sample taken (A-FT). In some instances the Agarose beads were washed and eluted as subsequently described for Dynabeads (A-E). The flow-through was combined with magnetic, non-porous, ProteinG-coupled Dynabeads (Invitrogen, use 50 % of translation reaction volume of bead slurry) pre-equilibrated with RNCW-1 and anti-FLAG antibody. After incubation for 30 min, the beads were collected and the flow-through set aside (D-FT). Beads were washed four times with 15 CVs each RNCW-1 followed by four washes with 15 CVs each RNCW-2. RNCW-1 and/or RNCW-2 washes were substituted with RNCW-H in certain protocols where necessary to produce the desired translation termination complexes. Bound RNC complexes were eluted using RNC-E (50 % of initial slurry volume) for 30 min at room temperature (D-E) and ribosome concentration was estimated by measuring absorbance at 260 nm.

Table 3.5: Buffers used in ribosome - nascent chain complex purifications.

buffer name	composition
RNCW-1	20 mM HEPES pH 7.4, 100 mM KCl, 5 mM MgCl ₂ , 1 mM DTT, 0.1 % (v/v) Triton-X100
RNCW-2	20 mM HEPES pH 7.4, 100 mM KCl, 5 mM MgCl ₂ , 1 mM DTT, 0.05 % (v/v) Triton-X100
RNCW-H	20 mM HEPES pH 7.4, 750 mM KCl, 15 mM MgCl ₂ , 1 mM DTT, 0.1 % (v/v) Triton-X100
RNC-E	20 mM HEPES pH 7.4, 100 mM KCl, 5 mM MgCl ₂ , 1 mM DTT, 200 ng/ μ l 3xFLAG peptide

3.2.5 Western blotting

After separating proteins using SDS-PAGE and stain-free visualization of proteins with supplemented 2,2,2-trichloroethanol (TCE, Acros Organics) (Ladner *et al.*, 2004), a wet transfer sandwich was assembled using PVDF Immobilon-PSQ Transfer Membrane (Millipore, pore size $0.2\ \mu\text{m}$) and Whatman-3MM sheets. The membrane was activated by incubation in 100% MeOH, followed by rinsing with ddH₂O. Protein was transferred for 90 min at 200 mA using ice-cold transfer buffer. The composition of the transfer buffer is listed in Table 3.6. Success of the transfer was monitored using TCE-based stain-free visualization.

Table 3.6: PVDF transfer buffer used in this study.

[final]	reagent
20 mM	Trizma base (Sigma)
150 mM	Glycine (Sigma)
0.01 %	SDS
5 %	MeOH (Sigma)
∅	ddH ₂ O

Following transfer, the membrane was blocked for 1 h at room temperature using 5% (v/v) non-fat dry milk in PBS-T buffer (1x PBS, Life Technologies and 0.1% Tween-20, Biorad). The blocking buffer was replaced with primary antibody solution (see subsection 3.1.1), followed by incubation for 1 h at 4°C. After washing the membrane four times for 5 min with PBS-T buffer, secondary antibody solution was added for incubation for 30 min at 4°C. Finally, the membrane was washed three times in PBS-T buffer and once in 1x PBS, before being rinsed in water and imaged using ECL-Prime reagent (GE Healthcare) and a LAS 4000 imager.

Chapter 4

Results

This chapter will present the published experimental results as well as additional unpublished key findings that form the basis of this thesis.

4.1 Structure of substrate-bound SMG1-8-9 kinase complex reveals the molecular basis for phosphorylation specificity

The manuscript presented in this section describes the structure of a SMG1-8-9 complex with a UPF1-derived peptide and AMPPNP bound to the SMG1 kinase active site. These structural data visualize for the first time how SMG1 specifically recognizes SQ motifs within UPF1 for phosphorylation, the key step in metazoan NMD. Together with mass spectrometry-based assays, the structure rationalizes the preference of SMG1 for phosphorylating SQ motifs with a Leu residue in the -1 position. More generally, this publication provides the first structural insight into substrate recognition by a PIKK family member. By comparison with other PIKKs, the structure facilitates broader insights into PIKK regulatory mechanisms (Langer *et al.*, 2020).

Structure of substrate-bound SMG1-8-9 kinase complex reveals molecular basis for phosphorylation specificity

Lukas M Langer, Yair Gat, Fabien Bonneau, Elena Conti*

Department of Structural Cell Biology, Max Planck Institute of Biochemistry, Martinsried, Germany

Abstract PI3K-related kinases (PIKKs) are large Serine/Threonine (Ser/Thr)-protein kinases central to the regulation of many fundamental cellular processes. PIKK family member SMG1 orchestrates progression of an RNA quality control pathway, termed nonsense-mediated mRNA decay (NMD), by phosphorylating the NMD factor UPF1. Phosphorylation of UPF1 occurs in its unstructured N- and C-terminal regions at Serine/Threonine-Glutamine (SQ) motifs. How SMG1 and other PIKKs specifically recognize SQ motifs has remained unclear. Here, we present a cryo-electron microscopy (cryo-EM) reconstruction of a human SMG1-8-9 kinase complex bound to a UPF1 phosphorylation site at an overall resolution of 2.9 Å. This structure provides the first snapshot of a human PIKK with a substrate-bound active site. Together with biochemical assays, it rationalizes how SMG1 and perhaps other PIKKs specifically phosphorylate Ser/Thr-containing motifs with a glutamine residue at position +1 and a hydrophobic residue at position -1, thus elucidating the molecular basis for phosphorylation site recognition.

Introduction

Family members of phosphatidylinositol 3-kinase-related kinases (PIKKs) activate distinct signaling pathways that promote cellular survival in different environmental and endogenous stress conditions (*Baretic and Williams, 2014; Imseng et al., 2018; Lempiäinen and Halazonetis, 2009*). Specifically, PIKKs oversee translation machinery in the cytoplasm (mTOR, SMG1), or regulate DNA damage repair in the nucleus (ATM, ATR and DNA-PK) (*Shimobayashi and Hall, 2014; Saxton and Sabatini, 2017; Yamashita, 2013; Yamashita et al., 2001; Blackford and Jackson, 2017; Elías-Villalobos et al., 2019*). With the exception of the enzymatically inactive TRAPP/Tra1, which serves as a scaffold in chromatin modification complexes, all other members of the PIKK family are Ser/Thr-protein kinases, and are among the largest proteins in the eukaryotic kinome. Recent publications have revealed the organization of PIKK active sites at better than 4 Å resolution (*Gat et al., 2019; Zhu et al., 2019; Yang et al., 2013; Jansma et al., 2020; Yates et al., 2020*), but the key question of how members of this kinase family recognize their substrates remains unanswered.

Human SMG1 is one of the largest PIKK family members (~410 kDa) and plays a crucial role in nonsense-mediated mRNA decay (NMD), a conserved pathway that regulates mRNA stability in the cytoplasm of eukaryotic cells (*Kurosaki and Maquat, 2016; Karousis and Mühlemann, 2019*). In its canonical surveillance function, the NMD pathway recognizes and degrades aberrant mRNAs containing premature translation termination codons, thus preventing the accumulation of truncated protein products. In addition, NMD also regulates the levels of a subset of normal, physiological transcripts, amounting to 5–10% of the transcriptome. In metazoans, SMG1 forms a stable complex with two additional proteins, SMG8 and SMG9, and specifically phosphorylates the RNA helicase UPF1. Phosphorylation of UPF1 is a crucial event in this pathway as it enables the recruitment of downstream NMD factors SMG5, SMG6 and SMG7, leading to ribonucleolytic cleavage of the RNA.

*For correspondence: conti@biochem.mpg.de

Competing interests: The authors declare that no competing interests exist.

Funding: See page 11

Received: 21 March 2020

Accepted: 28 May 2020

Published: 29 May 2020

Reviewing editor: Philip A Cole, Harvard Medical School, United States

© Copyright Langer et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

eLife digest The instructions for producing proteins in the cell are copied from DNA to molecules known as messenger RNA. If there is an error in the messenger RNA, this causes incorrect proteins to be produced that could potentially kill the cell. Cells have a special detection system that spots and removes any messenger RNA molecules that contain errors, which would result in the protein produced being too short.

For this error-detecting system to work, a protein called UPF1 must be modified by an enzyme called SMG1. This enzyme only binds to and modifies the UPF1 protein at sites that contain a specific pattern of amino acids – the building blocks that proteins are made from. However, it remained unclear how SMG1 recognizes this pattern and interacts with UPF1.

Now, Langer et al. have used a technique known as cryo-electron microscopy to image human SMG1 bound to a segment of UPF1. These images were then used to generate the three-dimensional structure of how the two proteins interact. This high-resolution structure showed that protein building blocks called leucine, serine and glutamine are the recognized pattern of amino acids. To further understand the role of the amino acids, Langer et al. replaced them one-by-one with different amino acids to see how each affected the interaction between the two proteins. This revealed that SMG1 preferred leucine at the beginning of the recognized pattern and glutamine at the end when binding to UPF1.

SMG1 is member of an important group of enzymes that are involved in various error detecting systems. This is the first time that a protein from this family has been imaged together with its target and these findings may also be relevant to other enzymes in this family. Furthermore, the approach used to determine the structure of SMG1 and the structural information itself could also be used in drug design to improve the accuracy with which drugs identify their targets.

SMG1 phosphorylates UPF1 specifically at Ser/Thr - Gln (SQ) motifs present in the unstructured N- and C-terminal regions that flank the helicase core (Yamashita et al., 2001; Denning et al., 2001). Specificity for a glutamine residue at the +1 position is shared by other PIKK family members, namely, ATM, ATR and DNA-PK kinases (Kim et al., 1999; Bannister et al., 1993). However, there is an additional layer of phosphorylation site specification. Among the 20 possible SQ motifs in UPF1, studies in vitro and in vivo have shown that only a selected few are effectively phosphorylated, including Ser1073, Ser1078, Ser1096 and Ser1116 (Yamashita et al., 2001; Ohnishi et al., 2003; Durand et al., 2016). Interestingly, these UPF1 phosphorylation sites share a Leu-Ser-Gln (LSQ) consensus sequence identical to the LSQ consensus motif identified in substrates of the ATM kinase (O'Neill et al., 2000; Kim et al., 1999). In this work, we studied the interaction between recombinant human SMG1-SMG8-SMG9 with UPF1 peptides using cryo-EM and mass spectrometry to identify the molecular basis with which SMG1, and potentially other PIKKs, recognizes specific phosphorylation sites in its substrate.

Results and discussion

Cryo-EM structure of the human SMG1-8-9 kinase complex bound to a UPF1 peptide

We used stably transfected HEK293T cells to express and purify a human wild-type SMG1-8-9 complex, as previously reported (Gat et al., 2019). The complex phosphorylated full-length recombinant UPF1 in a radioactive kinase assay (Figure 1—figure supplement 1A), confirming the enzymatic activity of purified SMG1 towards its physiological substrate. We selected a frequently phosphorylated site within UPF1 (Yamashita et al., 2001), Ser1078, and used a peptide spanning residues 1074–1084 (hereby defined as UPF1-LSQ) for subsequent structural and biochemical analysis (Figure 1—figure supplement 1B). We confirmed the ability of SMG1-8-9 to specifically phosphorylate UPF1-LSQ using a mass spectrometry-based phosphorylation assay (Figure 1—figure supplement 1C and D). This assay allowed us to monitor the relative amount of phosphorylation of a specific peptide over time. As a control, phosphorylation was abolished when Ser1078 was changed to Asp

(Figure 1—figure supplement 1C and D). Hence, the reconstitutions used in this study recapitulate specific phosphorylation site selection.

For structure determination, we incubated SMG1-8-9 with UPF1-LSQ and AMPPNP, a non-hydrolyzable ATP analogue, and subjected the sample to cryo-EM single particle analysis. The final reconstruction reached an overall resolution of 2.9 Å (Figure 1—figure supplements 2 and 3), and allowed us to further complete and refine the published model for SMG1-8-9 (Supplementary file 1; Gat et al., 2019). Briefly, SMG1 consists of an N-terminal solenoid ‘arch’ and a compact C-terminal ‘head’ region (Figure 1A and B). The C-terminal ‘head’ is formed by the tight interaction between the catalytic module, typical of Ser/Thr-kinases, and the so-called FAT and FATC domains (Imseng et al., 2018; Baretic and Williams, 2014; Bosotti et al., 2000). The N-terminal ‘arch’ provides binding sites for both SMG8 and SMG9 (Figure 1A and B). As we had previously reported, SMG9 contains an unusual G-fold domain that binds ATP rather than GTP or GDP (Gat et al., 2019).

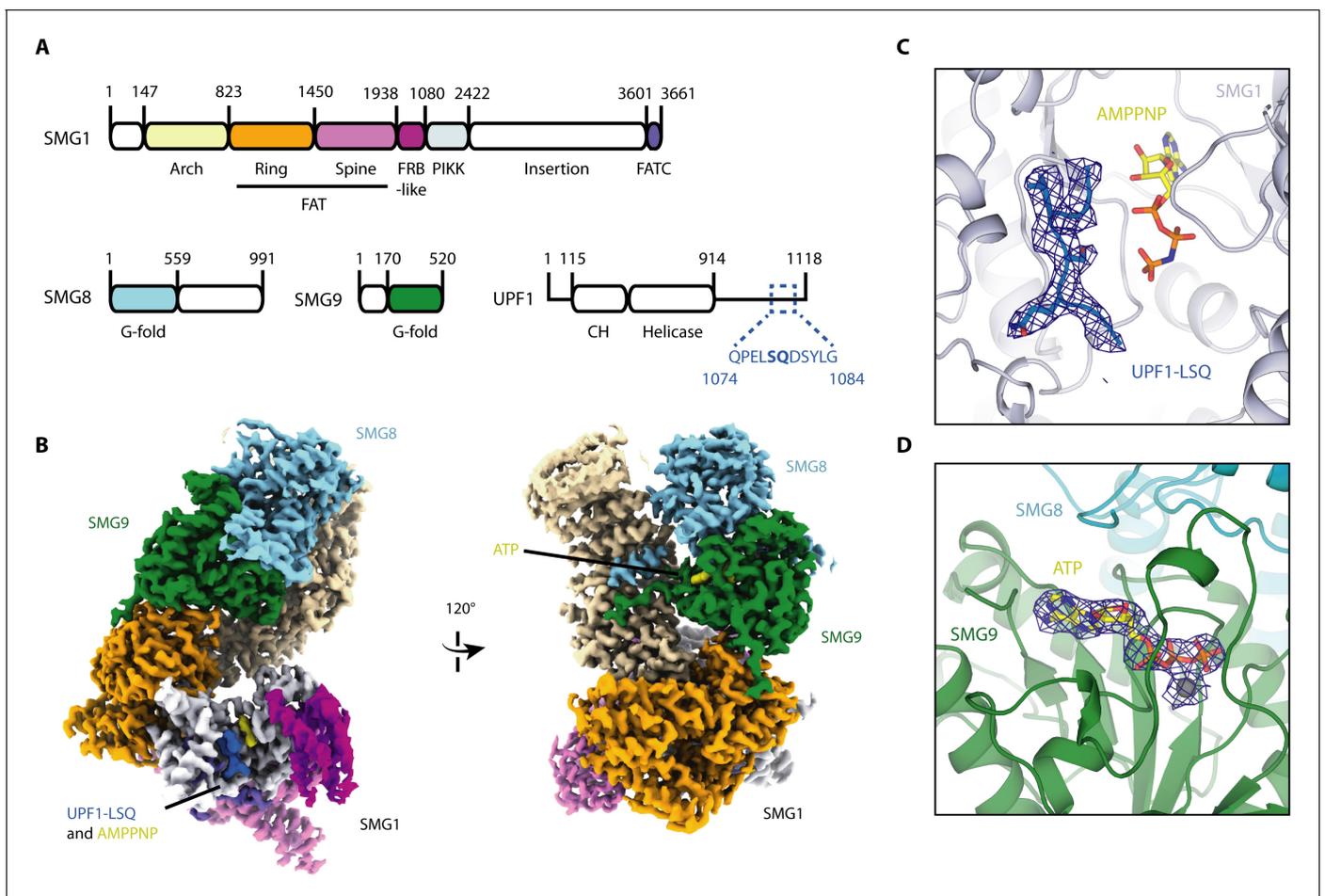


Figure 1. Cryo-EM reconstruction of SMG1-8-9 bound to UPF1-LSQ. (A) Domain organization of SMG1, SMG8, SMG9 and UPF1. White parts are not resolved in the reconstruction. The sequence and location of UPF1-LSQ is indicated with blue text and dotted lines. (B) Segmented cryo-EM reconstruction of substrate-bound SMG1-8-9. Two different views are shown; proteins and domains are colored as in A. (C) A zoomed-in view of SMG1 showing the kinase active site with bound AMPPNP and UPF1-LSQ. Reconstructed density for UPF1-LSQ is shown as a blue mesh. (D) Zoom-in showing ATP bound to SMG9 with reconstructed density displayed as a blue mesh.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. SMG1-8-9 activity and UPF1 SQ motifs.

Figure supplement 2. Cryo-EM analysis of SMG1-8-9 bound to UPF1-LSQ.

Figure supplement 3. Cryo-EM data processing scheme.

Figure supplement 4. SMG9 is a G-fold containing protein binding ATP and exhibits distinct differences to the *bona fide* GTPase RAS.

Figure supplement 5. Quality of the reconstructed density.

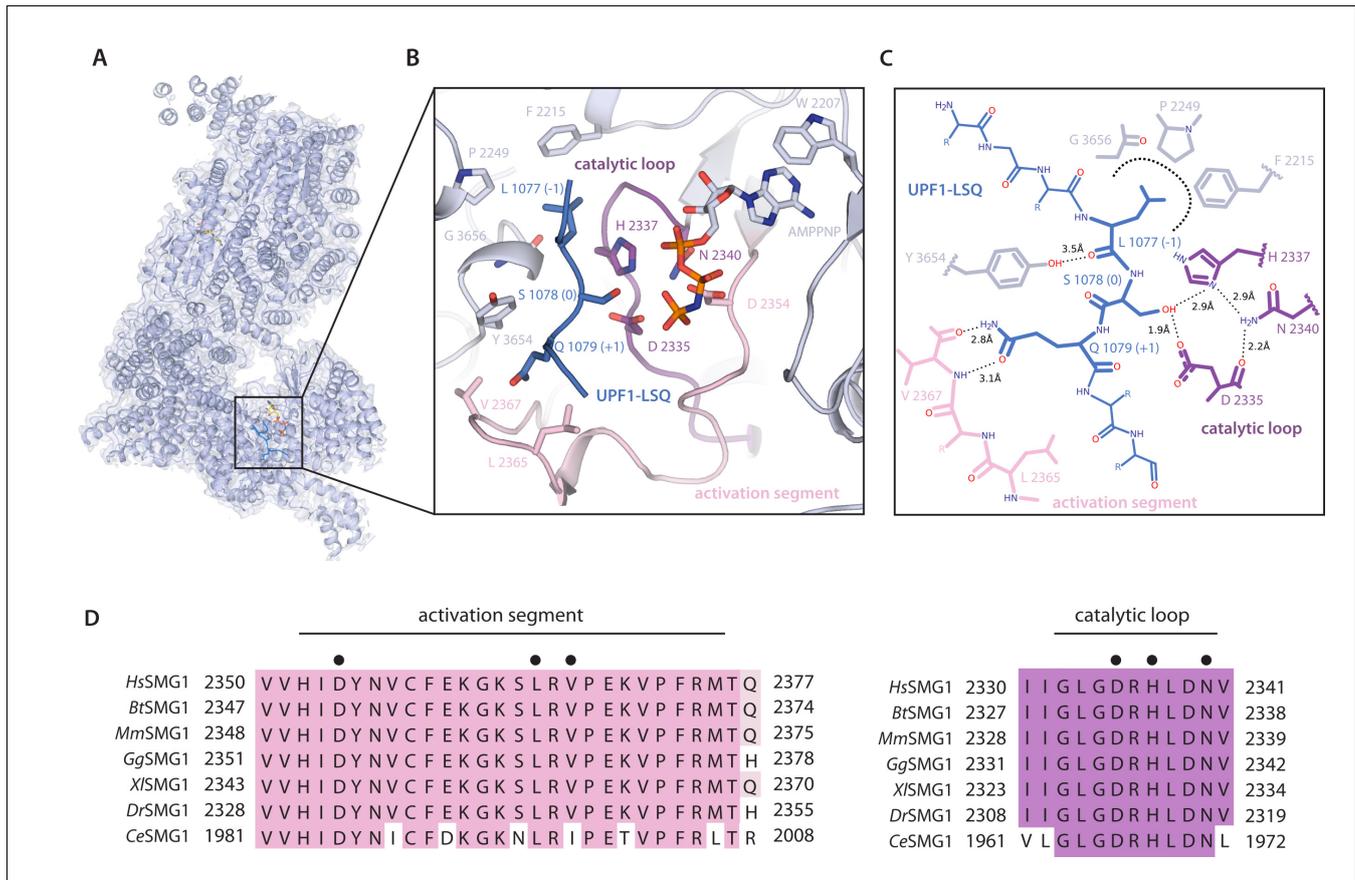


Figure 2. Organization of the substrate-bound kinase active site. (A) The structure of the entire complex is shown overlaid with transparent reconstructed density. The black box indicates the location of the kinase active site. (B) SMG1 active site with important residues shown as sticks. Activation segment and catalytic loop as indicated; the p-loop was omitted for clarity. UPF1-LSQ is shown in blue with positions of important residues highlighted. (C) Two-dimensional sketch of the SMG1 active site with key kinase-substrate interactions indicated. (D) Activation segment (pink) and catalytic loop (magenta) regions of a SMG1 sequence alignment are shown, indicating a high level of conservation across *Homo sapiens*, *Bos taurus*, *Mus musculus*, *Gallus gallus*, *Xenopus laevis*, *Danio rerio* and *Caenorhabditis elegans*. Key residues shown in B are highlighted by a black dot and activation segment and catalytic loop are indicated.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Comparison of SMG1 substrate-bound active site to other protein kinases.

Figure supplement 2. Recognition of UPF1-LSQ phospho-acceptor residue Ser1078 by SMG1 catalytic loop residues.

The local resolution of around 3 Å allowed us to model SMG9-bound ATP in the reconstructed density, revealing the molecular basis for how the adenosine nucleotide is recognized by this unusual G-fold domain (Figure 1B and D, Figure 1—figure supplement 4). Briefly, the G4 and G5 motifs responsible for the recognition of the base have rearranged to preferentially bind an adenine base rather than a guanine (Figure 1—figure supplement 4).

Importantly, compared to the previously published apo-SMG1-8-9 structure (Gat et al., 2019), the current reconstruction revealed extra density in the kinase active site accounting for both AMPPNP and UPF1-LSQ (Figure 1B and C, Figure 1—figure supplement 5A,B and C).

Positioning of UPF1 Ser1078 in the SMG1 active site for phosphoryl transfer

The active site of SMG1 shows excellent density for residues 1075–1081 of UPF1-LSQ (Figure 1C, Figure 1—figure supplement 5). The directionality of the bound substrate peptide is consistent with that of other kinase structures, such as CDK2, and the arrangement of key active site residues is well conserved (Figure 2—figure supplement 1; Bao et al., 2011). The geometry of the catalytic

loop (residues 2332 to 2340) and of the activation segment (residues 2352 to 2375) in the SMG1 kinase domain as well as orientation of important active site residues are very similar to those observed in mTOR (**Figure 2—figure supplement 1B**), indicative of an active kinase state (**Yang et al., 2013**). Specific recognition of the phosphorylation site is achieved via conserved residues contributed by the activation segment and the catalytic loop as well as by the FATC domain (**Figure 2**). The hydroxyl group of Ser1078, the phospho-acceptor residue in UPF1-LSQ, is positioned by residues of the catalytic loop, in particular Asp2335 and His2337 (**Figure 2B and C, Figure 2—figure supplement 2A and B**). Consistent with the structural observations, mutation of either of the corresponding residues in SMG1, mTOR and other protein kinases results in catalytically inactive enzyme (**Bao et al., 2011; Madhusudan et al., 2002; Yang et al., 2013; Brown et al., 1995; Yamashita et al., 2001; Denning et al., 2001**). Therefore, the overall architecture of the substrate-bound SMG1 catalytic module corroborates the structural conservation among PIKK active sites and reveals that positioning of the substrate phospho-acceptor is achieved by residues that are shared between a wide range of protein kinases.

Crucial recognition of a glutamine residue at +1 position of the UPF1 consensus motif

A glutamine residue following the phospho-acceptor site is the minimal requirement for UPF1 phosphorylation by SMG1 (**Figure 3A; Yamashita et al., 2001**). To validate the importance of this residue, we performed a mass spectrometry-based phosphorylation assay using a series of peptides based on UPF1-LSQ. We changed the residue at position +1 in the UPF1-LSQ peptide to test the effect of different side chain properties on phosphorylation. Only wildtype UPF1-LSQ was efficiently phosphorylated by SMG1 (**Figure 3B and Figure 1—figure supplement 1C**). In our structure, the glutamine at position +1 of UPF1-LSQ reaches into a hydrophobic cage formed by the SMG1 activation segment and FATC domain (**Figure 3C**). In particular, UPF1 Gln1079 stacks against Tyr3654 and Leu2365 and forms hydrogen bonds with the backbone of Val2367 (**Figure 3C**). The hydrophobic cage is highly conserved in other PIKKs that recognize SQ motifs (ATM, ATR and DNA-PK) but is different in mTOR (where Glu2369 is found at the equivalent position of SMG1 Leu2365) (**Figure 3D**). This difference is also apparent from the superposition of SMG1 with the 2.8 Å resolution structure of the ATM orthologue *Chaetomium thermophilum* (Ct) Tel1^{ATM} and with mTOR (**Figure 2—figure supplement 1C and D; Jansma et al., 2020; Yang et al., 2013**). While the geometry of the hydrophobic cage is highly similar between SMG1 and CtTel1^{ATM}, it deviates in mTOR due to the described Leu to Glu substitution. Indeed, mTOR has been found to prefer small or non-polar residues at position +1 of its phosphorylation consensus motif (**Hsu et al., 2011**). Taken together, these observations provide a rationale for the difference in phosphorylation site specificity between SMG1, ATM, ATR, DNA-PK and mTOR. Intriguingly, the structural superposition with CtTel1^{ATM} shows that its PIKK regulatory domain (PRD) places a Gln residue in the corresponding hydrophobic cage, effectively occupying the substrate Gln binding site (**Figure 2—figure supplement 1C**). This explains the autoinhibitory function of the ATM PRD domain (**Jansma et al., 2020; Yates et al., 2020**). The corresponding PRD domain in SMG1 is a ~ 1100 amino-acid long insertion (**Figure 1A**) that negatively impacts its kinase activity (**Deniaud et al., 2015**). However, there is no ordered density for this region in neither the previous apo-structure (**Gat et al., 2019**) nor in the current substrate-bound structure (**Figure 1A and B**).

Preferred recognition of a leucine residue at –1 position of the UPF1 consensus motif

Previous results have indicated that SQ motifs preceded by a hydrophobic residue in position –1 are preferentially phosphorylated by SMG1 (**Yamashita et al., 2001**). In our model, the Leu residue at position –1 in the substrate forms a C-H... π -interaction with SMG1 Phe2215 and is further stabilized by hydrophobic interactions with SMG1 Pro2249 and Gly3656. The binding pocket is also restricted by the catalytic loop residues His2337 and Asp2339 (**Figure 4A**). To biochemically characterize the importance of position –1, we assayed a peptide library based on UPF1-LSQ, in which we varied the residue in position –1 to represent all those found in the 20 different SQ motifs of human UPF1. Following phosphorylation of the peptides over time, we could observe that SQ motifs carrying a hydrophobic residue in position –1 were more efficiently phosphorylated. Notably, a Leu

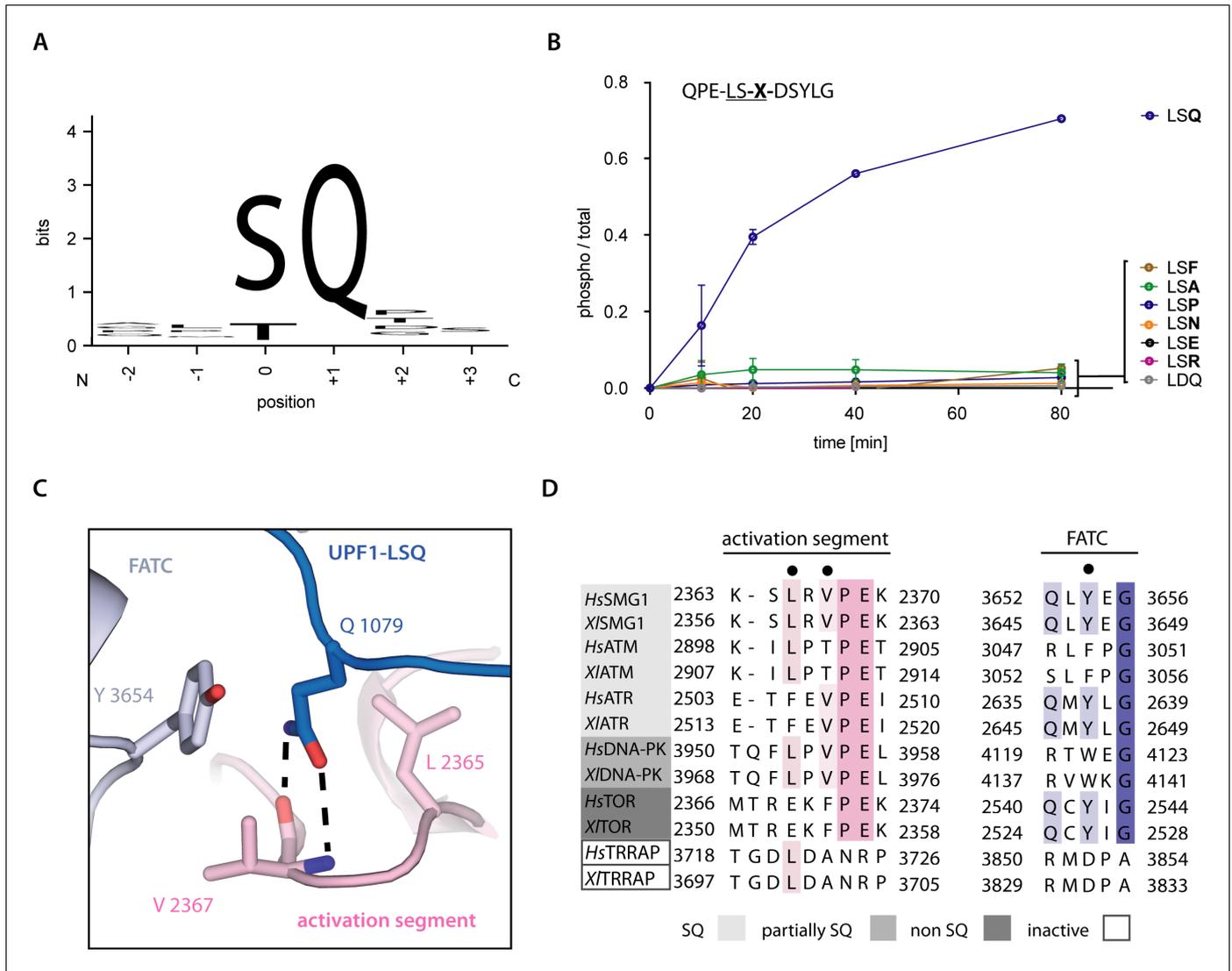


Figure 3. Recognition of position +1 glutamine of UPF1-LSQ. (A) Sequence logo derived from an alignment of all SQ motifs present in human UPF1 with the respective residue positions indicated. The heights of single letters correspond to the observed frequency at that position and the overall height of a stack of letters indicates the level of conservation (Figure 1—figure supplement 1B and Figure 4—figure supplement 2; Crooks et al., 2004). (B) Mass spectrometry-based phosphorylation assay comprising UPF1-LSQ and the indicated position +1 variations. The peptide sequence is indicated in the upper left with the varied position marked as ‘X’. Error bars representing standard deviations calculated from independent experimental triplicates are shown. (C) Zoom-in of the SMG1 active site showing the recognition of UPF1 position +1 glutamine by SMG1 residues located in the activation segment and FATC domain. Residues of interest are shown as sticks. Colors as in Figure 2. (D) Alignment of PIKK sequences from *Homo sapiens* and *Xenopus laevis* with the activation segment and FATC domain sequences shown and colored according to conservation. PIKKs are grouped by phosphorylation site specificity and residues highlighted in subfigure C are indicated by a black dot.

elicited the highest phosphorylation rate (Figure 4B). An end-point measurement experiment using single peptides confirmed these observations (Figure 4—figure supplement 1). We conclude that a Leu at position –1 is optimal for the interaction with SMG1 at the structural level. This is reflected at the biochemical level, whereby decreasing hydrophobicity of the residue at the –1 position negatively affects phosphorylation efficiency.

Interestingly, further analysis of the final time points in the time course phosphorylation experiment showed that the SQ motifs that carry rather hydrophobic residues at the –1 position (and are therefore more efficiently phosphorylated) reside exclusively in the UPF1 C-terminus (Figure 4C, Figure 4—figure supplement 2). To validate our hypothesis on the importance of position –1

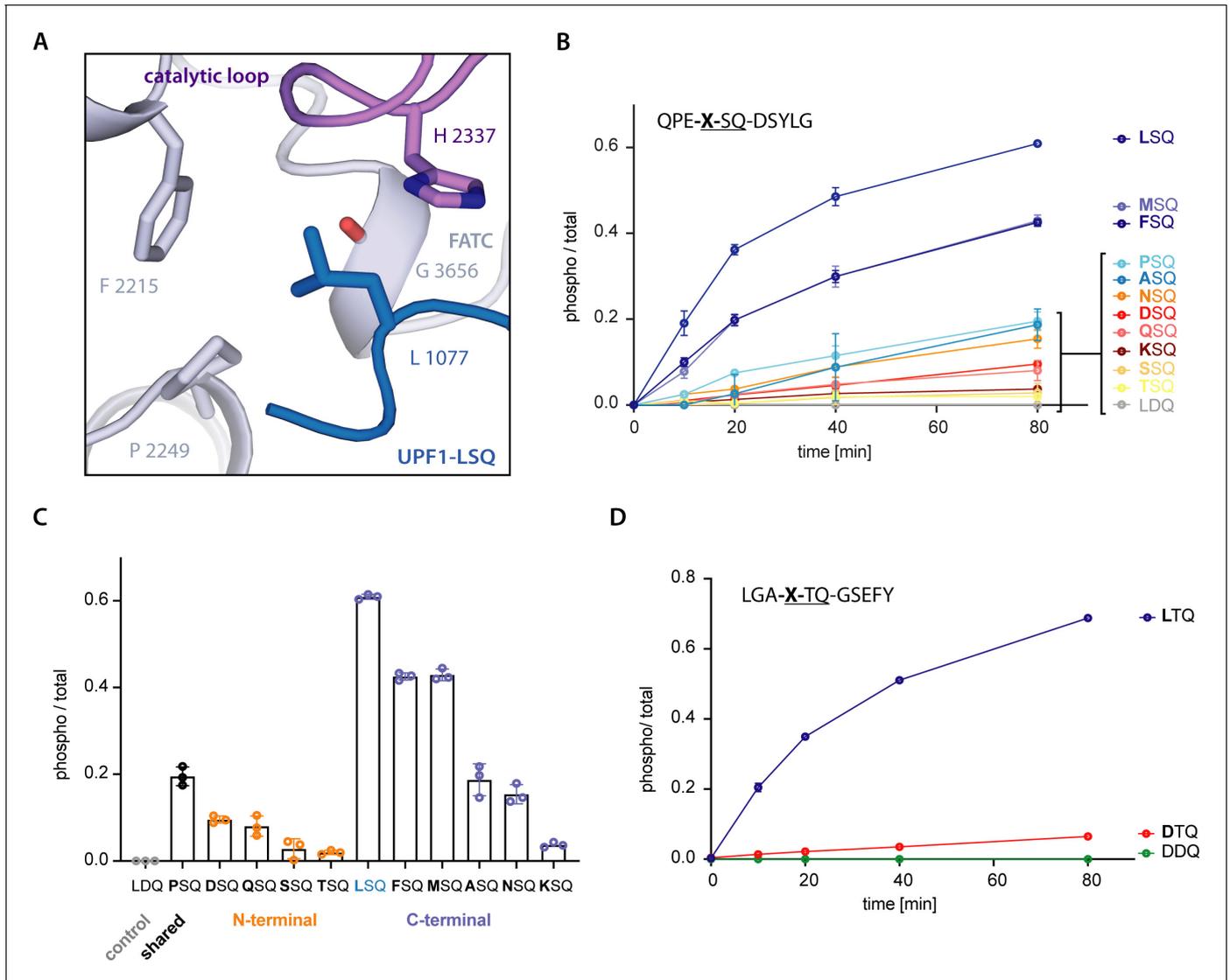


Figure 4. SMG1 preferentially selects for substrates with hydrophobic residues in position -1. (A) Recognition of position -1 residue of UPF1-LSQ by SMG1. Important residues are shown as sticks and colored as in **Figure 2**. (B) Mass spectrometry-based phosphorylation assay with UPF1-LSQ and derivatives varied in position -1. The peptide sequence is indicated in the upper left with the varying position marked as "X". Error bars representing standard deviations calculated from independent experimental triplicates are shown and curves are colored according to hydrophobicity of the position -1 residue (**Eisenberg et al., 1984**). The most hydrophobic peptides are in blue while non-hydrophobics are in red. (C) Final time points of experiment shown in B. Peptides grouped and colored according to the location of the respective position -1 residue in UPF1 N- or C-terminus. Individual data points are shown as circles and error bars representing standard deviations are indicated. (D) Mass spectrometry-based phosphorylation assay with UPF1 N-terminus phosphorylation site 28 and the indicated position -1 variation. The peptide sequence is shown in the upper left with the varied position marked as "X". Error bars represent standard deviations resulting from independent experimental triplicates. A tyrosine residue was added to the C-terminal end of the wildtype sequence to increase absorbance at $\lambda = 280$ nm.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. In vitro phosphorylation of UPF1-LSQ and derivatives.

Figure supplement 2. Alignment of UPF1 N- and C-terminal sequences from *Homo sapiens*, *Bos taurus*, *Canis lupus*, *Mus musculus*, *Gallus gallus*, *Xenopus tropicalis*, *Danio rerio* and *Caenorhabditis elegans*.

hydrophobicity, we turned to a known phosphorylation site in the N-terminus, Thr28. We tested whether SMG1-8-9 phosphorylation activity toward this motif could be enhanced by changing the residue at position -1 from the naturally occurring Asp to a Leu. Indeed, mutating the residue upstream of this SQ motif (Asp27Leu) resulted in a gain-of-function effect in the phosphorylation assay (**Figure 4D**). These findings are in agreement with data for ATM and ATR (**Kim et al., 1999**), although the residues involved in the recognition of UPF1 Leu1077 have diverged, suggesting that the details of -1 recognition will differ in other PIKKs. Finally, we do not observe extensive interactions between SMG1 and the peptide residues preceding or following the LSQ motif in our structure. Consistently, we did not detect a marked effect on phosphorylation in a time course experiment where we changed the residue at position $+2$ of UPF1-LSQ (**Figure 4—figure supplement 1B**).

