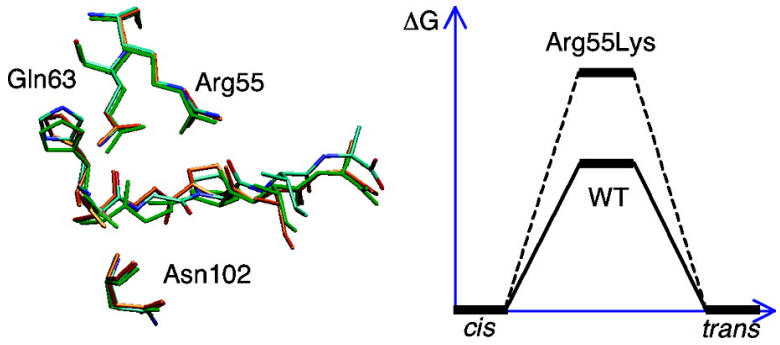
Article

What Is So Special about Arg 55 in the Catalysis of Cyclophilin A? Insights from Hybrid QM/MM Simulations

Guohui Li, and Qiang Cui

J. Am. Chem. Soc., **2003**, 125 (49), 15028-15038• DOI: 10.1021/ja0367851 • Publication Date (Web): 12 November 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 11/12/2003

What Is So Special about Arg 55 in the Catalysis of Cyclophilin A? Insights from Hybrid QM/MM Simulations

Guohui Li and Qiang Cui*

Contribution from the Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Avenue, Madison, Wisconsin 53706

Received June 19, 2003; E-mail: cui@chem.wisc.edu

Abstract: Potential of mean force (PMF) simulations with a hybrid QM/MM potential function were used to analyze the catalytic mechanism of human cyclophilin A (CypA). PMF calculations were performed for proline isomerization of peptides in solution, the wild-type CypA, and several CypA mutants. With an approximate density functional theory, the self-consistent-charge density functional tight binding (SCC-DFTB) as the QM level, and CHARMM 22 force field as MM, satisfactory energetics compared to available experiments were obtained. Calculations for the Arg55Ala and zero-charge-Arg55 mutants clearly indicated that Arg 55 significantly stabilizes the isomerization transition state through electrostatic interactions. However, the decrease in the average distance (thus the increase in interaction) between Arg 55 and the substrate amide N in going from the stable states to the transition state is mainly due to the pyramidalization of the amide N rather than motions associated with Arg 55. Although the nanosecond simulations cannot exclude the existence of sub-millisecond collective motions proposed on the basis of recent elegant NMR relaxation and line-shape analyses, the energetics obtained for the various enzyme systems here indicate that the contribution from motions of active site residues to catalysis is expected to be small. Instead, the present simulations support that the structural stability rather than mobility of the preorganized active site is more important. Through hydrogen-bonding interactions among the substrate, Arg 55, Gln 63, and Asn 102, the active site of the wild-type enzyme is structurally very stable and puts Arg 55 in a favorable position to perform its catalytic role in the transition state. This is further illustrated with the somewhat unexpected prediction that Arg55Lys is largely catalytically inactive, because Lys does not have the unique bifurcating construct of the guanidino group in Arg and thus the active site of Arg55Lys cannot accommodate Lys in a position capable of providing electrostatic stabilization of the isomerization transition state. Among all the enzyme systems studied, the wild-type CypA is the only one that selects the *syn*/*exo* transition state, while the *syn*/*endo* conformation is also present in the mutants, which is another reason for their higher barriers. Finally, the present analysis indicated that the population of near-attack-conformations (NAC) is not relevant to catalysis in CypA.

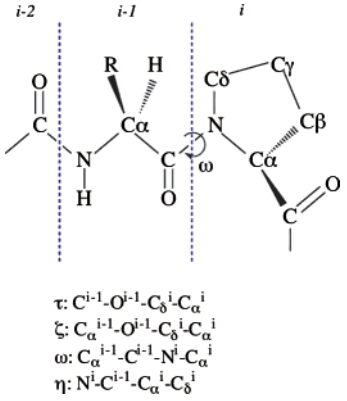
I. Introduction

Enzymes are efficient catalysts constructed by nature using the powerful mechanism of evolution.[1,2] It is therefore often thought that many enzymes are "perfect" [3] at least in terms of their active site construct, which determines the corresponding function (although recent studies also noted that many enzymes have "catalytic promiscuity", which can be taken advantage of to implement new functions[4,5]). The precise mechanism(s) through which the active site achieves extraordinary catalytic efficiency and specificity is likely to be system dependent, and recent experimental and theoretical analyses often propose that a number of factors contribute synergistically.[6−12] Among the many effects that have been suggested to be potentially important to catalysis,[1,2,13−17] well-positioned dipolar or charged residues at the active site have often been recognized as one of the dominant factors.[18] Another potential contribution that has attracted much attention lately involves the dynamical feature of enzymes.[19−31] Since the first molecular dynamics of proteins

(1) Fersht, A. *Enzyme structure, mechanism and protein foldings*, 3rd ed.; W. H. Freeman: New York, 1998.
(2) Jencks, W. P. *Catalysis in Chemistry and Biology*; Dover: New York, 1987.
(3) Knowles, J. R. *Nature* **1991**, *350*, 121−123.
(4) Begley, T. P.; Tsai, M. *Curr. Opin. Chem. Biol.* **2003**, *7*, 228−229.
(5) Copley, S. D. *Curr. Opin. Chem. Biol.* **2003**, *7*, 265−272.
(6) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035.
(7) Cui, Q.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 3093−3124.
(8) Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2002**, *106*, 1768.
(9) Cui, Q.; Karplus, M. *Adv. Protein Chem.* **2003**, *60*, 313−370.
(10) Gao, J. *Curr. Opin. Struct. Biol.* **2003**, *13*, 184−192.
(11) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.
(12) Cannon, W. R.; Benkovic, S. J. *J. Biol. Chem.* **1998**, *273*, 26257.
(13) Wu, N.; Mo, Y. R.; Gao, J. L.; Pai, E. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2017−2022.
(14) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem.* **1998**, *273*, 25529.
(15) Florian, J.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5950−5955.
(16) Miller, B. G.; Wolfenden, R. *Annu. Rev. Biochem.* **2002**, *71*, 847−885.
(17) Wolfenden, R.; Snider, M. J. *Acc. Chem. Res.* **2001**, *34*, 938−945.
(18) Warshel, A. *Computer modeling of chemical reactions in enzymes and solutions*; Wiley, John & Sons: New York, 1997.
(19) Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 444−448.
(20) Bruice, T. C.; Benkovic, S. J. *Biochemistry* **2000**, *39*, 6267−6274.
(21) Hammes-Schiffer, S. *Biochemistry* **2002**, *41*, 13335−13343.
(22) Villa, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 7887−7907.
(23) Cui, Q.; Elstner, M.; Karplus, M. *J. Phys. Chem. B* **2002**, *106*, 2721−2740.
(24) Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2002**, *106*, 7927−7947.
(25) Neria, E.; Karplus, M. *Chem. Phys. Lett.* **1997**, *267*, 23.

10.1021/ja0367851 CCC: $25.00

Scheme 1



25 years ago,[32] it has been well recognized that proteins are dynamical systems with motions spanning several orders of magnitudes ranging from femtoseconds up to milliseconds. Although it is relatively easy to appreciate motions that are required by substrate binding and product release,[33] it remains a significant challenge to reveal the microscopic nature of specific dynamical events that are directly connected to the chemical steps in the catalytic cycle,[34−36] i.e., whether there are dynamical contributions to the catalytic power of enzymes compared to reactions in solution.

Nuclear magnetic resonance (NMR) is a powerful technique for analyzing dynamical properties of macromolecules in solution spanning a wide range of time-scales with atomic details.[37,38] Several interesting relaxation measurements have been made recently on several enzyme systems including binase,[39] ribonuclease A,[40] and cyclophilin A (CypA).[41] A common theme is that specific residue(s) in the active site were found to undergo dynamical motions at a similar time-scale as the catalytic turnover. For CypA, which is an enzyme that catalyzes proline isomerization (Scheme 1) in peptides and proteins,[42−46] detailed relaxation and line-shape analyses[41] indicate that many residues have motions at the millisecond time-scale, which coincides with the chemical step in the catalytic cycle.[41,47] Although motions associated with many residues can be rationalized as the response to substrate binding and dissociation, the motion of Arg 55 was found to be different in terms of the substrate concentration dependence of the exchange rate (R_2); although other residues showed saturation behavior, R_2 of Arg 55 increased *monotonically* as a function of substrate concentration.[41] Since Arg 55 has been proposed to make essential contributions to catalysis through electrostatic interaction with the proline amide N,[48,49] it was natural for Eisenmesser et al.[41] to conclude that the dynamical character of Arg 55 could be an important factor that influences the catalytic efficiency of CypA. A molecular mechanism and quantitative contribution for such a motional coupling to catalysis, however, is not available.

