Allosteric effects in a catalytically impaired variant of the enzyme Cyclophilin A are unrelated to millisecond time scale motions

Pattama Wapeesittipan,[a] *Antonia S. J. S. Mey,*[a] *Malcolm D. Walkinshaw,*[b] *and Julien Michel*[a*]

[a] EaStCHEM School of Chemistry, David Brewster road, Joseph Black Building, The King's Buildings, Edinburgh, EH9 3FJ, UK.

[b] School of Biological Sciences, Michael Swann Building, Max Born Crescent, Edinburgh, Scotland, EH9 3BF, UK

E-mail: mail@julienmichel.net

Keywords: cyclophilins, catalysis, protein dynamics, molecular dynamics simulations

Abstract:

Robust catalyst design requires clear understanding of the mechanisms by which molecular motions influence catalysis. This work investigates the connection between molecular motions and catalysis for the much debated enzyme Cyclophilin A (CypA) in wild-type (WT) form, and a variant that features a distal serine to threonine (S99T) mutation. Previous biophysical studies have proposed that conformational exchange between a 'major' active and a 'minor' inactive state on millisecond timescales plays a key role in catalysis. Here this hypothesis was addressed with molecular dynamics simulation techniques. The simulations reproduce well NMR-derived measurements of changes of activation barriers for the *cis*/*trans* amide group isomerization of a model substrate, and support X-ray crystallography derived evidence for a shift in populations of major and minor active site conformations between the wild-type and S99T mutant forms. Strikingly, exchange between major and minor active site conformations occurs at a rate that is 5 to 6 orders of magnitude faster than previously proposed. Further analyses indicate that the decreased catalytic activity of the S99T mutant is a result of weakened hydrogen bonding interactions between the substrate and several active site residues in the transition state ensemble. Weakened hydrogen bonding interactions in the S99T mutant are due to an overall increase in fast positional fluctuations of active site residues caused by poorer packing of sidechains. Therefore the previously described millisecond time scale coupled motions are not necessary to explain allosteric effects in the S99T Cyclophilin A mutant.

1. Introduction

Robust catalyst design principles underpin progress in many areas of chemical sciences. A major goal of modern molecular biophysics is to clarify the connection between protein motions and catalysis in enzymes.[1–6] A wide range of experimental methods, e. g. neutron scattering, X-ray crystallography, NMR or vibrational spectroscopy have been used to characterize internal protein motions occurring from femtosecond to second timescales.[7–11] While there is broad consensus that protein motions are implicated in catalysis, there is much debate around the role of conformational changes occurring on a millisecond timescale, and several studies have linked changes in millisecond protein motions with changes in enzymatic function.[12–16] However, it remains unclear whether such motions have a causal link to catalysis, or are merely a manifestation of the inherent flexibility of proteins that covers a broad range of timescales.

There have been vigorous debates about the meaning of dynamics in the context of enzymatic catalysis.[17–20] In the framework of transition state theory, the reaction rate is given by equation 1:

$$k = A(T)e^{-\Delta G^{\ddagger}(T)/RT} \quad (1)$$

where T is the temperature and R the gas constant. The pre-exponential term $A(T)$ includes contributions from non-statistical motions such as recrossing or tunnelling, The exponential term involves the activation free energy of the chemical step $\Delta G^{\ddagger}(T)$. If transitions between reactant states are fast compared to the time scale of the chemical reaction, $\Delta G^{\ddagger}(T)$ is the free energy difference between the thermally equilibrated ensembles describing the reactant and transition states.[21–25] Non-statistical motions described by $A(T)$ have typically been found to make a small contribution to rate constants with respect to the exponential term that involves equilibrium fluctuations of the protein and solvent degrees of freedom.[26,27]

The current work is concerned with the connection between rates of thermally equilibrated motions, and catalysis in enzymes. Specifically, the focus is on clarifying the nature of protein motions implicated in catalysis for the well-studied enzyme Cyclophilin A (CypA). CypA is a member of the Cyclophilin family of peptidyl-prolyl isomerases which catalyzes the *cis*/*trans* isomerization of amide groups in proline residues.[28] CypA plays an essential role in protein folding and regulation, gene expression, cellular signaling and the immune system. Notably, CypA is involved in the infectious activity and the viral replication of HIV-1.[29–31] Accordingly, CypA has been the subject of structure-based drug design efforts for decades.[32–37] Because of its significance as a medical target, the catalytic mechanism of CypA has been the subject of extensive studies.[2,5,6,16,38–55] Computational studies have shown that the speedup of the rate of *cis*/*trans* isomerization rate of the prolyl peptide bond is a result of preferential transition-state stabilization through selective hydrogen bonding interactions in the active site of CypA.[42,48] Figure 1A depicts key interactions between the substrate and active site residues, whereas Figure 1B highlights the relevant ω angle of the substrate used to track the *cis*/*trans* isomerization reaction.

Elegant NMR relaxation experiments by Eisenmesser et al. have also characterized the existence of intrinsic motions in apo CypA that couple a 'major' state *M* with a 'minor' conformational state *m* with a rate constant $k_{M\rightarrow m}$=60 s^{-1}.[43] Fraser et al. later used ambient temperature X-ray crystallographic data to determine a high-resolution structure of this CypA state *m*, revealing an interconversion pathway with the major state *M* that involves coupled rotations of a network of side chains involving residues Ser99, Phe113, Met61 and Arg55. To establish the relevance of this minor state *m* to catalysis, the distal residue Ser99 was mutated to Thr99. Further X-ray and NMR measurements on the free enzyme confirmed that the S99T

mutant increased the population of the *m* state, while decreasing the conversion rate $k_{M\rightarrow m}$ to 1 s^{-1}.[56] Remarkably, additional NMR experiments established that this 60-fold decrease in conversion rate between *M* and *m* states in the S99T mutant correlates with a ca. 70-fold decrease in bidirectional isomerization rate ($k_{iso} = k_{cis\rightarrow trans} + k_{trans\rightarrow cis}$) of a model substrate with respect to wild type, comparable to rate decreases observed for mutations of key active site residues such as Arg55.[56]

While this body of work suggested a link between millisecond time scale motions and catalysis in enzymes, there was no apparent explanation for the decreased catalytic activity of S99T and there have been vigorous debates about the catalytic mechanism of CypA.[19] The present study uses a variety of extensive equilibrium and biased molecular dynamics (MD) simulations and Markov State Models (MSM) to clarify the link between catalytic activity and rates of molecular motions.

