

Ligand Binding

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 10172–10178

International Edition: doi.org/10.1002/anie.202016805

German Edition: doi.org/10.1002/ange.202016805

Ligand Strain and Its Conformational Complexity Is a Major Factor in the Binding of Cyclic Dinucleotides to STING Protein

Miroslav Smola⁺, Ondrej Gutten⁺, Milan Dejmek, Milan Kožíšek, Thomas Evangelidis, Zahra Aliakbar Tehrani, Barbora Novotná, Radim Nencka, Gabriel Birkuš, Lubomír Rulíšek,* and Evzen Boura*

Abstract: *STING* (stimulator of interferon genes) is a key regulator of innate immunity that has recently been recognized as a promising drug target. *STING* is activated by cyclic dinucleotides (CDNs) which eventually leads to expression of type I interferons and other cytokines. Factors underlying the affinity of various CDN analogues are poorly understood. Herein, we correlate structural biology, isothermal calorimetry (ITC) and computational modeling to elucidate factors contributing to binding of six CDNs—three pairs of natural (ribo) and fluorinated (2'-fluororibo) 3',3'-CDNs. X-ray structural analyses of six {*STING*:CDN} complexes did not offer any explanation for the different affinities of the studied ligands. ITC showed entropy/enthalpy compensation up to 25 kcal mol⁻¹ for this set of similar ligands. The higher affinities of fluorinated analogues are explained with help of computational methods by smaller loss of entropy upon binding and by smaller strain (free) energy.

Introduction

In a quest for high-affinity ligands, protein–ligand interactions are often thoroughly probed^[1–3] experimentally^[4–6] and computationally^[7,8] to evaluate physico-chemical factors determining their binding properties. This may provide further guidance for improving ligand affinity and ultimately

its biological activity in order to increase its potential therapeutic advantage.^[9] Standard approaches to quantify protein–ligand binding, including popular docking methods, are based on optimization of interactions in the protein cavity, filling “voids” in the binding site by attaching various functional groups to ligands, counting hydrogen bonds etc.^[1,10,11] Numerous studies,^[12] including our own efforts,^[13] pointed out that the ligand strain/deformation energies (the energy price for adapting the bound conformation),^[14,15] ligand/protein/solvent entropy changes^[16–18] and water thermodynamics^[19] might be of the same or even larger magnitude than the enthalpic gains provided by favourable intermolecular interactions. Herein, we set out to investigate these subtleties in ligand binding on the example of one of the prime targets of current medicinal chemistry: the *STING* (stimulator of interferon genes) protein.^[20]

STING is a member of the cGAS-*STING* pathway and a receptor of cyclic dinucleotides (CDNs) that act as second messengers.^[21–23] It transduces signal from the cytosolic dsDNA sensor cGAS (cyclic GMP-AMP synthase) to the transcription factor IRF3 (interferon regulatory factor 3) through the kinase TBK1 (TANK-binding kinase 1) and to the transcription factor NF-κB (nuclear factor kappa-light-chain-enhancer of activated B-cells) through the kinase IKK (I kappa B kinase). Stimulation of *STING* with CDNs eventually leads to induction of expression of type I interferons (INF) and other cytokines as TNFα or IL-6,^[21–23] which promote innate immune defenses against invading pathogens.^[24] Furthermore, cGAS-*STING* signaling pathway also plays a critical role in the inductions of spontaneous antitumor immunity and in a growing number of different autoimmune diseases,^[25] such as systemic lupus erythematosus (SLE), Aicardi-Goutières syndrome (AGS),^[25,26] or *STING*-associated vasculopathy with onset in infancy (SAVI).^[27]

Apart from the vertebrate CDN 2',3'-cGAMP, which is synthesized by cGAS, *STING* can also be activated by CDNs of bacterial origin, such as cyclic di-AMP (3',3'-c-di-AMP), cyclic di-GMP (3',3'-c-di-GMP) or 3',3'-cGAMP.^[25,28] These CDNs function as bacterial second messengers controlling biological processes such as biofilm development, motility, cell cycle and pathogenicity in bacteria.^[29] Furthermore, a plethora of non-natural CDNs with improved drug-like properties were synthesized; more or less complete list of the compounds can be found in the recent review.^[20]

Structurally, *STING* is a dimeric protein localized in the membrane of the endoplasmic reticulum (ER). The mono-

[*] Dr. M. Smola,^[†] Dr. O. Gutten,^[†] Dr. M. Dejmek, Dr. M. Kožíšek, Dr. T. Evangelidis, Dr. Z. A. Tehrani, B. Novotná, Dr. R. Nencka, Dr. G. Birkuš, Assoc. Prof. Dr. L. Rulíšek, Dr. E. Boura
 Gilead Sciences Research Centre at IOCB, Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences
 Flemingovo náměstí 2, 16610 Prague (Czech Republic)
 E-mail: rulisek@uochb.cas.cz
 boura@uochb.cas.cz

[†] These authors contributed equally to this work.

 Supporting information (protein purification details, data collection and refinement statistics, graphic representations of solved structures and computational details; Tables S1–S3; Figures S1–S7; all equilibrium geometries (QM regions of QM/MM calculations); and one full QM/MM model, including point charges on all atoms) and the ORCID identification number(s) for the author(s) of this article can be found under:

<https://doi.org/10.1002/anie.202016805>.

 © 2021 The Authors. *Angewandte Chemie International Edition* published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

mers, consisting of 379 amino acid residues, are composed of an N-terminal transmembrane domain with four transmembrane helices (1–154), a dimerization CDNs binding domain (155–342) and a C-terminal tail (343–379).^[30] Importantly, upon activation, STING changes its conformation from open to closed and a so called lid is formed above the cyclic dinucleotide binding site (Figure 1).^[22] This conformational change is transduced over the ER membrane and only in the closed conformation can STING recruit and activate protein kinases TBK1 and IKK.