Conclusions

In this manuscript, we report the first structure of a substrate-bound PIKK active site, thereby revealing the basis for phosphorylation site selection by SMG1 and other PIKK family members. The results elucidate the mechanism of phospho-acceptor recognition, and explain the specificity for Ser-containing substrates with a glutamine downstream residue at position $+1$ and an upstream hydrophobic residue at position -1 (particularly Leu). These findings can be extrapolated to other PIKK members, such as ATM and ATR, and suggest a specific mechanism for PRD function by acting as a pseudosubstrate. Our results provide molecular insights into a key step of the NMD pathway. Whether phosphorylation of full-length UPF1 by SMG1 involves additional elements of recognition and/or additional levels of regulation will be a subject for future studies.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Gene (<i>Homo sapiens</i>)	SMG1	Shigeo Ohno lab	Uniprot Q96Q15	
Gene (<i>Homo sapiens</i>)	SMG8	Shigeo Ohno lab	Uniprot Q8ND04	
Gene (<i>Homo sapiens</i>)	SMG9	Shigeo Ohno lab	Uniprot Q9H0W8	
Cell line (<i>Homo sapiens</i>)	HEK293T	ATCC		
Strain, strain background (<i>Escherichia coli</i>)	BL21 Star (DE3) pRARE	EMBL Heidelberg Core Facility		Electrocompetent cells
Peptide, recombinant protein	UPF1-LSQ (peptide 1078) and derivatives, UPF1-peptide 28 and derivatives	in-house as described in the Materials and methods section		
Chemical compound, drug	AMPPNP	Sigma-Aldrich		
Chemical compound, drug	ATP	Sigma-Aldrich		
Software, algorithm	RELION	DOI: 10.7554/eLife.42166	RELION 3.0	
Software, algorithm	Cryosparc	DOI: 10.1038/nmeth.4169	Cryosparc2	
Software, algorithm	CtfFind	DOI: 10.1016/j.jsb.2015.08.008	CtfFind4.1.9	
Software, algorithm	Cryosparc	DOI: 10.1038/nmeth.4169	Cryosparc2	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Software, algorithm	UCSF Chimera	UCSF, https://www.cgl.ucsf.edu/chimera/		
Software, algorithm	UCSF ChimeraX	UCSF, https://www.rbvi.ucsf.edu/chimerax/		
Software, algorithm	COOT	http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/		
Software, algorithm	PHENIX	https://www.phenix-online.org/	PHENIX 1.17	
Software, algorithm	PyMOL	PyMOL Molecular Graphics System, Schrodinger LLC	PyMOL 2.3.2	https://www.pymol.org/

Protein expression and purification

The SMG1-SMG8-SMG9 complex was expressed and purified as previously described (*Gat et al., 2019*). Briefly, a pool of HEK293T cells (obtained from ATCC) stably expressing full length SMG1 (N-terminally fused to a TwinStrep-tag and a 3C protease cleavage site), SMG8 and SMG9 was established using the *piggybac* method by initially transfecting the cells using polyethylenimine (*Yusa et al., 2011; Li et al., 2013*). The source cells were authenticated by genotyping (Eurofins) and tested negative for mycoplasma contamination (LookOut Mycoplasma PCR Detection Kit, Sigma-Aldrich). For SMG1-SMG8-SMG9 expression, cultures were adjusted to a density of 1×10^6 cells per mL in FreeStyle 293 Expression Medium (Gibco, Thermo Fisher Scientific). The cells were induced by addition of doxycycline and were harvested 48 hr after induction. After lysis by douncing the cells in 1xPBS, 1 mM MgCl₂ and 1 mM DTT supplemented with DNase I, Benzonase and EDTA-free cOmplete Protease Inhibitor Cocktail (Roche) the cleared lysate was applied to a StrepTrap HP column (Sigma-Aldrich) and the complex affinity purified using the N-terminal TwinStrep-tag on SMG1. After washing with 50 column volumes of 1xPBS, 1 mM MgCl₂ and 1 mM DTT the complex was eluted using wash buffer supplemented with 2.5 mM desthiobiotin. SMG1-8-9 was further purified by size-exclusion chromatography using a Superose 6 Increase 10/300 GL column (Sigma-Aldrich) equilibrated with 1xPBS, 1 mM MgCl₂ and 1 mM DTT (Aekta purifier FPLC system, GE Healthcare). Purified SMG1-8-9 was concentrated up to 6 μM and stored in gel filtration buffer. To obtain full-length unphosphorylated human UPF1, the protein was expressed in *Escherichia coli* BL21 STAR (DE3) pRARE fused to a C-terminal 6xHis-tag cleavable with Tobacco etch virus (TEV) protease, as described before (*Chakrabarti et al., 2011; Chakrabarti et al., 2014*). Bacteria were grown at 37°C in TB medium shaking at 180 rpm and induced using IPTG at an OD of 2 for overnight expression at 18°C. Harvested bacteria (6000 rpm, 10 min) were lysed by sonication in lysis buffer (50 mM Tris-Cl pH 7.5, 500 mM NaCl, 10 mM Imidazole, 1 mM β-mercaptoethanol, 10% (v/v) glycerol, 2 mM MgCl₂ and 0.2% (v/v) NP-40) supplemented with DNase I and EDTA-free cOmplete Protease Inhibitor Cocktail (Roche). The lysate was cleared by centrifugation (25,000 rpm, 30 min), filtered and combined with TALON resin (Takara) equilibrated with lysis buffer for gravity-flow affinity purification. After washing with 70 column volumes of lysis buffer, the protein was eluted with lysis buffer supplemented with 300 mM imidazole pH 7.5 and the eluate was combined with His-tagged TEV protease and dialyzed overnight against 20 mM HEPES pH 7.5, 85 mM KCl, 1 mM MgCl₂, 10% (v/v) glycerol and 2 mM DTT. The dialyzed sample was passed over another TALON column by gravity-flow, in order to separate cleaved protein from the cleaved-off His-tag, the His-tagged TEV protease and uncleaved UPF1 protein. The flow-through of this column contained cleaved UPF1 and was loaded on a HiTrap Heparin HP column (GE Healthcare). Following binding and washing with Heparin buffer A (as for dialysis), UPF1 was eluted by a gradient increasing salt concentration from 85 mM to 500 mM over 50 column volumes (Aekta prime FPLC system, GE Healthcare). The peak corresponding to full-length UPF1 was pooled and concentrated before a final sizing step using a Superdex 200 Increase 10/300 GL column (Sigma-Aldrich) equilibrated with Heparin buffer A (Aekta purifier FPLC system, GE Healthcare). Purified full-length UPF1 was pooled and concentrated up to 30 μM using an Amicon Ultra Centrifugal Filter (50 kDa MWCO, Merck). All described protein

purification steps were carried out at 4°C and all purified proteins were flash frozen in size-exclusion buffer using liquid nitrogen and stored at –80°C until further usage.

Cryo-EM sample preparation and data collection

A sample of 0.5 μM (final concentration) purified SMG1-SMG8-SMG9 was mixed with 0.5 mM of UPF1-LSQ, 1 mM AMPPNP, 2 mM MgCl₂, 2 mM DTT and 0.04% (v/v) n-octyl-beta-D-glucoside in 1xPBS and incubated for 30 min on ice. The UPF1-LSQ peptide (sequence: QPELSQDSYLG) was synthesized in-house as described for the mass spectrometry-based phosphorylation assay. A 4 μL sample was applied to a glow-discharged Quantifoil R1.2/1.3, Cu 200 mesh grid and incubated for 30 s at 4°C and approximately 100% humidity. Grids were subsequently plunge frozen directly after blotting using a liquid ethane/propane (37% ethane, temperature range when plunging: –170°C to –180°C) mixture and a ThermoFisher FEI Vitrobot IV set to a blot time of 3.5 s and a blot force of 4. Cryo-EM data were collected using a ThermoFisher FEI Titan Krios microscope operated at 300 kV equipped with a post-column GIF (energy width 20 eV) and a Gatan K3 camera operated in counting mode, the SerialEM software suite, and a beam-tilt based multi-shot acquisition scheme. Movies were recorded at a nominal magnification of 81,000x corresponding to a pixel size of 1.094 Å at the specimen level. The sample was imaged with a total exposure of 68.75 e⁻/Å² evenly spread over 5.5 s and 79 frames. The target defocus during data collection ranged between –0.8 and –2.9 μm.

Cryo-EM data processing

Data processing was carried out using RELION 3.0 (Zivanov *et al.*, 2018) unless stated otherwise. Beam-induced sample motions were corrected and dose-weighting was carried out using the RELION implementation of MotionCor2 (Zheng *et al.*, 2017). Particles were picked using Gautomatch (<https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/>) and CTF estimation was done using the RELION wrapper for CtfFind4.1 (Rohou and Grigorieff, 2015). After extraction (box size: 320 pix, 1.094 Å/pix) and downsampling (box size: 80 pix, 4.376 Å/pix), 4,368,586 particles were submitted to two rounds of reference-free 2D classification. A subset of the cleaned candidate particles was used to generate an initial model using CryoSPARC v2 (Punjani *et al.*, 2017). 1,524,355 selected particles were 3D classified before re-extracting 886,714 particles with original sampling followed by two additional rounds of 3D classification resulting in 481,754 final particles. All classification steps were carried out with the total amount of particles being distributed over multiple batches. After 3D auto-refinement, sharpening (b-factor = –119.5) and Ctf refinement in RELION 3.0, the final refined map (3D auto-refinement) was again submitted to RELIONs' post-processing routine for automatic B-factor weighting and high-resolution noise substitution (b-factor = –102.6). The final reconstruction (EMD-11063) reached an overall resolution of 2.9 Å with local resolution ranging from 2.8 Å to 4.5 Å as estimated by RELION 3.0.

Model building and refinement

The reconstructed density was interpreted using COOT (version 1.0) and our previously published model of SMG1-8-9 (PDB: 6SYT) (Emsley *et al.*, 2010). Model building was iteratively interrupted by real-space refinements using PHENIX (version 1.17) (Adams *et al.*, 2010; Liebschner *et al.*, 2019). Statistics assessing the quality of the final model (PDB ID 6Z3R) were generated using Molprobity (Chen *et al.*, 2010; Supplementary file 1). FSC curves were calculated using PHENIX and the 3D FSC online application (Tan *et al.*, 2017). Images of the calculated density and the built model were prepared using UCSF Chimera (Pettersen *et al.*, 2004), UCSF ChimeraX (Goddard *et al.*, 2018) and PyMOL (version 2.3.2).

Radioactive in vitro kinase assay

In vitro kinase assays were essentially carried out as before (Gat *et al.*, 2019). 1 μM of full-length UPF1 was mixed with 50 nM SMG1-8-9, 10mM MgCl₂ and 2mM DTT in 1xPBS. The reaction was started by adding 0.5mM ATP and 0.06 μM of γ-³²P-labeled ATP. The reaction was incubated at 30°C and samples were taken at different time points to follow phosphorylation over time. The samples were immediately quenched by adding SDS-containing sample buffer and initially analyzed by SDS gel electrophoresis followed by Coomassie-staining. Phosphoproteins were subsequently detected using autoradiography and a Typhoon FLA7000 imager (GE Healthcare).

Mass spectrometry-based in vitro kinase assay

All peptides were synthesized in-house using solid-phase peptide synthesis and the quality of the product was assessed by electrospray ionization mass spectrometry (ESI MS). For the purpose of this study, peptides were dissolved in 1xPBS supplemented with 500 mM HEPES pH 7.4. Two types of experiments were carried out. Firstly, several peptides (typically comprising a library) and a control were mixed, and their individual phosphorylation ratios were determined at several time points (0, 10, 20, 40 and 80 min). Secondly, one end-point measurement experiment was carried out. In this setup, a single peptide was mixed with a control and the relative phosphorylation ratio was determined at a single, final time point (80 min). This type of experiment was used to assess whether effects on phosphorylation ratios observed in time course assays are caused by competition between several peptides for SMG1 binding. In both cases, 0.5 μ M of kinase complex was combined with 0.1 mM of each peptide, 0.5 mM ATP, 20 mM MgCl₂ and 2 mM DTT in 1xPBS. The reaction was started by addition of kinase and incubated at 30°C. Samples were taken at desired time points and immediately quenched after collection by adding EDTA to a final concentration of 50 mM on ice.

In order to remove kinase complex and transfer the peptides into a compatible buffer, we made use of home-made StageTips (*Rappsilber et al., 2007*). Poly(styrene-divinylbenzene)copolymer (SDB-XC) was washed with methanol by centrifugation before being washed again with buffer B (0.1% (v/v) formic acid, 80% (v/v) acetonitrile). Buffer A (0.1% (v/v) formic acid) was used for equilibration of the SDB-XC material. Following sorbent equilibration, the sample was applied and the tips were washed using buffer A. Finally, the sample was eluted using buffer B. Using an Agilent 1290 HPLC, typically about 5 μ L of the sample in 70% (v/v) acetonitrile and 0.05% (v/v) trifluoroacetic acid were flow-injected (250 μ L/min) into a Bruker maXis II ETD mass spectrometer for ESI MS time-of-flight analysis. Peptides were ionized at a capillary voltage of 4500 V and an end plate offset of 500 V. Full scan MS spectra (200–1600 m/z) were acquired at a spectra rate of 1 Hz and a collision energy of 10 eV. All experiments were carried out as independent experimental triplicates. Raw data files were processed using Bruker Compass DataAnalysis software. The m/z spectra were deconvoluted by maximum entropy with an instrument resolving power of 10,000. The ¹²C peaks corresponding to individual peptides were identified in the resulting neutral spectra and integrated, both for masses accounting for unphosphorylated and phosphorylated peptides. To calculate a relative phosphorylation ratio, the area for phosphorylated peptide was divided by the sum of phosphorylated and unphosphorylated peptide. All time points were normalized to time point 0. Means of independent experimental triplicates and error bars indicating standard deviations were visualized using Prism (GraphPad).

Acknowledgements

We thank Daniel Bollschweiler and Tillman Schäfer at the MPIB cryo-EM facility for help with EM data collection and Elisabeth Weyher and Stefan Pettera at MPIB biochemistry core facility at MPIB for conducting mass spectrometry and synthesis of peptides used in this study, respectively. We thank Christian Benda and J Rajan Prabu for maintenance and development of computational infrastructure and Daniela Wartini for assistance in mammalian tissue culture. We are grateful to Courtney Long and members of the group for input and discussion on the manuscript. This study was supported by funding from the Max Planck Gesellschaft, the European Commission (ERC Advanced Investigator Grant EXORICO), and the German Research Foundation (DFG SFB1035, GRK1721, SFB/TRR 237) to EC and a Boehringer Ingelheim Fonds PhD fellowship to LL.

Additional information

Funding

Funder	Grant reference number	Author
Boehringer Ingelheim Fonds	PhD fellowship	Lukas M Langer
Max-Planck-Gesellschaft		Elena Conti
European Commission	ERC Advanced Investigator Grant EXORICO	Elena Conti

Deutsche Forschungsgemeinschaft	SFB1035	Elena Conti
Deutsche Forschungsgemeinschaft	GRK1721	Elena Conti
Deutsche Forschungsgemeinschaft	SFB/TRR 237	Elena Conti

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Lukas M Langer, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Yair Gat, Conceptualization, Validation, Investigation, Visualization, Writing - review and editing; Fabien Bonneau, Formal analysis, Validation, Methodology, Writing - review and editing; Elena Conti, Conceptualization, Resources, Supervision, Funding acquisition, Writing - original draft, Project administration, Writing - review and editing

Author ORCIDs

Lukas M Langer  <https://orcid.org/0000-0002-9977-2427>

Yair Gat  <https://orcid.org/0000-0002-2338-9384>

Fabien Bonneau  <http://orcid.org/0000-0001-8787-7662>

Elena Conti  <https://orcid.org/0000-0003-1254-5588>

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.57127.sa1>

Author response <https://doi.org/10.7554/eLife.57127.sa2>

Additional files

Supplementary files

- Supplementary file 1. Cryo-EM data collection, refinement and validation statistics.
- Transparent reporting form

Data availability

EM data have been deposited in EMDB under the accession code EMD-11063. The model has been deposited in PDB under the accession 6Z3R.

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Langer LM, Gat Y, Conti E	2020	Structure of SMG1-8-9 kinase complex bound to UPF1-LSQ	https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-11063	Electron Microscopy Data Bank, EMD-11063
Langer LM, Gat Y, Conti E	2020	Structure of SMG1-8-9 kinase complex bound to UPF1-LSQ	https://www.rcsb.org/structure/6Z3R	RCSB Protein Data Bank, 6Z3R

References

- Adams PD**, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography* **66**:213–221. DOI: <https://doi.org/10.1107/S0907444909052925>
- Bannister AJ**, Gottlieb TM, Kouzarides T, Jackson SP. 1993. c-Jun is phosphorylated by the DNA-dependent protein kinase in vitro; definition of the minimal kinase recognition motif. *Nucleic Acids Research* **21**:1289–1295. DOI: <https://doi.org/10.1093/nar/21.5.1289>, PMID: 8464713

- Bao ZQ**, Jacobsen DM, Young MA. 2011. Briefly bound to activate: transient binding of a second catalytic magnesium activates the structure and dynamics of CDK2 kinase for catalysis. *Structure* **19**:675–690. DOI: <https://doi.org/10.1016/j.str.2011.02.016>, PMID: 21565702
- Baretić D**, Williams RL. 2014. PIKKs—the solenoid nest where partners and kinases meet. *Current Opinion in Structural Biology* **29**:134–142. DOI: <https://doi.org/10.1016/j.sbi.2014.11.003>, PMID: 25460276
- Blackford AN**, Jackson SP. 2017. ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. *Molecular Cell* **66**:801–817. DOI: <https://doi.org/10.1016/j.molcel.2017.05.015>, PMID: 28622525
- Bosotti R**, Isacchi A, Sonnhammer EL. 2000. FAT: a novel domain in PIK-related kinases. *Trends in Biochemical Sciences* **25**:225–227. DOI: [https://doi.org/10.1016/S0968-0004\(00\)01563-2](https://doi.org/10.1016/S0968-0004(00)01563-2), PMID: 10782091
- Brown EJ**, Beal PA, Keith CT, Chen J, Bum Shin T, Schreiber SL. 1995. Control of p70 S6 kinase by kinase activity of FRAP in vivo. *Nature* **377**:441–446. DOI: <https://doi.org/10.1038/377441a0>
- Chakrabarti S**, Jayachandran U, Bonneau F, Fiorini F, Basquin C, Domcke S, Le Hir H, Conti E. 2011. Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Molecular Cell* **41**:693–703. DOI: <https://doi.org/10.1016/j.molcel.2011.02.010>, PMID: 21419344
- Chakrabarti S**, Bonneau F, Schüssler S, Eppinger E, Conti E. 2014. Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6. *Nucleic Acids Research* **42**:9447–9460. DOI: <https://doi.org/10.1093/nar/gku578>, PMID: 25013172
- Chen VB**, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* **66**:12–21. DOI: <https://doi.org/10.1107/S0907444909042073>, PMID: 20057044
- Crooks GE**, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Research* **14**:1188–1190. DOI: <https://doi.org/10.1101/gr.849004>, PMID: 15173120
- Deniaud A**, Karuppusamy M, Bock T, Masiulis S, Huard K, Garzoni F, Kerschgens K, Hentze MW, Kulozik AE, Beck M, Neu-Yilik G, Schaffitzel C. 2015. A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. *Nucleic Acids Research* **43**:7600–7611. DOI: <https://doi.org/10.1093/nar/gkv668>
- Denning G**, Jamieson L, Maquat LE, Thompson EA, Fields AP. 2001. Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein. *The Journal of Biological Chemistry* **276**:22709–22714. DOI: <https://doi.org/10.1074/jbc.C100144200>, PMID: 11331269
- Durand S**, Franks TM, Lykke-Andersen J. 2016. Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. *Nature Communications* **7**:12434. DOI: <https://doi.org/10.1038/ncomms12434>, PMID: 27511142
- Eisenberg D**, Schwarz E, Komaromy M, Wall R. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *Journal of Molecular Biology* **179**:125–142. DOI: [https://doi.org/10.1016/0022-2836\(84\)90309-7](https://doi.org/10.1016/0022-2836(84)90309-7), PMID: 6502707
- Eliás-Villalobos A**, Fort P, Helmlinger D. 2019. New insights into the evolutionary conservation of the sole PIKK pseudokinase Tra1/TRRAP. *Biochemical Society Transactions* **47**:1597–1608. DOI: <https://doi.org/10.1042/BST20180496>, PMID: 31769470
- Emsley P**, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of coot. *Acta Crystallographica Section D, Biological Crystallography* **66**:486–501. DOI: <https://doi.org/10.1107/S0907444910007493>, PMID: 20383002
- Gat Y**, Schuller JM, Lingaraju M, Weyher E, Bonneau F, Strauss M, Murray PJ, Conti E. 2019. InsP6 binding to PIKK kinases revealed by the cryo-EM structure of an SMG1-SMG8-SMG9 complex. *Nature Structural & Molecular Biology* **26**:1089–1093. DOI: <https://doi.org/10.1038/s41594-019-0342-7>, PMID: 31792449
- Goddard TD**, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Science* **27**:14–25. DOI: <https://doi.org/10.1002/pro.3235>, PMID: 28710774
- Hsu PP**, Kang SA, Rameseder J, Zhang Y, Ottina KA, Lim D, Peterson TR, Choi Y, Gray NS, Yaffe MB, Marto JA, Sabatini DM. 2011. The mTOR-Regulated Phosphoproteome Reveals a Mechanism of mTORC1-Mediated Inhibition of Growth Factor Signaling. *Science* **332**:1317–1322. DOI: <https://doi.org/10.1126/science.1199498>
- Imseong S**, Aylett CH, Maier T. 2018. Architecture and activation of phosphatidylinositol 3-kinase related kinases. *Current Opinion in Structural Biology* **49**:177–189. DOI: <https://doi.org/10.1016/j.sbi.2018.03.010>, PMID: 29625383
- Jansma M**, Linke-Winnebeck C, Eustermann S, Lammens K, Kostrewa D, Stakyte K, Litz C, Kessler B, Hopfner KP. 2020. Near-Complete structure and model of Tel1ATM from chaetomium thermophilum reveals a robust autoinhibited ATP state. *Structure* **28**:83–95. DOI: <https://doi.org/10.1016/j.str.2019.10.013>, PMID: 31740028
- Karousis ED**, Mühlemann O. 2019. Nonsense-Mediated mRNA decay begins where translation ends. *Cold Spring Harbor Perspectives in Biology* **11**:a032862. DOI: <https://doi.org/10.1101/cshperspect.a032862>, PMID: 29891560
- Kim ST**, Lim DS, Canman CE, Kastan MB. 1999. Substrate specificities and identification of putative substrates of ATM kinase family members. *Journal of Biological Chemistry* **274**:37538–37543. DOI: <https://doi.org/10.1074/jbc.274.53.37538>, PMID: 10608806
- Kurosaki T**, Maquat LE. 2016. Nonsense-mediated mRNA decay in humans at a glance. *Journal of Cell Science* **129**:461–467. DOI: <https://doi.org/10.1242/jcs.181008>, PMID: 26787741
- Lempäinen H**, Halazonetis TD. 2009. Emerging common themes in regulation of PIKKs and PI3Ks. *The EMBO Journal* **28**:3067–3073. DOI: <https://doi.org/10.1038/emboj.2009.281>, PMID: 19779456

- Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, Staber J, Wheelan SJ, Joung JK, McCray PB, Bushman FD, Sinn PL, Craig NL. 2013. piggyBac transposase tools for genome engineering. *PNAS* **110**:E2279–E2287. DOI: <https://doi.org/10.1073/pnas.1305987110>
- Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung L-W, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, et al. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D Structural Biology* **75**:861–877. DOI: <https://doi.org/10.1107/S2059798319011471>
- Madhusudan , Akamine P, Xuong NH, Taylor SS. 2002. Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase. *Nature Structural Biology* **9**:273–277. DOI: <https://doi.org/10.1038/nsb780>, PMID: 11896404
- O'Neill T, Dwyer AJ, Ziv Y, Chan DW, Lees-Miller SP, Abraham RH, Lai JH, Hill D, Shiloh Y, Cantley LC, Rathbun GA. 2000. Utilization of Oriented Peptide Libraries to Identify Substrate Motifs Selected by ATM. *Journal of Biological Chemistry* **275**:22719–22727. DOI: <https://doi.org/10.1074/jbc.M001002200>
- Ohnishi T, Yamashita A, Kashima I, Schell T, Anders KR, Grimson A, Hachiya T, Hentze MW, Anderson P, Ohno S. 2003. Phosphorylation of hUPF1 Induces Formation of mRNA Surveillance Complexes Containing hSMG-5 and hSMG-7. *Molecular Cell* **12**:1187–1200. DOI: [https://doi.org/10.1016/S1097-2765\(03\)00443-X](https://doi.org/10.1016/S1097-2765(03)00443-X)
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**:1605–1612. DOI: <https://doi.org/10.1002/jcc.20084>, PMID: 15264254
- Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**:290–296. DOI: <https://doi.org/10.1038/nmeth.4169>, PMID: 28165473
- Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* **2**:1896–1906. DOI: <https://doi.org/10.1038/nprot.2007.261>, PMID: 17703201
- Rohou A, Grigorieff N. 2015. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology* **192**:216–221. DOI: <https://doi.org/10.1016/j.jsb.2015.08.008>, PMID: 26278980
- Rosenthal PB, Henderson R. 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology* **333**:721–745. DOI: <https://doi.org/10.1016/j.jmb.2003.07.013>, PMID: 14568533
- Saxton RA, Sabatini DM. 2017. mTOR signaling in growth, metabolism, and disease. *Cell* **168**:960–976. DOI: <https://doi.org/10.1016/j.cell.2017.02.004>, PMID: 28283069
- Shimobayashi M, Hall MN. 2014. Making new contacts: the mTOR network in metabolism and signalling crosstalk. *Nature Reviews Molecular Cell Biology* **15**:155–162. DOI: <https://doi.org/10.1038/nrm3757>, PMID: 24556838
- Tan YZ, Baldwin PR, Davis JH, Williamson JR, Potter CS, Carragher B, Lyumkis D. 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods* **14**:793–796. DOI: <https://doi.org/10.1038/nmeth.4347>, PMID: 28671674
- Wang Z, Cole PA. 2014. Catalytic mechanisms and regulation of protein kinases. *Methods in Enzymology* **548**:1–21. DOI: <https://doi.org/10.1016/B978-0-12-397918-6.00001-X>, PMID: 25399640
- Yamashita A, Ohnishi T, Kashima I, Taya Y, Ohno S. 2001. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes & Development* **15**:2215–2228. DOI: <https://doi.org/10.1101/gad.913001>
- Yamashita A. 2013. Role of SMG-1-mediated Upf1 phosphorylation in mammalian nonsense-mediated mRNA decay. *Genes to Cells* **18**:161–175. DOI: <https://doi.org/10.1111/gtc.12033>, PMID: 23356578
- Yang H, Rudge DG, Koos JD, Vaidialingam B, Yang HJ, Pavletich NP. 2013. mTOR kinase structure, mechanism and regulation. *Nature* **497**:217–223. DOI: <https://doi.org/10.1038/nature12122>
- Yates LA, Williams RM, Hailemariam S, Ayala R, Burgers P, Zhang X. 2020. Cryo-EM structure of Nucleotide-Bound Tel1ATM unravels the molecular basis of inhibition and structural rationale for Disease-Associated mutations. *Structure* **28**:96–104. DOI: <https://doi.org/10.1016/j.str.2019.10.012>
- Yusa K, Zhou L, Li MA, Bradley A, Craig NL. 2011. A hyperactive piggyBac transposase for mammalian applications. *PNAS* **108**:1531–1536. DOI: <https://doi.org/10.1073/pnas.1008322108>, PMID: 21205896
- Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **14**:331–332. DOI: <https://doi.org/10.1038/nmeth.4193>, PMID: 28250466
- Zhu L, Li L, Qi Y, Yu Z, Xu Y. 2019. Cryo-EM structure of SMG1-SMG8-SMG9 complex. *Cell Research* **29**:1027–1034. DOI: <https://doi.org/10.1038/s41422-019-0255-3>, PMID: 31729466
- Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**:e42166. DOI: <https://doi.org/10.7554/eLife.42166>, PMID: 30412051

Figure supplements

Structure of substrate-bound SMG1-8-9 kinase complex reveals molecular basis for phosphorylation specificity

Lukas M Langer *et al*

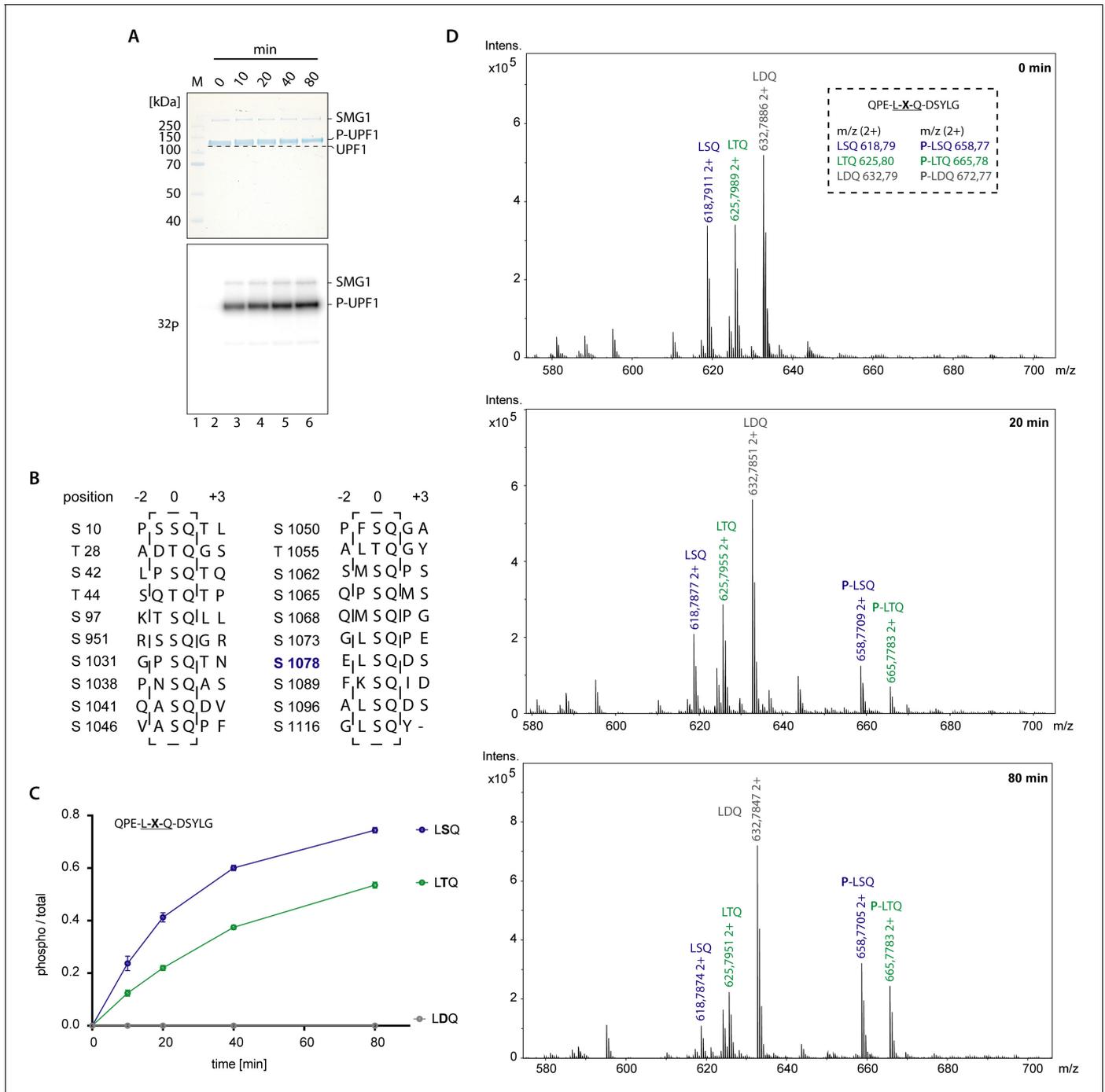


Figure 1—figure supplement 1. SMG1-8-9 activity and UPF1 SQ motifs. (A) Radioactive phosphorylation assay using SMG1-8-9 and full-length UPF1. Coomassie-stained SDS-PAGE showing a change in migration behavior for UPF1 over time as phosphorylation proceeds. The corresponding radioactive signal is shown in the lower panel indicating an increase of UPF1 phosphorylation over time. (B) Alignment showing all SQ motifs present in UPF1 N- and C-terminus including positions -2 to +3. Note the high variance amongst position -1 residues. (C) Mass spectrometry-based phosphorylation assay with UPF1-LSQ and the indicated position 0 variations. The peptide sequence is indicated with the varying position marked as 'X'. Error bars representing standard deviations calculated from independent experimental triplicates are shown. Phosphorylation was abolished when the phospho-acceptor residue was changed from Ser to Asp. This shows Ser1081 was not recognized for phosphorylation and confirms specificity toward the SQ motif. In addition, phosphorylation was decreased when Ser was changed to Thr, consistent with previous data for ATM, ATR and DNA-PK (Kim et al., 1999; O'Neill et al., 2000). (D). M/z spectra for representative single measurements at indicated time points of the experiment shown in Figure 1—figure supplement 1 continued on next page

Figure 1—figure supplement 1 continued

in C. The inset lists the expected sizes for the three peptides used in this experiment. Peaks corresponding to unphosphorylated and phosphorylated peptides are highlighted. Note the appearance and increase of intensity of peaks corresponding to phosphorylated peptides over time.

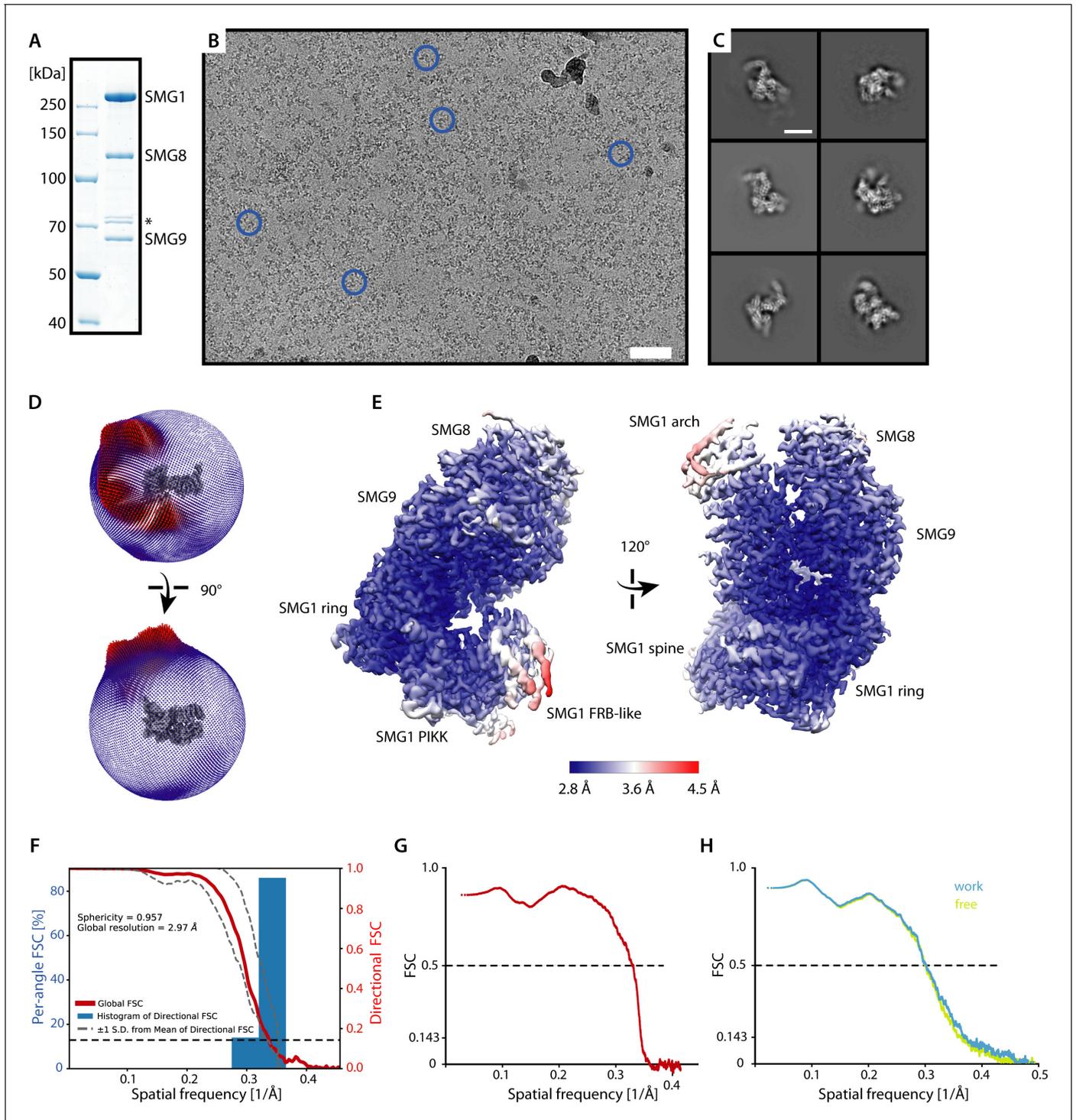


Figure 1—figure supplement 2. Cryo-EM analysis of SMG1-8-9 bound to UPF1-LSQ. (A) Coomassie-stained SDS-PAGE showing purified SMG1-8-9. The asterisk indicates contaminants. (B) Representative micrograph of the collected data set with some SMG1-8-9 particles indicated by blue circles. Scale bar ≈ 500 Å. (C) Representative 2D averages of picked particles. Scale bar ≈ 100 Å. (D) Spherical angular distribution of particles contributing to the final reconstruction with larger red rods indicating more prominent particle views and smaller blue rods indicating rarer particle views. (E) Map of SMG1-8-9 colored according to estimated local resolution shown in two different views as in **Figure 1B**. Large parts of the complex including the kinase domain and the bound UPF1 peptide are resolved to around 3 Å. Important features of the map are indicated. (F) Three-dimensional FSC plot and further analysis of orientation bias. The red line represents the estimated global masked half-map FSC curve indicating a nominal overall resolution of **Figure 1—figure supplement 2 continued on next page**

Figure 1—figure supplement 2 continued

2.97 Å according to the gold standard FSC cut off of 0.143 (**Rosenthal and Henderson, 2003**). The spread of directional resolution values is defined as $\pm 1\sigma$ (shown as dashed grey lines). Overall isotropy of the map is confirmed by a sphericity of 0.957 (out of 1) (**Tan et al., 2017**). (G) Model vs. map FSC plot for the real spaced refined model. (H) Model vs. map FSC plots for half map 1 ("work") used for real space refinement after displacing atoms of the final model ($\sigma = 0.5$ Å) and half map 2 not used for refinement ("free"). The good agreement of the two curves indicates that no substantial overfitting took place.

