

All-Atom Simulations Uncover Structural and Dynamical Properties of STING Proteins in the Membrane System

Rachel T. Payne¹, Silvia Crivelli², and Masakatsu Watanabe^{3*}

¹Boston University, Boston, MA 02215, United States

²Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States

³Fort Hays State University, Hays, KS 67601, United States

*To whom correspondence may be addressed. Email: m_watanabe@fhsu.edu

Abstract

Recent studies have shown that the stimulator of interferon gene (STING) protein plays a central role in the immune system by facilitating the production of Type I interferons in cells. The STING signaling pathway is also a prominent activator of cancer-killing T cells that initiates a powerful adaptive immune response. Since biomolecular signaling pathways are complicated and not easily identified through traditional experiments, molecular dynamics (MD) has often been used to study these biological pathways' structural and dynamical responses. Here, we carried out MD simulations for full-length chicken and human STING (chSTING and hSTING) proteins. Specifically, we investigated ligand-bound closed and ligand-unbound open forms of each STING in the membrane system by comparing conformational and dynamical differences among them. Our research provides clues for understanding the mechanism of the STING signaling pathway by uncovering some detailed insights for the examined systems: the residues from each chain in the binding pocket are strongly correlated to one another in the open STING structure compared with those in the closed STING structure. Ligand-bound closed STING displays approximately 170° rotation of the ligand-binding domain (LBD) relative to the open-STING structure. The detailed

dynamical analysis of residue Cys148 in the linker region of hSTING does not support the earlier hypothesis that Cys148 can form disulfide bonds between adjacent STING dimers. We also reveal that using the full-length proteins is critical as the MD simulations of the LBD portion only cannot properly describe the global conformational properties of multiple domain proteins, such as STING.

Introduction

The stimulator of interferon gene (STING) protein has become a focal point in immunological research and future drug discovery for its potential to enhance the ability to fight infection and kill cancerous cells. Recognition of aberrant DNA or cyclic dinucleotides (CDNs) in the cytosol activates cyclic GMP-AMP synthase (cGAS) to produce a second messenger ligand, cGAMP, which binds and activates STING¹. The cGAS-STING pathway triggers multiple signaling cascades leading to the production of Type I interferons (IFNs)².

The cGAS-STING signaling pathway can provide protection or resistance against infections; however, improper activation or overactivation may lead to autoinflammation and autoimmune diseases³. For instance, STING-associated vasculopathy with onset in infancy (SAVI) leads to perpetual STING activation and inflammation. Aicardi-Goutières syndrome causes mutations in DNA-degrading proteins, leading to the buildup of DNA inside cells, which can inadvertently trigger the activation of STING. These examples suggest that the cGAS-STING pathway plays an integral role in inflammation and autoimmune disease³. However, understanding the detailed mechanism of this signaling pathway remains a challenge.

Recently crystal structures of full-length STING proteins, i.e. ligand-unbound open forms (apo structure) for chicken and human as well as cGAMP ligand-bound closed form (holo structure) for chicken, were identified using cryogenic electron microscopy (cryo-EM)⁴. The structures for the apo and holo states are shown in Figure 1(a). STING exists as a homodimer transmembrane protein in nature, consisting of an N-terminal transmembrane (TM) region, linkers, a C-terminal ligand-binding domain (LBD), and a C-terminal tail (CTT). Additionally, the structure of the ligand-binding site (Figure 1 (b)) displays a V-shaped dimer. In the apo configuration, the CTT is tightly sequestered against the main body protein to auto inhibit

STING signaling⁵. The binding of cGAMP induces an inward rotation of both monomers toward the ligand-binding site with an accompanying four-stranded β -sheet ‘lid’ that encloses the pocket, shown in Figure 1(c)⁴⁻⁸.

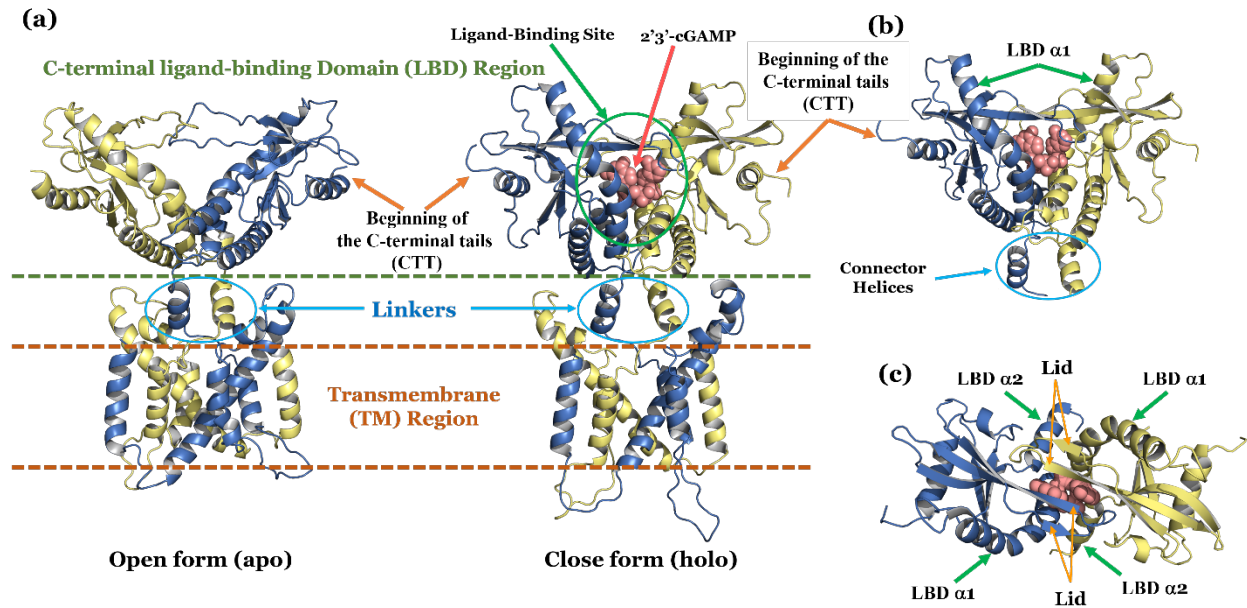


Figure 1 Structures of full-length hSTING. (a) apo (open) and holo (closed). Apo linker and LBD- $\alpha 1$ portions of the two subunits in the dimer create a right-handed crossover formation. In the holo state, this crossover is no longer formed. (b) The two LBD- $\alpha 1$ helices create the V-shape binding pocket. (c) Top view of STING LBD – β -sheets from chains A and B act as a lid for the binding site.

The recent cryo-EM experimental result⁴ also shows that LBD rotates clockwise relative to the TM domain upon ligand binding (Figure 1(a)). This rotation unwinds the right-handed crossover in the connector helix (as observed in the apo structures) and results in the parallel alignment of the connector helix in Figure 1(b). The study by Ergun et al.⁶ showed that the LBD rotational conformation change causes the CTT's to be released, which are otherwise sequestered in the apo structure. This conformational change of CTT's mediates the recruitment of Tank-binding kinase 1 (TBK1)^{2,9-12}.

