Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Selectivity of pyoverdine recognition by the FpvA receptor of *Pseudomonas aeruginosa* from molecular dynamics simulations – Supporting Information

Benjamin Bouvier*, Christine Cézard, Pascal Sonnet

Laboratoire de Glycochimie, des Antimicrobiens et des Agroressources, CNRS FRE3517 / Université de Picardie Jules Verne – 1, rue des Louvels, 80037 Amiens Cedex 1, France.

Stability of the FpvA pore structure in water

To ascertain the stability of the β-barrel structure of the FpvA transporter in water, we measured the root mean square deviation (RMSD) of the barrel backbone to the experimental structure during the 50 ns production simulations of *apo*-FpvA, FpvA/PVDI and FpvA/PVD_{G173}, after fitting to the barrel α carbons to remove rigid-body rotation and translation (Figure S1). The small average mean square deviation from the experimental structure (ca. 1 Å) and the absence of baseline drift with time reveal the pore structure to be stable in the absence of the bacterial outer membrane, at least on the timescales explored in this study.

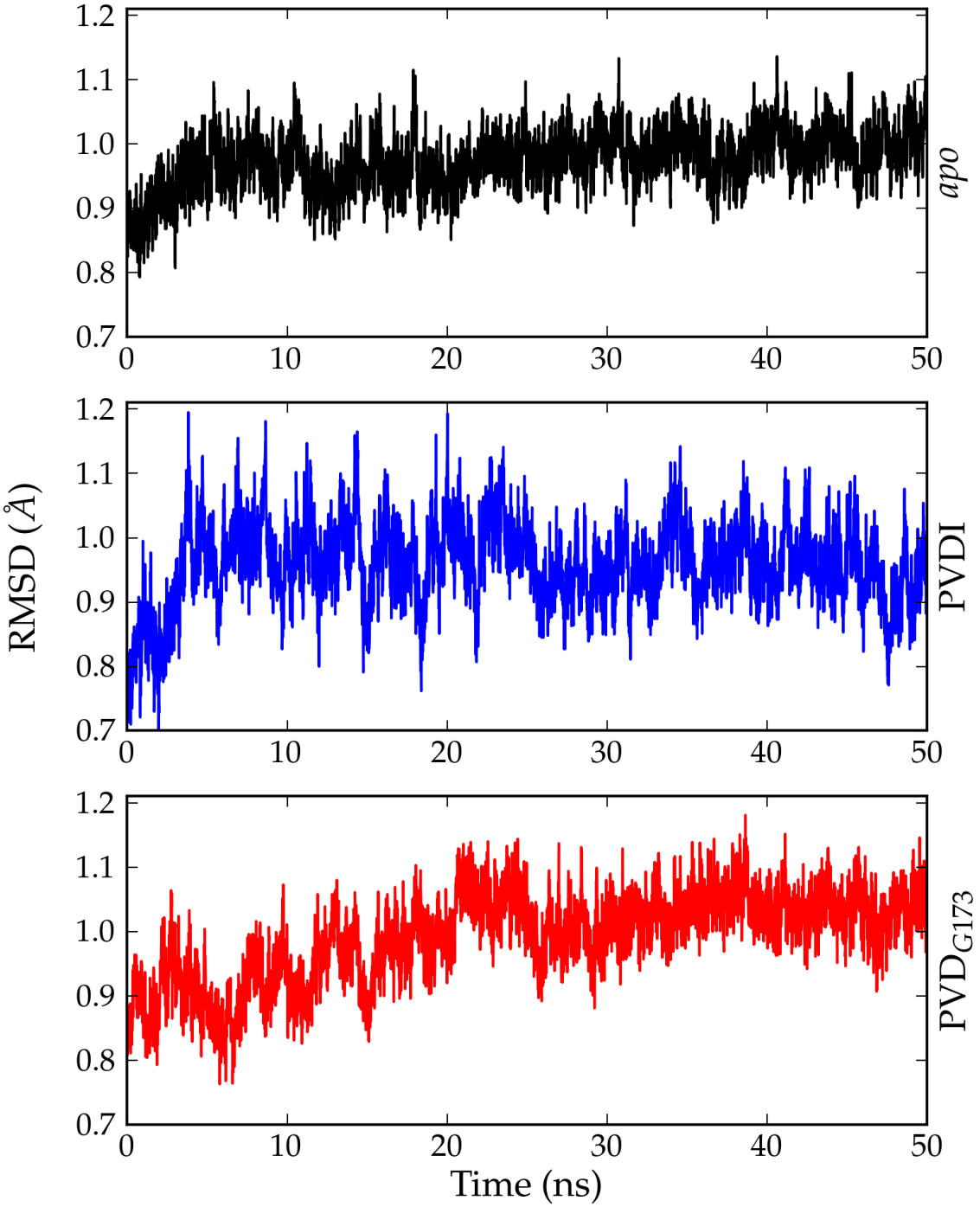

Fig. S1: RMSD of the β-barrel backbone of FpvA, in its *apo* form or in complex with PVDI or PVD_{G173}, to the corresponding experimental structure, as a function of simulation time.

Details of the alchemical decoupling simulations

Additional parameters and computational details

A total of 8 decoupling simulations were performed, corresponding to two siderophores (PVDI and PVD_{G173}), two environments (bound to FpvA or solvated in water), and two directions (creation or annihilation of interactions). The "alchemical" coordinate λ was discretized into a range of 13 different values (0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 1.0), where λ=0 (resp. λ=1) corresponds to the fully decoupled (resp. fully