Pathway for Large-Scale Conformational Change in Annexin V

Jana Sopkova-de Oliveira Santos,[†,‡,§] Stefan Fischer,[||] Christophe Guilbert,[‡] Anita Lewit-Bentley,[§] and Jeremy C. Smith*[,‡,||]

Section de Biophysique des Protéines et des Membranes, Département de Biologie Cellulaire et Moléculaire, CEA-Saclay, 91191 Gif-sur-Yvette Cedex, France, LURE, Centre Universitaire Paris-Sud, Bâtiment 209D, 91405 Orsay Cedex, France, and Lehrstuhl für Biocomputing, IWR, Universität Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany

Received March 22, 2000; Revised Manuscript Received July 13, 2000

ABSTRACT: Crystallographic studies have shown that the binding of calcium to domain III of annexin V is accompanied by a large conformational change involving surface exposure of Trp187. Here we examine this conformational transition using computer simulation. It is found that the burial of Trp187 is accompanied by a large increase in conformational strain, compensated by improved protein−protein interaction energies. A low energy pathway for the conformational change is determined using the conjugate peak refinement method [Fischer, S., and Karplus, M. (1992) *Chem. Phys. Lett. 194*, 252−261] with solvent effects taken into account using nonuniform charge scaling. The pathway obtained is complex, involving >300 dihedral angle transitions and the complete unwinding of one helix. Acidic residues play a key role in the conformational pathway, via a succession of direct hydrogen bonds with the indole ring of Trp187. This finding is discussed in the light of experimentally determined pH, calcium ion and mutational effects on the conformational transition [Sopkova, J., Vincent, M., Takahashi, M., Lewit-Bentley, A., and Gallay, J. (1998) *Biochemistry 37*, 11962−11970 and Beermann, B. B., Hinz, H.-J., Hofman, A., and Huber, H. (1998) *FEBS Lett. 423*, 265−269].

Annexins constitute a family of hydrophilic proteins all of which undergo reversible calcium-dependent binding to phospholipid bilayers and certain cellular membranes (*4, 5*). They are found in various species, tissues, and cell types, with about 18 members having been identified and characterized so far. They share a high sequence homology (∼40−60%). All annexins contain a 4- or 8-fold repeat of domains of about 70 residues that are highly conserved and a more variable N-terminal segment. The conserved repeats account for the major part of the protein (called the "core"), and it is in these regions that the calcium and phospholipid binding sites are situated. The N-terminal exhibits very little homology between annexins and might be involved in determining functional differences between individual annexins (*6*).

Annexins are abundant in most eukaryotic cells, comprising up to 1% of the total cell protein. Although their physiological functions are as yet unknown, their wide range of in vitro properties has led to an even wider range of proposed in vivo functions. Some may be involved in various types of membrane fusion events such as endo-, exo-, and phagocytosis, which may be related to their ability to bind to acidic phospholipid bilayers in a calcium-dependent manner (*4*). Several annexins exhibit antiinflammatory and anticoagulatory properties, while others display ion channel activity (*4, 6*).

All crystal structures of annexins solved to date show that the conserved sequence domains correspond to structural domains. Each structural domain, numbered I to IV, consists of five α-helices (A−E), four forming an antiparallel bundle (helices A, B, D, and E) to which the fifth (helix C) lies perpendicular (see Figure 1A). The four domains are arranged in a cyclic array, giving the molecule an overall flat, slightly curved shape with a convex and a concave face. The calcium binding sites are located on the convex, membrane-binding face of the protein, while the N- and C- termini lie on the concave face (*7*). Each structural domain contains one principal calcium-binding site. The calcium ion binds to carbonyl oxygens in the loop connecting helices A and B, and to a carboxyl group of a negatively charged amino acid side chain (Glu or Asp) about 40 residues downstream, in the loop connecting helices D and E. It has been shown that the presence of a negatively charged amino acid in this position is a necessary condition for the existence of a calcium-binding site (*8*).

Annexin V, due to its potential anticoagulant activity, has been the most studied member of the annexin family, and its 3D structure was the first to be solved (*7*). To date, several different X-ray structures of annexin V and its mutants have been obtained from different crystallization conditions. It was shown that when the calcium ion is not bound in domain III, the loop connecting helices A and B lies in the center of a cavity lined with antiparallel helices and the sole tryptophan (Trp187), which lies at the extremity of this loop, is buried in the interior of the domain (see Figure 1B) (*7, 9, 10*)—we call this structure here "TRP-IN".

The crystal structure of annexin V with a calcium ion bound in domain III shows significant structural changes compared to the protein with no calcium bound (*11−13*). The AB loop moves in a direction approximately parallel to

[†] J.S.O.S. was supported by EC Grant BIO4-96-0083 for a postdoctoral fellowship

* To whom correspondence should be addressed.
[‡] Section de Biophysique des Protéines et des Membranes.
[§] LURE.
[||] Lehrstuhl für Biocomputing.

10.1021/bi000659h CCC: $19.00 © 2000 American Chemical Society
Published on Web 10/24/2000

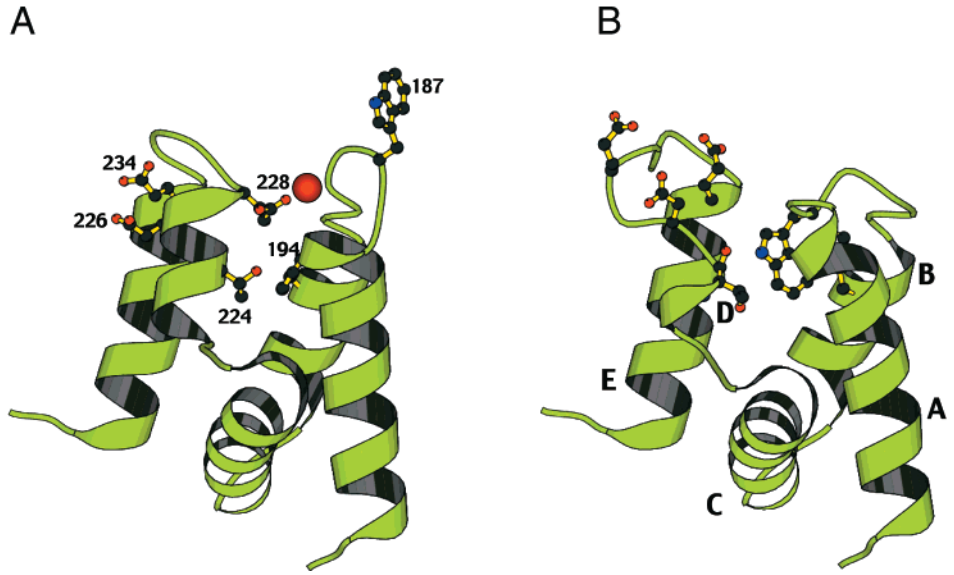

FIGURE 1: Crystal structures of the two conformational states of domain III of annexin V (A) with the calcium site (i.e., TRP-OUT) from ref *11* and (b) without the calcium-binding site (i.e., TRP-IN) from ref *28*. Important side chains are shown in a ball-and-stick representation. The large sphere in panel A represents the Ca^{2+} ion [figure produced using Molscript (*43*)].

the axes of helices A and B, thus bringing Trp187 onto the surface of the protein (see Figure 1A)—we call this structure "TRP-OUT". The position of the loop connecting helices D and E also changes significantly, allowing Glu228 to approach and ligate the calcium. A change in the arrangement of the helices is also observed: helix D changes from a 3_{10} helix of three or four residues [221(2)−224] in the structure without calcium to an α-helix of eight residues (221−228) in that with calcium bound. Fluorescence spectroscopy studies of the Trp187 in solution suggest that the conformational change from TRP-IN to TRP-OUT can occur under various conditions: in the presence of phospholipids and calcium (*14*, *15*), in acidic pH (*2*, *3*), or in the presence of high concentrations of calcium alone (*16*−*18*).

