The Loop Opening/Closing Motion of the Enzyme Triosephosphate Isomerase

Philippe Derreumaux* and Tamar Schlick#§

*Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Paris 75005, France; #Chemistry Department and Courant Institute of Mathematical Sciences, New York University, New York, New York 10012 USA; and §the Howard Hughes Medical Institute

ABSTRACT To explore the origin of the large-scale motion of triosephosphate isomerase's flexible loop (residues 166 to 176) at the active site, several simulation protocols are employed both for the free enzyme in vacuo and for the free enzyme with some solvent modeling: high-temperature Langevin dynamics simulations, sampling by a "dynamics driver" approach, and potential-energy surface calculations. Our focus is on obtaining the energy barrier to the enzyme's motion and establishing the nature of the loop movement. Previous calculations did not determine this energy barrier and the effect of solvent on the barrier. High-temperature molecular dynamics simulations and crystallographic studies have suggested a rigid-body motion with two hinges located at both ends of the loop; Brownian dynamics simulations at room temperature pointed to a very flexible behavior. The present simulations and analyses reveal that although solute/solvent hydrogen bonds play a crucial role in lowering the energy along the pathway, there still remains a high activation barrier. This finding clearly indicates that, if the loop opens and closes in the absence of a substrate at standard conditions (e.g., room temperature, appropriate concentration of isomerase), the time scale for transition is not in the nanosecond but rather the microsecond range. Our results also indicate that in the context of spontaneous opening in the free enzyme, the motion is of rigid-body type and that the specific interaction between residues Ala^{176} and Tyr^{208} plays a crucial role in the loop opening/closing mechanism.

INTRODUCTION

Triosephosphate isomerase (TIM) is a dimeric enzyme, with each subunit composed of 247 residues. The protein catalyzes the interconversion of dihydroxyacetone phosphate (DHAP) into D-glyceraldehyde 3-phosphate (GAP), the central reaction in the glycolytic pathway. A complex protein machinery has naturally evolved for optimal catalysis of this important reaction; even small mutations can lead to severe genetic diseases, ranging from hemolytic disorder to neuromuscular impairment.

A variety of experimental techniques have been used to elucidate the TIM-catalyzed isomerization process. These structural and spectroscopic studies have established the following facts. 1) "Natural" TIM (i.e., not manipulated by protein engineering; Borchert et al., 1994) is active only as a dimer. 2) The structures of the two subunits are nearly identical, except at the two amino-terminal ends (Banner et al., 1975; Lolis et al., 1990). 3) Upon substrate binding, an 11-residue loop (residues 166–176) moves ~7 Å as a rigid "lid" to close over the active site, and the carboxylate group of the catalytic base Glu^{165} moves ~2 Å toward the inhibitor from its position in the unbound enzyme. This motion has been deduced for many enzyme sources (e.g., chicken, Banner et al., 1975; yeast, Lolis et al., 1990; tripanosome, Wierenga et al., 1991), as well as for various substrate analogs, including phosphoglycolohydroxamate (PGH) (Zhang et al., 1994), 2-phosphoglycolate (PGA) (Lolis and Petsko, 1990), and 3-chloroacetol phosphate (CAP) (Yuksel et al., 1994).

Although the current hypothesis for the catalytic mechanism of TIM involves acid/base chemistry in which the Glu^{165} and His^{95} residues participate directly in proton transfer, other residues in the active site have been implicated in the catalysis (Bash et al., 1991; Lodi et al., 1994; Joseph-McCarthy et al., 1994; Sampson and Knowles, 1992a; Neria and Karplus, 1997). The following residues have been identified as key players.

1. Lys^{12}. This residue provides the positive charge required for substrate binding (Lodi et al., 1994). In the K12M structure, where Lys has been replaced by Met, the inhibitor PGH does not bind to the mutant isomerase, and the 11-residue loop always remains at the open state (Joseph-McCarthy et al., 1994).

2. Tyr^{208}. The hydroxyl of this residue forms a hydrogen bond with the amide nitrogen of Ala^{176} and is essential for proper closure of the active-site loop. Replacement of Tyr by Phe results in a 2400-fold decrease in catalytic activity (Sampson and Knowles, 1992a).

3. The 11-residue loop. The closed loop conformation stabilizes the charged intermediate (Joseph-McCarthy et al., 1994), protects the active site from contact with bulk water (Pompliano et al., 1990), and may prevent the phosphate elimination reaction (Lolis and Petsko, 1990). See Fig. 1 for an illustration of the open- and closed-loop forms, with key residues labeled.

Received for publication 11 July 1997 and in final form 23 September 1997.

Address reprint requests to Dr. Tamar Schlick, Chemistry Department and Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012. Tel.: 212-998-3116; Fax: 212-995-4152; E-mail: schlick@nyu.edu.

0006-3495/98/01/72/10 $2.00

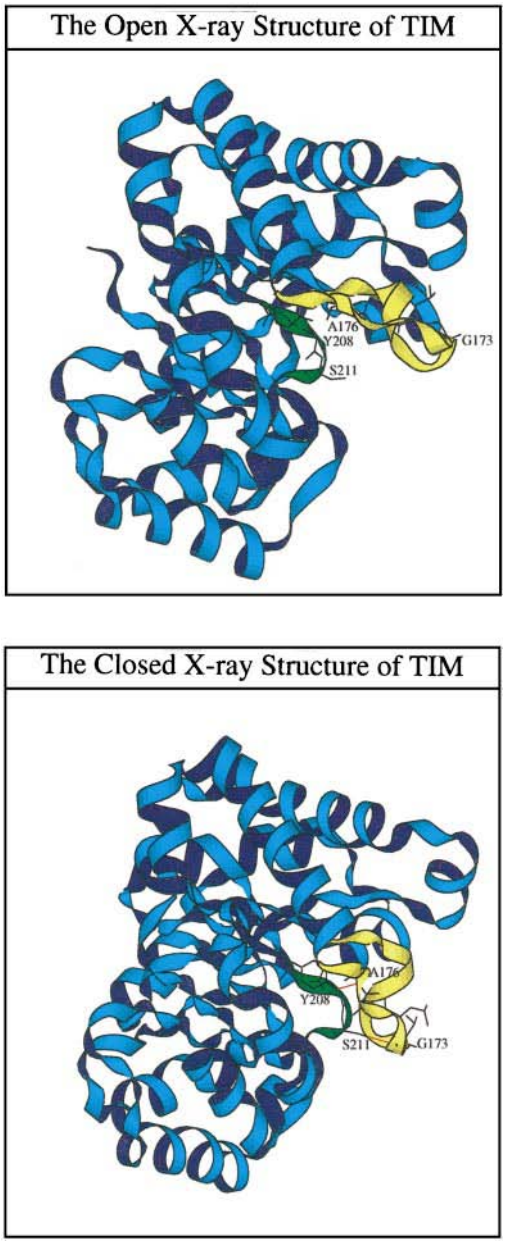

FIGURE 1 Schematic ribbon diagrams of the open (*top*) and closed (*bottom*) x-ray structures of subunit 1, created with MOLSCRIPT (Kraulis, 1991). The surface of loop residues is in yellow, that of residues 208–212 in green, and those of the others in blue. Residues Gly^{173}, Ala^{176}, Tyr^{208}, and Ser^{211} are identified, as well as the hydrogen bonds (*in red*) between Ser^{211} and Gly^{173}, and between Tyr^{208} and Ala^{176} in the closed form.