Methods

2.1 Systems preparation

Models for apo/substrate bound human CypA in WT or S99T forms were prepared for MD simulations from PDB structures 3K0N (*R*=1.39 Å) and 3K0O (*R*= 1.55 Å) respectively.[56] For WT the major conformation of 3K0N (altloc A, occupancy 0.58) was retained. The crystal structure of the CypA-HAGPIA peptide complex (PDB ID: 1AWR)[57] was used to obtain a suitable orientation for the CypA-AAPF complex. Structure 1AWR was aligned to the structure of WT and S99T mutant, and the substrate HAGPIA was converted to *trans* Ace-Ala-Ala-Pro-Phe-Nme using Schrödinger's Maestro.[58] The N-terminal and C-terminal ends of the proteins and substrate were capped and missing atoms were added using the same program. Table S1 in

the SI summarises all simulations conducted in this study.

2.2 apo WT and S99T MD simulations

Ten independent 200 ns MD trajectories of the apo WT and S99T proteins were generated using Gromacs 5.0.[59] WT and S99T were solvated in a rhombic dodecahedron box of TIP3P water molecules with edges extending 1 nm away from the proteins and chloride counter-ions were added to neutralise the overall net-charge. The AMBER99SB forcefield was used to describe protein atoms, with optimised ω angle parameters for amides as reported by Doshi and co-workers.[60] Each system was then energy minimized using the steepest descent method for 50,000 steps and equilibrated for 100 ps in an NVT ensemble, followed by a 100 ps NPT pre-equilibration. For the minimization and pre-equilibration processes, a position restraining force was applied to the heavy atom of the protein with the harmonic force constants at 1000 $kJ{\cdot}mol^{-1}{\cdot}nm^{-2}$. Each subsequent production run generated a 200 ns trajectory with a 2 fs time step. The first 5 ns were discarded to allow full equilibration of the unrestrained protein. The average temperature was maintained at 300 K with a stochastic Berendsen thermostat.[61] The Parrinello-Rahman barostat was used for pressure coupling at 1 bar.[62] The Particle Mesh Ewald scheme was used for long-range electrostatic interactions with a Fourier grid spacing of 0.16 nm, and fourth-order cubic interpolation.[63] Short-range van der Waals and electrostatic interaction were cutoff at 1 nm. The LINCS algorithm was used to constrain all bonds.[64]

2.3 Substrate-bound WT and S99T MD simulations

Models of CypA WT and S99T in complex with theAce-AAPF-Nme substrate were also prepared. The substrate was modelled in a *ts* (transition state) conformation. To generate a suitable *ts* model, a steered MD (SMD) protocol was used to rotate an equilibrated *trans* configuration (ω ca. 180º) of the substrate to a *ts configuration* (ω ca. 90º) using the software

PLUMED2.[65] Table S2 in the SI summarizes the protocol details. For the resulting models (WT+*ts*, S99T+*ts*) series of ten independent of 200 ns trajectories were generated using Gromacs 5.0 with simulation parameters identical to the apo simulations. The substrate proline ω angle was additionally restrained to 90º using a force constant of 400 kJ·mol^{-1}·rad^{-2}.

2.4 Umbrella sampling simulations

The above equilibrium MD simulations were complemented by a series of Umbrella Sampling (US) simulations,[66–68] to compute free energy profiles along ω.[39,42,44,69] A standard harmonic potential was used to bias the ω angle towards a series of target values ω_k spanning the interval [-180°,180º]. The force constants of the biasing potential and the spacing between ω_k values were adjusted by trial and error in order to obtain a good overlap between probability distributions of neighbouring ω_k values (Tables S3-S4 and Figure S1). The US calculations were either initiated from the substrate in a *trans* conformation taken from a 10-ns equilibration run (substrate in solution), or from an equilibrated *ts* conformation (substrate bound to WT or S99T). Simulations were performed serially initially for 7 ns, with the starting conformation for a given target angle ω_k taken from the preceding run performed at the neighbouring $\omega_{k+\Delta\omega}$ value. Each US was then extended to 20 ns, and the last 15 ns used for free energy estimation. The free energy profiles of three consecutive 5-ns blocks were calculated and then used to calculate the average free energy profile and the standard error of the mean. A total of 22 (substrate in solution) or 24 (substrate bound to WT or S99T) umbrellas were used. The entire procedure was repeated twice to estimate uncertainties in the resulting free energy profiles. All simulations were carried out using a PLUMED2 patched version of Gromacs 5.0 with simulation parameters identical to the previously described equilibrium MD simulation protocols unless otherwise mentioned. The

weighted histogram analysis method (WHAM) was used to produce a free energy profile from the pool of US simulations.[70]

2.5 Markov state models

In order to classify the dynamics of the key residues in the equilibrium apo WT and S99T mutant simulations a set of Markov state models (MSM) were built. MSMs allow the quantitative evaluation of equilibrium and dynamic properties of biomolecular systems.[71–74] For a detailed description of MSM construction and analysis see for example Bowman et al.[75] Here, a MSM is used to extract timescales of the movement of key CypA residues from the ten 200 ns simulations of apo WT and S99T mutant. To construct the MSM the cosine and sine of the following dihedral angles were used on the joint dataset of all apo simulations: χ_1 of Arg55, Met61, Phe113 and Ser/Thr99, χ_2 of Phe 113 and Met 61, χ_3 of Arg55. The 14 input coordinates were then clustered using k-means to assign each trajectory frame from both the mutant and WT trajectories to one of 100 clusters. From the set of discrete trajectories, transition matrices for the WT $T_{\mathrm{wt}}(\tau)$ and mutant $T_{\mathrm{S99T}}(\tau)$ were constructed using the Bayesian MSM option from the software package pyemma.[76] A transition matrix contains the conditional probabilities of 'jumping' from one microstate or cluster to the next such that the matrix elements are $T_{ij} = P(x_t \in j, x_{t+\tau} \in i)$, where i and j are microstates respectively and τ is a lagtime. Analysis of the eigenvalues/vectors of the transition matrix enables characterization of the slowest dynamic process occurring in the system. Using the same microstate description for the WT and mutant MSM facilitates direct comparison between the two MSMs. To classify the slow processes the PCCA method is used to coarse grain the MSM into two states for the WT, and four states for the mutant (see Figure S3).[77] Uncertainties in state populations and mean first passage times (MFPT) derived from the coarse-grain MSMs were estimated by calculating the standard deviation of the

mean from 1000 samples of the Bayesian MSM. All of the MSM analysis was carried out with the software package pyemma version 2.3.2. [76] Additional details and Figures (Figures S3-S7) describing the MSM validation are available in the supporting information.

