

Structural insights into the catalytic mechanism of cyclophilin A

Bruce R Howard¹, Felix F Vajdos¹, Su Li¹, Wesley I Sundquist & Christopher P Hill

Cyclophilins constitute a ubiquitous protein family whose functions include protein folding, transport and signaling. They possess both sequence-specific binding and proline *cis-trans* isomerase activities, as exemplified by the interaction between cyclophilin A (CypA) and the HIV-1 CA protein. Here, we report crystal structures of CypA in complex with HIV-1 CA protein variants that bind preferentially with the substrate proline residue in either the *cis* or the *trans* conformation. *Cis*- and *trans*-Pro substrates are accommodated within the enzyme active site by rearrangement of their N-terminal residues and with minimal distortions in the path of the main chain. CypA Arg55 guanidinium group probably facilitates catalysis by anchoring the substrate proline oxygen and stabilizing *sp*³ hybridization of the proline nitrogen in the transition state.

CypA is the prototypical member of the widespread cyclophilin family of enzymes, which catalyze the *cis-trans* isomerization of peptide bonds preceding proline residues^{1,2}. This activity accelerates protein folding *in vitro*^{3–5} and may underlie some of the many roles of cyclophilins⁶, which include signaling, mitochondrial function, chaperone activity, RNA splicing, stress response, gene expression and regulation of kinase activity. The biological activities of CypA include binding the HIV-1 CA protein and facilitating viral replication. Interaction with the N-terminal domain of the HIV-1 CA protein (CA^N) results in incorporation of CypA into viral particles at a CypA:CA ratio of ~1:10 (refs. 7–9). This seems to be required for replication of all main (M) and some outlier (O) HIV-1 strains^{10,11}, although the basis for the role of CypA in viral replication remains unclear and the CypA isomerase activity is reportedly not required for HIV-1 infectivity¹².

The recent demonstration that HIV-1 CA^N is a substrate for isomerization by CypA *in vitro*¹³ makes it an attractive system to determine the structural basis for this activity. This is important because the isomerase activity is essential for at least some biological functions of cyclophilins^{14–16} and because the catalytic mechanism remains controversial. Our previously determined structure of the CypA–CA^N_{NL43} complex (at a resolution of 2.4 Å)¹⁷ and associated binding studies¹⁸ revealed that all contacts to CypA are contained within a short exposed loop centered on CA^N Pro90. Essentially identical interactions were observed for CypA complexes with CA-derived peptides^{19,20}. Unlike other CypA–peptide complexes^{21–24}, the CypA–CA^N structure bound in the *trans* conformation. This seemed to result from the ability of Gly89, the residue preceding the isomeric peptide of Pro90, to bind deeply into the CypA active site cleft, whereas larger side chains would be excluded from this arrangement while in the *trans* conformation¹⁷. Comparison of the CypA–CA^N structure with CypA–peptide complexes further suggested that catalysis proceeds by rotation of groups

N-terminal to the isomeric proline. More recently, however, an NMR study of chemical shift and NMR relaxation rate changes in the presence and absence of a model substrate led to the conclusion that catalysis by CypA is achieved by rotation of the substrate's C-terminal residues while the N-terminal residues remain stationary²⁵.

In an effort to gain a better understanding of the catalytic mechanism, we have determined crystal structures of CypA complexes with a series of CA^N variants. The structures have been refined at high resolu-

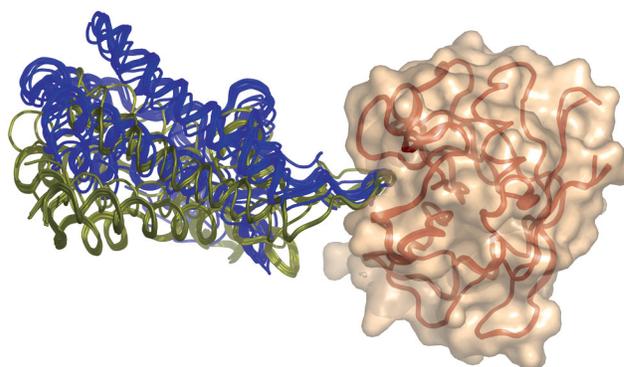


Figure 1 Overall structure of CypA–CA^N complexes. All 16 crystallographically independent structures reported in this paper are shown. Overlaps shown here and in all other figures were obtained by least-squares superposition of main chain atoms of CypA residues 3–165. The hinge angle depends on position in the asymmetric unit; A and A' CA^N molecules are colored green; B and B' CA^N, blue; CypA, red with a straw-colored molecular surface. The different hinge angles are accommodated by variation in ϕ and ψ angles over several residues either side of the CA^N Gly89–Pro90 peptide, with the largest change in the wild-type structure for CA^N Met96 ($\Delta\phi = 31^\circ$). **Figures 1, 2d** and **3** were made with PyMOL (DeLano Scientific; <http://www.pymol.org>). **Figures 2a–c** and **4** were made with MOLSCRIPT⁴⁰ and RASTER3D⁴¹.

Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, Utah 84132, USA. ¹These authors contributed equally to this work. Correspondence should be addressed to W.I.S. (wes@biochem.utah.edu) or C.P.H. (chris@biochem.utah.edu).