Clearly, loop closure is essential for the catalytic mechanism of TIM. However, the forces that drive this conformational change are unclear and pose natural questions for simulation work. To date, three scenarios have been proposed to explain the conformational transition of TIM. The first supports the hypothesis that the central part of the loop initiates the movement by long-range electrostatic interactions with the bound substrate (Brown and Kollman, 1987; Zhang et al., 1994). The second, based on studies with different viscogenic agents, proposes that either loop closure and deprotonation of the substrate by the carboxylate group of Glu^{165} are coupled (i.e., the proton is transferred as the loop closes), or these two processes happen sequentially (Sampson et al., 1992a). The third scenario is based on solid-state deuterium NMR of TIM (with and without substrate) and points to a loop closure that is not ligand gated, but rather a natural motion of the protein; it further suggests that the loop opens and closes on the 100-μs time scale (Williams and McDermott, 1995).

A goal of this study is to determine the energetics along the transitional pathway in TIM without the substrate. Understanding the nature of the energy barrier in free TIM should provide insights into the mechanism for loop motion. It is also of interest to establish the nature of the loop movement: does it move as a rigid body or as a flexible chain? Simulation work to date has not resolved this issue. From molecular dynamics (MD) simulations at 1000 K (Joseph et al., 1990) and crystallographic studies (Lolis and Petsko, 1990; Zhang et al., 1994), essentially a rigid-body motion with two hinges located at both ends of the loop has been suggested. However, Brownian dynamics simulations at room temperature point to a more flexible behavior, with dihedral angles spanning the full 360° range (Wade et al., 1993, 1994). Is one type of dynamic behavior correct, or does the TIM mechanism involve both? Details of the two simulations above are also quite different and might affect the results.

To gain insight into the loop motion of the first subunit, we perform high-temperature Langevin dynamics (LD) simulations at 1200 K, as well as simulations by the "dynamics driver" sampling approach (DA) for free TIM in vacuo, and for free TIM with some solvent account. The DA approach, originally developed on polypeptides, has been shown to be a useful tool for enhanced configurational sampling and exploring unfolding processes (Derreumaux and Schlick, 1995; see details below).

In the next section we detail the simulation protocol. To make computations tractable, we limit the number of residues that are free to move during the dynamics. (This protocol is used after energy minimization of the all-atom model). This approximation is reasonable, given the conformational similarity between the open and closed structures (see Fig. 1). Moreover, we found that allowing flexibility in more residues does not change our conclusion. The dynamics protocol consists of 20 ps of equilibration at the desired temperature with a standard explicit Langevin integrator, followed by a production run with LD or DA.

In the third section we detail our results. All of the simulations begin with the closed form. The first set of simulations evaluates the intrinsic dynamics of the loop in vacuo. We find that during the DA and LD simulations the loop opens and closes repeatedly. However, the in vacuo potential surface leads to an unphysical energy profile that would prevent the loop from opening and closing at room temperature. The second set of simulations focuses on the influence of solvent on the energetics and on loop motion. We find that the solvent/solute hydrogen bonds substan-

tially lower the barrier to transition; the loop hardly opens in the LD simulations, but opens completely in the DA simulation through a flexible motion. Questions concerning the flexibility of the chain and the nature of the forces driving the loop motion are addressed in the last section.

SIMULATION PROTOCOL

All energy minimization, MD, LD, and DA calculations were performed with the CHARMM program (Brooks et al., 1983). Standard PARAM19 charges and force constants were used for TIM. Electrostatic interactions were smoothly turned off at 12 Å with a shifting function, and van der Waals interactions were turned off at 9 Å with a switching function. A constant dielectric function ($\epsilon = 1$) and a distance-dependent dielectric function ($\epsilon = r$, where r is an interatomic distance) were used in the solvent and vacuum simulations, respectively. The nonbonded list was updated by a heuristic procedure (i.e., updates were performed only if atoms move by more than half of the width of the buffer region, 0.5 Å, since the last update). The minimizer used for all of the calculations was the adapted basis Newton-Raphson (ABNR). The TNPACK minimizer we have previously used (Derreumaux et al., 1994) is feasible for this size of system with the sparse Hessian storage form in the latest CHARMM version, as used by Barth and Schlick (manuscript submitted for publication).

Atomic models

The initial structure for TIM was taken from the x-ray complex of TIM-PGA, where the loop is in the closed form (Lolis and Petsko, 1990). We then added nonaliphatic hydrogens with the CHARMM routine Hbuild (Brünger and Karplus, 1988). For the TIM/solvent simulation, the water molecules were taken from a pure water simulation and centered about the hydroxyl group of Thr^{172}. All water molecules that were less than 2.2 Å or more than 3.2 Å away from any atoms in the 169–176 loop and in the region of Tyr^{208} and Ser^{211} were deleted. Therefore, only the first shell of hydration around the region of interest was included. Although very approximate, this modest solvent representation of 22 water molecules (modeled by the TIP3P potential) permits examination of the qualitative effects of solvent on the loop motion for a small increment in CPU time. We chose Thr^{172} for the centering because it is the most hydrophilic residue in the loop. The identities of residues 164–178 are Tyr-Glu-Pro-Val-Trp-Ala-Ile-Gly-Thr-Gly-Leu-Ala-Ala-Thr-Pro.

The all-atom models were then minimized by 5000 steps of ABNR with decreasing harmonic restraints. Many residues of the protein were then held at their minimized x-ray positions for dynamics production. Specifically, we allowed residues 168–176, 210, and 211 to move ("set 1" below). This represents a total of 79 free protein atoms. Together with the water atoms, we have 145 atoms for the TIM/solvent simulation.

Our set of free residues differs from that used by Joseph et al. (1990) in two respects. First, Joseph et al. allowed residue Glu^{129} to be flexible, but not residue Gly^{210}; we fix Glu^{129} and let Gly^{210} move. We believe that Glu^{129} does not play a critical role in the loop motion, although the hydrogen bond between Glu^{129} and Trp^{168} is important for stabilization of the closed form. The position of Glu^{129} is very similar in the open and closed structures: the (ϕ, ψ) values are (−80°, 150°) in the closed state versus (−81°, 179°) in the open state; the side-chain dihedral angle χ_1 is ~70° in both states, whereas the dihedral angle χ_2 is −130° in the close state and −170° in the open state. In contrast, as a result of the motion loop, the (ϕ, ψ) values of Gly^{210} change from (−91°,−155°) in the open state to (128°, 76°) in the closed form.

Second, Joseph et al. (1990) allowed residues 164–178 to move, whereas we only allowed residues 168–176 to move freely. Our set was established from the interactions between the loop and residue Ser^{211} in both the complexed and uncomplexed TIM structures. Ser^{211} is located at the entrance of the active site on the opposite side from the loop. As seen from the x-ray data given in Table 1, residues 165–167 and 177 do not move substantially from the β-carbon atom of Ser^{211}. This property is also reflected by the small changes in the α-carbon pseudodihedral angles 164–165 (20°) and 165–166 (3°) in going from the closed to the open crystal structures (Joseph et al., 1990). The pseudodihedral angle i to $i + 1$ is the dihedral angle defined by the four α-carbon atoms of residues $i - 1$, i, $i + 1$, and $i + 2$.

