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

Structural characterization of CD47 and STING in complex with the activating nanobody

Błażej Maciej Skrobek

aus

Łódź, Poland

2023

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Karl-Peter Hopfner betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

Łódź, den. 16.05.2023 .....

Błażej M. Skrobek

Dissertation eingereicht am 22.05.2023

1. Gutachter: Herr Prof. Dr. Karl-Peter Hopfner

2. Gutachter: Herr Prof. Dr. Veit Hornung

Mündliche Prüfung am 14.07.2023

Table of Contents

1. Summary

cGAS-STING pathway is one of the most important first line defense mechanisms in a human body. Its action begins when a dsDNA from a pathogen enters the cell. Upon its detection by cGAS, the small molecule ligand of STING, called 2',3'-cGAMP, is synthesized. It binds to STING which resides in form of a dimer in endoplasmic reticulum. This butterfly-shaped molecule then closes its wings and rotates. Afterwards, the partner protein, TBK1 binds at its top and the oligomerization process begins. Such complex migrates to Golgi apparatus and due to several phosphorylation events it activates IRF3 which migrates to the nucleus and initiates the expression of type I interferons. These molecules trigger the further cascade of events that leads to inflammation and defense against the pathogen. This important physiological process can also lead to various disease states and thus is tightly regulated. Despite long-term research efforts, this complex pathway is still not fully understood.

Here, I present the STING-activating nanobody which can be a tool to explore this protein's mysteries. The nanobody was generated using alpaca immunization and found to significantly increase expression of CXCL10 - one of the interferon-stimulated genes, and to cause IRF3 migration to the nucleus. The first insight on how it happens was provided by the crystal structure of the nanobody with a soluble domain of STING. The wings of STING do not close during this interaction, which resembles the ways in which bacterial cyclic dinucleotides activate this protein. For now we may only speculate on how exactly the nanobody activates STING and further structural work is required to explain it. Examples of potential ways would be: release of STING C-terminal tail, positive or negative influence on interaction with partner proteins and triggering an oligomerization.

CD47, the marker of self, is a membrane protein present on the surface of every human cell, making it a "friend" of the host immune system. "Foes", e.g. cancer cells, often use its power to evade detection and thrive. Such observations lead to development of more and more successful therapies based on antibodies that bind to CD47 and prevent its interaction with SIRPα on macrophages, priming cells for destruction. However, the problem is that no one fully determined how actually CD47 works.

Here, I present my multiple structural attempts to do that, using cryo-EM. I used multiple powerful tools, like nanodiscs, detergents, affinity binders to facilitate structure determination. Despite that, a high-enough resolution was not obtained, most probably because of too high inherent flexibility of this challenging target. Even recently published, crystal apo structure did not provide enough information to explain how this molecule perform its function. To decipher its mysteries multiple more structures of CD47 with its known and to-be-known partner proteins are needed.

2. Introduction to stimulator of interferon genes

2.1. Innate immune system

The mammalian immune system is an advanced creation of millions of years of evolution to efficiently defend organism against a wide variety of threats. It consists of the innate and adaptive immune systems. The former constitutes a non-specific, first line of defense that works instantly upon hazard detection. If it fails, the latter is activated to produce antibodies that provide a tailored, specific response against the pathogen. The essential feature of innate immune response is an ability to sense a presence of various foreign pathogen-associated molecular patterns (PAMPs), like nucleic acids from viruses, lipopolysaccharide, peptidoglycan and flagellins from bacteria, mannose from fungi, or self damage-associated molecular patterns (DAMPs), like DNA released from mitochondria, by the diverse array of pattern recognition receptors (PRRs) (Barrat et al., 2016, Schaefer 2014, Sukhithasri et al., 2013, Tan et al., 2018). Among PRRs there are cyclic GMP-AMP (cGAMP) synthase (cGAS), and a number of protein families that include: absent in melanoma 2 (AIM2)-like receptors (ALRs), retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs), C-type lectin receptors (CLRs), nucleotide-binding domain, leucine-rich repeat (LRR)-containing receptors (NLRs), also known as nucleotide-binding oligomerization domain (NOD)-like receptors, and most widely researched, toll-like receptors (TLRs) (Barrat et al., 2016, Brubaker et al., 2015).

cGAS is an intracellular double-stranded DNA (dsDNA) detector that is the first component of the stimulator of interferon genes (STING) pathway that leads to type I interferons (IFN) production (Barrat et al., 2016, Brubaker et al., 2015) (see Sections 2.2. and 2.3.).

The ALRs contain protein-binding pyrin and hematopoietic interferon-inducible nuclear (PYHIN) domain, and hematopoietic interferon-inducible nuclear (HIN)-200 domain responsible for detection of DNA that entered a cell. Two known members are IFI16 and AIM2. Both of these proteins are responsible for inflammasome formation. IFI16 was also shown to be able to activate STING (Barrat et al., 2016, Brubaker et al., 2015).

The RLRs reside in cytoplasm, contain three DExH/D box helicases, and search for foreign RNA. This RNA must contain characteristic viral features or exist in a form of viral replication intermediate to be detected. The family includes RIG-I, laboratory of genetics and physiology 2 (LPG2) and melanoma differentiation-associated protein 5 (MDA5) (Barrat et al., 2016, Brubaker et al., 2015).

The CLRs include a large group of dectins that have C-type lectin-like domain and recognize various building blocks of microorganisms, for example, bacterial and fungal glucans, mucins, and fungal mannans (Barrat et al., 2016, Brubaker et al., 2015).

The NLRs contain a number of NOD proteins, including two most well known members NOD1 and NOD2. They share multiple similar structural elements, such as C-terminal LRRs, NACHT-domain, and N-terminal caspase activation and recruitment domains (CARDs). They search for components of bacterial cell walls and outer membranes (Barrat et al., 2016, Brubaker et al., 2015).

Among the TLRs there are proteins specialized in detection of various parts of microorganisms, including their nucleic acids. TLR1, 2, 4, 5 and 6 detect a wide variety of pathogen surface molecules. TLR9 searches for DNA containing unmethylated CpG motifs. TLR3 recognizes double-stranded RNA (dsRNA), and TLR7 and 8 bind to single-stranded RNA (ssRNA) (Barrat et al., 2016, Brubaker et al., 2015).

The most important common goals of activation of PRRs are to induce production of interferons and proinflammatory cytokines. This triggers strong innate immune response and further leads to activation of adaptive immunity, in order to efficiently fight against infection (Brubaker et al., 2015).

2.2. Structure of cGAS

The first component of type I interferon pathway is cGAS that detects dsDNA. cGAS is a 60 kDa soluble protein built of 522 residues (aa). It starts with a small, unstructured amino-terminal domain (NTD) (aa 1-160) with the sequence weakly conserved among species (Civril et al., 2013, Wu et al., 2014). It is followed by a highly conserved carboxy-terminal domain (CTD) (aa 160-522) that comprises NTase core (aa 160-330) overlapping with male abnormal 21 (Mab21) domain (aa 213-513). Mab21 domain subdivides into two lobes that form an active site upon DNA binding. It contains zinc-thumb structure, also called a zinc ribbon motif (aa 390-404), that hosts a Zn^{2+} ion and enables recognition of dsDNA. Also, in its lobe 1 there is a motif containing a beta sheet, surrounded by two alpha helices, which harbors residues that coordinate magnesium ions, critical for 2',3'-cGAMP production (Civril et al., 2013, Hornung et al., 2014, Wu et al., 2014). This is a characteristic feature of nucleotidyl transferase protein family (NTases) that catalyses the reaction between a nucleotide triphosphate and an OH group of an acceptor molecule, which in this case is another nucleotide (Kuchta et al., 2009) (Figure 1A).

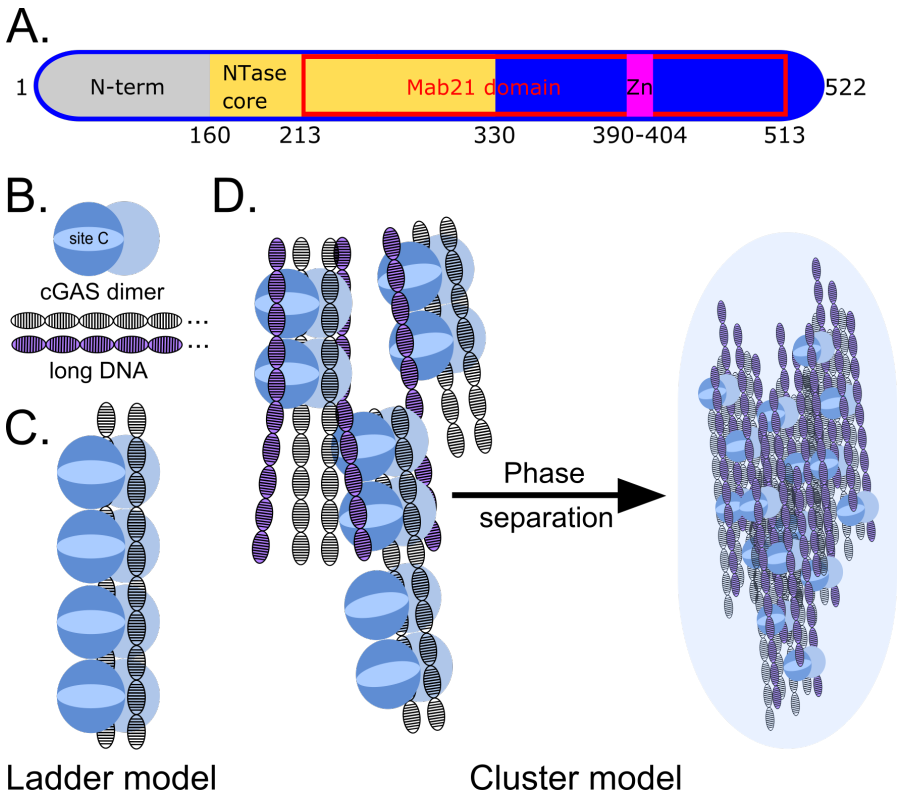

Figure 1. Domain organization and models of cGAS activation.
A. cGAS is composed of 513 residues. It starts with unstructured N-terminal domain (here "N-term") (aa 1-160), Ntase core (aa 160-330) that overlapps with Mab21 domain (aa 213-513). The last module contains a zinc thumb structure important for dsDNA recognition. B. cGAS dimerizes in contact with long, >45 bp dsDNA. It has three DNA-binding sites. Two of them, sites A and B, are in between the dimer interface (here not marked) and the last one, site C, is on the opposite side of the protein. C. Initially proposed cGAS activation model assumed that while binding with sites A and B to the dsDNA, cGAS dimers form a ladder-like structure to enhance production of cGAMP ("the ladder model"). D. Current activation model includes the third binding site, site C. cGAS molecules not only make ladder-like structures, but also cluster together with other "ladders" via binding site C. It is followed by phase separation and formation of efficient cGAMP-producing "membraneless organelles" ("the cluster model").

2.3. Mechanism of cGAS activation

Each cGAS monomer contains three DNA binding sites, called site A, B and C. These regions are rich in positively charged residues that attract negatively charged surfaces of dsDNA of any sequence. The first step of cGAS activation is its dimerization around two molecules of this nucleic acid. One monomer attaches to the DNA molecule through its binding site A, while the other monomer attaches to the opposite side of the target using its binding site B. Then more cGAS molecules attach to the string of the nucleic acid, and they form a structure resembling a ladder. This is why this model of activation was named the "ladder model" (Andreeva et al., 2017) (Figure 1B and C). After DNA binding, a conformational change occurs between the cGAS' two lobes that enables an active site

formation, and a proper interaction with ATP and GTP to synthesize 2',3'-cGAMP (Civril et al., 2013). The process is followed by a liquid phase separation event, mediated by positively charged cGAS N-terminal domain (Du et al., 2018). The phase separation is a process of migration of various molecules in aqueous solution into two phases, drastically differing in their concentrations, and thus forming a condensed environment that enhances signaling pathways inside the living cells (Xiao et al., 2021). Such concentrated liquid droplets forming in the cytoplasm are sometimes called the "membraneless organelles" (Marnik et al., 2019). The aforementioned model of activation does not explain how exactly it can result in this process. This fenomenon was explained later, after the third DNA binding site was identified. The site C is located on the opposite site of each monomer than other binding sites. This causes a possibility that the individual ladders can attach to each other via an additional dsDNA molecule to form larger and larger clusters. This is why currently accepted model of cGAS activation is called the "cluster model". The formation of such molecular structures was proved to initiate the phase separation event. High concentration of cGAS in membraneless organelles increases the stability of its dimers and facilitates production of mammalian signal molecule, 2',3'-cGAMP (Sun et al., 2013, Xie et al., 2019) (Figure 1D).

2.4. Factors important for cGAS activation

The first critical factor for cGAS activation is DNA length. The size of cGAS dimers enables them to bind to a very short DNA of less than 20 base pairs (bp). However, Andreeva et al. have shown in experiments on human monocyte cell line that such dimers have low stability and cGAS requires at least 45 bp of DNA to be efficiently activated (Andreeva et al., 2017).

Other factors without which the proper function of cGAS is impossible are zinc ions, critical for DNA binding, and magnesium ions needed for 2',3'-cGAMP formation (Civril et al., 2013). There is also an additional ion which was shown to be able to boost cGAS DNA sensing and activation - manganese. Mn^{2+} ions were even proved to activate cGAS themselves in absence of DNA, and due to that cause 2',3'-cGAMP synthesis. However, while Mg^{2+} ions are constantly available in the cytoplasm, Mn^{2+} ions are stored inside the Golgi apparatus and mitochondria, and are released only in special cases, for example, during a viral infection (Zhao et al., 2020).

2.5. Cellular location of cGAS

Since its discovery in 2013, cGAS was thought to be mainly a cytoplasmic protein, where DNA is present only in an event of genotoxic stress or intracellular infection, and this is how it was believed to be separated from the self DNA (Civril et al., 2013). However, the missing element in this assumption was the fact that most of the DNA viruses use nucleus for their replication and their DNA does not have actual contact with cytosol (Whittaker et al., 2000). Furthermore, multiple studies confirmed that cGAS activates cellular defense mechanisms in response to these viruses. Additionally, retroviral genetic material is reverse transcribed

inside a viral capsid upon entry into a cell and the resulting DNA is still recognized by cGAS (Volkman et al., 2019).

To explain this conundrum, Volkman et al. demonstrated, using fractination techniques and confocal microscopy, that most of the cellular cGAS is confined to the nuclear compartment, where it remains bound to chromatin (Volkman et al., 2019). This interaction was further characterized structurally by Michalski et al. to be mediated by evolutionary conserved loops that bind to the acidic patch regions on histones 2A and 2B. The catalytic activity of cGAS is inhibited in this form, because it has much higher binding affinity to histones than to DNA. One of the DNA binding sites, site B, is blocked by this interaction and it prevents cGAS dimerization while it resides in the nucleus (Michalski et al., 2020).

In contrast, another group reported that cGAS may be not a cytoplasmic protein at all and it is constantly attached to an internal side of the plasma membrane, thanks to interaction of its N-terminus with membrane lipid phosphatidylinositol 4,5-bisphosphate. This cellular location may facilitate detection of viruses directly after endocytosis, and thus may explain how this PRR differentiates between self and foreign DNA (Barnett et al., 2019).

Except for that, cGAS was identified as a direct detector of cyclic dinucleotides (CDNs) that enter cells upon endocytosis. cGAS localizes to endosomes, binds CDNs, dimerizes and forms complexes with STING leading to its activation and type I interferon pathway advancement (Liu et al., 2019).

It seems that cGAS resides in multiple cellular locations and more research is needed to fully explain the nature of its activity.

2.6. Cyclic dinucleotides as signal molecules

Earlier cyclic dinucleotides, like 3',5'-cyclic (c-) di-AMP, 3',5'-c-di-GMP and 3',5'-cGAMP were characterized only in bacteria where they constitute signal molecules for diverse processes, including expression of various genes, chemotaxis, motility, bacterial biofilm formation and other (Aline Dias da, 2020). Later, the mammalian counterpart, 2',3'-cGAMP was found. It has characteristic mixed phosphodiester bonds between 5'-phosphate of adenosine monophosphate (AMP) and 2'-OH of guanosine monophosphate (GMP), and on the other end, between 5'-phosphate of GMP and 3'-OH of AMP. This dinucleotide is the primary ligand of STING (Zhang et al., 2013).

2.7. Stimulator of interferon genes

2.7.1. Structure of STING

The stimulator of interferon genes (STING), previously described as ERIS, MPYS and MITA, is a transmembrane protein that in its inactive state resides in the endoplasmic reticulum (ER) in a dimeric form. Each 40 kDa monomer contains four N-terminal transmembrane helices (TMs) (aa 1-140) followed by the cytoplasmic carboxy-terminal domain (CTD) subdivided into the cyclic dinucleotide-binding domain (CBD) (aa 157-336) and the short carboxy-terminal tail (CTT) (aa 337-379). CTD is connected to the transmembrane part by a short connector helix (CHX) (aa 141-149), critical for protein oligomerization, and a connector loop (CLP) (aa 150-156). CTT comprises the regions important for signal transduction, which are: pLxIS motif (p - hydrophylic aa, x - any aa, S - phosphorylation site) (aa 362-366) which is critical for interferon regulatory factor 3 (IRF3) recruitment and activation, serine366 which is phosphorylated by TANK-binding kinase 1 (TBK1) or IκB kinase ε (IKKε), also described as IKK-inducible (IKK-i), and PLPLRT/SD motif to which TBK1 binds (aa 371-377) (Zhao et al., 2016, Zhao et al., 2019).

In the dimeric form, two CTDs form a butterfly-shaped structure with a ligand-binding site between its wings (Ishikawa et al., 2008, Sun et al., 2009, Sun et al., 2013, Zhong et al., 2008). Full-length cryoelectron microscopy (cryo-EM) structures of chicken STING with 2',3'-cGAMP bound have shown a characteristic domain-swapped architecture of transmembrane region where transmembrane helix 1 (TM1) of one monomer is inserted between the second monomer's TM2-4. The TMs of the dimer form a cylindrical structure with TMs2 and TMs4 in the core, surrounded by TMs1 and TMs3. There are four alpha helices and five beta strands in CBD. Its helix 1 is connected to TM4 by the connector helix and the connector loop (Shang et al., 2019) (Figure 2).

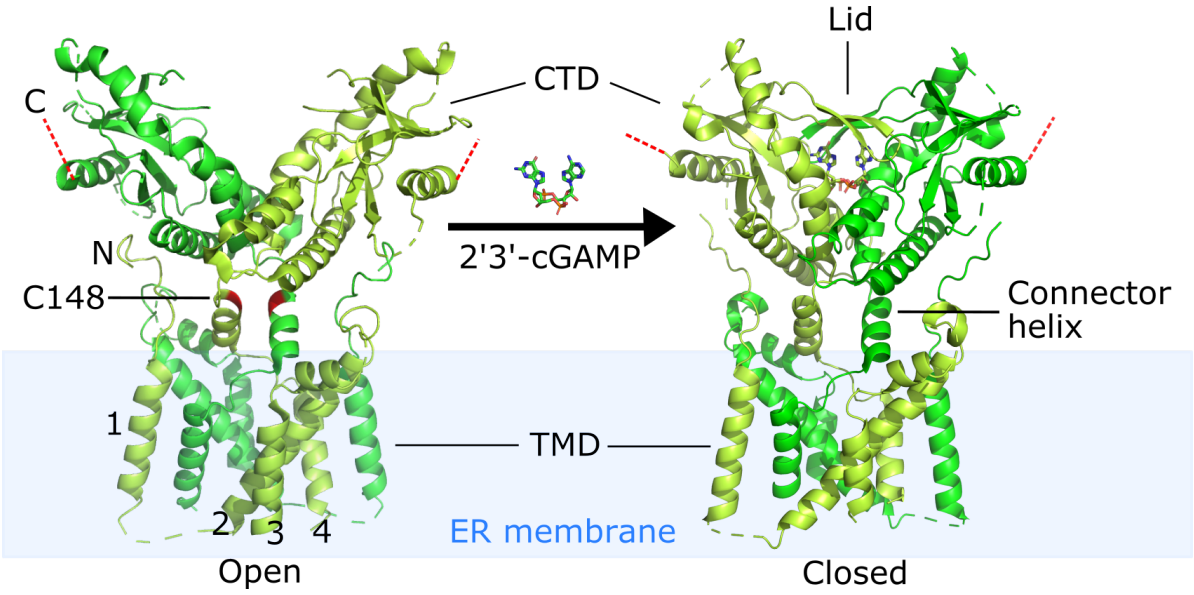

Figure 2. Structure of STING (previous page).
The models deposited in Protein Data Bank (PDB) with accession codes 6NT5 (human, open) and 6NT7 (chicken, closed) are shown as the cartoon representations. The monomers forming a single STING dimer are shown in limon and green, respectively. cGAMP is shown as sticks (C - green, H - grey, N - blue, O - red). STING is composed of the C-teminal domain (CTD) containing C-terminal tail, required for downstream signaling (represented by red, dashed lines) and the transmembrane domain (TMD), embedded in ER membrane, comprising four transmembrane helices per monomer, which are marked with Latin numbers. They are connected by the connector helix. Upon ligand binding, the CTD changes its conformation to a closed one, and a lid composed of four antiparallel beta sheets is formed. Except for that, the CTD makes 180 degrees turn relative to TMD, and cystein residues 148 (red), critical for protein oligomerization, get exposed. C-terminus and N-terminus is marked with letters C and N, respectively.

2.7.2. Mechanism of STING activation

In an inactive state, STING is kept in the ER by stromal interaction molecule 1 (STIM1), a Ca2+ sensor (Srikanth et al., 2019). Additionally, it is stabilized by toll-interacting protein (TOLLIP) to prevent its proteolytic degradation (Pokatayev et al., 2020). The CTT was shown to be another autoinhibitory element. It performs this function probably by sticking to a part of STING polymerization interface thus it is not accessible for partner proteins (Ergun et al., 2019, Shang et al., 2019).

Upon ligand binding, two wings of STING, formed by the adjacent CTDs, move towards each other and form a ligand binding pocket covered by a lid composed of four antiparallel beta sheets (Ergun et al., 2019, Shang et al., 2019). After this conformational change, the CTTs are released from their resting positions that makes them accessible for partner proteins binding. Also, the whole dimeric CTD makes a 180 degrees turn that exposes the cystein148 residues in the connector helices that causes protein oligomerization by disulfide bonds formation (Ergun et al., 2019, Shang et al., 2019) (Figure 3, 1-2). Mutations that prevent protein oligomerization were shown to block STING translocation from ER to Golgi apparatus and activation of TBK1, highlighting the importance of this occurence for the function of this protein (Ergun et al., 2019).

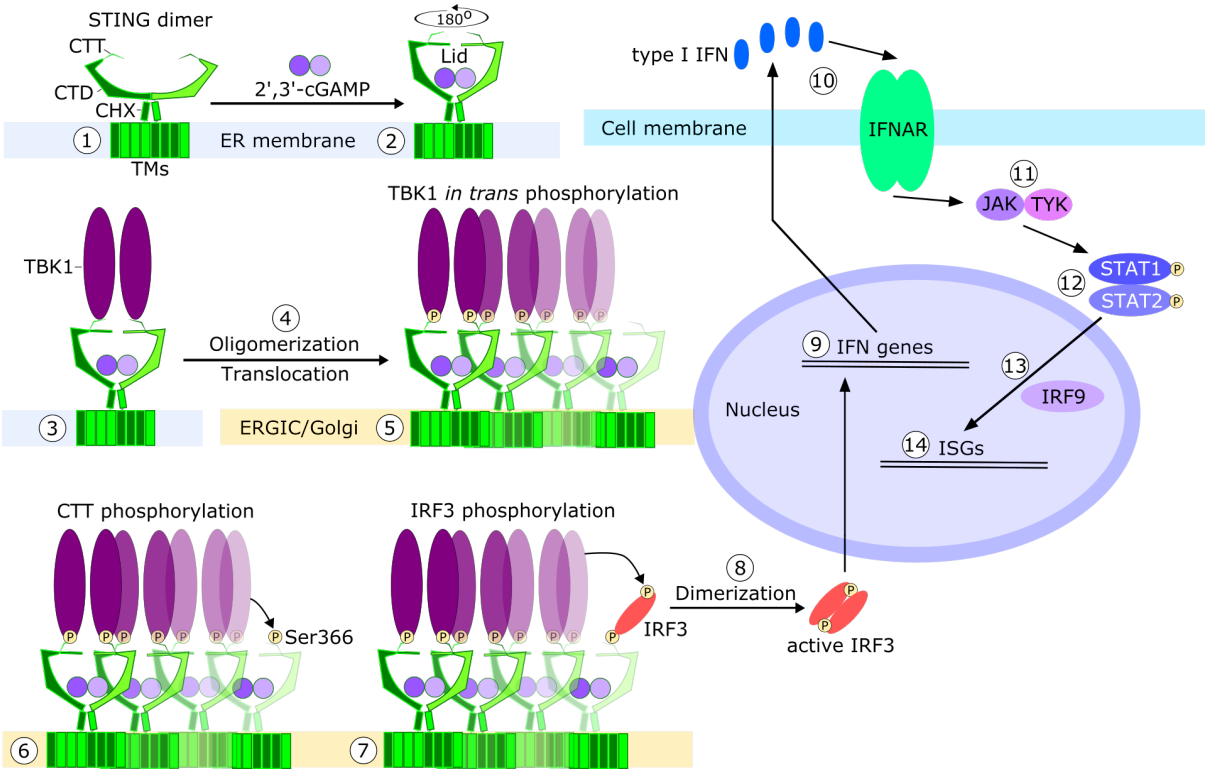

Figure 3. STING pathway.
1 - Inactive dimeric STING in an open conformation resides in the ER membrane, 2 - Upon 2',3'-cGAMP binding STING changes its conformation to closed. The lid is formed at the top and the CTD makes 180 degrees turn relative to TMs, 3 - Two TBK1 molecules bind to CTTs and dimerize, 4 - The complex is translocated through ERGIC to Golgi apparatus and there is an oligomerization event on the way, 5 - TBK1s phosphorylate each other *in trans*, 6 - TBK1 phosphorylate serine366 in the free CTT, 7 - IRF3 binds to phosphorylated CTT and gets phosphorylated by TBK1, 8 - Phosphorylated IRF3s dimerize and migrate to the nucleus, 9 - IFN genes transcription and expression starts, 10 - Type I interferons are transported outside of the cell and bind to IFNARs, 11 - JAK and TYK tyrosine kinases activation, 12 - STAT1 and STAT2 phosphorylation and dimerization, 13 - STAT1 and STAT2 dimer translocation to the nucleus, association with IRF9 and activation of ISGs transcription.

Additionally, due to 2',3'-cGAMP binding to STING, the interaction with STIM1 is compromised, and STING starts to be able to bind the SEC24C. It is one the proteins of the cellular trafficking complex - coat protein complex II (COPII). COPII is the evolutionary conserved system that forms vesicles and transports lipids and proteins from ER to different cellular locations. The activated, polymerized STING migrates together with COPII through ER-Golgi intermediate compartment (ERGIC) to Golgi apparatus where its cysteine88 and 91 residues, located in the proximity of the membrane, are palmitoylated. This modification further promotes protein oligomerization critical for its activity. However, the precise role of palmitoylation still remains a mystery (Bisnett et al., 2021, Mukai et al., 2016, Srikanth et al., 2019). During translocation STING couples with the TANK-binding kinase 1 (TBK1) through interaction with PLPLRT/SD motif. TBK1 forms a dimer and activates itself by phosphorylation (Figure 3, 3-5). Then TBK1 phosphorylates serine366 in pLxIS motif of

STING CTT which causes recruitment of the interferon regulatory factor 3 (IRF3). IRF3 dimerizes and is transported to the nucleus where it initializes a transcription of type I interferons (IFNs). The precise location of aforementioned interactions was not fully determined (Liu et al., 2015) (Figure 3, 6-9).

2.7.3. Type I interferon signaling

Human type I interferon family consists of the interferon beta (IFNβ), 13 types of interferon alpha (IFNα) and six other, less characterized interferons (McNab et al., 2015). Interferons alpha and beta bind to the transmembrane, cell surface interferon-α/β receptor (IFNAR), comprising IFNAR1 and IFNAR2 subunits. Activation of IFNAR leads to activation of two kinases - tyrosine kinase 2 (TYK2) and Janus kinase 1 (JAK1) which phophorylate signal transducer and activator of transcription 1 and 2 (STAT1 and STAT2). STAT1 and STAT2 then dimerize and translocate to the nucleus. Together with interferon regulatory factor 9 (IRF9) they create the interferon-stimulated gene factor 3 (ISGF3) complex that constitutes a transcription factor which binds to the interferon-stimulated response elements (ISRE) located in the interferon-stimulated gene promoters. This is how interferon induces transcription of hundreds of interferon-stimulated genes (ISGs) that initiate powerful innate immune response to, for instance, a viral infection (McNab et al., 2015) (Figure 3, 10-14).

2.8. Evolution of cGAS and STING

cGAS and STING proteins are highly evolutionary conserved and their homologs were found even in organisms as simple as bacteria.

2.8.1. Evolution of cGAS

cGAS, together with its homolog dinucleotide cyclase in Vibrio (DncV) that produces 3',3'-cGAMP in *Vibrio cholerae*, were found to form a large protein family, collectively called cGAS/DncV-like nucleotidyltransferase (CD-NTase) proteins. There is a large number of bacterial and eukaryotic enzymes within this family, including CdnD from *Enterobacter cloacae*, CdnE from *Escherichia coli* and metazoan OAS1 protein. They produce 3',3',3'-cAMP–AMP–GMP, 3',3'-cUMP–AMP and 2',5'-oligoadenylates, respectively (Kranzusch, 2019).

In *V. cholerae* cyclic dinucleotide, 3',3'-cGAMP, is produced upon bacteriophage infection that leads to initiation of cell destruction to prevent spreading of an infection among bacterial population. This shows that even first cGAS-like proteins evolved as a part of ancient host defense pathways (Cohen et al., 2019). However, there is still little knowledge on ligands and modes of action of these predecessors of cGAS (Kranzusch et al., 2013). According to current research, cGAS started to be able to bind dsDNA upon evolutionary addition of zinc-thumb domain to its CTD that occured in vertebrates (Kranzusch et al., 2013).

2.8.2. Evolution of STING

The bacterial homologs of STING were discovered using bioinformatic tools. Multiple crystal structures of bacterial STING from *Capnocytophaga granulosa* and a number of species of Flavobacteriaceae showed a characteristic V-shaped dimeric STING organization similar to eukaryotic protein. Comparing to metazoan counterparts, the bacterial STING was shown to be approximately 20% smaller, and to be able to bind cyclic dinuleotides, as 3',5'-cGAMP, after which it forms the lid, closing the ligand binding pocket (Morehouse et al., 2020). The structural alignement of human and bacterial STING is shown in Figure 4A.

The early form of eukaryotic STING was found in *Monosiga brevicollis*, which is a choanoflagellate organism that preceeds the metazoans and is their closest known living relative (Wu et al., 2014).

The primitive animal STING was described in the sea anemone that diverged from humans approximately 500 millions years ago, called *Nematostella vectensis*. The three dimensional structure of this protein is fairly similar to human homolog, despite having just 29% sequence homology. It was also shown to be able to bind 2',3'-cGAMP with high affinity of less than 1 nM (Kranzusch et al., 2015). Its structural alignement with human STING is shown in Figure 4B. The primary function of this protein is an activation of autophagy (Margolis et al., 2017). In invertebrates, STING is able to activate proinflammatory nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway in response to pathogenic infection. For instance, STING from *Drosophila melanogaster* was shown to detect cyclic dinucleotides from *Listeria monocytogenes* that elevates the expression of defense peptides, which is mediated by Relish transcription factor, a component of NF-κB pathway that increases chances of survival (Martin et al., 2018). In contrast, the most evolutionary advanced form of STING, present in the majority of vertebrate organisms, additionally contains a short, unstructured carboxy-terminal tail (CTT), which is essential for downstream signaling that leads to production of type I IFNs and other cytokines (Gui et al., 2019). But there are some exceptions, a number of amphibians, for instance, *Xenopus laevis* and *X. tropicalis* do not possess the above-mentioned motif in their STING molecules (Irazoqui et al., 2010). In contrast, roundworms, flatworms, several insect species and a small number of other animals, along with other eukaryoic organisms, like plants, protists and fungi, were shown not to possess any cGAS and STING homologs at all (Wu et al., 2014). Additionally, roundworms evolved without other innate immune components of vertebrates, like NF-κB pathway (Irazoqui et al., 2010).

Interestingly, analysis of bat CTT sequences by Xie at al., showed that the conserved serine358, which in humans is essential for IRF3 binding, is substituted in these animals to a range of other amino acids. The authors found that this mutation downregulates type I IFN response in their organisms. Bats require a fast cellular metabolism, as an adaptation for flight, which causes lots of DNA damage, which results in cytoplasmic DNA accummulation. They also constitute a reservoir of a large number of viruses with which they coexist mostly without suffering from any disease. The hypothesis is, that to survive, bats decreased innate

immune response to DNA and viruses by implementing the aforementioned mutations into their CTTs (Tanaka et al., 2012, Xie et al., 2018).

The evolution of components of cGAS-STING pathway in selected organisms is graphically summarized in Figure 4C.

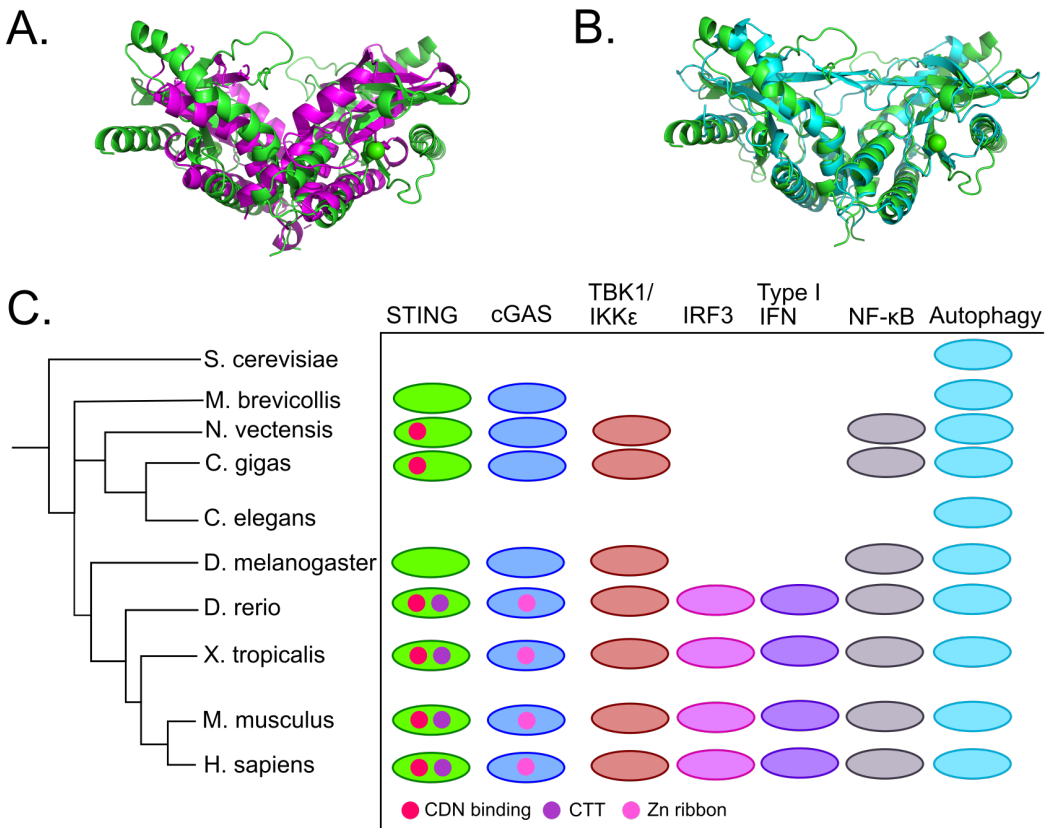

Figure 4. Evolution of cGAS-STING pathway.
A. and B. Alignements of crystal structures of dimeric soluble domains of human STING (green) (PDB accession code 4F5W) and *Capnocytophaga granulosa* STING (magenta) (PDB accession code 6WT5) (A) or *Nematostella vectensis* STING (cyan) (PDB accession code 5CFO) (B) show major structural similarities, especially in butterfly-shaped CTD, between bacterial, early metazoan and human STING. C. Conserved components of cGAS-STING pathway (based on Margolis et al., 2017).

2.9. Regulation of cGAS and STING

cGAS-STING pathway defends cells from pathogens by initializing type I interferons and proinflammatory cytokines expression which is critical for survival. On the other hand, prolonged overactivation of these proteins may lead to chronic inflammation and disease. This is why cGAS and STING are tightly regulated by interaction with other proteins, post-translational modifications (PTMs), proteolytic cleavage and by restricting ligand availability. For the long time, researchers considered also a cellular compartmentalization as a factor which regulates activity of cGAS by not allowing self-DNA to activate this protein.

However, recent findings showing that cGAS is not only a cytoplasmic protein have challenged this concept (Motwani et al., 2019a).

2.9.1. Ligand availability

One of the mechanisms of negative regulation of proteins is restriction of their ligand availability. For cGAS it is dsDNA of any sequence, and for STING, 2',3'-cGAMP, and other cyclic dinucleotides (Santa et al., 2021).

There is a plethora of DNA-digesting enzymes that reside in extra- and intracellular environment and prevent accumulation of DNA. Outside of the cell there are deoxyribonuclease 1 (DNase 1), DNase 1-like 1 and DNase 1-like 3 that cleave DNA released from dying cells and pathogens. Inside the cell, there are DNase 2a and 2b that digest DNA imported from the exterior in endolysosomes, and DNase 3, more commonly known as the three-prime repair exonuclease 1 (TREX1) that is a transmembrane resident of ER, which engages dsDNA present in the cytoplasm (Santa et al., 2021).

Cyclic dinucleotides are actively digested by phosphodiestrases. Extracellular example of a transmembrane protein proved to cleave 2',3'-cGAMP is ecto-nucleotide pyrophosphatase phosphodiestrase 1 (ENPP1). It was initially described as the protein that prevents bone mineralization. However, the same or even more important, intracellular counterpart, was not yet identified (Kato et al., 2018).

2'3'-cGAMP was also shown to be able to cross gap junctions and to spontaneously spread an innate immune response to the neighboring cells in order to prevent spreading of infection (Ablasser et al., 2013). Except for that it can be encapsulated into viral capsids, and sent to the far away destinations, where it can activate immune response much quicker, right upon viral entry (Bridgeman et al., 2015).

2.9.2. Post-translational modifications of cGAS

cGAS function is inhibited by a wide range of post-translational modifications (PTMs), including acetylation by an unknown enzyme (Dai et al., 2019), phosphorylation by Akt kinase (Seo et al., 2015), monoglutamylation by tubulin tyrosine ligase-like 4 (TTLL4), polyglutamylation by TTLL6 (Xia et al., 2016) and sumoylation by an unknown enzyme (Cui et al., 2017). Additionally, phosphorylation by B-lymphoid tyrosine kinase keeps cGAS in cytoplasmic compartment (Liu et al., 2018). Other set of modifications promotes cGAS proteolytic degradation, and includes desumoylation by sentrin/SUMO-specific protease 2 (SENP2) (Hu et al., 2016), K48-linked polyubiquitination by an unknown enzyme (Chen et al., 2016) and cleavage by the inflammasome-activated caspase 1 and 3 that inactivates the protein (Ning et al., 2019, Wang et al., 2017a).

cGAS function is promoted by deglutamylation by cytosolic carboxypeptidase 5 (CCP5) and CCP6 (Xia et al., 2016), K27-linked polyubiquitination by the E3 ubiquitin ligase RNF185 (Wang et al., 2017b) and monoubiquitynation by the E3 ubiquitin ligase tripartite motif containing 56 (TRIM56) (Seo et al., 2018). The protein is additionally stabilized by sumoylation by TRIM38 (Hu et al., 2016), and deubiquitination by TRIM14 and USP14 complex, which prevents its degradation by autophagy (Chen et al., 2016).

Post-translational modifications of cGAS are summarized in Table 1.

2.9.3. Post-translational modifications of STING

The STING function is inhibited by dephosphorylation by the protein phosphatase, Mg^{2+}/Mn^{2+} dependent 1A (PPM1A) (Li et al., 2015), and nitro-alkylation occuring in presence of nitro-fatty acids, which prevents the palmitoylation of the protein (Hansen et al., 2018). STING is sent for degradation by K48-linked polyubiquitination by the E3 ubiquitin ligase - ring finger protein 5 (RNF5) (Zhong et al., 2009) and desumoylation by SENP2 (Hu et al., 2016).

The STING function is promoted by palmitoylation in Golgi apparatus by the palmitoyl-transferases - DHHC cysteine-rich domains-containing protein 3 (DHHC3), DHHC7 and DHHC15 (Mukai et al., 2016), phosphorylation by TBK1 (Tanaka et al., 2012), K63-linked polyubiquitination by TRIM32 (Zhang et al., 2012), TRIM56 (Tsuchida et al., 2010) and mitochondrial E3 ubiquitin protein ligase 1 (MUL1) (Ni et al., 2017), and K27-linked polyubiquitination by the E3 ubiquitin ligase AMFR and insulin-induced gene 1 (INSIG1) complex (Wang et al., 2014). The protein is stabilized by deubiquitynation by CYLD that removes K48-linked polyubiquitin (Zhang et al., 2018), and by sumoylation by TRIM38, which also facilitates STING activation (Hu et al., 2016).

Post-translational modifications of STING are summarized in Table 1.

Table 1. Postranslational modifications of human cGAS and STING.

Target	PTM	Enzyme	Location*	Result	Ref.
cGAS	Acetylation	Unknown	K384/394/414	Inhibition	Dai et al., 2019
	Deacetylation	HDAC3	K384	Lack of inhibition	Dai et al., 2019
	Phosphorylation	Akt kinase	S305	Inhibition of enzymatic activity	Seo et al., 2015
	Phosphorylation	B-lymphoid tyrosine kinase	Y215	Promotion of cytoplasmic retention	Liu et al., 2018

	Monoglutamylation	TTLL4	E314[a]	Inhibition of enzymatic activity	Xia et al., 2016
	Polyglutamylation	TTLL6	E286[b]	Inhibition of ligand binding	Xia et al., 2016
	Deglutamylation	CCP5	E314[c]	Promotion of enzymatic function	Xia et al., 2016
	Deglutamylation	CCP6	E286[d]	Promotion of ligand binding	Xia et al., 2016
	Sumoylation	TRIM38	K231/479	Stabilization and prevention of degradation	Hu et al., 2016
	Sumoylation	Unknown	K347/384/394[e]	Inhibition	Cui et al., 2017
	Desumoylation	SENP2	K231/479	Proteolytic degradation	Hu et al., 2016
	Desumoylation	SENP7	K347/384/394[f]	Lack of inhibition	Cui et al., 2017
	Cleavage	Caspase 1	D140/157	Inactivation	Wang et al., 2017a
	Cleavage	Caspase 3	D319	Inactivation	Ning et al., 2019
	Polyubiquitination (K27)	RNF185	K173/384	Promotion of enzymatic function	Wang et al., 2017b
	Polyubiquitination (K48)	Unknown	K414	Proteolytic degradation	Chen et al., 2016
	Monoubiquitination	TRIM56	K335	Promotion of enzymatic function	Seo et al., 2018
	Deubiquitynation	TRIM14/ USP14	K414	Stabilization	Chen et al., 2016
STING	Palmitoylation	DHHC3/ DHHC7/ DHHC15	C88/91	Enabling activation	Mukai et al., 2016
	Nitro-alkylation	Nitro-fatty acids	C88/91	Palmitoylation inhibition	Hansen et al., 2018
	Phosphorylation	TBK1	S366	Enabling IRF3 binding	Tanaka et al., 2012
	Dephosphorylation	PPM1A	S358	Inhibition of oligomerization	Li et al., 2015

	Polyubiquitination (K63)	TRIM32	K20/150/ 224/236	Facilitation of TBK1 binding	Zhang et al., 2012
	Polyubiquitination (K27)	AMFR/ INSIG1	K137/150/ 224/236	Facilitation of TBK1 binding	Wang et al., 2014
	Polyubiquitination (K11)	RNF26	K150	Inhibition of K48 ubiquitination	Qin et al., 2014
	Polyubiquitination (K48)	RNF5	K150	Proteolytic degradation	Zhong et al., 2009
	Polyubiquitination (K63)	TRIM56	K150	Promotion of di- and oligomeri-zation	Tsuchida et al., 2010
	Polyubiquitination (K63)	MUL1	K224	Promotion of IRF3 phosphory-lation	Ni et al., 2017
	Deubiquitination	CYLD	K150	Stabilization	Zhang et al., 2018
	Sumoylation	TRIM38	K338	Stabilization and promoting activation	Hu et al., 2016
	Desumoylation	SENP2	K338	Promotion of degradation	Hu et al., 2016

[*] - locations normalized to amino acid number corresponding to human protein, [a] and [c] - mouse E302, [b] and [d] - mouse E272, [e] and [f] - mouse K335/372/382, respectively. Table based on Cheng et al., 2020, Hopfner et al., 2020, Motwani et al., 2019a.

2.9.4. Interactions of cGAS with other proteins

cGAS action was shown to be modulated by a range of partner proteins. The first of them is polyglutamine binding protein 1 (PQBP1) that was found to detect the DNA of human immunodeficiency virus type 1 (HIV-1), produced by reverse transcription machinery in dendritic cells, and to promote cGAS activation and cellular response to a viral infection (Yoh et al., 2015). Other example is non-POU domain-containing octamer-binding protein (NONO) that recognizes HIV-2 capsid proteins in the nucleus, and interacts with cGAS to facilitate its activation. Otherwise, encapsulated DNA from reverse transcription of HIV-2 genome constitutes a weak immunogen, and could enable the virus to evade detection (Lahaye et al., 2018). Researchers also discovered two proteins that enhance detection of dsDNA by cGAS. The first of them is zinc finger CCHC-type containing 3 (ZCCHC3) protein that was shown to cooperate with cGAS in dsDNA sensing and to promote its activation (Lian et al., 2018). The other is the G3BP stress granule assembly factor 1

(G3BP1), also known as Ras-GTPase-activating protein, SH3 domain-binding protein 1. When its gene was deleted from human monocytes their ability to express type I interferons was downregulated, despite exposing them to dsDNA source, suggesting that it is essential for proper activation of cGAS (Liu et al., 2019).

2.9.5. Interactions of STING with other proteins

As mentioned in Section 2.5.2., a resting state of STING in ER is mediated by the interaction of its transmembrane part with other transmembrane protein - STIM1, a Ca^{2+} sensor, maintaining high levels of calcium ions in the ER lumen. Srikanth et al. have shown that STING pathway gets activated in absence of infection in the cells that do not express STIM1, and that the resistance of such cells to viral infections is strongly elevated. The STING action was also shown to be inhibited by overexpression of STIM1, which additionally can mitigate the inflammatory effects of constatly activated STING mutants. Interstingly, STING was also found to influence STIM1 by preventing its transport to the proximity of plasma membrane calcium channels that normally causes influx of Ca^{2+} to the ER, when reserves get depleated (Srikanth et al., 2019).

Additionally, other ER resident - toll-interacting protein (TOLLIP) was shown to have a stabilizing effect on STING in its inactive state. In mice with a deficiency of TREX1 DNase, which causes a DNA accumulation and an inflammatory disease, the researchers have shown that knock out of the *Tollip* gene mitigated a development of such a disease and increased their survival. It highlights the TOLLIP importance in protecting STING from degradation by lysosomes and maintaining its physiological levels (Pokatayev et al., 2020).

Also, Jonsson at al. reported that other DNA sensor - interferon-gamma inducible factor 16 (IFI16) promotes binding of TBK1 to STING along with a resulting phosphorylation and protein activation (Jonsson et al., 2017).

2.10. STING-mediated activation of other pathways

STING was shown to be a versatile adaptor protein of innate immune system that is a critical element in type I interferon pathway. Except for that, it can activate or affect multiple other pathways, including NF-κB pathway, autophagy, apoptosis, lysosome-dependent cell death and necroptosis.

2.10.1. STING-mediated activation of NF-κB pathway

2.10.1.1. NF-κB pathway

Nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway is an immune cascade that leads to proinflammatory cytokines expression that results in inflammation. Additionally, it can modulate activation of inflammasome and influence

differentiation and activation of inflammatory T cells. Inbalance in NF-κB pathway may lead to inflammatory disorders, for instance, rheumatoid arthritis, systemic lupus erythematosus, inflammatory bowel disease and cancer (Dabek et al., 2010, Yu et al., 2020). There are two major types of NF-κB pathway - canonical and non-canonical (Liu et al., 2017). While activation of the former is fast and temporary, activation of the latter is slow and long-term. This is why, especially, overactivation of non-canonical pathway may lead to development of very serious inflammatory disorders (Sun, 2017).

2.10.1.1.1. Canonical pathway

The canonical pathway is activated in response to a range of ligands binding to B-cell and T-cell receptors, proteins from tumor necrosis factor (TNF) receptor (TNFR) superfamily, cytokine receptors and PRRs. Among them, there are growth factors, fragments of microorganisms, cytokines, and stress-inducing and mitogenic substances. Its main function is initiation of inflammatory reaction. The first component of this pathway is TGFβ- activated kinase 1 (TAK1), also called MAP3K7. This kinase activates IκB kinase (IKK) complex that is composed of three subunits: IKKα, IKKβ and a regulatory component IKKγ, also called NF-κB essential modulator (NEMO) (Liu et al., 2017, Sun, 2017, Verhelst et al., 2013). IKKγ can bind K63-linked ubiquitin present on cytoplasmic signal proteins, activated by various receptor molecules responding to inflammatory stimuli (Gilmore, 2006). Activated IKK complex phosphorylates IκBα, which is an inhibitory element of RelA-p50, also known as p65-NF-κB1, and c-Rel-p50 complexes. Afterwards, K48-linked polyubiquitination of IκBα sends it for proteosomal degradation, and causes RelA/p50 and c-Rel/p50 complexes activation, and their translocation to the nucleus where they initiate transcription of a plethora of inflammatory genes (Gilmore, 2006, Liu et al., 2017, Sun, 2017).

2.10.1.1.2. Non-canonical pathway

The non-canonical pathway is activated by ligands of B-cell-activating factor receptor (BAFFR), receptor activator for nuclear factor κB (RANK), cluster of differentiation 40 (CD40) and lymphotoxin β-receptor (LTβR), and multiple other receptors from TNFR superfamily. It supports canonical pathway in modulation of innate and adaptive immunity, and is also needed in B cells and lymphoid organs development, and osteoclasts differentiation (Sun, 2017). Two kinases take part in this cascade: NF-κB-inducing kinase (NIK) and IKKα, which is activated by the former protein (Sun, 2011). The resting levels of NIK are very low thanks to its interaction with TNFR-associated factor 2 (TRAF2) and TRAF3. The latter initiates K48-linked ubiquitination of NIK by cellular inhibitor of apoptosis 1 (cIAP1) and cIAP2, and their proteolytic degradation. When this pathway is activated, either disruption of TRAF2-TRAF3-cIAP1-cIAP2-NIK complex or degradation of TRAF3 occurs which increases cellular levels of NIK. Such degradation is mediated by ubiquitination of TRAF3 by cIAPs, which is initiated by their ubiquitination by TRAF2 (Sun, 2017). Then p100 processing occurs. This precursor of of NF-κB2 protein, also called p52, is first phosphorylated, and then ubiquitinated. Afterwards, p100 is processed to mature NF-

κB2 which forms complex with RelB. Then the complex migrates to the nucleus to start transcription of inflammatory genes (Liu et al., 2017, Sun, 2017, Verhelst et al., 2013, Yu et al., 2020).

There are also two other non-canonical IKKs - IκB kinase ε (IKKε), also described as IKK-inducible (IKK-i), and TANK-binding kinase 1 (TBK1), which are able to phosphorylate IκBα as well as interferon regulatory factors 3 and 7 (IRF3 and IRF7) (Verhelst et al., 2013).

2.10.1.2. STING-mediated activation of NF-κB pathway

It was noticed by multiple research groups that there is an interlink between NF-κB pathway and STING signaling. However, this topic requires further research, since multiple reports have been published that contradict each other.

On one hand, TBK1 was shown to be essential for STING-mediated activation of NF-κB pathway. Abe and Barber have shown in experiments on mouse embryonic fibroblast cells deficient in TBK1, or TBK1-silenced, that upon their stimulation with dsDNA, the expression of genes activated by NF-κB was much lower than in wild-type cells. This suggests that TBK1 is an essential element of STING-NF-κB signaling pathway (Abe and Barber, 2014).

On the other hand, the mouse models with STING that contained mutations of either serine365 or leucine373 (corresponding to human serine366 and leucine374, respectively) to alanine, and a one with deleted CTT, showed that when type I interferon pathway is compromised by either of these mutations, NF-κB pathway and autophagy can still be induced by STING. The serine365 mutation, which prevents IRF3 binding, but still allows TBK1 recruitment, does not compromise cell ability to fight herpes simplex virus type 1 (HSV-1) and cancer cells. However, the leucine373 mutation that prevents TBK1 binding to STING voids these abilities, showing that TBK1 is important in interferon-independent response. Interestingly, this mutant is able to initiate autophagy, which is not sufficient for the defense from viruses, since the mice with this mutation were not resistant to HSV-1 infection. It was also shown that there is an increase of transcription of NF-κB-stimulated genes: 4-1BB ligand (4-1BBL), C-X-C motif chemokine ligand 1 (CXCL1) and CXCL2 in cells with serine365 mutation, which can explain their resistance to viral infection. The mice possessing STING without CTT were not able to perform autophagy, and were not resistant to HSV-1 infection highlighting its importance in this pathway (Yum et al., 2021).

2.10.2. STING-mediated autophagy

2.10.2.1. Autophagy

Autophagy (from Greek - “self eating”) is a cellular recycling process that provides a cell with all necessary nutrients in case of stress in form of, for instance, starvation, infection,

oxidative stress, protein aggregation, hypoxia and decreased level of growth factors. It also recycles all not needed or broken organelles, not properly folded or mulfunctioning proteins, and other material, keeping cells safe. The process is mediated by a wide range of autophagy-related proteins (ATGs), and requires formation of membrane encapsulated vesicles that transport a cargo and expose it to lysosomes that contain acidic hydrolases that digest unwanted material. The autophagy is a strictly regulated and selective process, since it is connected to many deadly diseases, such as cancer and neurodegeneration. It also takes part in defense against infections by digesting material derived from pathogens.

The first step of autophagy is creation of a phagophore by the multiprotein complexes, called Unc-51-like kinase 1 (ULK1) complex (containing: ULK1, RB1-inducible coiled-coil protein 1 - FIP200, autophagy-related protein 13 (ATG13) and ATG101) and class III PI3K (PI3KC3) complex I (containing: PI3KC3, Beclin 1, vacuolar protein sorting 34 (VSP34), activating molecule in Beclin 1-regulated autophagy protein 1 - AMBRA1, ATG14, general vesicular transport factor - p115), which requires phophorylation of PI3KC3 complex I, that engages omegasome region on ER. Upon that, the synthesis of phosphatidylinositol-3-phosphate (PI3P) starts, and zinc-finger FYVE domain-containing protein 1 (DFCP1) together with WD repeat domain phosphoinositide-interacting protein 2 (WIPI2) bind to a piece of membrane isolated from omegasome. Afterwards, ATG5-ATG12-ATG16L1 complex is recruited together with ATG3 protein and ATG8 family proteins, which include two subfamilies: γ-aminobutyric acid receptor-associated proteins (GABARAPs) or light chain 3 proteins (LC3Ps) that conjugate with phosphatidylethanolamine, and expand the omegasome-derived membrane around the cargo. This ends with the vesicule sealing, and autophagosome formation. The membrane material for this occurence is also delivered by vesicles that contain ATG9, and other not precisely defined sources. Then the autophagosome matures, and fuses with lysosome to form autolysosome which lyses the transported material. Recycled material is then available for the cell machinery to synthesize new proteins and organelles (Dikic and Elazar, 2018).

2.10.2.2. STING-mediated autophagy

Mediating autophagy was reported to be most probably an ancient, initial function of the stimulator of interferon genes, since it occurs even in organisms, like *Nematostella vectensis* and *Xenopus tropicalis*, that do not have the C-terminal tail, necessary for type I interferon pathway dependent on TBK1 and IRF3. The major purpose of this type of autophagy seems to be elimination of pathogens and their components, during an infection, and DNA from other sources, released into the cytoplasm (Gui et al., 2019). In mammalian cells, autophagy mediated by STING was shown to be a defense mechanism that is initiated during infections with HSV-1 (Yamashiro et al., 2020), *Mycobacterium tuberculosis* (Watson et al., 2012) and various types of Gram-positive bacteria (Moretti et al., 2017).

STING-activated autophagy starts after STING is translocated to ERGIC by COP-II-dependent vesicles. This process does not require classical initiation by ULK1 complex and

class III PI3K complex I, and takes place at the moment of WIPI2 recruitment to ER. At this point LCP3 uses lipids from ERGIC, in which STING resides, for autophagosome formation. During this process, the harmful cargo is encapsulated into a membrane, and sent for degradation (Gui et al., 2019).

2.10.3. STING in cell death

A healthy homeostasis requires controlled removal of unnecessary, old, mulfunctioning or infected cells. Any inbalance in cell death processes can lead to potentially lethal cardiovascular, neurodegenerative, autoimmune and infectious diseases as well as cancer. Cell death is also important for proper development and functioning of every organism.

2.10.3.1. cGAS-STING pathway in apoptosis

2.10.3.1.1. Apoptotic pathways

The apoptosis is an important physiological process that clears out damaged, old or no longer necessary cells from an organism, so they can be replaced by new cells. This process is strictly regulated, and any inbalance may lead to deadly diseases where there is either too high or too low cell clearance. Among them, there are autoimmune, neurodegenerative, cardivascular and infectious disorders. The apoptosis is the most throughly characterized controlled cell death pathway which is performed by caspases (cysteine-aspartic proteases). These proteases cleave the components of a cell before it is phagocytized by macrophages. The process takes place without release of cell material to an extracellular space. There are two apoptotic pathways: the intrinsic and extrinsic. The former is mediated by mitochondria and the latter by death receptors (Singh et al., 2019).

2.10.3.1.1.1. Extrinsic pathway

The extrinsic apoptosis pathway is mediated by a wide range of membrane proteins from the TNFR superfamily, located on the cell surface, called the "death receptors", and their ligands - cytokines from the TNF protein family. The common feature of these proteins is a presence of the "death domain" in their cytoplasmic portion. Upon ligand binding they form oligomers, and bind to partner proteins. This is followed by activation of caspases 8 and 10 that initiate apoptosis, or by activation of NF-κB and mitogen-activated protein kinase (MAPK) pathways that promote cell survival. The cellular fate here, depends on type of death receptor and adaptor proteins. The best characterized examples of death receptors are: tumor necrosis factor receptor 1 (TNFR1), TNF-related apoptosis-inducing ligand receptor 1 (TRAIL-R1), TRAIL-R2 and Fas (Guicciardi et al., 2009).

2.10.3.1.1.2. Intrinsic pathway

The intrinsic apoptosis is initiated by various stressors, including, reactive oxygen species, ionizing radiation, lack of nutrients and growth factors, chemotherapeutic agents, and mitogens. This pathway is regulated by the members of B-cell lymphoma 2 (BCL-2) protein family that can decide, if the particular cell will or will not survive. Among them there are: BH3-only activators, including p53-upregulated modulator of apoptosis (PUMA), BCL-2-interacting mediator of cell death (BIM) and BH3-interacting domain death agonist (BID); BH3-only sensitizers that inhibit pro-survival proteins, including Noxa, BCL-2-associated agonist of cell death (BAD), activator of apoptosis harakiri (Hrk); pro-survival proteins, including BCL-2, B cell lymphoma extra large (BCL-X_L), B cell lymphoma W (BCL-W), myeloid cell leukaemia 1 (MCL1) and BCL-2-related isolated from fetal liver (BFL1); pore-forming proteins - BCL-2-associated X protein (BAX) and BCL-2 antagonist/killer (BAK).

The apoptosis starts when pro-apoptotic BH3-only proteins activate the two pore-forming proteins: BAX and BAK, residing on the surface of every mitochondrion. The activated proteins form oligomers, and initiate mitochondrial outer membrane permeabilization (MOMP). This ends with a release of certain mitochondrial proteins into the cytoplasm. The purpose of these proteins is to activate caspases - the macromolecules that execute apoptosis. Cytochrom c is the direct caspase activator, but there are also proteins, including the serine protease OMI, and second mitochondria-derived activator of caspases (SMAC) that act indirectly by eliminating proteins that inhibit caspases. The example of such inhibitor is X-linked inhibitor of apoptosis protein (XIAP). When cytochrom c is released, its multiple molecules bind to apoptotic protease-activating factor 1 (APAF1), and the apoptosome formes. This structure activates initiator caspase 9 which performs activation of effector caspases 3, 6 and 7 that cleave cellular components, and prime the cell for phagocytosis. During this process phosphatidylserine gets externalized, DNA is cleaved, and the cell is fragmented (Singh et al., 2019).

The proteins engaged in apoptosis, constitute a large group of potential targets for future drugs against aforementioned diseases related to inbalance in apoptotic processes. Initially, the caspase inhibitors were tested as means to prevent apoptosis, but they failed, and only lead to delay in apoptosis. It was discovered that inhibition of caspases cannot prevent cell death, because MOMP is irreversible, and it compromises mitochondrial function in supplying the cell with energy. Furthermore, it releases mitochondrial proteins that are toxic for the cell. Such form of cell death is often called caspase-independent cell death (Tait et al., 2013). Recently, alternative therapeutic approaches are tested. Their goal is to modulate acitivity of various apoptotic proteins from BCL-2 family. It is done by using the mimicking compounds that can either induce or inhibit apoptosis. For example, BH3 mimetics are currently tested in therapy of various blood cancers (Singh et al., 2019).

