Protein Structure

Deutsche Ausgabe: DOI: 10.1002/ange.201503698
Internationale Ausgabe: DOI: 10.1002/anie.201503698

A Structural Ensemble for the Enzyme Cyclophilin Reveals an Orchestrated Mode of Action at Atomic Resolution

*Celestine N. Chi, Beat Vögeli, Stefan Bibow, Dean Strotz, Julien Orts, Peter Güntert, and Roland Riek**

Abstract: *For enzyme activity, an exact structural and motional orchestration of the active site and its surroundings is believed to be key. In order to reveal such possible phenomena at atomic resolution on the basis of experimental evidence, an experimental restraint driven two-state ensemble of the prototypical enzyme cyclophilin was determined by using a recently introduced exact NOE approach. The ensemble description reveals the presence of an open and a closed state of cyclophilin, which is indicative of large-scale correlated motion. In the open state, the catalytic site is preorganized for catalysis, thus suggesting the mechanism of action to be conformational sampling, while the ligand-binding loop appears to act through an induced fit mechanism. This finding is supported by affinity measurements of a cyclophilin designed to be more open. Overall, more than 60–70 % of the side-chain conformations of cyclophilin appear to be correlated.*

The catalytic mechanisms of enzymes are believed to rely on a dynamic interplay between well-arranged structural states.[1] The magnitude of the conformational change may cover a large range in both space and time. The most relevant time scale for protein action is believed to be in the μs–ms range. Evidence of such dynamics has been found, for example, for the well-studied human cyclophilin A,[1a–c,2] a peptidylprolyl *cis–trans* isomerase. Peptidylprolyl *cis–trans* isomerases catalyze interconversion between the *cis* and *trans* isomers of the peptide bond of proline residue within a substrate.[3] For cyclophilin A, NMR relaxation experiments revealed ms motions both during catalysis and in the apo state that can be interpreted as a two-state interconversion process.[2c–e,4] In combination with room temperature X-ray crystallography[2d] and mutagenesis studies,[5] it has been suggested that the presence of a dynamic network encompassing the active site and its close neighborhood is key for activity. This finding has been complemented by a proposed mode of action of cyclophilin A derived from molecular dynamics simulation restricted by NMR restraints in combination with density functional theory calculations.[6] The calculations indicate that cyclophilin A acts through an electrostatic handle mechanism at the carbonyl of the residue preceding the proline in the substrate.

The traditional approaches dedicated to elucidating such slow conformational dynamics are R1ρ NMR relaxation measurements and fluorescence-based techniques.[7] However, it is difficult to represent the spatial sampling of these slow motions.[1d] New methodologies combining NMR probes with molecular dynamics simulations are being advanced to unravel this problem,.[1d,6] Recently, we introduced another concept that makes use of exact Nuclear Overhauser Effect (NOE)-derived distance restraints.[8] In order to obtain a plausible description of the various substates of cyclophilin A at atomic resolution, we employed an ensemble structure calculation with the use of eNOEs and residual dipolar couplings (RDCs).

Following an established protocol[8] with the software packages eNORA[9] and CYANA,[10] ensemble structure calculations were performed with a total of 3629 eNOE-based distance restraints, 396 H-N RDCs derived from four alignment media, 279 scalar couplings, and 128 angle restraints from $^{13}C\alpha$ chemical shifts (Table S1 in the Supporting Information). As a measure of the quality of the calculated structures, the CYANA target function, which is a weighted sum of all squared violations of the experimental restraints, is used. It drops significantly from one state to two states and levels off after three states (Figure 1c). This observation indicates that, in contrast to the single-state structure, multistate ensembles describe the experimental data well (Figure 1c and Table S1). In order to test for self-consistency of the experimental data, a cross validation test was performed with a jackknife procedure that repeats structure calculation ten times with 10 % of the experimental input data randomly deleted such that each distance restraint is omitted exactly once. The back-calculated target function of the omitted data then represents the entire data set. The decrease in this target function for higher-state ensembles (Figure 1c) indicates again that the experimental data are well described by two or more states. Similar cross-validations were also done with the RDCs and the $^{3}J_{HNHA}$ couplings (Figure 1d and 1e). Again, a significant and a moderate drop in the target function values for the $^{3}J_{HNHA}$ couplings and the RDCs is observed when increasing the number of states from one to two.

