Roles of Water Molecules in STING Activation: A Computational Perspective

Hannah Gates[1], Rachel T. Payne[2], Haley Pfeifer[3], Lyly Le[3],

Silvia Crivelli[4], and Masakatsu Watanabe[1*]

[1]Department of Chemistry, Fort Hays State University, Hays, KS 67601

[2]Boston University, Boston, MA 02215

[3]Kansas Academy of Mathematics and Science, Fort Hays State University, Hays, KS 67601,

[4]Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Present Address:

Hannah Gates: Bioinformatics & Computational Biology, Iowa State University, Ames, IA 50011, United States.

Rachel T. Payne: Computational Health Informatics Program, Boston Children's Hospital, Boston, MA 02115, United States

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

Abstract

The cyclic GMP-AMP synthase (cGAS) - stimulator of interferon genes (STING) pathway is crucial in the innate immune response, particularly in cancer immunotherapy. Despite promising preclinical results, 5,6-dimethylxanthenone-4-acetic acid (DMXAA) showed limited efficacy in human clinical trials due to species-specific differences in STING activation. This study investigates these differences by analyzing the binding dynamics and affinities of various STING-ligand complexes using molecular dynamics (MD) simulations and binding free energy calculations. We confirmed that specific point mutations, notably G230I and S162A/Q266I, significantly improve DMXAA's binding affinity to human STING (hSTING), replicating the

behavior observed in mouse STING (mSTING). Explicit solvent MD simulations revealed the essential role of water molecules in the binding site. Bridge water molecules, forming hydrogen bonds between the ligand and the protein, were significant in stabilizing the cyclic GMP-AMP (cGAMP) system, influencing local dielectric constants. Additionally, for DMXAA systems, we found that the mutations lowered the interaction energies required for ligand binding by reducing the number of water molecules and localizing them to specific locations within the binding site. These findings deepen our understanding of STING-DMXAA interactions and highlight potential pharmacological modifications required to enhance STING-targeted therapies. Integrating structural biology, computational simulations, and thermodynamic analyses offers a robust framework for advancing STING-based therapeutic development.

Introduction

The innate immune system is crucial for defending the body against nonspecific infectious pathogens. Central to this defense mechanism are receptor and adaptor proteins that activate signal transduction pathways upon detecting foreign DNA. One of the key pathways in this response is the cGAS-STING pathway (the cyclic GMP-AMP synthase - stimulator of interferon genes), for which the STING protein, a transmembrane protein located in the endoplasmic reticulum, plays a pivotal role[1].

When cGAS senses and binds cytosolic double-stranded DNA, it synthesizes cyclic GMP-AMP (cGAMP). STING is a direct sensor of cyclic dinucleotides (CDNs). When it senses a CDN, it binds to cGAMP, triggering a conformational change that allows STING to recruit TANK-binding kinase 1 (TBK1). This recruitment leads to the phosphorylation and activation of interferon regulatory factor 3 (IRF3) and nuclear factor-κB (NF-κB), which in turn stimulates the production of type 1 interferon (IFN1)[2,3].

Structural studies, particularly cryo-electron microscopy (cryo-EM)[4], have revealed that STING is a homodimer complex comprising a C-terminal ligand-binding domain (LBD), a C-terminal tail (CTT), linkers, and an N-terminal transmembrane (TM) region. The LBD forms a V-shaped dimer with a hydrophobic binding pocket, while the N-terminal TM contains eight helices within the dimer, including the phosphorylation site for TBK1. STING exists in two conformations: an open, ligand-unbound state (apo) and a closed, ligand-bound state (holo). In the holo state, ligand binding induces inward rotation and the formation of a β-sheet lid covering the binding site, as shown in Figure 1. This transition involves significant structural rearrangement, highlighted by a clockwise rotation of the LBD connector helices in the holo state[4,5].

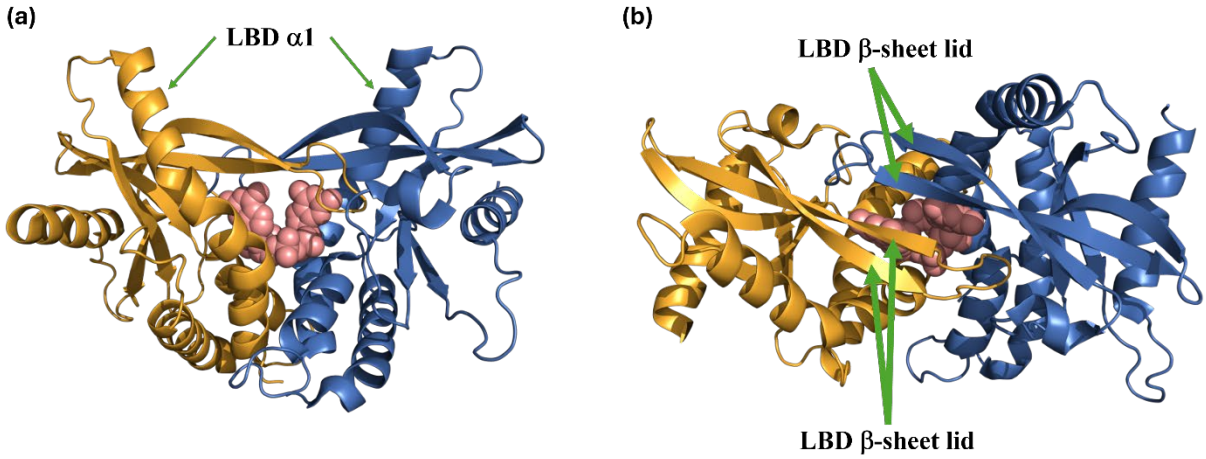

Figure 1 Ligand Binding Site of STING. *(a) The V-shaped ligand binding site is formed by the interaction of two α1 helices from the LBD (Ligand-Binding Domain). (b) A top-down view shows the β-sheets from Chains A and B acting as lids covering the binding site. The ligand is represented in pink, while the structural elements are illustrated with Chain A in orange and Chain B in blue, using a cartoon representation to distinctly differentiate between the two chains.*

We highlight this because the STING pathway is a critical focus in cancer therapy, with small-molecule STING-activating immunomodulators having been extensively studied as potential treatments for cancer[6–9]. An early example of a STING activator was the vascular disrupting agent 5,6-dimethylxanthenone-4-acetic acid (DMXAA), which demonstrated potential in mouse models[10]. DMXAA activates mouse STING (mSTING), leading to an aggressive antitumor immune response by activating tumor-associated macrophages and releasing cytokines. However, despite the high sequence identity (68.4%) and similarity (80.2%) between mSTING and human STING (hSTING), shown in Figure S1 of the supporting information, DMXAA failed to activate hSTING in human clinical trials[11]. This discrepancy highlights the need for a deeper investigation into the molecular interactions and conformational changes induced by STING agonists in humans.

Previous studies[12,13] have provided insights into the structural and functional disparities between hSTING and mSTING in response to DMXAA. Gao et al.[12] identified specific amino acid substitutions, such as S162A and Q266I, in hSTING that can mimic the sensitivity of mSTING to DMXAA. They also highlighted the critical role of the nonconserved residue I229 of mSTING, where the substitution G230I on hSTING facilitates a closed conformation, enhancing its binding affinity for DMXAA. Shih et al.[13] complemented these findings using implicit solvent molecular dynamics (MD) simulations, demonstrating that the G230I mutation creates a steric block, effectively preventing the escape of DMXAA, altering interaction dynamics, and stabilizing DMXAA within the binding pocket.

Given the failure of DMXAA in human trials, it is no longer considered a viable drug for human cancer treatment. However, pharmacological modification of DMXAA remains one of the promising therapeutic strategies[14–24]. Understanding these differences can guide the design of better-fitting DMXAA analogs to enhance hSTING's response and overcome the energy barrier for the transition to a closed conformation.

Our research utilized MD simulations to assist in decoding these complex pathways at the atomic level. MD simulations complement experimental studies by providing detailed insights into molecular interactions within biological systems. Recent MD studies[13,25–29] have focused on STING's structural and dynamic differences before and after ligand binding. Our research aims to validate and further elucidate the mechanisms underlying STING-ligand interactions using MD simulations and binding-affinity calculations. Importantly, we employed explicit solvent MD methods to provide a more accurate and detailed understanding of the solvation effects and the dynamic behavior of STING-ligand complexes. Explicit solvent simulations offer a more realistic representation of the biological environment, capturing critical

interactions between the solvent molecules and the protein-ligand complex that implicit solvent methods may overlook.

This study focuses on the conformational changes and binding dynamics induced by specific point mutations (S162A, Q266I, and G230I) in hSTING, aiming to provide a comprehensive molecular understanding of how these mutations impact STING's structure and function. Additionally, it evaluates and builds upon previous research, contributing to a deeper knowledge base that may guide the rational design of more effective therapeutic agents for human use, particularly in developing DMXAA analogs with enhanced efficacy for hSTING. The insights gained from this research will have broader implications for the field of immunotherapy and the development of treatments targeting diseases involving the STING pathway.

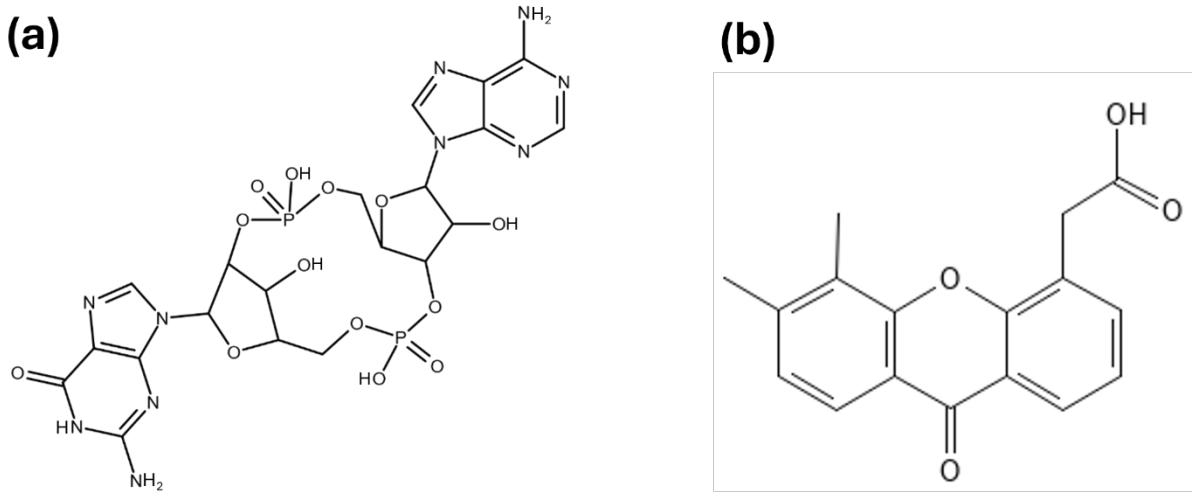

Figure 2. Structures of ligands used in this Study. (a) 2’3’-cGAMP and (b) DMXAA. (a) cGAMP molecule has 18 hydrogen bond acceptors and seven hydrogen bond donors. (b) DMXAA molecule has four hydrogen bond acceptors and one hydrogen bond donor.

Materials and Methods

Understanding the binding affinities of ligands to protein residues is crucial for uncovering the molecular mechanisms of their interactions and for designing effective therapeutic drugs. In this paper, we examine five different systems: cGAMP [wild-type human STING (wt-hSTING) with cGAMP ligand], DMXAA [wt-hSTING with DMXAA ligand], G230I [G230I mutant hSTING with DMXAA ligand], S162A [S162A/Q266I mutant hSTING with DMXAA ligand], and mDMXAA [wild-type mouse STING (wt-mSTING) with DMXAA ligand].

Selection and preparation of ligands

The 3D structures of 2'3'-cGAMP and DMXAA were obtained from the Protein Data Bank (PDB) using the codes 6NT7 and 4LOL, respectively. Their geometries were fully optimized employing density functional theory (DFT) with Becke's three-parameter hybrid exchange functional and the Lee-Yang-Parr correlation functional (B3LYP)[30,31] along with the 6-31G* basis set[32], as implemented in the Gaussian16 program[33]. The chemical structures of cGAMP and DMXAA are presented in Figure 2. Topologies and force-field parameters for cGAMP and DMXAA molecules were generated using the CGenFF program [34,35], with parameters for bonded and non-bonded interactions assigned according to the CHARMM36 force field [36].

Modeling of closed structure of wt-hSTING

In our previous study[28], we built full-human STING proteins for both apo and holo-structures in the membrane. Here, we extracted the ligand-binding domain (residues 154 to 341,

shown in Figure S1) from the holo full-human STING to be used as a closed form of *wt*-hSTING structure for our research.