However, because the carboxylate group of the catalytic base Glu^{165} moves ~2 Å upon substrate binding, we also examined high-temperature LD results with three sets of

TABLE 1 Distances (in Å) between the β-carbon atoms* of the loop and the Ser^{211} β-carbon in the open and closed X-ray structures

Residue	Open form	Closed form	Diff#
Glu^{165}	10.4	9.9	0.5
Pro^{166}	8.6	9.3	−0.7
Val^{167}	11.7	11.7	0.0
Trp^{168}	12.2	9.9	2.3
Ala^{169}	7.4	4.9	2.5
Ile^{170}	7.2	7.4	−0.2
Gly^{171}	9.7	5.4	4.3
Thr^{172}	10.8	6.3	4.5
Gly^{173}	11.2	4.5	6.7
Leu^{174}	11.2	6.1	5.1
Ala^{175}	8.6	4.3	4.3
Ala^{176}	8.0	7.7	0.3
Thr^{177}	11.4	11.8	−0.4

* In the case of glycine residues, it is the α-carbon atom instead of the β-carbon.

Each distance is calculated as the magnitude of the difference between the x-ray open and closed forms. As seen from the x-ray data, residues 165–167, 176, and 177 do not move much from the open to closed forms. They will thus be omitted in the upcoming tables.

flexible residues to guarantee that the dynamical properties of the loop are insensitive to the protocol used for chosen free residues. Recall that set 1 allowed flexibility in 11 residues: 168–176, 210, and 211. Set 2 added six residues to these, allowing flexibility in groups 164–178, 210, and 211. This set is similar to what was used by Joseph et al. (1990), as described above. Set 3 included the greatest number of flexible residues (55 in all): 127–130, 153–189, and 203–216; this set includes 13 residues (from each end) beyond the loop region (166–176) and five residues beyond the 208–211 region (which includes the two residues that can form hydrogen bonds with loop residues). We found our overall conclusions regarding the loop motion to be essentially the same in all protocols (see below).

Dynamics protocol

In practice, the residues are kept rigid (same coordinates) by the CONS FIX command in CHARMM. This command bypasses energy calculations involving only fixed atoms. The calculation consists of two phases. First, 20 ps of heating and equilibration at the desired temperature are performed with the Verlet-type discretization scheme of the Langevin equation known as BBK (Brünger et al., 1982), at a time step of 1 fs. Second, production is accomplished by either the Langevin dynamics or the DA procedure at a time step of 1 fs and 40 fs, respectively (see below).

The updating formulas for the positions and velocities for BBK are given by

$$\mathbf{X}^{n+1} = \mathbf{X}^n + \mathbf{V}^n \Delta t \left(\frac{1 - (1/2)\gamma \Delta t}{1 + (1/2)\gamma \Delta t} \right) + \left(\frac{\Delta t^2}{1 + (1/2)\gamma \Delta t} \right) \mathbf{M}^{-1} (\mathbf{R}^n - \mathbf{g}_E(\mathbf{X}^n)) \quad (1a)$$

$$(\mathbf{X}^{n+1} - \mathbf{X}^n)/\Delta t = \mathbf{V}^{n+1} \quad (1b)$$

In these equations, $\mathbf{X}$ and $\mathbf{V}$ are the collective coordinate and velocity vectors of the molecular system, respectively; γ is a collision frequency parameter; Δt is the time step; $\mathbf{M}$ is the diagonal mass matrix; $\mathbf{g}_E(\mathbf{X})$ is the gradient vector of the potential energy E; and the vector superscripts n refer to the approximate solutions at time $n\Delta t$. The random force is discretized according to the formulas

$$\langle \mathbf{R}^n \rangle = 0, \qquad \langle \mathbf{R}^n (\mathbf{R}^m)^T \rangle = 2\gamma k_B T \mathbf{M} (\delta_{nm}/\Delta t) \quad (2)$$

where δ is the Dirac symbol, T is the input temperature, and k_B is Boltzmann's constant. The collision frequency of all atoms, γ, is set to 50 ps^{-1}. This is a typical value for protein atoms exposed to solvent with a viscosity of 1 cp at room temperature (Pastor, 1994).

The DA approach involves minimization of a "dynamics" function, $\Phi(\mathbf{X})$, at each time step (Derreumaux and Schlick, 1995). The basic idea in DA is to counteract numerical damping of the high-frequency modes produced by the implicit integrator LI (Schlick et al., 1991) and to account in some way for the subdynamics that has taken place during the large 40-fs time step.

In DA, the new coordinate $\mathbf{X}^{n+1}$ is obtained as a minimum of $\Phi(\mathbf{X})$, where

$$\Phi(\mathbf{X}) = \tfrac{1}{2}(1 + \gamma \Delta t)(\mathbf{X} - \mathbf{X}_0^n)^T \mathbf{M} (\mathbf{X} - \mathbf{X}_0^n) + (\Delta t)^2 E(\mathbf{X}) \quad (2a)$$

$$\mathbf{X}_0^n = \mathbf{X}^n + [\Delta t/(1 + \gamma \Delta t)](\mathbf{V}^n + \Delta t \mathbf{M}^{-1} \mathbf{R}^{n+1}) \quad (2b)$$

The initial guess used for Φ minimization is

$$\mathbf{X} = \mathbf{X}^n + \mathbf{P} \quad (2c)$$

where the perturbation vector $\mathbf{P}$, of magnitude λ, is defined by components

$$\mathbf{P} = \lambda \sin\theta \cos\phi, \lambda \sin\theta \sin\phi, \lambda \cos\theta \quad (2d)$$

The angles θ and ϕ are randomly chosen in the interval $[-\pi, \pi]$.

Each step of DA consists of

1. Specification of a perturbation vector $\mathbf{P}$ to $\mathbf{X}^n$ (the position vector of the previous step) and rescaling of the initial velocity vector $\mathbf{V}^n$ by a factor of 3 (the stepsize λ is set to 0.35 Å)
2. Minimization of $\Phi(\mathbf{X})$ so that the final gradient norm is less than 0.5 kcal/(mol Å)
3. Application of two acceptance criteria involving both the energetic components associated with the high-frequency modes and the total kinetic energy. That is, we require the internal energy of bond lengths, bond angles, and improper angles to remain less than an energy threshold defined by 2/3 ($k_B T$) per internal degree of freedom. The bath (input) temperature is set to 300 K in all of the present calculations. We also enforce the DA kinetic energy to be less than half of the time-averaged kinetic energy of a MD simulation. In practice, the tentative move is rejected if one energetic component is larger than the associated energy threshold or if the kinetic energy condition is not verified; in such a case, the initial candidate $\mathbf{X}^n$ is used as a starting point for Φ minimization.
4. Determination of the velocity vector from Eq. 1b. For full details, see Derreumaux and Schlick (1995).

