

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/7040965>

# The planar conformation of a strained proline ring: A QM/MM study

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · AUGUST 2006

Impact Factor: 2.63 · DOI: 10.1002/prot.21006 · Source: PubMed

CITATIONS

14

READS

7

4 AUTHORS, INCLUDING:



Rik Wierenga

University of Oulu

207 PUBLICATIONS 8,418 CITATIONS

SEE PROFILE



André Juffer

University of Oulu

63 PUBLICATIONS 975 CITATIONS

SEE PROFILE

# The Planar Conformation of a Strained Proline Ring: A QM/MM Study

Serena Donnini,<sup>1</sup> Gerrit Groenhof,<sup>2</sup> Rik K. Wierenga,<sup>1</sup> and André H. Juffer<sup>1\*</sup>

<sup>1</sup>The Biocenter and the Department of Biochemistry, University of Oulu, FIN-90014 University of Oulu, Finland

<sup>2</sup>Max Planck Institute for Biophysical Chemistry, 37077 Goettingen, Germany

**ABSTRACT** QM and QM/MM energy calculations have been carried out on an atomic resolution structure of liganded triosephosphate isomerase (TIM) that has an active site proline (Pro168) in a planar conformation. The origin of the planarity of this proline has been identified. Steric interactions between the atoms of the proline ring and a tyrosine ring (Tyr166) on one side of the proline prevent the ring from adopting the up pucker ( $\chi_1$  is approximately  $-30^\circ$ ), while the side chain of a nearby alanine (Ala171) forbids the down pucker ( $\chi_1$  is approximately  $+30^\circ$ ). To obtain a proline conformation that is in agreement with the experimentally observed planar state, a quantum system of sufficient size is required and should at least include the nearby side chains of Tyr166, Ala171, and Glu129 to provide enough stabilization. It is argued that the current force fields for structure optimization do not describe strained protein fragments correctly. The proline is part of a catalytic loop that closes upon ligand binding. Comparison of the proline conformation in different TIM X-ray structures, indicates that in the closed conformation of TIM the proline is planar or nearly planar, while in the open conformation it is down puckered. This suggests that the planarity possibly plays a role in the overall catalytic cycle of TIM, presumable acting as a reservoir of energy that becomes available upon loop opening. *Proteins* 2006;64:700–710.

© 2006 Wiley-Liss, Inc.

**Key words:** MD; proline puckers; strain; triosephosphate isomerase; force field; active site

## INTRODUCTION

Proline residues can significantly affect the conformation of the polypeptide backbone, because of the constraints on the  $\varphi$  torsions imposed by the five-membered pyrrolidine ring.<sup>1</sup> These constraints are caused by torsional potentials around the C—C and C—N bonds of proline that favor large values of  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$ , but smaller values of  $\chi_4$  and  $\chi_5$ ,<sup>2</sup> where  $\chi_5$  is related to  $\varphi$  as they both describe the rotation around the same N—C $\alpha$  bond (Fig. 1). Recently, it was shown that the C $\gamma$ —C $\delta$ —N bond angle strain, rather than the torsional barriers, could be largely responsible for the observed conformations of the proline ring.<sup>3</sup> A second consequence of these constraints is the commonly observed puckering at the C $\beta$  and/or C $\gamma$  of the

proline ring. The proline side chain ring usually occurs in two predominant conformations.<sup>2,4–6</sup> The difference between these two conformations, known as the down pucker ( $\chi_1$  is approximately  $+30^\circ$ ) and the up pucker ( $\chi_1$  is approximately  $-30^\circ$ ), is defined according to the location of the fifth atom in the ring, C $\gamma$  with respect to the plane defined by C $\alpha$ , C $\beta$ , C $\delta$ , and N (Fig. 1).

With the advance of the technologies for X-ray protein structure determination and, consequently, the ever increasing resolution at which structures are solved, it has been observed that proline, in fact, does not always occur down or up puckered. In such alternative conformations, the pyrrolidine ring can adopt an almost planar conformation. In a recent study performed by Ho and coworkers, a set of 500 very high resolution (i.e.  $<1.8$  Å) protein structures was analyzed for planar prolines and as many as 235 of such prolines were found out of a total of 4525.<sup>3</sup> Despite this observation, the molecular mechanics force fields employed in many X-ray refinement software impose a too severe energy penalty for planarity. Most of the force fields parameters are, in fact, derived from an analysis of the properties of short peptides<sup>6–8</sup> and transferability to proteins is implicitly assumed. The protein environment, however, can induce strain in the structure and could favor particular conformations that are not commonly observed in peptides. It was already pointed out earlier that the puckering of the proline side chain applies to nearly all the compounds free of strain.<sup>2</sup> It is therefore not unlikely that in many low resolution X-ray structures the puckering of the proline is actually an artifact of the force field used in the optimization, while also force fields employed for computer simulation (e.g., molecular dynamics) could suffer from a similar artifact. To address this issue, one needs to understand the structural determinants for the planarity and use that knowledge to improve current force fields.

In this work, we report on a series of MM and QM/MM calculations that were carried out with three structures of wild-type triosephosphate isomerase (TIM). TIM is a dimeric glycolytic enzyme that catalyzes the interconversion of dihydroxyacetone phosphate (DHAP) and D-glyceraldehyde 3-phosphate (DGAP).<sup>9,10</sup> The protein con-

\*Correspondence to: André H. Juffer, The Biocenter and the Department of Biochemistry, University of Oulu, P.O. Box 3000, FIN-90014 University of Oulu, Finland. E-mail: Andre.Juffer@oulu.fi

Received 24 October 2005; Revised 27 January 2005; Accepted 16 February 2006

Published online 1 June 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21006

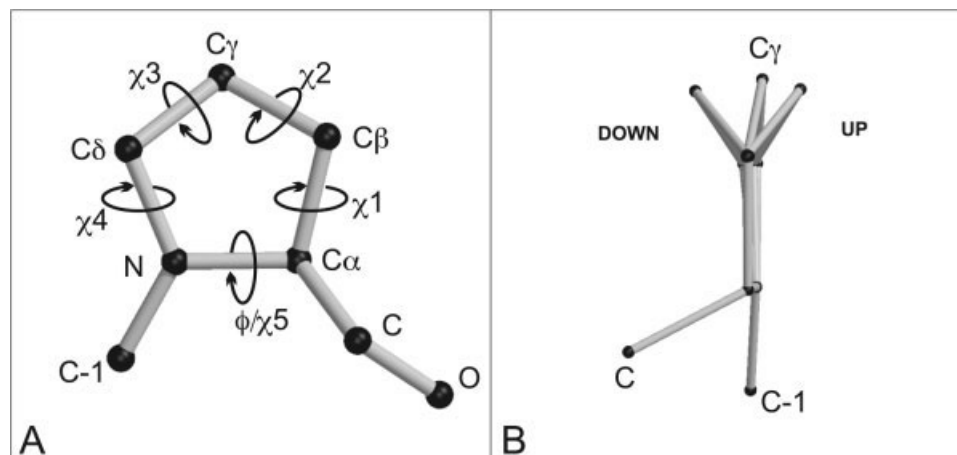


Fig. 1. (A) Schematic representation of proline with atom names and dihedral angles. The C-1 atom is the C atom of the previous amino acid. Dihedrals  $\phi$  and  $\chi_5$  are defined by the atoms C-1, N, C $\alpha$ , C, and C $\delta$ , N, C $\alpha$ , C $\beta$ , respectively. (B) Side view of down, planar, and up conformations of proline. The C $\alpha$  atom is towards the reader. The C $\gamma$  is on the proline ring (planar,  $\chi_1 \approx 0^\circ$ ) or is displaced towards left (down,  $\chi_1 \approx 30^\circ$ ), or right (up,  $\chi_1 \approx -30^\circ$ ). Molscrip<sup>46</sup> 2.1.2 was used for the figure.

tains a proline (Pro168) in close proximity to the active site. In a liganded atomic resolution structure of the enzyme (PDB code 1N55), this proline ring adopts a planar conformation. In the unliganded conformation, Pro168 is down pucker. To understand the determinants for the planarity of the proline ring, we have compared the conformational energies of Pro168 of TIM in vacuum and in the active site of the enzyme and investigated the interactions of this proline with its protein environment. Our calculations are consistent with a planar proline in the context of the closed, liganded active site and quantitatively explain why the planar conformation is favored. These calculations could support the efforts to further improve the quality of force fields to allow for the occurrence of a planar proline, induced by the protein environment.