2.10.3.1.2. cGAS-STING pathway in apoptosis

It was recently discovered that expanding BAX/BAK pores can also release mitochondrial DNA into the cytoplasm upon mitochondrial inner membrane rupture. In usual scenario, caspases inhibit cGAS by cleavage, but if caspases are not present, such DNA can activate cGAS-STING pathway, and lead to type I interferons production and inflammation. This observation could be used in cancer immunotherapy, where with use of caspase-blocking agents, interferon-mediated immunity, targeting cancer cells could be initiated (Riley et al., 2018, Singh et al., 2019).

The cGAS-STING pathway was also shown to be able to initiate apoptosis via intrinsic pathway that is dependent on TBK1 and IRF3. However, in this process, the transcriptional activity of IRF3 is not necessary. This protein was found to tether with BAX protein, and to promote apoptosis in response to a presence of reverse transcription intermediates from human T cell leukemia virus type 1 (HTLV-1) infection (Sze et al., 2013). This special apoptotic pathway was also found to be activated upon RNA virus detection by RIG-I-like receptors (RLRs) and this is why it is often called the RLR-induced, IRF3-mediated pathway of apoptosis (RIPA) or "viral apoptosis" (Chattopadhyay et al., 2010, Chattopadhyay et al., 2016).

The apoptosis was also shown to be activated by anti-tumor drug 5,6-dimethylxanthenone-4-acetic acid (DMXAA) - agonist of mouse STING. However, this drug turned out to bind to only mouse STING, and thus it is not suitable for human use (Conlon et al., 2013, Paludan et al., 2019).

Other research group have found that in T cells that have naturally high STING levels, the STING activation by its agonists may lead to expression of proapoptotic genes, which is p53- and IRF3-dependent, and triggers apoptosis. This was shown to occur in both wild-type and cancer mouse T cells, and may be used to create an anti-cancer therapy (Gulen et al., 2017). The pro-apoptotic action, but with different mechanism, was also described in healthy and malignant mouse B cells treated with STING agonist (Tang et al., 2016).

The cGAS-STING pathway was also shown to be able to induce extrinsic apoptosis pathway by induction of tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) expression by type I interferon. In this process, TRAIL activates death receptor 5, which triggers apoptosis (Paludan et al., 2019, Zhu et al., 2018).

2.10.3.2. STING in lysosome-dependent cell death

2.10.3.2.1. Lysosome-dependent cell death

The lysosome-dependent cell death pathway is a controlled cell death pathway performed by cathepsins that are released from permeabilized lysosomes. The permeabilization may be caused by detergents, pore forming BCL-2 family members (BAK and BAX), reactive

oxygen species and bacterial toxins. The cathepsins are proteases that, when released to the cytoplasm, cleave the cellular components. Their release can also lead to activation of proapoptotic proteins - BAX and BID that initiate cell death program. The lysosomal cell death is important in many physiological processes, including tissue remodelling and response to infection. Any inbalance in this pathway can lead to deadly diseases, for example, neurodegeneration, autoimmunity, lysosomal storage diseases and cancer (Wang et al., 2018).

2.10.3.2.2. STING in lysosome-dependent cell death

The research on human myeloid cells have shown that cGAS-STING axis can cause lysosomal cell death upon DNA detection. Activated STING moves to lysosomal membranes, and permeabilizes them releasing their components that initiate cell death pathway. It additionally causes activation of the NLR family pyrin domain containing 3 (NLRP3) inflammasome (Gaidt et al., 2017). However, this is not a case in mouse embryonic fibroblasts where STING is degraded upon translocation to lysosome, suggesting that lysosomal cell death pathway, mediated by STING, may be cell-specific (Gonugunta et al., 2017).

2.10.3.3. STING in necroptosis

2.10.3.3.1. Necroptosis

The necroptosis is a form of regulated cell death that occurs when the apoptotic pathway is blocked. It morphologically resembles both necrosis - an uncontrolled cell death, and apoptosis. Practically, necroptosis is a controlled form of necrosis that may be activated by death receptors (TNF superfamily receptors), interferon receptors (e.g. INFR1) and toll-like receptors (TLR3 and TLR7). In the classical pathway, mediated by death receptors, when caspase 8, which normally initiates the extrinsic apoptosis, is absent or inhibited, the cell switches to necroptosis. The necroptotic pathway starts with activation of receptor-interacting serine/threonine-protein kinase 1 and 3 (RIPK1 and RIPK3) that form a dimer. Afterwards, the kinases phosphorylate the mixed lineage kinase domain-like protein (MLKL) that forms oligomers that translocate to plasma membrane, and form a pore. Then cytokines, chemokines and damage-associated molecular patterns are released from the cell through the pore. The pore can also recruit Na^{+} and Ca^{2+} ion channels. These processes cause inflammation and cell destruction. The necroptosis occurs only in some cells that express required components, and in normal conditions, it is inhibited by caspase 8 (Dhuriya et al., 2018).

2.10.3.3.2. STING in necroptosis

It was shown in experiments on bone marrow-derivied macrophages that exposing them to STING agonists can initiate their necroptosis. Such necroptosis requires TNF and type I interferon signaling via cGAS-STING pathway. It was also shown that inflammatory disease

inflicted in mice by injection with STING agonist does not occur in animals with deleted components responsible for necroptosis, confirming the interlink between these two pathways (Brault et al., 2018). Additionally, in experiments on fibrosarcoma L929 cell line, other researchers observed that upon murine gammaherpesvirus-68 infection, a necroptosis occuring in cells is also mediated by TNF and cGAS-STING pathway (Schock et al., 2017).

2.11. cGAS-STING pathway in disease

Every defense system can be destructive for the host in case of mulfunction. cGAS-STING pathway was shown to mediate development of many severe diseases, including type I interferonopathies, neurodegenerative diseases and cancer.

2.11.1. Type I interferonopathies

Type I interferonopathies is a collective term describing a number of diseases with pathologically increased producion of type I interferons that causes chronic inflammation. Among them there are Aicardi-Goutières syndrome (AGS), systemic lupus erythematosus (SLE), type I interferonopathy caused by a mutation in DNase II, familial chilblain lupus (FCL) and STING-associated vasculopathy with onset in infancy (SAVI).

2.11.1.1. Aicardi-Goutières syndrome

Aicardi-Goutières syndrome is an inflammatory disease caused by mutations of the genes encoding multiple proteins, including three-prime repair exonuclease 1 (TREX), melanoma differentiation-associated protein 5 (MDA5), adenosine deaminase acting on RNA (ADAR), sterile alpha motif and histidine aspartate domain-containing protein 1 (SAMHD1), endonuclease complex (RNase H2A, H2B and H2C) and ribonuclease H2 (RNase H2). The connection to cGAS-STING pathway is through TREX. When this protein is absent or mulfunctioning, the DNA accumulates in the cytoplasm, and causes constant cGAS activation (Crow and Manel, 2015, Stetson et al., 2008). The DNA sources for this process are not clear, but they may include: derivatives of DNA repair as well as the DNA pieces, derived from replication pathways and retroelements activity (Motwani et al., 2019a). The resulting high interferon levels, circulating in serum and cerebrospinal fluid, trigger inflammatory symptoms similar to those present during inborn viral infection. In humans, it leads to chronic inflammation of neurons, and in mice to interferon-mediated myocarditis (Crow and Manel, 2015, Stetson et al., 2008). Gray et al. for the first time linked cGAS to inflammatory disease in mice, proving that in absence of cGAS such mice remain healthy. They suggested that cGAS inhibitors could be used to treat Aicardi-Goutières syndrome, and other similar diseases (Gray et al., 2015). Furthermore, it was also shown that mice with deleted genes encoding STING, IRF3 and IFNAR1 where not susceptible to autoimmune disorders triggered by a production of interferon (Gall et al., 2012, Stetson et al., 2008).

2.11.1.2. DNase II-related interferonopathy

DNase II cleaves DNA in endosomes and phagolysosomes during clean up of apoptotic debris by phagocytic cells. The mutations in the gene, encoding this enzyme, where shown in humans to cause symptoms, including high levels of type I interferones in monocytes and lymphocytes, severe neonatal anemia, arthritis, liver fibrosis, membranoproliferative glomerulonephritis (an inflammation of glomeruli that filter blood in kidneys, caused by build-up of antibodies) and increased expression of antinuclear antibodies (anti-DNA antibodies), which cause autoimmune disease (Imai et al., 1997, Rodero et al., 2017).

In contrast, in mice, deletion of DNase II-encoding gene causes an inflammatory disease, which is lethal to embryos. This is because, apoptosis, necessary for proper embryo growth, releases large amounts of DNA that in absence of DNase II accumulates in phagolysosomes, and activates DNA sensors that leads to production of type I interferons, mediated by STING (Ahn et al., 2012). What is interesting, the DNase II-deficient mice with deleted *Ifnar* gene were shown not to die during embryonal development, but instead to develop autoimmune disease with symptoms, including arthritis, elevated levels of autoantibodies and proinflammatory cytokines. Baum et al. have formed the mice with additional mutation in STING- or AIM2- or Unc93b-encoding genes. AIM2 is the cytosolic DNA sensor that forms inflammasome that processes interleukin 1β and 18 (IL-1β and IL-18) into mature forms, and causes an inflammatory cell death through pyroptosis. Unc93b is a chaperone protein that causes migration of toll-like receptors 7 and 9 (TLR7 and TLR9) to endosomes where they function as nucleic acid sensors for RNA and DNA, respectively. In triple mutant mice with *Sting* deletion, only a slight inflammation with no arthiritis occurs. In contrast, *Aim2* deletion significantly inhibits arthritis development in interferon- and STING-independent mode. Furthermore, both triple mutants have shown autoantibody expression that indicates that STING and AIM2 are not engaged in this process. Interestingly, Unc93b deficiency causes lack of autoantibody production, indicating importance of endosomal localization of TLR7 and TLR9 in autoimmune disease development (Baum et al., 2015).

2.11.1.3. Systemic lupus erythematosus

Systemic lupus erythematosus (SLE) is a progressive autoimmune disease in humans. It causes high levels of antinuclear antibodies, targeting nucleosomes and DNA. Its symptoms include: erythematoid rash, arthritis and accumulation of immune complexes in the blood vesels, joints and glomeruli that causes membranoproliferative glomerulonephritis (Napirei et al., 2000, Yasutomo et al., 2001).

The disease with symptoms similar to the ones observed in SLE was shown to be present in mice with DNase I deficiency, which is an enzyme that cleaves DNA in extracellular space (Napirei et al., 2000). To find out if this is also a case in humans, Yasutomo et al. analyzed material from patients with SLE, and found that all of them had a heterozygous mutation in gene encoding DNase I, which makes it unfunctional. Thus, the assumption is that such mutation can be one of the causes of this disease (Yasutomo et al., 2001).

Another enzyme that was linked to SLE in humans is DNase IL3 that clears out DNA in microparticles, released during apoptosis. A rare mutation in its gene was shown to cause SLE in children (Al-Mayouf et al., 2011). Similar disease was also observed in mice with such a mutation (Sisirak et al., 2016).

Type I interferons are also known to be engaged in development of SLE (Muskardin and Niewold, 2018). The increased levels of ISGs products were found in periferal blood mononuclear cells of 50% of SLE patients (Baechler et al., 2003). However, in this case, it is not clear how cGAS can be activated by extracellular DNA (Motwani et al., 2019a).

Thim-Uam et al. have shown in experiments on *Fcgr2b*$^{-/-}$ mice (lupus mouse model) that they have an increased STING-encoding mRNA levels in the spleen (Thim-Uam et al., 2020). Fc gamma receptor IIb (FCGR2b) is a member of family of receptors that enable clean-up of immune complexes through phagocytosis. Fc gamma receptors (FCGRs) are important in SLE pathogenesis, and their polymorphisms were shown to modulate human susceptibility to this disease (Zhu et al., 2016). STING deletion in *Fcgr2b*$^{-/-}$ mice increased overall survival, suggesting the interlink between SLE and STING axis. STING was also shown to be implicated in autoantibody expression and development of glomerulonephritis. Additionally, in absence of STING, multiple interferon regulatory factors and interferon stimulated genes, which are normally upregulated in kidneys of lupus mice, are downregulated. The authors also shown that SLE can develop in *Fcgr2b*$^{-/-}$ mice upon activation of STING pathway in dendritic cells. To conclude, the researchers suggested that STING may be a potential drug target to treat SLE (Thim-Uam et al., 2020).

2.11.1.4. Familial chilblain lupus

Familial chilblain lupus (FCL) is a rare autoimmune disease that starts in early age, and is caused by the mutations in TREX and SAMHD1 enzymes that cleave cytoplasmic DNA and deoxynucleoside triphosphates, respectively (Herrmann et al., 2018, König et al., 2017). Such mutation causes accumulation of DNA inside the cells, and can activate cGAS that leads to initiation of type I interferon pathway. How is cGAS-STING axis associated with this disease is not fully understood, but increased levels of type I interferons and elevated expression of interferon stimulated genes were described in patients with FCL (Berndt et al., 2022, Peschke et al., 2014). The disease was also described in patients with heterozygous mutation that causes constant activation of STING (König et al., 2017).

2.11.1.5. Gain-of-function mutations of STING

The autosomal dominant mutations of STING that make it constantly activated in absence of 2’,3’-cGAMP were shown to be linked to type I interferonopathy, called STING-associated vasculopathy with onset in infancy (SAVI) (Liu et al., 2014). Its main symptoms include systemic inflammation, ulcerative skin lesions, fever, cutaneous vasculopathy and interstitial

lung disease, which eventually causes fibrosis. First symptoms occur in a very early age, and the disease leads to premature death (Decout et al., 2021, Ergun et al., 2019).

The mutations that cause SAVI are located either in STING dimerization interface (V147L, N154S and V155M) or in its polymerization interface (G207E, R281Q, R284S and R284G) (Decout et al., 2021, Liu et al., 2014). Many of these residues, including positions 147, 154 and 155, are highly conserved across many species. The position 147 lies in the connector helix, while the residues 154 and 155 are located in the connector loop, thus there is a possibility that they anyhow cause 180 degrees turn of the STING CTD and trigger oligomerization. The same may be the case with the residues that are located in polymerization interface. However, how exactly these mutations activate STING is still not fully understood (Decout et al., 2021, Ergun et al., 2019).

What is interesting, the symptoms of SAVI were shown to differ between human individuals, and their severity depends on mutation. Similar observations were made in mice with either N153S or V154M mutations (corresponding to human N154S and V155M, respectively). The former develop only a skin vasculopathy with mild lung inflammation, and the latter experience a strong lung inflammation that leads to fibrosis. Thus, V154M mutation causes more severe outcomes than N153S (Motwani et al., 2019a, Motwani et al., 2019b). The surprising difference between humans with N154S and mice with N153S mutation, lies in much lower ISGs expression levels present in the latter (Warner et al., 2017). It was shown that in mice, IRF3, IFNAR and MLKL (the protein implicated in necroptosis) proteins have low or even none influence on development of this inflammatory disease (Motwani et al., 2019a, Motwani et al., 2019b, Warner et al., 2017). In N153S SAVI mice, T cells were found to be the factor promoting aforementioned lung inflammation (Luksch et al., 2019).

2.11.2. Parkinson's disease

Parkinson's disease is a neurodegenerative disorder with not fully explained etiology (Balestrino and Schapira, 2020). It was shown that mutations of the E3 ubiquitin ligase parkin and mitochondrial serine/threonine-protein kinase, called PTEN-induced kinase 1 (PINK1), can trigger this disease. Physiological action of both of these proteins leads to elimination of damaged mitochondria in mitophagic pathway. For instance, a mitochondrial stress may lead to release of mitochondrial DNA, and initiate innate immune response. This is why proper mitophagy is critical in preventing a development of inflammation. The mice with deletion of parkin and PINK1 proteins that were either subject to excessive physical excercise or had additional mutation that causes accumulation of mutations in mitochondrial DNA were shown to develop inflammation, which is not present when the STING gene is deleted. Additionally, a deletion of STING in mice that accumulate mitochondrial DNA mutations was shown to prevent loss of dopaminergic neurons. These observations link innate immunity and Parkinson's disease (Sliter et al., 2018).

2.11.3. Intestinal homeostasis and inflammation control

STING pathway is critical for maintaining intestinal homeostasis. Type I interferons regulate development and function of regulatory T cells that reside in the intestines. The production of these molecules is maintained by microbiota that constantly activate STING. In the experimental conditions it was shown that mouse deprived of STING are more prone to development of gut infections and colitis (large intestine inflammation) (Motwani et al., 2019a).

2.11.4. Non-alcoholic fatty liver disease

The non-alcoholic fatty liver disease is characterized by a pathological accumulation of fatty tissue of non-alcoholic origin in the liver (Benedict et al., 2017). It was found that STING is activated in macrophages that reside in the liver, which leads to expression of proinflammatory cytokines that cause steatosis and fibrosis of this organ (Luo et al., 2018, Motwani et al., 2019a).

2.11.5. Hutchinson-Gilford progeria syndrome

The Hutchinson-Gilford progeria syndrome is a disease in which aging is severely accellerated. It is caused by the expression of truncated variant of lamin A, called progerin. This protein was shown to induce replication stress and genomic instability that activates cGAS-STING pathway and interferon response regulated by STAT1. Administration of calcitrol and other Hutchinson-Gilford progeria syndrome symptoms-improving drugs, decreases replication stress and response mediated by type I interferons. These findings linked type I interferon pathway to this disease (Kreienkamp et al., 2018, Motwani et al., 2019a).

2.11.6. Bloom syndrome

The Bloom syndrome is caused by lack of Bloom syndrome protein (BLM) which responds to DNA damage and together with other proteins maintains genome integrity. When BLM is missing, the cells accumulate micronuclei with damaged DNA that predisposes Bloom syndrome patients to carcinogenesis. Furthermore, the patients' blood was shown to contain increased amounts of interferon stimulated and other proinflammatory genes products. In experiments on BLM-deficient cells, it was shown that cGAS-STING pathway is responsible for such content of the blood (Motwani et al., 2019a).

2.11.7. cGAS and genome instability

Mackenzie et al. linked innate immunity pathway, which includes cGAS, to genome instability. cGAS was shown to detect genomic instability by micronuclei surveillance that

results in ISGs activation. The micronuclei are the structures composed of a nuclear membrane fragment that encapsulates a piece of chromatin. They are formed during mitosis due to genome instability. In both normal mouse cells, exposed to ionising radiation, and human cancer cells, cGAS was found to localize to micronuclei, which results in elevated ISGs expression. It was also shown that in the cell division process, cGAS complexes with DNA damage marker - the histone γH2AX, and together they may be enclosed into the micronuclei (Mackenzie et al., 2017).

Furthermore, Liu et al., have demonstrated a role of cGAS in promotion of genome instability. They reported that cGAS blocks DNA double-strand breaks repair by homologous recombination while residing in the nuclear compartment. Its interaction with poly(ADP-ribose) polymerase-1 (PARP1) prevents its complexing with TIMELESS protein, critical for the repair. Also, cGAS was shown to form oligomers with the DNA, interfering with repair machinery, mediated by RAD51 protein. This occurence can promote accumulation of damaged DNA, micronuclei formation, cell death upon genomic stress as well as tumor growth (Jiang et al., 2019, Liu et al., 2018).

2.11.8. cGAS-STING pathway in cancer

The process of tumorigenesis progresses upon accumulation of chromosomal instability (CIN), damaged DNA and micronuclei. Such DNA sources were shown to activate cGAS-STING pathway that results in expression of interferon-stimulated genes and inflammatory cytokines. Generally, such outcome would suggest that this pathway has beneficial anti-tumor effects. However, multiple contrary reports were presented, in which cGAS-STING, on one hand, can prevent tumor growth, and on the other, promote tumor metastasis (Motwani et al., 2019a).

In a course of chemo and radiation therapy of cancer, cGAS-STING signaling was shown to be activated and to positively affect the therapy outcomes. The inhibition of DNA repair using PARP1 inhibitors that triggers micronuclei formation and activates cGAS in tumor environment was proved to initiate innate immune response that prevents tumor growth (Chabanon et al., 2019, Ding et al., 2018, Shen et al., 2019).

In contrary to this, Liu et al. characterized cGAS as a molecule that promotes tumorigenesis by inhibiting PARP1 and facilitating accumulation of DNA damage. Deletion of cGAS-encoding gene stops DNA damage and blocks tumor developement (Liu et al., 2018). The DNA from cancer cells was shown to promote induction of type I interferon pathway in dendritic cells that potentiates cancer antigen presentation and $CD8^+$ T cells activation which brings anti-tumor response (Woo et al., 2014).

What complicates this matter even further, is the report stating that cytokine expression, induced upon micronuclei detection by cGAS signaling pathway, can cause accumulation of CIN. Such accumulation was linked to cancer metastasis in a comparative study in which the

levels of CIN in metastatic tumors were shown to be much higher than in the primary brain tumors from which they originated. From this research it was deduced that metastasis is promoted by high-CIN levels. The observations were later confirmed in mouse models by exposing mice to high- and low-CIN tumor transplants. The former had much lower survival rates and developed more metastatic tumors than the latter ones. The cancer cells with high CIN levels were also shown to have more micronuclei, containing cGAS. In such cells, STING induces signaling through non-canonical NF-κB pathway (Bakhoum et al., 2018, Motwani et al., 2019a). Furthermore, metastasis promotion was shown to be stimulated by activation of both type I interferon and NF-κB pathways in astrocytes to which cGAMP was transferred through gap junctions from brain cancer cells. This activation is mediated by cGAS and STING. What is important from therapeutical point of view, inhibitors of gap junctions were shown to be able to inhibit brain metastasis in mice (Chen et al., 2016).

2.12. cGAS-STING pathway in medicine

cGAS-STING pathway influences many diseases, and thus constitutes a hot medical target, which is pursued by various pharmaceutical companies. STING is expressed in a wide variety of cell types and could be a universal therapeutical target for many diseases.

2.12.1. Cancer therapy

Many studies have shown that STING agonists are able to promote anti-tumor innate immune response that also enables tumor antigen presentation by dendritic cells and resulting proliferation of T cells. Furthermore, STING induces destruction of tumor vasculature and tumor cell death. The latter process, additionally, releases tumor antigens that can be targeted by immune cells (Motwani et al., 2019a).

Pharmacological activation of STING in mouse models has been proven to have significant therapeutic effects in fighting cancer. A well-known example is 5,6-dimethylxanthenone-4-acetic acid (DMXAA) - the vascular disrupting agent and mouse STING agonist. This experimental drug have shown promising effects in mice, but it failed in human clinical trials. It was most probably caused by the fact that this small molecule binds only to mouse, but not to human STING. However, the precise way of action of this drug is still not fully understood (Conlon et al., 2013, Daei Farshchi Adli et al., 2018). Other potential drugs that shown positive effects in mice include synthetic cyclic dinucleotide dervatives (Corrales et al., 2015) and c-di-GMP. They were found to improve effects of anti-cancer vaccination (Chandra et al., 2014).

The first generation of STING agonists in preclinical studies was built for direct injection into a tumor. The major disadvantage of this form of therapy is that the tumor must be easily accessible for injection. An intravenous injection of such compounds is risky, because of possibility of serious adverse effect - cytokines release syndrom (Motwani et al., 2019a). Two synthetic cyclic dinucleotides, MK-1454 and ADU-S100, agonists of STING, made

their way into human clinical trials. However, intratumoral injections of these drugs alone did not show promising effects. Only in combination with intravenous injection of other drug, pembrolizumab, they boosted its own destructive effects on head and neck cancer.

Recently, the new generation of small-molecule drugs was discovered, called amidobenzimidazole-based compounds (ABZIs). These STING agonists are not cyclic dinucleotides, and can be applied as intravenous injection, thus, they can be used also against tumors that are not easily accessible. In mouse studies, administration of ABZIs inhibited tumor growth, and caused anti-tumor immune response mediated by $CD8^+$ cytotoxic T cells (Ramanjulu et al., 2018).

The STING agonists could be also used in combination with, for example, an immunotherapy. Such compounds can increase immunogenicity of various tumors making them more susceptible to immune response, and thus, possibly, reinforce the results of such therapy. However, STING was also found to trigger production of immunosupressive proteins: C-C motif chemokine receptor 2 (CCR2) and programmed death-ligand 1 (PD-L1). These macromolecules may have positive effects on tumor growth (Motwani et al., 2019a). However, this dark side of STING can be reversed by adding inhibitors of CCR2 to the therapeutical mix (Liang et al., 2017).

Lastly, it is worth noted that STING was also found to promote metastasis of high-CIN tumors. In this case, STING antagonists in combination with other selected therapy would be needed (Motwani et al., 2019a).

2.12.2. Infectious diseases

2.12.2.1. Bacterial infections

The high priority for current healthcare is development of new better adjuvants that could improve vaccine efficiency. STING agonists as potent immunogens were shown to be superior over conventional adjuvants in experimental vaccines against multiple diseases in mouse models. These vaccines were tested against: *Mycobacterium tuberculosis*, influenza, *Streptococcus pneumoniae*, anthrax and multiple others. They are able to induce strong Th1, Th2, Th17, and CD8+ cytotoxic cell responses. The major challenge with these small molecules is their negative charge that makes them unable to easily cross the cell membrane. To tackle this issue, multiple currently known technologies can be used. They include liposomes, virus-like particles, emulsions and polymers (Gogoi et al., 2020).

2.12.2.2. Hepatitis B virus

In the simulated mouse liver environment, created by coupling hepatocyte and macrophage culture, administration of DMXAA was shown to trigger STING-mediated innate immune response against hepatitis B virus (HBV), and decrease viral replication. Therefore, STING

agonists could be potentialy used to enhance anti-viral therapy against this virus (Guo et al., 2015).

2.12.2.3. Human immunodeficiency virus

The cells that are latently-infected with HIV constitute the major problem with development of an effective therapy aginst this virus. This is because, despite lowering viral loads in serum by antiretroviral treatment these cells retain virus inside the body. The STING agonists may be good candidates to solve this problem. In experiments on cynomolgus monkeys infected with simian immunodeficiency virus (SIV), which constitute an animal model of latent HIV infection, it was found that levels of proinflammatory antiviral cytokines were increased in virus-carrying peripheral blood mononuclear cells (PBMCs) treated with STING ligands (c-di-AMP and 2',3'-cGAMP) and TLR7/8 ligand (R848). These substances were able to reactivate the latently-infected cells, and lead to immune response against them. Furthermore, c-di-AMP administered together with the combination antiretroviral therapy boosted response of cytotoxic T cells against them (Yamamoto et al., 2019).

2.12.2.4. Herpes simplex virus type 2

Local or intraperitoneal administration of 2',3'-cGAMP and other cyclic dinucleotides that activate STING was also found to have promising effects against genital herpes simplex virus type 2 (HSV-2) by triggering strong type I interferon response against this pathogen (Skouboe et al., 2018).

2.12.2.5. Severe acute respiratory syndrom coronavirus 2

Recently, in advent of coronavirus disease 2019 (COVID-19), which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-COVID-2), scientists explored various already known compounds to find a drug that could stop the pandemic. One of such potential medicines is STING agonist, diamidobenzimidazole 4 (diABZI-4). It was shown to efficiently inhibit replication of SARS-COVID-2. The results were confirmed in lung epithelial cell line and in transgenic mouse model. Intranasal application of diABZI-4 pre- or post-infection protected the animals from severe lung disease. diABZI-4 caused temporary, local STING activation that triggered proinflammatory cytokines expression and activation of lymphocytes that negatively affected viral replication (Humphries et al., 2021).

2.12.2.6. Baterial strategies against cGAS-STING pathway

In case of bacterial infections type I interferon signaling mediated by STING was shown to have either anti- or pro-microbial effects depending on type of bacteria. Activation of this pathway may increase susceptibility to some bacterial infections and promote their growth and survival. Some microorganisms are even able to compromise its function in order to remain undetected. Therapeutic approaches targeting cGAS-STING pathway should be then

customized to the specific type of bacterial infection (Liu et al., 2022, Nandakumar et al., 2019).

2.12.2.6.1. cGAS-STING pathway as defense means against immune cells

The type I interferon pathway, mediated by STING, was shown to promote anti-inflammatory state of macrophages that prevents clearance of infection with *Staphylococcus aureus*. Also, it facilitates the growth of these bacteria (Liu et al., 2022). Very interesting examples of other bacterial species that use cGAS-STING pathway to their advantage are: *Listeria monocytogenes*, *Legionella pneumophila* and *Francisella tularensis*. Their DNA was shown to be delivered to bystander cells by extracellular vesicles where it activates cGAS. By such means, the bacteria can cause apoptosis of T cells and inhibit their proliferation (Nandakumar et al., 2019).

2.12.2.6.2. Evasion strategies against cGAS-STING pathway

Mycobacterium tuberculosis uses multiple strategies to evade cGAS-STING pathway. This bacterium releases phosphodiestrase that degrades cyclic dinucleotides, including its own c-di-AMP, which prevents activation of STING. Except for that, the bacterial protein Rv0753c is able to bind to and disable STING function. It also primes it for autophagic degradation (Liu et al., 2022). Inhibitors of bacterial phosphodiestrase could be used as potential therapeutic agents (Karanja et al., 2021). Furthermore, the subunit vaccine with cyclic dinucleotide-based adjuvant was shown to have protecting effects against this bacterium in the mouse model (Van Dis et al., 2018).

2.12.3. Inflammatory diseases

Inhibitors of cGAS and STING could potentially be used against a wide range of inflammatory diseases. Currently, multiple compounds that inhibit cGAS function by either binding to its active site or to its DNA binding sites, were identified. Also, a number of STING-inhibiting small molecules was found. They can be applied alone or in combination with inhibitors of cGAS for similar purposes. They were shown to bind to the cyclic dinucleotides binding site, which prevents ligand binding, or to the cystein residues number 88 and 91, which prevents STING palmitoylation. However, there is still a long way from selection of appropriate candidates to development of effective therapies (Decout et al., 2021).

3. Introduction to cluster of differentiation 47

3.1. Cluster of differentiation 47

Cluster of differentiation 47 (CD47) was discovered and named multiple times. Only later on the researchers realized that they have all found the same protein. It was characterized as an integral part of Rh complex present on the surface of erythrocytes, and called CD47 (Mawby et al., 1994). It was also found to complex with integrin αVβ3, and named integrin associated protein (IAP) (Lindberg et al., 1993). Additionally, it was detected in elevated levels in ovarian cancer cells, and described as ovarian tumor marker OA3 (Campbell et al., 1992).

CD47 is a heavily glycosylated integral membrane protein with theoretical mass of approximately 47 kDa. However, because of glycosylation pattern, it appears on SDS-PAGE gels as a smeary band at approximately 60 kDa (Hayat et al., 2020).

The evolutionary conservation of CD47 between related species, including *Mus musculus*, *Rattus rattus* and *Bos taurus* is quite high - 60-70% (Oldenborg, 2013). It interacts with multiple partners and performs a wide variety of functions, but the mechanisms of many of them still remain a mystery. They are described in detail in the further sections.

3.1.1. Structure of CD47

Human CD47 is composed of the N-terminal extracellular domain (ECD) (aa 1-117), followed by the transmembrane domain (TMD) and short C-terminal tail (CTT) (Fenalti et al., 2021). The CTT length may be from 4 to 36 residues, and due to that the protein can be produced in form of four isoforms. Different isoforms are expressed in various tissues, and the most common is isoform 2 with 16 amino acids, which is present in endothelial, epithelial and blood cells. The second most common is isoform 4 with 34 amino acids, which is present mostly in neural tissue, testis and intestines (Reinhold et al., 1995). The functions of these different CTTs are not clear (Oldenborg, 2013). The ECD has an immunoglobulin variable (IgV)-like architecture with five glycosylation sites: N5, N16, N32, N55 and N93. It is composed of mostly parallel, six-stranded beta-sheet. It is connected to the TMD with the conserved ^{114}RVVSWF^{119} linker, called the SWF loop that resembles an antibody hinge region. It is located in the extracellular loop region (ECLR). ECLR also contains extracellular loops 1 and 2 (ECL1 - aa 179-185 and ECL2 - aa 243-250) (Fenalti et al., 2021). The connection is additionally stabilized by the conserved disufide bond between cystein15 of the ECD and cystein245 of the ECL2. This unique, long-range disulfide bond, along with other weaker interactions between ECLR loops, are critical for keeping ECD in proper spacial orientation, necessary for its biological function. The interdomain disulfide bond was, for example, found to be essential for the interaction with signal regulatory protein α (SIRPα). However, its absence does not affect binding to integrins (Fenalti et al., 2021, Rebres et al., 2001). The TMD forms a pentagonal helical bundle composed of five transmembrane helices

(helix I - aa 121-140, helix II - aa 160-178, helix III - aa 186-207, helix IV - aa 224-242, helix V - aa 251-278). The helix II is directed at almost 90 degrees relative to the surface of the cell membrane. The rest of the helices is gently tilted at various smaller angles. The helix III is in the center of the bundle and is surrounded by other helices. The CD47 TMD was shown to have a unique fold with only little similarities comparing to other currently known proteins (Fenalti et al., 2021). The annotated structure of CD47 is shown in Figure 5.

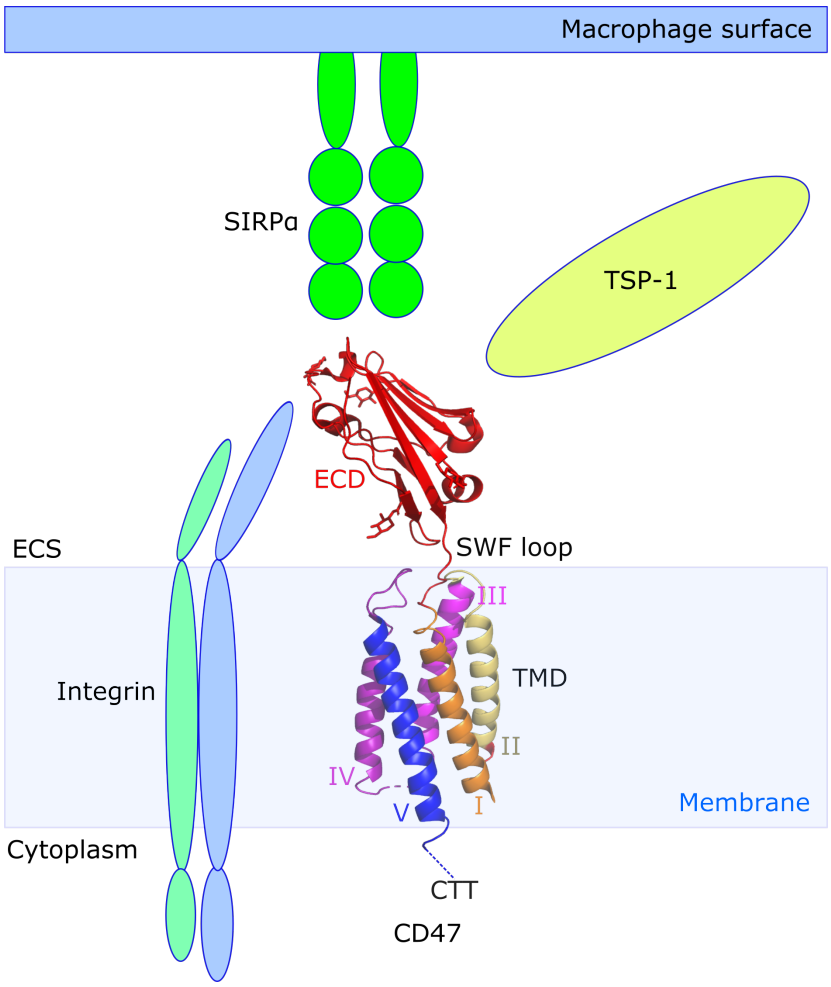

Figure 5. Structure of CD47 and its main interacting partners.
Human CD47 is shown as a cartoon representation of molecule deposited in PDB with accession code 7MYZ (Fenalti et al., 2021). It is composed of extracellular domain (ECD), exposed to extracellular space (ECS), transmembrane domain (TMD), embedded in the cell membrane, and a short C-terminal tail (CTT), which is represented as blue, dashed line. ECD and TMD are connected via SWF loop. TMD is composed of five transmembrane helices, marked with Roman numbers. CD47 interacts with multiple integrins *in cis*, within the same cell membrane. It also has two major interaction partners in the ECS - SIRPα, the membrane protein on the surface of macrophages, and thrombospondin 1 (TSP-1), the large resident of extracellular matrix. These are just schematic representations of partner proteins, their sizes are not to scale.

3.1.2. Functions of CD47

3.1.2.1. Main functions

CD47 is expressed on the surfaces of most of the cells building human body, and one of its best characterized functions is acting as a "don't eat me" signal for macrophages which makes it a "universal marker of self" (Eladl et al., 2020). Downregulation of CD47 expression leads to removal of unnecessary or old cells. It also plays important roles in various physiological and pathological processes, including regulation of immune system, carcinogenesis, atherogenesis and stress resistance (Hayat et al., 2019). The CD47 expression changes in various diseases, for instance, tumors use its overexpression to defend against phagocytosis (Jaiswal et al., 2009). Apart from that, its expression decreases in brains of individuals with multiple sclerosis (Koning et al., 2007).

CD47 to perform its many functions, interacts with a number of partner proteins, such as signal regulatory protein alpha (SIRPα), thrombospondin 1 (TSP-1) and integrins. The cell type dictates a result of CD47-partner protein interaction. The CD47-mediated signaling cascades are important in processes, like hydrogen sulfide and nitric oxide signaling, synthesis and homeostasis of Ca^{2+} ions, activation of trascription factors, CDNs signaling and other (Hayat et al., 2020).

3.1.2.2. Other functions

CD47 has an important role in connecting the cytoskeleton with the cell membrane by interaction with proteins linking IAP and cytoskeleton 1 and 2 (PLIC-1 and PLIC-2) (Wu et al., 1999). It was shown to be necessary for migration of epithelial cells of the intestine. Additionally, it is required for cyclooxygenase-2 (COX-2) production, initiated by collagen I, which is important in inflammatory processes (Broom et al., 2009). CD47 was also shown to be able to initialize apoptosis in various cells using different mechanisms. For example, mAbs against CD47 and TSP-1 are able to induce cell death similar to apoptosis that does not depend on caspases in leukemic B-cells (Mateo et al., 1999). Multiple monoclonal antibodies that bind to CD47 as well as TSP-1 and 4N1K peptide were also used to cause apoptosis in Jurkat T-cells (Leclair et al., 2018, Manna & Frazier, 2003). Manna and Frazier, have shown that CD47-dependent apoptosis requires heterotrimeric G protein (heterotrimeric Gi). The cell death is caused by significant decrease of cyclic adenosine monophosphate (cAMP) levels inside the cells, because of influence of hetrotrimeric Gi on protein kinase A (PKA) (Manna & Frazier, 2003). These authors also found a similar apoptotic pathway mediated by CD47 in multiple breast cancer cell lines (Manna & Frazier, 2004). CD47 is also a component of Rh complex in erythrocyte membrane which is connected to Band 3 complex that attaches cytoskeleton built of spectrin to the membrane. Its function in this complex is not clear (Oldenborg, 2013). CD47 was shown to be engaged in red blood cells clearance. In older erythrocytes a conformational change of CD47 of unknown origin occurs that makes it able to bind TSP-1. Such change is usually not possible in young erythrocytes. It converts

CD47 to an "eat me" signal for macrophages, and this is how such no longer necessary cells get destroyed (Burger et al., 2012).

3.2. Interactions with partner proteins

3.2.1. Interaction with SIRP proteins

Signal regulatory protein α (SIRPα), previously described as P84, BIT, MFR, CD172A and SHPS-1, is a heavily glycosylated membrane protein present on the surface of myeloid cells, endothelial cells, neurons and fibroblasts (Hayat et al., 2020, Oldenborg, 2013). It is worth noted that glycosylation patterns on CD47 and SIRPα do not have an influence on their mutual interaction and signal transduction (Lee et al. 2007, Subramanian et al., 2007). Their interaction was also shown to be highly species-specific (Oldenborg, 2013).

The carboxy-terminal domain of SIRPα, which resides in the cytoplasm, consists of two immunoreceptor tyrosine-based inhibitory motifs (ITIMs) that host four tyrosines that can be phosphorylated. Its N-terminal domain, exposed to the extracellular environment, is built of three immunoglobulin-like domains, and hosts CD47-binding site. Upon interaction with CD47, tyrosines in ITIMs get phosphorylated, and they activate phosphatases, called Src homology region 2-domain-containing phosphatase 1 and 2 (SHP-1 and SHP-2), or other phosphatases depending on cell. In macrophages, SHP-1 and SHP-2 dephosphorylate myosin 2a that impairs phagocytosis (Hayat et al., 2020).

Except for impacting phagocytosis, the CD47-SIRPα axis is crucial for many other important processes, including development of neurons in hippocampus region of the brain (Murata et al., 2006), and inhibition of dendritic cells and T cells function, which makes it an anti-inflammatory factor (Latour et al., 2001).

CD47 was also shown to bind to other protein from the SIRP family, called SIRPγ or SIRPβ2, present on the surface of T cells. This interaction improves proliferation of these cells in response to antigen, provided by antigen-presenting cells (Piccio et al., 2005).

3.2.2. Interaction with thrombospondin 1

Thrombospondin 1 (TSP-1) is a 480 kDa member of a family containing five large heavily glycosylated proteins that are secreted to extracellular space (ECS). With its multiple domains it interacts with a wide range of proteins and thus takes part in a plethora of signaling pathways (Kale et al., 2021). The main functions of this protein are: aggregation of thrombocytes, regulation of cell adhesion as well as cell migration, differentiation and proliferation (Chung et al., 1999, Oldenborg, 2013).

Extracellular domain of CD47 was shown to interact with C-terminal domain of TSP-1. A proper glycosylation pattern is critical for this binding to occur (Isenberg et al., 2009, Kaur et

al., 2011). There is only limited data about function of CD47-TSP-1 axis, and unfortunately, most of the research was done using TSP-1-derived, artificially modified peptide that binds to CD47, called 4N1K (KRFYVVMWKK) (Gao et al., 1996, Soto-Pantoja et al., 2015). The natural part of this peptide was shown to be located in the part of C-terminal domain of TSP-1 that is not accessible for CD47 (Kvansakul et al., 2004). A drastic conformational change would be required to make it available for interaction, but there is currently no evidence that such rearrangement could occur. Due to that, there is a lot of scepticism about a biological relevance of the results generated using this peptide (Soto-Pantoja et al., 2015).

The early-known function of CD47-TSP-1 signaling pathway is stimulation of thrombocytes to aggregate during the blood clot formation process (Chung et al., 1997). Thrombospondin 1 interaction with CD47 was also found to cause damage of red blood cells by triggering Ca^{2+} ions influx (Bissinger et al., 2020). Another result of their combined action is contraction of blood vessels, because of inhibition of NO signaling in smooth muscle cells (Bauer et al., 2010). CD47-TSP-1 interaction also inhibits phosphorylation of vascular endothelial growth factor receptor 2 (VEGFR2), which stops angiogenesis (Kaur et al., 2014). Other researchers found that this proteins couple can damage blood vessels by inducing production of reactive oxygen species that leads to oxidative stress. These findings suggest its potential role in development of diseases affecting blood vessels (Csányi et al., 2012). Signaling by these two proteins also promotes cellular senescence and ageing by inhibiton of cell cycle and damage of blood vessels (Gao et al., 2016, Ghimire et al., 2020).

3.2.3. Interactions with integrins

Integrins are transmembrane proteins that reside in cellular membrane in form of heterodimers and mediate cell adhesion, proliferation and migration (Hynes et al., 2002, Lindberg et al., 1996). CD47 was initially purified together with integrins, and due to that it was named the integrin associated protein (IAP). It was shown to associate with multiple integrins *in cis*, including $\alpha_V\beta_3$ integrin. This integrin is present on the surfaces of many cells, for instance, neutrophils. It was shown to require CD47 in order to be able to attach to its ligand vitronectin. First experiments with function blocking antibodies against CD47 have shown that they are able to inhibit activation of monocytes and neutrophils, and also to stop tissue migration of the latter ones. At this point the theory shown up that CD47 influences these cells by its association with $\alpha_V\beta_3$ integrin (Lindberg et al., 1996). Besides that, other teams have demonstrated that CD47 exerts an influence on various processes occurring in neutrophils, such as their response to infection and inflammation as well as their migration in endothelial and epithelial tissues. These findings make CD47 an important regulator of host defense, however, up to now there is no clear confirmation that these functions are related to its interaction with integrins (Liu et al., 2001, Parkos et al., 1996). Chung et al. have found that interaction of TSP-1 with CD47 can cause activation of αIIbβ3 integrin that causes thrombocytes to spread on plates coated with fibrinogen, and to aggregate in suspension. This way of activation requires other partner proteins, including G_i-like heterotrimeric G protein, spleen tyrosine kinase (SYK), focal adhesion kinase (FAK) and proto-oncogene tyrosine-

protein kinase Src (cSrc) (Chung et al., 1997). Additionally, CD47 in interaction with 4N1K peptide and α2β1 integrin was shown to initiate chemotaxis of smooth muscles, building blood vessels, to soluble form of collagen (Wang et al., 1998). The same complex was also proved to cause thrombocytes activation (Chung et al., 1999). Besides that, CD47 is important in B-cells migration where it interacts with integrin α4β1 (Yoshida et al., 2000). Except for integrins mentioned above, it was also observed to bind to other integrins, such as α6β1 and α5 integrin (Koenigsknecht et al., 2004, Orazizadeh et al., 2008).

3.3. CD47 in medicine

3.3.1. CD47 in cancer therapy

The interest to use CD47 as a target for potential cancer therapy shown up early after its discovery as an ovarian cancer antigen, and after finding the correlation between CD47 overexpression and progression of several types of tumors (Soto-Pantoja et al., 2015). The first partially successful attempts to develop anti-CD47 bi-specific antibodies for immunotherapy against cancer were made in 1994 (van Ravenswaay Claasen et al., 1994). However, later, its popularity decreased because of concerns related to discovery that this is not solely cancer antigen, but also the protein present on the surfaces of almost all human cells (Soto-Pantoja et al., 2015). CD47 entered the race again when it was found to help tumor evade detection by immune system, because of its ability to prevent phagocytosis. The researchers observed that antibodies bind preferentially to CD47 on the surfaces of acute myeloid leukemia cells which is followed by their destruction by macrophages (Majeti et al., 2009). The potential explanation of such preference lies in differences in mobility, distribution and densities of protein on the surfaces of healthy, and cancer cells (Jiang et al., 2021). Currently, CD47-SIRPα axis is considered to be an important immune checkpoint for a wide range of cancer cells. A wide variety of antibodies and other molecules that disrupt this interaction was developed, including monoclonal antibodies and bi-specific antibodies. Some of them were designed to also be able to activate Fc gamma receptor (FCGR) on the surface of macrophages to improve their phagocytic activity (Jiang et al., 2021). However, major disadvantage of such therapies, is the fact that they may cause severe adverse effects, including anemia and thrombocytopenia. They are caused by hemolytic reactions triggered by hemagglutination. Due to that, initially there was a lot of scepticism around this way of cancer treatment. But eventually, the promising drug with managable, decreased side effects was developed, and the hopes in CD47, as a useful cancer target, were brought back in 2019. The monoclonal antibody, Hu5F9-G4 administered alone or in combination with a small molecule - azacytidine, was shown to be effective against acute myeloid leukemia and myelodysplastic syndrome (Sallman et al., 2019). The question is why some antibodies that bind to the same small antigen may cause more or less severe side efects? The probable answer lays in differences in type of membrane proteins partnering with CD47 in different cells that somehow changes the nature of its function. Another potential explanation would be that this is due to variablities in CD47 glycosylation pattern present in different cells (Jiang et al., 2021).

Antibodies blocking CD47-SIRPα interaction can also be used in conjugation with other therapies such as classical chemotherapy, antibodies against other cancer targets and other immune checkpoint inhibitors (Chen et al., 2021, Upton et al., 2021). At the moment, a plethora of bi-specific antibodies against CD47 and various targets was and is being developed. For instance, an anti-CD47 antibody can be conjugated with anti-tumor antigen antibody to make it more specific to tumor cells that improves targeting (Dheilly et al., 2018). Recently, the apoptosis-inducing CC2C6 antibody was proposed to be used in conjugation with classical chemotherapy to treat leukemia (Leclair et al., 2018). In 2021 there were already 46 clinical trials on various anti-CD47 constructs in the USA (Jiang et al., 2021). More potential drugs are still in development that brings hope for thousands of cancer patients around the world.

4. Introduction to membrane proteins and tools for structural research

4.1. Membrane proteins

The biological membranes are dynamic structures composed of lipids and membrane proteins (MPs). While membranes divide the cell into compartments, build cellular organelles and separate the cell interior from external environment, membrane proteins are molecular machines that can be enzymes, performing biochemical reactions, receptors, transducing signals, transport proteins, enabling exchange of various substances, cell attachment means, and others (Aguayo-Ortiz et al., 2021, Alberts et al., 2002). The cell surface is additionally covered with various oligosaccharides attached to lipids and membrane proteins which are mainly resposible for cell recognition and adhesion (Eliot, 2022). The membrane proteins attach themselves to the membrane in a number of ways, and on this basis they are classified into several groups (Figure 6.). However, there are multiple discrepancies in the literature on their classification, and several of them are described here.

The first group of membrane proteins is composed of amphipathic transmembrane proteins that fully cross the membrane with their hydrophobic parts, and often have soluble domains sticking outside the membrane. They may be composed of a single or multiple transmembrane alpha helices or a β barrel, which is built of a beta sheet organized into a cylindrical structure. They are also called the integral membrane proteins. The proteins that do not span the entire membrane are called peripheral membrane proteins. Among them there are proteins that have an amphipathic alpha helix, attached only to one layer of the membrane or are anchored in internal or external side of the membrane by interactions that can be mediated by a lipid chain or an oligosaccharide, covalently attached to the protein, respectively. The lipid chain spontaneously inserts itself into the membrane, and the oligosaccharide can be attached to one of the membrane lipids. The members of this subgroup of periferal membrane proteins are called monotopic membrane proteins. And the last subgroup contains membrane-associated proteins that are just non-covalently attached to the other membrane proteins (Alberts et al., 2002, Boes et al., 2021).

Some sources also divide membrane proteins into integral membrane proteins - those permanently attached to the membrane that can be removed only with the use of detergents, and periferal membrane proteins - those weakly attached to the membrane that can be removed just by changing pH or ionic strength (Alberts et al., 2002). This classification is extended by some authors where integral membrane proteins are additionally divided into: polytopic - spanning the entire membrane multiple times, bitopic - spanning the entire membrane one time, and monotopic - spanning only one side of the membrane (Allen et al., 2018).

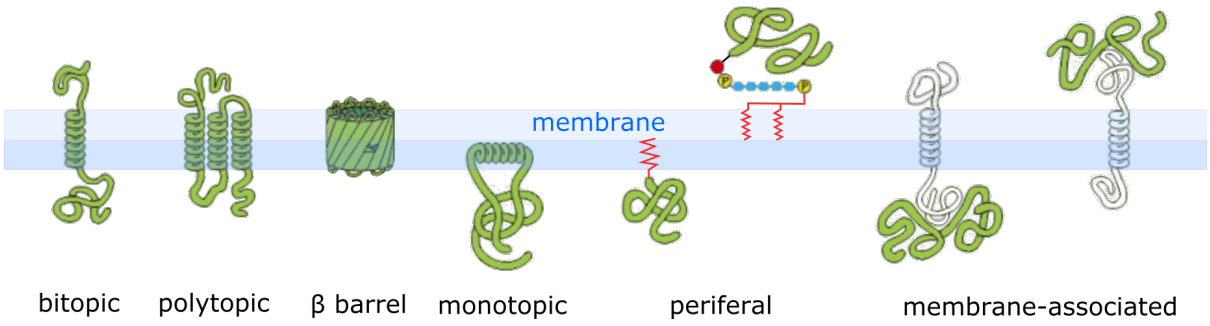

Figure 6. Classification of membrane proteins.
There are some discrepancies in literature on how to group membrane proteins, but for the purpose of this thesis, MPs were divided into: bitopic - spanning the entire membrane one time, polytopic - spanning the entire membrane multiple times, β barrel - cylindrical β sheet spanning the entire membrane, monotopic - spanning only one layer of the membrane, periferal - covalently attached to the membrane with a non-protein element, and membrane-associated proteins - attached to the membrane by the non-covalent interractions with other membrane proteins. All the above, except for periferal and membrane-associated MPs, are considered integral MPs. This figure uses elements derived and modified from Alberts et al., 2002.

4.2. Membrane-mimicking systems

The membrane proteins constitute approximately 30% of all proteins encoded by the human genome, and more than 60% of targets of current drugs. This is why their structural and functional characterization is extremely important for current medicine, and general understanding of the nature of living organisms (Aguayo-Ortiz et al., 2021). They are very challenging molecular targets due to their amphipatic nature, high flexibility, usual small size and low expression levels. This is why, the membrane proteins constitute only a small fraction of all structures deposited in Protein Data Bank (PDB). For example, in year 2017 there were only approximately 650 known structures of membrane proteins, and, in contrast, approximately 100,000 known structures of soluble proteins (Zabara et al., 2017).

4.2.1. Detergents for purification and characterization of membrane proteins

A membrane protein to be biochemically and structurally characterized needs to be extracted from the membrane, and incorporated into a proper membrane-mimicking system to make it stable and soluble in aquous buffering solution. One of the simplest ways to do that is to use a mild, non-ionic detergent that at appropriate concentration can displace most of the lipids, and surround the hydrophobic portion of the protein with a protective interior of the micelle. Every detergent is composed of the hydrophilic head and hydrophobic tail of various lengths and shapes. The micelle is usually a spherical or oblate structure composed of detergent molecules. The detergent tails are directed to its interior, and heads are outside, making it able to interact with water molecules. The micelles form spontaneously in aquous solution when the detergent concentration is at or above its critical micelle concentration (CMC). The

most commonly used and successful detergent for membrane protein purification and structure determination is n-dodecyl-β-D-maltopyranoside (DDM) with CMC of 0.17 mM or 0.0087%. Since many membrane proteins require cholesterol for their function, the detergent is often blended with cholesteryl hemisuccinate tris salt (CHS) which is a water-soluble derivative of cholesterol (Hardy et al., 2018, Stetsenko et al., 2017, https://www.anatrace .com/). The first step of usual purification, after cell lysis, is membrane solubilization. For this purpose, usually a very high concentration of detergent is used. Such concentration is determined experimentally, and may be even at 100 x CMC. For further purification steps much lower concentration is used, however it should never go below the CMC. A simple scheme of membrane protein purification using a detergent is shown in Figure 7.

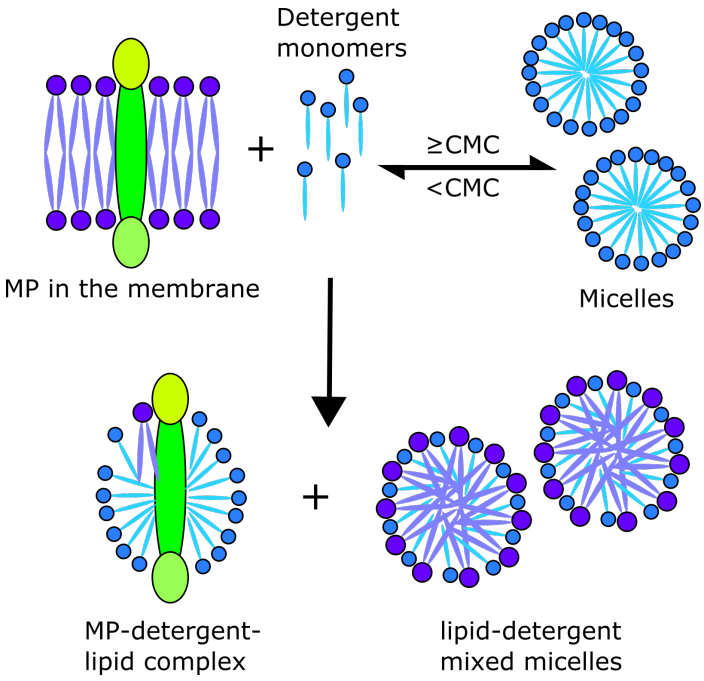

Figure 7. Purification of membrane proteins using a detergent.
The detergent molecules are composed of a hydrophilic head (dark blue) that makes them soluble in aquous solutions, and hydrophobic tail (light blue) that makes them able to form three-dimensional (3D) structures with all hydrophobic tails directed to their center, called the micelles. The micelles can be formed only at or above the certain, detergent-specific concentration, called CMC. At an appropriate concentration, usually much higher than CMC the detergents are able to solubilize biological membranes composed of lipids (dark and light violet). This is how one can purify MPs and keep them stable, surrounded by a protective layer of detergent molecules, in aquous solutions. Some natural lipids may also remain attached to the membrane protein during this process. The rest of the lipids from the membrane, solubilized during purification, forms lipid-detergent mixed micelles, which are 3D structures with all hydrophobic tails of detergent and lipid molecules directed to their center. These are just schematic representations, not to scale. All 3D structures are represented as 2D slices.

4.2.2. Other systems for characterization of membrane proteins

Detergents are convenient means for purification and characterization of membrane proteins, however, it sometimes turns out that these substances do not fully satisfy needs of these fragile macromolecules, and can destabilize them (Hardy et al., 2018). Furthermore, lipids often have regulatory and structural effects on membrane-embedded proteins which can be lost after striping them with detergents. This can cause, for example, that the protein would be researched in non-native conformation that can limit usefulness of obtained data (Gao et al., 2016). To overcome this, either proper detergent screening should be applied, or other system, suitable for selected membrane protein, should be chosen (Hardy et al., 2018). There is a range of such systems, including lipid nanodiscs (NDs) (Ritchie et al., 2009), amphipols (Tribet et al., 1996), Salipro (Frauenfeld et al., 2016) and styrene maleic-acid lipid particles (SMALPs) (Postis et al., 2015). The nanodiscs were used as a method of choice for this thesis, and thus only they will be described here in detail.

4.2.2.1. Nanodiscs

The nanodiscs are structures composed of a patch of lipid bilayer surrounded by two belts of amphipathic membrane scaffold protein (MSP) that interacts with lipids using its hydrophobic surfaces, and is soluble in water thanks to its hydrophilic surfaces. MSP is an engineered derivative of human apolipoprotein A-I that forms discoidal particles in the blood that transport cholesterol. All the domains, not responsible for nanodisc formation, were artificially removed from its sequence to form MSP. The great advantage of nanodiscs is that they can be customized to membrane proteins of any size and properties (Hagn et al., 2013).

Firstly, there is a wide variety of commercially available lipids that can be used to create nanodiscs. There are purified natural and artificial lipids that can be blended together in different ratios, and natural lipid extracts, from various sources, that can mimic any membrane environement (https://avantilipids.com/). Secondly, the size of nanodiscs can be adjusted by choosing the MSP of appropriate length. MSPs were designed to form nanodiscs from 6-17 nm in diameter (Inagaki et al., 2017).

In order to incorporate a membrane protein into the nanodiscs, it is first purified in the detergent of choice, and then it is mixed with the detergent-lipid mixed micelles and MSP. Afterwards, the detergent is slowly removed by the use of a polymeric absorbent, called the Bio-Beads. Upon detergent removal, the MSP spontaneously assembles into the nanodiscs. In order to prepare homogenous nanodiscs, thorough optimization of assembly ratios is needed. During the process, empty nanodiscs assemble additionally to protein-filled nanodiscs, and further purification is required to remove them. It is usually done by an affinity chromatography or size-exclusion chromatography (SEC) (Inagaki et al., 2017). The schematic representation of the process leading to incorporation of a MP into the NDs is shown in Figure 8.

The examples of the cryo-EM structures made with an application of nanodiscs are: human neurotensin receptor 1 (NTSR1)-heterotrimeric G_i1 complex (Zhang et al., 2021a), rat transient receptor potential cation channel subfamily V member 1 (TRPV1) (Gao et al., 2016) and rat type 1 inositol 1,4,5-trisphosphate receptor (IP3R1) (Baker et al., 2021).

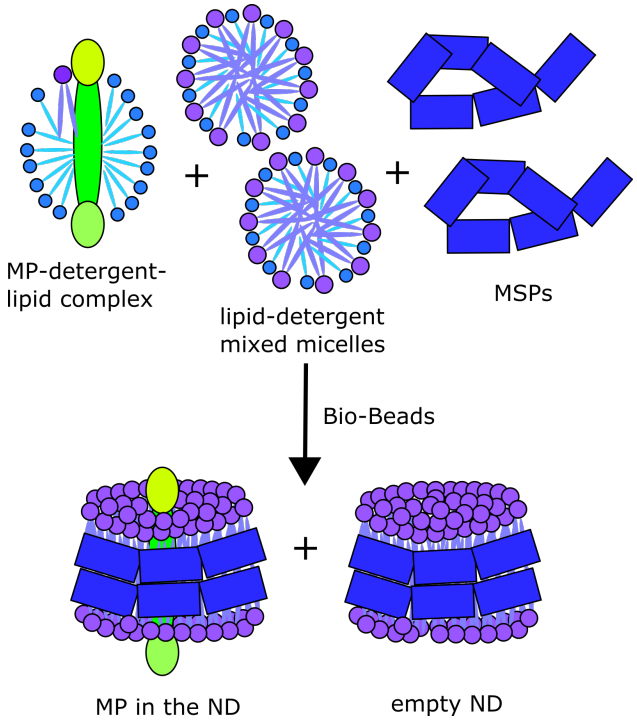

Figure 8. Incorporation of membrane proteins into nanodiscs.
A membrane protein (MP) (green) solubilized in the detergent solution is mixed with lipid-detergent mixed micelles and membrane scaffold protein (MSP). Upon detergent removal by the polymeric absorbent, called the Bio-Beads, the MSP spontaneously assembles around the lipids and MP, forming discoidal patches of artificial membrane, called the nanodiscs (NDs). During the process, also a population of empty NDs forms, and must be removed from the sample with a method of choice. Usually, it is achieved by affinity purification using an affinity tag on MP or size-exclusion chromatography. These are just schematic representations, not to scale.

4.3. Affinity binders as tools in science and medicine

The affinity binders constitute a group of natural and synthetic molecules that are able to bind a specific target. Among them, there are monoclonal antibodies (mAbs) and their antigen binding fragments (Fabs) and nanobodies (NBs), also called heavy-chain variable domains (VHHs), derived from camelid heavy-chain-only antibodies (hcAb). They are useful tools for structural biology, functional studies, diagnostics and medicine. The schematic representations of the structures of these molecules are depicted in Figure 9.