As a representative for the following discussion, the two-state ensemble described by a structural bundle of 2×20 conformers (PDB ID: 2n0t, Figure 1a,b) is used in order to

[*] Dr. C. N. Chi, Dr. B. Vögeli, Dr. S. Bibow, D. Strotz, Dr. J. Orts, Prof. Dr. P. Güntert, Prof. Dr. R. Riek
Laboratory of Physical Chemistry, ETH Zurich
ETH-Hönggerberg, 8093 Zürich (Switzerland)
E-mail: roland.riek@phys.chem.ethz.ch
Prof. Dr. P. Güntert
Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt am Main
Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201503698.

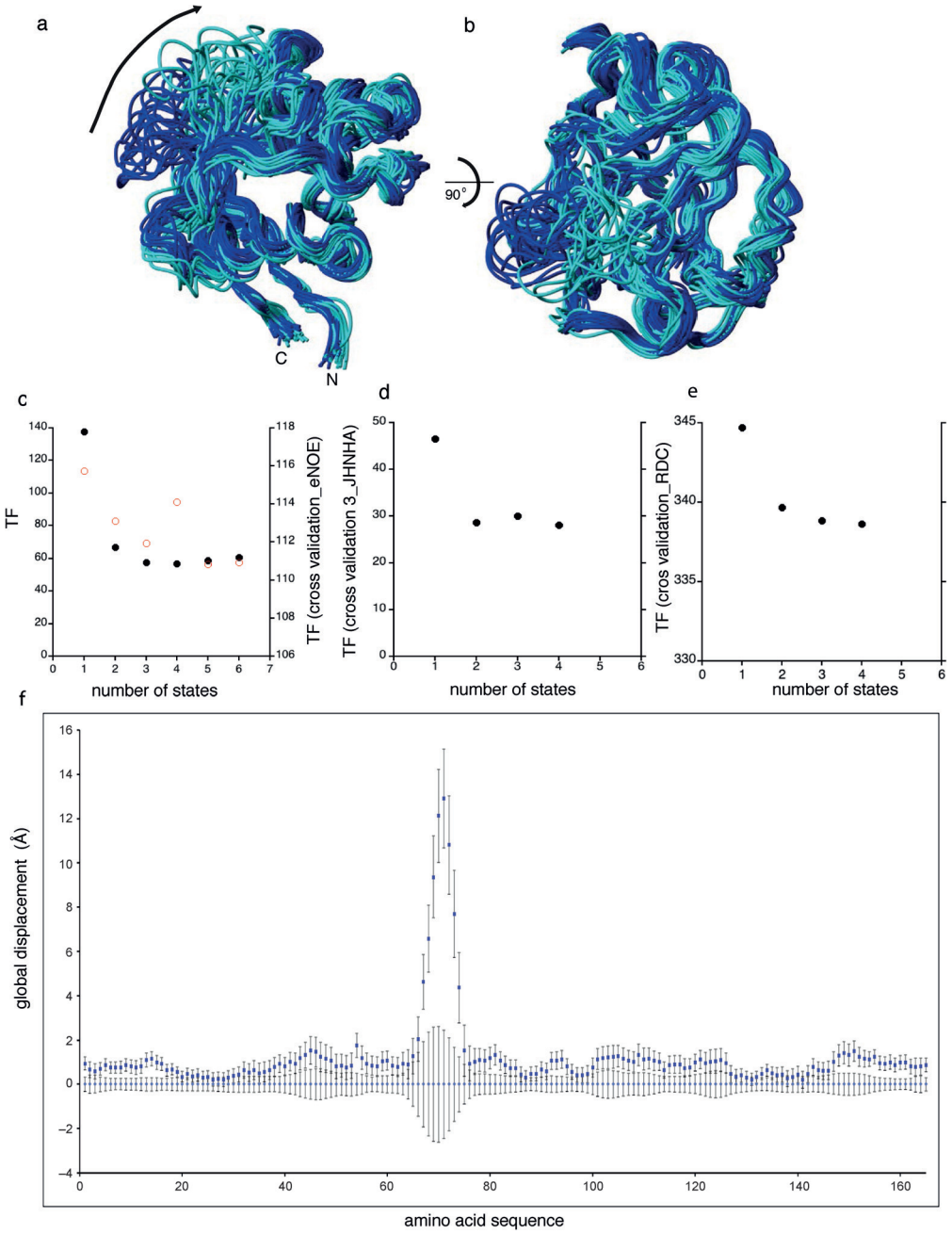

Figure 1. Structural ensemble of cyclophilin A in its apo form, highlighting the presence of two distinct states. a, b) Backbone trace of 20 structural ensembles, each representing two different states. States were color-coded as open (blue) and closed (cyan). Two distinct states are observed throughout most of the structure. The orientation shown in (b) is a 90° rotation of (a) as indicated. c) The dependence of the CYANA target function (TF, black) and the overall TF from the jackknife-type cross-validation (red) discussed in the main text are shown as functions of the number of states. There appears to be an outlier for the four-state ensemble of unknown origin. d, e) The dependence of the jackknife-type cross-validation TF of one RDC data set left-out calculation (i.e., the RDC data set obtained from negatively charged PEG media was left out) and left-out calculations of the $^{3}J_{HNHA}$ couplings, respectively, are shown as functions of the number of states. For (c), (d), and (e), the TF values drop from one to two states and level off thereafter, thus suggesting that the two-state representation of the ensembles reflects the experimental data. f) The backbone global displacement between the mean structure of the two substates is plotted against the amino acid sequence. The error bars are the weighted root mean square deviations (RMSD) of the individual substates of the backbone atoms of the residue of interest and its neigbouring residues. The RMSD was calculated with MolMol.