2.6 Other trajectory analyses

Additionally, interaction energies between binding site residues (Arg55, Ile57, Phe60, Met61, Gln63, Asn102, Gln111, Phe113, Trp121, Leu122 and His126) and all atoms of the substrate were analysed with the Gromacs g_energy module. The probability distribution of distances between key residues and substrate atoms during the simulations were computed using the MDAnalysis library.[78,79] The per-residue root mean square fluctuations (RMSF) of heavy atoms were measured for simulations around the transition state region (ω = 86º, 100º, 115º).

2. Results and Discussion

3.1 The proposed 'major' and 'minor' conformations exchange on timescales of nanoseconds to microseconds in apo CypA

Fraser et al. have described the proposed major *M* and minor *m* states according to sets of values of χ_1 (Phe113, Ser/Thr99), χ_2 (Met61) and χ_3 (Arg55) angles, as illustrated in Figure 2A.[56,80] The results from the MD simulations shown in Figure 2B suggest that in apo WT the Phe113 'in' (χ_1=+60º) orientation is dominant, which is consistent with the high-occupancy rotamer found in the X-ray structure (occupancy = 0.58). A marginal population is calculated for the Phe113 'out' (χ_1=-60º) rotamer, which is qualitatively consistent with this rotamer showing decreased occupancy in the X-ray structure (occupancy = 0.42).[56] In apo S99T there is a significant population shift towards the 'out' orientation (χ_1=-60º) with both rotamers sampled via multiple backwards and forwards transitions (see Figure 2B and Figure S2). This partially agrees with the X-ray structure of S99T where only the Phe113 'out' rotamer is observed

(occupancy = 1.0). The increased population of the χ_1=-60º rotamer of Phe113 in S99T is also consistent with simulation results of Papaleo et al. obtained with a range of different forcefields.[80] Figure 2C shows that Ser99 populates three χ_1 rotamers in WT, whereas only two rotamers were resolved by X-ray crystallography. In comparison with WT, there is an inversion in population preferences for χ_1 rotamers at -180º and -60º in the S99T structure. This is partially consistent with Thr99 being only observed in the χ_1=-180º rotamer by X-ray crystallography. Finally no significant changes are observed in the χ_2 values of Met61 (Figure 2D) and the χ_3 values of Arg55 (Figure 2E) which contrasts with the X-ray crystallography data that suggested that mutation of Ser99 into Thr99 also influences the conformation of Met61 and Arg55.

Overall the MD simulations significantly populate rotameric states that are observed in the X-ray structures of WT and S99T reported by Fraser et al., as well as other states that were not detected in the X-ray structures. The MD simulations also suggest a population shift in rotameric states of Phe133 and Ser/Thr99. However additional rotameric states for Met61 and Arg55 are also sampled in the simulations and no shift in conformational preferences is apparent for these two residues between the WT and S99T forms of CypA.

Figure 3 summarizes the kinetic data obtained from the Markov state modelling. Slow dynamics in WT can be satisfactorily described by a two state process (Figure 3A). This process involves a 180º rotation of the χ_2 dihedral of Phe113 (Figure 3B) with forward or backward rates of approximately 10 μs^{-1}. The slowest process detected in the S99T mutant MSM (Figure 3C) is similar to the one observed in WT, featuring a rotation of Phe113 around its χ_2 torsional angle (Figure 3D). However, the transition rate (ca. 0.8 μs^{-1}) is approximately one order of magnitude slower than in WT. This indicates that increased steric bulk upon mutation of nearby Ser99 into a threonine hinders rotation of Phe113. Additionally, a second slow process emerges in the

mutant which is associated with a transition rate of about 1 μs^{-1}, and highlighted by the transition between the green and cyan structures shown in Figure 3E. The underlying structural rearrangement responsible for this process is described by the changes in dihedral probability distributions in Figures 3F-I. Conversion between the green and cyan states involves rotation of Phe113 from χ_1 ca. +60º to χ_1 ca. -60º (Figure 3F). This rotation is coupled with rotation of Thr99 from χ_2 ca. -60º to ca. -180 º (Figure 3G), as well as a shift of Met61 χ_2 from ca. +90º to -180º (Figure 3H). This process also involves rotation of Phe113 such that χ_2 changes from ca. -90º to +90º (Figure 3I).

Overall it can be surmised that the second slowest process describes a concerted rotation motion of Thr99 and Phe113 that is similar to the proposed conversion mechanism between the CypA major *M* state into a minor *m* state proposed by Fraser et al.[56] However the concerted rotations do not involve the catalytically important residue Arg55, whose motions occur on faster timescales (Figure S5-S6). Crucially the rate of this concerted rotation process (ca. 1-10 μs^{-1}) is 5-6 orders of magnitude faster than the rates suggested by previous NMR relaxation dispersion experiments (ca. 1-100 s^{-1}).[56] Differences in temperatures between simulations (25 ºC) and experiments (10 ºC) are only expected to contribute few-fold variability in rates. While the accuracy of the modelling has its limitations, the large discrepancy in rates suggests that the coupled rotations of Thr99 and Phe113 inferred from the X-ray structures may not correspond to the motions of these residues detected by NMR relaxation dispersion experiments.

3.2 Differential catalytic activity is reproduced without exchange between major and minor conformations

Given that the timescales of concerted rotations of Phe113 and Ser/Thr99 in apo WT and S99T appear much faster than previously suggested, attention turned next to substrate bound CypA simulation to investigate the role of this coupled motion in catalysis. In complex with the peptide substrate AAPF no transitions between the major state *M* and the minor state *m* were observed. Simulations started in the *m* state rapidly revert to the *M* state (Figure S8) because the conformation Phe113 in the *m* state would cause a steric clash with the proline ring of the substrate.

Next the isomerization free energy profiles shown in Figure 4 were examined. Ladani and Hamelberg have previously shown that fixed-charge classical force fields reasonably well reproduce the energetics of amide bond rotation due to relatively small changes in intramolecular polarization during this process.[44] The calculated activation free energy for the uncatalyzed *cis*→*trans* isomerization process in water is consistent with experimental data (20.0 ± 0.4 kcal•mol^{-1} vs ca. 19.3 kcal•mol^{-1} for the related substrate Suc-AAPF-pNA at 283 K).[81,82] The free energy profile for the substrate bound to CypA WT shows that the enzyme catalyzes preferentially the isomerization reaction in both directions via a transition state with a positive ω value (ca. 90-100º). There is a more significant decrease in activation free energy for *trans*→*cis* (ca. -7 kcal•mol^{-1}) than for *cis*→*trans* (ca. -3 kcal•mol^{-1}), because the *cis* form is more tightly bound to CypA than the *trans* form. These results are consistent with previous calculations of Ladani and Hamelberg on WT CypA (that reported decreases of ca. -9 kcal/mol for *trans*→*cis* and ca. -4 kcal/mol for *cis*→*trans*).[44]

Additionally, the present calculations indicate that the S99T mutant is less effective than WT at the isomerization of *trans* or *cis* substrates. This is consistent with the experimental observation that the bidirectional isomerization rate k_{iso} is 70-fold smaller in S99T although the

effect is even more pronounced in the simulations (k_{iso} ca. 270-4700 smaller). Quantitative agreement was not expected as owing to the exponential relationship in eq 1 small differences in calculated and experimental activation free energies cause large variations in rate constants. Nevertheless the present calculations clearly demonstrate that the S99T mutant catalyzes the isomerization reaction less efficiently than WT which is consistent with experimental data. Crucially, no coupled rotations between Phe113 and Ser/Thr99 are observed in the ensembles of simulations used to construct the reaction free energy profiles, therefore this coupled motion does not appear to be linked to catalysis of the chemical step.