## METHODS

Potential energies were computed for different configurations of the pyrrolidine ring of Pro168 of triosephosphate isomerase (TIM). In addition, also the potential energy profiles for the down-to-up pucker transition were calculated. The computations were performed both for an isolated proline in vacuum and a proline embedded in its natural protein environment, which not only includes all the atoms of the protein, but also the solvating water molecules.

Because in the high-resolution X-ray structure this ring adopts a rather uncommon conformation, the computations provide a critical test for a biomolecular force field. We have therefore optimized the proline ring with various popular force fields for MM simulations including GROMOS96,<sup>8</sup> OPLS,<sup>11</sup> and AMBER99.<sup>12,13</sup> In addition, to avoid possible artefacts due to these MM force fields and to model the proline and its environment more correctly, another series of calculations were performed using the hybrid quantum mechanics/molecular mechanics (QM/

MM) approach,<sup>14</sup> for which a special version of Gromacs 3.3<sup>15,16</sup> with an interface to Gaussian03<sup>17</sup> was employed.<sup>18</sup>

In all QM/MM calculations, the proline, including a stretch of the backbone that ranged from the C $\alpha$  of Glu167 to the C $\alpha$  of Val169 and nearby residues, were described at the B3LYP/6-31G\* level of theory. These atoms constituted the QM subsystem, while all other atoms were described at the MM level. The computations were repeated with an increasing number of atoms in the QM subsystem [Fig. 2(B–F)]: 26 (QM1), 48 (QM2), 60 (QM3), 129 (QM4), and 201 (QM5) atoms. These roughly correspond to spheres centered at the proline with increasing radii up to about 0.5 nm in QM5. The remainder of the system (MM), consisting of the rest of the protein, the inhibitor (PGA, 2-phosphoglycolic acid), and approximately 20,000 SPC water molecules<sup>19</sup> in a truncated dodecahedron box with periodic boundary conditions, was modeled with the GROMOS96<sup>8</sup> force field. For comparison, the calculations on the QM1 and QM4 subsystems were also performed using the OPLS<sup>11</sup> force field. The bonds connecting the QM and MM subsystems were replaced by constraints,<sup>20</sup> and the QM part was capped with hydrogen atoms. The forces acting on these cap atoms were distributed over the atom pairs that formed the original bond. The QM system experienced the Coulomb field of all MM atoms and Lennard-Jones interactions between QM and MM atoms were also added.

In the energy minimizations, both at the MM and QM/MM level, only the positions of the protein atoms, including those of the PGA ligand, were allowed to relax. The water atoms were kept fixed, because minimizing that many water molecules is not only very difficult and computationally expensive, but would also lead to an unrealistic protein in an ice crystal. The minimizations were performed with the Broyden-Fletcher-Goldfarb-Shannon limited-memory quasi-Newton algorithm (BFGS), using a convergence threshold of 1 kJ mol<sup>-1</sup> nm<sup>-1</sup>. In the

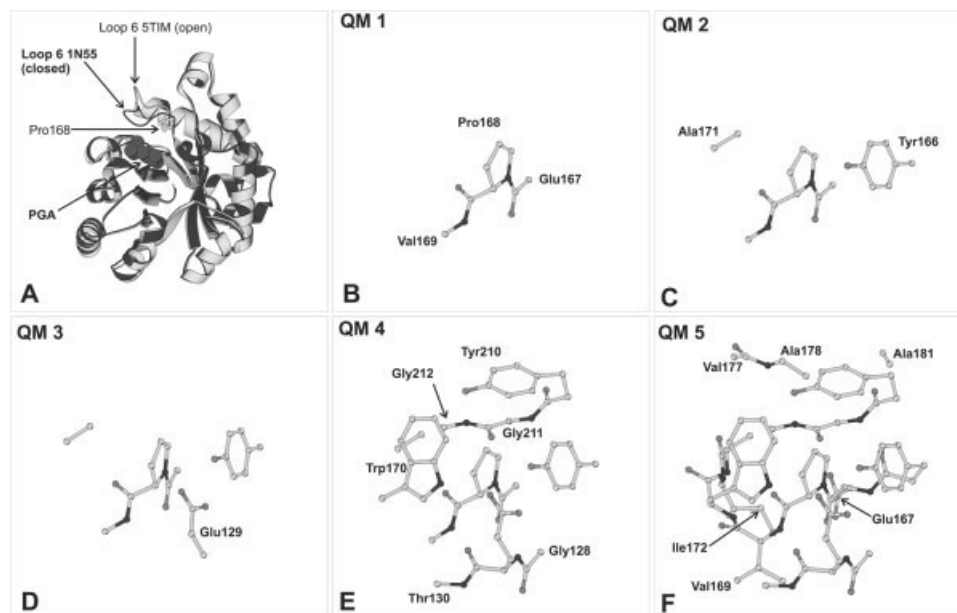


Fig. 2. (A) Superposition of 1N55A (black) and 5TIMA (gray) on the backbone atoms. Ligand of 1N55 is PGA. Side chain atoms of Pro168 of 5TIM are shown. (B–F) QM subsystems QM1, QM2, QM3, QM4, and QM5 consisted of 26, 48, 60, 129, and 201 atoms, respectively. Only residues not present in the previous QM subsystem are indicated. Molscript<sup>46</sup> 2.1.2 was used for the figure.

QM/MM calculations, the QM subsystem was further minimized using the more efficient Gaussian routines to a threshold of 0.00045 au. All MM atoms were kept frozen in this final minimization phase. The QM/MM minimizations were performed only on the QM1 and QM4 subsystems, with both the OPLS and GROMOS96 force field to describe the remainder of the system.

To estimate the contribution of each residue to the overall stabilization of the planar proline conformation, the QM/MM computations were repeated with a residue or the side chain of a residue left out completely. The latter roughly corresponds to a point mutation of the respective amino acid to glycine, although the structure was not allowed to relax following the mutation. To avoid potential force field artefacts, only atoms that are inside the QM subsystem were considered in these “mutation” studies.