Additionally, Shang et al.⁴ previously reported that the rotation of the LBD in full-length STING, upon binding cGAMP, is essential. This rotation facilitates the formation of STING

tetramers and higher-order oligomers through side-by-side packing. STING oligomerization allows for TBK1 clustering and trans-autophosphorylation, a process that is not possible without this configuration. Additionally, STING must be phosphorylated to continue the immune signaling response. Ser366, the major phosphorylation site in both human and chicken STING, is located too far away from the active site of TBK1 for STING to be phosphorylated by the same TBK1 to which it is bound². However, in large oligomers, the TBK1 that is tethered to the two CTT's of one STING dimer can phosphorylate the serine residues of CTT's in neighboring STING proteins, which are not bound to this TBK1^{2,4}.

The phosphorylated residues in the CTT provide a binding site for interferon regulatory factor 3 (IRF3), thereby recruiting IRF3 for phosphorylation via nearby TBK1. Phosphorylated IRF3 forms a dimer that translocate to the nucleus and induces the production of Type I IFNs and other cytokines². Specification of TBK1-mediated IRF3 activation is essential for the tight regulation of IFN production, which would otherwise lead to autoimmune diseases⁹.

Even though structural and dynamic effects associated with the biomolecular signaling pathways are complex and not easily recognized through traditional experiments, increasing structural biology and molecular dynamics (MD) capabilities provide powerful new scientific tools for decoding the signaling pathways. MD simulations facilitate and complement experimental studies by providing detail at the atomic level for molecular interactions in biological systems.

Recently MD simulations were applied to investigate structural and dynamic differences before and after the ligand binding of STING¹³⁻¹⁵. However, all previous studies used only the LBD portion of the protein. Here, we carried out MD simulations for full-length chicken and human STING (chSTING and hSTING) in the membrane environment to explore structural and

dynamical differences between them. We show that using the full-length protein is critical as the MD simulations of the LBD portion alone cannot properly describe the global conformational properties of multiple domain proteins, such as STING. The global conformational structures are necessary to understand the biological mechanism, including the signaling pathway.

chSTING was chosen for this research because the crystal structures for both full-length apo and holo states have been solved and published in the Protein Data Bank (PDB). According to the sequence alignment (supplemental information Figure 1S), chSTING shares 44.6% amino acid identity with hSTING. Additionally, the previous experimental study conducted by Shang et al.⁴ also confirmed that chSTING and hSTING have very similar apo structures. Thus, we homologically modeled a ligand-bound closed-form of hSTING for this study.

Using MD simulations, we answered how STING protein structures in the membrane environment at normal temperature might deviate from the cryo-EM and crystal structures published in the PDB. We looked at ligand-bound closed and ligand-unbound open forms to investigate the conformational and dynamical effects produced via ligand-binding in chSTING and hSTING. Additionally, we explored the dynamical and structural differences between chSTING and hSTING ligand binding sites.

Furthermore, Ergun and Li's previous study¹⁶ suggested that the uncrossing of the linker region in hSTING holo-state increases the availability of cysteine residues to form disulfide bonds between neighboring hSTING dimers for polymer stabilization. We investigated the accessibility of cysteine residue 148 (Cys148) located in the linker region of hSTING to explore its possibility of participation in polymer crosslinking.

Methods & Materials

Since STING is a homodimer TM protein (Figure 1(a)), we performed a series of MD simulations of full-length chSTING and hSTING proteins within a membrane environment. Here, we briefly describe our simulation protocols.

First, the complete structures of both full-length chSTING and hSTING were built. Since the published cryo-EM structures have missing residues, we used MODELLER¹⁷ to construct complete STING models. For building both apo-chSTING and apo-hSTING protein models, the cryo-EM structures from the PDB: 6NT6 and 6NT5, respectively, were used as initial configurations. Then, since the crystal structures of chSTING and hSTING are very similar to each other⁴, the crystal structure of the LBD portion of ligand-free hSTING [PDB ID: 4F5W] was used as a secondary template to construct the LBD portion more accurately. The PDB: 6NT7 was used as the initial structure for constructing both holo chSTING and hSTING. The crystal structure of the ligand-bound hSTING LBD [PDB ID: 4F5D] was used as a secondary template.

Once all STING proteins were completely constructed using MODELLER, the CHARMM-GUI web server¹⁸ was applied to build the initial STING membrane systems. The lipid bilayer membrane systems had dimensions of $125 \times 125 \times 25$ Å and consisted of 1,2-didecanoyl-sn-glycero-3-phosphocholine (DCPC) lipids. The topology and force-field parameters for the 2'3'-cGAMP molecule were built using CGenFF¹⁹.

The MD simulations were performed with the GROMACS 2020 software²⁰ and the CHARMM36 force field²¹. Each protein/membrane complex was initially placed in the center of a $125 \times 125 \times 155$ Å simulation box solvated by TIP3P water molecules²². Then, if necessary, potassium ions were added to neutralize the total system charge.

After assembling the protein/membrane complex, an equilibration was performed to relax the initial system from unrealistic high energy atom arrangements before MD production

simulations. First, the steepest descent algorithm was used to minimize each system for 5,000 steps. Then, six consecutive equilibrations were performed. Here, we had gradual equilibrations of the initially assembled system; various restraints were applied to the protein, ligand, water, ions, and lipid molecules during these equilibrations. The equilibration processes were similar to those used by Jo et al.²³

The last coordinates from the last equilibration were processed through a two-step final equilibration. The first step of this final equilibration was to apply a constant number, volume, and temperature (NVT) simulation using the V-rescale algorithm²⁴ without any position restraints for 200-ps. Another equilibration was conducted at a constant number, pressure, and temperature (NPT) for a 500-ps simulation via the Parrinello-Rahman algorithm²⁵ at 1 bar pressure and the V-rescale algorithm²⁴ at 30°C. After these additional equilibrations, the production MD simulations were run for 250-ns in the NPT condition (Table 1S).

In all cases, an integration time step of 2.0-fs was used. Periodic boundary conditions were applied in all directions throughout all simulations. The LINCS algorithm²⁶ was used only to constrain the bonds involving hydrogen atoms. The electrostatic interactions were calculated using a Particle mesh Ewald algorithm²⁷ with a 12Å cut-off. The trajectories were saved every 50-ps during each simulation for the analyses. A model for the holo hSTING MD simulation is presented in Figure 2S.

Results & Discussion

We inspected the condition and characteristics of the trajectories using the VMD software²⁸ and then proceeded to analyze the data further. We used our in-house developed Python and R code, GROMACS analysis tools²⁰, Bio3D²⁹, and PyMOL³⁰ to perform the analyses. To remove the bias of the initial states, we discarded the first 50-ns MD trajectories.