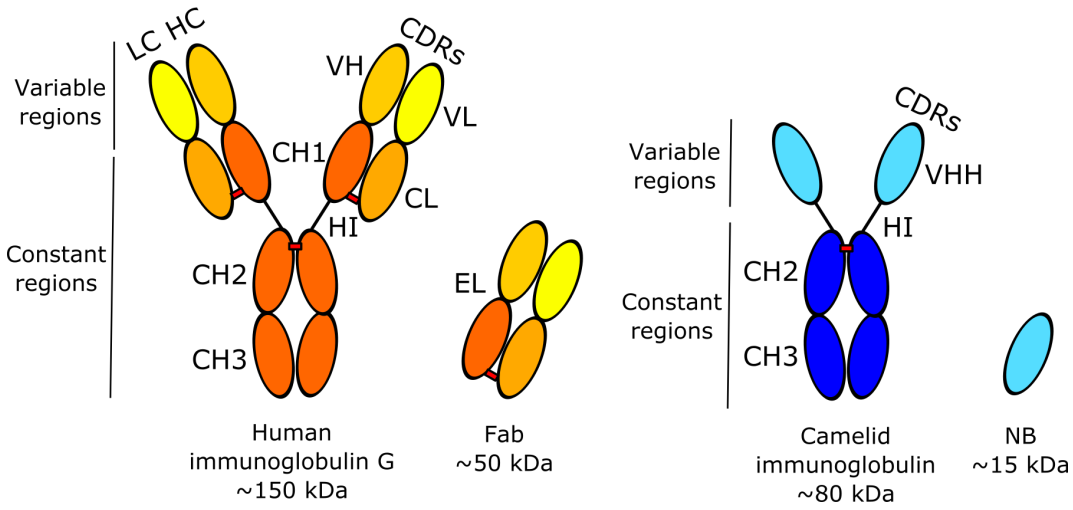

Figure 9. Immunoglobulins and their fragments.
The immunoglobulins are the protein complexes, produced by human and animal immune systems, able to bind to specific antigen. The Y-shaped human immunoglobulin G (IgG) has a molecular weight (M.W.) of approximately 150 kDa, and is composed of two heavy chains (HCs) and two light chains (LCs), attached via disulfide bonds (red rectangles). Each HC contains constant heavy 2 (CH2) and CH3 domains connected by a flexible hinge (HI) with CH1 domain, after which the constant region ends. It is further followed by the variable region, containing variable heavy (VH) region, hosting three CDRs that are able to bind to their antigen. Each LC is composed of constant light (CL) and variable light (VL) domains. The latter contains another three CDRs. The module that contains CH1 and VH, bound to CL and VL is called antigen binding fragment (Fab) with a M.W. of approximately 50 kDa. The partially flexible region connecting constant and variable regions of the Fab is called an elbow (EL). The camelid immunoglobulin is also called single-heavy chain antibody, because it contains only two heavy chains connected with disulfide bonds. Its molecular weight is approximately 80 kDa. The constant region has only CH2 and CH3, and is connected to a single variable region (VHH) by a flexible hinge. VHH contains CDRs that bind to the antigen. VHH alone is called a nanobody with approximate M.W. of 15 kDa. The internal disulfide bond of the nanobody is not shown. These are just schematic representations, not to scale.

4.3.1. Human immunoglobulins and Fabs

The human immunoglobulins (Igs) are approximately 150 kDa glycoproteins composed of two light chains (LCs) and two heavy chains (HCs). The LC can be made of either kappa or lambda chain, and contains variable light (VL) and constant light (CL) domains. The HC contains variable heavy (VH) domain and constant heavy regions CH1, CH2 and CH3 (in some classes of Igs, also CH4). Between CH1 and CH2 there is a flexible hinge (HI) region and disulfide bonds, holding both heavy chains together. There are five classes of Igs differing in structure of the heavy chain and length of the hinge region. They include IgG, IgA, IgD, IgE and IgM. Additionally, there are four subclasses of the most abundant IgG (IgG1-IgG4) and two subclasses of IgA (IgA1, IgA2). Each variable domain contains three complementarity determining regions (CDRs) that bind to the specific antigen. The whole antibody, also divides into two regions: antigen binding fragment (Fab) composed of light chain and part of a heavy chain, containing VH and CH1 domains, bound together via a

disulfide bond, and a so-called fragment crystallizable (Fc) that contains the hinge and the rest of an antibody. Its main function is to bind to Fc receptors on an effector cell surface, activate it and initialize further immune processes (Schroeder et al., 2010).

Igs can be obtained by using classical animal immunization and formation of hybridomas, or by applying artificial display techniques, like phage display, where the antibodies are exposed on surface of phages, and target is incubated against a large library of such bacteriophages with various CDRs combinations (Laustsen et al., 2021).

The antibodies are widely used in science and medicine. They can be diagnostic and therapeutic tools for many disorders, including cancer, inflammatory and autoimmune diseases (Shepard et al., 2017). The antibody-derived antigen binding fragments, Fabs, are important tools for structural biology. They may be used as mass boosters (+50 kDa) and fiducial marks for cryo-EM imaging (Bailey et al., 2017). Some Fabs can also decrease the inherent flexibility of the target protein or trap it in the specific conformation, desired for the research. In X-ray crystallography, Fabs may act as crystallization chaperones that mediate better contacts in the crystal lattice. The advantage of Fabs over full-length antibodies is that they are less flexible (Rader, 2009). However, Fabs also have a flexible region that sometimes causes problems with structure determination. It is called the elbow region, and is located between variable and constant regions. Bailey et al. used phage display technology to generate Fabs with modified, more rigid elbow regions. Such Fabs improved crystallization process as well as quality of crystals and the resolution of multiple protein structures. The other advantages of Fabs are their usual high stability and affinity to the target. Additionally, it is possible to express them in large quantities in bacteria that reduces effort and expense needed to produce them (Bailey et al., 2017, Rader, 2009).

The examples of proteins successfully crystallized with Fabs are: KcsA potassium channel from *Streptomyces lividans* (Zhou et al., 2001) and human β2-adrenergic receptor (Rasmussen et al., 2007). The several structures with Fabs were also obtained using cryo-EM. Among them there are: human ATP-binding cassette transporter ABCG5/G8 (Zhang et al., 2021b) and yeast ER membrane protein complex (EMC) (Miller-Vedam et al., 2020).

4.3.2. Nanobodies

The nanobodies (NBs), or single variable domains (VHHs), are fragments of camelid single-heavy-chain antibodies composed of their variable heavy chain, carrying three CDRs that specifically bind to their targets. Sharks also have similar single-chain antibodies that can be used for the same purposes as NBs. The nanobodies are tiny (approximately 15 kDa), very stable proteins that usually bind their targets with very high affinity. They can be obtained by immunization of alpacas, llamas, or other camelids with a target protein. The mRNA is then isolated from white blood cells of the animals, reverse transcribed, and cloned into phage display vectors to expose a range of isolated nanobodies on the surfaces of phages. Then the phage selection is performed using a target immobilized on beads or plates. The DNA of

phages, carrying candidates that are able to recognize the target, is then sequenced to determine the sequences of the corresponding nanobodies. Finally, the nanobodies are cloned into bacterial expression vectors, and can be easily expressed in bacterial periplasm in large quantities.

The nanobodies constitute very useful tools for structural biology. Because of their small size they can engage deep clefts, trapping the protein in a specific conformation and decreasing its flexibility that makes it more suitable for structural evaluation. The nanobodies can also be used as crystallization chaperones that mediate contacts in the crystal lattice (Dimitriev et al., 2016).

The examples of protein structures solved with nanobodies are: mouse serotonin 5-HT3 receptor (Hassaine et al., 2014), yeast proteasomal deubiquitylation module Rpn8-Rpn11 (Pathare et al., 2014) and human β2 adrenergic receptor (Rasmussen et al., 2011).

Except for structural biology applications, they can be also used in functional studies on target proteins. They may, for example, be used for tracing a cellular localization of a protein, or to research cellular mechanisms by trapping protein in the specific conformation. They can also activate or inhibit protein function (Dimitriev et al., 2016). What is worth noted, they can be applied as potential diagnostic and therapeutic molecules, because of possibility to express them in practically any living organism. Their advantage over classical antibodies is a small size, making them able to penetrate deeper into a tissue. There is ongoing research on the application of nanobodies in therapies against cancer as well as inflammatory and infectious diseases (Jovcevska et al., 2020).

5. Aim of the thesis

The aim of this thesis is to present two journeys pursued in order to structurally characterize two challenging molecular targets: STING and CD47. These proteins were selected because they constitute high-value medical targets, critical for development of many diseases, and essential for proper physiological functioning of human body. The first part of the thesis is focused on the structure of soluble domain of STING, crystallized in complex with activating nanobody. It was solved in order to characterize the nanobody binding interface and provide potential explanations on how this molecule is able to initiallize downstream signaling through STING. The second part of this thesis describes in detail multiple attempts made in order to obtain the structure of CD47 using cryo-EM. It also contains a comparison of the methodology presented here with the one used by a different group that published a medium resolution structure of this protein. It highlights the advantages of my scientific setup and contains suggestions on what could be done to improve the final results.

6. Materials

6.1. Chemicals

6.1.1. Common chemicals

All common chemicals have been purchased from Carl Roth (Karlsruhe, Germany) and Sigma-Aldrich (Darmstadt, Germany), unless otherwise stated. 2',3'-cGAMP was from Biolog (Hayward, CA, USA). TMB soluble reagent was from ScyTek Laboratories (Logan, UT, USA). GelRed stain was obtained from Biotium (San Francisco, CA, USA). SimplyBlue stain was from Invitrogen (Carlsbad, CA, USA). The detergents were purchased from Anatrace (Maumee, OH, USA). The lipids were obtained from Avanti Polar Lipids (Alabaster, AL, USA). The Bio-Beads SM-2 Resin was from Bio-Rad (Berkeley, CA, USA). The primers were synthesized by Metabion (Martinsried, Germany). The DNA was synthesized by Twist Bioscience (South San Francisco, CA, USA) and GenScript (Piscataway, NJ, USA).

6.1.2. Crystallization supplies

The crystallization screens were purchased from Molecular Dimensions (Holland, OH, USA), Hampton Research (Aliso Viejo, CA, USA) and Rigaku (Tokyo, Japan). MRC 2 96-well plates and Crystalgen SuperClear 24-well plates were obtained from Jena Bioscience (Jena, Germany).

6.2. Standards, enzymes, electrophoretic and chromatographic material

The molecular weight marker (PageRuler Prestained Protein Ladder, 10 to 180 kDa) and FastDigest DpnI enzyme were from Thermo Scientific (Waltham, MA, USA). Phusion DNA polymerase, T4 PNK, Q5 DNA polymerase and T4 DNA ligase were purchased from New England Biolabs (NEB) (Ipswich, MA, USA), Ni-NTA was from Macherey-Nagel (Düren, Germany), TALON resin was from Takara Bio (Shiga, Japan), Strep-Tactin XT Superflow resin and Strep-Tactin XT high-capacity resin were from IBA Lifesciences (Göttingen, Germany), SEC columns: Superdex 200 16/600, Superdex 75 16/600, Superdex 200 10/300, Superdex 200 5/150 and Superose 6 10/300 were purchased from GE Healthcare (Chicago, IL, USA), 4-20% gradient polyacrylamide gels were from Serva (Heidelberg, Germany), benzonase nuclease was from Merck (New York, NY, USA).

6.3. Cell culture

6.3.1. Mammalian cell culture supplies

The cells and other supplies for mammalian cell culture were purchased from Gibco (New York, NY, USA).

6.3.2. Bacterial cell culture media

Turbo Broth medium was from AthenaES (Baltimore, MD, USA). The ingredients of home-made LB and TB media are shown below.

LB - 1% (w/v) bacto trypton / 0.5% (w/v) yeast extract / 0.5% (w/v) NaCl. pH should be adjusted with NaOH to 7.0.

TB - For 1 L mix 900 mL of autoclaved solution 1 with 100 mL of autoclaved solution 2.
Solution 1 - 1.33% (w/v) bacto-trypton / 2.67% (w/v) bacto-yeast extract / 0.44% (v/v) glycerol
Solution 2 - 169.75 mM KH_2PO_4 / 719.92 mM K_2HPO_4 x $3H_2O$

6.3.3. Antibiotic stocks

The concentrations shown below are for 1000x concentrated stock solutions.
Ampicillin - 100 mg/mL in ddH_2O
Kanamycin - 50 mg/mL in ddH_2O
Chloramphenicol - 34 mg/mL in 100% ethanol

6.3.4. Bacterial strains

E. coli XL-1 blue cells (Agilent; Santa Clara, CA, USA)
E. coli Rosetta 2(DE3) (Sigma-Aldrich; Darmstadt, Germany)
E. coli WK6 (provided by Florian Schmidt)
E. coli SHuffle T7 cells (NEB; Ipswich, MA, USA)

6.3.5. Plasmids

Membrane scaffold proteins (MSPs)-harboring pET28a vectors were purchased from Addgene (Watertown, MA, USA). pET28M vector was already available in our laboratory. pHEN6 vector was provided by Florian Schmidt.

6.3.6. DNA purification and transfection

DNA purification kits were purchased from Macherey-Nagel (Düren, Germany). Polyethylenimine (PEI) MAX (M.W. 40,000 Da) transfection reagent was obtained from Polysciences (Warrington, PA, USA).

6.4. ELISA, WB and SPR

6.4.1. Supplies

Streptavidin magnetic particles were from Roche Diagnostics (Basel, Switzerland), high-binding, 96-well, clear, flat-bottom plates were from Corning (Corning, NY, USA), EZ-Link NHS-PEG4-Biotin reagent and NeutrAvidin were purchased from Thermo Scientific (Waltham, MA, USA), Trans-Blot Turbo Mini 0.2 µm PVDF Transfer Pack was from Bio-Rad (Berkeley, CA, USA), Ni-NTA SPR chip and HBS-P+ (10x) buffer were from Cytiva (Marlborough, MA, USA).

6.4.2. Antibodies

Anti-His antibody-HRP (further called: “anti-his-tag-HRP antibody”) was from Miltenyi Biotec (Bergisch Gladbach, Germany), anti-HA-tag (6E2) mouse mAb-HRP (further called: “anti-HA-HRP antibody”) was from Cell Signaling Technology (Danvers, MA, USA), anti-strep Tag II antibody-HRP (further called: “anti-strep-HRP antibody”) was from Sigma Aldrich (Darmstadt, Germany), goat anti-human IgG, $F(ab')_2$ fragment specific (further called “anti-Fab-HRP antibody”) and AffiniPure Goat Anti-Mouse IgG, Fcγ fragment specific HRP-conjugated antibody (further called “anti-mouse IgG-HRP antibody”) were from Jackson ImmunoResearch (Philadelphia, PA, USA), mouse anti-human STING (TMEM173) antibody was from BioLegend (San Diego, CA, USA), mouse anti-human CD47 antibody (B6H12) was from Invitrogen (Carlsbad, CA, USA).

6.5. Electron microscopy

The grids were purchased from Quantifoil (Großlöbichau, Germany).

6.6. Software and webpages

AutoBuild/Phenix (McCoy et al., 2007)
Avanti Polar Lipids - https://avantilipids.com/
Biacore Evaluation software - Cytiva (Marlborough, MA, USA)
CHAINSAW (Stein, 2008)
COOT (Emsley et al., 2010)
CryoSPARC - Structura Biotechnology (Toronto, Canada)

Inkscape - https://inkscape.org/
Keynote - Apple Inc. (Cupertino, California)
NEB Tm calculator - https://tmcalculator.neb.com/
OriginPro - OriginLab Corporation (Northampton, MA, USA)
Phaser/Phenix (McCoy et al., 2007)
Phenix.refine/Phenix (Adams et al., 2010)
POINTLESS (Evans, 2006)
PROCHECK (Laskowski et al., 1993)
https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/
ProtParam tool - https://web.expasy.org/protparam/
PyMOL - Schrödinger Inc. (New York, NY, USA)
XDS (Kabsch, 2010)
XSCALE (Kabsch, 2010)

6.7. Buffers and other solutions

6.7.1. Purification of MSPs

Buffer M1 - 40 mM HEPES, pH 8.0 / 300 mM NaCl / 1% (v/v) Triton X-100
Buffer M2 - 40 mM HEPES, pH 8.0 / 300 mM NaCl / 20 mM imidazole / 50 mM sodium cholate
Buffer M3 - 40 mM HEPES, pH 8.0 / 300 mM NaCl / 50 mM imidazole
Buffer M4 - 40 mM HEPES, pH 8.0 / 300 mM NaCl / 400 mM imidazole
Buffer M5 - 20 mM HEPES, pH 7.4 / 150 mM NaCl

6.7.2. Purification of Fabs and CBDs

Buffer F1 - 20 mM HEPES, pH 7.4 / 150 mM NaCl
Buffer F2 - 20 mM HEPES, pH 7.4 / 200 mM NaCl / 20 mM imidazole
Buffer F3 - 20 mM HEPES, pH 7.4 / 500 mM NaCl / 40 mM imidazole
Buffer F4 - 20 mM HEPES, pH 7.4 / 150 mM NaCl / 350 mM imidazole

6.7.3. Purification of nanobodies

Buffer N1 - 200 mM TRIS, pH 8.0 / 0.5 mM EDTA / 0.5 M sucrose
Buffer N2 - 20 mM HEPES, pH 7.4 / 200 mM NaCl / 20 mM imidazole
Buffer N3 - 20 mM HEPES, pH 7.4 / 500 mM NaCl / 40 mM imidazole
Buffer N4 - 20 mM HEPES, pH 7.4 / 150 mM NaCl / 350 mM imidazole
Buffer N5 - 20 mM HEPES, pH 7.4 / 150 mM NaCl

6.7.4. Purification of membrane proteins

Buffer D1 - 50 mM HEPES, pH 7.4 / 200 mM NaCl
Buffer D2 - 50 mM HEPES, pH 7.4 / 200 mM NaCl / 1% (w/v) DDM / 0.1% (w/v) CHS
Buffer D3 - 20 mM HEPES, pH 7.4 / 200 mM NaCl / 3% glycerol (v/v) / 0.1% (w/v) DDM / 0.01% (w/v) CHS
Buffer D4 - 20 mM HEPES, pH 7.4 / 200 mM NaCl / 3% glycerol (v/v) / 50 mM D-biotin / 0.1% (w/v) DDM / 0.01% (w/v) CHS

6.7.5. Binding assays

Buffer B1 - 20 mM HEPES, pH 7.4 / 200 mM NaCl / 3% glycerol (v/v) / 0.1% (w/v) DDM / 0.01% (w/v) CHS
Buffer B2 - 20 mM HEPES, pH 7.4 / 200 mM NaCl / 3% (v/v) glycerol / 50 mM D-biotin / 0.1% (w/v) DDM / 0.01% (w/v) CHS
Buffer B3 - 20 mM HEPES, pH 7.4 / 150 mM NaCl
Buffer B4 - 20 mM citric acid, pH 3.5 / 150 mM NaCl
Buffer B5 - 20 mM HEPES, pH 7.4 / 200 mM NaCl / 20 mM imidazole
Buffer B6 - 20 mM HEPES, pH 7.4 / 500 mM NaCl / 40 mM imidazole
Buffer B7 - 20 mM HEPES, pH 7.4 / 150 mM NaCl / 350 mM imidazole

6.7.6. Reconstitution of nanodiscs

Buffer S1 - 25 mM HEPES, pH 7.4 / 200 mM NaCl
Buffer S2 - 20 mM HEPES, pH 7.4 / 200 mM NaCl

6.7.7. Crystallization of proteins

Buffer X1 - 20 mM HEPES, pH 7.4 / 150 mM NaCl

6.7.8. Biotinylation and ELISA

Buffer E1 - 20 mM HEPES, pH 7.5 / 150 mM NaCl
Buffer E2 - 100 mM sodium carbonate, pH 9.6
Buffer E3 - 20 mM HEPES, pH 7.5 / 150 mM NaCl / 1% (w/v) BSA

6.7.9. SDS-PAGE

TGS buffer (10x) - 0.25 M TRIS / 1.92 M glycine / 1% (w/v) SDS
Gel loading dye (4x) - 0.11 M TRIS, pH 6.8 / 16% (v/v) glycerol / 4% (w/v) SDS / 5% (v/v) β- mercaptoethanol / 0.05% (w/v) bromophenol blue

Coomassie Brilliant Blue stain - 0.2% (w/v) Coomassie Brilliant Blue R250, 50% (v/v) ethanol, 7% (v/v) acetic acid

6.7.10. Agarose gel electrophoresis

TAE buffer (50x) - 2 M TRIS, pH 8.0 / 1 M acetic acid / 50 mM EDTA
Agarose gel loading dye (6x) - 10 mM TRIS-HCl (pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% (v/v) glycerol, 60 mM EDTA

6.7.11. SPR

HBS-P+ (10x) - 0.1 M HEPES, pH 7.4 / 1.5 M NaCl / 0.5% v/v Surfactant P20

6.7.12. Western blotting

PBST buffering solution - 137 mM NaCl / 2.7 mM KCl / 10 mM Na_2HPO_4 / 1.8 mM KH_2PO_4 / 0.1% (v/v) Tween 20
WB staining solution - 100 mM TRIS, pH 8.5 / 12.5 μM luminol / 2.25 μM cumaric acid, freshly add 3 μL of 30% (v/v) H_2O_2 per every 10 mL of staining solution. This solution can be stored in the fridge and used multiple times.

6.7.13. Protein purification additives

Protease inhibitors mix (100x) - 200 mM benzamidin HCl / 100 mM phenylmethylsulfonyl fluoride (PMSF) / 200 μM pepstatin A / 60 μM leupeptin hemisulfate in 100% ethanol.

7. Methods

7.1. Molecular cloning

7.1.1. Molecular cloning - general methods

The molecular cloning was peformed using mainly Gibson assembly method (Gibson et al., 2009) for insertion of selected genes into the plasmids, and around-the-horn polymerase chain reaction (ATH PCR) and assembly (not published) to remove, mutate or insert short fragments of DNA within a longer DNA sequence.

For Gibson assembly, standard polymerase chain reaction (PCR) was performed to generate an insert and linear vector with usually 20 bp overlaps with melting temperatures more than 50 °C. Primers were ordered from Metabion (Martinsried, Germany) and dissoved in purified deionized water (ddH_2O) to final concentration of 100 µM. For standard PCR, usually, 100 ng of template were mixed with 1 µL each of forward and reverse primers (10 µM each) and Phusion polymerase (NEB) master mix to final volume of 20 µL. The melting temperatures of the primers were calculated using NEB Tm calculator (https://tmcalculator.neb.com/). The elongation time was usually 30" per kilo base pair (kbp). Number of PCR cycles was 30, followed by the final elongation three times longer than a single elongation time.

For ATH PCR, the primers were designed to form a linear vector with blunt ends. One of the ends contained the modification. Firstly, the primers were phosphorylated using T4 polynucleotide kinase (T4-PNK) (NEB) at 10 U/µL according to manufacturer's guidelines, briefly, 1.25 µL of 100 µM forward and reverse primers each were mixed with 2.5 µL of T4 DNA ligase buffer (NEB) and 0.5 µL T4-PNK in final volume of 25 µL, adjusted with ddH_2O. The reaction mix was then incubated for 1 h at 37 °C, and the enzyme was inactivated by incubation for 20' at 65 °C. T4-PNK buffer contains the same ingredients as T4 DNA ligase buffer, except for ATP. This feature makes it possible to add standard or radioactive ATP to the former, and customize the reaction. T4 DNA ligase buffer was used here for convenience. The phosphorylated primers mix was used directly (it can be stored indefinitely at -20 °C) in 50 µL PCR reactions containing 2 mM each dNTPs (Carl Roth), Q5 DNA polymerase (NEB), Q5 reaction buffer (NEB) and GC enhancer (NEB), according to manufacturer's guidelines. The PCR was done in the same way as the standard PCR, but elongation time was increased by a single, additional 30".

After the PCR reactions, the original, methylated DNA template was digested using FastDigest DpnI enzyme (Thermo Scientific) according to manufacturer's guidelines. Then the products were supplemented with 6x agarose gel loading dye, and separated on 1% agarose gel supplemented with GelRed stain (Biotium). The gel has been run for approxinately 1 h at 120 V in 1x TAE buffer. The DNA was purified using NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel) according to manufacturer's guidelines.

Then, for Gibson, three-fold molar excess of insert DNA over the template was used in 10 µL reactions containg 5 µL of DNA mix and 5 µL of 2x Gibson master mix. The reactions were performed for 1 h at 50 °C.

The linear vector produced in ATH-PCR was circularized using ligation. The total volume of the reactions was 20 µL, and was adjusted with ddH_2O. The reactions contained 3-5 µL of ATH-PCR-derived product, 1 µL of T4 DNA ligase from 2,000,000 u/mL stock (NEB) and 2 µL of T4 DNA ligase buffer. They were performed for 2 h at 22 °C.

Finally, the chemically competent *E. coli* XL-1 blue cells (Agilent) were transformed with assembled DNA from either of the methods described above, using the heat shock method. The cells were incubated with 2-5 µL of DNA for 5' on ice which was followed by incubation for 45" in the water bath heated up to 42 °C and then 2' on ice. Then 700 µL of lysogeny broth (LB) medium were added, and the cells were recovered with shaking at 800 rpm, at 37 °C for 1 h. In the end, the cells were plated on antibiotic-containing agar plates (antibiotic(s) depending on vector used) and incubated overnight at 37 °C. On the following day, the single colonies were picked into 5 mL of LB medium, containing appropriate antibiotic, in 13 mL plastic, round-bottom tubes, and incubated overnight with shaking at 200 rpm at 37 °C. On the following day, the cells were spun down for 5' at 3,500 rpm. The plasmid DNA was then purified from the cell pellets using NucleoSpin Plasmid EasyPure kit (Macherey-Nagel) according to manufacturer's guidelines. The DNA concentration was measured using nanodrop (NanoPhotometer NP80, Implen) at 260 nm. The sequencing was done by Eurofins Genomics (Ebersberg, Germany).

7.1.2. Cloning of MSPs

The membrane scaffold proteins (MSPs)-harboring pET28a vectors with kanamycin resistance were purchased from Addgene (Watertown, MA, USA) in three length variants: MSP1D1dH5, MSP1D1 and MSP1E3D1. The first of them contains N-terminal six histidine-tag (6his-tag) and the rest possesses N-terminal 7his-tag; all followed by a Tobacco etch virus (TEV) protease cleavage site. They are used to form nanodiscs with three different diameters: 9 nm, 10 nm and 13 nm, respectively.

7.1.3. Cloning of Fabs

HuFab for cytoplasmic expression was cloned as a bicistronic gene with untagged light chain (LC) and C-terminally 6his-tagged heavy chain (HC). The genes were derived from pFUSE vector, containing the full-length Hu5F9 antibody (already available in the lab) and were cloned into pET28M vector with kanamycin resistance using the Gibson method. Hu5F9 antibody was developed by Liu et al. (Liu et al., 2015). Initially, by mistake, the Fab was cloned without the first aspartic acid in the LC (further called “HuFab-D”). It was later noticed and corrected using ATH PCR, and the product is further called “HuFab”. However, it does not seem to interfere anyhow with the proper function of this Fab.

HuFab-H12-P is a variant of HuFab with a modified elbow region $V^{111}FNQIKP^{118}$ in HC to make it more rigid (Bailey et al., 2018). The DNA insert containing the sequence of HuFab-H12-P with 6his-tag at the C-terminus of HC, codon-optimized for *E. coli* was ordered from Twist Bioscience (South San Francisco, CA, USA), and cloned into pET28M vector.

7.1.4. Cloning of nanobodies

7.1.4.1. Anti-Fab nanobody

The DNA insert containing the sequence of anti-Fab nanobody (aF-NB) (Ereño-Orbea et al., 2018) with C-terminal 6his-tag, codon-optimized for *E. coli* was synthesized by Twist Bioscience, and then it was cloned into pHEN6 vector with ampicillin resistance (provided by Florian Schmidt).

7.1.4.2. Nanobody 1829

Nanobody 1829 (NB1829) was generated using alpaca immunization in Florian Schmidt's laboratory at the University of Bonn, Germany. It was researched by Jennifer Deborah Wuerth. The original NB1829 construct in pHEN6 vector contained a short linker composed of two glycins followed by C-terminal hemaglutinin tag (HA-tag) and 6his-tag. For the purpose of this thesis, the HA-tag was removed from the original construct using ATH PCR.

7.1.5. Cloning of STING

Human STING cyclic dinucleotide-binding domain (STING CBD) (aa 139-345) is a cytoplasmic part of STING C-terminal domain (STING CTD) (aa 139-379) without C-terminal tail (CTT) (aa 346-379). Full-length human STING (STING FL) (aa 1-379) and its CBD, both with two mutations R220H and H232R, were cloned into pSecTag vector (for expression in mammalian cells) and pET28M vector (for expression in bacterial cells), respectively from the construct in pFBDM vector provided by Liudmila Andreeva. Later on R220H mutation in STING FL was modified to wild-type (WT) (further called STING FL) using ATH PCR. However, all of these mutations are considered natural, since the origin of these sequences is a mRNA library. STING FL contains C-terminal twin strep-tag without a cleavage site, and STING CBD contains N-terminal 6his-tag, followed by a thrombin site and SUMO1 protein sequence. Later in the study, two double mutants of original STING CBD were generated in the same vector using ATH PCR - D85K/L86E (further called CBD mutant 1) and E190R/R193A (further called CBD mutant 2) Here, for simplicity, the amino acid positions are numbered assuming the first amino acid of CBD as number 1. Thus, here, position 1 corresponds to 139^{th} amino acid of FL protein.

7.1.6. Cloning of CD47

Full-length human CD47 isoform 2 with C-terminal twin strep-tag was synthesized by GenScript (Piscataway, NJ, USA) and cloned into pSecTag vector with ampicillin resistance using Gibson assembly.

7.2. Expression and purification of proteins

7.2.1. Membrane scaffold proteins

7.2.1.1. Expression of MSPs

E. coli Rosetta 2(DE3) (Sigma-Aldrich) chemically competent cells harboring chloramphenicol resistant plasmid pRARE2, carrying the sequences coding tRNAs for 7 rare codones (AGA, AGG, CCC, GGA, CGG, AUA and CUA) were transformed with kanamycin resistant pET28a plasmids with respective MSP genes. 2 µL of DNA were incubated with 75 µL of competent cells for 5'. The transformation was performed using heat shock method by incubation at 42 °C for 45" followed by incubation for 2' on ice. The cells were supplemented with 700 µL of LB medium and recovered at 37 °C for 1 h with shaking at 800 rpm. The cells were then transferred into 100 mL of LB medium containing 50 µg/mL kanamycin (Carl Roth) and 34 µg/mL chloramphenicol (Carl Roth) and grown overnight at 37 °C with shaking at 200 rpm. On the following day, 10 mL of saturated cell suspension were used per 1 L of terrific broth (TB) medium supplemented with the same antibiotics. The cells were grown at 37 °C with shaking at 160 rpm until reaching an optical density at 600 nm (OD_{600}) of 1. The expression was then induced with 1 mL/L of 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) (Carl Roth) and performed for 4 h at the same parameters. The cells were then harvested by centrifugation in SLC-6000 rotor (Thermo Scientific) at 5000 rpm at 4 °C for 10'. The cell pellet was flash frozen in liquid nitrogen and stored at -80 °C.

7.2.1.2. Purification of MSPs

All MSP purification procedures were performed at 4 °C unless overwise stated. All amounts presented here are adjusted for cell pellet obtained from 1.5 L of bacterial culture. The MSP was purified as in (Bayburt et al., 2002) with several modifications. The pellet was thawed and re-suspended in 40 mL total of buffer M1 freshly supplemented with a home-made protease inhibitors mix, using a vortex and glass Dounce homogenizer. The cells were lysed by sonication (Branson sonicator) on ice, and the insoluble debris was separated from the solution by centrifugation in SS34 rotor (Thermo Scientific) at 17 krpm at 4 °C for 30'. The supernatant was incubated with 3 mL nickel-nitrilotriacetic acid agarose (Ni-NTA) (Macherey-Nagel) equillibrated with buffer M1 on rotator for 30'. The resin was separated from the flow-through (FT) by centrifugation at 2,500 rpm in the Centrifuge 5810 R (Eppendorf) at 4 °C for 5'. Then the resin was washed two times with 15 mL of buffer M1 using centrifugation at the same parameters. Afterwards, the resin was loaded onto a gravity column and washed with five column volumes (CVs) each of buffers M1 and M2, and 10

CVs of buffer M3. The protein was eluted with 2 x 3 CVs of buffer M4. The protein was supplemented with 1 mM final ethylenediaminetetraacetic acid (EDTA) (Carl Roth) and dialyzed overnight against 5 L of buffer M5 in 6-8 kDa molecular weight cut-off (M.W.C.O.) dialysis bags (Carl Roth) at 4 °C. On the following day, the protein, usually at approximately 3 mg/mL, was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

Some MSPs were also cleaved with TEV protease to remove his-tag. For this purpose, a home-made TEV protease was additionally added to the dialysis bag and buffer M5 was freshly supplemented with 2 mM final 2-mercaptoethanol (BME) (Carl Roth). On the following day, a subtractive immobilized metal affinity chromatography (IMAC) was performed to remove TEV protease and potential remaining his-tagged components. For this purpose, the cleavage mix was incubated with 3 mL of Ni-NTA equillibrated with buffer M5 on rotator for 30'. The resin was loaded onto a gravity column and washed with 5 CVs of buffer M5. The his-tagged components were eluted from the resin using 5 CVs of buffer M4. The FT and wash fractions were collected, concentrated using 10 kDa M.W.C.O. Amicon concentrators (Millipore) and ran on Superdex 200 16/600 column (GE Healthcare) equillibrated with buffer M5 or dialyzed overnight again as in the protocol without cleavage. The final product, usually at approximately 3 mg/mL, was flash frozen in liquid nitrogen and stored at -80 °C.

The samples were mixed with 4x gel loading dye, heated up to 95 °C for 5’ and analyzed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) on 4-20 % gradient polyacrylamide gel (Serva) ran at 300 V in Tris-Glycine-SDS (TGS) buffer. The gel was stained with a Coomassie Brilliant Blue stain.

7.2.2. Fabs

7.2.2.1. Expression of Fabs

E. coli SHuffle T7 chemically competent cells (NEB), designed to have partially oxidative cytoplasm, enabling disulfide bonds formation, were transformed with kanamycin resistant pET28M plasmids with respective bicistronic Fab sequences. All Fabs contain C-terminal 6his-tag on HC. 2 µL of DNA were incubated with 100 µL of competent cells for 5' on ice. The transformation was performed using heat shock method as described in the section above. The cells were supplemented with 700 µL of LB medium and recovered at 37 °C for 1 h with shaking at 800 rpm. The cells were then transferred into 100 mL of LB medium containing 50 µg/mL kanamycin and grown overnight at 30 °C with shaking at 200 rpm. On the following day, 30 mL of saturated cell suspension were used per 1 L of Turbo Broth medium (AthenaES) supplemented with 50 µg/mL final kanamycin. The cells were grown at 30 °C with shaking at 160 rpm until reaching OD_{600} of 1, cooled down to 20 °C at 140 rpm for 30' and induced with 1 mL/L of 1 M IPTG. The expression was performed overnight at the same parameters. The cells were then harvested by centrifugation in F9-6x1000 LEX rotor (Thermo Scientific) at 5,000 rpm at 4 °C for 10'. The cell pellet was frozen and stored at -80 °C.

7.2.2.2. Purification of Fabs

All Fabs purification procedures were performed at room temperature (RT) unless overwise stated. All collected fractions were stored on ice. All amounts presented here are adjusted for 6 L cell pellet, since expression level of the Fabs is low. The pellet was thawed and re-suspended in 150 mL total of buffer F1 freshly supplemented with a home-made protease inhibitors mix and 5 µL of benzonase nuclease (Merck) using a vortex and glass Dounce homogenizer. The cells were lysed by sonication on ice. The non-assembled Fab components and bacterial proteins were then denatured by incubation at 63 °C for 30'. Intact Fabs are usually stable up to 75 °C or higher. Afterwards, the lysate was cooled down for 10' on ice and the insoluble debris was separated from the solution by centrifugation in A27-8x50 rotor (Thermo Scientific) at 17 krpm at 4 °C for 30'. The supernatant was incubated with 6 mL of Ni-NTA resin equillibrated with buffer F1 on rotator for 30'. The resin was separated from the FT by centrifugation at 2,500 rpm in the Centrifuge 5810 R at 4 °C for 5'. Then the resin was loaded onto a gravity column and washed three times with 10 CVs of buffer F2 and buffer F3. In the end, the protein was eluted with 5 CVs of buffer F4, concentrated using 30 kDa M.W.C.O. Amicon concentrators, and ran on Superdex 200 16/600 column equillibrated with buffer F1 at 4 °C. The main-peak fractions were pooled, concentrated to, usually 3-5 mg/mL, frozen and stored at -80 °C.

The samples were mixed with 4x gel loading dye, heated up to 95 °C for 5’ and analyzed using SDS-PAGE on 4-20% gradient polyacrylamide gel ran at 300 V in TGS buffer. The gel was stained with a SimplyBlue stain (Invitrogen). Additionally, the thermal stability was analyzed using Tycho NT.6 device (Nanotemper) in 10 µL capillaries.

7.2.3. Nanobodies

7.2.3.1. Expression of nanobodies

The periplasmic expression of nanobodies was performed in chemically competent *E. coli* WK6 cells provided by Florian Schimdt. The cells were transformed with pHEN6 plasmid carrying the nanobody with pelB signal sequence sending the product to the periplasm. First, 2 µL of DNA were incubated with 100 µL of competent cells for 5'. Then the transformation was performed using heat shock method as described in the section above. The cells were supplemented with 700 µL of LB medium and recovered at 37 °C for 1 h with shaking at 800 rpm. The cells were then transferred into 100 mL of LB medium containing 100 µg/mL ampicillin final (Carl Roth) and grown overnight at 37 °C with shaking at 200 rpm. On the following day, 5 mL of saturated cell suspension were used per 1 L of Turbo Broth medium supplemented with the same antibiotic. The cells were grown at 37 °C with shaking at 160 rpm until reaching OD_{600} of 0.8, cooled down to 28 °C at 140 rpm for 30' and induced with 1 mL/L of 1 M IPTG. The expression was performed overnight at the same parameters. The cells were then harvested by centrifugation in F9-6x1000 LEX rotor at 5000 rpm at 4 °C for 10'. The cell pellet was flash frozen in liquid nitrogen and stored at -80 °C.

7.2.3.2. Purification of nanobodies

All NBs purification procedures were performed at RT unless overwise stated. All collected fractions were stored on ice. All amounts presented here are adjusted for 1 L cell pellet. The proteins were released from periplasmic space without lysing the whole cells using two-step osmotic shock. Firstly, the pellet was thawed and gently re-suspended in 15 mL of buffer N1 freshly supplemented with the home-made protease inhibitors mix using vortexing. The cells were incubated in this buffer with gentle rotation for 1 h at 4 °C. Then the suspension was incubated with additional 30 mL final of 4-fold diluted buffer N1 freshly supplemented with the protease inhibitors mix, with gentle rotation for 1 h to overnight at 4 °C. Afterwards, the insoluble debris was removed using centrifugation in A27-8x50 rotor at 17 krpm at 4 °C for 30'. The supernatant was incubated with 6 mL of Ni-NTA resin, equillibrated with buffer N1 on rotator for 30'. The resin was separated from the FT by centrifugation at 2,500 rpm in the Centrifuge 5810 R at 4 °C for 5'. Then the resin was loaded onto a gravity column and was washed three times with 10 CVs of buffer N2 and buffer N3. The protein was eluted with 5 CVs of buffer N4, concentrated in 10 kDa M.W.C.O. Amicon concentrators and ran on Superdex 75 16/600 column (GE Healthcare), equillibrated with buffer N5. The protein at 3-5 mg/mL was flash frozen in liquid nitrogen and stored at -80 °C.

The samples were mixed with 4x gel loading dye, heated up to 95 °C for 5' and analyzed using SDS-PAGE on 4-20% gradient polyacrylamide gel ran at 300 V in TGS buffer. The gel was stained with a SimplyBlue stain. Additionally, the thermal stability was analyzed using Tycho NT.6 device in 10 μL capillaries.

7.2.4. STING CBD and its mutants

7.2.4.1. Expression of STING CBD and its mutants

E. coli Rosetta 2(DE3) chemically competent cells were transformed with a kanamycin resistant pET28M plasmid carrying STING CBD (or its mutant) gene. 2 μL of DNA were incubated with 75 μL of competent cells for 5'. The transformation was performed using heat shock method as described above. The cells were supplemented with 700 μL of LB medium and recovered at 37 °C for 1 h with shaking at 800 rpm. The cells were then transferred into 100 mL of LB medium containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol final, and grown overnight at 37 °C with shaking at 200 rpm. On the following day, 10 mL of saturated cell suspension were used per 1 L of Turbo Broth medium supplemented with the same antibiotics. The cells were grown at 37 °C with shaking at 160 rpm until reaching OD_{600} of 1 and induced with 1 mL/L of 1 M IPTG. The expression was performed for 4 h at the same parameters. The cells were then harvested by centrifugation in F9-6x1000 LEX rotor at 5000 rpm at 4 °C for 10'. The cell pellet was flash frozen in liquid nitrogen and stored at -80 °C.

7.2.4.2. Purification of STING CBD and its mutants

All STING CBD purification procedures were performed at 4 °C unless overwise stated. All amounts presented here are adjusted for 3 L cell pellet. The pellet was thawed and re-suspended in 90 mL total of buffer F1 freshly supplemented with the home-made protease inhibitors mix and 5 µL of benzonase nuclease using a vortex and glass Dounce homogenizer. The cells were lysed by sonication on ice. The insoluble debris was separated from the solution by centrifugation in A27-8x50 rotor at 17 krpm at 4 °C for 30'. The supernatant was incubated with 5 mL of Ni-NTA resin equillibrated with buffer F1 on rotator for 30'. The resin was separated from the FT by centrifugation at 2,500 rpm in the Centrifuge 5810 R at 4 °C for 5'. The resin was loaded onto a gravity column, and washed four times with 5 CVs of buffer F2 and buffer F3. In the end, the protein was eluted with 5 CVs of buffer F4, concentrated with 10 kDa M.W.C.O. Amicon concentrators, and ran on Superdex 200 16/600 column equillibrated in buffer F1. The main-peak fractions were pooled, concentrated to 3-5 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C.

The samples were mixed with 4x gel loading dye, heated up to 95 °C for 5’ and analyzed using SDS-PAGE on 4-20% gradient polyacrylamide gel ran at 300 V in TGS buffer. The gel was stained with a SimplyBlue stain. Additionally, the thermal stability was analyzed using Tycho NT.6 device in 10 µL capillaries.

7.2.5. Full-length STING and CD47

7.2.5.1. Expression of full-length STING and CD47

DNA for transfection was prepared using NucleoBond PC 10,000 kit (Macherey-Nagel) using giga prep protocol according to manufacturer's guidelines. On the day preceeding the transfection, 450 mL of Expi293F cells (Gibco) at 2.0×10^6 cells/mL were prepared in Freestyle 293 medium (Gibco) supplemented with 1x Pluronic F-68 (Gibco) and 1x penicillin-streptomycin (Gibco), or in Expi293 medium (Gibco) without supplements. The former medium was used in the later phases of research to lower the costs. The transfection was perforemed using PEI MAX (M.W. 40,000 Da) transfection reagent (Polysciences). On the day of transfection the cells were counted again, and should be at approximately 3.0×10^6 cells/mL. First, 47 mL of FreeStyle 293 medium without additives (or Expi293 medium) were thoroughly mixed by inversion with 750 µg of plasmid DNA, containing full-length STING or CD47 gene. Then 3 mL of 1 g/L PEI MAX reagent solution in ddH_2O (pH 7.0; adjusted with NaOH) were added and mixed by inversion with the above. A 15' incubation at RT was performed. The transfection mix was then added to the cells along with additional amount of Pluronic F-68 and penicillin-streptomycin in case of FreeStyle 293 medium. The cells were then grown in ventilated flasks in an incubator with shaking at 95 rpm at 37 °C with 8% CO_2. After 24 h, 500 mL of fresh medium with appropriate supplements as above were added to the cells to sustain their growth and viability. The cells were then grown for additional 3 days and then harvested using centrifugation in F9-6x1000 LEX rotor at 3,000 rpm at 4 °C for 15'. The cell pellet was flash frozen in liquid nitrogen and stored at -80 °C.

7.2.5.2. Purification of full-length STING and CD47

All purification procedures were performed at 4 °C unless overwise stated. All amounts presented here are adjusted for 1 L cell pellet. The pellet was thawed and re-suspended in 120 mL total of buffer D1 freshly supplemented with a home-made protease inhibitors mix and 5 µL of benzonase nuclease using a vortex and glass Dounce homogenizer. The cells were lysed by sonication on ice. The cell membranes were pelleted using ultracentrifugation in type 45 Ti rotor (Beckman) at 40 krpm at 4 °C for 1 h. The supernatant was discarded, and the membrane pellet was re-suspended in the high-detergent buffer D2 freshly supplemented with protease inhibitors mix, using a paint brush and glass Dounce homogenizer. The membranes were solubilized for at least 2 h with gentle, horizontal rotation. Then the insoluble debris was removed by ultracentrifugation at the same parameters as above for 30'. The supernatant was incubated with 3 mL of Strep-Tactin XT high-capacity resin (IBA Lifesciences) equillibrated with buffer D2 for 30' with gentle, horizontal rotation. The resin was then loaded onto a gravity column and washed three times with 2 CVs of buffer D3. The protein was eluted with 3 CVs of buffer D4 applied in three equal portions. The elution was concentrated using 30 kDa M.W.C.O. Amicon concentrators and ran on Superdex 200 10/300 column (GE Healthcare) equillibrated with buffer D3 at 4 °C. The main-peak fractions were pooled, concentrated to 0.5-1.0 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C.

The samples were mixed with 4x gel loading dye and analyzed using SDS-PAGE on 4-20% gradient polyacrylamide gel ran at 300 V in TGS buffer. Heating up samples of membrane proteins before applying them on the gel is not recommended. Due to that they may degrade and become even more smeary. The gel was stained with the SimplyBlue stain. Additionally, the thermal stability was analyzed using Tycho NT.6 device in 10 µL capillaries.

7.3. Assays for binding and affinity evaluation

7.3.1. Binding assays with NB1829 and STING CBD mutants

7.3.1.1. Pull-down assay

TALON resin (Co^{2+}-NTA agarose) (Takara Bio) was used for the purpose of this assay since it was shown earlier that tagless CBD is able to unspecifically bind to Ni-NTA resin. 30 µg of CBD mutant 1, CBD mutant 2 and CBD WT were either used alone or mixed with NB1829 in 1:1 molar ratio (19 µg). Additional control was prepared with NB alone (30 µg). The proteins were incubated in 1.5 mL centrifugal tubes with 40 µL of TALON resin equillibrated with buffer B3 and incubated for 15’ on ice. The resin was spun down in benchtop centrifuge for 1’ at 2,000 rpm at 4 °C, and FT was discarded by gentle aspitation using a pipette. The resin was washed with 2 x 100 µL of buffers B5 and B6 using centrifugation as above. The last wash was collected. The proteins were eluted with 100 µL of buffer B7.

SDS-PAGE gel samples were prepared by mixing 30 µL of wash 4 and elution samples with 10 µL of 4x gel loading dye. The samples were heated up to 95 °C for 5' and ran on 4-20 % gradient polyacrylamide gel at 300 V in TGS buffer. The gel was stained with a SimplyBlue stain.

7.3.1.2. SEC binding test

100 µg of CBD mutant 1, CBD mutant 2 and CBD WT were either used alone or mixed with NB1829 in 1:1 molar ratio (63.5 µg). 100 µg of NB1829 alone were used as a control. The proteins were incubated on ice for 15' and injected onto Superdex 200 5/150 column (GE Healthcare) equillibrated with buffer B3. The peak fractions were analyzed on a SDS-PAGE gel (as above).

7.3.2. Binding assays for CD47, HuFab-H12-P and anti-Fab NB

In order to evaluate binding to CD47 of modified version of HuFab, called HuFab-H12-P, and anti-Fab nanobody (aF-NB) binding to this Fab, two small-scale pull-down assays in 1.5 mL centrifugal tubes have been done: The first on Strep-Tactin XT to check if the Fab in complex with the nanobody binds to CD47, and the second on CaptureSelect IgG-CH1 Affinity Matrix (Thermo Scientific), which is the affinity resin that binds CH1 of an antibody, to check NB binding to the Fab. All steps were done on ice unless overwise stated. All centrifugation steps were done in the benchtop centrifuge at 3,000 rpm at 4 °C for 1'. After each step, the buffer was gently removed from above the resin using a pipette.

7.3.2.1. Pull-down assay on Strep-Tactin XT resin

20 µg of CD47 were diluted to 80 µL total with buffer B1 and incubated for 10' with 40 µL of Strep-Tactin XT resin equillibrated with buffer B1. Then the several samples containing immobilized CD47, as above, were prepared: 1) Incubated with 2-fold molar excess of HuFab-H12-P over CD47 for 10', 2) Incubated with 2-fold molar excess of HuFab-H12-P over CD47 for 10' and then incubated with 2-fold molar excess of aF-NB over HuFab-H12-P for 10', 3) Incubated with 2-fold molar excess of aF-NB over CD47 for 10'. Two additional controls were also prepared: resin with Fab alone and with nanobody alone (amounts the same as in sample 1 and 3, respectively). The resin was then spun down and the FT was discarded. The resin was washed four times with 100 µL of buffer B1 using centrifugation. The wash 4 was collected. The proteins were eluted by 10' incubation with 80 µL of buffer B2 followed by centrifugation.

SDS-PAGE gel samples were prepared by mixing 30 µL of wash 4 or elution samples with 10 µL of 4x gel loading dye. The samples that did not contain CD47 FL were additionally heated up to 95 °C for 5'. All samples were ran on 4-20% gradient polyacrylamide gel at 300 V in TGS buffer. The gel was stained with a SimplyBlue stain.

7.3.2.2. Pull-down assay on CaptureSelect IgG-CH1 affinity matrix

20 µg of HuFab-H12-P were diluted to 80 µL total with buffer B3 and incubated for 10' with 40 µL of Strep-Tactin XT resin equillibrated with buffer B3. Then 2-fold molar excess of aF-NB was added and 10' incubation was performed. Additionally, the control sample was prepared containing the resin and the same amount of aF-NB alone as above. Afterwards, the resin was spun down, FT was discarded and four washes with 100 µL of buffer B3 were performed using centrifugation. The proteins were eluted using buffer B4.

SDS-PAGE gel samples were prepared by mixing 30 µL of wash 4 or elution samples with 10 µL of 4x gel loading dye. The samples were heated up to 95 °C for 5' and ran on 4-20% gradient polyacrylamide gel at 300 V in TGS buffer. The gel was stained with a SimplyBlue stain.

7.3.3. Enzyme-linked immunosorbent assay

7.3.3.1. Protein biotinylation

If for initial protein purification protocol tris-(hydroxymethyl)-aminomethane hydrochloride (TRIS-HCl) buffer was used, the proteins need to be desalted using, e.g. PD-10 desalting columns (Cytiva) in order to exchange this buffer into other buffer of choice. This is because TRIS-HCl inhibits chemical biotinylation process. In this case, STING CBD was in HEPES-containing buffer, so desalting was not needed. Firstly, 2 mg of EZ-Link NHS-PEG4-Biotin reagent (Thermo Scientific) were dissolved in 200 µL of 50% v/v dimethyl sulphoxide (DMSO) (Carl Roth) solution in ddH_2O. This solution has to be prepared freshly and cannot be stored. 1000 µg (0.0431 µmol) of STING CBD (stock concentration of 4.4 mg/mL) were thorougly mixed with buffer E1 and biotinylation reagent solution, added in this order, to final volume of 1000 µL (DMSO concentration cannot exceed 1.5% final, overwise the protein may precipitate). The biotinylation reagent was used in 10-fold molar excess over the protein. The reaction was performed for 2 h at 4 °C with gentle rotation. The biotinylation was quenched by addition of 1 M TRIS-HCl (Carl Roth) solution. TRIS was used in 7-fold molar excess over the biotinylation reagent. The biotinylation reagent excess was removed using SEC on Superdex 75 10/300 column (GE Healthcare) equillibrated with buffer E1 at 4 °C. The sample was concentrated using 10 k M.W.C.O. Amicon concentrator to 1.4 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C.

7.3.3.2. Biotinylation confirmation using pull-down assay on SAV-coated paramagnetic beads

The biotinylation was confirmed using a pull-down assay on streptavidin (SAV)-coated paramagnetic beads (further called “SAV beads”) (Roche Diagnostics). Liquid removal was done by immobilization of SAV beads using a magnet and pipetting. 100 µL of SAV beads suspension were transferred into 0.5 mL centrifugal tube, and washed three times with 100 µL of buffer E1. 2.5 µg of biotinylated STING CBD were diluted in 40 µL final of buffer E,

and incubated with drained SAV beads for 30’ at 18 °C with shaking at 600 rpm. The suspension was additionally mixed using gentle pipetting every 10’. It was done because shaking was insufficient with such a small volume. The FT was removed and collected, and the SAV beads were washed three times with 40 µL of buffer E1. The third wash was collected. The SAV beads were re-suspended in 40 µL of buffer E1. The elution was done using a temperature step during the gel sample preparation.

An additional input sample containing 2.5 µg of protein diluted in 40 µL final of buffer E1 was also prepared. 40 µL of each collected sample, including the sample containing re-suspended SAV beads with potential immobilized protein, were mixed with 15 µL of 4x gel loading dye and heated up to 95 °C for 5’. The sample containing SAV beads was briefly spun down and the liquid from above the beads was used for the gel sample. 10 µL of each gel sample were ran on SDS-PAGE 4-20% gel at 300 V in TGS buffer. The gel was stained with Coomassie Brilliant Blue stain. The additional gel was ran for western blot transfer using 5 µL of each gel sample per well. The samples were analyzed using western blot with 1:2,000 diluted anti-STING antibody (BioLegend), followed by 1:5,000 diluted anti-mouse IgG-HRP antibody (Jackson Immuno Research) (see Section 6.4.2.).

7.3.3.3. Enzyme-linked immunosorbent assay

Enzyme-linked immunosorbent assay (ELISA) was performed to evaluate NB1829 binding to STING CBD with or without 2’,3’-cGAMP. Only for the purpose of ELISA, the original nanobody with both HA- and his-tag was used (further called “NB1829-HA-his”). It was purified in the same way as other nanobodies described here. The high-binding, 96-well, clear, flat-bottom plates (Corning) were coated with NeutrAvidin (NAV) (Thermo Scientific) using 50 µL/well of 2 µg/mL NAV solution in buffer E2. The plates were incubated overnight at 4 °C. On the following day, the NAV solution was dropped, and the plate was washed three times with 100 µL/well of buffer E3 containing biotin-free bovine serum albumin (BSA) (Carl Roth). Then blocking was performed by incubation with 100 µL/well of buffer E3, for at least 1 h at RT with shaking at 600 rpm (or overnight at 4 °C, without shaking). The plates prepared this way can be stored in a sealed plastic bag at 4 °C for up to two weeks. The biotinylated STING CBD was diluted to 600 nM in buffer E3. Part of it was also supplemented with 2-fold molar excess of 2’,3’-cGAMP (Biolog). 2’,3’-cGAMP was earlier prepared in form of 1 mM stock in ddH_2O and stored at -20 °C. The 10’ incubation on ice was performed. The protein with or without ligand was immobilized on NAV by applying 100 µL of it per well and incubating for 1 h at RT with shaking at 600 rpm. Additional control was done with 600 nM NB1829-HA-his alone diluted in buffer E3. Then the plate was washed with 3 x 100 µL/well of buffer E3. Then the plate was incubated with 100 µL/well of 1,200 nM NB1829-HA-his diluted in buffer E3 for 1 h at RT with shaking at 600 rpm. Then the plate was washed as above. This was followed by incubation with 100 µL/well of the secondary anti-HA-horse radish peroxidase (HRP) antibody (Cell Signaling Technology) diluted 1:1,000 in buffer E3 for 1 h at RT with shaking at 600 rpm. The plates were washed as above. Then the plate was incubated with 50 µL/well of 3,3’,5,5’-

tetramethylbenzidine (TMB) soluble reagent (ScyTek Laboratories) 1:5 diluted in ddH_2O for 5'-10'. The reaction was quenched by addition of equal volume of 1 M HCl. The plate was incubated for at least 1' at RT with shaking at 600 rpm. The read-out was performed at 450 nm using Infinite M1000 microplate reader (Tecan).

7.3.4. Surface plasmon resonance

The surface plasmon resonans (SPR) experiments were done using Biacore X100+ (Cytiva) on Ni-NTA chip (Cytiva). The running buffer was HBS-P+ (Cytiva). The protein samples were diluted in the same buffer. 10 nM his-tagged NB1829 was immobilized on the chip in flow channel 2 (FC2) by 110" injection. It was followed by 180" injection of STING CBD at different concentration in each cycle (31.25 nM, 15.625 nM, 7.8125 nM and 3.9 nM) into both flow channels (FC1 and FC2). Each analyte injection was followed by 400" dissociation phase. In the end of each cycle, the chip was regenerated using standard EDTA, buffer wash, $NiCl_2$, buffer wash cycle. Biacore Evaluation software was used to evaluate the data with 1:1 binding model. Read-out from FC1 was used as a reference and its signal was subtracted from FC2 signal. The SPR measurement and evaluation were performed by Gregor Witte.

7.3.5. Western blotting

For western blotting (WB), protein samples were ran on standard 4-20 % SDS-PAGE polyacrylamide gel at 300 V in TGS buffer. The proteins were transferred on polyvinylidene fluoride (PVDF) membrane using commercial Trans-Blot Turbo Mini 0.2 µm PVDF Transfer Pack (Bio-Rad) and Trans-Blot Turbo Transfer System (Bio-Rad) according to manufacturer's guidelines, using standard transfer program running at up to 2 A and 25 V for 30'. After the transfer, the membrane was blocked using 3% milk (powdered milk, blotting grade) (Carl Roth) in phosphate buffer saline-tween 20 (PBST) buffering solution for minimum of 1 h at RT or overnight at 4 °C on nutator. Then the membrane was incubated with primary antibody diluted in 3% milk in PBST buffering solution for 1 h at 4 °C. Then three 10' washes in larger volumes of PBST buffering solution at RT, on nutator were done. The membrane was incubated with HRP-conjugated secondary antibody followed by three washes in the same conditions as above. For primary antibodies conjugated with HRP, a single staining step was used. In the end, the membrane was incubated for at least 1' in WB staining solution with freshly added H_2O_2 at RT, and the signal was read using Amersham Imager 600 (GE Healthcare) in bioluminescence mode.

7.4. Determination of protein concentration

The concentrations of all proteins were measured using nanodrop (Implen) at 280 nm and calculated using the extinction coefficients and molecular weights calculated using ProtParam tool (https://web.expasy.org/protparam/) according to universal guidelines described by Pace at al. (Pace et al., 1995).

7.5. X-ray crystallography

7.5.1. High-throughput crystallization screening

The high-throughput screening (HTS) of crystallization conditions was performed by the Max Planck Institute of Biochemistry (MPIB) Crystallization Facility (Martinsried, Germany) using vapour diffusion method. 100 nL + 100 nL sitting drops with 100 µL reservoir were set up on MRC 2, 96-well plates (Jena Bioscience) using Phoenix crystallization robot. The crystallization process was performed at 20 °C. A wide variety of commercial crystallization screens were used, including: JCSG-plus (Molecular Dimensions), PACT premier (Molecular Dimensions), Index (Hampton Research), Wizard Classic 1, 2, 3 and 4 (Rigaku). The crystallization complex was prepared by mixing STING CBD with 3-fold molar excess of NB1829 and incubation for 45' with gentle rotation at 4 °C. The complex was concentrated using 50 kDa M.W.C.O. Amicon concentrators and ran on Superdex 200 10/300 column equillibrated with buffer X1 at 4 °C. The main peak fractions were pooled and concentrated in the same way to 9.2 mg/mL. The sample was hard-spun in the benchtop centrifuge (Eppendorf) at 21,130 rcf for 5' at 4 °C before its application on the crystallization plates. The final sample was also analyzed using SDS-PAGE (as in the sections above). The crystals were photographed using M165 C light microscope (Leica).

7.5.2. Optimization of crystals

The crystallization complex was prepared in the same way as for the HTS, but in a larger scale. The complex was ran on Superdex 200 16/600 column equillibrated with buffer X1 at 4 °C. The main peak fractions were pooled and concentrated to 9 and 12 mg/mL. The optimization was performed using vapour diffusion method in a hanging drop format (1 µL + 1 µL droplets) in the Crystalgen SuperClear 24-well plates (Jena Bioscience) with 500 µL reservoir. The crystals from the original screen - Index (Hampton Research) position G10, containing 25% PEG 3350, 0.2 M magnesium chloride hexahydrate and 0.1 M BIS-TRIS, pH 5.5, were optimized to 20% PEG 3350, 0.25 M magnesium chloride hexahydrate, 0.1 M BIS-TRIS pH 5.5. The optimization temperature was 20 °C. The high-diffraction quality crystals were harvested after 20 days. The crystals were transferred into a droplet of the mother liquer, freshly mixed with 20% final 2,3-butanediol (Sigma-Aldrich), for the purpose of cryoprotection, and immediately frozen and stored in liquid nitrogen. The crystals were photographed using M165 C light microscope.

7.5.3. Data collection and processing

The data was collected at EMBL P14 beamline at PETRA III in Deutsches Elektronen-Synchrotron (DESY) (German Electron Synchrotron) located in Hamburg, Germany, using 0.9762 Å wavelength. The data indexing and integration was performed using XDS (Kabsch, 2010). The scaling was performed using XSCALE (Kabsch, 2010). The spacegroup determination was made using POINTLESS (Evans, 2006).

