Supporting Information:
Probing the dynamical interaction of 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
List of Figures

S1  PAF-sclx\_4 Initial Placements ................................. 5
S2  “Diffusion” Simulations: RMSD and SASA Plots ......................... 7
S3  MMPBSA Per-residue decomposition $\Delta\Delta G$ ......................... 9
S4  Lysine-sclx\_4 Coordination .................................. 10
S5  “Bound-State” Simulations Most Representative Clusters .............. 12
S6  PAF-sclx\_4 Distance Plots .................................. 13
S7  PAF-sclx\_4’s sulphur atoms Distance Distributions ..................... 14
S8  *exo* and *endo* sclx\_4 Distributions .......................... 16
S9  APR Starting Structure ....................................... 19
S10 PAF-sclx\_4 Binding Free Energy path ............................ 20
S11 SASA .......................................................... 22
S12 Lysine salt-bridge distributions ................................ 23
S13 RMSF .......................................................... 24

List of Tables

S1  PAF-sclx\_4 Initial Simulations Boxes ................................. 5
S2  PAF-sclx\_4’s Sulphur Atoms Distance Values ......................... 15
S3  MD and X-ray Distances Comparison ................................. 17
Experimental binding affinity

The binding equilibrium of para-sulfonato-calix[4]arene (L) interacting with a small antifungal protein (PAF, P)

\[ P + L \xrightleftharpoons[K_D]{K_A} PL \]  

was experimentally characterized by Crowley and coworkers\(^1\) by ITC experiments, with a dissociation constant \( K_D \) of 107 µM and a corresponding binding affinity equal to:

\[ \Delta G^o = RT \ln(K_D) = -5.45 \text{ kcal} \cdot \text{mol}^{-1} \]  

“Diffusion” Simulations

System Setup and Production Phase

We have sought to probe the interaction of the *Penicillium* antifungal protein (PAF) (PDB: 6HA4, 55 residues, overall charge +5) interacting with one sclx\(_4\) molecule. In PAF-sclx\(_4\) system, the sclx\(_4\) ligand and the PAF protein (distance between the centers of mass) were solvated in four different initial boxes (Box\(_1\), Box\(_2\), Box\(_3\) and Box\(_4\)) with TIP3P\(^2\) water molecules. Number of water molecules and simulation box dimensions are reported in Table S1, as well as PAF-sclx\(_4\) relative centers of mass distances. The ligand molecule was not placed in contact with the protein, as we aim at probing *de novo* binding site(s).
Table S1: Number of water molecules ($N_{\text{water}}$), box volume (Box Vol.) and initial distances in Å between the centers of mass of PAF and sclx$_4$.

<table>
<thead>
<tr>
<th>System</th>
<th>Box</th>
<th>$N_{\text{water}}$</th>
<th>Box Vol. (Å$^3$)</th>
<th>Prot–Lig</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAF–sclx$_4$</td>
<td>Box$_1$</td>
<td>20459</td>
<td>90 X 90 X 90</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Box$_2$</td>
<td>25029</td>
<td>96 X 96 X 96</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Box$_3$</td>
<td>27546</td>
<td>99 X 99 X 99</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Box$_4$</td>
<td>24604</td>
<td>96 X 96 X 96</td>
<td>34</td>
</tr>
</tbody>
</table>

Figure S1: PAF-sclx$_4$ initial systems. The protein is reported in green cartoon. For Box$_1$, Box$_2$, Box$_3$ and Box$_4$ the ligand is colored in orange, purple, green and red, respectively. PAF cysteines disulphur bonds are shown in lines, while water and sclx$_4$ hydrogen atoms are omitted for the sake of clarity.

The production phase for each PAF–sclx$_4$ simulation box (Box$_1$, Box$_2$, Box$_3$ and Box$_4$) was carried out in five replicates of 300 ns each, giving 1.5 µs simulation time per box, and
in total 6 µs for the system in total.

