Probing the dynamical interaction of the para-sulfonato-calix[4]arene with an antifungal protein

Alessio Bartocci*,† and Elise Dumont‡,¶

†*Institut de Chimie de Strasbourg, UMR 7177, CNRS, Université de Strasbourg, Strasbourg Cedex 67083, France*

‡*Université Côte d'Azur, CNRS, Institut de Chimie de Nice, UMR 7272 - 06108 Nice, France*

¶*Institut Universitaire de France, 5 rue Descartes, 75005 Paris, France*

E-mail: abartocci@unistra.fr

Abstract

Calixarenes are hallmark molecules in supramolecular chemistry as guest cages for small ligands. They have also conversely proved their interest as auxiliary ligands toward assisted co-crystallization of proteins. These functionalized macromolecular cages target positively-charges residues, and notably surface-exposed lysines, with a site-selectivity finely characterized experimentally, but that remains to be assessed. Relying on a tailored molecular dynamics simulations protocol, we explore the association of *para*-sulfonato-calix[4]arenes with an antifungal protein, as a small yet most competitive system with 13 surface-exposed

lysines. Our computational approach probe *de novo* the electrostatically-driven interaction, ruled out by a competition with salt bridges, not only corroborating the main binding site probed by X-ray, but also characterizing a second binding site that can act as a transient hub spot. The attach-pull-release (APR) method provides a very good assessment of the overall binding free energy measured experimentally (-6.42 $\pm$ 0.5 *vs.* -5.45 kcal.mol^{-1} by isothermal titration calorimetry). This work also probes dynamical allosteric modifications upon ligand binding, and our computational protocol could be generalized to situate the supramolecular forces ruling out the calixarene-assisted co-crystallization of proteins.

Introduction

Functionalizable ligands have attracted growing interest, mostly for their potential applications, spanning from chemistry[1,2], pharmaceutical fields [3–5] to material science[6,7]. Interaction of proteins with smaller ligands is a hallmark in Biochemistry, and has recently been implied in the conception of high-resolution tunable protein assemblies, with an increasing number of X-ray structures reported in the PDB databank. Three families of ligands have been proposed so far as "molecular glues" [8]: lanthanide-based coordination complexes[9–12], polyoxometalates[13] and calixarenes[14,15]. The three of them give rise to controlled and tuneable supramolecular bio-assemblies.

Calixarenes, due to their tunable size up to giant calixarenes[16] and to their large possibilities of functionalization[17–21], are associated with many applications in biological fields, as chemotherapy delivery carriers [22–27], protein inhibitors[28,29] and surface binders-recognizers[15,30–32]. The specific interaction with overall positively-charged proteins, like cytochrome-C[8,33,34], lysozyme[31] and small antifungal proteins (PAF)[35,36], makes calixarenes ideal ligand host partners[37] in encapsulating positively-charged residues such as lysines and arginines[38–41]. Electrostatic interactions lead protein-calixarene binding, and ligand "hot-spots" and affini-

ties can be experimentally detected through X-ray crystallography, isothermal calorimetry titration (ITC) and NMR, and even most recently solid-state NMR[42].

Computational approaches such as molecular dynamics (MD) simulations offer not only a complementary outlook into the genesis of the protein–calixarenes binding events, including their interaction patterns, mechanisms and affinities, but also molecular flexibility and allosteric effects. In this work, we challenge the genesis of $\mathbf{sclx}_4$ interaction with the small molecular mass *Penicillum* antifungal protein (*PAF*), for which a complete experimental characterization was obtained by Crowley and coworkers[35]. *PAF* is a 55-residues, rigid protein (three disulfide bridges, five β sheets). This cationic protein is a potent agent against dermatophytes[43–45], that interacts directly with anionic species and manages to permeate the membrane[43,46]. ITC experiments confirm the existence of one specific binding site, pointed out also by X-ray structures, delineating a single-site model for this complex[35] and a calixarene–assisted high-resolution crystal packing (1.33 Å). In this site, which is depicted in Figure 1B, K30 is encapsulated within the calixarene cage and two lysines K27 and K35 cooperatively form salt bridges with two sulfonate groups.

We rely on a computational approach, tailored recently to characterize *Cytc*–calixarene systems on cytochrome C protein[47,48] and sketched in Figure 1A. It consists in unbiased, diffusion MD simulations with many independent replica (20), followed by "Bound-state" and free energy MD simulations, based on the Attach-Pull-Release (APR)[49–53] method, to : (i) assign *de novo* plausible competitive binding sites, independently from the crystallographic ones, (ii) estimate the free energy of binding to be compared to ITC measurements and (iii) characterize the dynamics of binding and possible structural consequences on competitive salt bridges.