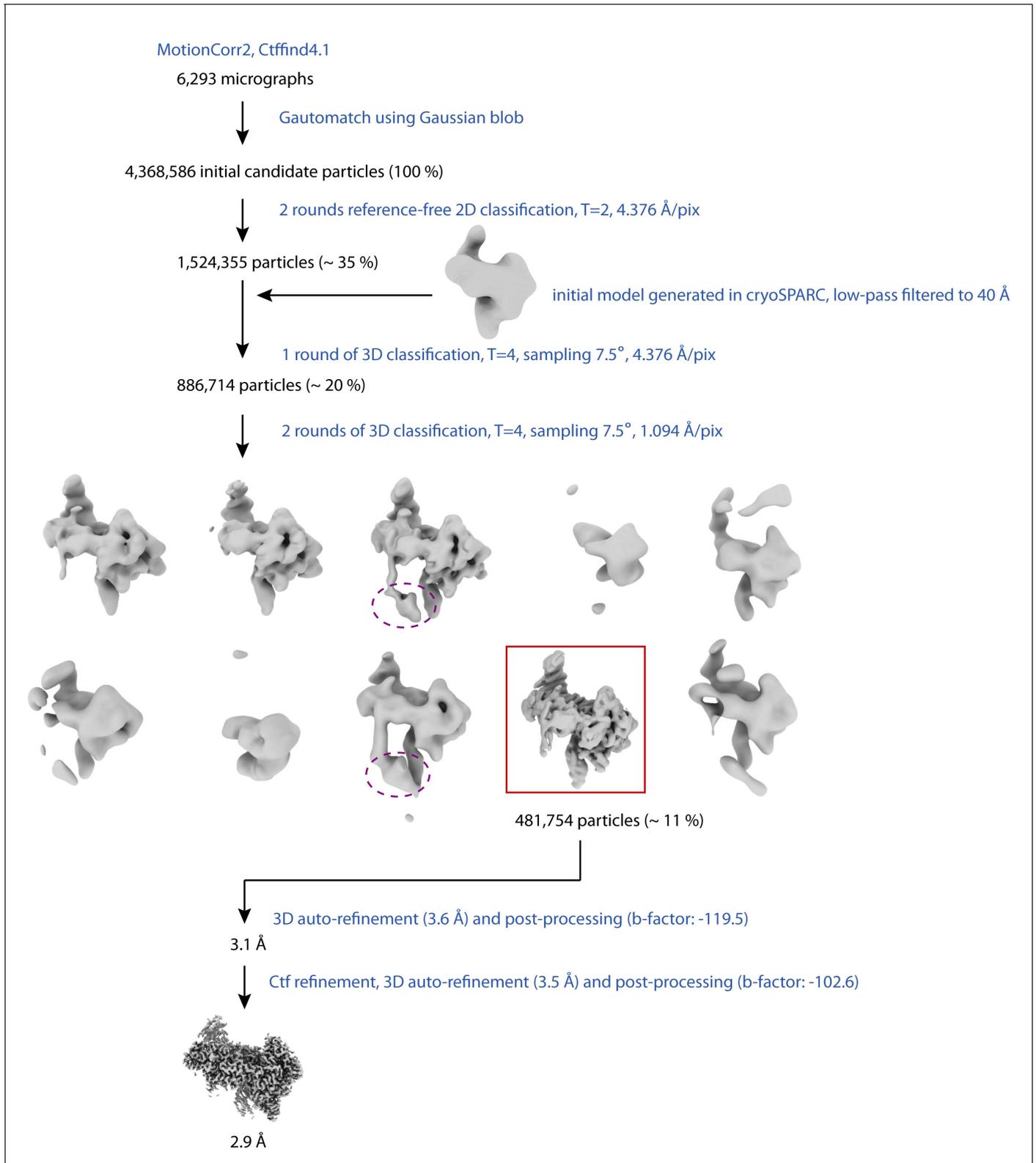


Figure 1—figure supplement 3. Cryo-EM data processing scheme. Processing steps are indicated in blue; particle numbers of classes used for downstream processing steps (in brackets: percentage with respect to initial candidate particles) and resolutions are in black. The class selected for the 3D refinement after the last step of 3D classification is indicated by a red rectangle. Density for the unmodelled, C-terminal part of SMG8 appearing in Figure 1—figure supplement 3 continued on next page

Figure 1—figure supplement 3 continued

two other 3D classes not used for the final reconstruction is highlighted in magenta. This part of SMG8 has been suggested to contribute to kinase regulation (Zhu et al., 2019).

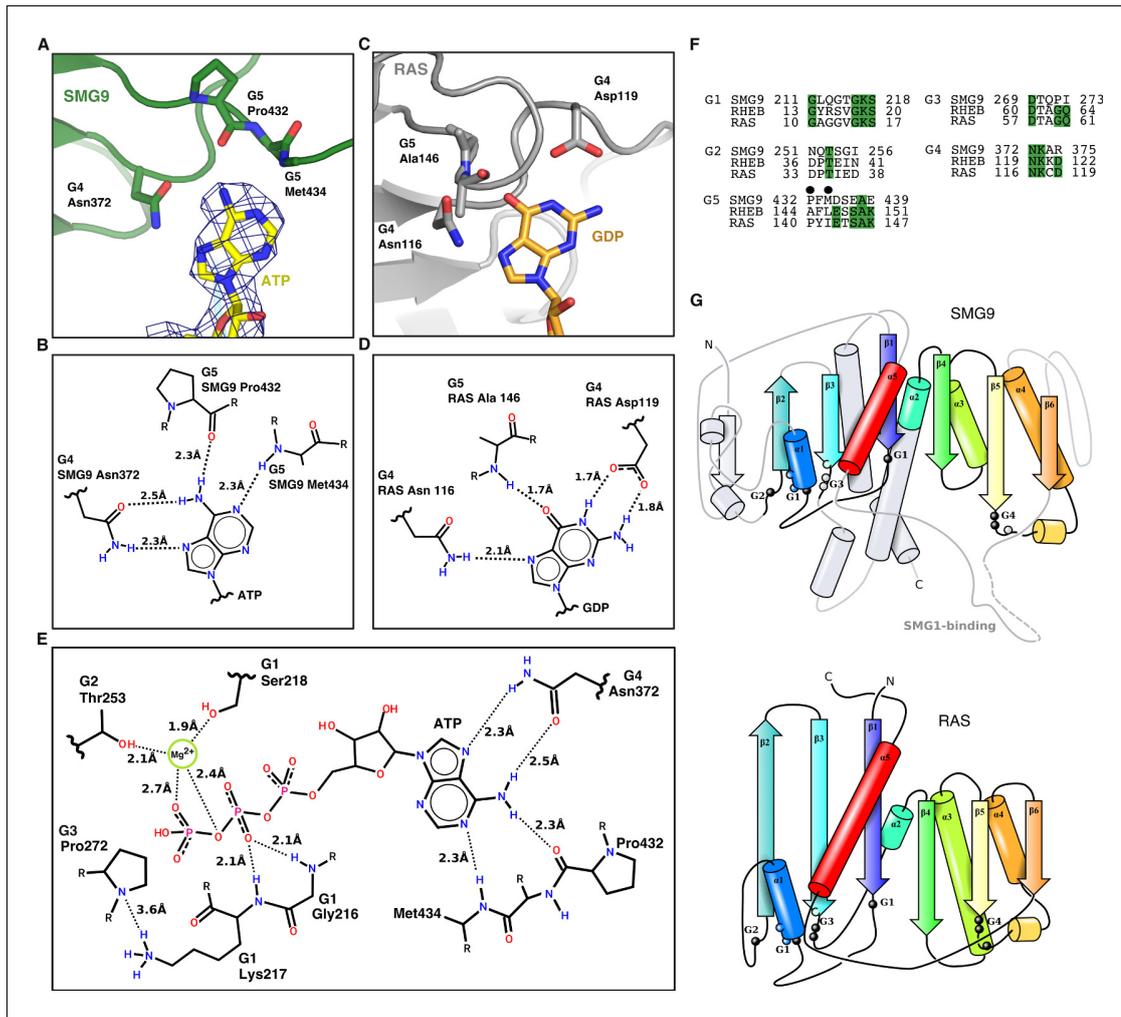


Figure 1—figure supplement 4. SMG9 is a G-fold containing protein binding ATP and exhibits distinct differences to the *bona fide* GTPase RAS. (A) Detailed view of ATP purine ring recognition by SMG9 with important residues highlighted and reconstructed density indicated. (B) Two-dimensional sketch of adenine base recognition by SMG9 shown in panel A with G motifs, key residues and distances indicated. (C) Detailed view of GTP recognition by the GTPase RAS with important residues highlighted (PDB ID 1ZVQ). (D) Two-dimensional sketch of guanine base recognition shown in panel C by RAS with G motifs, key residues and distances indicated for comparison with panel B (based on PDB ID 1ZVQ). (E) Two-dimensional sketch of the overall recognition of ATP by SMG9. The G1-G3 motifs of SMG9 coordinate the phosphate moieties of ATP, similarly to the corresponding motifs of canonical GTPases, but with an important exception, in that they lack the typical residues crucial for catalysis. Another difference with canonical G-fold is that the G4 and G5 motifs responsible for the recognition of the base have rearranged to preferentially bind an adenine base rather than a guanine. (F) Sequence alignment of G motifs of RHEB and RAS GTPases with SMG9. Residues highlighted in A are indicated by a black dot. In the G4 motif, the aspartic acid that specifically engages the guanine moiety in GTPases has diverged to an arginine residue. Instead, Asn372 in the SMG9 G4 motif has shifted so that it engages the side-chain carbonyl group to interact with the adenine amino group (compare panels A and B with C and D). In the G5 motif, the SAK consensus sequence in RAS and RHEB has diverged in SMG9, with the main-chain carbonyl and amine groups of the upstream residues, Pro432 and Met434, forming adenine-specific interactions (compare panels A and B with C and D). (G) Three-dimensional sketches of SMG9 (top) and RAS/RHEB (bottom) highlighting similarities in their overall topology. Positions of the G motifs are highlighted.

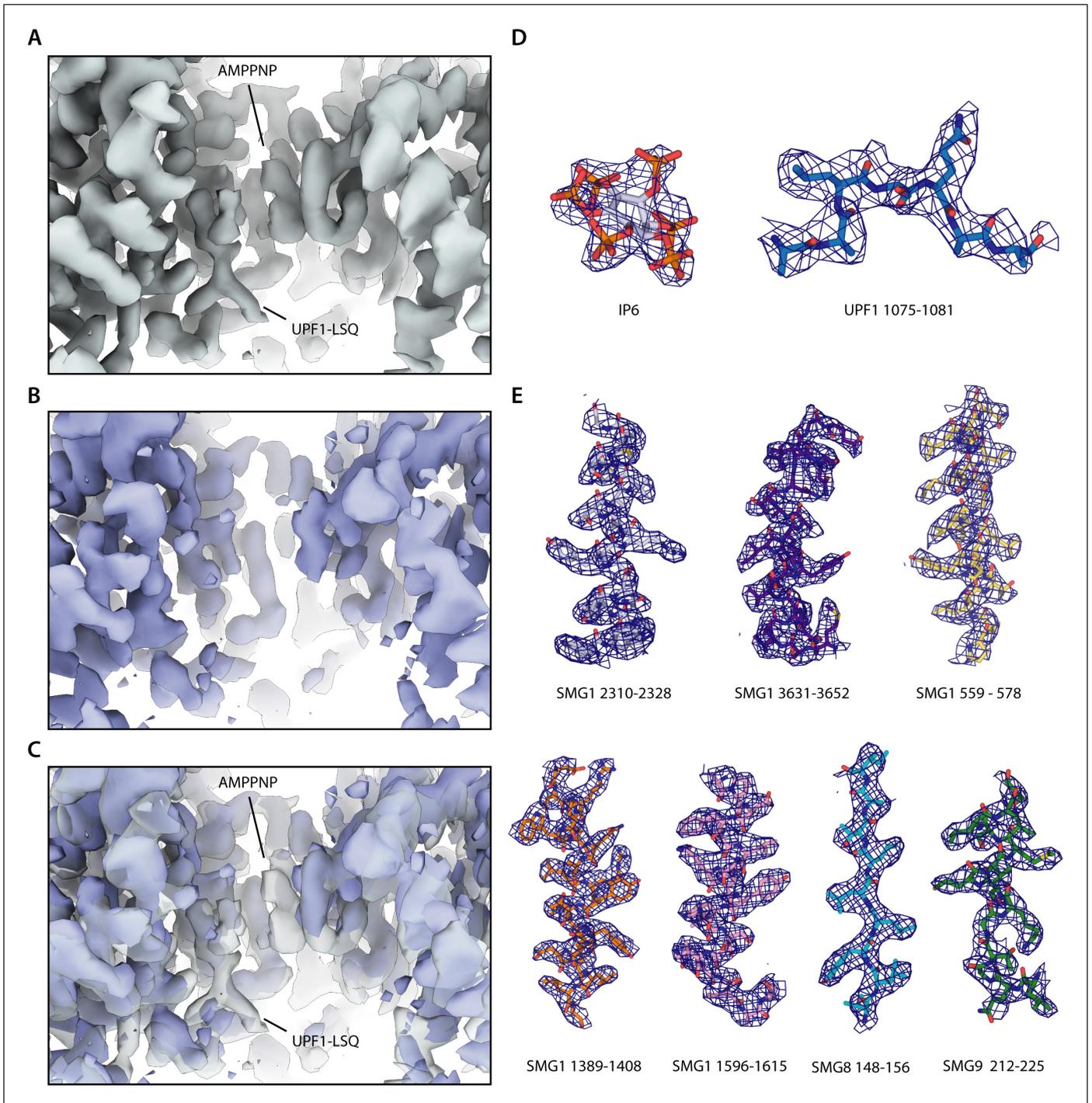


Figure 1—figure supplement 5. Quality of the reconstructed density. (A) Active site density reported in this study. Densities attributed to either AMPPNP or UPF1 peptide are indicated. (B) Active site density of the apo SMG1-8-9 reconstruction reported previously (EMDB 10347). (C) Overlay of the active site densities shown in A and B with the active site density reported in this study shown in transparent grey and density of apo SMG1-8-9 (EMDB 10347) shown in blue. Densities attributed to either AMPPNP or UPF1-LSQ are indicated. (D) Model of IP6 and UPF1-LSQ with the corresponding density shown as a blue mesh. Colors of the model as in **Figure 1**. (E) Additional parts of the model with the corresponding density shown as in B.

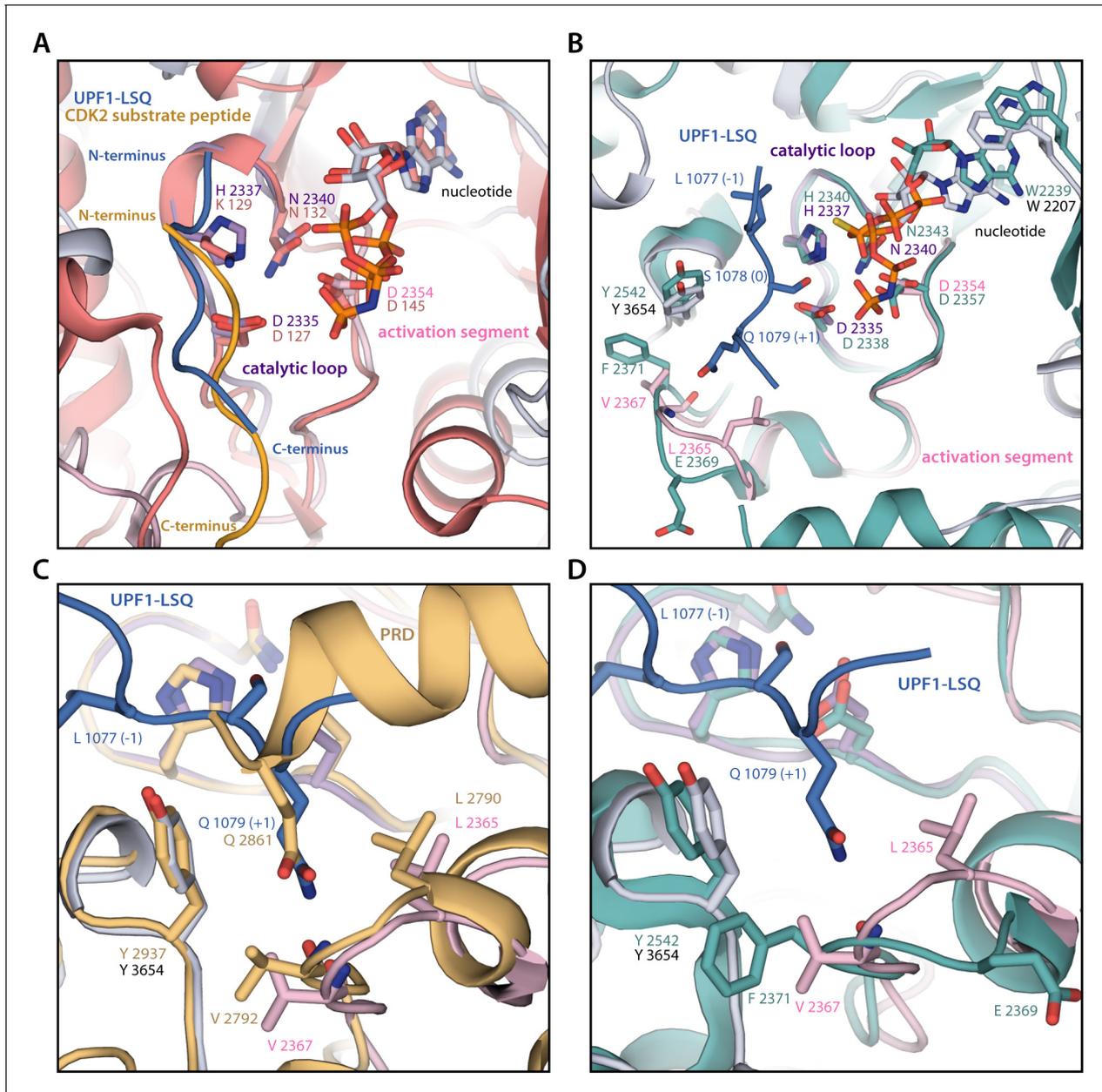


Figure 2—figure supplement 1. Comparison of SMG1 substrate-bound active site to other protein kinases. **(A)** Superposition of substrate-bound SMG1 active site with substrate-bound CDK2 active site (PDB ID, 3QHW). SMG1 catalytic loop is colored in magenta, SMG1 activation segment is in pink, CDK2 is shown in salmon and the CDK2 substrate peptide is in orange. Directionality of the substrates and important catalytic residues are indicated and colored as described above. P-loops were omitted for clarity. The SMG1 kinase contains features that have previously been associated with a model preferring the presence of a ‘dissociative’ transition state, such as a positive charge closely involved in β -phosphate coordination, namely K2155 (residue not shown) (*Wang and Cole, 2014*). **(B)** Superposition of substrate-bound SMG1 active site with nucleotide-bound mTOR active site (PDB ID, 4JSP). SMG1 active site is colored as described in panel A and mTOR is in green. Important active site residues are indicated. P-loops were omitted for clarity. **(C)** Close-up of the active site hydrophobic cage after superposing the model of the substrate-bound SMG1 active site with *Chaetomium thermophilum* Tel1^{ATM} (PDB ID, 6SKY). SMG1 active site is colored as described in panel A and CtTel1^{ATM} is colored in light orange. Key residues are indicated. **(D)** Close-up of the active site hydrophobic cage based on the superposition of SMG1 and mTOR shown in B. Color scheme as in B and important residues are indicated. View is similar to C.

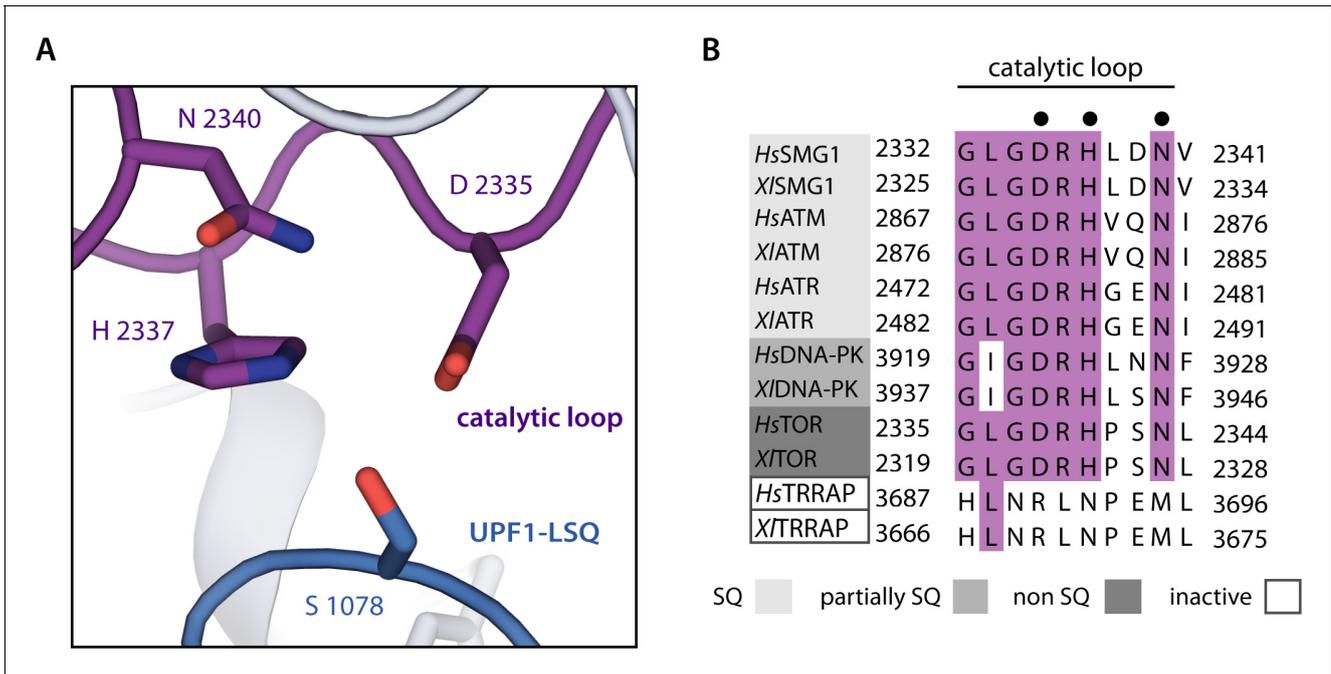


Figure 2—figure supplement 2. Recognition of UPF1-LSQ phospho-acceptor residue Ser1078 by SMG1 catalytic loop residues. (A) Close-up showing how residues of the SMG1 catalytic loop coordinate Ser1078 of UPF1-LSQ. (B) Alignment of catalytic loop sequences from *Homo sapiens* and *Xenopus laevis* for all PIKKs colored according to conservation. PIKKs are grouped by phosphorylation site specificity and residues highlighted in subfigure C are indicated by a black dot.

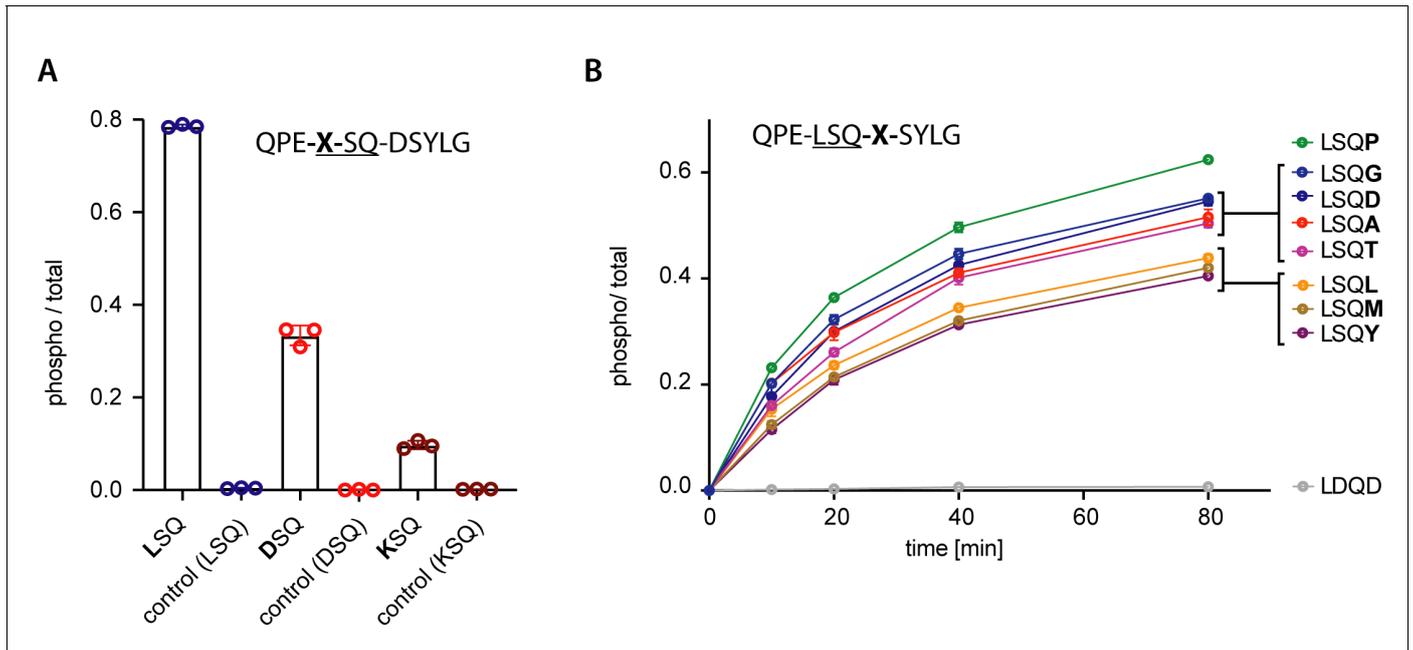


Figure 4—figure supplement 1. In vitro phosphorylation of UPF1-LSQ and derivatives. (A) Mass spectrometry-based end-point measurements of single peptides combined with a control. Data points corresponding to one experiment are shown in the same color. One preferred (LSQ) and two suboptimal (DSQ, KSQ) motifs were chosen to evaluate how the absence of more optimal, competing motifs affects phosphorylation efficiency toward single peptides. The results indicate that decreased phosphorylation of peptides with non-hydrophobic position -1 residues is not only caused by competition with more optimal substrates (compare assay shown in **Figure 4B**). A peptide with the sequence QPELDQDSYLG was used as a control in all three experiments. Reactions were quenched and analyzed after 80 min. The peptide sequence is shown with the varied position marked as 'X' (B) Mass spectrometry-based phosphorylation assay with UPF1-LSQ and derivatives varying in position $+2$. The peptide sequence is indicated with the varying position marked as 'X'. Error bars representing standard deviations calculated from independent experimental triplicates are shown.

Figure 4—figure supplement 2 continued

UPF1 N-terminus. SQ motifs are highlighted by orange boxes with the phosphorylation site position indicated on top. (B) As in A, but for UPF1 C-terminus and using red color to indicate SQ motifs. Location and range of UPF1-LSQ are indicated.

4.2 Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8

This section presents a paper describing structures of SMG1-9 and SMG1-8-9 complexes bound to AMPPNP or a SMG1-specific small molecule inhibitor. The resulting model of a SMG1 active site bound to the inhibitor shows that this compound acts in an ATP-competitive manner. It furthermore indicates interactions that might confer specificity to SMG1. This specificity is additionally shown using *in vitro* phosphorylation assays comparing the effect of the inhibitor on mTOR and SMG1 activity. Integrative modeling combining lower resolution cryo-EM data, crosslinking mass spectrometry and an AlphaFold2 prediction allowed to model the SMG8 C-terminal domain in the context of SMG1-8-9. Importantly, the structures - supported by biochemical assays - reveal that the SMG1 insertion domain contains a PRD that can block the substrate binding site within the SMG1 active site in an autoinhibited configuration likely stabilized by SMG8. These observations explain previously reported biochemical data on SMG1 activity regulation within NMD and reveal regulatory principles shared with related PIKKs (Langer *et al.*, 2021).

Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8

Lukas M Langer, Fabien Bonneau, Yair Gat, Elena Conti*

Max Planck Institute of Biochemistry, Martinsried, Germany

Abstract The PI3K-related kinase (PIKK) SMG1 monitors the progression of metazoan nonsense-mediated mRNA decay (NMD) by phosphorylating the RNA helicase UPF1. Previous work has shown that the activity of SMG1 is impaired by small molecule inhibitors, is reduced by the SMG1 interactors SMG8 and SMG9, and is downregulated by the so-called SMG1 insertion domain. However, the molecular basis for this complex regulatory network has remained elusive. Here, we present cryo-electron microscopy reconstructions of human SMG1-9 and SMG1-8-9 complexes bound to either a SMG1 inhibitor or a non-hydrolyzable ATP analog at overall resolutions ranging from 2.8 to 3.6 Å. These structures reveal the basis with which a small molecule inhibitor preferentially targets SMG1 over other PIKKs. By comparison with our previously reported substrate-bound structure (Langer et al., 2020), we show that the SMG1 insertion domain can exert an autoinhibitory function by directly blocking the substrate-binding path as well as overall access to the SMG1 kinase active site. Together with biochemical analysis, our data indicate that SMG1 autoinhibition is stabilized by the presence of SMG8. Our results explain the specific inhibition of SMG1 by an ATP-competitive small molecule, provide insights into regulation of its kinase activity within the NMD pathway, and expand the understanding of PIKK regulatory mechanisms in general.

*For correspondence:
conti@biochem.mpg.de

Competing interest: The authors declare that no competing interests exist.

Funding: See page 15

Preprinted: 28 July 2021

Received: 28 July 2021

Accepted: 25 October 2021

Published: 26 October 2021

Reviewing Editor: Philip A Cole, Harvard Medical School, United States

© Copyright Langer et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editor's evaluation

This study uses CryoEM and biochemical studies to uncover a new and potentially important conformational off-state of a key regulatory multi-subunit protein kinase, SMG1. The study was enabled by applying a small molecule ATP-site inhibitor to capture the structure. The work will be of wide interest to the signaling and structural biology communities.

Introduction

Nonsense-mediated mRNA decay (NMD) is a co-translational mRNA quality control pathway central to the detection and removal of mRNAs containing premature termination codons as well as to the regulation of many physiological transcripts (Kurosaki et al., 2019; Karousis and Mühlemann, 2019). In metazoans, translation termination at a premature stop codon triggers phosphorylation of the RNA helicase UPF1, which then enables the recruitment of downstream effectors, in turn resulting in the degradation of the targeted mRNA by ribonucleases (Ohnishi et al., 2003; Yamashita, 2013; Chakrabarti et al., 2014; Nicholson et al., 2014; Okada-Katsuhata et al., 2012; Kashima et al., 2006). UPF1 phosphorylation is thus a crucial point in metazoan NMD; it occurs specifically at particular Ser-containing motifs and is mediated by the SMG1 kinase (Denning et al., 2001; Yamashita et al., 2001).

SMG1 is a member of the phosphatidylinositol-3-kinase-related kinase (PIKK) family, which also includes the ATM, ATR, DNA-PKc, and mTOR kinases (Keith and Schreiber, 1995). Despite the

confounding name reflecting a possible evolutionary origin with a lipid kinase (Keith and Schreiber, 1995), and in retrospect the ability of some PIKK family members to bind inositol-6-phosphate (Gat et al., 2019), all active PIKKs are Ser-Thr protein kinases. These large, multidomain enzymes share an overall similar architecture: a C-terminal catalytic module (comprising the so-called FAT, kinase, and FATC domains) and an N-terminal α -solenoid (that typically serves as a protein-protein interaction module). Indeed, PIKKs bind to and are regulated by interacting proteins, with which they form large and dynamic assemblies. In addition, some PIKKs contain an autoregulatory element, particularly in the variable region connecting the kinase and FATC domains, that is known as the PIKK-regulatory domain (PRD) (Baretic and Williams, 2014; Imseng et al., 2018; Jansma and Hopfner, 2021). Finally, the kinase domains of PIKKs have been the target of numerous efforts in the development of specific inhibitors. Not surprisingly, given the central roles of these kinases in pathways surveilling cellular homeostasis, small molecules targeting specific PIKKs have been approved as therapeutics or are being evaluated in clinical trials (Janku et al., 2018; Zhang et al., 2011; Durant et al., 2018). Nevertheless, the high conservation of the PIKK kinase domain has complicated efforts to develop inhibitors specifically targeting only a selected member of this family of enzymes. Structural data rationalizing the binding specificity of such compounds to PIKKs are still sparse in general and entirely missing for the SMG1 kinase.

In the case of SMG1, the two interacting factors SMG8 and SMG9 have been linked to dysregulated NMD and neurodevelopmental disorders in humans (Alzahrani et al., 2020; Shaheen et al., 2016; Yamashita et al., 2009). Previous cryo-electron microscopy (cryo-EM) studies have revealed the molecular interactions that underpin the structure of the human SMG1-SMG8-SMG9 complex (Gat et al., 2019; Zhu et al., 2019) and the determinants with which it recognizes and phosphorylates UPF1 peptides with specific Leu-Ser-Gln (LSQ) motifs (Langer et al., 2020). SMG8 and SMG9 form an unusual G-domain heterodimer (Gat et al., 2019; Langer et al., 2020; Li et al., 2017). The G-domain of SMG9 binds both the α -solenoid and the catalytic module of SMG1 while the G-domain of SMG8 engages only the α -solenoid, thus rationalizing biochemical studies pointing to the crucial role of SMG9 in enabling the incorporation of SMG8 into the complex (Yamashita et al., 2009; Arias-Palomo et al., 2011). In turn, SMG8 appears to have a direct regulatory function on SMG1, as its removal or C-terminal truncation results in hyper-activation of SMG1 kinase activity (Yamashita et al., 2009; Deniaud et al., 2015; Arias-Palomo et al., 2011; Zhu et al., 2019). However, the C-terminal domain of SMG8 remains poorly defined in all available cryo-EM reconstructions (Zhu et al., 2019; Gat et al., 2019; Langer et al., 2020), hindering a molecular understanding of its regulatory role.

Another portion of the complex reported to downregulate SMG1 kinase activity is the so-called insertion domain, a large 1200-residue region connecting the SMG1 kinase and FATC domains. Removal of the SMG1 insertion domain causes hyper-activation of the kinase (Deniaud et al., 2015; Zhu et al., 2019), similarly to the effect reported for the PRDs of other PIKKs (McMahon et al., 2002; Edinger and Thompson, 2004; Xiao et al., 2019; Mordes et al., 2008). However, the SMG1 insertion domain shows no sequence similarity to the PRDs of other PIKKs and is remarkably larger in comparison. None of the current cryo-EM reconstructions of SMG1 or its complexes show density corresponding to the SMG1 insertion domain (Zhu et al., 2019; Gat et al., 2019; Langer et al., 2020), and it is thus unclear how it regulates kinase activity.

Here, we used cryo-EM and biochemical approaches to study the basis with which SMG1 is specifically inhibited by a small molecule compound (SMG1i), and in doing so we also identified the molecular mechanisms with which SMG1 is downregulated by its own insertion domain in cis and SMG8 in trans.

Results

The compound SMG1i specifically inhibits SMG1 kinase activity in vitro

We set out to study the inhibitory mechanism of SMG1i (Figure 1A), a small-molecule compound based on a pyrimidine derivative that has been reported to inhibit SMG1 catalytic activity (Gopalsamy et al., 2012; Mino et al., 2019). We expressed and purified human wild-type SMG1-8-9 complex from piggyBac transposase generated HEK293T stable cell pools, essentially as described before (Gat et al., 2019). We have previously shown that the use of a UPF1 peptide comprising the UPF1 phosphorylation site 1078 (UPF1-LSQ) allows monitoring of its specific phosphorylation by SMG1

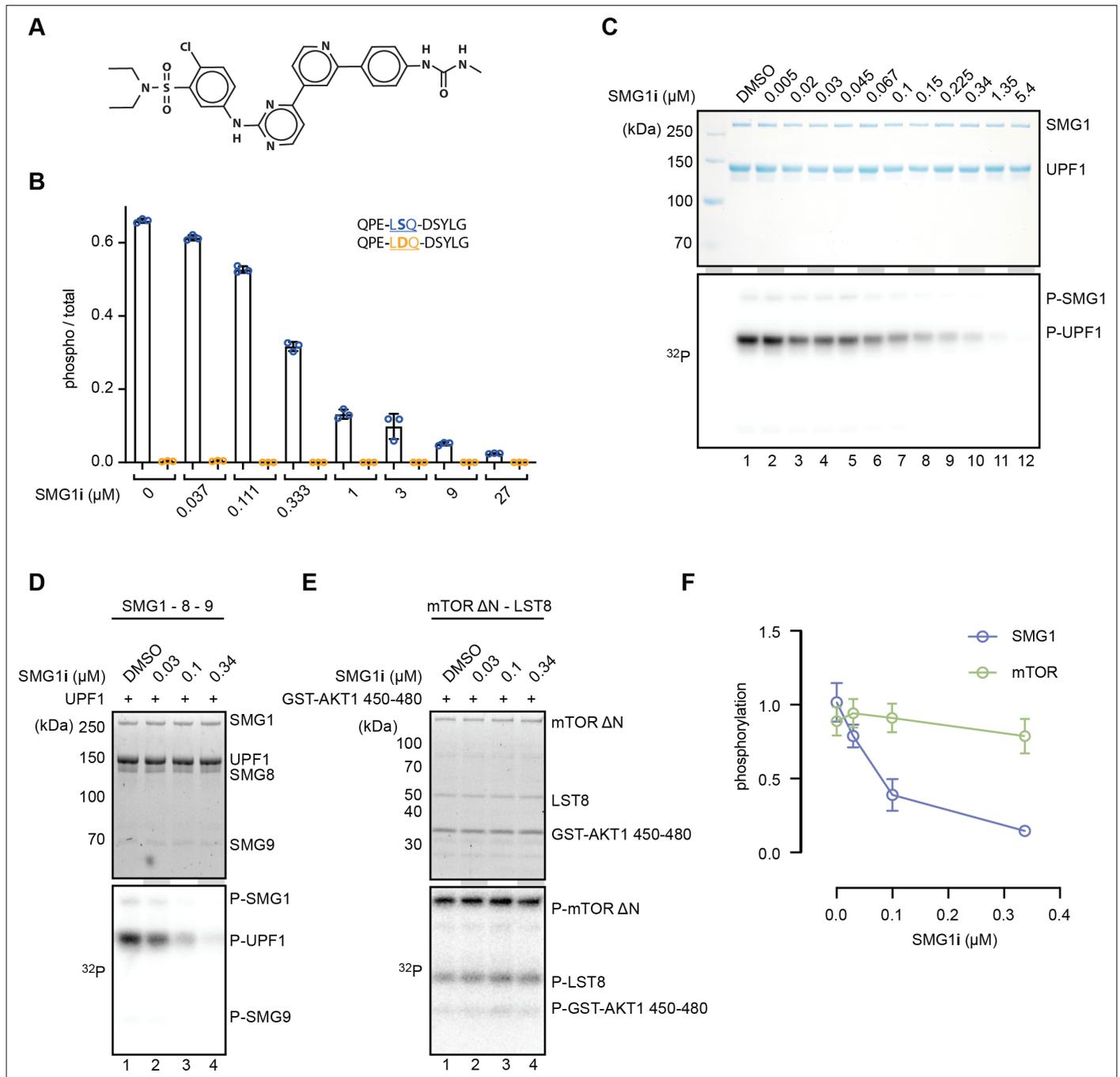


Figure 1. SMG1i specifically inhibits SMG1 kinase activity in vitro. **(A)** Structure of the SMG1 inhibitor (SMG1i). **(B)** Titration of SMG1i using a mass spectrometry-based phosphorylation assay with 500 nM SMG1-8-9 and the indicated UPF1-derived peptides as substrates. **(C)** Titration of SMG1i using a radioactivity-based phosphorylation assay with 100 nM SMG1-8-9 and full-length UPF1 as a substrate. The Coomassie-stained gel is shown on top, and the corresponding radioactive signal is below. **(D, E)** Titration of SMG1i using a radioactivity-based phosphorylation assay with either SMG1-8-9 and full-length UPF1 **(D)** or mTOR^{ΔN}-LST8 and GST-AKT1 **(E)** as a substrate. Stain-free gels are shown on top and the corresponding radioactive signal at the bottom (*Ladner et al., 2004*). **(F)** Quantification of normalized UPF1 or mTOR (auto-)phosphorylation in the presence of increasing amounts of SMG1i for both SMG1 and mTOR. For each data point, the mean is shown with standard deviations of the three replicates indicated.

The online version of this article includes the following figure supplement(s) for figure 1:

Source data 1. Unedited images for gels shown in **Figure 1**.

Figure supplement 1. Characterization of SMG1 inhibitor.

Figure 1 continued on next page

Figure 1 continued

Figure supplement 2. Radioactivity-based phosphorylation assays using SMG1 inhibitor with SMG1 and mTOR.**Figure supplement 2—source data 1.** Unedited images for gels shown in **Figure 1—figure supplement 2**.

over time using mass spectrometry (Langer et al., 2020). We made use of this method to assay the effect and potency of SMG1i (Figure 1, Figure 1—figure supplement 1). To avoid artifacts caused by ATP depletion, we performed an ATP titration and analyzed the end-point measurements (Figure 1—figure supplement 1B). As a control, we used a UPF1-LDQ peptide in which we exchanged the phosphor-acceptor Ser1078 of UPF1-LSQ to aspartic acid. Next, we repeated end-point measurements under conditions of stable ATP concentration and increasing amounts of SMG1i (Figure 1B). In this assay, we observed a significant reduction of UPF1-LSQ phosphorylation when adding SMG1i at concentrations equimolar to the enzyme (Figure 1B and Figure 1—figure supplement 1C). To corroborate these results, we repeated titration of SMG1i under conditions of stable ATP concentration using a radioactivity-based phosphorylation assay and full-length UPF1 as a substrate. Again, we observed a reduction in UPF1 phosphorylation in the presence of low micromolar amounts of SMG1i (Figure 1C).