3.3 Decreased hydrogen-bonding interactions with multiple active site residues explain transition state destabilization in the S99T mutant

Clarification of the mechanisms of allosteric inhibition of S99T as observed in the present calculations is provided by analysis of the umbrella sampling trajectories. The average Coulombic interaction energy of the substrate with binding site residues as a function of ω is shown in Figure 5A. It is apparent that around the main transition state region (ω ca. 90-100º) S99T stabilizes the substrate significantly less than WT (interaction energies more positive by up to 15 kcal•mol^{-1}). Figure 5B shows a decomposition of the substrate-protein electrostatic energies for each active site residue. It is apparent that five residues show weaker electrostatic interactions with the substrate (Arg55, Trp121, Asn102, His126, Gln63). These residues frequently form hydrogen-bonding interactions with the substrate in WT, but the probability of hydrogen bond formation decreases in all cases for S99T (Figure 5C, see also Figure S9 for probability distribution plots). Overall, this indicates that the poorer catalytic activity of S99T

may be attributed to weaker hydrogen-bonding interactions between the substrate and the entire active site of CypA.

3.4 Active site residues in the transition state ensemble of the S99T mutant are more flexible than in wild-type

It remains to clarify how mutation of the distal residue S99 could impact on hydrogen bonding interactions between active site residues and the substrate. Overall the mutant maintains a highly similar structure to wild type, but small increases in positional fluctuations over ps-ns timescales occur for patches of residues across the entire protein (Figure S10). Figure 6A presents root-mean squared fluctuations for the active site residues that make contacts with the substrate. Overall 10 out 11 residues show a significant increase in flexibility in S99T, including Phe113 that contacts the mutated Ser99 (Figure 1A). The increases in root mean square fluctuations in S99T correlates with small shifts and broadenings of ϕ, ψ, or χ torsional angle distributions around rotamers already sampled in WT, rather than jumps between new rotameric states (Figures S11-S12), highlighting the dynamic nature of allosteric effects in this mutant.

With this final piece of information, the puzzle of allosteric inhibition in the CypA S99T mutant can now be resolved. The extra methyl group introduced upon serine to threonine mutation at residue 99 disrupts packing with nearby Phe113 and His126, leading to small increases in amplitude of motions of these two residues (Figure 6A). Phe113 and His126 delineate the edges of the Pro pocket and pack against several other active site residues that become in turn more flexible. Notably His126 packs against Asn102, whereas Phe113 packs against Met61 and Leu122 packs against Trp121, whereas Met61 is in contacts with Gln63 and Arg55 (Figure 1A) The overall increases in flexibility of active site residues correlates with a decrease in probabilities of forming strong hydrogen bonds with the substrate at the *ts* region

(Figure 5C). Figure 6B summarizes the proposed mechanism of allosteric inhibition of CypA catalysis by the distal Ser/Thr99 mutation.

3. Conclusions

This work highlights the potential of detailed molecular simulation studies to guide the interpretation of biophysical measurements for the elucidation of protein allosteric mechanisms.[83] Previous work has suggested that exchange on millisecond timescales between conformational states in CypA are linked to its catalytic cycle,[43] leading to a proposal for a slow exchange between a major *M* and a minor *m* state of a set of side chain rotamers linking distal residue Ser99 to active-site residues.[43,56] The present results support the existence of concerted rotations of Ser/Thr99 and Phe113 side-chains between sets of distinct rotameric states in the free forms of CypA WT and S99T mutant, in agreement with earlier X-ray crystallography data.[53] However, coupling occurs on timescales five to six orders of magnitude faster than suggested by earlier NMR relaxation dispersion measurements.[56] Further, there is no evidence that such coupled motions affect catalysis in substrate bound simulations of CypA. Rather increased flexibility of the active site residues in the transition state ensemble of the S99T mutant is sufficient to explain decreased transition state stabilization through disruption of key hydrogen-bonding interactions. Thus, while there is strong evidence that large parts of CypA undergo fluctuations on millisecond timescales, the present results do not support the notion that such motions are causally linked to catalysis. Instead, the findings are consistent with reports that ps-ns time scale active site motions couple to the chemical step, as shown in other enzymes such as DHFR.[84] Thus this body of work has implications for the timescales that should be probed in experiments that aim to modulate enzymatic catalysis by tuning of protein motions.[41,85] This work also demonstrates that the transition state ensemble of CypA has been finely-tuned by

nature to promote prolyl amide bond isomerization by rigidifying the active site at the transition state, and that distal mutations can significantly influence catalytic efficiency by increasing active site flexibility. This has important implications for enzyme design and optimization strategies.

ASSOCIATED CONTENT

Supporting Information. Tables, figures movies datasets and extended description of protocols.

AUTHOR INFORMATION

Corresponding Author

* mail@julienmichel.net

Present Addresses

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

Gratitude is expressed to Fernanda Duarte for thoughtful discussions about this work.

Funding Sources

Julien Michel is supported by a University Research Fellowship from the Royal Society. The research leading to these results has received funding from the European Research Council under

the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement No. 336289. Pattama Wapeesittipan is supported by The Development and Promotion of Science and Technology Talents Project (DPST) Scholarship, Royal Thai Government. This project made use of time on ARCHER granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (http://hecbiosim.ac.uk), supported by EPSRC (grant no. EP/L000253/1)

REFERENCES

(1) Warshel, A. *Proc. Natl. Acad. Sci.* **1984**, *81* (2), 444.

(2) McGowan, L. C.; Hamelberg, D. *Biophys. J.* **2013**, *104* (1), 216.

(3) Callender, R.; Dyer, R. B. *Acc. Chem. Res.* **2015**, *48* (2), 407.