The known structures of members of the annexin family other than annexin V correspond either to the form with a principal calcium site in domain III [annexin II (*19*), annexin III (*20*); annexin VI (*21*), and annexin XII (*22*)] or to the form without it [annexin IV (*23*)]. However, unlike for annexin V, the alternative structures of these proteins have not been determined.

Annexin V system is an example of a protein for which ligand binding is accompanied by a large-scale conformational change. It is now recognized that conformational change is a common feature of protein function. Using X-ray crystallography, many cases have been recorded in which, in different functional states, proteins exhibit significant conformational differences, involving intra- and sometimes interdomain displacements (*24*). A "grand challenge" for biochemistry is to understand the mechanics of functionally related conformational change in macromolecules. This understanding requires characterization of the pathways between end-states, i.e., the sequence of structures accompanying the transition from one stable state to another, and the associated free energy variations. This presents a major problem as, while crystallography can provide valuable information on equilibrium end-state protein structures, the experimental characterization of conformational pathways is much more difficult to obtain, due primarily to their nonequilibrium, rarely populated nature.

Computer simulation techniques can in principle provide the required information. Using empirical energy functions, searches of protein conformational space can be used to determine plausible pathways. One of the techniques most heavily used for examining protein flexibility is molecular dynamics (MD) simulation. However, with present-day computer power, the sampling obtained with standard MD is insufficient if the conformational change to be studied occurs on a time scale longer than about a nanosecond, as is frequently the case. Therefore, specialized techniques are required. Although the development of these techniques is still in its infancy, a number do exist and have been used to examine important conformational changes in several proteins (*1*, *25*−*28*).

Here we use one of the most extensively tested methods, conjugate peak refinement (CPR) (*1*) to examine the annexin transition. CPR finds a continuous minimum energy path that follows valleys and crosses saddle-points of the high dimensional energy surface. In the present application, solvent effects are taken into account, using a novel nonuniform charge scaling method. The modeled pathway is presented as the sequence of molecular events leading from the reactant (TRP-OUT) to the product (TRP-IN) conformation of domain III. It is found that the burial of Trp187 is accompanied by a large increase in conformational strain, compensated by improved protein−protein interaction energies. The pathway obtained is complex, involving >300 dihedral angle transitions and the complete unwinding of one helix. Acidic residues are found to play a major role in the conformational change via a succession of direct hydrogen bonds with Trp187, suggesting experiments with which the pathway can be further characterized.

METHODS

All calculations were performed using version 23 of the program CHARMM (*29*), with the exception of some preliminary coordinate manipulations for which programs from the CCP4 package were used (*30*).

Model System and End State Structures. The coordinates of two X-ray structures of annexin V were used for the

calculations: TRP-IN (*32*) and TRP-OUT (*11*). In the X-ray analyses, both structures were refined using the same procedure to similar resolutions (l.9 and 2.0 Å for TRP-OUT and TRP-IN, respectively). The pathway determined corresponds to the conformational transition of the protein after dissociation of the calcium, i.e., the pathway between the structures with and without a calcium binding site, both with the calcium ion removed.

Force Field and Solvent Shielding. The model system consisted of all heavy atoms of the protein and of the polar and aromatic hydrogens [CHARMM parameter set 19 (*31*)]. The potential energy function contains bonded terms, representing bond length, valence angle and torsional (dihedral) angle variations and nonbonded (van der Waals and electrostatic) interactions. The van der Waals interactions were truncated using a switching function between 5.0 and 9 Å, and the electrostatic interactions were smoothly brought to zero at 9 Å using a shifting function (*29*).

Shielding of the electrostatic interactions in the protein by bulk water was taken into account here by nonuniform scaling of the partial atomic charges of the protein. This approach, which has been applied successfully to enzymatic (*33*) and ligand-docking (*34*) studies, consists of reducing the charges of localized groups of atoms by a group-specific scaling factor. Here the protein was partitioned into residues and each residue into a backbone group and a side-chain group. The scaling factors were determined for each group in such a manner that the standard Coulombic interaction energy, when computed with the scaled charges, optimally reproduced the calculated solvent-shielded interaction energy of each group with the rest of the protein.

The procedure for determining these scaling factors was as follows. The interaction energy of a group *i* with the rest of the protein can be written as

$$E_i = \sum_j \boldsymbol{q}_j \phi_i(\boldsymbol{r}_j) \qquad (1)$$

where the index *j* runs over all partial charges q_j of the protein (i.e., the probe charges) that do not belong to group *i*, and $\phi_i(\boldsymbol{r}_j)$ is the electrostatic potential generated at the probe charges by the partial charges of group *i* (i.e., the source charges). The vacuum Coulombic interaction energy E_i^{Vac} is obtained by replacing $\phi_i(\boldsymbol{r})$ in eq 1 with the Coulombic potential:

$$\phi_i^{\mathrm{Coul}}(r) = \sum_{k \in i} \frac{q_k}{|r - r_k|} \qquad (2)$$

where the index *k* runs over all the partial charges q_k belonging to group *i*. The interaction energy E_i^{Solv} of the solvated system is obtained when $\phi_i(\boldsymbol{r})$ in eq 1 is replaced by a solvated electrostatic potential ϕ_i^{Solv}. To compute E_i^{Solv}, the bulk solvent is modeled as a polarizable high-dielectric continuum. This is adequate for representing the solvent screening effect as this is essentially due to the reaction field of the polarized solvent. With this simplified representation of the solvent, E_i^{Solv} can be obtained by solving the linearized Poisson−Boltzmann (LPB), equation

$$\kappa\phi_i(\boldsymbol{r}) - \nabla\epsilon(\boldsymbol{r})\nabla\phi_i(\boldsymbol{r}) = \rho_i(\boldsymbol{r}) \qquad (3)$$

where $\rho_i(\boldsymbol{r})$ is the charge distribution of group *i*, $\epsilon(\boldsymbol{r})$ is the spatially dependent relative dielectric constant taken as $\epsilon(\boldsymbol{r}) = 80$ for bulk water outside the protein and as $\epsilon(\boldsymbol{r}) = 1$ for the protein interior (this value is consistent with the dielectric constant used for the derivation of the CHARMM charge parameters), and κ is a constant proportional to ionic strength.

The LPB equation was solved numerically for each source group *i*, by the finite difference method, as implemented in the program UHBD (*35*), using grid-focusing (*36*). The initial grid of 25 × 25 × 25 with a grid-spacing of 4 Å contains the whole protein surrounded by a layer of solvent at least 25 Å thick. The final focused grid had a size of 65 × 65 × 65 with a grid-spacing of 0.2 Å, and was centered onto source group *i*. The ionic strength was set to 145 mM (approximately physiological conditions). To allow the comparison between E_i^{Vac} and E_i^{Solv}, neither distance cutoff nor bonded-pair exclusions were used in the calculation of E_i^{Vac}, since these are not used for the calculation of E_i^{Solv} either. Taking the ratio for each group *i* between E_i^{Vac} and E_i^{Solv} yields a group-dependent dielectric constant $\epsilon_i = E_i^{\mathrm{Vac}}/E_i^{\mathrm{Solv}}$, which accounts for the difference in the solvent shielding of the interactions of different groups.