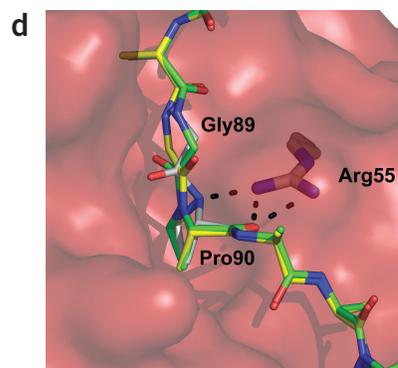
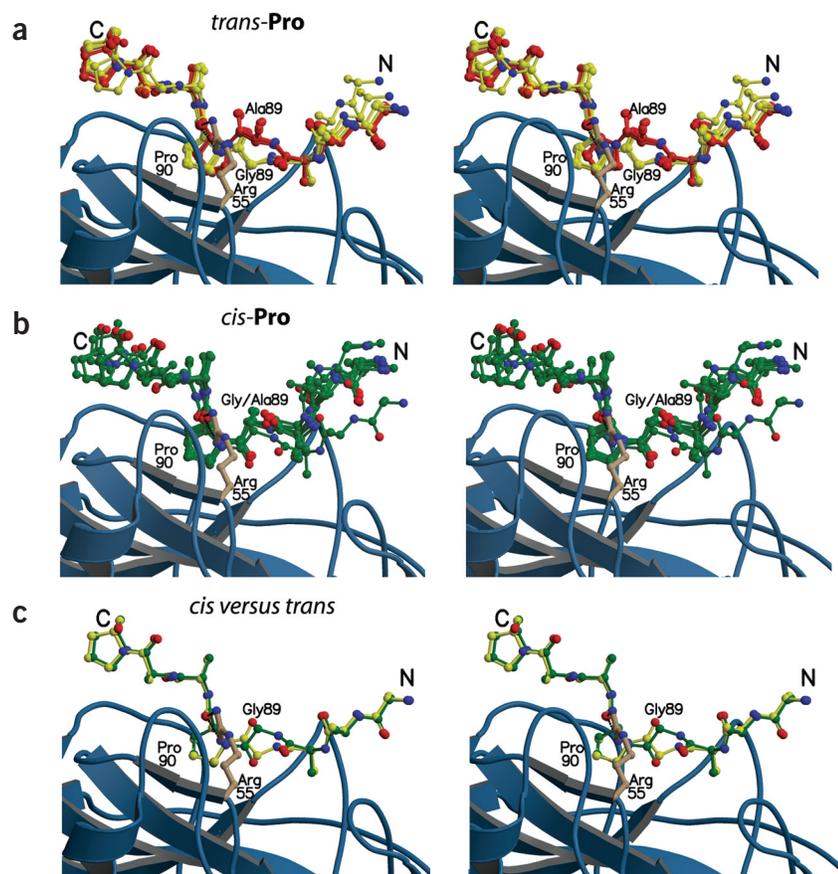


Figure 2 Comparison of CA^N loop conformations. CA^N residues 86–93 are shown as a stick representation with side chains truncated to the C β atom (except for proline) and carbon atoms colored yellow or orange (*trans*) and green (*cis*). For all figures, the minor (20% occupied) *cis* conformations of AMA-A and AMA-A' are not shown unless explicitly stated. CypA is shown in a ribbon representation with the Arg55 side chain shown explicitly. (a) Stereo view showing all eight CA^N structures that adopt the *trans* conformation. The four structures that contain Gly89 are colored yellow; the four Ala89 structures are colored orange. (b) Same as a but for all eight *cis* CA^N structures. (c) Comparison of AAG-A (*trans*, yellow) and AMG-A (*cis*, green). (d) Same as c, but top view. CypA molecular surface colored red. A model for the transition state is shown with the carbon atoms colored white. Hydrogen bonds between CypA Arg55 and CA^N N and O atoms are shown as dashed lines.

tion and include both *cis* and *trans* conformations. In one case, both conformations are seen in the same asymmetric unit with partial occupancy. The crystal structures support a mechanism in which substrate residues C-terminal to the isomeric proline remain stationary.

RESULTS

CypA–CA^N structures

We have refined crystal structures of the wild-type (WT) and five variant CA^N complexes with CypA that present views of 16 crystallographically independent complexes (Table 1) at resolutions of 1.7–2.0 Å and R_{free} of 22–26%. The crystals adopt space group $P2_1$ with two complexes in the asymmetric unit, except for two of the variants, which adopt space group $P1$ with four complexes in the asymmetric unit. The two crystal forms are very closely related: the $P1$ crystals differ from perfect $P2_1$ symmetry by rotation of $<1^\circ$ and translation of <0.5 Å between pseudo-equivalent molecules. The variants are named

Table 1. Constructs crystallized

	CA ^N sequences (residues 82–97)	A ^a	B	A'	B'
WT(HAG)	–R L H P V H A G P I A P G Q M R–	^b <i>T</i>	<i>T</i>		
AAG	–R L H P V <u>A</u> A G P I A P G Q M R–	<i>T</i>	<i>T</i>		
AMG	–R L H P V <u>A</u> <u>M</u> G P I A P G Q M R–	<i>C</i>	<i>C</i>		
AAA	–R L H P V <u>A</u> <u>A</u> <u>A</u> P I A P G Q M R–	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>
AMA	–R L H P V <u>A</u> <u>M</u> <u>A</u> P I A P G Q M R–	<i>T/C</i> ^c	<i>C</i>	<i>T/C</i> ^c	<i>C</i>
O-loop	–R <u>T</u> H P <u>P</u> <u>A</u> <u>M</u> G P <u>L</u> <u>P</u> P G Q <u>I</u> R–	<i>C</i>	<i>C</i>		

^aA, B, A', B' denote different CypA–CA^N complexes in the asymmetric unit. ^b*T*, *trans*; *C*, *cis*. ^cAMA-A and AMA(A') are predominantly *trans* with a minor *cis* conformation.

for their sequences (Table 1), and crystallographically independent structures are designated A or B for their position in the asymmetric unit, with the $P1$ crystals also containing A' and B' complexes.