In this work, we found it necessary to modify the original DA algorithm by the second part of step 1 above, namely velocity rescaling. This is because loop opening/closing is a slow process, involving a high activation energy barrier. Indeed, starting from the closed form, the loop never completely opened throughout the 2-ns DA simulation when the original algorithm was used, as well as with a 400-ps LD simulation at 800 K. Two other minor modifications should also be noted. The final accepted norm in step 2 is larger than the value of 10^{-2} kcal/(mol Å) used earlier; this relaxation saves computer time and gives very similar final minimized structure. The final accepted DA kinetic energy in step 3 is smaller than the value used earlier (6/5 of the expected kinetic energy); this lowers the effective simulation temperature.

DA properties

Our DA scheme is an enhanced sampling tool and not a dynamics method per se. First, the DA time step has no physical significance. Two successive conformations in the DA trajectory may resemble each other, and even have similar energies, but are not necessarily dynamically connected within 40 fs. This results from the arbitrary dynamic-based moves (via the application of random perturbation vector **P**; see Eq. 2c) to counteract quenching of the high-frequency modes. Second, the two acceptance criteria (step 3 above) are more lenient than the combination of Metropolis (Metropolis et al., 1953) and kinetic-energy criteria in Monte Carlo methods, or the total energy conservation in a Newtonian dynamics simulation. Our previous study on polypeptides showed that enforcement of the Metropolis and the kinetic energy criteria requires at least one order of magnitude more steps than enforcement of our two acceptance criteria. As a consequence, the energy fluctuations do not satisfy the Boltzmann distribution, and the time for crossing an energy barrier ΔE does not scale with $\exp(-\Delta E/k_B T)$ as it does in MD or Monte Carlo simulations. In fact, DA trajectories reveal large fluctuations in the potential energy because inexact instantaneous bond angles lead to a small increase in the torsional energy and a substantial increase in the van der Waals and electrostatic potentials. However, DA does not sample configurational space randomly, because of a dynamics-based component (first term of the dynamics function in Eq. 2a), and, moreover, generates Boltzmann-like distributions due to the two acceptance criteria. This makes it suitable for applications involving specific large-scale configurational moves. For more details, see Derreumaux and Schlick (1995).

RESULTS

The equilibrated structures for all models compare well to the observed closed structure. The two internal hydrogen bonds characteristic of the closed form (between the Ser hydroxyl oxygen of residue 211 and the amide nitrogen of Gly^{173} (hb1), and between the Tyr hydroxyl oxygen of residue 208 and the amide nitrogen of Ala^{176} (hb2)) are established. The dihedral angles in the nine-residue (168–176) and two-residue (210–211) loops deviate by less than 25° from their crystallographic positions. The distances between the main-chain atoms of the loop and the Ser^{211} β-carbon are in excellent agreement with the experimental values.

All simulations presented now use set 1 of flexible residues unless specified. We analyze the behavior during the simulations of the two models by the time histories of the following quantities: 1) distances between the β-carbon atoms of the loop (or, in the case of glycine residues, the α-carbon atom) and the Ser^{211} β-carbon (see Table 1); 2) characteristics of two hydrogen bonds in the closed form: hb1 and hb2; 3) loop α-carbon pseudodihedral angles. Recall that the pseudodihedral angle i to $i + 1$ is the dihedral angle defined by the four α-carbon atoms of residues $i - 1$, i, $i + 1$, and $i + 2$.

Free TIM in vacuo

At 1200 K, the 80-ps LD simulation reveals one opening/closing process starting from the closed form. The time history of the energy reveals a difference in potential energy between the closed and open forms of 40 kcal/mol. Table 2 reports one conformation that most resembles the experimental open form (column LD_{vac}). As seen, the two hydro-

TABLE 2 Distances (in Å) between the β-carbon atoms* of residues 168–175 and the Ser^{211} β-carbon in the open x-ray, in vacuo, and in solvent# LD and DA structures from the closed x-ray structure

Residue	X-ray	LD_{vac}§	DA_{vac}¶	LD‖	LD**	LD##	LD§§	DA¶¶	DA‖‖	DA***
Trp^{168}	2.3	0.4	0.2	0.1	1.4	1.1	0.1	0.3	0.5	0.0
Ala^{169}	2.5	0.6	2.4	0.6	1.5	1.2	0.6	0.5	0.1	0.1
Ile^{170}	−0.2	1.0	4.2	−0.4	−0.4	2.2	0.2	0.3	−1.0	0.1
Gly^{171}	4.3	2.9	6.6	0.8	2.9	−0.6	3.8	6.3	3.7	2.2
Thr^{172}	4.5	6.3	7.3	2.7	3.0	1.9	3.4	9.3	3.9	−2.1
Gly^{173}	6.7	5.6	7.8	3.5	2.3	2.4	2.8	8.5	7.6	4.2
Leu^{174}	5.1	3.4	6.2	2.7	2.2	2.8	2.5	7.0	5.3	5.2
Ala^{175}	4.3	4.1	4.0	0.9	1.2	0.2	1.5	4.9	2.8	2.1

* For glycine residues, the α-carbon atom is used instead. Each distance is calculated as the magnitude of the difference between the open-generated form and the closed x-ray form. The starting (equilibrated) structures of the LD and DA simulations agree very well with the closed x-ray structure.

Subscripts "vac" distinguish the vacuum from the solvent simulations. All LD simulations are performed at 1200 K. Set 1 allows flexibility in residues 168–176, 210, and 211; set 2 frees residues 164–178, 210, and 211; and set 3 increases the set of flexible residues to groups 127–130, 153–189, and 203–216.

§ One of the most open LD structures (at $t = 58$ ps) with set 1, having (hb1, hb2) = (11.2, 4.5) Å.

¶ The most open DA structure with set 1, having (hb1, hb2) = (12.5, 4.9) Å.

‖,** Two representative open LD structures with set 1; LD‖ at 39 ps has (hb1, hb2) = (7.3, 3.7) Å, and LD** at 175 ps has (8.0, 3.2) Å.

A representative open LD structure with set 2 having at 52 ps (hb1, hb2) = (5.9, 3.8) Å.

§§ A representative open LD structure with set 3 having at 37 ps (hb1, hb2) = (6.7, 4.4) Å.

¶¶,‖‖ Two of the most open DA conformations with set 1; DA¶¶ has (hb1, hb2) = (12.9, 5.7) Å, and DA‖‖ has (11.5, 3.9) Å.

*** The intermediate open DA conformation used as a starting point for MD with restraints (see Conclusions) having (hb1, hb2) = (7.1, 4.4) Å.

gen bonds hb1 and hb2 are broken, and deviations from the observed structure are 1.9, 1.9, 1.2, 1.1, and 1.7 Å for Trp^{168}, Ala^{169}, Gly^{171}, Gly^{173}, and Leu^{174}, respectively.

The DA trajectory captures four complete loop opening/closing processes during the 400-ps simulation started from the closed form. The DA potential energy in time depicts variations up to 75 kcal/mol, and no clear picture of the energy barrier to motion emerges. Table 2 reports distances for a representative of the DA-generated open forms (DA_{vac}). In all of the open structures, residues 170–172 tend to extend in vacuo more than they do in the crystal structure, and Trp^{168} is closer to Ser^{211}.