A combination rule for the shielding of the interaction between two residues *i* and *j* can be defined as $\epsilon_{ij} \equiv (\epsilon_i\epsilon_j)^{1/2}$. This is equivalent to dividing the charges of each group *i* by a factor $\gamma_i = \epsilon_i^{1/2}$. Computing the Coulombic interactions with these scaled protein charges yields for each group the interaction energy E_i^{Shield}, whose value is close to the actual solvated interaction E_i^{Solv}.

A difficulty in calculating charge scaling factors arises for certain groups when E_i^{Solv} and E_i^{Vac} are of opposite signs, in which case a meaningful scaling factor cannot be determined for these groups. For example when the protein has a net charge (−9 in the case of annexin), due to an imbalance in the number of acidic and basic side chains, most of the negatively charged side chains (Asp and Glu) have a positive vacuum interaction energy with the rest of the protein while their solvated interaction energies are negative. This behavior is not limited to charged side chains and occurs also for some neutral groups. While the Coulombic and solvated interactions between two single point charges will always have the same sign, an opposite sign can arise for the interaction between two collections of charges, when either collection is a mixture of positive and negative point charges. To avoid this problem, every scaling factor γ_i was determined by computing the interaction energies E_i^{Vac} and E_i^{Solv} with the absolute value of the partial charges.

Finally, the charge scaling factors γ_i were all divided by a single constant γ, which was chosen so as to compensate systematic deviation of E_i^{Shield} from E_i^{Solv}, caused, for example by the use of a distance cutoff. With the cutoff scheme used here (see above) the optimal value of γ was found to be 2.8. The resulting scaling factors γ_i for annexin in the product conformation have an average value of 1.56 for the backbone groups (γ_i ranging from 1.03 for Val 270 to 3.22 for Ala101) and of 2.41 for the side chains (γ_i ranging from 1.07 for Met 271 to 6.45 for Lys 27). The values E_i^{Shield} obtained with these scaling factors are plotted for each side chain against E_i^{Solv} in Figure 2B. The corresponding unshielded vacuum interactions E_i^{Vac} are plotted in Figure 2A.

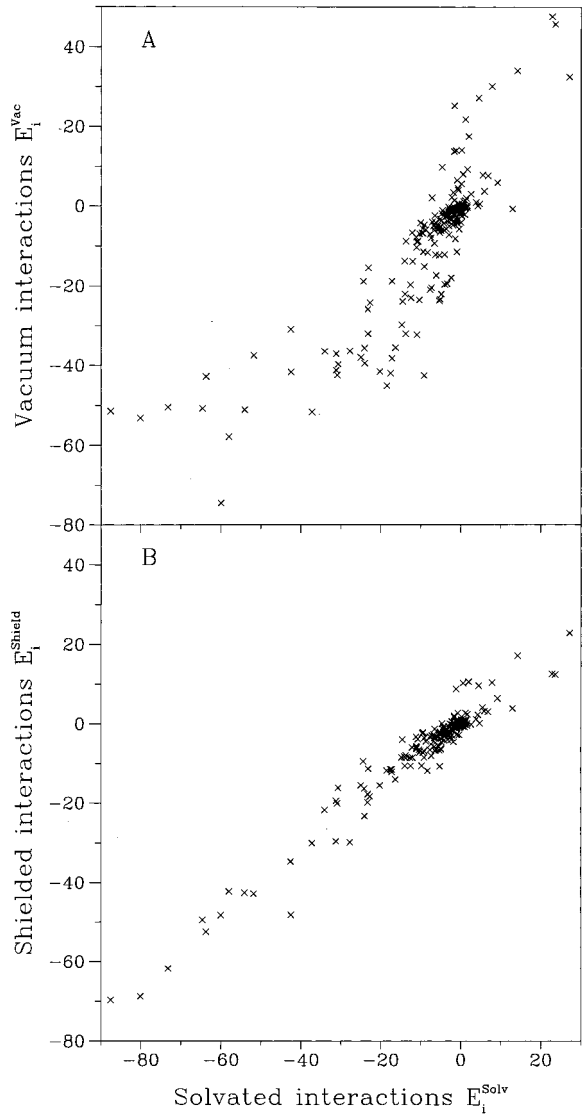

FIGURE 2: Interaction energy (kcal/mol) of each side-chain *i* with the rest of annexin V in the TRP IN conformation. E_i^{Solv}, the solvated interaction energy (continuum dielectric model of solvent, see Materials and Methods) is plotted versus the interaction of the respective side chain calculated as (A) E_i^{Vac}, the vacuum Coulombic interaction ($\epsilon = 2$); (B) E_i^{Shield}, the shielded Coulombic interaction obtained by scaling the partial charges with the residue dependent scaling factors γ_i. Both E_i^{Vac} and E_i^{Shield} interactions shown here were cut off with the same 9 Å shift function and bonded-pair exclusions as used in the path calculations.

Comparison of the two figures shows that the scaling of the partial charges significantly improves the agreement of the Coulombic interactions with the theoretical solvated interaction energies. To examine the conformation dependence of the scaling factors, we derived them for both protein end states. The difference between the two sets of γ_i was found to be small (3% on average, standard deviation 12%). The constant charge scaling factor used for each group during the path calculations was taken as the mean value between the end-states.

Energy Minimization. The coordinates of the atoms of domains I, II, and IV differ only slightly between the two states. Therefore, to render the pathway calculation computationally tractable the coordinates of these atoms were not allowed to change. Consequently, the atomic coordinates of these domains must be the same in the two end states. To achieve this, the structure of TRP-OUT was minimized with all atoms of domain III fixed. The resulting structure of domains I, II, and IV was then used to build two models, by combining it with domain III of either the TRP-OUT or the TRP-IN structures. Keeping the domains I, II, and IV fixed these structures were energy minimized to an RMS energy gradient of $<10^{-3}$ kcal/(mol Å).

As expected, the minimized structures differ slightly from the crystal structure coordinates: the RMSD between the coordinates of domain III before and after minimization is 2 Å for all heavy atoms of TRP-OUT and 0.6 Å for TRP-IN. The larger change in the TRP-OUT structure is due to the absence of the calcium ion. The side chains of amino acids proximal to the calcium binding site moved during the TRP-OUT energy minimization procedure, but the peptide backbones of the AB and DE loops preserved their crystallographic geometries. The TRP-IN structure showed one difference in the secondary structure: minimization changed helix D (four residues) from a 3_{10} helix into an α helix.

Calculation of the Transition Path. An initial approximate path between the two end-point structures was generated using the method of Path Exploration with Distance Constraints (PEDC) (*37*). The PEDC method uses a constraint potential added to the energy function that is harmonic in the distance from the target end state structure, expressed as the RMS difference between the Cartesian coordinates. Gradually changing the reference value of the constraint potential and minimizing brings the system from one end state to the other. In the present PEDC calculation, the trajectory ran from TRP-OUT to TRP-IN. Given that the RMS difference between the end structures is 4.1 Å, a step size of 0.1 Å was chosen for the PEDC calculation with a constraint force constant of 100 kcal/mol/Å^2. This method generated 42 intermediate structures, which were subsequently energy minimized.