The high-resolution structure of the M-type HIV-1_{NL43} CA^N–CypA complex is essentially identical to the medium-resolution structure reported earlier¹⁷, with the central CA Gly89–Pro90 peptide in the *trans* conformation. In contrast, the chimeric O-loop CA^N construct, in which the CypA-binding loop was replaced by the O-type HIV-1_{MVP5180} sequence, bound in the *cis* conformation (Table 1). To investigate this unexpected observation further, we made a series of chimeric loops in which O-type residues were substituted into the original M-type CA^N protein. This strategy allowed us to obtain single-residue variants that bind either in the *cis* or the *trans* conformation (Table 1).

In all structures reported here, the inherently flexible CypA-binding loop of CA (ref. 26) is ordered, with the central Pro90 residue adopting B -values that are comparable to those of surrounding residues in the CypA active site. The ensemble of structures includes cases in which both *cis*- and *trans*-Pro conformations occur in the same crystal lattice and even in equivalent asymmetric units (that is, as partial occupancy). This supports the assumption that the crystal structures represent on-pathway ground-state structures of the CypA-catalyzed isomerization. As expected from the database of known structures²⁷, all of the *cis*-Pro90 side chains adopt the *endo* pucker, whereas the *trans*-Pro90 side chains are found in both *endo* and *exo* pucker. Also as expected, the pucker of *trans*-Pro90 residues correlates with their ϕ angle; structures containing alanine at residue 89 have a relatively less negative ϕ angle and *exo* pucker, whereas the other *trans* structures

have a more negative ϕ angle and endo pucker. Despite these differences at the isomeric proline, all of the structures are quite similar, with the largest global difference corresponding to rigid-body rotation of the CA^N protein by $\sim 20^\circ$ in a hinge motion about the ends of the cyclophilin-binding loop (Fig. 1). This apparent motion does not significantly alter CA^N-CypA contacts, nor does it strongly correlate with *cis* or *trans* conformations. Rather, the hinge angle is determined by different packing interactions for the independent complexes in the asymmetric unit: all A and A' asymmetric unit complexes adopt one orientation, whereas all B and B' complexes adopt the alternate orientation.

Comparison of *cis* and *trans* conformations

An important finding from the crystal structures is that, relative to the central CA^N Pro90 residue, the C-terminal CA^N segment (Pro90–Pro93) adopts nearly the same conformation in all structures regardless of *cis* or *trans* conformation (Fig. 2). The largest difference between any of the 16 structures in a main chain ϕ or ψ angle over these residues is only 35° . The corresponding shifts in atomic positions are accommodated by variation in CypA–CA^N hydrogen bond lengths and solvent structure, including substitution of a direct hydrogen bond for a solvent-mediated interaction. The subsequent displacement of C-terminal residues seems to result from different positions of the loop ends caused by the variable CA hinge angle, and is allowed by the flat, open structure for this part of the active site cleft.

The situation is markedly different for residues N-terminal to CA^N Pro90. Although the overall path of the polypeptide is mostly unchanged, the switch between *cis* and *trans* conformations alters the main chain conformation for CA^N residues 88 and 89. The best overlap between *cis* and *trans* structures is observed for AAG-A (*trans*) and AMG-A (*cis*) (Fig. 2c). In going from *cis*- to *trans*-Pro90, CA^N residue 88 ψ angle and 89 ϕ and ψ angles change by 25° , 125° and 27° , respectively, whereas all other individual ϕ and ψ angles differ by no more than 10° for CA^N residues 75–105. Consequently, the structures overlap very closely both N- and C-terminal to the substrate proline, and the only substantial differences are confined to CA^N residues immediately N-terminal to the isomeric peptide. Conformational changes do not propagate to C-terminal CA^N residues because CypA binds Pro90 in a hydrophobic pocket and anchors the proline oxygen with two hydrogen bonds to the guanidinium of the essential CypA residue Arg55 (refs. 18,23,28) (Fig. 2d).

It is not obvious why the single substitution of A88M in CA^N results in a switch from *trans* to *cis* conformations for the AAG and AMG structures, because the CA^N residue 88 C α atoms are within 0.08 Å of each other after least-squares overlap on CypA structures. The switch may result because the CA^N Met88 side chain adopts a slightly ($\sim 10^\circ$) altered orientation on the CypA surface, and presumably reflects the fine balance in energy observed for the bound *cis* and *trans* conformations^{5,13}.