To gain insight into the energetics along interconversion, we also calculated the adiabatic three-dimensional potential energy surface of TIM in the hydrogen bond space involving the two particular interactions hb1 and hb2. This potential energy surface was obtained by a minimization process from the closed x-ray minimized form to the open form by varying the hydrogen bond distances with strong harmonic constraints. (The DA and LD results were not used in this analysis.) Specifically, the hb1 distance was varied between 3.0 and 10.0 Å, and the hb2 distance between 3.0 and 6.0 Å. The open PDB form is characterized by (hb1, hb2) = (9.77, 5.05 Å), and the closed PDB form by (2.7, 2.9 Å). We find that the closed form is more stable than the open state by 25 kcal/mol. (An identical energy difference is obtained by using the open and closed x-ray minimized structures.) The lowest energy barrier from the closed form is ~50 kcal/mol, which is independent of the number of atoms free to move. A similar energy profile is obtained with set 2 of flexible residues.

Clearly, a 25 kcal/mol energy difference between the open and closed states is unrealistic; unless the entropy gain in the open state is very large, the loop would not be expected to open at room temperature. What, then, produces this energy difference? Analysis of the intra- and interloop energies points to a hydrogen bond energy loss of 10 kcal/mol in going from the closed to open forms, associated with the breaking of the two hb1 and hb2 interactions. (Each hydrogen bond energy was calculated from the nonbonded interactions between the corresponding residues.) This suggests that if the open conformation can compensate for the loss of the two intramolecular hydrogen bonds by intermolecular hydrogen bonds (either with solvent or with the second monomer), the energy difference could be lowered to 15 kcal/mol. However, there are no interactions between the open 164–211 region of the first monomer and the second monomer. Results of the next section clearly support the main role of solute-solvent interactions in lowering the energy barrier between the open and closed forms of the enzyme.

TIM/solvent model

To assess the role of hydrogen bonding with solvent in lowering the energy difference, we calculated the three-dimensional adiabatic potential energy surface of the TIM/solvent model in the hydrogen bond space involving the two hb1 and hb2 interactions. The surface depicted in Fig. 2 shows that the closed form is more stable than the open form by 8 kcal/mol, and the lowest activation barrier from the closed form is ~15 kcal/mol. The energy surface also reveals one favorable pathway for loop opening. It involves the breaking of the two hydrogen bonds in turn. First, hb1 breaks and subsequently the hb1 distance increases to 6.9–7.2 Å, with the hb2 distance varying between 3.0 and 3.6 Å. Second, the hb2 bond breaks, followed by progressive arrangements of the hb1 and hb2 distances to their open form values. Thus hydrogen bonds with solvent lower the energy barrier heights substantially. Specifically, as seen in Fig. 3, which shows the closed and open minimized structures in solvent, three water molecules are able to compensate for the loss of the two hb1 and hb2 interactions: one forms a hydrogen bond with the Ser^{211} hydroxyl oxygen, one with the Ala^{176} amide nitrogen, and one with the Thr^{172} hydroxyl oxygen.

An important question is whether protein motion might affect the energy of the open and closed forms. To this end, Newtonian dynamics simulations ($\gamma = 0.0$) of 80 ps duration at 300 K of free TIM in solvent were performed for two distinct open forms and two distinct closed forms. The average energy for both forms was subsequently calculated. The open forms are characterized by (hb1, hb2) = (3.0, 3.0) and (3.6, 3.0), the closed forms by (8.1, 5.1) and (9.5, 5.0),

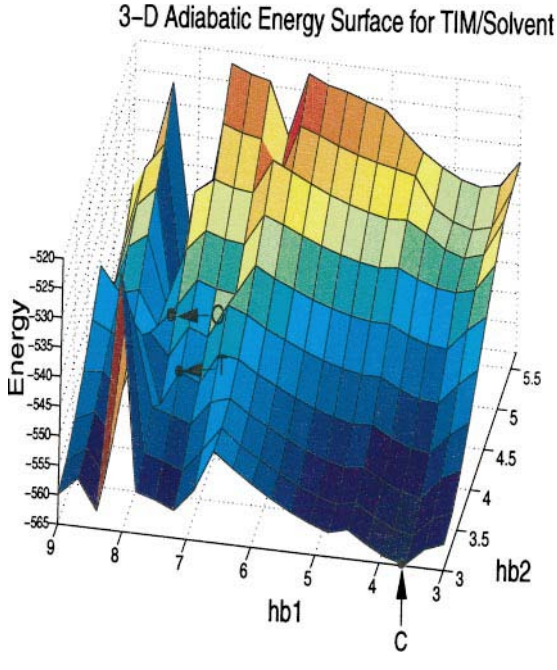

FIGURE 2 Three-dimensional adiabatic potential energy surface for the TIM/solvent model (in kcal/mol) in the space of the two hydrogen bonds: hb1, formed between Ser^{211} and Gly^{173}; and hb2, formed between Tyr^{208} and Ala^{176}. The closed, open, and "transition state" forms are characterized by the following triplets of values (hb1 in Å, hb2 in Å, E in kcal/mol): (3.6, 3.0, −565), (8.1, 5.1, −557), and (7.5, 3.9, −550), respectively. They are identified by C, O, and T.