7.5.4. Determination of the structure of STING CBD and NB1829 complex

The STING CBD and NB1829 complex structure was determined by molecular replacement using Phaser within Python-based Hierarchical ENvironment for Integrated Xtallography (Phenix) (McCoy et al., 2007). The search models for molecular replacement were selected from known structures deposited in PDB. Firstly, all the flexible loops were manually removed from the structure of apo human STING CTD (PDB accession code 4F5W). The human single-domain antibody (PDB accession code 3ZHK) was selected as a template for the nanobody. 3ZHK structure was converted to poly-alanine model using CHAINSAW (Stein, 2008). The two copies of each element of the complex, prepared as above, were used to solve the structure of STING CBD-NB1829 complex. The initial model was applied to AutoBuild within Phenix (McCoy et al., 2007). The structure was refined multiple times using manual model building in COOT (Emsley et al., 2010) and Phenix.refine within Phenix (Adams et al., 2010). The Ramachandran plot was generated in PROCHECK (Laskowski et al., 1993). The structure was visualized in PyMOL (Schrödinger, LLC). The structure was solved by Gregor Witte, and further refined by both of us.

7.6. Reconstitution of nanodiscs

7.6.1. DDM/CHS stock preparation

Firstly, 2 g of n-dodecyl-β-D-maltopyranoside (DDM) (Anatrace) were dissolved in 18 mL of ddH_2O in 50 mL falcon tube. Then 0.2 g of cholesteryl hemisuccinate tris salt (CHS) (Anatrace) was added. Afterwards, the volume was adjusted to 20 mL with ddH_2O, and the solution was sonicated using a microtip and constant sonication mode until very warm but not hot. Then the solution was cooled down until reaching RT, and the above procedure was repeated until the solution was totally clear. After addition of CHS, the solution should not be mixed, because undissolved CHS will stick to the walls of the tube or to the mixing tool, and it will not be efficiently dissolved. Then the solution was filtered using 0.45 µm syringe filter (Millipore), aliquoted, frozen and stored at -20 °C. The final concentrations were 10% DDM and 1% CHS.

7.6.2. Lipid-detergent mixed micelles stock preparation

Firstly, 310 µL of 25 mg/mL solution of polar soybean lipid extract in chloroform (Avanti Polar Lipids) were transferred to a small glass test tube with round bottom using 500 µL glass Hamilton syringe (Merck) with a blunt stainless steel needle. The lipids were thoroughly dryed under a nitrogen stream by constant rotation of the tube in near horizontal position to form as thin lipid film as possible (the tube should be made of glass since chloroform dissolves plastic; for storage of unused chloroform stocks at -20 °C, glass vials with teflon coated caps should be used, and additionally they should be closed with parafilm to restrict evaporation). Then the lipid film was desiccated under vacuum for at least 4 h to remove whole residual chloroform. Afterwards, 400 µL of ddH_2O were added and the tube was

vortexed until obtaining homogenous, milky solution without visible debris. The tube should be vortexed in almost vertical position, so no lipid debris sticks to the walls above the solution. Then 600 µL of DDM/CHS stock (10% DDM / 1% CHS) were added to form lipid-detergent mixed micelles soluble in water. Then the solution was gently but thoroughly mixed by pipetting up and down. Producing foam should be avoided during this process. Afterwards, the tube was closed using parafilm and the solution was incubated with stirring at RT using a tiny magnetic stirring bar until completely clear (usually 2 h; vortex should not be used to not foam the detergent). In the end, the solution was aliquoted into 1.5 mL centrifugal tubes and hard-spun in the benchtop centrifuge (Eppendorf) at 21,130 rcf at RT for 5' to remove residual non-solubilized lipids that may form a tiny white pellet. The pellet should be inexistent or present in trace quantities if lipids were solubilized properly, overwise the time of solubilization should be extended, or concentration of detergent should be increased. The lipid stock was then transferred to a new tube, aliquoted, frozen and stored at -80 °C. The detergent-lipid mixed micelles stock had the following final concentrations: lipids - 10 mM (7.75 mg/mL; assuming average mass of a lipid of 775 g/mol), DDM - 117.5 mM (6%) and CHS - 9.87 mM (0.6%). This type of lipid-detergent mixed micelles was used for all nanodisc preparations described in this thesis.

The classical method of lipid preparation uses lower percentage of detergent solution and requires sonication. However, sonication of this solution is problematic because of foam formation and frequent damaging of small tubes by vibrations. Additionally, in this method, foaming and excessive heating can cause lipid oxidation. After sonication, I also noticed unexplained black pellets after hard spins of mixed micelles. The method with solubilization of lipids in high percentage detergent solution is easier, more gentle and provides the same or even better results.

7.6.3. Bio-Beads activation

5 g of Bio-Beads SM-2 Resin (Bio-Beads) (Bio-Rad) were mixed with 100% methanol and incubated for 30' at RT on rotator. Then the methanol was decanted and the Bio-Beads were washed at least 15 times with 50 mL of ddH_2O by inversion and decantation until methanol was fully removed. Then the Bio-Beads were supplemented with 10 mL of ddH_2O and stored at 4 °C.

7.6.4. Empty nanodiscs - general protocol

The lipid-detergent mixed micelles were mixed with buffer S1 (to adjust final detergent concentration) and MSP, and incubated on ice for 10-30'. Then the solution was moved to a new tube containing drained, activated Bio-Beads in the amount of 0.5 g of drained Bio-Beads per 1 mL of solution. The Bio-Beads can be transferred to a new tube using a pipette with a cut 1 mL pipette tip and drained using 200 µL pipette. Its tip should be pressed against the tube wall during the insertion, to not accidentally aspirate the Bio-Beads that can clog it. The reconstitution mixture was incubated with Bio-Beads overnight at 4 °C with gentle,

horizontal rotation to not produce foam, and slowly remove the detergent that enables spontaneous nanodisc assembly. On the following day, the sample was separated from Bio-Beads using 200 µL pipette and additionally filtered using 0.5 mL, 0.45 µm centrifugal filter (Millipore) according to manufacturer's guidelines. Then the nanodiscs were ran onto Superdex 200 10/300 column equillibrated with buffer S1.

7.6.5. Empty nanodiscs - optimization

MSP1D1 was used without his-tag and the rest of MSPs was his-tagged. All nanodiscs have been done according to the general protocol (see the previous section).

The empty MSP1D1 nanodiscs (e1D1-NDs) samples for optimization were prepared in buffer S1 with the following MSP:lipids molar ratios: 1:20, 1:30, 1:40 and 1:50 with the final detergent concentration of 20 mM, and final lipids concentration of 1.7 mM. 88.2 µg of MSP were used for each sample.

The empty MSP1E3D1 nanodiscs (e1E3D1-NDs) samples for optimization were prepared in buffer S1 with the following MSP:lipids molar ratios: 1:20, 1:35, 1:40, 1:45, 1:65, 1:75 and 1:85 with the final detergent concentration of 10 mM and final lipids concentration of 0.9 mM. 130.9 µg of MSP were used for each sample.

The empty MSP1D1dH5 nanodiscs (edH5-NDs) samples for optimization were prepared in buffer S1 with the following MSP:lipids molar ratios: 1:10, 1:15, 1:20 and 1:25 with the final detergent concentration of 5 mM and final lipids concentration of 0.4 mM. 85.9 µg of MSP were used for each sample.

The main peak fractions were analyzed using negative stain electron microscopy (see Section 7.7.1.). The fractions from e1D1-NDs and e1E3D1-NDs were non-diluted and from edH5-NDs were diluted 1:5. The Quantifoil R3/3 + 2 nm carbon, Cu300 (Quantifoil) were used for e1D1-NDs and Quantifoil R2/1 +2 nm C, Cu200 (Quantifoil) were used for the other samples.

7.6.6. Membrane protein incorporation into nanodiscs - general protocol

The lipid-detergent mixed micelles prepared from polar soybean lipid extract solubilized in 6% DDM/0.6% CHS were mixed with buffer S1 or S2 (to adjust final detergent concentration before addition of membrane protein) and CD47. The mixture was incubated on ice for 15'. Then MSP was added and the reconstitution mix was incubated on ice for 10'. Afterwards, the solution was moved to the new tube containing drained, activated Bio-Beads in the amount of 0.5 g of drained Bio-Beads per 1 mL of solution. The reconstitution mixture was incubated with Bio-Beads overnight at 4 °C with gentle, horizontal rotation to not produce foam, and slowly remove the detergent that enables spontaneous nanodisc assembly. On the following day, the sample was separated from Bio-Beads using 200 µL pipette and

additionally filtered using 0.5 mL, 0.45 µm centrifugal filter according to manufacturer's guidelines. The sample was incubated with 200-400 µL of Strep-Tactin XT equillibrated in buffer S1 or S2 for 20-30' with gentle, horizontal rotation at 4 °C in 2 mL centrifugal tube. For nanodiscs, the standard Strep-Tactin XT should be used. The high-capacity version of Strep-Tactin XT significantly decreases the yield for unknown reasons. Then the resin was loaded onto a small gravity column and the FT was re-loaded three times. The resin was washed two times with 1 CV of buffer S1 or S2 and transferred into 2 mL centrifugal tube. The Fab was added and another incubation for 20-30' as above was performed. Then the resin was loaded again onto a gravity column and the FT was re-loaded three times. The resin was washed three times with 2 CVs of buffer S1 or S2. The elution was performed with 6 CVs of buffer S1 or S2, containing additionally 50 mM D-biotin, applied in three equal portions. Afterwards, the sample was concentrated in 30 kDa M.W.C.O. Amicon concentrator (Millipore) and ran onto Superose 6 10/300 column (GE Healthcare) equillibrated with buffer S1 or S2.

7.6.7. CD47 in complex with Fab10 in MSP1D1 nanodiscs

Fab10 is derrived from a custom-made antibody. The nanodiscs were prepared according to the general protocol (see Section 7.6.6.) in buffer S1 with slight modifications. Two 1 h 30' incubations with equal amount of fresh, activated Bio-Beads were used instead of overnight incubation. His-tagged MSP1D1 was used. The target:MSP:lipids molar ratio was 1:10:200 with the final detergent concentration of 18 mM and final lipids concentration of 1.4 mM. 301 µg of target protein and 1819 µg of MSP were used. 2.4-fold molar excess of Fab10 over CD47 was added to the sample. The sample was purified using 400 µL of Strep-Tactin XT.

The samples were prepared in two concentrations: 0.36 mg/mL and 0.15 mg/mL. Part of them was supplemented with 0.05% n-octyl-β-D-glucopyranoside (β-OG) (Anatrace) just before plunging (4 µL of the sample were mixed with 0.5 µL of 0.5% β-OG in ddH_2O). Firstly, the Quantifoil Cu200, R2/1 grids (Quantifoil) were glow discharged for 7". Then 4.5 µL of each sample were applied on the grids, the blotting was performed for 2.0-2.5", and the grids were plunge-frozen in liquid ethane using Leica plunger (Leica). The cryo-EM data was collected on 300 kV Titan Krios G3 transmission electron microscope (Thermo Scientific) with Gatan K2 camera.

All electron microscopy experiments in this thesis were done with help from Joseph Bartho and Katja Lammens.

7.6.8. CD47 in complex with HuFab-D in MSP1E3D1 nanodiscs

The nanodiscs were prepared according to the general protocol (see Section 7.6.6.) in buffer S1. MSP1E3D1 with his-tag was used. The target:MSP:lipids molar ratio was 1:10:300 with the final detergent concentration of 20 mM and final lipids concentration of 1.6 mM. 291 µg

of target protein and 2616.8 µg of MSP were used. 2.7-fold molar excess of HuFab-D over CD47 was used. The sample was purified using 400 µL of Strep-Tactin XT.

The sample was additionally analyzed using western blotting with anti-CD47 antibody (B6H12; Invitrogen) diluted 1:2,000, followed by anti-mouse IgG-HRP antibody (Jackson ImmunoResearch) diluted 1:5,000 and with anti-Fab-HRP antibody (Jackson ImmunoResearch) diluted 1:5,000 (see Sections 6.4.2. and 7.3.5.).

Before plunging the sample was analyzed using negative stain electron microscopy (see Section 7.7.1.). Grids Quantifoil R 2/1, +2 nm C, Cu200 were used. 0.15 mg/mL nanodisc sample was diluted 1:50 in buffer S1.

For cryo-EM, the samples were prepared in the following concentrations: 0.15 mg/mL, 0.075 mg/mL, 0.0375 mg/mL and 0.01875 mg/mL. Firstly, the Quantifoil R2/1, Cu200 grids were glow discharged for 7". Then 4.5 µL of each sample were applied on the grids, the blotting was performed for 2.0-2.5", and the grids were plunge-frozen in liquid ethane using Leica plunge freezer. The cryo-EM data was collected on 300 kV Titan Krios G3 transmission electron microscope with Gatan K2 camera.

7.6.9. CD47 in complex with HuFab in MSP1D1dH5 nanodiscs

The nanodiscs were prepared according to the general protocol (see Section 7.6.6.) in buffer S1. MSP1D1dH5 with his-tag was used. The target:MSP:lipids molar ratio was 1:10:100 with the final detergent concentration of 15 mM and final lipids concentration of 1.0 mM. 250 µg of target protein and 1481 µg of MSP were used. 2.5-fold molar excess of HuFab over CD47 was used. 300 µL of Strep-Tactin XT were used for purification.

The sample was additionally analyzed using western blotting with anti-CD47 antibody diluted 1:4,000, followed by anti-mouse IgG-HRP antibody diluted 1:5,000 and with anti-Fab-HRP antibody diluted 1:4,000 (see Sections 6.4.2. and 7.3.5.).

For cryo-EM, the samples were prepared in the following concentrations: 0.52 mg/mL, 0.4 mg/mL and 0.18 mg/mL. Firstly, the Quantifoil R2/1, Cu200 grids were glow discharged for 30". Then 4.5 µL of each sample were applied on the grids, the blotting was performed for 2.0-2.5", and the grids were plunge-frozen in liquid ethane using Vitrobot plunge freezer (Thermo Scientific). Part of the samples was supplemented with 0.1%, 0.05% or 0.025% fluorinated octyl maltoside (FOM) (Anatrace) just before plunging. The cryo-EM data was collected on 300 kV Titan Krios G3 transmission electron microscope with Gatan K2 camera.

7.6.10. CD47 in complex with HuFab-H12-P and anti-Fab NB in MSP1D1dH5 nanodiscs

The nanodiscs were prepared according to the general protocol (see Section 7.6.6.) in buffer S2. After incubation of the sample with Fab, another 15' incubation with anti-Fab NB was performed. MSP1D1dH5 with his-tag was used. The target:MSP:lipids molar ratio was 1:10:100 with the final detergent concentration of 15 mM and final lipids concentration of 1.1 mM. 339 µg of target protein and 2104 µg of MSP were used. 2.5-fold molar excess of HuFab-H12-P over CD47, and 4.5-fold molar excess of anti-Fab NB over Fab were used. Two 15' incubations were performed with gentle rotation at 4 °C - the first after addition of Fab and the second after addition of NB. 400 µL of Strep-Tactin XT were used for purification.

The sample was additionally analyzed on SDS-PAGE gel stained with SimplyBlue stain and using western blotting with anti-strep-tag-HRP antibody (Millipore) diluted 1:4,000 and with anti-his-tag-HRP antibody (Miltenyi Biotec) diluted 1:5,000 (see Sections 6.4.2. and 7.3.5.).

Before plunging the sample was analyzed using negative stain electron microscopy (see Section 7.7.1.). Grids Quantifoil R 2/1, +2 nm C, Cu200 were used. 0.49 mg/mL nanodisc sample was diluted 1:80 in buffer S2.

For cryo-EM, the samples were prepared in the following concentrations: 0.49 mg/mL, 0.3 mg/mL, 0.2 mg/mL and 0.1 mg/mL. Firstly, the Quantifoil R2/1, Cu200 grids were glow discharged for 7". Then 4.5 µL of each sample were applied on the grids, the blotting was performed for 2.0-2.5", and the grids were plunge-frozen in liquid ethane using Leica plunger. The cryo-EM data was collected on 300 kV Titan Krios G3 transmission electron microscope with Gatan K2 camera.

7.6.11. CD47 in complex with HuFab in DDM/CHS micelles

Frozen CD47 was thawed on ice and hard-spun in the benchtop centrifuge (Eppendorf) at 21,130 rcf at 4 °C for 5'. Two samples were prepared by mixing either 4 µL of 1.14 mg/mL CD47 and 1 µL of 6.48 mg/mL HuFab, or 4 µL of 0.32 mg/mL CD47 and 1 µL of 1.85 mg/mL HuFab (both diluted with buffer S2). In both samples, CD47:HuFab molar ratio was 1:1. The samples were incubated for 1 h on ice.

For cryo-EM, the Quantifoil R2/1, Cu200 grids were glow discharged for 7". 4.5 µL of each sample were applied on the grids, the blotting was performed for 2.5", and the grids were plunge-frozen in liquid ethane using Leica plunger. The cryo-EM data was collected on 300 kV Titan Krios G3 transmission electron microscope with Gatan K2 camera.

7.7. Electron microscopy

7.7.1. Negative stain electron microscopy

Firstly, the grids were glow discharged for 20". Then 3.5 µL of the protein sample were applied on the grid and incubated for 30". Afterwards, the grid was briefly flashed with ddH_2O, the excess liquid was dropped onto a filter paper, and 3.5 µL of uranyl acetate solution were applied two times with 30" incubation. After the second incubation the excess solution was blotted away using a filter paper, and the grid was additionally dried in air for at least 2'. The grids were observed using 100 kV transmission electron microscope Morgagni (FEI) at 56,000x magnification.

7.7.2. Cryoelectron microscopy data processing

The Cryo-EM data was processed using Cryo-EM Single Particle Ab-Initio Reconstruction and Classification (CryoSPARC) software (Punjani et al., 2017). The models were visualized in UCSF Chimera (Pettersen et al., 2004).

8. Results for stimulator of interferon genes

8.1. Expression and purification of proteins

8.1.1. Expression and purification of NB1829

Nanobody 1829 was expressed in periplasm of *E. coli* WK6 cells using pelB periplasmic signal sequence in order to enable a proper formation of disulfide bonds. It was purified from the periplasm using two-step osmotic shock, Ni-NTA chromatography and size exclusion chromatography using Superdex 75 16/600 column. The results are shown in Figure 10.

NB1829 migrated on SDS-PAGE gel at approximately 15 kDa (lanes E and SE in Figure 10A) and eluted from SEC column at approximately 100 mL. The main peak was preceded by a small aggregation peak, and followed by an additional small peak, possibly containing degradation products (Figure 10B). The approximate yield was 15 mg/L. The melting temperature was 74.3 °C (Figure 10C).

The protein concentration was calculated using the following values: ε = 34,505 $M^{-1}cm^{-1}$, M.W. = 14,814.31 Da.

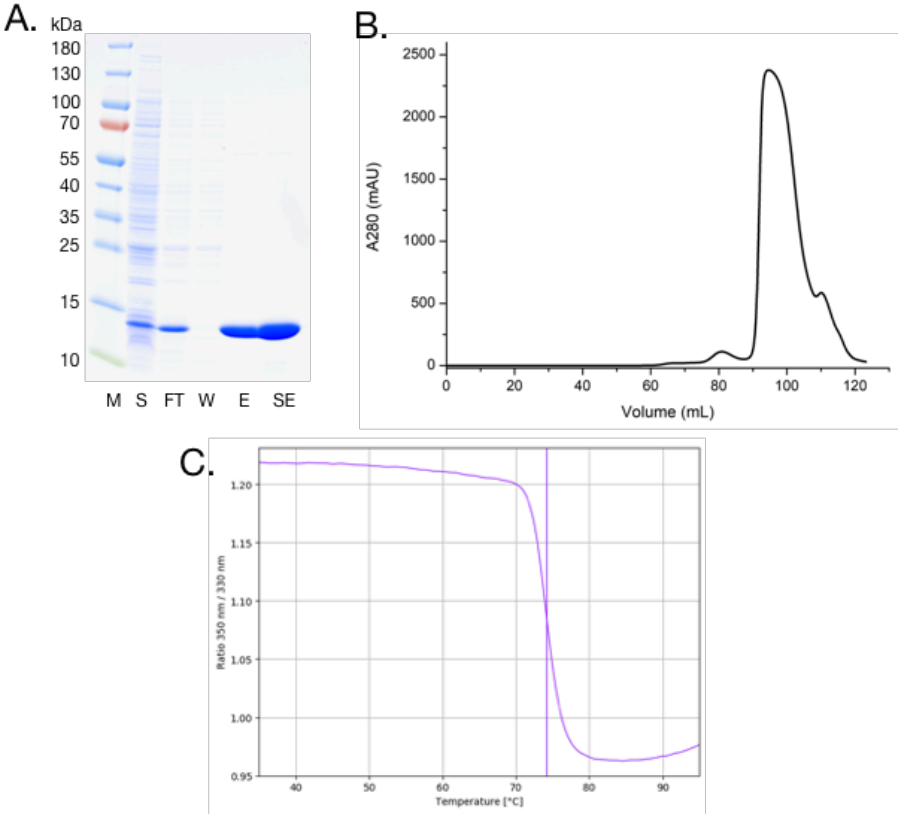

Figure 10. Purification of NB1829 (previous page).
A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, S - supernatant after centrifugation, FT - flow-through, W - final wash, E - elution, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superdex 75 16/600 column. C. Melting temperature curve from Tycho NT.6 shown as a ratio of measurements at 350 nm and 330 nm.

8.1.2. Expression and purification of STING CBD and its mutants

STING CBDs were expressed in *E. coli* Rosetta 2(DE3) and purified using Ni-NTA affinity chromatography. The his-tagged SUMO1 protein was cleaved using SENP2 protease and removed using subtractive IMAC on Ni-NTA resin. The protein was additionally purified using size-exclusion chromatography on Superdex 200 16/600 column. The results are shown in Figures 11 and 12.

The cleaved STING CBD migrated on the gel at approximately 23 kDa with slight degradation in form of lower molecular weight bands (lanes sF and SE in Figure 11A). It eluted from SEC column as a single monodisperse peak at approximately 90 mL (Figure 11B). The approximate yield was 30 mg/L. The melting temperature was 52.6 °C (Figure 11C).

The two mutants of STING CBD: CBD mutant 1 - D85K/L86E (the modification at the top of the butterfly-shaped domain) and CBD mutant 2 - E190R/R193A (the modification at the bottom) were purified in the same way as the wild-type protein. Both cleaved mutants migrated on the gel at approximately 23 kDa (lanes sF and SE in Figure 12 A and C). The CBD mutant 1 eluted from SEC as multiple peaks: the first with an additional aggregation shoulder at approximately 60 mL and the second at approximately 90 mL, followed by several small possibly degradation peaks (Figure 12B). The analysis on SDS-PAGE gel have shown that the second peak contained the majority of protein (data not shown) and this peak corresponded by elution volume to WT peak. This is why its fractions were collected and used for assays. The first peak possibly contained heavily aggregated protein. The CBD mutant 2 eluted from the SEC column as a single, monodisperse peak at approximately 90 mL and it was preceded by a tiny aggregation peak (Figure 12D). The yields were 1 mg/L and 24 mg/L for mutant 1 and 2, respectively. The mutant 1 had melting temperature of 48.5 °C, and the mutant 2 had two melting temperatures of: 45.3 °C and 51.4 °C (data not shown).

The concentrations were calculated using the following values: STING CBD (ε = 22,140 $M^{-1}cm^{-1}$; M.W. = 23,404.36 Da), CBD mutant 1 (ε = 22,140 $M^{-1}cm^{-1}$; M.W. = 23,418.38 Da), CBD mutant 2 (ε = 22,140 $M^{-1}cm^{-1}$; M.W. = 23,346.32 Da).

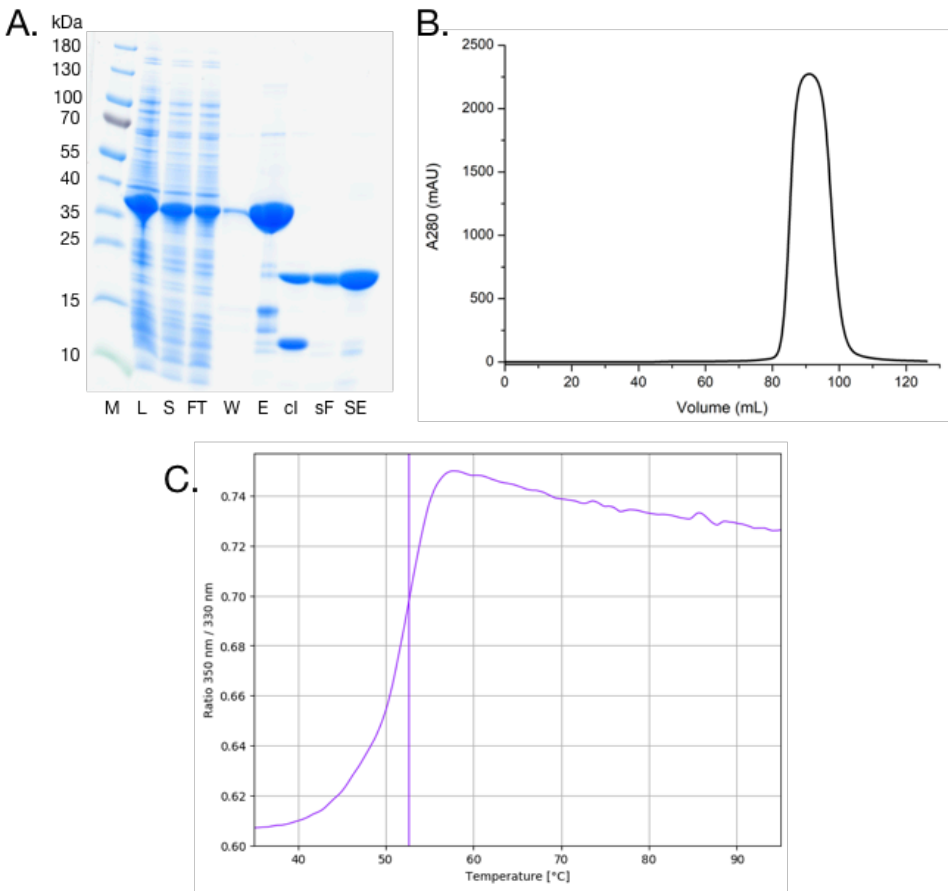

Figure 11. Purification of STING CBD.
A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, L - lysate, S - supernatant after centrifugation, FT - flow-through, W - final wash, E - elution, sI - subtractive IMAC input, sF - subtractive IMAC flow-through, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superdex 200 16/600 column. C. Melting temperature measurement curve from Tycho NT.6 shown as ratio of measurements at 350 nm and 330 nm.

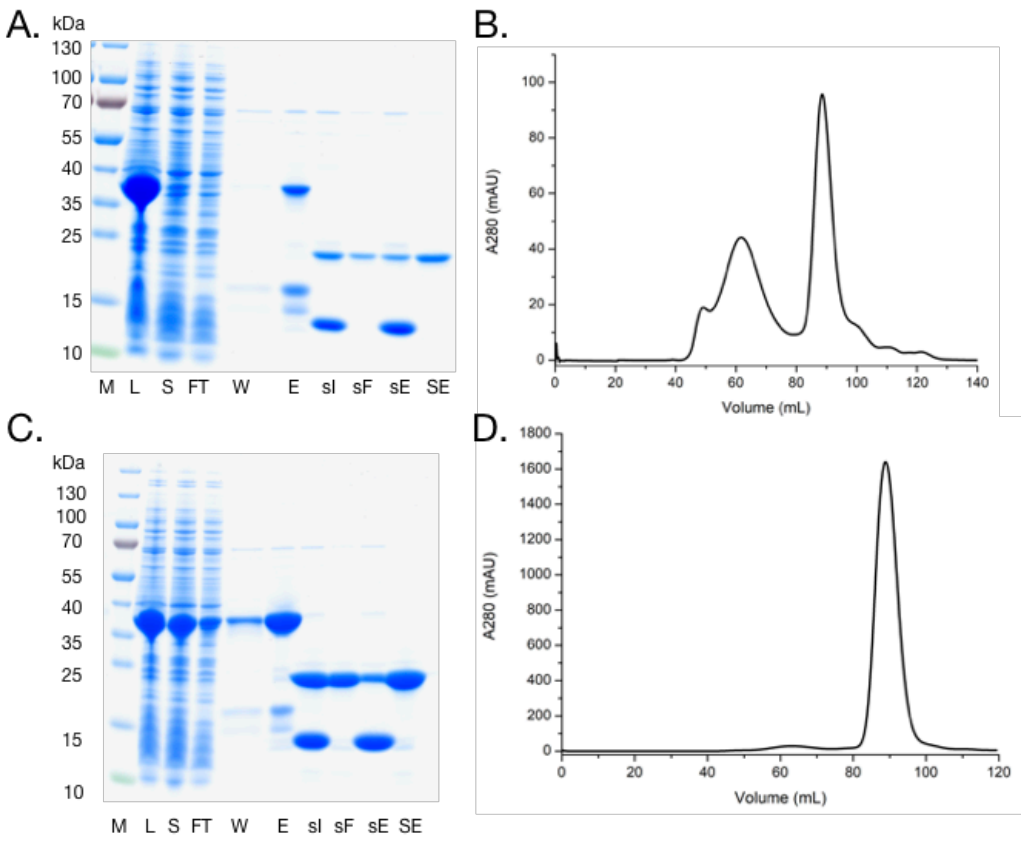

Figure 12. Purification of STING CBD mutants.
A. and C. SDS-PAGE gels with protein purification samples stained with SimplyBlue stain for CBD mutant 1 and mutant 2, respectively. M - molecular weight marker, L - lysate, S - supernatant after centrifugation, FT - flow-through, W - final wash, E - elution, sI - subtractive IMAC input, sF - subtractive IMAC flow-through, sE - subtractive IMAC elution, SE - SEC. B. and D. Size-exclusion chromatography protein elution profiles from Superdex 75 16/600 column for CBD mutants 1 and 2, respectively.

8.1.3. Expression and purification of full-length STING

STING FL was expressed in HEK293 Expi cells transfected using PEI reagent. The protein was extracted from membranes using a highly concentrated mild detergent DDM supplemented with CHS. The purification was performed using affinity chromatography on Strep-Tactin XT high-capacity resin (IBA Lifesciences) and size-exclusion chromatography on Superose 6 10/300 column. The purified protein was analyzed on SDS-PAGE gel. The results are shown in Figure 13.

STING FL migrated on the gel as a single band at approximately 43 kDa (lane SE in Figure 13A) and eluted from SEC column as a single peak at approximately 16 mL, followed by a tiny additional peak, containing possibly degradation products or other impurities (Figure 13B). The approximate yield was 0.09 mg/L.

The concentrations were calculated using the following values: Ɛ = 117,910 $M^{-1}cm^{-1}$, M.W. = 90,765.9 Da.

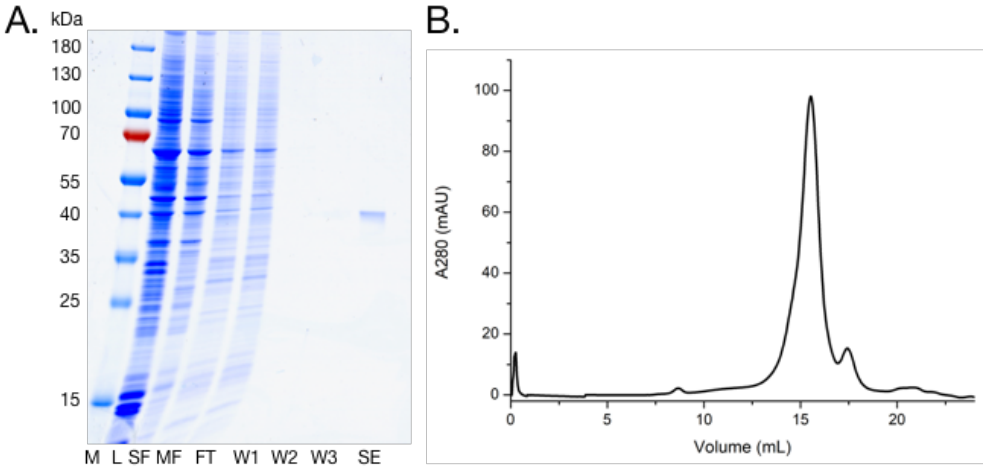

Figure 13. Purification of full-length STING.
A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, L - lysate, SF - soluble fraction after first ultracentrifugation, MF - solubilized membrane fraction after second ultracentrifugation, FT - flow-through, W1-W3 - washes 1-3, E - elution, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superose 6 10/300 column.

8.2. Evaluation of binding and binding affinity of proteins

8.2.1. Enzyme-linked immunosorbent assay

8.2.1.1. Biotinylation evaluation

STING CBD biotinylation was confirmed using pull-down assay on SAV-coated paramagnetic beads. The results were analyzed on SDS-PAGE gel (Figure 14A). Input fraction contained the same amount of protein as the sample used for the assay. All of the protein bound to the SAV beads during the experiment, since no band was observed in the FT fraction. The potential sample excess was fully removed by washing because no band was found in final wash fraction (W). During the washes approximately 60% of the sample was removed. This sample was probably weakly or unspecifically bound to the beads and is considered not biotinylated. By comparing the bands corresponding to STING CBD (marked with black triangle in figure 14A) that migrated on the gel at 20 kDa in input and elution fractions, the amount of successfully biotinylated protein was estimated to have been approximately 40%. The ladder-like pattern in the elution fraction contained most possibly streptavidin (approximate M.W. of a monomer is 16 kDa), released from the paramagnetic beads during heating step, and its higher order oligomers. The presence of STING CBD was additionally confirmed, using western blotting with anti-STING antibody, followed by anti-mouse IgG-HRP antibody, as two bands at approximately 30 kDa (Figure 14B).

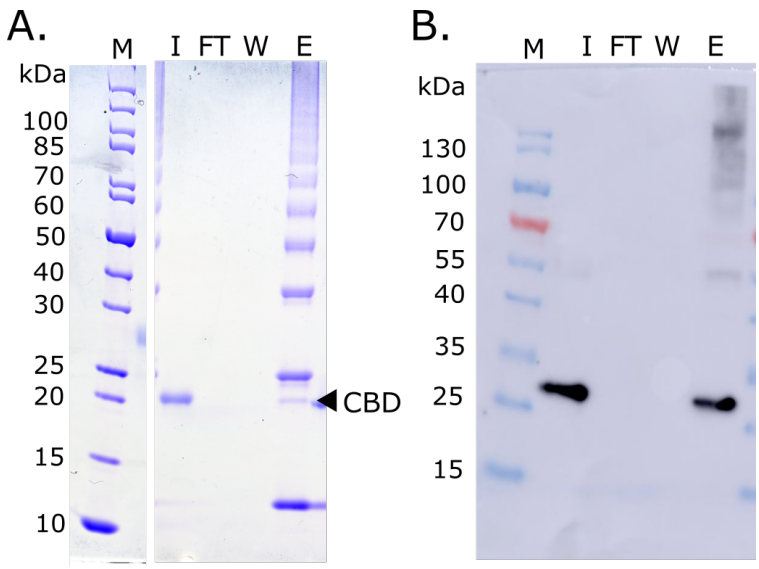

Figure 14. Evaluation of protein biotinylation.
The biotinylation of STING CBD was confirmed using pull-down assay on SAV-beads. M - marker, I - input, FT - flow-through, W - final wash, E - elution. A. SDS-PAGE gel with samples from pull-down assay. The bands corresponding to STING CBD were marked with a black triangle. B. Western blot from pull-down assay using anti-STING antibody followed by anti-mouse IgG-HRP antibody.

8.2.1.2. Enzyme-linked immunosorbent assay

ELISA was done on NAV-coated plates using biotinylated STING CBD with and without its ligand, 2',3'-cGAMP. The NB1829 with HA-6his-tag was used as a primary antibody, and anti-HA-tag-HRP antibody was used as a secondary antibody. The signal was developed using HRP substrate reagent and read at 450 nm. There were two repetitions of each sample. Their signal was averaged and shown as a column plot. The error bars reflect standard deviation (Figure 15).

ELISA confirmed NB1829 binding to STING CBD both without and with its ligand. Nanobody alone was used as a control and its signal is considered a baseline (marked with a black, dashed line in Figure 14). The signal from CBD alone with and without ligand is considered negative because it is below the baseline.

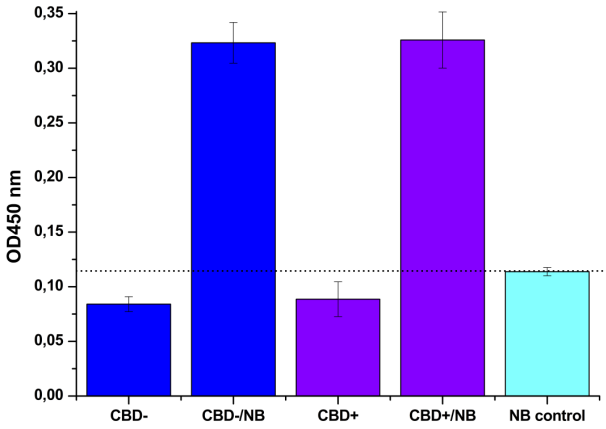

Figure 15. ELISA assay for STING CBD and NB1829.
Signals from STING CBD without ligand, alone (CBD-) and with NB (CBD-/NB) are shown as the blue columns. Signals from STING CBD with ligand, alone (CBD+) and with NB (CBD+/NB) are shown as the violet columns. Signal from NB alone is shown as a cyan column. The error bars correspond to standard deviation. The black dashed line is a baseline below which the signal is considered negative.

8.2.2. Surface plasmon resonans

The SPR was done using Biacore X100+ on Ni-NTA chip. 10 nM his-tagged NB1829 was immobilized on the chip and subjected to STING CBD injections at different concentration in each cycle (31.25 nM, 15.625 nM, 7.8125 nM and 3.9 nM). The data was evaluated using 1:1 binding model in Biacore Evaluation software. The results are presented in Figure 16. The binding affinity was 1.015 nM, association rate constant was 6.672 x 10^{-5} 1/(M*sec) and dissociation rate constant was 6.774 x 10^{-4} 1/sec. The halflife time of the complex was approximately 17'.

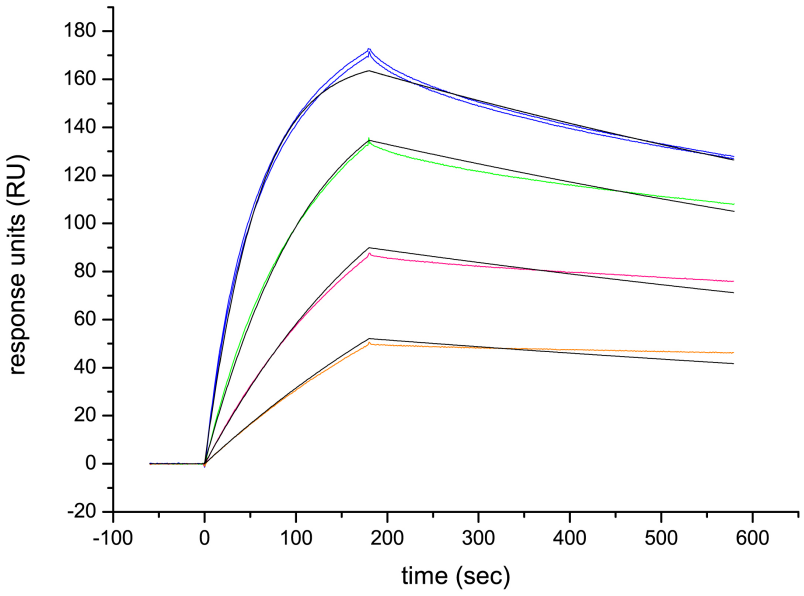

Figure 16. SPR measurement for STING CBD and NB1829 (previous page).
Injections of NB1829: 31.25 nM (blue), 15.625 nM (green), 7.8125 nM (magenta) and 3.9 nM (orange). Black lines show the best fit calcultated by Biacore Evaluation software. The resuts are shown in reference-subtracted (FC2-1) response units vs. time in seconds.

8.3. X-ray crystallography

8.3.1. HTS of crystallization conditions and optimization of crystals

STING CBD and NB1829 complex was purified using size-exclusion chromatography on Superdex 200 10/300 column. The proteins eluted in form of two monodisperse peaks - the first at approximately 14 mL, containing the complex, and the second at approximately 20 mL, containing the excess of NB (Figure 17B). The pooled main peak fractions were analyzed on SDS-PAGE gel which confirmed 1:1 complex formation with STING CBD and NB detected as single bands at 26 and 17 kDa, respectively (Figure 17A). The main peak fractions containing the complex were pooled and concentrated to 9.2 mg/mL. The complex was crystallized using vapour diffusion method in 200 nL sitting drops. The high-throughput screening of crystallization conditions gave rise to crystals of various morphologies (Figure 18). The diamond-shaped crystals from the condition containing 25% PEG 3350, 0.1 M BIS-TRIS, pH 5.5 and 0.2 M magnesium chloride (Figure 18B) were further optimized.

The optimization complex was prepared in larger scale in the same way as HTS complex and purified using Superdex 200 16/600 column. It eluted as two monodisperse peaks at approximately 85 mL (peak containing the complex) and 110 mL (peak containing an excess of NB) (Figure 17D). The pooled fractions were analyzed on SDS-PAGE gel which confirmed 1:1 complex formation. Its components migrated similarly as those of HTS complex (Figure 17C). The complex was concentrated to 9 and 12 mg/mL. The crystallization was performed in 2 μL hanging drops. The optimization gave rise to bigger crystals of various morfologies (Figure 19). The high diffraction quality crystals were obtained from 12 mg/mL protein incubated in condition containing 20 % PEG 3350, 0.1 M BIS-TRIS, pH 5.5 and 0.25 M magnesium chloride (Figure 19F).

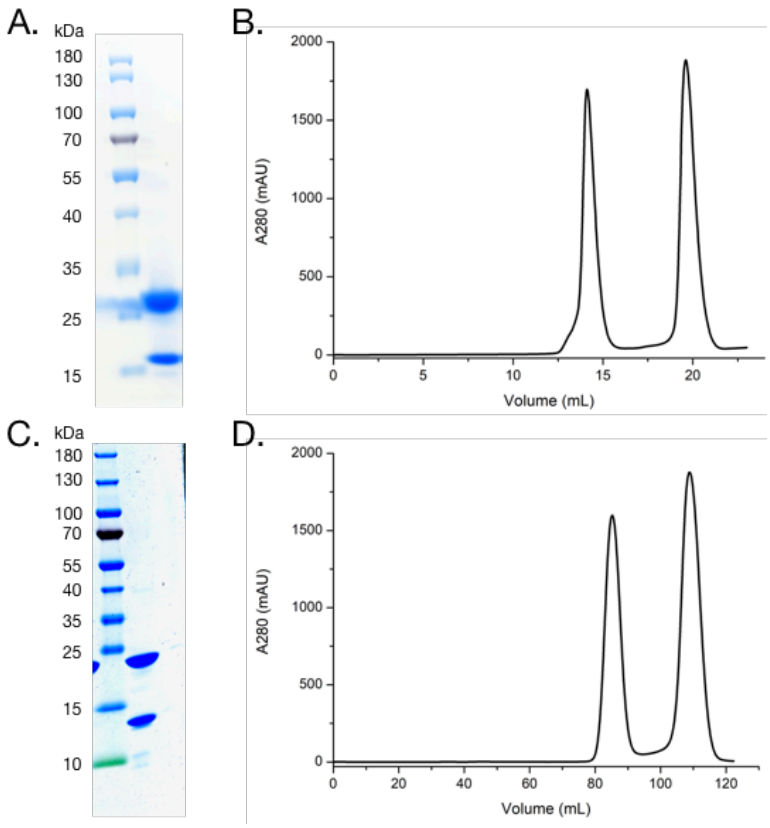

Figure 17. Purification of crystallization complexes.
A. and C. SDS-PAGE gel for STING CBD-NB1829 crystallization complexes prepared for HTS of crystallization conditions and optimization of crystallization, respectively. B. and D. Size-exclusion chromatography elution profiles for STING CBD-NB1829 crystallization complexes for HTS of crystallization conditions (Superdex 200 10/300) and optimization of crystallization (Superdex 200 160/600), respectively.

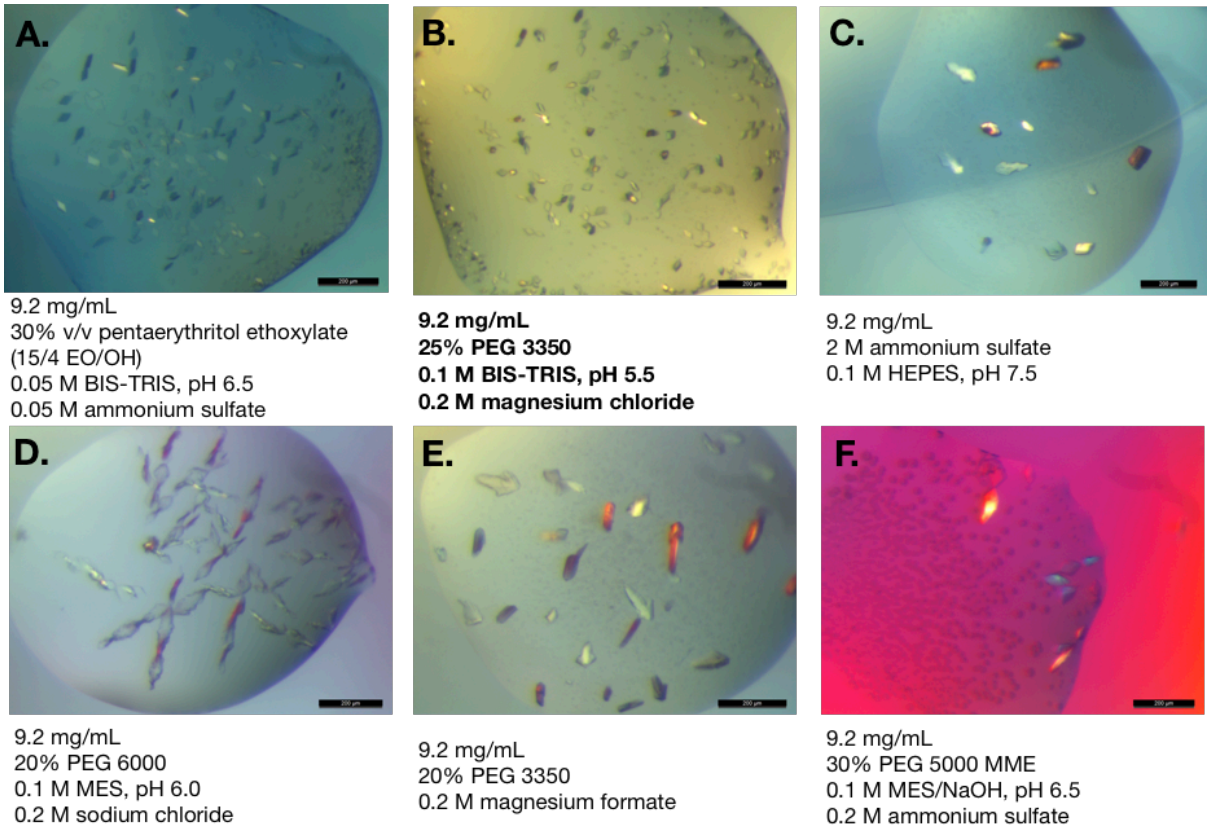

Figure 18. HTS of crystallization conditions for STING CBD and NB1829 complex.
A.-F. The representative crystals from HTS. The concentration of the complex and components of crystallization conditions are shown below the photos. B. The condition used to prepare initial crystals that were further optimized, and gave the high-resolution structure presented in this thesis.

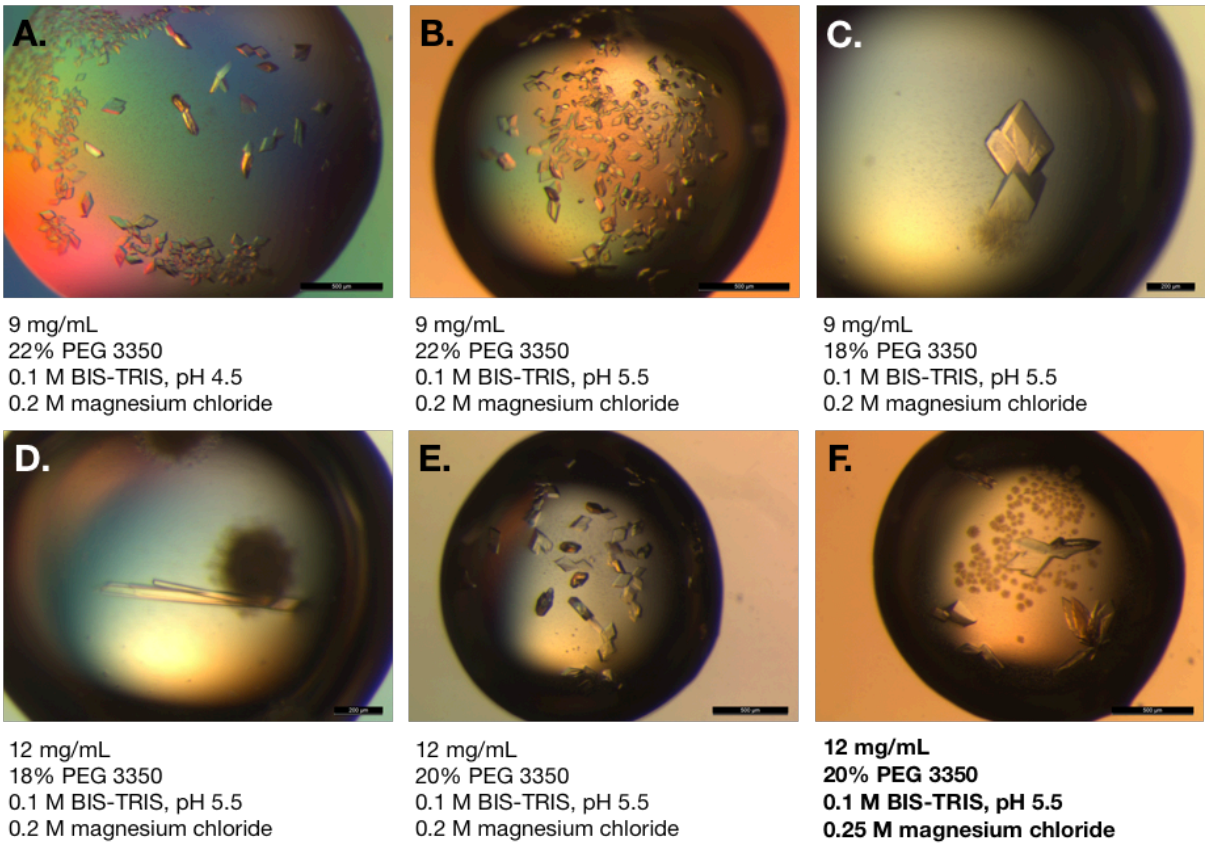

Figure 19. Optimization of crystallization conditions for STING CBD and NB1829 complex. A.-F. The representative crystals from optimization of crystallization. The concentrations of the complex and components of crystallization conditions are shown below the photos. F. The crystalls that gave rise to the high-resolution structure presented in this thesis.

8.3.2. Determination of STING CBD and NB1829 complex structure

The structure was solved using molecular replacement with application of the ligandless human STING CTD structure (PDB accession code 4F5W) with manually removed flexible loops and the structure of human single domain antibody (PDB accession code 3ZHK), converted to poly-alanine model, as search models. The crystal was in space group P1 and diffracted to 1.95 Å resolution. The diffraction pattern is shown in Figure 20. The data collection statistics are shown in Table 2.

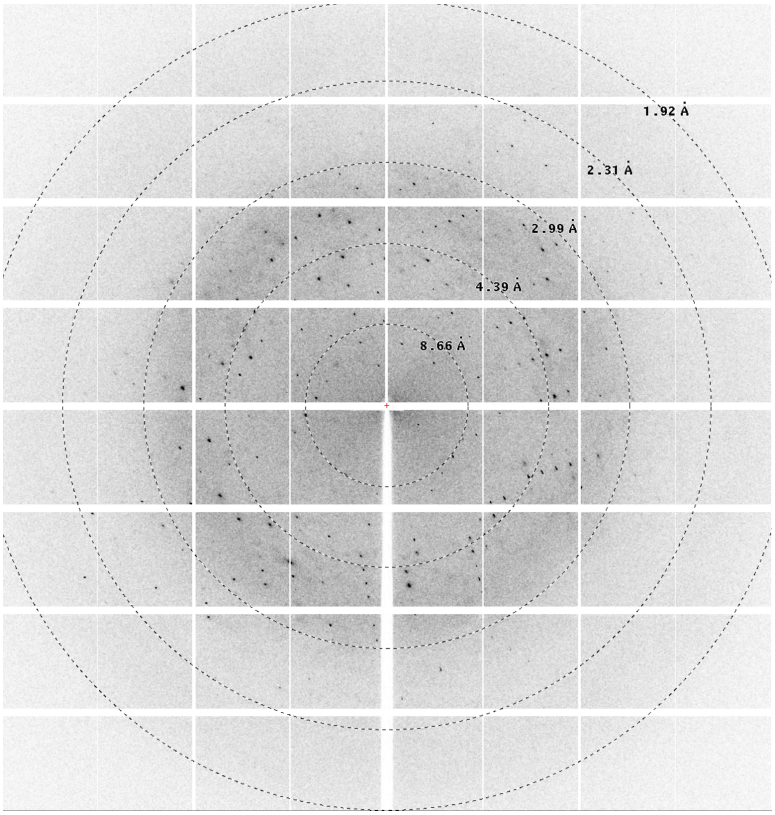

Figure 20. Diffraction pattern.
The diffraction pattern obtained from a crystal of STING CBD and NB1829 complex.

Table 2. Data collection and refinement statistics.

Data collection	
Beamline	EMBL P14 at PETRA III, DESY Hamburg
Wavelength (Å)	0.9762
Resolution range (Å)	999-1.95 (2.00-1.95)
Space group	P1
a,b,c (Å)	51.86, 52.32, 60.16
α,β,γ (°)	97.32, 102.17, 84.95
Total reflections	153852 (9428)
Unique reflections	39834 (2434)
Multiplicity	3.86 (3.87)
Completeness (%)	89.2 (73.3)
I/σ(I)	17.1 (1.77)
Wilson B factor	56.35
R-meas (%)	3.8 (83.4)
CC1/2	99.9 (80.7)
Refinement	
Resolution (Å)	51.79-1.95
Reflections used in refinement	39828

R_{work} (%)	21.46
R_{free} (%)	25.86
Molecules in ASU	4
Macromolecules	4787
Solvent	64
RMSD (bond lengths) (Å)	0.09
RMSD (bond angles) (°)	1.092
Ramachandran favored (%)	96.45
Ramachandran allowed (%)	3.19
Ramachandran outliers (%)	0.35
Rotamer outliers (%)	0.41
Average B factors (macromol./solv.)	65.45

Values in parentheses are for the last shell. ASU - asymmetric unit.

The asymmetric unit (ASU) contains a STING CBD dimer (chains A and B) and two molecules of nanobody (chains C and D) in proximity of the N-terminal side of the dimer (at the bottom of butterfly-shaped molecule) (Figure 21A). The tightly packed crystal lattice is shown in Figure 21B.

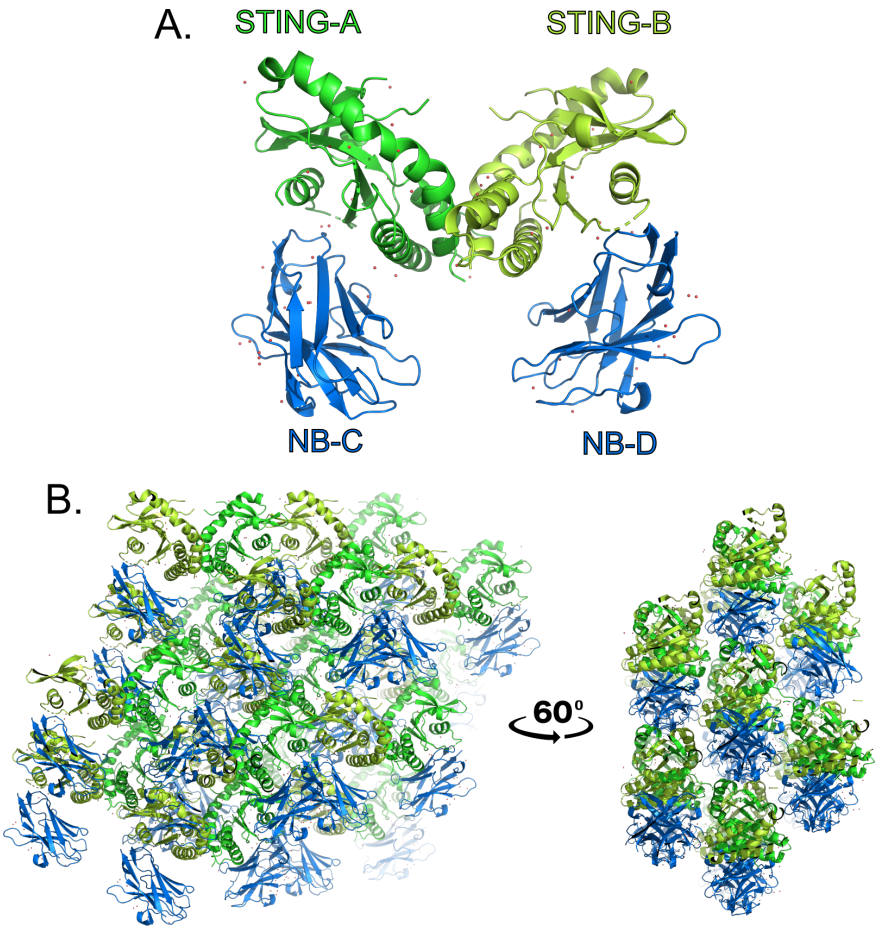

Figure 21. Structure of STING CBD and NB1829 complex (previous page).
A. Asymmetric unit containing STING CBD dimer and two nanobodies at the N-terminal side of the dimer. Chain IDs marked with A, B, C and D letters added to the protein's name. B. Part of the crystal lattice, shown at two angles. Water molecules are shown as red spheres.

After multiple rounds of model building and refinement, the model had overall resolution of 1.95 Å, R_{work} = 21.46% and R_{free} = 25.86%. The Ramachandran statistics were within acceptable limits with only two residues in generously allowed regions (Gln138 in chain B, and Glu166 in chain A) and one residue in disallowed region (Ser167 in chain A). The Ramachandran plot in shown in Figure 22. The refinement statistics are listed in Table 2.

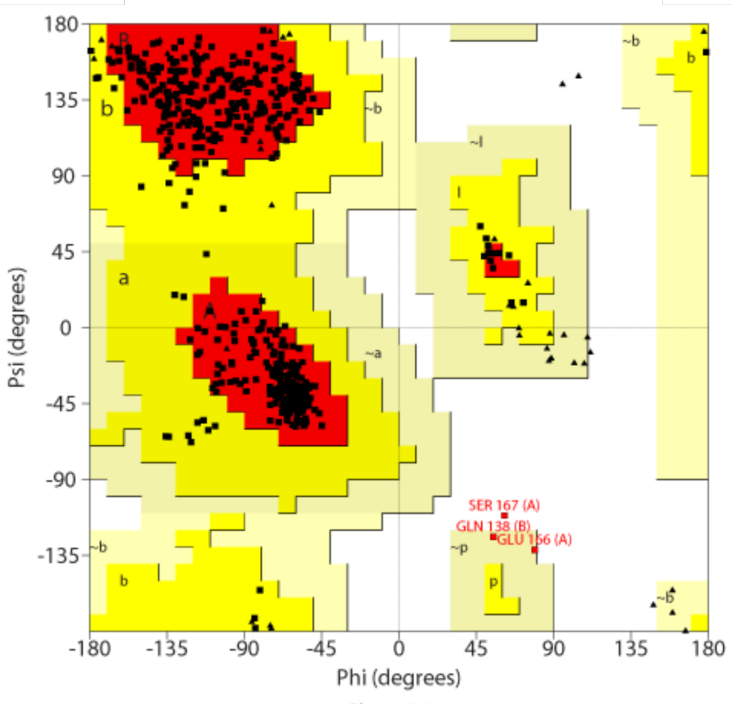

Figure 22. Ramachandran plot.
The Ramachandran plot of STING CBD-NB1829 crystal structure. Red - most favored regions (marked with A, B and L), yellow - additional allowed regions (marked with a, b, l and p), pale yellow - generously allowed regions (marked with ~a, ~b, ~l and ~p), white - disallowed regions. Black squares represent all amino acids except for glycines and prolines. Glycine residues are represented by black triangles. Red squares represent disallowed amino acids. The Ramachandran plot was generated in PROCHECK (Laskowski et al., 1993).

8.3.3. Determination of actual nanobody binding site

Because of quite tight crystal lattice there was a hesitation about where the nanobody binding site on the surface of STING is actually located. To clarify this, the two STING CBD mutants were generated - CBD mutant 1 with mutations at the top of the wings of butterfly-shaped CBD, and CBD mutant 2 with mutations at the bottom (D85K/K86E and E190R/R193A, respectively). The mutations were engineered to compromise binding between the complex

components. The locations of mutations in a slice of crystal lattice were shown in Figure 23. The mutants were used in two binding tests to evaluate which mutations were able to compromise NB binding, and thus identify the binding site location. One test was based on pull-down assay and the other on SEC.

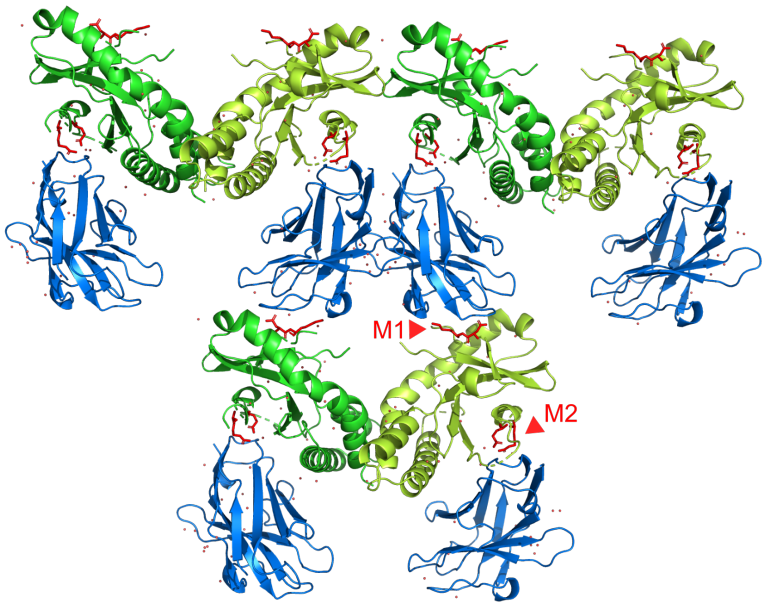

Figure 23. Overview of STING mutants.
Part of the crystal lattice composed of STING (green and limon) and NB1829 (blue) molecules was shown as a cartoon with mutated amino acids shown as sticks (red), and additionally marked with red triangles (M1 - CBD mutant 1, M2 - CBD mutant 2). Water molecules are shown as red spheres.