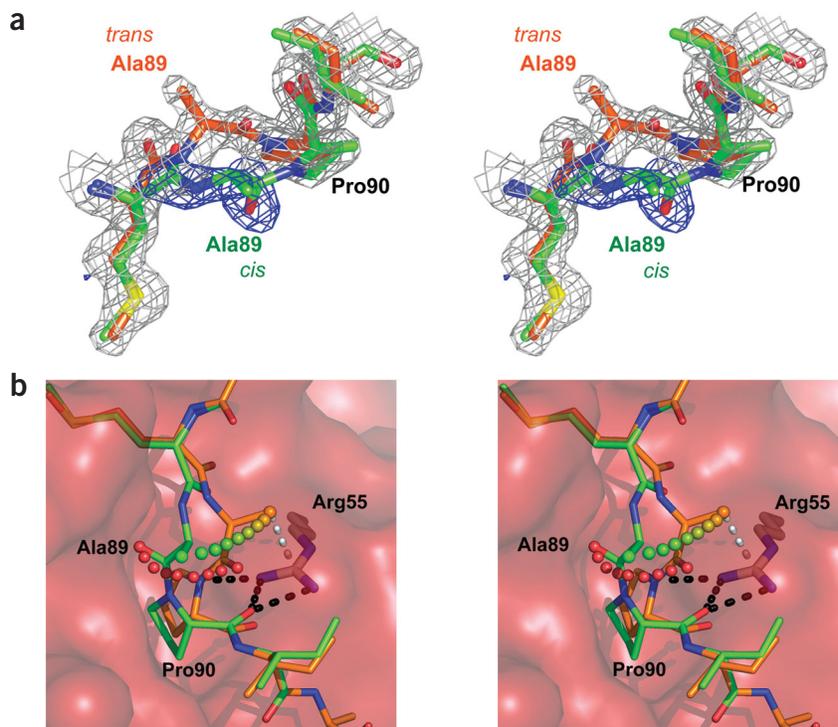


Figure 3 Proposed reaction pathway. (a) Mixed *trans* (80%) and *cis* (20%) structures of AMA-A. Maps were calculated before inclusion of the minor *cis* conformation in the model. $2F_o - F_c$ (silver) and $F_o - F_c$ (blue) maps are contoured at 1.0 and $2.0 \times$ r.m.s. deviation, respectively. Final refined coordinates for the two partially occupied conformations are shown in orange (*trans*) and green (*cis*). (b) Top view of AMA-A *trans* (orange, 80% occupied) and *cis* (green, 20% occupied) conformations. Series of red and orange/green spheres show path of CA^N Ala89 O and C β atoms for intermediate conformations. This path would keep the Ala89 side chain clear of CypA protein and maintains a staggered conformation. White dashed line; contact between the side chain of CA^N Ala89 and CypA Arg55 that prevents CA^N Pro90 from binding fully into the active site when in the *trans* conformation. Black dashed lines represent the hydrogen bonds between CypA Arg55 and CA^N Pro90 O that prevent propagation of conformational changes to C-terminal residues and the hydrogen bond to CA^N Pro90 N that promotes catalysis.

Proposed reaction pathway

We favor a reaction pathway that requires minimal deviation from the ground-state crystal structures. Specifically, we propose that the proline remains essentially fixed relative to the enzyme, whereas the oxygen of the preceding residue rotates 180° in a clockwise direction (moving *cis* to *trans*) when viewed from the proline nitrogen atom (Fig. 2d). Catalysis will be enhanced by variables that disfavor the double-bond character of the peptide bond and favor the less polar and freely rotating single bond²⁹. Thus, desolvation and the absence of compensating interactions for the peptide partial charges presumably enhance CypA catalysis by destabilizing the ground states. Additionally, formation of a hydrogen bond between Arg55 and the proline nitrogen would stabilize a pyramidal sp^3 hybridization state for the CA^N Pro90 nitrogen atom and resulting single-bond character for the peptide^{23,30}. The distance between CypA Arg55 N η 1 and CA^N Pro90 main chain N atoms ranges from 3.3 to 4.4 Å in the ground-state complexes reported here, and is expected to shorten in the transition state to form a hydrogen bond between CypA Arg55 and the pyramidal sp^3 -hybridized CA^N Pro90 N atom (Fig. 2d)²³.

A steric contribution to catalysis is also suggested by our structures of CA^N G89A complexes. Our decision to mutate CA^N Gly89 to alanine in the AAA construct was prompted by the observation that *trans* structures bind with CA^N Gly89 deep in the CypA active site cleft,

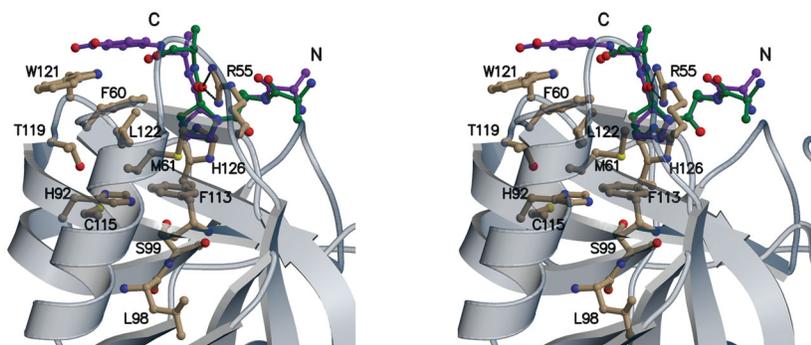


Figure 4 Superposition of CypA complexes with the *cis* conformations of CA^N and Suc-Ala-Ala-Pro-Phe-NA (ref. 23). Residues 88–91 of CA^N AAA-B, green; Suc-Ala-Ala-Pro-Phe-NA, purple. Both ligands are shown truncated to C β , except for proline and the tetrapeptide C-terminal nitrophenyl group. Residues of CypA that presumably move to allow the *trans* peptide substrate to contact Leu98 and Ser99 (bottom of figure) are shown explicitly²⁵. These CypA residues and the protein substrate residues C-terminal to the isomeric proline are essentially unchanged in the *cis* and *trans* complexes of CA.

apparently with no space for a side chain and with (ϕ , ψ) angles (+123°, +168°) that are favored only for glycine¹⁷. Consistent with this, the CA^N G89A substitutions weaken the binding considerably; the K_d increases from ~20 μ M to ~500 μ M (ref. 18), with only slight variation in these values observed in a variety of sequence contexts (data not shown). The G89A substitution also shifts the equilibrium of the bound AAA species toward the *cis* conformation, with the result that both *cis* and *trans* structures are present in the same crystal (Table 1) and the *trans* conformation buries less deeply into the active site.

Remarkably, the AMA-A and AMA-A' complexes show a mixture of 80% *trans* and 20% *cis* conformations, each of which show close similarity to the fully occupied single conformation structures (Fig. 3a). Contact of the CA^N *trans*-Ala89 C β with CypA Arg55 (absent in the Gly89 structures) seems to hold the CA^N Pro90 side chain ~1.5 Å away from its preferred binding positions for the Gly89 and *cis*-Ala89 structures (Figs. 2a and 3b). Thus, for a non-glycine residue preceding the substrate proline, movement from the *trans* conformation toward the transition state would relieve the steric clash with CypA Arg55 and allow optimal binding of the proline side chain (Fig. 3b).