To produce the two mutants STING proteins (G230I and S162A/Q266I), shown in Figure 3, point mutations were performed utilizing the Mutagenesis Wizard tool[37] provided by PyMOL[38]. Throughout the mutation process, the lowest energy conformation was consistently chosen for the side-chain conformation to ensure the stability and accuracy of the mutated proteins. This approach minimizes potential distortions and energetically unfavorable interactions, providing a reliable basis for subsequent structural and functional analyses.

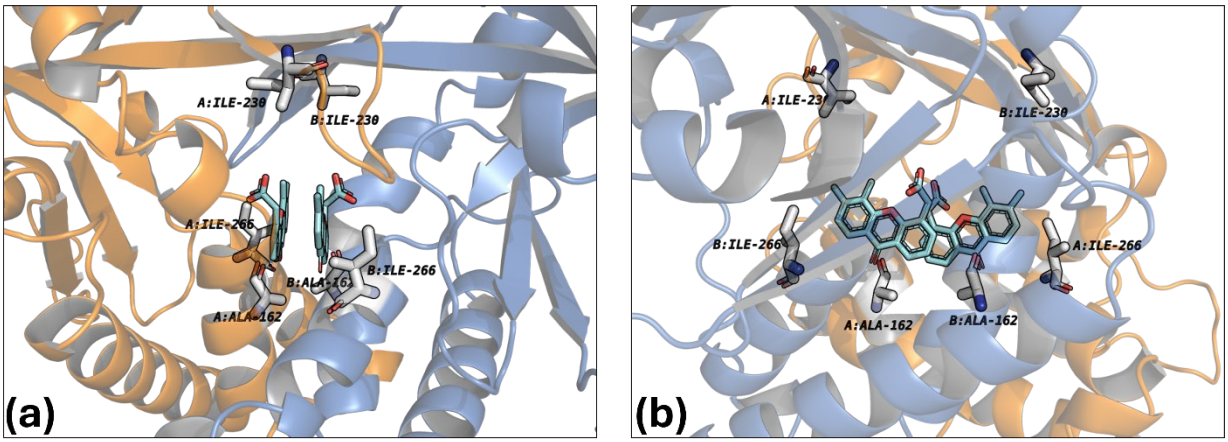

Figure 3. Structural representation of G230I and double mutations (S162A and Q266I) in hSTING. (a) The left panel highlights the G230I mutation in both chains (A and B) of hSTING, localized in the lid region of the protein. Additionally, the S162A and Q266I double mutations are shown at the bottom of the binding site. (b) The right panel provides a different perspective of the mutation sites. The structural elements are depicted with Chain A in orange and Chain B in blue, using a cartoon representation to distinguish the chains clearly.

Modeling wt-mSTING

To construct a closed form of *wt*-mSTING, we obtained the coordinates for mSTING [PDB ID: 4LOL] from the PDB. Since the PDB coordinate structures contained missing residues in the loop and turn regions, we used MODELLER[39], a protein homology modeling program, to

resolve this issue. Additionally, since the structures of mSTING and hSTING are very similar to each other, the crystal structure of the LBD portion of holo-hSTING [PDB ID: 4F5D] was used as a secondary structural template in MODELLER[39] to more accurately refine the LBD portions.

The final model was selected based on the GA341[40] and DOPE[41] scores. GA341 ranges between 0 (worst score) and 1 (best score), while the lowest DOPE value is considered to represent the best model. Although the DOPE score has been shown to be an extremely accurate model assessment score in a number of studies[42], it is generally recommended to use multiple assessment scores for a comprehensive evaluation. Since the GA341 score is best used for ruling out substandard models, any models whose values were less than 0.6 were discarded. Among the remaining models, we selected the structure with the lowest DOPE score as our final model.

Simulation Preparations

Once all STING protein models were completely constructed, we used the CHARMM-GUI web server[24] to build the initial solvated STING and ligand complex systems. The MD simulations were performed with the GROMACS 2020 software[43] and the CHARMM36 force field[36]. Each protein-ligand complex was initially positioned within a rectangular box, ensuring that each side of the box was at least 1.0 nm away from any atom of the protein. The simulation box was then solvated using TIP3P water molecules[29]. Then, if necessary, potassium or chloride ions were added to neutralize the total system charge.

After assembling the protein-ligand complex, several pre-production steps were 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 by applying the position restraints to a ligand molecule and non-hydrogen atoms

of STING protein. Then, an equilibration was performed by applying the same position restraints. 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, as shown in Table S2 of the supporting information.

Production MD Simulations

The last coordinates from the previous equilibration were further processed for the production runs. The production simulations were conducted at a constant number, pressure, and temperature (NPT) for a 1.0-μs simulation via the Parrinello-Rahman algorithm[44] at 1 bar pressure and the V-rescale algorithm[45] at 30ºC.

In all production cases, an integration time step of 2.0-fs was used. Periodic boundary conditions were applied in all directions throughout all simulations. The LINCS algorithm[46] was used only to constrain the bonds involving hydrogen atoms. The electrostatic interactions were calculated using a Particle mesh Ewald algorithm[47] with a 1.2 nm cut-off. The trajectories were saved every 50-ps during each simulation for analysis.

Reproducibility and reliability are essential features of the scientific method, whether experimental or computational. Scientific methods should reproduce the results in a statistical sense regardless of who performs them. However, chaotic dynamical systems, such as MD simulations, are extremely sensitive to initial conditions[48,49]. For instance, in this study, "*replicas*" refers to MD simulations consisting of identical initial structures, identical parameters, and varying initial velocities randomly assigned via the Maxwell-Boltzmann velocity distribution. These replicas often produce different trajectories caused by the roughness of the potential energy surface, which contains many local minima frequently separated by high energy

barriers. Thus, the free energy landscape of these simulations is often rugged, and minor differences in the initial conditions can cause replicas to take different paths on the phase-space surface[50]. As a result, we expected to see some differences in dynamical details between the replicas. Thus, conclusions obtained from a single MD simulation are usually insufficient. So, the reproducibility and reliability of MD simulations should be based on an averaged statistical sense.

Table 1 *Listing of Simulation systems that we studied.*

	Descriptions of System	
System Name	**STING Protein**	**Ligand**
cGAMP	wt-hSTING	2'3'-cGAMP
DMXAA	wt-hSTING	DMXAA
G230I	hSTING with G230I mutation	DMXAA
S162A	hSTING with S162A/Q266I mutations	DMXAA
mDMXAA	wt-mSTING	DMXAA

In this study, five replicas were utilized for each system, as detailed in Tables 1 and S2. Each replica underwent a production run of 1.0 μs. The primary objective was to identify the underlying similarities among the replicas. To this end, statistical properties for each system were calculated on the combined trajectories obtained from five replica simulations.

Binding free energy calculations

Free energy calculations are pivotal for understanding molecular interactions and stability in biological systems. These calculations have become invaluable for numerous applications in computational biology, such as drug design and protein structure determination. Several methodologies exist for calculating free energies, including Free Energy Perturbation (FEP)[51,52], Replica Exchange Free Energy Perturbation (REFEP)[53,54], and Thermodynamic Integration (TI)[55,56]. Although these methods are theoretically rigorous, they are computationally intensive and can become prohibitively expensive as system size increases[57].

The Molecular Mechanics/Poisson−Boltzmann (Generalized Born) Surface Area (MM/PB(GB)SA) method has emerged as one of the most popular approaches for estimating binding free energies since its introduction in the late 1990s [58,59]. This method achieves a favorable balance between accuracy and computational efficiency, offering greater precision than most scoring functions, particularly in protein−protein and protein−nucleic acid systems, while being less computationally demanding than alchemical free energy methods [59,60].

The gmx_MMPBSA tool[61] integrates the molecular dynamics package GROMACS with the MM/PB(GB)SA approach to compute the free energy of binding between biomolecules, such as proteins and ligands. Once the MD simulation is complete, gmx_MMPBSA extracts snapshots from the trajectory at specified intervals. These snapshots are used to calculate the free energy of binding using the MM/PB(GB)SA method. The approach combines molecular mechanics energies with solvation terms derived from Poisson-Boltzmann and Surface Area calculations. Specifically, it decomposes the free energy into contributions from van der Waals interactions, electrostatic interactions, polar solvation energy, and non-polar solvation energy.

The gmx_MMPBSA tool includes calculations of entropic contributions to free energy changes using the Interaction Entropy (IE)[62] or second-order cumulant approximation (C2)[63] methods, which evaluate the entropy associated with molecular motions and fluctuations. In this study, we employed the C2 method to estimate the binding entropy contribution, as it demonstrated better convergence, albeit with a higher standard deviation of the interaction energies (σ_{IE}) compared to the IE method[64]. Specifically, the C2 method showed reliable entropy convergence when σ_{IE} was less than 6 kcal/mol, whereas the IE method required σ_{IE} to be below 3 kcal/mol to achieve reliable convergence[64].

To ensure the convergence of the free energy calculations, the last 500 ns of trajectories were divided into 20 segments, each 25 ns in length. This segmentation allows for the investigation of the convergence of the calculations, ensuring that the computed free energies are reliable.

Results and Discussion

We inspected the conditions and characteristics of the trajectories of the protein-ligand complex by VMD[65] and PyMOL[38] and then proceeded with further analysis. For this, we utilized our in-house developed codes in Python and R along the GROMACS suite.

First, we analyzed the α-carbon (Cα) root mean square deviation (RMSD) and ligand RMSD. The results are depicted in Figures S2 and S3 in the supporting information. The RMSD was used to analyze the structural evolution of the protein-ligand complex over time.

Figure S2 displays the Cα-RMSD for five different systems (cGAMP, DMXAA, G230I, S162A, mDMXAA) over a 1.0 μs simulation period, showing how each system deviates from its initial conformation. Initially, RMSD values increase for all proteins, reflecting early

conformational changes, before stabilizing, which suggests that each protein reaches a stable conformation. cGAMP and S162A exhibit slightly higher RMSD values and fluctuations. In contrast, DMXAA and G230I maintain lower, more consistent RMSD values, indicating greater stability. mDMXAA shows an initial rise in RMSD but quickly stabilizes, maintaining lower deviation throughout. Generally, RMSD values remain below 5.0 Å, indicating no drastic conformational changes. The consistency across multiple lines within each protein's plot further supports the reproducibility of the results.

Figure S3 shows the RMSD of the ligand in the various systems to check the structural stability of the ligand within proteins over time. In contrast to the Cα-RMSD, which showed relatively consistent trends, the ligand RMSD exhibits greater variability in all cases. cGAMP and G230I exhibit relative stability with consistent RMSD values, indicating stable conformations post-initial adjustment. DMXAA exhibits greater instability and more frequent conformational changes, as indicated by higher RMSD values. While S162A and mDMXAA remain flexible, they demonstrate slightly more stability and less severe fluctuations. However, these variants still show slightly higher fluctuations, indicating that the ligand undergoes some conformational changes in the binding site. These RMSD results already indicate that the DMXAA system (wt-hSTING with DMXAA) might have a weaker binding affinity than the other systems. Also, based on the observation of RMSD graphs, we discarded the first 500 ns of the MD trajectory to eliminate the bias from the initial states for the rest of the analyses in this paper.

The root-mean-square fluctuation (RMSF) of Cα-atoms of protein residues was calculated using the last 500 ns of the simulations. Each replica of the system shows some differences in the detailed fluctuations. However, the primary objective was to identify the

underlying similarities among the replicas. To this end, statistical properties for each system were calculated on the combined trajectories obtained from five replica simulations. The average RMSFs for all systems are represented in Figure S4.

Since the flexible protein tail (residues beyond 330) shows very high RMSF values, we focus on examining ligand-induced alterations within the protein's core domain. RMSF calculations focused on Cα atoms across all models reveal a consistent distribution of peak intensities within this region. While both chains A and B exhibit correlated motion, with similar fluctuation patterns, subtle disparities in their RMSF values suggest distinct self-correlation behaviors despite the protein's homodimeric nature.

Characteristics of Ligands: cGAMP and DMXAA

Lipophilicity, quantified by log P (logarithm of the partition coefficient)[66], is a measure of how a compound distributes between octanol and water, indicating its hydrophilicity or hydrophobicity. A higher log P value signifies greater hydrophobicity. Additionally, log S (ESOL)[67] describes how well it dissolves in water. ESOL (Estimated SOLubility) predicts the solubility of organic compounds based on their molecular structure. Here, a higher log S(ESOL) value indicates greater hydrophilicity. In our study, the log P and log S(ESOL) values for DMXAA, its dimer, and cGAMP were calculated using the SwissADME web tool[68].

