doi:10.1006/jmbi.2001.4808 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 310, 1039–1053

JMB

Investigation of the Mechanism of Domain Closure in Citrate Synthase by Molecular Dynamics Simulation

Danilo Roccatano[1], Alan E. Mark[1] and Steven Hayward[2]*

[1]*BIOSON Research Institute Laboratory of Biophysical Chemistry, University of Groningen, Groningen, The Netherlands*

[2]*Royal Society Wolfson Bioinformatics Laboratory School of Information Systems University of East Anglia Norwich NR2 7TJ, UK*

**Corresponding author*

Six, 2 ns molecular dynamics simulations have been performed on the homodimeric enzyme citrate synthase. In three, both monomers were started from the open, unliganded X-ray conformation. In the remaining three, both monomers started from a closed, liganded X-ray conformation, with the ligands removed. Projecting the motion from the simulations onto the experimental domain motion revealed that the free-energy profile is rather flat around the open conformation, with steep sides. The most closed conformations correspond to hinge-bending angles of 12-14° compared to the 20° that occurs upon the binding of oxaloacetate. It is also found that the open, unliganded X-ray conformation is situated at the edge of the steep rise in free energy, although conformations that are about 5° more open were sampled. A rigid-body essential dynamics analysis of the combined open trajectories has shown that domain motions in the direction of the closed X-ray conformation are compatible with the natural domain motion of the unliganded protein, which has just two main degrees of freedom. The simulations starting from the closed conformation suggest a free-energy profile with a small barrier in going from the closed to open conformation. A combined essential dynamics and hinge-bending analysis of a trajectory that spontaneously converts from the closed to open state shows an almost exact correspondence to the experimental transition that occurs upon ligand binding. The simulations support the conclusion from an earlier analysis of the experimental transition that the β-hairpin acts as a mechanical hinge by attaching the small domain to the large domain through a conserved main-chain hydrogen bond and salt-bridges, and allowing rotation to occur via its two flexible termini. The results point to a mechanism of domain closure in citrate synthase that has analogy to the process of closing a door.

Keywords: hinge bending; hinge axis; dynamic domains; essential dynamics analysis; rigid-body analysis

Introduction

Citrate synthase catalyses an important step in the citric acid cycle, the Claisen condensation of acetyl-coenzyme A with oxaloacetate, to form citrate and coenzyme A.[1] In animals, plants, fungi, archaebacteria, and gram-positive bacteria it is comprised of two identical subunits. The structures of citrate synthase from pig heart and chicken heart have been determined by X-ray crystallography.[2] The structures of the unliganded protein, as well as the protein bound to citrate and coenzyme A were solved. The monomer is comprised of two domains, a large domain and a small domain. The unliganded structure is called the open form. In the protein liganded to citrate and coenzyme A, the small domain is closed upon the large domain. This structure is called the closed form, a hydrolytic form, induced upon formation of citryl-coenzyme A. Structures of the enzyme liganded to its natural substrate, oxaloacetate, and inhibitor analogues of aceteyl-coenzyme A have been solved.[3,4] These structures have been called the nearly closed form.[1] Although there are import-

Present address: Danilo Roccatano, Department of Chemistry, University of Rome ''La Sapienza'', P. le Aldo Moro, 5, 00185, Roma, Italy.

Abbreviations used: MSF, mean-square fluctuation; PDB, protein data bank.

E-mail address of the corresponding author: sjh@sys.uea.ac.uk

0022-2836/01/051039–15 $35.00/0

ant differences between the closed and nearly closed forms, in comparison to the open to closed domain movement there is very little difference. These structural studies, along with kinetics experiments, indicate the following mechanism. The binding of oxaloacetate causes the small domain to close upon the large domain by about 18-20°, the difference in hinge-bending angle between the open structure and the closed structures. A binding site for acetyl-coenzyme A is subsequently formed that spans both domains. After the final stage of the reaction is completed, coenzyme A leaves first, followed by citrate.

Lesk & Chothia[5] studied the mechanism of domain closure in citrate synthase and concluded that the domain motion occurred through the accumulation of small shifts between packed segments at the domain interface. The motion was classified as a shear motion by Gerstein *et al*.[6] Nevertheless, it was clear from an early comparison of the two forms that a ''hinge point'' is located at His274 and Gly275.[1] This is significant, as His274 binds to oxaloacetate and is one of the proposed catalytic residues.[3] A domain and hinge-bending study of the open and closed forms has been made using the program DynDom,[7] which determines regions that act as ''mechanical hinges''.[8] These are defined to be regions that are both bending, in that a rotational transition occurs at these regions, and located in the vicinity of the interdomain screw axis. The DynDom analysis of citrate synthase revealed that the β-hairpin in this largely helical protein acts as a mechanical hinge by way of its two flexible termini and attaching the large domain to the small domain through a main-chain hydrogen bond and salt-bridges at its tip. The α-helix formed from residues 375-383 is also a mechanical hinge, and it therefore appears that these two separated hinging regions are largely responsible for the location and orientation of the hinge axis. A study of 24 domain proteins, for which two or more X-ray conformers are known, has shown that many domain proteins have two or more separated mechanical hinges creating a stable hinge axis for control of domain closure.[8] Citrate synthase apparently conforms to this ''door closing model'' of domain closure. This makes sense, as the binding site of acetyl-coenzyme A spans both domains in the closed conformation, implying that closure must be controlled precisely in order that the groups that bind acetyl-coenzyme A from the two domains are positioned correctly with respect to each other. The region containing residues 274 and 275, however, is distant from the hinge axis and it was suggested therefore that the term hinge point was inappropriate.

Due to its large size, there have been relatively few computational studies on citrate synthase that address internal motion. A normal mode analysis that used a block Lanczos algorithm to calculate the first ten normal modes has been reported.[9] The inner product between the lowest-frequency mode and the mode that represents the transition between the open and the closed crystallographic structures was 0.49. A domain analysis of the experimental transition was made and compared with the results of a domain analysis of the normal modes of citrate synthase.[10] Similar domain decomposition was reported. Finally, a study has been made of the possible path taken between the crystallographic open and closed structure. Two methods were compared: the Elber-Karplus[11] method, which determines the path with the lowest average potential energy, and the directed dynamics method, which restarts a molecular dynamics simulation in an iterative procedure, each time with a new set of velocities that are in a direction to bring it to the other structure.[12] The work focussed on particular dihedral angles that showed large variations along the path.

Here, we present a molecular dynamics study of citrate synthase. Three 2 ns simulations were performed on citrate synthase starting from the open X-ray structure, and a further three 2 ns simulations starting from the closed X-ray structure. Using projection techniques, essential dynamics analysis, and the domain and hinge-bending analysis implemented in the program, DynDom, new insights into the functional domain movement in citrate synthase are presented.

Results and Discussion

Standard checks on the simulations

As mentioned in the Introduction, three independent 2 ns simulations were performed on the dimer with both subunits starting from the open conformation, and a further three simulations, again each of 2 ns on the dimer with each subunit starting from the closed conformation. In order to check that the simulations did not contain obvious artefacts, a series of standard analyses were performed. Plots of the RMSD from the starting structures and of the secondary structure as a function of time indicated no significant changes in secondary structure or other unexpected behaviour. Table 1 gives the average backbone RMSD of the monomers in each simulation.

Rigid-body RMSD and projections

The domain decomposition used in this analysis was taken from an earlier DynDom analysis of the open to closed experimental transition, which used a window length of 11 residues.[7] Domain 1, the large domain comprised residues 1-55, 67-278, and 378-437, and domain 2, the small domain, residues 56-66, 278-377. The results presented below are insensitive to the inclusion or exclusion of small numbers of residues into the large or small domains.