Following the work of Eisenmesser et al.,[41] a classical molecular dynamics (MD) study has been carried out by Bruice and co-workers.[50] It was found that Arg 55 did not move much relative to the rest of the protein in 600−700 ps MD simulations in which a transition state (TS) model was docked into the active site instead of the reactant or product. The average distance between Arg 55 and the proline N did become shorter, however, in the TS simulation because the proline amide N has a pyramidal instead of planar structure as found in the reactant or product states. It was argued, therefore, that the motion of Arg 55 does not make an obvious contribution to catalysis. Since pure classical force field (CHARMM)[51] was used in that MD study,[50] detailed energetics associated with the catalysis have not been analyzed.

In the present work, we present analyses on the catalytic mechanism and energetics of CypA using a combination of molecular dynamics and hybrid QM/MM potentials.[11,30,52−54] The energetics of proline isomerization have been determined with potential of mean force[55] simulations for the wild-type CypA and several mutants constructed in silico. In so doing, we were able to evaluate semiquantitatively the contribution of various factors, such as the positive charge of Arg 55, to the catalytic efficiency of CypA, and therefore provide further insights into the catalytic mechanism. Moreover, we *predict* that the efficiency of CypA will be substantially quenched if Arg55 is replaced by a Lys residue, despite the same overall charges and similar sizes of the two amino acids. This is because the bifurcated feature of the guanidino group in Arg makes it capable of interacting effectively with both the isomerizing proline and other amino acids (Gln 63) in the active site; i.e., it is actually the preorganized position of Arg 55 rather than its mobility that makes it catalytically effective. This difference between the patterns of hydrogen-bonding interactions in which Arg and Lys participate could be a general principle that enzymes employ to fine-tune their active site for catalytic proficiency. In the following, we first describe the simulation setup in section II and then present the results of various

(26) Kohen, A.; Klinman, J. *Acc. Chem. Res.* **2000**, *31*, 397.
(27) Kiefer, P. M.; Hynes, J. T. *J. Phys. Chem. A* **2002**, *106*, 1834.
(28) Kiefer, P. M.; Hynes, J. T. *J. Phys. Chem. A* **2002**, *106*.
(29) Caratzoulas, S.; Schwartz, S. D. *J. Chem. Phys.* **2001**, *114*, 2910.
(30) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443.
(31) Warshel, A.; Parson, W. W. *Q. Rev. Biophys.* **2001**, *34*, 563−679.
(32) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267*, 585−590.
(33) Joseph, D.; Petsko, G. A.; Karplus, M. *Science* **1990**, *249*, 1425−1428.
(34) Lu, H. P.; Xun, L.; Xie, X. *Science* **1998**, *282*, 1877.
(35) Xie, X. *J. Chem. Phys.* **2002**, *117*, 11024−11032.
(36) Xie, X.; Lu, H. P. *J. Biol. Chem.* **1999**, *274*, 15967.
(37) Akke, M. *Curr. Opin. Struct. Biol.* **2002**, *12*, 642−647.
(38) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *339*, 204−238.
(39) Wang, L.; Pang, Y.; Holder, T.; Brender, J. R.; Kurochkin, A. V.; Zuiderweg, E. R. P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7684−7689.
(40) Cole, R.; Loria, J. P. *Biochemistry* **2002**, *41*, 6072−6081.
(41) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. *Science* **2002**, *295*, 1520−1523.
(42) Bosco, D. A.; Eisenmesser, E. Z.; Pochapsky, S.; Sundquist, W. I.; Kern, D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5247−5252.
(43) Zydowsky, L. D.; Etzkorn, F. A.; Chang, H. Y.; Ferguson, S. B.; Stolz, L. A.; Ho, S. I.; Walsh, C. T. *Protein Sci.* **1992**, *1*, 1092−1099.
(44) Kofron, J. L.; Kuzmic, P.; Kishore, V.; Colon-Bonilla, E.; Rich, D. H. *Biochemistry* **1991**, *30*, 6127−6134.
(45) Gamble, T. R.; Vajdos, F. F.; Yoo, S.; Worthylake, D. K.; Houseweart, M.; Sundquist, W. I.; Hill, C. P. *Cell* **1996**, *87*, 1285−1294.
(46) Schiene, C.; Fischer, G. *Curr. Opin. Struct. Biol.* **2000**, *10*, 40.
(47) Kern, D.; Kern, G.; Scherer, G.; Fischer, G.; Drakenberg, T. *Biochemistry* **1995**, *34*, 13594−13602.
(48) Fischer, S.; Michnick, S.; Karplus, M. *Biochemistry* **1993**, *32*, 13830−13837.
(49) Zhao, Y.; Ke, H. *Biochemistry* **1996**, *35*.
(50) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2002**, *124*, 7303−7313.
(51) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Kraplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(52) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298−305.
(53) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.
(54) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523.
(55) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, & Thermodynamics*; John Wiley & Sons: 1988; Vol. LXXI.

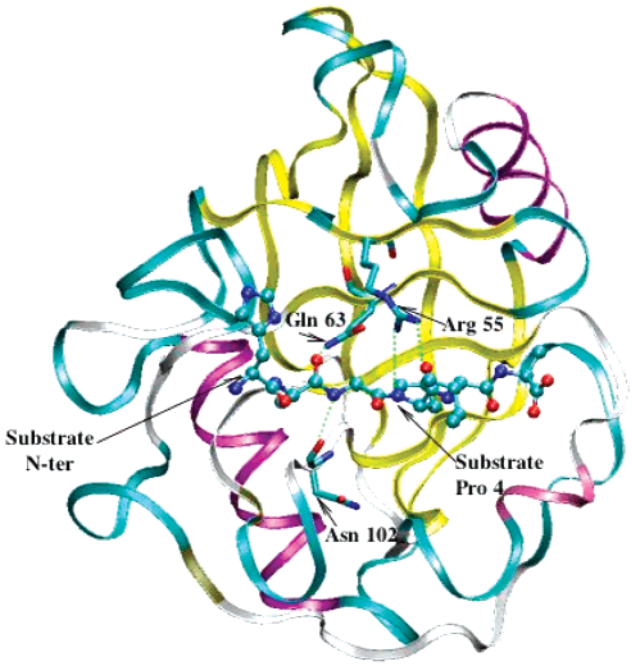


Figure 1. The structure of the cyclophilin A−peptide complex based on the X-ray structure (PDB code 1AWQ[64]). The protein is shown in ribbon format and is color-coded based on secondary structures (helix, purple; sheet, yellow; loops, blue and white). The substrate peptide (His-Ala-Gly-Pro-Ile-Ala) is shown in ball-and-stick and is color-coded by atom types. Three essential active site residues discussed extensively in the text, Arg 55, Gln 63, and Asn 102, are shown in line forms. The graphic was made using VMD (Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33).

calculations in section III. Discussion of results and a few conclusions are summarized in section IV.

II. Computational Methods

The QM/MM simulations[11,52−54] were set up using the standard stochastic boundary potential protocol.[56,57] The QM region included the proline that undergoes isomerization (Pro 4) and the preceding residue in the substrate peptide (Gly 2); see Figure 1. The rest of substrate and protein atoms were treated with the CHARMM 22 force field;[51] the water molecules were described with a modified TIP3P model.[58,59] Link atoms were introduced to saturate the valence of the QM boundary atoms; the link atoms were fully flexible and interact with the MM atoms, except the "link host" MM atom (e.g., the Cα atoms in this case), through electrostatic terms. No van der Waals interactions were included. This scheme has been shown to be a satisfactory way to treat the QM/MM interface, particularly when the charges of atoms in the neighborhood of the link atom are small,[60] which is true in the present case. No cutoff was used for QM/MM electrostatic interactions, while a force-shift based cutoff scheme was used for MM interactions (the inner and outer distances involved in the force shift are 8.0 and 12.0 Å, respectively); the same cutoff was used for all van der Waals interactions (i.e., between QM and MM atoms as well as among MM atoms). The QM method is the Self-Consistent-Charge Density Functional Tight Binding (SCC-DFTB) approach,[61] which was introduced into CHARMM[62] recently in a combined QM/MM framework.[63] As shown in Table 1, the SCC-DFTB approach reproduced B3LYP/6-311+G(d,p)//B3LYP/6-31+G(d,p) and MP2/6-311+G(d,p)//B3LYP/6-31+G(d,p) calculations for proline isomerization reaction in the gas phase, and therefore is expected to be appropriate for the current study. The van der Waals parameters used for SCC-DFTB atoms are the standard CHARMM parameters; a recent study indicated that relative free energies are less sensitive to the choice of van der Waals parameters compared to potential energies in QM/MM simulations (Riccardi, Cui, not published). An alternative QM/MM scheme in which the critical Arg 55 was also included in the QM region was tested with adiabatic mapping studies and was found to give very similar barrier heights for the isomerization; the partition with the smaller QM region was therefore used throughout the study to make balanced treatments for the wild-type and various mutant enzymes. We note that since SCC-DFTB is very fast for the relatively small QM region (∼30 atoms), the bottleneck of the calculations actually concerns the MM interactions (∼12 000 atoms).