In this study, we present structural, biochemical and computational characterization of human wild-type STING in complex with natural CDNs: 3',3'-cGAMP, 3',3'-c-di-GMP and 3',3'-c-di-AMP and their fluorinated analogues at both 2' positions of the ribose ring, where the fluorine atom replaces the hydroxyl group (Figure 2). The fluorinated analogues were shown to be at least one order of magnitude more potent in inducing interferon type I expression than the corresponding non-fluorinated compounds.^[33,34]

First, we show that fluorinated ligands exhibit a significantly higher stabilization effect on STING in the thermal stability assay than the non-fluorinated ligands. This observation has been independently confirmed by measuring K_D values pertinent to the STING-CDN interaction, employing the isothermal titration calorimetry (ITC). To explain the effect of fluorine substitution, we compared crystal structures of STING with all six above-mentioned ligands (i.e., we solved five CDN-STING crystal structures, whereas the sixth—STING in complex with the 3',3'-c-di-GMP—was solved by us previously, PDB entry 6S86, ref. [38]). Surprisingly, all the ligands were positioned in the binding pocket in (almost) exactly the same way; also, the protein seemed to adopt the identical conformation. As such, it has not provided any plausible explanation for the profound difference between potency of fluorinated and non-fluorinated ligands. Quite surprising data, though not entirely unprecedented in the literature,^[35] were obtained by ITC. Entropy/enthalpy compensation observed amounted up to 25 kcal mol⁻¹. Computational analyses (QM/MM modeling, DFT-D3/COSMORS calculations of ligand strain energies and ligand entropy calculations) provided a plausible rationale for the observed thermodynamics of binding. In summary, we explain in detail the complexity of the binding process and interplay of various

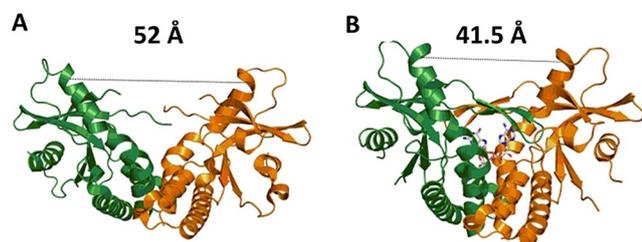


Figure 1. C-terminal CDN binding domain of human wild-type STING in open and closed conformations. A) Unliganded STING in an open conformation (PDB entry 4EMU).^[31] B) STING in complex with its natural agonist 2',3'-cGAMP in a closed conformation (PDB entry 4K-SY).^[32] Displayed distances are measured between C α atoms of Gln184 of each monomer.

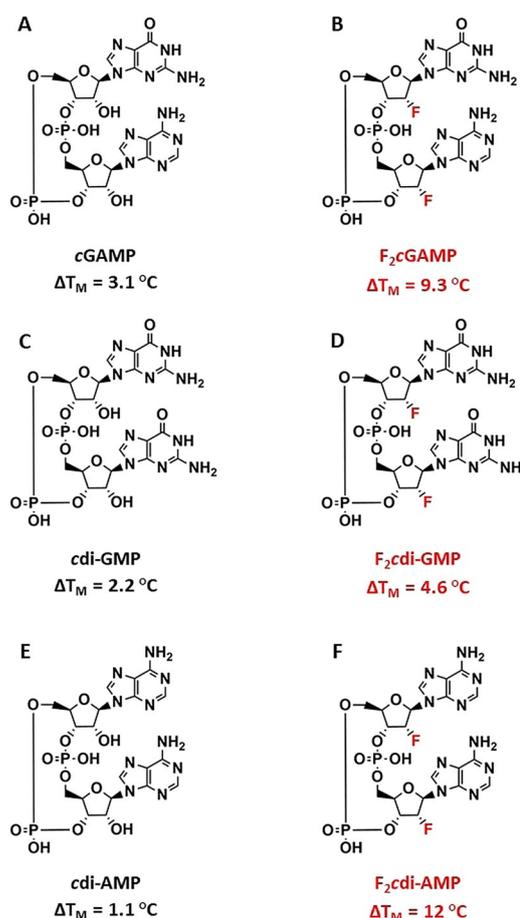


Figure 2. Cyclic dinucleotides studied in this work: A) cGAMP, B) F₂cGAMP, C) cdi-GMP, D) F₂cdi-GMP, E) cdi-AMP, F) F₂cdi-AMP. ΔT_M values represent melting temperature differences of protein:ligand complexes in comparison with the unliganded STING.

physico-chemical contributions to the overall binding free energy (ΔG_{bind}). At the same time, we show that ΔG_{bind} is amenable to quantitative analysis by combining structural biology, thermodynamic measurements and advanced computer modeling. We also argue that use of quantum mechanics is imperative in quantitative characterization of ligand-binding in biomolecules, as has been demonstrated before.^[7,8,36] This is especially true in cases such as CDNs—charged macrocycles remain a formidable challenge for force-fields, accuracy of which underlie success of methods such as MM/PBSA or MM/FEP.

Results and Discussion

Fluorine substitution at the ribose 2'-position improves thermal stability of CDN:STING complexes. Differential scanning fluorimetry (DSF) was employed to investigate the thermal stability of STING in complex with 3',3'-cGAMP (cGAMP), 3',3'-c-di-GMP (cdi-GMP), 3',3'-c-di-AMP (cdi-AMP) and their 2' fluorinated analogues (further denoted as F₂cXMP, X = di-A, di-G, GA). It has been shown that the difference between melting temperatures (ΔT_M values) of the

unliganded STING and STING in complex with ligand strongly correlate with binding affinities of ligands towards the STING protein and can be used as reliable estimates of binding free energy.^[37] All three difluorinated analogues— F_2c di-GMP, F_2c di-AMP, and F_2c GAMP—thermally stabilize the STING:ligand complex to a greater degree than their unsubstituted counterparts (Figure 2). In comparison with the unliganded STING whose melting temperature (T_M) is 44.0 ± 0.3 °C, the ΔT_M is 3.1 °C for STING: c GAMP complex, 2.2 °C for the STING: c di-GMP and 1.1 °C for the STING: c di-AMP complexes. However, all three fluorinated analogues stabilize STING dramatically more: ΔT_M is 9.3 °C for F_2c GAMP, 4.6 °C for F_2c di-GMP and 12 °C for F_2c di-AMP.