Table 1. Summary of average backbone RMSDs from 500-2000 ps

Structure	Monomer	Simulation	From open (nm)	From closed (nm)
Open	A	I	0.26	0.34
		II	0.20	0.29
		III	0.20	0.25
	B	I	0.29	0.27
		II	0.22	0.25
		III	0.21	0.24
Closed	A	I	0.28	0.18
		II	0.20	0.26
		III	0.33	0.21
	B	I	0.30	0.17
		II	0.27	0.21
		III	0.35	0.23

Open simulations

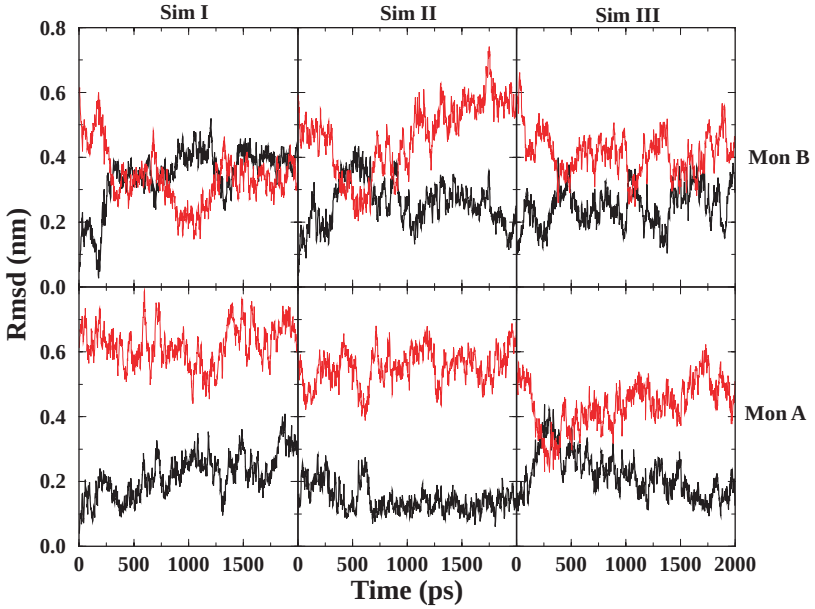
Figure 1(a) shows the rigid-body RMSDs for the simulations started from the open conformation. It is clear that on occasions the domain conformation is closer to the closed conformation than the open. Figure 2(a) shows the trajectories of the projections $\eta^o(t)$, $\hat{\eta}^o(t)$. These confirm the assessment given on the basis of the rigid-body RMSD. $\hat{\eta}^o(t)$ is seen to achieve values very close to 1.0 over a significant proportion of the total combined trajectory time (treating two monomers within a single simulation as independent trajectories), indicating that the domain movement is directly on the path between the open and closed experimental conformations. However, $\eta^o(t)$ achieves a maximum value of 0.78, indicating that the open conformation never reaches the closed conformation.

The quantity $\eta^o(t)$ can be used as a reaction coordinate. Merging the six open trajectories (treating the two trajectories in the dimer as independent), one can estimate the free-energy profile along this reaction coordinate by calculating the probability histogram determined from the distribution of values for $\eta^o(t)$ from the six trajectories. Figure 3(a) shows the resulting free-energy profile, where the reaction coordinate is in fact $\eta^o(t) \times 20$. The two vertical lines in this Figure show the locations of the open and closed experimental conformations. The free energy has a rather flat profile with steep sides. Interestingly, the open, unliganded experimental conformation, at 0° by definition, is located just before the steep rise, although even more open conformations with rotation angles of up to -5° were sampled. This result also shows that there are a large number of other unliganded conformations, all of which are as accessible as the experimental one, some achieving closure angles of 12-14° before the steep barrier is reached. The closed, liganded conformation is beyond this barrier, and it is likely that the binding of oxaloacetate provides energy to achieve closure.

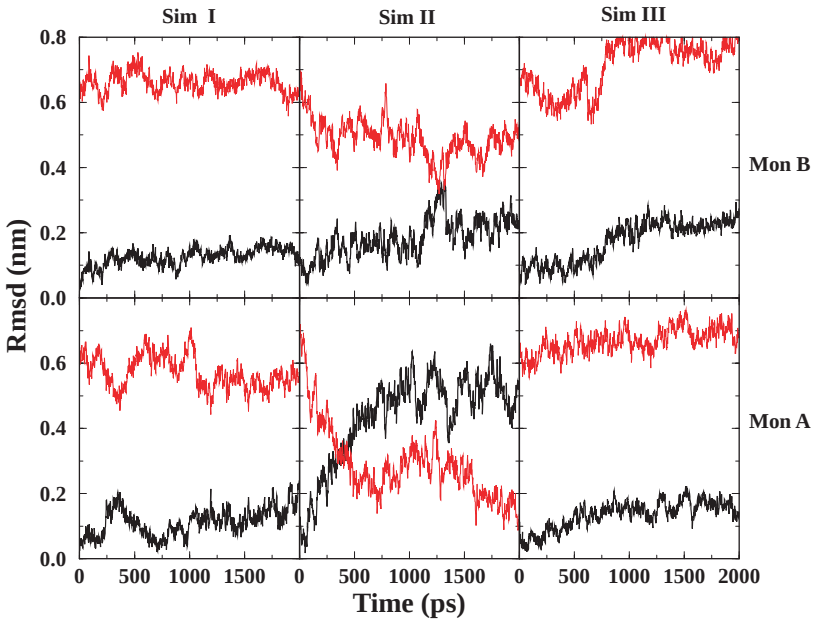
Closed simulations

Figure 1(b) shows the rigid-body RMSDs for the simulations that start from the closed conformation. It is clear that in all but one case the domains remain largely closed. The exception is in simulation II, where monomer A shows a dramatic movement towards the open conformation. Again the projections $\eta^c(t)$, and $\hat{\eta}^c(t)$ in Figure 2(b) confirm this. In this ''transitional trajectory,'' $\eta^c(t)$, and $\hat{\eta}^c(t)$ both achieve values well above 0.75 over a large portion of the simulation, sometimes achieving values close to 1.0, simultaneously. This indicates that the domain conformation (the relative position of the two domains irrespective of the internal conformations) in this trajectory superimposes almost perfectly on the experimental open domain conformation. The nature of the trajectory indicates a relatively rapid transition away from the closed conformation towards the open, followed by a final period during which it remains relatively near to the open conformation. The form and location of the transitional trajectory seen in Figure 2(b) is consistent with it descending the steep free-energy gradient that forms the barrier between the open and the closed conformations seen in Figure 3(a). After the transition, there is no discernible difference in its trajectory from those that started from the open conformation. Although the other subunit from this simulation remains more closed than open, the direction of movement is towards the open conformation. Obviously, more simulations would be necessary to confirm intersubunit cooperativity, but the experimental evidence is that there is none.[13]

In analogy to the analysis on the open simulations, the continuous line in Figure 3(b) shows the free-energy profile along the reaction coordinate from the merged closed trajectories, excluding the transitional trajectory. The broken line in Figure 3(b) shows the result when the trajectories of both monomers from simulation II are eliminated. The X-ray conformation is not situated at the minimum of the free energy, but slightly towards the open conformation. A difference would be expected, given that the X-ray conformation is liganded to citrate and co-enzyme A. It shows that the more closed conformations, with hinge-bending angles of up to 26-27° or more, are thermally accessible, although rare. Clearly, the lack of statistics limits detailed analysis but the fact



(a)

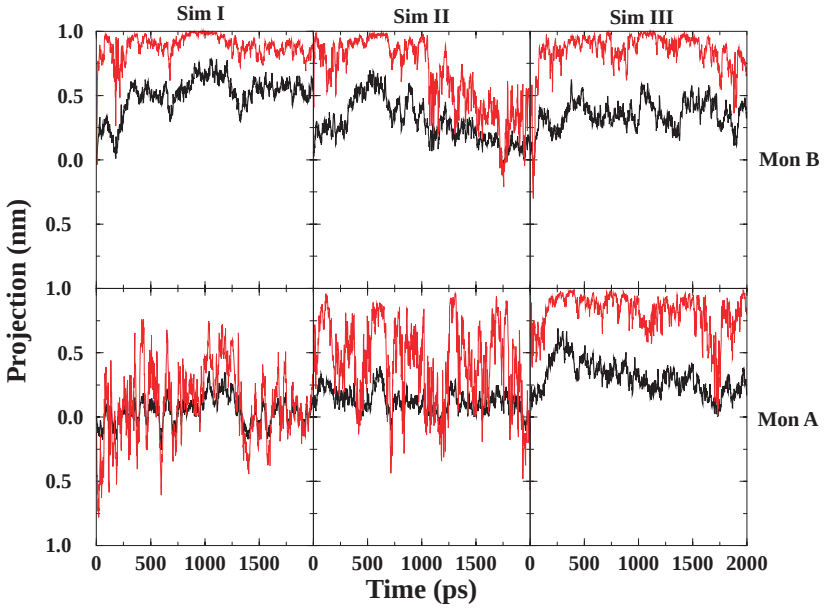


(b)