### PDB Structures

We have used three RCSB-PDB<sup>21</sup> structures of (TIM) in the QM/MM calculations. These are identified by the codes: 1N55,<sup>22</sup> 1NEY,<sup>23</sup> and 5TIM.<sup>24</sup> The resolutions of the structures were 0.83, 1.2, and 1.83 Å, respectively. Structures refined at 0.83 Å have an average error in atomic position of approximately 0.03 Å for well-defined regions of its structure, whereas for 1.8-Å structures the corresponding position error is approximately 0.3 Å. The out-of-plane puckering concerns a shift of the C $\gamma$  atom of approximately 0.5 Å, which is much larger than the positional accuracy at 0.83-Å resolution but comparable to that of the lower resolution structures. For the numbering of the residues, we always followed the numbering scheme of 1N55. 1N55 and 1NEY are both liganded wild-type TIM structures, but

from different organisms. 1N55's source organism is *Leishmania mexicana* and its ligand is the transition state analog PGA (2-phosphoglycolic acid),<sup>25</sup> while 1NEY's source organism is *Saccharomyces cerevisiae* and its ligand is the substrate DHAP (dihydroxyacetone phosphate). Upon ligand binding, TIM adopts a closed conformation in which loop 6 (from residues 166 to 176) closes over the active site. In the unliganded form of the enzyme, loop 6 is in an “open” conformation.<sup>26,27</sup> TIM is a dimer; Pro168, and its environment were described at the QM level in the monomer that corresponds to subunit A of the PDB structures. The other monomer (subunit B) was modeled completely by the MM force field. Subunits A of 1N55 and 1NEY are liganded (closed active site), while subunit A of 5TIM is unliganded (open active site).

For a comparison of the conformation of Pro168 in other liganded and unliganded X-ray structures of TIM, eight unliganded (PDB codes and chain identifier 5TIMA, 1YPIA, 8TIMA, 1I45A, 1TPFA, 1TPEA, 1AG1A, 6TIMA) and 12 liganded TIM structures (PDB codes and chain identifier 1IF2A, 1LYXA, 1N55A, 1NEYA, 1NEYB, 1TPHA, 1AMKA, 1IIGB, 4TIMB, 1IHHB, 6TIMB, 1AG1B) were analyzed.

### Vacuum

The vacuum calculations were performed on a stretch of TIM atoms that ranged from the C $\alpha$  of Glu167 to the C $\alpha$  of Val169 (QM1). Calculations were repeated on this system with down, up, and planar conformations of the proline ring fitted onto the backbone of Pro168. Gaussian03<sup>17</sup> was used for all calculations and the QM subsystems were described at the B3LYP/6-31G\* level of theory.



The starting coordinates for the down, up, and planar conformations of the side chain of Pro168 were taken from the side chain atoms of Pro58 (down), Pro49 (up), and Pro168 (planar) itself of the TIM PDB structure 1N55. These prolines were fitted onto the backbone atoms of Pro168 of TIM 1N55. Hydrogens were added with Molden.<sup>28</sup> Each system was optimized in vacuum. The coordinates of the heavy backbone atoms (except for the hydrogens) were kept fixed. During the optimization of the planar conformation, the dihedrals of the side chain atoms of the proline were kept fixed as well. This procedure of generating the down, up, and planar conformations and computing their energies was repeated with the Pro168 of the other PDB structures (1NEY and 5TIM).

The energy profile for the down-to-up transition was calculated for the smallest QM subsystem (QM1) using the Synchronous Transit-Guided Quasi-Newton (STQN) method of Schlegel and coworkers.<sup>29</sup> A total of 12 intermediate points were employed to construct a suitable minimum energy path connecting the down and up minima. Because atoms cannot be kept fixed during this type of optimization, it was necessary to rely on a somewhat smaller system than the original system QM1. The reduced system consisted of the proline side chain, the C $\alpha$ , and C=O of the preceding residue Glu167 and a methyl group that substituted the C=O carbonyl of the proline. Using the original QM1 system without constraining the positions of the backbone atoms, would have resulted in deformations of that backbone, rendering a direct comparison to the protein system more difficult. Instead, the atoms that were removed from the system prior to the path calculations, were added afterwards to the optimized structures of the path and the energies were reevaluated. In all calculations the system was described at the B3LYP/6-31G\* level.

To test the effect of solvation in media with different dielectric properties, the polarizable continuum model (PCM)<sup>30,31</sup> was employed with Gaussian03.<sup>17</sup>

## Protein

The starting structures for the QM/MM computations were the X-ray structures of the three different TIM proteins (1N55, 1NEY, and 5TIM) around which water was placed and equilibrated. The up, down, and planar conformations of Pro168 in these proteins were created by replacing the Pro168 side chain in each of the X-ray structures by the down, up, and planar conformations optimized in vacuum (see above) after a least-squares fit onto the proline backbone atoms.

The water equilibration simulations were performed with the molecular dynamics (MD) program Gromacs<sup>15,16</sup> together with the GROMOS96 force field.<sup>8</sup> The force field parameters for the phosphate moiety of the ligands were derived from the force field parameters of phosphoserine.<sup>32,33</sup> TIM was solvated in a dodecahedral box with approximately 20,000 SCP (Simple Point Charge) water molecules.<sup>19</sup> Crystallographic water molecules were included. The waters were minimized during 100 steps using a steepest descent approach, followed by a 10-ps MD

simulation with the heavy atoms of the protein position frozen. A twin range cutoff was used for the Coulomb and Lennard-Jones interactions. For the electrostatic energy, a reaction field correction was applied. Interactions between atoms within 1.0 nm were evaluated every step, while interactions between atoms within 1.8 and 1.6 nm for the Coulomb and Lennard-Jones interactions, respectively, were evaluated every 10 steps. The bond distances and the bond angle of the water molecules were constrained using the SETTLE algorithm.<sup>34</sup> All other bond distances were constrained using the LINCS algorithm.<sup>20</sup> Constant pressure  $p$  and temperature  $T$  were maintained by weakly coupling the system to an external bath at 1 bar and 298 K using the Berendsen barostat and thermostat,<sup>35</sup> with coupling times of 1.0 and 0.1 ps, respectively. The final frames of the simulations were used as the input structure for the QM/MM calculations (see above).

The energy profile for the down-to-up transition of the proline side chain in the protein was computed by calculating QM/MM energies for every intermediate conformation on the minimum energy path connecting the down and up puckers. These intermediate conformations were obtained from the vacuum energy profile calculation (see above) after a least-squares fit onto the proline 168 backbone atoms. No optimization was performed. In these calculations all atoms belonging to QM subsystem 4 were described at the B3LYP/6-31G\* level while for the remaining atoms the GROMOS96 force field was used.

## Database analysis

Proline residues from a 500 very high-resolution nonhomologous structures subset of the PDB, provided by the Richardson laboratory,<sup>36</sup> were employed to generate a frequency distribution of  $\chi_1$  and  $\chi_2$  dihedral angles (degrees) of proline residues.

## RESULTS

In the atomic resolution structure of TIM (triosephosphate isomerase) from *L. mexicana* (PDB code 1N55) Pro168 has a planar conformation of the pyrrolidine ring. The  $\varphi$  and  $\psi$  dihedral angles measure 175 and 133 degrees, respectively, and fall therefore in the permitted region of the Ramachandran plot. The peptide bond between this proline and the preceding glutamate is in a *trans* configuration.

Table I reports the three torsion angles defining the conformation of this proline ring after the energy minimizations. None of the force fields employed in this work (GROMOS96, OPLS, and AMBER99) is able to reproduce the experimentally observed planar conformation, for which the torsion angles  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  measure  $-2.6$ ,  $6.5$ , and  $-7.6$  degrees, respectively. Even with the smallest QM subsystem [QM1, 26 atoms, Fig. 2(B)], the QM/MM minimization did not result in the correct conformation. Only if those residues that are within 0.45 nm from the proline ring are explicitly included in the QM subsystem [QM4, 129 atoms, Fig. 2(E)], does one find the observed planar ring structure. It is noted that the last observation does not depend on whether one employs GROMOS96 or OPLS to model the MM part of the system.