interacting) siderophore. Starting from either end of the λ range, 12 simulations were performed in opposite directions, to assess the free energy variation between λ values n and n+1 (or n and n-1). To avoid discontinuity issues and particle overlap around λ=0, a soft-core potential scheme was used,[1] and electrostatic (resp. Van der Waals) interactions were maintained fully decoupled (resp. fully coupled) for λ values lower than 0.05 (resp. greater than 0.95). For the FpvA/PVD complexes, each λ window lasted a total of 11 ns, of which 1 ns was discarded for equilibration, amounting to an effective 240 ns for each of the complexes. For the unbound solvated PVDs, each window lasted 9 ns of which the last 8 ns were used for the analysis, amounting to an effective total of 192 ns. The derivatives of the potential energy with respect to λ were recorded every 10 fs in each case, and were integrated using the Bennett acceptance ratio method.[2] The associated errors were computed using the ParseFEP plugin inside VMD.[3]

Assessment of convergence

The evolution of ΔG versus λ with the window length is presented on Figure S2 for the decoupling of PVDI and PVD_{G173} from water and the FpvA binding pocket; it is represented as its cumulated sum along the forward (λ=0 to 1) and backward (λ=1 to 0) transformations. Since the ΔG values are defined relative to an arbitrary reference, the final ΔG value for the forward path has been chosen to coincide with the initial ΔG value for the backward path at λ=1. The importance of hysteresis effects can hence be measured by comparing the ΔG values of both paths at λ=0. As can be seen, short simulation windows give rise to spurious hysteresis effects between the λ=0 to 1 and the λ=1 to 0 transitions; this issue vanishes for longer window lengths, guaranteeing that the overall ΔG values reported in this study are properly converged.