Next, we analyzed the C α root mean square deviation (RMSD) and the C α root mean square fluctuation (RMSF). Each RMSD was calculated with reference to the initial cryo-EM structures of open form STING (PDB: chSTING: 6NT6, hSTING: 6NT5) and closed form STING (PDB: chSTING: 6NT7, hSTING: MODELLER generated initial structure). Those results are presented in Figures 3S-6S. Additional analyses of simulation trajectories are presented next.

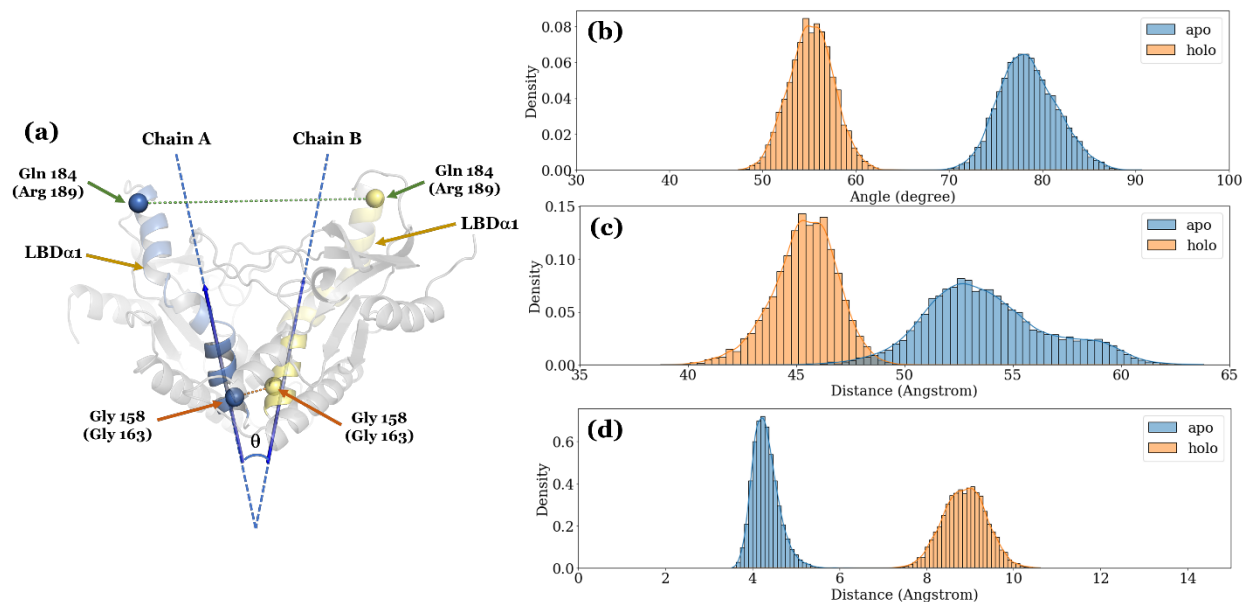


Figure 2 Geometry of hSTING LBD- α 1 helices. (a) The schematic LBD figure shows the definitions of the angle between the two α 1-helices, the “tip distance,” and the distance between closed-end α 1-helices. (b) Distributions of the angles in apo and holo states. (c) and (d) Distributions of distances between Gln184 residues and between Gly158 residues, respectively.

Analysis for Ligand-binding domain (LBD) structures

Previous crystal structural studies show that both chSTING and hSTING display open LBD conformations in the apo state and that cGAMP binding prompts the LBD to adopt a closed conformation^{4,6,8}. To confirm these observations quantitatively, we analyzed the angle between the two ligand-binding domain α -helices (LBD- α 1) from the two monomer chains. These α -helices consist of residues 160-190 for chSTING and 155-185 for hSTING (Figure 1S). However, it was not helpful to use all residues of this helix to define the axis in the angle

calculation because during the simulation, the long LBD- α 1 helices broke into two separate α -helices. Since the breaking point was around residue 165 for hSTING and residue 170 for chSTING, we used residues 155-165 for hSTING and residues 160-170 for chSTING to define the helix axis for each LBD- α 1. Then the angle between these two helix axes were calculated (Figure 2 (a), Table 2S).

The distributions of the angles in apo and holo-structures of chSTING and hSTING simulations are presented in Figure 2 (b) and Figure 7S (a). There are clear differences in the distributions between the apo and holo-structures: The average angles between LBD- α 1 helices in the apo states for hSTING and chSTING are $78.5^\circ (\pm 3.1^\circ)$ and $69.8^\circ (\pm 4.1^\circ)$, respectively. The average angles in the holo states for hSTING and chSTING are $55.2^\circ (\pm 2.4^\circ)$ and $53.4^\circ (\pm 2.6^\circ)$, respectively. This difference in consistency between residue pairs reinforces the idea that strong intermolecular bonds are formed between cGAMP and LBD, enforcing a more rigid composition, ultimately resulting in greater structural stability. These results can be used to classify the basic conformations of the LBD region.

To get more insight into the STING proteins' open vs. closed conformational change upon ligand binding, we monitored the “*tip distances*” between Gln184 in hSTING and between Arg189 in chSTING at the wide-end of the two LBD- α 1 helices. According to Figure 2 (b), the “*tip distance*” distributions for apo hSTING and chSTING have peaks at $53.9 \text{ \AA} (\pm 2.9 \text{ \AA})$ and $50.6 \text{ \AA} (\pm 4.5 \text{ \AA})$, respectively. Apo state chSTING (Figure 7S (b)), shows rather large deviations in the “*tip distance*” distribution compared to hSTING. The average value of apo hSTING agrees with 55.5 \AA of the “*tip distance*” measured from the cryo-EM hSTING structure (PDB: 6NT5)⁴.

After binding cGAMP, this region closes to produce a tight V-shape conformation with resulting tip distance peaks of 45.4 Å (± 1.5 Å) for hSTING and 44.8 Å (± 1.7 Å) for chSTING (Figure 2 (c)). The value observed in hSTING is slightly different from that observed in the previous experiments⁸. However, our observations for the apo and holo states support the general understanding of the V-shape conformational changes.

In the apo structures of both hSTING and chSTING, the connector helix and LBD- α 1 in the dimer form a right-handed crossover formation that packs closely at one end of the LBD- α 1. We measured the distances between conserved glycine residues (hSTING: Gly158; chSTING: Gly163) in the LBD- α 1. The average distance between Gly158 in both chains for apo hSTING is 4.3 Å (± 0.3 Å). The average distance between Gly163 residues in apo chSTING is 4.8 Å (± 1.0 Å). According to Figure 2(d), the distribution of the distance between Gly158 in apo hSTING shows relatively small fluctuations throughout the simulations and is similar to the one observed from PDB: 6NT5⁴. However, the chSTING value is slightly larger than the one observed from PDB: 6NT6⁴. Additionally, similar to the “*tip distance*,” apo chSTING also displays larger deviations in the Gly163 distance distribution (Figure 7S (c)) than those of hSTING.