Crystal structures of human STING in complex with 3',3'- c GAMP, 3',3'- c di-GMP, 3',3'- c di-AMP and their difluorinated analogues. Being intrigued by the observed increase of the thermal stability (ΔT_M) of ligand:protein complexes of difluorinated CDNs, we solved the crystal structures of all but one of the six studied STING:CDN complexes, (one $-3',3'$ - c -di-GMP-STING complexed that was solved and published previously, ref. [38]). In analogy with previous crystallographic studies on the STING protein, we were able to obtain crystals for the studied CDNs complexed with the STING-CDN binding domain (residues 140–343). The crystals belonged to the tetragonal space groups $P4_12_12$ or $P4_32_12$ and diffracted to 2.2 – 2.9 Å resolution. The structures were solved by molecular replacement using STING in complex with $2',3'$ - c GAMP (PDB entry 4KSY, ref. [32]) as a search model. The structures were refined by standard procedures to good R factors and geometry, as summarized in Table S1†. We were able to trace the whole protein chain except for approximately 13 amino acids at N-terminus, a disordered loop between Pro317 and Phe323 and a disordered loop between Gln184 and Ala193. Overall, our structures are in a good agreement with the previously published structures of the STING with CDNs.^[32,37,39–41] They reveal the STING-CDN binding domain in the closed conformation where the lid is formed above the ligand binding site and the distance between the two main alpha helices is shorter by approximately 10 Å in comparison to the unliganded STING (Figures 3 and S1†, S2† and S3†). To reveal conformational changes of STING when bound to common $3',3'$ CDN and their fluorinated analogues we aligned the corresponding crystal structures using also the structure of wild-type STING in complex with c di-GMP (PDB entry 6S86).^[38] Surprisingly, we did not observe any significant conformational changes among the six structures, neither for ligands nor for the protein (Figure 3). Correspondingly, the RMSD values between each pair of structures are very low, ranging between 0.32 to 0.64 Å.

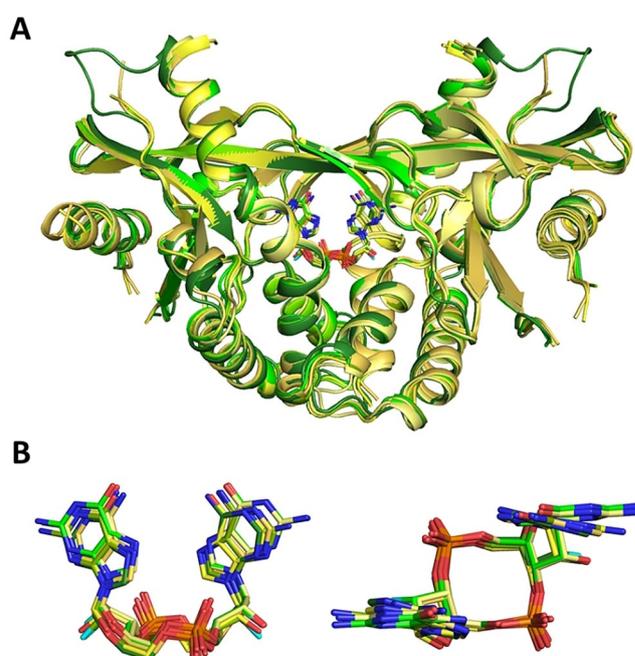


Figure 3. Structural alignment of all structures of interest: A) STING in complex with c di-GMP (PDB entry 6S86), F_2c di-GMP (PDB entry 6YDB), c GAMP (PDB entry 6YDZ), F_2c GAMP (PDB entry 6YEA), c di-AMP (PDB entry 6Z15), or F_2c di-AMP (PDB entry 6Z0Z). Complexes with natural ligands are represented in yellow, while their fluorinated analogues are in green. B) Structural alignment of studied ligands in bound conformations.

Isothermal titration calorimetry. To understand differences in binding of natural $3',3'$ -CDNs and their difluorinated analogues, we performed thermodynamic characterization of STING:ligand complex formation using isothermal titration calorimetry (ITC). The experiments were carried out in three buffers (Tris, HEPES, PIPES) to account for a proton exchange between buffer, protein and ligand.^[42] The three buffers used cover the whole spectrum of deprotonation free energies.^[43,44] For c di-AMP or F_2c di-AMP we were not able to observe any significant change of enthalpy which has been already reported before.^[35] Final results, to a great extent independent of (de)protonation events upon ligand binding, are summarized in Table 1. Representative measurements are depicted in Figure 4. In full accordance with the DSF values presented above, the observed K_D values are distinctly lower for fluorinated CDNs, up to two orders of magnitude for c GAMP/ F_2c GAMP. Of a much greater surprise are individual ΔH_{bind} and $-T\Delta S_{\text{bind}}$ terms. For very similar ligands which bind to STING in an almost identical manner (c.f. X-ray structures discussed above), the observed entropy/enthalpy

Table 1: Thermodynamic characterization of STING:CDN complex formation by isothermal titration calorimetry.