(4) Ma, B.; Nussinov, R. *Curr. Opin. Chem. Biol.* **2010**, *14* (5), 652.

(5) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. *Science.* **2002**, *295* (5559), 1520.

(6) Agarwal, P. K.; Geist, A.; Gorin, A. *Biochemistry* **2004**, *43* (33), 10605.

(7) Rader, S. D.; Agard, D. a. *Protein Sci.* **1997**, *6* (7), 1375.

(8) Benkovic, S. J.; Hammes-Schiffer, S. *Science.* **2003**, *301* (5637), 1196.

(9) Bu, Z.; Neumann, D. a; Lee, S. H.; Brown, C. M.; Engelman, D. M.; Han, C. C. *J. Mol. Biol.* **2000**, *301*, 525.

(10) Tehei, M.; Madern, D.; Pfister, C.; Zaccai, G. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (25), 14356.

(11) Gabel, F.; Bicout, D.; Lehnert, U.; Tehei, M.; Weik, M.; Zaccai, G. *Q Rev Biophys* **2002**, *35* (4), 327.

(12) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450* (7172), 964.

(13) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. *Nat. Struct. Mol. Biol.* **2004**, *11* (10), 945.

(14) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzym.* **2001**, *339* (2001), 204.

(15) Doucet, N.; Watt, E. D.; Loria, J. P. *Biochemistry* **2009**, *48* (30), 7160.

(16) Schlegel, J.; Armstrong, G. S.; Redzic, J. S.; Zhang, F.; Eisenmesser, E. Z. *Protein Sci.*

2009, *18* (4), 811.

(17) Pisliakov, A. V; Cao, J.; Kamerlin, S. C. L.; Warshel, A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (41), 17359.

(18) Olsson, M. H. M.; Parson, W. W.; Warshel, A. *Chem. Rev.* **2006**, *106* (5), 1737.

(19) Kamerlin, S. C. L.; Warshel, A. *Proteins Struct. Funct. Bioinforma.* **2010**, *78* (6), 1339.

(20) Warshel, A.; Bora, R. P. *J. Chem. Phys.* **2016**, *144* (18).

(21) Masgrau, L.; Truhlar, D. G. *Acc. Chem. Res.* **2015**, *48* (2), 431.

(22) Kohen, A. *Acc. Chem. Res.* **2015**, *48* (2), 466.

(23) Wolfenden, R. *Nature* **1969**, *223*, 704.

(24) Schramm, V. L. *Curr. Opin. Struct. Biol.* **2005**, *15* (6), 604.

(25) Lienhard, G. E.; Secemski, I. I.; Koehler, K. A.; Lindquist, R. N. *Cold Spring Harb. Symp. Quant. Biol.* **1972**, *36*, 45.

(26) Delgado, M.; Görlich, S.; Longbotham, J. E.; Scrutton, N. S.; Hay, S.; Moliner, V.; Tuñón, I. *ACS Catal.* **2017**, *7* (5), 3190.

(27) Luk, L. Y. P.; Javier Ruiz-Pernia, J.; Dawson, W. M.; Roca, M.; Loveridge, E. J.; Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J.; Tunon, I.; Moliner, V.; Allemann, R. K. *Proc. Natl. Acad. Sci.* **2013**, *110* (41), 16344.

(28) Go, S. F.; Marahiel, M. A.; Gothel, S. F. *Cell Mol Life Sci* **1999**, *55* (3), 423.

(29) Li, J.; Tan, Z.; Tang, S.; Hewlett, I.; Pang, R.; He, M.; He, S.; Tian, B.; Chen, K.; Yang, M. *Bioorganic Med. Chem.* **2009**, *17* (8), 3177.

(30) Zhao, Y.; Chen, Y.; Schutkowski, M.; Fischer, G.; Ke, H. *Structure* **1997**, *5* (1), 139.

(31) Yoo, S.; Myszka, D. G.; Yeh, C.; McMurray, M.; Hill, C. P.; Sundquist, W. I. *J. Mol. Biol.* **1997**, *269* (5), 780.

(32) Dornan, J.; Taylor, P.; Walkinshaw, M. *Curr. Top. Med. Chem.* **2003**, *3* (12), 1392.

(33) Gallay, P. A. *Immunol. Res.* **2012**, *52* (3), 200.

(34) Yang, Y.; Moir, E.; Kontopidis, G.; Taylor, P.; Wear, M. A.; Malone, K.; Dunsmore, C. J.; Page, A. P.; Turner, N. J.; Walkinshaw, M. D. *Biochem. Biophys. Res. Commun.* **2007**, *363* (4), 1013.

(35) Daum, S.; Schumann, M.; Mathea, S.; Aumüller, T.; Balsley, M. A.; Constant, S. L.; De Lacroix, B. F.; Kruska, F.; Braun, M.; Schiene-Fischer, C. *Biochemistry* **2009**, *48* (26), 6268.

(36) Ahmed-Belkacem, A.; Colliandre, L.; Ahnou, N.; Nevers, Q.; Gelin, M.; Bessin, Y.; Brillet, R.; Cala, O.; Douguet, D.; Bourguet, W.; Krimm, I.; Pawlotsky, J.-M.; Guichou, J.-F. *Nat. Commun.* **2016**, *7*, 12777.

(37) Georgiou, C.; McNae, I. W.; Wear, M. A.; Ioannidis, H.; Michel, J.; Walkinshaw, M. D. *bioRxiv* **2017**.

(38) Doshi, U.; Holliday, M. J.; Eisenmesser, E. Z.; Hamelberg, D. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (17), 4735.

(39) Agarwal, P. K. *Proteins Struct. Funct. Genet.* **2004**, *56* (3), 449.

(40) Camilloni, C.; Sahakyan, A. B.; Holliday, M. J.; Isern, N. G.; Zhang, F.; Eisenmesser, E. Z.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (28), 10203.

(41) Holliday, M. J.; Camilloni, C.; Armstrong, G. S.; Vendruscolo, M.; Eisenmesser, E. Z. *Structure* **2017**, *25* (2), 276.

(42) Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2009**, *131* (1), 147.

(43) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. a; Skalicky, J. J.; Kay, L. E.; Kern, D. *Nature* **2005**, *438* (7064), 117.

(44) Ladani, S. T.; Hamelberg, D. *J. Phys. Chem. B* **2012**, *116* (35), 10771.

(45) Howard, B. R.; Vajdos, F. F.; Li, S.; Sundquist, W. I.; Hill, C. P. *Nat. Struct. Biol.* **2003**, *10* (6), 475.

