Supporting Information

Partial opening of cytochrome P450cam (CYP101A1) is driven by allostery and putidaredoxin binding

Simon P Skinner[1*], Alec H Follmer[2], Marcellus Ubbink[3], Thomas L Poulos[2,4,5], Jeanine J Houwing-Duistermaat[6] and Emanuele Paci[1*]

[1] School or Molecular and Cellular Biology and Astbury Centre, University of Leeds, Leeds LS2 9JT, UK

[2] Department of Chemistry, University of California, Irvine, California 92697-3900, United States

[3] Leiden University, Institute of Chemistry, Einsteinweg 55, 2333 CC Leiden Netherlands

[4] Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697-3900, United States

[5] Department of Pharmaceutical Sciences, University of California, Irvine, California 92697-3900, United States

[6] School of Mathematics, University of Leeds, Leeds LS2 9JT, UK

Simon P Skinner simon.p.skinner@gmail.com

Emanuele Paci e.paci@leeds.ac.uk

Methods

Structure based clustering

All structures of the P450cam/Pdx set and the 3cam set were clustered using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm as implemented in the CPPTRAJ software[1]. The minimum number of structures to form a cluster was 25 and the RMSD cut-off between points for forming a cluster, epsilon, was determined using a K-dist plot. A K-dist plot shows for each frame, the Kth farthest distance, sorted by decreasing distance. Epsilon was determined from the initial part where the curve begins to flatten out, indicating an increase in density.

Results

Assessing the validity of the Δχ tensors

The magnitudes of the axial and rhombic tensor components for the P450cam/Pdx set, fitted to PCS measured in the absence of Pdx and presence of 1 mM camphor, for the E195C/A199C/C334A mutant of P450cam linked to CLaNP7-(Yb^{3+}), were determined to be $7.6 \pm 0.1 \times 10^{-32}$ m^3 and $6.8 \pm 0.1 \times 10^{-32}$ m^3, respectively, and $7.1 \pm 0.1 \times 10^{-32}$ m^3 and $6.5 \pm 0.1 \times 10^{-32}$ m^3, for the corresponding tensor components for the 3cam set. Previously published values for tensor component magnitudes of Yb-CLaNP-7. show that the axial component can range from 4.2×10^{-32} m^3 to 6.4×10^{-32} m^3 and the rhombic component can range from 5.6×10^{-32} m^3 to 9.6×10^{-32} m^3. We therefore concluded that the tensor component magnitudes observed for both simulations are in line with previously published values for CLaNP-7 attached to P450cam, and other protein systems.

It has been previously shown for CLaNP-7, that the distances between the Cα atoms of the attachment site and the position of the lanthanoid ion must fall between 7 Å and 9 Å[2–4]. This is due to the steric constraints imposed by attachment of the probe to the two cysteines and geometry of CLaNP-7. These distances are critical for assessing the validity of a Δχ tensor fit, since deviation from this distance range would indicate a sterically impossible conformation of the CLaNP-7, the protein, or both. The distributions of attachment site distances within the P450cam/Pdx simulations are unimodal, and minima of the C195Cα-Ln and C199Cα-Ln distance both exceed 9 Å (Figure S3). On the other hand, average distances between the lanthanoid and the attachment site Cα atoms, for the 3cam simulation, were found to be 8.62

Å and 9.48 Å, for Cys195 and Cys199, respectively, which are less inconsistent with results of other studies using this and other CLaNP tags. The distance distribution of the paramagnetic center for the 3cam set contains physically feasible positions, whereas that of the P450cam/Pdx set does not.

To determine if a physically feasible tensor could be determined for the P450cam/Pdx set, a second fit was performed by restraining the position of the paramagnetic center to within 2 Å sphere. The center of this sphere was obtained by superposing the G helix the P450cam/Pdx set with the G helix of the closed form of P450cam (PDB code 3L63[5]), thereby translating the previously published lanthanoid position into the frame of the P450cam/Pdx set. The ensemble-averaged Qa score of this fit was determined to be 0.21, indicating a worse fit than with the lanthanoid unrestrained. Consequently, the PCS predicted for the P450cam/Pdx set were not analyzed further.