Table 2 *Log P and Log S values for ligands.*

Ligand	Lipophilicity (Log P)	Water Solubility (Log S(ESOL))	Water Solubility Class
DMXAA Monomer	2.99	−3.94	Soluble
DMXAA Dimer	4.60	−7.55	Poorly Soluble
cGAMP	−4.25	−0.19	Very Soluble

Table 2 shows the log P and log S values for different ligands. For instance, the DMXAA monomer has a log P of 2.99 and a log S of −3.94, classifying it as soluble. However, the DMXAA dimer is poorly soluble, with a log P of 4.60 and a log S of −7.55. cGAMP, with a log P of −4.25 and a log S of −0.19, is highly soluble in water.

Our calculations corroborate the established hydrophobicity of DMXAA, attributed to its aromatic structure and lack of polar groups. Similarly, the hydrophilicity of cGAMP, enhanced by its polar phosphate and hydroxyl groups as well as its cyclic dinucleotide structure, was confirmed by our findings. Hence, the poor solubility of the DMXAA dimer suggests that an additional process is required to displace water molecules from the ligand-binding site. This process demands extra energy, which must be compensated by favorable interactions between DMXAA and the protein to achieve a lower binding free energy.

Dynamical characteristics of STING protein

Ligands and mutations can affect the protein's stability and dynamics. To describe the dynamical characteristics of STING in the various systems, we monitored pairwise cross-correlation coefficients, representing how the atomic fluctuations/displacements of a system are correlated. We investigated the Linear Mutual Information (LMI) maps[69] of all five systems. Although dynamical cross-correlation (DCC) maps are commonly used, they have limitations. Specifically, if two atoms move in perpendicular directions simultaneously, DCC cannot measure their correlation due to the dot product involved in its calculation. Conversely, LMI does not have this unwanted dependency on the relative orientation of fluctuations[70]. Despite expected differences among system replicas with different initial velocities, our focus is on the underlying similarities. Thus, we present the averaged LMI maps for all five systems in Figure 4.

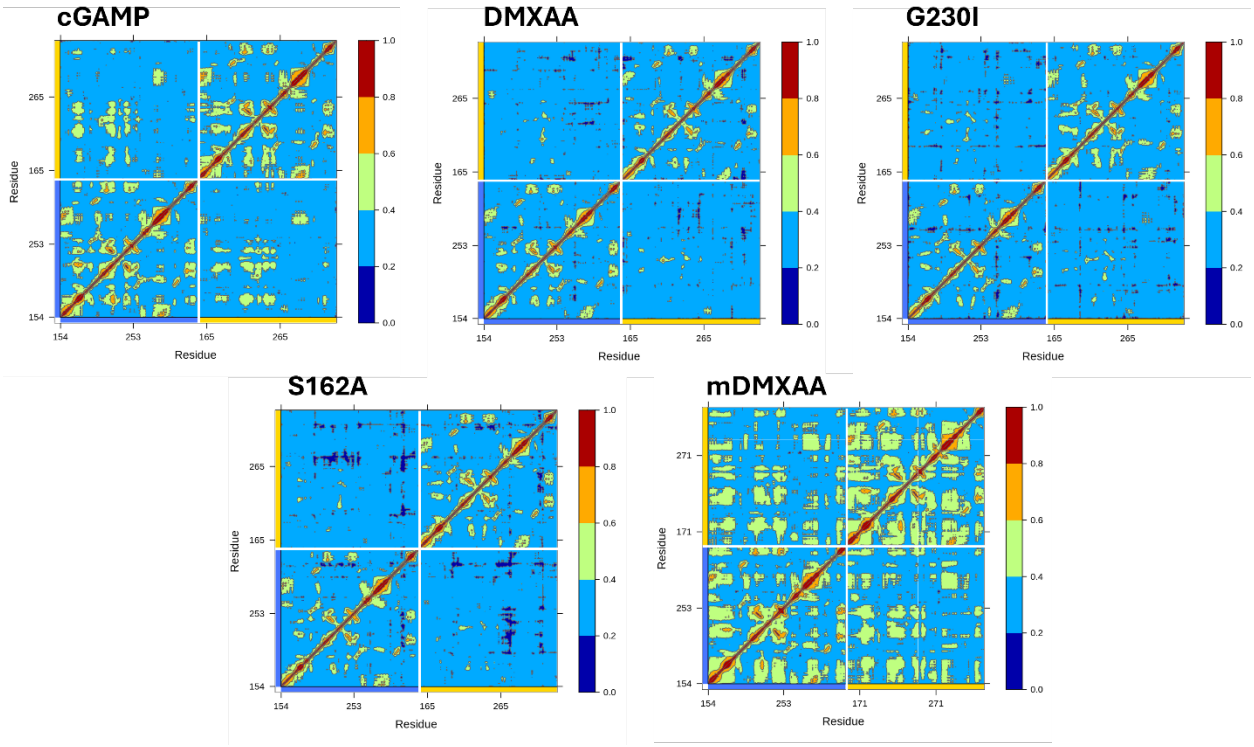

Figure 4 ***Averaged linear mutual information maps for chains A and B across five systems***. The figure displays Linear Mutual Information (LMI) maps comparing inter-residue correlations between chains A and B for the five systems. The x- and y-axes represent the residues of Chains A and B, respectively. The color scale, ranging from blue to red, indicates the degree of correlation, with blue representing the lowest and red the highest levels of inter-residue correlation. The blue and yellow regions on the axes correspond to Chains A and B, respectively.

The cGAMP and mDMXAA systems demonstrate a higher degree of inter-residue correlations compared to the DMXAA, S162A, and G230I systems. The mDMXAA system especially exhibits extensive correlations, indicating significant inter-residue communication throughout the protein. This observation aligns with the findings of Shih et al.[13], who, in their dynamic network analysis of the mDMXAA system, reported a denser network, further supporting the extensive correlations within the system.

In contrast, the DMXAA, S162A, and G230I systems display fewer and more dispersed significant correlations, indicative of weaker interaction networks. Particularly, no significant

off-diagonal correlations exist in any of these three systems, suggesting limited intra-chain residue interaction. Furthermore, G230I and S162A/Q266I mutations do not introduce notable enhancement or disruptions in inter-residue interactions, maintaining similar dynamic behavior to the DMXAA-bound state of *wt*-hSTING. The difference LMI maps, presented in Figure S5, also support these observations. In addition, these correlation maps imply that species differences, beyond simple mutations, may contribute to the enhanced susceptibility of mSTING to the DMXAA binding.

Roles of Water Molecules in Binding Site

Water molecules in biological systems can play a significant role in ligand binding, which is essential for understanding the mechanisms of biological processes and drug development[71–73]. One key aspect is that water molecules can act as a bridge between the ligand and the protein. This bridging facilitates the formation of hydrogen bonds, thereby enhancing the stability of the ligand-protein complex, which is vital for the proper function of the protein[74].

In addition to their role in hydrogen bonding, water molecules can also influence ligand binding by modifying the local environment of the binding site. Specifically, water molecules in the active site can increase the overall dielectric constant by providing a highly polarizable environment[75].

The increase in the dielectric constant can reduce the electrostatic interactions between charged groups in the active site and the ligand. This reduction in electrostatic interactions makes it easier for the ligand to bind to the protein, thereby facilitating the binding process[72]. Understanding these roles of water molecules can significantly aid in designing more effective drugs that target specific proteins[57,76]. For this reason, we examined the influence of water

molecules on protein-ligand interactions by analyzing their contribution to the overall binding process withing the binding site.

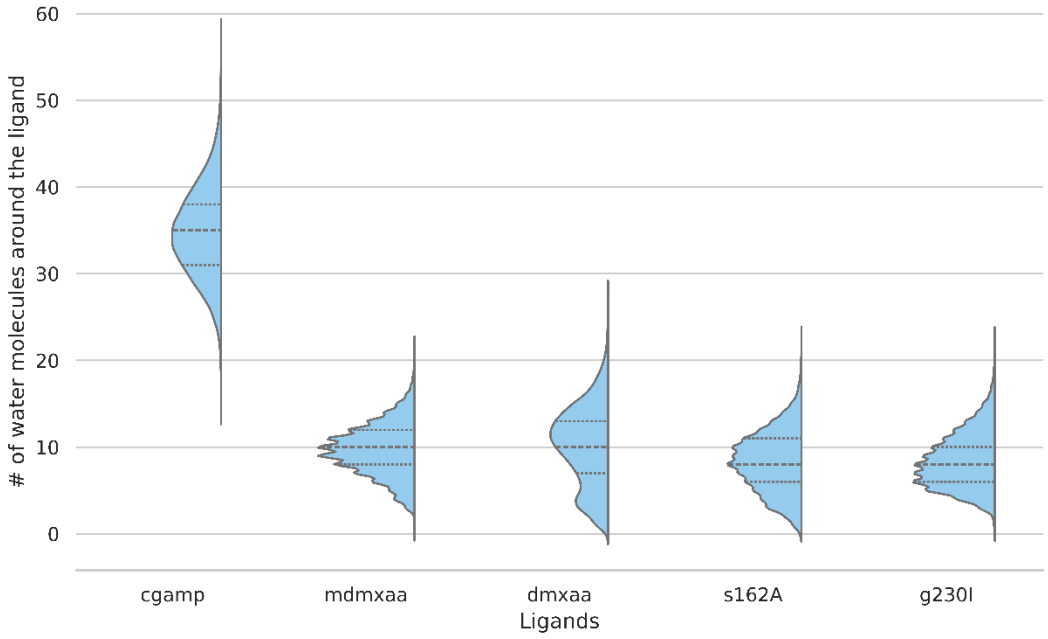

Figure 5 Distribution of the number of water molecules within 4 Å of each ligand molecule, based on the last 500 ns of trajectory data for each system.

(a) Numbers of water molecules in the binding site

To understand the influence of water molecules within the binding site, we evaluated the number of water molecules within 4 Å of each ligand molecule. According to Figure 5, the analysis revealed significant differences among these systems. cGAMP exhibited the highest average interaction value (34.8) with a standard deviation of 5.2. Other systems—DMXAA, G230I, S162A, and mDMXAA—had significantly lower averages, ranging from 8.2 to 10.1.

These observations suggest that in the cGAMP system, water molecules play a crucial role in modulating ligand binding by significantly altering the local environment of the binding site. In contrast, the other systems involving the DMXAA ligand exhibit fewer surrounding water

molecules, indicating a relatively dry and less hydrated environment. This dryness may result in a more hydrophobic binding site with a stronger dependence on direct ligand-protein interactions rather than water-mediated interactions. These observations are vital for the accurate calculation of ligand binding energy using the gmx_MMPBSA tool, as discussed in detail in a subsequent section. Additionally, these findings corroborate the characteristics of the ligands presented in the previous section: cGAMP is hydrophilic, whereas DMXAA molecules are hydrophobic.

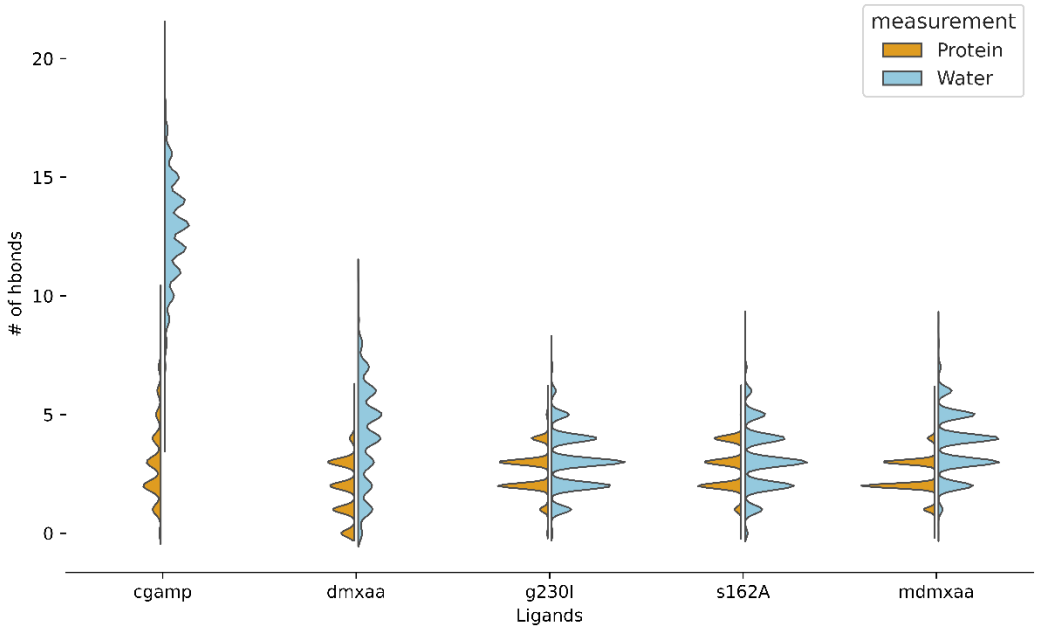

Figure 6 **Distribution of the number of hydrogen bonds for each system during the last 500 ns of simulation.** The orange bars (left) represent the number of hydrogen bonds between the protein and ligand, while the blue bars (right) indicate the number of hydrogen bonds between the ligand and water molecules.