The PEDC method cannot find transition states, nor a continuous minimum energy path, but serves to provide an initial guess for the conjugate peak refinement (CPR) method (*1*). The CPR method, starting from the initial guess, finds a minimum-energy path connecting the end-states and locates the saddle points on the energy surface that are transition states along the path. CPR has been extensively tested on a range of small molecule and macromolecule conformational transitions (*1*, *33*, *34*, *38*−*42*).

We briefly summarize how convergence to a saddle point-(s) is achieved with CPR. At a true saddle point, the Hessian matrix $\boldsymbol{H}_s$ (the matrix of the second derivatives of the energy), whose dimension is $D = 3N - 6$, has exactly one negative eigenvalue. This means that in the vicinity of a saddle point, for any set of vectors s_j conjugate with respect to $\boldsymbol{H}_s$, there will be one direction s_o along which the energy has a local maximum, and $D - 1$ directions along which the energy has a local minimum. Once a path segment crossing a saddle-region along which the energy has a local maximum is found, which is achieved by a heuristic procedure of adding and removing path points, this direction is taken as s_o and the rest of the conjugate basis (s_j, $j > 0$) is built recursively, minimizing the energy along each direction s_j.

The path is described by a series of structures which, displayed in sequence, yield a "movie" of the conformational transition. The reaction coordinate, λ, is defined as the sum

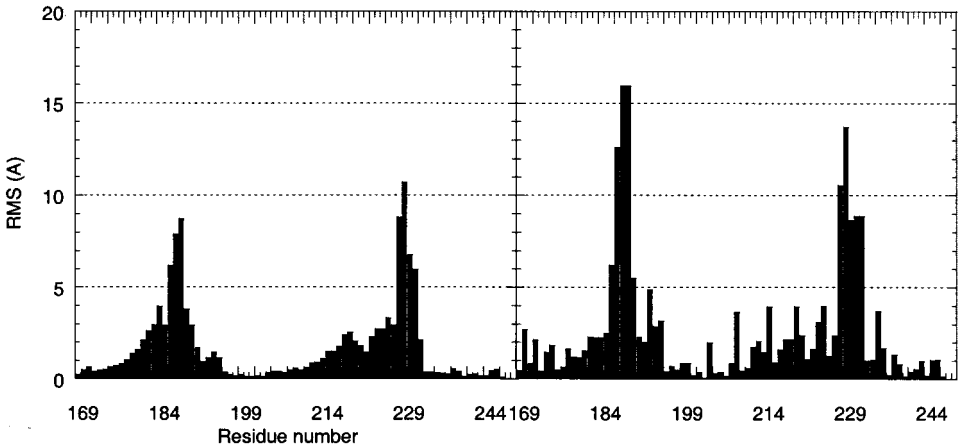

FIGURE 3: RMS coordinate difference (RMSD) for each residue (i.e., residues 168−247) between the TRP-OUT and TRP-IN structures: right panel, side chains; and left panel, backbone.

up to a given point of the path, of the RMS difference between the coordinates of successive pairs of structures along the path. It is normalised, so that $\lambda = 1$ represents the end of the reaction. The CPR algorithm is implemented in the Trajectory Refinement Algorithm (TRAvel) module of the CHARMM program.

Convergence of the whole path refinement is achieved when all the local energy maxima along the path are exact saddle points. All saddle-points were refined here until their RMS gradient was less than 10^{-4} kcal/(mol Å). The final path obtained consists of a sequence of 4153 structures, of which 189 are saddle points.

RESULTS

End State Structures. The conformational transition involves the movement of the loops over the protein surface, accompanied by a change in the volume of the domain by ∼13%, from 8860 Å^3 in TRP-OUT to 7860 Å^3 in TRP-IN. The root-mean-square deviation (RMSD) between the two end-states is 4.1 Å. Eighty residues are involved in the difference. The residue dependence of the RMSD (Figure 3) shows that the two regions where the structures differ most are residues 185−187 and residues 227−230, corresponding to two loops, between helices A and B and between helices D and E, respectively. In both loops, the RMSD is ∼8 Å for the backbone and ∼10−15 Å for the side chains. The Cα atoms of Glu228 and Trp187 move 12 and 9 Å, respectively.

The potential energy difference between the minimized end-states $\Delta E = E_{\mathrm{TRP-IN}} - E_{\mathrm{TRP-OUT}}$ is only 0.5 kcal/mol. However, further analysis involving the decomposition, for individual secondary structural elements, of the self-energy (all terms in the energy function representing interactions within the element) and interaction energy (the terms representing interactions of the element with the rest of the protein) reveals some interesting energetical differences between the two states. In particular, it was found that the sum of the self-energies of the individual secondary structure elements of domain III is more favorable in the TRP-OUT structure, $\Delta E^{\mathrm{III}}_{\mathrm{SELF}} = 29.2$ kcal/mol, and that this effect is closely compensated by the interaction energies between the secondary structure elements $\Delta E^{\mathrm{III}}_{\mathrm{INTE}} = -29.6$ kcal/mol.

In Figure 4 are shown the self- and interaction energy differences between the minimized end-state structures for each secondary structural element. Most elements exhibit compensation of the self- and interaction energies, although not always in the same direction, e.g., the terms for the DE loop ($\Delta E^{\mathrm{DE}}_{\mathrm{SELF}} = -6.2$ kcal/mol, $\Delta E^{\mathrm{DE}}_{\mathrm{INTE}} = 12.9$ kcal/mol) are of opposite sign to the overall effect. The AB loop has a particularly large self-energy difference, $\Delta E^{\mathrm{AB}}_{\mathrm{SELF}} = 17.6$ kcal/mol, favoring the TRP-OUT state, compensated by a large interaction energy favoring TRP-IN ($\Delta E^{\mathrm{AB}}_{\mathrm{INTE}} = -23.2$ kcal/mol). The major component of $\Delta E^{\mathrm{III}}_{\mathrm{SELF}}$ is the dihedral energy (20 kcal/mol) which is due mainly to strain on the AB loop and associated secondary structure elements in TRP-IN accompanying the burial of Trp187. The compensating interaction energy comes in part from strong hydrogen bonding between the $N_{\epsilon 1}$ atom of the Trp187 indole ring and the carbonyl and $O_{\gamma 1}$ oxygens of Thr224. A further contribution is from nonpolar interactions between the tryptophan side chain and residues in the cavity in which it is enclosed, i.e., two aromatic rings, Phe194 and Phe198, as well as Leu237 and Ile225. In addition, the strained AB conformation is stabilized by hydrogen bonds between backbone atoms (O of Ala182 with N of Trp187 and O of Leu185 with N of Gly188).

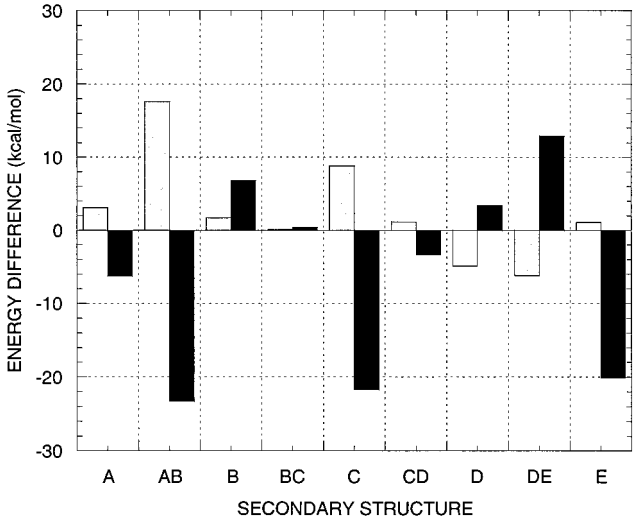

FIGURE 4: Changes in the self-energies (gray) and the interaction energies (black) of the individual structural elements of domain III between the end states, $E_{(\mathrm{TRP-IN})} - E_{(\mathrm{TRP-OUT})}$.