Model-based clustering

The BIC's for models with four, five and six clusters are 422347.8, 422347.8 and 432358.8, showing that a model with six clusters provides the best fit. The distance between the cluster means varied from 0.02 to 0.49, for a model with six clusters, from 0.02 to 0.46 for a model with five clusters and from 0.11 to 0.39 for a model with four clusters. Due to the fact that the smallest mean distance range was observed for the model with four clusters, this was chosen as the most representative of the observed PCS.

Cluster	$\Delta\chi_{ax}$ ($\times 10^{-32}$ m^3)	$\Delta\chi_{rh}$ ($\times 10^{-32}$ m^3)	Number of members
1	7.7 ± 0.04	7.9 ± 0.09	1306
2	7.1 ± 0.01	6.0 ± 0.03	990
3	6.3 ± 0.03	5.6 ± 0.02	1889
4	11.2 ± 0.20	9.1 ± 0.20	815

Table S1 Magnitudes of the axial and rhombic components of the magnetic susceptibility tensors fitted for each cluster of the 3cam set

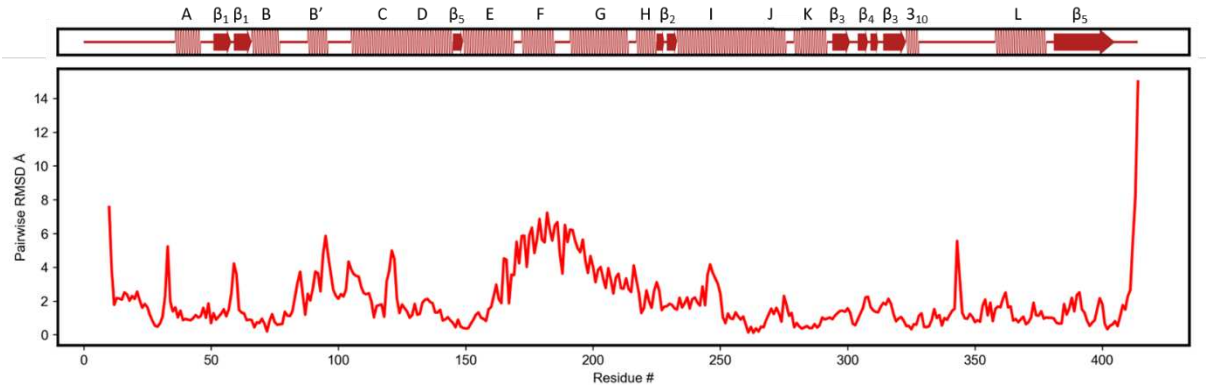

Figure S1 Pairwise RMSD between the Cα atoms of the cluster centroids of the P450cam/Pdx set and the 3cam set.

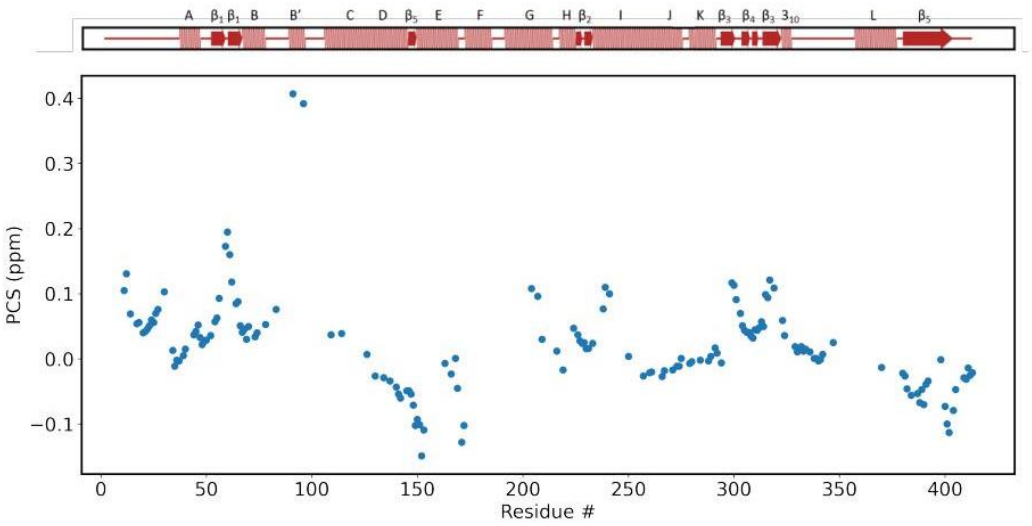

Figure S2 Observed PCS for the E195C/A199C mutant of P450cam with Yb-CLaNP-7 attached.