(b) Hydrogen bonds

To further understand the interactions between water molecules and ligands, we calculated the number of hydrogen bonds between the ligand and water molecules, as well as between the ligand and protein, using GROMACS[43]. The formation of hydrogen bonds is

determined by the following geometric criteria in GROMACS,: (1) The distance between the donor (D) and acceptor (A) atoms must be less than or equal to 3.5 Å, and (2) The angle formed between the donor-hydrogen-acceptor (D-H···A) must be less than or equal to 30 degrees.

Figure 6 displays significant differences among the systems in the number of hydrogen bonds between the protein and ligand. cGAMP exhibited the highest average value (3.1), slightly higher than the other systems. The G230I, mDMXAA, and S162A systems displayed intermediate average values (2.6, 2.4, and 2.7, respectively) with similar distribution patterns, while DMXAA had the lowest average number (1.9).

For hydrogen bonds between water and ligand, cGAMP again had the highest average value (12.8), significantly higher than all other ligands. DMXAA followed with a lower average (4.1), but still higher than G230I, mDMXAA, and S162A (3.0, 3.7, and 3.1, respectively).

This analysis clearly demonstrates that DMXAA ligand's interactions with *wt*-hSTING differ significantly from those observed in other systems involving DMXAA as a ligand. These differences could also be evidence of the ineffectiveness of DMXAA in binding to *wt*-hSTING.

Ligand Binding Free Energy

(a) Full-length vs LBD hSTING Binding Free Energies

In our previous work[28], we investigated the full-length wild-type human STING (wt-hSTING) protein in both its apo and holo states within the membrane. We reused these trajectories to compare the binding free energy of cGAMP with both the full-length hSTING in the membrane and the ligand-binding domain (LBD) only simulations aiming to validate the use of the LBD alone in assessing the binding affinity of cGAMP. To enhance the robustness of our findings, we extended the three replica simulations of the holo full-length wild-type hSTING with cGAMP in the membrane, conducted in a previous study[28], from 1.0 μs to 1.5 μs.

The binding free energy (ΔG), enthalpy change (ΔH), and entropy contribution (-TΔS) of cGAMP with full-length *wt*-hSTING and LBD *wt*-hSTING were calculated using gmx_MMPBSA[61]. To ensure the reliability and accuracy of the computed free energy calculations, the last 500 ns of trajectories were divided into 20 segments, every 25 ns in length. The results for both full-length and LBD *wt*-hSTING proteins with cGAMP are presented in Figures S6 and S7.

Since the convergence of the entropic term is often problematic in the MM/PB(GB)SA approach[64], it is necessary to monitor the standard deviation of the interaction energies (σ_{IE}). For the C2 method used, the threshold is 6.0 kcal/mol. According to Figures S6 and S7, the σ_{IE} values for full-length STING and cGAMP simulations demonstrate good convergence across multiple replicas. In both datasets, σ_{IE} values are generally below the threshold required for reliable entropy calculations using the C2 method. All full-length STING replicas and most cGAMP replicas exhibit stable interaction energies, with an exception in cGAMP Replica 2 around window 14, where the standard deviation exceeds the threshold, indicating a potential need for further investigation. Overall, both simulations show consistent and reproducible results, confirming the reliability of the interaction energy measurements. Based on these observations, we used the last 150 ns of each replica's trajectory to calculate the combined average free energies for both systems.

Table 3 *Ligand binding freeeEnergies of cGAMP with full-length and only LBD hSTING.*

Model	ΔG (kcal/mol)	ΔH (kcal/mol)	-TΔS (kcal/mol)
Full-length hSTING	-34.84 ± 3.62	-46.28 ± 3.53	11.45 ± 0.81
LBD hSTING	-37.21 ± 3.47	-49.02 ± 3.35	11.81 ± 0.88

The results, presented in Table 3, show that the LBD *wt*-hSTING (-37.21 ± 3.47 kcal/mol) has a slightly lower ΔG than the full-length wt-hSTING (-34.84 ± 3.62 kcal/mol). This slight difference, within the standard deviation range, indicates that both forms have similar binding affinities. Both models' values and standard deviations of ΔH and -TΔS further support this conclusion. Therefore, the study justifies using only the LBD region for binding energy calculations, as it provides a reasonable approximation to the full-length protein, simplifying the analysis without significant loss of accuracy.

Table 4: *Thermodynamic analysis of binding interactions in cGAMP, DMXAA, G230I, S162A, and mDMXAA Systems.*

System ID	ΔG (kcal/mol)	ΔH (kcal/mol)	-TΔS (kcal/mol)
cGAMP	-37.21 ± 3.47	-49.02 ± 3.35	11.81 ± 0.88
DMXAA	-5.48 ± 6.53	-28.23 ± 6.31	22.75 ± 1.69
G230I	-31.76 ± 5.12	-48.72 ± 4.94	16.96 ± 1.35
S162A	-28.65 ± 5.25	-47.37 ± 5.03	18.72 ± 1.48
mDMXAA	-33.00 ± 5.65	-52.92 ± 5.45	19.92 ± 1.47

(b) Binding Free Energies of five systems

In this section, we compare the binding free energies of five systems: cGAMP, DMXAA, G230I, S162A, and mDMXAA. Prior to our calculations, we adjusted the internal dielectric constant, "*indi*," in the gmx_MMPBSA calculations based on the observed differences in the number of water molecules in the binding site. According to Figure 2, the number of water molecules in the binding site of the cGAMP system is approximately three times higher than in the DMXAA systems. This discrepancy necessitated the use of different "*indi*" values between

the cGAMP and other systems during the gmx_MMPBSA calculations. The "*indi*" parameter adjusts the internal dielectric constant based on the environment, requiring higher values to reflect environments with more water molecules or polar groups accurately. Consequently, we reduced the "*indi*" value from 2.5 for cGAMP to 1.0 for the other systems.

Then, to ensure the reliability of our calculations, the last 500 ns of trajectories were again divided into 20 segments, each 25 ns in length, to monitor the convergence of the calculations. The results for individual replicas of each system are presented in Figures S7 to S11. With the exception of a few segments in some replicas, the σ_{IE} values are below the threshold required for reliable entropy calculations using the C2 method. This consistency across multiple simulations indicates that the interaction energy measurements are robust and reproducible. Consequently, we conclude that the simulations provide reliable data, reinforcing the validity of the interaction energy assessments.

We used the last 150 ns of each replica's trajectory to calculate the combined average free energies for all systems, which are presented in Table 4. cGAMP exhibits the most favorable binding ($\Delta G = -37.21 \pm 3.47$ kcal/mol) with a highly favorable binding enthalpy ($\Delta H = -49.02 \pm 3.35$ kcal/mol) and moderate entropy decrease ($-T\Delta S = 11.81 \pm 0.88$ kcal/mol).

DMXAA shows the least favorable binding ($\Delta G = -5.48 \pm 6.53$ kcal/mol) with significant variability. It has a less favorable binding enthalpy ($\Delta H = -28.23 \pm 6.31$ kcal/mol) and the highest entropy decrease ($-T\Delta S = 22.75 \pm 1.69$ kcal/mol), suggesting a complex and less favorable binding process.

G230I and S162A display similar thermodynamic profiles, as their values fall within overlapping standard deviations. G230I ($\Delta G = -31.76 \pm 5.12$ kcal/mol) and S162A ($\Delta G =$

−28.65 ± 5.25 kcal/mol) both show favorable binding and favorable enthalpy (ΔH= − 48.72 ± 4.94 kcal/mol for G230I and ΔH = −47.37 ± 5.03 kcal/mol for S162A). Their entropy contributions (−TΔS = 16.96 ± 1.35 kcal/mol for G230I and −TΔS = 18.72 ± 1.48 kcal/mol for S162A) are also similar, indicating comparable binding characteristics. mDMXAA shows significant favorable binding (ΔG = −33.00 ± 5.65 kcal/mol) and has the most favorable binding enthalpy (ΔH = −52.92 ± 5.45 kcal/mol).

While cGAMP's comparison with other systems is complex due to different dielectric constants influenced by water molecules, comparisons among DMXAA, G230I, S162A, and mDMXAA are valid. DMXAA shows less favorable binding conditions, whereas G230I and S162A present similar favorable binding behaviors. mDMXAA and cGAMP demonstrate the most favorable binding energetics. In the next section, we investigated detailed studies to understand the specific interactions and structural features contributing to the observed stability, entropy, and enthalpy changes.

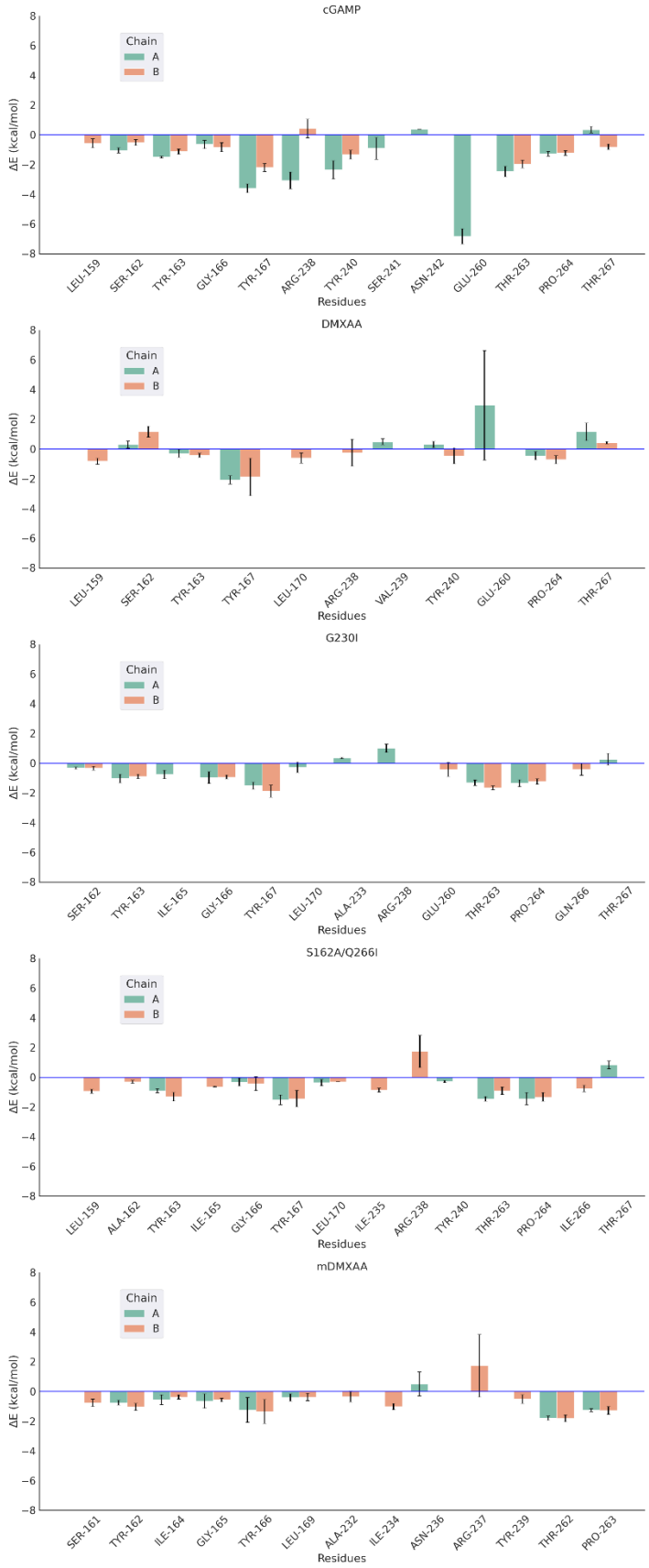

Figure 7. Comparative free energy contributions per residue in different Systems. The graphs depict the free energy contributions (ΔE) per residue for five different systems: cGAMP, DMXAA, G230I, S162A, and mDMXAA. Only residues within 4Å of the ligand and with contributions greater than 0.25 kcal/mol are included. The ligand, being the highest energy contributor, is excluded from the graphs. As we indicated earlier, the numbering systems of human and mouse STING proteins are offset by one residue,

Specific interactions – Decompositions of free energy contribution per residue

In our earlier study[28], we identified the residues interacting with cGAMP in the holo state of full-length STING. Based on these findings, we analyzed the binding free energy contributions of each residue. Figure 7 highlights residues contributing more than 0.25 kcal/mol and located within 4 Å of the ligand, revealing crucial insights into their stability factors. Additionally, the heatmap for the average free energy contributions per residue across different systems is presented in Figure S12.