The X-ray structure used in the simulation is the one with a bound piece of peptide (sequence: His-Ala-Gly-Pro-Ile-Ala) from HIV capsid protein (PDB code 1AWQ), where the peptide is in the *trans* conformation at Pro 4.[64] Simulations have also been carried out with a X-ray structure (PDB code 1RMH) with a different substrate peptide sequence (Suc-Ala-Ala-Pro-Phe-4-NA), where the peptide is in the *cis* conformation at Pro 4.[65] The results are in general very similar, and therefore only 1AWQ simulations will be presented because the corresponding X-ray structure has a higher resolution (1.58 Å) compared to that of 1RMH (2.4 Å). In the stochastic boundary setup, a 30 Å water sphere centered at the substrate proline amide and the corresponding deformable boundary potentials were used.[56] The final system

Table 1. Energetics of Proline Isomerization in Different Systems from Full QM and QM/MM Calculations[a]

system[b]	method[c]	TS[d]	*cis*
gas	SCC-DFTB	14.7; 15.7; 16.4; 18.4	0.9
	B3LYP/II//B3LYP/I	17.4; 20.7; 21.1; 25.7	0.9
	MP2/II//B3LYP/I	15.4; 19.0; 19.7; 23.4	0.7
solution[e]	SCC-DFTB/TIP3P	18.5; 18.9 (20.1)	1.2 (0.6)
CypA, WT[e]	SCC-DFTB/CHARMM	11.8 (12.6−12.7)	−0.1(0.3−0.1)
CypA, R55 no charge	SCC-DFTB/CHARMM	14.6	3.5
CypA, R55A	SCC-DFTB/CHARMM	18.1	3.9
CypA, R55K[f]	SCC-DFTB/CHARMM	21.3/20.8	3.7/0.4

[a] All energies are given in kcal/mol relative to the case where the proline is in the *trans* conformation. The gas-phase barriers do not include vibrational contributions, while those in solution and enzyme include classical vibrational contributions because potential of mean force calculations were performed. [b] In the gas-phase calculations, the model is a Gly-Pro dipeptide without the terminal charged groups. For the solution and enzyme calculations, the peptide His-Ala-Gly-Pro-Ile-Ala studied in the X-ray crystallography work (ref 64) was used. For other details, see section II of the main text. [c] In the gas-phase calculations, basis set I is 6-31+G(d,p) and basis set II is 6-311+G(d,p). In solution and enzyme studies, proline and the preceding residue in the substrate peptide were treated with SCC-DFTB, and all other atoms were treated with a molecular mechanics force field (i.e., TIP3P for water, and CHARMM 22 for the protein atoms). [d] The sequence of the numbers are *syn/exo*, *anti/endo*, *anti/exo*, *syn/endo*. For solution studies, only *syn/exo* and *anti/endo* were calculated; for the enzyme systems, only the *syn/exo* was calculated, and *anti/endo* was found to have a substantially higher barrier. [e] The numbers in parentheses are converted from kinetic data (refs 41 and 47) for SUC-Ala-Phe-Pro-Phe-*p*NA in CypA using Transition State Theory with a prefactor of k_BT/h ($T = 298$ K for ref 41, $T = 283$ K for ref 47), which is appropriate for a thermally activated unimolecular isomerization process. [f] Two independent sets of PMF simulations were carried out, whose results are given before and after the slants.

(56) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(57) Brooks, C. L., III; Brünger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843.
(58) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(59) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902−1921.
(60) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.
(61) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seigert, G. *Phys. Rev.* **1998**, *B58*, 7260−7268.
(62) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(63) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(64) Vajdos, F. F.; Yoo, S.; Houseweart, M.; Sundquist, W. I.; Hill, C. P. *Protein Sci.* **1997**, *6*, 2297.
(65) Zhao, Y.; Ke, H. *Biochemistry* **1996**, *35*, 7356−7361.

included 2568 protein atoms and 3078 water molecules. On the basis of the wild-type protein structures, several mutant calculations were also carried out; these included Arg55Ala, Arg55Lys, and a system where the partial charges on the side chain of Arg 55 were set to zero. The last artificial "mutant" and Arg55Ala were used to probe the effect of electrostatic interactions on catalysis due to the positive charge of Arg 55, and the Arg55Lys system was designed to explore if there are important features of Arg 55 other than the net positive charge.

In addition to the protein system, simulations were also carried out for a model dipeptide (Gly-Pro) in the gas phase and for the actual substrate peptide in solution; these calculations serve as references for the enzyme simulations. For the gas-phase model, SCC-DFTB,[61] B3LYP,[66−68] and MP2[69] calculations were carried out (see Table 1); for the solution system, SCC-DFTB/MM simulations were performed using a periodic boundary setup with a box size of 38 Å.

For the gas-phase system, the geometries for the *cis*, *trans* isomers of the dipeptide and the four types of isomerization transition states[70] (*syn-exo*, *syn-endo*, *anti-exo*, *anti-endo*) were determined at both SCC-DFTB and B3LYP/6-31+G(d,p)[71] levels. Single-point energy calculations were then performed at the B3LYP geometries using B3LYP/6-311+G(d,p) and MP2/6-311+G(d,p)[72] calculations. To study the isomerization reaction in solution and in CypA, potential of mean force (PMF) calculations were carried out. As discussed in previous studies,[48,70] the ω dihedral angle (Scheme 1) is not an appropriate reaction coordinate for adiabatic mapping or PMF calculations because the pyramidalization of the proline amide N is also an important part of the isomerization. Therefore, a different dihedral angle, τ ($C^{i-1}-O^{i-1}-C_\delta{}^i-C_\alpha{}^i$, Scheme 1), was used in the current work as the approximate reaction coordinate, which is very similar to the ζ angle ($C_\alpha{}^{i-1}-O^{i-1}-C_\delta{}^i-C_\alpha{}^i$, Scheme 1) suggested in previous studies;[48,70] adiabatic mapping calculations for the Gly-Pro dipeptide in the gas phase gave essentially identical energetics using either dihedral angle as the mapping coordinate (not shown).

In the solution PMF calculations, τ was sampled in the range of −180 to 180° using the adaptive umbrella sampling technique;[73] the umbrella potential was expanded using a 12-order trigonometric polynomial. The umbrella potentials were updated every 20 ps, and totally 1 ns of production MD simulations were carried out; the average difference between the computed potential of mean force and the umbrella potential is less than 0.1 kcal/mol, which indicates that satisfactory convergence (i.e., near-uniform sampling) has been reached. For the simulation of CypA, test calculations indicate that counterclockwise rotation is higher in energy than the clockwise rotation, and therefore τ was only sampled from 0 to 180° in 10 windows of umbrella sampling calculations[74] following 300 ps of equilibrations for the CypA−*trans*-peptide complex. Each window was equilibrated with 20 ps of hybrid Langevian Dynamics (LD) and Molecular Dynamics (MD) simulation as in the stochastic boundary approach,[57] which was followed by 50 ps of production run; e.g., the entire PMF calculation employed ∼1 ns of MD simulations for each enzyme−peptide system (wild-type, zero-charge-Arg55, Arg55Ala, and Arg55Lys). For Arg 55K, two sets of PMF calculations were carried out with different initial conditions to confirm the results. Distributions of τ were inverted to obtain PMF with the Weighted Histogram Analysis Method (WHAM).[75] An integration time-step of 1 fs was used, with all the bonds involving hydrogen atoms constrained using SHAKE.[76] The LD/MD boundary in the stochastic boundary simulations[56,57] was set to be 25 Å from the proline amide in the substrate and was updated every 25 steps of integration.

III. Results

In this section, we first present PMF for the isomerization reaction in the wild-type CypA and compare the results to those in the gas-phase and solution cases. We will then discuss mutations that eliminate the net charge on Arg 55. Finally, we explore if Arg 55 can be replaced by a Lys residue, which has the same net positive charge and a similar size of side chain. Results from these calculations will be discussed in the context of the recent NMR[41] and X-ray analysis[77] as well as previous biochemical studies[43,44,47,78] on proline isomerization catalyzed by CypA.

Proline Isomerization in the Gas-Phase, Solution, and Wild-Type Cyclophilin A. The proline isomerization reaction has been discussed extensively in the past.[70,79] It is well established that there are four possible routes for the *cis*−*trans* isomerization, depending on the puckering of the proline ring and the direction of rotation. Among the four transition states optimized for the model Gly-Pro dipeptide, *syn-exo* is the one with the lowest barrier height, while *syn*/*endo* has the highest barrier. At the MP2/6-311+G(d,p)//B3LYP/6-31+G(d,p) level, the barriers for the two extreme cases are 15.4 and 23.4 kcal/mol, respectively. The SCC-DFTB approach reproduced the trends in the barrier heights, although the differences among various cases are smaller (see Table 1). It reproduces very well the barrier for *syn*/*exo* (14.7 kcal/mol vs 15.4 kcal/mol at the MP2 level) and the energy difference between the *cis* and *trans* isomers (0.9 kcal/mol vs 0.7 kcal/mol at the MP2 level). Therefore, it is appropriate to use SCC-DFTB as the QM level to study the isomerization in solution and enzyme environment with QM/MM simulations.