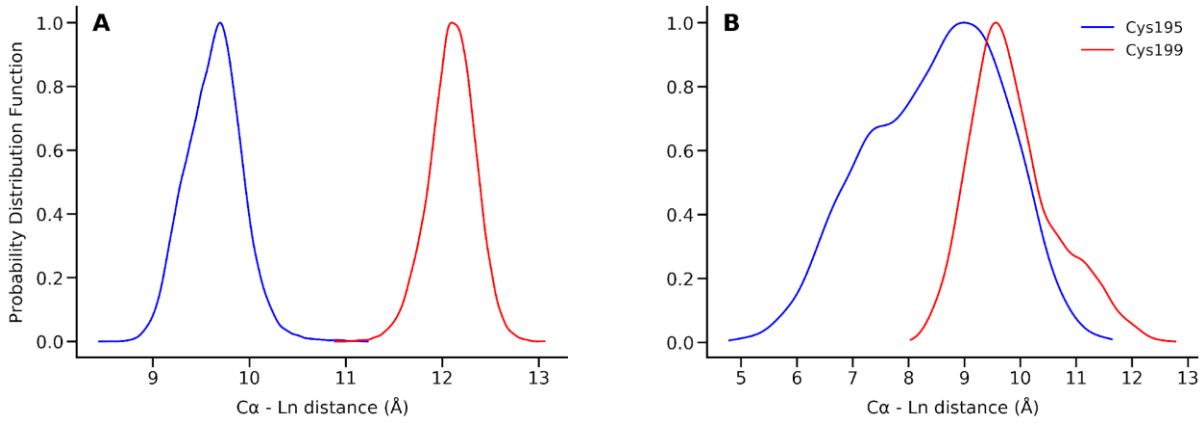

Figure S3 E195Cα-Ln and A199Cα-Ln distance distributions measured for the P450cam/Pdx (A) and 3cam sets (B). Each distance distribution was fit to a Gaussian KDE using "Silverman's rule of thumb" for bandwidth selection.