First, common stabilizing and destabilizing factors are identified by comparing aligned residues in the G230I, S162A, and mDMXAA systems. Two tyrosine residues, TYR-163 and TYR-167, and their equivalents, TYR-162 and TYR-166 for mDMXAA, consistently exhibit negative free energy contributions, indicating their stabilizing effects. These residues likely engage in critical hydrogen bonding and hydrophobic interactions reinforcing system stability. Additionally, residues THR-263 (mDMXAA: THR-262), as well as PRO-264 (mDMXAA: PRO-263), also show negative energy contributions. These findings suggest that these threonine and proline residues contribute to stabilizing interactions, possibly through maintaining structural integrity and forming favorable local interactions.

In contrast, ARG-238 (mDMXAA: ARG-237) shows positive energy contributions, indicating destabilizing effects. The positive contributions from these arginine residues suggest involvement in less favorable interactions or potential disruption of local stability, possibly due to unfavorable electrostatic interactions or steric hindrance.

The cGAMP system stands out for its high stability among the systems analyzed, with an overall ΔG value of -37.21 ± 3.47 kcal/mol. This substantial negative free energy indicates robust stabilization. Key residues such as TYR-163, TYR-167, THR-263, and PRO-264

contribute significantly to the system's stability, showing strong negative energy contributions similar to those in the G230I, S162A, and mDMXAA systems. Additionally, the cGAMP system exhibits a favorable contribution from ARG-238 and GLN-260, differing from other systems. These residues likely form critical hydrogen bonds and hydrophobic interactions, reinforcing the system's structural integrity. Unlike DMXAA, cGAMP has fewer residues with positive energy contributions, indicating minimal destabilizing interactions. The cGAMP system's stability can be attributed to a network of strong stabilizing interactions and the absence of significant destabilizing residues.

The DMXAA system exhibits unique stability characteristics compared to the other systems. With an overall ΔG value of -5.48 ± 6.53 kcal/mol, DMXAA is notably less stable. Key residues such as TYR-163 and TYR-167 in DMXAA contribute less negatively to the free energy compared to their equivalents in other systems, indicating weaker stabilizing interactions. Furthermore, GLU-260 in DMXAA provides a significant unfavorable contribution, differing from its counterparts in the other systems. Figure S12 shows that the DMXAA system features several residues with positive contributions, suggesting more destabilizing interactions overall. Additionally, the error bars for DMXAA are generally larger, suggesting less consistent interactions and higher variability in binding. These variations in residue interactions and contributions result from differences in local environments, residue positioning, and structural conformation.

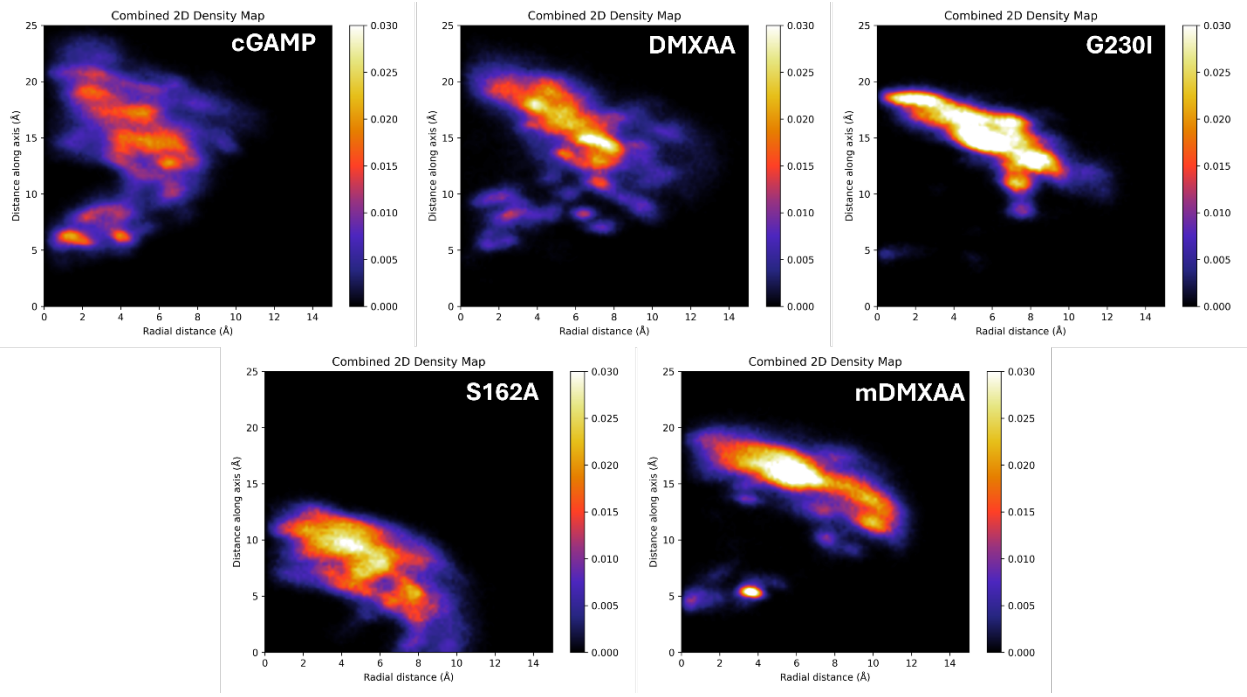

Figure 8 Averaged 2D Density maps of water molecule distributions for five different systems. The maps illustrate the spatial distribution of water molecules projected onto a plane parallel to the axis defined by the midpoints of the centers of mass of the GLY-158 (GLY-157 for mSTING) and GLN-184 (GLN-183 for mSTING) residues in chains A and B. The x-axis represents the radial distance (Å) from the defined axis, while the y-axis represents the distance along the axis (Å). The color scale indicates the water molecule density, with white areas representing higher density and blue areas representing lower density.

Spatial distributions of water molecules in the binding site

An interesting aspect of water molecules within the binding site of the closed-form STING is their spatial distribution. To investigate this, we first monitored water molecules within 3.5 Å of both the ligand and the protein to identify those located within the binding site. We then analyzed the spatial distribution by defining an axis that connects the midpoint of the centers of mass of the GLY-158 residues (mSTING: GLY-157) in chains A and B to the midpoint of the GLN-184 (mSTING: GLN-183) residues in the chains A and B. This axis serves as a reference for measuring radial distances, as illustrated in Figure S13. The water molecule density was then projected onto a plane parallel to this axis to analyze the spatial distribution relative to the

defined axis and radial distances. The midpoint between the centers of mass of the GLY-158 residues was set as the origin for these calculations.

Additionally, the STING's ligand-binding domain (LBD) features a V-shaped binding pocket. The vertex of this pocket, formed where two long helices converge and are closest, creates the apex of the V-shape near the residue GLY-158. Conversely, the outer side of the binding pocket, where the two helices are farthest apart, corresponds to the broader opening of the V-shape around GLY-184.

The analysis of 2D density maps for cGAMP, DMXAA, G230I, S162A, and mDMXAA, shown in Figure 8 and S14, reveals intriguing variations in the spatial distribution and concentration of water molecules within the binding site of each system. Figure S14 highlights distinct high-density regions and illustrates the variations in water molecule distribution and interaction within different molecular environments. These patterns offer valuable insights into the structural and interactive properties of water molecule distribution within these binding sites. Additionally, the density maps in Figure 8 show that the data from all replicas of each system were averaged for analysis.

The 2D density maps, presented in Figure 8, reveal water molecule distributions within the binding sites of different STING-ligand systems, highlighting key interactions influencing stability. For cGAMP with wt-hSTING, the two distinct high-density water regions are consistent with cGAMP’s hydrophilic nature. These regions indicate favorable water interactions that contribute to a structured and stable binding environment, consistent with the known stability of this system.

In the *wt*-hSTING with DMXAA system, the elongated high-density water region suggests an unstable interaction for the hydrophobic DMXAA, as it likely disfavors extensive

water interaction, contributing to instability. In contrast, the G230I and S162A/Q266I mutations with DMXAA display more localized high-density water regions, indicating a more stable binding environment. These mutations likely reduce water spread, creating a favorable environment for DMXAA, which aligns with observed stability in these systems.

In the mDMXAA system, which involves the *wt*-mSTING bound to DMXAA, two distinct high-density areas were observed. One peak is elongated, resembling the pattern seen in the G230I and S162A systems, while the other is much more localized and concentrated. This very localized high-density spot suggests the presence of one or two water molecules on the vertex side of the binding pocket.

Overall, the water density maps correlate with the hydrophobic or hydrophilic properties of the ligands and their stability in the binding sites. DMXAA's instability in wt-hSTING is marked by extensive water distribution, while G230I, S162A, and mSTING provide more localized water distribution in the binding sites. The stable binding of hydrophilic cGAMP is well-supported by distinct high-density water regions, emphasizing the crucial role of water interactions in ligand stability within STING binding sites. In the G230I, S162A, and mSTING systems, high-density regions are predominantly concentrated near the outer side of the binding pocket, particularly in the area between the lid regions of the protein and the ligand. This observation suggests that water molecules may contribute to stabilizing the ligand within the binding site. This mechanism contrasts with previous studies that primarily attributed ligand stabilization to the lid alone[13].

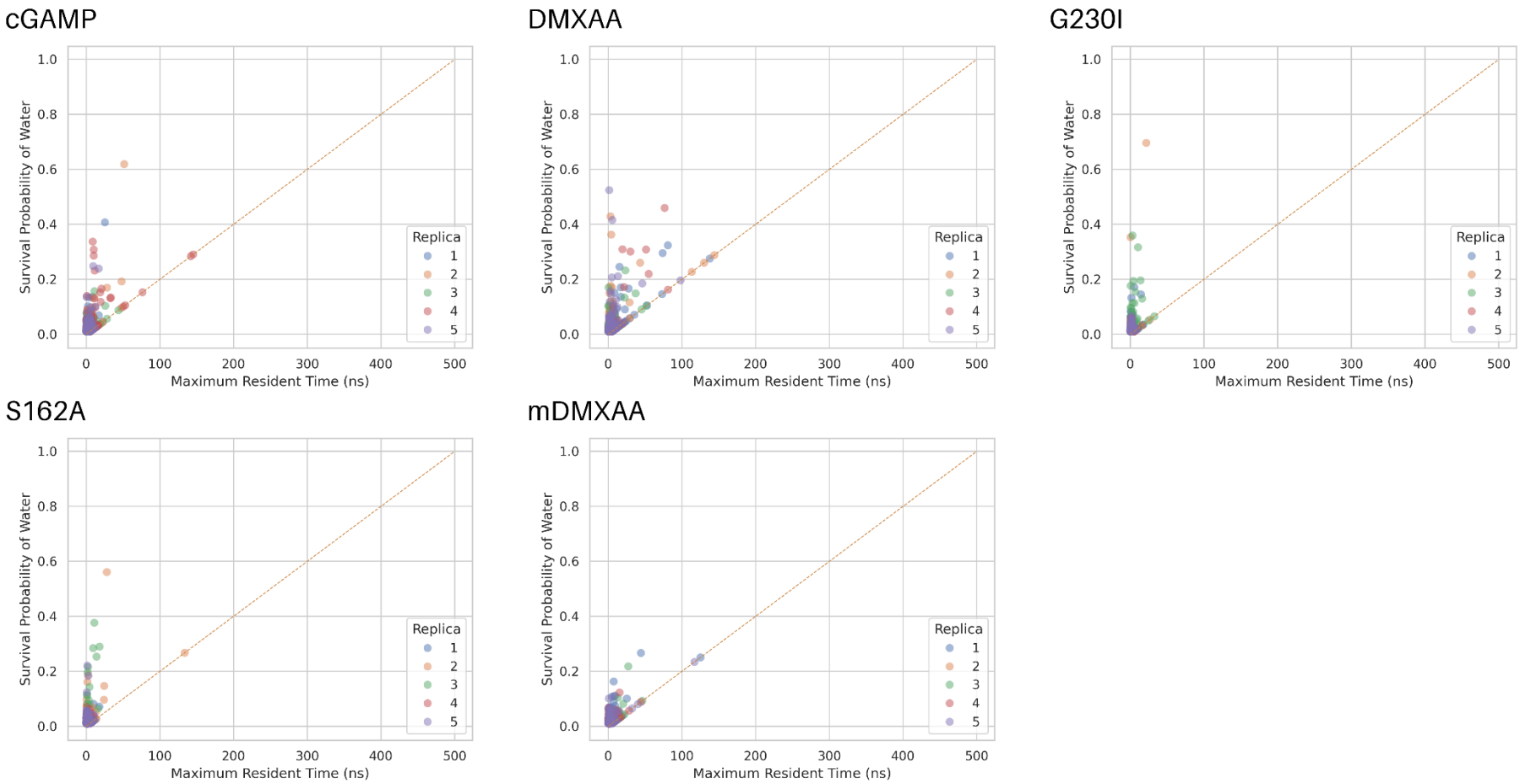

Figure 9 The behavior of water molecules in the binding site for five systems. The graphs compare the behavior of water molecules in five systems (cGAMP, DMXAA, G230I, S162A, and mDMXAA) over the final 500 ns of 1.0 μs simulations. Each point on the graph represents a specific water molecule in the binding site; each color indicates a different replica of the simulation. The X-axis (Maximum Resident Time) indicates the maximum consecutive residence time of water molecules in the binding site. The Y-axis (Survival Probability) represents the fraction of the simulation time that a particular water molecule spends in the binding site. For instance, a Y-axis value of 0.6 means a water molecule is in the binding site for 300 ns (or 60%) of the 500-ns simulation, not necessarily consecutively.