Cluster Analysis

To obtain representative structures for the sclx₄ molecule interacting in the different binding sites, and a good starting point for bound-state and binding free-energy simulations, cluster analysis was performed by concatenating the last 100 ns of each trajectory reported in the previous section. Clustering of the MD trajectories was carried out using the hierarchical average-linkage clustering algorithm³ implemented in the cpptraj module of AmberTools18 with a random sieving frequency of 10 and a cutoff of 10 Å on the protein and ligand heavy atoms. The most representative structures of PAF-sclx₄ are shown in Figure 2 (panels A and B) in the main article.
RMSD Analysis

![SAS and RMSD plots](image)

Figure S2: sclx₄ SASA and RMSD plots in function of the simulation time for PAF-sclx₄ system. The last frame of each replicate (bound state) was taken as RMSD reference. SASA is reported in Å² (red lines), while RMSD in Å (green lines). The darker red and green colors, respectively, highlight the simulation part in which the post-process analysis was carried out (last 100 ns of each replicate). The unbound sclx₄ SASA is reported as a dashed line (902.00 ± 12.50 Å²) as obtained in ref.⁴.

MMPBSA Energy Analysis

The last 10 ns of each replicate run per system, concatenated into a 200 ns long trajectory, were analyzed using the MM-PBSA⁵-⁷ post-processing approach. A per-residue free energy additive decomposition analysis was carried out, with the aim of identifying the residues which mediate PAF-sclx₄ binding.

The values of the free energy of ligand binding ΔG_{bind} were calculated according to the
MM-PBSA equation S3:

\[ \Delta G_{\text{bind}} = \langle G_{\text{compl}} \rangle - \langle G_{\text{rec}} \rangle - \langle G_{\text{lig}} \rangle \]  

(S3)

where compl, rec, and lig stand for complex, receptor, and ligand, respectively. The free energy contribution of each species of the binding reaction was estimated as a sum of four terms as in eq. S4

\[ G = \langle E_{MM} \rangle + \langle G_{\text{psolv}} \rangle + \langle G_{\text{npsolv}} \rangle - T \langle S \rangle \]  

(S4)

where \( E_{MM} \) is the molecular mechanics energy of the molecule expressed as the sum of the internal energy of the molecule plus the electrostatics and van der Waals interactions in gas-phase, \( G_{\text{psolv}} \) is the polar contribution to the solvation free energy of the molecule (Poisson Boltzmann PB), \( G_{\text{npsolv}} \) is the non-polar contribution to the solvation free energy, T is the absolute temperature, and S is the entropy. The inclusion of the entropic contribution in eq. S4 was not considered here in the standard MM-PBSA calculations. As recommended in ref. 6, we used the default value for calculating nonpolar solvation free energies (inp=2) and for its associated variables cavity_surften and cavity_offset (0.0378 and -0.5692, respectively). A 1.4 Å probe radius (keyword prbrad, default value) has been used, as well as the optimized radii set (radiopt=1).
Figure S3: Per-residue decomposition $\Delta \Delta G$ of the total interaction energy estimated by the MM-PBSA approach, in kcal.mol$^{-1}$, as a function of the residue number. The colored boxes in red and green correspond respectively to Site 1 and Site 2. The vertical dashed green line correspond to K6, as part of Site 2.
Figure S4: Snapshots extracted from “Diffusion Simulations” to show supplementary lysines (K9, K11 and K38) capable of coordinating sclx₄, as well as those present in the main binding sites.