(46) Leone, V.; Lattanzi, G.; Molteni, C.; Carloni, P. *PLoS Comput. Biol.* **2009**, *5* (3).

(47) Trzesniak, D.; van Gunsteren, W. F. *Protein Sci.* **2006**, *15* (11), 2544.

(48) Ladani, S. T.; Hamelberg, D. *Mol. Simul.* **2014**, *40* (10–11), 765.

(49) Nagaraju, M.; McGowan, L. C.; Hamelberg, D. *J. Chem. Inf. Model.* **2013**, *53* (2), 403.

(50) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2002**, *124* (25), 7303.

(51) Kofron, J. L.; Kuzmic, P.; Kishore, V.; Colón-Bonilla, E.; Rich, D. H. *Biochemistry* **1991**, *30* (25), 6127.

(52) Kern, D.; Kern, G.; Scherer, G.; Fischer, G.; Drakenberg, T. *Biochemistry* **1995**, *34* (41), 13594.

(53) Zhao, Y.; Ke, H. *Biochemistry* **1996**, *35* (23), 7356.

(54) Doshi, U.; McGowan, L. C.; Ladani, S. T.; Hamelberg, D. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (15), 5699.

(55) Li, G.; Cui, Q. *J. Am. Chem. Soc.* **2003**, *125* (49), 15028.

(56) Fraser, J. S.; Clarkson, M. W.; Degnan, S. C.; Erion, R.; Kern, D.; Alber, T. *Nature* **2009**, *462* (7273), 669.

(57) Vajdos, F. F.; Yoo, S.; Houseweart, M.; Sundquist, W. I.; Hill, C. P. *Protein Sci.* **1997**, *6* (11), 2297.

(58) Schrödinger Release 2017-1: Maestro, Schrödinger, LLC, New York, NY, 2017. .

(59) Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindah, E. *SoftwareX* **2015**, *1–2*, 19.

(60) Doshi, U.; Hamelberg, D. *J. Phys. Chem. B* **2009**, *113* (52), 16590.

(61) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126* (1).

(62) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52* (12), 7182.

(63) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103* (19), 8577.

(64) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18* (12), 1463.

(65) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185* (2), 604.

(66) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23* (2), 187.

(67) Hooft, R. W. W.; van Eijck, B. P.; Kroon, J. *J Chem Phys* **1992**, *97* (9), 6690.

(68) Northrup, S. H.; Pear, M. R.; Lee, C. Y.; McCammon, J. A.; Karplus, M. *Proc. Natl. Acad. Sci.* **1982**, *79* (13), 4035.

(69) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127* (6), 1969.

(70) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13* (8), 1011.

(71) Chodera, J. D.; Noé, F. *Current Opinion in Structural Biology*. 2014, pp 135–144.

(72) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (45), 19011.

(73) Plattner, N.; Noé, F. *Nat. Commun.* **2015**, *6*, 7653.

(74) Lane, T. J.; Bowman, G. R.; Beauchamp, K.; Voelz, V. A.; Pande, V. S. *J. Am. Chem. Soc.* **2011**, *133* (45), 18413.

(75) Bowman, G. R.; Pande, V. S.; Noé, F. *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*; 2014; Vol. 797.

(76) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J. H.; Noé, F. *J. Chem. Theory Comput.* **2015**, *11* (11), 5525.

(77) Röblitz, S.; Weber, M. *Adv. Data Anal. Classif.* **2013**, *7* (2), 147.

(78) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. *J. Comput. Chem.* **2011**, *32* (10), 2319.

(79) Naveen Michaud-Agrawal, Elizabeth J. Dening, Thomas B. Woolf, O. B. *J. Comput. Chem.* **2010**, *31* (16), 2967.

(80) Papaleo, E.; Sutto, L.; Gervasio, F. L.; Lindorff-Larsen, K. *J. Chem. Theory Comput.* **2014**, *10* (9), 4169.

(81) Harrison, R. K.; Stein, R. L. *J. Am. Chem. Soc.* **1992**, *114* (9), 3464.

(82) Dugave, C.; Demange, L. *Chem. Rev.* **2003**, *103* (7), 2475.

(83) Michel, J. *Phys. Chem. Chem. Phys* **2014**, *16* (16), 4465.

(84) Ruiz-Pernia, J. J.; Luk, L. Y. P.; García-Meseguer, R.; Martí, S.; Loveridge, E. J.; Tuñón, I.; Moliner, V.; Allemann, R. K. *J. Am. Chem. Soc.* **2013**, *135* (49), 18689.

(85) Liu, C. T.; Wang, L.; Goodey, N. M.; Hanoian, P.; Benkovic, S. J. *Biochemistry* **2013**, *52* (32), 5332.

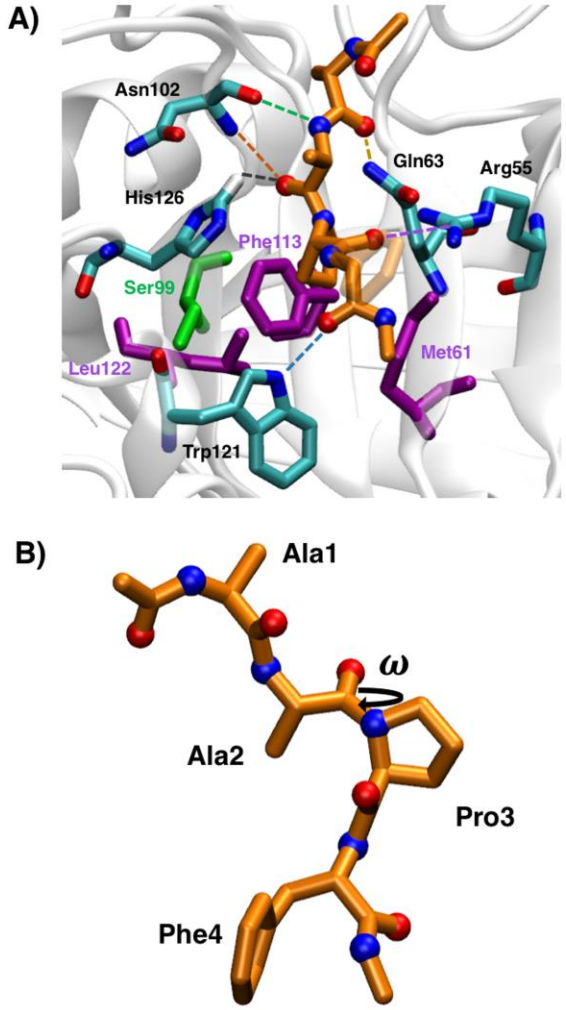


Figure 1. (**A**) Key residues in the active site of Cyclophilin A that form hydrogen bonds (cyan sticks, dashed lines) or are in contact (purple sticks) with the transition-state form of the peptide Ace-AAPF-Nme (orange sticks). For clarity the Phe sidechain of the substrate is represented as transparent sticks. The distal residue Ser99 is depicted (green sticks). (**B**) The ω torsional angle of -Ala2-Pro3- is used to track the progression of the isomerization reaction between *cis* and *trans* forms.