Resident time of water molecules in the protein binding site

Here, we aim to determine whether these water molecules in the binding site are transient or permanently positioned. To elucidate the residence times of water molecules within the binding site, we analyzed the behavior of specific water molecules during the simulation period by quantifying two key metrics: (1) the Maximum Residence Time (MRT) and (2) the Survival Probability (SP). The MRT refers to the longest consecutive duration that a particular water molecule remains within the binding site, providing insight into its stability and interaction with the site. In contrast, the SP measures the proportion of the total simulation time that a particular water molecule occupies the binding site, irrespective of whether this occupancy is continuous or intermittent, offering a broader perspective on the water molecule's overall engagement with the binding site.

Figure 9 reveals a common pattern across all systems: most water molecules exhibit low maximum residence times and low survival probabilities, clustering near the origin. This consistency suggests that transient water molecules dominate the behavior in these environments. The data points for each system, differentiated by color, show similar distributions across different replicas, underscoring the reproducibility of the results.

However, notable differences are evident. DMXAA exhibits a broader distribution, with some water molecules having residence times exceeding 100 ns, indicating a more diverse interaction landscape. In contrast, G230I, S162A, and mDMXAA systems show more pronounced clustering near the origin, suggesting that water molecules in these systems tend to have very short residence times, indicative of transient water molecules. Similarly, the cGAMP system also shows pronounced clustering near the origin, with fewer water molecules exhibiting high residence times, indicating a slightly more constrained environment and the presence of

transient water molecules. The slightly longer residence times in cGAMP are likely due to its hydrophilic nature.

Subsequently, the appearance of water molecules within the binding site was characterized by an indicator function, $h_i(t)$, which assigns a value of 0 if a specific water molecule is absent from the binding site at time *t* and a value of 1 if it is present. To assess the temporal dynamics and relaxation times of these water molecules, we computed the normalized autocorrelation function of the indicator function. This function, $C(t)$, is defined as follows:

$$C(t) = \langle \frac{h_{i(t_0)} \cdot h_{i(t_0+t)}}{h_{i(t_0)} \cdot h_{i(t_0)}} \rangle = \frac{\sum_{t_0} \sum_{i=1}^{N} h_{i(t_0)} h_{i(t_0+t)}}{\sum_{t_0} \sum_{i=1}^{N} h_{i(t_0)} h_{i(t_0)}}$$

In this equation, $h_i(t)$ represents the indicator function for the *i*-th water molecule at time *t*, where $h_i(t) = 0$ if the water molecule is not present in the binding site and $h_i(t) = 1$ if it is present. The angle brackets $\langle \cdot \rangle$ denote averaging over both the different water molecules *i* and the initial time points t_0. The normalizing factors related to the number of water molecules *N* and the number of time origins t_0 have been omitted from the equation, as they cancel out in the ratio of the numerator to the denominator. *N* represents the total number of water molecules observed in the binding site during the simulation period.

Table 5 *Relaxation time of water molecules in the binding site in five systems.*

System	Relaxation Time (ns)	Standard Deviation (ns)
cGAMP	0.165	0.026
DMXAA	0.127	0.035
G230I	0.110	0.031
S162A	0.120	0.035
mDMXAA	0.097	0.016

The autocorrelation functions of water molecules within the binding site across the five systems, as shown in Figure S15, exhibit a rapid decay, indicating a swift loss of correlation.

This behavior is consistent across multiple replicas for all systems, with the autocorrelation functions generally decaying to near zero within 2 to 4 nanoseconds. The relaxation times derived from these autocorrelation functions, presented in Table 5, further indicate that diffusion of water molecules from the binding site is a relatively fast process, highlighting the transient nature of these molecules in the binding site. Understanding this transient behavior is essential for elucidating the role of water dynamics in the ligand-binding process, as these fleeting water molecules facilitate the structural adjustments necessary for the ligand to align within the protein binding site properly.

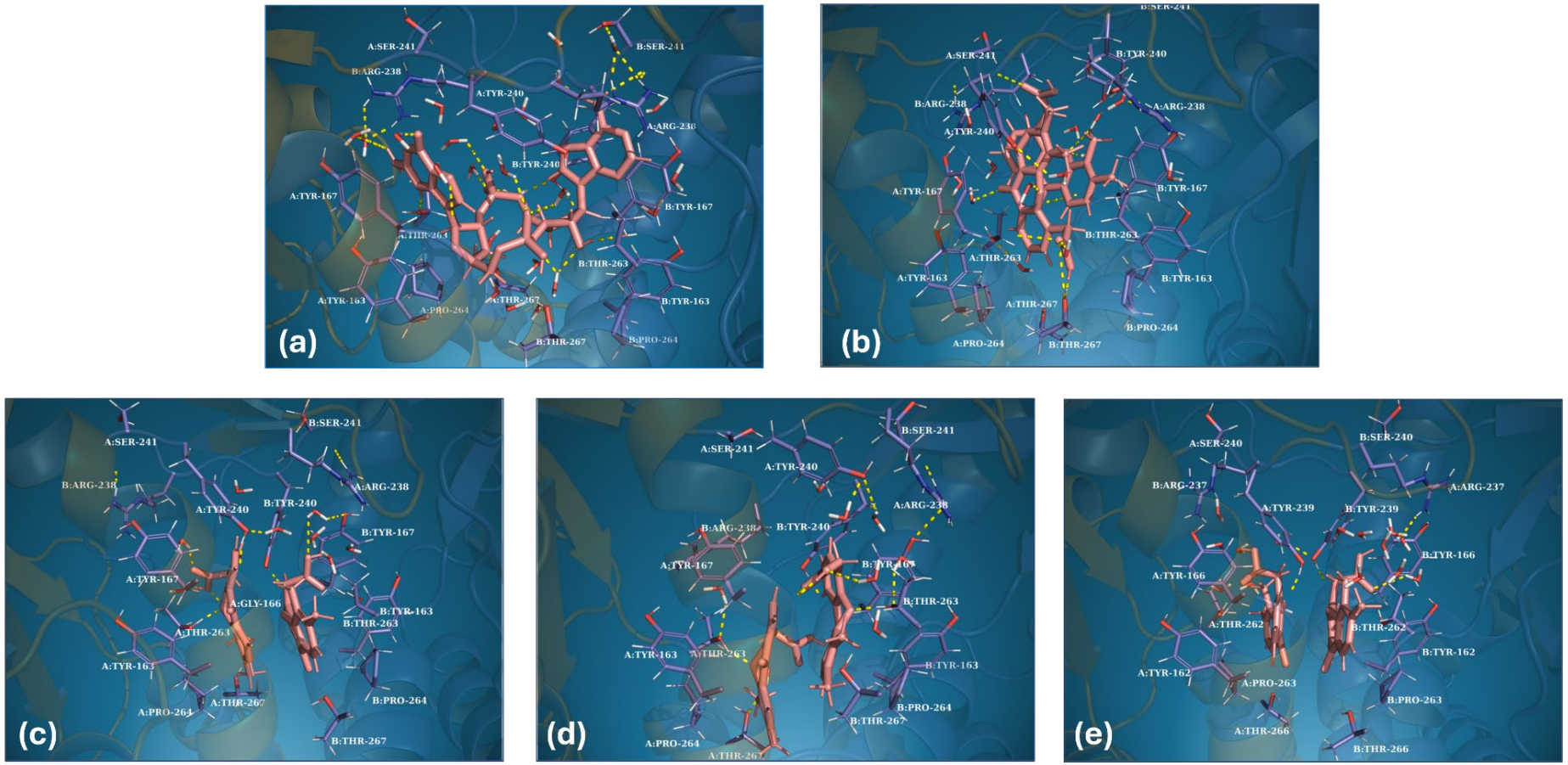

Figure 10. Snapshots of the binding site regions at 750 ns from MD simulations of five different systems: (a) cGAMP with wild-type hSTING, (b) DMXAA with wild-type hSTING, (c) DMXAA with the G230I mutant of hSTING, (d) DMXAA with the S162A/Q266I double mutant of hSTING, and (e) DMXAA with wild-type mSTING. In these representations, the ligand is shown in pink, hydrogen bonds are indicated by yellow dotted lines, and key residues in the binding site are highlighted in purple. Chain A is depicted in orange, while Chain B is shown in blue.

Structural features contributing to the stability of the ligand

Building on the previous sections, we now focus on the residues that significantly impact binding free energies, either favorably or unfavorably, by examining the local geometries of the binding site. This analysis enables us to understand how these residues and water molecules influence the ligand-binding environment.

According to Figure 10(a), TYR-167 forms a π-π interaction with cGAMP, contributing to the stability of the complex. ARG-238 forms hydrogen bonds with water molecules, which in turn create hydrogen bonds with the ligand, acting as bridging waters. Additionally, ARG-238 residues on both chains interact directly or indirectly (via bridging waters) with TYR-167 and SER-241 residues on the opposite chains, facilitating the rigid closure of the lids. TYR-240 residues are also involved in interactions with bridge waters, further stabilizing the structure. The phosphodiester linkages of cGAMP form hydrogen bonds with water molecules, which also serve as bridge waters. Collectively, the active site exhibits a distinct hydrogen bond arrangement, forming a stable network that not only stabilizes the active site but also enhances the interactions between the protein and cGAMP.

On the other hand, based on Figure 10(b), at least one of the TYR-167 residues no longer exhibits a π-π interaction with the ligand in the DMXAA system. ARG-238 residues have lost their hydrogen bonds to the opposite chains, resulting in a slightly opened conformation of the lids. Additionally, water molecules near THR-267 disrupt the π-π interaction between dimers by forming multiple hydrogen bonds with DMXAA. These changes create an unfavorable environment for DMXAA due to its reliance on hydrophobic interactions.

For G230I, S162A, and mDMXAA, TYR-167 (mDMXAA: TYR-166) no longer contributes to the π-π interaction with the ligand. The π-π interactions between DMXAA

molecules likely contribute to their own stability of the dimer. The hydrogen bond network involving water is localized to the lid area. Additionally, ARG-238 (mDMXAA: ARG-237) residues on both chains form hydrogen bonds with residues on the opposite chains, resulting in a rigid closure of the lids. However, these arginine residues exhibit slightly more steric hindrance than the same residues in the cGAMP system. This hindrance could, in fact, explain the slightly unfavorable energy contributions from ARG-238 observed in the G230I, S162A, and mDMXAA systems mentioned earlier in Figure 7.

For the G230I mutation in the hSTING system, no water molecules were observed near the vertex of the binding site, creating an environment favorable for hydrophobic interactions between DMXAA and the protein residues. The previous study by Gao et al.[12] hypothesized that mutating glycine to isoleucine at position 230 in both chains of hSTING might stabilize DMXAA binding by allowing the isoleucine side chains to become buried in a hydrophobic pocket. Our study of the G230I mutation supports their hypothesis, demonstrating that the isoleucine side chains rotate inward, bringing the sides of the β-sheet lid inward. These rotations induce significant local structural changes around the mutation site, displacing water molecules and enhancing hydrophobic interactions. This structural alteration likely contributes to the observed lower binding energy.

The S162A/Q266I mutations in hSTING remove polar hydroxyl and amide groups from these positions, which alters local hydrogen bonding networks and hydrophobic interactions. Consequently, our simulations observed no water molecules near these mutation points. This absence of water molecules increases hydrophobic interactions with DMXAA in the binding site, stabilizing ligand binding and lowering binding free energy values. Additionally, according to Figures 8 and 10(d), the dimer of DMXAA in the S162A/Q266I variant was located much closer

to the vertex of the binding site than in the G230I variant, confirming that the S162A and Q266I mutations created a stronger hydrophobic environment, which were favorable to DMXAA.

On the other hand, in the mSTING system (mDMXAA), a few water molecules interacted with SER-161 and GLN-265. This observation correlates with the very localized high-density point in the spatial distribution of DMXAA shown in Figure 8. We believe that the presence of few water molecules in the vertex of the binding site slightly increases the dielectric constant of the area, potentially reducing the electrostatic interactions between charged groups in the binding site and DMXAA, thus facilitating better hydrophobic interactions between protein residues and DMXAA.