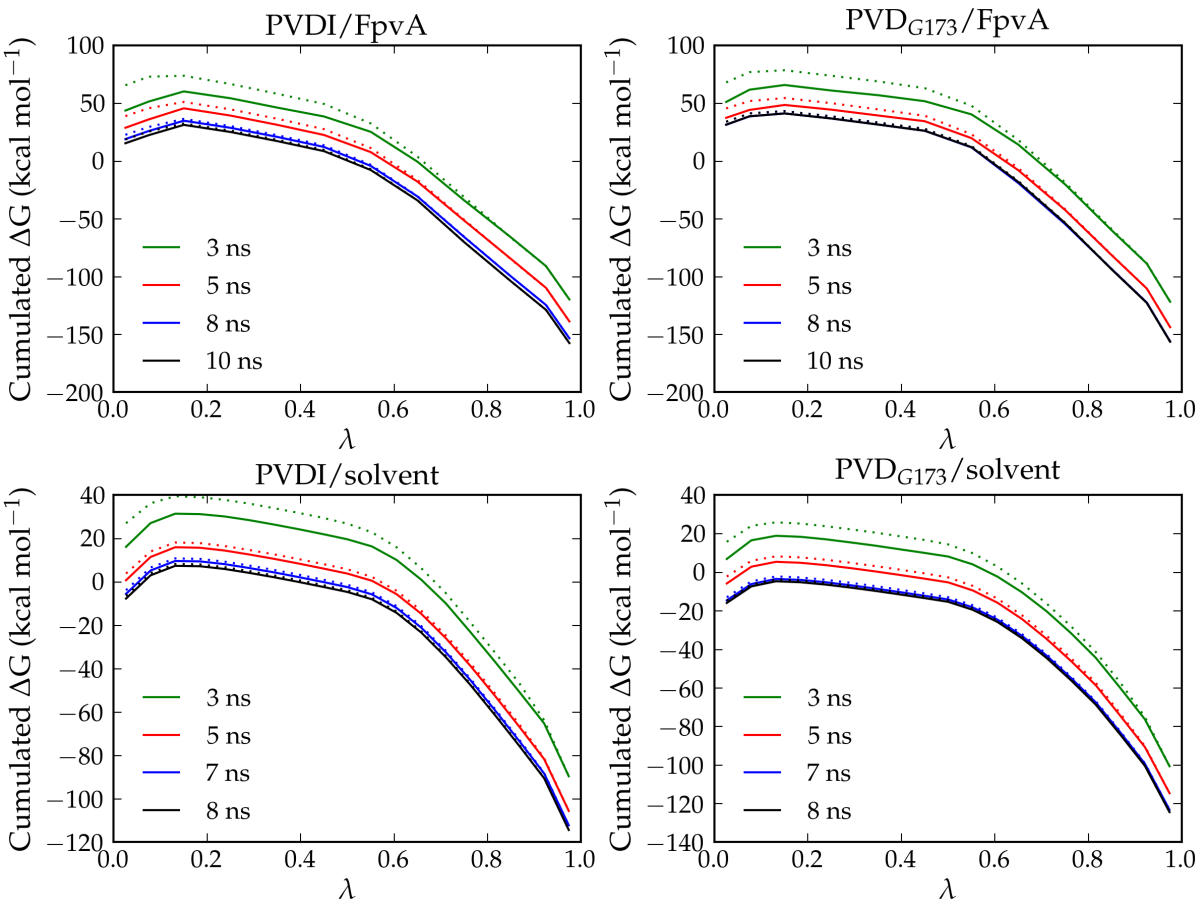

Fig. S2: Evolution of the cumulated ΔG along the coupling (forward, λ=0 to 1 – continuous lines) and decoupling (backward, λ=1 to 0 – dotted lines) pathways of the PVDI and PVD_{G173} siderophores to either the FpvA receptor binding pocket or to water with the length of individual simulation windows (curve colors).

Free energy profiles of FpvA/PVD dissociation: WHAM accuracy and consistency tests

Evaluation of statistical error

Zhu and Hummer[4] have shown that the cumulative statistical error for the potential of mean force G(x) along a reaction coordinate x, sampled using a series of umbrella windows i with harmonic biasing potentials of the form $K\left(x - r_i\right)^2$ centered at

$r_i = r_0 + i\Delta r$, can be expressed as the square root of the variance in the free energy estimator:

$$\mathrm{var}\left(G\left(x\right)\right) = \left(K\Delta r\right)^2 \sum_{i=1}^{\left(x-r_0\right)/\Delta r} \mathrm{var}\left(\overline{x}_i\right)$$

where $\mathrm{var}\left(\overline{x}_i\right)$ is the squared error in the estimate of the mean position of the reaction coordinate x in umbrella simulation i. This can be obtained from a straightforward block averaging of the value of x over the corresponding window:

$$\mathrm{var}\left(\overline{x}_i\right) = \frac{1}{n\left(n-1\right)} \sum_{b=0}^{n-1} \left[\frac{1}{m} \sum_{a=bm+1}^{(b+1)m} x_{ia} - \overline{x}_i \right]^2$$

Each umbrella window is split into n blocks of size m such that the averages of x over the blocks can be considered independent from each other. In the present cases, values of n=5-10 were found give the best results and an intermediate value n=8 was selected for the computation of the statistical error. The resulting error bars are shown on Fig. 3 in the main text of the article.