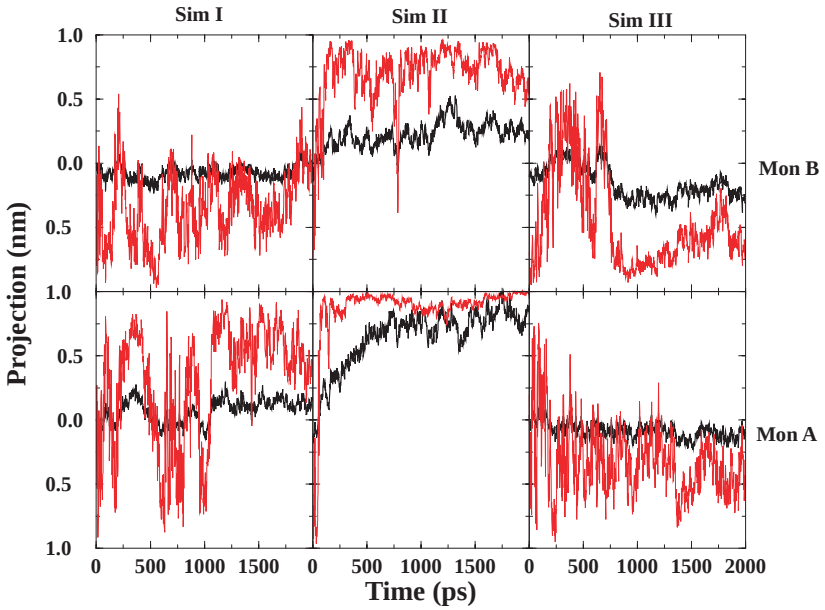
Figure 1. (a) Rigid-body RMSD plots for the open simulations. A column shows the result for each monomer from the same simulation. The thick line shows the rigid-body RMSD trajectory of the small domain from its position in the open X-ray conformation, the thin line from its position in the closed X-ray conformation. (b) Equivalent analysis for closed simulations, where the thick line is the rigid-body RMSD from the closed X-ray conformation, and the thin line from the open X-ray conformation.

that in five out of the six simulations the conformations remain closer to the closed, suggests that there is a free-energy barrier to be overcome in going from the closed to the open conformation.

Figure 4 illustrates a free-energy profile that is consistent with our results. The closed conformations have a higher free energy than the open, as the transition from the closed to open is apparently a downhill process, occurring at values of the reaction coordinate where the barrier determined from the open simulations is located. To verify this profile from free simulations, a much greater degree of sampling would be necessary in which multiple crossings between these two levels are observed. This is impractical and different techniques such as constrained molecular dynamics, or umbrella sampling are required. However, the profile illustrated in Figure 4 is consistent with the results from the simulations and is biochemically reasonable. The observation that the infusion of oxaloacetate into crystals of citrate synthase in the open form, causes the crystal to crack[1] suggests



(a)



(b)

Figure 2. (a) Trajectories of the projections $\eta^o(t)$ and $\hat{\eta}^o(t)$ (see equations (3) derived from the trajectories of the simulations that start from the open conformations, shown as thick lines and thin lines, respectively. A column shows the result for each monomer from the same simulation. (b) Trajectories of the projections $\eta^c(t)$ and $\hat{\eta}^c(t)$ derived from the trajectories of the simulations that start from the closed conformations, shown as thick lines and thin lines, respectively. A column shows the result for each monomer from the same simulation.

that oxaloacetate induces the conformational transition from the open to the closed form and therefore provides the energy to overcome the proposed free-energy barrier.

Rigid-body essential dynamics analysis of merged trajectories

All trajectories and the experimental conformations have been projected onto the space defined by the first two eigenvectors from the rigid-body essential dynamics analysis of the combined trajectories from the simulations starting from the open conformation. The result is shown in Figure 5.

Open simulations

In all there are six, 2 ns simulations of the monomer starting from the open conformation. These have been combined into a single trajectory and the rigid-body essential dynamics analysis made.

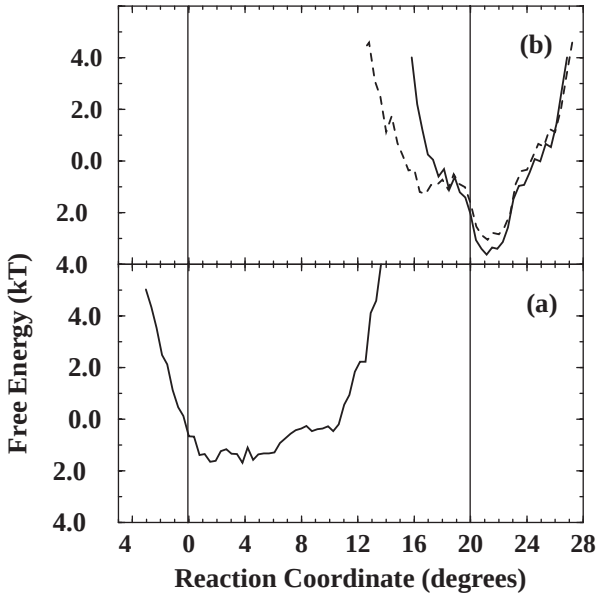


Figure 3. (a) The free-energy profile estimated from the open simulations, where the reaction coordinate is $\eta^{o}(t) \times 20$. Populations in bins of width 0.02 were calculated by creating histograms from the data presented in Figure 2(a). Bins that contained less than ten points (out of a total of 6000 created by dividing the combined trajectories in 2 ps intervals) were not included in the free-energy calculation. The vertical line at 0° is the value of the reaction coordinate at the open X-ray conformation, the vertical line at 20° is the value of the reaction coordinate at the closed X-ray conformation. (b) Free-energy profile derived in identical manner from data presented in Figure 2(b). The continuous line represents the predicted free-energy profile within the closed well derived from all the closed trajectories apart from the transitional one. The broken line represents the profile predicted when the trajectories of both monomers from the simulation in which the transition occurred are eliminated. The histogram parameters to create these were the same as for the open case, but the total number of points was 5000 and 4000, respectively. Note that we do not know the relative positions of these two free-energy profiles and their zero points are arbitrary.

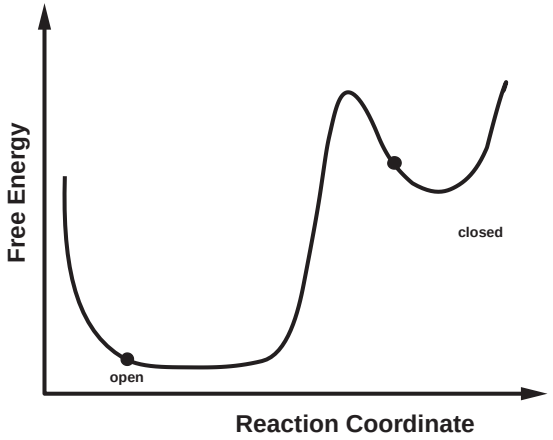


Figure 4. An illustration of the free-energy profile that is consistent with our results, showing the closed conformations at a higher free energy than the open conformations. The main evidence for this comes from the transitional trajectory, which suggests that a barrier is crossed, implying a lower free energy in the open conformation.

Table 2 shows the contribution of each essential mode to the total mean-square fluctuation (MSF), the inner product values $\gamma_n^{RG/o}$, and to what degree the experimental transition is represented by the modes. The first essential mode contributes 68 % to the total MSF, the second, 21 %, and the remaining four modes relatively little. The comparatively large contribution of the first essential mode is partly due to the variant of the essential dynamics analysis we have used, which builds a covariance matrix of fluctuations from the starting conformation rather than the average. The value of 68 %, therefore, also reflects the fact that the average of the distribution does not coincide with the starting conformation (see Figure 5). The experimental transition from the open to the closed conformation is represented to 98 % by just the first two eigenvectors, 87 % coming from the first essential mode, 11 % from the second. This shows that domain movements in the direction of the closed experimental conformation are part of the natural domain motion of the unliganded protein (which has just two degrees of freedom) when it starts from open experimental conformation.