Conclusions

In this study, we investigated the molecular interactions and conformational dynamics of human STING (hSTING) in the presence of the agonist DMXAA using two computational approaches: molecular dynamics simulations and binding free energy calculations. The aim was to elucidate the key structural and dynamic differences between wild-type and mutant forms of hSTING bound to DMXAA, as well as to compare these findings with wild-type hSTING bound to cGAMP and mouse STING (mSTING) bound to DMXAA. The molecular dynamics (MD) simulations employed in this study provided a comprehensive view of the dynamic behavior of STING-ligand complexes. The explicit solvent simulations offered a realistic representation of the biological environment, capturing critical interactions between solvent molecules and the protein-ligand complexes.

Our findings reiterate the importance of specific amino acid residues in modulating the binding affinity and conformational stability of the STING-ligand complexes. As previously

reported [12,13,77], we further confirmed that mutations of G230I together with S162A/Q266I in hSTING significantly improve the binding ability toward DMXAA, closely mimicking the behavior of mSTING with DMXAA. Also, based on our findings, these mutations are key to change to a closed conformation and provide a steric block, such as a unique water molecule network located between the ligand and the β-sheet lids, that stabilizes DMXAA within the binding pocket.

Our findings also suggest that the removal of water molecules near the vertex side of the binding pocket, combined with increasing the hydrophobic character of the binding site, are two critical factors that enhance the binding affinity of DMXAA for hSTING. Our study suggests that G230I and S162A/Q226I mutations in hSTING minimize water-mediated perturbations within the binding site. An earlier study[13] suggested that the instability of DMXAA binding in *wt*-hSTING was due to reduced interactions between DMXAA and polar residues. Our results further reveal that water molecules played a crucial role in disrupting these interactions within the binding site, resulting in weaker hydrophobic interactions between DMXAA and *wt*-hSTING.

Previous studies[72,73,78] have demonstrated that water molecules significantly influence the dynamics of ligand binding. In particular, transient bridge water molecules within the binding site have been shown to modulate the kinetics of this process[78,79]. Our observations further suggest that these transient bridge waters are crucial in mediating the interaction between the STING protein and cGAMP. Additionally, the dynamics of these transient water molecules appear to influence the binding of DMXAA in both hSTING and mSTING. Therefore, we can conclude that water molecules within the STING binding site likely play a vital role in

facilitating the environmental changes necessary for the effective association and dissociation of ligands.

Our study also highlighted the importance of considering the local dielectric environment in binding free energy calculations. Protein's internal dielectric constants, influenced by the presence of water molecules and polar groups, necessitated adjustments in the MM/PBSA calculations to accurately reflect the true binding energetics.

In conclusion, this work broadens our understanding of STING–ligand interactions at the molecular level, providing crucial insights for the rational design of hSTING agonists. We identified key environmental factors that optimize ligand binding by addressing the structural and dynamic differences between human and mouse STING. A key factor in increasing DMXAA's affinity for hSTING is the removal of water molecules around the vertex side of the binding pocket, thereby creating a more hydrophobic environment. From a drug development perspective, it is crucial to explore modifications to the DMXAA molecule that either strengthen its interactions with water molecules around the vertex area or promote their displacement with minimal free energy cost. These findings open new avenues for the development of highly efficient and selective immunotherapeutic agents, significantly enhancing the potential for successful cancer treatments. Future research should prioritize structural modifications of drugs and the optimization of water-mediated interactions to fully exploit STING-targeted therapies.

Acknowledgment

This work was 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. This work was also

partially funded by the Visiting Faculty Program (VFP) and the Science Undergraduate Laboratory Internships (SULI) at the US Department of Energy's (DOE) Office of Science, the Exascale Computing Project (ECP), and the Sustainable Research Pathways (SRP) program, a partnership between the Sustainable Horizons Institute and the Lawrence Berkeley National Laboratory. The authors would like to thank the National Energy Research Scientific Computing (NERSC) Center for the allocation of computing hours on Perlmutter Supercomputers.

References

(1) Zhang, X.; Bai, X.; Chen, Z. J. Structures and Mechanisms in the cGAS-STING Innate Immunity Pathway. *Immunity* **2020**, *53* (1), 43–53. https://doi.org/10.1016/j.immuni.2020.05.013.

(2) Prabakaran, T.; Troldborg, A.; Kumpunya, S.; Alee, I.; Marinković, E.; Windross, S. J.; Nandakumar, R.; Narita, R.; Zhang, B.; Carstensen, M.; Vejvisithsakul, P.; Marqvorsen, M. H. S.; Iversen, M. B.; Holm, C. K.; Østergaard, L. J.; Pedersen, F. S.; Pisitkun, T.; Behrendt, R.; Pisitkun, P.; Paludan, S. R. A STING Antagonist Modulating the Interaction with STIM1 Blocks ER-to-Golgi Trafficking and Inhibits Lupus Pathology. *EBioMedicine* **2021**, *66*, 103314. https://doi.org/10.1016/j.ebiom.2021.103314.

(3) Paludan, S. R.; Bowie, A. G. Immune Sensing of DNA. *Immunity* **2013**, *38* (5), 870–880. https://doi.org/10.1016/j.immuni.2013.05.004.

(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) Liu, S.; Yang, B.; Hou, Y.; Cui, K.; Yang, X.; Li, X.; Chen, L.; Liu, S.; Zhang, Z.; Jia, Y.; Xie, Y.; Xue, Y.; Li, X.; Yan, B.; Wu, C.; Deng, W.; Qi, J.; Lu, D.; Gao, G. F.; Wang, P.; Shang, G. The Mechanism of STING Autoinhibition and Activation. *Mol. Cell* **2023**, *83* (9), 1502-1518.e10. https://doi.org/10.1016/j.molcel.2023.03.029.

(6) Su, T.; Zhang, Y.; Valerie, K.; Wang, X.-Y.; Lin, S.; Zhu, G. STING Activation in Cancer Immunotherapy. *Theranostics* **2019**, *9* (25), 7759–7771. https://doi.org/10.7150/thno.37574.

(7) Sun, X.; Zhou, X.; Lei, Y. L.; Moon, J. J. Unlocking the Promise of Systemic STING Agonist for Cancer Immunotherapy. *J. Controlled Release* **2023**, *357*, 417–421. https://doi.org/10.1016/j.jconrel.2023.03.047.

(8) Zhao, K.; Huang, J.; Zhao, Y.; Wang, S.; Xu, J.; Yin, K. Targeting STING in Cancer: Challenges and Emerging Opportunities. *Biochim. Biophys. Acta BBA - Rev. Cancer* **2023**, *1878* (6), 188983. https://doi.org/10.1016/j.bbcan.2023.188983.

(9) Li, L. Stimulating STING for Cancer Therapy: Taking the Extracellular Route. *Cell Chem. Biol.* **2024**, *31* (5), 851–861. https://doi.org/10.1016/j.chembiol.2024.04.004.

(10) Prantner, D.; Perkins, D. J.; Lai, W.; Williams, M. S.; Sharma, S.; Fitzgerald, K. A.; Vogel, S. N. 5,6-Dimethylxanthenone-4-Acetic Acid (DMXAA) Activates Stimulator of Interferon Gene (STING)-Dependent Innate Immune Pathways and Is Regulated by Mitochondrial Membrane Potential. *J. Biol. Chem.* **2012**, *287* (47), 39776–39788. https://doi.org/10.1074/jbc.M112.382986.

(11) Lara, P. N.; Douillard, J.-Y.; Nakagawa, K.; von Pawel, J.; McKeage, M. J.; Albert, I.; Losonczy, G.; Reck, M.; Heo, D.-S.; Fan, X.; Fandi, A.; Scagliotti, G. Randomized Phase III Placebo-Controlled Trial of Carboplatin and Paclitaxel With or Without the Vascular Disrupting Agent Vadimezan (ASA404) in Advanced Non–Small-Cell Lung Cancer. *J. Clin. Oncol.* **2011**, *29* (22), 2965–2971. https://doi.org/10.1200/JCO.2011.35.0660.

(12) Gao, P.; Zillinger, T.; Wang, W.; Ascano, M.; Dai, P.; Hartmann, G.; Tuschl, T.; Deng, L.; Barchet, W.; Patel, D. J. Binding-Pocket and Lid-Region Substitutions Render Human STING Sensitive to the Species-Specific Drug DMXAA. *Cell Rep.* **2014**, *8* (6), 1668–1676. https://doi.org/10.1016/j.celrep.2014.08.010.

(13) Shih, A. Y.; Damm-Ganamet, K. L.; Mirzadegan, T. Dynamic Structural Differences between Human and Mouse STING Lead to Differing Sensitivity to DMXAA. *Biophys. J.* **2018**, *114* (1), 32–39. https://doi.org/10.1016/j.bpj.2017.10.027.

(14) Zhang, Y.; Sun, Z.; Pei, J.; Luo, Q.; Zeng, X.; Li, Q.; Yang, Z.; Quan, J. Identification of α-Mangostin as an Agonist of Human STING. *ChemMedChem* **2018**, *13* (19), 2057–2064. https://doi.org/10.1002/cmdc.201800481.

(15) Temizoz, B.; Shibahara, T.; Hioki, K.; Hayashi, T.; Kobiyama, K.; Lee, M. S. J.; Surucu, N.; Sag, E.; Kumanogoh, A.; Yamamoto, M.; Gursel, M.; Ozen, S.; Kuroda, E.; Coban, C.; Ishii, K. J. 5,6-Dimethylxanthenone-4-Acetic Acid (DMXAA), a Partial STING Agonist, Competes for Human STING Activation. *Front. Immunol.* **2024**, *15*, 1353336. https://doi.org/10.3389/fimmu.2024.1353336.
(16) Suzuki, Y.; Sato, T.; Sugimori, M.; Kanemaru, Y.; Onodera, S.; Tsuchiya, H.; Nakamori, Y.; Tsuyuki, S.; Ikeda, A.; Ikeda, R.; Goda, Y.; Kaneko, H.; Irie, K.; Sue, S.; Maeda, S. Activation of STING in Pancreatic Cancer-Associated Fibroblasts Exerts an Antitumor Effect by Enhancing Tumor Immunity. *Sci. Rep.* **2024**, *14* (1), 17071. https://doi.org/10.1038/s41598-024-68061-y.
(17) Tan, J.; Wang, M.; Ding, B.; Ma, P.; Lin, J. Advanced Nanomaterials Targeting Activation of STING for Enhanced Cancer Immunotherapy. *Coord. Chem. Rev.* **2023**, *493*, 215316. https://doi.org/10.1016/j.ccr.2023.215316.
(18) Chang, J.; Hou, S.; Yan, X.; Li, W.; Xiao, J. Discovery of Novel STING Inhibitors Based on the Structure of the Mouse STING Agonist DMXAA. *Molecules* **2023**, *28* (7), 2906. https://doi.org/10.3390/molecules28072906.
(19) Chen, X.; Xu, Z.; Li, T.; Thakur, A.; Wen, Y.; Zhang, K.; Liu, Y.; Liang, Q.; Liu, W.; Qin, J.-J.; Yan, Y. Nanomaterial-Encapsulated STING Agonists for Immune Modulation in Cancer Therapy. *Biomark. Res.* **2024**, *12* (1), 2. https://doi.org/10.1186/s40364-023-00551-z.
(20) Wang, Y.; Liu, Y.; Zhang, J.; Peng, Q.; Wang, X.; Xiao, X.; Shi, K. Nanomaterial-Mediated Modulation of the cGAS-STING Signaling Pathway for Enhanced Cancer Immunotherapy. *Acta Biomater.* **2024**, *176*, 51–76. https://doi.org/10.1016/j.actbio.2024.01.008.
(21) Song, X.; Xi, Y.; Dai, M.; Li, T.; Du, S.; Zhu, Y.; Li, M.; Li, Y.; Liu, S.; Ding, X.; Yao, X.; Lai, Y.; Liu, X. STING Guides the STX17-SNAP29-VAMP8 Complex Assembly to Control Autophagy. *Cell Insight* **2024**, *3* (2), 100147. https://doi.org/10.1016/j.cellin.2024.100147.
(22) Lanng, K. R. B.; Lauridsen, E. L.; Jakobsen, M. R. The Balance of STING Signaling Orchestrates Immunity in Cancer. *Nat. Immunol.* **2024**, *25* (7), 1144–1157. https://doi.org/10.1038/s41590-024-01872-3.
(23) Hines, J. B.; Kacew, A. J.; Sweis, R. F. The Development of STING Agonists and Emerging Results as a Cancer Immunotherapy. *Curr. Oncol. Rep.* **2023**, *25* (3), 189–199. https://doi.org/10.1007/s11912-023-01361-0.
(24) Gu, J.; Liu, X.; Ji, Z.; Shen, M.; Zhu, M.; Ren, Y.; Guo, L.; Yang, K.; Liu, T.; Yi, X. Tumor Vascular Destruction and cGAS-STING Activation Induced by Single Drug-Loaded Nano-Micelles for Multiple Synergistic Therapies of Cancer. *Small* **2023**, *19* (46), 2303517. https://doi.org/10.1002/smll.202303517.
(25) Chen, L.; Zhao, S.; Zhu, Y.; Liu, Y.; Li, H.; Zhao, Q. Molecular Dynamics Simulations Reveal the Modulated Mechanism of STING Conformation. *Interdiscip. Sci. Comput. Life Sci.* **2021**. https://doi.org/10.1007/s12539-021-00446-3.
(26) Aliakbar Tehrani, Z.; Rulíšek, L.; Černý, J. Molecular Dynamics Simulations Provide Structural Insight into Binding of Cyclic Dinucleotides to Human STING Protein. *J. Biomol. Struct. Dyn.* **2021**, 1–15. https://doi.org/10.1080/07391102.2021.1942213.
(27) Vavřina, Z.; Gutten, O.; Smola, M.; Zavřel, M.; Aliakbar Tehrani, Z.; Charvát, V.; Kožíšek, M.; Boura, E.; Birkuš, G.; Rulíšek, L. Protein–Ligand Interactions in the STING Binding Site Probed by Rationally Designed Single-Point Mutations: Experiment and Theory. *Biochemistry* **2021**, *60* (8), 607–620. https://doi.org/10.1021/acs.biochem.0c00949.
(28) Payne, R. T.; Crivelli, S.; Watanabe, M. All-Atom Simulations Uncover Structural and Dynamical Properties of STING Proteins in the Membrane System. *J. Chem. Inf. Model.* **2022**, *62* (18), 4486–4499. https://doi.org/10.1021/acs.jcim.2c00595.