To assess the specificity of SMG1i, we tested its effect on the mTOR kinase using a similar radioactive kinase assay. In this experiment, we used a truncated mTOR^{ΔN}-LST8 complex previously shown to be constitutively active and responsive to mTOR inhibitors (Yang et al., 2013). As an mTOR kinase substrate, we used a GST-AKT1^{450–480} fusion protein. We observed that SMG1i only affected mTOR activity in the highest concentration range, well beyond concentrations that had an observable effect

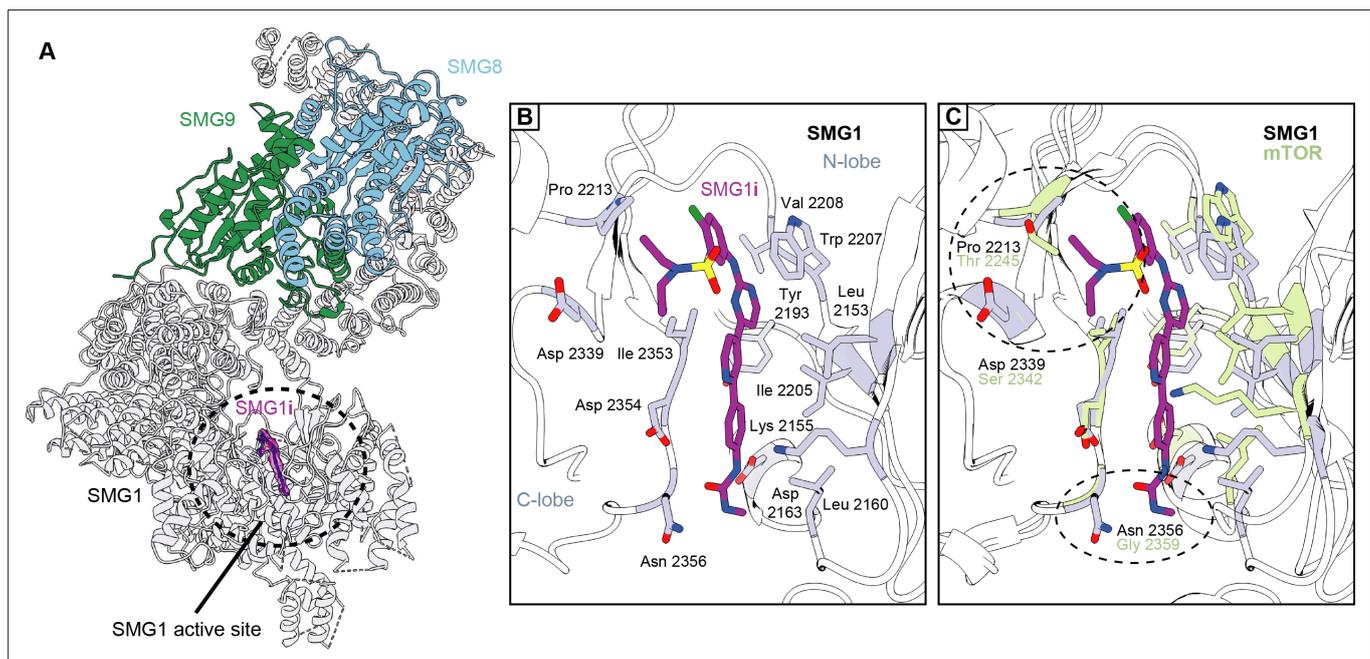


Figure 2. Structural basis for selective targeting of SMG1 by the SMG1 inhibitor. (A) Model of the SMG1-8-9 kinase complex bound to SMG1i. SMG1 is in gray, SMG8 is in blue, and SMG9 is shown in green. SMG1i is shown as a magenta model overlaid with the isolated transparent density. Approximate location of the SMG1 active site is indicated by a black circle. (B) Key interactions of SMG1i with SMG1 active site residues. Important residues located in either the N- or the C-lobe of the SMG1 kinase domain are colored gray. Other parts of SMG1 are transparent and interactions of SMG1i with SMG1 backbone are not shown. (C) Superposition of SMG1i-bound SMG1 with the mTOR active site (PDB identifier: 4JSP) over the catalytic loops of both kinases. Key SMG1 residues indicated in (B) are shown alongside the respective mTOR residues colored in green. Regions possibly accounting for preferential interaction of SMG1i with SMG1 over mTOR are circled and the relevant residues are labeled.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Resolution distribution and isotropy of SMG1-centered cryo-EM maps bound to SMG1i.

Figure supplement 2. Cryo-EM data processing of SMG1i data set.

Figure supplement 3. Further details of SMG1i binding and specificity.

on SMG1 activity (**Figure 1—figure supplement 2A**). We selected four different concentrations of SMG1i based on these assays (**Figure 1C, Figure 1—figure supplement 2A**) and repeated the radioactivity-based kinase assay for both SMG1 and mTOR in triplicates (**Figure 1D and E, Figure 1—figure supplement 2B, C**). Upon quantifying the levels of phosphorylation using densitometry and normalizing them to the amount of protein loaded in each lane, we found that SMG1i robustly inhibited SMG1 while mTOR activity was only weakly affected (**Figure 1F, Figure 1—figure supplement 2D, E,F**). In agreement with previous results (*Gopalsamy et al., 2012*), we concluded that SMG1i displays high potency and specificity in inhibiting the SMG1 kinase in vitro.

Cryo-EM structure of SMG1-8-9 bound to the SMG1i inhibitor

To understand the molecular basis for the SMG1i mode of action, we reconstituted purified wild-type SMG1-8-9 with an excess of SMG1i and subjected the sample to single-particle cryo-EM analysis. The resulting reconstruction reached an overall resolution of 3.1 Å, and we observed clear density for the SMG1 inhibitor bound to the kinase active site (**Figure 2A, Figure 2—figure supplements 1 and 2** and **Figure 2—figure supplement 3B, C**). Superposition with our previously published AMPPNP- and substrate-bound model (*Langer et al., 2020*) revealed that SMG1i exploits essentially the same binding site as the ATP analog. However, SMG1i wedges deeper in between the N- and C-lobe of the kinase domain and is engaged in more extensive interactions (**Figure 2B** and **Figure 2—figure supplement 3A and D**). These observations identify SMG1i as a potent ATP-competitive inhibitor.

While many of the SMG1 residues interacting with SMG1i are conserved within the PIKK family, two contact sites appear to be specific: SMG1 Pro2213 and Asp2339 contact the sulfonamide moiety of the inhibitor while Asn2356 forms hydrogen bonds with the phenyl urea/carbamide group (**Figure 2B**). Superposition of the SMG1 active site in the SMG1i-bound structure with the active site of mTOR^{ΔN} shows that the corresponding residues have diverged and would not be able to mediate analogous interactions with the SMG1i sulfonamide group (mTOR Thr2245 and Ser2342 at the positions of SMG1 Pro2213 and Asp2339) or with the urea group (mTOR Gly2359 at the position of SMG1 Asn2356) (**Figure 2C**). The same regions of the small molecule are diverging or missing in the mTOR inhibitor Torin 2 (**Figure 2—figure supplement 3E**). Taken together, these observations rationalize the specificity of SMG1i for SMG1 over mTOR^{ΔN} that we observed in the in vitro assays (**Figure 1D-F** and **Figure 1—figure supplement 2D, E and F**).

Superposition with other PIKKs suggests the presence of potentially similar discriminatory interactions: like mTOR, DNA-PK lacks a favorable interaction site for the SMG1i urea group (Gly3943 at the position of SMG1 Asn2356) (**Figure 2—figure supplement 3F**). In addition, DNA-PK PRD residue Lys4019 would sterically clash with the SMG1i urea group in the inactive conformation of the DNA-PK active site. In the case of ATR, the SMG1i sulfonamide site has diverged (Gly2385 at the position of SMG1 Pro2213) and residues in the N-lobe would result in a steric clash with the urea group of the inhibitor (Lys2329 and Asp2330 corresponding to SMG1 Leu2157 and Glu2158) (**Figure 2—figure supplement 3G**). Finally, mTOR, DNA-PK, and ATR active sites have a different relative orientation of the kinase N- and C-lobes as compared to SMG1, changing the overall chemical environment at the SMG1i-binding site (**Figure 2C** and **Figure 2—figure supplement 3F and G**). We conclude that subtle differences in the SMG1 structure underpin the specificity of the SMG1i inhibitor.

The SMG1 insertion domain contains a PRD that blocks substrate binding in the presence of SMG8

In the reconstruction of the SMG1i-bound SMG1-8-9 complex, we observed additional density near the end of the kinase domain, protruding from the position where the insertion domain is expected to start (**Figure 3, Figure 3—figure supplement 1A and B**). While the lower resolution for this part of the map was not sufficient to unambiguously build an atomic model, this density clearly docks at the substrate-binding site of the kinase domain (**Figure 3A, Figure 3—figure supplement 1A, B and C**). Superposition with the previous cryo-EM structure of SMG1-8-9 complex bound to a UPF1 peptide (*Langer et al., 2020*) revealed that the additional density in the current reconstruction is mutually exclusive with the position of a bound substrate (**Figure 3I, Figure 3—figure supplement 1A**). This observation is in good agreement with the reported positions of the PRDs in structures of the auto-inhibited forms of yeast Tel1^{ATM} and human DNA-PK (*Jansma et al., 2020; Yates et al., 2020; Chen et al., 2021*). We concluded that the SMG1 insertion domain contains a PRD that can directly occupy

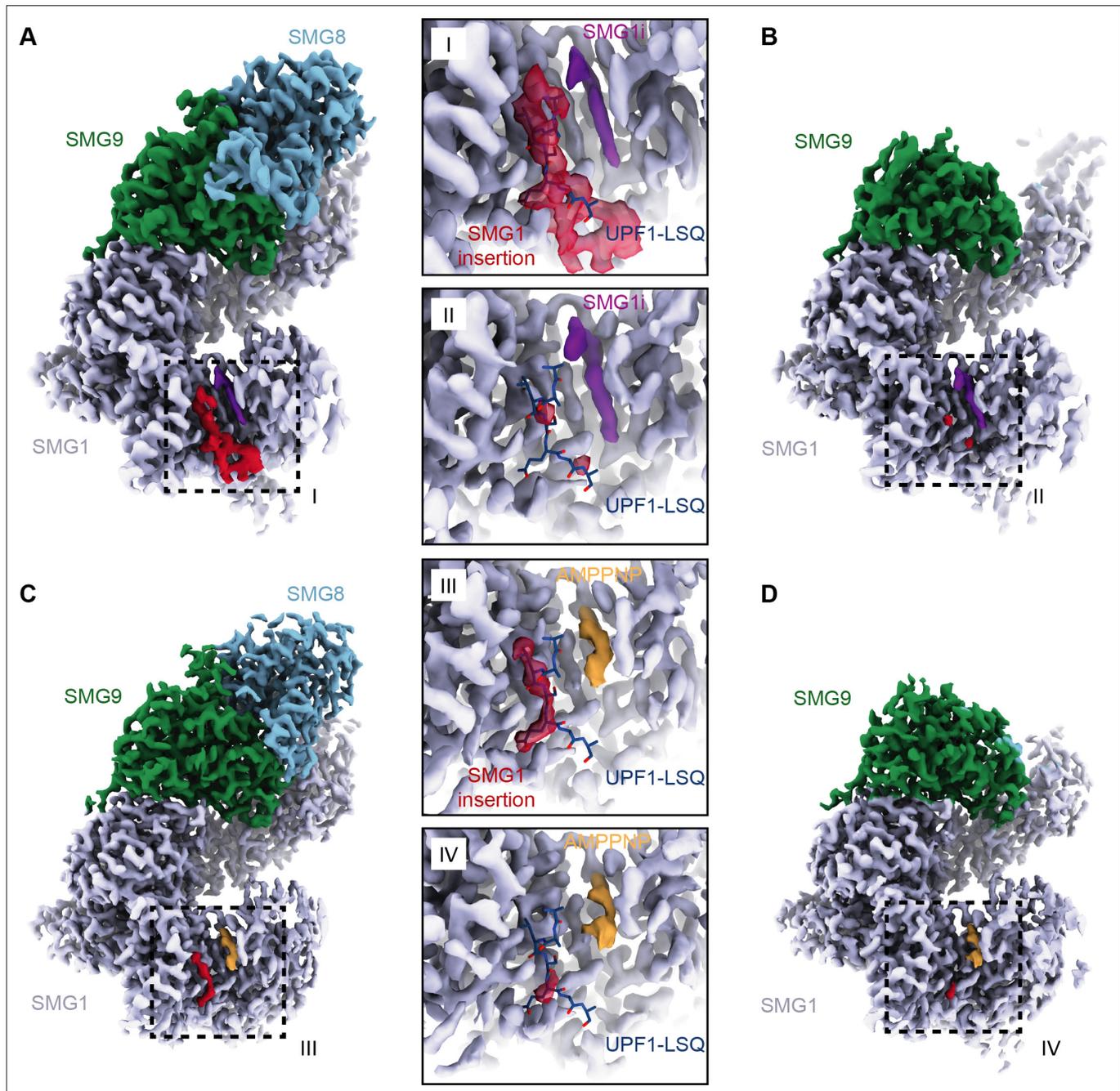


Figure 3. Structures of SMG1-8-9 and SMG1-9 complexes reveal that the SMG1 insertion domain can block the substrate-binding path in the presence of SMG8. (A) Cryo-EM density of SMG1-8-9 bound to SMG1i. Density for the inhibitor is in magenta, the N-terminus of the SMG1 insertion is in red, and all other parts as indicated. (B) Cryo-EM density of SMG1-9 complex bound to the SMG1 inhibitor. Everything else as in (A). (C) Cryo-EM density of SMG1-8-9 bound to AMPPNP. Density for AMPPNP is in orange and extra density attributed to the SMG1 insertion is in red. (D) As in (C), but for the SMG1-9 complex bound to AMPPNP. Insets I–IV show close-ups of the indicated kinase active site densities superimposed with the model for the UPF1-LSQ substrate shown as blue sticks (PDB identifier: 6Z3R). All maps segmented. cryo-EM, cryo-electron microscopy.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Details of the SMG1 insertion N-terminus.

Figure supplement 2. Resolution distribution and isotropy of SMG1-centered cryo-EM maps bound to AMPPNP.

Figure supplement 3. Cryo-EM data processing of AMPPNP data set.

Figure supplement 4. Details of the SMG1-9 complex.

and block the substrate-binding path in the kinase active site, analogous to other PIKKs. The SMG1 PRD would have to re-arrange upon substrate binding, as observed in the case of the inactive-to-active transition of DNA-PK (Chen et al., 2021).

Within the cryo-EM data set collected on an SMG1-8-9 complex incubated with SMG1i, we isolated a subset of particles missing any density for SMG8, hence representing an SMG1i-bound SMG1-9 complex (Figure 3B, Figure 2—figure supplement 1C, Figure 3—figure supplement 4). The corresponding 3.6 Å reconstruction not only showed a similar mode of interaction of SMG1i, but also several features distinct from SMG1 in isolation and the SMG1-8-9 ternary complex. Globally, the α -solenoid of SMG1 rearranges, consistent with the notion of a stepwise compaction of SMG1 upon binding first SMG9 and then SMG8 (Figure 3—figure supplement 4A and B; Melero et al., 2014; Zhu et al., 2019). Locally, the mode of interaction between SMG9 and the SMG1 α -solenoid differs between the two complexes. In the binary SMG1-9 complex, the portion of SMG9 that interacts with the N-terminal HEAT repeat helices of SMG1 changes conformation, exploiting part of a binding site that was occupied by a short N-terminal segment of SMG8 in the ternary complex (Figure 3—figure supplement 4C, D and E). Indeed, two hydrophobic residues of SMG9 (Leu456 and Leu457) take the place normally reserved for two hydrophobic residues of SMG8 (Leu351 and Leu352) (Figure 3—figure supplement 4F and G). Hence, the conformation of the SMG9 segment observed in the SMG1-9 complex is incompatible with the binding of SMG8 to this part of SMG1. Importantly, the reconstruction of SMG1-9 showed no ordered density at the active site for the PRD (Figure 3B and II, Figure 3—figure supplement 1A), which is in contrast to the ordered PRD density visualized in the SMG1i-bound SMG1-8-9 complex.

Next, we asked whether the autoinhibitory state of the SMG1 insertion domain in the SMG1-8-9 complex is due specifically to the presence of the SMG inhibitor or whether it is a more general feature. To address this question, we reconstituted purified wild-type SMG1-8-9 with an excess of AMPPNP (instead of SMG1i) and subjected the sample to a similar single-particle cryo-EM analysis routine. We obtained maps for SMG1-8-9 and SMG1-9 at overall resolutions of 2.8 Å and 3.1 Å, respectively (Figure 3C and D, Figure 3—figure supplements 2 and 3). Both reconstructions showed clear density for AMPPNP in the SMG1 active site. In the reconstruction of the ternary SMG1-8-9 complex, we observed an additional density at the substrate-binding site, in the same position as that observed when in the presence of SMG1i (Figure 3III, IV). However, in the context of AMPPNP this additional density is less prominent and does not connect to the SMG1 kinase domain. It thus appears that the SMG1 inhibitor has a stabilizing effect on the autoinhibitory conformation of the SMG1 PRD. In the reconstruction of the binary SMG1-9 complex in the presence of AMPPNP, we observed no ordered density at the substrate-binding site, consistent with the reconstruction obtained in the presence of SMG1i. We concluded that the autoinhibitory conformation of the PRD in the insertion domain of SMG1 is connected to the presence of a nucleotide in the ATP-binding site and is stabilized by the presence of SMG8.

The SMG1 insertion domain can block overall access to the kinase active site and interacts with the SMG8 C-terminus

The findings above suggested a possible role for SMG8 in stabilizing an autoinhibited state of SMG1. Indeed, previous work implicated SMG8 and, in particular, its C-terminal domain in the downregulation of SMG1 activity (Zhu et al., 2019; Arias-Palomo et al., 2011; Deniaud et al., 2015; Yamashita et al., 2009). SMG8 has a modular domain organization: the N-terminal G-domain is followed by a helical stalk that protrudes into solvent with well-defined density, but the remaining C-terminal region (amounting to about 45% of the molecule) is flexible and poorly resolved in the published cryo-EM studies (Gat et al., 2019; Zhu et al., 2019; Langer et al., 2020). By further processing of the SMG1i-bound SMG1-8-9 data, we could achieve improved density for the C-terminal domain of SMG8, showing a knob-like feature directly connected to the end of the stalk and in turn connecting to a larger globular mass (Figure 4A). Concurrently, we observed an additional density extending from the location of the PRD that we attributed to the SMG1 insertion domain (see below). On one side, this density wraps along the catalytic module of SMG1, occludes the access to the kinase active site (Figure 4A), and reaches toward the IP₆-binding site (Figure 4A and C). On the other side, the density projects into solvent and approaches the globular SMG8 C-terminal region. Although the low resolution of this part of the map did not allow model building, the close proximity suggested a physical

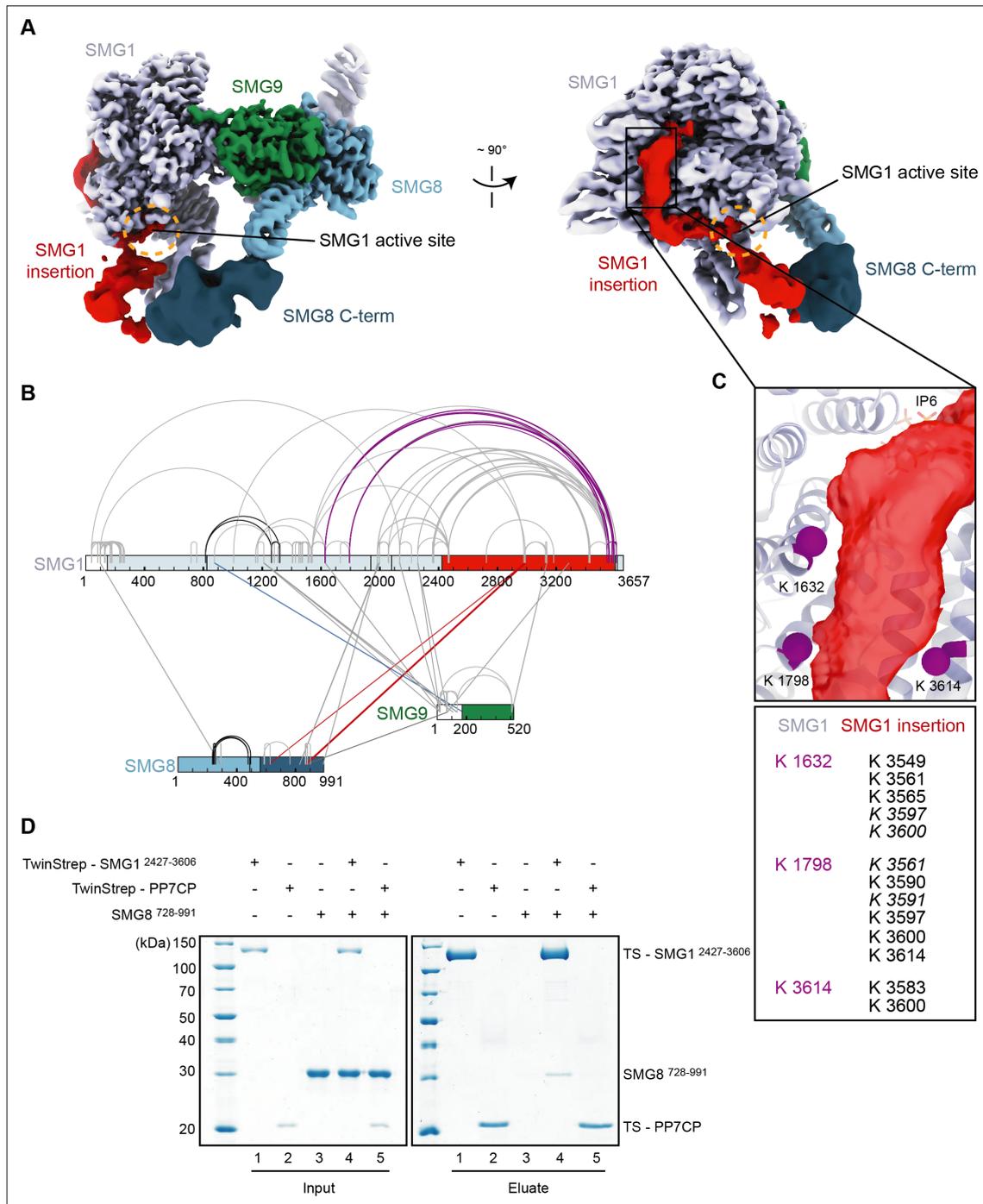


Figure 4. The SMG1 insertion domain can block overall access to the kinase active site. **(A)** Cryo-EM map after 3D variability analysis filtered by resolution and segmented. Two different views displaying extra density for SMG8 C-terminus (dark blue) and SMG1 insertion domain. The location of the SMG1 kinase active site is indicated by an orange circle. **(B)** Inter and intra cross-links of the SMG1-8-9 kinase complex. Proteins are colored as in **(A)**. Intra-links shown in **Figure 4—figure supplement 1B, C** are in black, the inter-link shown in **Figure 4—figure supplement 1D** is in blue, and inter-links between SMG1 insertion and SMG1 C-terminus are in red. **(C)** Zoom-in highlighting positions of Lys residues (magenta spheres) crosslinking to the C-terminus of the SMG1 insertion domain. A table listing the cross-linked residues is shown below. Cross-links found in only one of two samples are in italics. **(D)** Coomassie-stained SDS-PAGE analysis of pull-down experiment showing an interaction between SMG1 insertion domain (SMG1^{2427–3606}) and SMG8 C-terminus (SMG8^{728–991}). cryo-EM, cryo-electron microscopy.

The online version of this article includes the following source data and figure supplement(s) for figure 4:

Figure 4 continued on next page

Figure 4 continued

Source data 1. Table with cross-linking data used in **Figure 4**.

Source data 2. Unedited images for gels shown in **Figure 4**.

Figure supplement 1. Cross-linking mass spectrometry of SMG1-8-9.

Figure supplement 1—source data 1. Unedited images for gels shown in **Figure 4—figure supplement 1**.

Figure supplement 2. Selected spectra of detected SMG1-8-9 intra-links.

Figure supplement 3. Integration of cryo-EM, cross-linking MS, and AlphaFold data reveals a model for the SMG8 C-terminus.

Figure supplement 4. Further characterization of SMG1-8-9—centered interactions.

Figure supplement 4—source data 1. Unedited images for gels shown in **Figure 4—figure supplement 4**.

contact between the SMG1 insertion domain and the SMG8 C-terminal domain. We proceeded to test the interplay between these two parts of the SMG1-8-9 complex.

We first used cross-linking mass spectrometry (XL-MS) analysis (**Figure 4B**, **Figure 4—figure supplements 1 and 2**). Upon treatment of purified SMG1-8-9 with Bis(sulfosuccinimidyl)suberate (BS³), we detected intra and inter cross-links consistent with parts of the complex for which an atomic model was available and measuring well within the expected range of distance for BS³ cross-links (**Figure 4B**, **Figure 4—figure supplements 1 and 2**). We also observed reproducible intra cross-links between the end of the SMG1 insertion and the catalytic module, in particular the FAT (residues 823–1938) and kinase (residues 2080–2422) domains (**Figure 4B and C**). These cross-links mapped in close proximity to the aforementioned density extending from the PRD. We thus attribute this density to the C-terminus of the SMG1 insertion (**Figure 4A, B and C**). By combining our low-resolution EM density and our cross-linking data with an AlphaFold prediction of SMG8 (**Jumper et al., 2021**), we were able to produce a more complete model of SMG1-8-9, now containing large parts of the SMG8 C-terminus (**Figure 4—figure supplement 3**). Importantly, we detected inter cross-links between the SMG1 insertion domain and the SMG8 C-terminal domain (**Figure 4B**), consistent with the proximity of the EM densities (**Figure 4A**) and previously reported XL-MS data (**Deniaud et al., 2015**).

Next, we tested whether these parts of the complex can interact directly in biochemical assays with recombinant proteins. We purified SMG1 insertion domain (SMG1 residues 2427–3606) with an N-terminal TwinStrep-tag from transiently transfected HEK293T cells and SMG8 C-terminus (SMG8 residues 728–991) from *Escherichia coli*. Since the structural analysis is indicative of a dynamic or weak contact, we performed pull-down assays in low-salt conditions. We observed that the SMG1 insertion domain indeed co-precipitated the SMG8 C-terminus in these conditions (**Figure 4D**). As a control, an unrelated TwinStrep-tagged protein (PP7 coat protein) did not pull down SMG8 C-terminus under the same conditions (**Figure 4D**). As anticipated, increasing the salt concentration had a detrimental effect on the SMG8 C-terminus—SMG1 insertion co-precipitation (**Figure 4—figure supplement 4A**).

Taken together, these experiments indicate the existence not just of proximity but of a direct physical link between the SMG1 insertion domain and SMG8 C-terminus, rationalizing the observed dependency of SMG1 autoinhibition on the presence of SMG8 (**Figure 3**).

Conclusions

In this manuscript, we reported a reconstruction of the SMG1-8-9 complex bound to an SMG1 inhibitor. We showed that this compound binds to the ATP-binding site within the kinase active site. By comparison with structural data of related kinases, we identified two functional groups within the small molecule that possibly confer specificity to SMG1. Using the SMG1-related mTOR kinase as an example, we confirmed the specific action of the inhibitor biochemically.

In addition to the SMG1 inhibitor, we observed density for the N-terminal part of the SMG1 insertion domain in the SMG1-8-9 reconstruction. This part of the insertion domain acts as a PRD and occupied the substrate-binding path within the SMG1 active site. Consistent with data for other PIKKs, a PRD within the SMG1 insertion domain can therefore restrict access to the kinase active site.

The structure of an SMG1i-bound SMG1-9 complex reconstructed from the same data set as SMG1i-bound SMG1-8-9 did not show ordered density for the SMG1 PRD in the SMG1 active site. This suggested that an interaction between the SMG1 insertion and SMG8 was important for stabilizing the autoinhibited state of the complex. Consistently, our biochemical analysis indicated a direct physical link between parts of the SMG1 insertion domain and the C-terminal domain of SMG8.

While autoinhibition was not observed in previously published apo-structures, our reconstructions of SMG1-9 and SMG1-8-9 bound to AMPPNP showed that blockage of the substrate-binding path was also possible when the SMG1-8-9 complex active site was bound to a nucleotide. Accordingly, the described autoinhibition was not a phenomenon limited to the inhibitor-bound structures. We conclude that at least two layers of regulation of SMG1 kinase activity exist: first, access to the SMG1 active site can be restricted in cis by a PRD within the SMG1 insertion domain. Second, this autoinhibition is stabilized in trans by the SMG8 C-terminus.

How is the autoinhibited complex activated? The position of the SMG1 insertion domain in the inhibitor-bound SMG1-8-9 complex reconstruction raised the possibility that the autoinhibited complex might not be able to interact with its substrate, UPF1 (Figures 3I, III and 4A). To test this hypothesis, we incubated TwinStrep-tagged SMG1-8-9 with increasing concentrations of SMG1i before adding UPF1 and performed a pull-down experiment (Figure 4—figure supplement 4B). The efficiency of UPF1 co-precipitation was unaffected by the presence of SMG1i, suggesting that either UPF1 can overcome blockage of the SMG1 active site by the SMG1 insertion domain or has secondary binding sites independent of the observed SMG1 autoinhibition. Indeed, superposition of the autoinhibited complex reconstruction with a previously reported negative-stain map of a cross-linked UPF1-bound SMG1-8-9 complex raised the possibility that UPF1-binding and SMG1 autoinhibition might be able to occur in parallel (Figure 4—figure supplement 4C). An alternative explanation would be that optimal positioning of UPF1 with respect to the SMG1 kinase will allow the unstructured N- and C-terminal ends of UPF1 to efficiently compete with the SMG1 insertion domain for binding to the SMG1 active site in vivo. This would be mimicked by the high protein concentrations used in in vitro assays (such as Figure 4—figure supplement 4B).

Whether SMG8 or an SMG8-9 complex dissociates from SMG1 at any point in the kinases' catalytic cycle during NMD in vivo is an outstanding question. Early biochemical experiments have suggested that SMG8 plays a crucial role in recruiting SMG1 complex to the NMD-competent messenger ribonucleoprotein particle (Yamashita et al., 2009). Given that the SMG1-8-9 complex actively phosphorylated UPF1 in in vitro experiments using purified proteins and our observation that SMG1i binding did not impair the interaction of SMG1 with UPF1, it is tempting to speculate that UPF1 itself might be sufficient to overcome the effect of SMG8 on stabilizing SMG1 autoinhibition within an SMG1-8-9 complex. In this model, nucleotide binding and kinase activation could occur in two distinct steps during NMD: upon ATP binding, the kinase complex would adopt an autoinhibited conformation and

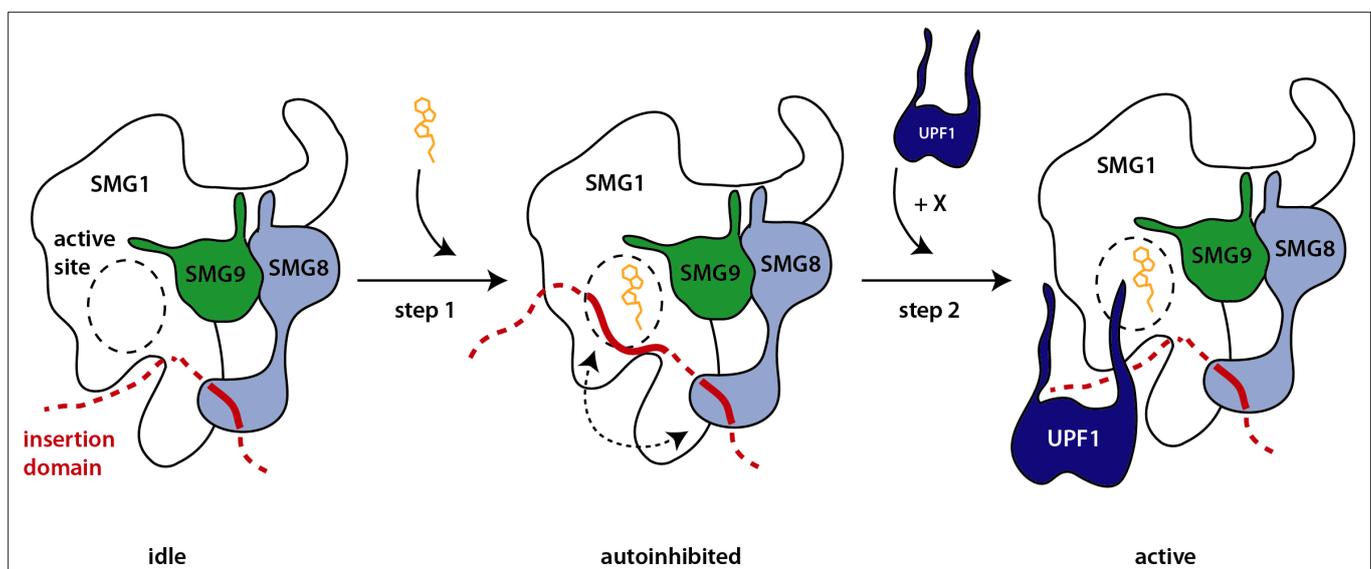


Figure 5. Hypothetical model of SMG1 kinase regulation. Structural and biochemical data suggest different layers of regulation on SMG1 kinase activity. Upon ATP binding (orange), the SMG1 kinase adopts an autoinhibited conformation (step 1)—mediated by the concerted action of the SMG1 insertion domain in cis and the SMG8 C-terminus in trans (dotted arrow). The presence of the correct substrate and likely other factors/cues (indicated by X) then trigger the release of the autoinhibition and activity of the kinase toward UPF1 (step 2).

only be licensed for specific phosphorylation of its target by the substrate itself (**Figure 5**). In addition, autophosphorylation of the kinase complex may play a crucial role in kinase activation.

Taken together, our structures will help to improve the design of selective PIKK active site inhibitors, for example, specific for SMG1. Our data suggest a unifying structural model for concerted action of the SMG1 insertion domain in cis and SMG8 in trans to tune SMG1 kinase activity, shedding light on the intricate regulation of this kinase complex in metazoan NMD.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Gene (<i>Homo sapiens</i>)	SMG1	Shigeo Ohno lab	Uniprot Q96Q15	
Gene (<i>H. sapiens</i>)	SMG8	Shigeo Ohno lab	Uniprot Q8ND04	
Gene (<i>H. sapiens</i>)	SMG9	Shigeo Ohno lab	Uniprot Q9H0W8	
Cell line (<i>H. sapiens</i>)	HEK293T	ATCC		
Strain, strain background (<i>Escherichia coli</i>)	BL21 Star (DE3) pRARE	EMBL Heidelberg Core Facility		Electrocompetent cells
Peptide, recombinant protein	UPF1-LSQ (peptide 1078) and derivative	In-house as described in doi: https://elifesciences.org/articles/57127		
Chemical compound, drug	SMG1 inhibitor	Robert Bridges, Rosalind Franklin University of Medicine and Science, and the Cystic Fibrosis Foundation		
Chemical compound, drug	AMPPNP	Sigma-Aldrich		
Chemical compound, drug	ATP	Sigma-Aldrich		
Software, algorithm	SerialEM	https://bio3d.colorado.edu/SerialEM/		
Software, algorithm	Focus	https://focus.c-cina.unibas.ch/wiki/doku.php	v1.1.0	
Software, algorithm	RELION	doi: 10.7554/eLife.42166	RELION 3.0	
Software, algorithm	Cryosparc	doi: 10.1038/nmeth.4169	Cryosparc2	
Software, algorithm	UCSF Chimera	UCSF, https://www.cgl.ucsf.edu/chimera/		
Software, algorithm	UCSF ChimeraX	UCSF, https://www.rbvi.ucsf.edu/chimerax/		
Software, algorithm	COOT	http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/		
Software, algorithm	PHENIX	https://www.phenix-online.org/	PHENIX 1.17	
Software, algorithm	Molprobit	Duke Biochemistry, http://molprobit.biochem.duke.edu/		
Software, algorithm	PyMOL	PyMOL Molecular Graphics System, Schrodinger LLC	PyMOL 2.3.2	

Protein expression and purification

TwinStrep-tagged SMG1-8-9 complex was expressed and purified as described before (**Gat et al., 2019; Langer et al., 2020**). The source cells were authenticated by genotyping (Eurofins) and tested negative for mycoplasma contamination (LookOut Mycoplasma PCR Detection Kit).

N-terminally TwinStrep-tagged SMG1 insertion domain (SMG1^{2427–3606}) was expressed and purified essentially identical to the SMG1-8-9 full-length complex, with the exception that lysis and affinity purification was carried out in a buffer based on 2× phosphate-buffered saline (PBS). The protein was frozen in gel filtration buffer at concentrations ranging from 15 to 40 μM depending on the purification.

Untagged full-length UPF1 expressed in *E. coli* was purified as reported previously (**Langer et al., 2020**).

N-terminally TwinStrep-tagged full-length UPF1 was expressed using a stable pool of HEK293T cells as described for SMG1-8-9. Approximately 400×10^6 cells were lysed with a Dounce homogenizer in buffer containing $1 \times$ PBS, 5 mM $MgCl_2$, 1 μ M $ZnCl_2$, 10% (v/v) glycerol, 1 mM DTT supplemented for lysis with Benzonase and DNase I as well as EDTA-free Complete Protease Inhibitor Cocktail (Roche). The lysate was cleared by centrifugation at 25,000 rpm for 30 min at $10^\circ C$. The supernatant was filtered and applied to a StrepTrap HP column (Sigma-Aldrich). After passing 25 column volumes of lysis buffer without supplements, the column was further washed with 25 column volumes of buffer Hep A (20 mM HEPES/NaOH pH 7.4, 85 mM KCl, 2 mM $MgCl_2$, 1 μ M $ZnCl_2$, 10% (v/v) glycerol, and 1 mM DTT) and bound protein was eluted directly onto a HiTrap Heparin column (GE Healthcare) using buffer Hep A supplemented with 2.5 mM Desthiobiotin. After washing the Heparin column with 10 column volumes of buffer Hep A, bound protein was eluted using a gradient increasing the salt concentration to 500 mM KCl over 100 ml while collecting fractions (AEKTA prime FPLC system, GE Healthcare). Fractions were analyzed by SDS-PAGE and those containing pure full-length TwinStrep-UPF1 were collected, concentrated to 30–80 μ M, and flash-frozen without buffer exchange using liquid nitrogen until further usage.

The SMG8 C-terminus (SMG8^{728–991}) was N-terminally fused to a 6xHis-SUMO tag and expressed in *E. coli* BL21 STAR (DE3) pRARE cells. Bacteria were grown in TB medium at $37^\circ C$ shaking at 180 rpm. At an OD of 2, cultures were cooled down to $18^\circ C$ and overnight protein expression was induced by adding IPTG. After harvesting (6000 rpm, 10 min), bacteria were lysed by sonication in a buffer containing 20 mM Tris-HCl pH 7.5, 400 mM NaCl, 10 mM Imidazole pH 7.5, 1 mM β -mercaptoethanol, Benzonase, DNase I and 1 mM PMSF. After centrifugation (25,000 rpm, 30 min, $10^\circ C$) and filtration, the cleared lysate was applied to a HisTrap FF column (Cytiva) and the column was washed with 30 column volumes of wash buffer (20 mM Tris-HCl pH 7.5, 400 mM NaCl, 40 mM Imidazole pH 7.5, and 1 mM β -mercaptoethanol). Bound protein was eluted with wash buffer supplemented with 340 mM Imidazole pH 7.5 and the elution was combined with His-tagged SUMO protease for cleavage of the N-terminal tag and dialyzed overnight against 20 mM Tris-HCl pH 7.5, 400 mM NaCl, and 1 mM β -mercaptoethanol. The dialyzed, protease treated sample was passed again over a 5 ml HisTrap FF column to remove cleaved tags and protease, the flow-through was diluted to 150 mM NaCl and $10^\circ C$ (v/v) glycerol and applied to a HiTrap SP HP column (Cytiva). After washing with 10 column volumes of 20 mM Tris-Cl pH 7.5, 150 mM NaCl, $10^\circ C$ (v/v) glycerol, and 1 mM β -mercaptoethanol, proteins were eluted using a gradient increasing the salt concentration to 500 mM NaCl. The peak fractions containing SMG8^{728–991} were pooled, concentrated, and injected onto a Superdex 75 10/300 GL (Cytiva) equilibrated with 20 mM HEPES/NaOH pH 7.4, 200 mM NaCl, and 1 mM DTT. Again, the peak fraction containing SMG8^{728–991} was pooled and concentrated to 100–200 μ M and flash-frozen in gel filtration buffer using liquid nitrogen. mTOR^{AN}-LST8 and GST-AKT1^{450–480} were prepared as described before (Gat *et al.*, 2019).