Curiously, substitution of Gly89 to alanine in the AMA construct prompted a partial shift from the *cis* toward the *trans* conformation (Table 1). Comparison of the structures suggests that this results from intramolecular interactions within the CA^N protein loop. The ψ angle of Gly89 is unfavorable for non-glycine residues at this position in the *cis*-AMG structure (where $\psi = -163^\circ$ and -176° for AMG-A and AMG-B, respectively). This potentially strained conformation is relieved in AMA-A/A' by partial switch to *trans* and by smaller adjustments to a ψ angle to +158° in *cis*-AMA-B/B'. The structures are therefore consistent with our earlier finding that CypA is a Gly-Pro-specific binding protein and support our proposal that, conversely, catalytic rates may be greatest for suboptimal binding sequences such as Ala-Pro^{17,18}.

Comparison to earlier proposals

The catalytic pathway proposed here for the CA protein substrate differs from the pathway for peptide substrates reported recently²⁵ on the basis of chemical shift and NMR relaxation

rate changes of main chain amides in the presence and absence of a model tetrapeptide substrate, Suc-Ala-Phe-Pro-Phe-NA, where Suc is succinyl and NA is *p*-nitroanilide. Catalytic pathways for both substrates seem to involve the same *cis* conformation, because all known CypA-*cis*-peptide structures overlap closely with the CypA-*cis*-Pro90 CA^N structures described here, including a Suc-Ala-Ala-*cis*-Pro-Phe-NA substrate²³ that is closely related to the substrate used by Eisenmesser *et al.*²⁵ (Fig. 4). There are important differences, however, upon moving to the *trans* conformation. As described above, the CA^N protein-substrate complex shows minimal conformational changes that are localized N-terminal to the isomeric proline. In contrast, the observation of catalysis-correlated relaxation rate changes at the main chain nitrogen atoms of CypA residues Leu98 and

Ser99 indicated that formation of the *trans* Suc-Ala-Phe-Pro-Phe-NA peptide complex involved substantial conformational changes in substrate and enzyme that position the substrate's C-terminal residues adjacent to CypA Leu98 and Ser99 (ref. 25). Curiously, the different pathway for Suc-Ala-Phe-Pro-Phe-NA does not seem to be simply the result of using a peptide substrate rather than protein, because structures of CypA bound to CA-derived hexapeptides^{19,20} overlap closely with the *trans* CypA-CA^N complexes.

A recent molecular dynamics study³¹ of four tetrapeptides proposed the same transition-state geometry suggested here for the CA substrate. The crystallographic and dynamics studies are not in complete agreement, however, because the simulation suggested that both *cis* and *trans* ground-state structures bind with the peptide twisted ~20° from planarity, whereas we find that the X-Pro peptides of the protein substrate are not strongly twisted in the CypA crystal complexes with CA^N (Table 2), nor even in the peptide used as a starting point in the simulations (see Methods). The average ω -angle for the eight CA^N *trans*-X-Pro90 peptides is just 0.2° away from a planar 180° angle (range 173°–190°), and this value increased to only 1.0° after refinement to convergence (sparse matrix, iterative conjugate gradient) in the absence of peptide planarity restraints. The eight *cis* conformations (excluding the 20% occupied structures) all show some twist (average 9.5°; range 2°–18°) toward the transition state, with an average value of 10.0° after refinement in the absence of peptide planarity restraints. Thus, the *trans*-X-Pro90 peptides are planar in CypA-CA^N

Table 2. ω angles of CA^N Gly/Ala89-Pro90 peptide bond^a

<i>Trans</i>		HAG-A	HAG-B	AAG-A	AAG-B	AAA-A	AAA-A'	AMA-A	AMA-A'	Average
		190.1	178.4	188.9	185.1	175.8	175.7	174.5	172.9	180.2
		197.1	174.9	189.5	185.6	176.4	177.3	174.8	172.6	181.0
<i>Cis</i>		O-loop-A	O-loop-B	AMG-A	AMG-B	AAA-B	AAA-B'	AMA-B	AMA-B'	Average
		2.2	9.7	18.1	9.1	10.9	11.3	6.6	7.9	9.5
		2.7	11.1	18.3	9.3	11.4	12.3	6.7	8.0	10.0

^aValues shown are ω angles in degrees as determined by PROCHECK⁴². Upper values are in the presence of standard stereochemical restraints. Lower values are after 30 cycles of refinement in the absence of planar restraints. In agreement with expected values⁴³, the r.m.s. deviation from planarity of all peptides in the model ranges from 5.0° to 5.9° for the different refined structures.

crystal structures, and the *cis*-peptides are twisted by only $\sim 10^\circ$ toward the transition state. These results support the idea that CypA catalyzes proline isomerization by stabilizing the transition state of protein substrates rather than selectively binding 'near attack' conformations³¹.

CONCLUSIONS

The structures reported here provide mechanistic insight to the isomerization of protein substrates by CypA. Tight binding of the proline side chain and main chain oxygen atom require that conformational changes resulting from isomerization occur at residues N-terminal to the isomeric bond. CypA Arg55 carries out a twin function by anchoring the proline oxygen and activating the proline amide of the isomeric peptide bond. Both *cis* and *trans* structures are accommodated in the same active site with minimal changes in the path of the polypeptide. The peptide bond is planar when bound in the *trans* conformation and

only slightly twisted in the *cis* conformation. Finally, steric clash of the side chain in the residue preceding the proline prevents optimal binding of *trans*-proline, and the main chain geometry further appears to destabilize optimal binding of Ala-Pro sequences in the ground state *cis* and *trans* conformations. This explains why CypA preferentially binds Gly-Pro sequences and predicts that suboptimal binding sequences will be better substrates for isomerization.

METHODS

Protein purification and crystallization. Site-directed mutagenesis and protein expression and purification were as described^{18,26}. Crystals were grown in sitting drops at 21 °C as described¹⁷, with slight modification. The higher resolution data available in the present study resulted from deletion of five disordered residues from the C terminus of the original CA^N construct, with the result that the proteins used here comprise CA residues 1–146. The protein solution contained 0.4 mM CypA–CA^N, 1 mM β ME and 10 mM Tris, pH 8.0.