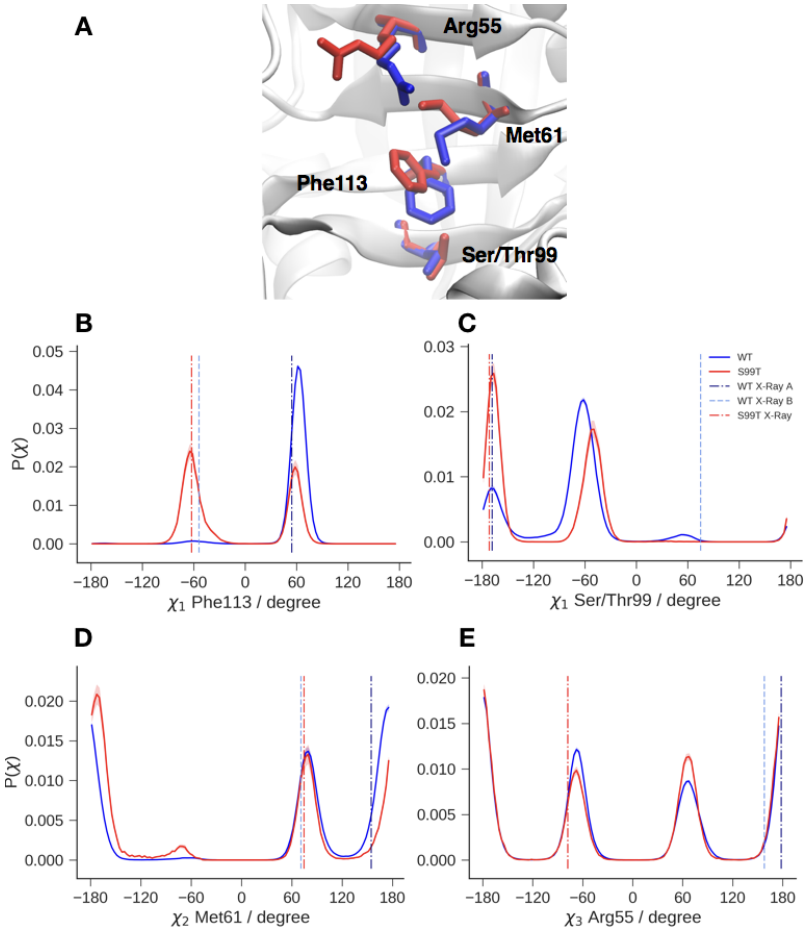


Figure 2. (**A**) The orientations of side chains Ser99/Thr99, Phe113, Met61 and Arg55 define the 'major' conformation of WT (blue) and stabilized 'minor' conformation of S99T (red), taken from crystallography data.[56] The MSM-derived probability distributions of selected dihedral angles for WT (blue) or S99T (red) (**B**) χ_1 in Phe113, (**C**) χ_1 in Ser/Thr99, (**D**) χ_2 in Met6, (**E**) χ_3 in Arg55. The error bars indicate the standard deviation of the mean. The X-ray crystallography derived data of Fraser et al.[56] is depicted for WT with dark blue dashed-dotted lines (high occupancy conformation), or light blue dashed lines (low occupancy conformation), and for S99T in red dashed-doted lines.

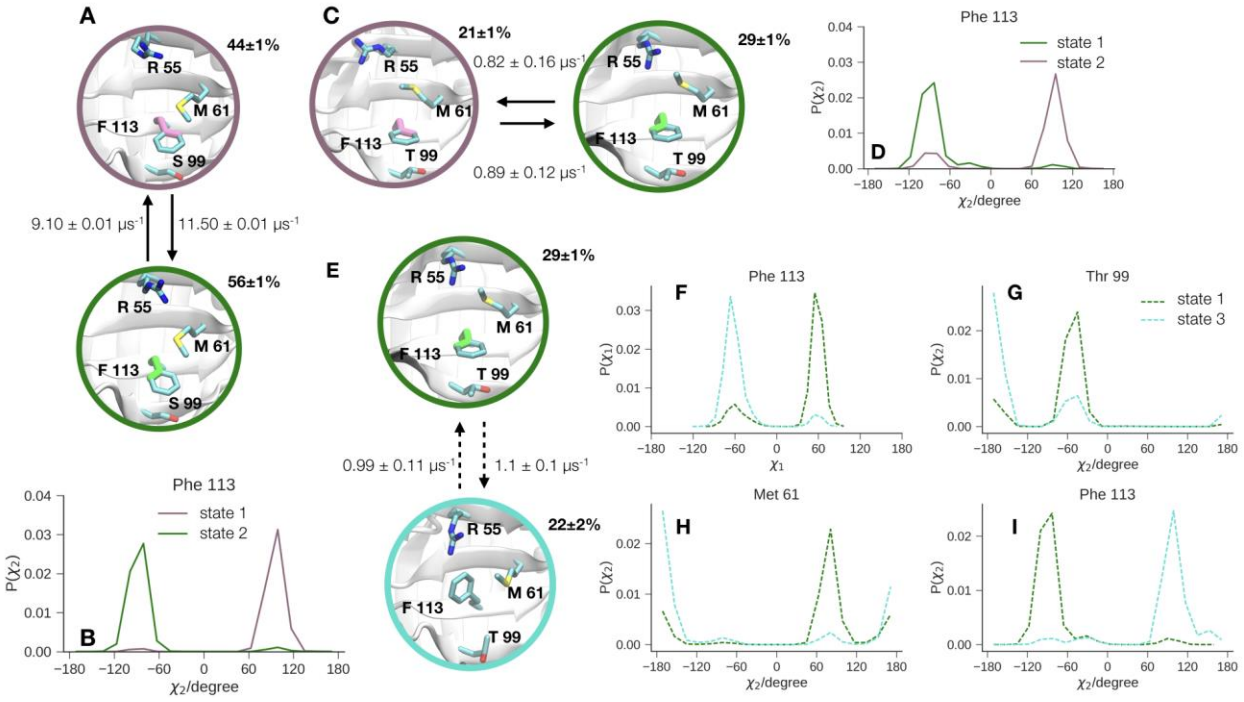
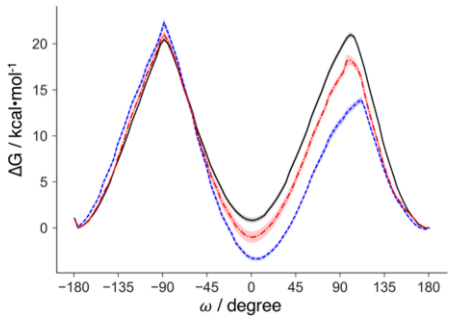