Checking for sampling consistency

The FpvA/PVD dissociation free energy profile obtained from the biased simulations using the weighted histogram analysis method (WHAM) is a function of the chosen reaction coordinate, with all other degrees of freedom integrated out. This implies that these orthogonal degrees of freedom are adequately sampled in each window, which might require much longer timescales than typical window lengths if they feature local minima separated by high free energy barriers. In addition, an individual window trajectory where the system remains in a minimum without visiting any other basins will not show symptoms of insufficient sampling, making the identification of the problem difficult. Zhu and Hummer[4] have proposed that insufficient sampling can be detected by checking the consistency of histograms in neighboring simulation windows: if different states of the orthogonal coordinates are visited in adjacent windows, inconsistent probability distributions of the reaction coordinate will ensue.

Consider a virtual simulation window halfway between two adjacent windows labeled 1 and 2, centered at $d^*=(d_1+d_2)/2$ and with biasing potential $E^*=k(d-d^*)^2$. The corresponding probability distribution $p^*(d)$ can be computed from both $p_1(d)$ and $p_2(d)$:

$$p_i^*(d) = \frac{p_i(d)\exp\left\{\left[E_i(d) - E^*(d)\right]/kT\right\}}{\int_{-\infty}^{+\infty} p_i(d)\exp\left\{\left[E_i(x) - E^*(x)\right]/kT\right\}dx}, \; i \in \{1,2\}$$

The inconsistency between the two distributions, based on a Kolmogorov-Smirnov test, can be written as:

$$\theta_{1,2} = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \max_d \left| \int_{-\infty}^{d} \left[p_1^*(x) - p_2^*(x) \right] dx \right|$$

where the number of independent samples N_1 and N_2 can be evaluated from the variance of the distribution of d and its average over subsets of the corresponding window trajectories.[4]

Results

Figure S3 presents graphs of the PMF, the statistical error and the consistency measure ϑ as a function of interpartner distance for the two FpvA/PVD complexes under study, using simulation windows of various lengths for WHAM. As can be seen, the PMF and statistical error are satisfactorily converged for windows longer than 5 ns. However, the inter-window inconsistency measure ϑ remains high until 6 ns in the 4.0-5.0 Å region, and to a lesser extent in the 3.0-3.5 Å region for PVDI. As explained in the main text, the conformational volume accessible to the ligand on the basis of rigid-body translation and rotation alone quickly increases with the separation distance, limiting the practicality of the minimum distance separation method at large interpartner distance; however, the contained increase seen for window lengths of 6 and 8 ns shows that sampling consistency is not an issue, in the relevant range of interpartner distances, for the complexes under study.