In the holo structure, the direct binding of cGAMP to the middle section of LBD- α 1 pushes the two LBD- α 1 helices in the dimer away from each other, increasing the average distance between the two glycine residues to 8.9 Å (± 0.5 Å) in human (Figure 2 (d)) and 8.1 Å (± 0.5 Å) in chicken STING (Figure 7S (c)). The differences in holo hSTING and chSTING are not significant; a small deviation in distance is expected due to the natural dynamics of biological systems. The value for chSTING is consistent with the previous experimental result presented by Shang et al.⁴

The increase in distance between glycine residues from the apo to holo states is caused by the conformational changes induced by cGAMP driving the LBD rotation. If the crossover conformation of the connector and LBD- α 1 helices of the two chains, seen in the apo state, was maintained, the outward movement of LBD- α 1 would lead to clashes in the connector loop. Thus, the observations of increased distance in this region are evidence of conformational changes between apo and holo states.

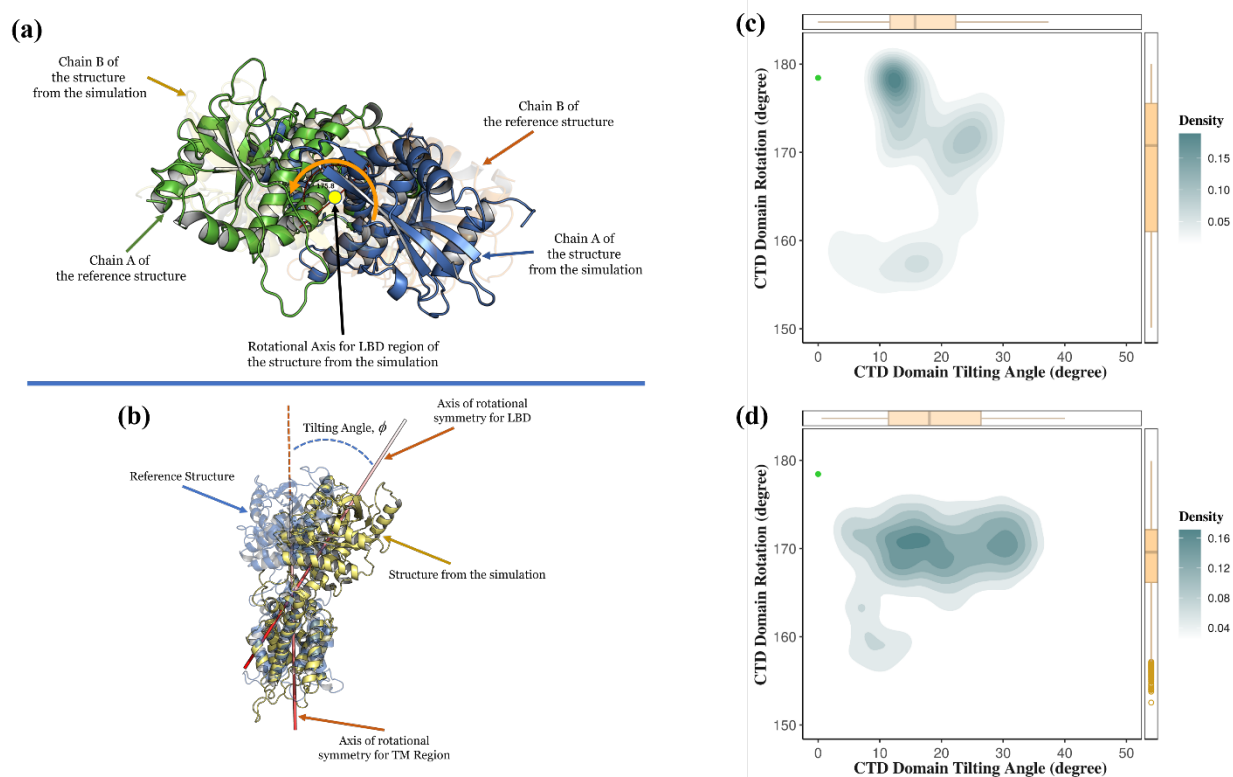


Figure 3 Rotational and Tilting Dynamics of STING from last 150-ns simulation trajectories. (a) Rotational angle of the LBD with respect to the initial cryo-EM apo structure. (b) Tilting angle is defined as the angle between two axes of rotational symmetry – one for the LBD and another for the TM region. (c) and (d) 2-D histogram of chSTING and hSTING LBD tilting and rotational angle. In (c), the green dot indicates angle values from apo and holo chSTING cryo-EM structures (PDB: 6NT6 and 6NT7). In (d), the green dot indicates angle values from apo (PDB 6NT5) and homologically modeled initial holo hSTING structure. Note: the box-and-whisker plots on the sides of the 2-D histogram describe the distributions of tilting and LBD rotational angles.

Analysis for C-terminus domain rotation and tilting movements

One advantage of carrying out full-length STING simulations is to monitor relative motions between the LBD and TM regions exhibited by holo and apo STING proteins. Additionally, since the segmental flexibility is likely a dominant feature of numerous proteins composed from multiple connected domains, we were interested in understanding the flexibility between the LBD and TM regions present in the membrane environment. Therefore, we examined the rotation and tilting movements of STING with respect to the reference structures. The cryo-EM apo state model (chSTING:6NT6, hSTING: 6NT5) was set as the reference structure. First, the TM regions of both reference and simulated structures were aligned by performing a sequence alignment followed by a structural superposition. Then, because STING is a homodimer, the axes of rotational symmetry for the TM and LBD regions were separately evaluated for both the reference and simulated structures.

The rotational angle of the LBD (Figure 3 (a), Table 3S), was extracted from the transformation matrix between the LBD of the reference and simulated structures. The tilting angle of the LBD region with respect to the TM region was defined as the angle between two axes of rotational symmetry – one for the LBD and another for the TM regions (Figure 3(b)). Results for holo chSTING and hSTING are presented in Figure 3(c) and (d). (The apo results for both STING proteins are shown in Figure 8S.)

Our results showed that holo chSTING had an average LBD rotation relative to the apo structures of $168.8^{\circ} (\pm 7.9^{\circ})$. The equivalent rotation for hSTING was $168.7^{\circ} (\pm 5.0^{\circ})$. These results differ from the 180° rotation observed in the cryo-EM experiment⁴. According to Figure 3(c), in our simulations, most chSTING conformational samples were located around 178° . Thus, the holo chSTING had a skewed rotational angle distribution, with a median value of 170.8° , which was slightly higher than the 168.8° average angle. The rotational angle distribution in the

hSTING model was also slightly skewed with a median angle of 169.5°. The difference between our observations and reported findings⁴ could be a result of cryogenic vs. membrane environment research systems.