**TABLE I. Proline168 Dihedral Angles (Degrees) after Geometry Optimization**

	$\chi_1$	$\chi_2$	$\chi_3$
Molecular mechanics			
AMBER99	20.7	-26.8	28.2
GROMOS96	-26.0	33.8	-28.9
OPLS	11.0	-11.9	7.9
QM(B3LYP/6-31G*)/MM(GROMOS)			
QM1	-23.0	33.1	-30.2
QM4	-1.7	8.7	-11.9
QM(B3LYP/6-31G*)/MM(OPLS)			
QM1	17.4	-14.9	6.4
QM4	-2.5	8.1	-10.5
Experiment (TIM)			
	-2.6	6.5	-7.6

**TABLE II. Relative Energies (kJ mol<sup>-1</sup>) of Down, Up, and Planar Conformations of Pro168 of 1N55, 1NEY, and 5TIM Structures in Vacuum (QM1: 26 Atoms) and in the Protein (QM4: 129 Atoms)**

PDB	Down	Up	Planar
Vacuum (QM1)			
1N55	0.0	0.75	13.4
1NEY	0.0	1.2	13.6
5TIM	0.0	1.4	15.6
Protein (QM4 + MM)			
1N55	0.0	-1.0	-10.0
1NEY	0.0	-18.0	-16.0
5TIM	0.0	18.0	29.0

The zero energy for each system is always the down pucker energy.

Table II lists the energies for proline in vacuum and in the protein environment. The vacuum calculations were performed for QM1 [Fig. 2(B)], which consists of the proline side chain and a stretch of the backbone ranging from the C $\alpha$  of Glu167 to the C $\alpha$  of Val169 (a total of 26 atoms). The calculations in the protein environment, where the proline is embedded in a fully solvated protein, were carried out for QM4 + MM, where the QM subsystem consisted of QM1 plus a number of nearby residues [a total of 129 atoms, Fig. 2(E)] and the MM part included the rest of the system. The zero energy for each system is taken as the down pucker energy. It is seen in Table II that in vacuum the planar conformation always has the highest energy (13.4, 13.6, and 15.6 kJ mol<sup>-1</sup> for 1N55, 1NEY, and 5TIM, respectively), whereas in the protein it has the lowest energy for 1N55 (-10.0 kJ mol<sup>-1</sup>). In vacuum, there is no clear preference for either the down or up pucker. The QM4 + MM computations for 1NEY show that the energies of the up and planar conformations are almost the same (-18 and -16 kJ mol<sup>-1</sup>, respectively), but clearly lower than the down conformation, while for 5TIM, the energy of the down conformation is 18 kJ mol<sup>-1</sup> more favorable than the up and 29 kJ mol<sup>-1</sup> more favorable than the planar conformation.

The  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  torsions that characterize the down, up, and planar prolines in these calculations are listed in Table III and are the result of a partial optimization of the proline in vacuum, in which the backbone was fixed, as

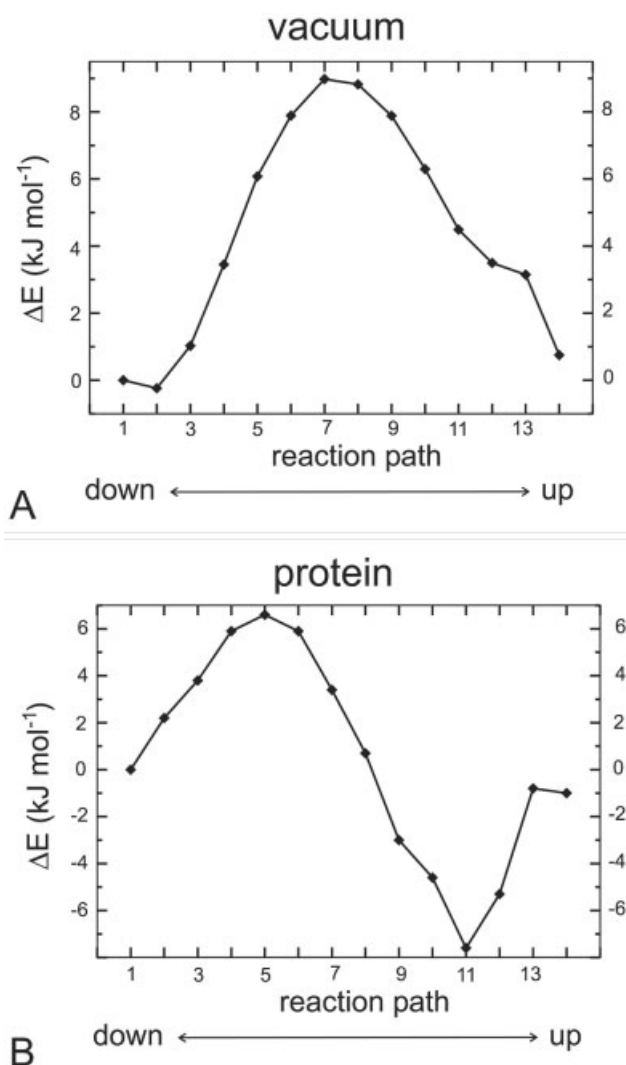
well as the side chain torsions in the case of the planar conformation (see Methods). In general, the down pucker is characterized by having  $\chi_1 > 0$  and the up pucker by  $\chi_1 < 0$ . In the planar conformation, however, the absolute value of all three dihedral angles is less than 10 degrees. The structures of both the up and down conformations of the proline are similar in 1N55 and 1NEY, whereas they differ from 5TIM. This correlates with a similar main chain trace for 1N55 and 1NEY (liganded, closed) which is different from 5TIM (unliganded, open). Upon ligand binding, loop 6 of TIM (residues 167–169) “closes” on the active site.<sup>26,27</sup> The tip of this loop moves by about 0.8 nm upon loop closure. Pro168 is positioned at the N-terminal end of the loop and the conformation of the backbone at this site significantly differs in the structures of the liganded (e.g., 1N55 and 1NEY) and unliganded enzyme (e.g., 5TIM) [Figs. 2(A) and 5(B)]. Therefore, during optimization of the proline side chain, it is not expected that the down and up puckers of these structures are equal.

The energy profile for the transformation between the down and up pucker has been calculated in vacuum and in 1N55. In vacuum (isolated QM1), the transition state for the transformation between the down and up pucker corresponds to point 7 along the reaction path [Fig. 3(A)]. The energy of this transition state was almost 9 kJ mol<sup>-1</sup> higher than the down pucker energy (point 1). The  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  torsions for this structure measure 22, -8, and -9 degrees, respectively. Of all structures along the reaction path, point 7 is the only one in which more than one dihedral measures less than 10 degrees; therefore, it is the most strained of the structures. The five pyrrolidine ring dihedrals of Pro168 in 1N55 all measure less than 10 degrees ( $\chi_1 = -2.6$ ,  $\chi_2 = 6.5$ , and  $\chi_3 = -7.6$  degrees). To establish which of the points on the minimum energy path connecting the up and down puckers resembles the planar conformation of Pro168 in 1N55 most closely, the root-mean-square deviation (RMSD) between the ring atoms of the points and the original 1N55 planar Pro168 were computed. The lowest RMSD (0.0107 nm) was found for point 11 ( $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  of 1.5, 17, and -28 degrees, respectively). The  $\chi_1$  of the conformation corresponding to point 11 and the planar proline of 1N55 differ the least of all structures. This dihedral is the one that defines the position of the proline's C $\gamma$  (on the tip of the ring; Fig. 1).