prevent over-fitting of the data (i.e., minimalizing the number of degrees of freedom) and because the same two states are also observed in the higher-state structure calculations (compare Figure 1 and Figure 2 with Figures S3 and S9 in the Supporting Information). It is important to note that the latter finding excludes the potential issue that the presence of two states might be an artifact of constraining the number of states to two in the structure calculation. Moreover, the deletion of 120 distance restraints violated in the single state structure calculation (Table S3) resulted in a loss of two distinct structural states in the two-state structure calculation (Figure S5b).

Inspection of the Cyclophilin A two-state ensemble reveals several remarkable features. Most prominently, the ligand-binding loop comprising residues 64–74 samples two spatially well-separated states (Figure 1a, b, f, Figure 2 and Table S2). The two states are referred to as "open" and "closed" (Figure 1 and Figure 2) because the closed state (cyan) is slightly more compact (average total surface area: 8937 Å^2 calculated by the program MolMol[11]) compared to the open state (blue; average total surface area: 9211 Å^2; Figure 1). The two states are also distinct at the active site and in surrounding regions, thus indicating long-range correlations (Figure 1–3 and Table S2). In the backbone, the two states are considerably different for residues 9–16, 34–42, 54–57, 64–78, 89–94, 101–107,

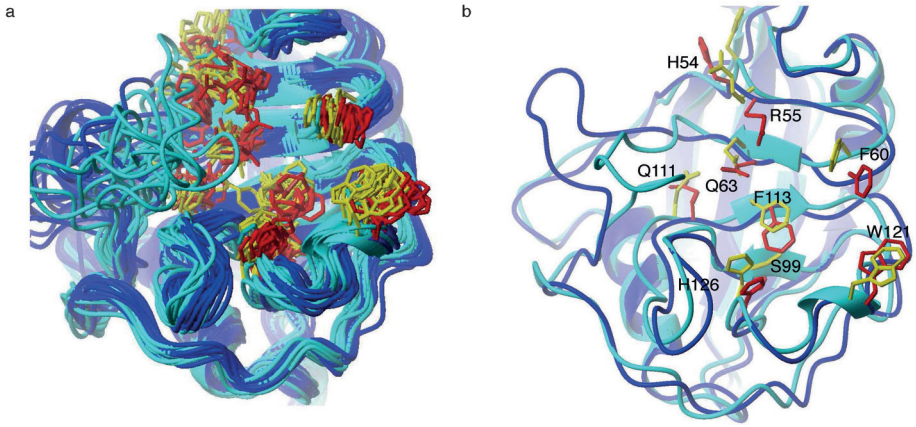

Figure 2. The two-state ensemble of the active-site residues of cyclophilin. a) A ribbon representation of the 2×20 structural ensembles is shown color coded individually for the two conformational substates: the closed state is shown in cyan for the backbone and yellow for the side chains of the active site, while the closed state is shown in blue for the backbone and red for the side chains in the active site, respectively. b) A single representative of (a) with the side chains labeled. The lowest energy two-state conformers were selected.