Conformational Pathway. The potential energy profile along the pathway is plotted in Figure 5A. It shows many

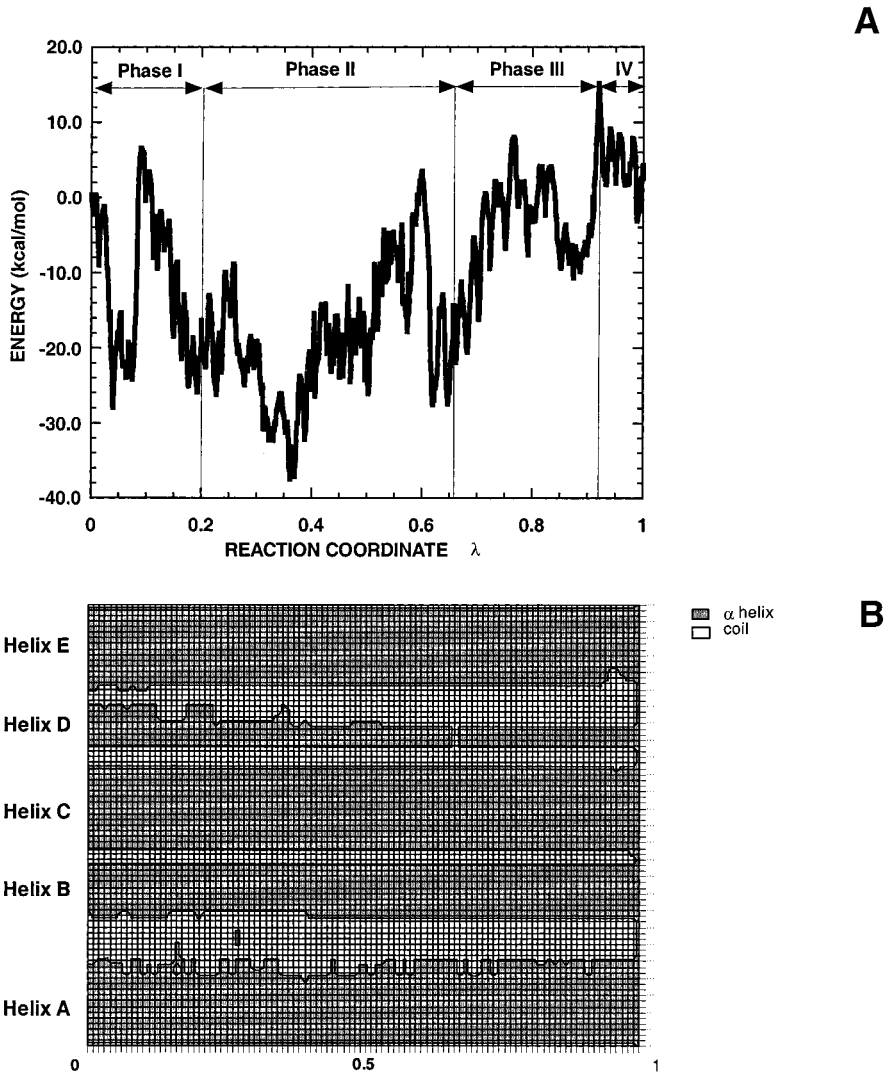

FIGURE 5: (A) Potential energy profile along the conformational pathway. The reaction coordinate λ is the sum of all successive coordinate RMS differences λ along the curvilinear path. It is normalised here with respect to the total path length so that $\lambda = 0$ represents the TRP-OUT conformation and $\lambda = 1$ represents the TRP-IN conformation (shown in Figure 1). (B) The evolution of the secondary structure along the pathway. On the *y*-axis each square corresponds to one residue, on the *x*-axis is the reaction coordinate λ. The tertiary structure (α helix, coil) was determined using MOLMOL program (*44*).

energy fluctuations locally along the reaction coordinate of ~5−15 kcal/mol together with slower energy variations of ~20−40 kcal/mol. The local fluctuations correspond to the energy barriers crossed during the many small discrete conformational rearrangements that are required during the transition, for example, side-chain reorientations.

Visual inspection of the pathway using molecular graphics revealed that it can be conveniently divided into four phases indicated in Figure 5A, delimited by the structures shown in Figure 6. Important associated hydrogen bonds are depicted in Figure 7. In the first phase, from $\lambda = 0-0.2$, Trp187 approaches the entrance of the cavity at the center of the domain, accompanied by deformation of the C-terminal part of helix A, which is adjacent to the loop containing Trp187. The entrance to the cavity is still blocked by the DE loop, which contains the side chain of Glu228, with which the Trp187 indole nitrogen forms a hydrogen bond (Figures 6A and 7A).

In the second phase ($\lambda = 0.2-0.65$) the DE loop opens the domain cavity (Figure 6B). This is achieved through the partial unravelling of the C-terminal turn of helix D. The hydrogen bond between Glu228 and the indole is broken. The insertion of Trp187 into the cavity is, however, still hindered, this time by the formation of a hydrogen bond between the indole and the side chain of Glu234 (Figure 7B). The third phase ($\lambda = 0.65-0.93$) starts by breaking this hydrogen bond, followed by a further rearrangement of the DE loop and of the C-terminal part of helix D. The partial insertion of Trp187 into the protein pocket takes place during this phase. However, insertion is once more halted, now by the formation of another hydrogen bond, between the indole and the side chain of Asp226 (Figures 6C and 7C). During the preceding phases, Asp226 had remained at the interface between domains III and IV, maintained by electrostatic interactions between its side chain and Glu234. In the third phase, this hydrogen bond breaks and Asp226 turns toward the domain cavity, allowing it to interact with Trp187 (Figure 7C). The structure at the end of the third phase is that of highest energy along the pathway.

In the final, fourth phase ($\lambda = 0.93-1.0$), the hydrogen bond with Asp226 breaks, and Trp187 completes its trajectory into the protein cavity, where it forms hydrogen bonds