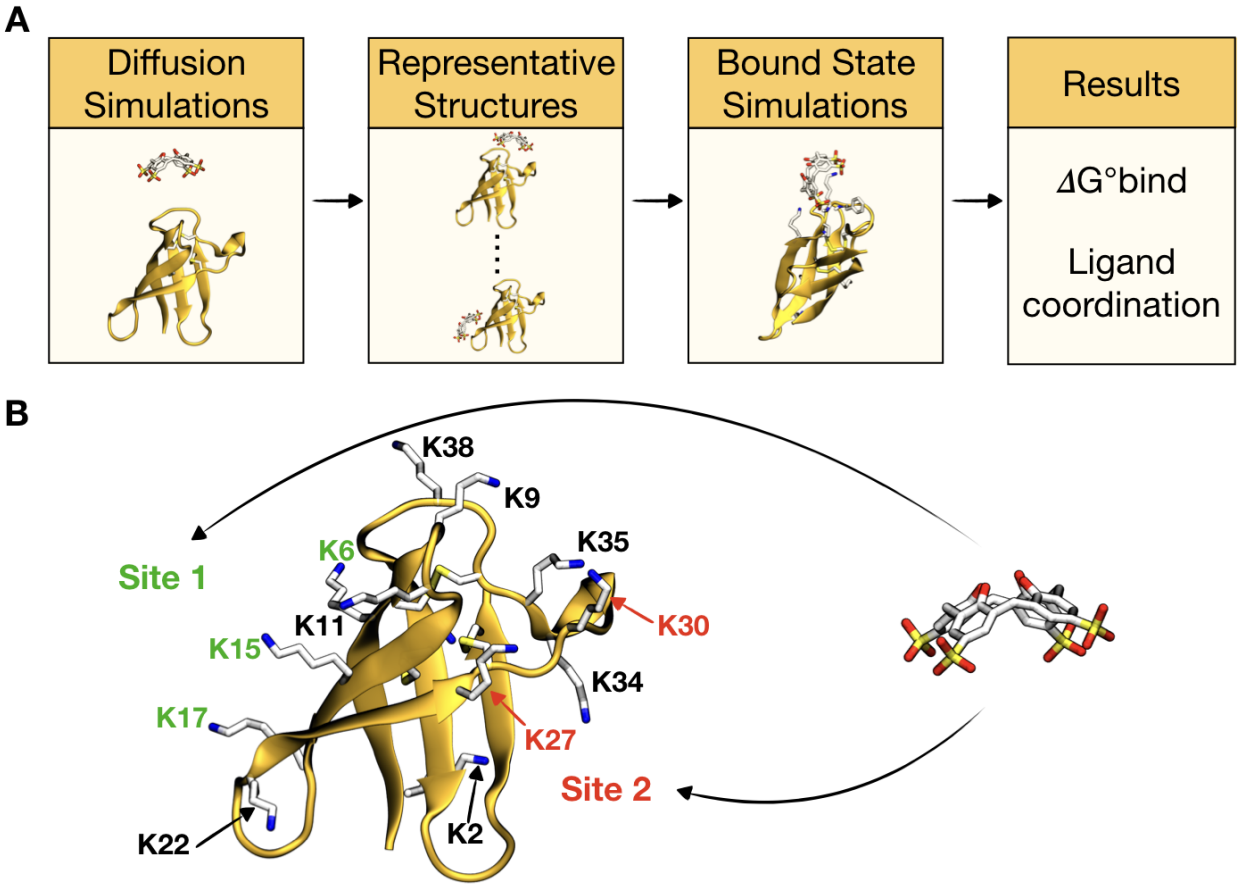

Figure 1: **A)** *MD simulations protocol stages carried out in this work. MD "Diffusion" simulations are implemented to capture the dynamics of ligand binding events and extract PAF–*$\mathbf{sclx}_4$ *representative structures. Subsequent "Bound State" simulations allow to assess coordination and free binding energy* ΔG^0_{bind} *between protein and ligand.* **B)** *Two main modes of approaches for* para*-sulfonato-calix-[4]-arene (*$\mathbf{sclx}_4$*, in colored sticks) towards PAF (in orange cartoon, PDB ID 6HA4) identified in our MD simulations. The two binding sites at the protein surface,* **Site 1** *and* **Site 2***, which are labeled in green and in red throughout this work, are shown together with all protein lysines indicated in colored sticks and labeled in black. Hydrogen atoms are not reported for the sake of clarity.*

Material and Methods

Force field parameters

Classical all-atom molecular dynamics (MD) simulations were performed using the AMBER18 package[54]. *Penicillum* antifugal protein "*PAF*" protein (PDB: 6HA4, 55 residues) was used as interacting partner with *para*-sulfonato-calix[4]arene ($\mathbf{sclx}_4$). The protonation state of PAF titratable residues was determined via the H++ server[55] at the experimental pH of 6.0[35]. For the $\mathbf{sclx}_4$ ligand, present with one deprotonated phenolic hydroxyl (net charge of -5) at the current pH, was used a previous topology described in ref.[48]. Force field parameters for the $\mathbf{sclx}_4$ ligand were taken from GAFF2[56], while the protein was described with the AMBER/ff14SB[57] force field (including the three disulfide bridges C7–C36, C14–C43 and C28–C54).

MD Simulations Protocol

Diffusion Simulations

To start our simulations, the calixarene ligand was placed far from the proteins (at more than 20 Å between the centers of mass, see Figure S1), to probe molecular association with no bias. The systems were initially minimized for 10000 steps (5000 of steepest descent and 5000 of conjugate gradient), and then heated up from $0\,K$ to $300\,K$ (with an integration time step t_{step} of 1.0 fs) for a total of 30 ps using the Langevin thermostat ($\gamma_{coll} = 1\,ps^{-1}$) in the isothermal-isochoric ensemble (NVT). Equilibration was carried out for 1 ns, using an integration timestep of 2 fs in the isothermal-isobaric (NPT) ensemble ($P = 1\,atm$ and $T = 300\,K$). During these phases, position restraints were applied on the protein heavy atoms, with a force constant of 5 kcal/mol.Å^2 during the minimization and heating processes, and of 2.5 kcal/mol.Å^2 during the equilibration. A subsequent equilibration of

1 ns followed with no position restraints applied. Pressure control in NPT simulations was achieved using the Berendsen barostat. For each system (see SI section "Diffusion Simulations"), the production phase was carried out in five replicates of 300 ns each with different initial velocities, giving in total 6 μs per system, with the `pmemd.cuda` module of AMBER18, through the use of graphics processing units (GPUs). The last 100 ns of each production run were used to post-processing analysis. In the equilibration and production runs, a cutoff of 10 Å was applied for the van der Waals, for electrostatic interactions and for the real space of the electrostatic interaction. Long-range electrostatic interactions were computed using Particle Mesh Ewald (PME) algorithm[58,59].

Bound-State Simulations

The most representative structures of the bound-states, extracted from cluster-analysis (see section "Bound-State Simulations" in SI), were simulated to dynamically characterize the interaction. In order to keep the ligand bound in the binding sites, position restraints were applied both on the protein and ligands heavy atoms. The pre-production simulation stages occurred as described previously. Three 500 ns plus two 1 μs long replicates (with different initial velocities) per bound state were performed in order to monitor and assess the binding site stability.

PAF–$\mathbf{sclx_4}$ Binding Affinity

PAF–$\mathbf{sclx_4}$ binding affinity was obtained using the attach-pull-release (APR) method[49–53]. Starting from the bound-state structure, reported in Figure 2 and depicting $\mathbf{sclx_4}$ interacting in **Site 2**, the ligand, subject to restraints, was pulled away from the protein contacts towards the bulk along the distance reaction coordinate. In this phase ("attachment phase", in total 14 windows), the ligand restraints are increased up to the maximum value defined by the force constants. Then $\mathbf{sclx_4}$ is pulled away from the protein (total of 70 windows) from

0 Å to 28 Å, with an increment of 0.4 Å to ensure good overlap. For each window of the entire process, the system was first energy minimized by 5000 initial steps of steepest descent and 15000 of conjugate gradient, and then was slowly heated for 100 ps from 10 to 300 K at constant volume (NVT), and equilibrated for 2.5 ns in the NPT ensemble. Subsequent production runs were performed for 110 ns with the `pmemd.cuda` module of `Amber18`. The simulation analysis started discarding the first 10 ns. The total simulation time amounts to 8.4 μs (1.4 μs for the attachment phase and 7.0 μs for the pull phase).

Results and Discussion

Among the total number of 55 residues, *PAF* features 13 lysines (K2, K6, K9, K11, K15, K17, K22, K27, K30, K34, K35, K38 and K42), 7 aspartates (D19, D23, D32, D39, D46, D53 and D55) and one glutamate (E13) as charged residues, which corresponds to an overall positive charge (+5) at pH 6. As for other proteins prone to co-crystallize with *para*-sulfonato-calixarenes, the overall positive charge makes the protein a perfect interacting partner for $\mathbf{sclx}_4$, delineating a high competition. Salt bridges between lysines and negatively-charged residues make a good point for investigating their accessibility and selectivity to bind $\mathbf{sclx}_4$.