In addition, according to Figure 8S (a) and (b), the average LBD rotations from the initial cryo-EM structures are 32.1° ($\pm 5.9^\circ$) and 17.5° ($\pm 5.9^\circ$) for apo chSTING and hSTING, respectively. These observations suggest that there are large deviations from the cryo-EM structures and that STING proteins are flexible in the membrane environment.

The tilting dynamics for STING are unique to our work. Previous simulation and many experimental research studies used only the LBD structures and did not include the membrane system. Since the LBD region is located in the cytosol of the cell, the movement of the LBD with respect to the membrane surface is important to understand the flexibility of STING in the cytosol. According to our results, the average tilting angles are 35.7° ($\pm 9.7^\circ$) and 13.6° ($\pm 6.3^\circ$) for the apo states of chSTING and hSTING, respectively. Apo chSTING has a larger average tilting angle, implying that its LBD tends to lean to one side. However, we do not believe the tilting motion observed has a preferred side.

For the holo states of chSTING and hSTING, we observed 16.7° ($\pm 7.0^\circ$) and 18.7° ($\pm 9.0^\circ$) tilts. Figures 3(c) and (d) indicate large fluctuations present in holo chSTING and hSTING. These large fluctuations demonstrate the ‘*swaying*’ nature of the LBD, resulting from the flexibility of the LBD in the membrane environment observed in our simulations.

Analysis of STING chains by shape factor

To understand how the ligand-binding influences the general shape of each chain in STING dimers, we evaluated a shape factor, S , which describes the relative anisotropic shape of

the molecule³¹. The shape factor is defined by three gyration tensors of all C α coordinates in each chain. The factor ranges from 0 to 1, where $S = 0$ occurs if all C α atoms in the chain are spherically symmetric and $S = 1$ occurs if all C α atoms lie on a line. Furthermore, the shape factor illustrates how the structures of chains in the STING dimer differ from each other.

According to Figure 9S and Table 4S, both chSTING and hSTING exhibit asymmetry in the monomer chains. Although the STING protein is a homodimer, the results from the shape factor reveal that the conformational dynamics in chain A differ from those in chain B. Additionally, chSTING shows larger fluctuations in shape factors compared to hSTING. This result indicates that chSTING is more flexible than hSTING in the membrane environment.

Furthermore, shape factors of the apo states of chSTING and hSTING display no substantial differences between chains during our simulations. If there were any large conformational changes, such as separation of chains, during the simulation, it would reflect significant changes in the shape factors. We did not observe any large or sudden deviations in shape factors, indicating that our simulation models did not display any separation of monomer chains during the entire 250-ns production. This observation is further supported by Figure 4S, which shows there were no significant increases in LBD's RMSD values during the simulations. However, this result conflicts with the earlier simulation study by Shih et al.¹³, which reported the separation of monomer chains in apo hSTING. The origin of this conflict should be explored further in future research.

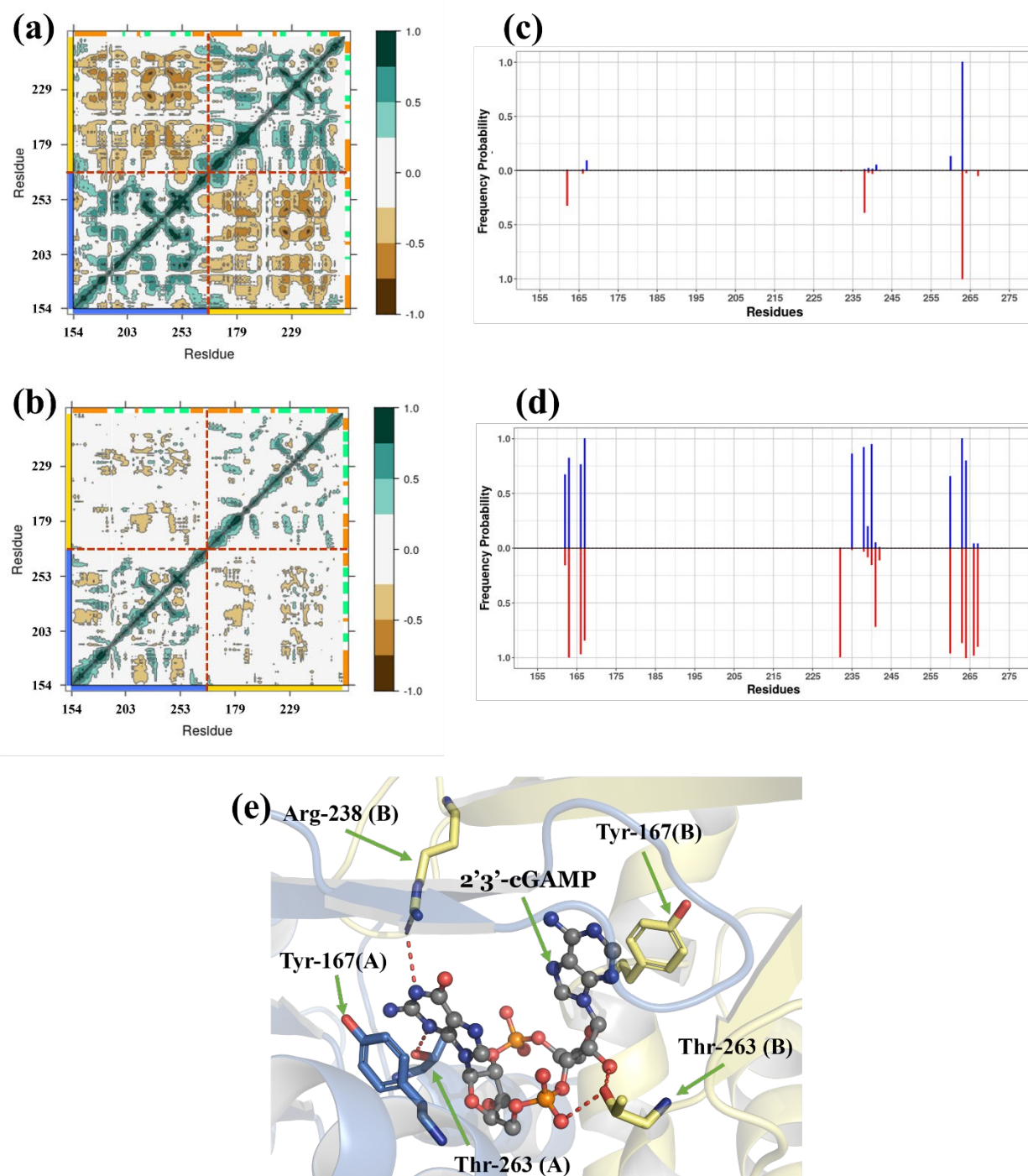


Figure 4 Dynamical cross-correlation maps between residues 154 and 277 of chains A and B. (a) apo hSTING (b) holo hSTING. The blue and yellow axes indicate residues in chains A and B, respectively. Secondary structures are presented by orange blocks (α -helices) and green blocks (β -sheets). All correlations in the holo state are significantly weaker than those of the apo state. (c) Histogram for hydrogen bonding between ligand and residues in hSTING – blue and red indicate residues in chains A and B, respectively. (d) Close contact histogram: the probability of residues within 4.0 Å of ligand during

the simulation – blue and red indicate residues in chains A and B, respectively. (e) Snapshot (at 200-ns) of the ligand binding to STING. The red dotted lines indicate hydrogen bonds.