In solution, the adaptive umbrella sampling calculations found that the *cis* conformer is about 1.2 kcal/mol higher in potential of mean force (PMF) than the *trans* conformer, in agreement with the general trend observed experimentally for peptide isomerizations.[80] Two maxima were found along the PMF as a function of τ (Scheme 1), which occurred at 65° and −67°, respectively. The average peptide structure from MD simulations where τ was restrained at these two values using large harmonic constraints (force constants of 2000 kcal/mol·Å^{-2}) found that the two transition states correspond to *syn*/*exo* and *anti*/*endo*, respectively. In the gas phase, it is also these two transitions states that have lower barriers (Table 1). The solution barriers (SCC-DFTB/CHARMM) are close to be 18.0 kcal/mol, which are about 3 kcal/mol higher than the corresponding gas-phase (SCC-DFTB) values. As discussed in previous literature,[79] such an increase of barrier height upon solvation can be understood by the fact that solvent molecules increase the double-bond character of the isomerizing C−N bond via hydrogen-bonding interactions with the corresponding carbonyl group (Scheme 1).

(66) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(67) Becke, A. D. *Phys. Rev. A* **1998**, *38*, 3098−3100.
(68) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(69) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to advanced electronic structure theory*; Dover: New York, 1989.
(70) Fischer, S.; Dunbrack, R. L.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931−11937.
(71) Hariharan, P. C.; Pople, J. A. *Theoret. Chim. Acta* **1973**, *28*, 213−222.
(72) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.
(73) Bartels, C.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 865−880.
(74) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(75) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(76) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327−341.
(77) Howard, B. R.; Vajdos, F. F.; Li, S.; Sundquist, W. I.; Hill, C. P. *Nat. Struct. Biol.* **2003**, *10*, 475−481.
(78) Stein, R. L. *Adv. Protein Chem.* **1993**, *44*, 1−24.
(79) Scheiner, S.; Kern, C. W. *J. Am. Chem. Soc.* **1977**, *99*, 7042−7050.
(80) Albers, M. W.; Walsh, C. T.; Schreiber, S. L. *J. Org. Chem.* **1990**, *55*, 4984.

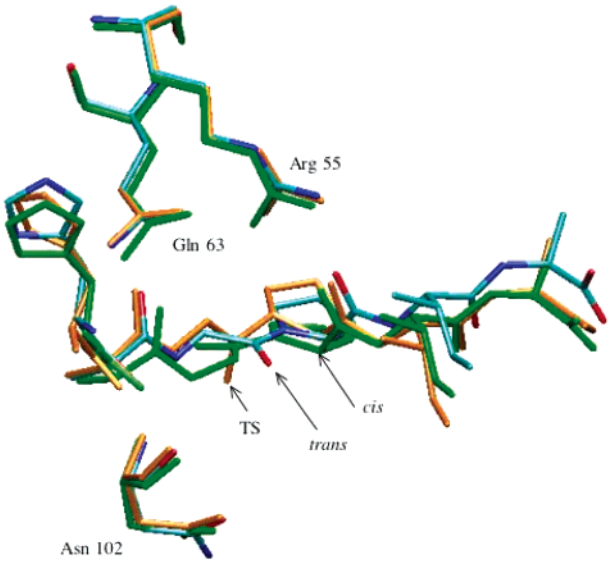


Figure 2. Superposition of average structures of the substrate peptide and three active site residues from MD simulations of the wild-type CypA−peptide system where the enzyme system is in the *trans* (color-coded based on atom types), transition state (in orange), and *cis* (in green) conformation, respectively. For the protocol used to generate these structures, see the footnote of Table 2.

For the *trans* isomer, for example, the average C−N bond length from solution MD simulations is about 1.312 Å, which is substantially shorter than the optimized value of 1.378 Å in the gas phase.

Having reproduced the expected trends in the gas phase and in solution, SCC-DFTB/CHARMM calculations were used to analyze isomerizations in the wild-type CypA. Explorative calculations indicated that the *syn*/*exo* is the only pathway that is low in activation barrier; the *anti*/*syn* pathway, for example, has a significantly higher barrier (see below). The *syn*/*exo* isomerization barrier is 11.8 kcal/mol above the *trans* peptide−CypA complex, and the isomerization is almost isoenergetic (Table 1). The value of the barrier is similar to experimental measurements (∼12.7 kcal/mol) for the isomerization of a slightly different peptide (Suc-Ala-Phe-Pro-Phe-*p*NA) in CypA;[41,47] it was known that the barrier is fairly insensitive to the peptide sequence, although the enthalpic and entropic components vary.[78] The similar energetics associated with the *cis*- and *trans*-peptide−CypA complexes are also consistent with experimental measurements; e.g., it was found that isomerizations starting from either isomer have very similar rate constants.[41]

As shown in Figure 2, the average structure of the CypA active site changes very little during the isomerization. Both the main chain and side chain of Arg 55, for example, superimpose very well when the substrate is in the *trans*, *cis* and transition state conformation, respectively. Other polar residues close to the isomerizing proline, such as Gln 63 and Asn 102, also undergo only small displacements in their average positions during the PMF calculations; the RMS fluctuations of the heavy atoms in these residues are below 0.50 Å. The fact that these residues are fairly rigid during the proline isomerization is not unexpected considering the strong interactions they are involved in. As shown in Figure 3 and Table 2, the -NH_2 groups in Arg 55 maintain stable interactions with the C═O group in the isomerizing proline (the average N−O distances are ∼2.8 Å) throughout the reaction; one of the -NH_2 groups is also involved in a short hydrogen bond with the side chain carbonyl in Gln 63 (the average N−O distance is ∼2.8 Å). The -NH_2 group in Gln63 is in turn hydrogen bonded with the nearby C═O in the substrate peptide (O^2) throughout the reaction, where the interaction is slightly weaker in the *cis* complex as indicated by the longer average N−O distance (Table 2, Figure 3).

During the isomerization, the C-terminal of the substrate peptide undergoes slightly larger changes compared to the N-terminal, which is also the conclusion reached by NMR studies of a similar peptide−CypA system.[47] The absolute magnitude of the displacement, however, is quite modest (Figure 2) and much smaller than that implied in the proposal raised in ref 41. The small displacements in the substrate during the isomerization observed in the current MD study are more consistent with the findings from a recent X-ray study[77] that we became aware of after our work was submitted for review. High-resolution X-ray structures were solved for a number of complexes between CypA and the N-terminal domain of HIV-1 CA protein variants that bind preferentially with the substrate proline in either the *cis* or *trans* conformation. Comparison between these structures suggested an isomerization pathway in which the C-terminal of the substrate remains stationary while the N-terminal undergoes small displacements.[77] The pyramidalization of the amide N in the isomerization transition state makes the average distance between the amide N and the guanidino group of Arg 55 shorter by 0.4−0.5 Å compared to the stable isomers (Figure 3, Table 2), which was speculated to be an important factor for the catalysis in CypA. Although it is conceivable that the stronger interaction between Arg 55 and the amide N at the transition state can make a substantial stabilizing contribution, it is interesting to point out that it is the substrate geometry that changes during the reaction rather than the position of Arg 55 itself. The similar behavior was also found in the previous classical MD simulations by Bruice and co-workers,[50] although their "TS" structure was modeled based on gas-phase calculations and then docked into the active site, as opposed to the more self-consistent PMF approach adopted here.

Along the *syn*/*exo* isomerization pathway, the C═O in the neighboring Gly undergoes clockwise rotation (looking from the N-terminal to the C-terminal of the substrate peptide). At the transition state, where τ is close to be 80°, this C═O group is in a favorable position to form a decent hydrogen bond with the main chain NH of Asn 102; the average distance is ∼3.0 Å (Figure 3, Table 2). The carbonyl group in the same residue is involved in a hydrogen bond with the Ala 2 main chain NH of the peptide, forming another anchor on the opposite side of Arg 55 and Gln 63 (Figure 3). Apparently, in addition to the intrinsically lower barrier associated with the *syn*/*exo* pathway (i.e., it is favored in both gas phase and in solution, see Table 1), the interaction with Asn 102 is another major driving force for selecting this isomerization route in CypA. This nature of the transition state is consistent with previous secondary kinetic isotope measurements,[78] which recorded a secondary KIE of 1.13 that is only consistent with a *syn*/*exo* transition state.