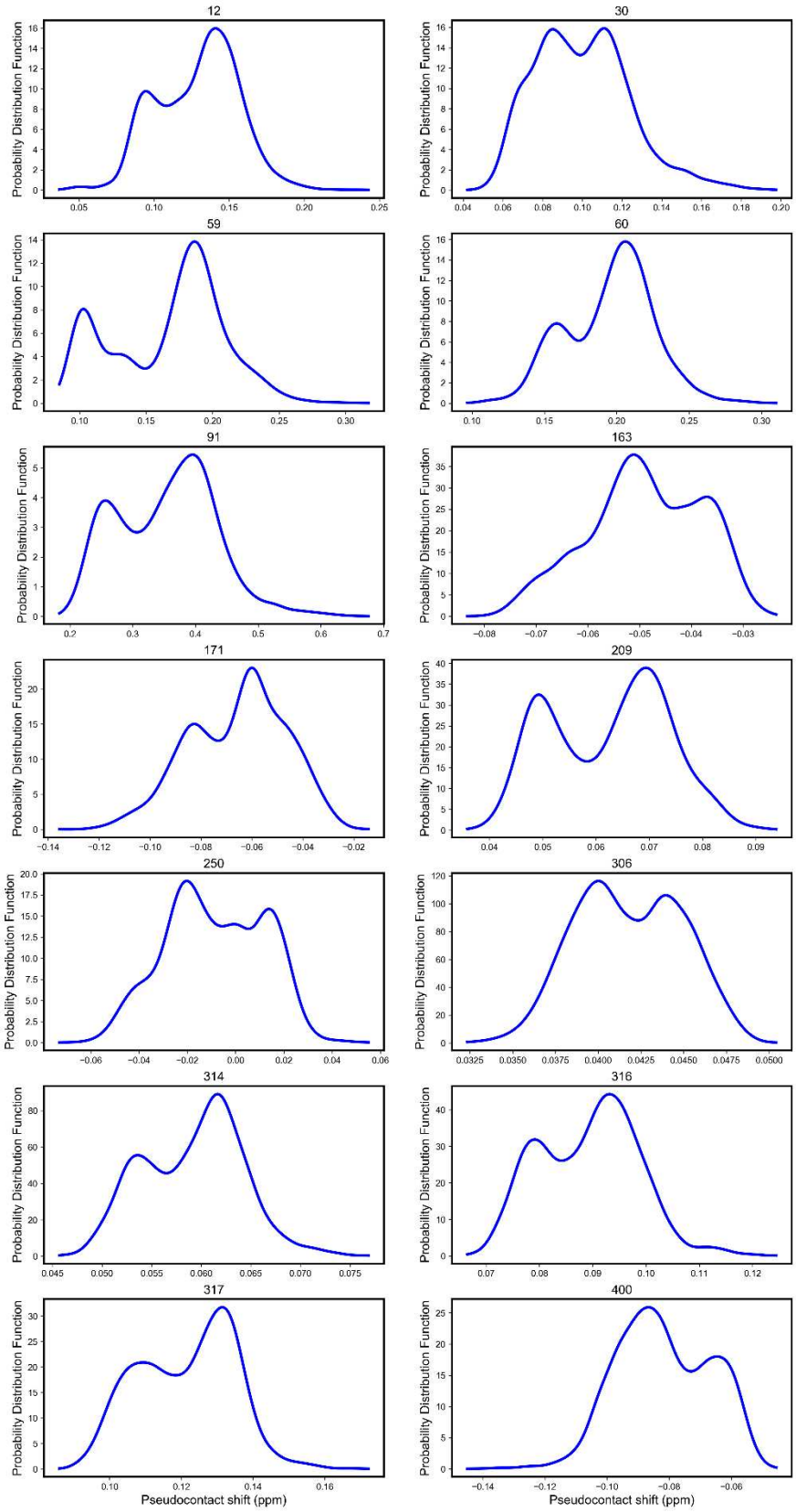

Figure S4 PCS distributions of the fourteen residues that multimodal distribution of PCS.