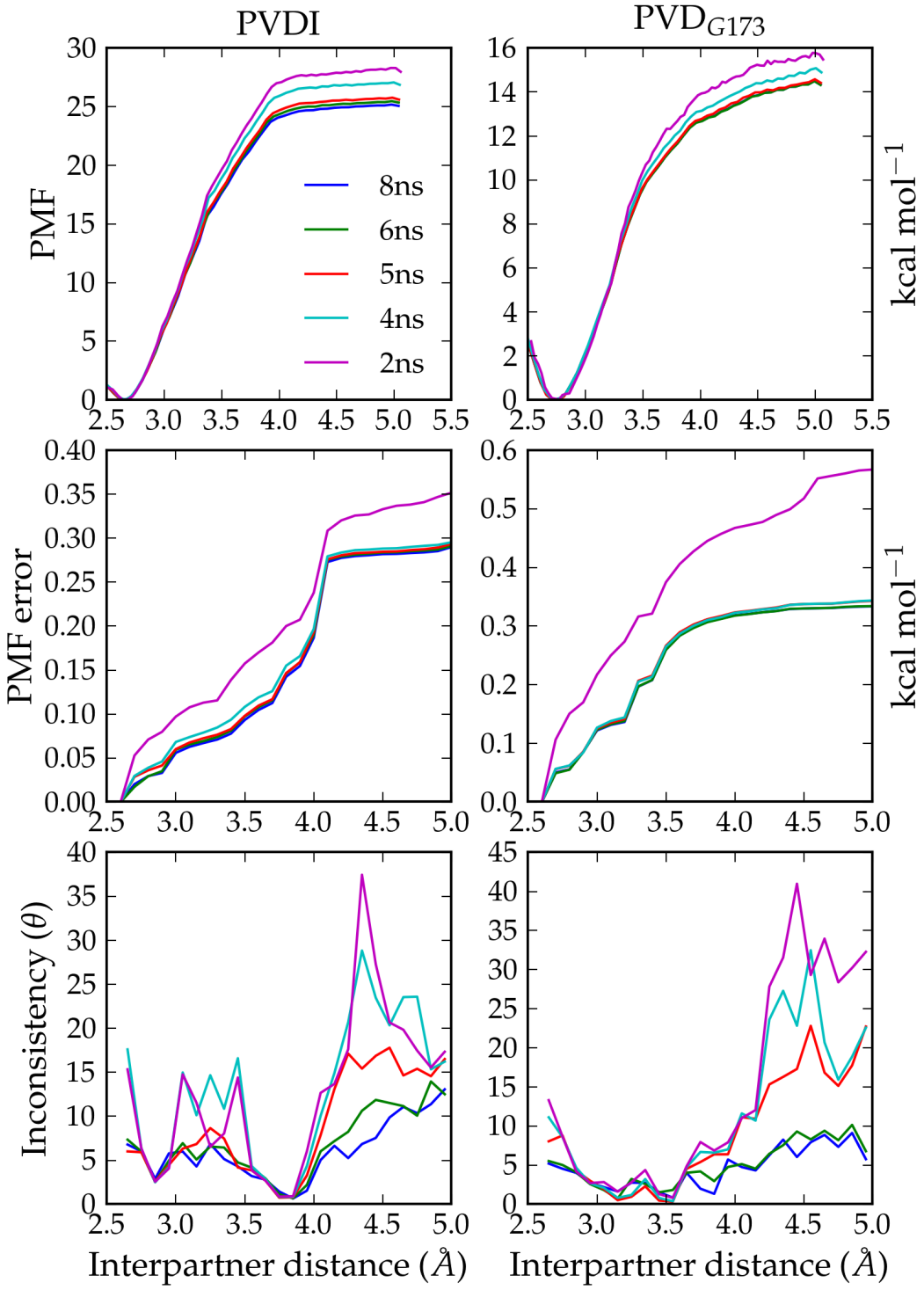

Fig. S3: PMF (top), PMF statistical error (middle) and inter-window inconsistency (bottom) as functions of the FpvA/PVD distance, for several lengths of individual umbrella windows. Left: PVDI; right: PVD$_{G173}$.