"Diffusion Simulations": *de novo* identification of binding sites

Extensive unbiased MDs (20 independent trajectories of 300 ns each) were performed to identify *de novo* the main ligand binding site(s) at the protein surface. Each trajectory was found to converge toward a bound state after at most 200 ns ligand diffusion (see Figure S2) : inspection of our MD followed by cluster analysis reveal two main protein binding sites (**Site 1** and **Site 2**, shown in Figure 2) serving as anchor points for $\mathbf{sclx}_4$. In **Site 2**, $\mathbf{sclx}_4$ encapsulates in its hydrophobic cage lysine K30, whose side chain adopts an inchworm conformation, identified in the co-crystallized X-ray structure of PAF[35] and for

other proteins[8,33]. In addition, $CH\cdots\pi$ interactions play also a role between the methylene groups of K30 and the **sclx**$_4$ phenolic rings. Through their oxygen atoms, the sulfonate groups form direct HB with the protein residues: two of them with K30 -NH_3^+ terminal group at an average distance of 2.8 Å, one with K27 -NH_3^+ terminus (2.9 Å) and one with the backbone amide NHs of K30 and F31 (2.9 and 3.4 Å, respectively). This structure lies in very good agreement with the X-ray[35] distances (2.93, 3.05 and 3.02 Å, respectively, for contacts with K30 -NH_3^+ terminus, and K30 and F31 peptidic groups).

In contrast, in the interaction at **Site 1** shown in Figure 2B, the ligand approaches the surface through its *exo* conformation, developing only sulfonate–lysine HB interactions. Three lysines (K6, K15 and K17) cooperatively establish HBs with the oxygens of the ligand's sulfonate groups, appearing as coordinating residues in an elongated conformation, at a distance of 2.8 Å. An additional stabilization is driven by $CH\cdots\pi$ interaction between the lysines alkyl side chain and the **sclx**$_4$ phenyl rings.

Per-residue decomposition interactions

The last 10 ns of each replicate were analyzed to obtain a per-residue decomposition ($\Delta\Delta G$) of the protein-ligand interaction energy with the Molecular Mechanics/Poisson–Boltzmann surface area (MM–PBSA) approach[60–62]. This approach provides in turn a per-residue decomposition, which highlights the most relevant per-residue contributions $\Delta\Delta G$, and provides in turn a map-view of the interaction profile (see Figure S3). The overall lysines contribution is estimated at -27.5 ± 9.7 kcal.mol^{-1} (see Table 1), which can be further decomposed as -4.7 ± 4.5 kcal.mol^{-1} in **Site 1** (K6, K15 and K17) and -16.3 ± 8.0 kcal.mol^{-1} in **Site 2** (K27 and K30). Repulsive interactions with negatively-charged residues, aspartates and glutamates, stand for 4.3 ± 1.6 kcal.mol^{-1} and 0.7 ± 0.2 kcal.mol^{-1}, respectively.

A non-negligible role in **Site 1**, as already pointed out by the cluster analysis in Figure 2B, is associated to the coordination of the **sclx**$_4$ ligand by K6, K15 and K17, leading to three nearly

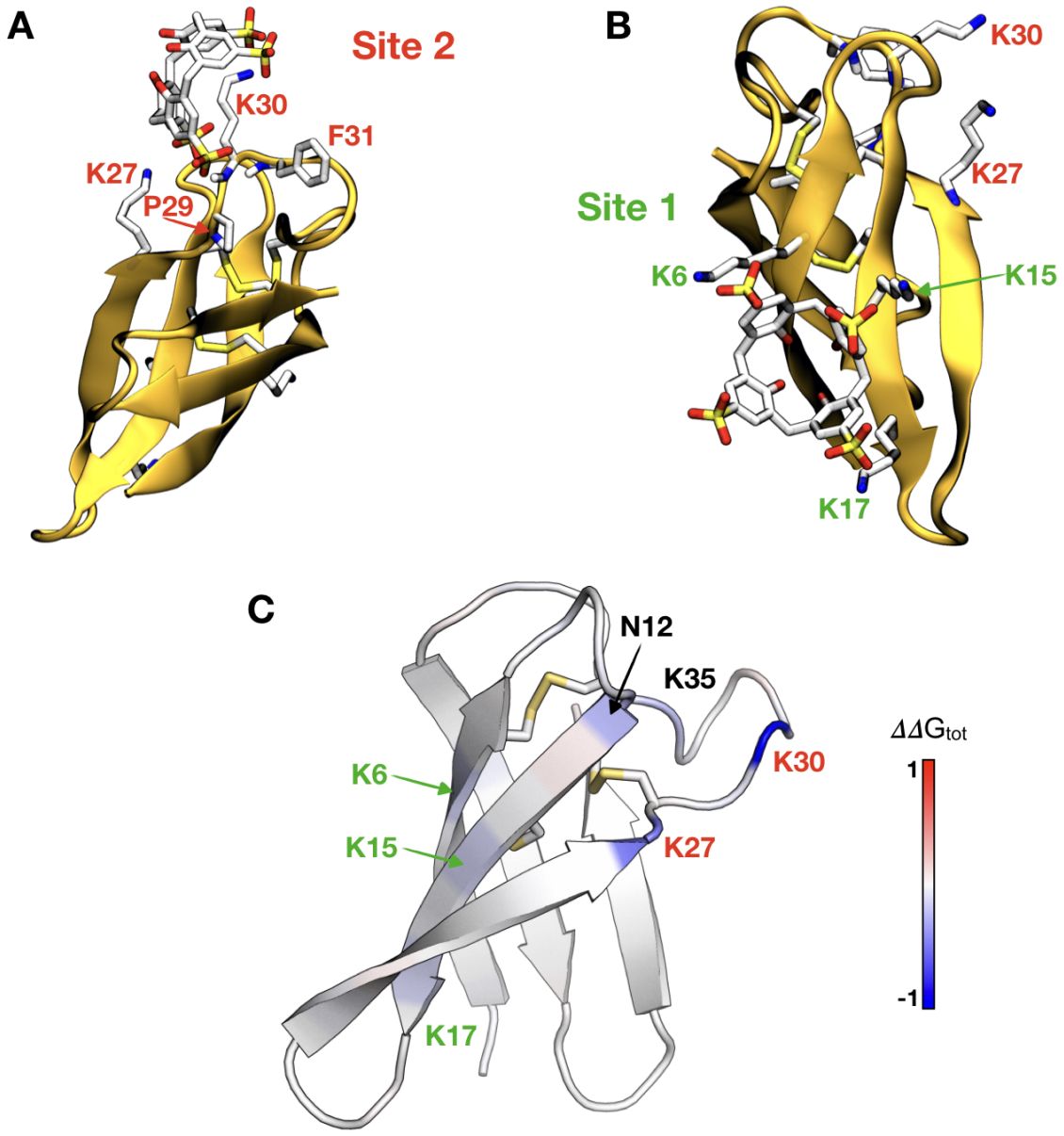

Figure 2: **Panels A and B**: *the two most representative structures with* $\mathbf{sclx_4}$ *interacting in* **Site 2** *(70%) and* **Site 1** *(25%) obtained upon cluster analysis.* $\mathbf{sclx_4}$ *binds through its endo and exo conformations, respectively. PAF is reported as orange cartoons, while* $\mathbf{sclx_4}$ *and the binding residues are shown in licorice tubes. K30 (***Site 2***) adopts an inchworm conformation inside the* $\mathbf{sclx_4}$ *cage.* **Panel C**: *color map of the per-residue decomposition* $\Delta\Delta G$ *(see Figure S3). The color scale is defined by dividing each contribution by the maximal absolute one (*$\Delta\Delta G$ *(K30)).*