Analysis for Ligand-Binding Sites

Ligands can affect the protein's stability and conformation. To further describe the formational characteristics of STING in the presence of a ligand, we monitored pairwise cross-correlation coefficients, which represent how the atomic fluctuations/displacements of a system are correlated with one another. The dynamical cross-correlation (DCC) maps of the ligand-binding sites in apo and holo hSTING are shown in Figure 4. Additional DCC maps of chSTING and hSTING are shown in Figures 10S and 11S. The apo state shows significant correlations between chains A and B (Figure 4 (a)); however, the holo state (Figure 4 (b)) displays notably fewer correlations between chains. These observations suggest that the ligand heavily regulates the dynamics of the ligand binding site in holo STING structures. Our observations also indicate that ligand binding enforces a more rigid composition in the LBD. The DCC results also support our previous shape factor analysis results of differences in self-correlation between chains A and B despite their homodimer nature.

To get better insight into how the ligand binds to STING and influences its dynamics, we identified the residues in contact with the ligand as those within a 4.0 Å distance of the cGAMP atoms. A hydrogen bond (h-bond) was considered to be formed when the donor (D) – acceptor (A) distance was within 3.5 Å, and the D-H-A angle was less than 45° (Figure 12S and 13S).

Figure 4 (c) characterizes the frequencies of h-bonds between residues in cGAMP-bound hSTING during the MD runs. There are consistent and stable h-bonds between Thr263 in hSTING (chSTING: Ser268) and cGAMP. Additionally, Arg238, located in the β -sheet lid and occasionally forms h-bonds with cGAMP. However, based on this analysis, h-bonds from the β -

sheet “lids” are not consistent interactions. Our other study³² showed that the binding free energy of cGAMP in hSTING was -34.8 ± 2.0 kcal/mol, which indicates stable and strong binding. The number of stable h-bonds found between the ligand and protein are shown in Figure 4 (c). However, these hydrogen bonds do not provide enough binding energy to explain the stability found in the ligand-binding site. Other intermolecular interactions are necessary to stabilize the ligand in the binding site.

According to Figure 4(d), hSTING has a fair number of residues in contact with cGAMP. This demonstrates that conserved residues between chSTING and hSTING, such as Ser162, Tyr163, Gly166, Tyr167, Arg232, Arg238, Tyr230, and Pro264 (in hSTING, sequence shown in Figure 1S) play significant roles in protein-ligand contact; forming van der Waals and Coulomb interactions with cGAMP. Additionally, in hSTING a few non-conserved residues, such as Ser241 and Thr263 are also involved in cGAMP interactions.

Figure 4(e) displays important intermolecular interactions between hSTING and cGAMP, specifically, the close contacts between the ligand and Tyr167 residues in chains A and B. The close contact of Tyr167 elucidates the π - π stacking patterns of the benzene-like ring of the ligand and the aromatic residue of hSTING. In other words, cGAMP is held between these two tyrosine residues by the π - π stacking interaction. These interactions combined with h-bonds, van der Waals, and electrostatic interactions are enough to keep cGAMP in the binding site. Very similar π - π stacking interactions by Tyr172 in chSTING were also observed.

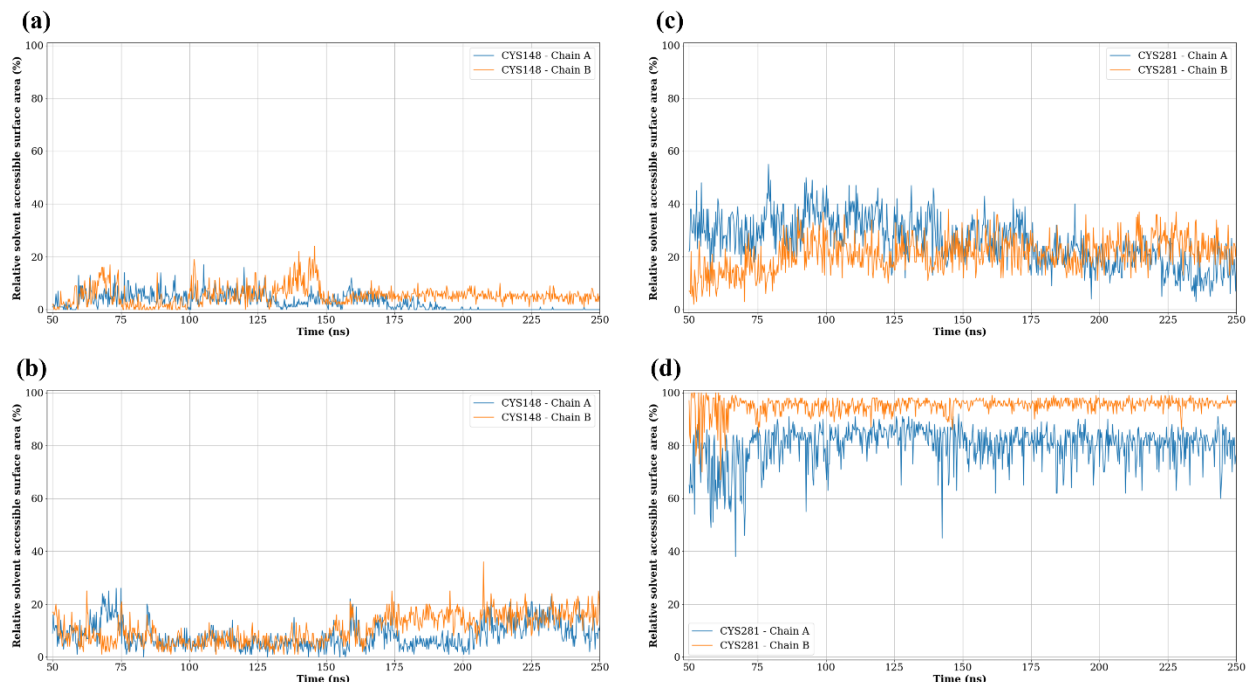


Figure 5 Relative Solvent Accessible Surface Area (SASA). Calculated relative per-residue SASA values for selected cysteine residues to determine the potential capability to form disulfide bonds between neighboring tetramers. The value is relative to full exposure of the residue, calculated by removing all other residues except its two next neighbors if present. Values vary between 0.0% (fully buried) and 100.0% (fully exposed). Cys148 in human STING (a) apo state and (b) holo state. Relative SASA of Cys281, located on the ligand-binding domain loop in chSTING, (c) apo, and (d) holo state.