Kinase assays

The mass spectrometry-based kinase assays were carried out as previously reported (Langer *et al.*, 2020). For experiments involving SMG1i or Torin 2, the assembled reactions were incubated in the absence of ATP for 30 min at $4^\circ C$. All reactions involving SMG1i were performed with equal amounts of DMSO present.

Kinase assay based on the use of γ -³²P-labeled ATP was carried out as reported. For all assays, 100 nM of kinase complex and 1 μ M of substrate were used. All reactions were supplemented with equal amounts of DMSO and incubated without ATP for 30 min at $4^\circ C$ before starting the experiment. After 30 min at $30^\circ C$, phosphorylation reactions were stopped by adding SDS-containing sample buffer and samples were analyzed using an SDS-polyacrylamide gel supplemented with 2,2,2-trichloroethanol (TCE, Acros Organics) for stain-free visualization (Ladner *et al.*, 2004). The TCE-tryptophan reaction was started by exposing the protein gel to UV light for 2 min and proteins were visualized and quantified using a Bio-Rad gel imaging system and the Image Lab software (version 6.1, Bio-Rad) or by Coomassie-staining. Subsequently the gel was washed several times with water to remove residual γ -³²P-labeled ATP and phosphoproteins were detected using autoradiography and an Amersham Typhoon RGB biomolecular imager (GE Healthcare). Phosphoprotein signal (P-UPF1 or P-mTOR^{AN}) was quantified by densitometry using Fiji and normalized to the total amount of protein

determined using TCE. Independent triplicates of each condition were normalized to the DMSO-only sample and plotted using Prism (GraphPad).

Cross-linking mass spectrometry

For cross-linking mass spectrometry, 1 μM of SMG1-8-9 complex was incubated with 0.5 mM BS³ for 30 min on ice in a buffer containing 1 \times PBS, 5 mM MgCl₂, and 1 mM DTT. The reaction was quenched by adding 40 mM Tris-HCl pH 7.9 and incubating for 20 min on ice. The sample was spun for 15 min at 18,000 \times g. For denaturation of the crosslinked proteins, 4 M Urea and 50 mM Tris were added to the supernatant and the samples were sonicated using a Bioruptor Plus sonication system (Diogenode) for 10 \times 30 s at high intensity. For reduction and alkylation of the proteins, 40 mM 2-chloroacetamide (CAA, Sigma-Aldrich) and 10 mM Tris (2-carboxyethyl)phosphine (TCEP; Thermo Fisher Scientific) were added. After incubation for 20 min at 37°C, the samples were diluted 1:2 with MS grade water (VWR). Proteins were digested overnight at 37°C by addition of 1 μg of trypsin (Promega). Thereafter, the solution was acidified with trifluoroacetic acid (TFA; Merck) to a final concentration of 1%, followed by desalting of the peptides using Sep-Pak C18 1 cc vacuum cartridges (Waters). The elution was vacuum dried.

Enriched peptides were loaded onto a 30 cm analytical column (inner diameter: 75 μm ; packed in-house with ReproSil-Pur C18-AQ 1.9 μm beads, Dr. Maisch GmbH) by the Thermo Easy-nLC 1000 (Thermo Fisher Scientific) with buffer A (0.1% (v/v) Formic acid) at 400 nl/min. The analytical column was heated to 60°C. Using the nanoelectrospray interface, eluting peptides were sprayed into the benchtop Orbitrap Q Exactive HF (Thermo Fisher Scientific)(*Hosp et al., 2015*). As gradient, the following steps were programmed with increasing addition of buffer B (80% Acetonitrile, 0.1% Formic acid): linear increase from 8% to 30% over 60 min, followed by a linear increase to 60% over 5 min, a linear increase to 95% over the next 5 min, and finally maintenance at 95% for another 5 min. The mass spectrometer was operated in data-dependent mode with survey scans from m/z 300 to 1650 Th (resolution of 60k at $m/z=200$ Th), and up to 15 of the most abundant precursors were selected and fragmented using stepped Higher-energy C-trap Dissociation (HCD with a normalized collision energy of value of 19, 27, and 35). The MS2 spectra were recorded with dynamic m/z range (resolution of 30k at $m/z=200$ Th). AGC target for MS1 and MS2 scans were set to 3×10^6 and 10^5 , respectively, within a maximum injection time of 100 and 60 ms for the MS1 and MS2 scans, respectively. Charge state 2 was excluded from fragmentation to enrich the fragmentation scans for cross-linked peptide precursors.

The acquired raw data were processed using Proteome Discoverer (version 2.5.0.400) with the XlinkX/PD nodes integrated (*Klykov et al., 2018*). To identify the crosslinked peptide pairs, a database search was performed against a FASTA containing the sequences of the proteins under investigation. DSS was set as a crosslinker. Cysteine carbamidomethylation was set as fixed modification and methionine oxidation and protein N-term acetylation were set as dynamic modifications. Trypsin/P was specified as protease and up to two missed cleavages were allowed. Furthermore, identifications were only accepted with a minimal score of 40 and a minimal delta score of 4. Otherwise, standard settings were applied. Filtering at 1% false discovery rate (FDR) at peptide level was applied through the XlinkX Validator node with setting simple. A file listing identified cross-links can be found as **Figure 4—source data 1**.

Pull-down assays

Pull downs including the SMG8 C-terminus were performed using Strep-TactinXT Superflow beads (IBA) equilibrated with 20 mM HEPES/NaOH pH 7.4, 50 mM NaCl, 10% (w/v) glycerol, 1 mM DTT, and 0.1% (w/v) NP-40 substitute (Fluka) unless stated otherwise. 1.5 μM TwinStrep-tagged protein was combined with 15 μM untagged protein in the buffer described above and incubated at 4°C for 30 min before adding equilibrated resin. The complete reaction was incubated for 1.5 hr, beads were washed four times with 20 \times resin volume of buffer, and bound protein eluted by adding buffer supplemented with 50 mM Biotin.

To analyze the effect of SMG1i binding to SMG1 on the interaction between the SMG1-8-9 kinase complex and UPF1, 0.12 μM TwinStrep-tagged SMG1-8-9 was combined with DMSO, or the respective amounts of SMG1i in buffer containing 1 \times PBS, 5 mM MgCl₂, 1 mM DTT, and 0.1% (v/v) NP-40. After incubation for 30 min at 4°C, 0.4 μM of untagged UPF1 was added and the reactions were combined

with MagStrep 'type3' XT beads (IBA). Samples were incubated for 30 min before washing four times with 20× resin volume of buffer and precipitation of bound proteins using SDS-containing sample buffer.

For all pull-down assays, input and elution samples were analyzed by SDS-PAGE and stained with Coomassie.

Cryo-EM sample preparation, data collection, and data processing

Grids were prepared as described before ([Langer et al., 2020](#)), with the difference that SMG1-8-9 was incubated with either 4 μM SMG1i or 1 mM AMPPNP for 30 min on ice in 1× PBS, 5 mM MgCl₂, and 1 mM DTT before adding 0.04% (v/v) n-octyl-β-D-glucoside and plunging using an ethane/propane mixture and a Thermo Fisher FEI Vitrobot IV.

Cryo-EM data were essentially acquired as reported previously using a Thermo Fisher FEI Titan Krios G3 microscope equipped with a post-column GIF (energy width 20 eV). The Gatan K3 camera was used in counting mode and data were acquired using SerialEM ([Mastronarde, 2005](#)) and a beam-tilt based acquisition scheme. The nominal magnification during data collection for both data sets was 105,000×, corresponding to a pixel size of 0.8512 Å at the specimen level. The SMG1i sample was imaged with a total exposure of 89.32 e⁻/Å² evenly spread over 4 s and 40 frames. The AMPPNP sample was imaged with a total exposure of 60.99 e⁻/Å² evenly spread over 5.7 s and 40 frames at CDS mode. For both data collections, the target defocus ranged between -0.5 and -2.9 μm.

Movies were pre-processed on the fly using Focus ([Biyani et al., 2017](#)), while automatically discarding images of poor quality. Picked candidate particles were extracted in RELION 3.1 ([Zivanov et al., 2018](#); [Scheres, 2012](#)). After two rounds of reference-free 2D classification, particles were imported to CryoSPARC v2 ([Punjani et al., 2017](#)) for further processing. After additional sorting in 2D, particles were iteratively classified in 3D to separate SMG1-8-9 and SMG1-9. Particles were distributed over multiple batches for all classification steps. To visualize highly flexible parts of the SMG1 insertion domain in the SMG1i-bound complex, the particle stack containing SMG1-8-9 data was submitted to 3D variability analysis into six clusters. Filtering according to local resolution was employed to better visualize parts of the complex with different resolutions after refinement. Following Ctf refinement, a mask was generated based on the more rigid SMG1 body to focus refinement of the cleaned SMG1i-bound SMG1-8-9 particles on the inhibitor-bound SMG1 active site. For details refer to [Figure 2—figure supplement 2](#) and [Figure 3—figure supplement 3](#). 3D FSC curves were calculated using the 3D FSC online application ([Tan et al., 2017](#)). Map versus model FSCs were calculated within the PHENIX software suit ([Adams et al., 2011](#); [Liebschner et al., 2019](#)).

Model building

Model building was carried out using Coot (version 0.8.9.2) ([Emsley et al., 2010](#); [Emsley and Cowtan, 2004](#)) and iterative rounds of real-space refinement in the PHENIX software suit ([Adams et al., 2011](#); [Liebschner et al., 2019](#)) and was based on our previously published models of SMG1-8-9 (PDB identifiers: 6SYT, 6Z3R). Geometric restraints for the SMG1i molecule were calculated with eLBOW ([Moriarty et al., 2009](#)). For details see [Supplementary file 1](#). Structure visualization and analysis were carried out using UCSF ChimeraX (version 1.2.5) ([Goddard et al., 2018](#)) and PyMOL (version 2.3.2).

To include the AlphaFold derived SMG8 C-terminus model, we downloaded the predicted model for full-length SMG8 from the AlphaFold Protein Structure Database ([Jumper et al., 2021](#)). We removed all parts with a per-residue confidence score below 50 and superimposed the prediction with our EM data-derived model for the SMG8 N-terminal half (residues 4–559) using Coot. After fusing the roughly positioned AlphaFold-predicted SMG8 C-terminal half (560–991) to our SMG1i-bound SMG1-8-9 model (PDB identifier: 7PW4), we rigid body-fitted the SMG8 C-terminus into the EM density (resulting in PDB 7PW5). It is worth noting that the somewhat different orientation of the SMG8 C-terminal domain with respect to the N-terminal part of the molecule observed when comparing the EM density to the AlphaFold model before rigid body-fitting is in agreement with the flexibility of this domain generally present in the cryo-EM data. An overview of refinement statistics is shown in [Supplementary file 1](#).

Acknowledgements

The authors are grateful to Robert Bridges, Rosalind Franklin University of Medicine and Science, and the Cystic Fibrosis Foundation for providing the SMG1 inhibitor. Daniel Bollschweiler and Tillman

Schäfer at the MPIB cryo-EM facility for help with EM data collection and Barbara Steigenberger and Elisabeth Weyher at MPIB biochemistry core facility for carrying out mass spectrometry. The authors thank Christian Benda and J Rajan Prabu for maintenance and development of computational infrastructure for EM data processing, Marcela Cueto for TwinStrep-PP7CP, Daniela Wartini for help with tissue culture, and Courtney Long and members of the lab for help with the preparation of the manuscript. This work was supported by funding from the Max Planck Gesellschaft, the European Commission (ERC Advanced Investigator Grant EXORICO), the Novo Nordisk Foundation (Exo-Adapt Center) and the German Research Foundation (DFG SFB1035, GRK1721, SFB/TRR 237) to EC and a Boehringer Ingelheim Fonds PhD fellowship to LL.

Additional information

Funding

Funder	Grant reference number	Author
Boehringer Ingelheim Fonds		Lukas M Langer
Max-Planck-Gesellschaft		Elena Conti
European Commission	ERC Advanced Investigator Grant EXORICO	Elena Conti
Deutsche Forschungsgemeinschaft	SFB1035	Elena Conti
Deutsche Forschungsgemeinschaft	GRK1721	Elena Conti
Deutsche Forschungsgemeinschaft	SFB/TRR 237	Elena Conti
Novo Nordisk Foundation	Exo-Adapt Center	Elena Conti

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Lukas M Langer, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review and editing; Fabien Bonneau, Data curation, Formal analysis, Methodology, Writing - review and editing; Yair Gat, Investigation, Writing - review and editing; Elena Conti, Conceptualization, Project administration, Resources, Supervision, Writing - original draft, Writing - review and editing

Author ORCIDs

Lukas M Langer  <http://orcid.org/0000-0002-9977-2427>

Fabien Bonneau  <http://orcid.org/0000-0001-8787-7662>

Yair Gat  <http://orcid.org/0000-0002-2338-9384>

Elena Conti  <http://orcid.org/0000-0003-1254-5588>

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.72353.sa1>

Author response <https://doi.org/10.7554/eLife.72353.sa2>

Additional files

Supplementary files

- Supplementary file 1. Cryo-EM data collection, refinement and validation statistics.
- Transparent reporting form

Data availability

Models have been deposited in the PDB under the accession codes 7PW4 (SMG1-8-9 bound to SMG1i), 7PW5 (SMG1-8-9 bound to SMG1i, with SMG8 C-terminus), 7PW6 (SMG1 body bound to SMG1i), 7PW7 (SMG1-9 bound to SMG1i), 7PW8 (SMG1-8-9 bound to AMPPNP) and 7PW9 (SMG1-9 bound to AMPPNP). The corresponding EM maps have been deposited in the EMDB under the accession codes EMD-13674, EMD-13675, EMD-13676, EMD-13677, EMD-13678 and EMD-13679.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex bound to a SMG1 inhibitor	https://www.rcsb.org/structure/7PW4	RCSB Protein Data Bank, 7PW4
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex bound to a SMG1 inhibitor with predicted SMG8 C-terminus	https://www.rcsb.org/structure/7PW5	RCSB Protein Data Bank, 7PW5
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex bound to a SMG1 inhibitor - SMG1 body	https://www.rcsb.org/structure/7PW6	RCSB Protein Data Bank, 7PW6
Langer LM, Conti E	2021	Human SMG1-9 kinase complex bound to a SMG1 inhibitor	https://www.rcsb.org/structure/7PW7	RCSB Protein Data Bank, 7PW7
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex bound to AMPPNP	https://www.rcsb.org/structure/7PW8	RCSB Protein Data Bank, 7PW8
Langer LM, Conti E	2021	Human SMG1-9 kinase complex bound to AMPPNP	https://www.rcsb.org/structure/7PW9	RCSB Protein Data Bank, 7PW9
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex bound to a SMG1 inhibitor	https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-13674	Electron Microscopy Data Bank, EMD-13674
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex with AlphaFold predicted SMG8 C-terminus, bound to a SMG1 inhibitor	https://www.ebi.ac.uk/emdb/error/entry/EMD-13675	Electron Microscopy Data Bank, EMD-13675
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex bound to a SMG1 inhibitor - SMG1 body	https://www.ebi.ac.uk/emdb/error/entry/EMD-13676	Electron Microscopy Data Bank, EMD-13676
Langer LM, Conti E	2021	Human SMG1-9 kinase complex bound to a SMG1 inhibitor	https://www.ebi.ac.uk/emdb/error/entry/EMD-13677	Electron Microscopy Data Bank, EMD-13677
Langer LM, Conti E	2021	Human SMG1-8-9 kinase complex bound to AMPPNP	https://www.ebi.ac.uk/emdb/error/entry/EMD-13678	Electron Microscopy Data Bank, EMD-13678
Langer LM, Conti E	2021	Human SMG1-9 kinase complex bound to AMPPNP	https://www.ebi.ac.uk/emdb/error/entry/EMD-13679	Electron Microscopy Data Bank, EMD-13679

References

- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Echols N, Headd JJ, Hung L-W, Jain S, Kapral GJ, Grosse Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner RD, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2011. The Phenix software for automated determination of macromolecular structures. *Methods* **55**: 94–106. DOI: <https://doi.org/10.1016/j.ymeth.2011.07.005>, PMID: 21821126
- Alzahrani F, Kuwahara H, Long Y, Al-Owain M, Tohary M, AlSayed M, Mahnashi M, Fathi L, Alnemer M, Al-Hamed MH, Lemire G, Boycott KM, Hashem M, Han W, Al-Maawali A, Al-Mahrizi F, Al-Thihli K, Gao X, Alkuraya FS. 2020. Recessive, Deleterious Variants in SMG8 Expand the Role of Nonsense-Mediated Decay in

- Developmental Disorders in Humans. *American Journal of Human Genetics* **107**: 1178–1185. DOI: <https://doi.org/10.1016/j.ajhg.2020.11.007>, PMID: 33242396
- Arias-Palomo E**, Yamashita A, Fernández IS, Núñez-Ramírez R, Bamba Y, Izumi N, Ohno S, Llorca O. 2011. The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes & Development* **25**: 153–164. DOI: <https://doi.org/10.1101/gad.606911>, PMID: 21245168
- Baretić D**, Williams RL. 2014. PIKKs--the solenoid nest where partners and kinases meet. *Current Opinion in Structural Biology* **29**: 134–142. DOI: <https://doi.org/10.1016/j.sbi.2014.11.003>, PMID: 25460276
- Biyani N**, Righetto RD, McLeod R, Caujolle-Bert D, Castano-Diez D, Goldie KN, Stahlberg H. 2017. Focus: The interface between data collection and data processing in cryo-EM. *Journal of Structural Biology* **198**: 124–133. DOI: <https://doi.org/10.1016/j.jsb.2017.03.007>, PMID: 28344036
- Chakrabarti S**, Bonneau F, Schüssler S, Eppinger E, Conti E. 2014. Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6. *Nucleic Acids Research* **42**: 9447–9460. DOI: <https://doi.org/10.1093/nar/gku578>, PMID: 25013172
- Chen X**, Xu X, Chen Y, Cheung JC, Wang H, Jiang J, de Val N, Gellert M, Yang W. 2021. Structure of an activated DNA-PK and its implications for NHEJ. *Molecular Cell* **81**: 801–810. DOI: <https://doi.org/10.1016/j.molcel.2020.12.015>
- Deniaud A**, Karuppusamy M, Bock T, Masiulis S, Huard K, Garzoni F, Kerschgens K, Hentze MW, Kulozik AE, Beck M, Neu-Yilik G, Schaffitzel C. 2015. A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. *Nucleic Acids Research* **43**: 7600–7611. DOI: <https://doi.org/10.1093/nar/gkv668>, PMID: 26130714
- Denning G**, Jamieson L, Maquat LE, Thompson EA, Fields AP. 2001. Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein. *The Journal of Biological Chemistry* **276**: 22709–22714. DOI: <https://doi.org/10.1074/jbc.C100144200>, PMID: 11331269
- Durant ST**, Zheng L, Wang Y, Chen K, Zhang L, Zhang T, Yang Z, Riches L, Trinidad AG, Fok JHL, Hunt T, Pike KG, Wilson J, Smith A, Colclough N, Reddy VP, Sykes A, Janefeldt A, Johnström P, Varnäs K, et al. 2018. The brain-penetrant clinical ATM inhibitor AZD1390 radiosensitizes and improves survival of preclinical brain tumor models. *Science Advances* **4**: eaat1719. DOI: <https://doi.org/10.1126/sciadv.aat1719>, PMID: 29938225
- Edinger AL**, Thompson CB. 2004. An activated mTOR mutant supports growth factor-independent, nutrient-dependent cell survival. *Oncogene* **23**: 5654–5663. DOI: <https://doi.org/10.1038/sj.onc.1207738>, PMID: 15133498
- Emsley P**, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography* **60**: 2126–2132. DOI: <https://doi.org/10.1107/S0907444904019158>, PMID: 15572765
- Emsley P**, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallographica. Section D, Biological Crystallography* **66**: 486–501. DOI: <https://doi.org/10.1107/S0907444910007493>, PMID: 20383002
- Gat Y**, Schuller JM, Lingaraju M, Weyher E, Bonneau F, Strauss M, Murray PJ, Conti E. 2019. InsP6 binding to PIKK kinases revealed by the cryo-EM structure of an SMG1-SMG8-SMG9 complex. *Nature Structural & Molecular Biology* **26**: 1089–1093. DOI: <https://doi.org/10.1038/s41594-019-0342-7>, PMID: 31792449
- Goddard TD**, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* **27**: 14–25. DOI: <https://doi.org/10.1002/pro.3235>, PMID: 28710774
- Gopalsamy A**, Bennett EM, Shi M, Zhang WG, Bard J, Yu K. 2012. Identification of pyrimidine derivatives as hSMG-1 inhibitors. *Bioorganic & Medicinal Chemistry Letters* **22**: 6636–6641. DOI: <https://doi.org/10.1016/j.bmcl.2012.08.107>, PMID: 23021994
- Hosp F**, Scheltema RA, Eberl HC, Kulak NA, Keilhauer EC, Mayr K, Mann M. 2015. A Double-Barrel Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) System to Quantify 96 Interactomes per Day. *Molecular & Cellular Proteomics* **14**: 2030–2041. DOI: <https://doi.org/10.1074/mcp.O115.049460>, PMID: 25887394
- Imseng S**, Aylett CH, Maier T. 2018. Architecture and activation of phosphatidylinositol 3-kinase related kinases. *Current Opinion in Structural Biology* **49**: 177–189. DOI: <https://doi.org/10.1016/j.sbi.2018.03.010>, PMID: 29625383
- Janku F**, Yap TA, Meric-Bernstam F. 2018. Targeting the PI3K pathway in cancer: are we making headway? *Nature Reviews. Clinical Oncology* **15**: 273–291. DOI: <https://doi.org/10.1038/nrclinonc.2018.28>, PMID: 29508857
- Jansma M**, Linke-Winnebeck C, Eustermann S, Lammens K, Kostrewa D, Stakyte K, Litz C, Kessler B, Hopfner KP. 2020. Near-Complete Structure and Model of Tel1ATM from *Chaetomium thermophilum* Reveals a Robust Autoinhibited ATP State. *Structure* **28**: 83–95. DOI: <https://doi.org/10.1016/j.str.2019.10.013>
- Jansma M**, Hopfner KP. 2021. Structural basis of the (in)activity of the apical DNA damage response kinases ATM, ATR and DNA-PKcs. *Progress in Biophysics and Molecular Biology* **163**: 120–129. DOI: <https://doi.org/10.1016/j.pbiomolbio.2020.10.009>
- Jumper J**, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583–589. DOI: <https://doi.org/10.1038/s41586-021-03819-2>, PMID: 34265844
- Karousis ED**, Mühlemann O. 2019. Nonsense-Mediated mRNA Decay Begins Where Translation Ends. *Cold Spring Harbor Perspectives in Biology* **11**: a032862. DOI: <https://doi.org/10.1101/cshperspect.a032862>

- Kashima I**, Yamashita A, Izumi N, Kataoka N, Morishita R, Hoshino S, Ohno M, Dreyfuss G, Ohno S. 2006. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes & Development* **20**: 355–367. DOI: <https://doi.org/10.1101/gad.1389006>, PMID: 16452507
- Keith CT**, Schreiber SL. 1995. PIK-related kinases: DNA repair, recombination, and cell cycle checkpoints. *Science* **270**: 50–51. DOI: <https://doi.org/10.1126/science.270.5233.50>, PMID: 7569949
- Klykov O**, Steigenberger B, Pektaş S, Fasci D, Heck AJR, Scheltema RA. 2018. Efficient and robust proteome-wide approaches for cross-linking mass spectrometry. *Nature Protocols* **13**: 2964–2990. DOI: <https://doi.org/10.1038/s41596-018-0074-x>, PMID: 30446747
- Kurosaki T**, Popp MW, Maquat LE. 2019. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nature Reviews. Molecular Cell Biology* **20**: 406–420. DOI: <https://doi.org/10.1038/s41580-019-0126-2>, PMID: 30992545
- Ladner CL**, Yang J, Turner RJ, Edwards RA. 2004. Visible fluorescent detection of proteins in polyacrylamide gels without staining. *Analytical Biochemistry* **326**: 13–20. DOI: <https://doi.org/10.1016/j.ab.2003.10.047>, PMID: 14769330
- Langer LM**, Gat Y, Bonneau F, Conti E. 2020. Structure of substrate-bound SMG1-8-9 kinase complex reveals molecular basis for phosphorylation specificity. *eLife* **9**: e57127. DOI: <https://doi.org/10.7554/eLife.57127>, PMID: 32469312
- Li L**, Lingaraju M, Basquin C, Basquin J, Conti E. 2017. Structure of a SMG8-SMG9 complex identifies a G-domain heterodimer in the NMD effector proteins. *RNA* **23**: 1028–1034. DOI: <https://doi.org/10.1261/rna.061200.117>, PMID: 28389433
- Liebschner D**, Afonine PV, Baker ML, Bunkóczy G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, et al. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica. Section D, Structural Biology* **75**: 861–877. DOI: <https://doi.org/10.1107/S2059798319011471>, PMID: 31588918
- Mastronarde DN**. 2005. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* **152**: 36–51. DOI: <https://doi.org/10.1016/j.jsb.2005.07.007>, PMID: 16182563
- McMahon LP**, Choi KM, Lin TA, Abraham RT, Lawrence JC. 2002. The rapamycin-binding domain governs substrate selectivity by the mammalian target of rapamycin. *Molecular and Cellular Biology* **22**: 7428–7438. DOI: <https://doi.org/10.1128/MCB.22.21.7428-7438.2002>, PMID: 12370290
- Melero R**, Uchiyama A, Castaño R, Kataoka N, Kurosawa H, Ohno S, Yamashita A, Llorca O. 2014. Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. *Structure* **22**: 1105–1119. DOI: <https://doi.org/10.1016/j.str.2014.05.015>, PMID: 25002321
- Mino T**, Iwai N, Endo M, Inoue K, Akaki K, Hia F, Uehata T, Emura T, Hidaka K, Suzuki Y, Standley DM, Okada-Hatakeyama M, Ohno S, Sugiyama H, Yamashita A, Takeuchi O. 2019. Translation-dependent unwinding of stem-loops by UPF1 licenses Regnase-1 to degrade inflammatory mRNAs. *Nucleic Acids Research* **47**: 8838–8859. DOI: <https://doi.org/10.1093/nar/gkz628>, PMID: 31329944
- Mordes DA**, Glick GG, Zhao R, Cortez D. 2008. TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. *Genes & Development* **22**: 1478–1489. DOI: <https://doi.org/10.1101/gad.1666208>, PMID: 18519640
- Moriarty NW**, Grosse-Kunstleve RW, Adams PD. 2009. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica. Section D, Biological Crystallography* **65**: 1074–1080. DOI: <https://doi.org/10.1107/S0907444909029436>, PMID: 19770504
- Nicholson P**, Josi C, Kurosawa H, Yamashita A, Mühlemann O. 2014. A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic Acids Research* **42**: 9217–9235. DOI: <https://doi.org/10.1093/nar/gku645>, PMID: 25053839
- Ohnishi T**, Yamashita A, Kashima I, Schell T, Anders KR, Grimson A, Hachiya T, Hentze MW, Anderson P, Ohno S. 2003. Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. *Molecular Cell* **12**: 1187–1200. DOI: [https://doi.org/10.1016/S1097-2765\(03\)00443-X](https://doi.org/10.1016/S1097-2765(03)00443-X), PMID: 14636577
- Okada-Katsuhata Y**, Yamashita A, Kutsuzawa K, Izumi N, Hirahara F, Ohno S. 2012. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Research* **40**: 1251–1266. DOI: <https://doi.org/10.1093/nar/gkr791>, PMID: 21965535
- Punjani A**, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**: 290–296. DOI: <https://doi.org/10.1038/nmeth.4169>, PMID: 28165473
- Rosenthal PB**, Henderson R. 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology* **333**: 721–745. DOI: <https://doi.org/10.1016/j.jmb.2003.07.013>, PMID: 14568533
- Scheres SHW**. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* **180**: 519–530. DOI: <https://doi.org/10.1016/j.jsb.2012.09.006>, PMID: 23000701
- Shaheen R**, Anazi S, Ben-Omran T, Seidahmed MZ, Caddle LB, Palmer K, Ali R, Alshidi T, Hagos S, Goodwin L, Hashem M, Wakil SM, Abouelhoda M, Colak D, Murray SA, Alkuraya FS. 2016. Mutations in SMG9, Encoding an Essential Component of Nonsense-Mediated Decay Machinery, Cause a Multiple Congenital Anomaly Syndrome in Humans and Mice. *American Journal of Human Genetics* **98**: 643–652. DOI: <https://doi.org/10.1016/j.ajhg.2016.02.010>, PMID: 27018474

- Tan YZ**, Baldwin PR, Davis JH, Williamson JR, Potter CS, Carragher B, Lyumkis D. 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods* **14**: 793–796. DOI: <https://doi.org/10.1038/nmeth.4347>, PMID: 28671674
- Xiao J**, Liu M, Qi Y, Chaban Y, Gao C, Pan B, Tian Y, Yu Z, Li J, Zhang P, Xu Y. 2019. Structural insights into the activation of ATM kinase. *Cell Research* **29**: 683–685. DOI: <https://doi.org/10.1038/s41422-019-0205-0>, PMID: 31320732
- Yamashita A**, Ohnishi T, Kashima I, Taya Y, Ohno S. 2001. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes & Development* **15**: 2215–2228. DOI: <https://doi.org/10.1101/gad.913001>, PMID: 11544179
- Yamashita A**, Izumi N, Kashima I, Ohnishi T, Saari B, Katsuhata Y, Muramatsu R, Morita T, Iwamatsu A, Hachiya T, Kurata R, Hirano H, Anderson P, Ohno S. 2009. SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. *Genes & Development* **23**: 1091–1105. DOI: <https://doi.org/10.1101/gad.1767209>, PMID: 19417104
- Yamashita A**. 2013. Role of SMG-1-mediated Upf1 phosphorylation in mammalian nonsense-mediated mRNA decay. *Genes to Cells* **18**: 161–175. DOI: <https://doi.org/10.1111/gtc.12033>, PMID: 23356578
- Yang H**, Rudge DG, Koos JD, Vaidialingam B, Yang HJ, Pavletich NP. 2013. mTOR kinase structure, mechanism and regulation. *Nature* **497**: 217–223. DOI: <https://doi.org/10.1038/nature12122>, PMID: 23636326
- Yates LA**, Williams RM, Hailemariam S, Ayala R, Burgers P, Zhang X. 2020. Cryo-EM Structure of Nucleotide-Bound Tel1(ATM) Unravels the Molecular Basis of Inhibition and Structural Rationale for Disease-Associated Mutations. *Structure* **28**: 96–104. DOI: <https://doi.org/10.1016/j.str.2019.10.012>
- Zhang Y-J**, Duan Y, Zheng XFS. 2011. Targeting the mTOR kinase domain: the second generation of mTOR inhibitors. *Drug Discovery Today* **16**: 325–331. DOI: <https://doi.org/10.1016/j.drudis.2011.02.008>, PMID: 21333749
- Zhu L**, Li L, Qi Y, Yu Z, Xu Y. 2019. Cryo-EM structure of SMG1-SMG8-SMG9 complex. *Cell Research* **29**: 1027–1034. DOI: <https://doi.org/10.1038/s41422-019-0255-3>, PMID: 31729466
- Zivanov J**, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**: e42166. DOI: <https://doi.org/10.7554/eLife.42166>, PMID: 30412051



Figure supplements

Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8

Lukas M Langer *et al*

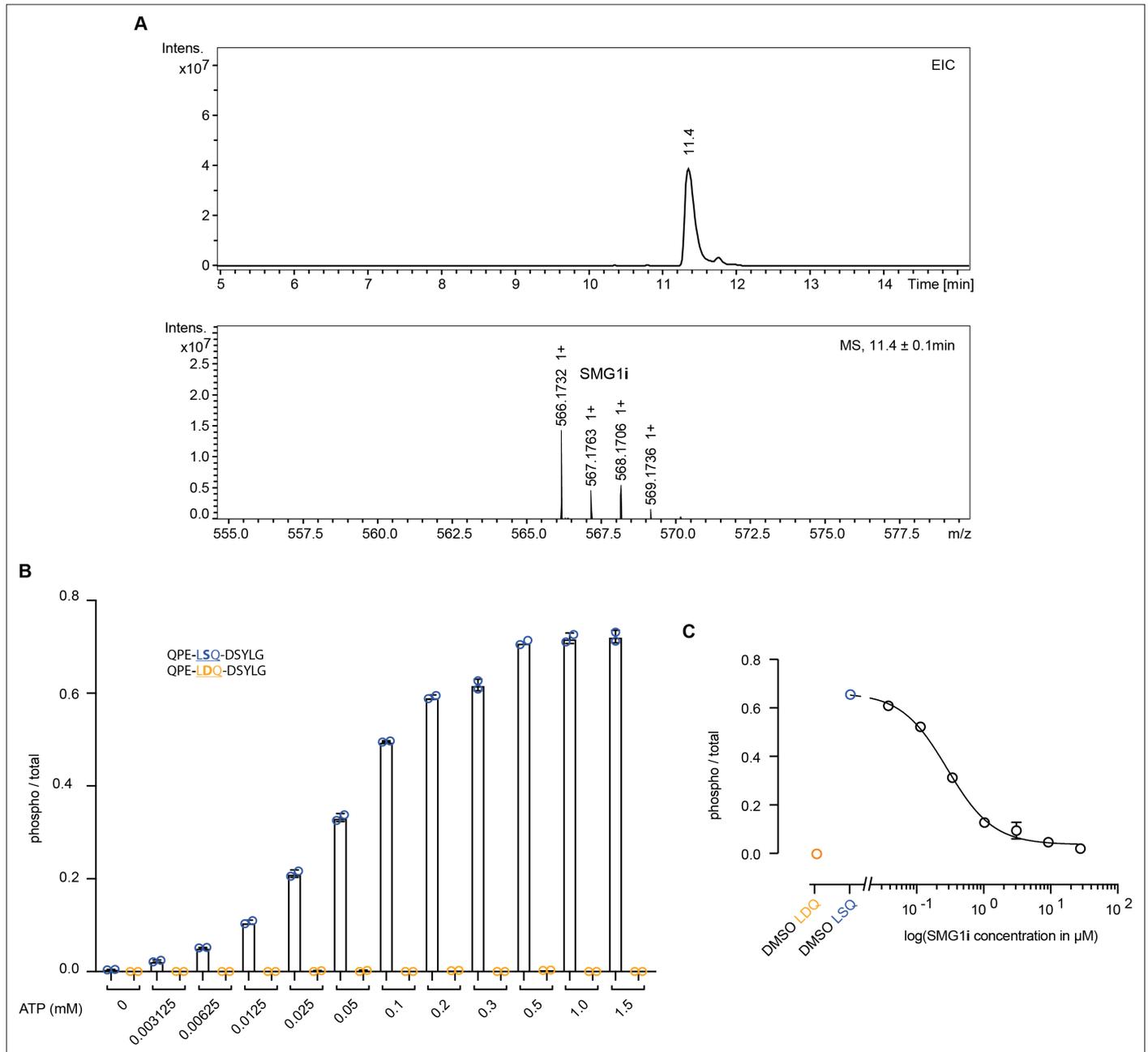


Figure 1—figure supplement 1. Characterization of SMG1 inhibitor. **(A)** Liquid chromatography-mass spectrometry (LC-MS) experiment with the SMG1 inhibitor sample used throughout this study. The expected mass for SMG1i is 566.13 Da. Differences of +1 are caused by the varying composition of naturally occurring carbon isotopes. **(B)** Titration of ATP requirement in mass spectrometry-based phosphorylation assay using 500 nM SMG1-8-9 and the indicated UPF1-derived peptides as substrates. 0.75 mM ATP was chosen for later experiments. **(C)** Dose-response curve for SMG1i in mass spectrometry-based phosphorylation assay (using triplicate data points shown in **Figure 1B**). Phosphorylation ratios of substrate (blue) and control peptide (orange) in the absence of SMG1i are shown as mean values with error bars indicating standard deviation (were large enough to visualize). The control peptide was included for reference but was not used for fitting.

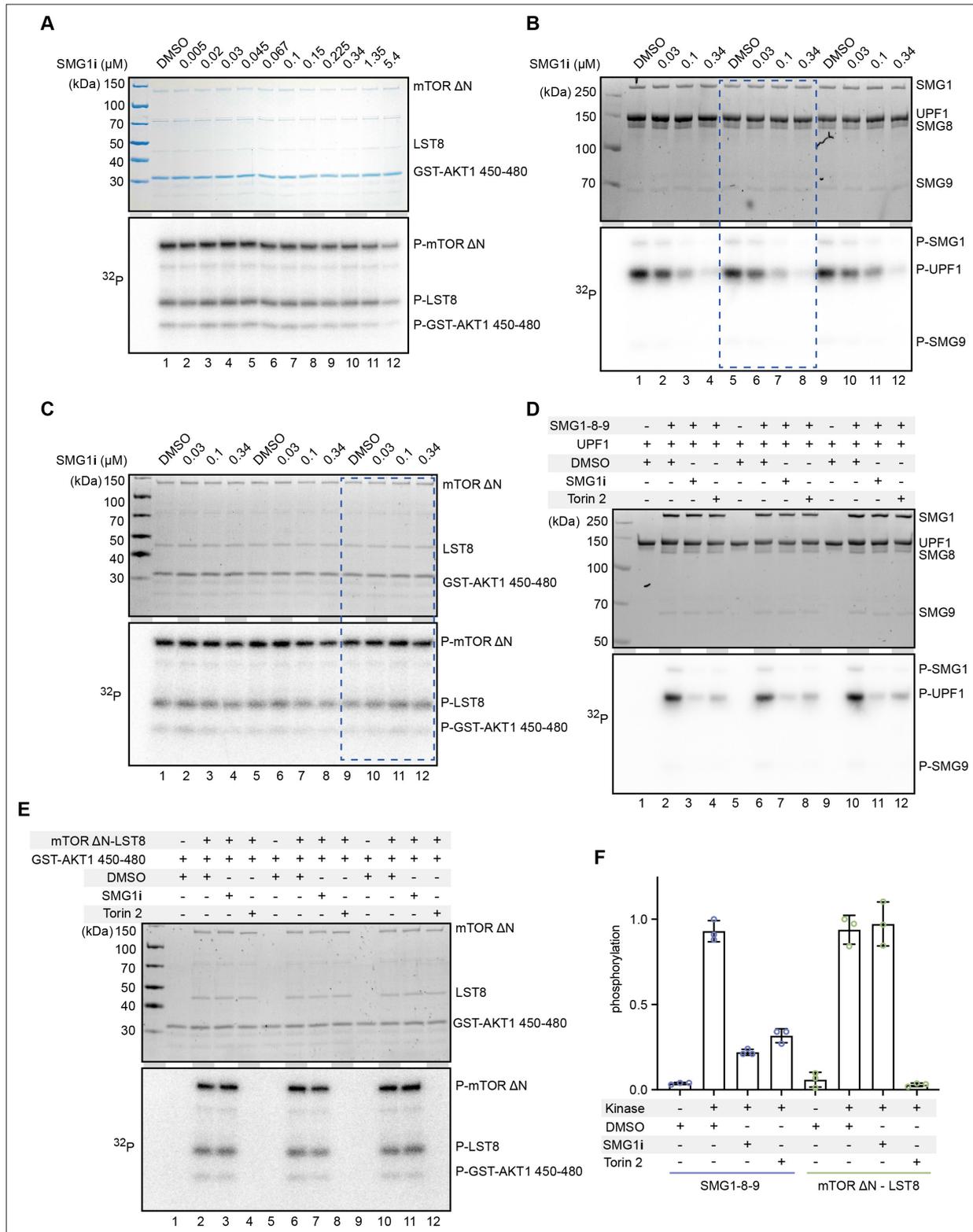


Figure 1—figure supplement 2. Radioactivity-based phosphorylation assays using SMG1 inhibitor with SMG1 and mTOR. **(A)** Titration of SMG1i using 100 nM mTOR-LST8 and GST-AKT1 as a substrate. The Coomassie-stained gel is shown on top and the radioactive signal on the bottom. **(B, C)** SMG1-8-9 or mTOR-LST8 kinase assay with four selected concentrations of SMG1i. Assays were performed in triplicates and used for quantification. The blue rectangle indicates the replicate of the assays shown in **Figure 1**. **(D, E)** SMG1-8-9 or mTOR-LST8 kinase assay supplemented with either DMSO, SMG1i, **Figure 1—figure supplement 2 continued on next page**

Figure 1—figure supplement 2 continued

or Torin 2. A control without kinase was included. The highest concentration of SMG1i (0.34 μM) was used in panels **(B, C)** and the same amount of Torin 2 was used. Both assays were performed in triplicates. **(F)** Quantification of normalized UPF1 or mTOR (auto-)phosphorylation of the experimental triplicates shown in panels **(D, E)**.