Table 1: *Per-residue decomposition contributions $\Delta\Delta G$ for the residues interacting with* $\mathbf{sclx}_4$*(in kcal.mol^{-1}). $\Delta\Delta G_{per-site}$ indicates, for each binding site, the total ($\Delta\Delta G_{tot}$) and per-lysine ($\Delta\Delta G_K$) contributions. The protein residues, identified in the experimental crystal structure[35], are boldfaced.*

	RES	$\Delta\Delta G$ **(kcal.mol^{-1})**	$\Delta\Delta G_{per-site}$ **(kcal.mol^{-1})**
Site 1	K6	-1.8 ± 2.9	$\Delta\Delta G_K$ = -4.7 ± 4.5
	K15	-1.5 ± 2.5	
	K17	-1.4 ± 2.3	
Site 2	**K27**	-5.4 ± 3.9	$\Delta\Delta G_{tot}$ = -17.6 ± 8.0
	P29	-0.8 ± 0.7	$\Delta\Delta G_K$ = -16.3 ± 8.0
	K30	-10.9 ± 7.0	
	F31	-0.5 ± 0.6	
Other	N12	-2.9 ± 1.8	
	K34	-0.7 ± 0.3	
	K35	-2.7 ± 2.7	

equivalent contributions of -1.8 ± 2.9, -1.5 ± 2.5 and -1.4 ± 2.3 kcal.mol^{-1}, respectively. The direct encapsulation of a lysine residue, as in **Site 2** for K30, gives rise to a per-residue contribution of -10.9 ± 7.0 kcal.mol^{-1}, underlining how the singularity of **sclx**$_4$ as a host species is particularly appealing for more complex applications as mediators of protein–protein contacts. At the same time, P29 and F31 offer additional contacts to **sclx**$_4$ (-0.8 ± 0.7 and -0.5 ± 0.6 kcal.mol^{-1}, respectively), revealing the ligand ability to also interact with not–charged residues. N12 and K34-K35 tweezer, located close to **Site 2**, contribute for -2.9 ± 1.8, -0.7 ± 0.3 and -2.7 ± 2.7 kcal.mol^{-1}, respectively, forming a bridge between the two protein sites, a role that can also be ensured by other residues as arginine[47,48]. The overall contributions are shown by the color map provided in Figure 2C, which visibly distinguishes the two different binding sites with contrasted affinities. We note in passing that K35's contribution is absent from the X-ray structure, at a distance of 3.85 Å between the ammonium termini and the closest sulfonate.

Binding free energy assessment for two stable binding modes

Starting from the stable *PAF*–**sclx**$_4$ bound structures shown in Figure 2, unbiased MD simulations were carried out (3.5 μs in total for each structure) to assess the stability of the two binding sites at the microsecond time range, to dynamically characterize the ligand coordination, and compute overall binding free energies. The two most representative resulting structures obtained from the cluster analysis are shown in Figure S4. In **Site 1**, **sclx**$_4$ binds to the protein surface through the *exo* conformation and coordinates *PAF* through a triangle of three lysines K6, K15 and K17. From **Site 1**, **sclx**$_4$ could migrate to the more thermodynamically stable **Site 2** as can be seen from Figure S5 (upper panel), where the distances between the centers of mass of **sclx**$_4$ and the residues implied in the binding sites are monitored. Ligand local rearrangements occur and along different replicates the ligand binds back to **Site 2** after unbinding from **Site 1**, suggesting this site as an intermediate one

to **Site 2**. Our MD simulations confirm the ligand ability to hop between the lysine-driven sites at the protein surface.

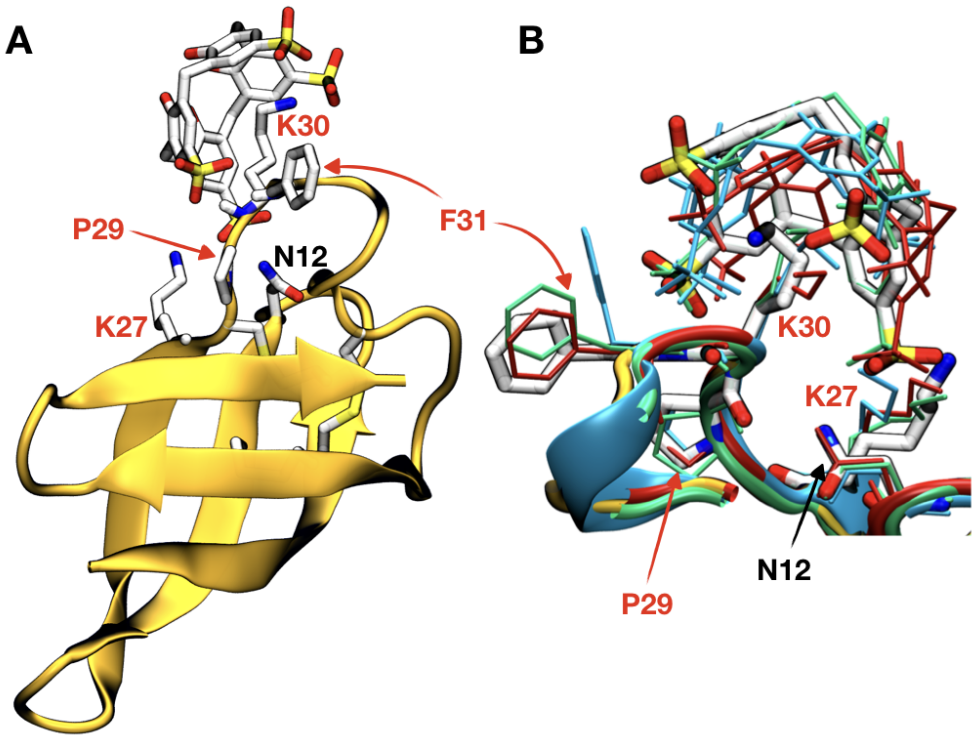

Figure 3: **A***: the most representative structure of* **sclx**$_4$ *interacting with PAF at* **Site 2** *obtained from "Bound State" simulations of Figure 2A structure.* **B***: zoom of the superimposition of the most representative MD-extracted structures with the X-ray structure*[35] *(cyan). The second most representative structure from cluster analysis of the "Bound State" simulations is also superimposed in red, while the one from "Diffusion Simulations" (see Figure 2A) is sketched in green. RMSD values are reported in Table 2.*