Analysis of Cysteine Residues

A previous study by Ergun et al.⁶ suggested that Cys148 in hSTING plays an important role in forming inter-dimer disulfide bonds and stabilizing STING tetramers and higher-order oligomers. This previous study only used the LBD of STING. Using full-length hSTING in the membrane environment (as it occurs naturally in the body), we measured the relative per-residue solvent accessible surface area (SASA) of Cys148 to determine if it would be accessible to form disulfide bonds between neighboring tetramers. According to Figure 5 (a) and (b), our results show that residue Cys148 has less than 25% SASA in both apo and holo states. Our simulations also showed that Cys148, located in the linker region, is buried deep within the protein's main body, rendering Cys148 unable to form disulfide bonds between STING tetramers at this

location in a membrane environment. We cannot support the hypothesis that STING forms disulfide bonds via Cys148, based on the fact that STING occurs naturally as a transmembrane protein and is unable to form such bonds under these conditions unless the TM region has undergone conformational changes, such as the reorganization of the TM helices.

Moreover, cysteine residues are absent from the linker region in chSTING, yet it successfully forms STING tetramers in the holo state⁴. Assuming the cysteine residues in the linker region were required for disulfide bonding between adjacent STING dimers to be activated, then chSTING would not be able to form tetramers or be activated. Therefore, we conclude that Cys148 in hSTING does not contribute to stabilizing the STING tetramer. However, we do believe that Cys148 plays an important role in activating STING.

On the other hand, Cys281 in chSTING (Figure 14S) is located in the LBD loop and is situated in a way that is likely to form disulfide bonds between neighboring STING dimers and could contribute to the formation of STING tetramers. In support of this, the simulation results (Figures 5(c) and (d)), display significant changes in relative SASA between the apo and holo states. Due to the optimal location, large SASA, and lack of other cysteine residues present in the LBD loop that could act similarly, we conclude that Cys281 in chSTING is an ideal candidate for disulfide bonds between adjacent chSTING dimers.

Since there are no cysteine residues in the vicinity of the LBD loop in hSTING, other cysteine residues besides Cys148 should be further examined to determine if disulfide bonds play a crucial role in the stabilization of STING. The recent study conducted by Ergun et al.⁶ suggested that Asp301 of one STING dimer is positioned between Arg281 or Arg284 from a neighboring dimer, allowing for salt bridge formation. The structures obtained from our MD simulations support the idea that the salt bridge could be an essential factor for stabilizing

STING tetramers and higher-order oligomers. This hypothesis should be examined in future research.

Conclusion

In summary, using MD simulations, we explored the conformational and dynamical characteristics of full-length apo and holo STING as well as the interactions between STING and cGAMP within a membrane environment. Our research provides mechanistic insights into the conformation of STING modulated by the cGAMP ligand in a membrane environment. Based on our MD simulation studies, we conclude the following:

First, cGAMP-bound holo STING displays a median value of 170.8° and 169.5° rotation of the LBD relative to the TM region in chSTING and hSTING, respectively. These values are notably different from the previously proposed value of 180°^{4,16}. The MD simulations showed the LBD swaying back and forth within the cytosol in the cell, suggesting that the LBD regions of STING are relatively flexible in a membrane environment.

Second, the residues from each chain in the binding pocket are strongly correlated to one another in apo STING compared with those in holo STING. Ligand-binding alters the chain interactions within the LBD, forming a less correlated but more rigid structure.

Third, although both chSTING and hSTING display small numbers of steady hydrogen bonds between cGAMP and STING, the binding energy produced is not enough to stabilize the ligand in the binding site. Other intermolecular bonds, such as π - π stacking interactions, are required for stabilization.

Fourth, our analysis of the relative SASA of Cys148 in hSTING reveals that the location of the residues in the linker region would render it impossible to form a disulfide bond with a

neighboring STING dimer under the current simulated conditions. Thus, our results do not support the previous hypothesis⁶ that Cys148 residues need to form disulfide bonds between adjacent STING dimers to stabilize the STING polymer. Therefore, we have proposed that Cys148 in hSTING does not play a major role in stabilizing STING tetramers but does play an important role in the activation of STING.

Some of our results obtained from our MD simulations and subsequent analyses differ from the findings reported in the previous simulation studies^{6,13–16}. We attribute these differences to the research model used in those studies, which was composed of only the STING LBD. Since STING is naturally a transmembrane protein, it is imperative that the full-length protein in a membrane system be used to ensure realistic and accurate results.

The information gathered from this research provides an initial look at the structural and dynamical effects induced via cGAMP binding in full-length human and chicken STING in a membrane environment using MD simulations. It also provides clarification regarding previous studies and proposed hypotheses, increasing our overall understanding of the STING signaling pathway. Our future research will include using computational methods to understand the STING dimer diffusion process in a membrane and how STING forms tetramers.

Acknowledgment

This work was funded in part by the Visiting Faculty Program (VFP) and Science Undergraduate Laboratory Internships (SULI) of the U.S. Department of Energy's (DOE) Office of Science and Sustainable Research Pathways (SRP) program from Sustainable Horizons Institute to R.P. and M.W. This work was also supported in part by National Institutes of Health (NIH) grants P20 GM103418 and P20 RR016475 from the Kansas Idea Network of Biomedical Research

Excellence (K-INBRE) program of the National Institute of General Medical Sciences to M. W. The authors also thank National Energy Research Scientific Computing (NERSC) for allocation of computing hours on the Cori Supercomputer and in particular Dr. Ramana Madupu of DOE for the additional allocation of computing hours.