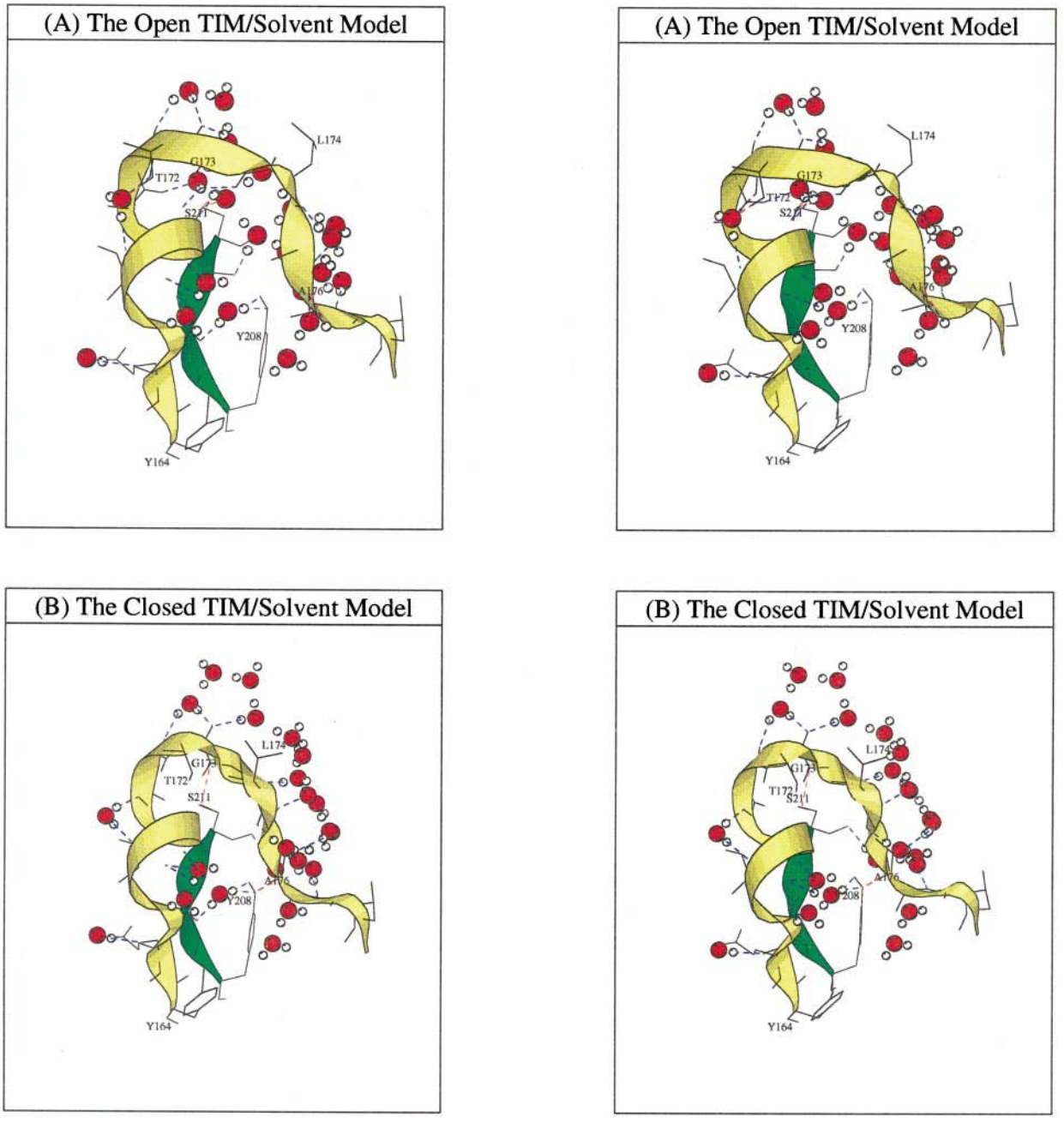

FIGURE 3 The stereo open (*A*) and closed (*B*) minimized forms of subunit 1 in solvent model, drawn with MOLSCRIPT (Kraulis, 1991). These structures are obtained from the adiabatic potential energy surface (Fig. 2). The surface of loop 164–177 residues is in yellow, that of residues 208–212 in green, and the water molecules are in red and white. Residues Tyr^{164}, Thr^{172}, Gly^{173}, Leu^{174}, Ala^{176}, Tyr^{208}, and Ser^{211} are identified. The two hydrogen bonds hb1 and hb2 of the closed form (Fig. 3 *B*) are shown dashed in red; the three specific solvent-solute hydrogen bonds that compensate for the loss of the two hb1 and hb2 interactions of the closed form are shown dashed in red (Fig. 3 *A*), and the other solute-solvent hydrogen bonds in both forms are shown dashed in blue.

all measured in Ångstroms. Over the last 50 ps of the trajectories, we find that protein dynamics reduces the energy gap to 4.5 kcal/mol.

The effects of solvent on the LD and DA trajectories are now examined in turn. We first analyze the 220-ps LD trajectory at 1200 K, starting from the closed form. Table 2 reports distances for two representative open forms (columns $LD^{\|}$ and LD**). The α-carbon atom of Gly^{173} moves in the open direction, but only by ~3 Å (6.7 Å is recorded experimentally); the α-carbon atom of Ala^{175} remains at its closed position because the hydrogen bond hb2 exerts a weak attractive force. We also note that the displacement of Trp^{168} from its closed position is less restricted than in the vacuo calculation, almost reaching the experimental value.

To guarantee that the opening process does not depend on the number of atoms allowed to move, we repeated the 1200

K LD simulation by using sets 2 and 3 of flexible residues instead of set 1 (see Atomic Models under Simulation Protocol). Again, we find that the loop hardly opens. Although the more open structures generated with sets 2 (LD[##]) and 3 (LD[§§]) deviate from those generated with set 1 (LD[‖] and LD**) (see the hb1 and hb2 distances in Table 2), the overall picture does not depend on the protocol used, in the sense that the peptide groups that appear to be static by comparison of the open and closed x-ray structures fluctuate about their observed positions. Furthermore, the rms fluctuations of the pseudodihedral angles 167–175 from the average LD structure generated with sets 2 and 3 are essentially identical to those obtained with set 1. Plots of the open LD minus open x-ray pseudodihedral angles and of the rms fluctuations of the pseudodihedral angles from the average LD structure generated with set 1 in Fig. 4 clearly show that the loop motion is primarily of the rigid-body type, although main-chain dihedral angle transitions occur.

In contrast to the three LD simulations, the 400-ps DA simulation succeeds in capturing one complete loop motion, starting from the closed structure. Information about the energetics along the pathway is obtained by minimizing the DA trajectory with harmonic restraints of 0.5 kcal/(mol Å^2) for positions of the α-carbons of the loop. This minimization procedure thus preserves the overall displacement of the loop from its closed-form position. We find that the potential energy difference of 5–10 kcal/mol between the closed and open forms agrees well with the adiabatic energy estimate; the energy barrier to the motion is ~25 kcal/mol, 10 kcal/mol higher than our adiabatic potential energy estimate.

Two of the open DA conformations are compared to the open x-ray structure in Table 2 (DA[¶¶] and DA[‖‖]). One open conformation extends more in solution than in the crystal, and the other agrees fairly well with the observed structure in terms of the distances between the β-carbon atoms of the loop and the Ser^{211} β-carbon. This agreement, however, is not shared by the values of the pseudodihedral angles, as seen in Fig. 4 *B*.

The plot of the rms fluctuations of the pseudodihedral angles from the average DA structure in Fig. 4 *A* shows that DA results point to a flexible motion distributed over residues 169–175, with two large rms fluctuations about the angles 170–171 and 174–175. A detailed understanding of the loop opening mechanism emerges from the evolution in time of the distances between the β-carbon atoms of the loop and the Ser^{211} β-carbon. The first step in the conformational transition from closed to open involves the motion of Gly^{173}, which breaks hydrogen bond hb1. Then residue 174 comes into play, leading to a conformation where Gly^{173} and Leu^{174} have moved 4 Å from their closed positions, and the other residues are relatively rigid, and hb1 = 7.0 Å and hb2 = 3.3 Å. (Note the analogy between the first two DA steps and the first step of the pathway on the adiabatic energy surface.) Subsequently, Gly^{171} and Thr^{172} reorient, and Ala^{175} moves to break hb2. Finally, small variations in the orientation of residues 171–175 stabilize the loop in the open state.