Figure 3. (**A**) Populations and exchanges rates of the two states WT MSM. The χ_2 torsional angle of Phe113 is depicted as thick sticks. **(B)** χ_2 probability distribution of Phe113 from the WT simulations showing that the slowest process involves rotation of Phe113 about this torsional angle. **(C)** Populations and exchange rates for the slowest process found in the S99T mutant MSM. **(D)** χ_2 probability distribution of Phe113 from the S99T simulations showing that the slowest process involves rotation of Phe113 about this torsional angle. (**E**) Populations and exchange rates for the second slowest process in the S99T mutant MSM. **(F-I)** Per state dihedral distributions highlighting the conformational changes occurring as part of the second slowest process in S99T mutant. Most striking the change in χ_1 of Phe113 defining the 'minor' and 'major' conformations.

A)



B)

Systems	$\Delta G^{\ddagger}$(cis→trans)	$\Delta G^{\ddagger}$(trans→cis)
AAPF	20.0±0.4	21.0±0.3
WT-AAPF	17.3±0.3	13.9±0.3
S99T-AAPF	19.2±0.8	18.2±0.5

Figure 4. (A) Isomerization free energy profiles (in kcal.mol^{-1}) for the substrate AAPF in water (black), and bound to WT (blue) or S99T mutant (red) forms of Cyp A. The free energies of the *trans* conformation were set to zero at $\omega = 180°$. Error bars represent one standard error of the mean. **(B)** The table shows the activation free energies for both directions of the isomerization reaction.

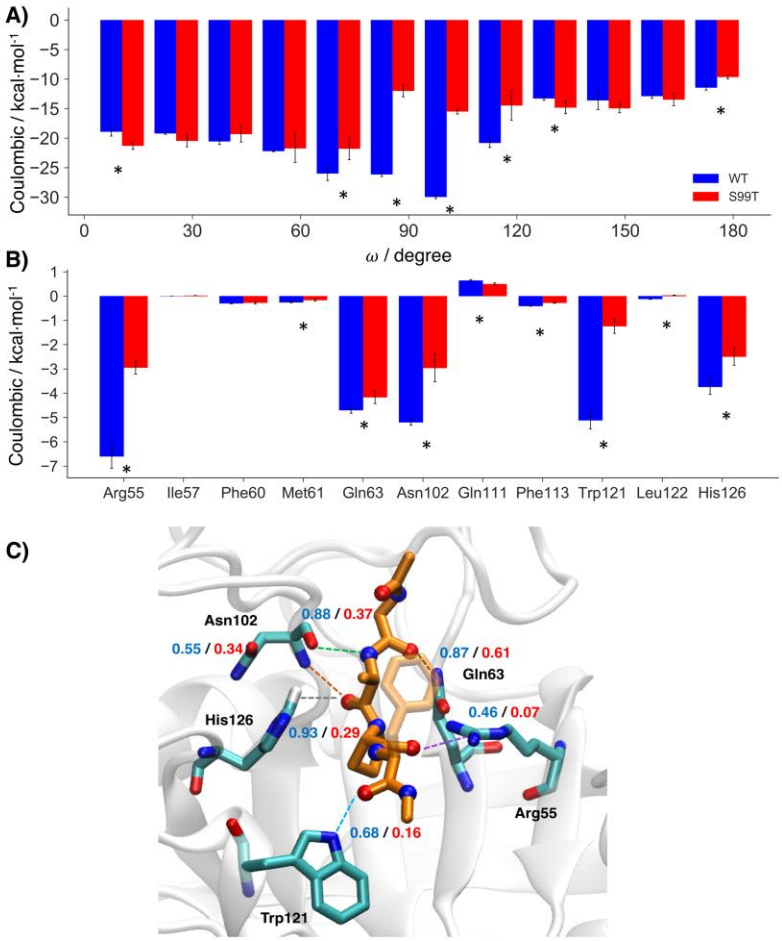


Figure 5. (**A**) Electrostatic energies between the substrate and active site residues as a function of the ω angle in WT (blue) and the S99T systems (red). (**B**) Average electrostatic energies per-active site residues at the transition state region. (**C**) Hydrogen-bonding probability for active site residues at the transition state ensemble. Error bars denote the standard error of the mean. The stars indicate differences in energies between WT and S99T greater than $2\sigma_E$.

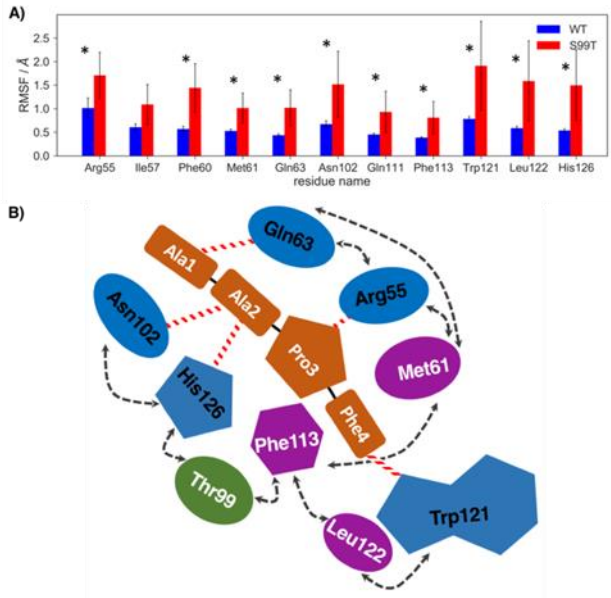


Figure 6. (A) The average root mean square fluctuations (RMSFs) of active site residues in WT (blue) and S99T (red) CypA at the transition state region. Error bars represent one standard error of the mean. Star symbols indicate differences in RMSF between WT and S99T greater than $2\sigma_E$. **(B)** Schematic depiction of the proposed model of allosteric inhibition. The Ser99Thr mutation increases fast vibrational motions of Phe113 and His126, leading to increased vibrational motions of neighboring residues (grey arrows) and thus decreased probability of strong hydrogen-bonding interactions (red dotted lines) with the substrate residues Ala1, Ala2, Pro3, Phe4.

Graphical TOC