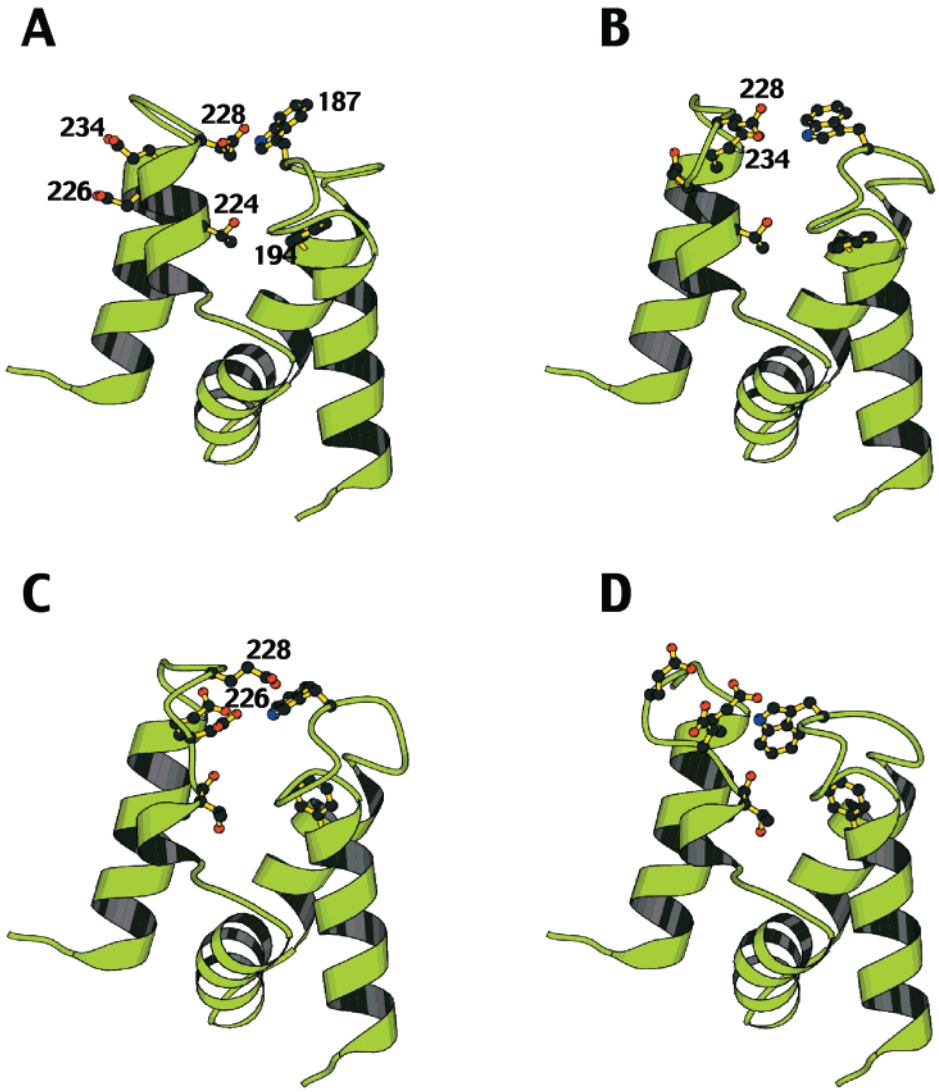

FIGURE 6: Structures at successive stages along the pathway (see also Figure 1 for the two end-states of the reaction, $\lambda = 0$ and $\lambda = 1$). The Trp187, Thr224, Asp226, Glu234, Glu228, and Phe194 side chains are shown in a ball-and-stick representation [figure produced using Molscript (*39*)]. (A) $\lambda = 0.2$, end of phase I, (B) $\lambda = 0.65$, end of phase II, (C) $\lambda = 0.92$ end of phase III and the highest energy structure, (D) $\lambda = 0.94$, start of phase IV.

with the oxygens of Thr224 and the remaining interactions of the TRP-IN state. The DE loop moves to the interface of domains III and IV, and the formation of a hydrogen bond between Thr229 and Asp226 stabilizes the loop conformation.

During the transition, collective motion of the whole domain was found to occur. This involves a contraction and expansion of the domain, performed by the four parallel helices A, B, D, and E moving toward and out of the center of the domain, accompanied by a torsion of helix C. The conformational transition also involves considerable variations in the secondary structure, the evolution of which along the pathway is plotted in Figure 5B. The fluctuations of the helix lengths permit the motion of the loops, and thus facilitate the conformational change. The largest changes are observed in helix D which, at one point of the pathway (at $\lambda = 0.66$), even disappears completely. The second most fluctuating helix is A, which transiently loses about 20% of its length. Helices B and E show minor fluctuations in length along the pathway, while helix C is the most stable.

Over 100 dihedral transitions (defined here as involving torsional variations of at least 60°) were observed along the pathway. Table 1A summarises these for the backbone in which 106 transitions were found, involving 21 dihedrals all confined to loops AB (76 transitions) and DE (30 transitions). Gly188 φ alone undergoes as many as 15 transitions, Thr189 φ undergoes 12, and Asp190 ψ 11. A total of 241 side-chain dihedral transitions were also found.

DISCUSSION

The motions involved in the calculated annexin V conformational transition are complex, involving transient complete unwinding of an α-helix (D) and the lid-closing of the Trp-bearing loop AB. They include hundreds of dihedral transitions of backbone and side-chain torsion angles. The path complexity, of similar magnitude to that recently derived using CPR for the escape of the prosthetic group from the retinoic acid receptor (*42*), is such that potential errors should be carefully considered. First, computational limitations do not permit all the possible pathways for large-scale conformational change in proteins to be found. Therefore, the pathway presented here should be considered as a plausible, energetically reasonable estimate. Second, free energy effects along the pathway cannot easily be included.

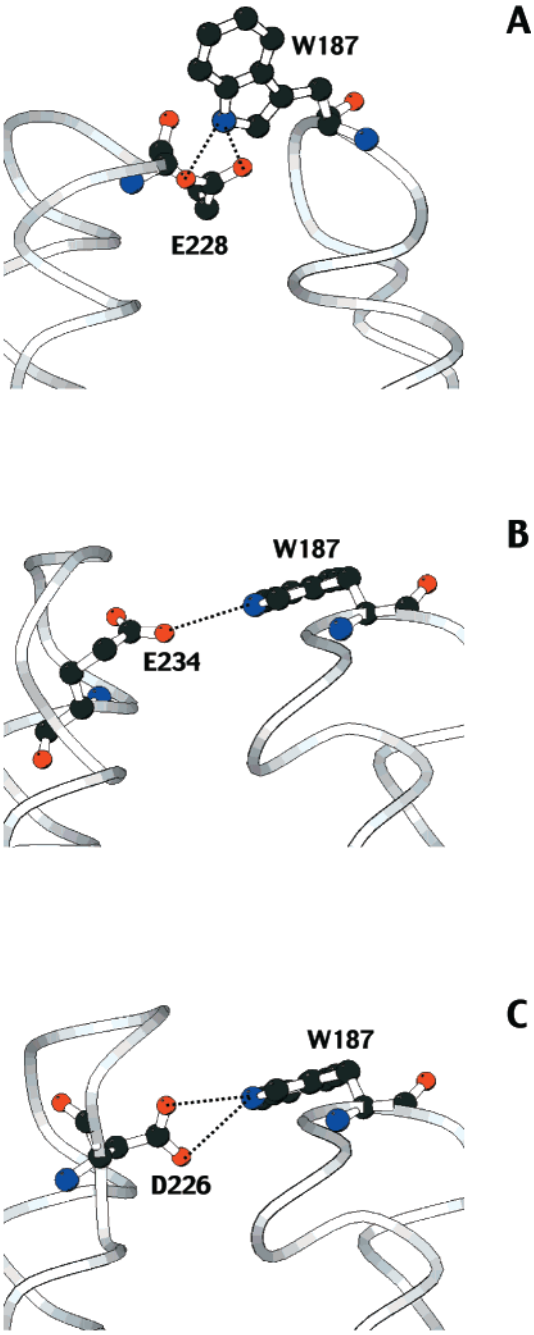

FIGURE 7: Interaction of Trp187 with acidic residues along the pathway. The Trp187, Asp226, Glu234, and Glu228 are shown in a ball-and-stick representation [figure produced using Molscript (*39*)]. (A) Interaction between Glu228 and Trp187 (phase I), (B) interaction between Glu234 and Trp187 (phase II), (C) interaction between Asp226 and Trp187 (phase III).