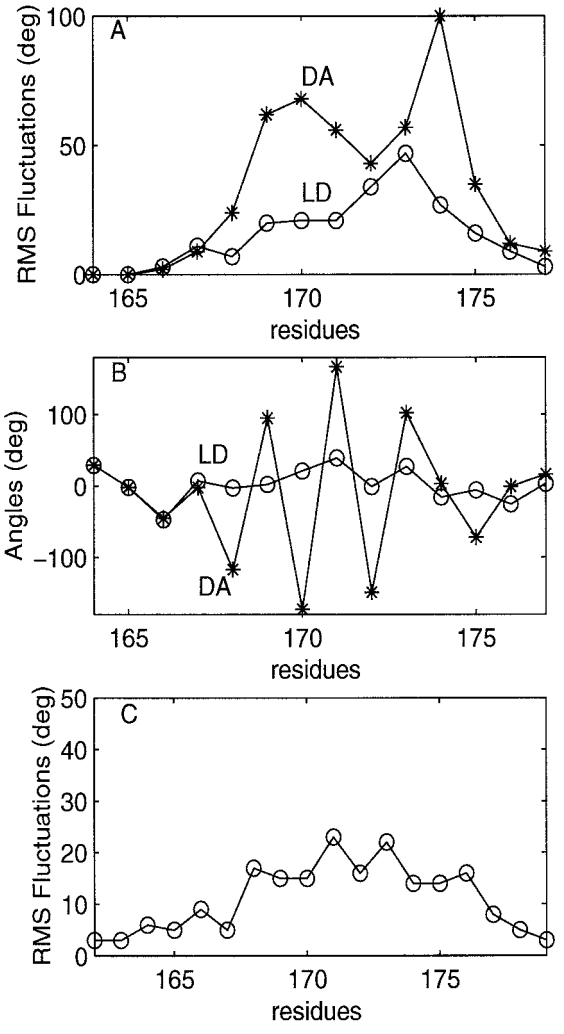

FIGURE 4 Plots of the α-carbon pseudodihedral loop angles. The angle i corresponds to rotation about atoms i, $i + 1$. (*A*) The rms fluctuations of angles from the average LD and DA forms generated with set 1 of flexible residues in solvent simulations. (*B*) The open LD structure (see LD[‖] in Table 2) minus the open crystal structure, and the DA open structure (see DA[¶¶] in Table 2) minus the open crystal structure. (*C*) The rms fluctuations of angles from the average 300 K MD form in solvent simulation with harmonic constraints of 2 kcal/(mol Å^2) on hb1 and hb2.

DISCUSSION AND CONCLUSIONS

The first goal of this work was to obtain an energetic description for the barrier to the loop motion in free TIM to examine the various mechanisms for the loop opening/closing process presented in the literature. Specifically, we questioned the nature of the loop motion: Is loop closure triggered by the interactions with the bound ligand (electrostatic interactions or charge transfer), or is it an internal dynamic property of the enzyme?

Very little is known about the magnitude and the origin of the barrier to the loop motion (Wierenga et al., 1992). The

present calculations of free TIM in vacuo and free TIM with some solvent account indicate that the hydrogen bonds with the solvent, and to a lesser extent protein motion, compensate for the loss of energy within the enzyme in going from the closed to the open forms. However, there still remains a potential energy barrier of 15 kcal/mol from the closed form (30% of the barrier originating from the van der Waals component and 70% from the electrostatic component) and a "free energy" difference of 4 kcal/mol between the closed form (the predominant form) and the open form. In contrast, recent solid-state deuterium NMR work on TIM points to an energy barrier of 12 kcal/mol from the open form and an energy difference of 1.8 kcal/mol between the open form and the closed form, although the NMR data cannot distinguish which loop conformation is more populated (Williams and McDermott, 1995).

Because our hydration model is too simple and entropic effects are either ignored (in the transition state) or too crude (in the open and closed forms), disagreement between our data and NMR-derived data is expected. Certainly a better understanding of the energetics and thus of the true scenario for loop closure might emerge from stochastic boundary simulations of fully hydrated TIM for different values of the hydrogen bonds hb1 and hb2 in the open, closed, and transition forms. However, from the magnitude of the energy barrier, it is evident that, if the loop opens and closes in the free enzyme at standard conditions, the relevant lower bound for the transition time is more than the nanosecond range suggested by Brownian dynamics simulations (Wade et al., 1993), and is more likely in the microsecond time frame. It is possible that the motion of the loop as observed in Brownian dynamics simulations is faster because the model for the potential energy surface did not take into account the two hydrogen bonds that the loop can make with nonloop residues (hb1,hb2).

The second goal of this work was to establish the nature of the loop motion: Does it move as a rigid lid with two hinges, or as a flexible chain? To this end, high-temperature Langevin dynamics (LD) and a "dynamics driver" approach (DA) were applied to the study of loop motion in partial solvent. In an attempt to make computations tractable, we limited the number of flexible residues during the simulations and examined the impact of three sets of flexible residues (containing 11, 17, and 55 free residues) on the LD trajectory. Although the three LD trajectories differ from one another, the overall picture does not depend on the set of flexible residues chosen. The LD simulations suggest a rigid-body motion, although the loop hardly moves. The DA simulation (with set 1), in contrast, captures the complete opening process. A reorientation of the chain is suggested, as in all four in vacuo DA-generated transitions, but we believe this to be an artifact of the DA procedure for two reasons.

First, our lowest energy DA pathway involves an energy barrier of 25 kcal/mol, one that would prevent opening at room temperature. Moreover, the open DA structure deviates substantially from the open crystal structure in terms of pseudodihedral angles, and thus DA with velocity rescaling might produce unrealistic motion for the low-frequency modes.

Second, starting from the closed form, MD simulations of free TIM in solvent with hb1 and hb2 restrained to the open-form configuration converge rapidly (a few picoseconds) to the open form without any pseudodihedral angle transitions. In these 300 K Newtonian dynamics simulations, residues 162–180, 210, and 211 were free to move, and several sets of force constants on the hb1 and hb2 distances were used to explore the dependence of the trajectory on the constraints. Fig. 4 *C* reports the rms fluctuations of the pseudodihedral angles from the average MD conformation. As seen, the motion is of the rigid-body type (rms fluctuations $< 25°$). Interestingly, the lowest energy pathway on the (hb1, hb2) potential energy surface with the same set of residues free to move also involves a rigid-body motion and an energy barrier of $\sim$15 kcal/mol.

Among all of the interactions that would drive loop opening in free enzyme, is it possible to detect one major interaction?

Analysis of the LD and DA trajectories in partial solvent points to the critical role of residues 173, 174, and 175 in initiating loop opening. This agrees with the prediction of Joseph et al. (1990). The interaction between Ser^{211} and Gly^{173} (hb1) is easily disrupted in the initial steps of loop opening in all simulations, as well as in the 300 K LD simulation. This interaction also occurs very easily in the final steps of loop closure (see below). We thus conclude that this hydrogen bond is not essential in the loop opening/closing process, although it is important for stabilization of the closed form and thus for optimal catalysis. This finding is consistent with the small decrease (30-fold) in catalytic activity resulting from the replacement of Ser^{211} by Ala^{-211} (Sampson and Knowles, 1992b).

We do not find any evidence that the region 164–167 plays a major role in the loop opening/closing motion. Results of the three LD simulations with different sets of flexible residues are globally very similar. The possibility that the pseudodihedral angle 166–167 moves $\sim 50°$ from its open to closed-form positions (Joseph et al., 1990) and the carboxylate group of Glu^{165} moves $\sim$2 Å before proton transfer from the substrate to Glu^{165} occurs has already been suggested (Sampson and Knowles, 1992a); this interpretation avoids physical coupling between the loop motion and that of Glu^{165}.