(29) Li, Z.; Yue, C.; Xie, S.; Shi, S.; Ye, S. Computational Insights into the Conformational Transition of STING: Mechanistic, Energetic Considerations, and the Influence of Crucial Mutations. *J. Mol. Graph. Model.* **2024**, *129*, 108764. https://doi.org/10.1016/j.jmgm.2024.108764.
(30) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652. https://doi.org/10.1063/1.464913.
(31) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789. https://doi.org/10.1103/PhysRevB.37.785.
(32) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. 6-31G* Basis Set for Atoms K through Zn. *J. Chem. Phys.* **1998**, *109* (4), 1223–1229. https://doi.org/10.1063/1.476673.
(33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16 Rev. C.01, 2016.
(34) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31* (4), 671–690. https://doi.org/10.1002/jcc.21367.
(35) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell Jr., A. D. Extension of the CHARMM General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33* (31), 2451–2468. https://doi.org/10.1002/jcc.23067.
(36) Huang, J.; MacKerell Jr, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–2145. https://doi.org/10.1002/jcc.23354.
(37) Rosignoli, S.; Paiardini, A. Boosting the Full Potential of PyMOL with Structural Biology Plugins. *Biomolecules* **2022**, *12* (12), 1764. https://doi.org/10.3390/biom12121764.
(38) Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 2.5, 2015.
(39) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, *54* (1), 5.6.1-5.6.37. https://doi.org/10.1002/cpbi.3.
(40) John, B. Comparative Protein Structure Modeling by Iterative Alignment, Model Building and Model Assessment. *Nucleic Acids Res.* **2003**, *31* (14), 3982–3992. https://doi.org/10.1093/nar/gkg460.
(41) Shen, M.; Sali, A. Statistical Potential for Assessment and Prediction of Protein Structures. *Protein Sci.* **2006**, *15* (11), 2507–2524. https://doi.org/10.1110/ps.062416606.
(42) Eramian, D.; Eswar, N.; Shen, M.-Y.; Sali, A. How Well Can the Accuracy of Comparative Protein Structure Models Be Predicted? *Protein Sci.* **2008**, *17* (11), 1881–1893. https://doi.org/10.1110/ps.036061.108.
(43) 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.
(44) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190. https://doi.org/10.1063/1.328693.

(45) 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.
(46) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.
(47) 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.
(48) Knapp, B.; Ospina, L.; Deane, C. M. Avoiding False Positive Conclusions in Molecular Simulation: The Importance of Replicas. *J. Chem. Theory Comput.* **2018**, *14* (12), 6127–6138. https://doi.org/10.1021/acs.jctc.8b00391.
(49) Wan, S.; Sinclair, R. C.; Coveney, P. V. Uncertainty Quantification in Classical Molecular Dynamics. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **2021**, *379* (2197), rsta.2020.0082, 20200082. https://doi.org/10.1098/rsta.2020.0082.
(50) Chong, S.-H.; Im, H.; Ham, S. Explicit Characterization of the Free Energy Landscape of pKID–KIX Coupled Folding and Binding. *ACS Cent. Sci.* **2019**, *5* (8), 1342–1351. https://doi.org/10.1021/acscentsci.9b00200.
(51) Hirono, S.; Kollman, P. A. Calculation of the Relative Binding Free Energy of 2'GMP and 2'AMP to Ribonuclease T1 Using Molecular Dynamics/Free Energy Perturbation Approaches. *J. Mol. Biol.* **1990**, *212* (1), 197–209. https://doi.org/10.1016/0022-2836(90)90315-D.
(52) Clark, A. J.; Gindin, T.; Zhang, B.; Wang, L.; Abel, R.; Murret, C. S.; Xu, F.; Bao, A.; Lu, N. J.; Zhou, T.; Kwong, P. D.; Shapiro, L.; Honig, B.; Friesner, R. A. Free Energy Perturbation Calculation of Relative Binding Free Energy between Broadly Neutralizing Antibodies and the Gp120 Glycoprotein of HIV-1. *J. Mol. Biol.* **2017**, *429* (7), 930–947. https://doi.org/10.1016/j.jmb.2016.11.021.
(53) Jiang, W.; Roux, B. Free Energy Perturbation Hamiltonian Replica-Exchange Molecular Dynamics (FEP/H-REMD) for Absolute Ligand Binding Free Energy Calculations. *J. Chem. Theory Comput.* **2010**, *6* (9), 2559–2565. https://doi.org/10.1021/ct1001768.
(54) Fratev, F.; Sirimulla, S. An Improved Free Energy Perturbation FEP+ Sampling Protocol for Flexible Ligand-Binding Domains. *Sci. Rep.* **2019**, *9* (1), 16829. https://doi.org/10.1038/s41598-019-53133-1.
(55) Wu, K.-W.; Chen, P.-C.; Wang, J.; Sun, Y.-C. Computation of Relative Binding Free Energy for an Inhibitor and Its Analogs Binding with Erk Kinase Using Thermodynamic Integration MD Simulation. *J. Comput. Aided Mol. Des.* **2012**, *26* (10), 1159–1169. https://doi.org/10.1007/s10822-012-9606-6.
(56) Lawrenz, M.; Baron, R.; Wang, Y.; McCammon, J. A. Independent-Trajectory Thermodynamic Integration: A Practical Guide to Protein-Drug Binding Free Energy Calculations Using Distributed Computing. *Methods Mol. Biol. Clifton NJ* **2012**, *819*, 469–486. https://doi.org/10.1007/978-1-61779-465-0_27.
(57) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119* (16), 9478–9508. https://doi.org/10.1021/acs.chemrev.9b00055.
(58) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate−DNA Helices. *J. Am. Chem. Soc.* **1998**, *120* (37), 9401–9409. https://doi.org/10.1021/ja981844+.
(59) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33* (12), 889–897. https://doi.org/10.1021/ar000033j.
(60) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* **2015**, *10* (5), 449–461. https://doi.org/10.1517/17460441.2015.1032936.

(61) Valdés-Tresanco, M. S.; Valdés-Tresanco, M. E.; Valiente, P. A.; Moreno, E. gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17* (10), 6281–6291. https://doi.org/10.1021/acs.jctc.1c00645.
(62) Duan, L.; Liu, X.; Zhang, J. Z. H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein–Ligand Binding Free Energy. *J. Am. Chem. Soc.* **2016**, *138* (17), 5722–5728. https://doi.org/10.1021/jacs.6b02682.
(63) Menzer, W. M.; Li, C.; Sun, W.; Xie, B.; Minh, D. D. L. Simple Entropy Terms for End-Point Binding Free Energy Calculations. *J. Chem. Theory Comput.* **2018**, *14* (11), 6035–6049. https://doi.org/10.1021/acs.jctc.8b00418.
(64) Ekberg, V.; Ryde, U. On the Use of Interaction Entropy and Related Methods to Estimate Binding Entropies. *J. Chem. Theory Comput.* **2021**, *17* (8), 5379–5391. https://doi.org/10.1021/acs.jctc.1c00374.
(65) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.
(66) Pliška, V.; Testa, B.; van de Waterbeemd, H. Lipophilicity: The Empirical Tool and the Fundamental Objective. An Introduction. In *Lipophilicity in Drug Action and Toxicology*; Methods and Principles in Medicinal Chemistry; 1996; pp 1–6. https://doi.org/10.1002/9783527614998.ch1.
(67) Delaney, J. S. ESOL: Estimating Aqueous Solubility Directly from Molecular Structure. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (3), 1000–1005. https://doi.org/10.1021/ci034243x.
(68) Daina, A.; Michielin, O.; Zoete, V. SwissADME: A Free Web Tool to Evaluate Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules. *Sci. Rep.* **2017**, *7* (1), 42717. https://doi.org/10.1038/srep42717.
(69) Lange, O. F.; Grubmüller, H. Generalized Correlation for Biomolecular Dynamics. *Proteins Struct. Funct. Bioinforma.* **2006**, *62* (4), 1053–1061. https://doi.org/10.1002/prot.20784.
(70) Tekpinar, M.; Neron, B.; Delarue, M. Extracting Dynamical Correlations and Identifying Key Residues for Allosteric Communication in Proteins by *Correlationplus*. *J. Chem. Inf. Model.* **2021**, *61* (10), 4832–4838. https://doi.org/10.1021/acs.jcim.1c00742.
(71) Mahmoud, A. H.; Masters, M. R.; Yang, Y.; Lill, M. A. Elucidating the Multiple Roles of Hydration for Accurate Protein-Ligand Binding Prediction via Deep Learning. *Commun. Chem.* **2020**, *3* (1), 19. https://doi.org/10.1038/s42004-020-0261-x.
(72) Schiebel, J.; Gaspari, R.; Wulsdorf, T.; Ngo, K.; Sohn, C.; Schrader, T. E.; Cavalli, A.; Ostermann, A.; Heine, A.; Klebe, G. Intriguing Role of Water in Protein-Ligand Binding Studied by Neutron Crystallography on Trypsin Complexes. *Nat. Commun.* **2018**, *9* (1), 3559. https://doi.org/10.1038/s41467-018-05769-2.
(73) Maurer, M.; Oostenbrink, C. Water in Protein Hydration and Ligand Recognition. *J. Mol. Recognit. JMR* **2019**, *32* (12), e2810. https://doi.org/10.1002/jmr.2810.
(74) Bitencourt-Ferreira, G.; Veit-Acosta, M.; de Azevedo, W. F. Hydrogen Bonds in Protein-Ligand Complexes. In *Docking Screens for Drug Discovery*; de Azevedo Jr., W. F., Ed.; Springer New York: New York, NY, 2019; pp 93–107. https://doi.org/10.1007/978-1-4939-9752-7_7.
(75) Wong, S.; Amaro, R. E.; McCammon, J. A. MM-PBSA Captures Key Role of Intercalating Water Molecules at a Protein−Protein Interface. *J. Chem. Theory Comput.* **2009**, *5* (2), 422–429. https://doi.org/10.1021/ct8003707.
(76) Zhang, J.; Zhang, H.; Wu, T.; Wang, Q.; van der Spoel, D. Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents. *J. Chem. Theory Comput.* **2017**, *13* (3), 1034–1043. https://doi.org/10.1021/acs.jctc.7b00169.
(77) 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.
(78) Huang, Y. M.; Raymundo, M. A. V.; Chen, W.; Chang, C. A. Mechanism of the Association Pathways for a Pair of Fast and Slow Binding Ligands of HIV-1 Protease. *Biochemistry* **2017**, *56* (9), 1311–1323. https://doi.org/10.1021/acs.biochem.6b01112.
(79) Chen, W.; He, H.; Wang, J.; Wang, J.; Chang, C. A. Uncovering Water Effects in Protein–Ligand Recognition: Importance in the Second Hydration Shell and Binding Kinetics. *Phys. Chem. Chem. Phys.* **2023**, *25* (3), 2098–2109. https://doi.org/10.1039/D2CP04584B.