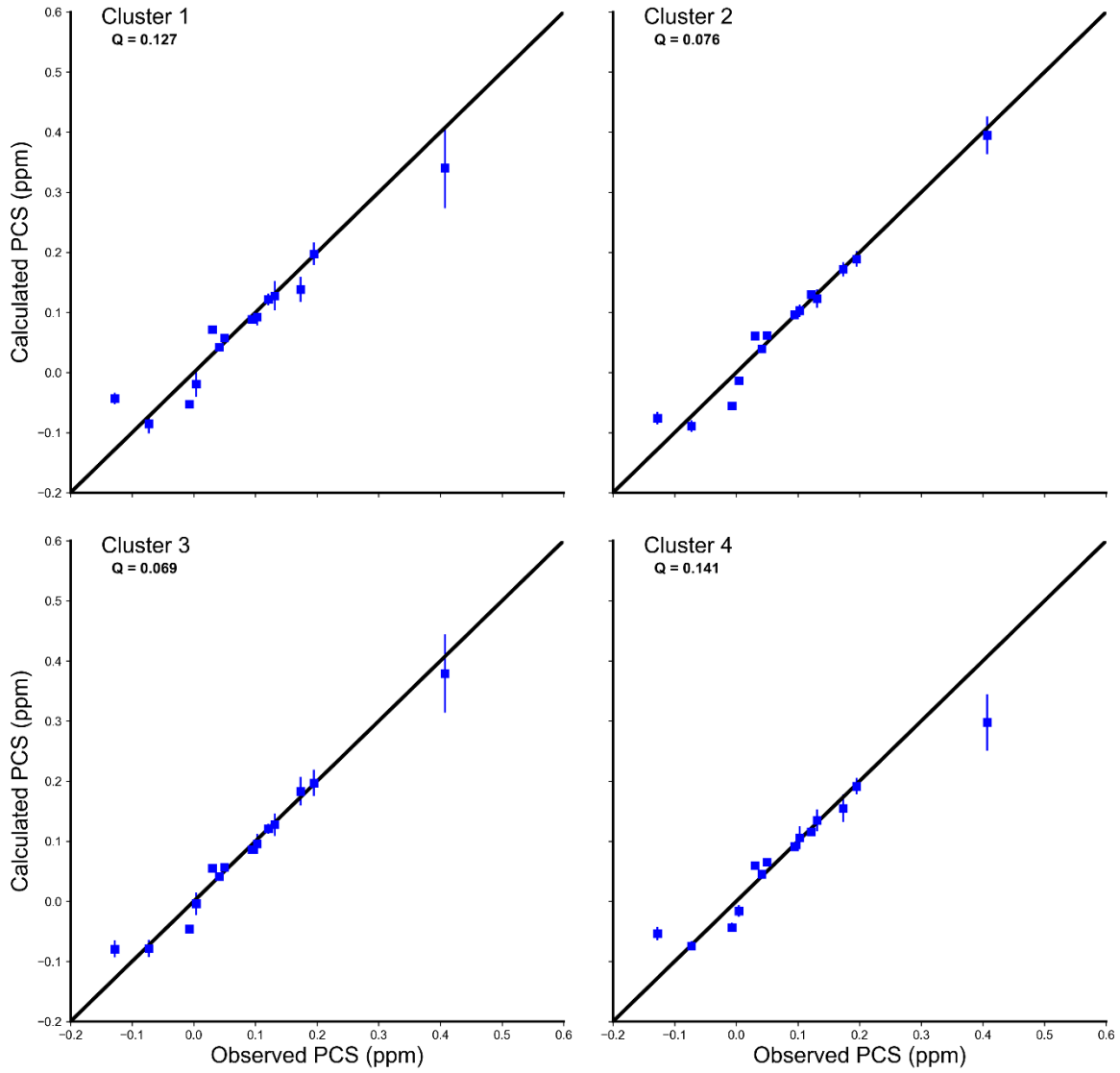

Figure S5 Plot of average calculated vs. observed PCS for the 14 residues identified by diptest for the four clusters of the 3cam ensemble. The errors in the observed PCSs were estimated to be ±0.008 ppm and the errors in the calculated PCS are the standard deviations of the predicted PCS determined from all simulation frames. Qa scores were calculated using the equation of Ubbink and co-workers[6].