In Figure 3A, the most representative structure from cluster analysis of "Bound State" simulations of Figure 2A structure, shows **sclx**$_4$ in **Site 2** encapsulating K30, and coordinated by HBs with K27, F31 and N12. The ligand is interacting through the *endo* surface, and the stability of this binding site is highlighted in the lower panel of Figure S5. While all over the simulations K30 remains encapsulated in the hydrophobic **sclx**$_4$ cavity and K27 coordinates the ligand through via -NH_3^+–sulfonate interactions, at around 1.8 μs it jumps to encapsulate K27 yet comes back after a few nanoseconds (see Figure S5). The ability of

the calixarene to jump between binding sites was also surmised by Crowley and coworkers[8,33], and is corroborated by our MD simulations, which restore the flexibility of the binding site in solution compared to the X-ray structure. In addition, HBs between $\mathbf{sclx_4}$ and K30 and F31 amide -NH moieties contribute weakly yet cooperatively to stabilize the complex. Figure 3B, zooming in **Site 2**, superimposes the X-ray structure with relevant structures obtained from cluster analysis. The associated RMSD values, given in Table 2, indicate a neat agreement with the X-ray reference structure. The time-averaged distances between residues nitrogens and the $\mathbf{sclx_4}$ sulphur atoms (see lower panel of Table 2 as MD) shows a good overlap with the X-ray values. The dynamic local rearrangements and the short jump (around 1.8 μs) allow $\mathbf{sclx_4}$ to interchange the interacting sulphur atoms with the nitrogens. This is pointed out also by the distance distributions and their average values, reported in Figure S6 and Table S2, respectively. The binding in **Site 2** turns out to be specific not only for the *endo* $\mathbf{sclx_4}$ surface, but also rotational insensitive, around the ligand principal axis, for the 4-fold symmetric sulfonate groups.

Our computational approach also brings information on the *endo* vs. *exo* approaching conformational preference as $\mathbf{sclx_4}$ binds PAF, which is reflected by the Θ angle (see SI for definition). In **Site 2**, Θ is centered at a value higher than 100° (average 123.9 ± 13.1° , Table 2), indicating a stable *endo* conformation, while for **Site 1** a broader angle distribution can be seen in Figure S7 : for this site, unbinding events and local flexibility allow the ligand to interact from multiple sides. We note that *endo* vs. *exo* conformations as the calixarene ligands walks on the protein surface were also observed for calixarenes interacting with cytochrome C[48].

The overall *PAF*–$\mathbf{sclx_4}$ binding free energy (-5.45 kcal.mol^{-1}) was also assessed by ITC experiments[35], showing a clear $\mathbf{sclx_4}$ selectivity for **Site 2**. Our calculated ΔG^o_{bind} by the APR approach (-6.42 ± 0.46 kcal.mol^{-1}) lies in very good agreement with the experimental findings with a difference lower than 1 kcal.mol^{-1}(0.97), underlining its applicability in

Table 2: **Upper table**: *RMSD values are reported in Å for the most representative structures obtained from the cluster analysis shown in Figures 2A and 3. The* $\mathbf{sclx_4}$ *ligand is interacting in* **Site 2** *in the X-ray structure (PDB ID 6AH4) taken as reference. "Protein" refers to RMSD values calculated for the protein backbone atoms.* **Lower table**: *distances (in Å) between the nitrogen atom of the -NH_3^+ group and the closest* $\mathbf{sclx_4}$ *sulphur atom (d_{S-NZ}) and between the amide nitrogen atom and the closest* $\mathbf{sclx_4}$ *sulphur atom (d_{S-NH}), and* Θ *angle (in degrees °), formed between the centers of mass of the protein, the sulfonate's sulphur and hydroxyl oxygen atoms (see section "Interacting Surface of* $\mathbf{sclx_4}$*" in SI for further details). X-ray refers to the value in the crystallographic structure , MD to the average value obtained from the "Bound State" simulations, while in both tables* $Bound^1_{Site2}$, $Bound^2_{Site2}$ *and* $Diff_1$ *to the values extracted from the structures shown in Figures 2A and 3.*

	RMSD		
	$\mathbf{Bound}^1_{Site2}$	$\mathbf{Bound}^2_{Site2}$	$\mathbf{Diff}_1$
Protein	0.67	0.74	0.84
K27	1.63	1.10	1.50
P29	0.45	0.44	0.78
K30	1.20	1.00	0.90
F31	3.32	3.67	3.46
$\mathbf{sclx_4}$	2.90	3.30	2.98

Res.		**X-ray**[35]	**MD**	$\mathbf{Bound}^1_{Site2}$	$\mathbf{Bound}^2_{Site2}$	$\mathbf{Diff}_1$
K27	d_{S-NZ}	5.10	5.34 ± 2.46	3.90	3.91	4.20
	d_{C-NZ}	6.81	6.28 ± 0.43	6.81	6.40	6.23
K30	d_{S-NZ}	3.90	4.27 ± 1.02	3.80	3.90	3.44
	d_{S-NH}	2.70	4.61 ± 1.62	3.60	4.13	4.10
	d_{C-NZ}	4.87	5.00 ± 0.82	4.90	4.80	5.64
F31	d_{S-NH}	3.34	5.10 ± 1.21	4.00	5.86	4.25
$\mathbf{sclx_4}$	Θ	140.68	123.90 ± 13.10	118.00	126.25	114.65

host-guest–type protein-ligand interactions[48] for ligands binding surface-exposed residues, whereas the MMPBSA approach suffers from a well-known overshooting.

Competition with salt bridges and $\mathbf{sclx_4}$ binding allosteric implications

Surface lysine residues are decisive to strongly bind $\mathbf{sclx_4}$ at the protein surface. Their solvent-accessible surface area differs strongly amongst the 13 lysines (Figure S12) and impacts the propensity to bind calixarenes: K30 is the most solvent-exposed lysine residue (108.05 ± 0.58 Å^2), whereas other lysines are significantly more buried. The dynamic salt-bridge formation–deformation equilibrium is also a factor tuning the $\mathbf{sclx_4}$ binding. As surmised experimentally[35], our MD simulations corroborate that salt bridges modulate the lysine availability in solution. For instance, the K34–K35 tweezer, which is a favorable motif for calixarenes toward cytochrome C, is not identified as a favorable binding site for PAF. It adopts a V-shaped conformation where the two side chains extremities are distant by 13.5 ± 0.9 Å and only weakly solvent-exposed. Our *PAF*–$\mathbf{sclx_4}$ unbound MD simulations provide an exhaustive view of the dynamics of salt bridges in absence of $\mathbf{sclx_4}$ (see section "Salt Bridge, RMSF and SASA Analysis" in SI). Salt bridges are identified between the solvent-exposed K2 and the D46 and D53 carboxylate groups, and a HB between K6 terminus and the N40 side-chain amide carbonylic group (see Figure S10). -NH_3^+ groups of K9 and K11, present in a flexible loop, are in equilibrium with the peptidic oxygen of T37 and with E13 carboxylate, respectively. In **Site 1**, N40 provides a coordinating point for K6 by its side-chain amide carbonylic group, while E23 acts as a bridge between K15 and K17 on the structured β–sheet. A flexible loop, as the one where **Site 2** is located, allows an higher flexibility of the lysines side–chains. Close to this site, K34 and K35 form stable interactions with E13, D53 and D32, hence preventing K34–K35 to act as a tweezer. Their vicinity to

Site 2, as well as the presence of K27, makes the region lysine-rich and potential anchor point for $\mathbf{sclx}_4$. The binding of the anionic $\mathbf{sclx}_4$ ligand, as noticed previously for its interaction with cytochrome C[48], leads to a global rearrangement of the protein salt-bridge network all around the binding site and, perhaps more surprinsingly, allosteric effects in specific protein regions opposite to it. The region, highlighted by the residues 14 to 24 and distinct from **Site 1** (demarcated by K6, K15 and K17) denotes a behavior sensitive to the $\mathbf{sclx}_4$ binding in **Site 2**, undergoing an overall reduction of flexibility as pointed out by the degree of RMS fluctuations with respect to the protein unbound system (see Figure S11). For instance, Figure 4 shows that Y3 and Y16 residues span a larger conformational space in absence of $\mathbf{sclx}_4$ (blue lines), underlining that the transient T-shaped conformation between Y3 and Y16 is absent when the calixarene is bound. Beyond the evident electrostatic complementarity, such a macrocyclic ligand hence also acts by modifying specifically some remote protein regions. This allosteric outcome is probably a key feature to tune protein–protein contacts, and thus favour the formation of more complex protein aggregates in solution towards crystallization and possibly in the solid state by surface entropy reduction[63].