“Bound-State” Simulations

Setup

In the two “Bound-State” representative structures, extracted from cluster-analysis and reported in Figure 2A and B in the main article, sclx₄ interacts in Site 1 and Site 2. The two systems were solvated with 11708 and 10871 TIP3P² water molecules, leading to 86 X 86 X 86 and 76 X 76 X 76 Å³ simulation boxes, respectively. The systems were called Bound_site1 and Bound_site2.
“Bound-State” Cluster Analysis

Cluster analysis of “Bound State” simulations was performed on the entire trajectories (in total 3.5 µs) for each system, using the hierarchical average-linkage clustering algorithm via the cpptraj module of AmberTools18, with a random sieving frequency of 10 and a cutoff of 10 Å on the protein and ligand heavy atoms. The most representative structures of Bound\textsubscript{Site1} and Bound\textsubscript{Site2} are reported in Figure 3 in the main article and Figure S5.
Figure S5:  

A: most representative structure (80 %) of sclx$_4$ interacting with PAF in Site 1 obtained from “Bound State” simulations (Bound$_{Site 1}$, shown in Figure 2A in the main article).

B: superimposition of the first (Cluster 1, sclx$_4$ in Site 2, 51.4 %) and second (Cluster 2, sclx$_4$ in Site 1, 45 %) most representative structures obtained from cluster analysis of Bound$_{Site 2}$ system simulations. PAF is colored in orange cartoon; sclx$_4$ in Cluster 1, and the binding site residues K6, K15 and K17 in white licorice tubes, while sclx$_4$ in Cluster 2, and the binding site residues K27, P29, K30 and F31 in red lines.
Distance Analysis

Figure S6: Distance ($d, \text{Å}$) between the centers of mass of sclx$_4$ interacting in Site 1 (upper panel, Bound$_{Site1}$ system) and in Site 2 (lower panel, Bound$_{Site2}$ system) and the coordinating partners in function of the simulation time. For the entire binding site (Site(i)-sclx$_4$, light red line) the center of mass is defined by the coupled residues K27, P29, K30 and F31 for Site 1 and K6, K15 and K17 for Site 2, while it corresponds to the center of mass of the specific lysine in the other cases.
Figure S7: Distance distribution (Prob, %) between \textit{sclx}_4's sulphur atoms and K27 (left panel) and K30 (right panel) -NH$_3^+$ groups (Upper panels) and between \textit{sclx}_4's sulphur atoms and K30 (left panel) and F31 (right panel) -NH amide groups (Lower panels), extracted from the “Bound State” simulations of \textit{sclx}_4 interacting in Site 1 (BoundSite1 system).
Table S2: Average distances, reported as probability distributions in Figure S7 and extracted from the “Bound State” simulations of sclx₄ interacting in Site 1 (Boundₜₙ₁ system), between the nitrogen atom of the -NH₃⁺ group and the sclx₄ sulphur atoms (dₛ₋ₐₙ) and between the amide nitrogen atom and the sclx₄ sulphur atoms (dₛ₋ₐₙ). S_CM refers to consider the center of mass of the four sulphur atoms instead of a single sulphur. Distance values are given in Å.

<table>
<thead>
<tr>
<th>Res.</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
<th>S_CM</th>
</tr>
</thead>
<tbody>
<tr>
<td>K27</td>
<td>dₛ₋ₐₙ</td>
<td>13.10 ± 2.27</td>
<td>9.71 ± 3.00</td>
<td>5.34 ± 2.46</td>
<td>10.13 ± 1.85</td>
</tr>
<tr>
<td>K30</td>
<td>dₛ₋ₐₙ</td>
<td>4.27 ± 1.02</td>
<td>4.28 ± 0.84</td>
<td>6.76 ± 1.15</td>
<td>6.44 ± 1.25</td>
</tr>
<tr>
<td></td>
<td>dₛ₋ₐₙ</td>
<td>8.68 ± 1.04</td>
<td>8.38 ± 1.51</td>
<td>5.48 ± 1.03</td>
<td>4.61 ± 1.62</td>
</tr>
<tr>
<td>F31</td>
<td>dₛ₋ₐₙ</td>
<td>7.63 ± 1.46</td>
<td>8.51 ± 1.20</td>
<td>7.20 ± 1.40</td>
<td>5.10 ± 1.21</td>
</tr>
</tbody>
</table>