Closed simulations

A similar analysis has been performed on the trajectories starting from the closed conformation but with the transitional trajectory eliminated. The transitional trajectory is treated separately below. Table 2 shows the contribution of each essential mode to the total MSF, the inner product values $\gamma_n^{RG/c}$, and to what degree the experimental transition is represented by the modes. The first essential mode contributes 47 % to the total MSF and the second essential mode 30 %. The experimental transition is represented to 79 % by just the first two eigenvectors, 77 % coming from the first essential mode, and only 2 % from the second. This is a lower figure than that found for the open simulations. This seems reasonable, because the motion of the two domains in the closed conformation will be influenced by interdomain interactions. The elliptical shape of the closed distribution seen in Figure 5, with its main axis pointing in the direction towards the open conformation, indicates that the predominant motion within the proposed energy well of the closed conformation, is in a direction that would bring it back to the open structure.

Table 2. Results of rigid-body essential dynamics analysis on merged trajectories in comparison with experimental transition

	Open simulations			Closed simulations[a]		
Mode number n	Contribution to total MSF (%)	$\gamma_n^{RG/o}$	Percentage to which first n modes represent experimental transition	Contribution to total MSF (%)	$\gamma_n^{RG/c}$	Percentage to which first n modes represent experimental transition
1	68.1	−0.93	87.2	47.2	0.88	77.5
2	21.4	−0.32	97.7	30.0	0.13	79.3
3	5.7	0.13	99.5	11.9	0.16	81.9
4	1.9	0.02	99.5	7.3	−0.38	96.7
5	1.5	−0.05	99.8	2.1	0.14	98.8
6	1.3	−0.02	99.9	1.5	0.01	98.8

[a] The transitional trajectory is not included in this analysis (see the text).

Essential dynamics analyses and DynDom analysis of the transitional trajectory

The application of the rigid-body essential dynamics technique is described above. In this section, we report on the results of a conventional essential dynamics analysis, i.e. one where intradomain fluctuations are not removed on a prior assumption of the domain decomposition. Table 3 shows the contribution of the first ten modes to the total MSF, their inner products γ_n^o with the experimental vector, and the extent to which the experimental transition can be represented by the modes, from an essential dynamics analysis of the combined trajectories from the open simulations. Table 3 shows the results of an equivalent analysis of the combined closed simulations, again omitting the transitional trajectory. The values indicate a reasonable correspondence in the first few modes for the merged open trajectories, but a much poorer correspondence for the merged closed trajectories. Table 3 shows the result of an equivalent analysis of the transitional trajectory alone. It shows a relatively high value of 0.73 for the inner product of the first mode with the experimental vector. The first mode dominates all other modes, with an 84 % contribution to the total MSF compared with 3 % for the second, and so on.

The maximum and minimum structures from eigenvectors indicated in Table 3 were generated as described in Methods, and used with the corresponding experimental structure as input for the program DynDom, in order to determine dynamics domains, hinge axes and hinge-bending regions. The results of this analysis were compared to the DynDom analysis of the X-ray conformers.

For the combined open analysis and the combined closed (excluding the transitional trajectory) no reasonable domain decomposition was found for any of the movements implied by the eigenvectors. In creating a domain decomposition, DynDom measures the ratio of interdomain displacement to intradomain displacement as defined by Hayward & Berendsen.[7] For a domain decomposition to be accepted for the hinge axis analysis, this ratio must be larger than 1.0, i.e. there must be more interdomain displacement than intradomain displacement. The fact that we do not obtain a

Table 3. Results of conventional essential dynamics analysis on merged trajectories and transitional trajectory in comparison with experimental transition

	Open simulations			Closed simulations[a]			Transitional trajectory		
Mode number n	Contribution to total MSF (%)	γ_n^o	Percentage to which first n modes represent experimental transition	Contribution to total MSF (%)	γ_n^c	Percentage to which first n modes represent experimental transition	Contribution to total MSF (%)	γ_n^c	Percentage to which first n modes represent experimental transition
1	46.6	0.50	25.2	45.4	−0.05	0.0	83.8	0.73	53.2
2	9.0	−0.32	35.2	13.3	−0.27	7.5	3.3	−0.01	53.2
3	7.8	0.20	39.2	8.1	0.31	17.0	2.8	−0.17	56.1
4	5.5	0.10	40.2	6.0	−0.27	24.2	1.5	0.20	60.0
5	4.4	−0.38	54.7	4.8	0.22	29.0	0.6	0.10	61.0
6	3.8	0.16	57.4	2.1	−0.01	29.0	0.5	−0.16	63.6
7	2.1	0.09	58.1	1.5	−0.08	29.0	0.4	−0.04	63.8
8	1.6	0.05	58.4	1.2	0.19	32.5	0.4	−0.20	67.8
9	1.4	−0.13	60.3	0.9	−0.30	41.4	0.3	0.18	71.2
10	1.0	−0.30	69.1	0.9	0.11	42.6	0.3	0.11	72.3

[a] The transitional trajectory is not included in this analysis (see the text).

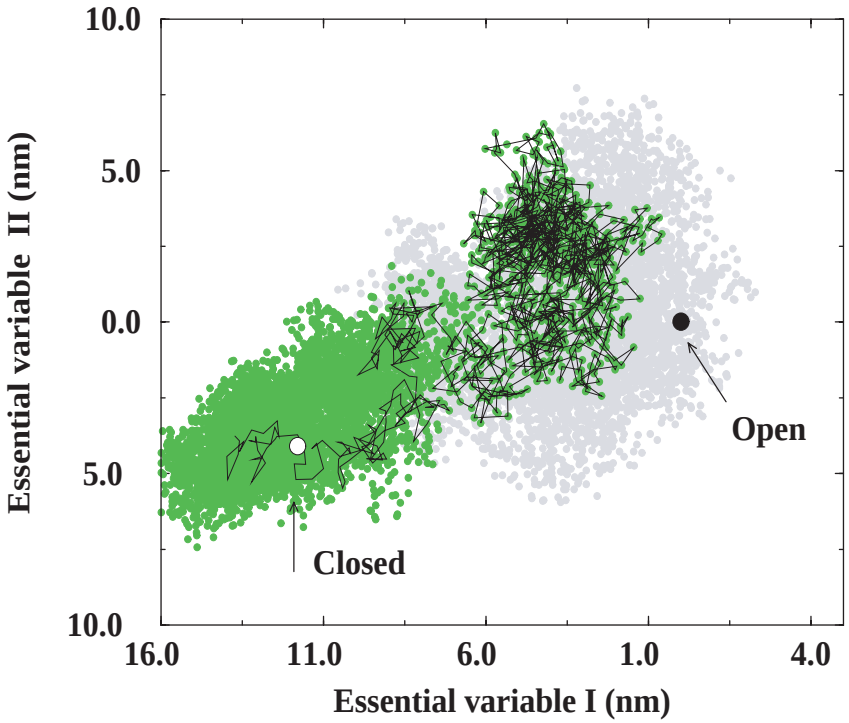


Figure 5. Projections of the trajectories onto the first two eigenvectors of the rigid-body essential dynamics analysis of the merged open trajectories. The faint grey dots represent points along the open trajectories, the darker grey dots, points along the closed trajectories. The joined dots show the transitional trajectory. The X-ray open and closed conformations are indicated in this projection. The line (not shown) between the two X-ray conformations represents the experimentally determined domain movement, and is represented to 98 % by the two eigenvectors that define this plane. The filled circles indicate the experimental conformations.

result with DynDom means that there is significant internal deformation within the domains selected. However, for the first eigenvector from the essential dynamics analysis of the transitional trajectory, DynDom makes a successful domain and hinge-bending analysis, using default parameter settings. The input into DynDom was the closed X-ray starting conformation, and the maximum structure from the first eigenvector, calculated as described in equations (1a) and (1b). In order to make an appropriate comparison to the experimental data, the X-ray conformers were re-analysed, also using default parameters. Table 4 compares the domain decomposition and the hinge-bending residues from the simulation and experiment. Figure 6 shows the comparison graphically. The correspondence between experiment and simulation is remarkable. A very similar domain decomposition is found, and a good correspondence in the hinge-bending residues. The simulation results confirm that the β-hairpin acts as a mechanical hinge with the hinge axis passing through its N terminus. The bending regions are situated at the termini of the β-hairpin in both cases, although the bending region of the C terminus involves residues 69 and 70 in the simulation case, but residues 64-67 in the experimental case. For the N-terminal side, residues 56 and 57 are assigned as bending residues in both cases. Residues 274-285 are assigned as bending residues from the simulation, whereas residues

Table 4. Comparison between experiment and simulation of domain decomposition and hinge-bending residues from DynDom analysis using a window length of five residues

	Experiment	First essential mode of transitional trajectory
Large domain residues	1-55, 68-274, 384-437	1-55, 70-273, 380-437
Small domain residues	58-63, 286-329, 347-372	58-68, 285-294, 299-328, 331-343, 352-377
Bending residues	56-57, 64-67, 275-285, 330-346, 373-383	56-57, 69-70, 274-284, 295-298, 329-330, 344-351, 378-379

DynDom sometimes eliminates small regions (usually one to three residues) from the analysis, but here these have been included into the domains of the residues that they neighbour for clarity. In those cases where the neighbour is a bending region, they are assigned as bending regions. The residues 275-278 form such a region for the experimental case (see the text for more details).