Ligand	c GAMP	F_2c GAMP	c di-GMP	F_2c di-GMP
K_s [M^{-1}]	$(5.32 \pm 0.94) \times 10^5$	$(3.99 \pm 1.75) \times 10^7$	$(4.06 \pm 0.90) \times 10^5$	$(2.74 \pm 0.37) \times 10^6$
K_D [nM]	1900 ± 400	25 ± 14	2500 ± 600	370 ± 50
ΔG [kcal mol $^{-1}$]	-7.8 ± 0.1	-10.4 ± 0.3	-7.7 ± 0.1	-8.8 ± 0.1
ΔH_{bind} [kcal mol $^{-1}$]	-2.8 ± 0.1	6.0 ± 0.1	-20.4 ± 0.4	-15.7 ± 0.1
$-T\Delta S_{\text{bind}}$ [kcal mol $^{-1}$]	-5.0 ± 0.2	-16.4 ± 0.4	12.7 ± 0.6	6.7 ± 0.2
Δn_{H^+}	-0.18 ± 0.01	-0.19 ± 0.01	≈ 0	≈ 0

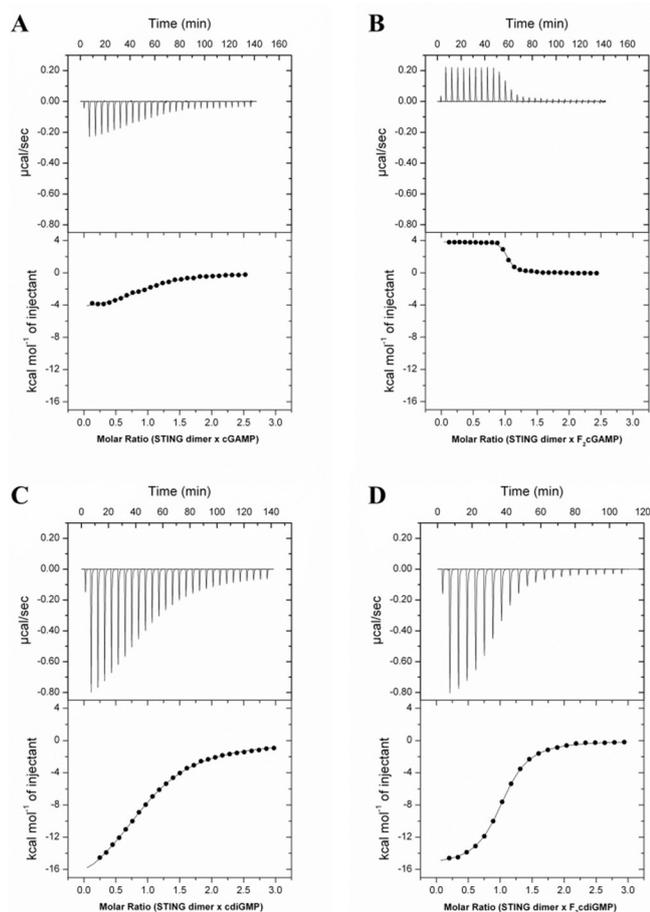


Figure 4. Thermodynamic characterization of STING:ligand complex forming process by ITC. Raw and integrated data of the titration of STING in Tris buffer with A) *c*GAMP, B) F_2c GAMP, C) *cdi*-GMP, D) F_2cdi -GMP. The *c*-values for titrations: STING protein with ligand *c*GAMP ($c=6$), F_2c GAMP ($c=480$), *cdi*-GMP ($c=5$), and F_2cdi -GMP ($c=33$).

compensations amount up to 25 kcal mol^{-1} . Considering natural vs. difluoro CDN pairs, binding of the latter is entropically $6\text{--}11 \text{ kcal mol}^{-1}$ more favourable. This more than compensates for enthalpy contributions, which tend to favour natural ($2'$ -hydroxy) CDNs by $5\text{--}9 \text{ kcal mol}^{-1}$. For F_2c GAMP this even leads to endothermic process, $\Delta H_{\text{bind}} = 6.0 \text{ kcal mol}^{-1}$. Interestingly, in case of *cdi*-GMP and F_2cdi -GMP, titrations in different buffers yielded similar binding enthalpies indicating no proton transfer is involved in protein:ligand complex formation. For *c*GAMP and F_2c GAMP, ITC titrations in varying buffers predicted that 0.2 protons are released from the complex upon formation.

Computational analysis. Our structural analysis did not reveal any significant difference between the STING protein structures bound to fluorinated ($2'$ -fluororibo) or natural (ribo) $3',3'$ CDNs that could, *per se*, explain the observed higher activity of the fluorinated CDNs (Figure 3). Moreover, the binding modes of all six ligands are remarkably similar (Figure 5).

In order to gain insight into the stronger binding of fluorinated analogues, we decided to model the process

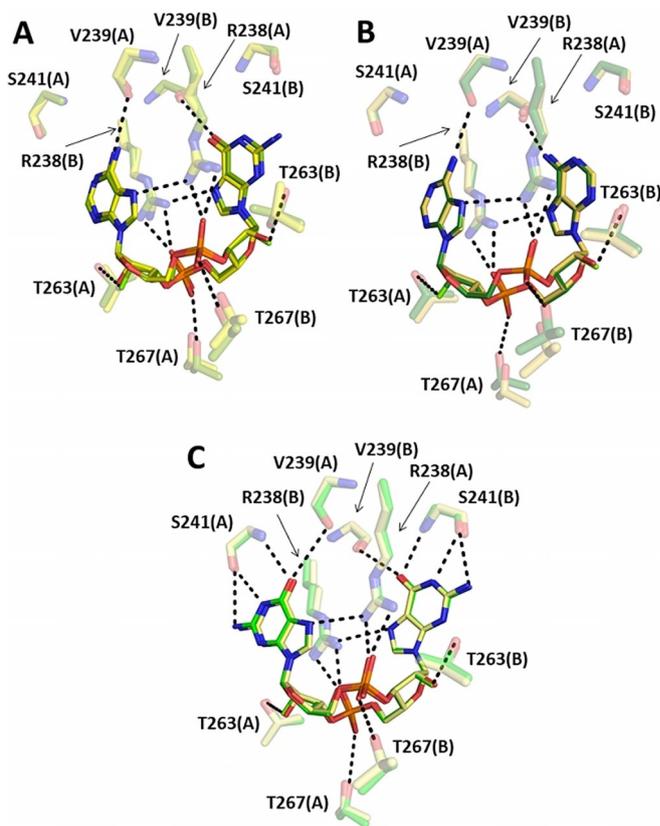


Figure 5. Binding modes of ligands of interest and their interaction with STING: A) *c*GAMP vs. F_2c GAMP; B) *cdi*-AMP vs. F_2cdi -AMP; C) *cdi*-GMP and F_2cdi -GMP (note that only these two form additional three H-bonds with Ser241).