Interacting Surface of sclx₄

The angle Θ is formed between the centers of mass of the protein (P_CM), sclx₄ sulfonate’s sulphur (S_CM) and hydroxyl oxygen atoms (O_CM). The normalized distribution, extracted from the “Bound State” simulations, is reported in Figure S8. The main exo and endo conformations have an average Θ < 50° and ≃ 125°, respectively.
Figure S8: A schematic visualization of $\Theta$ angle is reported in the upper panel. In the lower panel, normalized distributions of $\Theta$ angle (in degrees $^\circ$), formed between the centers of mass of the protein, $\text{sclx}_4$ sulfonate’s sulphur and hydroxyl oxygen atoms. $\text{sclx}_4$ interacting in Site 1 ($\text{Bound}_{\text{Site1}}$ system) and in Site 2 ($\text{Bound}_{\text{Site2}}$ system) are reported as red and green lines, respectively.
Table S3: Distances (in Å) between the nitrogen atom of the -NH$_3^+$ group and the closest sclx$_4$ sulphur atom ($d_{S-NZ}$) and between the amide nitrogen atom and the closest sclx$_4$ sulphur atom ($d_{S-NH}$), and $\Theta$ 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 sclx$_4$”). X-ray refers to the value in the crystallographic structure (asymmetric unit in Figure 1 in the main article), MD to the average value obtained from the “Bound State” simulations, while Diff$_1$, Bound$_{Site2}^1$ and Bound$_{Site2}^2$ to the values extracted from the structures shown in Figures 2A and 3B in the main article.

<table>
<thead>
<tr>
<th>Res.</th>
<th>X-ray$^1$</th>
<th>MD</th>
<th>Bound$_{Site2}^1$</th>
<th>Bound$_{Site2}^2$</th>
<th>Diff$_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>K27</td>
<td>$d_{S-NZ}$ 5.10</td>
<td>5.34 ± 2.46</td>
<td>3.90</td>
<td>3.91</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>$d_{C-NZ}$ 6.81</td>
<td>6.28 ± 0.43</td>
<td>6.81</td>
<td>6.40</td>
<td>6.23</td>
</tr>
<tr>
<td>K30</td>
<td>$d_{S-NZ}$ 3.90</td>
<td>4.27 ± 1.02</td>
<td>3.80</td>
<td>3.90</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>$d_{S-NH}$ 2.70</td>
<td>4.61 ± 1.62</td>
<td>3.60</td>
<td>4.13</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>$d_{C-NZ}$ 4.87</td>
<td>5.00 ± 0.82</td>
<td>4.90</td>
<td>4.80</td>
<td>5.64</td>
</tr>
<tr>
<td>F31</td>
<td>$d_{S-NH}$ 3.34</td>
<td>5.10 ± 1.21</td>
<td>4.00</td>
<td>5.86</td>
<td>4.25</td>
</tr>
<tr>
<td>sclx$_4$</td>
<td>$\Theta$ 140.68</td>
<td>123.90 ± 13.10</td>
<td>118.00</td>
<td>126.25</td>
<td>114.65</td>
</tr>
</tbody>
</table>

Binding Affinity

Within the APR (attach-pull-release)$^8$–$^{11}$ protocol, the standard (absolute) binding free energy between the protein ($PAF$) and a ligand (sclx$_4$) is computed as the result of the reversible work of transferring the ligand from the binding site to solution (unbound state). The free energy difference is calculated as a sum of the work $W$ required to attach restraints on the ligand ($W_{attach}$), to pull the ligand away from the binding site ($W_{pull}$), and then to release the attached restraints and to place the ligand at the standard concentration ($W_{release-std}$).

$$\Delta G_{bind}^o = - (W_{attach} + W_{pull} + W_{release-std}) \quad (S5)$$
For further details on the protocol implementation, we refer the reader to the references\textsuperscript{8–11}.