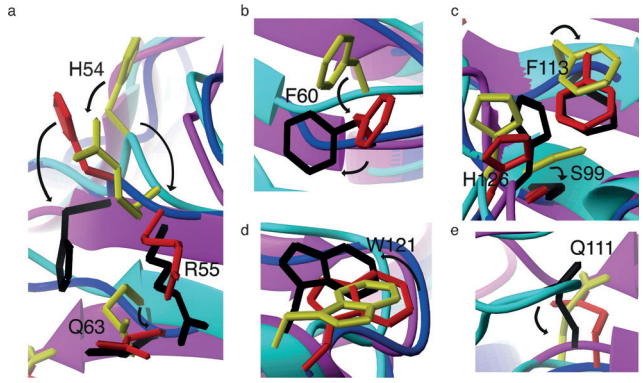

Figure 3. Proposed mechanism of action of cyclophilin at atomic resolution. The X-ray structure of Cyclophilin A in complex with the HIV-1 capsid protein (PDB ID: 1ak4) was superimposed with the presented two-state ensemble, which highlights the fact that the open state matches the ligand-bound state well. The closed state is shown in cyan for the backbone ribbon and yellow for the side chains, the open state is shown in blue for the backbone ribbon and red for the side chains, and the X-ray structure is shown in magenta for the backbone ribbon and black for the side chains. Individual close-ups of the superposition are shown. The potential modes of action for catalysis of the individual residues are indicated by arrows.

and 118–127. The presence of two distinct interchanging states at these locations concurs with documented slow conformational exchange measured by NMR relaxation data both in the apo state and during catalysis (Figure S1).[2c,e]

In concert with the backbone, many side chains show two distinct states. This is primarily caused by a propagation of different peptide plane orientations into the side chains since the side-chain rotamer angles between the two states are similar (Figure S7). Of particular interest is that also the side chains at the active site show two distinct states, that is, the side chains H54 R55, F60, Q63, S99, Q111, F113, W121 and H126 (Figure 2, Figure 3, and Table S2). These include side chains important for ligand binding (i.e., H54, F60, Q111, F113, W121 and H126[12]), as well as side chains essential for activity (i.e., Q63, S99, F113, and R55[5]). The presented two-state correlation of the latter side chains resolves the proposed activity-related dynamic network at atomic resolution,[2d] which guides the charged side chain of R55 into position to create an electrostatic potential that acts on the carbonyl group of the proline-preceding residue of the ligand.[6] Interestingly, the orientations of the discussed side chains of the open state closely match those of the crystal structure of cyclophilin A in complex with the HIV-1 capsid protein (Figure 3; PDB ID: 1ak4.[12]), thus highlighting that cyclophilin A closely samples the ligand-bound state at the active site in its apo state. This finding implies that the mechanism of conformational sampling[13] is key for its ligand recognition as suggested earlier.[2c,e] By contrast, the ligand-binding loop comprising residues 64–74 is not in its ligand-bound state. Thus, an induced fit mechanism[7,14] must prevail for this part of the protein upon complex formation (Figure S2).

In contrast to the open state, the closed state appears to be far from the active ligand-bound conformation. The loop residues 64–74 penetrate the space the bound ligand would cover (Figure S2) and the side chains of R55, Q63, F113, and S99 are not well positioned for activity (Figure 3 and Figure S2). Overall, large conformational differences between the two states are observed for residues 64–74, the side chains of H54, F60, S99, Q111, W121, and H126, and the side-chain conformation and configurations of R55 (Figure 2 and Figure 3). Furthermore, subtle dissimilarities in the side-chain conformations and configurations are apparent in 60–70% of all the side chains that are not solvent exposed, thus indicating that the differences between the two states are spread over almost the entire protein structure (Figure S4). Interestingly, each of the individual states of the six-state ensemble can be grouped to one of the two substates determined in the two-state ensemble (Figures S3b and S9). This finding allows a rough estimate of the population of the two states by counting the number of each of the two states in the six-state structure calculation (Figure S3b). With this analysis, the closed state appears to be less populated (between 15–40%) than the open state, which is in line with previous findings based on relaxation measurements (i.e., ca. 15% population at 10 °C[2c]). In summary, from the observation that cyclophilin A interchanges on the ms timescale as deduced from a single set of cross-peaks in the spectra,[2e] relaxation measurements,[2e] and the presented two-state ensemble, the presence of large-scale concerted motion between two states is suggested for the apo form of cyclophilin A.