Miscellaneous

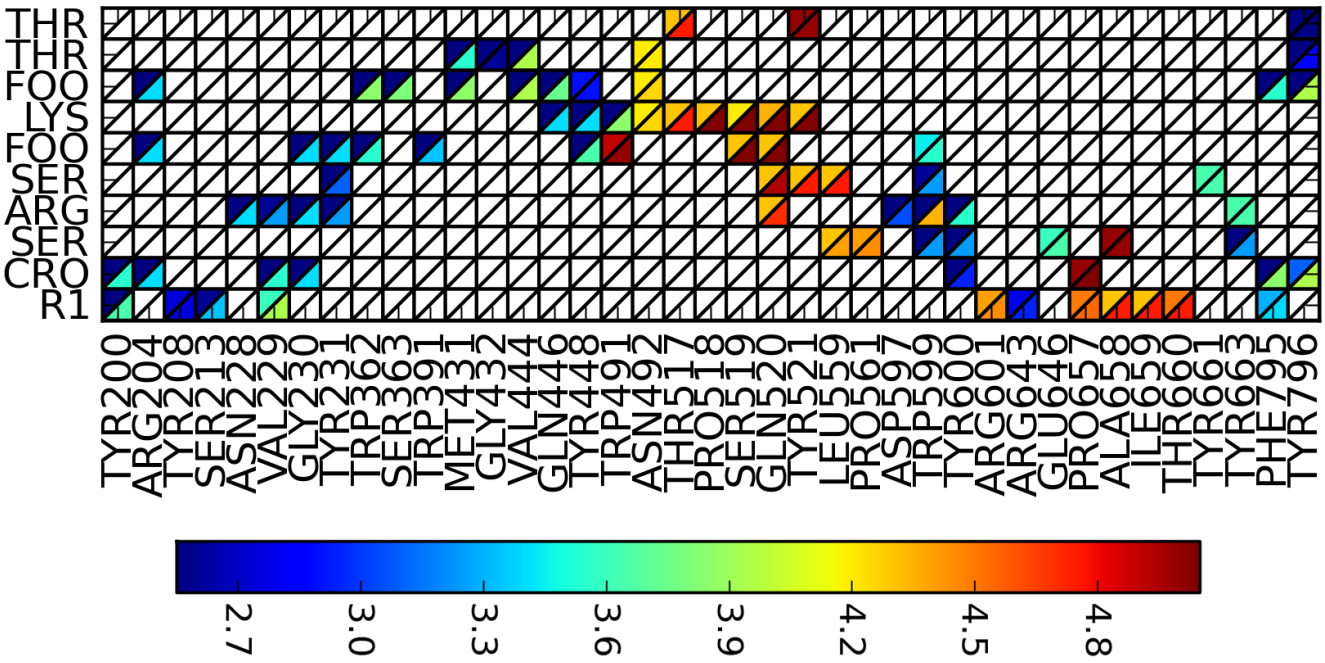

Fig. S4: Contact map along the binding/unbinding pathway of PVDI to FpvA. FpvA aminoacids are shown on the horizontal axis, PVDI residues on the vertical axis (CRO: chromophore; R1: chromophore dicarboxylic acid C3 substituent chain). The upper left (resp. lower right) triangle in each cell denotes the smallest (resp. largest) interpartner distance (color-coded, in Å) at which the corresponding contact is detected. A contact is considered present if it is found in at least 95% of the simulation snapshots having interpartner distances in a 0.05 Å window centered on the relevant distance.