Effect of Electrostatic Stabilization from Arg 55. When the partial charges on Arg 55 were set to zero, the side chain is no longer anchored to the substrate or Gln 63 as reflected by

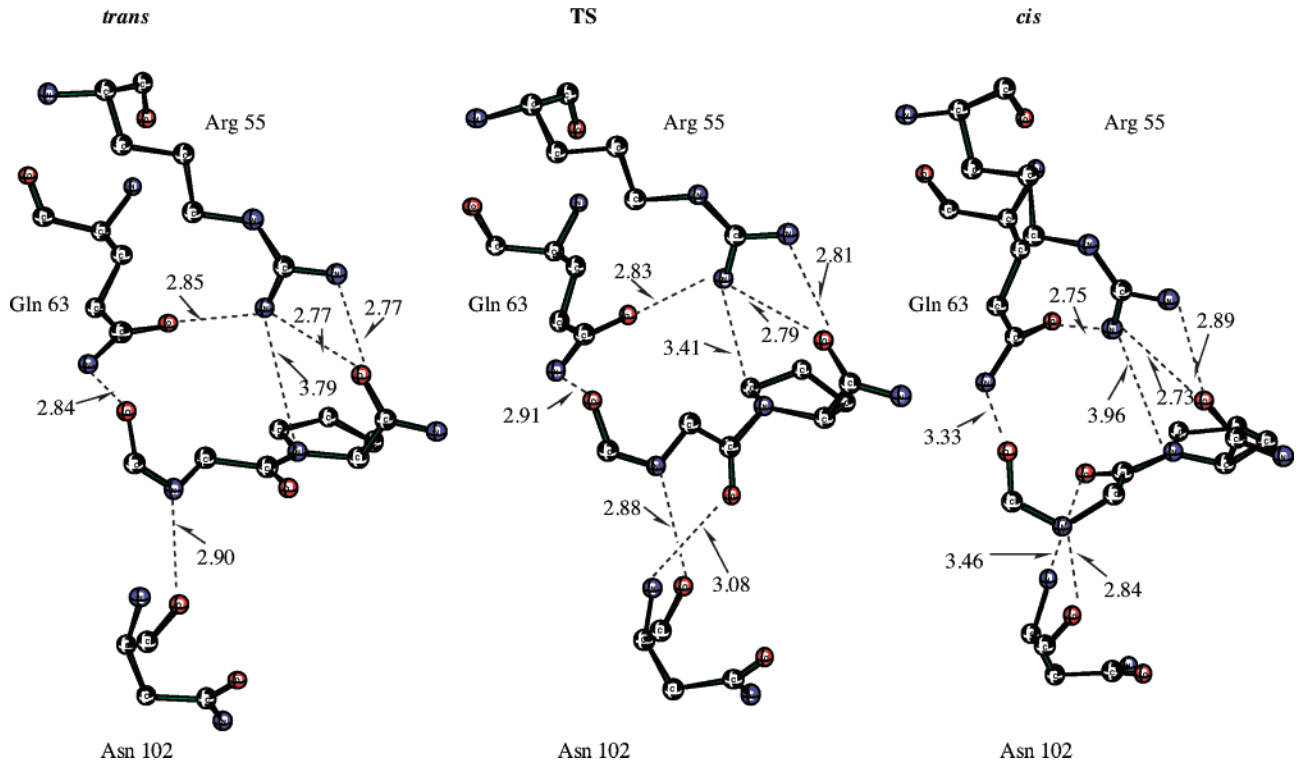


Figure 3. Superposition of average structures of the substrate peptide and three active site residues from MD simulations of the wild-type CypA−peptide system, with detailed representation of hydrogen-bonding interactions between the substrate peptide and active site residues in the *trans*, transition state, and *cis* conformations. Distances are given in Å. Note that the distances given here are between atoms in the average structures from MD simulations, while Table 2 shows the average distances from snapshots in the MD trajectories.

Table 2. Important Distances and Dihedrals and Their Fluctuations from MD Simulations of the trans-, cis- and Transition State (TS) Peptide-CypA Complexes[a]

		distances (Å)							dihedrals (deg)		
systems		#55–N^4	#55–O^4	#55–O^2	#55–Q63	Q63–O^2	N102–O^3	N102–N^3	τ	ω	η^c
WT	*trans*	3.82/0.21	2.72/0.09		2.88/0.16	2.85/0.15	5.75/0.46	2.95/0.21	179/15	179/9	0/9
	TS	**3.44/0.22**	2.74/0.10		2.86/0.16	2.93/0.17	**3.10/0.22**	2.90/0.16	*84*		**−28/8**
	cis	3.98/0.27	2.72/0.09		2.79/0.13	3.36/0.43	3.50/0.42	2.88/0.15	10/7	24/14	−13/9
R55	*trans*	4.05/0.35	3.50/0.35		3.65/0.39	3.07/0.27	5.58/0.37	2.97/0.22	173/13	178/9	−2/8
no charge	TS	**4.42/0.47**	4.08/0.66		4.13/0.59	2.96/0.33	**2.96/0.20**	2.98/0.19	*76*		**12/21**
	cis	6.36/0.64	3.80/0.58		5.65/0.63	6.40/0.64	2.99/0.19	3.22/0.25	7/5	13/13	−5/9
R55A	*trans*					3.02/0.34	5.73/0.45	2.95/0.19	179/14	183/10	−3/8
	TS					2.86/0.17	**3.54/0.65**	2.97/0.20	*74*		**−5/25**
	cis					3.06/0.27	3.12/0.25	2.84/0.13	11/6	23/12	−12/8
R55K	*trans*	4.18/0.17	2.70/0.09	2.71/0.09	4.01/0.34	2.85/0.14	6.48/0.36	2.87/0.16	182/13	184/9	−4/8
	TS	**4.82/0.24**	2.78/0.13	2.77/0.12	**3.97/0.61**	2.85/0.13	**3.70/0.68**	2.95/0.19	*74*		**−7/26**
	cis	4.60/0.24	2.68/0.09	2.98/0.26	2.86/0.25	3.06/0.40	3.16/0.28	2.88/0.17	9/6	20/12	−11/8

[a] To accumulate statistics for the TS complex, τ was restrained at the angle where the potential mean force reaches the maxima (shown at italics) using a large harmonic force (force constant of 2000 kcal/mol·rad^2). For the *cis* and *trans* complexes, unconstrained MD simulations were used. For all three states, the statistics were collected from 100 ps production runs, respectively. Note that these calculations were preceded by 500 ps−1 ns of MD simulations for PMF calculations. For each quantity, the values before and after the slashes are the average and standard deviations, respectively. The bold numbers are these of particular interests and discussed extensively in the text. [b] The distances are labeled by the residue number in CypA and atomic names$^{\text{residue number}}$ for the substrate atoms. For the definition of dihedral angles, see Scheme 1. [c] Also see Figure 6 for the illustration of transitions between positive and negative values during MD simulations for mutant transition states.

the longer average distances and larger fluctuations (see Table 2). Other aspects of the active site, however, remained similar to those of the wild-type enzyme during the nanosecond simulations. It is possible that the substrate has a significantly altered binding affinity and structure if the zero-charge "mutant" can be engineered in reality (e.g., by replacing Arg 55 with a Leu), but our goal here is to use this somewhat artificial construct to explore the contribution of electrostatic interactions from Arg 55 to catalysis. PMF calculations show that the barrier measured relative to the *trans*-peptide−CypA complex becomes substantially higher by 2.8 kcal/mol, which is consistent with the conventional hypothesis that electrostatic stabilization of the amide lone pair is important for the catalytic proficiency of CypA. The *cis*-peptide−CypA complex becomes slightly less stable than the *trans* complex by nearly 4 kcal/mol upon removing the charge on Arg 55. Similarly, the Arg55Ala mutant exhibits a destabilized *cis* complex and a higher isomerization barrier (Table 1). The barrier measured from the *trans* complex is 18.1 kcal/mol, which is more than 6.3 kcal/mol higher than that for the wild-type enzyme. This result is qualitatively

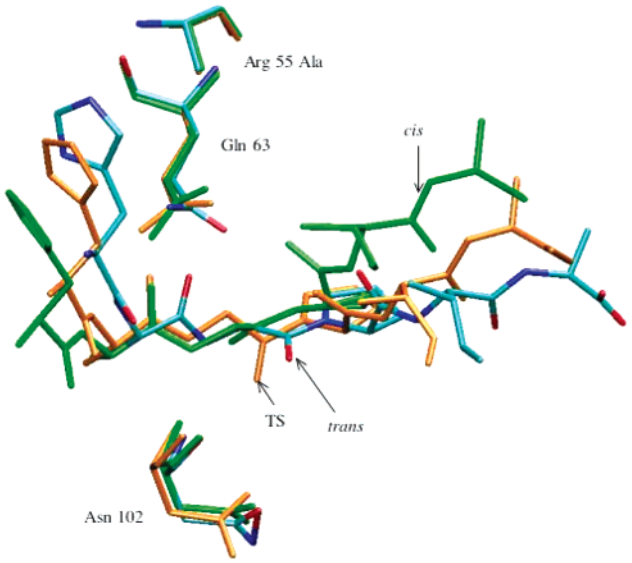


Figure 4. Similar to Figure 2, but for the Arg55Ala mutant.

consistent with the previous mutation study, which showed that the Arg55Ala mutant has less than 1% catalytic activity (k_{cat}/K_m) compared to the wild type.[43]