In contrast, the interaction between Tyr^{208} and Ala^{176} (hb2) breaks in the final steps of loop opening in the DA simulation and on the adiabatic potential energy surface. Interestingly, the 300 K MD simulation starting from an intermediate open structure (see structure DA*** in Table 2) shows that loop closure occurs rapidly once the hb2 interaction is formed. Specifically, we find that at 10 ps the loop starts to close, at 59 ps the hb2 interaction forms with the hb1 distance varying between 6 to 7 Å, and once the hb2 interaction is formed, it takes 10 ps for the loop to adopt its closed position. We conclude that this interaction contrib-

utes substantially to the driving force for loop closure. This finding is in accord with the 2000-fold decrease in catalytic activity of the mutant Y208F enzyme (Sampson and Knowles, 1992a).

We are indebted to Martin Karplus for many valuable discussions. We thank Diane Joseph-McCarthy for providing the coordinates for the closed and open forms of TIM; Bernie Brooks at the National Institutes of Health, Division of Computer Research and Technology, for generous computing support; and Wei Xu for the MOLSCRIPT figures.

PD is grateful to TS for financial support, from November 1993 to August 1994 while on a leave of absence from France, and to the French CNRS since November 1994. TS acknowledges support from the National Institutes of Health (National Center for Research Resources, Award RR08102), the National Science Foundation (Presidential Young Investigator Award, and NSF GCAG Award ASC-9318159), and the Alfred P. Sloan Foundation. TS is an investigator of the Howard Hughes Medical Institute.

REFERENCES

Banner, D. W., A. C. Bloomer, G. A. Petsko, D. C. Phillips, C. I. Pogson, I. A. Wilson, P. H. Corran, A. J. Furth, J. D. Milman, R. E. Offord, J. D. Priddle, and S. G. Waley. 1975. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino sequence data. *Nature.* 255:609–614.

Bash, P. A., M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe, and M. Karplus. 1991. Computer simulation and analysis of the reaction pathway of triosephosphate isomerase. *Biochemistry.* 30:5826–5832.

Borchert, V. T., R. Abagyan, R. Jaenicke, and R. K. Wienega. 1994. Design, creation, and characterization of a stable, monomeric triosephosphate isomerase. *Proc. Natl. Acad. Sci. USA.* 91:1515–1518.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Brown, F. K., and P. A. Kollman. 1987. Molecular dynamics simulations of loop closing in the enzyme triose phosphate isomerase. *J. Mol. Biol.* 198:533–546.

Brünger, A. T., C. L. Brooks III, and M. Karplus. 1982. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.* 105:495–500.

Brünger, A. T., and M. Karplus. 1988. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins Struct. Funct. Genet.* 4:148–156.

Derreumaux, P., and T. Schlick. 1995. Long timestep dynamics of peptides by the dynamics driver approach. *Proteins Struct. Funct. Genet.* 21: 282–302.

Derreumaux, P., G. Zhang, B. Brooks, and T. Schlick. 1994. A truncated Newton minimizer adapted for CHARMM and biomolecular applications. *J. Comput. Chem.* 15:532–552.

Joseph, D., G. A. Petsko, and M. Karplus. 1990. Anatomy of a conformational change: hinged lid motion of the triosephosphate isomerase loop. *Science.* 249:1425–1428.

Joseph-McCarthy, D., D. Lolis, E. A. Komives, and G. A. Petsko. 1994. Crystal structure of the K12M/G15A triosephosphate isomerase double mutant and electrostatic analysis of the active site. *Biochemistry.* 33: 2815–2823.

Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.

Lodi, P. J., L. C. Chang, J. R. Knowles, and E. A. Komives. 1994. Triosephosphate isomerase requires a positively charged active site: the role of lysine-12. *Biochemistry.* 33:2809–2814.

Lolis, E., T. Alber, R. C. Davenport, D. Rose, F. C. Hartman, and G. A. Petsko. 1990. Structure of yeast triosephosphate isomerase at 1.9 Å resolution. *Biochemistry.* 29:6609–6618.

Lolis, E., and G. A. Petsko. 1990. Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5-Å resolution: implications for catalysis. *Biochemistry.* 29:6619–6625.

Metropolis, N. S., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1092.

Neria, E., and M. Karplus. 1997. Molecular dynamics of an enzyme reaction: proton transfer in TIM. *Chem. Phys. Lett.* 267:23–30.

Pastor, R. W. 1994. Techniques and applications of Langevin dynamics simulations. *In* The Molecular Dynamics of Liquid Crystals. G. R. Luckhurst and C. A. Veracini, editors. Kluwer Academic Publishers, Dordrecht, The Netherlands. 85–138.

Pompliano, D. L., A. Peyman, and J. R. Knowles. 1990. Stabilization of a reaction intermediate as a catalytic device: definition of the functional role of the flexible loop in triosephosphate isomerase. *Biochemistry.* 29:3186–3194.

Sampson, N. S., and J. R. Knowles. 1992a. Segmental motion in catalysis: investigation of a hydrogen bond critical for loop closure in the reaction of triosephosphate isomerase. *Biochemistry.* 31:8488–8494.

Sampson, N. S., and J. R. Knowles. 1992b. Segmental motion in catalysis: definition of the structural requirements for loop closure in catalysis by triosephosphate isomerase. *Biochemistry.* 31:8482–8487.

Schlick, T., S. Figueroa, and M. Mezei. 1991. A molecular dynamics simulation of a water droplet by the implicit-Euler/Langevin scheme. *J. Chem. Phys.* 94:2118–2129.

Wade, R. C., M. E. Davis, B. A. Luty, J. D. Madura, and J. A. McCammon. 1993. Gating of the active site of triose phosphate isomerase: Brownian dynamics simulations of flexible peptide loops in the enzyme. *Biophys. J.* 64:9–15.

Wade, R. C., B. A. Luty, E. Demchuk, J. D. Madura, M. E. Davis, J. M. Briggs, and J. A. McCammon. 1994. Simulation of enzyme-substrate encounter with gated active sites. *Struct. Biol.* 1:65–69.

Wierenga, R. K., T. V. Borchert, and M. E. M. Noble. 1992. Crystallographic binding studies with triosephosphate isomerases: conformational changes induced by substrate and substrate-analogues. *FEBS Lett.* 307: 34–39.

Wierenga, R. K., M. E. M. Noble, G. Vriend, S. Nauche, and W. G. J. Hol. 1991. Refined 1.83 Å structure of the trypanosomal triosephosphate isomerase crystallized in the presence of 2.4 m-ammonium sulfate. *J. Mol. Biol.* 220:995–1015.

Williams, J. C., and A. E. McDermott. 1995. Dynamics of the flexible loop of triosephosphate isomerase: the loop motion is not ligand gated. *Biochemistry.* 34:8309–8319.

Yuksel, K. U., A. Sun, R. W. Gracy, and K. D. Schnackerz. 1994. The hinged lid of yeast triose-phosphate isomerase. *J. Biol. Chem.* 269: 5005–5508.

Zhang, Z., S. Sugio, E. A. Komives, K. D. Liu, J. R. Knowles, G. A. Petsko, and D. Ringe. 1994. Crystal structure of recombinant chicken triosephosphate isomerase-phosphoglycolohydroxamate complex at 1.8-Å resolution. *Biochemistry.* 33:2830:2837.