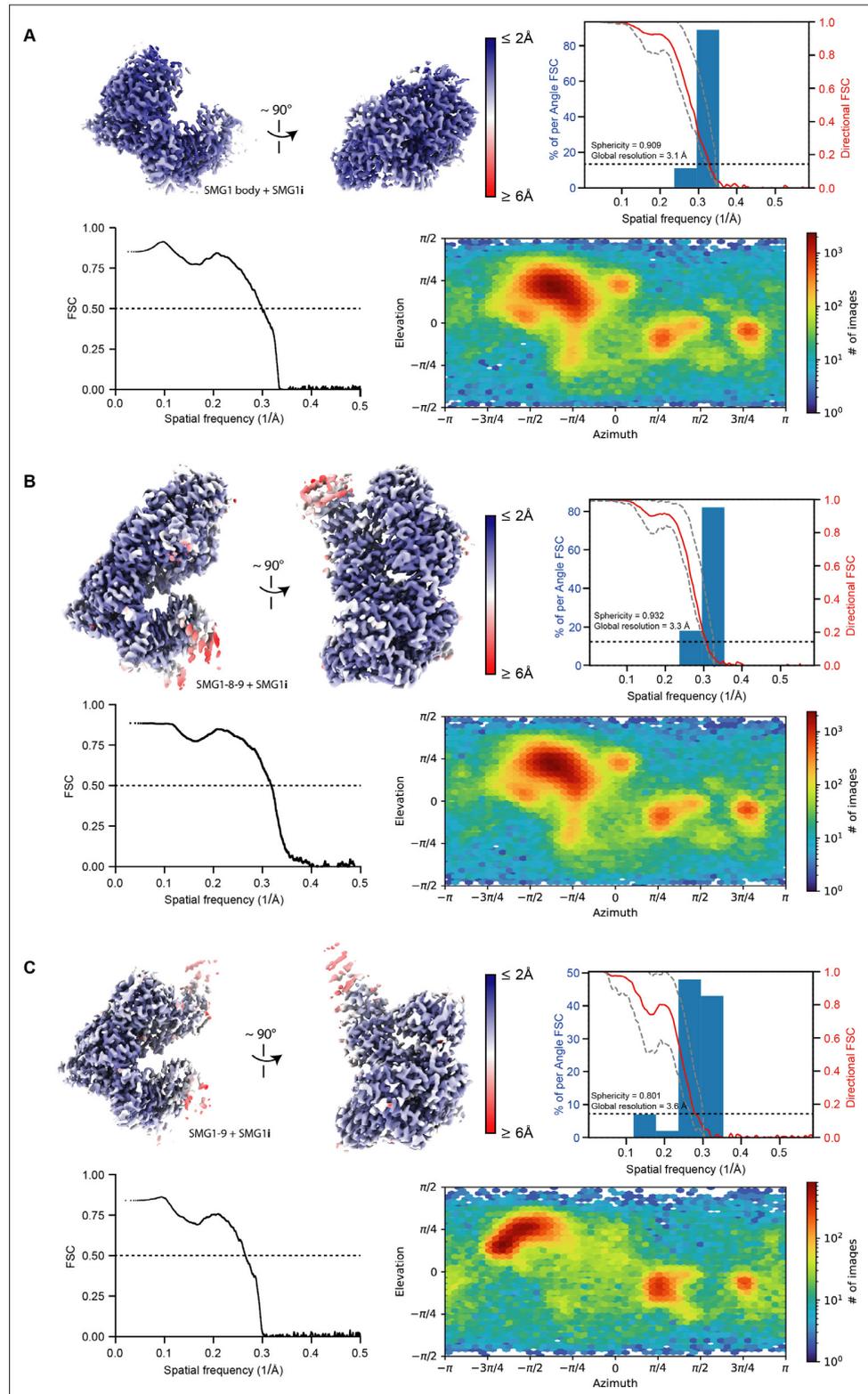


Figure 2—figure supplement 1. Resolution distribution and isotropy of SMG1-centered cryo-EM maps bound to SMG1i. SMG1i-bound reconstructions used in this study for model building are colored according to estimated local resolution shown in two different orientations. A three-dimensional Fourier shell correlation (FSC) plot is included for each reconstruction (Tan et al., 2017). The red line represents the estimated global masked half

Figure 2—figure supplement 1 continued on next page

Figure 2—figure supplement 1 continued

map FSC. The resolutions according to the gold standard FSC cutoff of 0.143 are indicated and shown as a black dashed line (**Rosenthal and Henderson, 2003**). The spread of directional resolution values is defined as $\pm 1\sigma$ (dashed gray lines). Overall isotropy of the maps is indicated by the given sphericity values (out of 1). In the bottom row of each panel, a model versus map FSC is shown alongside a plot visualizing the distribution of particle views. **(A)** SMG1 body bound to SMG1i after focused refinement (EMD-13676, PDB identifier: 7PW6). **(B)** SMG1-8-9 bound to SMG1i (EMD-13674, PDB identifier: 7PW4). **(C)** SMG1-9 bound to SMG1i (EMD-13677, PDB identifier: 7PW7). cryo-EM, cryo-electron microscopy; FSC, XXX.

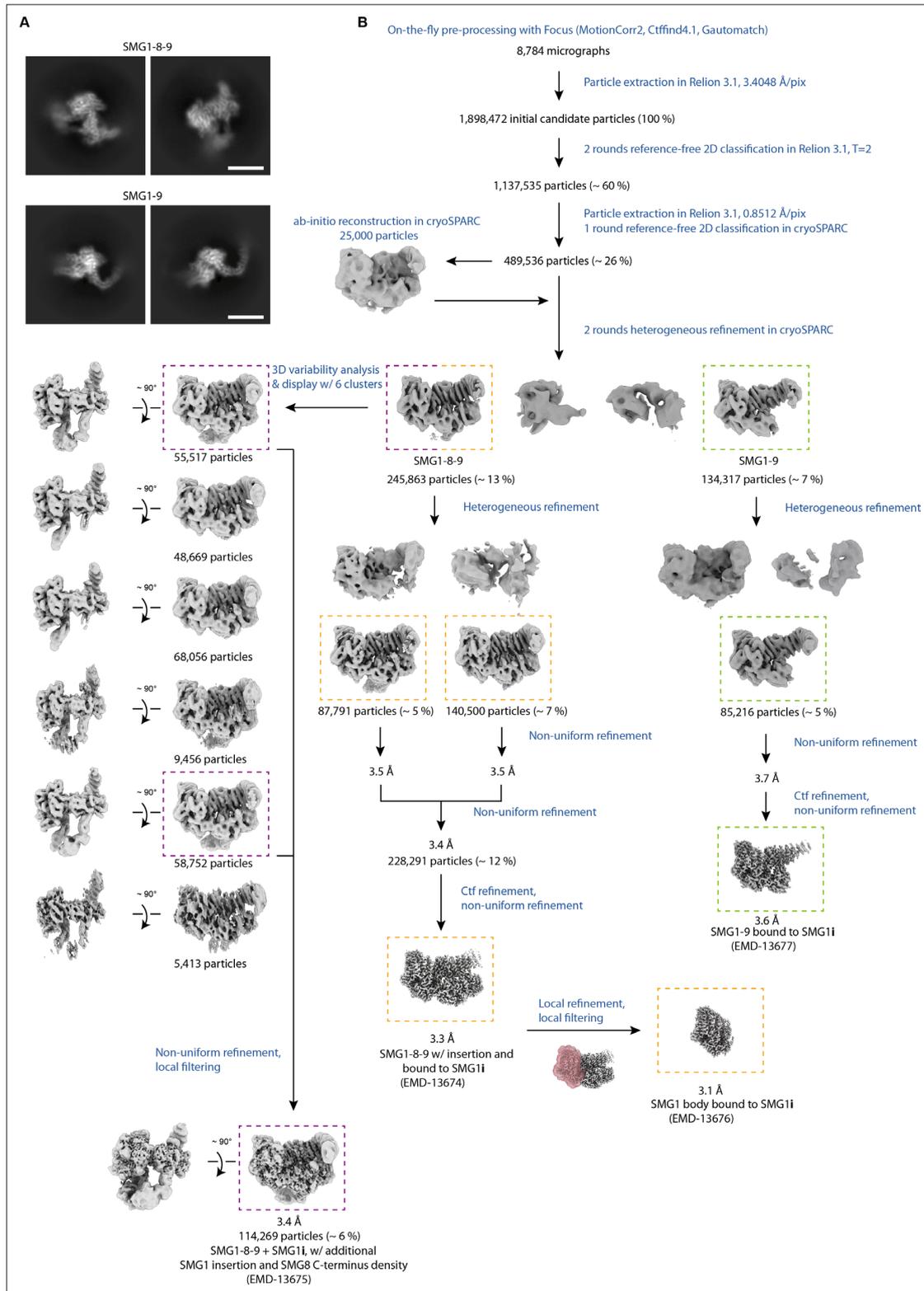


Figure 2—figure supplement 2. Cryo-EM data processing of SMG1i data set. **(A)** 2D class averages of SMG1-8-9 and SMG1-9 are calculated from the final particle stacks. Scale bars ≈ 100 Å. **(B)** Processing scheme. Processing steps are indicated in blue; particle numbers and percentages with respect to initial candidate particles are shown for relevant classes. Colored, dashed rectangles indicate the different final reconstructions and the respective classes obtained from the data set collected in the presence of SMG1i. cryo-EM, cryo-electron microscopy.

Figure 2—figure supplement 3 continued

magenta. **(D)** Superposition of SMG1i-bound active site with AMPPNP-bound active site (PDB identifier: 6Z3R). Same view as in **(C)**. **(E)** Superposition of SMG1i-bound active site with Torin 2-bound mTOR active site (PDB identifier: 4JXP). Similar view as in **(C)**. **(F)** Superposition of SMG1i-bound SMG1 with DNA-PK active site in the inactive conformation (PDB identifier: 7K11). Figure prepared and labeled as in **Figure 2C**, with DNA-PK residues colored in red. The asterisk indicates a residue only visible in the DNA-PK structure due to structural rearrangements. **(G)** Superposition of SMG1i-bound SMG1 with ATR active site (PDB identifier: 5YZO). Figure prepared and labeled as in **Figure 2C**, with ATR residues colored in orange. An asterisk indicates corresponding residues in SMG1 and ATR that are separated due to conformational divergence. cryo-EM, cryo-electron microscopy.

Figure 3—figure supplement 1 continued

active site of the SMG1i-bound SMG1-8-9 complex connects to the modeled N-terminus of the SMG1 insertion domain. The left panel highlights the last modeled N- and C-terminal residues of the insertion domain. The right panel shows the same view, superimposed with the isolated extra density. (C) Multiple sequence alignment of the N-terminal 100 residues of the SMG1 insertion domain and the PRDs of the other human PIKK family members colored by identity. The beginning of the unmodeled part of the SMG1 insertion is indicated by a black arrow (compare B). EM, electron microscopy.

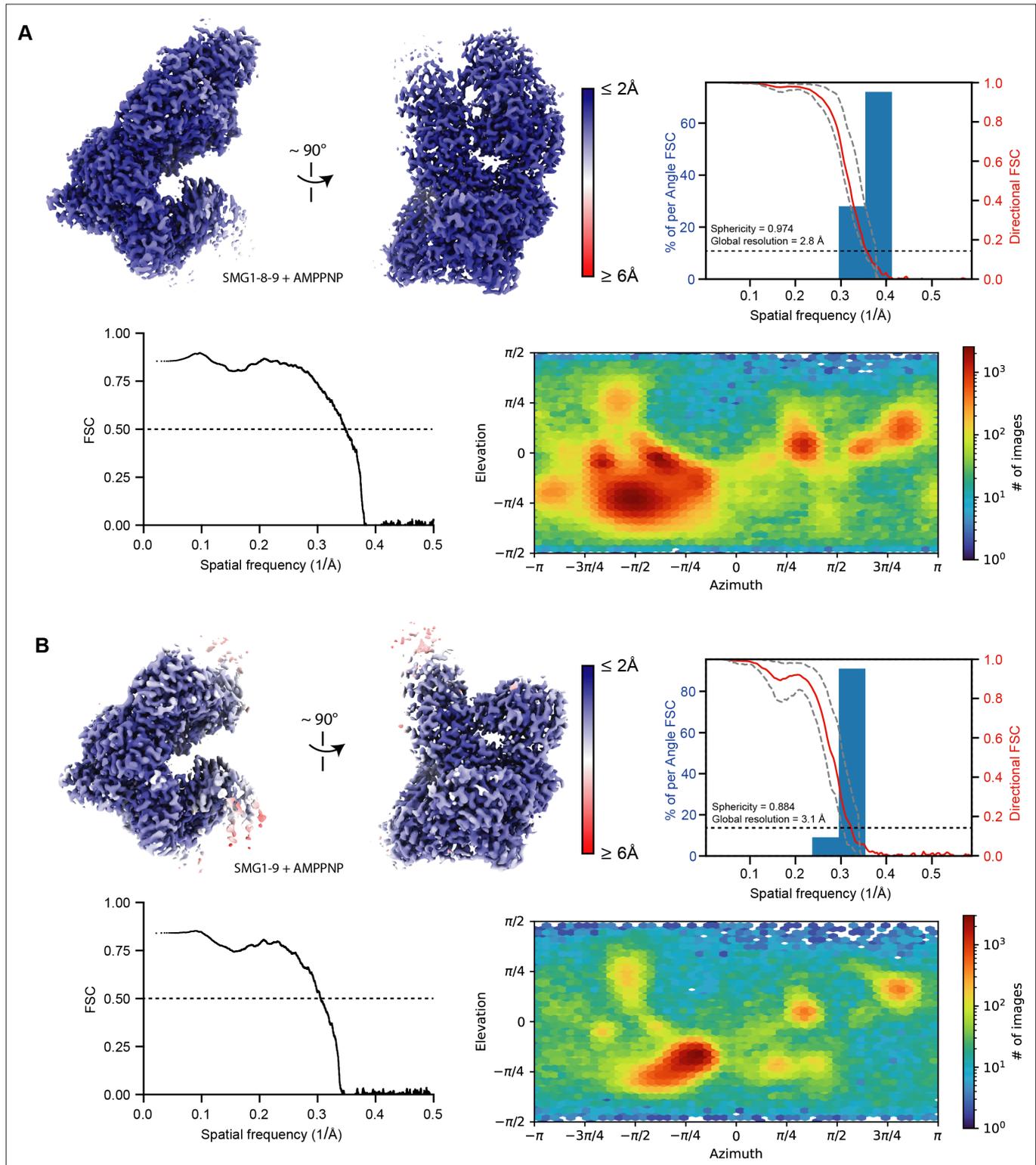


Figure 3—figure supplement 2. Resolution distribution and isotropy of SMG1-centered cryo-EM maps bound to AMPPNP. AMPPNP-bound reconstructions used in this study for model building are colored according to estimated local resolution shown in two different orientations. A three-dimensional FSC plot is included for each reconstruction (Tan et al., 2017). The red line represents the estimated global masked half map FSC. The resolutions according to the gold standard FSC cutoff of 0.143 are indicated and shown as a black dashed line (Rosenthal and Henderson, 2003).
Figure 3—figure supplement 2 continued on next page

Figure 3—figure supplement 2 continued

spread of directional resolution values is defined as $\pm 1\sigma$ (dashed gray lines). Overall isotropy of the maps is indicated by the given sphericity values (out of 1). In the bottom row of each panel, a model versus map FSC is shown alongside a plot visualizing the distribution of particle views. **(A)** SMG1-8-9 bound to AMPPNP (EMD-13678, PDB identifier: 7PW8). **(B)** SMG1-9 bound to AMPPNP (EMD-13679, PDB identifier: 7PW9). cryo-EM, cryo-electron microscopy; FSC, XXX.

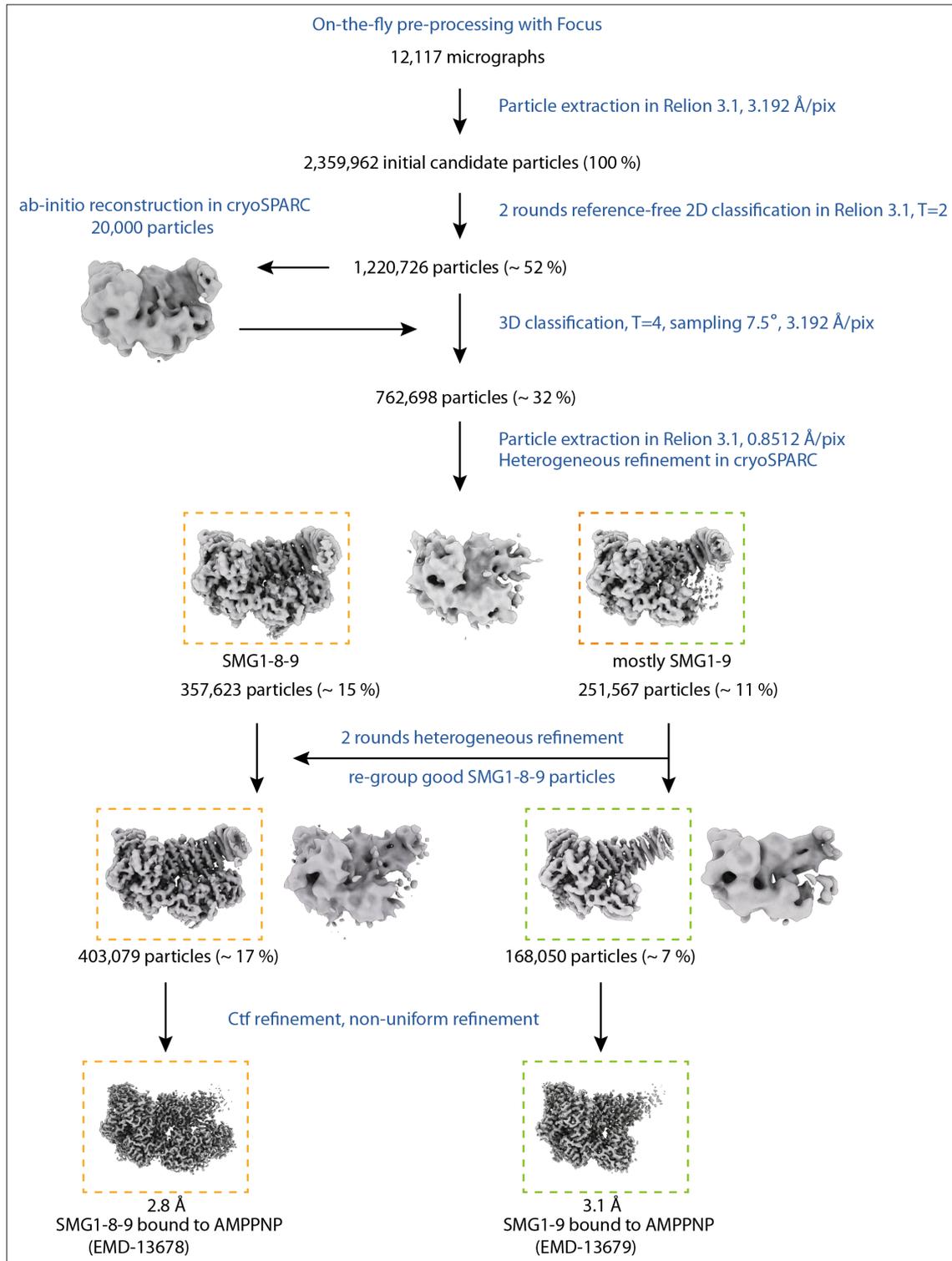


Figure 3—figure supplement 3. Cryo-EM data processing of AMPPNP data set. Processing steps are indicated in blue; particle numbers and percentages with respect to initial candidate particles are shown for relevant classes. Colored, dashed rectangles indicate the different final reconstructions and the respective classes obtained from the data set collected in the presence of AMPPNP. cryo-EM, cryo-electron microscopy.

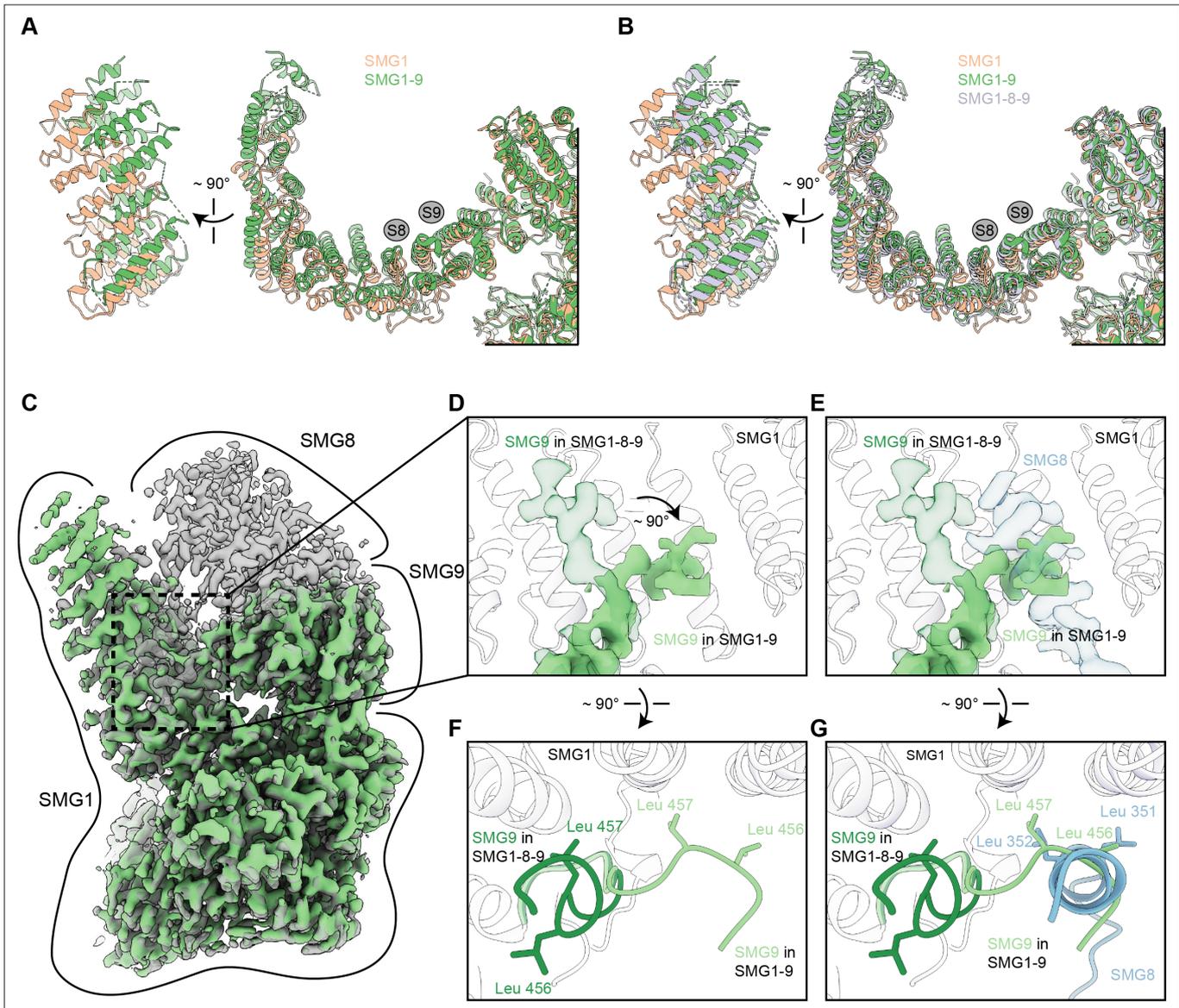


Figure 3—figure supplement 4. Details of the SMG1-9 complex. **(A)** Overlay of SMG1 (PDB identifier: 6L53) and SMG1-9 detailing movements of the N-terminal HEAT repeats. A front and a side view are shown and binding sites for SMG8 and SMG9 are indicated by gray circles. **(B)** as in **(A)**, but including SMG1-8-9 (PDB identifier: 6Z3R). Models of SMG8 and SMG9 were occluded. **(C)** Overlay of densities for SMG1-9 and SMG1-8-9 complexes, with SMG1-9 density in light green and SMG1-8-9 density in transparent gray. Approximate location of the single proteins within the densities are indicated. **(D)** Close-up showing the rearrangement of an SMG9 segment in the SMG1-9 complex compared to the SMG1-8-9 complex. Density of the segment in SMG1-9 is shown in light green, superimposed with the SMG9 segment in SMG1-8-9 displayed in transparent dark green. The model for the interacting region of the SMG1 arch is shown as a transparent cartoon, and SMG8 is not shown. **(E)** same as **(D)**, but with density for an SMG1-interacting SMG8 segment in the SMG1-8-9 complex shown in transparent blue, that would clash with the SMG9 segment conformation observed in the SMG1-9 complex. **(F)** Overlay of the model of the SMG9 segment in the SMG1-9 (light green) and the SMG1-8-9 (dark green) complex. Two Leu residues undergoing rearrangement between the two complexes are shown. Position of SMG1 is indicated, SMG8 is not shown. **(G)** same as in **(F)**, but shown with the SMG1-interacting SMG8 segment (blue). In the SMG1-9 complex, the highlighted pair of Leu residues in SMG9 substitutes for a pair of SMG8 Leu residues on the SMG1 arch.

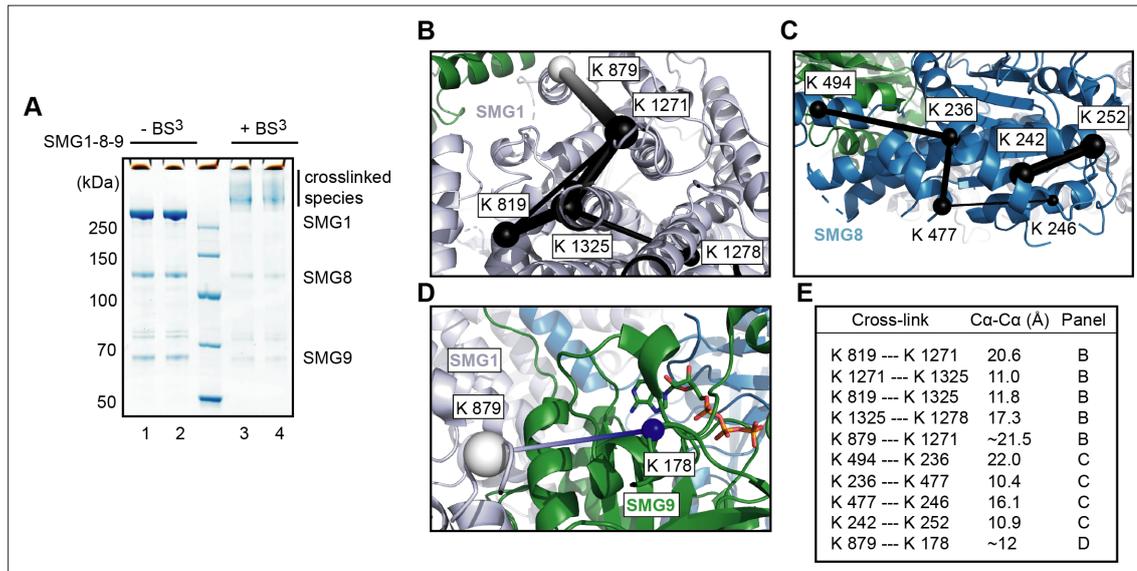


Figure 4—figure supplement 1. Cross-linking mass spectrometry of SMG1-8-9. **(A)** Two samples of SMG1-8-9 (lanes 1 and 2) were incubated with BS³ (lanes 3 and 4) and analyzed using SDS-PAGE and Coomassie staining. **(B)** Exemplary intra cross-links detected for SMG1 mapped on the model of the apo complex (PDB identifier: 6SYT). Intra cross-links are shown in black with the thickness of the line indicating their score (thicker=higher score), and the respective residues indicated as spheres. A cross-link to an unmodeled region is shown in white (placed in the middle of the segment whose ends are the closest visible C α atoms). **(C)** same as **(B)**, but for SMG8. **(D)** Inter cross-link between SMG1 and SMG9. The Mg²⁺-ion complexed with SMG9 was omitted. **(E)** Table listing measured distances for cross-links visualized in panels **(B–D)**.

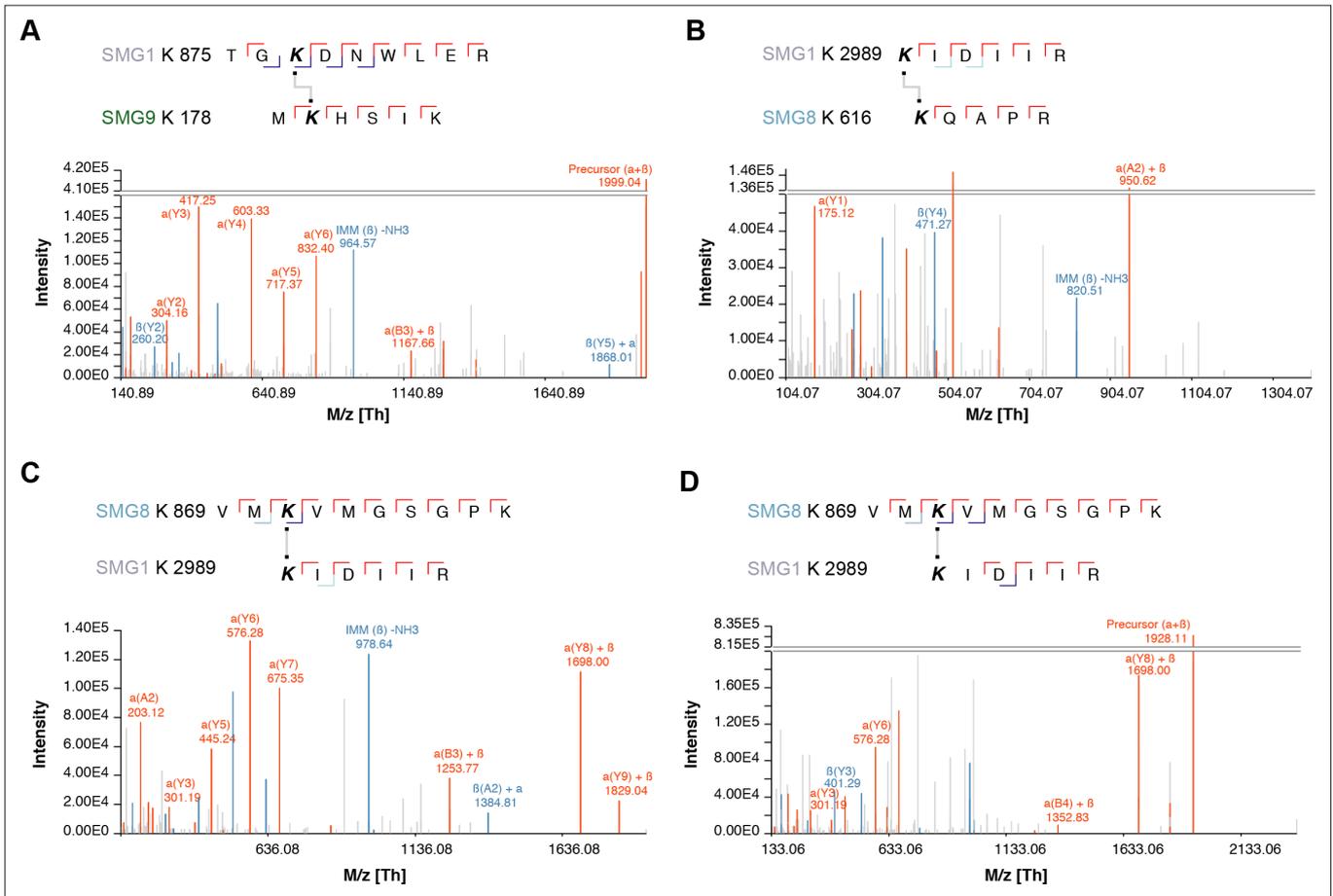


Figure 4—figure supplement 2. Selected spectra of detected SMG1-8-9 intra-links. (A–D) Cross-links are shown above each panel. All spectra showed good sequence coverage with full y-ion series, many b-ions, and highly specific fragments.

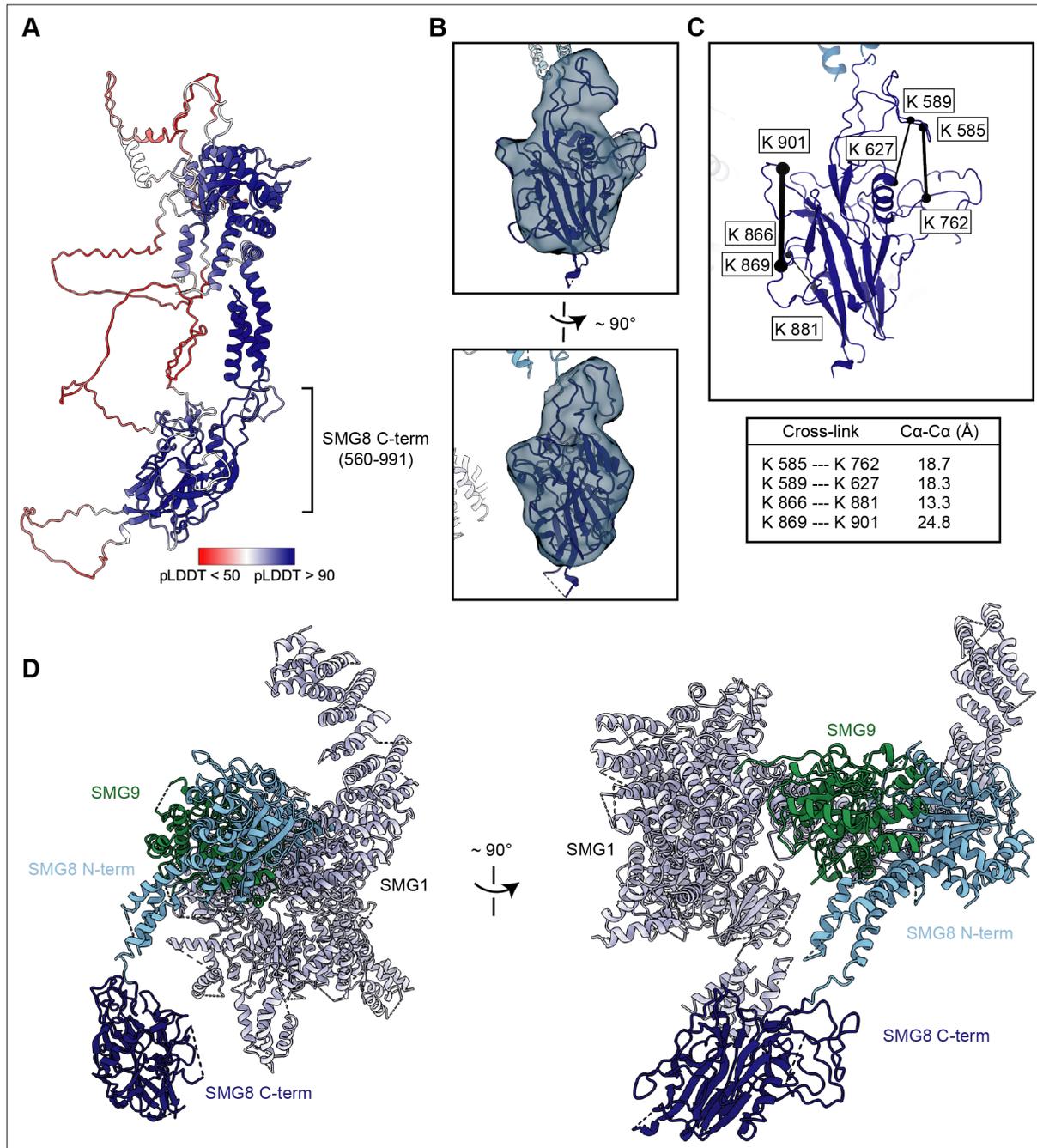


Figure 4—figure supplement 3. Integration of cryo-EM, cross-linking MS, and AlphaFold data reveals a model for the SMG8 C-terminus. **(A)** AlphaFold model of full-length SMG8 (UniProt: Q8ND04) colored by per-residue confidence score (pLDDT) (*Jumper et al., 2021*). The previously unmodeled SMG8 C-terminus is indicated. **(B)** AlphaFold model of the SMG8 C-terminus (dark blue) rigid-body fitted into the isolated cryo-EM density and fused to the modeled N-terminal part of SMG8 (light blue), shown in two related views. Regions with a pLDDT < 50 were deleted from the model. **(C)** Intra cross-links of the SMG8 C-terminus mapped on the model, as described in *Figure 4—figure supplement 1*. A table shows the corresponding measured distances. **(D)** New model of the SMG1-8-9 complex, now including the SMG8 C-terminal region, shown from two different perspectives (PDB identifier: 7PW5). cryo-EM, cryo-electron microscopy.