As is usual for this type of calculation, the ~20−40 kcal/mol potential energy variations along the pathway in Figure 5A are considerably larger than the expected variation in free energy along the path if the transition were to occur on a physiologically relevant time-scale. Consequently, the exact location and height of the energy barriers must be interpreted with caution. The potential energy changes may be somewhat overestimated, due to the use of a simplified solvation correction by charge scaling, which assumes that all solvent shielding effects remain constant along the path. Moreover, the potential energy fluctuations are likely to be compensated by entropic effects. Due to the disordering involved in the partial unwinding of secondary structures, the associated changes of the conformational entropy are significant. Solvent entropy changes along the pathway may also play a role.

Table 1: Torsion Angle Transitions Observed along the Path

(A) All 106 Backbone Dihedral Angle Transitions

AB loop				DE loop			
residue	dih. angle	no. of trans.	net[b] rotation	residue	dih. angle	no. of trans.	net rotation
Ala182	φ	4	0	Ile225	ψ	1	−120
Ala182	ψ	6	0	Ile225	ω	2	0
Leu185	ψ	2	0	Asp226	φ	4	0
Lys186	φ	2	−120	Asp226	ψ	4	120
Lys186	ψ	5	120	Arg227	φ	6	0
Trp187	φ	5	−120	Arg227	ψ	2	0
Trp187	ψ	8	0	Glu228	φ	2	0
Gly188	φ	15	−120	Ser230	ψ	5	120
Gly188	ψ	4	−120	Gly231	φ	4	−120
Thr189	φ	2	0				
Thr189	ψ	12	0				
Asp190	φ	11	0				

(B) Side-Chain Torsion Angle Transitions of Acidic Amino Acids

residue	tors. angle	no.[a] of trans.	net[b] rotation
Glu169	$C_\alpha-C_\beta$	1	120
Glu169	$C_\beta-C_\gamma$	2	120
Glu169	$C_\gamma-C_\delta$	1	120
Glu184	$C_\gamma-C_\delta$	2	0
Asp190	$C_\alpha-C_\beta$	1	−120
Glu191	$C_\gamma-C_\delta$	3	120
Glu192	$C_\alpha-C_\beta$	13	−120
Asp211	$C_\beta-C_\gamma$	4	0
Glu222	$C_\gamma-C_\delta$	4	0
Glu226	$C_\alpha-C_\beta$	4	0
Glu226	$C_\beta-C_\gamma$	3	−120
Glu228	$C_\alpha-C_\beta$	1	120
Glu228	$C_\beta-C_\gamma$	6	0
Glu228	$C_\gamma-C_\delta$	4	120
Glu234	$C_\alpha-C_\beta$	1	−120
Glu234	$C_\beta-C_\gamma$	1	120
Glu234	$C_\gamma-C_\delta$	3	−120

[a] Every relative angle change by 60° is counted as one transition. [b] The net rotation corresponds to the angle difference between the end states.

However, even taking into account the above limitations, certain structural and energetical properties of the pathway clearly emerge. One of these is the build up of conformational strain, compensated by improved interaction energies as Trp187 is buried. The interactions between the DE and AB loops at the end of phase III are strong enough to cause considerable strain in both loops, inducing deviations of some peptide bonds by more than 20° from planarity (Lys186, Trp187, Gly188, Asp190, Asp226, and Glu228) and that of Asp190 by 35°. One peptide bond (Ile225) changes its conformation from trans to cis at $\lambda = 0.40$, and returns to *trans* at $\lambda = 0.90$, two transitions present in the initial path generated by PEDC. The gas-phase barrier in the present force field for a peptide group *cis*−*trans* transition is 19kcal/mol. The calculated pathway suggests that strain energy of this magnitude can be accommodated by interaction energy compensation—this is indeed found to reach ~20 kcal/mol for a single secondary structural element, the AB loop. Whether the peptide group actually flips (and in large-scale protein conformational transitions in general) or whether the

flexibility of the protein backbone is sufficient to avoid this energetic cost in an alternative, and perhaps even more complicated, pathway will be the subject of future studies.

It is important that any theoretical or computational study should give suggestions for further experiments which can test or be used to improve the calculations. The inaccessibility to experiment of molecular geometries connecting stable states renders direct experimental testing of conformational pathways difficult. However, the present results do raise a number of questions that can be examined experimentally. First, the question arises as to what extent the conformational changes in the present calculation are also sampled during experimentally detectable equilibrium thermal fluctuations of the end states. Results obtained by time-resolved fluorescence decay on Trp187 (*2*, *15*) showed that for annexin V both in the presence and absence of calcium, three motions exist: the Brownian motion of the entire molecule (correlation time $\approx$ 14 ns), the fast rotation motion of the indole ring around its C_α-C_β and C_β-C_γ bond (correlation time $\sim$200−800ps) and another unidentified intermediate rotational motion characterized by a correlation time of about 1−5 ns. During the present calculated transition collective motion involving contraction and expansion of the whole domain occurs. This collective motion may be related to the as yet unidentified 1−5 ns dynamics. Further examination of this would require computationally expensive extensive MD simulation of the end states. Both in free and in calcium-bound annexin V the fluorescence decay is described by three lifetime populations, showing that Trp interacts with three different chemical environments in slow exchange on the fluorescence time scale. The possibility exists that this conformational heterogeneity may also be related to the existence of collective motions and the accompanying distinct transitions of Trp187 found in the present pathway calculation.

Experimental results have shown that the mechanism of the conformational change is pH-dependent (*2*, *15*). It is therefore interesting to examine in more detail the role of charged residues in the conformational pathway. We find that certain acidic residues play crucial roles at various stages. There is a total of 13 aspartic and glutamic acids. Side-chain dihedral transitions along the pathway were observed for 10 of these residues (see Table 1B), of which seven have a final conformation that is different from the start. Four undergo more than 10 transitions along the pathway (Asp190, Glu192, Asp226, and Glu228) and three of these four residues—Asp190, Asp226, and Glu228—also undergo main-chain transitions (Table 1A). Moreover, the insertion of the indole moiety of Trp187 into the protein cavity is blocked by hydrogen-bonding interactions with three successive acidic residues, Glu234, Glu228, and Asp226, depicted in Figure 7.

The above considerations based on the calculated pathway indicate that Asp226 and Glu228, Asp190, Glu192, and Glu234 play important roles in the conformational change. Consequently, changing the protonation states of these residues might be expected to affect the pathway, by modifying their hydrogen-bonding and other electrostatic properties. With respect to Asp226 it is interesting that an early fluorescence spectroscopy study of annexin III showed that its domain III, which also contains a Trp residue in the AB loop, remains in the TRP-OUT conformation even in the absence of calcium (*20*). Comparison of the sequence of these two annexins reveals a lysine at position 226 in annexin III, i.e., a residue with opposite charge to that in annexin V.