computationally. Throughout, we used “CCSD(T)-calibrated” protocol that combines DLPNO-CCSD(T), MP2-F12 and DFT-D3(BP-86)^[45] complemented with COSMO-RS implicit solvation model.^[46,47]

The overall binding process is encompassed by large conformational changes of the protein, making it very difficult to model directly. Instead, we used different approximations to model individual steps of the process. The thermodynamic cycle used is presented in Figure 6. In order to gain additional support for our analysis, we complemented our study with calculations of $2'$ -deoxy analogues (see Figure S4), which are presented in the Supporting Information material (Tables S2 and S3).

Step 1: Ligand Conformational Entropies. The first step of the thermodynamic cycle shown in Figure 6 restricts the conformational ensemble of an unbound ligand into a single, lowest-energy conformer. The conformational flexibility of all ligands can be inspected visually (Figure S5†). Alternatively, the number of unique clusters (and their energies) is used to estimate entropic cost [see Equations (S2) and (S3) in Computational Details] of restricting the ensemble of ligands in solvent to a single structure. We approximate the overall free energy cost of this step with this entropic term, i.e. $-TS_{\text{conf}} \approx G_{\text{conf}}$. This cost, summarized in Table 2, is consistently lower for fluorinated ligands.

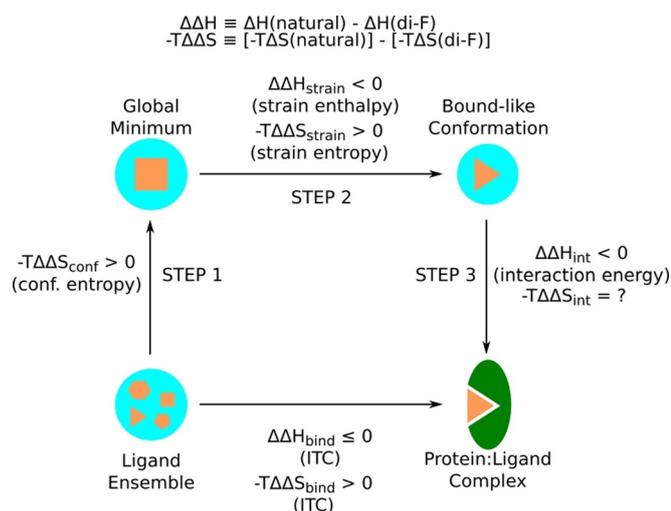


Figure 6. Thermodynamic cycle used to analyze the binding of CDNs to STING. The direct path (bottom) corresponds to the values measured by ITC. For computational analysis, we split this process into three steps. Step 1 represents restricting a conformational ensemble of free ligand in solvent to a single structure—a global minimum. Step 2 represents changing the conformation from global minimum to a bound-like structure. Step 3 represents replacing the solvent environment with a cluster model of protein–ligand complex. Double difference values are defined as $\Delta\Delta X = \Delta X(\text{natural}) - \Delta X(\text{di-F})$.

Table 2: Calculated values of conformational entropy and strain free energy. $-T\Delta S_{\text{conf}}$ represents conformational entropy cost of restricting a free-ligand ensemble to a single conformation. $\Delta G_{\text{strain}}/\Delta H_{\text{strain}}/ -T\Delta S_{\text{strain}}$ represents conformational free energy/enthalpy/entropy difference of a bound-like conformation in solvent and a global minimum. All values are in kcal mol^{-1} .

Ligand	$-T\Delta S_{\text{conf}}$	ΔG_{strain}	ΔH_{strain}	$-T\Delta S_{\text{strain}}$
cGAMP	1.0	9.9	7.4	2.5
F ₂ cGAMP	0.7	9.1	12.7	-3.6
cdi-GMP	1.1	11.4	7.1	4.3
F ₂ cdi-GMP	0.6	10.5	10.4	0.1
cdi-AMP	1.2	8.8	5.5	3.2
F ₂ cdi-AMP	0.7	6.0	7.1	-1.1

Step 2: Ligand-Strain Free Energies. In the second step, we account for the strain that has to be applied to the ligand to change from the most probable conformation in solvent (referred to as the global minimum, see Computational Details for detailed definition) to the bound-like conformation (obtained via refining X-ray models with QM/MM energy minimization). The strain Gibbs free energies are summarized in Table 2. Despite high structural similarity of both the global minima and the bound-like conformations, the strain free energies are consistently lower for fluorinated ligands.

Next, we may dissect the strain free energies into enthalpic and entropic contributions. The partitioning is obtained from temperature dependence of free energies as modelled by the COSMO-RS solvation model. The enthalpic contribution represents the energetic cost of changing the ligand conformation as well as enthalpic change in the interaction with solvent. The entropic contribution reflects

only the changes in solvent, as the conformational entropy of the ligand itself has been discussed separately in the first step. The strain enthalpies are lower for natural ligands in all three cases. At the same time, the entropic term consistently disfavors natural ligands. The 2'-hydroxyl can form both intramolecular and solute-solvent hydrogen bonds. The hydrogen bond formed with the solvent does impose restrictions on organization of solvent, resulting in higher entropic penalty of non-fluorinated ligands. On the other hand, the fluorinated analogues lack these solute-solvent interactions, resulting in comparatively lower entropic penalty. The same holds true for the 2'-deoxy analogues, see Table S2.