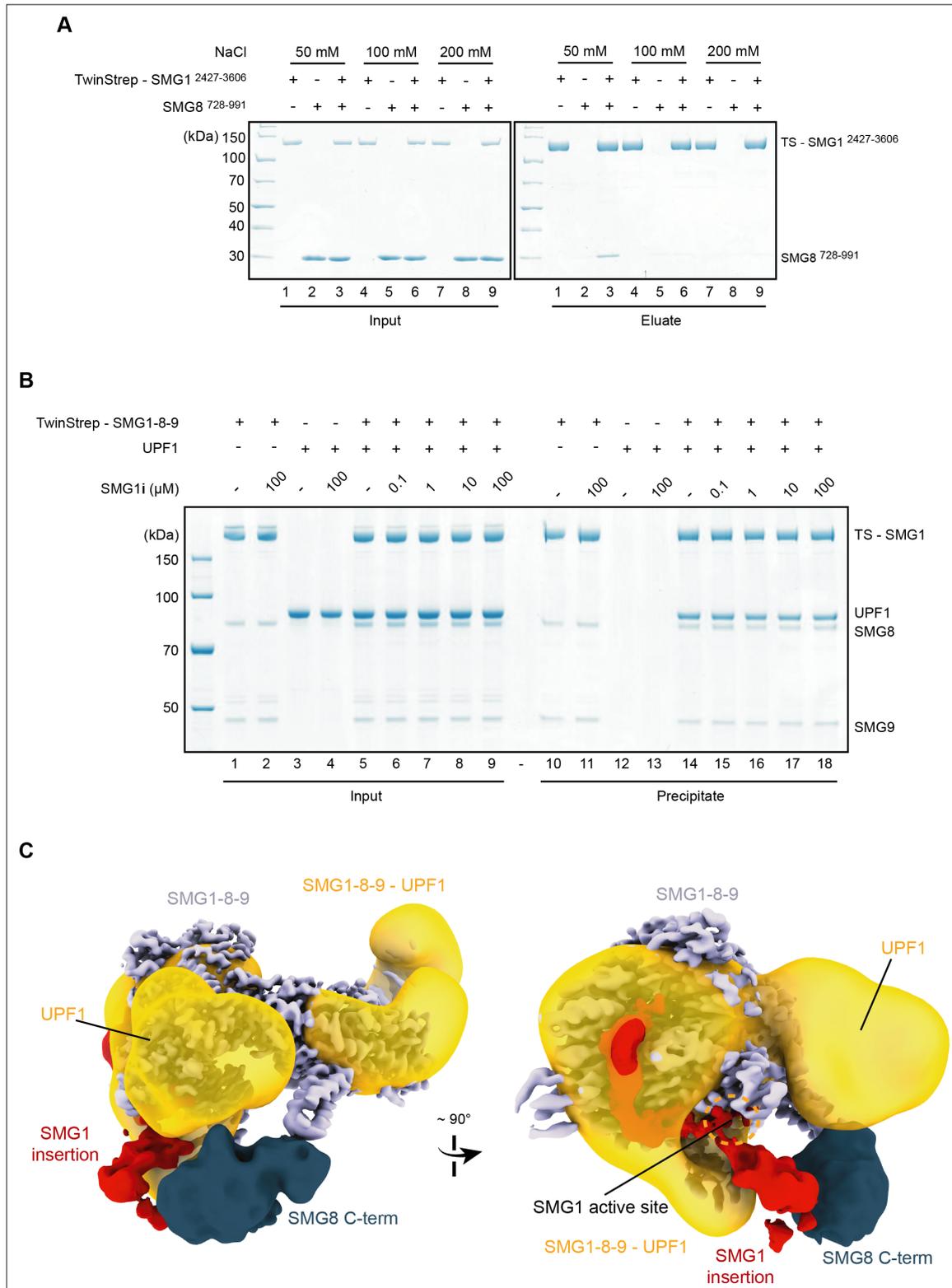


Figure 4—figure supplement 4. Further characterization of SMG1-8-9—centered interactions. **(A)** Coomassie-stained SDS-PAGE analysis of pull-down experiment showing that the interaction between SMG1 insertion domain (SMG1²⁴²⁷⁻³⁶⁰⁶) and SMG8 C-terminus (SMG8⁷²⁸⁻⁹⁹¹) is dependent on low-salt conditions. **(B)** Coomassie-stained SDS-PAGE analysis of pull-down experiment showing that SMG1-8-9—SMG1i complex formation does not prevent interaction with UPF1. 0.12 μ M of TS-SMG1-8-9 and 0.4 μ M of UPF1 were used throughout. **(C)** The resolution-filtered, segmented cryo-EM density of Figure 4—figure supplement 4 continued on next page

Figure 4—figure supplement 4 continued

SMG1i-bound autoinhibited SMG1-8-9 complex fitted in a negative-stain reconstruction of a cross-linked SMG1-8-9—UPF1 complex (EMD-2664) shown in orientations similar to **Figure 4A**. The density within the negative-stain reconstruction assigned to UPF1 is indicated.

4.3 Reconstitution of stalled ribosome translation termination complexes

NMD is thought to be elicited by the recognition of a premature translation termination event by NMD factors SMG1-8-9, UPF1 and possibly others (Karousis & Mühlemann, 2019). This part of the thesis presents unpublished, newly established tools facilitating the reconstitution and purification of four distinct ribosome translation termination complexes.

According to the prevailing NMD model, the discrimination between a premature and a canonical translation termination event is facilitated by the interaction of NMD factors with a terminating ribosome, which in turn is enabled through the correct mRNP environment, for example a downstream EJC. How NMD factors directly contact the translation termination machinery remains unclear (Kashima *et al.*, 2006; Karousis & Mühlemann, 2019). To facilitate *in vitro* reconstitutions of specific NMD-initiating complexes, this work aimed to establish protocols for preparing defined ribosome translation termination complexes. The employed strategy is based on a rabbit reticulocyte cell-free translation system programmed with *in vitro* transcribed RNA and purified mutants of eukaryotic release factors to stall ribosome - nascent chain (RNC) translation termination complexes (Feng & Shao, 2018). These RNC complexes can be detected using immunoblotting against an affinity tag present in the N-terminus of the nascent chain (NC) and/or enriched by affinity purification (Brown *et al.*, 2015; Shao *et al.*, 2016). The RNA used in the assays discussed here has been previously described (Brown *et al.*, 2015) and encodes for a nascent chain with an N-terminal 3xFLAG affinity tag (see Figure 4.1 A). Translation of this RNA can be monitored by immunoblotting, where the appearance of a ~17 kDa band indicates NC released from ribosomes (see Figure 4.1 B, lane 2). Addition of a recombinant, purified, well established GGQ-to-AAQ motif mutant of eRF1 incapable of hydrolyzing the tRNA-NC ester bond leads to a ~25 kDa shift of the NC signal, indicative of a tRNA-NC adduct stalled on a eRF1-ABCE1-80S complex (see Figure 4.1 B, lane 3) (Frolova *et al.*, 1999; Brown *et al.*, 2015). As a control, addition of recombinant, purified wildtype eRF1 does not lead to entrapment of the NC on the ribosome (see Figure 4.1 B, lane 4). To stall translation termination at an earlier stage, namely upon eRF1-eRF3a binding to a ribosome (see Figure 2.4), an eRF3a GTPase mutant (His299 to Leu299) previously described to efficiently stall translation termination in yeast was used (Salas-Marco & Bedwell, 2004; Shao *et al.*, 2016). Within the same experimental setup as described for the eRF1 mutant,

the eRF3a mutant was indeed able to stall translation termination in the *in vitro* translation system as indicated by the appearance of a tRNA-NC adduct (see Figure 4.1 C). Purification of RNC complexes is commonly hampered by the presence of free NC, which effectively competes with the much larger RNC complexes for binding to the beads (see Figure 4.1 A and D, lanes 4 to 6) (Feng & Shao, 2018). This problem can be mitigated by the use of excessive amounts of affinity beads or by separating free NC from ribosome-bound NC by pelleting the entire ribosome population (data not shown). Still, pelleting and resuspension puts considerable stress on the stalled ribosomes and bound factors, while also reducing the yield and increasing preparation time (Feng & Shao, 2018). Based on the observation that in a one-step pull down the porous beads mainly capture free NC (see Figure 4.1 D, lanes 1 to 6) whereas non-porous beads also capture RNC complexes (see Figure 4.1 D, lanes 7 to 12), a two-step pull down protocol exploiting the size selection displayed by the used magnetic beads was established (see Figure 4.1 E). This new two-step purification overcomes the aforementioned limitations. In the new purification scheme, the translation reaction is combined with a titrated amount of porous Agarose beads which preferentially bind NC, thereby removing it effectively from the sample (see Figure 4.1 F Agarose-FT). In a second step, the flow-through of the first step is then combined with non-porous ProteinG-Dynabeads, which have been pre-bound to an anti-FLAG antibody. The non-porous beads now efficiently capture RNC complexes with yields responsive to the amounts of beads used, with a significant reduction of free NC and a strong increase in bound RNC complexes as compared to a sample not subjected to the first step pull down (see Figure 4.1 F Dyna-E, sample 1 to 4). Furthermore, analysis of the purified eRF3a-stalled ribosome-nascent chain complexes using western blot analysis revealed the presence of endogenous eRF1 in the purified complexes (see Figure 4.1 F, lane 12). This is in agreement with a model in which the GTPase-inactive eRF3a mutant can still bind terminating ribosomes as part of an eRF1-eRF3a complex, trapping endogenous eRF1 in its unaccommodated inactive conformation, thereby preventing nascent chain release (compare Figure 2.4).

4.3 Reconstitution of stalled ribosome translation termination complexes 105

Building on the results described above, the size-selective double pull down protocol was further modified to also allow for purification of two additional, different translation termination complexes: first, stalled RNC complexes with a stop codon in the A-site and without further factors bound. Second, stalled eRF1-RNC complexes with eRF1 in its accommodated conformation, but without eRF3a or ABCE1 (compare Figure 2.4). To this end, high salt washes were used on beads-bound RNC complexes stalled with eRF1 or eRF3a mutants, either for all wash steps or only after binding to anti-FLAG coated ProteinG-Dynabeads (second step, see Figure 4.2 A). Immunoblotting was used to monitor the success of the respective pull downs (see Figure 4.2 B): anti-FLAG was used to monitor enrichment of tRNA-nascent chain adducts. Anti-RPS24 was used as a loading control and to monitor potential loss of 40S subunits. eRF3a could only be detected in eRF3a-stalled RNC complexes, which also carried eRF1 (see Figure 4.2 B, lane 1). Washing with high salt throughout or after the second binding step removed both eRF1 and eRF3a, resulting in factor-free RNC complexes with a stop codon in the A-site (see Figure 4.2 B, lanes 2 and 3). As described above, eRF1-stalled RNC complexes also captured ABCE1 (see Figure 4.2 B, lane 4). Washing eRF1-ABCE1-RNCs with high salt throughout mainly yielded factor-free RNCs with a stop codon in the A-site, with small amounts of eRF1 still present (see Figure 4.2 B, lane 5). However, washing eRF1-ABCE1-RNC complexes with high salt only after the second binding step removed ABCE1 but not eRF1 (see Figure 4.2 B, lane 6). In comparison to the eRF3a-stalled RNC complexes, eRF1 is more stably associated with the ribosomes in these RNC complexes because it already underwent the conformational change accommodating the mutated GGQ motif in the ribosome active site (see section 2.2.2). The resulting eRF1-RNC complexes thus represent a translation termination complex after eRF3a release and eRF1 accommodation, but before ABCE1 binding (see Figure 4.2 B). Cryo-EM screening data sets confirmed the successful programming of an eRF1-eRF3a-RNC complex using the eRF3a mutant, as well as of an eRF1-ABCE1-RNC complex using the eRF1 mutant (see Figure 4.2 C). Collectively, using release factor mutants, the described protocol allows rapid reconstitution of major intermediates along the translation termination pathway: a translating 80S encountering a stop codon before release factor binding, an eRF1-eRF3a-80S complex upon stop codon recognition (eRF1 not accommodated), an eRF1-80S complex poised for nascent chain release after eRF3a GTP hydrolysis but before ABCE1 binding (eRF1 accommodated) and an eRF1-ABCE1-80S complex poised for ribosomal subunit recycling (see Figure 4.2 D).

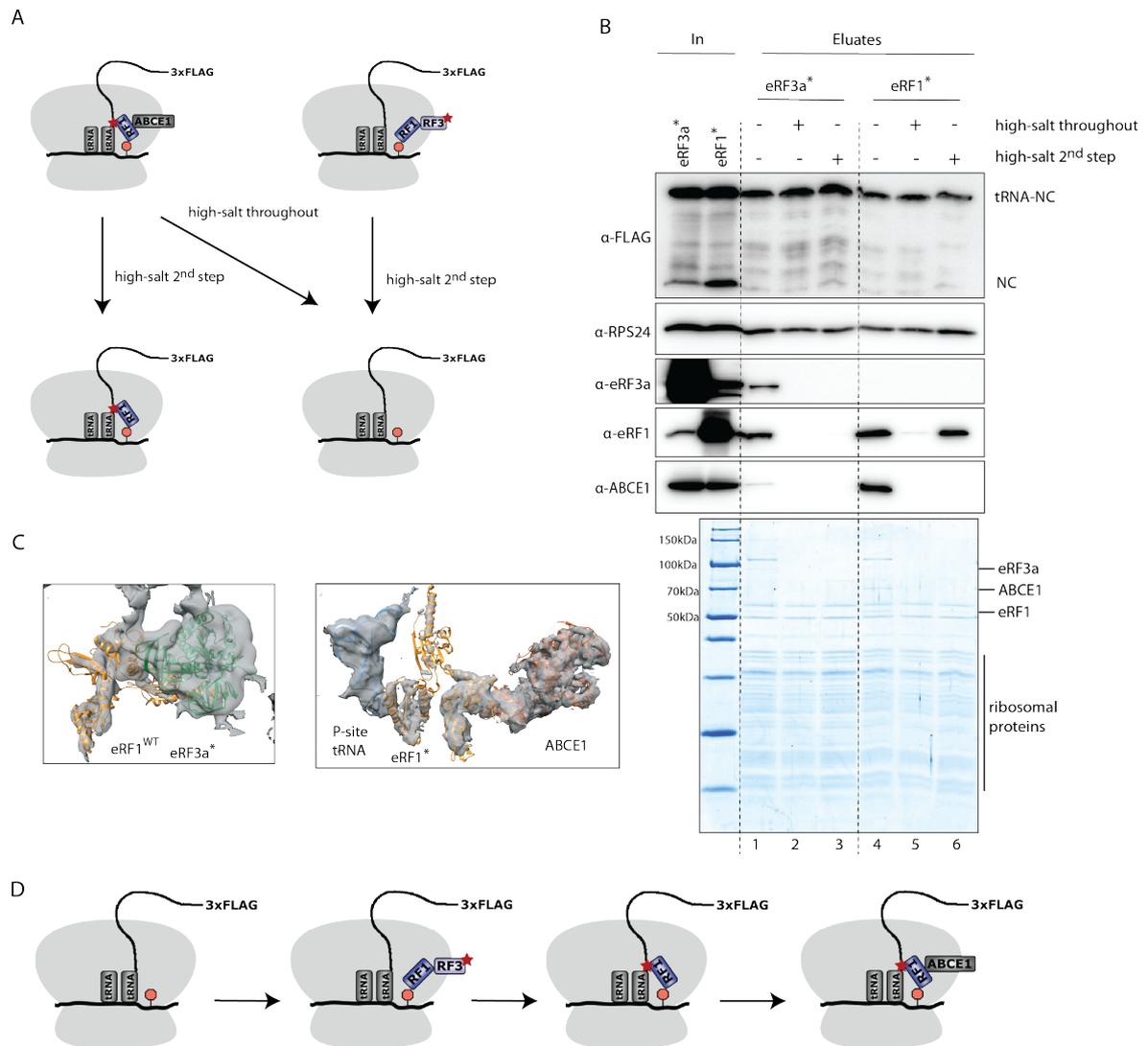


Figure 4.2: Size-selective double pull down to enrich ribosome-nascent chain translation termination complexes. Asterisks indicate mutant proteins. A. Scheme illustrating modifications of double pull down washing strategy to obtain different complex compositions. B. Top: Western blots analyzing translation termination complexes prepared as illustrated in A. B. Bottom: Coomassie-stained SDS-PAGE showing the corresponding protein contents. C. Cryo-EM densities obtained for eRF1-eRF3a- and eRF1-ABCE1-RNC complexes essentially prepared as described above. Densities located in the ribosomal intersubunit space are shown with the corresponding models fitted. (PDB left: 5LZT; PDB right: 3JAH.) D. The presented protocol allows to reconstitute major intermediates of eukaryotic translation termination (compare panel B). Mutant eRFs indicated by red asterisk.

Chapter 5

Discussion

5.1 SMG1-8-9 in NMD

Insights provided by biochemical and high-resolution structural data have vastly improved our understanding of SMG1 activity in NMD: central questions regarding substrate specificity and regulation have been answered by the work presented here. Nevertheless, some problems remain unsolved and several new features of the SMG1-8-9 complex have emerged that warrant further investigation (Gat *et al.*, 2019; Zhu *et al.*, 2019; Langer *et al.*, 2020, 2021).

Insights into UPF1 phosphorylation by SMG1-8-9

The structure of a UPF1-derived peptide bound to the SMG1 active site has visualized a central step of metazoan NMD (Langer *et al.*, 2020). Together with biochemical data, this structure showed that in addition to specificity for SQ motifs, SMG1 preferentially selects motifs with larger hydrophobic residues, in particular Leu, in the -1 position. This observation is in agreement with data that reported several Leu-Ser-Gln motifs in the C-terminus of UPF1 to be particularly important for phosphorylation (Yamashita *et al.*, 2001; Ohnishi *et al.*, 2003; Kashima *et al.*, 2006; Fukuhara *et al.*, 2005; Okada-Katsuhata *et al.*, 2012). Counterintuitively though, another important UPF1 phosphorylation site known to be crucial for SMG6 recruitment displays a Asp-Thr-Gln sequence (Okada-Katsuhata *et al.*, 2012; Durand *et al.*, 2016). Phosphorylation assays showed that substrate properties of this peptide were vastly improved when substituting Asp for a Leu, suggesting that the wildtype Asp-Thr-Gln phosphorylation site would be at a significant disadvantage when directly competing for binding to the SMG1 active site. Intriguingly, this motif as well as

almost all other less efficiently phosphorylated motifs are located in UPF1s N-terminus, whereas efficiently phosphorylated motifs are generally found in the C-terminus (Langer *et al.*, 2020). Importantly, efficiency of phosphorylation by SMG1 was monitored in an experimental setup using only short peptides carrying the motifs of interest as substrates, but not in the context of full-length UPF1. Considering the described well-established importance of the suboptimal N-terminal Asp-Thr-Gln motif in NMD, it is tempting to speculate that steric arrangements in the context of a full-length SMG1-8-9 - UPF1 complex position this motif closer to the active site compared to the more optimal C-terminal Leu-Ser-Gln motifs. Alternatively, phosphorylation of UPF1 N- and C-terminus might represent temporally separated events in NMD and not occur concomitantly. Notably hyperphosphorylation of UPF1 over time has been suggested to act as a molecular clock, reinforcing the notion that phosphorylation could occur in steps (Durand *et al.*, 2016). Both hypotheses could be connected to the observation that the unphosphorylated UPF1 C-terminus can interact intramolecularly with the helicase domain, therefore potentially being less accessible for SMG1 at certain timepoints in early NMD (Fiorini *et al.*, 2013). Similarly rearrangement of the CH-domain - which directly neighbours the unstructured N-terminus - upon UPF2 binding might impact phosphorylation site accessibility (Clerici *et al.*, 2009; Chakrabarti *et al.*, 2011).

Whether the downstream effectors SMG5-7 and SMG6 display binding preferences for phosphorylated SQ motifs with certain residues in the -1 position warrants investigation (Fukuhara *et al.*, 2005; Okada-Katsuhata *et al.*, 2012).

Intra- and intermolecular principles governing SMG1 regulation

In vitro studies have previously shown SMG1 kinase hyperactivation upon removal of either the SMG1 insertion/PRD or SMG8 (C-terminal domain), with a cumulative effect observed for removing both (Yamashita *et al.*, 2009; Arias-Palomo *et al.*, 2011; Deniaud *et al.*, 2015; Zhu *et al.*, 2019). The results described here, using integrated data from both cryo-EM as well as biochemical experiments with purified factors, have revealed an intricate interplay between the globular SMG8 C-terminal domain and the SMG1 insertion domain/PRD to directly block SMG1 active site accessibility (Langer *et al.*, 2020, 2021). Notably, the autoinhibited state of the SMG1-8-9 complex could still allow for binding of UPF1 to the periphery of the complex, as judged by superposition of the autoinhibited state with a negative-stain reconstruction of a SMG1-8-9 - UPF1 full-length complex (Melero *et al.*, 2014; Langer *et al.*, 2021). This observation raises the possibility that binding of the

UPF1 body to SMG1 is a prerequisite for overcoming SMG1 autoinhibition by efficient competition for SMG1 active site binding by the UPF1 phosphorylation sites located in its N- and C-terminal regions (see Figure 5.1).

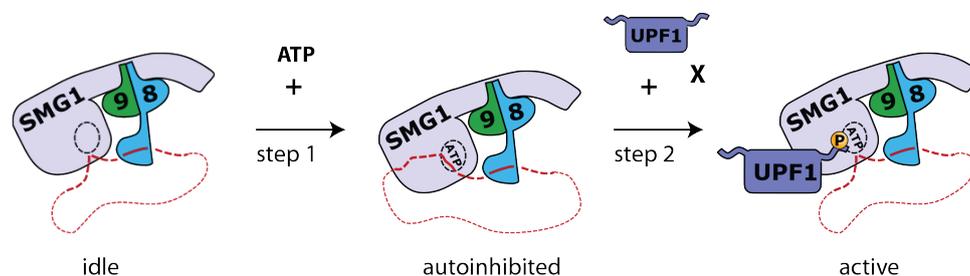


Figure 5.1: Hypothetical model for SMG1 activity regulation. The idle SMG1-8-9 complex adopts an autoinhibited conformation upon nucleotide-binding, mediated by the SMG1 insertion PRD (red) and the SMG8 C-terminal domain. Only in the context of the correct substrate (UPF1) and possibly other factors, the kinase complex becomes active to drive NMD by UPF1 phosphorylation.

The work discussed here has therefore provided a unifying mechanistic model for keeping the SMG1 kinase in an autoinhibited conformation, thereby preventing off-target effects before its timely activation in the context of an NMD mRNP (Langer *et al.*, 2021). Data showing that SMG8 is required for correct recruitment of SMG1-8-9 to NMD mRNPs via the EJC *in vivo* point to a possible mechanism connecting SMG1-8-9 recruitment to a premature termination event and kinase activation (Yamashita *et al.*, 2009; Kashima *et al.*, 2006). Hence, while not harboring any known enzymatic activity, SMG1 insertion/PRD and SMG8 C-terminus are critical for SMG1-8-9 regulation in NMD. The autoinhibitory potential of the SMG1 insertion domain/PRD is in agreement with the function described for PRDs of other PIKKs (Edinger & Thompson, 2004; Mordes *et al.*, 2008; Jansma *et al.*, 2020; Chen *et al.*, 2021b; Langer *et al.*, 2021). Due to its extraordinary size compared to other PRDs, the SMG1 insertion domain is likely to have additional important functions with respect to SMG1 action in NMD.

The combination of a low-resolution EM density for the SMG8 C-terminus, crosslinking mass spectrometry data and a model derived by AlphaFold2 has facilitated the description of a close to full-length model for SMG8 in the context of SMG1 and SMG9, providing a starting point for further biochemical analysis (see Figure 5.2) (Langer *et al.*, 2021; Jumper *et al.*, 2021).

Whether the SMG1-9 complex computationally identified within a SMG1-8-9 preparation represents a physiologically relevant complex or an experimental artefact is currently un-

clear. It might represent an assembly intermediate of SMG1-8-9 or a reaction intermediate with a currently not understood function. Interestingly, the SMG9 segment interacting with the SMG1 HEAT repeats bends by almost 90° in the SMG1-9 complex, exploiting a binding site occupied by SMG8 in the SMG1-8-9 structures (Langer *et al.*, 2021). Taking into consideration the observation that SMG1 purifies as a stable and stoichiometric complex with both SMG8 and SMG9 (Gat *et al.*, 2019), SMG8 and SMG9 having been described to be knocked down in an effort to purify SMG1 in isolation (Melero *et al.*, 2014) and finally SMG8-9 forming a stable heterodimer (Li *et al.*, 2017), it seems more likely that the SMG1-9 complex is not necessarily physiologically relevant. Of note, the N-terminal HEAT repeats of SMG1 seem to compact upon binding of SMG9 and SMG8-9 (Zhu *et al.*, 2019; Langer *et al.*, 2021).

Paths to understanding the recognition of premature translation termination events by SMG1-8-9 and UPF1

NMD is dependent on active translation in general and has been linked to translation termination in particular (Kashima *et al.*, 2006; Karousis & Mühlemann, 2019). While many NMD factor screens suffer from the limitation of not being able to differentiate between components specifically important for NMD and those generally required for translation, a current model suggests an eRF1-eRF3a-ribosome translation termination complex as the unit recognized by NMD factors. The GTPase eRF3a serves as a chaperone delivering eRF1 to a ribosome with a stop codon in its A-site, where the latter undergoes a conformational change to release the nascent chain by hydrolyzing the tRNA-nascent chain ester bond. Recent single-molecule fluorescence data showed that eRF3a association with the 80S ribosome occurs transiently, whereas eRF1-80S association is more stable (Lawson *et al.*, 2021). Hence, given the short-lived nature of the eRF1-eRF3a-80S complex, it remains to be seen whether it really provides the sought-after link between the translation machinery and NMD factors. The presented results describing the successful reconstitution of major intermediates along the translation termination pathway now provide a framework for a systematic testing of interactions between early acting NMD factors - such as SMG1-8-9 and UPF1 - and specific ribosomal termination complexes (see section 4.3). A limitation of the presented protocol is the use of *in vitro* transcribed RNA as a template for translation. As such, the RNA is missing mRNP components known to be important for the initiation of NMD, such as an EJC downstream of the termination codon. mRNPs more closely resembling the endogenous NMD substrates could potentially be obtained by

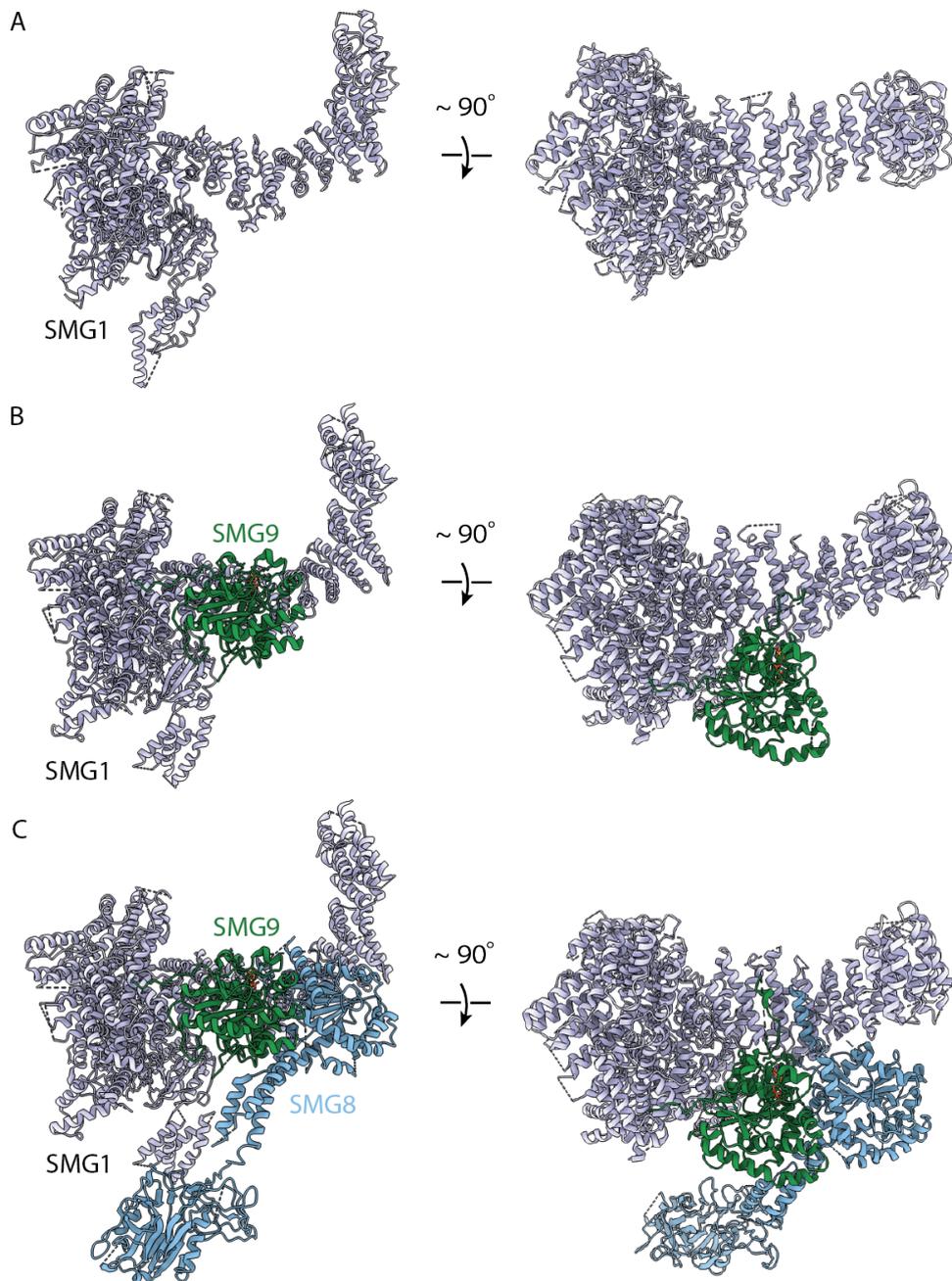


Figure 5.2: Structures of human SMG1, SMG1-9 and SMG1-8-9. Different SMG1-centered structures shown in two different orientations. A. Structure of SMG1 in isolation (PDB: 5LZT). B. Structure of SMG1 bound to SMG9 (PDB: 7PW9, this work) C. Structure of SMG1 bound to the SMG8-9 heterodimer (PDB: 7PW5, this work).

using *in vitro* splicing reactions to deposit EJCs on *in vitro* transcribed mRNA, or the purification of mRNPs from cells (Mayeda & Krainer, 1999; Fritz *et al.*, 2018).

5.2 SMG1-8-9 and other PIKKs

Recent years have seen numerous new structures of all (human) PIKKs being published, some of them detailing regulatory mechanisms, some describing sophisticated higher-order complexes in which PIKKs take a central place, and others detailing the mechanism of small molecule binding to the kinase active site (Rao *et al.*, 2018; Gat *et al.*, 2019; Langer *et al.*, 2020; Scaiola *et al.*, 2020; Jansma *et al.*, 2020; Yates *et al.*, 2020; Chen *et al.*, 2021b,a; Chaplin *et al.*, 2021; Langer *et al.*, 2021; Stakyte *et al.*, 2021; Liang *et al.*, 2022; Warren & Pavletich, 2022).

Small molecule inhibitors targeting PIKKs

The growing number of higher resolution structures of PIKKs bound to inhibitory compounds has provided an initial glimpse into small molecule as well as kinase active site features that can be exploited to improve targeting specificity. To date, inhibitor-bound active site structures have been determined for human mTOR, DNA-PKcs, ATM and SMG1 (Yang *et al.*, 2013; Langer *et al.*, 2021; Stakyte *et al.*, 2021; Liang *et al.*, 2022). Superposition of these recently determined structures of PIKK inhibitors with the SMG1i-bound SMG1 active site revealed both shared and distinct features of the SMG1 inhibitor (see Figure 5.3). All of these active site-targeting inhibitors work in an ATP-competitive manner, exploiting the same binding site in the catalytic cleft. Importantly, the phenylurea-group of SMG1i extends further along the catalytic cleft than any other compound. It therefore exploits binding sites not used by the other PIKK-targeting compounds. This is in agreement with the reported biophysical and structural observations describing the phenylurea-group as a potentially important part for achieving specificity of the compound towards SMG1 (Gopalsamy *et al.*, 2012; Langer *et al.*, 2021). In contrast to observations made for DNA-PKcs inhibitors (Liang *et al.*, 2022), binding of SMG1i to the SMG1 active site did not result in significant structural rearrangements of the SMG1 kinase active site (Langer *et al.*, 2021).

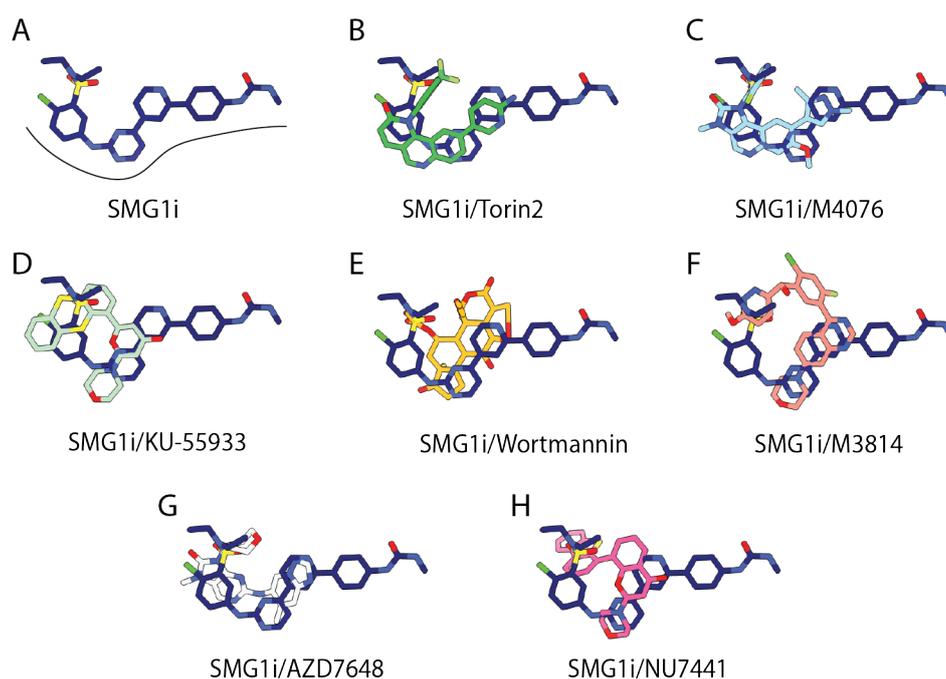


Figure 5.3: Superposition of PIKK active site-bound inhibitors with SMG1i. All superpositions were performed using the catalytic loops of the omitted kinase domains. A. Structure of SMG1i when bound to SMG1 (PDB: 7PW6). The interface with the SMG1 active site is indicated by a black line. B. mTOR inhibitor Torin2 (PDB: 4JSX) C. ATM inhibitor M4076 (PDB: 7NI4). D. ATM inhibitor KU-55933 (PDB: 7NI5). E. PIKK inhibitor Wortmannin bound to DNA-PKcs (PDB: 7OTV). F. DNA-PKcs inhibitor M3814 (PDB: 7OTY). G. DNA-PKcs inhibitor AZD7648 (PDB: 7OTW). H. DNA-PKcs inhibitor NU7441 (PDB: 7OTM).

Implications for substrate recognition and activity regulation across the PIKK family

The determination of the first structure of a substrate- and nucleotide-bound PIKK active site in conjunction with the report of many other nucleotide-bound PIKK structures has also provided invaluable insights into regulation of the kinases and rationalized previously reported biochemical data. Next to SMG1, specificity for SQ motifs has also been reported for ATM, ATR and DNA-PKcs (O'Neill *et al.*, 2000; Kim *et al.*, 1999; Bannister *et al.*, 1993). ATM and DNA-PKcs have previously been described to prefer a Leu residue in position -1, as shown here for SMG1. Hence, superposition of the SMG1-UPF1 peptide structure with high-resolution structures of these kinases allows an initial comparison of substrate recognition features conserved between these related enzymes (see Figure 5.4).

In the SMG1 structure, the substrate position +1 Gln residue is recognized by a pocket formed by the activation loop via hydrogen bonds to the peptide backbone as well as hydrophobic interactions. This pocket is also present in ATM and DNA-PKcs, suggesting a highly similar mode of recognition for the SQ specificity motif. Notably, it has diverged in mTOR, consistent with the reported differences in phosphorylation motif specificity (Hsu *et al.*, 2011; Langer *et al.*, 2020; Battaglioni *et al.*, 2022). Interaction with the -1 Leu residue is apparently less conserved across SMG1, ATM and DNA-PKcs, possibly reflecting its less strict requirement with regards to selectivity (Langer *et al.*, 2020).

Importantly, superposition of the UPF1 substrate peptide with the human ATM active site revealed that the ATM PRD folding back into the catalytic cleft would sterically clash with a bound substrate peptide (see Figure 5.4 A). Intriguingly, the ATM PRD inserts a Gln residue into the SQ motif Gln recognition pocket (see Figure 5.4 B). As proposed previously, it thereby effectively mimicks substrate binding and stably locks the kinase in an autoinhibited state, a feature conserved across species (Jansma *et al.*, 2020; Yates *et al.*, 2020; Langer *et al.*, 2020; Stakyte *et al.*, 2021; Warren & Pavletich, 2022). Similarly, superposition with structures of inactive and active DNA-PK revealed that a substrate bound to the active site is mutually exclusive with the PRD conformation in the inactive state of the kinase (see Figure 5.4 C). Only upon activation the PRD rearranges, making the substrate binding site accessible (see Figure 5.4 D) (Chen *et al.*, 2021b).

The observation that the SMG1 insertion domain contains a PRD that can physically block the substrate binding site as well therefore revealed an additional regulatory feature shared across different PIKKs (Langer *et al.*, 2021). The lower resolution of the SMG1 insertion within the kinase active site suggests flexibility and less stable binding as compared to the ATM PRD. This might also point towards different mechanisms controlling activation, where in the case of SMG1 correct positioning of UPF1 with respect to the kinase could be enough to allow the SQ motifs located in its unstructured N- and C-termini to efficiently compete with the insertion domain PRD for binding to the active site (see above). Such a model would be in agreement with the high SMG1 activity levels observed in *in vitro* assays (Gat *et al.*, 2019; Langer *et al.*, 2020, 2021). A more detailed dissection of SMG1 insertion PRD function has been hindered by the limited resolution, not allowing to model residues for this part of the molecule. Whether autophosphorylation plays a role in SMG1-8-9 activation as observed for other PIKKs awaits further investigation (Liu *et al.*, 2022).

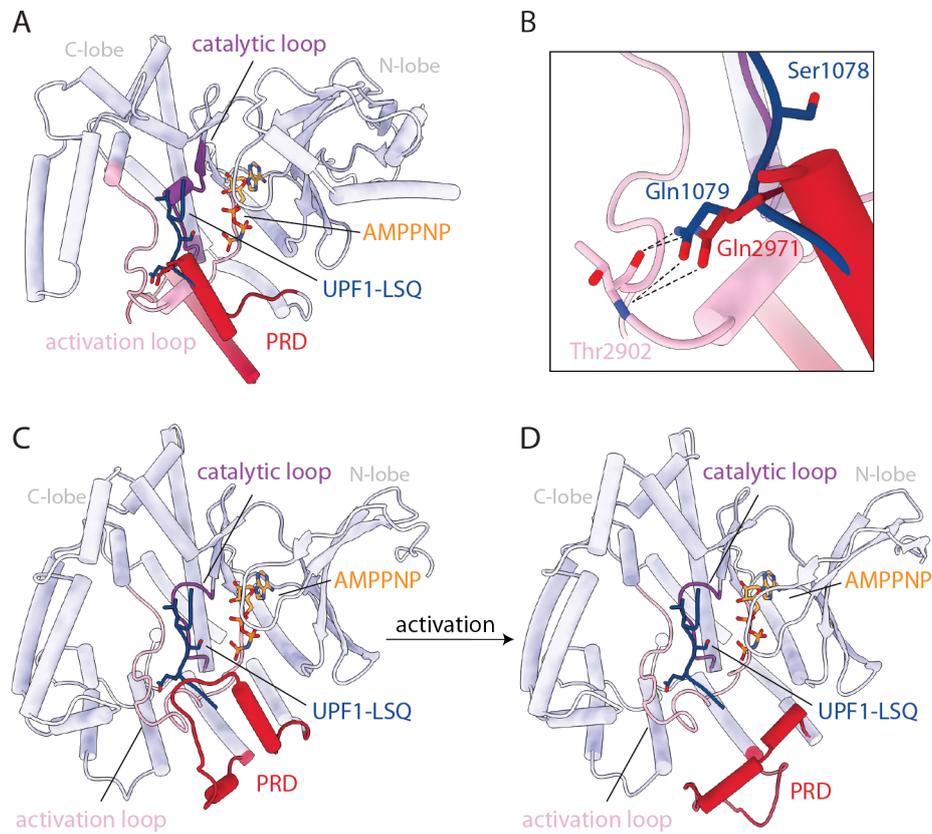


Figure 5.4: Superposition of UPF1-LSQ peptide and AMPPNP with human ATM and DNA-PKcs active sites. All superpositions were done using the catalytic loops, the SMG1 active site was omitted for clarity. A. Human ATM kinase domain (PDB: 7NI6) is shown in light blue, with the activation loop in pink, the catalytic loop in magenta and the PRD in red. The SMG1 kinase domain is not shown, SMG1-bound AMPPNP is in orange and the substrate peptide in dark blue (PDB: 6Z3R). PRD and substrate peptide would sterically clash. B. Zoom-in into the superposition shown in A., detailing the Gln residue of the ATM PRD mimicking substrate Gln binding. C. Inactive human DNA-PK kinase domain (PDB: 7K11), otherwise as in A., PRD and substrate peptide would sterically clash. D. Activated human DNA-PK kinase domain (PDB: 7K10), otherwise as in A. After rearrangement, PRD and substrate peptide do not clash.