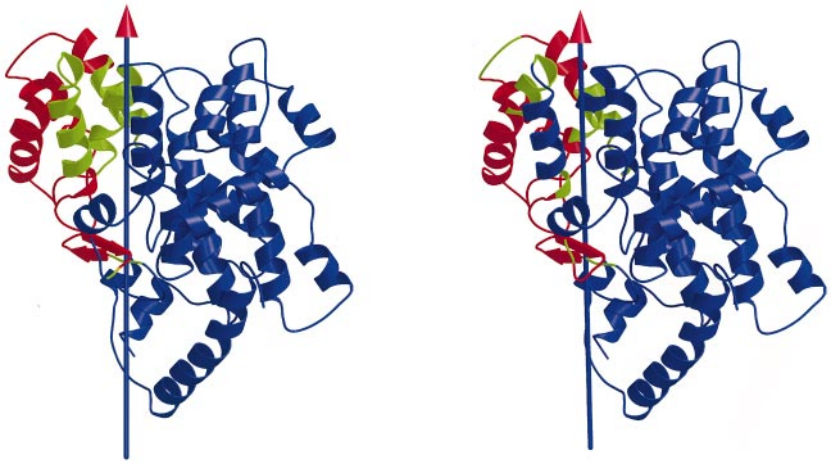


Figure 6. Comparison of the DynDom analysis of the experimental transition on the left, and the first essential mode of the transitional trajectory, on the right. In both cases, default parameters for DynDom were used, see Hayward & Berendsen[7] for details. The large domain is coloured blue, the small red, and the bending residues green (see Table 4 for details). The arrows indicate the hinge axes with the direction of rotation that would take this closed domain conformation of citrate synthase (PDB entry: 2cts) to the open domain conformation given by the right-hand rule. In both cases, the angle of rotation is 20°.

279-285 from experiment are assigned as bending residues. DynDom does not assign the residues 275-278 to either domain, or as a bending region. Visual inspection of the rotation vectors corresponding to residues 275-278 shows that it is most appropriately assigned as a bending region. In Figure 6 and Table 4, residues 275-278 and 343-346 have been assigned as bending residues, according to the rule that unassigned regions that neighbour a bending region are most appropriately assigned as a bending region. Another bending region spans residues 373-383 in the experimental case, but is less extended in the simulation case, involving just residues 378 and 379. The region 330-346 is also assigned as a bending region in the experimental case, whereas, the region 331-343 is assigned to the large domain, with residues 329-330 and 344-351 as bending residues in the simulation case. The hinge axes pass at an angle of 10° at a distance of 2.8 Å. In the experimental case, the motion is 80% closure, in the simulation case it is 86%. The rotation angle is 20° in the experimental case, and 20° in the simulation case. The translation along the hinge axis is −0.1 Å in the experimental case, −1.1 Å for the simulation case. The ratios of interdomain to intradomain displacements are 2.5 and 1.1 for the experiment and simulation cases, respectively.

The fact that we find the same general domain decomposition from the transitional trajectory, as from the experiment justifies our approach of imposing a domain decomposition from the outset determined on the basis of the experimental data.

Analysis of bending region 274-285 in both experiment and simulation

Analysis of the rotation vectors from the experimental open to closed transition, associated with the N-C^{α}-C^{β}-C tetrahedra of individual residues, or segments from short sliding windows (results not shown), has shown that the region 274-277 is primarily responsible for the bending between the two domains, with a lesser contribution coming from residues 278-285. Figure 7(a) shows the ϕ and ψ-dihedral angle values of these residues for the open conformation, and Figure 7(b) in the closed conformation. Figure 7(c) shows the dihedral angles of the maximum conformation from the first eigenvector of the essential dynamics analysis of the transitional trajectory. As this trajectory goes from the closed to an open conformation, this conformation corresponds to an open conformation. An analysis of the dihedral angle changes of residues 274-285 shows that the largest rotation occurs at the ψ-dihedral of His274 in both the experimental and simulated transition. A good correspondence is found for the dihedral angle charges ''between'' His274 and Gly275, with $\Delta\psi_{274}$ equal to 83° and $\Delta\phi_{275}$ equal to −26° in the experimental case, and $\Delta\psi_{274}$ equal to 76° and $\Delta\phi_{275}$ equal to −22° in the simulation case. The correspondence is not so good in the other dihedral angles. Comparing Figure 7(a) and (b), it is evident that the region 275-277 is more α-helical in nature in the closed form than the open. In fact the program DSSP[14] assigns Ala277 to the α-helix conformation in the closed X-ray structure, but not the open. In going from the open to closed structure, the carbonyl group of His274 has rotated about ψ_{274} such that the oxygen moves away from the negatively charged oxaloacetate molecule. The reverse of this process is seen in the simulated transition. The C^{α}-C bond, which is the ψ-dihedral axis, makes an angle of 42° and 33° to the hinge axis in the experimental and simulated transition, respectively. This shows how rotation about this bond can contribute to the hinge bending. However, the hinge axis is not coincident with this dihedral axis, but lies at a distance of roughly 12 Å. In Conclusions, we present a model that makes sense of these results.

Mutation studies where Gly275 has been substituted by alanine and valine show that the enzyme's activity is greatly reduced by these substitutions. This has been attributed to a hindering of the conformational transition destabilising the

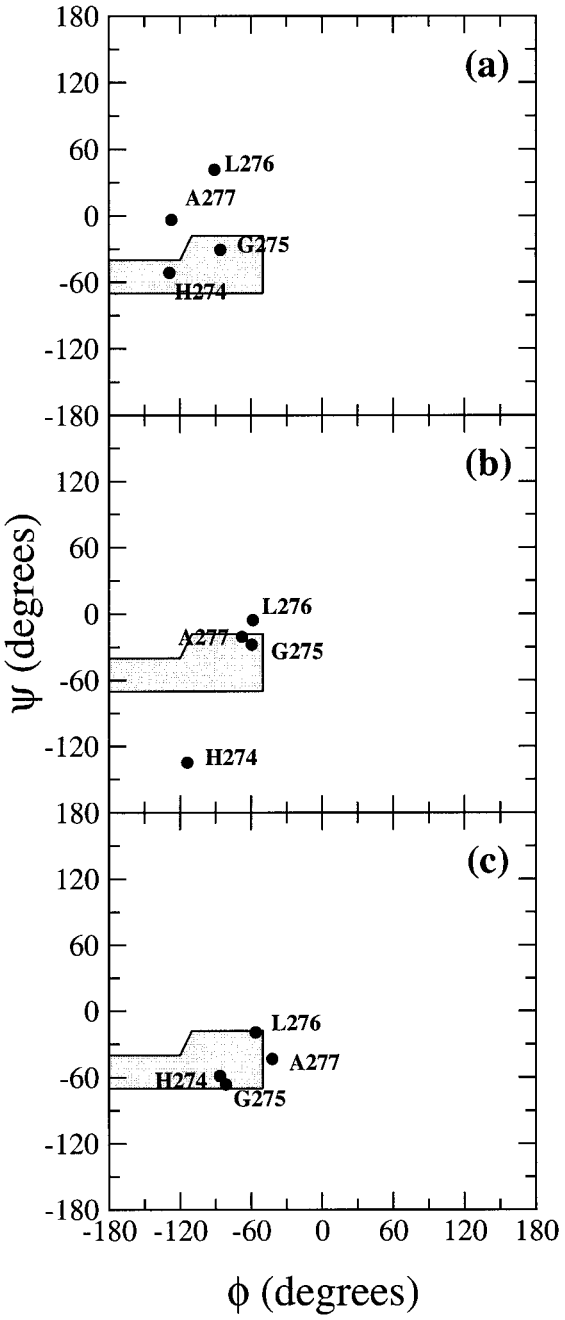


Figure 7. Ramachandran plots for the residues 274-277. (a) The ϕ,ψ-angles of the open X-ray conformation. (b) The ϕ,ψ-angles of the closed X-ray conformation. (c) The ϕ,ψ-angles of the maximum conformation from the first essential mode of the transitional trajectory. This corresponds to an open structure in its domain conformation, but started in the simulation from the closed experimental conformation with dihedral angles as in (b).