8.3.3.1. Pull-down assay with NB1829 and STING CBD mutants

His-tagged nanobody 1829 (NB1829) was mixed with non-tagged STING CBD and its mutants, and incubated with TALON resin, which was used instead of Ni-NTA to prevent unspecific binding of CBD to nickel resin. STING CBD WT, its mutants and NB1829 were also used alone as controls. The results are shown in Figure 24.

STING CBD WT and its mutants were shown not to bind to the TALON resin (wells E - elution in sections: M1 - CBD mutant 1, M2 - CBD mutant 2 and WT - CBD wild-type, all not containing any bands). The nanobody was confirmed to bind to the resin (well E in section N - nanobody, containing a single band at 15 kDa for nanobody). The mutations in CBD mutant 1 were shown to compromise its binding to the nanobody (well E - elution in section M1N - CBD mutant 1 with nanobody, containing only a band at 15 kDa for nanobody). This assay confirmed that the NB1829 binds at the top of butterfly-shaped STING CBD. CBD mutant 2 and STING CBD WT were shown to be able to bind to NB1829 (wells E - elution in sections M2N - CBD mutant 2 with nanobody and WTN - STING CBD wild-type with nanobody, both containing two bands: at 25 kDa for CBD mutant 2 or STING CBD WT, and 15 kDa for NB1829). All wash fractions in the samples that bound to the resin

contained a small amount of residual protein (wells W - final wash in sections N, M1N, M2N and WTN in Figure 24).

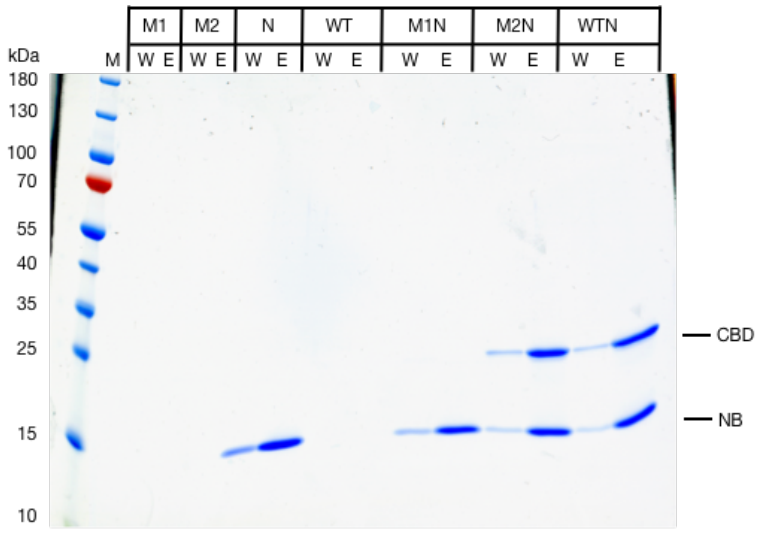

Figure 24. Pull-down assay with STING CBD mutants and NB1829.
SDS-PAGE gel stained with SimplyBlue stain showing results of pull-down assay done on TALON resin. M - molecular weight marker, M1 - CBD mutant 1 alone, M2 - CBD mutant 2 alone, N - nanobody alone, WT - STING CBD WT alone, M1N - CBD mutant 1 + nanobody, M2N - CBD mutant 2 + nanobody, WTN - STING CBD WT + nanobody, W - final wash, E - elution.

8.3.3.2. SEC binding test with NB1829 and STING CBD mutants

STING CBD WT and its mutants were used either alone or mixed with NB1829. NB alone was used as an additional control. After incubation, the samples were analyzed using size-exclusion chromatography on Superdex 200 5/150 column and SDS-PAGE. The results are shown in Figure 25.

CBD mutant 1 mixed with nanobody eluted as two peaks: the first at approximately 2.25 and the second at approximately 3.00 mL (red curve in Figure 25A). The first one had the same elution volume as CBD mutant 1 alone (black curve in Figure 25A), and was confirmed on SDS-PAGE gel to migrate as a single band at 25 kDa, corresponding to the size of CBD mutant 1 alone (well 1-3 in section M1N - mutant 1 with NB in Figure 25C). The analysis of the second peak resulted in two bands at 25 kDa and 15 kDa, corresponding to the size of CBD mutant 1 and nanobody, respectively (well 4 in section M1N in Figure 25C). Thus, possibly, the capacity of the column was slightly exceeded. Nevertheless, the second peak contained all nanobody used in this sample proving that CBD mutant 1 did not bind to the nanobody. The CBD mutant 2 with nanobody also eluted as two peaks at approximately 2.25 and 3.25 mL (green curve in Figure 25B). However, the peak of CBD mutant 2 alone (purple curve in Figure 25B) was slightly shifted towards higher molecular weight, comparing to complex-containing peak 1 for unexplained reasons. SDS-PAGE has shown that both peaks contained both CBD mutant 2 and NB1829 confirming the nanobody binding ability (bands

at 25 and 15 kDa, respectively in wells 1-3 for peak 1, and 4 for peak 2 in section M2N - mutant 2 with nanobody in Figure 25C). The SEC curves for CBD WT alone, nanobody alone and CBD WT with nanobody are not shown here. However, the SDS-PAGE gel presented in Figure 25C contains the gel samples from these runs with a single band at 25 kDa for STING CBD WT and its mutants, and at 15 kDa for NB (wells M1 - CBD mutant 1, M2 - CBD mutant 2, WT - STING CBD WT and N - nanobody). STING CBD WT was confirmed to bind to the nanobody, and together they eluted from the column as two peaks (data not shown). Both of them ran on the gel as two bands at 25 kDa for STING CBD WT and 15 kDa for NB (wells 1 and 2 in section WTN - STING CBD WT with NB).

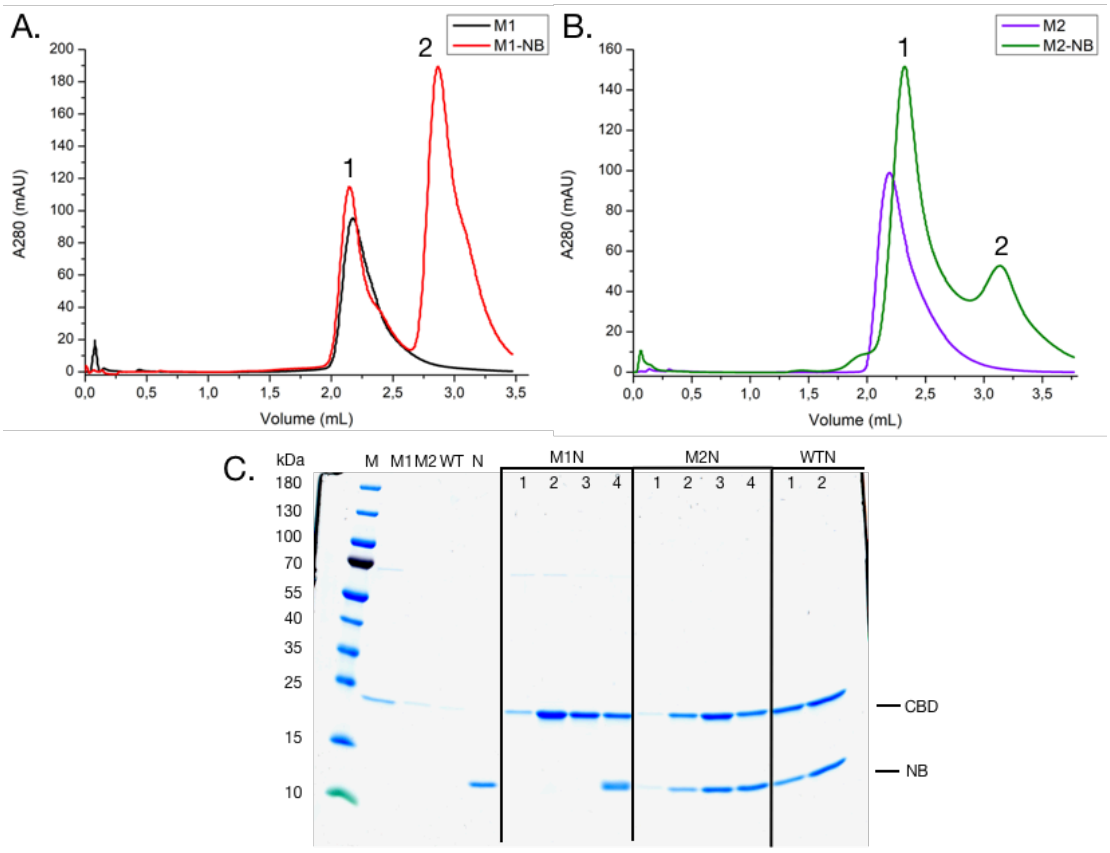

Figure 25. SEC binding test with STING CBD mutants and NB1829.
A. and B. Size-exclusion protein elution profiles. M1 - CBD mutant 1 alone, M1-NB - CBD mutant 1 + nanobody, M2 - CBD mutant 2, M2-NB - CBD mutant 2 + nanobody. C. SDS-PAGE gel stained with SimplyBlue stain with selected SEC fractions. M - molecular weight marker, M1 - CBD mutant 1 alone, M2 - CBD mutant 2 alone, WT - STING CBD WT alone, N - nanobody alone, M1N - CBD mutant 1 + nanobody, M2N - CBD mutant 2 + nanobody, WTN - STING CBD WT + nanobody. M1N lane 1-3 - peak 1 (from A.), M1N lane 4 - peak 2 (from A.), M2N lane 1-3 - peak 1 (from B.), M2N lane 4 - peak 2 (from B.).

8.3.4. General architecture of STING-NB1829 complex

Despite that asymmetric unit contains two nanobodies sticking to the N-terminal surface of the STING CBD dimer, it was shown experimentally that the real interaction surface is at the top of the butterfly-shaped dimer, where two nanobodies from adjacent units bind (Figure 26A). The locations of the CDRs in the nanobody sequence were determined basing on nanobody sequences analysis presented by Melarkode Vattekatte et al. (Melarkode Vattekatte et al., 2020), and are as follows: CDR1 - residues 26-33, CDR2 - residues 52-59 and CDR3 - residues 97-114 (shown in magenta, yellow and orange, respectively in Figure 26A).

The binding interface was analyzed using PDBsum database (Laskowski et al., 2001). Each STING monomer has a binding interface with a total surface area of 640 Å^2, and contains 12 residues (Tyr44-Leu52, Phe83-Val101 and Arg115). The nanobody binding interface has a total surface area of 601 Å^2, and contains 14 residues which are within CDR2 (Gly55-Tyr59) and CDR3 (His99-Asp113). CDR1 is not engaged in this interaction. There are ten hydrogen bonds, one salt bridge and 87 non-bonded contacts between the STING monomer and the nanobody. The representative part of the binding interface is shown in Figure 26B.

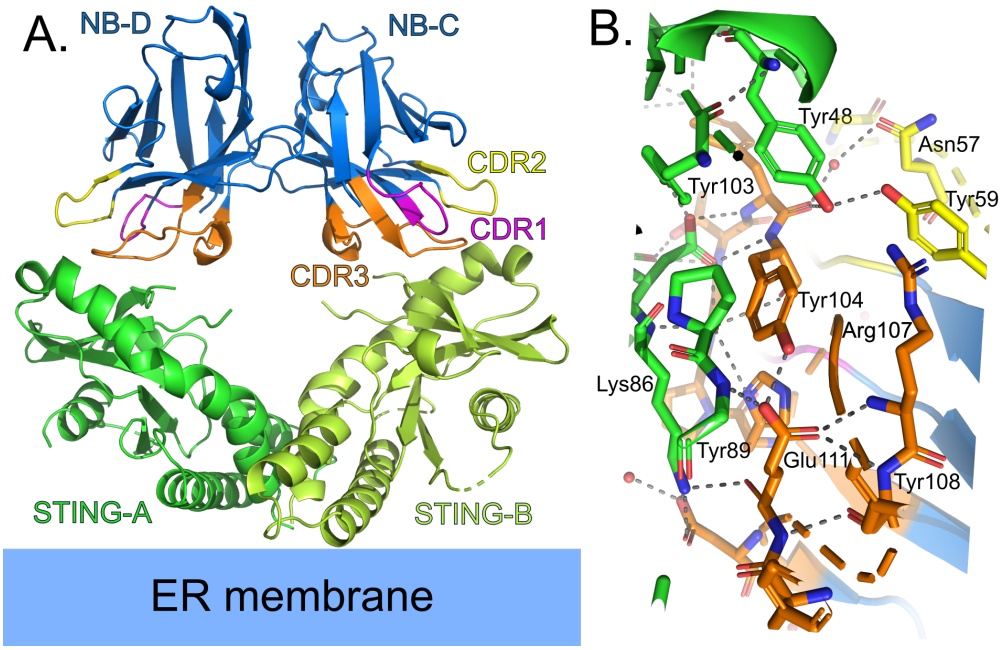

Figure 26. Binding interface of STING CBD and NB1829 complex.
A. The STING CBD dimer is shown in two shades of green (chain A - green, chain B - limon). The nanobodies are shown in blue (chain C and D). The CDRs are colored in magenta, yellow and orange, for CDR1, 2 and 3, respectively. Water molecules were removed for clarity. B. The representative part of STING CBD-NB1829 binding interface with interacting amino acids shown as sticks. Color code is the same as in (A.) for the parts represented as a cartoon and for carbon atoms. Nitrogen atoms are shown in blue, oxygen atoms are shown in red. Polar contacts, generated in PyMOL are shown as dashed lines in dark grey. Water molecules are represented as red spheres.

8.3.5. Conformation of STING-NB1829 complex

NB1829 was shown by Jennifer Deborah Wuerth and Florian Schmidt to gently activate STING by increasing interferon stimulated gene CXCL10 expression and slightly increasing interferon beta expression. Additionally, intracellular expression of NB1829 was shown to cause IRF3 accumulation in the nucleus which normally occurs after STING activation (data not shown). The structure presented here was made in attempt to explain these observations. Surprisingly, there was just a slight conformational change of the STING CBD dimer interacting with nanobody 1829, comparing with the open (apo) conformation (PDB accession code 4F5W) (Shang et al., 2012) with RMSD = 1.108 Å for 308 atoms. In contrast, for alignment of open and closed (with 2',3'-cGAMP bound; PDB accession code 5BQX) (Shi et al., 2015) conformations, RMSD = 3.197 Å for 314 atoms. After comparison of distances between Cα atoms of Gln184 residues between monomers building each STING CBD dimer, it was found that the dimer complexed with the nanobody was even in slightly more open conformational state than the apo protein. The distances were 59.1 Å, 56.5 Å and 41.0 Å for complexed with nanobody, open and closed states, respectively (Figure 27).

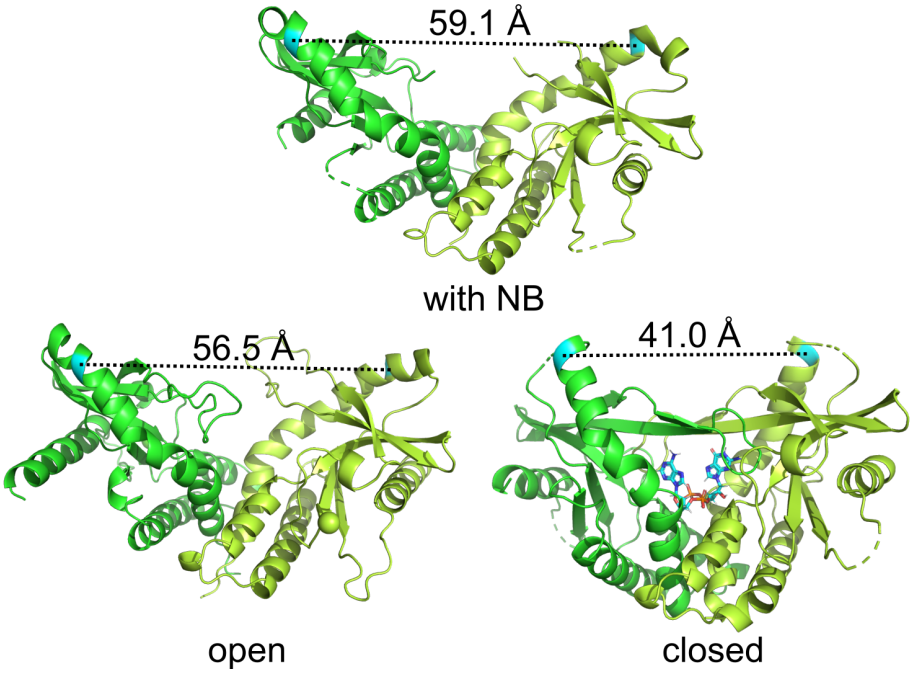

Figure 27. Comparison of STING CBD conformations.
STING CBD monomers are shown as cartoon in green and limon, respectively. Gln184 residues are shown in cyan. 2',3'-cGAMP is shown as sticks (carbon - cyan, nitrogen - blue, oxygen - red, phosphorus - orange). Water molecules were removed for clarity. The sphere in limon in the structure of STING in its open conformation represents a Ca^{2+} ion. The distances are shown between Cα atoms of Gln184 (amino acid numbering corresponds to the full-length protein). Open and closed states represent models deposited in PDB with accession codes 4F5W and 5BQX, respectively.

8.3.6. General architecture of NB1829

NB1829 has a very similar molecular architecture to other nanobodies deposited in the PDB. To illustrate that, the NB1829 was aligned with other selected nanobodies using PyMOL program. The biggest similarity was observed between NB1829 and the template structure used for molecular replacement to solve its structure (PDB accession code 3ZHK) (Kim et al., 2014). The R.M.S.D. was 0.365 Å for 86 atoms. The alignments of NB1829 with two nanobodies obtained from *Vicugna pacos* (alpaca) and two others derrived from *Lama glama* (llama) have resulted in R.M.S.D.s below 0.7 Å. Among them there were structures with PDB accession codes: 7KJH (R.M.S.D. 0.629 Å for 90 atoms) (Dietrich et al., 2021), 7X2L (R.M.S.D. 0.427 Å for 90 atoms) (not published), 7NMU (R.M.S.D. 0.555 Å for 92 atoms) (Slater et al., 2021) and 7R63 (R.M.S.D. 0.420 Å for 93 atoms) (Srinivasan et al., 2022). The least similar to NB1829 is the structure of synthetic anti-Fab nanobody with PDB accession code 6WW2 (R.M.S.D. 0.951 Å for 103 atoms) (Tsutsumi et al., 2020). The alignments are presented in Figure 28.

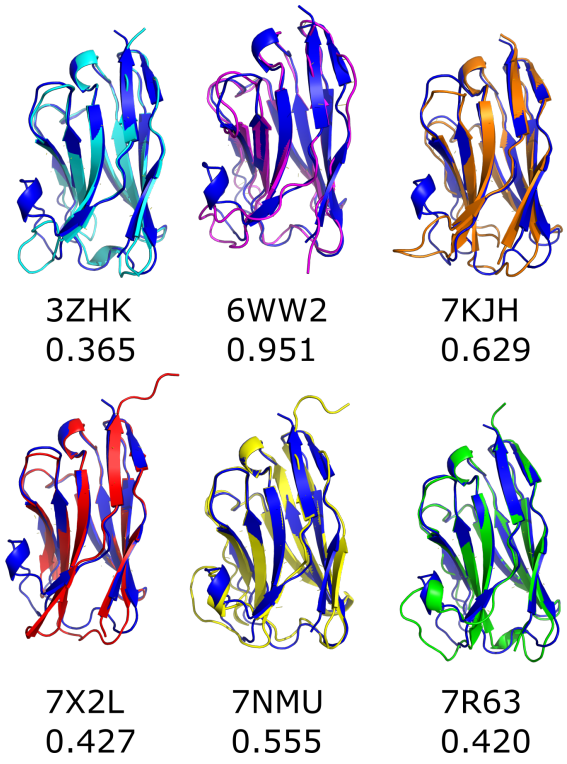

Figure 28. Alignements of selected known nanobody structures with NB1829.
NB1829 was aligned in PyMOL with other selected structures of nanobodies deposited in PDB. All structures are shown as a cartoon with removed waters, ions and partner protein components. NB1829 is shown in blue, human single domain antibody 2x1 scaffold (3ZHK) is shown in cyan, synthetic anti-Fab nanobody (6WW2) is shown in magenta, *V. pacos* nanobodies: nanobody B9 (7KHJ) and nanobody 3-2A2-4 (7X2L) are shown in orange and red, respectively. *L. glama* nanobodies: nanobody 2 (7NMU) and nanobody 82 (7R63) are shown in yellow and green, respectively. The descriptions contain PDB accession codes and numerical values representing R.M.S.D. between NB1829 and selected nanobody in Ångströms (Å).

9. Results for cluster of differentiation 47

9.1. Expression and purification of proteins

9.1.1. Expression and purification of MSPs

MSPs were expressed in *E. coli* Rosetta 2(DE3) and purified using Ni-NTA affinity chromatography according to protocol published by Bayburt et al. with several modifications (Bayburt et al., 2002). The mild detergent sodium cholate was present during the washing step to strip any *E. coli* lipids sticking to the amphipatic MSP. Optional step was the cleavage of his-tag using TEV protease, followed by the subtractive IMAC on Ni-NTA (data not shown). The results are shown in Figure 29.

His-tagged MSP1D1 migrated as a single band on the SDS-PAGE gel at approximately 25 kDa. Slight degradation was present in form of lower molecular weight bands (lane E in Figure 29A). His-tagged MSP1E3D1 migrated as a single band at approximately 28 kDa with the highest amount of degradation among all MSPs (lane E in Figure 29B). His-tagged MSP1D1dH5 migrated as a single band at approximately 20 kDa and did not show any significant degradation (lane E in Figure 29C). A protein degradation is typical for MSPs, it does not influence nanodisc formation and is cleared out during NDs purification. The usual protocol of MSP purification does not include SEC because of amphipatic nature of this protein and its tendency to form oligomers in absence of lipids. The yield was approximately 40 mg of each MSP type per liter of culture. The concentrations were calculated using the following values: his-tagged MSP1D1 (Ɛ = 21,430 $M^{-1}cm^{-1}$; M.W. = 24,792.86 Da), his-tagged MSP1E3D1 (Ɛ = 29,910 $M^{-1}cm^{-1}$; M.W. = 32,730.86 Da), his-tagged MSP1D1dH5 (Ɛ = 19,940 $M^{-1}cm^{-1}$; M.W. = 21,468.12 Da).

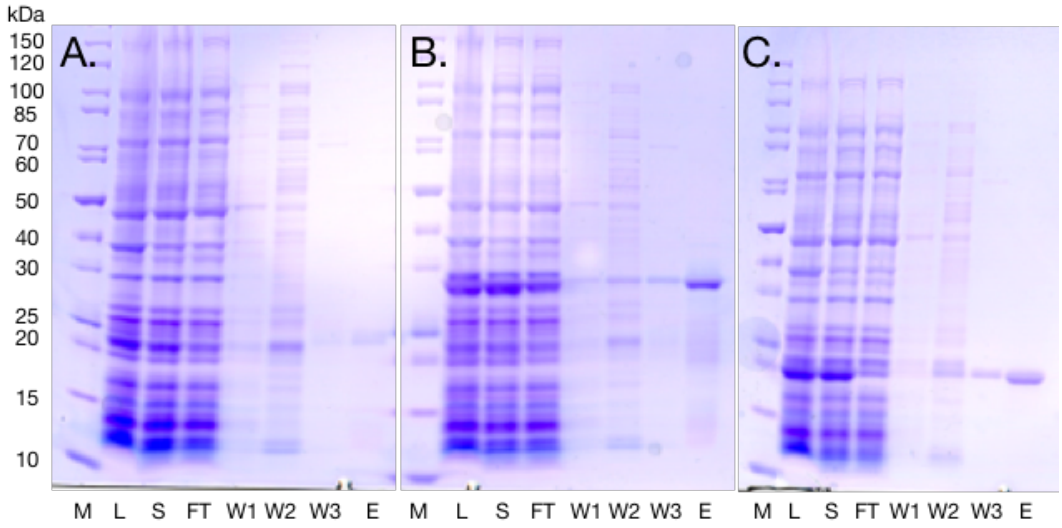

Figure 29. Purification of MSPs (previous page).
MSPs were purified using Ni-NTA affinity chromatography followed by dialysis. In this purification the his-tag was not cleaved. The purification samples were analyzed on SDS-PAGE gel stained with Coomassie Brilliant Blue stain. M - molecular weight marker, L - lysate, S - supernatant after centrifugation, FT - flow-through, W1-W3 - washes 1-3, E - elution. A. MSP1D1. B. MSP1E3D1. C. MSP1D1dH5.

9.1.2. Expression and purification of Fabs

The Fabs were expressed in partially oxidative cytoplasm of *E. coli* SHuffle T7 cells that enables disulfide bonds formation. Since Fabs are very stable proteins with melting temperatures usually exceeding 75 °C, during purification the cell lysates were heated up to 63 °C for 30' to denature bacterial proteins and not assembled fragments of Fabs. Afterwards, the Fabs were purified using Ni-NTA affinity chromatography and size-exclusion chromatography. Initially, HuFab was cloned without the first aspartic acid in the light chain (HuFab-D) which was later modified (HuFab). However, such modification seems not to affect the Fab stability or binding function. HuFab-H12-P is the Fab with a rigidified elbow region. The results are shown in Figures 30-32.

HuFab-D migrated on SDS-PAGE gel as a single band at approximately 26 kDa with slight degradation in form of lower molecular weight bands (lanes E and SE in Figure 30A). HuFab-D eluted as a single, monodisperse peak at approximately 16 mL from Superdex 200 10/300 column. The main peak was preceded by the tiny peak containing aggregated protein (Figure 30B). The melting temperature was measured using Tycho NT.6 device, and it was 90.8 °C, which was even higher than in case of HuFab (Figure 30C). The yield of HuFab-D was very low, approximately 0.12 mg/L.

HuFab was purified in the same way as HuFab-D but in larger scale. The column Superdex 200 16/600 was used. It migrated on the gel as a single band at approximately 26 kDa with a slight degradation (lanes E and SE in Figure 31A) and eluted from SEC column as a single, monodisperse peak at approximately 93 mL. The main peak was preceded by a tiny peak containing aggregated protein (Figure 31B). The approximate yield was 0.5 mg/L. The melting temperature was 87.9 °C (Figure 31C).

HuFab-H12-P migrated on the gel as a single band at approximately 26 kDa with slight degradation (Figure 32A) and eluted from Superdex 200 16/600 column as a single, monodisperse peak at approximately 93 mL. The main peak was preceded by the tiny peak containing aggregated protein (Figure 32B). The approximate yield was 0.5 mg/L. The melting temperature was 80.7 °C (Figure 32C).

The concentrations were calculated using the following values: HuFab-D (ε = 67,770 $M^{-1}cm^{-1}$; M.W. = 49003.74 Da), HuFab (ε = 67,770 $M^{-1}cm^{-1}$; M.W. = 49,381.21 Da), HuFab-H12-P (ε = 67,770 $M^{-1}cm^{-1}$; M.W. = 49,228.05 Da).

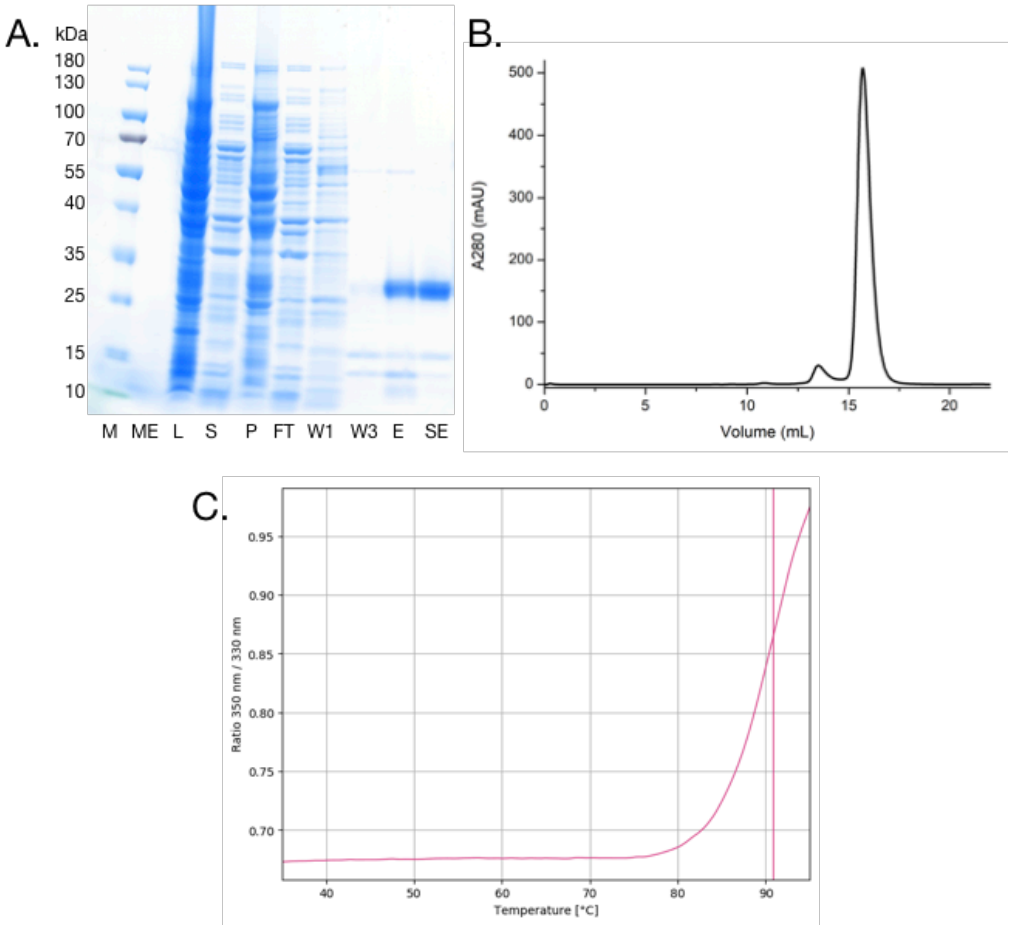

Figure 30. Purification of HuFab-D.

A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, ME - medium, L - lysate, S - supernatant after centrifugation, P - pellet, FT - flow-through, W1, W3 - washes 1 and 3, E - elution, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superdex 200 10/300 column. C. Melting temperature curve from Tycho NT.6 shown as a ratio of measurements at 350 nm and 330 nm.

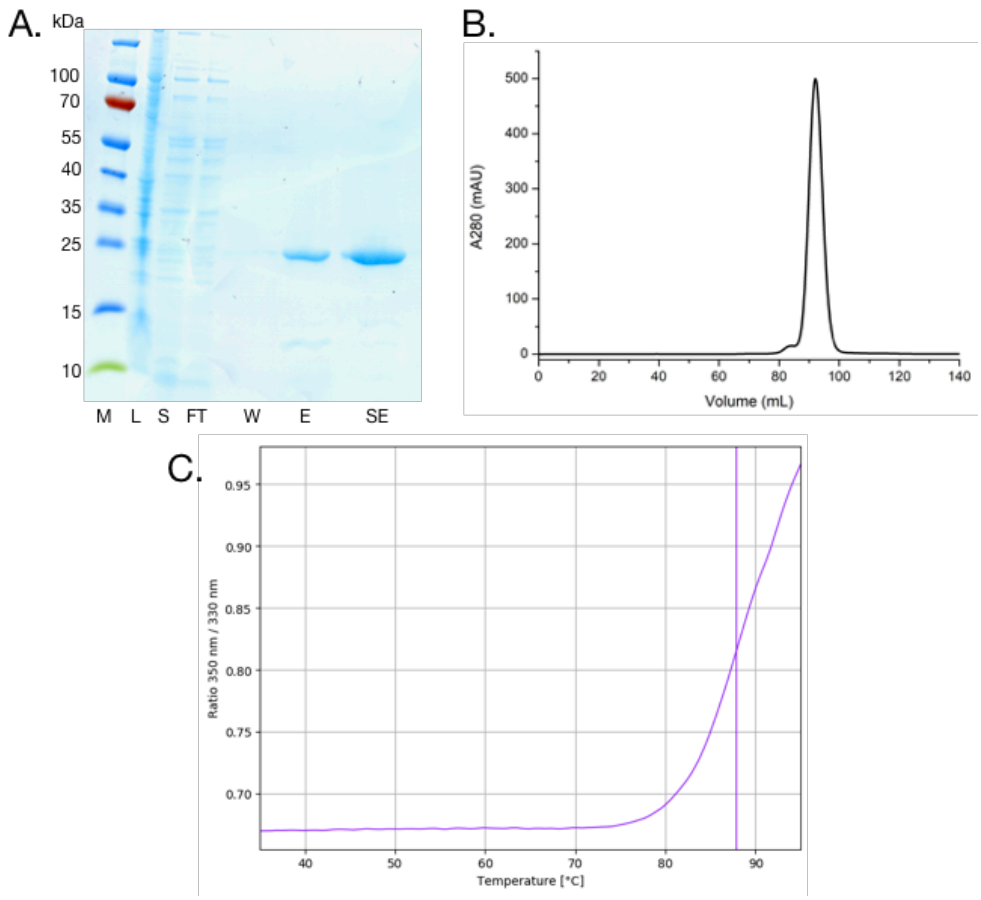

Figure 31. Purification of HuFab.
A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, L - lysate, S - supernatant after centrifugation, FT - flow-through, W - final wash, E - elution, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superdex 200 16/600 column. C. Melting temperature curve from Tycho NT.6 shown as ratio of measurements at 350 nm and 330 nm.

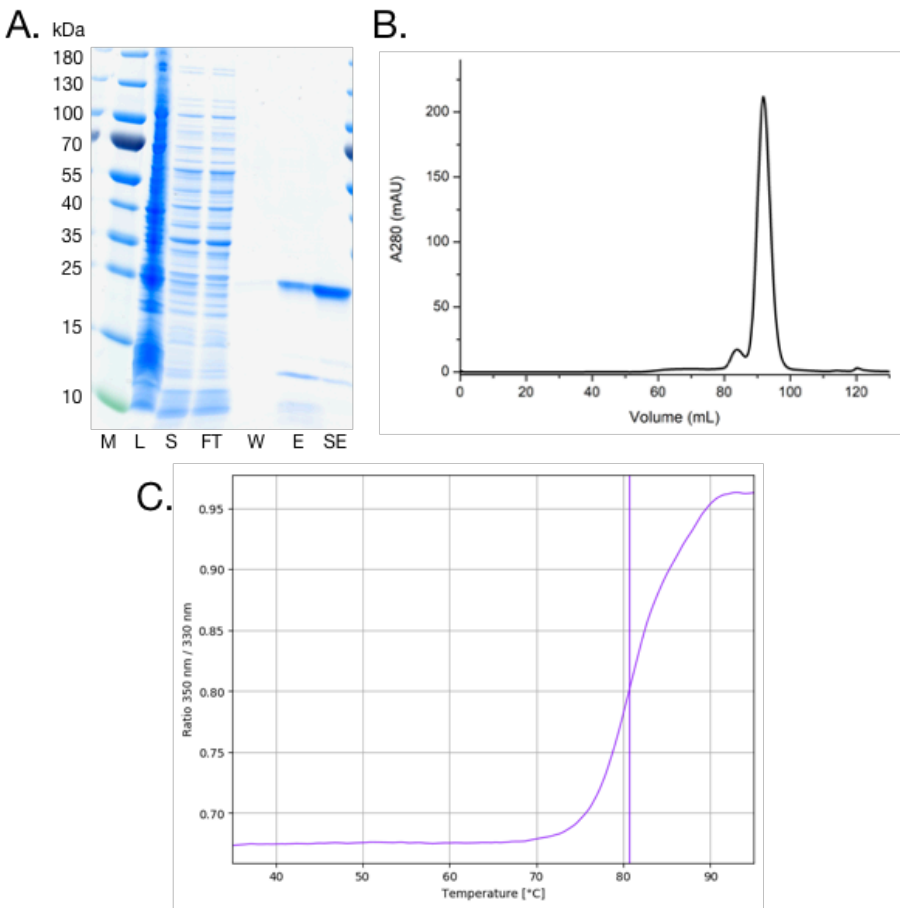

Figure 32. Purification of HuFab-H12-P.
A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, L - lysate, S - supernatant after centrifugation, FT - flow-through, W - final wash, E - elution, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superdex 200 16/600 column. C. Melting temperature curve from Tycho NT.6 shown as ratio of measurements at 350 nm and 330 nm.

9.1.3. Expression and purification of anti-Fab NB

The anti-Fab nanobody was expressed in periplasm of *E. coli* WK6 cells using pelB periplasmic signal sequence to properly form disulfide bonds. It was purified from periplasm using two-step osmotic shock, Ni-NTA chromatography and size exclusion chromatography using Superdex 75 16/600 column. The results are shown in Figure 33.

His-tagged anti-Fab nanobody migrated on the SDS-PAGE gel as a double band at approximately 15 kDa (Figure 33A) and eluted from SEC column as a single, monodisperse peak at approximately 90 mL (Figure 33B). The melting temperature could not be measured using Tycho NT.6 device, possibly because of a very low tryptophan and tyrosine content (1.6 % and 6.5 %, respectively). The approximate yield was 23 mg/L.

The concentrations were calculated using the following values: Ɛ = 23,045 $M^{-1}cm^{-1}$; M.W. = 14,171.51 Da.

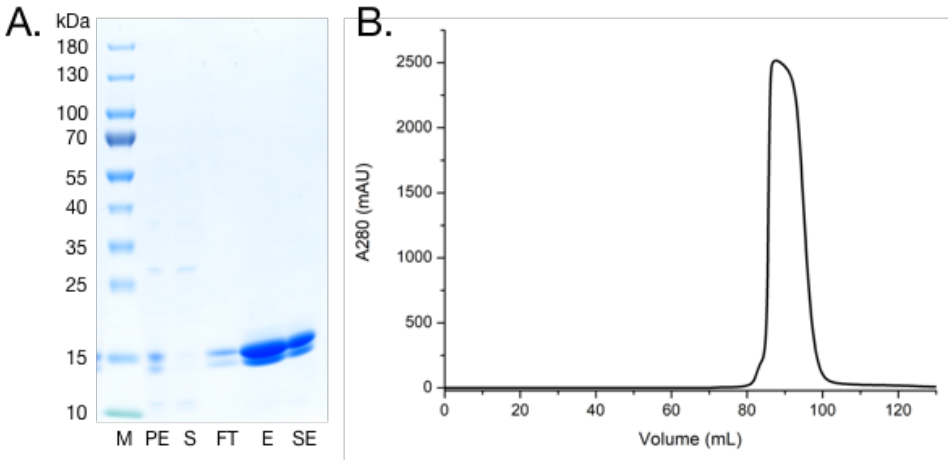

Figure 33. Purification of anti-Fab NB.
A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, PE - periplasmic extract after second osmotic shock, S - supernatant after centrifugation, FT - flow-through, E - elution, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superdex 75 16/600 column.

9.1.4. Expression and purification of full-length CD47

CD47 was expressed in HEK293 Expi cells transfected using PEI reagent. The proteins were extracted from the membranes using a highly concentrated mild detergent DDM supplemented with CHS. The purification was performed using affinity chromatography on Strep-Tactin XT high-capacity resin and size-exclusion chromatography on Superose 6 10/300 column. The purified protein was analyzed on SDS-PAGE gel. The results are shown in Figure 34.

CD47, due to heavy glycosylation migrated on the gel as a smear at approximately 55 kDa (Figure 34A) and eluted from SEC column as a single monodisperse peak at approximately 12 mL which was preceded by a tiny aggregation peak (Figure 34B). The approximate yield was 1 mg/L. The melting temperature was 71.4 °C (Figure 34C).

The concentrations were calculated using the following values: Ɛ = 42,650 $M^{-1}cm^{-1}$; M.W. = 34,618.47 Da.

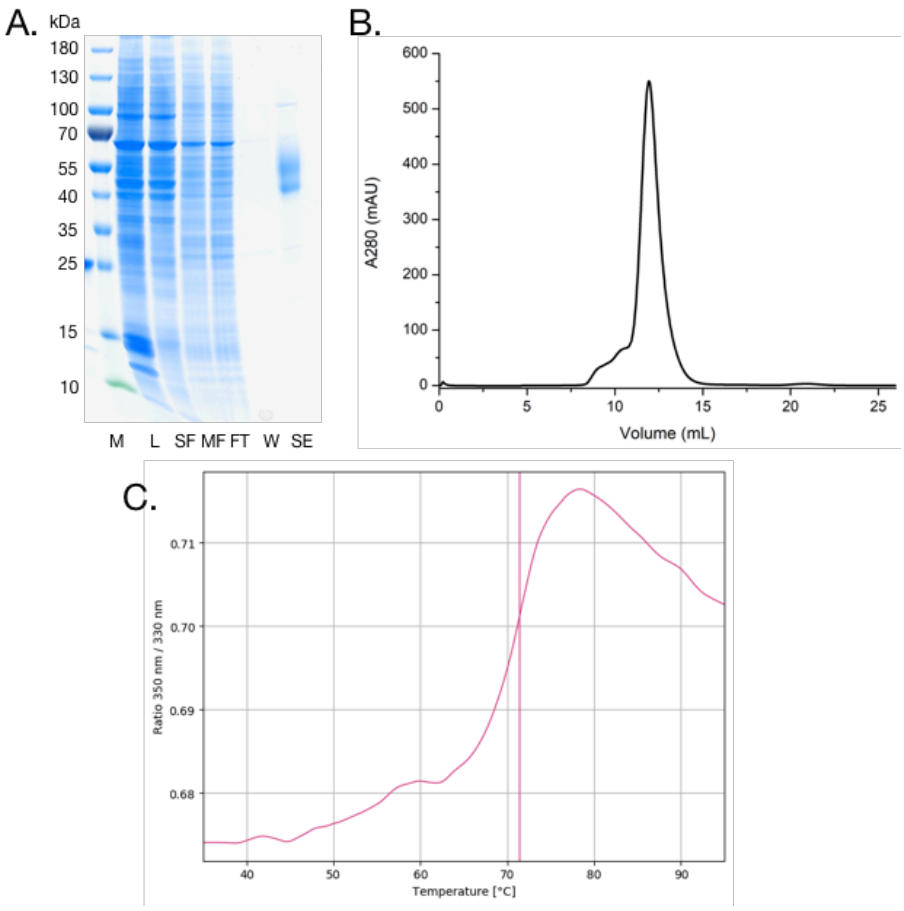

Figure 34. Purification of CD47.
A. SDS-PAGE gel with protein purification samples stained with SimplyBlue stain. M - molecular weight marker, L - lysate, SF - soluble fraction after first ultracentrifugation, MF - solubilized membrane fraction after second ultracentrifugation, FT - flow-through, W - final wash, E - elution, SE - SEC. B. Size-exclusion chromatography protein elution profile from Superose 6 10/300 column. C. The melting temperature curve from Tycho NT.6 shown as a ratio of measurements at 350 nm and 330 nm.

9.2. Evaluation of binding between proteins

9.2.1. Evaluation of binding between CD47, HuFab-H12-P and anti-Fab NB

To evaluate the binding of newly prepared, engineered HuFab-H12-P to CD47 and to anti-Fab nanobody (aF-NB), two pull-down assays were performed - one on Strep-Tactin XT resin that binds strep-tagged CD47, and the other on anti-IgG-CH1 resin that binds HuFab-H12-P. In the first assay, CD47 was immobilized on the resin and mixed with the Fab and NB or Fab alone or NB alone. Additional controls with Fab alone and NB alone without CD47 were also performed. The results are shown in Figure 35. In the second assay the Fab was immobilized on the resin and mixed with nanobody. An additional control with nanobody alone was also performed. The results are shown in Figure 35.

Controls in the first experiment have shown that neither Fab nor NB bound unspecifically to the Strep-Tactin XT resin. Trace amounts of Fab were present in elution fraction, but comparing to large amounts identified in positive samples, its slight, non-specific binding can be ignored. The control has shown that the NB does not bind to CD47 (wells E - elution in sections CN - CD47 with NB, containing only a smear above 43 kDa for CD47; F - Fab only, containing trace amounts of Fab as a band at 25 kDa; and N - NB, not containing any bands, in Figure 35). Fab was confirmed to bind to CD47 and the nanobody was confirmed to bind to the Fab (wells CF - CD47 with Fab, containing a smear above 40 kDa for CD47 and a single band at 25 kDa for Fab; and CFN - CD47 with Fab and NB with bands as in CF and additional double band at 15 kDa for NB in Figure 35).

In the second experiment the nanobody was additionally confirmed to bind to the Fab (well FN - Fab with NB with a band at 25 kDa for Fab and a double band at 15 kDa for NB in Figure 35). The control has shown that nanobody does not unspecifically bind to the anti-IgG-CH1 resin (well N - NB, not containing any bands in Figure 35). The washes were sufficient for all samples (wells W - wash - not containing any bands in Figure 35).

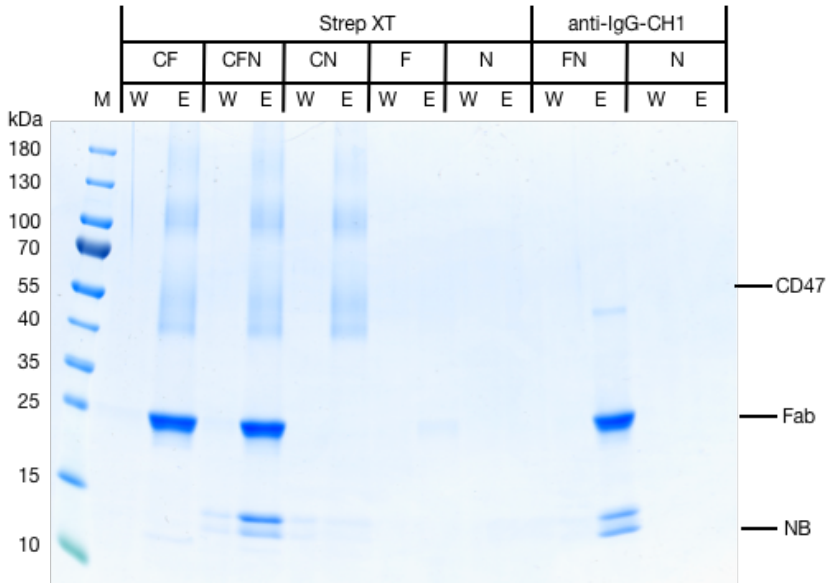

Figure 35. Binding test with CD47, HuFab-H12-P and anti-Fab NB.
SDS-PAGE gel stained with SimplyBlue stain with the samples from binding tests performed on Strep-Tactin XT and anti-IgG-CH1 resins. M - molecular weight marker, CF - CD47 + Fab, CFN - CD47 + Fab + nanobody, CN - CD47 + nanobody, F - Fab only, N - nanobody only, FN - Fab + nanobody, W - final wash, E - elution.

9.3. Optimization of nanodiscs

9.3.1. MSP1D1 nanodiscs

Empty nanodiscs were prepared using the general protocol by mixing membrane scaffold protein of a desired size with detergent-lipid mixed micelles prepared from the mixture of DDM, CHS and polar soybean lipid extract. Multiple MSP:lipids molar ratios were tested in order to find the proper one, from which the most homogenous population of nanodiscs can be formed. The nanodiscs were analyzed using SEC (Superdex 200 10/300) and negative stain EM. Non-tagged empty MSP1D1 nanodiscs (e1D1-NDs) were formed with MSP:lipids molar ratios of 1:20, 1:30, 1:40 and 1:50. The SEC curves are presented in Figure 36A. The most efficient ratio for MSP1D1 nanodiscs was 1:20. These NDs formed a relatively monodisperse peak on SEC (black peak marked with a black triangle in Figure 36A). Except for that, the homogenous, round nanodiscs with diameter of approximately 10 nm were observed using negative stain EM (Figure 36B). The higher the amount of the lipids used, the more polydisperse SEC peaks, shifted more and more towards the column void volume (approximately 8 mL; red, blue and green peaks for 1:30, 1:40 and 1:50 ratios, respectively in Figure 36A). The higher lipids ratios lead to formation of bigger, non-homogenous nanodiscs and liposomes which was confirmed using negative stain EM (1:30 ratio and 1:40 ratio, Figure 36B and D, respectively; the EM image for 1:50 ratio is not shown). The additional control without MSP contained approximately 30 nm liposomes and larger lipid aggregates (Figure 36E).

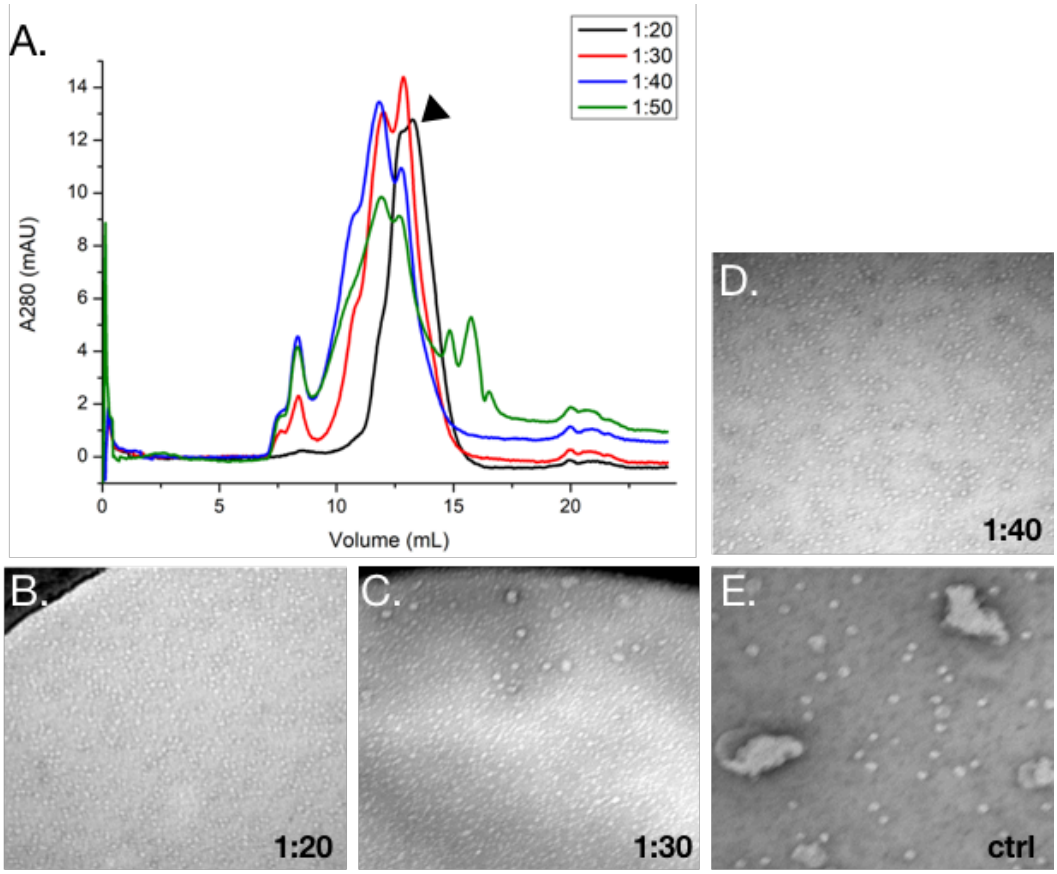

Figure 36. Optimization of empty MSP1D1 NDs (previous page).
A. Comparison of size-exclusion chromatography elution profiles from Superdex 200 10/300 column for nanodiscs prepared with different MSP:lipids molar ratios. The curve representing the most monodisperse nanodiscs is marked with the black triangle. B.-D. Negative stain electron micrographs showing nanodiscs prepared with three different MSP:lipids molar ratios. E. Negative stain electron micrograph of control sample prepared without MSP.

9.3.2. MSP1E3D1 nanodiscs

His-tagged empty 1E3D1 nanodiscs (e1E3D1-NDs) were prepared using MSP:lipids molar ratios of 1:20, 1:35, 1:40, 1:45, 1:65, 1:75 and 1:85. The most efficient ratio which gave rise to nanodiscs running as a monodisperse peak on SEC was 1:35 (red peak, marked with a red triangle in Figure 37A). This ratio was later optimized to 1:30, which was used in further experiments (data not shown). The other ratios are shown in Figure 37A, and similarly as in case of MSP1D1, the higher ratios cause the formation of nanodiscs that migrate as more and more polydisperse peaks shifted towards the column's void volume. With increasing amount of lipids also a peak at the void volume, containing liposomes and lipid aggregates, becomes larger. The SEC peak of the nanodiscs prepared using the lowest ratio (1:20; black peak in Figure 37A) was strongly shifted towards lower molecular weight, comparing to peak from 1:35 nanodiscs, and gently tilted for unknown reasons. The negative stain images of nanodiscs formed using three selected ratios: 1:35, 1:40 and 1:45 are shown in Figure 37B, C and D, respectively. The nanodiscs with 1:35 ratio were the most homogenous with diameter of approximately 10 nm. Even slight increase of amount of lipids in ratio 1:40 caused a formation of larger and less homogenous nanodiscs.

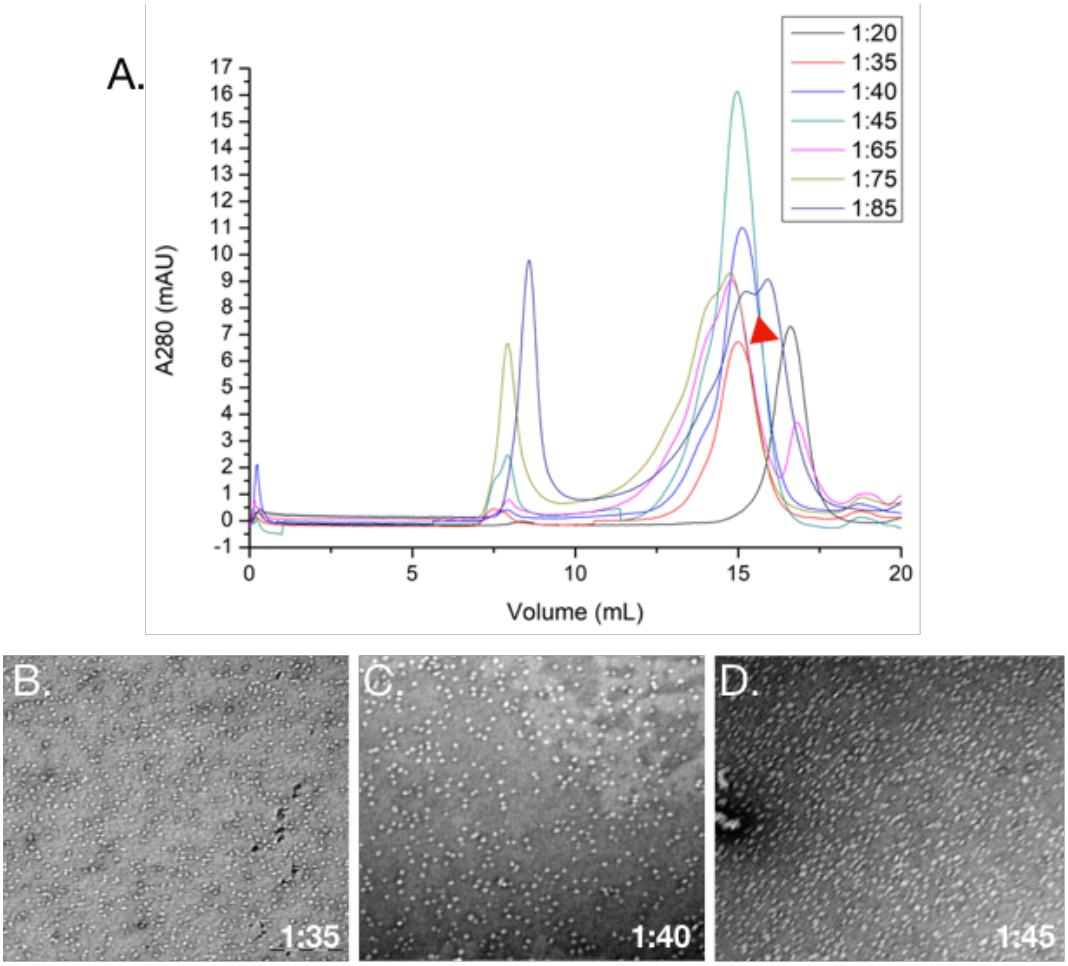

Figure 37. Optimization of empty MSP1E3D1 NDs.
A. Comparison of size-exclusion chromatography elution profiles from Superdex 200 10/300 column for nanodiscs prepared with different MSP:lipids molar ratios. The curve representing the most monodisperse nanodiscs is marked with red triangle. B.-D. Negative stain electron micrographs showing nanodiscs prepared with three different MSP:lipids molar ratios.

9.3.3. MSP1D1dH5 nanodiscs

His-tagged empty MSP1D1dH5 nanodiscs (edH5-NDs) were formed using MSP:lipids molar ratios of 1:10, 1:15, 1:20 and 1:25. The most monodisperse SEC peak was obtained using 1:10 ratio, and this ratio was used in further experiments (black peak, marked with the black triangle in Figure 38A). The ratios 1:15 and 1:20 gave fairly similar results with tiny, additional aggregation peaks (blue and red peaks, respectively in Figure 38A). Surprisingly, the SEC peak for the highest ratio - 1:25 (green peak in Figure 38A) was slightly shifted towards lower molecular weigth comparing to the other peaks. However, it had the largest additional, aggregation peak. The nanodiscs prepared with MSP:lipids ratios 1:10, 1:15 and 1:20 were also fairly similar in EM images (Figure 38B, C and D).

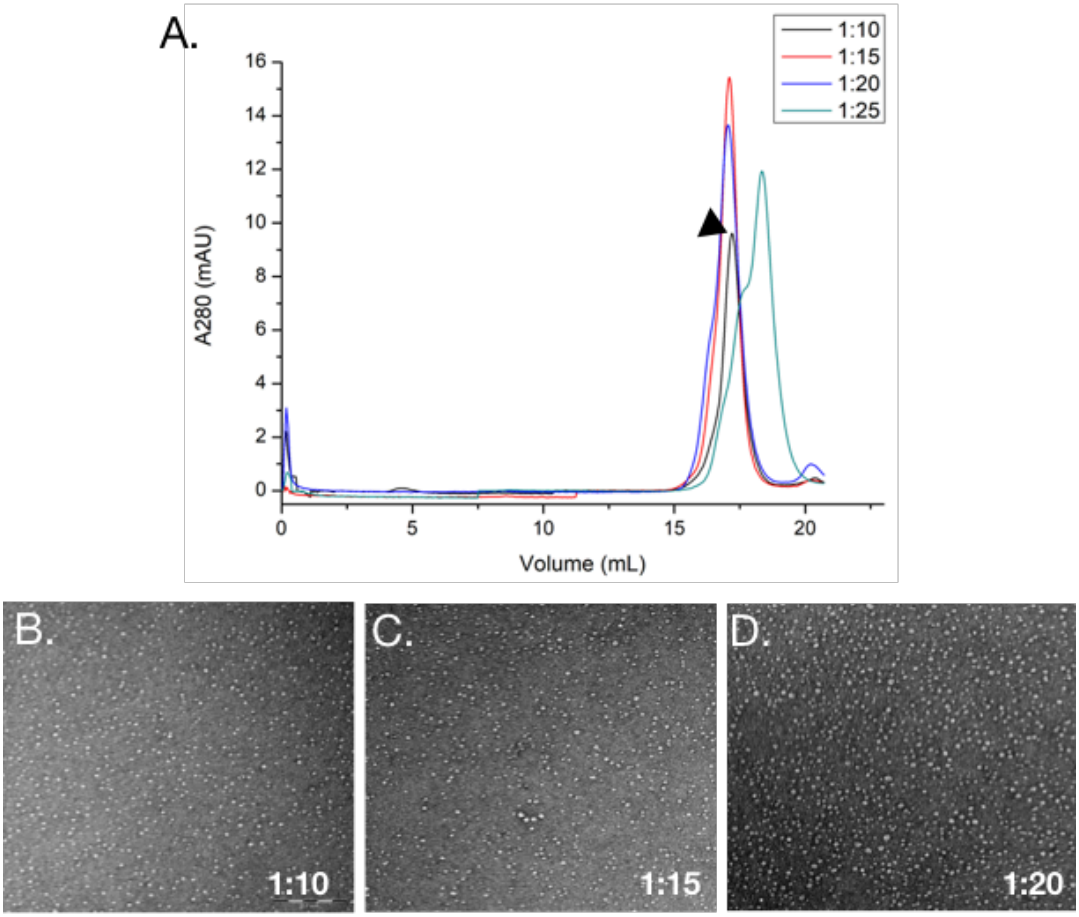

Figure 38. Optimization of empty MSP1D1dH5 NDs.
A. Comparison of size-exclusion chromatography elution profiles from Superdex 200 10/300 column for nanodiscs prepared with different MSP:lipids molar ratios. The curve representing the most monodisperse nanodiscs is marked with the black triangle. B.-D. Negative stain electron micrographs showing nanodiscs prepared with three different MSP:lipids molar ratios.

9.4. Cryoelectron microscopy

9.4.1. CD47 in complex with Fab10 in MSP1D1 nanodiscs

9.4.1.1. Sample preparation

The MSP1D1 nanodiscs were reconstituted using 1:10:200 CD47:MSP:lipids molar ratio. They were complexed with Fab10 using 2.4-fold molar excess of Fab over CD47. The complex was purified using Strep-Tactin XT and SEC on Superose 6 10/300 column (data not shown). The 0.15 mg/mL was applied on the grid, blotted for 2.5" and plunge frozen using Leica plunger.