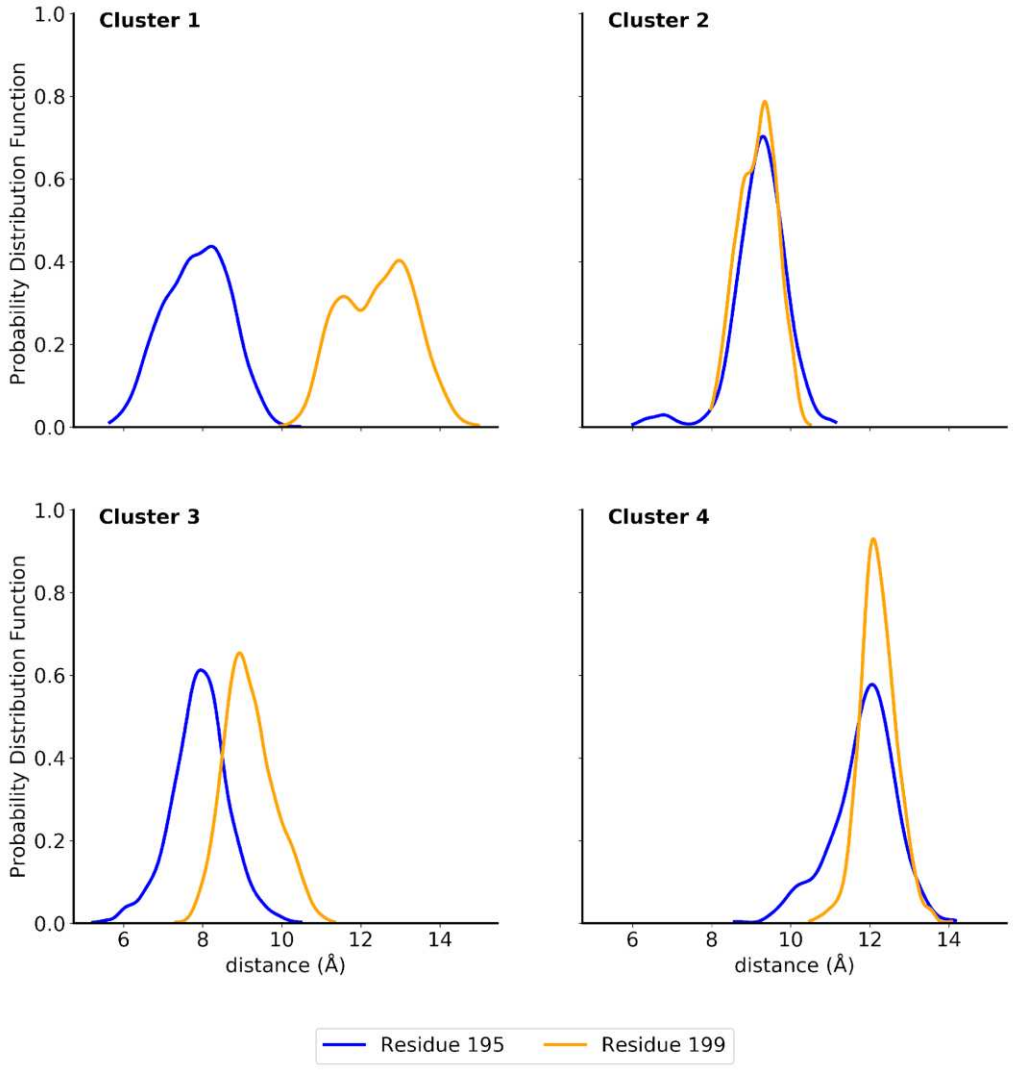

Figure S6 E195Cα-Ln and A199Cα-Ln distance distributions measured for the four clusters of the 3cam ensemble. Each distance distribution was fit to a Gaussian KDE using Silverman's Rule of Thumb for bandwidth selection.

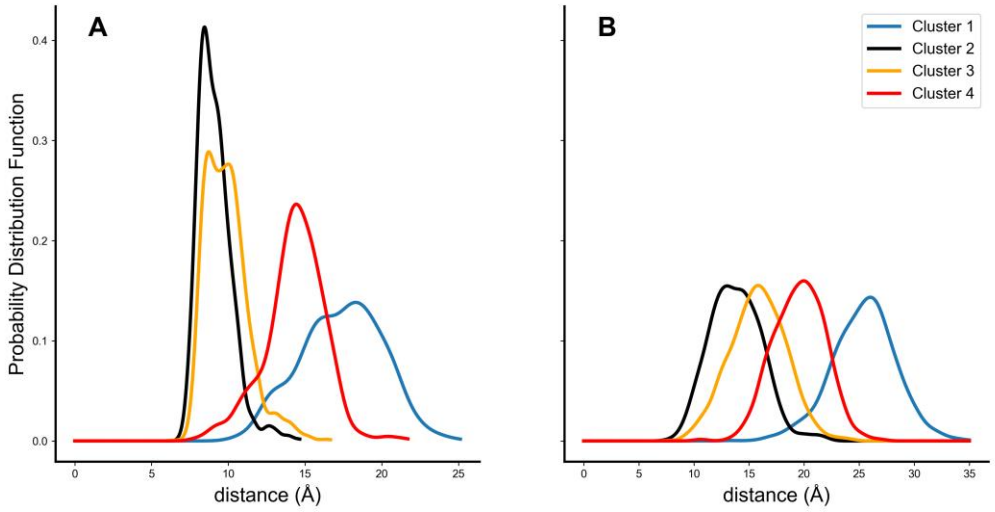

Figure S7 E91Cα–S192Cα (A) and N59Cα–G189Cα (B) distance distributions measured for the four clusters of the 3cam ensemble. Each distance distribution was fit to a Gaussian KDE using Silverman's Rule of Thumb for bandwidth selection.