The energetics of the Arg55Ala mutant are fairly close to the solution values (Table 1), which clearly underline the importance of Arg 55 to the catalytic power of CypA. The fact that the barrier for the mutant where Arg 55 has zero charge but standard van der Waals parameters is lower than that for Arg55Ala suggests that an additional role of Arg 55 is to maintain the structural stability of the active site. Indeed, the substrate peptide undergoes substantially larger average displacements in both the N- and C-termini during the isomerization in Arg55Ala compared to the wild type (Figure 4). As a result, the interaction between the substrate C═O[3] and the main chain NH in Asn 102 is weaker in the Arg55Lys transition state, as indicated by the longer average N−O distance of ∼3.5 Å (Figure 5, Table 2) compared to ∼3.1 Å in the wild type (Figure 3, Table 2). In the zero-charge-Arg 55 system, the Asn 102−O[3] interaction remains stable in the transition state, with an average distance of ∼3.0 Å; this is consistent with the fact that the corresponding barrier is lower than that for the Arg55Ala mutant. It is also interesting to note that the pyramidalization of the amide N in the *average* transition state structure is less significant in both the Arg55Ala mutant and the zero-charge Arg55 system compared to the wild-type enzyme; the η (N−C−Cα−Cδ) dihedral angles are −5°, +12°, and −28°, respectively (Table 2). The instantaneous values of η along the MD trajectories reveal that the small average values for the two mutants are due to the fact that η fluctuates between positive and negative values (Figures 6b, 6c); η remains close to −30° in the wild-type simulations, although positive values are also sampled occasionally (Figure 6a). In other words, although the *syn*/*exo* transition state is essentially the only one that is populated in the wild-type enzyme, both *syn*/*exo* and *syn*/*endo* transition states can be populated once the positive charge associated with Arg 55 is absent. Since the *syn*/*endo* isomerization pathway has a higher intrinsic barrier based on gas-phase calculations (Table 1), the lack of selection in TS conformation is another reason that the two mutants have higher activation barriers than those of the wild type.

Can Arg 55 Be Replaced by a Lysine Residue? Although the above analysis confirmed the previous hypothesis that Arg 55 makes an essential contribution to catalysis through electrostatic interactions, the difference between Arg55Ala and zero-charge Arg55 calculations indicated that the stability of the active site is also important. Therefore, we decided to explore whether Arg 55 can be replaced by a Lys residue. As shown in Table 1, the QM/MM PMF calculations predicted that the barrier for the Arg55Lys mutant would be substantially higher than not only the wild-type CypA but also all other mutants studied here; the transition state is ∼21 kcal/mol above the *trans* complex, while the *cis* complex is higher in free energy than the *trans* to different degrees in the two sets of independent simulations.

As shown in Figure 7 and Figure 8, a key difference between Lys 55 and Arg 55 is the interaction with Gln 63 and the substrate peptide. In the wild-type enzyme, Arg 55 interacts with both Gln 63 and the proline carbonyl in the substrate in a rather persistent fashion *throughout* the reaction; it is the unique bifurcating construct of the guanidino group that makes this possible. In the Arg55Lys mutant, by contrast, the hydrogen-bonding pattern changes substantially during the isomerization, which was consistently observed in two independent sets of PMF simulations. In the *trans* and transition state structures, Lys 55 interacts mainly with the carbonyl of Ala2 and Pro4 in the substrate peptide; the interaction with Gln 63 is much weaker by comparison as indicated by the longer average distances (∼4.0 Å) and large fluctuations (Figure 7, Table 2). In the *cis* complex, the interaction between Lys 55 and Gln 63 is stronger as reflected by the shorter average distance of ∼3.0 Å, although there are still substantial fluctuations of 0.40 Å (Table 2). Due to the particular pattern of interactions, the proline amide N is far from Lys 55, and the rotating carbonyl (O[3]) is not stabilized by the main chain NH in Asn 102; the relevant average distances in the transition state are ∼4.8 Å and 3.7 Å, respectively, which can be compared to the values in the wild-type enzyme of ∼3.4 Å and 3.1 Å, respectively (Table 2). Moreover, as in the Arg55Ala and zero-charge Arg55 simulations, the pyramidal angle of the amide N in the Arg55Lys transition state fluctuates between positive and negative values (Figure 6d), which indicates the population of both *syn*/*exo* and *syn*/*endo* conformers. All these factors contribute to the substantially higher barrier found for Arg55Lys relative to the wild-type CypA, and the even weaker interaction with Asn 102 makes the barrier higher than that for Arg55Ala. Since the Arg55Lys calculations were performed with the X-ray structure for the wild-type enzyme−peptide complex, the PMF results should be taken as less quantitative compared to the wild-type data; this is reflected by the variation in the relative stabilities of the *cis* and *trans* conformers in different Arg55Lys simulations (Table 2). Nevertheless, we expect that the trend in the isomerization barrier is meaningful.[91]

IV. Concluding Discussions

In the present work, the mechanism and energetics of proline isomerization catalyzed by cyclophilin A (CypA) have been analyzed with potential of mean force (PMF) calculations employing a hybrid QM/MM potential function, where an approximate density functional theory, SCC-DFTB, was used in combination with the CHARMM force field. Such calculations in solution and wild-type CypA as well as in several

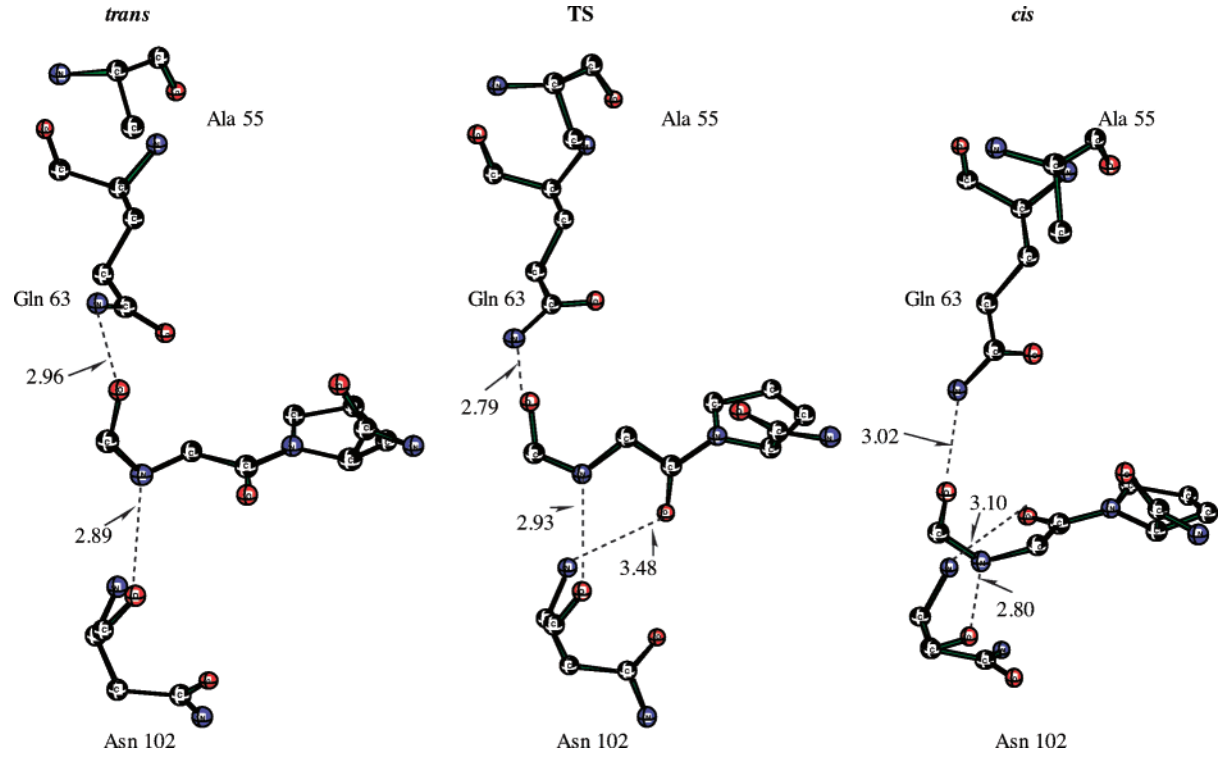


Figure 5. Similar to Figure 3, but for the Arg55Ala mutant.