Step 3: Protein–ligand interactions. A cluster model of the binding site derived from an optimized QM/MM model (equivalent to QM region and comprising ≈ 600 atoms) was used for evaluating interaction of ligands with the protein. Calculation of interaction energy requires a reference state for both the ligand and the protein. The reference state for ligand has been obtained in the first two steps of our thermodynamic cycle. The reference state of the protein is identical for all ligands (unliganded protein structure). Thus, we focus on $\Delta\Delta G_{\text{int}}$, that is, the difference between the interaction energies for a pair of ligands (natural and fluorinated), in which the term for the reference state of the protein cancels out [Eq. (1)].

$$\Delta\Delta G_{\text{int}} = \Delta G_{\text{int}}(\text{natural}) - \Delta G_{\text{int}}(\text{di-F}) \quad (1)$$

In this convention, positive values imply preference for fluorinated ligands, while negative values imply preference for natural ligands. The values of $\Delta\Delta G_{\text{int}}$ are summarized in Table 3. The negative values of $\Delta\Delta G_{\text{int}}$ can be interpreted as natural ligands being favoured as a result of specific interactions of the hydroxyl group with the protein, which are absent in the case of the fluorinated ligands. Similarly, the absence of specific interactions in this region for 2'-deoxy analogues results in values of $\Delta\Delta G_{\text{int}}$ similar to those of fluorinated ligands, see Table S3.

Table 3: Comparison of calculated ($\Delta\Delta G_{\text{conf}}^{\text{calc}}$, $\Delta\Delta G_{\text{strain}}^{\text{calc}}$, $\Delta\Delta G_{\text{int}}^{\text{calc}}$) values with measured $\Delta\Delta G_{\text{ITC}}^{\text{exp}}$, $\Delta\Delta H_{\text{ITC}}^{\text{exp}}$, $-T\Delta\Delta S_{\text{ITC}}^{\text{exp}}$, and $\Delta\Delta T_m$ values. $\Delta G_{\text{conf}}^{\text{calc}} = -T\Delta S_{\text{conf}}$ represents conformational cost of restricting a free-ligand ensemble to a single conformation. $\Delta G_{\text{strain}}^{\text{calc}}$ represents conformational free energy difference of a bound-like conformation in solvent and a global minimum. $\Delta G_{\text{int}}^{\text{calc}}$ represents free energy of interaction of a protein–ligand complex and a standalone ligand. Each double-difference quantity is defined as $\Delta\Delta X = \Delta X(\text{natural}) - \Delta X(\text{di-F})$. All energy values are in kcal mol^{-1} .

Ligand	cGAMP/ F ₂ cGAMP	cdi-GMP/ F ₂ cdi-GMP	cdi-AMP/ F ₂ cdi-AMP
$\Delta\Delta G_{\text{conf}}^{\text{calc}}$	0.3	0.5	0.6
$-T\Delta\Delta S_{\text{strain}}^{\text{calc}}$	6.1	4.1	4.3
$\Delta\Delta H_{\text{strain}}^{\text{calc}}$	-5.3	-3.3	-1.6
$\Delta\Delta G_{\text{int}}^{\text{calc}}$	-7.1	-3.2	-3.1
$\Delta\Delta H_{\text{ITC}}^{\text{exp}}$	-8.8	-4.7	N/A
$-T\Delta\Delta S_{\text{ITC}}^{\text{exp}}$	11.4	6.0	N/A
$\Delta\Delta G_{\text{ITC}}^{\text{exp}}$	2.6	1.1	N/A
$\Delta\Delta T_m^{\text{exp}} [^{\circ}\text{C}]$	-6.2	-2.4	-10.9

It is important to reiterate that this step of the cycle is modeled by a stationary cluster model. As a result, the $\Delta\Delta G_{\text{int}}^{\text{calc}}$ term is mostly enthalpic and, thus, incomplete. The entropy contributions from the protein:ligand complex are not accounted for in our model and incredibly difficult to model in general. However, we speculate that similarly to the case of the strain free energies (the second step of our thermodynamic cycle discussed above), the entropic and enthalpic components partially compensate each other. Thus, the specific interactions of the 2'-hydroxyl that favour the natural ligands enthalpically may be expected to favour the fluorinated ligands entropically.

The calculated $\Delta\Delta G^{\text{calc}}$ values and the experimental $\Delta\Delta G^{\text{exp}}$ and $\Delta\Delta T_{\text{M}}$ values are summarized in Table 3. The mere summation of the contributions from the individual steps should be interpreted with caution, as it lacks the entropic part of the protein:ligand interaction. Moreover, each of the three calculated steps is addressed using different set of approximations and, hence, we do not expect the values for individual steps to scale consistently.

The analysis above shows that the difluorinated ligands are favoured by lower conformational flexibility of the free ligand as well as lower entropy cost imposed by specific interactions with the solvent and the STING. This preference is only partially countered by stronger enthalpic interactions of the parent ligands with the protein. The observed higher preference for difluorinated ligands implies that it is the entropic term that prevails. This is in full agreement with the large and non-intuitive enthalpy/entropy compensation revealed by ITC measurements.

Conclusion

A series of six ligands –3',3'-cGAMP, 3',3'-cdi-GMP, 3',3'-cdi-AMP and their difluorinated analogues—and their binding to human STING were studied from several perspectives. The ΔT_{M} values obtained using differential scanning fluorimetry (DSF), support previously published data presenting higher potency of 2',2'-fluororibo 3',3' CDNs in interferon type I induction.^[33,34] Therefore, we solved three pairs of crystal structures of STING in complex with natural and difluorinated ligands to understand the high potency of difluorinated CDNs. However, comparison of these six crystal structures did not reveal any significant structural differences that could explain the extraordinary properties of difluorinated ligands. The thermodynamic characterization of the STING:ligand complex formation by ITC provided K_{D} values distinctly lower for difluorinated CDNs, in accordance with the DSF data. Intriguingly, the ΔH values favour natural CDNs, suggesting a huge entropy compensation for the difluorinated ligands.