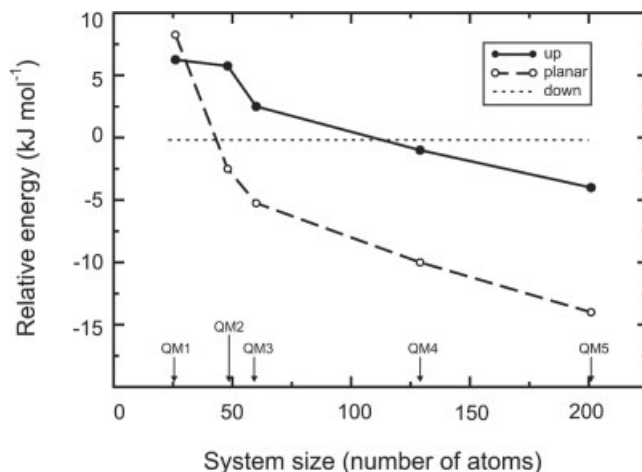
When the same transformation from down to up pucker occurs in the protein, the energy profile changes dramatically [Fig. 3(B)]. Because the proline structures in the protein and in vacuum are the same, the internal energy of the proline ring was the same as well. The deviations of the curve are therefore solely due to the interaction of the proline with the protein environment. The difference in energy between the vacuum transition state (point 7) and the down pucker has decreased to about 6 kJ mol<sup>-1</sup> in the protein environment. In the protein, the minimum of the energy profile corresponds to point 11 along the reaction path. This structure resembles the original conformation of Pro168 of 1N55 the most, based on the RMSD analysis described above.

**TABLE III. Dihedral Angles (Degrees)  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  of the Down, Up, and Planar Conformations of Pro168 Used in the Calculations**

PDB	$\chi_1$			$\chi_2$			$\chi_3$		
	Down	Up	Planar	Down	Up	Planar	Down	Up	Planar
1N55	27.8	-19.6	-2.6	-36.2	34.2	6.5	30.1	-36.2	-7.6
1NEY	29.1	-19.4	-2.6	-37.4	34.9	6.5	30.8	-37.1	-7.6
5TIM	32.1	-3.9	-2.6	-33.0	25.1	6.5	20.7	-36.4	-7.6

**Fig. 3.** Energy profile for the down/up pucker transition of (A) proline in vacuum (isolated QM1) and (B) in the protein (QM4 + MM). The zero energy for each system is taken as the down pucker energy.

To estimate the effect of the size of the QM subsystem (i.e., the number of atoms that is considered at the the QM level in the calculations) and the convergence of the overall QM/MM energy with respect to QM subsystem size, we have repeated the protein QM/MM calculations for TIM's structure 1N55 with an increasing number of atoms in the QM subsystem (Fig. 2). The results of the calculations with the different QM subsystem sizes are shown in Figure 4. They clearly show that it is necessary to at least include the side chains of Glu129, Tyr166, and Ala171 [QM3, 60

**Fig. 4.** Relative QM/MM energies (kJ mol<sup>-1</sup>) of the down, up, and planar conformations of Pro168 of 1N55 TIM structure with inclusion of different number of atoms in the QM subsystem (QM1-5). The zero energy for each system is always the down pucker energy.**TABLE IV. Relative Energies (kJ mol<sup>-1</sup>) of the Down, Up, and Planar Conformations of Pro168 in 1N55 Using QM Subsystem 4, Described at the B3LYP/6-31G\* Level of Theory**

Deletion	Down	Up	Planar
—	0.0	-4.1	-11.2
Ala 171 <sub>sc</sub>	0.0	12.0	4.9
Tyr 166 <sub>sc</sub>	0.0	-20.6	-11.0
Glu 129 <sub>sc</sub>	0.0	3.4	-4.7
Glu 129 <sub>bb</sub>	0.0	-5.1	-12.2
Trp 170 <sub>sc</sub>	0.0	-6.4	-11.7
Tyr210 Gly 211 Gly212	0.0	-3.8	-10.1
Trp 170 <sub>sc</sub> Glu129 <sub>bb</sub>	0.0	-7.7	-18.5
Trp 170 <sub>sc</sub> Tyr210 Gly 211 Gly212	0.0	-27.3	-27.5
Glu 129 <sub>bb</sub>			

The zero energy for each system is always the down pucker energy. bb, backbone atoms; sc, side chain atoms.

atoms, Fig. 2(D)] to reach a significant stabilization of the planar conformation with respect to the down and up puckers. In addition, it is seen that the stabilization of the planar conformation with respect to the down and up puckers increases with system size.

To estimate the contribution of individual residues or chemical groups to the stability of proline, a series of deletion experiments was carried out for QM4 in 1N55 [Fig. 2(E)]. The results are presented in Table IV. The MM system was modeled with the GROMOS96 force field. Decomposing the QM and MM contributions shows that



the MM system has only a minor destabilizing effect on the planar and up conformation. This observation also suggests that the stabilization of the planar conformation is due to short range interactions only (compare the first row of Table IV with Table II), which is also supported by the results obtained about the effects of system size presented above. The results in Table IV indicate that the up and planar conformations were respectively  $-4.1$  and  $-11.2$  kJ mol $^{-1}$  more stable than the down pucker compared to the  $-1.0$  and  $-10.0$  kJ mol $^{-1}$  stabilization as calculated for the completely solvated protein QM/MM system. The subsystem originally included all atoms in a sphere of about 0.45 nm around the proline (129 atoms). By deleting groups of atoms in this sphere, and repeating the calculations, the effect of the individual groups of atoms can be estimated. However, because the calculations do not take into account the fact that interactions between groups of atoms may be correlated, these computations can only give an indication of the contribution of specific groups. When deleting a complete side chain, a rough approximation of the effect of a point mutation of the respective amino acid to glycine is obtained.

The most important residues for stabilizing the planar proline ring are Ala171, located on the down pucker side of the ring, and Tyr166 and Glu129, whose side chains are located on the other side of the ring (up pucker side). These three residues are highly conserved in the TIM protein family.<sup>38</sup> When the side chain of Ala171 is left out of the calculation, the down pucker becomes the most favorable conformation with relative energies of 12.0 and 4.9 kJ mol $^{-1}$  for the up and planar conformations, respectively (Table IV). When the side chain of Tyr166 is deleted, the system favors the up pucker conformation:  $-20.6$  and  $-9.6$  kJ mol $^{-1}$  more favorable than the down and planar conformations, respectively. The side chain of Glu129 also appears to affect the stabilization of the planar proline: the planar conformation is only  $-4.7$  kJ mol $^{-1}$  more favorable than the down pucker when it is deleted. Glu129 does not interact directly with Pro168, but it is hydrogen bonded to Tyr166. The energy of the up pucker is also less favorable with respect to the down pucker when this interaction is missing: it is 3.4 kJ mol $^{-1}$  more positive than the down pucker. Deletion of Trp170 side chain or of the stretch of atoms from Tyr210 to Gly212 [Fig. 2(E)] does not have a large influence on the energies of the three conformations of proline. Interestingly, deletion of Trp170 and Glu129 backbone makes the planar conformation  $-18.5$  kJ mol $^{-1}$  more favorable than the down pucker. As already indicated above, the individual contribution of groups is not easily identified. Deleting all other groups of atoms in the QM subsystem, except for the side chains of Ala171, Tyr166, and Glu129 stabilizes both the up and planar conformation by about the same amount, indicating that those other groups play an essential role in the specific stabilization of the planar conformation as well.

To investigate if the polarity of the surrounding medium affects the relative energies of the three conformations of the proline ring, the computations were repeated for an isolated proline [QM1, Fig. 2(B)] in different dielectric

**TABLE V. Relative Energies (kJ mol $^{-1}$ ) of the Down, Up, and Planar Conformations of the Isolated Pro168 (QM1: 26 Atoms) from 1N55 Calculated at Different Relative Dielectric Constant ( $\epsilon$ ) Values**

$\epsilon$	Down	Up	Planar
1	0.0	0.7	13.4
2	0.0	0.4	13.5
4	0.0	-0.1	12.8
10	0.0	-0.4	12.3
50	0.0	-0.6	12.0
80	0.0	-0.5	12.1

The zero energy for each system is always the down pucker energy.

media. Table V shows that an increase of the dielectric constant (i.e., the solvent becoming more polarizable), does not significantly change the relative energies of the down, up, and planar ring conformations. The down conformation becomes slightly less stable with respect to the other two (about 1 kJ mol $^{-2}$ , Table V). This supports the notion that the stabilization is due to specific interactions within the protein, as was asserted earlier.