open conformation and, in the case of the valine residue, causing substantial disruption of the active site.[15]

Bonding of the β-hairpin

The X-ray structures show that the β-hairpin is bonded to the rest of the small domain *via* one or two salt-bridges and a main-chain hydrogen bond between Pro60 and Arg324. Focussing on the open simulations only, this bond is formed (in the DSSP definition) 79% of the total simulation time (summing all simulations excluding the first 500 ps in each). In the closed simulations, it is formed 77% of the total simulation time. This shows that this interaction is probably unaffected by the domain closure and helps to keep the tip of the β-hairpin firmly attached to the rest of the small domain. Table 5 shows the data regarding the potential salt-bridge forming residues in this region. We see that Asp61 and Arg324 remain within bridging distance in both the open and closed simulations. In slight contrast to what is observed in the open X-ray structure, Glu62 makes only a weak interaction with Lys325 in the open simulations. This bridge is not formed in any of the closed simulations nor is it observed in the X-ray structure of the closed form. Overall, the main difference with the experimental data is that the Glu62, Lys325

Table 5. Summary of possible salt-bridge partners in β-hairpin from 500-2000 ps

Structure	Monomer	Simulation/ Experiment	Asp61-Arg324 (nm)	Glu62-Lys325 (nm)	Asp61-Lys325 (nm)
Open	A	I	0.33	0.57	0.82
		II	0.65	0.45	0.54
		III	0.36	0.53	0.83
	B	I	0.39	0.61	0.72
		II	0.41	0.57	0.53
		III	0.33	0.47	0.79
		Exp	0.23	0.24	0.65
Closed	A	I	0.32	0.62	0.71
		II	0.44	0.84	0.62
		III	0.64	0.70	0.56
	B	I	0.37	0.70	0.69
		II	0.43	0.60	0.82
		III	0.41	0.69	0.59
		Exp	0.82	0.73	0.23

All distances between the charged side-chain atoms (C,O,N and H) of lysine, arginine, glutamic acid and aspartic acid were calculated at each time-frame and the minimum distance selected. The distances presented here are the averages over time of these minimum distances.

interaction is weaker in the open conformation, Asp61 has a stronger interaction with Arg324 in the closed conformation, and we do not observe the salt-bridge between Asp61 and Lys325 in the closed conformation. We do not see, therefore, Asp61 change bridging partner from Arg324 to Lys325 in the transition from open to closed. Despite this difference, salt-bridges are found at the tip of the β-hairpin in both open and closed simulations. These, together with the Pro60 to Arg324 hydrogen bond, keep this region attached to the small domain.

Conclusions

A large-scale simulation study of citrate synthase has been performed and the trajectories analysed using a variety of techniques. The simulations started from the open structure give a free-energy profile for the domain closure that has a steep free-energy barrier between the most closed conformation observed and the experimentally determined closed conformation bound to citrate and coenzyme A. All but one of the monomers from the simulations starting from the closed conformation remained closer to the closed conformation than to the open conformation. The exception showed a motion consistent with it crossing a barrier to a lower free energy. This transition occurs at values of the reaction coordinate where the steep free-energy barrier is predicted from the open simulations. Opening occurs very rapidly (within 1 ns). The comparison of this motion with the experimentally determined functional movement, which is induced by the binding of oxaloacetate, is remarkably good. The domain decomposition and hinge-bending regions are similar, and so are the location and orientation of the hinge axis, and the extent of the rotation. The implication, therefore, is that this free-energy barrier is overcome upon binding oxaloacetate, and the energy for this process comes from the interaction of oxaloacetate with the protein. The energy most likely comes primarily from the interaction of oxaloacetate with His274 and its neighbouring residues. His274 is one of the catalytic residues and is involved in the interdomain bending. The ψ-dihedral angle of His274 is seen to undergo a large change in both experiment and simulation, which in going from the open to closed conformation would rotate the carbonyl group such that the oxygen atom moves away from the negatively charged oxaloacetate molecule, which probably binds in a conformation different from that found in the closed structure.[16] This rotation is part of the mechanism that results in a more α-helical structure in residues 274-277, and may drive domain closure. Our results suggest that if coenzyme A and citrate have diffused away from the binding pocket, the protein can open, once a small free-energy barrier is overcome. However, the process of product release is thought to be a complicated one, due to the tight binding of citrate, and the presence of the products is sure to influence the energetics of the opening process.[1,2] The nature of the free-energy barrier between the open and closed conformations is unknown, but it has been pointed out that the ϕ, ψ-dihedral angles of His274 in the closed X-ray conformation put it well outside the ''extreme limiting contact'' region in a Ramachandran plot (see Figure 7(b)),[2] whereas in the open X-ray conformation the ϕ-ψ-dihedral angles put it near the α-helix region. This suggests that the energy resides in His274 itself. However, the dihedral angles of His274 are often located in low-energy regions in the closed simulations, implying that this interpretation is too simplistic.

The region 274-285 is too distant from the hinge axis to control the domain closure.[8] This is the role of mechanical hinges. The mechanical hinges are formed from the termini of the β-hairpin, and the α-helix 375-383, with further stabilisation coming from the region 330-346. With a region distant from the hinge axis that drives the domains open or closed, but the actual motion being under the

control of two separated hinges, the process of domain closure in citrate synthase suggests a model of domain closure that has an analogy with the process of closing a door, an ''arm-closing-door'' model, where the region 274-285 acts as the arm. Just like an arm, the region 274-285 undergoes a complex movement with multiple hinging regions, such that the actual rotation of the ''door'' can take place about an axis distant from the ''arm''. There is a good correspondence between experiment and simulation for the change in ψ- and ϕ-dihedral angles between His274 and Gly275, but there is not such a good correspondence for residues 276-285.

The ligand-induced domain movement was found to be an intrinsic motion of the unliganded protein. By this, it is meant that the ligand-induced domain movement is in a direction that lies in the space within which most of the fluctuation of the domains occurs for the unliganded protein. In the arm-closing-door model as presented above, the movement that occurs upon ligand binding is under precise control of the mechanical hinges. However, although the mechanical hinges restrict the space within which the domain motion occurs, the actual direction of the ligand-induced domain movement within this space may be influenced by the conformational change at the arm itself.

The results presented here, as with those of a previous study,[32] suggest that even simulations of just a few nanoseconds in length can, in certain circumstances, be used to identify functional movements in proteins. The results of this particular study show that a ligand-induced conformational change can be determined by simulating the reverse process, that is the reopening from the closed liganded conformation with the ligands removed. Further work will show whether this kind of approach can be used to predict functional movements in other enzymes.

Materials and Methods

MD simulations

The starting structures for the simulations were the unliganded, open structure (PDB entry 1cts[2]), and the closed structure in complex with coenzyme A and citrate (PDB entry 2cts[2]). The ligands were removed from the closed structure, and Ala32 in the closed structure was changed to valine so that the sequences of the open and closed structures were identical. The dimers of the open and closed structures were generated using a 2-fold symmetry operation, as given in the corresponding PDB file.