Computational analysis provided plausible rationalization of this phenomenon. The difluorinated analogues seem to be favoured entropically due to their lower conformational flexibility and organization of solvent around the molecules. This advantage is partially compensated by weaker interaction with the protein. Comparing the magnitude of conformational and strain free energies, $\Delta\Delta G_{\text{conf}}^{\text{calc}} + \Delta\Delta G_{\text{strain}}^{\text{calc}}$ (i.e.

the ligand-only part of the $\Delta\Delta G_{\text{bind}}^{\text{calc}}$) with $\Delta\Delta G_{\text{int}}^{\text{calc}}$ shows that the properties of the ligand itself do play a prominent role, but the resulting affinity of the ligand to this protein is a complex interplay of both the ligand and protein-ligand properties. This highlights that design of ligands can benefit heavily from thorough conformational analysis performed at high level of theory—using standard, but advanced methods of quantum chemistry.

Acknowledgements

The project was supported by the Grant Agency of the Czech Republic (20-08772S to L.R.), European Regional Development Fund; OP RDE (Project No. CZ.02.1.01/0.0/0.0/16_019/0000729) and the Ministry of Education, Youth and Sports from the Large Infrastructures for Research, Experimental Development and Innovations project „e-Infrastructure CZ – LM2018140“.

Conflict of interest

The authors declare no conflict of interest.

Keywords: conformational analysis · cyclic dinucleotides · entropy · quantum chemistry · strain energy

- [1] Y. B. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, D. E. Shaw, *J. Am. Chem. Soc.* **2011**, *133*, 9181–9183.
- [2] U. Ryde, P. Soderhjelm, *Chem. Rev.* **2016**, *116*, 5520–5566.
- [3] “What Next for Quantum Mechanics in Structure-Based Drug Discovery?”: R. A. Bryce in *Quantum Mechanics in Drug Discovery* (Ed.: A. Heifetz), Springer US, New York, **2020**, pp. 339–353.
- [4] A. Sandner, T. Hufner-Wulsdorf, A. Heine, T. Steinmetzer, G. Klebe, *J. Med. Chem.* **2019**, *62*, 9753–9771.
- [5] Y. Wang, J. Kim, C. Hilty, *Chem. Sci.* **2020**, *11*, 5935–5943. RETURN TO ISSUEPREVARTICLENEXT
- [6] I. Mejdrová, D. Chalupská, P. Plačková, C. Müller, M. Šála, M. Klíma, A. Baumlová, H. Hřebabecý, E. Procházková, M. Dejmek, D. Strunin, J. Weber, G. Lee, M. Matoušová, H. Mertlíková-Kaiserová, J. Ziebuhr, G. Birkus, E. Boura, R. Nencka, *J. Med. Chem.* **2017**, *60*, 100–118.
- [7] S. M. Eyrlmez, C. Kopruluoglu, J. Rezac, P. Hobza, *ChemPhysChem* **2019**, *20*, 2759–2766.
- [8] A. Pecina, J. Brynda, L. Vrzal, R. Gnanasekaran, M. Hořejší, S. M. Eyrlmez, J. Řezáč, M. Lepšík, P. Řezáčová, P. Hobza, P. Majer, V. Veverka, J. Fanfrlík, *ChemPhysChem* **2018**, *19*, 873–879.
- [9] G. C. Wu, T. Zhao, D. W. Kang, J. Zhang, Y. N. Song, V. Namasivayam, J. Kongsted, C. Pannecouque, E. De Clercq, V. Poongavanam, X. Y. Liu, P. Zhan, *J. Med. Chem.* **2019**, *62*, 9375–9414.
- [10] L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood, D. L. Romero, C. Masse, J. L. Knight, T. Steinbrecher, T. Beuming, W. Damm, E. Harder, W. Sherman, M. Brewer, R. Wester, M. Murcko, L. Frye, R. Farid, T. Lin, D. L. Mobley, W. L. Jorgensen,