For comparison purposes, the relative energies of the down, up, and planar conformations of Pro168 in the QM subsystem 4 (129 atoms) were also computed for 1NEY and 5TIM using the QM/MM calculations (Table II). The resolution of these two structures is lower than 1N55. In TIM from yeast (1NEY, Table II), the up pucker is about 2 kJ mol $^{-1}$  more stable than the planar conformation and about 18 kJ mol $^{-1}$  more stable than the down pucker. Pro168 in 1NEY has adopted an almost planar conformation; the torsions  $\chi_1$  and  $\chi_2$  are  $-7.9$  and  $12.5$  degrees, respectively. The C $\gamma$  points slightly to the phenyl ring of Tyr166. In Figure 5(A), a comparison of the proline environment of 1N55 and 1NEY is shown, while Figure 5(C) displays the differences in the conformation of the prolines in the three structures of TIM, 1N55, 1NEY, and 5TIM. The structures of 1N55 and 1NEY have been superimposed using a least-squares fit of the proline backbone atoms. A crystallographic water is present at a distance of 0.35 nm from the C $\gamma$  of the proline 168 in 1NEY. This water was included in the QM subsystem, so that for 1NEY the total number of atoms in QM4 was 132 instead of 129. If this water was described at the MM level, the relative energies calculated with respect to the down pucker were  $-17.0$  and  $-17.0$  kJ mol $^{-1}$  for the up and planar conformation of the proline, which is very similar to the numbers given in Table II.

The third TIM considered, 5TIM, has no ligand bound in the active site. Pro168 is located nearby the active site and, due to the absence of the ligand, its local environment is different in this conformation of the protein [Fig. 5(B)]. In the unliganded 5TIM structure, the down pucker is the most stable of the three conformations of the proline (Table II). In full agreement with this result, Pro168 in the X-ray structure of 5TIM is down puckered, with torsion angles  $\chi_1$  and  $\chi_2$  of  $41.0$  and  $-39.6$  degrees, respectively. Figure 5(B) shows a superposition of the 5TIM and 1N55 structures, using the proline backbone atoms for the least squares fit. Residues 166 to 176 belong to loop 6 [Fig. 2(A)]. When the



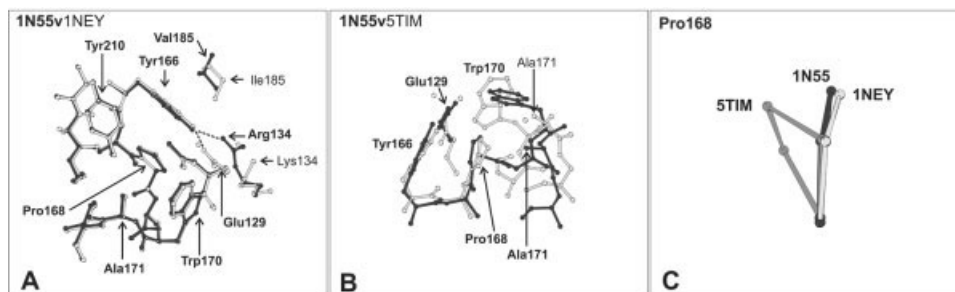


Fig. 5. (A) 1N55A and 1NEYA superposed by the proline backbone atoms. (B) 1N55A and 5TIMA superposed by the proline backbone atoms. (C) Superposition of Pro168 from the X-ray structure of 1N55A, 1NEYA, and 5TIMA. 1N55 is always colored black while the other structures are gray. Gray labels indicate residues of 1NEY and 5TIM when different from the ones labeled in 1N55 (bold). Spheres represent O atom of crystallographic water molecules. Molscript<sup>46</sup> 2.1.2 was used for the figure.

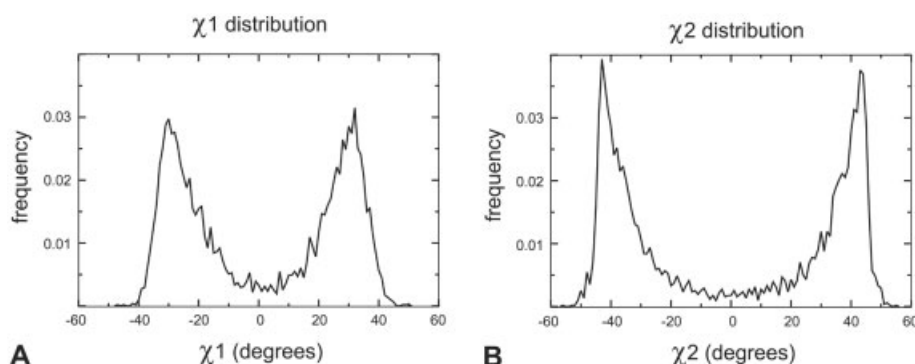


Fig. 6. Frequency distribution of  $\chi_1$  and  $\chi_2$  dihedral angles (degrees) of proline residues from a 500 high-resolution structures subset of the PDB.

loop is open (5TIM), Ala171 is slightly further away from the proline ring with respect to its position in the closed liganded (1N55) conformation [Fig. 5(B)] and the tip of the ring ( $C\gamma$ ) can relax there.

A comparison of eight unliganded and 12 liganded TIM structures with resolution ranging from 0.83 to 2.6 Å showed that in all unliganded structures Pro168 is down pucker (average  $\chi_1$  is 26 degrees), whereas in the liganded structures, the  $\chi_1$  measures 10 degrees on average, indicating a near planar conformation of the proline ring [Fig. 1(B)].

Proline residues (both *trans* and *cis*) from the set of 500 high resolution protein structures were employed to generate a frequency plot of the  $\chi_1$  and  $\chi_2$  dihedral angles (degrees; Fig. 6). Values of  $\chi_2 < -10$ ,  $\chi_2 > 10$  and  $-10 < \chi_2 < 10$  degrees are used to identify the down, up, and planar conformations of proline, respectively.<sup>3,5</sup> Figure 6 identifies two clusters of points corresponding to the down and up puckers. In contrast, prolines with a planar conformation are not grouped into a separate cluster, but rather represent a continuum of conformations between the down and up puckers and one observes a smooth transition between the down and up puckers.

## DISCUSSION

We have examined about 5000 prolines from a high resolution ( $<1.8$  Å) subset of the PDB.<sup>36</sup> From the  $\chi_1$  and

$\chi_2$  distribution (Fig. 6), we identified a continuum of conformations of the ring going from the down to the up pucker. This suggests that the occurrence of conformations of the proline ring possibly can be explained by a destabilization of the up and down pucker due to specific interactions with the protein environment. To validate this hypothesis, we investigated a specific planar proline from an atomic resolution (0.83 Å) triosephosphate isomerase (TIM) structure (PDB code 1N55). This proline (Pro168) is positioned in close proximity to the active site of the enzyme [Fig. 2(A)]. The nearest residues around the proline are an alanine (Ala171) and a tyrosine (Tyr166) that are located on the two sides of the planar ring [Fig. 2(C)]. Ala171 destabilizes the down conformation that would pucker toward Ala171, while Tyr166 destabilizes the up conformation of the ring that would pucker towards Tyr166 (Table IV). No long-range interactions were involved in the stabilization of the planar conformation. Including all atoms in a sphere with radius of about 0.45 nm around the proline ring was already sufficient to reproduce the relative energies of the three conformations that were computed with all protein and solvent atoms taken into account (Table II and IV). Therefore, the origin of TIM's planar conformation appears to be the steric interaction of proline with its local environment.