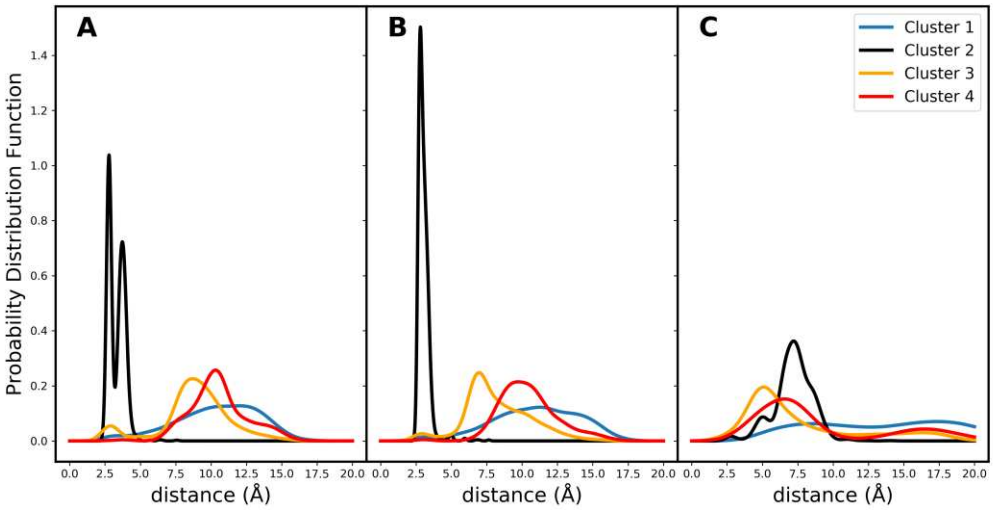

Figure S8 R186Hη1–D251Oδ1 (A), R186Hη2–D251Oδ2 (B) and K178Nζ–D251Oδ2 (C) distance distributions measured for the four clusters of the 3cam ensemble PCS. Each distance distribution was fit to a Gaussian KDE using Silverman's Rule of Thumb for bandwidth selection.