References

- (1) de Oliveira Mann, C. C.; Orzalli, M. H.; King, D. S.; Kagan, J. C.; Lee, A. S. Y.; Kranzusch, P. J. Modular Architecture of the STING C-Terminal Tail Allows Interferon and NF- κ B Signaling Adaptation. *Cell Reports* **2019**, 27 (4), 1165–1175.e5. <https://doi.org/10.1016/j.celrep.2019.03.098>.
- (2) Zhang, C.; Shang, G.; Gui, X.; Zhang, X.; Bai, X.; Chen, Z. J. Structural Basis of STING Binding with and Phosphorylation by TBK1. *Nature* **2019**, 567 (7748), 394–398. <https://doi.org/10.1038/s41586-019-1000-2>.
- (3) Zhou, R.; Xie, X.; Li, X.; Qin, Z.; Wei, C.; Liu, J.; Luo, Y. The Triggers of the CGAS-STING Pathway and the Connection with Inflammatory and Autoimmune Diseases. *Infection, Genetics and Evolution* **2020**, 77, 104094. <https://doi.org/10.1016/j.meegid.2019.104094>.
- (4) Shang, G.; Zhang, C.; Chen, Z. J.; Bai, X.; Zhang, X. Cryo-EM Structures of STING Reveal Its Mechanism of Activation by Cyclic GMP–AMP. *Nature* **2019**, 567 (7748), 389–393. <https://doi.org/10.1038/s41586-019-0998-5>.
- (5) Yin, Q.; Tian, Y.; Kabaleeswaran, V.; Jiang, X.; Tu, D.; Eck, M. J.; Chen, Z. J.; Wu, H. Cyclic Di-GMP Sensing via the Innate Immune Signaling Protein STING. *Molecular Cell* **2012**, 46 (6), 735–745. <https://doi.org/10.1016/j.molcel.2012.05.029>.
- (6) Ergun, S. L.; Fernandez, D.; Weiss, T. M.; Li, L. STING Polymer Structure Reveals Mechanisms for Activation, Hyperactivation, and Inhibition. *Cell* **2019**, 178 (2), 290–301.e10. <https://doi.org/10.1016/j.cell.2019.05.036>.
- (7) Zhang, X.; Shi, H.; Wu, J.; Zhang, X.; Sun, L.; Chen, C.; Chen, Z. J. Cyclic GMP-AMP Containing Mixed Phosphodiester Linkages Is An Endogenous High-Affinity Ligand for STING. *Molecular Cell* **2013**, 51 (2), 226–235. <https://doi.org/10.1016/j.molcel.2013.05.022>.
- (8) Gao, P.; Ascano, M.; Zillinger, T.; Wang, W.; Dai, P.; Serganov, A. A.; Gaffney, B. L.; Shuman, S.; Jones, R. A.; Deng, L.; Hartmann, G.; Barchet, W.; Tuschl, T.; Patel, D. J. Structure-Function Analysis of STING Activation by c[G(2',5')PA(3',5')p] and Targeting by Antiviral DMXAA. *Cell* **2013**, 154 (4), 748–762. <https://doi.org/10.1016/j.cell.2013.07.023>.
- (9) 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. Phosphorylation of Innate Immune Adaptor Proteins MAVS, STING, and TRIF Induces IRF3 Activation. *Science* **2015**, 347 (6227), aaa2630–aaa2630. <https://doi.org/10.1126/science.aaa2630>.
- (10) Wu, J.; Chen, Z. J. Innate Immune Sensing and Signaling of Cytosolic Nucleic Acids. *Annu. Rev. Immunol.* **2014**, 32 (1), 461–488. <https://doi.org/10.1146/annurev-immunol-032713-120156>.
- (11) Tanaka Yasuo; Chen Zhijian J. STING Specifies IRF3 Phosphorylation by TBK1 in the Cytosolic DNA Signaling Pathway. *Science Signaling* **2012**, 5 (214), ra20–ra20. <https://doi.org/10.1126/scisignal.2002521>.
- (12) Zhao, B.; Shu, C.; Gao, X.; Sankaran, B.; Du, F.; Shelton, C. L.; Herr, A. B.; Ji, J.-Y.; Li, P. Structural Basis for Concerted Recruitment and Activation of IRF-3 by Innate Immune Adaptor Proteins. *Proc Natl Acad Sci USA* **2016**, 113 (24), E3403–E3412. <https://doi.org/10.1073/pnas.1603269113>.
- (13) Shih, A. Y.; Damm-Ganamet, K. L.; Mirzadegan, T. Dynamic Structural Differences between Human and Mouse STING Lead to Differing Sensitivity to DMXAA. *Biophysical Journal* **2018**, 114 (1), 32–39. <https://doi.org/10.1016/j.bpj.2017.10.027>.

- (14) Tehrani, Z. A.; Rulišek, L.; Černý, J. Molecular Dynamics Simulations Provide Structural Insight into Binding of Cyclic Dinucleotides to Human STING Protein. *Journal of Biomolecular Structure and Dynamics* **2021**, 0 (0), 1–15. <https://doi.org/10.1080/07391102.2021.1942213>.
- (15) Chen, L.; Zhao, S.; Zhu, Y.; Liu, Y.; Li, H.; Zhao, Q. Molecular Dynamics Simulations Reveal the Modulated Mechanism of STING Conformation. *Interdisciplinary Sciences: Computational Life Sciences* **2021**. <https://doi.org/10.1007/s12539-021-00446-3>.
- (16) Ergun, S. L.; Li, L. Structural Insights into STING Signaling. *Trends in Cell Biology* **2020**, 30 (5), 399–407. <https://doi.org/10.1016/j.tcb.2020.01.010>.
- (17) Šali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology* **1993**, 234 (3), 779–815. <https://doi.org/10.1006/jmbi.1993.1626>.
- (18) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, 12 (1), 405–413. <https://doi.org/10.1021/acs.jctc.5b00935>.
- (19) Kim, S.; Lee, J.; Jo, S.; Brooks III, C. L.; Lee, H. S.; Im, W. CHARMM-GUI Ligand Reader and Modeler for CHARMM Force Field Generation of Small Molecules. *Journal of Computational Chemistry* **2017**, 38 (21), 1879–1886. <https://doi.org/10.1002/jcc.24829>.
- (20) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, 1–2, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>.
- (21) Huang, J.; MacKerell Jr, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *Journal of Computational Chemistry* **2013**, 34 (25), 2135–2145. <https://doi.org/10.1002/jcc.23354>.
- (22) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, 105 (43), 9954–9960. <https://doi.org/10.1021/jp003020w>.
- (23) Jo, S.; Kim, T.; Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS ONE* **2007**, 2 (9), e880. <https://doi.org/10.1371/journal.pone.0000880>.
- (24) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, 126 (1), 014101. <https://doi.org/10.1063/1.2408420>.
- (25) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *Journal of Applied Physics* **1981**, 52 (12), 7182–7190. <https://doi.org/10.1063/1.328693>.
- (26) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *Journal of Computational Chemistry* **1997**, 18 (12), 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H).
- (27) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, 103 (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
- (28) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, 14, 33–38.

- (29) Grant, B. J.; Skjærven, L.; Yao, X. The Bio3D Packages for Structural Bioinformatics. *Protein Science* **2021**, *30* (1), 20–30. <https://doi.org/10.1002/pro.3923>.
- (30) Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 2.5, 2015.
- (31) Theodorou, D. N.; Suter, U. W. Shape of Unperturbed Linear Polymers: Polypropylene. *Macromolecules* **1985**, *18* (6), 1206–1214. <https://doi.org/10.1021/ma00148a028>.
- (32) Payne, R.; Crivelli, S.; Pfeifer, H.; Le, L.; Watanabe, M. Computational Study of Structural and Energetic Effects on Human STING Protein by CGAMP and DMXAA Ligands. *Manuscript in preparation*.