Based on the two-state ensemble presented herein, as well as other data,[5,6,12] it is suggested that cyclophilin A interchanges between a closed and an open substate on the millisecond timescale both in its apo state and during catalysis. Owing to the intrinsic exchange between the two states, there is a time window during which the ligand can bind to the open state. In detail, the loop opens up accompanied by an associated outward movement of H54 and an inward movement of R55 to its place for catalysis. F113 turns about 60° degrees downwards and S99 rotates 180° downwards (Figure 3a,c). F60, W121, and H126 flip outwards, while Q63 moves upward to keep R55 in place (Figure 3a,b,d,e). These rearrangements, accompanied by many subtle changes throughout most of the protein structure, create the ligand-binding cavity, position the active site residue R55 for activity, and prepare the active site for ligand binding. Upon binding of the ligand, only small, subtle side-chain rearrangements at the active side occur (Figure 3). The switch back to the closed state requires release of the peptide by a mode of action reminiscent of a catapult, with the ligand-binding loop being the handle. In part, this process may be entropy driven because of the shrinking of the protein surface accompanied by conformational changes in the side chains at the active site.

In order to find further support for the role of the loop motion in protein activity and the presence of two states, we attempted to

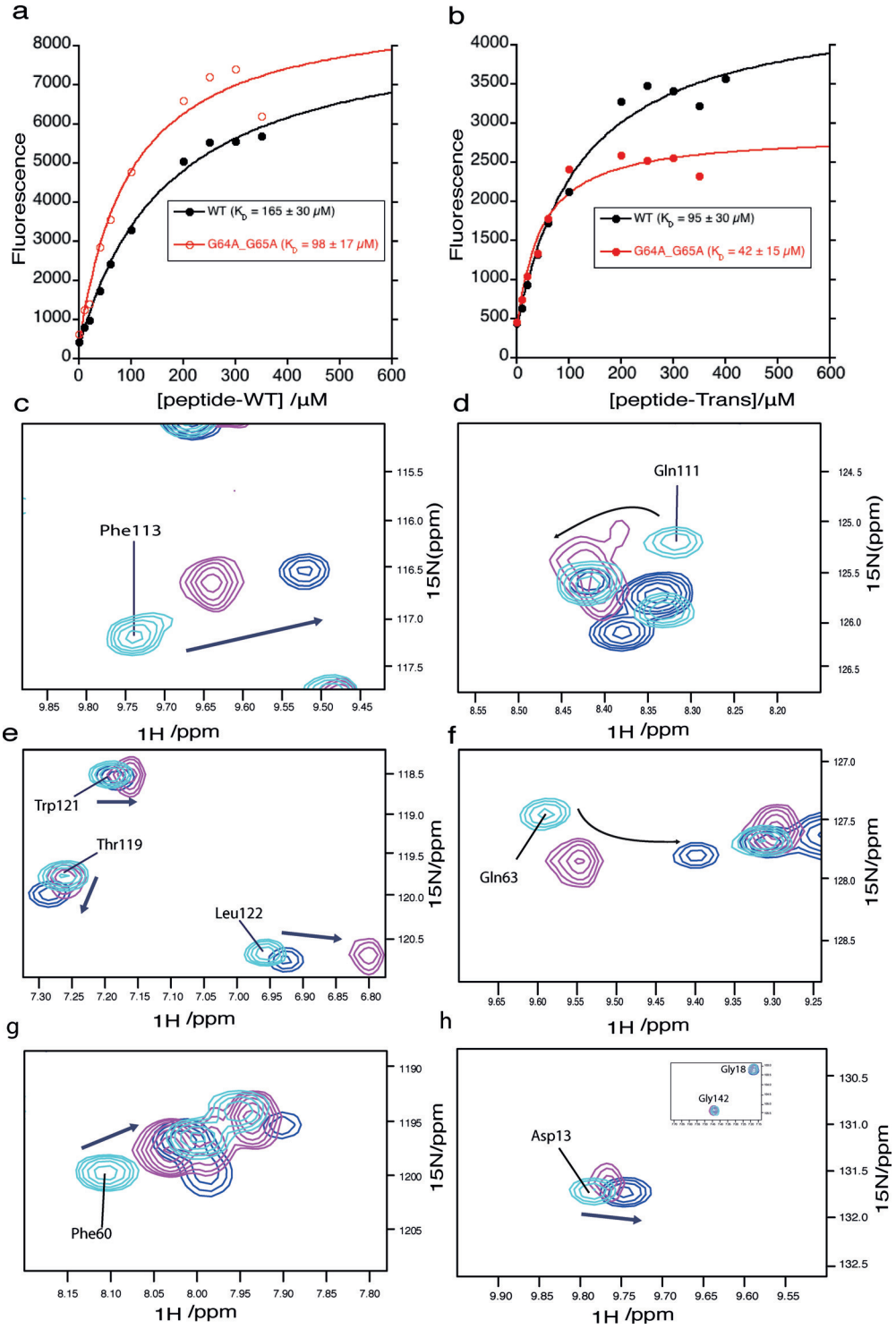

Figure 4. A cyclophilin double mutant shifts the equilibrium towards the peptide-bound form. a,b) For determination of the binding constant, wild-type (WT) cyclophilin A and its double mutant, denoted G64A G65A, were mixed with peptide ligands (in (a) with wild-type and in (b) with the *trans* peptide) at various concentrations as described in the Supporting Information. The ligand binding was measured by fluorescence.