All simulations were performed using the GROMACS software package.[17] Protonation of the starting conformations resulted in an N-terminal $NH3^+$ group, and a C-terminal COO^- group. Histidine residues were protonated at either the δ or ε nitrogen atom so as to maximise their interaction with neighbouring groups. In the case of His274, the δ nitrogen atom was protonated in accordance with the proposed catalytic mechanism.[3]

The proteins were solvated in rectangular boxes with water, by stacking equilibrated boxes of the solvent to form a box large enough to contain the protein and 0.8 nm of solvent on all sides. All solvent molecules with any atom within 0.15 nm of the protein were removed. Since the resulting protonation state (at pH 7) of the molecule gives a total charge of −2, two sodium counterions were added to provide a neutral simulation cell. Each simulation box contained roughly 27,000 water molecules, resulting in about 90,000 atoms. During the productive phase of the simulation, constant pressure and temperature were maintained at one bar (1 bar = 10^5 Pa) and 300 K using the weak-coupling algorithm[18] with a coupling constant of 0.1 ps for the temperature and 0.5 ps for the pressure. The protein was coupled to the temperature-bath separately from the rest of the system. The GROMOS96 forcefield[19] was used. For the solvent, the SPC[20] water model was used. The LINCS algorithm[21] was used to constrain all the bond lengths. For the water molecules, the SETTLE algorithm was used to constrain bond lengths and the bond angle[22]. The dielectric constant was set to a value of 1.0. By using dummy hydrogen atoms[23] a time-step of 4 fs could be chosen. A twin-range cut-off method was used for non-bonded interactions. Non-bonded interactions within 0.8 nm were calculated every step, whereas interactions at a distance of between 0.8 nm and 1.4 nm were updated every ten steps. After construction, the potential energy of each system was minimised using 100 steps of steepest descent. The water was allowed to adjust to the presence of the protein by performing 10 ps of molecular dynamics with position restraints on the protein. Atoms were given initial velocities from the Maxwell distribution at 300 K. Each system was equilibrated for 10 ps before trajectory data was produced for analysis. This relatively short equilibration time was chosen to be long enough to remove any arteacts of the setting-up procedure, but not so long as to equilibrate the system fully, as we are interested in the process of equilibration itself. Three 2 ns production runs were performed on the open dimer and a further three 2 ns production runs were performed on the closed dimer.

Analysis of MD trajectories

Essential dynamics

Essential dynamics[24] and related analysis techniques have been widely applied and we refer to the original literature for a detailed description.[25–29] In this case, we are interested in the motions away from the open or closed X-ray structure, and have therefore used a variation of the standard prescription, whereby the fluctuations are not calculated with respect to the average structure, but from the starting structure, which in this case is one of the experimental structures. This type of analysis is useful where one is interested in the movement away from the starting conformation, as would be the case if the starting conformation were a non-equilibrium one. It also makes comparisons with the experimental data more straightforward. Each eigenvector describes a collective motion. This motion can be represented by two structures, one at the maximum value, the other at the minimum value of the collective variable when the trajectory is projected onto it. The ith atomic coordinate, x_{in}^{max} of the structure corresponding to the maximum value attained by the nth collective variable over the whole simulation time T is given by:

$$\Delta X_n^{\max} = \max_T\left(\sum_i V_{in}\Delta x_i(t)\right) \quad (1a)$$

$$x_{in}^{\max} = x_i^{\exp} + V_{in}\Delta X_n^{\max} \quad (1b)$$

and similarly for the minimum value. Here V_{in} is the ith component of the nth eigenvector, $\Delta X_n^{\max}$ is the maximum value attained by the nth collective variable over time, $x_i^{\exp}$ is the value of the ith atomic coordinate in the experimental structure, and $\Delta x_i(t)$ is the displacement of the ith atomic coordinate at time t from $x_i^{\exp}$. The maximum or minimum structure, together with the starting structure, can be used by the program DynDom,[7] to determine dynamic domains, hinge axes and hinge-bending residues.

In order to compare with the experimental transition from open to closed, we calculate the following quantity:

$$\gamma_n^{\circ} = \sum_i V_{in}\Delta\hat{x}_i^{\mathrm{o-c}} \quad (2)$$

where $\Delta\hat{x}_i^{\mathrm{o-c}}$ is the ith component of the unit-vector in the direction of the displacement from the open to the closed conformation (i.e. it is the eigenvector corresponding to this two-point distribution). We call this the experimental vector. γ_n^{o} quantifies the similarity between the collective mode described by the nth eigenvector and the actual experimental eigenvector. The extent to which the experimental transition can be ''represented by the first m eigenvectors'', is just the sum of the square of this quantity over the first m eigenvectors. An analogous quantity, γ_n^{c} exists for the closed to open transition. This method to compare modes of motion derived from simulation to those derived from multiple X-ray conformations has been used in a number of previous studies.[9,30–32]

Rigid-body domain analysis

Rigid-body RMSD. If one can identify distinct domains, internal motions within the domains can be removed by a least-squares best fit of a reference structure of each domain to those generated in the course of a simulation.[31] Figure 8 illustrates the procedure involved. It involves first superimposing the generated conformations of domain 1 at each time-frame on a single reference conformation of domain 1, repositioning the whole protein accordingly. The reference conformation is the experimental one, which is the starting conformation of the simulation. This will yield a trajectory of the motion of domain 2 relative to domain 1, but domain 2 will still contain internal deformations. To remove the internal deformations, the reference conformation of domain 2 is superimposed on each conformation of domain 2. One then has the rigid-body motion of domain 2 relative to domain 1. One can determine the trajectory of the RMSD between domain 2 as it moves as a rigid body, and domain 2 at the start of the simulation (i.e. the experimental conformation). Calculating this rigid-body RMSD relative to both the experimental open and closed conformations gives an indication of the movement away from the open conformation towards the closed conformation, or *vice versa*.

Rigid-body projection. The conformations along the rigid-body trajectory described above that start from the open structure are compared to the open to closed transition, by taking the inner products as:

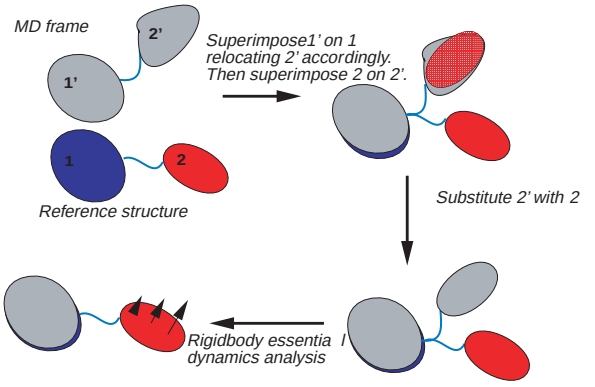


Figure 8. An illustration of the procedure to remove internal motions within the domains. This procedure is used to calculate the rigid-body RMSD, the projections onto the rigid-body projections, and the rigid-body essential dynamics.

$$\hat{\eta}^{\mathrm{o}}(t) = \sum_i \Delta\hat{x}_i^{\mathrm{RG/o}}(t)\Delta\hat{x}_i^{\mathrm{RG/o-c}} \quad (3a)$$

and:

$$\eta^{\mathrm{o}}(t) = \frac{\sum_i \Delta x_i^{\mathrm{RG/o}}(t)\Delta\hat{x}_i^{\mathrm{RG/o-c}}}{\sqrt{\sum_i (\Delta x_i^{\mathrm{RG/o-c}})^2}} \quad (3b)$$

and analogously for the closed to open transition, to give $\hat{\eta}^{\mathrm{c}}(t)$ and $\eta^{\mathrm{c}}(t)$. Here, $\Delta x^{\mathrm{RG/o}}{}_i(t)$ is the ith component of the rigid-body displacement vector from the experimental open conformation of domain 2, to its conformation at time t for the simulations that start from the open experimental structure, and $\Delta\hat{x}_i^{\mathrm{RG/o}}(t)$ is the corresponding unit-vector component. $\Delta\hat{x}_i^{\mathrm{RG/o-c}}$ is the ith component of the unit vector of the rigid-body displacement from the experimental open conformation of domain 2 to its conformation in the experimental closed conformation. The quantity $\hat{\eta}^{\mathrm{o}}(t)$ measures the extent to which domain 2 is on the path from the open to closed experimental conformations and achieves a value of 1.0 when it lies directly on the path. $\eta^{\mathrm{o}}(t)$ is a measure of the extent to which domain 2 has reached its position in the closed experimental conformation. It is a projection, however, and only when both these measures attain a value of 1.0, will domain 2 from the simulation be in the same position as found in the closed experimental conformation. An identical argument applies to $\hat{\eta}^{\mathrm{c}}(t)$ and $\eta^{\mathrm{c}}(t)$ for the closed to open simulations. It is important to realise that this analysis does not assume that the domains are rigid, but simply projects the motion onto a collective coordinate that describes the experimentally observed relative motion of the two domains.