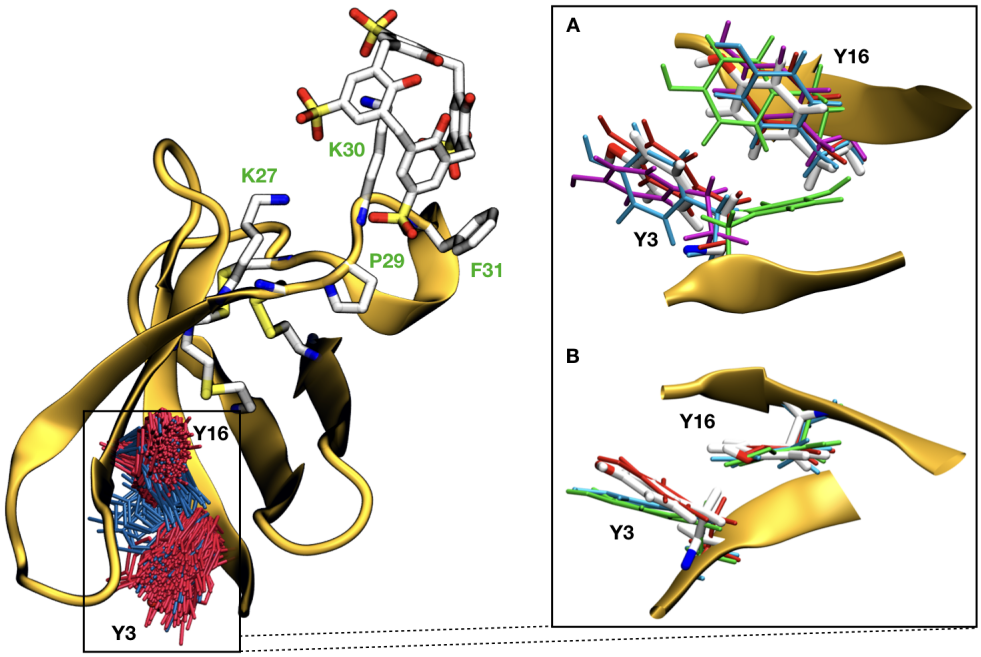

Figure 4: **Left panel**: *superimposition of Y3 and Y16 spatial fluctuations during the "Bound State" simulations of* $\mathbf{sclx}_4$ *interacting in* **Site 2** *(*$Bound^2_{Site2}$*, red lines) and PAF simulation unbound (blue lines) on the X-ray structure*[35] *(PAF in orange cartoon,* $\mathbf{sclx}_4$ *and residues in* **Site 2** *in white licorice tubes).* **Right panel**: *zoom on Y3/Y16 region reports structures obtained by cluster analysis (considering the two residues Y3 and Y16 as cluster center) for the "Bound State" (panel A) and PAF simulation unbound (panel B). The most representative structure is reported in white lines, while the other representative structures in colored lines.*

Conclusions

The properties of small ligands to tune protein–protein interfaces have been investigated computationally taking advantage of available experimental data. In this work, the interaction of $\mathbf{sclx}_4$ as an organic macromolecule with a small cationic antifungal protein (*PAF*) is investigated by MD simulations, and corroborates one major binding site featuring three lysines playing a central role in ligand binding, as already identified experimentally. Our MD simulations restores a dynamic view of the calixarene–ligand interaction beyond the X-ray

structure, but also allow to probe a second, also lysine-driven, stable binding site which can act as a transient site and was not identified experimentally. Beyond the sole structural and dynamical insights, a semi-quantitative per-residue energetic contribution allows to distinguish lysines and pinpoints the most important ones for the binding of calixarenes. Our computational approach predicts *de novo* the selective encapsulation of K30 and offers perspectives for a general mapping of calixarenes onto proteins, to decipher protein candidates and propose mutations. The agreement for the overall binding free energy from ITC experiments requires to rely on most rigorous schemes such the APR method. Finally, the $\mathbf{sclx}_4$ binding effect was also studied in relation to protein rearrangements and residues local movements. In fact, a specific surface region opposite to the binding site becomes more rigid, favoring the subsequent increase of protein–protein contacts for a more complex and controlled structure assembly according to the surface entropy reduction.

Acknowledgements

AB is grateful for the support from the University of Strasbourg. Computational HPC resources from the Mesocentre of the University of Reims, ROMEO, are gratefully acknowledged. ED is grateful for the support from the Institut Universitaire de France. AB would like to thank Gilberto Pereira for fruitful discussions.

Rendering

All pictures were rendered using Visual Molecular Dynamics (VMD)[64]. Figure 2C was rendered using Pymol[65].

References

(1) Sharma, V. K.; Patel, U.; Thakar, S.; Rathod, S. L.; Sharma, A.; Shrivastav, P. S.; Athar, M. Fluorescein appended calixarene functionalized supramolecular AIE-active liquid crystalline materials for self-assembly and bio-imaging applications. *New J. Chem.* **2022**,

(2) Salassi, S.; Simonelli, F.; Bartocci, A.; Rossi, G. A Martini coarse-grained model of the calcein fluorescent dye. *J. Phys. D: Appl. Phys.* **2018**, *51*, 384002.

(3) Zhao, Z.; Bourne, P. E. Harnessing systematic protein–ligand interaction fingerprints for drug discovery. *Drug Discovery Today* **2022**,

(4) Sabin, C.; Sam, S.; Hrishikes, A.; Salin, B.; Vigneshkumar, P.; George, J.; John, F. Supramolecular Drug Delivery Systems Based on Host-Guest Interactions for Nucleic Acid Delivery. *ChemistrySelect* **2022**, *7*, e202203644.

(5) Kjølbye, L. R.; Pereira, G. P.; Bartocci, A.; Pannuzzo, M.; Albani, S.; Marchetto, A.; Jiménez-García, B.; Martin, J.; Rossetti, G.; Cecchini, M.; Wu, S.; Monticelli, L.; Souza, P. C. T. Towards design of drugs and delivery systems with the Martini coarse-grained model. *QRB Discovery* **2022**, *3*, e19.