9.4.1.2. Data collection and processing

The data was collected using Titan Krios microscope with K2 detector. The pixel size was 1.059 Å, accellerating voltage 300 kV, spherical aberration 2.7 mm and total exposure dose of 45.1 e/Å^2. The micrographs were collected from the regions containing thin ice. Part of the samples was supplemented with β-OG, which totally dissolved the nanodiscs (data not shown). This is why the sample without detergent was used for data collection. After data processing using CryoSPARC software, these regions were found to host mostly dissociated Fabs, their aggregates and possibly aggregated components of dissociated nanodiscs. The motion-corrected micrograph from this data collection, containing elongated Fab molecules is shown in Figure 39A. Several Fabs were marked with the black triangles.

Later on, during cryo-EM screening of different regions of the grid containing a sample of CD47 in MSP1D1dH5 NDs in complex with HuFab, using Titan Krios microscope with K2 detector, it was found that nanodiscs remain intact after plunge freezing only in regions of ice of medium (Figure 39C and D) and high thickness. Despite remaining intact, the nanodiscs in these regions were heavily aggregated, and a large number of them sticked to the grid edges (figure 39D). The thin ice region from the grid containing the same sample is shown in Figure 39B for comparison. Fabs are marked with the black triangles and nanodiscs and their aggregates are marked with the white triangles.

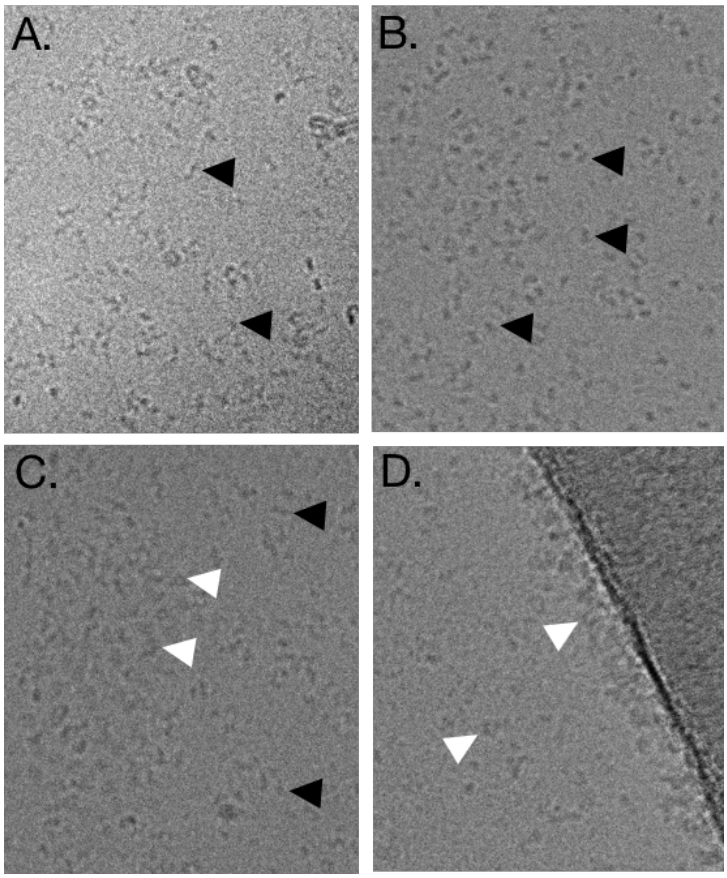

Figure 39. Samples of NDs in different grid regions.
A. Representative raw image from cryo-EM data collection from a grid containing CD47 in MSP1D1 NDs in complex with Fab10. B.-D. The cryo-EM screening images of three different regions with thin (B.) and medium-thick ice (C., D.) of a grid containing CD47 in MSP1D1dH5 NDs in complex with HuFab. Fabs are marked with the black triangles, and nanodiscs and their aggregates are marked with the white triangles.

9.4.2. CD47 in complex with HuFab-D in MSP1E3D1 nanodiscs

9.4.2.1. Sample preparation and evaluation

The MSP1E3D1 nanodiscs were reconstituted using 1:10:300 CD47:MSP:lipids molar ratio. They were complexed with HuFab-D using 2.7-fold molar excess of Fab over CD47. The complex was purified using Strep-Tactin XT and SEC on Superose 6 10/300 column.

The nanodisc complex eluted from SEC column as a monodisperse peak at approximately 15 mL (Figure 40B). The presence of target and Fab in the complex was confirmed using a western blot. CD47 was identified as a smeary band at approximately 55 kDa using anti-CD47 antibody, followed by anti-mouse IgG-HRP antibody (left panel in Figure 40C). HuFab-D was identified as a single band at approximately 30 kDa using anti-Fab-HRP antibody (right panel in Figure 40C). The nanodiscs assembly and Fab presence were additionally proved on negative stain EM image (Figure 40A). The micrograph has shown

clearly a homogenous population of round objects with a diameter of approximately 10 nm, containing the elongated extensions in their central areas, representing nanodiscs and Fabs, respectively. The nanodisc portion of the complex was marked with the blue traingle and the Fab portion with the green triangle in the magnified part of the micrograph in Figure 40A. The sample quality was very high without any significant aggregation.

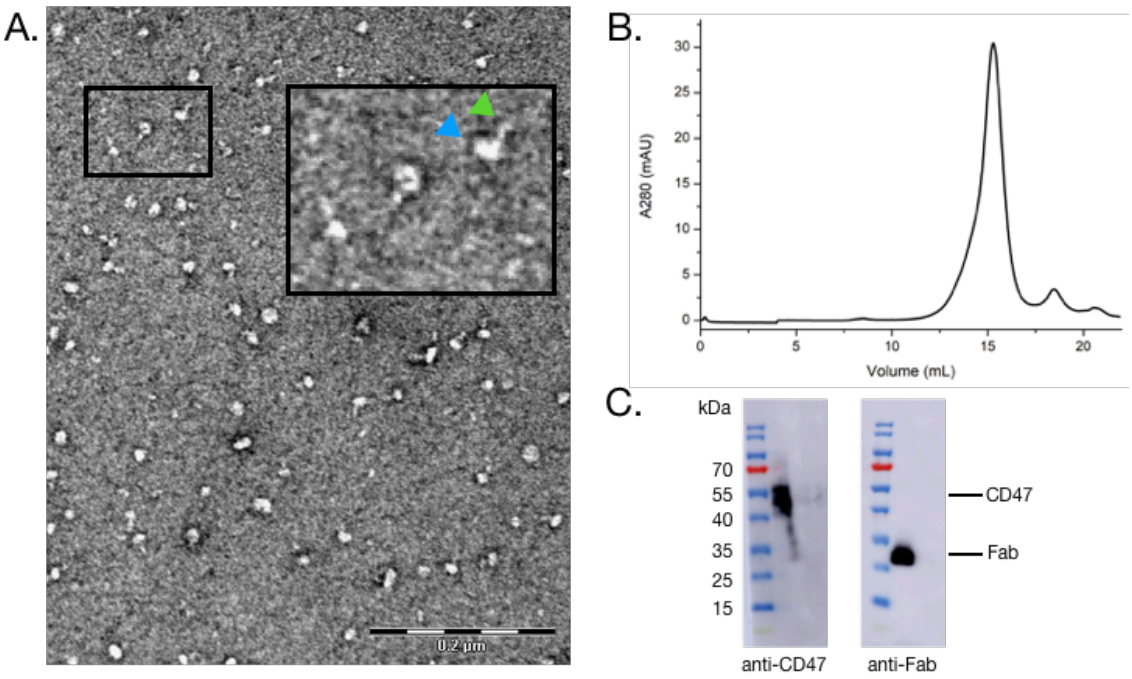

Figure 40. Quality control of MSP1E3D1 NDs containing CD47-HuFab-D complex.
A. Negative stain electron micrograph showing assembled nanodiscs. The representative region was magnified. Blue triangle marks the nanodisc and green triangle marks the Fab. B. Size-exclusion chromatography elution profile from Superose 6 10/300 column of the nanodisc sample. C. Western blot analysis of the nanodisc sample with anti-CD47 (left) and anti-Fab (right) antibodies.

The final 0.075 mg/mL sample was applied on the cryo-EM grid, blotted for 2.5" and plunge frozen using Leica plunger for cryo-EM analysis.

9.4.2.2. Data collection and processing

The data was collected using Titan Krios microscope with K2 detector. The pixel size was 1.059 Å, accellerating voltage 300 kV, spherical aberration 2.7 mm and total exposure dose of 43 e/Å^2. The data was analyzed using CryoSPARC software using C1 symmetry. The data processing flow chart is presented in Figure 41.

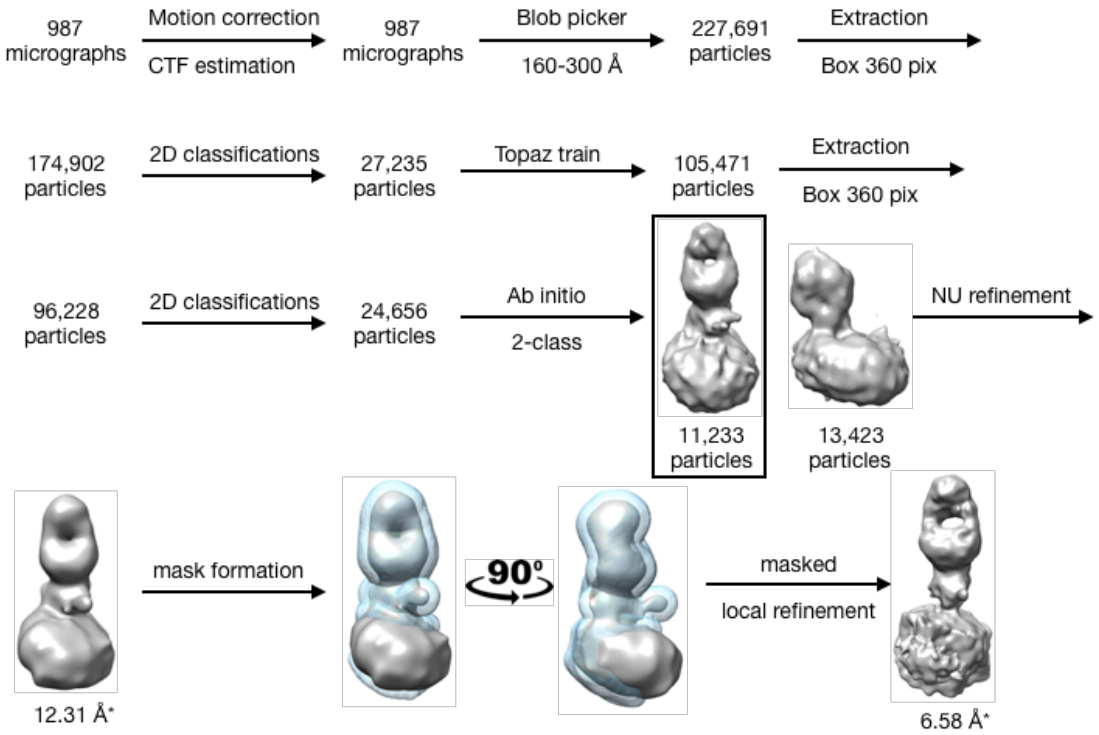

Figure 41. Cryo-EM data processing flow chart for CD47 in complex with HuFab-D in MSP1E3D1 NDs.
The data was processed using CryoSPARC software. The model selected for further processing is enclosed in the black box. Mask applied in final refinement is shown as semi-transparent cyan surface. *Most probably highly inaccurate resolutions estimated by CryoSPARC.

985 micrographs were collected in total. The motion correction was performed using MotionCor2 (Zheng et al., 2017), The CTF estimation was done using CTFFIND4 (Rohou et al., 2015). The circular blob picker with a diameter range 160-300 Å was used to pick 227,691 particles. The particles were extracted using box size of 360 pixels. Multiple 2D classifications with 40 EM iterations and batchsize per class equal 200 were performed which reduced number of particles to 27,235. The best particles were used as templates for Topaz train (Bepler et al., 2019) particle picking. After duplicates removal and another particle extraction with the same box size, 96,228 particles were found. Then another multiple 2D classifications were performed which reduced number of particles to 24,656. Afterwards, 2-class ab initio reconstitution was performed and the class containing 11,233 particles was selected (enclosed in the black box in Figure 41). The other class contained junk particles and was excluded. The successful class was refined using a non-uniform refinement (Punjani et al., 2020). The model had 12.31 Å resolution according to estimates from CryoSPARC. Then the nanodisc regions surrounding potential transmembrane domains were removed in UCSF Chimera (Pettersen et al., 2004) and the mask was formed using Volume tools in CryoSPARC with dilation radius and soft padding of 5 Å each (shown as a semi-transparent cyan surface around initial model from non-uniform refinement in Figure 41). The mask was used in local refinement and improved model to 6.58 Å resolution according to CryoSPARC

estimates. However, judging on the overall quality of the obtained models, all resolution estimates performed by CryoSPARC during this data processing are highly inaccurate.

9.4.2.3. Data quality and 3D model assessment

The representative, motion corrected micrograph image is shown in Figure 42A. It contains low density, round particles and some aggregates. The quality of the sample and particles distribution was sufficient for successful particle picking. The representative 2D classes are shown in Figure 42B. They clearly contain round nanodisc and elongated Fab on even, featureless background. They seem to contain particles in a wide variety of spacial orientations. In some classes a characteristic round space between light and heavy chains of the Fab is visible. There seem to be a slight variability of the nanodisc size between some of the classes. The Fab molecules in most of the classes are oriented 90 degrees relative to the nanodisc surface. However, there are two features in the images that may suggest some degree of flexibility of either Fab or target molecule. In some classes Fab seems to be gently tilted at different angles and the ends of all Fabs seem to be gently blurry. Additionally, a small excluded population of 2D classes contained two or even three molecules of Fab per nanodisc (Figure 42C). The resulting 3D density was of very low quality and was insufficient for model building and interpretation of secondary structure (Figure 42D). Despite that, its size and overall geometry fitted fairly well into existing models of Fab and CD47 from Protein Data Bank (PDB accession code 7MYZ; Figure E). The transmembrane helices did not provide enough signal to be resolved to even the lowest resolution and in their place there is an irregular, solid density with a thickness of a cellular membrane. Additionally, the 3D models contain unexplained artifacts on the sides of the part corresponding to CD47 extracellular domain (Figure 42D).

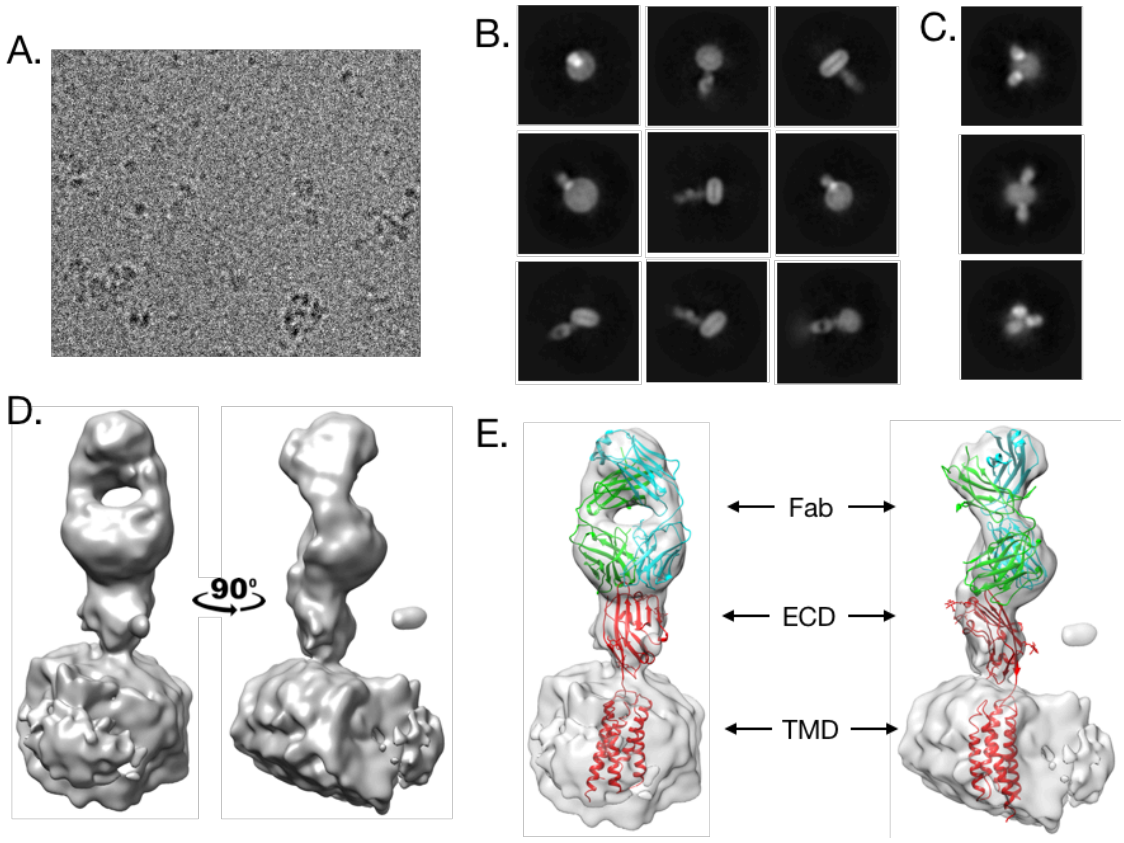

Figure 42. Cryo-EM data processing results for CD47 in complex with HuFab-D in MSP1E3D1 NDs.
A. Representative, motion corrected micrograph containing vitrified nanodisc sample. B. The representative 2D classes from the final round of 2D classification. C. Representative excluded 2D classes containing two or three CD47 molecules. D. Final low resolution model. E. CD47 (red) and Fab (heavy chain - green, light chain - cyan) structures available in PDB with accession code 7MYZ, splitted and fitted manually into the final low resolution model.

9.4.3. CD47 in complex with HuFab in MSP1D1dH5 nanodiscs

9.4.3.1. Sample preparation and evaluation

The MSP1D1dH5 nanodiscs were reconstituted using 1:10:100 CD47:MSP:lipids molar ratio. They were complexed with HuFab using 2.5-fold molar excess of HuFab over CD47. They were purified using the same method as the nanodiscs in the previous section. The nanodiscs eluted from the SEC column as a monodisperse peak at approximately 16 mL, without any significant aggregation (Figure 43A). The integrity of the complex was confirmed using a western blot. CD47 was visualized as a smeary band at approximately 55 kDa using anti-CD47 antibody, followed by anti-mouse IgG-HRP antibody (left panel in Figure 43B). The Fab migrated as a single band at 34 kDa and was identified using anti-Fab-HRP antibody (right panel in Figure 43B).

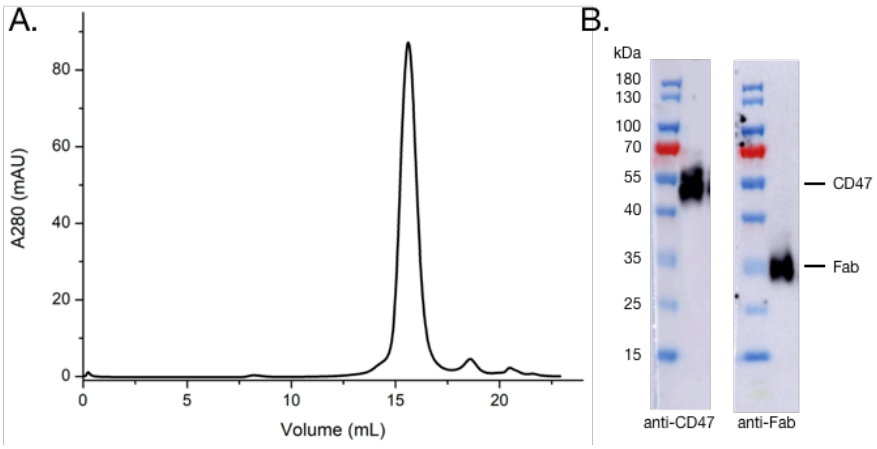

Figure 43. Quality control of MSP1D1dH5 NDs containing CD47-HuFab complex.
A. Size-exclusion chromatography elution profile from Superose 6 10/300 column of the nanodisc sample. B. Western blot analysis of the nanodisc sample with anti-CD47 (left) and anti-Fab (right) antibodies.

The final 0.18 mg/mL sample was applied on the cryo-EM grid, blotted for 2" and plunge frozen using Leica plunger for cryo-EM analysis.

9.4.3.2. Data collection and processing

The data was collected using Titan Krios microscope with K2 detector. The pixel size was 1.059 Å, accellerating voltage was 300 kV, spherical aberration was 2.7 mm and the total exposure dose was 45.6795 e/Å^2. Part of the samples was supplemented with various concentrations of fluorinated octyl maltoside (FOM). FOM is a detergent with fluoride instead of hydrogen atoms in hydrophobic tail that makes it not lipofillic. This is why its molecules do not solubilize lipid membranes, but remain hydrophobic, and were shown to prevent sample aggregation at air-water interface and grid edges (Kampjut et al., 2021). However, here, it caused an opposite effect and the aggregation of particles increased (data not shown). This is why the data was collected from the grid containing the sample without FOM. The data was analyzed using CryoSPARC software using C1 symmetry. The data processing flow chart is presented in Figure 44.

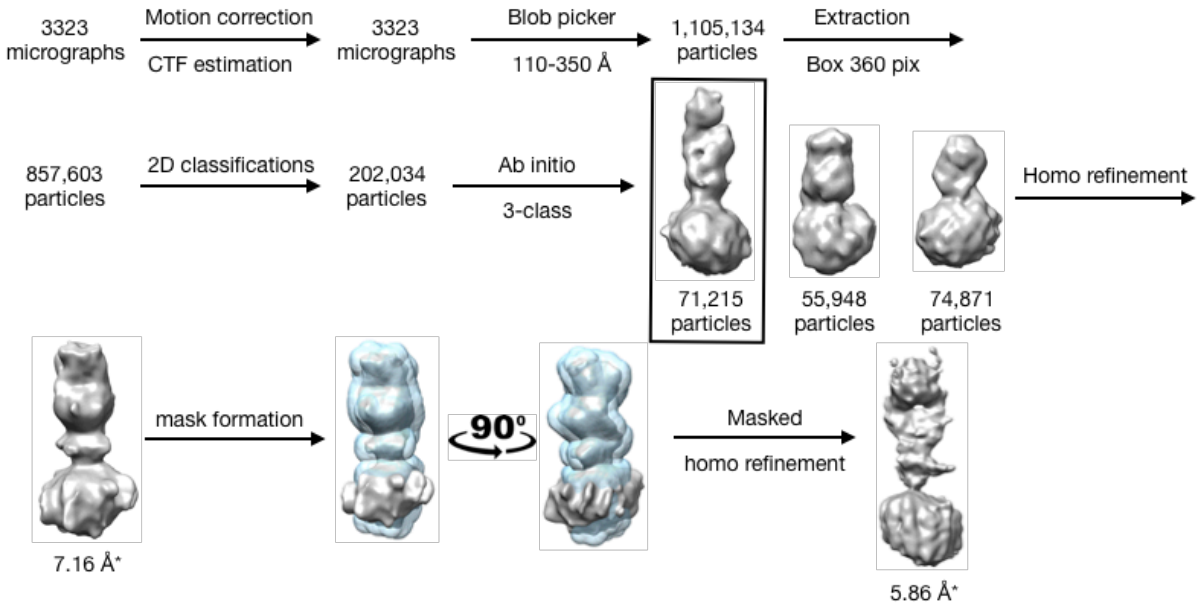

Figure 44. Cryo-EM data processing flow chart for CD47 in complex with HuFab in MSP1D1dH5 NDs.
The data was processed using CryoSPARC software. The model selected for further processing is enclosed in the black box. Mask applied in final refinement is shown as a semi-transparent cyan surface. *Most probably highly inaccurate resolutions estimated by CryoSPARC.

3323 micrographs were collected. The motion correction was performed using MotionCor2 and CTF estimation was done using Patch CTF estimation. Then the circular blob picker with diameter range of 110-350 Å was used to pick 1,105,134 particles. The particles were extracted using box size of 360 pixels. After multiple 2D classifications the amount of particles was reduced to 202,034. They were used for 3-class ab initio reconstitution. Two junk classes were excluded. The successful class contained 71,215 particles (enclosed in the black box in Figure 44). It was refined using homogenous refinement that yielded in a model with CryoSPARC estimated resolution of 7.16 Å. Then the external parts of nanodisc were removed in UCSF Chimera and the mask with dilation radius and soft padding of 5 Å each was formed using Volume tools in CryoSPARC (shown as a semi-transparent cyan surface around the initial model from the homogenous refinement in Figure 44). The mask was used in another homogenous refinement that gave rise to a model with estimated resolution of 5.86 Å. Basing on overall quality of these models the resolution estimates provided by CryoSPARC are highly inaccurate.

9.4.3.3. Data quality and 3D model assessment

The representative, motion corrected micrograph is shown in Figure 45A. It contains quite dense but pickable round particles, some aggregates and slight contamination with ethane. The representative 2D classes are shown in Figure 45B. They contain clearly visible round nanodiscs and elongated Fabs oriented at 90 degrees relative the nanodisc surface with high signal on uniform, featureless background. In some classes a characteristic space inside the

Fab is visible. All parts of particles are fairly sharp. The particles seem to be in a wide variety of spacial orientations. The model building and interpretation of secondary structure could not be performed because of not sufficient model quality (Figure 45C). However, general geometry and size are in accordance with known models of Fab and CD47 from PDB (PDB accession code 7MYZ). Only transmembrane domains could not be resolved to even the lowest resolution and in their place there is an irregular, solid density with a thickness of a cellular membrane. Additionally, the 3D models also contain unexplained artifacts on the sides of the part corresponding to CD47 extracellular domain (Figure 45C).

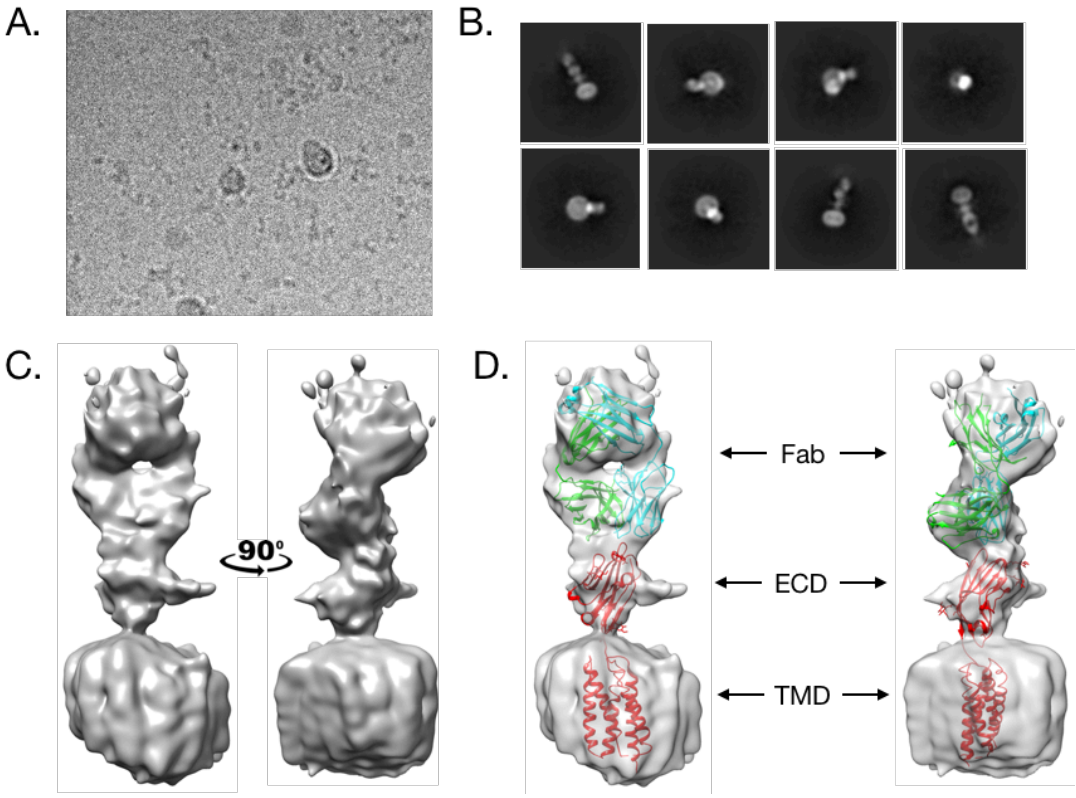

Figure 45. Cryo-EM data processing results for CD47 in complex with HuFab in MSP1D1dH5 NDs.
A. Representative, motion corrected micrograph containing vitrified nanodisc sample. B. Representative 2D classes from the final round of 2D classification. C. Final low resolution model. D. CD47 (red) and Fab (heavy chain - green, light chain - cyan) structures available in PDB with accession code 7MYZ, splitted and fitted manually into the final low resolution model.

9.4.4. CD47 in complex with HuFab-H12-P and anti-Fab NB in MSP1D1dH5 nanodiscs

9.4.4.1. Sample preparation and evaluation

The MSP1D1dH5 nanodiscs were reconstituted using 1:10:100 CD47:MSP:lipids molar ratio. They were complexed with HuFab-H12-P (HuFab with rigidified elbow region between variable and constant regions of the heavy chain) and anti-Fab nanobody, using 2.5-fold

molar excess of the Fab over CD47, and 4.5-fold molar excess of the NB over the Fab. They were purified using the same method as the nanodiscs in the previous section. They eluted from SEC column as a monodisperse peak at 16 mL without any signs of aggregation (Figure 46B). The presence of the complex components was confirmed using a SDS-PAGE and western blot. The smeary band at approximately 55 kDa for CD47 was present in SDS-PAGE gel and western blot with anti-strep-tag-HRP antibody (left and middle panel in Figure 46C, respectively). The Fab was present as a single band at approximately 28 kDa, and the nanobody migrated as a double band at approximately 20 kDa in both SDS-PAGE gel and western blot using anti-his-tag-HRP antibody (left and right panel in Figure 46C, respectively). The complex assembly was additionally confirmed using negative stain EM. Round particles with elongated part in their central regions corresponded to the nanodisc and the Fab with nanobody, respectively (marked with blue and purple triangle, respectively, in magnified part of an image in Figure 46A). The micrograph contained also a potential population of nanodiscs without the Fab attached.

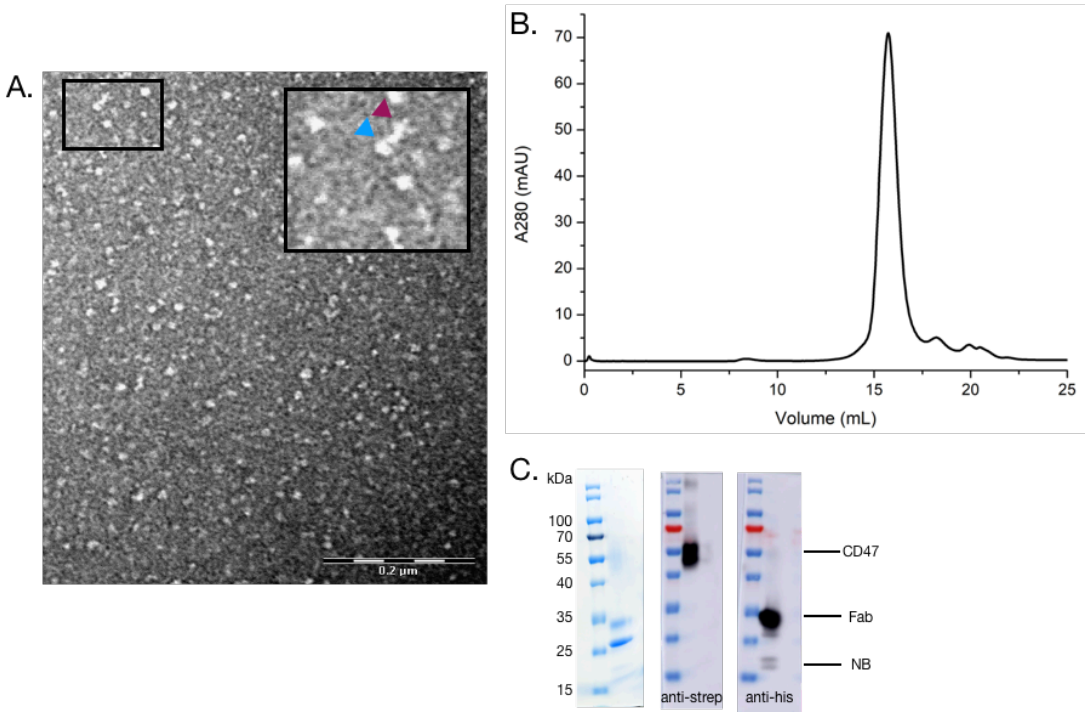

Figure 46. Quality control of MSP1D1dH5 NDs containing CD47 in complex with HuFab-H12-P and anti-Fab NB.
A. Negative stain electron micrograph showing assembled nanodiscs. The representative region was magnified. The blue triangle marks the nanodisc and the purple triangle marks the Fab-NB complex. B. Size-exclusion chromatography elution profile from Superose 6 10/300 column of the nanodisc sample. C. SDS-PAGE gel stained with SimplyBlue stain (left) and western blot analysis of the nanodisc sample with anti-strep-tag (middle) and anti-his-tag (right) antibodies.

The final 0.2 mg/mL sample was applied on the cryo-EM grid, blotted for 2.5” and plunge frozen using Leica plunger for cryo-EM analysis.

9.4.4.2. Data collection and processing

The data was collected using Titan Krios microscope with K2 detector. The pixel size was 1.059 Å, accellerating voltage 300 kV, spherical aberration 2.7 mm and total exposure dose of 46.527 e/Å^2. The data was analyzed using CryoSPARC software. The data processing flow chart is presented in Figure 47.

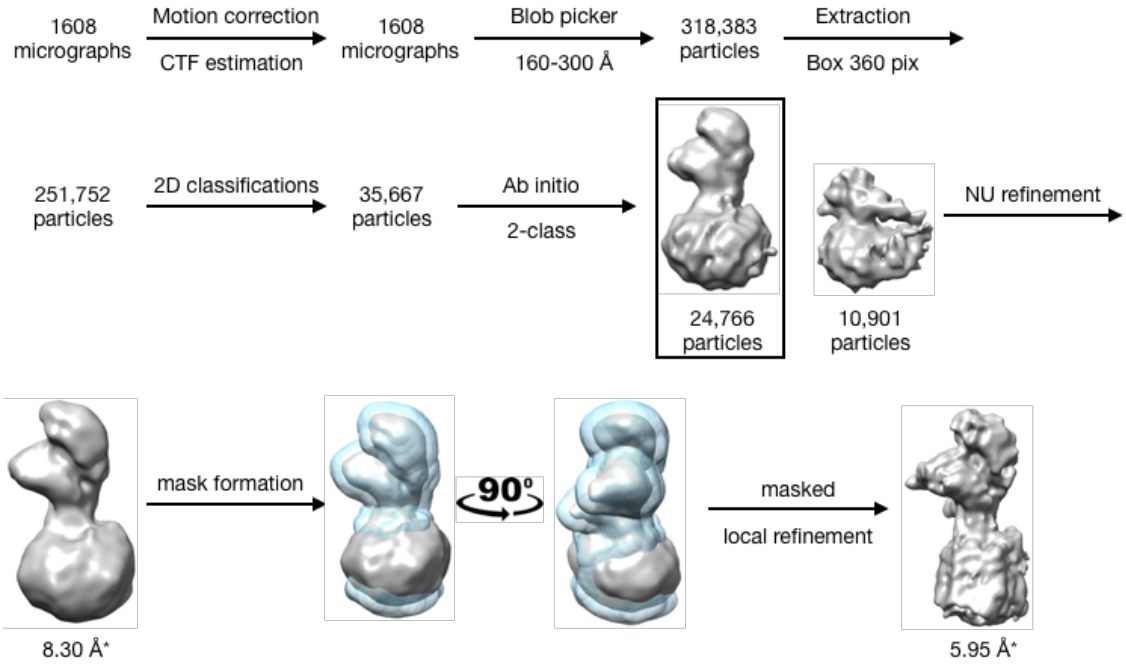

Figure 47. Cryo-EM data processing flow chart for CD47 in complex with HuFab-H12-P and anti-Fab NB in MSP1D1dH5 NDs.
The data was processed using CryoSPARC software. The model selected for further processing is enclosed in the black box. Mask applied in the final refinement is shown as semi-transparent cyan surface. *Most probably highly inaccurate resolutions estimated by CryoSPARC.

1608 micrographs were collected. Motion correction was done using MotionCor2 and the CTF estimation was performed using CTFFIND4. Round blob picker with a diameter range of 160-300 Å was used to pick 318,383 particles. The particles were extracted using box size of 360 pixels. Multiple rounds of 2D classification reduced the amount of particles to 35,667. Then 2-class ab initio reconstitution was performed and the junk model was excluded. The selected model contained 24,766 particles (enclosed in the black box in Figure 47). It was refined to estimated 8.30 Å model using non-uniform refinement. The nanodisc's edges surrounding potential transmembrane helices were removed in UCSF Chimera and the mask with dilation radius and soft padding of 5 Å was created using Volume tools in CryoSPARC (shown as a semi-transparent cyan surface around initial model from non-uniform refinement in Figure 47). The mask was used in local refinement that yielded the volume with estimated resolution of 5.95 Å. Judging on overall quality of these models the resolution estimates done by CryoSPARC are highly inaccurate.

9.4.4.3. Data quality and 3D model assessment

The representative, motion corrected micrograph is shown in Figure 48A. It contains quite dense, but pickable round particles, quite large amount of aggregates and a slight contamination with ethane (Figure 48A). The representative 2D classes are shown in Figure 48B. The classes contain clear particles with strong signal on uniform, featureless background. The particles contain a round, nanodisc portion and elongated portion, representing the Fab. In some classes, the characteristic space inside the Fab is visible. In some classes also small additional density sticking form the Fab that represents the nanobody is visible (marked with the purple triangle on the magnified 2D class in Figure 48B). The 2D classes seem to contain particles in a wide variety of spatial orientations. Attempts to use Topaz train to find more particles, for unknown reasons, yielded with 2D classes of much lower quality than shown here (data not shown). The final model was of not sufficient resolution to perform model building and analyze the secondary structure (Figure 48C). However, its size and general geometry fitted fairly well into known models of Fab and CD47 from PDB (PDB accession code 7MYZ) and nanobody from this thesis (NB1829). The transmembrane domains could not be resolved to even the lowest resolution and in their place there is an irregular, solid density with a thickness of a cellular membrane. There is also a clearly visible, unexplained additional density on the side of the fragment corresponding to CD47 ECD (Figure 48C).

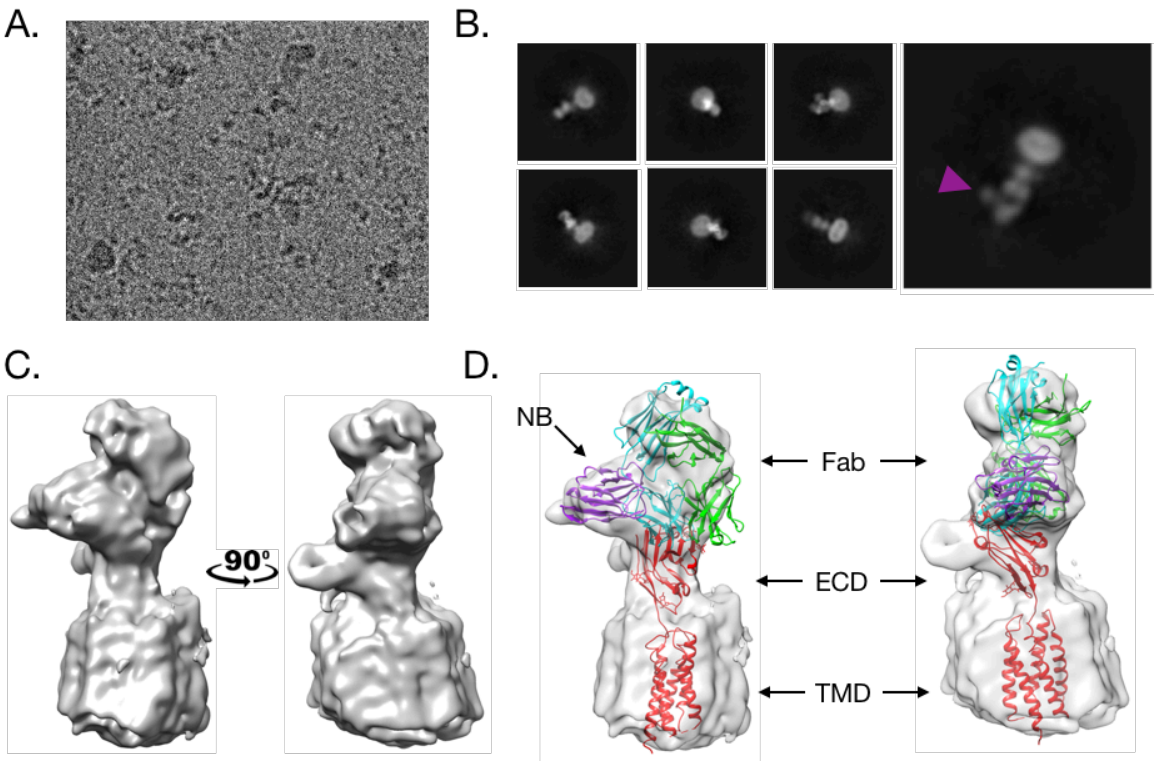

Figure 48. Cryo-EM data processing results for CD47 in complex with HuFab-H12-P and anti-Fab NB in MSP1D1dH5 NDs (previous page).
A. Representative, motion corrected micrograph containing vitrified nanodisc sample. B. Representative 2D classes from the final round of 2D classification. One of the classes, depicting the nanobody bound to the Fab is magnified. The nanobody is marked with the violet triangle. C. Final low resolution model. D. CD47 (red) and Fab (heavy chain - green, light chain - cyan) structures available in PDB with accession code 7MYZ, splitted and fitted manually, together with nanobody structure (violet) from STING-NB1829 structure, into the final low resolution model.

9.4.5. CD47 in complex with HuFab in DDM/CHS micelles

9.4.5.1. Sample preparation and evaluation

CD47 in DDM/CHS micelles was complexed with HuFab and the sample, prepared from 1.14 mg/mL CD47, was applied on the grid, blotted for 2.5", and plunged frozen using Leica plunger.

9.4.5.2. Data collection and processing

The data was collected using Titan Krios microscope with K2 detector. The pixel size was 1.046 Å, accellerating voltage was 300 kV, spherical aberration was 2.7 mm and the total exposure dose was 43.479 e/Å^2. In contrary to nanodiscs, in case of micelles, the useful data could be collected from the thin ice regions of the grid. The data was processed using CryoSPARC software. The data processing flow chart is presented in Figure 49.

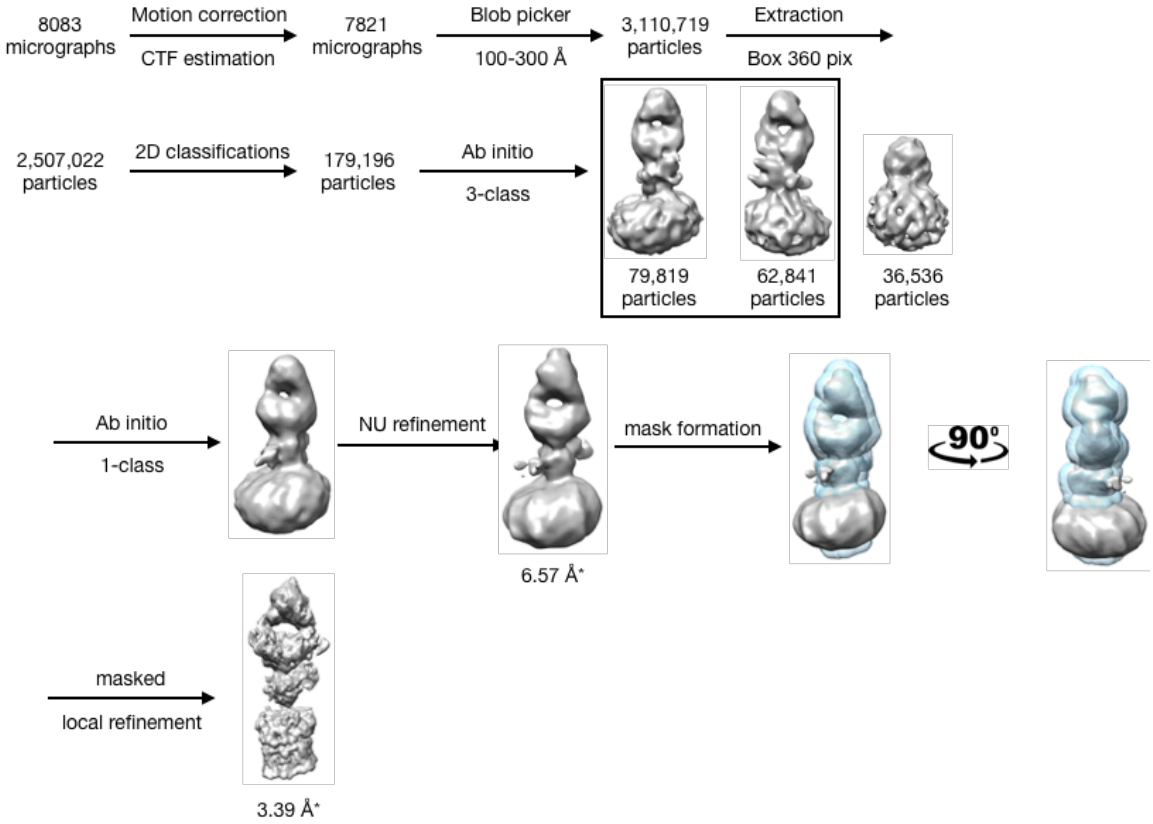

Figure 49. Cryo-EM data processing flow chart for CD47 in complex with HuFab in DDM/CHS micelles (previous page).
The data was processed using CryoSPARC software. The models selected for further processing are enclosed in the black box. The mask applied in the final refinement is shown as a semi-transparent cyan surface. *Surely highly inaccurate resolutions estimated by CryoSPARC.

8083 micrographs were collected. Motion correction and CTF estimation was performed using MotionCor2 and CTFFIND4, respectively. Circular blob picker with the diameter range of 100-300 Å was used to pick 3,110,719 particles. The particles were extracted using box size of 360 pixels. After multiple 2D classifications the amount of particles was reduced to 179,196. They were used for 3-class ab initio reconstitution. The particles from the two classes, resembling a proper complex (enclosed in the black box in Figure 49) were merged using 1-class ab initio reconstitution, containing in total 142,660 particles. The model was refined using non-uniform refinement to estimated resolution of 6.57 Å. The side edges of nanodisc surrounding potential transmembrane helices were removed, along with artifacts on the sides of the region corresponding to CD47 ECD, using UCSF Chimera. The model prepared in this way was used for mask formation with dilation radius and soft padding of 5 Å each, using Volume tools in CryoSPARC (shown as a semi-transparent cyan surface around initial model from non-uniform refinement in Figure 49). This mask was used for local refinement which gave rise to the model of estimated resolution of 3.39 Å. Basing on the overall model quality, the resolutions calculated by CryoSPARC are highly inaccurate.

9.4.5.3. Data quality and 3D model assessment

The representative, motion corrected micrograph is shown in Figure 50A. It contains quite dense but pickable round particles without a significant aggregation. The representative 2D classes are shown in Figure 50B. They contain clear particles with a round ND portion and elongated region representing the Fab. The background is uniform and featureless. The classes seem to contain particles in a wide range of spatial orientations. The densities representing transmembrane helices inside the nanodiscs are visible for the first time. In some classes the characteristic space inside the Fab is present. In most of the classes, the Fab is oriented 90 degrees relative to the surface of nanodisc. However, there are also some, in which it is gently tilted. The end of the Fab is blurry and contains some additional density in most of the classes. The final model quality was too low for model building and determination of secondary structure (Figure 50C). However, the size and overall geometry fits fairly well into known models of Fab and CD47 from PDB (PDB accession code 7MYZ). Only the transmembrane helices could not be resolved even to a lowest resolution and they constitute an irregular, solid density with a thickness of a cellular membrane. The initial model contained unexplained, additional densities on the sides of the fragement corresponding to CD47 ECD, which were removed by masking. Similar artifacts were present in all other models presented here.

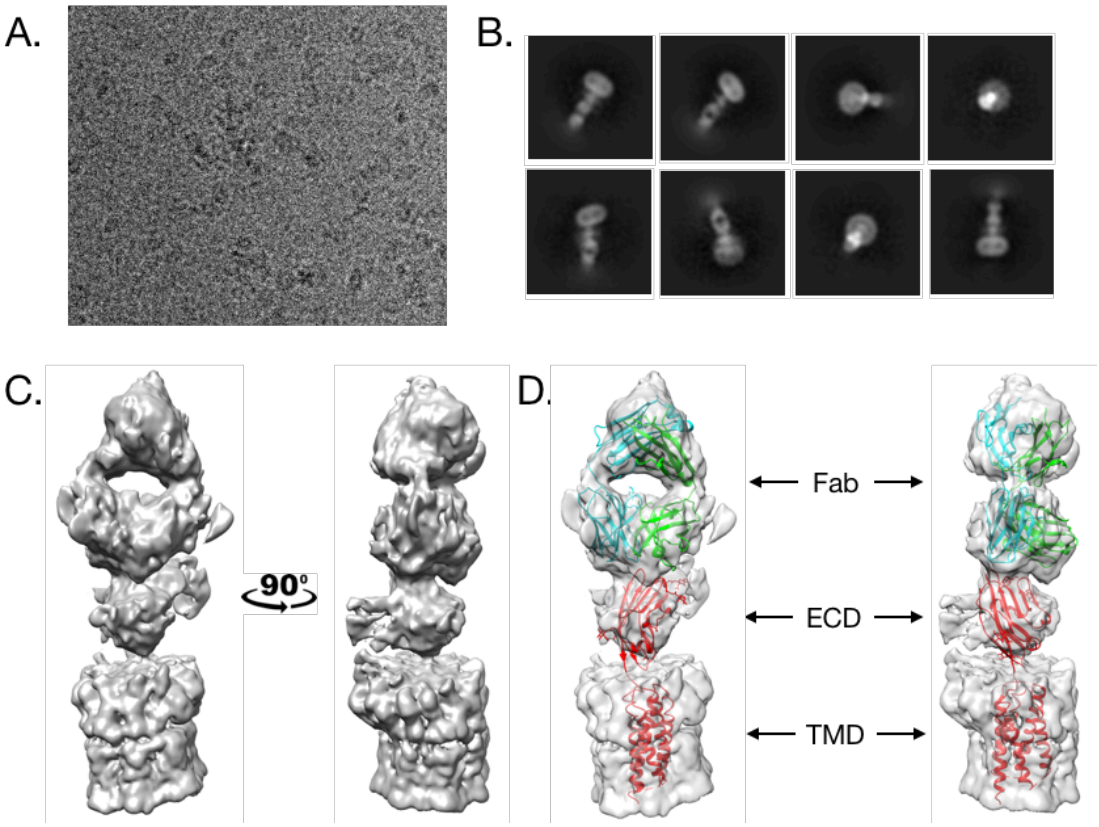

Figure 50. Cryo-EM data processing results for CD47 in complex with HuFab in DDM/CHS micelles.

A. Representative, motion corrected micrograph containing vitrified protein sample. B. Representative 2D classes from the final round of 2D classification. C. Final low resolution model. D. CD47 (red) and Fab (heavy chain - green, light chain - cyan) structures available in PDB with accession code 7MYZ, splitted and fitted manually into the final low resolution model.

10. Discussion

10.1. STING-activating nanobody

Florian Schmidt and Jennifer Deborah Wuerth used alpaca immunization and phage display to generate a plethora of nanobodies specific to the soluble domain of human STING. Then they used human monocytes cell line THP-1, expressing STING, which was modified to be able to express nanobodies with HA tag upon doxycycline induction, to characterize them. After induction they observed a slight increase of interferon beta and a significant increase of CXCL10, one of the interferon stimulated genes, in the cell supernatants in case of one nanobody, called NB1829. This showed that this nanobody is able to gently activate STING. This was further confirmed with another experiment in which they used the same cell line, modified to be able to express EGFP-tagged nanobodies upon doxycycline induction. After induction, the cells were stained with anti-IRF3 antibody and Hoechst stain to visualize IRF3 and nuclei, respectively. In this assay it was proved that the nanobody presence inside the cell causes accumulation of IRF3 inside the nucleus, which happens usually after STING activation (data not published).

10.2. How does NB1829 activate STING?

One of the purposes of this thesis was to explain how NB1829 is able to activate STING. This is why their complex was crystallized and mapped using X-ray crystallography. Despite that, the precise answer to this question still remains a mystery. The high-resolution crystal structure contained a soluble butterfly-shaped STING dimer with two nanobodies on top of the wings at C-terminal surface. To our surprise, the nanobody caused only a slight conformational change with R.M.S.D. equal to 1.108 Å, for 308 atoms when aligned with the structure of STING in its open conformation (PDB accession code: 4F5W) (Shang et al., 2012). Upon nanobody binding, the dimer still remained in an open conformation. By comparing the distances between Cα atoms of Gln184 residues that are located at the top of the wings of STING, it was found that the angle between the monomers was even slightly bigger than in the apo conformation. The distances were 59.1 Å and 56.5 Å, respectively (see Figure 27 in Section 8.3.5.). The nanobody did not cause the butterfly-shaped structure to close its wings as it happens during an interaction with the mammalian ligand 2’,3’-cGAMP (Shi et al., 2015). This is why, one could only speculate on how the nanobody is able to activate its target. Only a high-resolution, full-length protein structure could perhaps provide us with enough information to determine the detailed activation mechanism, since CBD alone does not contain most of the parts critical for activation. They include the C-terminal tail where the partner proteins bind to perform downstream signaling, and transmembrane domain without which the protein oligomerization is not possible.

However, the lack of significant conformational change, which is normally induced upon 2',3'-cGAMP binding, does not exclude activation of STING. Multiple researchers in early

studies on STING have shown that other cyclic dinucleotides, for example, cyclic diguanylate monophosphate (c-di-GMP) are able to bind to and activate STING while keeping it in almost an open conformation without any significant conformational rearrangements. For c-di-GMP, RMSD between protein with and without ligand varies between 0.5 Å and 1.2 Å (Ouyang et al., 2012, Yin et al., 2012, Shang et al., 2012 and Shu et al., 2012). For my structure, the conformational change was similar. The comparison of distances between Cα atoms of Gln184 between NB1829- and c-di-GMP-bound structures (PDB accession code 4EF4) (Ouyang et al., 2012) is shown in Figure 51. They were: 59.1 Å for the former and 56.2 Å for the latter.

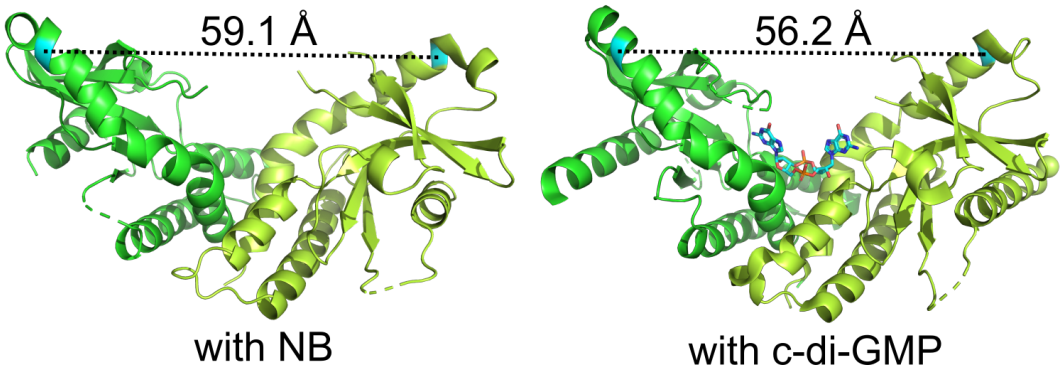

Figure 51. STING with NB vs. STING with c-di-GMP
STING CBD monomers are shown as a cartoon in green and limon, respectively. Gln184 residues are shown in cyan. c-di-GMP is shown as sticks (carbon - cyan, nitrogen - blue, oxygen - red, phosphorus - orange). The water molecules were removed for clarity. The distances are shown between Cα atoms of Gln184 (amino acid numbering corresponds to the full-length protein). The model of STING with c-di-GMP has PDB accession code 4EF4.

There were multiple speculations on how c-di-GMP can activate STING that may also apply to NB1829. Ouyang et al. have found that c-di-GMP and c-di-AMP binding to STING dimer increases binding affinity of TBK1 to this protein. However, they did not explain how this property could lead to increased type I interferon production upon ligand binding (Ouyang et al., 2012). Yin et al. suggested that STING is autoinhibited by its own C-terminal tail sticking to the CBD and that c-di-GMP binding can disrupt this interaction, and this is how it can activate STING. This was done by mutational and intracellular analysis where STING with and without CTT with FLAG-tag was co-transfected together with free HA-tagged CTT into HEK293 cells. Co-immunoprecipitation experiments have shown that isolated CTT was able to bind to STING and this binding was decreased in presence of c-di-GMP. They also predicted that protein oligomerization required for activation may be mediated by released CTTs (Yin et al., 2012). However, Shu et al. have found that c-di-GMP binding to soluble domain of STING was not able to induce phosphorylation and activation of IRF3 by TBK1. They concluded that this process may require unknown partner proteins or membrane domain of STING (Shu et al., 2012). It is worth noted that all above observations and speculations were made before the full-length structure of this protein was known, and the novel mechanism of activation was proposed. It happened in 2019, when the medium-resolution,

full-length cryo-EM structures of human apo and chicken ligand-bound STING were solved. The authors showed that STING may be activated by 180 degrees turn made by CTD relative to transmembrane part that leads to protein oligomerization (Shang et al., 2019). The same group also confirmed that in an inactive state, STING CTTs, which are critical for partner proteins binding and downstream signaling, remain tightly attached to protein surface keeping it in an autoinhibited state. Upon ligand biding, the CTTs get released, and the protein becomes ready for activation (Zhang et al., 2019). The speculation is that NB1829 can anyhow impose such changes. It could, for example, locally pull CTT from its resting position, which is located in the proximity of nanobody binding site. More difficult would be to find an explanation on how the nanobody could cause the turn of the whole molecule. Another theory would be that nanobody binding itself causes oligomerization of STING thanks to nanobody-nanobody interactions. However, this does not occur in case of isolated soluble domain, since the elution peaks from size-exclusion chromatography were monodisperse so they cannot contain higher order oligomers. But, what if, a gentle interaction between nanobody molecules that is too weak to be seen on chromatograms from SEC, and proximity of transmembrane domains in full-length STING molecules could anyhow initiate and amplify oligomerization? It is also not excluded that other proteins may take part in such activation mechanism. For instance, the nanobody could destabilize STING interraction with inhibiting STIM1 protein. However, this protein resides in the membrane which is far away from nanobody binding site. Could the slight conformational change, caused by the nanobody, detach STIM1 and activate STING? These questions require further research to be answered, and this endeavour can be very challenging.

What is interesting, it was also shown that NB1829 is able to bind to STING which is already coupled with 2',3'-cGAMP. In the ligandless crystal structure presented here, the nanobody molecules on top of STING wings are very close to each other and after ligand binding and formation of closed conformation they could simply clash. What if in such case only one nanobody molecule binds to the STING dimer? To explain it, another crystal structure would be needed.

10.3. Why further research on STING activation mechanism may be challenging?

First of all, because of a small size of STING (approximately 80 kDa) that still remains an obstacle for cryo-EM. Secondly, because of a potential oligomerization event that may occur after nanobody binding, and may cause difficulties with sample preparation and image processing. And lastly, because such activation may require some partner proteins without which it cannot be observed.

To have a first insight into oligomerization, it would be worth to perform multi angle light scattering (MALS) on STING soluble domain-nanobody and full-length protein-nanobody complexes, in order to determine if they are able to form oligomers. However, in aquous solution where the protein is isolated in micelles or nanodiscs, oligomerization that requires

transmembrane domains may not occur because of lack of continuous membrane that would enable proper interaction between them. Another idea would be to use liposomes or very large nanodiscs where multiple STING molecules could be accommodated, and analyze them using cryo-EM.

10.4. Potential applications of STING-activating nanobody

STING-activating nanobody could find multiple uses in research and medicine. Because it can be easily expressed and tracked inside the living cells, it could be applied in assays leading to better understanding of cGAS-STING pathway activation and function. In medicine, one day, it could be used as a drug that by activation of type I interferon signaling could lead to innate immune response in, for example, cancer or pathogen-infected cells. The technology used in mRNA vaccines against SARS-COVID-2 virus could, for instance, be used as delivery means (Chaudhary et al., 2021). NB1829-encoding mRNA could also be used similarly to cyclic dinucleotides as an adjuvant in future vaccines by boosting immune response to selected antigens.

10.5. Challenges in structural research on CD47

CD47, the marker of self, is the protein of many important, not fully understood functions. Despite multiple trials and various experimental strategies my attempts to obtain medium- or high-resolution structure of this protein using cryo-EM did not end with full success. In 2021 the first full-length structure of this protein was solved but it still did not provide an explanation on how it actually functions (Fenalti et al., 2021). Further structures with partner proteins, of which some may still not be known, are required to decipher this difficult mystery. The additional problem with research on CD47 is the fact that currently there are no developed *in vitro* assays to study and monitor activity of this protein. This causes issues even with such simple thing as quality control. We cannot answer an important question: Is our protein active and in its native conformation in our scientific setup?

10.6. Challenging amphipathic nature of membrane proteins

The first challenge with CD47 is its amphipathic nature and providing it with proper near-native membrane environment to keep it stable in aquous solution after extraction from a membrane. I solved it by using mild, non-ionic detergent DDM supplemented with CHS for purification, and nanodiscs with natural lipids for structural experiments. CHS was shown to successfully imitate cholesterol (Kulig et al., 2014) that is important for this protein function (Green et al., 1999). There are also some authors that claim that CHS is not a good imitator of cholesterol (Augustyn et al., 2019). However, currently, there is no other choice of more satisfactorily, water soluble cholesterol alternative, and the protein tested here seemed to behave very well in presence of CHS. In such detergent solution, CD47 elutes as a single, monodisperse peak and its melting temperature is relatively high, suggesting pretty good

stability. However, because detergents can sometimes destabilize membrane proteins, I started my journey with its structural studies in nanodiscs, which were previously proved successful to facilitate structure determination of multiple proteins (Baker et al., 2021, Gao et al., 2016, Zhang et al., 2021a).