Chapter 6

Outlook

Building on the results of this thesis, the reconstitution and structural analysis of larger SMG1-8-9 centered complexes can facilitate further important insights into SMG1 regulation and UPF1 phosphorylation in NMD. The RNA helicase DHX34 has been suggested to support recruitment of unphosphorylated UPF1 to SMG1-8-9 (Melero *et al.*, 2016). DHX34 has also been shown to form a complex with RuvBL1/2, which is known to be involved in the assembly of other macromolecular complexes and interacts with SMG1 by itself (López-Perrote *et al.*, 2020). In contrast to DHX34, UPF2 binding to UPF1 has been suggested to release it from SMG1-8-9, potentially after phosphorylation (Deniaud *et al.*, 2015). Given the multi-faceted intra- and intermolecular interactions governing UPF1 activity, it seems plausible that different factors will contribute to its correct recruitment to SMG1-8-9. Regardless, understanding the interaction between full-length UPF1 and SMG1-8-9 to an extent reaching beyond SQ motifs will provide important insights into NMD.

The work described in this thesis also provides a framework to further dissect NMD-initiating events in a broader context. A major obstacle to understanding NMD is the unresolved question regarding correct localization and activation of SMG1-8-9. This concerns both spatial and temporal aspects. Addressing how SMG1-8-9 is recruited to premature translation termination events in a timely manner would fill a major gap in the general mechanistic picture of NMD. Based on the existing body of literature it is likely that an interplay of translation termination machinery and other mRNP components such as EJC and UPF1 play an important role, but the details remain almost entirely unclear. Evidence has emerged that NMD is a leaky process, with statistically only every eighth ribosome translating an NMD mRNA inducing its degradation, at least for selected reporter transcripts. In this setup, every ribosome had the same probability of inducing NMD

(Hoek *et al.*, 2019). Furthermore, recent biochemical data suggest that NMD is induced independently of stable ribosome stalling, an important difference to other co-translational mRNA quality control pathways (Karousis *et al.*, 2020).

Gaining structural insights into recruitment and activation of SMG1-8-9 in an NMD mRNP environment will likely require the simultaneous usage of different tools, ranging from targeted manipulation of translation termination over specific inhibition of high-quality recombinant NMD factors at defined stages of the pathway to the use of correctly processed mRNPs as a substrate. Given the advances presented in this work, revealing an integrated picture of SMG1-8-9 action in NMD is now within reach.

Bibliography

- Adami, A., García-Álvarez, B., Arias-Palomo, E., Barford, D., & Llorca, O. Structure of TOR and its complex with KOG1. *Molecular cell* **27**(3):509–516 (2007).
- Alkalaeva, E. Z., Pisarev, A. V., Frolova, L. Y., Kisselev, L. L., & Pestova, T. V. In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. *Cell* **125**(6):1125–1136 (2006).
- Alzahrani, F., Kuwahara, H., Long, Y., Al-Owain, M., Tohary, M., AlSayed, M., Mahnashi, M., Fathi, L., Alnemer, M., Al-Hamed, M. H. *et al.* Recessive, deleterious variants in SMG8 expand the role of nonsense-mediated decay in developmental disorders in humans. *The American Journal of Human Genetics* **107**(6):1178–1185 (2020).
- Amrani, N., Ganesan, R., Kervestin, S., Mangus, D. A., Ghosh, S., & Jacobson, A. A faux 3-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432**(7013):112–118 (2004).
- Andersen, C. B., Ballut, L., Johansen, J. S., Chamieh, H., Nielsen, K. H., Oliveira, C. L., Pedersen, J. S., Séraphin, B., Hir, H. L., & Andersen, G. R. Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science* **313**(5795):1968–1972 (2006).
- Anderson, J. S. J. & Parker, R. The 3 to 5 degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3 to 5 exonucleases of the exosome complex. *The EMBO journal* **17**(5):1497–1506 (1998).
- Arias-Palomo, E., Yamashita, A., Fernández, I. S., Núñez-Ramírez, R., Bamba, Y., Izumi, N., Ohno, S., & Llorca, O. The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes & development* **25**(2):153–164 (2011).
- Bannister, A. J., Gottlieb, T. M., Kouzarides, T., & Jackson, S. P. c-Jun is phosphorylated by the DNA-dependent protein kinase in vitro; definition of the minimal kinase recognition motif. *Nucleic acids research* **21**(5):1289–1295 (1993).
- Baretić, D. & Williams, R. L. PIKKs—the solenoid nest where partners and kinases meet. *Current opinion in structural biology* **29**:134–142 (2014).

- Battaglion, S., Benjamin, D., Wälchli, M., Maier, T., & Hall, M. N. mTOR substrate phosphorylation in growth control. *Cell* (2022).
- Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., & Izaurralde, E. A conserved role for cytoplasmic poly (A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *The EMBO journal* **26**(6):1591–1601 (2007).
- Belgrader, P., Cheng, J., & Maquat, L. E. Evidence to implicate translation by ribosomes in the mechanism by which nonsense codons reduce the nuclear level of human triosephosphate isomerase mRNA. *Proceedings of the National Academy of Sciences* **90**(2):482–486 (1993).
- Bhuvanagiri, M., Schlitter, A. M., Hentze, M. W., & Kulozik, A. E. NMD: RNA biology meets human genetic medicine. *Biochemical Journal* **430**(3):365–377 (2010).
- Boehm, V., Kueckelmann, S., Gerbracht, J. V., Kallabis, S., Britto-Borges, T., Altmüller, J., Krüger, M., Dieterich, C., & Gehring, N. H. SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity. *Nature communications* **12**(1):1–19 (2021).
- Boersma, S., Khuperkar, D., Verhagen, B. M., Sonneveld, S., Grimm, J. B., Lavis, L. D., & Tanenbaum, M. E. Multi-color single-molecule imaging uncovers extensive heterogeneity in mRNA decoding. *Cell* **178**(2):458–472 (2019).
- Bono, F., Ebert, J., Lorentzen, E., & Conti, E. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* **126**(4):713–725 (2006).
- Brandman, O. & Hegde, R. S. Ribosome-associated protein quality control. *Nature structural & molecular biology* **23**(1):7–15 (2016).
- Brown, A., Shao, S., Murray, J., Hegde, R. S., & Ramakrishnan, V. Structural basis for stop codon recognition in eukaryotes. *Nature* **524**(7566):493–496 (2015).
- Brumbaugh, K. M., Otterness, D. M., Geisen, C., Oliveira, V., Brognard, J., Li, X., Lejeune, F., Tibbetts, R. S., Maquat, L. E., & Abraham, R. T. The mRNA surveillance protein hSMG-1 functions in genotoxic stress response pathways in mammalian cells. *Molecular cell* **14**(5):585–598 (2004).
- Buchwald, G., Ebert, J., Basquin, C., Sauliere, J., Jayachandran, U., Bono, F., Le Hir, H., & Conti, E. Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. *Proceedings of the National Academy of Sciences* **107**(22):10050–10055 (2010).
- Bühler, M., Steiner, S., Mohn, F., Paillusson, A., & Mühlemann, O. EJC-independent degradation of nonsense immunoglobulin- μ mRNA depends on 3' UTR length. *Nature structural & molecular biology* **13**(5):462–464 (2006).

- Campofelice, A., Lentini, L., Di Leonardo, A., Melfi, R., Tutone, M., Pace, A., & Pibiri, I. Strategies against nonsense: oxadiazoles as translational readthrough-inducing drugs (TRIDs). *International Journal of Molecular Sciences* **20**(13):3329 (2019).
- Carter, M. S., Doskow, J., Morris, P., Li, S., Nhim, R. P., Sandstedt, S., & Wilkinson, M. F. A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. *Journal of Biological Chemistry* **270**(48):28995–29003 (1995).
- Chakrabarti, S., Bonneau, F., Schüssler, S., Eppinger, E., & Conti, E. Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5–SMG7 and SMG6. *Nucleic acids research* **42**(14):9447–9460 (2014).
- Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., & Conti, E. Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Molecular cell* **41**(6):693–703 (2011).
- Chamieh, H., Ballut, L., Bonneau, F., & Le Hir, H. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nature structural & molecular biology* **15**(1):85–93 (2008).
- Chandrasekaran, V., Juszkievicz, S., Choi, J., Puglisi, J. D., Brown, A., Shao, S., Ramakrishnan, V., & Hegde, R. S. Mechanism of ribosome stalling during translation of a poly (A) tail. *Nature structural & molecular biology* **26**(12):1132–1140 (2019).
- Chang, J. C. & Kan, Y. W. beta 0 thalassemia, a nonsense mutation in man. *Proceedings of the National Academy of Sciences* **76**(6):2886–2889 (1979).
- Chaplin, A. K., Hardwick, S. W., Stavridi, A. K., Buehl, C. J., Goff, N. J., Ropars, V., Liang, S., De Oliveira, T. M., Chirgadze, D. Y., Meek, K. *et al.* Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Molecular cell* **81**(16):3400–3409 (2021).
- Chapman, J. H., Craig, J. M., Wand, C. D., Gundlach, J. H., Neuman, K. C., & Hogg, J. R. UPF1 mutants with intact ATPase but deficient helicase activities promote efficient nonsense-mediated mRNA decay. *bioRxiv* (2022).
- Chen, S., Lee, L., Naila, T., Fishbain, S., Wang, A., Tomkinson, A. E., Lees-Miller, S. P., & He, Y. Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature* **593**(7858):294–298 (2021a).
- Chen, X., Xu, X., Chen, Y., Cheung, J. C., Wang, H., Jiang, J., de Val, N., Fox, T., Gellert, M., & Yang, W. Structure of an activated DNA-PK and its implications for NHEJ. *Molecular cell* **81**(4):801–810 (2021b).
- Cheng, Z., Muhrad, D., Lim, M. K., Parker, R., & Song, H. Structural and functional insights into the human Upf1 helicase core. *The EMBO journal* **26**(1):253–264 (2007).

- Chu, V., Feng, Q., Lim, Y., & Shao, S. Selective destabilization of polypeptides synthesized from NMD-targeted transcripts. *Molecular biology of the cell* **32**(22):ar38 (2021).
- Clerici, M., Mourão, A., Gutsche, I., Gehring, N. H., Hentze, M. W., Kulozik, A., Kadlec, J., Sattler, M., & Cusack, S. Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *The EMBO journal* **28**(15):2293–2306 (2009).
- Dabrowski, M., Bukowy-Bieryllo, Z., & Zietkiewicz, E. Advances in therapeutic use of a drug-stimulated translational readthrough of premature termination codons. *Molecular medicine* **24**(1):1–15 (2018).
- Dehecq, M., Decourty, L., Namane, A., Proux, C., Kanaan, J., Le Hir, H., Jacquier, A., & Saveanu, C. Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes. *The EMBO journal* **37**(21):e99278 (2018).
- Deniaud, A., Karuppasamy, M., Bock, T., Masiulis, S., Huard, K., Garzoni, F., Kerschgens, K., Hentze, M. W., Kulozik, A. E., Beck, M. *et al.* A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. *Nucleic acids research* **43**(15):7600–7611 (2015).
- Denning, G., Jamieson, L., Maquat, L. E., Thompson, E. A., & Fields, A. P. Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein. *Journal of Biological Chemistry* **276**(25):22709–22714 (2001).
- Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**(7083):561–564 (2006).
- Durand, S., Cougot, N., Mahuteau-Betzer, F., Nguyen, C.-H., Grierson, D. S., Bertrand, E., Tazi, J., & Lejeune, F. Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. *The Journal of cell biology* **178**(7):1145–1160 (2007).
- Durand, S., Franks, T. M., & Lykke-Andersen, J. Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. *Nature communications* **7**(1):1–12 (2016).
- Durand, S. & Lykke-Andersen, J. Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nature structural & molecular biology* **20**(6):702–709 (2013).
- Eberle, A. B., Stalder, L., Mathys, H., Orozco, R. Z., & Mühlemann, O. Posttranscriptional gene regulation by spatial rearrangement of the 3 untranslated region. *PLoS biology* **6**(4):e92 (2008).
- Edinger, A. L. & Thompson, C. B. An activated mTOR mutant supports growth factor-independent, nutrient-dependent cell survival. *Oncogene* **23**(33):5654–5663 (2004).

- Feng, Q. & Shao, S. In vitro reconstitution of translational arrest pathways. *Methods* **137**:20–36 (2018).
- Fiorini, F., Boudvillain, M., & Le Hir, H. Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains. *Nucleic acids research* **41**(4):2404–2415 (2013).
- Franks, T. M., Singh, G., & Lykke-Andersen, J. Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. *Cell* **143**(6):938–950 (2010).
- Frischmeyer, P. A., Van Hoof, A., O'Donnell, K., Guerrerío, A. L., Parker, R., & Dietz, H. C. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* **295**(5563):2258–2261 (2002).
- Fritz, S. E., Haque, N., & Hogg, J. R. Highly efficient in vitro translation of authentic affinity-purified messenger ribonucleoprotein complexes. *RNA* **24**(7):982–989 (2018).
- Frolova, L., Le Goff, X., Rasmussen, H. H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A.-L., Celis, J. E., Phillippe, M. *et al.* A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. *Nature* **372**(6507):701–703 (1994).
- Frolova, L., Seit-Nebi, A., & Kisselev, L. Highly conserved NIKS tetrapeptide is functionally essential in eukaryotic translation termination factor eRF1. *Rna* **8**(2):129–136 (2002).
- Frolova, L. Y., Tsivkovskii, R. Y., Sivolobova, G. F., Oparina, N. Y., Serpinsky, O. I., Blinov, V. M., Tatkov, S. I., & Kisselev, L. L. Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *Rna* **5**(8):1014–1020 (1999).
- Fukuhara, N., Ebert, J., Unterholzner, L., Lindner, D., Izaurralde, E., & Conti, E. SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Molecular cell* **17**(4):537–547 (2005).
- Gaba, A., Jacobson, A., & Sachs, M. S. Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay. *Molecular cell* **20**(3):449–460 (2005).
- Gat, Y., Schuller, J. M., Lingaraju, M., Weyher, E., Bonneau, F., Strauss, M., Murray, P. J., & Conti, E. InsP6 binding to PIKK kinases revealed by the cryo-EM structure of an SMG1–SMG8–SMG9 complex. *Nature structural & molecular biology* **26**(12):1089–1093 (2019).

- Ge, Z., Quek, B. L., Beemon, K. L., & Hogg, J. R. Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. *Elife* **5**:e11155 (2016).
- Gehen, S. C., Stavarsky, R. J., Bambara, R. A., Keng, P. C., & O'Reilly, M. A. hSMG-1 and ATM sequentially and independently regulate the G1 checkpoint during oxidative stress. *Oncogene* **27**(29):4065–4074 (2008).
- Gehring, N. H., Lamprinaki, S., Kulozik, A. E., & Hentze, M. W. Disassembly of exon junction complexes by PYM. *Cell* **137**(3):536–548 (2009).
- Gewandter, J. S., Bambara, R. A., & o'Reilly, M. A. The RNA surveillance protein SMG1 activates p53 in response to DNA double-strand breaks but not exogenously oxidized mRNA. *Cell Cycle* **10**(15):2561–2567 (2011).
- Glavan, F., Behm-Ansmant, I., Izaurralde, E., & Conti, E. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *The EMBO journal* **25**(21):5117–5125 (2006).
- Glover, M. L., Burroughs, A. M., Monem, P. C., Egelhofer, T. A., Pule, M. N., Aravind, L., & Arribere, J. A. NONU-1 encodes a conserved endonuclease required for mRNA translation surveillance. *Cell reports* **30**(13):4321–4331 (2020).
- Gopalsamy, A., Bennett, E. M., Shi, M., Zhang, W.-G., Bard, J., & Yu, K. Identification of pyrimidine derivatives as hSMG-1 inhibitors. *Bioorganic & medicinal chemistry letters* **22**(21):6636–6641 (2012).
- Gowravaram, M., Schwarz, J., Khilji, S. K., Urlaub, H., & Chakrabarti, S. Insights into the assembly and architecture of a Staufien-mediated mRNA decay (SMD)-competent mRNP. *Nature communications* **10**(1):1–14 (2019).
- Hodgkin, J., Papp, A., Pulak, R., Ambros, V., & Anderson, P. A new kind of informational suppression in the nematode *Caenorhabditis elegans*. *Genetics* **123**(2):301–313 (1989).
- Hoek, T. A., Khuperkar, D., Lindeboom, R. G., Sonneveld, S., Verhagen, B. M., Boersma, S., Vermeulen, M., & Tanenbaum, M. E. Single-molecule imaging uncovers rules governing nonsense-mediated mRNA decay. *Molecular cell* **75**(2):324–339 (2019).
- Hogg, J. R. & Goff, S. P. Upf1 senses 3' UTR length to potentiate mRNA decay. *Cell* **143**(3):379–389 (2010).
- Hsu, P. P., Kang, S. A., Rameseder, J., Zhang, Y., Ottina, K. A., Lim, D., Peterson, T. R., Choi, Y., Gray, N. S., Yaffe, M. B. *et al.* The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* **332**(6035):1317–1322 (2011).

- Hug, N. & Cáceres, J. F. The RNA helicase DHX34 activates NMD by promoting a transition from the surveillance to the decay-inducing complex. *Cell reports* **8**(6):1845–1856 (2014).
- Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., & Izaurralde, E. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *Rna* **14**(12):2609–2617 (2008).
- Imseeng, S., Aylett, C. H., & Maier, T. Architecture and activation of phosphatidylinositol 3-kinase related kinases. *Current Opinion in Structural Biology* **49**:177–189 (2018).
- Isken, O. & Maquat, L. E. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes & development* **21**(15):1833–3856 (2007).
- Jankowsky, A., Guenther, U.-P., & Jankowsky, E. The RNA helicase database. *Nucleic acids research* **39**(suppl_1):D338–D341 (2010).
- Jansma, M. & Hopfner, K.-P. Structural basis of the (in) activity of the apical DNA damage response kinases ATM, ATR and DNA-PKcs. *Progress in Biophysics and Molecular Biology* **163**:120–129 (2021).
- Jansma, M., Linke-Winnebeck, C., Eustermann, S., Lammens, K., Kostrewa, D., Stakyte, K., Litz, C., Kessler, B., & Hopfner, K.-P. Near-complete structure and model of Tel1ATM from *Chaetomium thermophilum* reveals a robust autoinhibited ATP state. *Structure* **28**(1):83–95 (2020).
- Jonas, S., Weichenrieder, O., & Izaurralde, E. An unusual arrangement of two 14-3-3-like domains in the SMG5–SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. *Genes & development* **27**(2):211–225 (2013).
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**(7873):583–589 (2021).
- Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., & Hegde, R. S. ZNF598 is a quality control sensor of collided ribosomes. *Molecular cell* **72**(3):469–481 (2018).
- Juszkiewicz, S. & Hegde, R. S. Initiation of quality control during poly (A) translation requires site-specific ribosome ubiquitination. *Molecular cell* **65**(4):743–750 (2017).
- Kadlec, J., Guilligay, D., Ravelli, R. B., & Cusack, S. Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1. *Rna* **12**(10):1817–1824 (2006).

- Karousis, E. D., Gurzeler, L.-A., Annibaldis, G., Dreos, R., & Mühlemann, O. Human NMD ensues independently of stable ribosome stalling. *Nature communications* **11**(1):1–12 (2020).
- Karousis, E. D. & Mühlemann, O. Nonsense-mediated mRNA decay begins where translation ends. *Cold Spring Harbor Perspectives in Biology* **11**(2):a032862 (2019).
- Karousis, E. D. & Mühlemann, O. The broader sense of nonsense. *Trends in biochemical sciences* (2022).
- Karousis, E. D., Nasif, S., & Mühlemann, O. Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. *Wiley Interdisciplinary Reviews: RNA* **7**(5):661–682 (2016).
- Kashima, I., Jonas, S., Jayachandran, U., Buchwald, G., Conti, E., Lupas, A. N., & Izauralde, E. SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes & development* **24**(21):2440–2450 (2010).
- Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., & Ohno, S. Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes & development* **20**(3):355–367 (2006).
- Kebaara, B. W. & Atkin, A. L. Long 3-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in *Saccharomyces cerevisiae*. *Nucleic acids research* **37**(9):2771–2778 (2009).
- Keith, C. T. & Schreiber, S. L. PIK-related kinases: DNA repair, recombination, and cell cycle checkpoints. *Science* **270**(5233):50–50 (1995).
- Kertesz, S., Kerenyi, Z., Merai, Z., Bartos, I., Palfy, T., Barta, E., & Silhavy, D. Both introns and long 3-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. *Nucleic acids research* **34**(21):6147–6157 (2006).
- Kim, S.-T., Lim, D.-S., Canman, C. E., & Kastan, M. B. Substrate specificities and identification of putative substrates of ATM kinase family members. *Journal of Biological Chemistry* **274**(53):37538–37543 (1999).
- Kim, Y. J., Nomakuchi, T., Papaleonidopoulou, F., Yang, L., Zhang, Q., & Krainer, A. R. Gene-specific nonsense-mediated mRNA decay targeting for cystic fibrosis therapy. *Nature communications* **13**(1):1–13 (2022).
- Kim, Y. K. & Maquat, L. E. UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. *Rna* **25**(4):407–422 (2019).

- Kishor, A., Fritz, S. E., & Hogg, J. R. Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. *Wiley Interdisciplinary Reviews: RNA* **10**(6):e1548 (2019).
- Kögel, A., Keidel, A., Bonneau, F., Schäfer, I. B., & Conti, E. The human SKI complex regulates channeling of ribosome-bound RNA to the exosome via an intrinsic gatekeeping mechanism. *Molecular cell* **82**(4):756–769 (2022).
- Kononenko, A. V., Mitkevich, V. A., Atkinson, G. C., Tenson, T., Dubovaya, V. I., Frolova, L. Y., Makarov, A. A., & Hauryliuk, V. GTP-dependent structural rearrangement of the eRF1: eRF3 complex and eRF3 sequence motifs essential for PABP binding. *Nucleic acids research* **38**(2):548–558 (2010).
- Kowalinski, E., Kögel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B., & Conti, E. Structure of a cytoplasmic 11-subunit RNA exosome complex. *Molecular cell* **63**(1):125–134 (2016).
- Kuroha, K., Tatematsu, T., & Inada, T. Upf1 stimulates degradation of the product derived from aberrant messenger RNA containing a specific nonsense mutation by the proteasome. *EMBO reports* **10**(11):1265–1271 (2009).
- Kurosaki, T., Popp, M. W., & Maquat, L. E. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nature reviews Molecular cell biology* **20**(7):406–420 (2019).
- Ladner, C. L., Yang, J., Turner, R. J., & Edwards, R. A. Visible fluorescent detection of proteins in polyacrylamide gels without staining. *Analytical biochemistry* **326**(1):13–20 (2004).
- Langer, L. M., Bonneau, F., Gat, Y., & Conti, E. Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8. *Elife* **10**:e72353 (2021).
- Langer, L. M., Gat, Y., Bonneau, F., & Conti, E. Structure of substrate-bound SMG1-8-9 kinase complex reveals molecular basis for phosphorylation specificity. *Elife* **9**:e57127 (2020).
- Lavysh, D. & Neu-Yilik, G. UPF1-mediated RNA decay—Danse Macabre in a cloud. *Biomolecules* **10**(7):999 (2020).
- Lawson, M. R., Lessen, L. N., Wang, J., Prabhakar, A., Corsepis, N. C., Green, R., & Puglisi, J. D. Mechanisms that ensure speed and fidelity in eukaryotic translation termination. *Science* **373**(6557):876–882 (2021).
- Le Hir, H., Gatfield, D., Izaurralde, E., & Moore, M. J. The exon–exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *The EMBO journal* **20**(17):4987–4997 (2001).

- Le Hir, H., Izaurralde, E., Maquat, L. E., & Moore, M. J. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *The EMBO journal* **19**(24):6860–6869 (2000).
- Le Hir, H. & Séraphin, B. EJCs at the heart of translational control. *Cell* **133**(2):213–216 (2008).
- Leeds, P., Wood, J., Lee, B., & Culbertson, M. Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*. *Molecular and cellular biology* **12**(5):2165–2177 (1992).
- Leubitz, A., Frydman-Marom, A., Sharpe, N., van Duzer, J., Campbell, K. C., & Vanhoutte, F. Safety, tolerability, and pharmacokinetics of single ascending doses of ELX-02, a potential treatment for genetic disorders caused by nonsense mutations, in healthy volunteers. *Clinical pharmacology in drug development* **8**(8):984–994 (2019).
- Li, L., Lingaraju, M., Basquin, C., Basquin, J., & Conti, E. Structure of a SMG8–SMG9 complex identifies a G-domain heterodimer in the NMD effector proteins. *RNA* **23**(7):1028–1034 (2017).
- Liang, S., Thomas, S. E., Chaplin, A. K., Hardwick, S. W., Chirgadze, D. Y., & Blundell, T. L. Structural insights into inhibitor regulation of the DNA repair protein DNA-PKcs. *Nature* **601**(7894):643–648 (2022).
- Lindeboom, R. G., Vermeulen, M., Lehner, B., & Supek, F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nature genetics* **51**(11):1645–1651 (2019).
- Liu, L., Chen, X., Li, J., Wang, H., Buehl, C. J., Goff, N. J., Meek, K., Yang, W., & Gellert, M. Autophosphorylation transforms DNA-PK from protecting to processing DNA ends. *Molecular Cell* **82**(1):177–189 (2022).
- Llorca, O., Rivera-Calzada, A., Grantham, J., & Willison, K. Electron microscopy and 3D reconstructions reveal that human ATM kinase uses an arm-like domain to clamp around double-stranded DNA. *Oncogene* **22**(25):3867–3874 (2003).
- Loh, B., Jonas, S., & Izaurralde, E. The SMG5–SMG7 heterodimer directly recruits the CCR4–NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes & development* **27**(19):2125–2138 (2013).
- Longman, D., Plasterk, R. H., Johnstone, I. L., & Cáceres, J. F. Mechanistic insights and identification of two novel factors in the *C. elegans* NMD pathway. *Genes & development* **21**(9):1075–1085 (2007).
- López-Perrote, A., Hug, N., González-Corpas, A., Rodríguez, C. F., Serna, M., García-Martín, C., Boskovic, J., Fernandez-Leiro, R., Cáceres, J. F., & Llorca, O. Regulation

- of RUVBL1-RUVBL2 AAA-ATPases by the nonsense-mediated mRNA decay factor DHX34, as evidenced by Cryo-EM. *Elife* **9**:e63042 (2020).
- Losson, R. & Lacroute, F. Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proceedings of the National Academy of Sciences* **76**(10):5134–5137 (1979).
- Lovejoy, C. A. & Cortez, D. Common mechanisms of PIKK regulation. *DNA repair* **8**(9):1004–1008 (2009).
- Lykke-Andersen, J., Shu, M.-D., & Steitz, J. A. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103**(7):1121–1131 (2000).
- Lykke-Andersen, S. & Jensen, T. H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nature reviews Molecular cell biology* **16**(11):665–677 (2015).
- Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A., & Ross, J. Unstable β -globin mRNA in mRNA-deficient β 0 thalassemia. *Cell* **27**(3):543–553 (1981).
- Maquat, L. E., Tarn, W.-Y., & Isken, O. The pioneer round of translation: features and functions. *Cell* **142**(3):368–374 (2010).
- Marzluff, W. F. & Koreski, K. P. Birth and death of histone mRNAs. *Trends in Genetics* **33**(10):745–759 (2017).
- Mayeda, A. & Krainer, A. R. Mammalian in vitro splicing assays. In *RNA-Protein Interaction Protocols*, pages 315–321 (Springer, 1999).
- McMahon, S. B., Van Buskirk, H. A., Dugan, K. A., Copeland, T. D., & Cole, M. D. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell* **94**(3):363–374 (1998).
- Melero, R., Buchwald, G., Castaño, R., Raabe, M., Gil, D., Lázaro, M., Urlaub, H., Conti, E., & Llorca, O. The cryo-EM structure of the UPF–EJC complex shows UPF1 poised toward the RNA 3' end. *Nature structural & molecular biology* **19**(5):498–505 (2012).
- Melero, R., Hug, N., López-Perrote, A., Yamashita, A., Cáceres, J. F., & Llorca, O. The RNA helicase DHX34 functions as a scaffold for SMG1-mediated UPF1 phosphorylation. *Nature communications* **7**(1):1–12 (2016).
- Melero, R., Uchiyama, A., Castaño, R., Kataoka, N., Kurosawa, H., Ohno, S., Yamashita, A., & Llorca, O. Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. *Structure* **22**(8):1105–1119 (2014).

- Mino, T., Iwai, N., Endo, M., Inoue, K., Akaki, K., Hia, F., Uehata, T., Emura, T., Hidaka, K., Suzuki, Y. *et al.* Translation-dependent unwinding of stem-loops by UPF1 licenses Regnase-1 to degrade inflammatory mRNAs. *Nucleic Acids Research* **47**(16):8838–8859 (2019).
- Mordes, D. A., Glick, G. G., Zhao, R., & Cortez, D. TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. *Genes & development* **22**(11):1478–1489 (2008).
- Naeger, L. K., Schoborg, R. V., Zhao, Q., Tullis, G., & Pintel, D. J. Nonsense mutations inhibit splicing of MVM RNA in cis when they interrupt the reading frame of either exon of the final spliced product. *Genes & development* **6**(6):1107–1119 (1992).
- Nagy, E. & Maquat, L. E. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends in biochemical sciences* **23**(6):198–199 (1998).
- Neu-Yilik, G., Raimondeau, E., Eliseev, B., Yeramala, L., Amthor, B., Deniaud, A., Huard, K., Kerschgens, K., Hentze, M. W., Schaffitzel, C. *et al.* Dual function of UPF3B in early and late translation termination. *The EMBO journal* **36**(20):2968–2986 (2017).
- Nicholson, P., Josi, C., Kurosawa, H., Yamashita, A., & Mühlemann, O. A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic acids research* **42**(14):9217–9235 (2014).
- Nyikó, T., Sonkoly, B., Mérai, Z., Benkovics, A. H., & Silhavy, D. Plant upstream ORFs can trigger nonsense-mediated mRNA decay in a size-dependent manner. *Plant molecular biology* **71**(4):367–378 (2009).
- Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K. R., Grimson, A., Hachiya, T., Hentze, M. W., Anderson, P., & Ohno, S. Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. *Molecular cell* **12**(5):1187–1200 (2003).
- Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., & Ohno, S. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5: SMG-7 during NMD. *Nucleic acids research* **40**(3):1251–1266 (2012).
- O’Neill, T., Dwyer, A. J., Ziv, Y., Chan, D. W., Lees-Miller, S. P., Abraham, R. H., Lai, J. H., Hill, D., Shiloh, Y., Cantley, L. C. *et al.* Utilization of oriented peptide libraries to identify substrate motifs selected by ATM. *Journal of Biological Chemistry* **275**(30):22719–22727 (2000).
- Peltz, S. W., Brown, A. H., & Jacobson, A. mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. *Genes & development* **7**(9):1737–1754 (1993).

- Pochopien, A. A., Beckert, B., Kasvandik, S., Berninghausen, O., Beckmann, R., Tenson, T., & Wilson, D. N. Structure of Gen1 bound to stalled and colliding 80S ribosomes. *Proceedings of the National Academy of Sciences* **118**(14):e2022756118 (2021).
- Rao, Q., Liu, M., Tian, Y., Wu, Z., Hao, Y., Song, L., Qin, Z., Ding, C., Wang, H.-W., Wang, J. *et al.* Cryo-EM structure of human ATR-ATRIP complex. *Cell research* **28**(2):143–156 (2018).
- Rufener, S. C. & Mühlemann, O. eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nature structural & molecular biology* **20**(6):710–717 (2013).
- Salas-Marco, J. & Bedwell, D. M. GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. *Molecular and cellular biology* **24**(17):7769–7778 (2004).
- Saulière, J., Murigneux, V., Wang, Z., Marquet, E., Barbosa, I., Le Tonquèze, O., Audic, Y., Paillard, L., Crollius, H. R., & Le Hir, H. CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nature structural & molecular biology* **19**(11):1124–1131 (2012).
- Saxton, R. A. & Sabatini, D. M. mTOR signaling in growth, metabolism, and disease. *Cell* **168**(6):960–976 (2017).
- Scaiola, A., Mangia, F., Imseng, S., Boehringer, D., Berneiser, K., Shimobayashi, M., Stutfeld, E., Hall, M. N., Ban, N., & Maier, T. The 3.2-Å resolution structure of human mTORC2. *Science advances* **6**(45):eabc1251 (2020).
- Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A., & Mühlemann, O. Nonsense-mediated mRNA decay—mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1829**(6-7):612–623 (2013).
- Shaheen, R., Anazi, S., Ben-Omran, T., Seidahmed, M. Z., Caddle, L. B., Palmer, K., Ali, R., Alshidi, T., Hagos, S., Goodwin, L. *et al.* Mutations in SMG9, encoding an essential component of nonsense-mediated decay machinery, cause a multiple congenital anomaly syndrome in humans and mice. *The American Journal of Human Genetics* **98**(4):643–652 (2016).
- Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., & Hegde, R. S. Decoding mammalian ribosome-mRNA states by translational GTPase complexes. *Cell* **167**(5):1229–1240 (2016).
- Shiloh, Y. & Ziv, Y. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nature reviews Molecular cell biology* **14**(4):197–210 (2013).

- Shoemaker, C. J. & Green, R. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences* **108**(51):E1392–E1398 (2011).
- Shoemaker, C. J. & Green, R. Translation drives mRNA quality control. *Nature structural & molecular biology* **19**(6):594–601 (2012).
- Sibanda, B. L., Chirgadze, D. Y., & Blundell, T. L. Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats. *Nature* **463**(7277):118–121 (2010).
- Simms, C. L., Yan, L. L., & Zaher, H. S. Ribosome collision is critical for quality control during no-go decay. *Molecular cell* **68**(2):361–373 (2017).
- Singh, G., Rebbapragada, I., & Lykke-Andersen, J. A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. *PLoS biology* **6**(4):e111 (2008).
- Song, H., Mugnier, P., Das, A. K., Webb, H. M., Evans, D. R., Tuite, M. F., Hemmings, B. A., & Barford, D. The crystal structure of human eukaryotic release factor eRF1—mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell* **100**(3):311–321 (2000).
- Stakyte, K., Rotheneder, M., Lammens, K., Bartho, J., Grädler, U., Fuchß, T., Pehl, U., Alt, A., van de Logt, E., & Hopfner, K. Molecular basis of human ATM kinase inhibition. *Nature structural & molecular biology* **28**(10):789–798 (2021).
- Udy, D. B. & Bradley, R. K. Nonsense-mediated decay utilizes complementary mechanisms to suppress mRNA and protein accumulation (2021).
- Van Hoof, A., Frischmeyer, P. A., Dietz, H. C., & Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**(5563):2262–2264 (2002).
- Wallmeroth, D., Lackmann, J.-W., Kueckelmann, S., Altmüller, J., Dieterich, C., Boehm, V., & Gehring, N. H. Human UPF3A and UPF3B enable fault-tolerant activation of nonsense-mediated mRNA decay. *The EMBO Journal* **41**(10):e109191 (2022).
- Wang, J., Zhang, Y., Mendonca, C. A., Yukselen, O., Muneeruddin, K., Ren, L., Liang, J., Zhou, C., Xie, J., Li, J. *et al.* AAV-delivered suppressor tRNA overcomes a nonsense mutation in mice. *Nature* **604**(7905):343–348 (2022).
- Warren, C. & Pavletich, N. P. Structure of the human ATM kinase and mechanism of Nbs1 binding. *Elife* **11**:e74218 (2022).
- Welch, E. M., Barton, E. R., Zhuo, J., Tomizawa, Y., Friesen, W. J., Trifillis, P., Paushkin, S., Patel, M., Trotta, C. R., Hwang, S. *et al.* PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **447**(7140):87–91 (2007).

-
- Xue, G., Maciej, V. D., de Amorim, A. M., Pak, M., Jayachandran, U., & Chakrabarti, S. Modulation of RNA binding properties of the RNA helicase UPF1 by its activator UPF2. *bioRxiv* (2022).
- Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T. *et al.* SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. *Genes & development* **23**(9):1091–1105 (2009).
- Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., & Ohno, S. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes & development* **15**(17):2215–2228 (2001).
- Yang, H., Jiang, X., Li, B., Yang, H. J., Miller, M., Yang, A., Dhar, A., & Pavletich, N. P. Mechanisms of mTORC1 activation by RHEB and inhibition by PRAS40. *Nature* **552**(7685):368–373 (2017).
- Yang, H., Rudge, D. G., Koos, J. D., Vaidialingam, B., Yang, H. J., & Pavletich, N. P. mTOR kinase structure, mechanism and regulation. *Nature* **497**(7448):217–223 (2013).
- Yates, L. A., Williams, R. M., Hailemariam, S., Ayala, R., Burgers, P., & Zhang, X. Cryo-EM structure of nucleotide-bound Tel1ATM unravels the molecular basis of inhibition and structural rationale for disease-associated mutations. *Structure* **28**(1):96–104 (2020).
- Yi, Z., Arvola, R. M., Myers, S., Dilsavor, C. N., Abu Alhasan, R., Carter, B. N., Patton, R. D., Bundschuh, R., & Singh, G. Mammalian UPF3A and UPF3B can activate nonsense-mediated mRNA decay independently of their exon junction complex binding. *The EMBO journal* **41**(10):e109202 (2022).
- Yip, C. K., Murata, K., Walz, T., Sabatini, D. M., & Kang, S. A. Structure of the human mTOR complex I and its implications for rapamycin inhibition. *Molecular cell* **38**(5):768–774 (2010).
- Yip, M. C. & Shao, S. Detecting and rescuing stalled ribosomes. *Trends in Biochemical Sciences* **46**(9):731–743 (2021).
- Zhu, L., Li, L., Qi, Y., Yu, Z., & Xu, Y. Cryo-EM structure of SMG1–SMG8–SMG9 complex. *Cell research* **29**(12):1027–1034 (2019).
- Zou, L. & Elledge, S. J. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300**(5625):1542–1548 (2003).

Acknowledgements

I would like to thank Elena for an exciting project and for a research environment full of opportunities. For the scientific freedom and trust allowing me to venture into different directions during my PhD and for creating an atmosphere that fosters participation.

I would like to thank all people in the lab - past and present - for support throughout my PhD studies. You made the lab a special place I always looked forward going to. A true privilege in general, but especially invaluable during a pandemic. My special thanks to Fabien for ideas, discussions and invaluable insights into NMD and biochemistry. Yair for initial guidance on the SMG1 structure project and many stimulating and entertaining conversations. Felix for endless discussions on assays, projects, science, politics and on how to save the world in general. Sebastian for introducing me to the lab in the first place. Jana, Steffi, Giulia, Michaela, Marcela, Sofia, Achim, Alex, Dirk, Julian, Christian and Mahesh for sharing the ups and downs of the PhD journey and for loads of fun, in- and outside the lab.

At the MPIB, I would like to thank Daniel, Tilman and Rajan for support and help with all things cryo-EM. Lissy for mass spectrometry measurements and an unstoppable enthusiasm for science even after retirement. Stefan for synthesizing countless peptides.

I would like to mention Elisa Izaurralde and Kiyoshi Nagai, I feel deeply grateful to have met and worked with them.

Most importantly, I would like to thank Ksenia, my parents, their partners, my brothers and my entire family for unconditional support and the everyday reminder of the things that truly matter.