(6) Hecht, S.; Fréchet, J. M. Dendritic encapsulation of function: applying nature's site isolation principle from biomimetics to materials science. *Angew. Chem. Int. Ed.* **2001**, *40*, 74–91.

(7) Rodik, R.; Cherenok, S.; Kalchenko, O.; Yesypenko, O.; Lipkowski, J.; Kalchenko, V. Functional calixarenes for material and life science. *Curr. Org. Chem.* **2018**, *22*, 2200–2222.

(8) McGovern, R. E.; Fernandes, H.; Khan, A. R.; Power, N. P.; Crowley, P. B. Protein camouflage in cytochrome c–calixarene complexes. *Nature Chem.* **2012**, *4*, 527.

(9) Pompidor, G.; d'Aleo, A.; Vicat, J.; Toupet, L.; Giraud, N.; Kahn, R.; Maury, O. Protein crystallography through supramolecular interactions between a lanthanide complex and arginine. *Angew. Chem. Int. Ed.* **2008**, *120*, 3436–3439.

(10) Roux, A.; Talon, R.; Alsalman, Z.; Engilberge, S.; d'Aléo, A.; Di Pietro, S.; Robin, A.; Bartocci, A.; Pilet, G.; Dumont, E.; T., W.; Seigo, S.; Riobé, F.; Girard, E.; Maury, O. Influence of divalent cations in the protein crystallization process assisted by lanthanide-based additives. *Inorg. Chem.* **2021**, *60*, 15208–15214.

(11) Dumont, E.; Pompidor, G.; d'Aléo, A.; Vicat, J.; Toupet, L.; Kahn, R.; Girard, E.; Maury, O.; Giraud, N. Exploration of the supramolecular interactions involving tris-dipicolinate lanthanide complexes in protein crystals by a combined biostructural, computational and NMR study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 18235–18242.

(12) Denis-Quanquin, S.; Bartocci, A.; Szczepaniak, F.; Riobé, F.; Maury, O.; Dumont, E.; Giraud, N. Capturing the dynamic association between a tris-dipicolinate lanthanide complex and a decapeptide: a combined paramagnetic NMR and molecular dynamics exploration. *Phys. Chem. Chem. Phys.* **2021**, *23*, 11224–11232.

(13) Bijelic, A.; Rompel, A. The use of polyoxometalates in protein crystallography – An attempt to widen a well-known bottleneck. *Coord. Chem. Rev.* **2015**, *299*, 22–38.

(14) Engilberge, S.; Rennie, M. L.; Dumont, E.; Crowley, P. B. Tuning Protein Frameworks via Auxiliary Supramolecular Interactions. *ACS Nano* **2019**, *13*, 10343–10350.

(15) Crowley, P. B. Protein–Calixarene Complexation: From Recognition to Assembly. *Acc. Chem. Res.* **2022**, *55*, 2019–2032.

(16) Guérineau, V.; Rollet, M.; Viel, S.; Lepoittevin, B.; Costa, L.; Saint-Aguet, P.; Laurent, R.; Roger, P.; Gigmes, D.; Martini, C.; Huc, V. The synthesis and characterization of giant Calixarenes. *Nature Comm.* **2019**, *10*, 113.

(17) Giuliani, M.; Morbioli, I.; Sansone, F.; Casnati, A. Moulding calixarenes for biomacromolecule targeting. *Chem. Commun.* **2015**, *51*, 14140–14159.

(18) Perret, F.; Lazar, A. N.; Coleman, A. W. Biochemistry of the para-sulfonato-calix [n] arenes. *Chem. Commun.* **2006**, 2425–2438.

(19) Español, E. S.; Villamil, M. M. Calixarenes: Generalities and Their Role in Improving the Solubility, Biocompatibility, Stability, Bioavailability, Detection, and Transport of Biomolecules. *Biomolecules* **2019**, *9*, 90.

(20) Danil de Namor, A. F.; Cleverley, R. M.; Zapata-Ormachea, M. L. Thermodynamics of calixarene chemistry. *Chem. Rev.* **1998**, *98*, 2495–2526.

(21) Polli, F.; Cianfoni, G.; Elnahas, R.; Mangiardi, L.; Scaramuzzo, F. A.; Cammarone, S.; Quaglio, D.; Calcaterra, A.; Pierini, M.; Mazzei, F.; Zanoni, R.; Botta, B.; Ghirga, F. Resorc[4]arene modifiers for supramolecular site-directed immobilization of antibodies on multi-walled carbon nanotubes. *Chem. Bio. Chem.* **2023**, e202300030.

(22) Wang, L.; Li, L.-l.; Fan, Y.-s.; Wang, H. Host–guest supramolecular nanosystems for cancer diagnostics and therapeutics. *Adv. Mater.* **2013**, *25*, 3888–3898.

(23) Basilotta, R.; Mannino, D.; Filippone, A.; Casili, G.; Prestifilippo, A.; Colarossi, L.; Raciti, G.; Esposito, E.; Campolo, M. Role of Calixarene in Chemotherapy Delivery Strategies. *Molecules* **2021**, *26*, 3963.

(24) Pan, Y.-C.; Hu, X.-Y.; Guo, D.-S. Biomedical applications of calixarenes: state of the art and perspectives. *Angew. Chem. Int. Ed.* **2021**, *60*, 2768–2794.

(25) Guo, D.-S.; Liu, Y. Supramolecular chemistry of p-sulfonatocalix [n] arenes and its biological applications. *Acc. Chem. Res.* **2014**, *47*, 1925–1934.

(26) Siddharthan, A.; Kumar, V.; Barooah, N.; Chatterjee, S.; Bhasikuttan, A. C.; Mohanty, J. Supramolecular interaction of ofloxacin drug with p-sulfonatocalix [6] arene: Metal-ion responsive fluorescence behavior and enhanced antibacterial activity. *J. Mol. Liq.* **2022**, 121047.

(27) Yi, W. Y.; Supian, F. L.; Musa, M.; Abd Karim, N. F. N.; Naim, A. F. Calixarenes as Host Molecules for Drug Carriers in the Cosmetic and Medical Field. *Macromol. Res.* **2022**, 1–10.

(28) Buldenko, V. M.; Trush, V. V.; Kobzar, O. L.; Drapailo, A. B.; Kalchenko, V. I.; Vovk, A. I. Calixarene-based phosphinic acids as inhibitors of protein tyrosine phosphatases. *Bioorg. Med. Chem. Lett.* **2019**, *29*, 797–801.

(29) Xia, S.; Jiang, Y.; Guo, X.; Wang, Y.; Xu, W. Effects of p-sulfonatocalixarene and p-sulfonatocalixarene/sulfobetaine surfactant complex on the activities of bromelain and polyphenol oxidase. *J. Mol. Liq.* **2022**, *367*, 120398.

(30) Douteau-Guével, N.; Coleman, A. W.; Morel, J.-P.; Morel-Desrosiers, N. Complexation of basic amino acids by water-soluble calixarene sulphonates as a study of the possible-mechanisms of recognition of calixarene sulphonates by proteins. *J. Phys. Org. Chem.* **1998**, *11*, 693–696.