10.7. Nanodiscs as troublesome, but useful membrane protein carriers

The nanodiscs, despite their previous successful history, turned out to be a challenge themselves. Initially, I tried to assemble nanodiscs using porcine brain polar lipid extract (Anatrace), because of its high similarity to average composition of human membranes (van Meer et al., 2011, avantilipids.com). For some reasons, numerous optimization attempts did not result in satisfactorily, single, monodisperse peak on size-exclusion chomatography using these lipids. Afterwards, I successfully optimized nanodiscs with soybean polar lipid extract, which also has a lipid content similar to the one in human cells (van Meer et al., 2011, avantilipids.com). It has been used previously to obtain multiple structures of membrane proteins, such as transient receptor potential cation channel subfamily V member 1 (TRPV1) (Gao et al., 2016) and calcium homeostasis modulators 1 and 2 (CALHM1 and CALHM2) (Syrjanen et al., 2020). One of the disadvantages of nanodiscs is the fact that they are too fragile to be frozen so they need to be prepared freshly every time, at least for structural studies that require high-quality samples. However, it is worth noted that usually there is a possibility to freeze them in presence of cryoprotectant, such as glycerol or sucrose, for standard biochemical assays. Furthermore, multiple preparation and purification steps required to incorporate a membrane protein into nanodiscs, consume a lot of protein which has a limited availability. This is caused because of low expression levels in mammalian cells. This fact additionally increases effort and expense.

My nanodiscs also did not behave well on the grids used for cryo-EM. The best regions of grids for data collection, containing thin ice, contained only aggregated Fab molecules and possibly diassembled, aggregated parts of nanodiscs. The nanodiscs preferred to localize to ice of medium and high thickness. Also, they had a tendency to aggregate and stick to the edges of the holes of the grids. The availability of such regions is limited and the thick ice increases noise levels and decreases data quality that can limit a resolution of resulting models. To overcome this, I tried to increase the concentration of nanodiscs to push them into the better grid regions, but it only caused more aggregation, and presence of more disassembled parts that could make particle picking impossible. Therefore, the data must have been collected from samples with relatively low concentrations that limited the amount of available, high quality particles. Other strategy to at least partially solve this problem, was to use the grids with bigger holes to lengthen the ice gradient, however, it brought questionable improvement. Shortening the blotting time also did not bring satisfactorily ice quality.

10.8. Attempts to obtain structure of CD47

In the initial attempt with MSP1D1 nanodiscs, the data was collected from the thin ice only areas, and it did not contain nanodiscs at all. At this point we were not aware that nanodiscs reside in thicker ice, because it was difficult to recognize such small particles in low contrast, raw images. So the initial assumption was that they disassembled because of too small nanodiscs size. This is why I opted for bigger MSP1E3D1 nanodiscs that, this time, turned out to be too large. Except for nanodiscs with a single CD47 molecule, the population of particles containing two or three CD47 molecules per nanodisc, which had to be excluded from the data processing, was present. Too big diameter of nanodiscs and the possibility of too excessive movement of the protein across the nanodisc could also cause problems with proper particle alignement. In the end, I started to use the smallest nanodiscs that I had available, prepared from MSP1D1dH5. Despite successful assembly, they behaved much worse on the grids than their bigger counterparts. They produced much more disassembled parts and they had higher aggregation levels. Usually, detergents are added to the samples to prevent aggregation and sticking to air-water interface. However, in case of nanodiscs, standard detergent, β-OG that is used for this purpose could not have been applied, because of its ability to solubilize membranes, and thus also nanodiscs (Kampjut et al., 2021). The attempt to add this detergent even in small quantities ended up in complete lack of particles in vitrous ice. However, there is in existence a special detergent that is safe for nanodiscs, called FOM. It has fluoride atoms instead of hydrogen atoms in its hydrophobic tail. This modification causes that it remains hydrophobic while it is not lipophillic anymore, and therefore does not dissolve lipid membranes (Efremov et al., 2017). Nevertheless, my attempts to use this detergent at various concentrations (0.025-0.1%) caused even more aggregation. In the end, I started to use samples in detergent micelles that solved many problems. High aggregation and complex disassembly issues were still present, but the distribution of molecules in ice was more or less equal across each hole, and even the thinnest ice contained particles that enabled collection of higher quality images with more useful particles available.

10.9. Affinity binders as mass boosters and fiducial marks

Other major challenges with CD47 were its small size (approximately 50 kDa) and the fact that it is relatively featureless. Furthermore, part of it is embeded into a membrane or detergent micelle, leaving only the small extracellular domain, for which a proper particle alignement by currently available data processing software is practically impossible.

To solve these obstacles, I attempted to use anti-CD47 Fab derived from high-affinity Hu5F9 antibody (Liu et al., 2015). The Fab adds approximately 50 kDa of mass and acts as a fiducial mark for proper particles alignement. However, this Fab turned out to have some significant disadvantages. First, was that despite its high affinity to the target, a noticable percent of it was disassembling from the complex during plunging. The second, more important issue, was the fact that it may be flexible to some extent. Some degree of flexiblity is natural for Fabs

and its extent differs among them. It occurs mostly in the region that links variable and constant parts, called the elbow (Bailey et al., 2018). Some of my 2D classes contained blurry densities at the top of the Fab molecules that may suggest alignement problems and flexibility. I tried to overcome this issue by using elbow region modification, shown to rigidify it and improve crystallization results (Bailey et al., 2018). Additionally, I applied the anti-Fab nanobody that binds between variable and constant parts of human kappa light chain and may additionally stabilize the Fab (Ereño-Orbea et al., 2018). This nanobody was successfully used in structure determination of human Frizzled5 (Tsutsumi et al., 2020). Unfortunately, this approach did not increase my model's quality and resolution anyhow.

In my work, I also attempted to prepare another fiducial mark - a megabody. For this purpose, I tried to transplant CDRs of anti-CD47 nanobody published by Ma et al. (Ma et al., 2020) into the megabody c7HopQ developed by Uchański et al. (Uchański et al., 2021). Unfortunately, such protein turned out to be insoluble and the production was not possible (data not shown).

10.10. Overcoming inherent flexibility of CD47 may be the key

Because affinity binders did not improve the data quality, the biggest challenge in case of CD47 may not be the Fab, but its own high levels of inherent flexibility. In 2021 the medium resolution structure of full-length CD47 was solved with assistance of B6H12 Fab. However, such challenging endeavor required completely different experimental approach and technique. I assumed that the authors, initially have also did not achieve sufficient data quality in attempts to solve the structure of this protein using cryo-EM. In their publication they presented only an example of low resolution 2D class of CD47 in complex with the different Fab and low-resolution 3D model that fits to known structures available in PDB fairly well. To solve the final structure, they used other method - the lipid cubic phase crystallization (Fenalti et al., 2021). To make protein suitable for this purpose and to boost the resolution of the model, they modified CD47 by adding a small, rigid and stable, modified *E. coli* apocytochrome b562RIL (M7W, H102I, and R106L), called "BRIL", into intracellular loop 1 (Chu et al., 2002, Fenalti et al., 2021). This was done to rigidify the protein and to provide additional crystal contacts for proper crystal lattice formation (Fenalti et al., 2021). BRIL was used previously by multiple research groups to assist structural determination of various challenging proteins (Chun et al., 2012, Mukherjee et al., 2020, Tsutsumi et al., 2020). These results also suggest that inherent flexibility was a major obstacle on the way to medium resolution structure of CD47.

10.11. What were the advantages of my research approach over Fenalti et al.'s?

I believe that I was very close to success, because my 2D classes of CD47 were of far higher quality then the one published by Fenalti at al. Similarly, I was able to present the series of

low resolution models that fitted quite well by size and shape to previous structures deposited in PDB. My advantage was that I expressed this protein in human cells, while Fenalti et al. needed to use insect cells for this purpose, because of much higher demand for protein needed for crystallization trials (Fenalti et al., 2021). CD47 is a highly glycosylated protein and when it is produced in insect cells, the glycosylation pattern will differ significantly from the human one, and thus the final structure may not fully reflect the actual physiological state of this protein in human cells (Shi & Jarvis, 2007). Additionally, my protein did not contain any internal modifications, except for C-terminal tag. Fenalti et al. fused their protein with BRIL and this may possibly affect its natural conformation to unpredictable extent. However, of course, without such tools, solving structures of many proteins using current technology may not be possible.

10.12. If I would start this project again today, I would...

If I would start this project again today, I would implement some changes in the methodology and try a couple of alternative approaches. First of all, I would screen more Fabs to find a potentially more suitable one with lower flexibility and better affinity. I would also try to use different elbow region modifications published by Bailey et al. (Bailey et al., 2018). Because CD47 is a very small protein, I would use even smaller nanodiscs, applying MSP1dH4-H6. Such nanodiscs have an average diameter of 4 nm. It could further decrease variability resulting from potential movements of target molecules inside the nanodiscs (Hagn et al., 2013). I would also try other membrane-mimicking systems such as Salipro (Frauenfeld et al., 2016), amphipols (Tribet et al., 1996) and styrene maleic-acid lipid particles (SMALPs) (Hardy et al., 2018). It would be also beneficial to screen more detergents to find the one that would bring better data quality. I would also screen more freezing conditions for nanodiscs to anyhow cause better particle distribution and less aggregation, possibly by using different fluorinated detergent or other additives (e.g. amphipols). It would be also useful to attempt to use anti-MSP nanobody that could be used as a fiducial mark for the nanodisc itself. For sure it would be a very flexible connection, but probably it would still improve particle alignement to some extent (Uchański et al., 2021). I would also attempt to solve CD47 structure with a partner protein, for instance, an integrin. On one hand, it would boost the mass and add more features, but on the other, adding complexity could increase flexibility, and result in additional problems with sample preparation and data processing. However, for sure, it would provide us with more useful biological information on how does CD47 function than the apo structure. But this is only the beginning...

11. References

Abe, T., & Barber, G. N. (2014). Cytosolic-DNA-mediated, STING-dependent proinflammatory gene induction necessitates canonical NF-κB activation through TBK1. *Journal of virology*, *88*(10), 5328–5341. https://doi.org/10.1128/JVI.00037-14

Ablasser, A., Schmid-Burgk, J. L., Hemmerling, I., Horvath, G. L., Schmidt, T., Latz, E., & Hornung, V. (2013). Cell intrinsic immunity spreads to bystander cells via the intercellular transfer of cGAMP. *Nature*, *503*(7477), 530–534. https://doi.org/10.1038/nature12640

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 2), 213–221. https://doi.org/10.1107/S0907444909052925

Aguayo-Ortiz, R., Creech, J., Jiménez-Vázquez, E. N., Guerrero-Serna, G., Wang, N., da Rocha, A. M., Herron, T. J., & Espinoza-Fonseca, L. M. (2021). A multiscale approach for bridging the gap between potency, efficacy, and safety of small molecules directed at membrane proteins. *Scientific reports*, *11*(1), 16580. https://doi.org/10.1038/s41598-021-96217-7

Ahn, J., Gutman, D., Saijo, S., & Barber, G. N. (2012). STING manifests self DNA-dependent inflammatory disease. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(47), 19386–19391. https://doi.org/10.1073/pnas.1215006109

Al-Mayouf, S. M., Sunker, A., Abdwani, R., Abrawi, S. A., Almurshedi, F., Alhashmi, N., Al Sonbul, A., Sewairi, W., Qari, A., Abdallah, E., Al-Owain, M., Al Motywee, S., Al-Rayes, H., Hashem, M., Khalak, H., Al-Jebali, L., & Alkuraya, F. S. (2011). Loss-of-function variant in DNASE1L3 causes a familial form of systemic lupus erythematosus. *Nature genetics*, *43*(12), 1186–1188. https://doi.org/10.1038/ng.975

Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002. Membrane Proteins. Available from: https://www.ncbi.nlm.nih.gov/books/NBK26878/

Aline Dias da, P., Nathalia Marins de, A., Gabriel Guarany de, A., Robson Francisco de, S., & Cristiane Rodrigues, G. (2020). The World of Cyclic Dinucleotides in Bacterial Behavior. *Molecules (Basel, Switzerland)*, *25*(10), 2462. https://doi.org/10.3390/molecules25102462

Allen, K. N., Entova, S., Ray, L. C., & Imperiali, B. (2019). Monotopic Membrane Proteins Join the Fold. *Trends in biochemical sciences*, *44*(1), 7–20. https://doi.org/10.1016/j.tibs.2018.09.013

Andreeva, L., Hiller, B., Kostrewa, D., Lässig, C., de Oliveira Mann, C. C., Jan Drexler, D., Maiser, A., Gaidt, M., Leonhardt, H., Hornung, V., & Hopfner, K. P. (2017). cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein-DNA ladders. *Nature*, *549*(7672), 394–398. https://doi.org/10.1038/nature23890

Augustyn, B., Stepien, P., Poojari, C., Mobarak, E., Polit, A., Wisniewska-Becker, A., & Róg, T. (2019). Cholesteryl Hemisuccinate Is Not a Good Replacement for Cholesterol in Lipid Nanodiscs. *The journal of physical chemistry. B*, *123*(46), 9839–9845. https://doi.org/10.1021/acs.jpcb.9b07853

Baechler, E. C., Batliwalla, F. M., Karypis, G., Gaffney, P. M., Ortmann, W. A., Espe, K. J., Shark, K. B., Grande, W. J., Hughes, K. M., Kapur, V., Gregersen, P. K., & Behrens, T. W. (2003). Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proceedings of the National*

Academy of Sciences of the United States of America, *100*(5), 2610–2615. https://doi.org/10.1073/pnas.0337679100

Bailey, L. J., Sheehy, K. M., Dominik, P. K., Liang, W. G., Rui, H., Clark, M., Jaskolowski, M., Kim, Y., Deneka, D., Tang, W. J., & Kossiakoff, A. A. (2018). Locking the Elbow: Improved Antibody Fab Fragments as Chaperones for Structure Determination. *Journal of Molecular Biology*, *430*(3), 337–347. https://doi.org/10.1016/j.jmb.2017.12.012

Baker, M. R., Fan, G., Seryshev, A. B., Agosto, M. A., Baker, M. L., & Serysheva, I. I. (2021). Cryo-EM structure of type 1 IP_3R channel in a lipid bilayer. *Communications biology*, *4*(1), 625. https://doi.org/10.1038/s42003-021-02156-4

Bakhoum, S. F., Ngo, B., Laughney, A. M., Cavallo, J. A., Murphy, C. J., Ly, P., Shah, P., Sriram, R. K., Watkins, T., Taunk, N. K., Duran, M., Pauli, C., Shaw, C., Chadalavada, K., Rajasekhar, V. K., Genovese, G., Venkatesan, S., Birkbak, N. J., McGranahan, N., Lundquist, M., … Cantley, L. C. (2018). Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature*, *553*(7689), 467–472. https://doi.org/10.1038/nature25432

Balestrino, R., & Schapira, A. (2020). Parkinson disease. *European journal of neurology*, *27*(1), 27–42. https://doi.org/10.1111/ene.14108

Barnett, K. C., Coronas-Serna, J. M., Zhou, W., Ernandes, M. J., Cao, A., Kranzusch, P. J., & Kagan, J. C. (2019). Phosphoinositide Interactions Position cGAS at the Plasma Membrane to Ensure Efficient Distinction between Self- and Viral DNA. *Cell*, *176*(6), 1432–1446.e11. https://doi.org/10.1016/j.cell.2019.01.049

Barrat, F. J., Elkon, K. B., & Fitzgerald, K. A. (2016). Importance of Nucleic Acid Recognition in Inflammation and Autoimmunity. *Annual review of medicine*, *67*, 323–336. https://doi.org/10.1146/annurev-med-052814-023338

Bauer, E. M., Qin, Y., Miller, T. W., Bandle, R. W., Csanyi, G., Pagano, P. J., Bauer, P. M., Schnermann, J., Roberts, D. D., & Isenberg, J. S. (2010). Thrombospondin-1 supports blood pressure by limiting eNOS activation and endothelial-dependent vasorelaxation. *Cardiovascular research*, *88*(3), 471–481. https://doi.org/10.1093/cvr/cvq218

Baum, R., Sharma, S., Carpenter, S., Li, Q. Z., Busto, P., Fitzgerald, K. A., Marshak-Rothstein, A., & Gravallese, E. M. (2015). Cutting edge: AIM2 and endosomal TLRs differentially regulate arthritis and autoantibody production in DNase II-deficient mice. *Journal of immunology (Baltimore, Md. : 1950)*, *194*(3), 873–877. https://doi.org/10.4049/jimmunol.1402573

Bayburt T. H., Grinkova Y. V., & Sligar S. G. (2002). Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters*, 2, 8, 853–856. https://doi.org/10.1021/nl025623k

Benedict, M., & Zhang, X. (2017). Non-alcoholic fatty liver disease: An expanded review. *World journal of hepatology*, *9*(16), 715–732. https://doi.org/10.4254/wjh.v9.i16.715

Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J., & Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature methods*, *16*(11), 1153–1160. https://doi.org/10.1038/s41592-019-0575-8

Berndt, N., Wolf, C., Fischer, K., Cura Costa, E., Knuschke, P., Zimmermann, N., Schmidt, F., Merkel, M., Chara, O., Lee-Kirsch, M. A., & Günther, C. (2022). Photosensitivity and cGAS-Dependent IFN-1 Activation in

Patients with Lupus and TREX1 Deficiency. *The Journal of investigative dermatology*, *142*(3 Pt A), 633–640.e6. https://doi.org/10.1016/j.jid.2021.04.037

Bisnett, B. J., Condon, B. M., Lamb, C. H., Georgiou, G. R., & Boyce, M. (2021). Export Control: Post-transcriptional Regulation of the COPII Trafficking Pathway. *Frontiers in cell and developmental biology*, *8*, 618652. https://doi.org/10.3389/fcell.2020.618652

Bissinger, R., Petkova-Kirova, P., Mykhailova, O., Oldenborg, P. A., Novikova, E., Donkor, D. A., Dietz, T., Bhuyan, A., Sheffield, W. P., Grau, M., Artunc, F., Kaestner, L., Acker, J. P., & Qadri, S. M. (2020). Thrombospondin-1/CD47 signaling modulates transmembrane cation conductance, survival, and deformability of human red blood cells. *Cell communication and signaling : CCS*, *18*(1), 155. https://doi.org/10.1186/s12964-020-00651-5

Boes, D. M., Godoy-Hernandez, A., & McMillan, D. (2021). Peripheral Membrane Proteins: Promising Therapeutic Targets across Domains of Life. *Membranes*, *11*(5), 346. https://doi.org/10.3390/membranes11050346

Brault, M., Olsen, T. M., Martinez, J., Stetson, D. B., & Oberst, A. (2018). Intracellular Nucleic Acid Sensing Triggers Necroptosis through Synergistic Type I IFN and TNF Signaling. *Journal of immunology (Baltimore, Md. : 1950)*, *200*(8), 2748–2756. https://doi.org/10.4049/jimmunol.1701492

Bridgeman, A., Maelfait, J., Davenne, T., Partridge, T., Peng, Y., Mayer, A., Dong, T., Kaever, V., Borrow, P., & Rehwinkel, J. (2015). Viruses transfer the antiviral second messenger cGAMP between cells. *Science (New York, N.Y.)*, *349*(6253), 1228–1232. https://doi.org/10.1126/science.aab3632

Broom, O. J., Zhang, Y., Oldenborg, P. A., Massoumi, R., & Sjölander, A. (2009). CD47 regulates collagen I-induced cyclooxygenase-2 expression and intestinal epithelial cell migration. *PloS one*, *4*(7), e6371. https://doi.org/10.1371/journal.pone.0006371

Bürckstümmer, T., Baumann, C., Blüml, S., Dixit, E., Dürnberger, G., Jahn, H., Planyavsky, M., Bilban, M., Colinge, J., Bennett, K. L., & Superti-Furga, G. (2009). An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. *Nature immunology*, *10*(3), 266–272. https://doi.org/10.1038/ni.1702

Burger, P., de Korte, D., van den Berg, T. K., & van Bruggen, R. (2012). CD47 in Erythrocyte Ageing and Clearance - the Dutch Point of View. Transfusion medicine and hemotherapy: offizielles Organ der Deutschen Gesellschaft fur Transfusionsmedizin und Immunhamatologie, 39(5), 348–352. https://doi.org/10.1159/000342231

Campbell, I. G., Freemont, P. S., Foulkes, W., & Trowsdale, J. (1992). An ovarian tumor marker with homology to vaccinia virus contains an IgV-like region and multiple transmembrane domains. *Cancer research*, *52*(19), 5416–5420.

Canesso, M., Lemos, L., Neves, T. C., Marim, F. M., Castro, T., Veloso, É. S., Queiroz, C. P., Ahn, J., Santiago, H. C., Martins, F. S., Alves-Silva, J., Ferreira, E., Cara, D. C., Vieira, A. T., Barber, G. N., Oliveira, S. C., & Faria, A. (2018). The cytosolic sensor STING is required for intestinal homeostasis and control of inflammation. *Mucosal immunology*, *11*(3), 820–834. https://doi.org/10.1038/mi.2017.88

Chabanon, R. M., Muirhead, G., Krastev, D. B., Adam, J., Morel, D., Garrido, M., Lamb, A., Hénon, C., Dorvault, N., Rouanne, M., Marlow, R., Bajrami, I., Cardeñosa, M. L., Konde, A., Besse, B., Ashworth, A., Pettitt, S. J., Haider, S., Marabelle, A., Tutt, A. N., … Postel-Vinay, S. (2019). PARP inhibition enhances tumor

cell-intrinsic immunity in ERCC1-deficient non-small cell lung cancer. *The Journal of clinical investigation*, *129*(3), 1211–1228. https://doi.org/10.1172/JCI123319

Chandra, D., Quispe-Tintaya, W., Jahangir, A., Asafu-Adjei, D., Ramos, I., Sintim, H. O., Zhou, J., Hayakawa, Y., Karaolis, D. K., & Gravekamp, C. (2014). STING ligand c-di-GMP improves cancer vaccination against metastatic breast cancer. *Cancer immunology research*, *2*(9), 901–910. https://doi.org/10.1158/2326-6066.CIR-13-0123

Chattopadhyay, S., Kuzmanovic, T., Zhang, Y., Wetzel, J. L., & Sen, G. C. (2016). Ubiquitination of the Transcription Factor IRF-3 Activates RIPA, the Apoptotic Pathway that Protects Mice from Viral Pathogenesis. *Immunity*, *44*(5), 1151–1161. https://doi.org/10.1016/j.immuni.2016.04.009

Chattopadhyay, S., Marques, J. T., Yamashita, M., Peters, K. L., Smith, K., Desai, A., Williams, B. R., & Sen, G. C. (2010). Viral apoptosis is induced by IRF-3-mediated activation of Bax. *The EMBO journal*, *29*(10), 1762–1773. https://doi.org/10.1038/emboj.2010.50

Chaudhary, N., Weissman, D., & Whitehead, K. A. (2021). mRNA vaccines for infectious diseases: principles, delivery and clinical translation. *Nature reviews. Drug discovery*, *20*(11), 817–838. https://doi.org/10.1038/s41573-021-00283-5

Cheloha, R. W., Harmand, T. J., Wijne, C., Schwartz, T. U., & Ploegh, H. L. (2020). Exploring cellular biochemistry with nanobodies. *The Journal of biological chemistry*, *295*(45), 15307–15327. https://doi.org/10.1074/jbc.REV120.012960

Chen, M., Meng, Q., Qin, Y., Liang, P., Tan, P., He, L., Zhou, Y., Chen, Y., Huang, J., Wang, R. F., & Cui, J. (2016). TRIM14 Inhibits cGAS Degradation Mediated by Selective Autophagy Receptor p62 to Promote Innate Immune Responses. *Molecular cell*, *64*(1), 105–119. https://doi.org/10.1016/j.molcel.2016.08.025

Chen, Q., Boire, A., Jin, X., Valiente, M., Er, E. E., Lopez-Soto, A., Jacob, L., Patwa, R., Shah, H., Xu, K., Cross, J. R., & Massagué, J. (2016b). Carcinoma-astrocyte gap junctions promote brain metastasis by cGAMP transfer. *Nature*, *533*(7604), 493–498. https://doi.org/10.1038/nature18268

Chen, S. H., Dominik, P. K., Stanfield, J., Ding, S., Yang, W., Kurd, N., Llewellyn, R., Heyen, J., Wang, C., Melton, Z., Van Blarcom, T., Lindquist, K. C., Chaparro-Riggers, J., & Salek-Ardakani, S. (2021). Dual checkpoint blockade of CD47 and PD-L1 using an affinity-tuned bispecific antibody maximizes antitumor immunity. *Journal for immunotherapy of cancer*, *9*(10), e003464. https://doi.org/10.1136/jitc-2021-003464

Cheng, Z., Dai, T., He, X., Zhang, Z., Xie, F., Wang, S., Zhang, L., & Zhou, F. (2020). The interactions between cGAS-STING pathway and pathogens. *Signal transduction and targeted therapy*, *5*(1), 91. https://doi.org/10.1038/s41392-020-0198-7

Chiu, Y. H., Macmillan, J. B., & Chen, Z. J. (2009). RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell*, *138*(3), 576–591. https://doi.org/10.1016/j.cell.2009.06.015

Chu, R., Takei, J., Knowlton, J. R., Andrykovitch, M., Pei, W., Kajava, A. V., Steinbach, P. J., Ji, X., & Bai, Y. (2002). Redesign of a four-helix bundle protein by phage display coupled with proteolysis and structural characterization by NMR and X-ray crystallography. *Journal of molecular biology*, *323*(2), 253–262. https://doi.org/10.1016/s0022-2836(02)00884-7

Chun, E., Thompson, A. A., Liu, W., Roth, C. B., Griffith, M. T., Katritch, V., Kunken, J., Xu, F., Cherezov, V., Hanson, M. A., & Stevens, R. C. (2012). Fusion partner toolchest for the stabilization and crystallization of G

protein-coupled receptors. *Structure (London, England : 1993)*, *20*(6), 967–976. https://doi.org/10.1016/j.str.2012.04.010

Chung, J., Gao, A. G., & Frazier, W. A. (1997). Thrombspondin acts via integrin-associated protein to activate the platelet integrin alphaIIbbeta3. *The Journal of biological chemistry*, *272*(23), 14740–14746. https://doi.org/10.1074/jbc.272.23.14740

Chung, J., Wang, X. Q., Lindberg, F. P., & Frazier, W. A. (1999). Thrombospondin-1 acts via IAP/CD47 to synergize with collagen in alpha2beta1-mediated platelet activation. *Blood*, *94*(2), 642–648.

Civril, F., Deimling, T., de Oliveira Mann, C. C., Ablasser, A., Moldt, M., Witte, G., Hornung, V., & Hopfner, K. P. (2013). Structural mechanism of cytosolic DNA sensing by cGAS. *Nature*, *498*(7454), 332–337. https://doi.org/10.1038/nature12305

Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., & Sorek, R. (2019). Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature*, *574*(7780), 691–695. https://doi.org/10.1038/s41586-019-1605-5

Conlon, J., Burdette, D. L., Sharma, S., Bhat, N., Thompson, M., Jiang, Z., Rathinam, V. A., Monks, B., Jin, T., Xiao, T. S., Vogel, S. N., Vance, R. E., & Fitzgerald, K. A. (2013). Mouse, but not human STING, binds and signals in response to the vascular disrupting agent 5,6-dimethylxanthenone-4-acetic acid. *Journal of immunology (Baltimore, Md. : 1950)*, *190*(10), 5216–5225. https://doi.org/10.4049/jimmunol.1300097

Conlon, J., Burdette, D. L., Sharma, S., Bhat, N., Thompson, M., Jiang, Z., Rathinam, V. A., Monks, B., Jin, T., Xiao, T. S., Vogel, S. N., Vance, R. E., & Fitzgerald, K. A. (2013). Mouse, but not human STING, binds and signals in response to the vascular disrupting agent 5,6-dimethylxanthenone-4-acetic acid. *Journal of immunology (Baltimore, Md. : 1950)*, *190*(10), 5216–5225. https://doi.org/10.4049/jimmunol.1300097

Corrales, L., Glickman, L. H., McWhirter, S. M., Kanne, D. B., Sivick, K. E., Katibah, G. E., Woo, S. R., Lemmens, E., Banda, T., Leong, J. J., Metchette, K., Dubensky, T. W., Jr, & Gajewski, T. F. (2015). Direct Activation of STING in the Tumor Microenvironment Leads to Potent and Systemic Tumor Regression and Immunity. *Cell reports*, *11*(7), 1018–1030. https://doi.org/10.1016/j.celrep.2015.04.031

Crow, Y. J., & Manel, N. (2015). Aicardi-Goutières syndrome and the type I interferonopathies. *Nature reviews. Immunology*, *15*(7), 429–440. https://doi.org/10.1038/nri3850

Csányi, G., Yao, M., Rodríguez, A. I., Al Ghouleh, I., Sharifi-Sanjani, M., Frazziano, G., Huang, X., Kelley, E. E., Isenberg, J. S., & Pagano, P. J. (2012). Thrombospondin-1 regulates blood flow via CD47 receptor-mediated activation of NADPH oxidase 1. *Arteriosclerosis, thrombosis, and vascular biology*, *32*(12), 2966–2973. https://doi.org/10.1161/ATVBAHA.112.300031

Cui, Y., Yu, H., Zheng, X., Peng, R., Wang, Q., Zhou, Y., Wang, R., Wang, J., Qu, B., Shen, N., Guo, Q., Liu, X., & Wang, C. (2017). SENP7 Potentiates cGAS Activation by Relieving SUMO-Mediated Inhibition of Cytosolic DNA Sensing. *PLoS pathogens*, *13*(1), e1006156. https://doi.org/10.1371/journal.ppat.1006156

Dąbek, J., Kułach, A., & Gąsior, Z. (2010). Nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB): a new potential therapeutic target in atherosclerosis?. *Pharmacological reports : PR*, *62*(5), 778–783. https://doi.org/10.1016/s1734-1140(10)70338-8

Daei Farshchi Adli, A., Jahanban-Esfahlan, R., Seidi, K., Samandari-Rad, S., & Zarghami, N. (2018). An overview on Vadimezan (DMXAA): The vascular disrupting agent. *Chemical biology & drug design*, *91*(5), 996–1006. https://doi.org/10.1111/cbdd.13166

Dai, J., Huang, Y. J., He, X., Zhao, M., Wang, X., Liu, Z. S., Xue, W., Cai, H., Zhan, X. Y., Huang, S. Y., He, K., Wang, H., Wang, N., Sang, Z., Li, T., Han, Q. Y., Mao, J., Diao, X., Song, N., Chen, Y., … Li, T. (2019). Acetylation Blocks cGAS Activity and Inhibits Self-DNA-Induced Autoimmunity. *Cell*, *176*(6), 1447–1460.e14. https://doi.org/10.1016/j.cell.2019.01.016

Davies, B. W., Bogard, R. W., Young, T. S., & Mekalanos, J. J. (2012). Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for V. cholerae virulence. *Cell*, *149*(2), 358–370. https://doi.org/10.1016/j.cell.2012.01.053

Decout, A., Katz, J. D., Venkatraman, S., & Ablasser, A. (2021). The cGAS-STING pathway as a therapeutic target in inflammatory diseases. *Nature reviews. Immunology*, *21*(9), 548–569. https://doi.org/10.1038/s41577-021-00524-z

Dheilly, E., Majocchi, S., Moine, V., Didelot, G., Broyer, L., Calloud, S., Malinge, P., Chatel, L., Ferlin, W. G., Kosco-Vilbois, M. H., Fischer, N., & Masternak, K. (2018). Tumor-Directed Blockade of CD47 with Bispecific Antibodies Induces Adaptive Antitumor Immunity. *Antibodies (Basel, Switzerland)*, *7*(1), 3. https://doi.org/10.3390/antib7010003

Dhuriya, Y. K., & Sharma, D. (2018). Necroptosis: a regulated inflammatory mode of cell death. *Journal of neuroinflammation*, *15*(1), 199. https://doi.org/10.1186/s12974-018-1235-0

Dietrich, M. H., Chan, L. J., Adair, A., Keremane, S., Pymm, P., Lo, A. W., Cao, Y. C., & Tham, W. H. (2021). Nanobody generation and structural characterization of Plasmodium falciparum 6-cysteine protein Pf12p. *The Biochemical journal*, *478*(3), 579–595. https://doi.org/10.1042/BCJ20200415

Dikic, I., & Elazar, Z. (2018). Mechanism and medical implications of mammalian autophagy. *Nature reviews. Molecular cell biology*, *19*(6), 349–364. https://doi.org/10.1038/s41580-018-0003-4

Diner, E. J., Burdette, D. L., Wilson, S. C., Monroe, K. M., Kellenberger, C. A., Hyodo, M., Hayakawa, Y., Hammond, M. C., & Vance, R. E. (2013). The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. *Cell reports*, *3*(5), 1355–1361. https://doi.org/10.1016/j.celrep.2013.05.009

Ding, L., Kim, H. J., Wang, Q., Kearns, M., Jiang, T., Ohlson, C. E., Li, B. B., Xie, S., Liu, J. F., Stover, E. H., Howitt, B. E., Bronson, R. T., Lazo, S., Roberts, T. M., Freeman, G. J., Konstantinopoulos, P. A., Matulonis, U. A., & Zhao, J. J. (2018). PARP Inhibition Elicits STING-Dependent Antitumor Immunity in Brca1-Deficient Ovarian Cancer. *Cell reports*, *25*(11), 2972–2980.e5. https://doi.org/10.1016/j.celrep.2018.11.054

Dmitriev, O. Y., Lutsenko, S., & Muyldermans, S. (2016). Nanobodies as Probes for Protein Dynamics in Vitro and in Cells. *The Journal of biological chemistry*, *291*(8), 3767–3775. https://doi.org/10.1074/jbc.R115.679811

Du, M., & Chen, Z. J. (2018). DNA-induced liquid phase condensation of cGAS activates innate immune signaling. *Science (New York, N.Y.)*, *361*(6403), 704–709. https://doi.org/10.1126/science.aat1022

Duhoo, Y., Roche, J., Trinh, T., Desmyter, A., Gaubert, A., Kellenberger, C., Cambillau, C., Roussel, A., & Leone, P. (2017). Camelid nanobodies used as crystallization chaperones for different constructs of PorM, a component of the type IX secretion system from Porphyromonas gingivalis. *Acta crystallographica. Section F, Structural biology communications*, *73*(Pt 5), 286–293. https://doi.org/10.1107/S2053230X17005969

References

Efremov, R. G., Gatsogiannis, C., & Raunser, S. (2017). Lipid Nanodiscs as a Tool for High-Resolution Structure Determination of Membrane Proteins by Single-Particle Cryo-EM. *Methods in enzymology*, *594*, 1–30. https://doi.org/10.1016/bs.mie.2017.05.007

Eladl, E., Tremblay-LeMay, R., Rastgoo, N., Musani, R., Chen, W., Liu, A., & Chang, H. (2020). Role of CD47 in Hematological Malignancies. *Journal of hematology & oncology*, *13*(1), 96. https://doi.org/10.1186/s13045-020-00930-1

Eliot G (2022) Different Types of Oligosaccharides and Their Functions. *J Glycomics Lipidomics*. 11: 321

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 4), 486–501. https://doi.org/10.1107/S0907444910007493

Ereño-Orbea, J., Sicard, T., Cui, H., Carson, J., Hermans, P., & Julien, J. P. (2018). Structural Basis of Enhanced Crystallizability Induced by a Molecular Chaperone for Antibody Antigen-Binding Fragments. *Journal of molecular biology*, *430*(3), 322–336. https://doi.org/10.1016/j.jmb.2017.12.010

Ergun, S. L., Fernandez, D., Weiss, T. M., & Li, L. (2019). STING Polymer Structure Reveals Mechanisms for Activation, Hyperactivation, and Inhibition. *Cell*, *178*(2), 290–301.e10. https://doi.org/10.1016/j.cell.2019.05.036

Evans P. (2006). Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography*, *62*(Pt 1), 72–82. https://doi.org/10.1107/S0907444905036693

Fenalti, G., Villanueva, N., Griffith, M., Pagarigan, B., Lakkaraju, S. K., Huang, R. Y., Ladygina, N., Sharma, A., Mikolon, D., Abbasian, M., Johnson, J., Hadjivassiliou, H., Zhu, D., Chamberlain, P. P., Cho, H., & Hariharan, K. (2021). Structure of the human marker of self 5-transmembrane receptor CD47. *Nature communications*, *12*(1), 5218. https://doi.org/10.1038/s41467-021-25475-w

Frauenfeld, J., Löving, R., Armache, J. P., Sonnen, A. F., Guettou, F., Moberg, P., Zhu, L., Jegerschöld, C., Flayhan, A., Briggs, J. A., Garoff, H., Löw, C., Cheng, Y., & Nordlund, P. (2016). A saposin-lipoprotein nanoparticle system for membrane proteins. *Nature methods*, *13*(4), 345–351. https://doi.org/10.1038/nmeth.3801

Gaidt, M. M., Ebert, T. S., Chauhan, D., Ramshorn, K., Pinci, F., Zuber, S., O'Duill, F., Schmid-Burgk, J. L., Hoss, F., Buhmann, R., Wittmann, G., Latz, E., Subklewe, M., & Hornung, V. (2017). The DNA Inflammasome in Human Myeloid Cells Is Initiated by a STING-Cell Death Program Upstream of NLRP3. *Cell*, *171*(5), 1110–1124.e18. https://doi.org/10.1016/j.cell.2017.09.039

Gall, A., Treuting, P., Elkon, K. B., Loo, Y. M., Gale, M., Jr, Barber, G. N., & Stetson, D. B. (2012). Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease. *Immunity*, *36*(1), 120–131. https://doi.org/10.1016/j.immuni.2011.11.018

Gao, A. G., Lindberg, F. P., Finn, M. B., Blystone, S. D., Brown, E. J., & Frazier, W. A. (1996). Integrin-associated protein is a receptor for the C-terminal domain of thrombospondin. *The Journal of biological chemistry*, *271*(1), 21–24. https://doi.org/10.1074/jbc.271.1.21

Gao, Q., Chen, K., Gao, L., Zheng, Y., & Yang, Y. G. (2016b). Thrombospondin-1 signaling through CD47 inhibits cell cycle progression and induces senescence in endothelial cells. *Cell death & disease*, *7*(9), e2368. https://doi.org/10.1038/cddis.2016.155

Gao, Y., Cao, E., Julius, D., & Cheng, Y. (2016). TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature*, *534*(7607), 347–351. https://doi.org/10.1038/nature17964

Ghimire, K., Li, Y., Chiba, T., Julovi, S. M., Li, J., Ross, M. A., Straub, A. C., O'Connell, P. J., Rüegg, C., Pagano, P. J., Isenberg, J. S., & Rogers, N. M. (2020). CD47 Promotes Age-Associated Deterioration in Angiogenesis, Blood Flow and Glucose Homeostasis. *Cells*, *9*(7), 1695. https://doi.org/10.3390/cells9071695

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318

Gilmore T. D. (2006). Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene*, *25*(51), 6680–6684. https://doi.org/10.1038/sj.onc.1209954

Gogoi, H., Mansouri, S., & Jin, L. (2020). The Age of Cyclic Dinucleotide Vaccine Adjuvants. *Vaccines*, *8*(3), 453. https://doi.org/10.3390/vaccines8030453

Gonugunta, V. K., Sakai, T., Pokatayev, V., Yang, K., Wu, J., Dobbs, N., & Yan, N. (2017). Trafficking-Mediated STING Degradation Requires Sorting to Acidified Endolysosomes and Can Be Targeted to Enhance Anti-tumor Response. *Cell reports*, *21*(11), 3234–3242. https://doi.org/10.1016/j.celrep.2017.11.061

Gray, E. E., Treuting, P. M., Woodward, J. J., & Stetson, D. B. (2015). Cutting Edge: cGAS Is Required for Lethal Autoimmune Disease in the Trex1-Deficient Mouse Model of Aicardi-Goutières Syndrome. *Journal of immunology (Baltimore, Md. : 1950)*, *195*(5), 1939–1943. https://doi.org/10.4049/jimmunol.1500969

Green, J. M., Zheleznyak, A., Chung, J., Lindberg, F. P., Sarfati, M., Frazier, W. A., & Brown, E. J. (1999). Role of cholesterol in formation and function of a signaling complex involving alphavbeta3, integrin-associated protein (CD47), and heterotrimeric G proteins. *The Journal of cell biology*, *146*(3), 673–682. https://doi.org/10.1083/jcb.146.3.673

Gui, X., Yang, H., Li, T., Tan, X., Shi, P., Li, M., Du, F., & Chen, Z. J. (2019). Autophagy induction via STING trafficking is a primordial function of the cGAS pathway. *Nature*, *567*(7747), 262–266. https://doi.org/10.1038/s41586-019-1006-9

Guicciardi, M. E., & Gores, G. J. (2009). Life and death by death receptors. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *23*(6), 1625–1637. https://doi.org/10.1096/fj.08-111005

Gulen, M. F., Koch, U., Haag, S. M., Schuler, F., Apetoh, L., Villunger, A., Radtke, F., & Ablasser, A. (2017). Signalling strength determines proapoptotic functions of STING. *Nature communications*, *8*(1), 427. https://doi.org/10.1038/s41467-017-00573-w

Guo, F., Han, Y., Zhao, X., Wang, J., Liu, F., Xu, C., Wei, L., Jiang, J. D., Block, T. M., Guo, J. T., & Chang, J. (2015). STING agonists induce an innate antiviral immune response against hepatitis B virus. *Antimicrobial agents and chemotherapy*, *59*(2), 1273–1281. https://doi.org/10.1128/AAC.04321-14

Hagn, F., Etzkorn, M., Raschle, T., & Wagner, G. (2013). Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *Journal of the American Chemical Society*, *135*(5), 1919–1925. https://doi.org/10.1021/ja310901f

Hansen, A. L., Buchan, G. J., Rühl, M., Mukai, K., Salvatore, S. R., Ogawa, E., Andersen, S. D., Iversen, M. B., Thielke, A. L., Gunderstofte, C., Motwani, M., Møller, C. T., Jakobsen, A. S., Fitzgerald, K. A., Roos, J., Lin,

R., Maier, T. J., Goldbach-Mansky, R., Miner, C. A., Qian, W., … Holm, C. K. (2018). Nitro-fatty acids are formed in response to virus infection and are potent inhibitors of STING palmitoylation and signaling. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(33), E7768–E7775. https://doi.org/10.1073/pnas.1806239115

Hardy, D., Desuzinges Mandon, E., Rothnie, A. J., & Jawhari, A. (2018). The yin and yang of solubilization and stabilization for wild-type and full-length membrane protein. *Methods (San Diego, Calif.)*, *147*, 118–125. https://doi.org/10.1016/j.ymeth.2018.02.017

Hassaine, G., Deluz, C., Grasso, L., Wyss, R., Tol, M. B., Hovius, R., Graff, A., Stahlberg, H., Tomizaki, T., Desmyter, A., Moreau, C., Li, X. D., Poitevin, F., Vogel, H., & Nury, H. (2014). X-ray structure of the mouse serotonin 5-HT3 receptor. *Nature*, *512*(7514), 276–281. https://doi.org/10.1038/nature13552

Hayat, S., Bianconi, V., Pirro, M., Jaafari, M. R., Hatamipour, M., & Sahebkar, A. (2020). CD47: role in the immune system and application to cancer therapy. *Cellular oncology (Dordrecht)*, *43*(1), 19–30. https://doi.org/10.1007/s13402-019-00469-5

Hemmi, H., Takeuchi, O., Kawai, T., Kaisho, T., Sato, S., Sanjo, H., Matsumoto, M., Hoshino, K., Wagner, H., Takeda, K., & Akira, S. (2000). A Toll-like receptor recognizes bacterial DNA. *Nature*, *408*(6813), 740–745. https://doi.org/10.1038/35047123

Herrmann, A., Wittmann, S., Thomas, D., Shepard, C. N., Kim, B., Ferreirós, N., & Gramberg, T. (2018). The SAMHD1-mediated block of LINE-1 retroelements is regulated by phosphorylation. *Mobile DNA*, *9*, 11. https://doi.org/10.1186/s13100-018-0116-5

Hopfner, K. P., & Hornung, V. (2020). Molecular mechanisms and cellular functions of cGAS-STING signalling. *Nature reviews. Molecular cell biology*, *21*(9), 501–521. https://doi.org/10.1038/s41580-020-0244-x

Hornung, V., Hartmann, R., Ablasser, A., & Hopfner, K. P. (2014). OAS proteins and cGAS: unifying concepts in sensing and responding to cytosolic nucleic acids. *Nature reviews. Immunology*, *14*(8), 521–528. https://doi.org/10.1038/nri3719

Hu, M. M., Yang, Q., Xie, X. Q., Liao, C. Y., Lin, H., Liu, T. T., Yin, L., & Shu, H. B. (2016). Sumoylation Promotes the Stability of the DNA Sensor cGAS and the Adaptor STING to Regulate the Kinetics of Response to DNA Virus. *Immunity*, *45*(3), 555–569. https://doi.org/10.1016/j.immuni.2016.08.014

Humphries, F., Shmuel-Galia, L., Jiang, Z., Wilson, R., Landis, P., Ng, S. L., Parsi, K. M., Maehr, R., Cruz, J., Morales-Ramos, A., Ramanjulu, J. M., Bertin, J., Pesiridis, G. S., & Fitzgerald, K. A. (2021). A diamidobenzimidazole STING agonist protects against SARS-CoV-2 infection. *Science immunology*, *6*(59), eabi9002. https://doi.org/10.1126/sciimmunol.abi9002

Hynes, R. O., Lively, J. C., McCarty, J. H., Taverna, D., Francis, S. E., Hodivala-Dilke, K., & Xiao, Q. (2002). The diverse roles of integrins and their ligands in angiogenesis. *Cold Spring Harbor symposia on quantitative biology*, *67*, 143–153. https://doi.org/10.1101/sqb.2002.67.143

Imai, H., Hamai, K., Komatsuda, A., Ohtani, H., & Miura, A. B. (1997). IgG subclasses in patients with membranoproliferative glomerulonephritis, membranous nephropathy, and lupus nephritis. *Kidney international*, *51*(1), 270–276. https://doi.org/10.1038/ki.1997.32

Inagaki, S. & Ghirlando, R. (2017). Nanodisc characterization by analytical ultracentrifugation. Nanotechnology Reviews, 6(1), 3-14. https://doi.org/10.1515/ntrev-2016-0082

Irazoqui, J. E., Urbach, J. M., & Ausubel, F. M. (2010). Evolution of host innate defense: insights from Caenorhabditis elegans and primitive invertebrates. *Nature reviews. Immunology*, *10*(1), 47–58. https://doi.org/10.1038/nri2689

Isenberg, J. S., Annis, D. S., Pendrak, M. L., Ptaszynska, M., Frazier, W. A., Mosher, D. F., & Roberts, D. D. (2009). Differential interactions of thrombospondin-1, -2, and -4 with CD47 and effects on cGMP signaling and ischemic injury responses. *The Journal of biological chemistry*, *284*(2), 1116–1125. https://doi.org/10.1074/jbc.M804860200

Ishikawa, H., & Barber, G. N. (2008). STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature*, *455*(7213), 674–678. https://doi.org/10.1038/nature07317

Jaiswal, S., Jamieson, C. H., Pang, W. W., Park, C. Y., Chao, M. P., Majeti, R., Traver, D., van Rooijen, N., & Weissman, I. L. (2009). CD47 is upregulated on circulating hematopoietic stem cells and leukemia cells to avoid phagocytosis. *Cell*, *138*(2), 271–285. https://doi.org/10.1016/j.cell.2009.05.046

Jiang, H., Xue, X., Panda, S., Kawale, A., Hooy, R. M., Liang, F., Sohn, J., Sung, P., & Gekara, N. O. (2019). Chromatin-bound cGAS is an inhibitor of DNA repair and hence accelerates genome destabilization and cell death. *The EMBO journal*, *38*(21), e102718. https://doi.org/10.15252/embj.2019102718

Jiang, Z., Sun, H., Yu, J., Tian, W., & Song, Y. (2021). Targeting CD47 for cancer immunotherapy. *Journal of hematology & oncology*, *14*(1), 180. https://doi.org/10.1186/s13045-021-01197-w

Jin, L., Waterman, P. M., Jonscher, K. R., Short, C. M., Reisdorph, N. A., & Cambier, J. C. (2008). MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals. *Molecular and cellular biology*, *28*(16), 5014–5026. https://doi.org/10.1128/MCB.00640-08

Jønsson, K. L., Laustsen, A., Krapp, C., Skipper, K. A., Thavachelvam, K., Hotter, D., Egedal, J. H., Kjolby, M., Mohammadi, P., Prabakaran, T., Sørensen, L. K., Sun, C., Jensen, S. B., Holm, C. K., Lebbink, R. J., Johannsen, M., Nyegaard, M., Mikkelsen, J. G., Kirchhoff, F., Paludan, S. R., … Jakobsen, M. R. (2017). IFI16 is required for DNA sensing in human macrophages by promoting production and function of cGAMP. *Nature communications*, *8*, 14391. https://doi.org/10.1038/ncomms14391

Jovčevska, I., & Muyldermans, S. (2020). The Therapeutic Potential of Nanobodies. *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy*, *34*(1), 11–26. https://doi.org/10.1007/s40259-019-00392-z

Kabsch W. (2010). XDS. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 2), 125–132. https://doi.org/10.1107/S0907444909047337

Kale, A., Rogers, N. M., & Ghimire, K. (2021). Thrombospondin-1 CD47 Signalling: From Mechanisms to Medicine. *International journal of molecular sciences*, *22*(8), 4062. https://doi.org/10.3390/ijms22084062

Kampjut, D., Steiner, J., & Sazanov, L. A. (2021). Cryo-EM grid optimization for membrane proteins. *iScience*, *24*(3), 102139. https://doi.org/10.1016/j.isci.2021.102139

Karanja, C. W., Yeboah, K. S., & Sintim, H. O. (2021). Identification of a *Mycobacterium tuberculosis* Cyclic Dinucleotide Phosphodiesterase Inhibitor. *ACS infectious diseases*, *7*(2), 309–317. https://doi.org/10.1021/acsinfecdis.0c00444

References

Kato, K., Nishimasu, H., Oikawa, D., Hirano, S., Hirano, H., Kasuya, G., Ishitani, R., Tokunaga, F., & Nureki, O. (2018). Structural insights into cGAMP degradation by Ecto-nucleotide pyrophosphatase phosphodiesterase 1. *Nature communications*, *9*(1), 4424. https://doi.org/10.1038/s41467-018-06922-7

Kaur, S., Chang, T., Singh, S. P., Lim, L., Mannan, P., Garfield, S. H., Pendrak, M. L., Soto-Pantoja, D. R., Rosenberg, A. Z., Jin, S., & Roberts, D. D. (2014). CD47 signaling regulates the immunosuppressive activity of VEGF in T cells. *Journal of immunology (Baltimore, Md. : 1950)*, *193*(8), 3914–3924. https://doi.org/10.4049/jimmunol.1303116

Kaur, S., Kuznetsova, S. A., Pendrak, M. L., Sipes, J. M., Romeo, M. J., Li, Z., Zhang, L., & Roberts, D. D. (2011). Heparan sulfate modification of the transmembrane receptor CD47 is necessary for inhibition of T cell receptor signaling by thrombospondin-1. *The Journal of biological chemistry*, *286*(17), 14991–15002. https://doi.org/10.1074/jbc.M110.179663

Kim, D. S., Song, H. N., Nam, H. J., Kim, S. G., Park, Y. S., Park, J. C., Woo, E. J., & Lim, H. K. (2014). Directed evolution of human heavy chain variable domain (VH) using in vivo protein fitness filter. *PloS one*, *9*(6), e98178. https://doi.org/10.1371/journal.pone.0098178

Koenigsknecht, J., & Landreth, G. (2004). Microglial phagocytosis of fibrillar beta-amyloid through a beta1 integrin-dependent mechanism. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *24*(44), 9838–9846. https://doi.org/10.1523/JNEUROSCI.2557-04.2004

König, N., Fiehn, C., Wolf, C., Schuster, M., Cura Costa, E., Tüngler, V., Alvarez, H. A., Chara, O., Engel, K., Goldbach-Mansky, R., Günther, C., & Lee-Kirsch, M. A. (2017). Familial chilblain lupus due to a gain-of-function mutation in STING. *Annals of the rheumatic diseases*, *76*(2), 468–472. https://doi.org/10.1136/annrheumdis-2016-209841

Koning, N., Bö, L., Hoek, R. M., & Huitinga, I. (2007). Downregulation of macrophage inhibitory molecules in multiple sclerosis lesions. *Annals of neurology*, *62*(5), 504–514. https://doi.org/10.1002/ana.21220

Kranzusch P. J. (2019). cGAS and CD-NTase enzymes: structure, mechanism, and evolution. *Current opinion in structural biology*, *59*, 178–187. https://doi.org/10.1016/j.sbi.2019.08.003

Kranzusch, P. J., Lee, A. S., Berger, J. M., & Doudna, J. A. (2013). Structure of human cGAS reveals a conserved family of second-messenger enzymes in innate immunity. *Cell reports*, *3*(5), 1362–1368. https://doi.org/10.1016/j.celrep.2013.05.008

Kranzusch, P. J., Wilson, S. C., Lee, A. S., Berger, J. M., Doudna, J. A., & Vance, R. E. (2015). Ancient Origin of cGAS-STING Reveals Mechanism of Universal 2',3' cGAMP Signaling. *Molecular cell*, *59*(6), 891–903. https://doi.org/10.1016/j.molcel.2015.07.022

Kreienkamp, R., Graziano, S., Coll-Bonfill, N., Bedia-Diaz, G., Cybulla, E., Vindigni, A., Dorsett, D., Kubben, N., Batista, L., & Gonzalo, S. (2018). A Cell-Intrinsic Interferon-like Response Links Replication Stress to Cellular Aging Caused by Progerin. *Cell reports*, *22*(8), 2006–2015. https://doi.org/10.1016/j.celrep.2018.01.090

Kuchta, K., Knizewski, L., Wyrwicz, L. S., Rychlewski, L., & Ginalski, K. (2009). Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. *Nucleic acids research*, *37*(22), 7701–7714. https://doi.org/10.1093/nar/gkp854

Kulig, W., Tynkkynen, J., Javanainen, M., Manna, M., Rog, T., Vattulainen, I., & Jungwirth, P. (2014). How well does cholesteryl hemisuccinate mimic cholesterol in saturated phospholipid bilayers?. *Journal of molecular modeling*, *20*(2), 2121. https://doi.org/10.1007/s00894-014-2121-z

Kvansakul, M., Adams, J. C., & Hohenester, E. (2004). Structure of a thrombospondin C-terminal fragment reveals a novel calcium core in the type 3 repeats. *The EMBO journal*, *23*(6), 1223–1233. https://doi.org/10.1038/sj.emboj.7600166

Lahaye, X., Gentili, M., Silvin, A., Conrad, C., Picard, L., Jouve, M., Zueva, E., Maurin, M., Nadalin, F., Knott, G. J., Zhao, B., Du, F., Rio, M., Amiel, J., Fox, A. H., Li, P., Etienne, L., Bond, C. S., Colleaux, L., & Manel, N. (2018). NONO Detects the Nuclear HIV Capsid to Promote cGAS-Mediated Innate Immune Activation. *Cell*, *175*(2), 488–501.e22. https://doi.org/10.1016/j.cell.2018.08.062

Laskowski R. A. (2001). PDBsum: summaries and analyses of PDB structures. *Nucleic acids research*, *29*(1), 221–222. https://doi.org/10.1093/nar/29.1.221

Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283-291. https://doi.org/10.1107/S0021889892009944

Latour, S., Tanaka, H., Demeure, C., Mateo, V., Rubio, M., Brown, E. J., Maliszewski, C., Lindberg, F. P., Oldenborg, A., Ullrich, A., Delespesse, G., & Sarfati, M. (2001). Bidirectional negative regulation of human T and dendritic cells by CD47 and its cognate receptor signal-regulator protein-alpha: down-regulation of IL-12 responsiveness and inhibition of dendritic cell activation. *Journal of immunology (Baltimore, Md. : 1950)*, *167*(5), 2547–2554. https://doi.org/10.4049/jimmunol.167.5.2547

Laustsen, A. H., Greiff, V., Karatt-Vellatt, A., Muyldermans, S., & Jenkins, T. P. (2021). Animal Immunization, in Vitro Display Technologies, and Machine Learning for Antibody Discovery. *Trends in biotechnology*, *39*(12), 1263–1273. https://doi.org/10.1016/j.tibtech.2021.03.003

Leclair, P., Liu, C. C., Monajemi, M., Reid, G. S., Sly, L. M., & Lim, C. J. (2018). CD47-ligation induced cell death in T-acute lymphoblastic leukemia. *Cell death & disease*, *9*(5), 544. https://doi.org/10.1038/s41419-018-0601-2

Lee, W. Y., Weber, D. A., Laur, O., Severson, E. A., McCall, I., Jen, R. P., Chin, A. C., Wu, T., Gernert, K. M., & Parkos, C. A. (2007). Novel structural determinants on SIRP alpha that mediate binding to CD47. *Journal of immunology (Baltimore, Md. : 1950)*, *179*(11), 7741–7750. https://doi.org/10.4049/jimmunol.179.11.7741

Li, Z., Liu, G., Sun, L., Teng, Y., Guo, X., Jia, J., Sha, J., Yang, X., Chen, D., & Sun, Q. (2015). PPM1A regulates antiviral signaling by antagonizing TBK1-mediated STING phosphorylation and aggregation. *PLoS pathogens*, *11*(3), e1004783. https://doi.org/10.1371/journal.ppat.1004783

Lian, H., Wei, J., Zang, R., Ye, W., Yang, Q., Zhang, X. N., Chen, Y. D., Fu, Y. Z., Hu, M. M., Lei, C. Q., Luo, W. W., Li, S., & Shu, H. B. (2018). ZCCHC3 is a co-sensor of cGAS for dsDNA recognition in innate immune response. *Nature communications*, *9*(1), 3349. https://doi.org/10.1038/s41467-018-05559-w

Liang, H., Deng, L., Hou, Y., Meng, X., Huang, X., Rao, E., Zheng, W., Mauceri, H., Mack, M., Xu, M., Fu, Y. X., & Weichselbaum, R. R. (2017). Host STING-dependent MDSC mobilization drives extrinsic radiation resistance. *Nature communications*, *8*(1), 1736. https://doi.org/10.1038/s41467-017-01566-5

Lindberg, F. P., Gresham, H. D., Reinhold, M. I., & Brown, E. J. (1996). Integrin-associated protein immunoglobulin domain is necessary for efficient vitronectin bead binding. *The Journal of cell biology*, *134*(5), 1313–1322. https://doi.org/10.1083/jcb.134.5.1313

Lindberg, F. P., Gresham, H. D., Schwarz, E., & Brown, E. J. (1993). Molecular cloning of integrin-associated protein: an immunoglobulin family member with multiple membrane-spanning domains implicated in alpha v beta 3-dependent ligand binding. *The Journal of cell biology*, *123*(2), 485–496. https://doi.org/10.1083/jcb.123.2.485

Liu, H., Moura-Alves, P., Pei, G., Mollenkopf, H. J., Hurwitz, R., Wu, X., Wang, F., Liu, S., Ma, M., Fei, Y., Zhu, C., Koehler, A. B., Oberbeck-Mueller, D., Hahnke, K., Klemm, M., Guhlich-Bornhof, U., Ge, B., Tuukkanen, A., Kolbe, M., Dorhoi, A., … Kaufmann, S. H. (2019b). cGAS facilitates sensing of extracellular cyclic dinucleotides to activate innate immunity. *EMBO reports*, *20*(4), e46293. https://doi.org/10.15252/embr.201846293

Liu, H., Zhang, H., Wu, X., Ma, D., Wu, J., Wang, L., Jiang, Y., Fei, Y., Zhu, C., Tan, R., Jungblut, P., Pei, G., Dorhoi, A., Yan, Q., Zhang, F., Zheng, R., Liu, S., Liang, H., Liu, Z., Yang, H., … Ge, B. (2018). Nuclear cGAS suppresses DNA repair and promotes tumorigenesis. *Nature*, *563*(7729), 131–136. https://doi.org/10.1038/s41586-018-0629-6

Liu, J., Wang, L., Zhao, F., Tseng, S., Narayanan, C., Shura, L., Willingham, S., Howard, M., Prohaska, S., Volkmer, J., Chao, M., Weissman, I. L., & Majeti, R. (2015). Pre-Clinical Development of a Humanized Anti-CD47 Antibody with Anti-Cancer Therapeutic Potential. *PloS one*, *10*(9), e0137345. https://doi.org/10.1371/journal.pone.0137345

Liu, N., Pang, X., Zhang, H., & Ji, P. (2022). The cGAS-STING Pathway in Bacterial Infection and Bacterial Immunity. *Frontiers in immunology*, *12*, 814709. https://doi.org/10.3389/fimmu.2021.814709

Liu, S., Cai, X., Wu, J., Cong, Q., Chen, X., Li, T., Du, F., Ren, J., Wu, Y. T., Grishin, N. V., & Chen, Z. J. (2015). Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. *Science (New York, N.Y.)*, *347*(6227), aaa2630. https://doi.org/10.1126/science.aaa2630

Liu, T., Zhang, L., Joo, D., & Sun, S. C. (2017). NF-κB signaling in inflammation. *Signal transduction and targeted therapy*, *2*, 17023–. https://doi.org/10.1038/sigtrans.2017.23

Liu, Y., Jesus, A. A., Marrero, B., Yang, D., Ramsey, S. E., Sanchez, G., Tenbrock, K., Wittkowski, H., Jones, O. Y., Kuehn, H. S., Lee, C. R., DiMattia, M. A., Cowen, E. W., Gonzalez, B., Palmer, I., DiGiovanna, J. J., Biancotto, A., Kim, H., Tsai, W. L., Trier, A. M., … Goldbach-Mansky, R. (2014). Activated STING in a vascular and pulmonary syndrome. *The New England journal of medicine*, *371*(6), 507–518. https://doi.org/10.1056/NEJMoa1312625