The maximum of the energy profile for the transition from down to up was estimated to be about 9 kJ mol<sup>-1</sup> in

vacuum, in good agreement with previous calculations,<sup>2,39</sup> from which our calculations differ in the fact that a (frozen) stretch of backbone atoms from the residues adjacent the proline was included. We observed a dramatic change in the energy profile when the protein environment is taken into account. Because the structures corresponding to the points on the minimum energy path connecting the down and up puckers were the same in the vacuum and the protein calculations, the difference between the two can only arise from the interactions of the proline ring with the rest of the protein. The lowest energy structures along the reaction path in the protein all had values of the  $\chi_1$  below 10 degrees. In the protein, repulsive steric interactions of the aliphatic pyrrolidine ring, in particular those of the tip of the ring, with nearby residues overcome the internal strain caused by planarity. In the transition state (point 7) on the minimum energy path [Fig. 3(A)], the  $\chi_1$  torsion measures 22 degrees. Because this torsion determines the position of the C $\gamma$  at the tip of the loop, the overall similarity of this structure to the Pro168 structure of 1N55 is actually somewhat smaller than for points 10 to 12.

The environment of the Pro168 of TIM changes upon ligand binding.<sup>27</sup> This proline is located at the base of loop 6 (residues 166 to 176). Loop 6 acts like a “lid” that in the liganded form of TIM is “closed” on the active site and interacts directly with the ligand, while in the unliganded form, it is “opened” towards the solvent<sup>26,27</sup> [Fig. 2(A)]. The distance between the tip of the loop in the liganded and unliganded conformation is about 0.8 nm. Ala171 is in this loop. In the unliganded conformation, the distance between the Pro168 ring and the Ala171 side chain is sufficiently large for the proline ring to adopt the down puckered conformation [Fig. 5(B)]. This conformation was also found to be the most stable in the series of calculations for the unliganded 5TIM structure (Table II). However, upon ligand binding, the loop closes and Ala171 is pushed closer to Pro168. The resulting steric repulsion is sufficiently high to compensate for the energy that is required for the ring to adopt a planar conformation. Upon loop closure, the proline could store (part of) the energy, which, after completion of the enzymatic conversion, is readily available to open the loop again and facilitate product release. This possibly suggests an important functional role for the proline, conserved in all TIM proteins,<sup>40</sup> in the overall catalysis process. This hypothesis is supported by the observation that the proline is planar or near planar in all closed structures of liganded TIM (on average  $\chi_1$  10 degrees), and down puckered when TIM is in its open unliganded conformation ( $\chi_1$  averages to about 26 degrees).

The functional importance of geometric distortion, particularly in the proximity of active sites, is a notion that has been introduced previously.<sup>41,42</sup> In particular, Heringa and Argos have suggested that ligand binding could induce nonrotamericity, and therefore strain in protein–ligand complexes, and that the increased internal energy might play a role in the formation and the release of reaction products.<sup>42</sup> A recent QM/MM study by Torrent et

al.<sup>43</sup> confirmed that the interaction of the active site with the rest of the protein induces strain in the active site.

Interestingly, the backbone dihedrals apparently do not play a role in the stabilization of the planar conformation of the ring. This conjecture was recently suggested by Ho et al. to account for the occurrence of different conformations of the pyrrolidine ring of prolines.<sup>3</sup> However, our work shows that for a certain configuration of the backbone, the ring can, in fact, adopt a number of different conformations in vacuum and in the protein (Table II), and is thus an effect due to specific interactions with the immediate environment. Therefore, the algorithm developed by Ho et al. to predict the conformation of the proline ring based on the information of the backbone of the protein alone may not always apply.

Finally, in the QM/MM calculations, the stabilization of the planar conformation with respect to the up and down puckers increased with the size of the QM subsystem (Fig. 4), a trend that was already visible with only 60 atoms in the QM subsystem (QM system 3, Fig. 4). Also, describing only the proline side chain and backbone at the QM level did not show any stabilization. In contrast, removing the torsion potentials of the ring altogether in the fully MM minimizations, yielded a perfectly planar proline ring. These three observations suggest that the force fields we used here (AMBER99, GROMOS96, and OPLS) fail to describe the balance between the internal interactions of the ring (torsions and angles) and the short-range interactions with the local protein environment and thus fail to capture the apparent strain that exists in the proline pocket where the side chains are packed densely together. Improving the force fields to better model the proline under such conditions as in TIM, therefore, should not be restricted to finding better torsion potentials of the pyrrolidine ring, but should also involve the short-range non-bonded interactions. Because the structural data that were originally employed for the parameterization of the force field did not include such strained structural entities, it is no surprise that the force field fails to model the proline pocket appropriately.

Along similar lines, active sites are generally not very well described by the current force fields. To accelerate the reaction of the substrate, the residues in the active site have been shown in more cases to be oriented somewhat unfavorably from a structural point of view, as to induce strain.<sup>41–43</sup> It might therefore prove essential to use a QM/MM approach to refine X-ray structures in which strain is suspected. The use of ab initio calculations has been already shown to be an important tool in the refinement of atomic resolution X-ray structures.<sup>44,45</sup>

While computing the energy of the different proline conformations the environment was not allowed to relax. As a consequence, the energies reported here are overestimations, especially those of the up and down puckering. The latter two conformations were forced into the TIM structure, and additional relaxation of the nearby residues would lower the steric repulsion. However, because the proline is planar due to the packing, forcing the ring up or down and subsequently performing an all atom energy

minimization would result in an unphysical overall protein structure. In fact, as we mentioned already, the proline is puckered downwards in the open TIM structure, and keeping the proline down during minimization should eventually lead to this open structures, provided the barriers are not too high to be overcome in the minimization. In practice, reaching the open conformation is not feasible due to the large number of atoms and the quality of the minimization algorithm. Thus, the stabilization of the planar proline in reality will be somewhat smaller than we calculated. However, considering that the diffraction pattern of 1N55 was taken at 100 K, only a small stabilization of about 4 kJ mol<sup>-1</sup> is already sufficient to find the planar conformation almost exclusively populated (>99%). At higher temperatures, we therefore expect to find a superposition of the down, planar, and up ring conformations.

## CONCLUSIONS

In proteins, the conformation of the proline side chain ring is not always restricted to the up or down puckers. Depending on the specific interactions with the environment inside the protein, the ring can adopt almost any conformation and even be completely planar. To experimentally observe such conformations in protein structures with sufficient reliability, the resolution of the X-ray structure needs to be very high (<1 Å). The current force fields for structure refinement tend to overestimate the stability of the two preferred puckers. As a consequence, the occurrence of such conformations could be severely underestimated in lower resolution protein structures, and this could lead to misinterpretation of the functional importance of such strained conformations. In this article, we have demonstrated that by using a QM/MM strategy, it is not only possible to reproduce the correct planar conformation of proline 168 in TIM, but also to identify the interactions responsible for the planarity. The short-range steric repulsive interaction of Pro168 with the side chain of Tyr166 on one side of the proline ring, prohibit the up pucker, while the side chain of Ala171 prevents the ring from puckering down.

## ACKNOWLEDGMENTS

The authors wish to thank Marco Casteleijn and Olli Niemitalo from the University of Oulu and Alessandra Villa from the University of Frankfurt for useful discussions. We acknowledge the MGrid Cluster Project of the University of Oulu, Finland, for computing power.