Table 3. Crystallographic data and refinement^a

CA ^N protein	WT(HAG)	O-loop	AAG	AMG	AAA	AMA
Crystallographic data						
Space group	<i>P</i> 2 ₁	<i>P</i> 2 ₁	<i>P</i> 2 ₁	<i>P</i> 2 ₁	<i>P</i> 1	<i>P</i> 1
Unit cell dimensions						
<i>a</i> (Å)	38.5	38.9	38.3	38.5	38.4	38.5
<i>b</i> (Å)	113.2	109.3	111.0	110.9	111.2	111.1
<i>c</i> (Å)	67.0	67.7	67.7	67.9	67.8	67.9
α (°)	90	90	90	90	90.0	89.9
β (°)	100.5	99.7	101.0	101.4	101.4	101.6
γ (°)	90	90	90	90	89.7	89.9
Resolution (Å) ^a	38–2.00 (2.03–2.00)	67–1.90 (1.93–1.90)	26–1.72 (1.82–1.72)	20–1.73 (1.78–1.73)	20–1.90 (1.93–1.90)	20–1.70 (1.73–1.70)
Number of observations						
Total	89,358	219,914	188,143	135,448	173,271	261,634
Unique	31,588	37,052	41,047	56,085	84,527	117,639
Complete (%)	83 (60)	94 (82)	78 (31)	96 (93)	98 (96)	97 (95)
R_{sym}^b	6.3 (19.5)	5.4 (35.2)	5.6 (30.4)	5.7 (36.6)	4.0 (25.1)	5.3 (35.7)
$\langle I / \sigma(I) \rangle$	14.8 (2.8)	21.1 (3.0)	9.4 (2.3)	15.9 (2.5)	19.0 (2.9)	15.8 (2.3)
Beamline ^c	NSLS	SSRL 9-1	SSRL 1-5	ALS	SSRL 9-1	SSRL 9-1
Detector ^d	Q4 CCD	Mar345	Q4 CCD	Q4 CCD	Mar345	Mar345
Refinement						
R -factor ^e	0.201	0.170	0.178	0.172	0.169	0.172
R_{work}^f	0.195	0.163	0.173	0.167	0.162	0.165
R_{free}^g	0.259	0.232	0.226	0.220	0.234	0.230
R.m.s. deviations						
Bond lengths (Å)	0.018	0.018	0.018	0.018	0.018	0.019
Bond angles (°)	2.7	2.9	3.0	2.8	2.6	2.5
ϕ/ψ angles ^h						
Most favored (%)	86.7	88.1	88.2	88.8	88.1	88.2
Additional allowed (%)	13.1	11.5	11.4	10.6	11.4	11.4
Average B -factors (Å ²)						
A: CypA–CA ^N	33/32	28/28	21/22	28/26	24/26	21/24
B: CypA–CA ^N	26/41	24/30	18/30	21/34	25/43	19/33
A': CypA–CA ^N					43/34	31/28
B': CypA–CA ^N					21/35	16/29
Solvent molecules	32	38	27	38	38	36

^aValues in parentheses are for the high-resolution shell. ^b $R_{\text{sym}} = 100 \times \sum_{\text{hkl}} |I - \langle I \rangle| / \sum \langle I \rangle$. ^cNSLS, National Synchrotron Light Source, beamline X12C, Brookhaven, New York; SSRL, Stanford Synchrotron Radiation Laboratory beamlines 9-1 or 1-5; ALS, Advanced Light Source beamline 5.0.2. ^dQ4 CCD, ADSC Quantum4. Mar345 image plate, Mar Research. ^e R -factor = $\sum_{\text{hkl}} |F_o - F_c| / \sum |F_o|$. ^f R_{work} is the R -factor for 90% of data used during refinement. ^g R_{free} is the R -factor for 10% of the data not used in refinement. ^hFor non-Gly and non-Pro residues only.

The reservoir solution was 15–25% (w/v) PEG 8K, 1 M LiCl and 100 mM Bicine, pH 7.0–9.0. Protein and reservoir solutions were mixed at a 1:1 ratio and microseeded. Crystals typically grew in radial clusters as elongated plates or prisms, reaching a maximal size ($\sim 0.1 \times 0.5 \times 0.5 \text{ mm}^3$ or $0.1 \times 0.1 \times 0.5 \text{ mm}^3$) within 2 weeks. Single crystals were harvested by transferring briefly to a cryoprotectant solution, suspended in a rayon loop and cooled by plunging into liquid nitrogen. The cryoprotectant consisted of 11–26% (w/v) PEG 8K, 1 M LiCl, 100 mM Bicine, pH 7.0–8.0, and 10% (v/v) butanediol.

Data collection and processing. All diffraction data were collected at 100 K. Data for WT, AAG, AMG, AAA, AMA and O-loop structures were processed with DENZO and SCALEPACK³². Data for AAG were processed with MOSFLM and SCALA³³. Data for AAA and AMA did not merge well in space group $P2_1$, giving R_{sym} values of 12.3% and 10.6%, respectively, in the low-resolution bins (16.0% and 13.2% overall). Reprocessing of these data in space group $P1$ gave much better statistics. The corresponding reduction to ~ 2.1 -fold data redundancy does not account for the drastically lowered R_{sym} values in $P1$, because scaling in $P2_1$ of truncated data sets that have ~ 2 -fold redundancy results in R_{sym} values in the low-resolution bins that remain at 10.9% and 8.8%, respectively (15.0% and 11.1% overall). Processing of data for WT, AAG, AMG and O-loop structures in space group $P1$ did not result in a significant reduction in R_{sym} values. Crystallographic statistics are given in Table 3.