- B. J. Berne, R. A. Friesner, R. Abel, *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.
- [11] G. Klebe, *Drug Discovery Today* **2019**, *24*, 943–948.
- [12] M. Sarter, D. Niether, B. W. Koenig, W. Lohstroh, M. Zamponi, N. H. Jalarvo, S. Wiegand, J. Fitter, A. M. Stadler, *J. Phys. Chem. B* **2020**, *124*, 324–335.
- [13] C. Barinka, Z. Novakova, N. Hin, D. Bim, D. V. Ferraris, B. Duvall, G. Kabarriti, R. Tsukamoto, M. Budesinsky, L. Motlova, C. Rojas, B. S. Slusher, T. A. Rokob, L. Rulíšek, T. Tsukamoto, *Bioorg. Med. Chem.* **2019**, *27*, 255–264.
- [14] N. Foppe, I. J. Chen, *Bioorg. Med. Chem.* **2016**, *24*, 2159–2189.
- [15] M. Sitzmann, I. E. Weidlich, I. V. Filippov, C. Liao, M. L. Peach, W. D. Ihlenfeldt, R. G. Karki, Y. V. Borodina, R. E. Cachau, M. C. Nicklaus, *J. Chem. Inf. Model.* **2012**, *52*, 739–756.
- [16] M. L. Verteramo, O. Stenstrom, M. M. Ignjatovic, O. Caldararu, M. A. Olsson, F. Manzoni, H. Leffler, E. Oksanen, D. T. Logan, U. J. Nilsson, U. Ryde, M. Akke, *J. Am. Chem. Soc.* **2019**, *141*, 2012–2026.
- [17] S. R. Tzeng, C. G. Kalodimos, *Nature* **2012**, *488*, 236–240.
- [18] R. Gaspari, A. E. Prota, K. Bargsten, A. Cavalli, M. O. Steinmetz, *Chem* **2017**, *2*, 102–113.
- [19] S. Geschwindner, J. Ulander, *Expert Opin. Drug Discovery* **2019**, *14*, 1221–1225.
- [20] H. Zhang, Q. D. You, X. L. Xu, *J. Med. Chem.* **2020**, *63*, 3785–3816.
- [21] Y. Tanaka, Z. J. J. Chen, *Sci. Signaling* **2012**, *5*, 11.
- [22] X. Cai, Y. H. Chiu, Z. J. J. Chen, *Mol. Cell* **2014**, *54*, 289–296.
- [23] R. Fang, C. G. Wang, Q. F. Jiang, M. Z. Lv, P. F. Gao, X. Y. Yu, P. Mu, R. Zhang, S. Bi, J. M. Feng, Z. F. Jiang, *J. Immunol.* **2017**, *199*, 3222–3233.
- [24] J. Ahn, S. Son, S. C. Oliveira, G. N. Barber, *Cell Rep.* **2017**, *21*, 3873–3884.
- [25] G. N. Barber, *Nat. Rev. Immunol.* **2015**, *15*, 760–770.
- [26] J. Papinska, H. Bagavant, G. B. Gmyrek, M. Sroka, S. Tummala, K. A. Fitzgerald, U. S. Deshmukh, *J. Dent. Res.* **2018**, *97*, 893–900.
- [27] Q. Chen, L. J. Sun, Z. J. J. Chen, *Nat. Immunol.* **2016**, *17*, 1142–1149.
- [28] D. L. Burdette, K. M. Monroe, K. Sotelo-Troha, J. S. Iwig, B. Eckert, M. Hyodo, Y. Hayakawa, R. E. Vance, *Nature* **2011**, *478*, 515–518.
- [29] R. Tamayo, J. T. Pratt, A. Camilli, *Annu. Rev. Microbiol.* **2007**, *61*, 131–148.
- [30] Y. Tsuchiya, N. Jounai, F. Takeshita, K. J. Ishii, K. Mizuguchi, *EBioMedicine* **2016**, *9*, 87–96.
- [31] C. Shu, G. H. Yi, T. Watts, C. C. Kao, P. W. Li, *Nat. Struct. Mol. Biol.* **2012**, *19*, 722–724.
- [32] X. Zhang, H. P. Shi, J. X. Wu, X. W. Zhang, L. J. Sun, C. Chen, Z. J. J. Chen, *Mol. Cell* **2013**, *51*, 226–235.
- [33] T. Lioux, M.-A. Mauny, A. Lamoureux, N. Bascoul, M. Hays, F. Vernejoul, A.-S. Baudru, C. Boullaran, J. Lopes-Vicente, G. Qushair, G. Tiraby, *J. Med. Chem.* **2016**, *59*, 10253–10267.
- [34] “Fluorinated Cyclic Dinucleotides for Cytokine Induction”: INVIVOGEN, WO2016096174A, **2016**.
- [35] S. Y. Ouyang, X. Q. Song, Y. Y. Wang, H. Ru, N. Shaw, Y. Jiang, F. F. Niu, Y. P. Zhu, W. C. Qiu, K. Parvatiyar, Y. Li, R. G. Zhang, G. H. Cheng, Z. J. Liu, *Immunity* **2012**, *36*, 1073–1086.
- [36] J. Antony, S. Grimme, D. G. Liakos, F. Neese, *J. Phys. Chem. A* **2011**, *115*, 11210–11220.
- [37] B. Novotná, L. Vaneková, M. Zavřel, M. Buděšínský, M. Dejmek, M. Smola, O. Gutten, Z. A. Tehrani, M. P. Polidarová, A. Brázdová, R. Liboska, I. Štěpánek, Z. Vavřina, T. Jandůšek, R. Nencka, L. Rulíšek, E. Bouřa, J. Brynda, O. Páv, G. Birkuš, *J. Med. Chem.* **2019**, *62*, 10676–10690.
- [38] M. Smola, G. Birkus, E. Boura, *Acta Crystallogr. Sect. F* **2019**, *75*, 593–598.
- [39] S. L. Ergun, D. Fernandez, T. M. Weiss, L. Y. Li, *Cell* **2019**, *178*, 290–301.
- [40] H. P. Shi, J. X. Wu, Z. J. J. Chen, C. Chen, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8947–8952.
- [41] P. Gao, M. Ascano, T. Zillinger, W. Y. Wang, P. H. Dai, A. A. Serganov, B. L. Gaffney, S. Shuman, R. A. Jones, L. Deng, G. Hartmann, W. Barchet, T. Tuschl, D. J. Patel, *Cell* **2013**, *154*, 748–762.
- [42] S. G. Krimmer, G. Klebe, *J. Comput.-Aided Mol. Des.* **2015**, *29*, 867–883.
- [43] O. Bastiansen, J. J. Christensen, L. D. Hansen, R. M. Izatt, *Handbook of Proton Ionization Heats and Related Thermodynamic Quantities*, Wiley, New York **1976**.
- [44] H. Fukada, K. Takahashi, *Proteins: Struct. Funct. Genet.* **1998**, *33*, 159–166.
- [45] J. Řezáč, D. Bím, O. Gutten, L. Rulíšek, *J. Chem. Theory Comput.* **2018**, *14*, 1254–1266.
- [46] A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224–2235.
- [47] A. Klamt, V. Jonas, T. Bürger, J. C. W. Lohrenz, *J. Phys. Chem. A* **1998**, *102*, 5074–5085.

Manuscript received: December 18, 2020

Accepted manuscript online: February 22, 2021

Version of record online: March 24, 2021