## REFERENCES

- MacArthur MW, Thornton JM. Influence of proline residues on protein conformations. *J Mol Biol* 1991;218:397–412.
- Madison V. Flexibility of pyrrolidine ring in proline peptides. *Biopolymers* 1977;16:2671–2692.
- Ho BK, Coutasias EA, Seok C, Dill KA. The flexibility in the proline ring couples to the protein backbone. *Protein Sci* 2005;14:1011–1018.
- Ramachandran GN, Lakshminarayanan AV, Balasubramanian R, Tegner G. Studies on the conformation of amino acids, XII: energy calculations on propyl residue. *Biochim Biophys Acta* 1970;221:165–181.
- DeTar DF, Luthra NP. Conformations of proline. *J Am Chem Soc* 1977;99:1232–1244.
- Nemethy G, Gibson KD, Palmer KA, Yoon CN, Paterlini G, Zagari A, Rumsey S, Scheraga HA. Energy Parameters in Polypeptides. 10. Improved geometric parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides. *J Phys Chem* 1992;96:6472–6484.
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comp Chem* 1983;4:187–217.
- van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG. Biomolecular simulation: GROMOS96 manual and user guide. Zürich, Groningen: BIOMOS b.v.; 1996.
- Rose IA. Mechanism of C—H bond cleavage in aldolase and isomerase reactions. *Brookhaven Symp Biol* 1962;15:293–309.
- Knowles JR. Enzyme catalysis: not different, just better. *Nature* 1991;350:121–124.
- Jorgensen WL, Maxwell S, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 1996;118:11225–11236.
- Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comp Chem* 2004;25:1157–1174.
- Sorin EJ, Pande VS. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys J* 2005;88:2472–2493.
- Warshel A, Levitt M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 1976;103:227–249.
- Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J Mol Model* 2001;7:306–317.
- van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark A, Berendsen H. Gromacs: fast, flexible and free. *J Comp Chem* 2005;26:1701–1718.
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian 03, Revision C.02. Wallingford, CT: Gaussian Inc.; 2004.
- Groenhof G, Bouxin-Cademartory M, Hess B, de Visser S, Berendsen H, Olivucci M, Mark A, Robb M. Photoactivation of the photoactive yellow protein: why photon absorption triggers a *trans-to-cis* isomerization of the chromophore in the protein. *J Am Chem Soc* 2004;126:4228–4233.
- Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, editor. *Intermolecular forces*. Dordrecht: Reidel; 1981. p 331–342.
- Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. *J Comp Chem* 1997;18:1463–1472.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov I, Bourne P. The Protein Data Bank. *Nucleic Acids Res* 2000;28:235–242.
- Kursula I, Wierenga RK. Crystal structure of triosephosphate isomerase complexed with 2-phosphoglycolate at 0.83-Å resolution. *J Biol Chem* 2003;278:9544–9551.
- Jogl G, Rozovsky S, McDermott AE, Tong L. Optimal alignment for enzymatic proton transfer: structure of the Michaelis complex of triosephosphate isomerase at 1.2-Å resolution. *Proc Natl Acad Sci USA* 2003;100:50–55.

24. Wierenga RK, Noble MEM, Vriend G, Nauche S, Hol WGJ. Refined 1.83 Angstroms structure of trypanosomal triosephosphate isomerase, crystallized in the presence of 2.4 M-ammonium sulphate. A comparison with the structure of the trypanosomal triosephosphate isomerase-glycerol-3-phosphate complex. *J Mol Biol* 1991;220:995–1015.
25. Wolfenden R. Transition state analogues for enzyme catalysis. *Nature* 1969;223:704–705.
26. Davenport RC, Bash PA, Seaton BA, Karplus M, Petsko GA, Ringe D. Structure of triosephosphate isomerase–phosphoglycolohydroxamate complex: an analogue of the intermediate on the reaction pathway. *Biochemistry* 1991;30:5821–5826.
27. Wierenga RK, Noble ME, Davenport RC. Comparison of the refined crystal structures of liganded and unliganded chicken, yeast and trypanosomal triosephosphate isomerase. *Biochemistry* 1992;38:4389–4397.
28. Schaftenaar G, Noordik JH. Molden: a pre- and post-processing program for molecular and electronic structures. *J Comp-Aided Mol Des* 2000;14:123–134.
29. Peng C, Schlegel H. Combining synchronous transit and Quasi-Newton methods for finding transition state. *Isr J Chem* 1994;33:449–454.
30. Miertus S, Scrocco E, Tomasi J. Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for a prevision of solvent effects. *Chem Phys* 1981;55:117–129.
31. Mennucci B, Tomasi J. Continuum solvation models. A new approach to the problem of solute's distribution and cavity boundaries. *J Chem Phys* 1997;106:5151–5158.
32. [http://www.gromacs.org/topologies/force\\_fields.php#ffG43a1p](http://www.gromacs.org/topologies/force_fields.php#ffG43a1p): uploaded May 15, 2002 by Graham Smith.
33. Hansson T, Nordlund P, Åqvist JA. Energetics and nucleophile activation in a protein tyrosine phosphatase. *J Mol Biol* 1997;265:118–127.
34. Miyamoto S, Kollman PA. SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J Comp Chem* 1992;13:952–962.
35. Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A. Molecular dynamics with coupling to an external bath. *J Chem Phys* 1984;81:3684–3690.
36. Lovell SC, Davis IW, III WBA, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by  $\alpha$  geometry:  $\varphi$ ,  $\psi$  and  $C\beta$  deviation. *Proteins* 2003;50:437–450.
37. Deb K, Anand A, Joshi D. A computationally efficient evolutionary algorithm for real-parameter optimization. In KanGAL, editor. Report no. 2002003. Tsutsui Kanpur, India: Indian Institute of Technology Kanpur; 2002.
38. Maes D, Zeelen JP, Thanki N, Beaucamp N, Alvarez M, Thi MH, Backmann J, Martial JA, Wyns L, Jaenicke R, Wierenga RK. The crystal structure of triosephosphate isomerase (TIM) from *Thermotoga maritima*: a comparative thermostability structural analysis of ten different TIM structures. *Proteins* 1999;37:441–453.
39. Kang YK, Choi HY. *cis-trans* isomerization and puckering of proline residues. *Biophys Chem* 2004;111:135–142.
40. Kursula I, Salin M, Sun J, Norledge BV, Haapalainen AM, Sampson NS, Wierenga RK. Understanding protein lids: structural analysis of active hinge mutants in triosephosphate isomerase. *PEDS* 2004;17:375–382.
41. Herzberg O, Moult J. Analysis of the steric strain in the polypeptide backbone of protein molecules. *Proteins* 1991;11:223–229.
42. Heringa J, Argos P. Strain in protein structures as viewed through nonrotameric side chains: II. Effects upon ligand binding. *Proteins* 1999;37:44–55.
43. Torrent M, Vreven T, Musaev DG, Morokuma K. Effects of the protein environment on the structure and energetics of active sites of metalloenzymes. ONIOM study of methane monooxygenase and ribonuclease reductase. *J Am Chem Soc* 2002;124:192–193.
44. Schmidt A, Jelsch C, Ostergaard P, Rypniewski W, Lamzin VS. Trypsin revisited: crystallography at (sub) atomic resolution and quantum chemistry revealing details of catalysis. *J Biol Chem* 2003;278:43357–43362.
45. Schmidt A, Lamzin VS. Veni, vidi, vici—atomic resolution unravelling the mysteries of protein function. *Curr Opin Struct Biol* 2002;12:698–703.
46. Kraulis P. Molscript—a program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 1991;24:946–950.