[15] The affinity increased by almost two-fold for the double mutant compared to wild-type cyclophilin for both peptides. c–h) Details of [^{15}N,^{1}H]-TROSY spectra (shown in Figure S8) of free wild-type (cyan) cyclophilin, cyclophilin in presence of a four-fold excess of ligand (purple), and the double mutant (blue). Cross-peaks of residues at or close to the active site are labeled. The changes in chemical shift introduced either by mutagenesis or by the addition of the ligand are indicated by arrows. In the insert of Figure 4 h, cross-peaks of Gly18 and Gly142 are shown that do not experience any change in position upon addition of ligand or upon mutagenesis. The finding that the double mutant promotes chemical shift changes in a similar direction as when the ligand is added indicates that the equilibrium population of the glycine double mutant is shifted towards the ligand-binding open form compared to free cyclophilin A.

lock the loop into the open state by replacing the Gly residues of the N-terminal hinge of the peptide-binding loop (i.e. Gly64-Gly65) with the more conformationally restricted residue Ala. It is thus expected that the double mutant is close to a ligand-bound state (open state) and concomitantly should show increased activity. Indeed, the well-folded double mutant (Figure S8) shows an increased affinity for both peptide ligands studied when compared to wild-type cyclophilin A (Figure 4a,b). Furthermore, the mutation introduces chemical shifts of ^{15}N–^{1}H moieties in the same direction as the addition of ligand to a sample of wild-type cyclophilin A (Figure 4c–h), thus indicating that the mutant is indeed in a more open state than free wild-type enzyme.

The presented ensemble structure calculation of the apo state of the enzyme cyclophilin, complemented by mutagenesis and affinity measurements, reveals a long-range well-orchestrated conformational interchange between substrates important for its catalytic activity and highlights a synergistic induced fit and conformational sampling mechanism of action. The complexity unraveled reflects the adaptation and optimization power of evolution, as well as the beauty of these types of biological machineries, which are composed of several hundreds of atoms that move in concert.

Acknowledgements

This work was support in part by the Swedish research council (Ventenskap Rådet VR-2011-13) and Wenner-Gren Stiftelsern WG-13, Postdoctoral fellowship to C. N. C, B. V. was supported by the Swiss National Science Foundation with Grant 140214 and by the ETH Research Grant ETH-04 13-1.