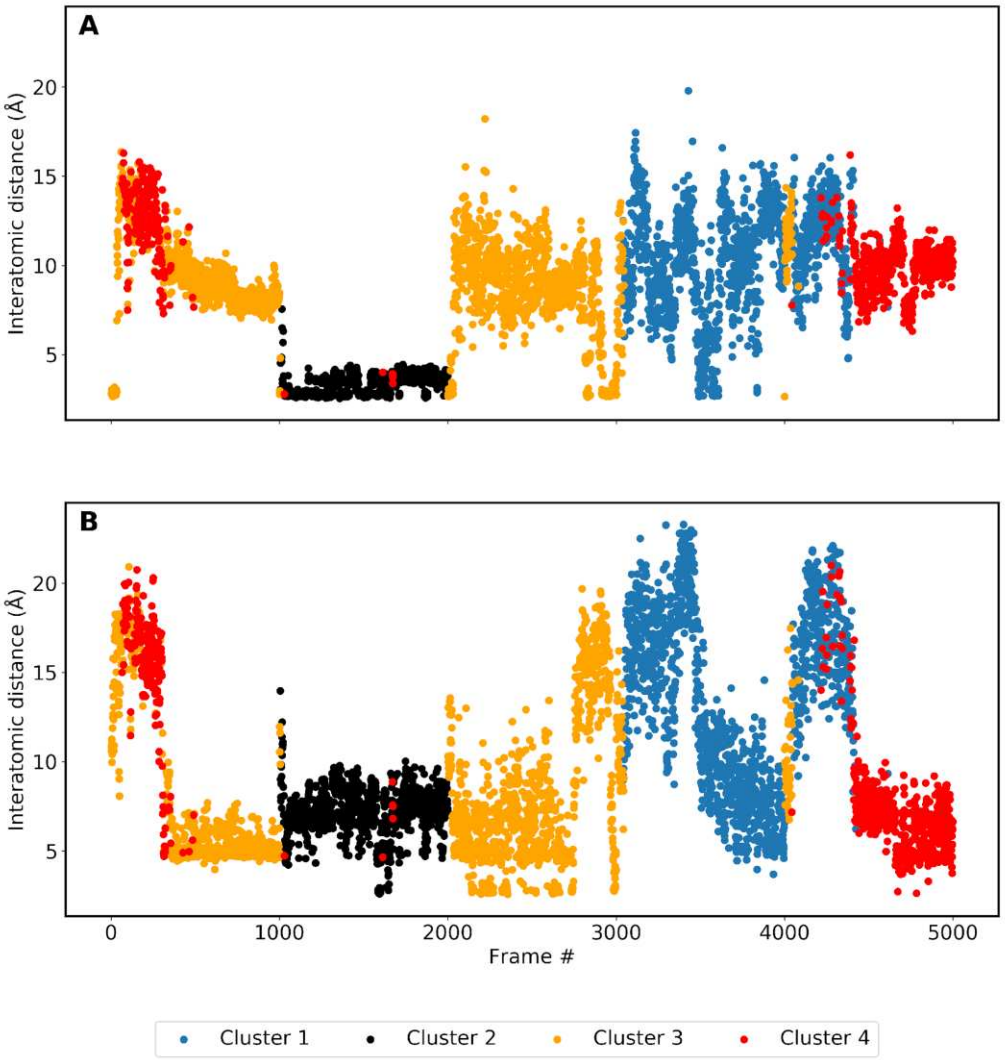

Figure S9 R186Hη1–D251Oδ1 (A) and K178Nζ–D251Oδ2 (B) distances plotted against frame number.

References

(1) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput* **2013**, *9* (7), 3084–3095. https://doi.org/10.1021/ct400341p.

(2) Skinner, S. P.; Liu, W.-M.; Hiruma, Y.; Timmer, M.; Blok, A.; Hass, M. A. S.; Ubbink, M. Delicate conformational balance of the redox enzyme cytochrome P450cam. *Proc Natl Acad Sci U.S.A* **2015**, *112* (29), 9022–9027. https://doi.org/10.1073/pnas.1502351112.

(3) Hiruma, Y.; Hass, M. A. S.; Kikui, Y.; Liu, W.-M.; Ölmez, B.; Skinner, S. P.; Blok, A.; Kloosterman, A.; Koteishi, H.; Löhr, F.; Schwalbe, H.; Nojiri, M.; Ubbink, M. The Structure of the Cytochrome P450cam–Putidaredoxin Complex Determined by Paramagnetic NMR Spectroscopy and Crystallography. *J. Mol. Biol* **2013**, *425* (22), 4353–4365. https://doi.org/10.1016/j.jmb.2013.07.006.

(4) Liu, W.-M.; Keizers, P. H. J.; Hass, M. A. S.; Blok, A.; Timmer, M.; Sarris, A. J. C.; Overhand, M.; Ubbink, M. A pH-Sensitive, Colorful, Lanthanide-Chelating Paramagnetic NMR Probe. *J. Am. Chem. Soc.* **2012**, *134* (41), 17306–17313. https://doi.org/10.1021/ja307824e.

(5) Lee, Y.-T.; Wilson, R. F.; Rupniewski, I.; Goodin, D. B. P450cam Visits an Open Conformation in the Absence of Substrate. *Biochemistry* **2010**, *49* (16), 3412–3419. https://doi.org/10.1021/bi100183g.

(6) Bashir, Q.; Volkov, A. N.; Ullmann, G. M.; Ubbink, M. Visualization of the Encounter Ensemble of the Transient Electron Transfer Complex of Cytochrome c and Cytochrome c Peroxidase. *J. Am. Chem. Soc.* **2010**, *132* (1), 241–247. https://doi.org/10.1021/ja9064574.