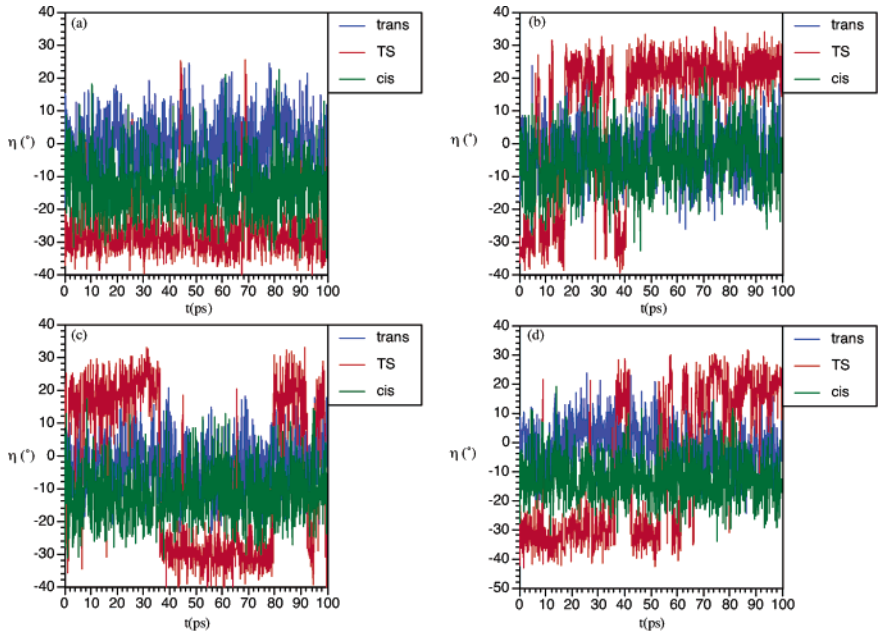


Figure 6. The instantaneous values of the dihederal angle (η, see Scheme 1) that characterize the pyramidalization of the proline amide N in MD simulations: (a) wild-type, (b) zero-charge-Arg55, (c) Arg55Ala, and (d) Arg55Lys systems.

mutants, some of which are difficult to construct experimentally (e.g., zero-charge Arg 55), made it possible to analyze possible contributions of various effects to catalysis.

Compared to proline isomerization in solution, the barrier in the wild-type CypA was found to be nearly 6 kcal/mol lower by QM/MM-PMF calculations, which corresponds to a catalytic power of 10^{4-5} at 300 K; both the absolute and relative barrier values are fairly close to estimates based on experimental kinetic measurements for a series of proline containing peptides (Table 1),[41,43,47,78] which confirmed the reliability of the simulation protocols used here.

In agreement with the previous hypothesis,[43,48,49,65] Arg 55 was found to play a dominant role in CypA catalysis. Calculations for the Arg55Ala mutant and an artificial system with the partial charges of Arg 55 set to zero confirmed that the contribution of Arg 55 is largely electrostatic in nature. One of the new insights from the present analysis is on the mechanism through which Arg 55 implements its electrostatic contribution. It was shown that a key feature of Arg 55 is its bifurcating construct of the guanidino group, which allows Arg 55 to form stable hydrogen-bonding interactions with both Gln 63 and the substrate peptide throughout the reaction (Figure 2); stabilization

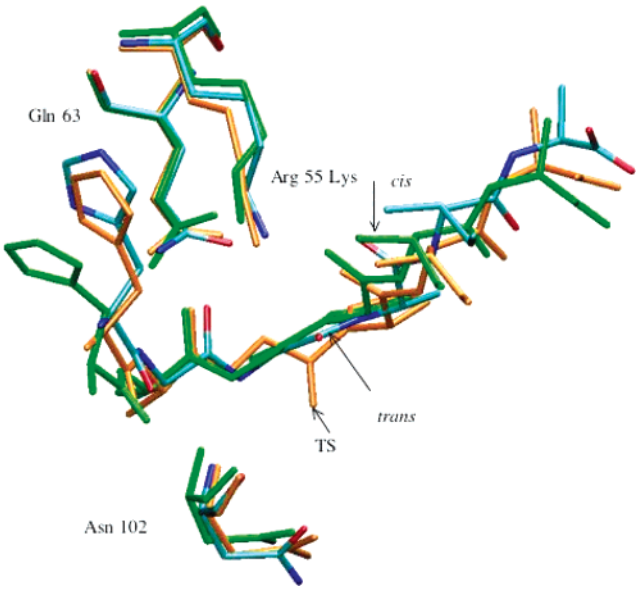


Figure 7. Similar to Figure 2, but for the Arg55Lys mutant.

of the transition state arises from the fact that the distance between the amide N and Arg 55 becomes substantially shorter due not to the displacement of Arg 55 but to the pyramidalization of the substrate amide N. A demonstration for the importance of active site stability is the somewhat unexpected prediction that the Arg55Lys mutation significantly quenches the catalytic power of CypA; this is due to the fact that the side chain of Lys does not form the enduring hydrogen-bond-network with Gln63 and the proline carbonyls and therefore does not have the appropriate spatial arrangement that would achieve transition state stabilization. The preorganized position of Arg 55 was also found to be important in selecting the low-energy *syn*/*exo* transition state for isomerization. All mutant calculations (Arg55Ala, Arg55Lys, and zero-charge Arg 55) populated both *syn*/*exo* and syn/*endo* conformers in the transition state; the latter corresponds to a substantially higher intrinsic barrier according to gas-phase calculations (Table 1). In addition to Arg 55, another major player is Asn 102, whose main chain NH forms favorable interactions with the rotating substrate carbonyl in the transition state. It is this interaction that further locks *syn*/*exo* to be the favorable route for isomerization in the wild-type enzyme, in addition to the intrinsically low barrier associated with this path as found in the gas-phase and solution calculations. We note that the interpretation that Arg 55 stabilizes the active site through hydrogen-bonding interactions with the protein (Gln 63) and substrate is not inconsistent with the notion that functionally important residues are often found to destabilize the protein (i.e., relative to the unfolded state). The latter is highlighted by several recent mutation studies which demonstrated that mutating key amino acids in the active site led to energetically more stable albeit inactive enzymes;[81−83] computational studies with dielectric continuum models applied to a fairly large set of proteins also supported such an argument.[84] In our study, however, we are not addressing the issue of whether preorganizing Arg 55 is associated with a significant energetic penalty that has to be paid during protein folding, although this might be qualitatively true. Rather, we emphasize that the electrostatic interaction between Arg 55 with its environment helps to establish an enduring hydrogen-bonding network that contributes favorably to the proline isomerization. It would be interesting in the future to explore whether mutating Arg 55 to Lys induces a significant change in the stability of CypA.

It is interesting to compare CypA with another peptidylproline isomerase that has been studied previously in details, the FK506 binding protein (FKBP).[80] Fischer and co-workers[48] used a series of molecular mechanics calculations (augmented by ab initio results to correct energetics for the isomerization) to study the minimum energy path associated with the isomerization. Analysis of the results indicated that several factors contributed to the catalytic power of FKBP, which corresponds to ~6 kcal/mol reduction in the activation free energy.[44,80] First, the proline carbonyl in FKBP is not hydrogen bonded with any strong hydrogen-bonding donor groups throughout the reaction; thus the double bond character of the isomerizing C−N bond is lower than that in solution.[79] Such a desolvation effect[48,79] is believed to contribute substantially to the reduction of the isomerization barrier in FKBP compared to that in solution. Another unique feature of the FKBP system is that the NH group adjacent to the isomerizing proline becomes hydrogen bonded to the lone pair of the pyramidalized amide nitrogen at the transition state; in other words, the substrate has an autocatalytic feature[48] due to the type VIa turn conformation stabilized by the enzyme. Finally, it was found[48] that the substrate was distorted substantially around proline due to the interaction with the enzyme in both the *cis* and *trans* conformations; e.g., the ζ angle (Scheme 1) goes from 0° to 25° in the *cis* complex and from 180° to 137° in the *trans* complex.

Compared to FKBP,[48] the catalytic mechanism in CypA is very different. The substrate proline carbonyl is hydrogen-bonded with Arg 55 throughout the reaction and therefore desolvation is not a significant factor in CypA. Due to the specific conformation of the substrate peptide, the autocatalytic contribution observed in FKBP is not possible in CypA. As to the contribution from substrate distortion, the effect in CypA is also less significant, based on the observation that the amide torsional angle deviates much less from the ideal values (i.e., 0° for *cis* and 180° for *trans*, see Table 1) than in FKBP. Therefore, the catalytic power of CypA is dominated by the electrostatic contribution from Arg 55 and Asn 102; it is interesting to note that Fischer et al. have already speculated in their study[48] that Arg 55 would contribute to catalysis in CypA on the basis of its different active site compared to FKBP.

One of the factors that motivated the current work is a recent NMR relaxation study of CypA,[41] which suggested that various enzyme residues undergo motions at the same time scale as the chemical step. Since the chemical steps are rare events and occur on the millisecond time scale, it is difficult to directly observe the motions discussed in ref 41 in regular nanosecond MD simulations. In the present work, although umbrella sampling was used to circumvent the rare nature of the barrier crossing events, we still can only hope to observe motions of protein atoms that respond to the chemical step at the subnanosecond time scale. We note that the umbrella sampling technique

(81) Meiering, E. M.; Serrano, L.; Fersht, A. R. *J. Mol. Biol.* **1992**, *225*, 585−589.
(82) Shoichet, B. K.; Baase, W. A.; Kuroki, R.; Matthews, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 452−456.
(83) Kanaya, S.; Oobatake, M.; Liu, Y. *J. Biol. Chem.* **1996**, *271*, 32729−32736.
(84) Elcock, A. H. *J. Mol. Biol.* **2001**, *312*, 885−896.