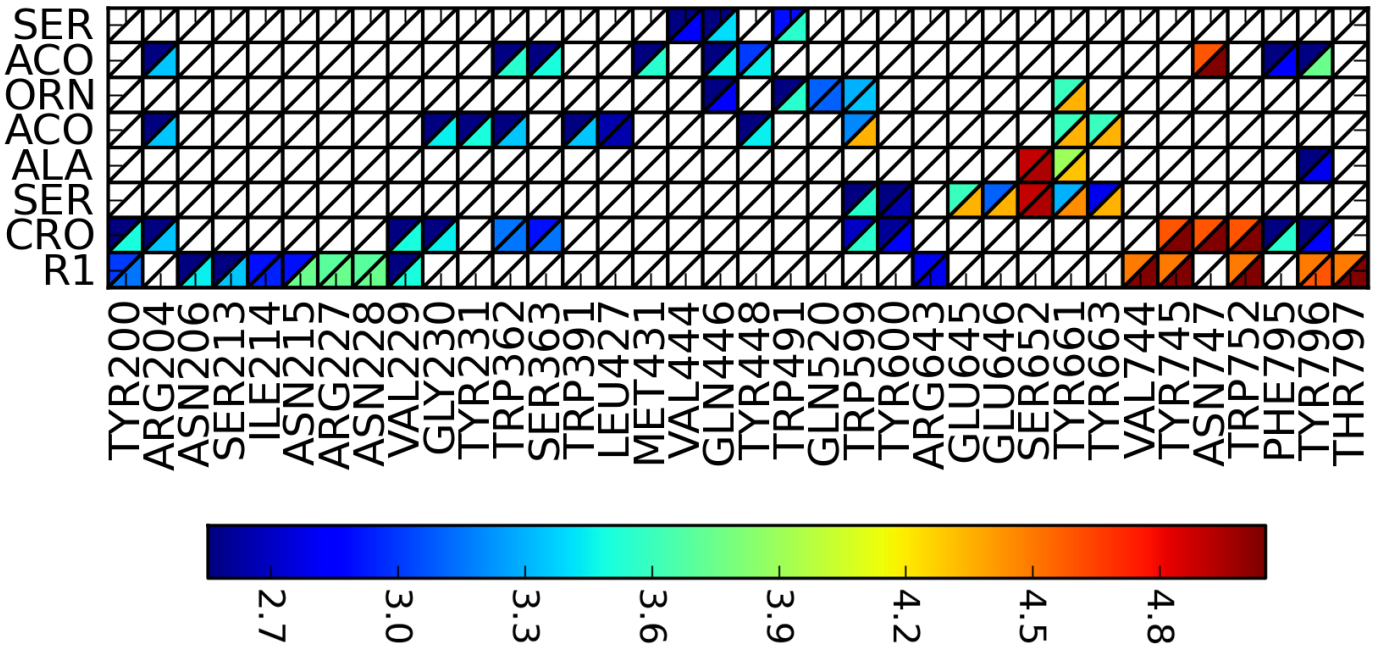

Fig. S5: Contact map along the binding/unbinding pathway of PVD_{G173} to FpvA. FpvA aminoacids are shown on the horizontal axis, PVD_{G173} residues on the vertical axis (CRO: chromophore; R1: chromophore dicarboxylic acid C3 substituent chain). The upper left (resp. lower right) triangle in each cell denotes the smallest (resp. largest) interpartner distance (color-coded, in Å) at which the corresponding contact is detected. A contact is considered present if it is found in at least 95% of the simulation snapshots having interpartner distances in a 0.05 Å window centered on the relevant distance.