Fluorescence spectroscopic work is underway on mutants of annexin V chosen in the light of the present pathway calculation. Of the acidic residues mentioned above, only Glu228 cannot be replaced without a perturbation of the calcium binding site in this domain (and with it probably the physiological function of annexin V). Glu228 is one of the calcium-ion ligands in domain III, and it has been shown experimentally that its presence is a necessary condition for the existence of a calcium-binding site (*8*). Initial results on the mutant D226K show that Trp187 is exposed to the solvent both in the absence of calcium and at neutral pH (Sopkova et al., submitted for publication), confirming the important role of Asp 226 in the pathway mechanism. The spectroscopic characteristics of this mutant at pH 4 also indicate the participation of further acidic residue(s) in the conformational change. The conformational change is induced by high concentrations of calcium or by acidic pH (*2*, *15*). The question arises as to whether the effect of calcium ions at high concentration is nonspecific or due to direct interaction with particular residues on the protein. The nonspecific effect could be via a general influence on the electrostatic interactions on the surface of the protein and associated protonation states. In this case both the pH and calcium dependence would act via protonation and deprotonation of acidic residues. Experiments to test this possibility using the mutations D190K, D192K, and E234K, are underway.

ACKNOWLEDGMENT

We thank Jacques Gallay, LURE, Orsay, for helpful discussions.

NOTE ADDED AFTER ASAP POSTING

The fourth sentence of the text on the first page was in error when this article was released ASAP on October 24, 2000. The text incorrectly read: “All annexins contain a 14- or 8-fold repeat...”. It now reads correctly: “All annexins contain a 4- or 8-fold repeat...”.

REFERENCES

1. Fischer, S., and Karplus, M. (1992) *Chem. Phys. Lett. 194*, 252−261.
2. Sopkova, J., Vincent, M., Takahashi, M., Lewit-Bentley, A., and Gallay, J. (1998) *Biochemistry 37*, 11962−11970.
3. Beermann, B. B., Hinz, H.-J., Hofman, A., and Huber, H. (1998) *FEBS Lett. 423*, 265−269.
4. Raynal, P., and Pollard, H. B. (1994) *Biochim. Biophys. Acta 1197*, 63−93.
5. Swairjo, M. A., and Seaton, B. A. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 193-213.
6. Gerke, V., and Moss, S. (1997) *Biochim. Biophys. Acta 1357*, 129−154.
7. Huber, R., Berendes, R., Burger, A., Schneider, M., Karshikov, A., Luecke, H., Römisch, J., and Paques, E. (1992) *J. Mol. Biol. 223*, 683−704.
8. Weng, X., Luecke, H., Song, I. S., Kang, D. S., Kim, S.-H., and Huber, R. (1993) *Protein Sci. 2*, 448−458.
9. Lewit-Bentley, A., Morera, S., Huber, R., and Bodo, G. (1992) *Eur. J. Biochem. 210*, 73−77.
10. Bewley, M. C., Boustead, C. M., Walker, J. H., Waller, D. A., and Huber, R. (1993) *Biochemistry 32*, 3923−3929.

11. Sopkova, J., Renouard, M., and Lewit-Bentley, A. (1993) *J.Mol. Biol. 234*, 816−825.
12. Concha, N. O., Head, J. F., Kaetzel, M. A., Dedman, J. R., and Seaton, B. A. (1993) *Science 261*, 1321−1324.
13. Berendes, R., Vosges, D., Demange, P., Huber, R., and Burger, A. (1993) *Science 262*, 427−430.
14. Follenius-Wund, A., Piémont, E., Freyssinet, J.-M., Gérard, D., and Pigault, C. (1997) *Biochem. Biophys. Res. Commun 234*, 111−116.
15. Sopkova, J., Vincent, M., Takahashi, M., Lewit-Bentley, A., and Gallay, J. (1999) *Biochemistry 38* (17), 5447−58.
16. Meers, P. (1990) *Biochemistry 29*, 3325−3330.
17. Meers, P., and Mealy, T. (1993) *Biochemistry 32*, 5411−5418.
18. Sopkova, J., Gallay, J., Vincent, M., Pancoska, P., and Lewit-Bentley, A. (1994) *Biochemistry 33*, 4490−4499.
19. Burger, A., Berendes, R., Liemann, S., Benz, J., Hofmann, A., Göttig, P., Huber, R., Gerke, V., Thiel, C., Römisch, J., and Weber, K. (1996) *J. Biol. Chem. 257*, 839−847.
20. Favier-Perron, B., Lewit-Bentley, A., and Russo-Marie, F. (1996) *Biochemistry 35*, 1740−1744.
21. Benz, J., Bergner, A., Hofmann, A., Demange, P., Göttig, P., Liemann, S., Huber, H., and Voges, D. (1996) *J. Mol. Biol. 260*, 638−643.
22. Luecke, H., Chang, B. T., Maillard, W. S., Schlaepfer, D. D., and Haigler, H. T. (1995) *Nature 378*, 512−515.
23. Zanotti, G., Malpeli, G., Gliubich, F., Folli, C., Stoppini, M., Olivi, L., Savoia, A., and Berni, R. (1998) *Biochem. J. 329*, 101−106.
24. Gerstein, M., Lesk, A., and Chothia, C. (1994) *Biochemistry 33* (22), 6739−6749.
25. Schlitter, J., Engels, M., and Kruger, P. (1994) *J. Mol. Graphics 12*, 84−89.
26. Elber, R. and Karplus, M. (1987) *Chem. Phys. Lett. 139*, 375−380.
27. Wroblowski, B., Diaz, J. F., Schlitter, J., and Engelborghs, Y. (1997) *Protein Eng. 10*, 1163−1128.
28. Ech-Cherif el-Kettani, M. A., and Durup, J. (1992) *Biopolymers 32* (5), 561−574.
29. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
30. CCP4. (1994) *Acta Crystallogr., Sect. D 50*, 760−763.
31. Neria, E., Fischer, S., and Karplus, M. (1996) *J. Chem. Phys. 105*, 1902−1921.
32. Lewit-Bentley, A., Bentley, G. A., Favier, B., L'Hermite, G., and Renouard, M. (1994) *FEBS Lett. 345*, 38−42.
33. Fischer, S., Michnick, S., and Karplus, M. (1993) *Biochemistry 32*, 13830−13837.
34. Caflisch, A., Fischer, S., and Karplus, M. (1997) *J. Comput. Chem. 18*, 723−743.
35. Davis, M. E., Madura, J. D., Luty, B. A., and McCammon, J. A. (1991) *Comput. Phys. Commun. 62*, 187−197.
36. Mohan, V., Davis, M. E., McCammon, J. A., and Pettitt, B. M. (1992) *J. Phys. Chem. 96*, 6428.
37. Guilbert, C., Perahia, D., and Mouawad, L. (1995) *Comput. Phys. Commun. 91*, 263−273.
38. Fischer, S., Dunbrack. R., and Karplus. M. (1994) *J. Am. Chem. Soc. 116*, 11931−11937.
39. Fischer, S. Grootenhuis, P. D. J., Groenen, L. C. vanHoorn, W. P., vanVeggel, F., Reinhoudt D., and Karplus. M. (1995) *J. Am. Chem. Soc. 117*, 1611−1620.
40. Verma, C. S., Fischer, S., Caves, L., Roberts G. C. K., and Hubbard R. E. (1996) *J. Phys. Chem. 100*, 2510−2518.
41. Fischer, S., Verma C. S., and Hubbard, R. E. (1998) *J. Phys. Chem.* 102, 1797−1805.
42. Blondel, A., Renaud, J. P., Fischer, S., Moras, D., and Karplus, M. (1999) *J. Mol. Biol. 291*, 101−115.
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
44. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.

BI000659H