Rigid-body essential dynamics analysis. This analysis was introduced in an application to a simulation of hen lysozyme.[29] It involves applying essential dynamics analysis to the rigid-body trajectory of domain 2 relative to domain 1 (see Figure 8). Here, fluctuations were also calculated from the starting structure rather than from the average. Only six eigenvalues are non-zero (a rigid body has only six degrees of freedom), with their associated eigenvectors describing the rigid-body motion of the

moving domain. One can quantify the similarity between the essential modes of motion and the experimental transition by calculating the following inner product in analogy to equation (2):

$$\gamma_n^{\mathrm{RG/o}} = \sum_i V_{in}^{\mathrm{RG/o}} \Delta \hat{x}_i^{\mathrm{RG/o-c}} \quad (4)$$

where $V_{in}^{\mathrm{RG/o}}$ is the ith component of the nth eigenvector of the rigid-body essential dynamics analysis on the simulation, or simulations, starting from the open conformation. An analogous quantity, $\gamma_n^{\mathrm{RG/c}}$ exists for simulations that start from the closed conformation. Again, the extent to which the experimental transition can be represented by the first m eigenvectors is just the sum of the square of this quantity over the first m eigenvectors. The sum of the square over the first six eigenvectors will equal 1.0.

In all the essential dynamics analyses, only backbone atoms were used.

Acknowledgments

This work was supported by the Netherlands National Computer Facility (SC-530). We thank Dr Moe Razaz and Professor Alfredo Di Nola for their support. D.R. was partly supported by an EU-TMR project on Structure and Dynamics of Intermediate States in Protein Folding.

References

1. Wiegand, G. & Remington, S. J. (1986). Citrate synthase, structure, control, and mechanism. *Annu. Rev. Biophys. Biophys. Chem.* **15**, 97-117.
2. Remington, S., Wiegand, G. & Huber, R. (1982). Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. *J. Mol. Biol.* **158**, 111-152.
3. Karpusas, M., Branchaud, B. & Remington, S. (1990). Proposed mechanism for the condensation reaction of citrate synthase: 1.9 Å structure of the ternary complex with oxaloacetate and carboxymethyl coenzyme A. *Biochemistry,* **29**, 2213-2219.
4. Wiegand, G., Remington, S., Deisenhofer, J. & Huber, R. (1984). Crystal structure analysis and molecular model of a complex of citrate synthase with oxaloacetate and s-acetonyl-coenzyme A. *J. Mol. Biol.* **174**, 205-219.
5. Lesk, A. M. & Chothia, C. (1984). Mechanisms of domain closure in proteins. *J. Mol. Biol.* **174**, 175-191.
6. Gerstein, M., Lesk, A. M. & Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry,* **33**, 6739-6749.
7. Hayward, S. & Berendsen, H. J. C. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins: Struct. Funct. Genet.* **30**, 144-154.
8. Hayward, S. (1999). Structural principles governing domain motions in proteins. *Proteins: Struct. Funct. Genet.* **36**, 425-435.
9. Marques, O. & Sanejouand, Y.-H. (1995). Hinge-bending motion in citrate synthase arising from normal mode calculations. *Proteins: Struct. Funct. Genet.* **23**, 557-560.
10. Hinsen, K., Thomas, A. & Field, M. J. (1999). Analysis of domain motions in large proteins. *Proteins: Struct. Funct. Genet.* **34**, 369-382.
11. Elber, R. & Karplus, M. (1987). A method for determining reaction paths in large molecules: application to myoglobin. *Chem. Phys. Letters,* **139**, 375-380.
12. Ech-Cherif El-Kettani, M. A. & Durup, J. (1992). Theoretical determination of conformational paths in citrate synthase. *Biopolymers,* **32**, 561-574.
13. Myers, J. A. & Boyer, P. D. (1984). Oxygen and deuterium exchanges show reversal of catalytic steps of citrate synthase: catalytic cooperativity is not observed. *Biochemistry,* **23**, 1264-1269.
14. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers,* **22**, 2577-2637.
15. Evans, C. T., Kurz, L. C., Remington, S. J. & Srere, P. A. (1996). Active site mutants of pig citrate synthase: effect of mutations on the enzyme catalytic and structural properties. *Biochemistry,* **35**, 10661-10672.
16. Mulholland, A. J. & Richards, W. G. (1998). Calculations on the substrates of citrate synthase I. Oxalacetate. *J. Mol. Struct. (Theochem.),* **429**, 13-21.
17. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. (1995). GROMACS: a message passing parallel molecular dynamics implementation. *Comput. Phys. Commum.* **91**, 43-56.
18. Berendsen, H. J. C., Postma, J. P. M., Gunsteren v., W. F. & Nola, A. D. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690.
19. van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hunenberger, P. H., Kruger, P. & Mark, A. E., *et al.* (1996). *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, ETH Zurich, vdf Hochschlverlang, Switzerland.
20. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. & Hermans, H. J. (1981). Interaction models for water in relation to protein hydration. In *Intermolecular Forces* (Pullman, B., ed.), pp. 331-342, , Reidel, Dordrecht, Holland.
21. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraije, J. G. E. M. (1997). LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472.
22. Miyamoto, S. & Kollman, P. A. (1992). SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J. Comput. Chem.* **13**, 952-962.
23. Feenstra, K. A., Hess, B. & Berendsen, H. J. C. (1999). Improving efficiency of large time-scale molecular dynamics simulations of hydrogen rich systems. *J. Comput. Chem.* **20**, 786-798.
24. Amadei, A., Linssen, A. B. M. & Berendsen, H. J. C. (1993). Essential dynamics of proteins. *Proteins: Struct. Funct. Genet.* **17**, 412-425.
25. Levy, R. M., Srinivasan, A. R., Olson, W. K. & McCammon, J. A. (1984). Quasi-harmonic method for studying very low frequency modes in proteins. *Biopolymers,* **23**, 1099-1112.
26. Kitao, A., Hirata, F. & Go, N. (1991). The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulation of melittin in water and in vacuum. *Chem. Phys.* **158**, 447-472.

27. Garcia, A. E. (1992). Large-amplitude nonlinear motions in proteins. *Phys. Rev. Letters*, **68**, 2696-2699.
28. Hayward, S., Kitao, A., Hirata, F. & Go, N. (1993). Effect of solvent on collective motions in globular protein. *J. Mol. Biol.* **234**, 1207-1217.
29. Hayward, S. & Go, N. (1995). Collective variable description of native protein dynamics. *Annu. Rev. Phys. Chem.* **46**, 223-250.
30. Ma, J. & Karplus, M. (1997). Ligand-induced conformational changes in ras p21: a normal mode and energy minimization analysis. *J. Mol. Biol.* **274**, 114-131.
31. van Aalten, D. M. F., Conn, D. A., de Groot, B. L., Berendsen, H. J. C., Findlay, J. B. C. & Amadei, A. (1997). Protein dynamics derived from clusters of crystal structures. *Biophys. J.* **73**, 2891-2896.
32. de Groot, B. L., Hayward, S., van Aalten, D. M. F., Amadei, A. & Berendsen, H. J. C. (1998). Domain motions in bacteriophage T4 lysozyme: a comparison between molecular dynamics and crystallographic data. *Proteins: Struct. Funct. Genet.* **31**, 116-127.
33. Hayward, S., Kitao, A. & Berendsen, H. J. C. (1997). Model free methods to analyse domain motions in proteins from simulation: a comparison of a normal mode analysis and a molecular dynamics simulation of lysozyme. *Proteins: Struct. Funct. Genet.* **27**, 425-437.

Edited by R. Huber

(Received 31 October 2000; received in revised form 25 May 2001; accepted 29 May 2001)