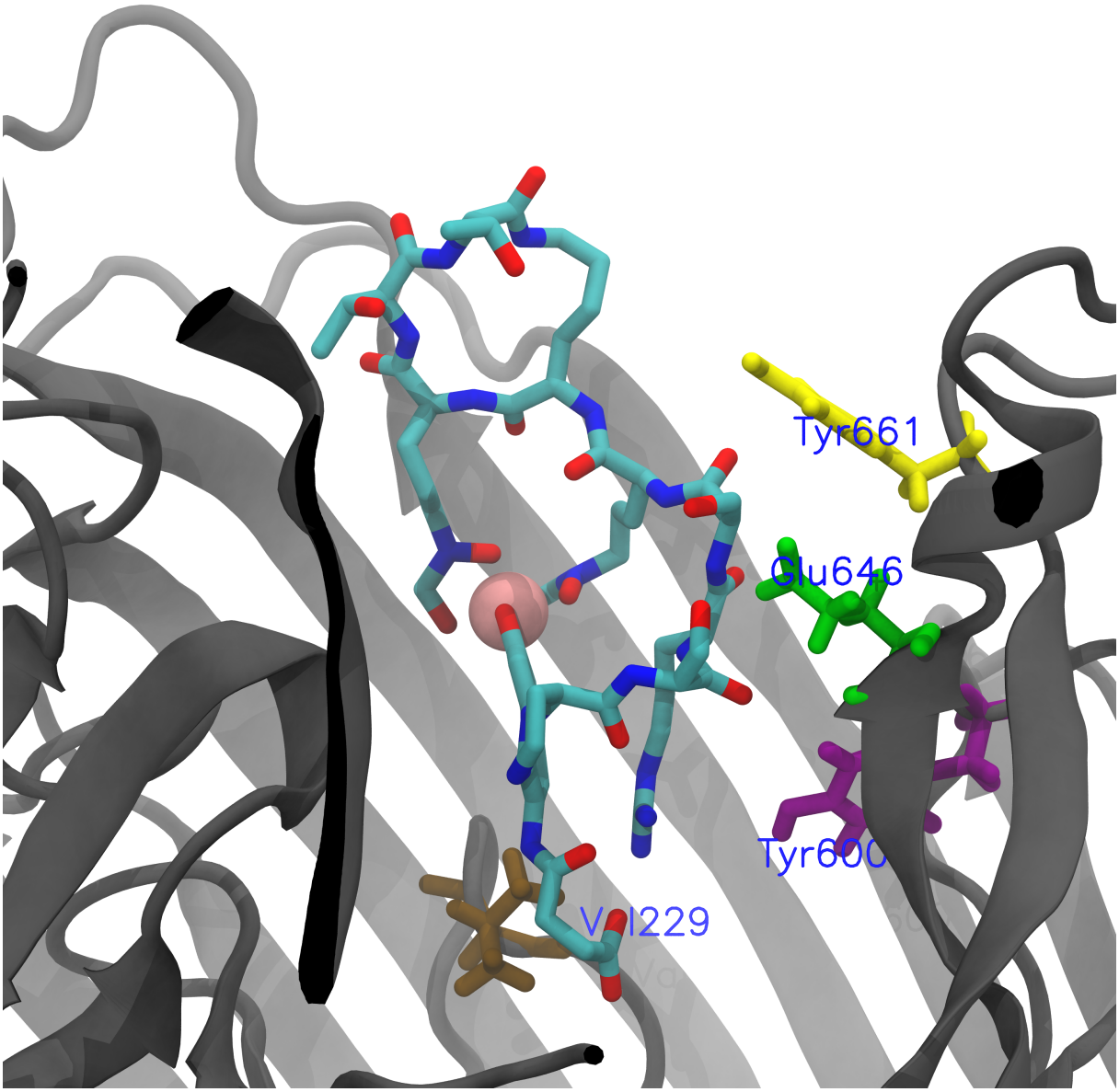

Fig. S6: Representation of the conformational bottleneck structure for PVDI (blue and red sticks, Fe^{3+} as pink sphere) along the FpvA (grey cartoon) binding/unbinding pathways. Important aminoacids on FpvA are represented as sticks and labeled on the figure. See main text for details.

Strain	Peptide chain sequence†	Exp. K_d (nM)‡	Docking score (kcal mol^{-1})
PA01	Ser–Arg–Ser–FoOHOrn–[Lys–FoOHOrn–Thr–Thr]	0.1	-10.9
G173	Ser–Ala–AcOHOrn–[Orn–Asp–AcOHOrn–Ser]	>10 000	-11.2
DSM50106	Ser–Lys–Gly–FoOHOrn–Ser–Ser–Gly–[Orn–FoOHOrn–Ser]	2.7	-9.0
ATCC13525	Ser–Lys–Gly–FoOHOrn–[Lys–FoOHOrn–Ser]	2.7	-9.9
Pfl18.1	Ser–Lys–Gly–FoOHOrn–Ser–Ser–Gly–[Lys–FoOHOrn–Ser]	0.65	-9.6
Pa6	Ser–Dab–FoOHOrn–Gln–Gln–FoOHOrn–Gly	>10 000	-9.2

†: square brackets delimitate cycles
‡: see table 1 in article[5] for references to the each experimental study

Table S1: Docking scores of PVDs from six different Pseudomonads to the PA01 FpvA receptor, along with the corresponding experimental dissociation constants.

References

1 M. Zacharias, T. P. Straatsma and J. A. McCammon, *J. Chem. Phys.,* 1994, **100**, 9025.

2 C. H. Bennett, *J. Comput. Phys.,* 1976, **22**, 245.

3 P. Liu, F. O. Dehez, W. Cai and C. Chipot, *J. Chem. Theory Comput.,* 2012, **8**, 2606.

4 F. Zhu and G. Hummer, *J. Comput. Chem.,* 2012, **33**, 453.

5 J. Greenwald, M. Nader, H. Celia, C. Gruffaz, V. Geoffroy, J. M. Meyer, I. J. Schalk and F. Pattus, *Mol.. Microbiol..,* 2009, **72**, 1246.