Crystallographic refinement. Refinements were initiated with rigid-body minimization of the WT CypA-CA^N(1–151) model (PDB accession code 1AK4)¹⁷, from which water molecules were deleted. CA^N residues 85–95 were given zero occupancy, and CA^N and CypA molecules were treated as independent units. Crystallographic computing made extensive use of the CCP4 suite³⁴. Initial refinement was done with X-PLOR³⁵. For all structures, inspection of $F_o - F_c$ difference maps revealed clearly interpretable electron density for the omitted loop, which was built into the electron density using O³⁶, and water added. For AMG, AAA and AMA, solvent structure was built using ARP³⁷ and REFMAC³⁸. Final refinement of all structures was completed using REFMAC5 (ref. 39) with a maximum-likelihood target function and a Babinet-type bulk solvent correction. Riding hydrogen atoms were placed during refinement, but not written to the output coordinate files. During final refinement rounds of the AMA structure, a minor CA^N Ala89-Pro90 *cis*-proline conformation was apparent for the two loops that had been initially modeled as *trans*. These were modeled as alternate conformations for CA^N residues 88–91 and refined using a version of REFMAC5, modified for this purpose. Residues 89 and 90 of the minor *cis* conformations overlap closely with the fully occupied *cis* structures of the AMG and O-loop structures although, presumably because of the low occupancy, the minor conformations were not very stable in refinement. The occupancies were adjusted manually so that the temperature factors of residues 88 and 91 refined to similar values for the two conformations. The two loops showing this disorder in AMA were estimated to be $\sim 80\%$ *trans* and 20% *cis*. Difference density for AAA-A was also suggestive of a similar minor conformation, although in that case the occupancy was too low for reliable model building.

We further investigated the crystal structure of a CypA-tetrapeptide complex that has been used as the starting point for a molecular dynamics simulation³¹. This structure was reported to show twists of 21° and 14° from *cis* planarity for the two copies in the asymmetric unit²³, although these values are expected to be overestimates because this structure, which was published some years ago, was refined at that time in the presence of weak restraints toward the *trans* conformation²³. We subsequently refined this structure with conventional weights using the same procedure as for the CypA-CA^N complexes, and found that both X-Pro peptides differ by only 13° from perfect *cis* planarity after one round (50 cycles) of refinement. Unfortunately, these data are not of sufficient resolution to permit a reasonable unrestrained refinement.

Coordinates. The refined atomic coordinates and processed structure factor amplitudes have been deposited in the Protein Data Bank with the following accession codes: WT, 1M9C; O-loop, 1M9D; AAG, 1M9E; AMG, 1M9F; AAA, 1M9Y; AMA, 1M9X.

ACKNOWLEDGMENTS

We thank D.R. Davis, S.L. Alam, T.E. Cheatham III and members of the Hill and Sundquist labs for comments on this manuscript; H. Ke for providing processed

structure factor amplitudes for the CypA-tetrapeptide structure; H.L. Schubert for help with refinement; and G.N. Murshudov for modifying REFMAC5 to allow the refinement of peptide structures containing prolines with partial *cis* and partial *trans* conformations. Operations of the Advanced Light Source, National Synchrotron Light Source and Stanford Synchrotron Radiation Laboratory are supported by the U.S. Department of Energy, Office of Basic Energy Sciences, and by the National Institutes of Health. This work was supported by NIH grants to W.I.S. and C.P.H.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 15 October 2002; accepted 28 March 2003

Published online 5 May 2003; corrected 12 May 2003 (details online);