Keywords: biophysics · enzymes · enzyme mechanisms · NMR spectroscopy · protein structures

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 11657–11661
Angew. Chem. **2015**, *127*, 11823–11827

[1] a) K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964–972; b) K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus, D. Kern, *Nature* **2007**, *450*, 913–916; c) M. Wolf-Watz, V. Thai, K. Henzler-Wildman, G. Hadjipavlou, Z. E. Eisenmesser, D. Kern, *Nat. Struct. Mol. Biol.* **2004**, *11*, 945–949; d) O. F. Lange, N.-A. Lakomek, C. Fares, G. F. Schroder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmuller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471–1475.

[2] a) D. Kern, E. Z. Eisenmesser, M. Wolf-Watz, in *Methods in Enzymology, Vol. 394* (Ed.: L. J. Thomas), Academic Press, San Diego, **2005**, pp. 507–524; b) K. A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M. A. Wilson, G. A. Petsko, M. Karplus, C. G. Hubner, D. Kern, *Nature* **2007**, *450*, 838–844; c) E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, D. Kern, *Nature* **2005**, *438*, 117–121; d) J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern, T. Alber, *Nature* **2009**, *462*, 669–673; e) E. Z. Eisenmesser, D. A. Bosco, M. Akke, D. Kern, *Science* **2002**, *295*, 1520–1523.

[3] a) R. E. Handschumacher, M. W. Harding, J. Rice, R. J. Drugge, D. W. Speicher, *Science* **1984**, *226*, 544–547; b) E. M. de Villiers, D. Wagner, A. Schneider, H. Wesch, H. Miklaw, J. Wahrendorf, U. Papendick, H. Zur Hausen, *Lancet* **1987**, *330*, 703–706; c) G. Fischer, B. Wittmann-Liebold, K. Lang, T. Kiefhaber, F. X. Schmid, *Nature* **1989**, *337*, 476–478; d) R. T. Clubb, S. B. Ferguson, C. T. Walsh, G. Wagner, *Biochemistry* **1994**, *33*, 2761–2772.

[4] A. V. Finkelstein, O. B. Ptitsyn, *Protein Physics*, Academic Press, Amsterdam, **2002**.

[5] L. D. Zydowsky, F. A. Etzkorn, H. Y. Chang, S. B. Ferguson, L. A. Stolz, S. I. Ho, C. T. Walsh, *Protein Sci.* **1992**, *1*, 1092–1099.

[6] C. Camilloni, A. B. Sahakyan, M. J. Holliday, N. G. Isern, F. Zhang, E. Z. Eisenmesser, M. Vendruscolo, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10203–10208.

[7] a) C. N. Chi, A. Bach, A. Engstrom, H. Wang, K. Stromgaard, S. Gianni, P. Jemth, *Biochemistry* **2009**, *48*, 7089–7097; b) C. N. Chi, L. Elfstrom, Y. Shi, T. Snall, A. Engstrom, P. Jemth, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4679–4684.

[8] a) B. Vögeli, S. Kazemi, P. Güntert, R. Riek, *Nat. Struct. Mol. Biol.* **2012**, *19*, 1053–1057; b) B. Vögeli, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 1–46.

[9] J. Orts, B. Vögeli, R. Riek, *J. Chem. Theory Comput.* **2012**, *8*, 3483–3492.

[10] P. Güntert, C. Mumenthaler, K. Wuthrich, *J. Mol. Biol.* **1997**, *273*, 283–298.

[11] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 29–32.

[12] T. R. Gamble, F. F. Vajdos, S. Yoo, D. K. Worthylake, M. Houseweart, W. I. Sundquist, C. P. Hill, *Cell* **1996**, *87*, 1285–1294.

[13] J. Monod, J. Wyman, J. P. Changeux, *J. Mol. Biol.* **1965**, *12*, 88–118.

[14] D. E. Koshland, Jr., G. Nemethy, D. Filmer, *Biochemistry* **1966**, *5*, 365–385.

[15] C. N. Chi, A. Engstrom, S. Gianni, M. Larsson, P. Jemth, *J. Biol. Chem.* **2006**, *281*, 36811–36818.

Received: April 22, 2015
Revised: June 5, 2015
Published online: August 10, 2015