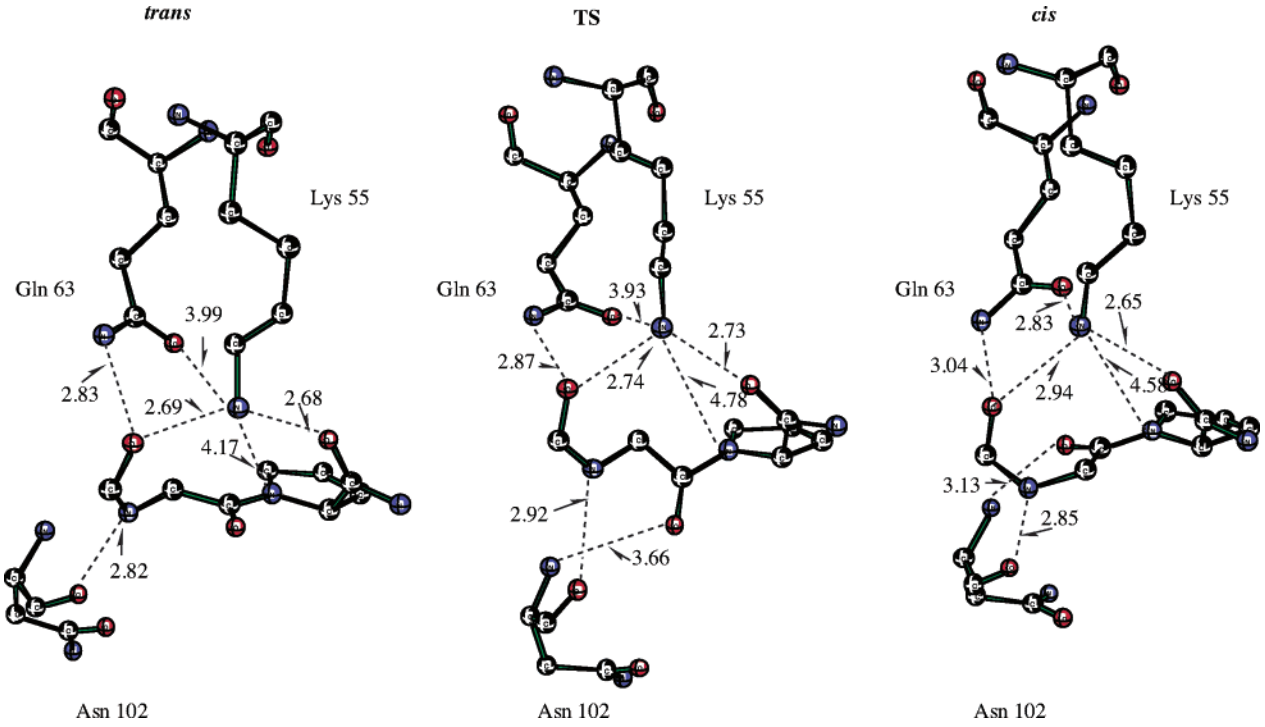


Figure 8. Similar to Figure 3, but for the Arg55Lys mutant.

addresses equilibrium thermodynamic effects during the reaction,[24,85−88] rather than nonequilibrium dynamical effects,[19,22,25,31] which are believed to make relatively small contributions to the rate of enzyme-catalyzed reactions at room temperature.[22,31,89] In both previous[50] and present nanosecond MD simulations, very small displacements were found for the average positions of active site residues (e.g., Arg 55), which suggests that the response of the enzyme active site to the isomerization is not significant in the subnanosecond regime. These nanosecond MD simulations, however, certainly cannot rule out the existence of millisecond time-scale motions during catalysis. Nevertheless, the computed energetics associated with the various enzyme systems in the present study argue that the contribution due to the sub-millisecond collective motion found in ref 41 (which is certainly a great triumph of the NMR analysis) to catalysis is likely to be small. Both previous classical MD and the present QM/MM-PMF calculations found that the decrease in the average distance (∼0.5 Å) between the proline amide N and Arg 55 in the transition state is not due to the displacement of Arg 55 but to the pyramidalization of the proline amide N. Therefore, it is not necessary for Arg 55 to move significantly to achieve its stabilizing effect in the TS. Rather, it is the structural stability of its side chain that makes it "ready" to be catalytic. Indeed, Lys 55 is far more "dynamical" throughout the reaction (Figure 7) but is not locked in a favorable position to stabilize the transition state; therefore, Arg55Lys has a substantially higher barrier than the wild-type enzyme despite the same overall charges and similar sizes of Arg and Lys residues. If the Lys was locked into the same position of Arg 55, we speculate that the same catalytic effect would arise.

Another interesting concept that has been proposed to rationalize enzyme catalysis is the Near Attack Conformer (NAC) population,[90] which somewhat vaguely refers to structures close to the stable state but has certain transition state characteristics; in CypA, for example, a NAC was defined in ref 50 as conformers in which the amide carbonyl group is distorted by ≥20° from the planar geometry.[92] It was proposed that an enzyme shifts the population distribution of the substrate such that more NAC population is found compared to that of the solution case. In CypA, for example, it was found that the ω angle associated with the substrate proline deviates substantially from 0° and 180° in the *cis*- and *trans*-peptide−CypA complexes, respectively.[50] The present work offers a unique opportunity to examine the importance of NAC population because the wild-type and mutant results can be directly compared and correlated with the variations in barrier heights. As shown in Table 2, both the τ and ω dihedral angles, which characterize the peptide conformation somewhat differently, have very similar average values and fluctuations in the wild-type, Arg55Ala, and Arg55Lys enzymes. For example, in the *trans*-peptide−enzyme complexes, the average τ values are 179°, 179°, and 182° for the wild-type, Arg55Ala, and Arg55Lys, respectively, and the fluctuations are 15°, 14°, and 14°, respectively. In the *cis*-peptide−enzyme complexes, the averages (fluctuations) are 10° (7°), 11° (6°), and 9° (6°), respectively. Very similar trends in the ω and η angles can be

(85) Radkiewicz, J. L.; Brooks, C. L., III. *J. Am. Chem. Soc.* **2000**, *122*, 225−231.
(86) Rod, T. H.; Radkiewicz, J. L.; Brooks, C. L., III. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6980−6985.
(87) Agarwal, P. K.; Billeter, S. R.; Rajagopalan, P. T. R.; Benkovic, S. J.; Hammes-Schiffer, S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2794−2799.
(88) Hammes, G. G. *Biochemistry* **2002**, *41*, 8221−8228.
(89) Karplus, M. *J. Phys. Chem. B* **2000**, *104*, 11−27.
(90) Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 139−148.
(91) After the present work was nearly ready for submission, we were informed that the Arg55Lys mutant was found experimentally to be largely inactive toward the proline isomerization (Kern, D., private communication).
(92) The authors also included the condition that the Pro(N)···Arg55(NH2) distance must be smaller than 4.0 Å. This was not included here because clearly no such condition can be defined for proline isomerization in solution.

found as well (Table 2). We note that the averaged values from our study are very consistent with the recent X-ray study on eight complexes between CypA and HIV-1 CA protein variants,[77] which found that ω varied from 173° to 190° with an average of 180.2° for the *trans* complexes, and from 2° to 18° with an average of 9.5° for the *cis* complexes. The fact that large deviations from planarity were found for both *cis* and *trans* complexes in previous MD simulations[50] could be an artifact due to the empirical potential function used in that study. On the basis of the definition of NACs given in ref 50, none of the *trans*-peptide−enzyme complexes in our study exhibit more NACs compared to that in solution; although the substrate in the *cis*-peptide−enzyme complexes deviate substantially from planarity (∼20° in ω) compared to those in solution (∼5°), the magnitude of deviations is very similar in the wild-type and all mutant enzymes. Clearly, NAC populations are not correlated with the trends found in the catalytic properties of various CypA's studied here. Thus the present analysis indicates that the catalytic power of CypA is dictated by the electrostatic contribution from Arg 55 and Asn 102.

The present study illustrates once again the power of computer simulations in understanding enzyme functions,[11,31,89] because artificial systems and mutants can be constructed and analyzed in a highly flexible manner. Artificial systems, such as CypA with a zero-charge Arg 55, are difficult to construct experimentally (although it is somewhat similar to a Leu residue) but are uniquely useful because they can be employed to isolate the contribution of a specific effect. Mutants that can be measured in experiments, such as Arg55Ala and Arg55Lys of CypA, are equally valuable because they can be used to examine the reliability of simulations. It is these flexibilities that make computer simulation a powerful approach for analyzing various contributions to enzyme catalysis without losing close connections to reality.

Acknowledgment. Q.C. acknowledges discussions from Professor D. Kern and her communication on their mutation results after the completion of the current work. The Q.C. group is partially supported by a start-up fund from the Department of Chemistry and College of Letters and Science at the University of Wisconsin, Madison, a PRF-G grant administered by the American Chemical Society, and a Research Innovation Award from the Research Corporation.

JA0367851