doi:10.1038/nsb927

- Fischer, G., Whittmann-Liebold, B., Lang, K., Kiefhaber, T. & Schmid, F.X. Cyclophilin and peptidyl-prolyl *cis-trans* isomerase are probably identical proteins. *Nature* **337**, 476–478 (1989).
- Takahashi, N., Hayano, T. & Suzuki, M. Peptidyl-prolyl *cis-trans* isomerase is the cyclosporin A-binding protein cyclophilin. *Nature* **337**, 473–475 (1989).
- Schmid, F.X., Mayr, L.M., Mücke, M. & Schönbrunner, E.R. Prolyl isomerases: role in protein folding. *Adv. Protein Chem.* **44**, 25–66 (1993).
- Kern, G., Kern, D., Schmid, F.X. & Fischer, G. A kinetic analysis of the folding of human carbonic anhydrase II and its catalysis by cyclophilin. *J. Biol. Chem.* **270**, 740–745 (1995).
- Kern, D., Kern, G., Scherer, G., Fischer, G. & Drakenberg, T. Kinetic analysis of cyclophilin-catalyzed prolyl *cis/trans* isomerization by dynamic NMR spectroscopy. *Biochemistry* **34**, 13594–13602 (1995).
- Göthel, S.F. & Marahiel, M.A. Peptidyl-prolyl *cis-trans* isomerases, a superfamily of ubiquitous folding catalysts. *Cell. Mol. Life Sci.* **55**, 423–436 (1999).
- Luban, J., Bossolt, K.L., Franke, E.K., Kalpana, G.V. & Goff, S.P. Human immunodeficiency virus type 1 gag protein binds to cyclophilins A and B. *Cell* **73**, 1067–1078 (1993).
- Franke, E.K., Yuan, H.E.H. & Luban, J. Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* **372**, 359–362 (1994).
- Thali, M. *et al.* Functional association of cyclophilin A with HIV-1 virions. *Nature* **372**, 363–365 (1994).
- Braaten, D., Franke, E.K. & Luban, J. Cyclophilin A is required for the replication of group M human immunodeficiency virus type 1 (HIV-1) and simian immunodeficiency virus SIV_{CPZ}GAB but not group O HIV-1 or other primate immunodeficiency viruses. *J. Virol.* **70**, 4220–4227 (1996).
- Wieggers, K. & Kräusslich, H.-G. Differential dependence of the infectivity of HIV-1 group O isolates on the cellular protein cyclophilin A. *Virology* **294**, 289–295 (2002).
- Saphire, A.C., Bobardt, M.D. & Galloway, P.A. *Trans*-complementation rescue of cyclophilin A-deficient viruses reveals that the requirement for cyclophilin A in human immunodeficiency virus type 1 replication is independent of its isomerase activity. *J. Virol.* **76**, 2255–2262 (2002).
- Bosco, D.A., Eisenmesser, E.Z., Pochapsky, S., Sundquist, W.I. & Kern, D. Catalysis of *cis/trans* isomerization in native HIV-1 capsid by human cyclophilin A. *Proc. Natl. Acad. Sci. USA* **99**, 5247–5252 (2002).
- Brazin, K.N., Mallis, R.J., Fulton, D.B. & Andreotti, A.H. Regulation of the tyrosine kinase Itk by the peptidyl-prolyl isomerase cyclophilin A. *Proc. Natl. Acad. Sci. USA* **99**, 1899–1904 (2002).
- Arévalo-Rodríguez, M., Cardenas, M.E., Wu, X., Hanes, S.D. & Heitman, J. Cyclophilin A and Ess1 interact with and regulate silencing by the Sin3-Rpd3 histone deacetylase. *EMBO J.* **19**, 3739–3749 (2000).
- Schmid, F.X. Prolyl isomerases. *Adv. Prot. Chem.* **59**, 243–282 (2002).
- Gamble, T.R. *et al.* Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **87**, 1285–1294 (1996).
- Yoo, S. *et al.* Molecular recognition in the HIV-1 capsid/cyclophilin A complex. *J. Mol. Biol.* **269**, 780–795 (1997).
- Zhao, Y., Chen, Y., Schutkowski, M., Fischer, G. & Ke, H. Cyclophilin A complexed with a fragment of HIV-1 gag protein: insights into HIV-1 infectious activity. *Structure* **5**, 139–146 (1997).
- Vajdos, F., Yoo, S.-H., Houseweart, M., Sundquist, W.I. & Hill, C.P. Crystal structure of cyclophilin A complexed with a binding site peptide from the HIV-1 capsid protein. *Protein Sci.* **6**, 2297–2307 (1997).
- Kallen, J. *et al.* Structure of human cyclophilin and its binding site for cyclosporin A determined by X-ray crystallography and NMR spectroscopy. *Nature* **353**, 276–279 (1991).
- Ke, H., Mayrose, D. & Cao, W. Crystal structure of cyclophilin A complexed with substrate Ala-Pro suggests a solvent-assisted mechanism of *cis-trans* isomerization. *Proc. Natl. Acad. Sci. USA* **90**, 3324–3328 (1993).
- Zhao, Y. & Ke, H. Crystal structure implies that cyclophilin predominantly catalyzes the *trans* to *cis* isomerization. *Biochemistry* **35**, 7356–7361 (1996).
- Zhao, Y. & Ke, H. Mechanistic implications of crystal structures of the cyclophilin-dipeptide complexes. *Biochemistry* **35**, 7362–7368 (1996).
- Eisenmesser, E.Z., Bosco, D.A., Akke, M. & Kern, D. Enzyme dynamics during catalysis. *Science* **295**, 1520–1523 (2002).
- Gitti, R.K. *et al.* Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273**, 231–235 (1996).

27. Vitagliano, L., Berisio, R., Mastrangelo, A., Mazzarella, L. & Zagari, A. Preferred proline puckerings in *cis* and *trans* peptide groups: implications for collagen stability. *Protein Sci.* **10**, 2627–2632 (2001).
28. Zydowsky, L.D. *et al.* Active site mutants of human cyclophilin A separate peptidyl-prolyl isomerase activity from cyclosporin A binding and calcineurin inhibition. *Protein Sci.* **1**, 1092–1099 (1992).
29. Stein, R.L. Mechanism of enzymatic and nonenzymatic prolyl *cis-trans* isomerization. *Adv. Protein Chem.* **44**, 1–24 (1993).
30. Fischer, S., Michnick, S. & Karplus, M. A mechanism for rotamase catalysis by the FK506 binding protein (FKBP). *Biochemistry* **32**, 13830–13837 (1993).
31. Hur, S. & Bruice, T.C. The mechanism of *cis-trans* isomerization of prolyl peptides by cyclophilin. *J. Am. Chem. Soc.* **124**, 7303–7313 (2002).
32. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
33. Leslie, A.G.W. *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography* Vol. 26 (Daresbury Laboratory, Warrington, 1992).
34. Collaborative Computational Project, Number 4. CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
35. Brünger, A.T. *X-PLOR Version 3.843: A System for X-ray Crystallography and NMR* (Yale University, New Haven, Connecticut, 1996).
36. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
37. Lamzin, V.S. & Wilson, K.S. Automated building of solvent structure combined with standard restrained refinement. *Methods Enzymol.* **277**, 269–305 (1997).
38. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
39. Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. & Dodson, E.J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D* **55**, 247–255 (1999).
40. Kraulis, P.J. Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
41. Merritt, E.A. & Bacon, D.J. Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).
42. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
43. Morris, A.L., MacArthur, M.W., Hutchinson, E.G. & Thornton, J.M. Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345–364 (1992).

Structural insights into the catalytic mechanism of cyclophilin A

Bruce R Howard, Felix F Vajdos, Su Li, Wesley I Sundquist & Christopher P Hill

Nature Structural Biology doi: 1038.10/nsb927

In the version of this article initially published online, the sequence for the model tetrapeptide substrate contains a mistake. This incorrect sequence is listed in two places, on page 4 line 21 and line 25 of the right hand column. The correct sequence should be Suc-Ala-Phe-Pro-Phe-NA. This mistake has been corrected for the HTML and print versions of the article.