(31) McGovern, R. E.; McCarthy, A. A.; Crowley, P. B. Protein assembly mediated by sulfonatocalix [4] arene. *Chem. Commun.* **2014**, *50*, 10412–10415.

(32) Ramberg, K. O.; Engilberge, S.; Skorek, T.; Crowley, P. B. Facile Fabrication of Protein–Macrocycle Frameworks. *J. Am. Chem. Soc.* **2021**, *143*, 1896–1907.

(33) Doolan, A. M.; Rennie, M. L.; Crowley, P. B. Protein recognition by functionalized sulfonatocalix [4] arenes. *Chem. Eur. J.* **2018**, *24*, 984–991.

(34) Volpi, S.; Doolan, A.; Baldini, L.; Casnati, A.; Crowley, P. B.; Sansone, F. Complex Formation between Cytochrome c and a Tetra-alanino-calix [4] arene. *Int. J. Mol. Sci.* **2022**, *23*, 15391.

(35) Alex, J. M.; Rennie, M. L.; Engilberge, S.; Lehoczki, G.; Dorottya, H.; Fizil, A.; Batta, G.; Crowley, P. B. Calixarene-mediated assembly of a small antifungal protein. *IUCrJ* **2019**, *6*, 238–247.

(36) Guagnini, F.; Huber, A.; Alex, J. M.; Marx, F.; Crowley, P. B. Porous assembly of an antifungal protein mediated by zinc and sulfonato-calix [8] arene. *Journal of Structural Biology* **2021**, *213*, 107711.

(37) Nag, R.; Rao, C. P. Calixarene-mediated host–guest interactions leading to supramolecular assemblies: visualization by microscopy. *Chem. Comm.* **2022**, *58*, 6044–6063.

(38) Kim, H. J.; Lee, M. H.; Mutihac, L.; Vicens, J.; Kim, J. S. Host–guest sensing by calixarenes on the surfaces. *Chem. Soc. Rev.* **2012**, *41*, 1173–1190.

(39) Tian, H.-W.; Chang, Y.-X.; Hu, X.-Y.; Shah, M. R.; Li, H.-B.; Guo, D.-S. Supramolecular imaging of spermine in cancer cells. *Nanoscale* **2021**,

(40) Masoumifeshani, E.; Chojecki, M.; Rutkowska-Zbik, D.; Korona, T. Association Complexes of Calix [6] arenes with Amino Acids Explained by Energy-Partitioning Methods. *Molecules* **2022**, *27*, 7938.

(41) Leśniewska, B.; Coleman, A. W.; Suwińska, K. Host–Guest Complexation of the para-Sulfonato-calix [4] arene Receptor with Biomolecules. *Cryst. Growth Des.* **2022**, *22*, 5947–5957.

(42) Cerofolini, L.; Ramberg, K. O.; Padilla, L. C.; Antonik, P.; Ravera, E.; Luchinat, C.; Fragai, M.; Crowley, P. B. Solid-state NMR â a complementary technique for protein framework characterization. *Chem. Commun.* **2023**, *59*, 776–779.

(43) Binder, U.; Oberparleiter, C.; Meyer, V.; Marx, F. The antifungal protein PAF interferes with PKC/MPK and cAMP/PKA signalling of Aspergillus nidulans. *Mol. Microbiol.* **2010**, *75*, 294–307.

(44) Palicz, Z.; Gáll, T.; Leiter, É.; Kollár, S.; Kovács, I.; Miszti-Blasius, K.; Pócsi, I.; Csernoch, L.; Szentesi, P. Application of a low molecular weight antifungal protein from Penicillium chrysogenum (PAF) to treat pulmonary aspergillosis in mice. *Emerging microbes & infections* **2016**, *5*, 1–9.

(45) Giner-Llorca, M.; Del Sol, F. G.; Marcos, J. F.; Marina, A.; Manzanares, P. Rationally designed antifungal protein chimeras reveal new insights into structure-activity relationship. *Int. J. Biol. Macromol.* **2023**, *225*, 135–148.

(46) Kvansakul, M.; Lay, F. T.; Adda, C. G.; Veneer, P. K.; Baxter, A. A.; Phan, T. K.; Poon, I. K.; Hulett, M. D. Binding of phosphatidic acid by NsD7 mediates the formation of helical defensin–lipid oligomeric assemblies and membrane permeabilization. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 11202–11207.

(47) Bartocci, A.; Gillet, N.; Jiang, T.; Szczepaniak, F.; Dumont, E. Molecular Dynamics Approach for Capturing Calixarene–Protein Interactions: The Case of Cytochrome C. *J. Phys. Chem. B* **2020**, *124*, 11371–11378.

(48) Bartocci, A.; Pereira, G.; Cecchini, M.; Dumont, E. Capturing the Recognition Dynamics of para-Sulfonato-calix[4]arenes by Cytochrome c: Toward a Quantitative Free Energy Assessment. *J. Chem. Inf. Model.* **2022**, *62*, 6739–6748.

(49) Velez-Vega, C.; Gilson, M. K. Overcoming dissipation in the calculation of standard binding free energies by ligand extraction. *J. Comput. Chem.* **2013**, *34*, 2360–2371.

(50) Henriksen, N. M.; Fenley, A. T.; Gilson, M. K. Computational calorimetry: high-precision calculation of host–guest binding thermodynamics. *J. Chem. Theory Comput.* **2015**, *11*, 4377–4394.

(51) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Gilson, M. K. The SAMPL5 host–guest challenge: computing binding free energies and enthalpies from explicit solvent simulations by the attach-pull-release (APR) method. *J. Comput. Aided Mol. Des.* **2017**, *31*, 133–145.

(52) Heinzelmann, G.; Henriksen, N. M.; Gilson, M. K. Attach-pull-release calculations of ligand binding and conformational changes on the first BRD4 bromodomain. *J. Chem. Theory Comput.* **2017**, *13*, 3260–3275.

(53) Heinzelmann, G.; Gilson, M. K. Automation of absolute protein-ligand binding free energy calculations for docking refinement and compound evaluation. *Sci. Rep.* **2021**, *11*, 1–18.

(54) Case, D.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham III, T.; Cruzeiro, V.; Darden, T.; Duke, R.; Ghoreishi, D.; Gilson, M., et al. AMBER 2018; 2018. *University of California, San Francisco*

(55) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40*, W537–W541.

(56) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(57) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(58) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(59) 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*, 8577–8593.

(60) 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*, 9478–9508.

(61) Homeyer, N.; Gohlke, H. Free energy calculations by the molecular mechanics Poisson-Boltzmann surface area method. *Molecular Informatics* **2012**, *31*, 114–122.

(62) Wang, C.; Greene, D.; Xiao, L.; Qi, R.; Luo, R. Recent Developments and Applications of the MMPBSA Method. *Front. Mol. Biosci.* **2018**, *4*, 87.

(63) Cooper, D. R.; Boczek, T.; Grelewska, K.; Pinkowska, M.; Sikorska, M.; Zawadzki, M.; Derewenda, Z. Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr. D* **2007**, *63*, 636–645.

(64) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38.

(65) Schrödinger, L.; DeLano, W. PyMOL. *http://www.pymol.org/pymol* **2020**,