**Simulations Setup**

As starting $PAF$-$sclx_4$ snapshot, the most representative structure obtained from cluster analysis of the “Diffusion” simulations, reported in Figure 2A in the main article and in which $sclx_4$ interacts in Site 1, was used. $PAF$-$sclx_4$ complex was solvated with 10000 TIP3P\textsuperscript{2} water molecules.

The protein is kept fixed and the ligand’s distance and angle restraints set up by the use of three non-interacting anchor dummy particles with zero charge, zero LJ radius and well-depth, mass of 220 Da and subject to positional restraints of 50 kcal/(mol Å\textsuperscript{2}) (see Figure S9). The force constants of the distance and angle restraints, applied between the anchor particles and $PAF$ and between the anchor particles and $sclx_4$ were, respectively, 5 kcal/(mol Å\textsuperscript{2}) and 100 kcal/(mol rad\textsuperscript{2}). The reaction coordinate adopted for the pulling phase is represented by the distance between a dummy particle (P1 in Figure S9) and the $sclx_4$ atom C18, and the pull force constant is 5 kcal/(mol Å\textsuperscript{2}).
Figure S9: APR starting structure for PAF-scl₄ system, with scl₄ interacting in Site 1. PAF is represented as orange cartoon, while the ligand molecule and the binding site protein residues in white licorice tubes. Dummy atoms (P1, P2 and P3) are reported as red spheres, while the protein (C7, Y16 and V52 Cα carbons) and the ligand anchor’s atoms (C18 and C20) as green and cyan spheres, respectively.
Figure S10: $W_{\text{pull}}$ is represented by the potential of mean force (PMF) along the coordinate reaction, while $W_{\text{release-std}}$ corresponds to the work of releasing the ligand at the standard concentration (1M) and it is evaluated semi-analytically. The x-axis is defined as the pulling reaction, but the pull is shown starting at 1 Å. The points from 0 to 1 Å must be considered as the attaching phase $W_{\text{attach}}$ values (increasing of the constraints values).
Salt Bridge, RMSF and SASA Analysis

PAF unbound was simulated in a box of 72 X 72 X 72 Å³, solvated by 9778 TIP3P water molecules. Cl⁻ counterions were added to neutralize the system total charge. The system was 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 ($\gamma_{coll} = 1 \text{ ps}^{-1}$) 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 ($P = 1 \text{ atm}$ and $T = 300 \text{ K}$). During these phases, position restraints were applied on the protein heavy atoms, with a force constant of 5 kcal/mol·Å² during the minimization and heating processes and of 2.5 kcal/mol·Å² during the equilibration. A subsequent equilibration of 1ns followed with no position restraints applied. The production run was carried out for 1 μs, while for the analysis the first 100 were discarded.
Figure S11: SASA (in Å²) obtained from PAF unbound simulation. The lines indicate the average obtained by block-averaging the simulations: PAF unbound was decomposed in 3 blocks of 300ns each, while the first 100 ns were discarded. The filled blue light area indicate the corresponding standard deviations.
Figure S12: Lysine salt-bridge distributions (Prob, %) extracted from PAF unbound simulation (see text for details).
Figure S13: RMSF (in Å) obtained from PAF unbound simulation (PAF, blue line) and PAF-sclx$_4$ system (red line), with sclx$_4$ interacting in Site 1. The lines indicate the average obtained by block-averaging the simulations: PAF unbound was decomposed in 3 blocks of 300ns each, while PAF-sclx$_4$ in 3 blocks of 500 ns and 2 blocks of 1 µs. The filled blue and red light areas indicate the corresponding standard deviations.
References


(8) Velez-Vega, C.; Gilson, M. K. Overcoming dissipation in the calculation of standard binding free energies by ligand extraction. *Journal of computational chemistry* 2013, 34, 2360–2371.