Liu, Y., Merlin, D., Burst, S. L., Pochet, M., Madara, J. L., & Parkos, C. A. (2001). The role of CD47 in neutrophil transmigration. Increased rate of migration correlates with increased cell surface expression of CD47. *The Journal of biological chemistry*, *276*(43), 40156–40166. https://doi.org/10.1074/jbc.M104138200

Liu, Z. S., Cai, H., Xue, W., Wang, M., Xia, T., Li, W. J., Xing, J. Q., Zhao, M., Huang, Y. J., Chen, S., Wu, S. M., Wang, X., Liu, X., Pang, X., Zhang, Z. Y., Li, T., Dai, J., Dong, F., Xia, Q., Li, A. L., … Li, T. (2019). G3BP1 promotes DNA binding and activation of cGAS. *Nature immunology*, *20*(1), 18–28. https://doi.org/10.1038/s41590-018-0262-4

Luksch, H., Stinson, W. A., Platt, D. J., Qian, W., Kalugotla, G., Miner, C. A., Bennion, B. G., Gerbaulet, A., Rösen-Wolff, A., & Miner, J. J. (2019). STING-associated lung disease in mice relies on T cells but not type I

interferon. *The Journal of allergy and clinical immunology*, *144*(1), 254–266.e8. https://doi.org/10.1016/j.jaci.2019.01.044

Luo, X., Li, H., Ma, L., Zhou, J., Guo, X., Woo, S. L., Pei, Y., Knight, L. R., Deveau, M., Chen, Y., Qian, X., Xiao, X., Li, Q., Chen, X., Huo, Y., McDaniel, K., Francis, H., Glaser, S., Meng, F., Alpini, G., … Wu, C. (2018). Expression of STING Is Increased in Liver Tissues From Patients With NAFLD and Promotes Macrophage-Mediated Hepatic Inflammation and Fibrosis in Mice. *Gastroenterology*, *155*(6), 1971–1984.e4. https://doi.org/10.1053/j.gastro.2018.09.010

Ma, L., Zhu, M., Gai, J., Li, G., Chang, Q., Qiao, P., Cao, L., Chen, W., Zhang, S., & Wan, Y. (2020). Preclinical development of a novel CD47 nanobody with less toxicity and enhanced anti-cancer therapeutic potential. *Journal of nanobiotechnology*, *18*(1), 12. https://doi.org/10.1186/s12951-020-0571-2

Mackenzie, K. J., Carroll, P., Martin, C. A., Murina, O., Fluteau, A., Simpson, D. J., Olova, N., Sutcliffe, H., Rainger, J. K., Leitch, A., Osborn, R. T., Wheeler, A. P., Nowotny, M., Gilbert, N., Chandra, T., Reijns, M., & Jackson, A. P. (2017). cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature*, *548*(7668), 461–465. https://doi.org/10.1038/nature23449

Majeti, R., Chao, M. P., Alizadeh, A. A., Pang, W. W., Jaiswal, S., Gibbs, K. D., Jr, van Rooijen, N., & Weissman, I. L. (2009). CD47 is an adverse prognostic factor and therapeutic antibody target on human acute myeloid leukemia stem cells. *Cell*, *138*(2), 286–299. https://doi.org/10.1016/j.cell.2009.05.045

Manna, P. P., & Frazier, W. A. (2003). The mechanism of CD47-dependent killing of T cells: heterotrimeric Gi-dependent inhibition of protein kinase A. *Journal of immunology (Baltimore, Md. : 1950)*, *170*(7), 3544–3553. https://doi.org/10.4049/jimmunol.170.7.3544

Manna, P. P., & Frazier, W. A. (2004). CD47 mediates killing of breast tumor cells via Gi-dependent inhibition of protein kinase A. *Cancer research*, *64*(3), 1026–1036. https://doi.org/10.1158/0008-5472.can-03-1708

Margolis, S. R., Wilson, S. C., & Vance, R. E. (2017). Evolutionary Origins of cGAS-STING Signaling. *Trends in immunology*, *38*(10), 733–743. https://doi.org/10.1016/j.it.2017.03.004

Marnik, E. A., & Updike, D. L. (2019). Membraneless organelles: P granules in Caenorhabditis elegans. *Traffic (Copenhagen, Denmark)*, *20*(6), 373–379. https://doi.org/10.1111/tra.12644

Martin, M., Hiroyasu, A., Guzman, R. M., Roberts, S. A., & Goodman, A. G. (2018). Analysis of Drosophila STING Reveals an Evolutionarily Conserved Antimicrobial Function. *Cell reports*, *23*(12), 3537–3550.e6. https://doi.org/10.1016/j.celrep.2018.05.029

Mateo, V., Lagneaux, L., Bron, D., Biron, G., Armant, M., Delespesse, G., & Sarfati, M. (1999). CD47 ligation induces caspase-independent cell death in chronic lymphocytic leukemia. *Nature medicine*, *5*(11), 1277–1284. https://doi.org/10.1038/15233

Mawby, W. J., Holmes, C. H., Anstee, D. J., Spring, F. A., & Tanner, M. J. (1994). Isolation and characterization of CD47 glycoprotein: a multispanning membrane protein which is the same as integrin-associated protein (IAP) and the ovarian tumour marker OA3. *The Biochemical journal*, *304 (Pt 2)*(Pt 2), 525–530. https://doi.org/10.1042/bj3040525

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography*, *40*(Pt 4), 658–674. https://doi.org/10.1107/S0021889807021206

McNab, F., Mayer-Barber, K., Sher, A., Wack, A., & O'Garra, A. (2015). Type I interferons in infectious disease. *Nature reviews. Immunology*, *15*(2), 87–103. https://doi.org/10.1038/nri3787

Melarkode Vattekatte, A., Shinada, N. K., Narwani, T. J., Noël, F., Bertrand, O., Meyniel, J. P., Malpertuy, A., Gelly, J. C., Cadet, F., & de Brevern, A. G. (2020). Discrete analysis of camelid variable domains: sequences, structures, and in-silico structure prediction. *PeerJ*, *8*, e8408. https://doi.org/10.7717/peerj.8408

Michalski, S., de Oliveira Mann, C. C., Stafford, C. A., Witte, G., Bartho, J., Lammens, K., Hornung, V., & Hopfner, K. P. (2020). Structural basis for sequestration and autoinhibition of cGAS by chromatin. *Nature*, *587*(7835), 678–682. https://doi.org/10.1038/s41586-020-2748-0

Miller-Vedam, L. E., Bräuning, B., Popova, K. D., Schirle Oakdale, N. T., Bonnar, J. L., Prabu, J. R., Boydston, E. A., Sevillano, N., Shurtleff, M. J., Stroud, R. M., Craik, C. S., Schulman, B. A., Frost, A., & Weissman, J. S. (2020). Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. *eLife*, *9*, e62611. https://doi.org/10.7554/eLife.62611

Morehouse, B. R., Govande, A. A., Millman, A., Keszei, A., Lowey, B., Ofir, G., Shao, S., Sorek, R., & Kranzusch, P. J. (2020). STING cyclic dinucleotide sensing originated in bacteria. *Nature*, *586*(7829), 429–433. https://doi.org/10.1038/s41586-020-2719-5

Moretti, J., Roy, S., Bozec, D., Martinez, J., Chapman, J. R., Ueberheide, B., Lamming, D. W., Chen, Z. J., Horng, T., Yeretssian, G., Green, D. R., & Blander, J. M. (2017). STING Senses Microbial Viability to Orchestrate Stress-Mediated Autophagy of the Endoplasmic Reticulum. *Cell*, *171*(4), 809–823.e13. https://doi.org/10.1016/j.cell.2017.09.034

Motwani, M., Pawaria, S., Bernier, J., Moses, S., Henry, K., Fang, T., Burkly, L., Marshak-Rothstein, A., & Fitzgerald, K. A. (2019b). Hierarchy of clinical manifestations in SAVI N153S and V154M mouse models. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(16), 7941–7950. https://doi.org/10.1073/pnas.1818281116

Motwani, M., Pesiridis, S., & Fitzgerald, K. A. (2019a). DNA sensing by the cGAS-STING pathway in health and disease. *Nature reviews. Genetics*, *20*(11), 657–674. https://doi.org/10.1038/s41576-019-0151-1

Mukai, K., Konno, H., Akiba, T., Uemura, T., Waguri, S., Kobayashi, T., Barber, G. N., Arai, H., & Taguchi, T. (2016). Activation of STING requires palmitoylation at the Golgi. *Nature communications*, *7*, 11932. https://doi.org/10.1038/ncomms11932

Mukherjee, S., Erramilli, S. K., Ammirati, M., Alvarez, F., Fennell, K. F., Purdy, M. D., Skrobek, B. M., Radziwon, K., Coukos, J., Kang, Y., Dutka, P., Gao, X., Qiu, X., Yeager, M., Eric Xu, H., Han, S., & Kossiakoff, A. A. (2020). Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryoEM structure determination of membrane proteins. *Nature communications*, *11*(1), 1598. https://doi.org/10.1038/s41467-020-15363-0

Murata, T., Ohnishi, H., Okazawa, H., Murata, Y., Kusakari, S., Hayashi, Y., Miyashita, M., Itoh, H., Oldenborg, P. A., Furuya, N., & Matozaki, T. (2006). CD47 promotes neuronal development through Src- and FRG/Vav2-mediated activation of Rac and Cdc42. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *26*(48), 12397–12407. https://doi.org/10.1523/JNEUROSCI.3981-06.2006

Muskardin, T., & Niewold, T. B. (2018). Type I interferon in rheumatic diseases. *Nature reviews. Rheumatology*, *14*(4), 214–228. https://doi.org/10.1038/nrrheum.2018.31

Nandakumar, R., Tschismarov, R., Meissner, F., Prabakaran, T., Krissanaprasit, A., Farahani, E., Zhang, B. C., Assil, S., Martin, A., Bertrams, W., Holm, C. K., Ablasser, A., Klause, T., Thomsen, M. K., Schmeck, B., Howard, K. A., Henry, T., Gothelf, K. V., Decker, T., & Paludan, S. R. (2019). Intracellular bacteria engage a STING-TBK1-MVB12b pathway to enable paracrine cGAS-STING signalling. *Nature microbiology*, *4*(4), 701–713. https://doi.org/10.1038/s41564-019-0367-z

Napirei, M., Karsunky, H., Zevnik, B., Stephan, H., Mannherz, H. G., & Möröy, T. (2000). Features of systemic lupus erythematosus in Dnase1-deficient mice. *Nature genetics*, *25*(2), 177–181. https://doi.org/10.1038/76032

Ni, G., Konno, H., & Barber, G. N. (2017). Ubiquitination of STING at lysine 224 controls IRF3 activation. *Science immunology*, *2*(11), eaah7119. https://doi.org/10.1126/sciimmunol.aah7119

Ning, X., Wang, Y., Jing, M., Sha, M., Lv, M., Gao, P., Zhang, R., Huang, X., Feng, J. M., & Jiang, Z. (2019). Apoptotic Caspases Suppress Type I Interferon Production via the Cleavage of cGAS, MAVS, and IRF3. *Molecular cell*, *74*(1), 19–31.e7. https://doi.org/10.1016/j.molcel.2019.02.013

Oldenborg P. A. (2013). CD47: A Cell Surface Glycoprotein Which Regulates Multiple Functions of Hematopoietic Cells in Health and Disease. *ISRN hematology*, *2013*, 614619. https://doi.org/10.1155/2013/614619

Orazizadeh, M., Lee, H. S., Groenendijk, B., Sadler, S. J., Wright, M. O., Lindberg, F. P., & Salter, D. M. (2008). CD47 associates with alpha 5 integrin and regulates responses of human articular chondrocytes to mechanical stimulation in an in vitro model. *Arthritis research & therapy*, *10*(1), R4. https://doi.org/10.1186/ar2350

Ouyang, S., Song, X., Wang, Y., Ru, H., Shaw, N., Jiang, Y., Niu, F., Zhu, Y., Qiu, W., Parvatiyar, K., Li, Y., Zhang, R., Cheng, G., & Liu, Z. J. (2012). Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding. *Immunity*, *36*(6), 1073–1086. https://doi.org/10.1016/j.immuni.2012.03.019

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein science : a publication of the Protein Society, 4(11), 2411–2423. https://doi.org/10.1002/pro.5560041120

Paludan, S. R., Reinert, L. S., & Hornung, V. (2019). DNA-stimulated cell death: implications for host defense, inflammatory diseases and cancer. *Nature reviews. Immunology*, *19*(3), 141–153. https://doi.org/10.1038/s41577-018-0117-0

Parkos, C. A., Colgan, S. P., Liang, T. W., Nusrat, A., Bacarra, A. E., Carnes, D. K., & Madara, J. L. (1996). CD47 mediates post-adhesive events required for neutrophil migration across polarized intestinal epithelia. *The Journal of cell biology*, *132*(3), 437–450. https://doi.org/10.1083/jcb.132.3.437

Pathare, G. R., Nagy, I., Śledź, P., Anderson, D. J., Zhou, H. J., Pardon, E., Steyaert, J., Förster, F., Bracher, A., & Baumeister, W. (2014). Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(8), 2984–2989. https://doi.org/10.1073/pnas.1400546111

Peschke, K., Friebe, F., Zimmermann, N., Wahlicht, T., Schumann, T., Achleitner, M., Berndt, N., Luksch, H., Behrendt, R., Lee-Kirsch, M. A., Roers, A., & Günther, C. (2014). Deregulated type I IFN response in TREX1-associated familial chilblain lupus. *The Journal of investigative dermatology*, *134*(5), 1456–1459. https://doi.org/10.1038/jid.2013.496

References

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Piccio, L., Vermi, W., Boles, K. S., Fuchs, A., Strader, C. A., Facchetti, F., Cella, M., & Colonna, M. (2005). Adhesion of human T cells to antigen-presenting cells through SIRPbeta2-CD47 interaction costimulates T-cell proliferation. *Blood*, *105*(6), 2421–2427. https://doi.org/10.1182/blood-2004-07-2823

Pokatayev, V., Yang, K., Tu, X., Dobbs, N., Wu, J., Kalb, R. G., & Yan, N. (2020). Homeostatic regulation of STING protein at the resting state by stabilizer TOLLIP. *Nature immunology*, *21*(2), 158–167. https://doi.org/10.1038/s41590-019-0569-9

Postis, V., Rawson, S., Mitchell, J. K., Lee, S. C., Parslow, R. A., Dafforn, T. R., Baldwin, S. A., & Muench, S. P. (2015). The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochimica et biophysica acta*, *1848*(2), 496–501. https://doi.org/10.1016/j.bbamem.2014.10.018

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

Punjani, A., Zhang, H., & Fleet, D. J. (2020). Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nature methods*, *17*(12), 1214–1221. https://doi.org/10.1038/s41592-020-00990-8

Qin, Y., Zhou, M. T., Hu, M. M., Hu, Y. H., Zhang, J., Guo, L., Zhong, B., & Shu, H. B. (2014). RNF26 temporally regulates virus-triggered type I interferon induction by two distinct mechanisms. *PLoS pathogens*, *10*(9), e1004358. https://doi.org/10.1371/journal.ppat.1004358

Rader C. (2009). Overview on concepts and applications of Fab antibody fragments. *Current protocols in protein science*, *Chapter 6*, . https://doi.org/10.1002/0471140864.ps0609s55

Ramanjulu, J. M., Pesiridis, G. S., Yang, J., Concha, N., Singhaus, R., Zhang, S. Y., Tran, J. L., Moore, P., Lehmann, S., Eberl, H. C., Muelbaier, M., Schneck, J. L., Clemens, J., Adam, M., Mehlmann, J., Romano, J., Morales, A., Kang, J., Leister, L., Graybill, T. L., … Bertin, J. (2018). Design of amidobenzimidazole STING receptor agonists with systemic activity. *Nature*, *564*(7736), 439–443. https://doi.org/10.1038/s41586-018-0705-y

Rasmussen, S. G., Choi, H. J., Fung, J. J., Pardon, E., Casarosa, P., Chae, P. S., Devree, B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I., & Kobilka, B. K. (2011). Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. *Nature*, *469*(7329), 175–180. https://doi.org/10.1038/nature09648

Rasmussen, S. G., Choi, H. J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R., Sanishvili, R., Fischetti, R. F., Schertler, G. F., Weis, W. I., & Kobilka, B. K. (2007). Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*, *450*(7168), 383–387. https://doi.org/10.1038/nature06325

Rebres, R. A., Vaz, L. E., Green, J. M., & Brown, E. J. (2001). Normal ligand binding and signaling by CD47 (integrin-associated protein) requires a long range disulfide bond between the extracellular and membrane-spanning domains. *The Journal of biological chemistry*, *276*(37), 34607–34616. https://doi.org/10.1074/jbc.M106107200

Reinhold, M. I., Lindberg, F. P., Plas, D., Reynolds, S., Peters, M. G., & Brown, E. J. (1995). In vivo expression of alternatively spliced forms of integrin-associated protein (CD47). *Journal of cell science*, *108 (Pt 11)*, 3419–3425. https://doi.org/10.1242/jcs.108.11.3419

Riley, J. S., Quarato, G., Cloix, C., Lopez, J., O'Prey, J., Pearson, M., Chapman, J., Sesaki, H., Carlin, L. M., Passos, J. F., Wheeler, A. P., Oberst, A., Ryan, K. M., & Tait, S. W. (2018). Mitochondrial inner membrane permeabilisation enables mtDNA release during apoptosis. *The EMBO journal*, *37*(17), e99238. https://doi.org/10.15252/embj.201899238

Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I. G., Zolnerciks, J. K., Atkins, W. M., & Sligar, S. G. (2009). Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in enzymology*, *464*, 211–231. https://doi.org/10.1016/S0076-6879(09)64011-8

Rodero, M. P., Tesser, A., Bartok, E., Rice, G. I., Della Mina, E., Depp, M., Beitz, B., Bondet, V., Cagnard, N., Duffy, D., Dussiot, M., Frémond, M. L., Gattorno, M., Guillem, F., Kitabayashi, N., Porcheray, F., Rieux-Laucat, F., Seabra, L., Uggenti, C., Volpi, S., … Crow, Y. J. (2017). Type I interferon-mediated autoinflammation due to DNase II deficiency. *Nature communications*, *8*(1), 2176. https://doi.org/10.1038/s41467-017-01932-3

Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of structural biology*, *192*(2), 216–221. https://doi.org/10.1016/j.jsb.2015.08.008

Sallman DA, Donnellan WB, Asch AS, et al. (2019). The first-in-class anti-CD47 antibody Hu5F9-G4 is active and well tolerated alone or with azacitidine in AML and MDS patients: initial phase 1b results. *Journal of Clinical Oncology* 37, no. 15_suppl (May 20, 2019) 7009-7009

Santa, P., Garreau, A., Serpas, L., Ferriere, A., Blanco, P., Soni, C., & Sisirak, V. (2021). The Role of Nucleases and Nucleic Acid Editing Enzymes in the Regulation of Self-Nucleic Acid Sensing. *Frontiers in immunology*, *12*, 629922. https://doi.org/10.3389/fimmu.2021.629922

Schaefer L. (2014). Complexity of danger: the diverse nature of damage-associated molecular patterns. *The Journal of biological chemistry*, *289*(51), 35237–35245. https://doi.org/10.1074/jbc.R114.619304

Schock, S. N., Chandra, N. V., Sun, Y., Irie, T., Kitagawa, Y., Gotoh, B., Coscoy, L., & Winoto, A. (2017). Induction of necroptotic cell death by viral activation of the RIG-I or STING pathway. *Cell death and differentiation*, *24*(4), 615–625. https://doi.org/10.1038/cdd.2016.153

Schroeder, H. W., Jr, & Cavacini, L. (2010). Structure and function of immunoglobulins. *The Journal of allergy and clinical immunology*, *125*(2 Suppl 2), S41–S52. https://doi.org/10.1016/j.jaci.2009.09.046

Seo, G. J., Kim, C., Shin, W. J., Sklan, E. H., Eoh, H., & Jung, J. U. (2018). TRIM56-mediated monoubiquitination of cGAS for cytosolic DNA sensing. *Nature communications*, *9*(1), 613. https://doi.org/10.1038/s41467-018-02936-3

Seo, G. J., Yang, A., Tan, B., Kim, S., Liang, Q., Choi, Y., Yuan, W., Feng, P., Park, H. S., & Jung, J. U. (2015). Akt Kinase-Mediated Checkpoint of cGAS DNA Sensing Pathway. *Cell reports*, *13*(2), 440–449. https://doi.org/10.1016/j.celrep.2015.09.007

Shang, G., Zhang, C., Chen, Z. J., Bai, X. C., & Zhang, X. (2019). Cryo-EM structures of STING reveal its mechanism of activation by cyclic GMP-AMP. *Nature*, *567*(7748), 389–393. https://doi.org/10.1038/s41586-019-0998-5

Shang, G., Zhu, D., Li, N., Zhang, J., Zhu, C., Lu, D., Liu, C., Yu, Q., Zhao, Y., Xu, S., & Gu, L. (2012). Crystal structures of STING protein reveal basis for recognition of cyclic di-GMP. *Nature structural & molecular biology*, *19*(7), 725–727. https://doi.org/10.1038/nsmb.2332

Shen, J., Zhao, W., Ju, Z., Wang, L., Peng, Y., Labrie, M., Yap, T. A., Mills, G. B., & Peng, G. (2019). PARPi Triggers the STING-Dependent Immune Response and Enhances the Therapeutic Efficacy of Immune Checkpoint Blockade Independent of BRCAness. *Cancer research*, *79*(2), 311–319. https://doi.org/10.1158/0008-5472.CAN-18-1003

Shepard, H. M., Phillips, G. L., D Thanos, C., & Feldmann, M. (2017). Developments in therapy with monoclonal antibodies and related proteins. *Clinical medicine (London, England)*, *17*(3), 220–232. https://doi.org/10.7861/clinmedicine.17-3-220

Shi, H., Wu, J., Chen, Z. J., & Chen, C. (2015). Molecular basis for the specific recognition of the metazoan cyclic GMP-AMP by the innate immune adaptor protein STING. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), 8947–8952. https://doi.org/10.1073/pnas.1507317112

Shi, X., & Jarvis, D. L. (2007). Protein N-glycosylation in the baculovirus-insect cell system. *Current drug targets*, *8*(10), 1116–1125. https://doi.org/10.2174/138945007782151360

Shu, C., Yi, G., Watts, T., Kao, C. C., & Li, P. (2012). Structure of STING bound to cyclic di-GMP reveals the mechanism of cyclic dinucleotide recognition by the immune system. *Nature structural & molecular biology*, *19*(7), 722–724. https://doi.org/10.1038/nsmb.2331

Singh, R., Letai, A., & Sarosiek, K. (2019). Regulation of apoptosis in health and disease: the balancing act of BCL-2 family proteins. *Nature reviews. Molecular cell biology*, *20*(3), 175–193. https://doi.org/10.1038/s41580-018-0089-8

Sisirak, V., Sally, B., D'Agati, V., Martinez-Ortiz, W., Özçakar, Z. B., David, J., Rashidfarrokhi, A., Yeste, A., Panea, C., Chida, A. S., Bogunovic, M., Ivanov, I. I., Quintana, F. J., Sanz, I., Elkon, K. B., Tekin, M., Yalçınkaya, F., Cardozo, T. J., Clancy, R. M., Buyon, J. P., … Reizis, B. (2016). Digestion of Chromatin in Apoptotic Cell Microparticles Prevents Autoimmunity. *Cell*, *166*(1), 88–101. https://doi.org/10.1016/j.cell.2016.05.034

Skouboe, M. K., Knudsen, A., Reinert, L. S., Boularan, C., Lioux, T., Perouzel, E., Thomsen, M. K., & Paludan, S. R. (2018). STING agonists enable antiviral cross-talk between human cells and confer protection against genital herpes in mice. *PLoS pathogens*, *14*(4), e1006976. https://doi.org/10.1371/journal.ppat.1006976

Slater, A., Di, Y., Clark, J. C., Jooss, N. J., Martin, E. M., Alenazy, F., Thomas, M. R., Ariëns, R. A. S., Herr, A. B., Poulter, N. S., Emsley, J., & Watson, S. P. (2021). Structural characterization of a novel GPVI-nanobody complex reveals a biologically active domain-swapped GPVI dimer. *Blood*, *137*(24), 3443–3453. https://doi.org/10.1182/blood.2020009440

Sliter, D. A., Martinez, J., Hao, L., Chen, X., Sun, N., Fischer, T. D., Burman, J. L., Li, Y., Zhang, Z., Narendra, D. P., Cai, H., Borsche, M., Klein, C., & Youle, R. J. (2018). Parkin and PINK1 mitigate STING-induced inflammation. *Nature*, *561*(7722), 258–262. https://doi.org/10.1038/s41586-018-0448-9

Soto-Pantoja, D. R., Kaur, S., & Roberts, D. D. (2015). CD47 signaling pathways controlling cellular differentiation and responses to stress. *Critical reviews in biochemistry and molecular biology*, *50*(3), 212–230. https://doi.org/10.3109/10409238.2015.1014024

Srikanth, S., Woo, J. S., Wu, B., El-Sherbiny, Y. M., Leung, J., Chupradit, K., Rice, L., Seo, G. J., Calmettes, G., Ramakrishna, C., Cantin, E., An, D. S., Sun, R., Wu, T. T., Jung, J. U., Savic, S., & Gwack, Y. (2019). The Ca^{2+} sensor STIM1 regulates the type I interferon response by retaining the signaling adaptor STING at the endoplasmic reticulum. *Nature immunology*, *20*(2), 152–162. https://doi.org/10.1038/s41590-018-0287-8

Srinivasan, L., Alzogaray, V., Selvakumar, D., Nathan, S., Yoder, J. B., Wright, K. M., Klinke, S., Nwafor, J. N., Labanda, M. S., Goldbaum, F. A., Schön, A., Freire, E., Tomaselli, G. F., Amzel, L. M., Ben-Johny, M., & Gabelli, S. B. (2022). Development of high-affinity nanobodies specific for $Na_V1.4$ and $Na_V1.5$ voltage-gated sodium channel isoforms. *The Journal of biological chemistry*, *298*(4), 101763. https://doi.org/10.1016/j.jbc.2022.101763

Stein N. (2008). CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. J Appl Crystallogr. 41:641–643.

Stetsenko, A., & Guskov, A. (2017). An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization. *Crystals*, *7*(7), 1-16. [197]. https://doi.org/10.3390/cryst7070197

Stetson, D. B., Ko, J. S., Heidmann, T., & Medzhitov, R. (2008). Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell*, *134*(4), 587–598. https://doi.org/10.1016/j.cell.2008.06.032

Subramanian, S., Boder, E. T., & Discher, D. E. (2007). Phylogenetic divergence of CD47 interactions with human signal regulatory protein alpha reveals locus of species specificity. Implications for the binding site. *The Journal of biological chemistry*, *282*(3), 1805–1818. https://doi.org/10.1074/jbc.M603923200

Sukhithasri, V., Nisha, N., Biswas, L., Anil Kumar, V., & Biswas, R. (2013). Innate immune recognition of microbial cell wall components and microbial strategies to evade such recognitions. *Microbiological research*, *168*(7), 396–406. https://doi.org/10.1016/j.micres.2013.02.005

Sun S. C. (2011). Non-canonical NF-κB signaling pathway. *Cell research*, *21*(1), 71–85. https://doi.org/10.1038/cr.2010.177

Sun S. C. (2017). The non-canonical NF-κB pathway in immunity and inflammation. *Nature reviews. Immunology*, *17*(9), 545–558. https://doi.org/10.1038/nri.2017.52

Sun, L., Wu, J., Du, F., Chen, X., & Chen, Z. J. (2013). Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science (New York, N.Y.)*, *339*(6121), 786–791. https://doi.org/10.1126/science.1232458

Sun, W., Li, Y., Chen, L., Chen, H., You, F., Zhou, X., Zhou, Y., Zhai, Z., Chen, D., & Jiang, Z. (2009). ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(21), 8653–8658. https://doi.org/10.1073/pnas.0900850106

Syrjanen, J. L., Michalski, K., Chou, T. H., Grant, T., Rao, S., Simorowski, N., Tucker, S. J., Grigorieff, N., & Furukawa, H. (2020). Structure and assembly of calcium homeostasis modulator proteins. *Nature structural & molecular biology*, *27*(2), 150–159. https://doi.org/10.1038/s41594-019-0369-9

Sze, A., Belgnaoui, S. M., Olagnier, D., Lin, R., Hiscott, J., & van Grevenynghe, J. (2013). Host restriction factor SAMHD1 limits human T cell leukemia virus type 1 infection of monocytes via STING-mediated apoptosis. *Cell host & microbe*, *14*(4), 422–434. https://doi.org/10.1016/j.chom.2013.09.009

Tait, S. W., & Green, D. R. (2013). Mitochondrial regulation of cell death. *Cold Spring Harbor perspectives in biology*, *5*(9), a008706. https://doi.org/10.1101/cshperspect.a008706

Takaoka, A., Wang, Z., Choi, M. K., Yanai, H., Negishi, H., Ban, T., Lu, Y., Miyagishi, M., Kodama, T., Honda, K., Ohba, Y., & Taniguchi, T. (2007). DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature*, *448*(7152), 501–505. https://doi.org/10.1038/nature06013

Tan, X., Sun, L., Chen, J., & Chen, Z. J. (2018). Detection of Microbial Infections Through Innate Immune Sensing of Nucleic Acids. *Annual review of microbiology*, *72*, 447–478. https://doi.org/10.1146/annurev-micro-102215-095605

Tanaka, Y., & Chen, Z. J. (2012). STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway. *Science signaling*, *5*(214), ra20. https://doi.org/10.1126/scisignal.2002521

Tang, C. H., Zundell, J. A., Ranatunga, S., Lin, C., Nefedova, Y., Del Valle, J. R., & Hu, C. C. (2016). Agonist-Mediated Activation of STING Induces Apoptosis in Malignant B Cells. *Cancer research*, *76*(8), 2137–2152. https://doi.org/10.1158/0008-5472.CAN-15-1885

Thim-Uam, A., Prabakaran, T., Tansakul, M., Makjaroen, J., Wongkongkathep, P., Chantaravisoot, N., Saethang, T., Leelahavanichkul, A., Benjachat, T., Paludan, S., Pisitkun, T., & Pisitkun, P. (2020). STING Mediates Lupus via the Activation of Conventional Dendritic Cell Maturation and Plasmacytoid Dendritic Cell Differentiation. *iScience*, *23*(9), 101530. https://doi.org/10.1016/j.isci.2020.101530

Tribet, C., Audebert, R., & Popot, J. L. (1996). Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(26), 15047–15050. https://doi.org/10.1073/pnas.93.26.15047

Tsuchida, T., Zou, J., Saitoh, T., Kumar, H., Abe, T., Matsuura, Y., Kawai, T., & Akira, S. (2010). The ubiquitin ligase TRIM56 regulates innate immune responses to intracellular double-stranded DNA. *Immunity*, *33*(5), 765–776. https://doi.org/10.1016/j.immuni.2010.10.013

Tsutsumi, N., Mukherjee, S., Waghray, D., Janda, C. Y., Jude, K. M., Miao, Y., Burg, J. S., Aduri, N. G., Kossiakoff, A. A., Gati, C., & Garcia, K. C. (2020). Structure of human Frizzled5 by fiducial-assisted cryo-EM supports a heterodimeric mechanism of canonical Wnt signaling. *eLife*, *9*, e58464. https://doi.org/10.7554/eLife.58464

Uchański, T., Masiulis, S., Fischer, B., Kalichuk, V., López-Sánchez, U., Zarkadas, E., Weckener, M., Sente, A., Ward, P., Wohlkönig, A., Zögg, T., Remaut, H., Naismith, J. H., Nury, H., Vranken, W., Aricescu, A. R., Pardon, E., & Steyaert, J. (2021). Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nature methods*, *18*(1), 60–68. https://doi.org/10.1038/s41592-020-01001-6

Unterholzner, L., Keating, S. E., Baran, M., Horan, K. A., Jensen, S. B., Sharma, S., Sirois, C. M., Jin, T., Latz, E., Xiao, T. S., Fitzgerald, K. A., Paludan, S. R., & Bowie, A. G. (2010). IFI16 is an innate immune sensor for intracellular DNA. *Nature immunology*, *11*(11), 997–1004. https://doi.org/10.1038/ni.1932

Upton, R., Banuelos, A., Feng, D., Biswas, T., Kao, K., McKenna, K., Willingham, S., Ho, P. Y., Rosental, B., Tal, M. C., Raveh, T., Volkmer, J. P., Pegram, M. D., & Weissman, I. L. (2021). Combining CD47 blockade with trastuzumab eliminates HER2-positive breast cancer cells and overcomes trastuzumab tolerance. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(29), e2026849118. https://doi.org/10.1073/pnas.2026849118

Van Dis, E., Sogi, K. M., Rae, C. S., Sivick, K. E., Surh, N. H., Leong, M. L., Kanne, D. B., Metchette, K., Leong, J. J., Bruml, J. R., Chen, V., Heydari, K., Cadieux, N., Evans, T., McWhirter, S. M., Dubensky, T. W., Jr, Portnoy, D. A., & Stanley, S. A. (2018). STING-Activating Adjuvants Elicit a Th17 Immune Response and Protect against Mycobacterium tuberculosis Infection. *Cell reports*, *23*(5), 1435–1447. https://doi.org/10.1016/j.celrep.2018.04.003

van Meer, G., & de Kroon, A. I. (2011). Lipid map of the mammalian cell. *Journal of cell science*, *124*(Pt 1), 5–8. https://doi.org/10.1242/jcs.071233

van Ravenswaay Claasen, H. H., Eggermont, A. M., Nooyen, Y. A., Warnaar, S. O., & Fieuren, G. J. (1994). Immunotherapy in a human ovarian cancer xenograft model with two bispecific monoclonal antibodies: OV-TL 3/CD3 and OC/TR. *Gynecologic oncology*, *52*(2), 199–206. https://doi.org/10.1006/gyno.1994.1031

Verhelst, K., Verstrepen, L., Carpentier, I., & Beyaert, R. (2013). IκB kinase ε (IKKε): a therapeutic target in inflammation and cancer. *Biochemical pharmacology*, *85*(7), 873–880. https://doi.org/10.1016/j.bcp.2013.01.007

Verhelst, K., Verstrepen, L., Carpentier, I., & Beyaert, R. (2013). IκB kinase ε (IKKε): a therapeutic target in inflammation and cancer. *Biochemical pharmacology*, *85*(7), 873–880. https://doi.org/10.1016/j.bcp.2013.01.007

Volkman, H. E., Cambier, S., Gray, E. E., & Stetson, D. B. (2019). Tight nuclear tethering of cGAS is essential for preventing autoreactivity. *eLife*, *8*, e47491. https://doi.org/10.7554/eLife.47491

Wang, F., Gómez-Sintes, R., & Boya, P. (2018). Lysosomal membrane permeabilization and cell death. *Traffic (Copenhagen, Denmark)*, *19*(12), 918–931. https://doi.org/10.1111/tra.12613

Wang, Y., Ning, X., Gao, P., Wu, S., Sha, M., Lv, M., Zhou, X., Gao, J., Fang, R., Meng, G., Su, X., & Jiang, Z. (2017a). Inflammasome Activation Triggers Caspase-1-Mediated Cleavage of cGAS to Regulate Responses to DNA Virus Infection. *Immunity*, *46*(3), 393–404. https://doi.org/10.1016/j.immuni.2017.02.011

Wang, Q., Huang, L., Hong, Z., Lv, Z., Mao, Z., Tang, Y., Kong, X., Li, S., Cui, Y., Liu, H., Zhang, L., Zhang, X., Jiang, L., Wang, C., & Zhou, Q. (2017b). The E3 ubiquitin ligase RNF185 facilitates the cGAS-mediated innate immune response. *PLoS pathogens*, *13*(3), e1006264. https://doi.org/10.1371/journal.ppat.1006264

Wang, Q., Liu, X., Cui, Y., Tang, Y., Chen, W., Li, S., Yu, H., Pan, Y., & Wang, C. (2014). The E3 ubiquitin ligase AMFR and INSIG1 bridge the activation of TBK1 kinase by modifying the adaptor STING. *Immunity*, *41*(6), 919–933. https://doi.org/10.1016/j.immuni.2014.11.011

Wang, X. Q., & Frazier, W. A. (1998). The thrombospondin receptor CD47 (IAP) modulates and associates with alpha2 beta1 integrin in vascular smooth muscle cells. *Molecular biology of the cell*, *9*(4), 865–874. https://doi.org/10.1091/mbc.9.4.865

Warner, J. D., Irizarry-Caro, R. A., Bennion, B. G., Ai, T. L., Smith, A. M., Miner, C. A., Sakai, T., Gonugunta, V. K., Wu, J., Platt, D. J., Yan, N., & Miner, J. J. (2017). STING-associated vasculopathy develops independently of IRF3 in mice. *The Journal of experimental medicine*, *214*(11), 3279–3292. https://doi.org/10.1084/jem.20171351

Watson, R. O., Manzanillo, P. S., & Cox, J. S. (2012). Extracellular M. tuberculosis DNA targets bacteria for autophagy by activating the host DNA-sensing pathway. *Cell*, *150*(4), 803–815. https://doi.org/10.1016/j.cell.2012.06.040

Whiteley, A. T., Eaglesham, J. B., de Oliveira Mann, C. C., Morehouse, B. R., Lowey, B., Nieminen, E. A., Danilchanka, O., King, D. S., Lee, A., Mekalanos, J. J., & Kranzusch, P. J. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature*, *567*(7747), 194–199. https://doi.org/10.1038/s41586-019-0953-5

Whittaker, G. R., Kann, M., & Helenius, A. (2000). Viral entry into the nucleus. *Annual review of cell and developmental biology*, *16*, 627–651. https://doi.org/10.1146/annurev.cellbio.16.1.627

Woo, S. R., Fuertes, M. B., Corrales, L., Spranger, S., Furdyna, M. J., Leung, M. Y., Duggan, R., Wang, Y., Barber, G. N., Fitzgerald, K. A., Alegre, M. L., & Gajewski, T. F. (2014). STING-dependent cytosolic DNA sensing mediates innate immune recognition of immunogenic tumors. *Immunity*, *41*(5), 830–842. https://doi.org/10.1016/j.immuni.2014.10.017

Wu, A. L., Wang, J., Zheleznyak, A., & Brown, E. J. (1999). Ubiquitin-related proteins regulate interaction of vimentin intermediate filaments with the plasma membrane. *Molecular cell*, *4*(4), 619–625. https://doi.org/10.1016/s1097-2765(00)80212-9

Wu, X., Wu, F. H., Wang, X., Wang, L., Siedow, J. N., Zhang, W., & Pei, Z. M. (2014). Molecular evolutionary and structural analysis of the cytosolic DNA sensor cGAS and STING. *Nucleic acids research*, *42*(13), 8243–8257. https://doi.org/10.1093/nar/gku569

Xia, P., Ye, B., Wang, S., Zhu, X., Du, Y., Xiong, Z., Tian, Y., & Fan, Z. (2016). Glutamylation of the DNA sensor cGAS regulates its binding and synthase activity in antiviral immunity. *Nature immunology*, *17*(4), 369–378. https://doi.org/10.1038/ni.3356

Xiao, Q., McAtee, C. K., & Su, X. (2021). Phase separation in immune signalling. *Nature reviews. Immunology*, 10.1038/s41577-021-00572-5. Advance online publication. https://doi.org/10.1038/s41577-021-00572-5

Xie, J., Li, Y., Shen, X., Goh, G., Zhu, Y., Cui, J., Wang, L. F., Shi, Z. L., & Zhou, P. (2018). Dampened STING-Dependent Interferon Activation in Bats. *Cell host & microbe*, *23*(3), 297–301.e4. https://doi.org/10.1016/j.chom.2018.01.006

Xie, W., Lama, L., Adura, C., Tomita, D., Glickman, J. F., Tuschl, T., & Patel, D. J. (2019). Human cGAS catalytic domain has an additional DNA-binding interface that enhances enzymatic activity and liquid-phase condensation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(24), 11946–11955. https://doi.org/10.1073/pnas.1905013116

Yamamoto, T., Kanuma, T., Takahama, S., Okamura, T., Moriishi, E., Ishii, K. J., Terahara, K., & Yasutomi, Y. (2019). STING agonists activate latently infected cells and enhance SIV-specific responses ex vivo in naturally SIV controlled cynomolgus macaques. *Scientific reports*, *9*(1), 5917. https://doi.org/10.1038/s41598-019-42253-3

Yamashiro, L. H., Wilson, S. C., Morrison, H. M., Karalis, V., Chung, J. J., Chen, K. J., Bateup, H. S., Szpara, M. L., Lee, A. Y., Cox, J. S., & Vance, R. E. (2020). Interferon-independent STING signaling promotes resistance to HSV-1 in vivo. *Nature communications*, *11*(1), 3382. https://doi.org/10.1038/s41467-020-17156-x

Yasutomo, K., Horiuchi, T., Kagami, S., Tsukamoto, H., Hashimura, C., Urushihara, M., & Kuroda, Y. (2001). Mutation of DNASE1 in people with systemic lupus erythematosus. *Nature genetics*, *28*(4), 313–314. https://doi.org/10.1038/91070

Yin, Q., Tian, Y., Kabaleeswaran, V., Jiang, X., Tu, D., Eck, M. J., Chen, Z. J., & Wu, H. (2012). Cyclic di-GMP sensing via the innate immune signaling protein STING. *Molecular cell*, *46*(6), 735–745. https://doi.org/10.1016/j.molcel.2012.05.029

Yoh, S. M., Schneider, M., Seifried, J., Soonthornvacharin, S., Akleh, R. E., Olivieri, K. C., De Jesus, P. D., Ruan, C., de Castro, E., Ruiz, P. A., Germanaud, D., des Portes, V., García-Sastre, A., König, R., & Chanda, S. K. (2015). PQBP1 Is a Proximal Sensor of the cGAS-Dependent Innate Response to HIV-1. *Cell*, *161*(6), 1293–1305. https://doi.org/10.1016/j.cell.2015.04.050

Yoshida, H., Tomiyama, Y., Ishikawa, J., Oritani, K., Matsumura, I., Shiraga, M., Yokota, T., Okajima, Y., Ogawa, M., Miyagawa, J. i., Nishiura, T., & Matsuzawa, Y. (2000). Integrin-associated protein/CD47 regulates motile activity in human B-cell lines through CDC42. *Blood*, *96*(1), 234–241.

Yu, H., Lin, L., Zhang, Z., Zhang, H., & Hu, H. (2020). Targeting NF-κB pathway for the therapy of diseases: mechanism and clinical study. *Signal transduction and targeted therapy*, *5*(1), 209. https://doi.org/10.1038/s41392-020-00312-6

Yum, S., Li, M., Fang, Y., & Chen, Z. J. (2021). TBK1 recruitment to STING activates both IRF3 and NF-κB that mediate immune defense against tumors and viral infections. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(14), e2100225118. https://doi.org/10.1073/pnas.2100225118

Zhang, C., Shang, G., Gui, X., Zhang, X., Bai, X. C., & Chen, Z. J. (2019). Structural basis of STING binding with and phosphorylation by TBK1. *Nature*, *567*(7748), 394–398. https://doi.org/10.1038/s41586-019-1000-2

Zhang, M., Gui, M., Wang, Z. F., Gorgulla, C., Yu, J. J., Wu, H., Sun, Z. J., Klenk, C., Merklinger, L., Morstein, L., Hagn, F., Plückthun, A., Brown, A., Nasr, M. L., & Wagner, G. (2021a). Cryo-EM structure of an activated GPCR-G protein complex in lipid nanodiscs. *Nature structural & molecular biology*, *28*(3), 258–267. https://doi.org/10.1038/s41594-020-00554-6

Zhang, H., Huang, C. S., Yu, X., Lee, J., Vaish, A., Chen, Q., Zhou, M., Wang, Z., & Min, X. (2021b). Cryo-EM structure of ABCG5/G8 in complex with modulating antibodies. *Communications biology*, *4*(1), 526. https://doi.org/10.1038/s42003-021-02039-8

Zhang, J., Hu, M. M., Wang, Y. Y., & Shu, H. B. (2012). TRIM32 protein modulates type I interferon induction and cellular antiviral response by targeting MITA/STING protein for K63-linked ubiquitination. *The Journal of biological chemistry*, *287*(34), 28646–28655. https://doi.org/10.1074/jbc.M112.362608

Zhang, L., Wei, N., Cui, Y., Hong, Z., Liu, X., Wang, Q., Li, S., Liu, H., Yu, H., Cai, Y., Wang, Q., Zhu, J., Meng, W., Chen, Z., & Wang, C. (2018). The deubiquitinase CYLD is a specific checkpoint of the STING antiviral signaling pathway. *PLoS pathogens*, *14*(11), e1007435. https://doi.org/10.1371/journal.ppat.1007435

Zhang, X., Shi, H., Wu, J., Zhang, X., Sun, L., Chen, C., & Chen, Z. J. (2013). Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Molecular cell*, *51*(2), 226–235. https://doi.org/10.1016/j.molcel.2013.05.022

Zhao, B., Du, F., Xu, P., Shu, C., Sankaran, B., Bell, S. L., Liu, M., Lei, Y., Gao, X., Fu, X., Zhu, F., Liu, Y., Laganowsky, A., Zheng, X., Ji, J. Y., West, A. P., Watson, R. O., & Li, P. (2019). A conserved PLPLRT/SD motif of STING mediates the recruitment and activation of TBK1. *Nature*, *569*(7758), 718–722. https://doi.org/10.1038/s41586-019-1228-x

Zhao, B., Shu, C., Gao, X., Sankaran, B., Du, F., Shelton, C. L., Herr, A. B., Ji, J. Y., & Li, P. (2016). Structural basis for concerted recruitment and activation of IRF-3 by innate immune adaptor proteins. *Proceedings of the*

National Academy of Sciences of the United States of America, *113*(24), E3403–E3412. https://doi.org/10.1073/pnas.1603269113

Zhao, Z., Ma, Z., Wang, B., Guan, Y., Su, X. D., & Jiang, Z. (2020). Mn^{2+} Directly Activates cGAS and Structural Analysis Suggests Mn^{2+} Induces a Noncanonical Catalytic Synthesis of 2'3'-cGAMP. *Cell reports*, *32*(7), 108053. https://doi.org/10.1016/j.celrep.2020.108053

Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y., & Agard, D. A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods*, *14*(4), 331–332. https://doi.org/10.1038/nmeth.4193

Zhong, B., Yang, Y., Li, S., Wang, Y. Y., Li, Y., Diao, F., Lei, C., He, X., Zhang, L., Tien, P., & Shu, H. B. (2008). The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. *Immunity*, *29*(4), 538–550. https://doi.org/10.1016/j.immuni.2008.09.003

Zhong, B., Zhang, L., Lei, C., Li, Y., Mao, A. P., Yang, Y., Wang, Y. Y., Zhang, X. L., & Shu, H. B. (2009). The ubiquitin ligase RNF5 regulates antiviral responses by mediating degradation of the adaptor protein MITA. *Immunity*, *30*(3), 397–407. https://doi.org/10.1016/j.immuni.2009.01.008

Zhou, Y., Morais-Cabral, J. H., Kaufman, A., & MacKinnon, R. (2001). Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 A resolution. *Nature*, *414*(6859), 43–48. https://doi.org/10.1038/35102009

Zhu, Q., Man, S. M., Karki, R., Malireddi, R., & Kanneganti, T. D. (2018). Detrimental Type I Interferon Signaling Dominates Protective AIM2 Inflammasome Responses during Francisella novicida Infection. *Cell reports*, *22*(12), 3168–3174. https://doi.org/10.1016/j.celrep.2018.02.096

Zhu, X. W., Wang, Y., Wei, Y. H., Zhao, P. P., Wang, X. B., Rong, J. J., Zhong, W. Y., Zhang, X. W., Wang, L., & Zheng, H. F. (2016). Comprehensive Assessment of the Association between FCGRs polymorphisms and the risk of systemic lupus erythematosus: Evidence from a Meta-Analysis. *Scientific reports*, *6*, 31617. https://doi.org/10.1038/srep31617

Zoghbi, M. & Altenberg, G. (2017). Membrane protein reconstitution in nanodiscs for luminescence spectroscopy studies. Nanotechnology Reviews, 6(1), 33-46. https://doi.org/10.1515/ntrev-2016-0078eksy

12. Abbreviations

4-1BBL - 4-1BB ligand
aa - residues
ABZI - amidobenzimidazole-based compound
ADAR - adenosine deaminase acting on RNA
aF-NB - anti-Fab nanobody
AGS - Aicardi-Goutières syndrome
AIM2 - absent in melanoma 2
ALR - absent in melanoma 2-like receptor
AMBRA1 - activating molecule in Beclin 1-regulated autophagy protein 1
AMP - adenosine monophosphate
APAF1 - apoptotic protease-activating factor 1
ASU - asymmetric unit
ATG - autophagy-related protein
ATG13 - autophagy-related protein 13
ATH PCR - around-the-horn polymerase chain reaction
BAD - BCL-2-associated agonist of cell death
BAFFR - B-cell-activating factor receptor
BAK - BCL-2 antagonist/killer
BAX - BCL-2-associated X protein
BCL-2 - B-cell lymphoma 2
BCL-W - B cell lymphoma W
BCL-X_L - B cell lymphoma extra large
β-OG - n-octyl-β-D-glucopyranoside
BFL1 - BCL-2-related isolated from fetal liver
BID - BH3-interacting domain death agonist
BIM - BCL-2-interacting mediator of cell death
BLM - Bloom syndrome protein
BME - 2-mercaptoethanol
bp - base pair(s)
BSA - bovine serum albumin
c-di-GMP - cyclic diguanylate monophosphate
CALHM1 - calcium homeostasis modulator 1
cAMP - cyclic adenosine monophosphate
CARD - caspase activation and recruitment domain
CBD - cyclic dinucleotide-binding domain
CCP5 - cytosolic carboxypeptidase 5
CCR2 - C-C motif chemokine receptor 2
CD-NTase - cGAS/DncV-like nucleotidyltransferase
CD40 - cluster of differentiation 40
CD47 - cluster of differentiation 47
CDR - complementarity determining region
CDN - cyclic dinucleotide
cGAMP - cyclic guanosine monophosphate–adenosine monophosphate
cGAS - cyclic GMP-AMP synthase
CH - constant heavy

CHS - cholesteryl hemisuccinate tris salt
CHX - connector helix
cIAP1 - cellular inhibitor of apoptosis 1
CIN - chromosomal instability
CL - constant light
CLP - connector loop
CLR - C-type lectin receptor
CMC - critical micelle concentration
COPII - coat protein complex II
COVID-19 - coronavirus disease 2019
COX-2 - cyclooxygenase 2
cryo-EM - cryoelectron microscopy
CryoSPARC - Cryo-EM Single Particle Ab-Initio Reconstruction and Classification
cSrc - proto-oncogene tyrosine-protein kinase Src
CTD - carboxy-terminal domain
CTT - carboxy-terminal tail
CV - column volume
CXCL1 - C-X-C motif chemokine ligand 1
DAMP - damage-associated molecular pattern
ddH_2O - purified, deionized water
DDM - n-dodecyl-β-D-maltopyranoside
DESY - Deutsches Elektronen-Synchrotron (German Electron Synchrotron)
DFCP1 - zinc-finger FYVE domain-containing protein 1
DHHC3 - DHHC cysteine-rich domains-containing protein 3
diABZI-4 - diamidobenzimidazole 4
DMSO - dimethyl sulphoxide
DMXAA - 5,6-dimethylxanthenone-4-acetic acid
DNA - deoxyribonucleic acid
DNase 1 - deoxyribonuclease 1
DncV - dinucleotide cyclase in Vibrio
ds - double stranded
e1D1-NDs - non-tagged empty MSP1D1 nanodiscs
e1E3D1-NDs - his-tagged empty 1E3D1 nanodiscs
ECD - extracellular domain
ECL1 - extracellular loop 1
ECLR - extracellular loop region
ECS - extracellular space
edH5-NDs - his-tagged empty MSP1D1dH5 nanodiscs
EDTA - ethylenediaminetetraacetic acid
EL - elbow
ELISA - enzyme-linked immunosorbent assay
EM - electron microscopy
EMC - ER membrane protein complex
ENPP1 - ecto-nucleotide pyrophosphatase phosphodiestrase 1
ER - endoplasmic reticulum
ERGIC - ER-Golgi intermediate compartment
Fab - antigen binding fragment
FAK - focal adhesion kinase

FC1 - flow channel 1
FCGR - Fc gamma receptor
FCGR2b - Fc gamma receptor IIb
FCL - familial chilblain lupus
FIP200 - RB1-inducible coiled-coil protein 1
FLIP - FLICE-inhibitory protein
FOM - fluorinated octyl maltoside
FT - flow-through
G3BP1 - G3BP stress granule assembly factor 1
GABARAP - γ-aminobutyric acid receptor-associated protein
GMP - guanosine monophosphate
HBV - hepatitis B virus
HC - heavy chain
hcAb - heavy-chain-only antibody
HI - hinge
HIN - hematopoietic interferon-inducible nuclear
his-tag - histidine-tag
HIV-1 - human immunodeficiency virus type 1
Hrk - activator of apoptosis harakiri
HRP - horse radish peroxidase
HSV-1 - herpes simplex virus type 1
HTLV-1 - human T cell leukemia virus type 1
HTS - high-throughput screening
IAP - integrin associated protein
ICAM-1 - intercellular adhesion molecule 1
ICL1 - intracellular loop 1
IFI16 - interferon-gamma inducible factor 16
IFN - interferon
IFNAR - interferon-α/β receptor
Ig - immunoglobulin
IgV - immunoglobulin variable
IKK - IκB kinase
IL-1β - interleukin 1β
IMAC - immobilized metal affinity chromatography
IMPRS-LS - International Max Planck Research School for Molecular Life Sciences
iNOS - inducible nitric oxide synthase
INSIG1 - insulin-induced gene 1
IP-10 - interferon gamma-induced protein 10
IP3R1 - type 1 inositol 1,4,5-trisphosphate receptor
IPTG - isopropyl β-D-1-thiogalactopyranoside
IRF3 - interferon regulatory factor 3
ISG - interferon-stimulated gene
ISGF3 - interferon-stimulated gene factor 3
ISRE - interferon-stimulated response elements
ITIM - immunoreceptor tyrosine-based inhibitory motif
JAK1 - Janus kinase 1
kbp - kilobasepair(s)
kDa - kilodalton(s)

LB - Lysogeny broth
LC - light chain
LC3P - light chain 3 protein
LPG2 - laboratory of genetics and physiology 2
LRR - leucine-rich repeat
LTβR - lymphotoxin β-receptor
M.W. - molecular weight
M.W.C.O. - molecular weight cut-off
mAb - monoclonal antibody
Mab21 - male abnormal 21
MALS - multi angle light scattering
MCL1 - myeloid cell leukaemia 1
MCP-1 - monocyte chemoattractant protein 1
MDA5 - melanoma differentiation-associated protein 5
MIP-1 - macrophage inflammatory protein 1
MLKL - mixed lineage kinase domain-like protein
MOMP - mitochondrial outer membrane permeabilization
MP - membrane protein
MPIB - Max Planck Institute of Biochemistry
MSP - membrane scaffold protein
MUL1 - mitochondrial E3 ubiquitin protein ligase 1
NAV - NeutrAvidin
NB - nanobody
ND - nanodisc
NEB - New England Biolabs
NEMO - NF-κB essential modulator
NF-κB - nuclear factor kappa-light-chain-enhancer of activated B cells
Ni-NTA - nickel-nitrilotriacetic acid agarose
NIK - NF-κB-inducing kinase
NLR - nucleotide-binding domain, leucine-rich repeat-containing receptor
NLRP3 - NLR family pyrin domain containing 3
NOD - nucleotide-binding oligomerization domain
NONO - non-POU domain-containing octamer-binding protein
NTD - amino-terminal domain
NTSR1 - neurotensin receptor 1
OAS - oligoadenylate synthase
OD_{600} - optical density at 600 nm
PAMP - pathogen-associated molecular pattern
PARP1 - poly(ADP-ribose) polymerase-1
PBMC - peripheral blood mononuclear cell
PBST - phosphate buffer saline-tween 20
PCR - polymerase chain reaction
PD-L1 - programmed death-ligand 1
PDB - Protein Data Bank
PEI - polyethylenimine
PI3KC3 - class III PI3K
PI3P - phosphatidylinositol-3-phosphate
PINK1 - PTEN-induced kinase 1

PKA - protein kinase A
PLIC-1 - protein linking IAP and cytoskeleton 1
PMSF - phenylmethylsulfonyl fluoride
PPM1A - protein phosphatase, Mg^{2+}/Mn^{2+} dependent 1A
PQBP1 - polyglutamine binding protein 1
PRR - pattern recognition receptor
PTM - post-translational modification
PUMA - p53-upregulated modulator of apoptosis
PVDF - polyvinylidene fluoride
PYHIN - pyrin and hematopoietic interferon-inducible nuclear
RANK - receptor activator for nuclear factor κB
RIG-I - retinoic acid-inducible gene I
RIPA - RLR-induced IRF3-mediated pathway of apoptosis
RIPK1 - receptor-interacting serine/threonine-protein kinase 1
RLR - retinoic acid-inducible gene I-like receptor
RNA - ribonucleic acid
RNase H2 - ribonuclease H2
RNF5 - ring finger protein 5
RT - room temperature
SAMHD1 - sterile alpha motif and histidine aspartate domain-containing protein 1
SARS-COVID-2 - severe acute respiratory syndrome coronavirus 2
SAV - streptavidin
SAV beads - streptavidin-coated paramagnetic beads
SAVI - STING-associated vasculopathy with onset in infancy
SDS-PAGE - sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC - size-exclusion chromatography
SENP2 - sentrin/SUMO-specific protease 2
SHP-1 - Src homology region 2-domain-containing phosphatase 1
SIRPα - signal regulatory protein α
SIV - simian immunodeficiency virus
SLE - systemic lupus erythematosus
SMAC - second mitochondria-derived activator of caspases
SMALP - styrene maleic-acid lipid particle
SPR - surface plasmon resonance
ss - single stranded
STAT1 - signal transducer and activator of transcription 1
STIM1 - stromal interaction molecule 1
STING - stimulator of interferon genes
SYK - spleen tyrosine kinase
T4 PNK - T4 polynucleotide kinase
TAK1 - TGFβ- activated kinase 1
TB - terrific broth
TBK1 - TANK-binding kinase 1
TEV - Tobacco etch virus
TGS - Tris-Glycine-SDS
TLR - toll-like receptor
TM - transmembrane helix
TMB - 3,3’,5,5’-tetramethylbenzidine

TMD - transmembrane domain
TNF - tumor necrosis factor
TNFR - tumor necrosis factor receptor
TOLLIP - toll-interacting protein
TRAF2 - TNFR-associated factor 2
TRAIL - tumor necrosis factor-related apoptosis-inducing ligand
TRAIL-R1 - TNF-related apoptosis-inducing ligand receptor 1
TREX1 - three-prime repair exonuclease 1
TRIM56 - tripartite motif containing 56
TRIS-HCl - Tris-(hydroxymethyl)-aminomethane hydrochloride
TRPV1 - transient receptor potential cation channel subfamily V member 1
TSP-1 - thrombospondin 1
TTLL4 - tubulin tyrosine ligase-like 4
TYK2 - tyrosine kinase 2
ULK1 - Unc-51-like kinase 1
UMP - uridine monophosphate
VEGFR2 - vascular endothelial growth factor receptor 2
VH - variable heavy
VHH - heavy-chain variable domain
VL - variable light
VSP34 - vacuolar protein sorting 34
WB - western blotting
WIPI2 - WD repeat domain phosphoinositide-interacting protein 2
XIAP - X-linked inhibitor of apoptosis protein
ZCCHC3 - zinc finger CCHC-type containing 3

13. Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor Karl-Peter Hopfner for giving me the opportunity to gain invaluable experience and pursue my doctorate in Gene Center Munich. I thank for his kindness and guidance that made me able to successfully accomplish this work and pursue my interests.

My closest lab colleagues, Mariia and Aysenur, for always supporting and motivating me. Liudmila and Sebi, the colleagues who introduced me to the lab and gave me plenty of valuable advice. Carina, for our numerous discussions and her countless brilliant ideas.

All current and previous members of Hopfner's lab for their kindness, lots of scientific discussions, and their help with various things. I would like to thank Gregor Witte for guidance and help with X-ray crystallography and SPR. Joe Bartho and Katja Lammens for guidance and help with electron microscopy. Stephi Wolf for assistance with various administrative issues. And also the members of the Antibody Subgroup - my sister group, for their help and support.

All members of Schulman's lab for always welcoming me warmly. Especially, I would like to thank Bastian for our many scientific discussions and his suggestions on how to improve my research set up.

Many members of Gene Center community who gave me good advice and helped me a lot in various aspects of my work. Especially, the members of Beckmann's lab for letting me use their equipment.

The members of my Thesis Advisory Committee: Brenda Schulman and Veit Hornung for their advice, kindness and helping me to achieve my goals. Brenda, for believing in me, since our very first meeting, and for the words "I prefer direct communication" and being totally clear in whatever she says. I needed to go for my recent adventure in the Philippines, the land of direct people, to really learn to appreciate these words in our indirect world full of unnecessary confusion.

The members of my Examination Committee: Lucas Jae, Roland Beckmann, Julius Stingele and Johannes Stigler for their time and kindness.

The members of the International Max Planck Research School for Molecular Life Sciences (IMPRS-LS) Coordination Office for organizing this PhD program, and their kindness and help with various issues. And especially, for organizing the unique interview for this program that unlike other interviews that I attended, was a really pleasant experience. Initially, I was surprised with how interview can last five days. And now I know, this is all about finding a "chemistry" between people, and it takes time, and it really matters. Thanks for this valuable concept.

The members of MPIB Crystallization Facility for setting up HTS crystallization trials, essential for my research.

My collaborators from the University of Bonn - Florian Schmidt and Jennifer Deborah Wuerth for letting me participate in their exciting project.

All my Friends and Family, especially Gosia, Jakub and Asia for being always close to me, motivating and supporting me in all my endeavors. Additionally, Asia for being my amazing hiking buddy in the “land of mountains”. Magdalena, Piotr, Dominika and Ania for their support and filling me up with positivity.

And last but not least, I would like to thank my Best Friend Dominika for bringing light to my life, even in the moments of the thickest darkness.