Mechanics and dynamics of B1 domain of protein G: role of packing and surface hydrophobic residues

MA Ceruso, A Amadei and A Di Nola

Protein Sci. 1999 8: 147-160
Access the most recent version at doi:10.1110/ps.8.1.147

Protein Science (1999), *8*:147–160. Cambridge University Press. Printed in the USA.

Mechanics and dynamics of B1 domain of protein G: Role of packing and surface hydrophobic residues

MARC A. CERUSO, ANDREA AMADEI, AND ALFREDO DI NOLA
Department of Chemistry, University of Rome "La Sapienza," Rome 00185, Italy

(RECEIVED July 1, 1998; ACCEPTED September 23, 1998)

Abstract

The structural organization of the B1 domain of streptococcal protein G (PGA) has been probed using molecular dynamics simulations, with a particular emphasis on the role of the solvent exposed Ile6 residue. In addition to the native protein (WT-PGA), three single-mutants (I6G-PGA, I6F-PGA, and I6T-PGA), one double-mutant (I6T,T53G-PGA), and three isolated peptide fragments (corresponding to the helix and the two β-hairpins) were studied in the presence of explicit water molecules. Comparative analysis of the various systems showed that the level of perturbation was directly related to the hydrophobicity and the size of the side chain of residue 6, the internal rigidity of the proteins decreasing in the order I6T-PGA > I6G-PGA > WT-PGA > I6F-PGA. The results emphasized the importance of residue 6 in controlling both the integrity of the sheet's surface and the orientation of the helix in relation to the sheet by modulation of surface/core interactions. The effects of mutations were delocalized across the structure, and glycine residues, in particular, absorbed most of the introduced strain. A qualitative structural decomposition of the native fold into elementary building-blocks was achieved using principal component analysis and mechanical response matrices. Within this framework, internal motions of the protein were described as coordinated articulations of these structural units, mutations affecting mostly the amplitude of the motions rather than the structure/location of the building-blocks. Analysis of the isolated peptidic fragments suggested that packing did not play a determinant role in defining the elementary building-blocks, but that chain topology was mostly responsible.

Keywords: alpha-helix; beta-sheet; hydrophobic residues; protein folding; protein packing

Defining the relationship between an amino acid sequence and the three-dimensional configuration of the corresponding protein remains one of the great challenges in biology. There is now increasing evidence that the three-dimensional organization of proteins is hierarchical, arising from the packing or juxtaposition of a limited number of variably-sized components (Richardson, 1985; Finkelstein & Ptitsyn, 1987; Efimov, 1994; Doolittle, 1995; Brenner et al., 1997). Thus, a possible way of addressing the protein folding problem is to identify the elements that characterize a given protein fold and to determine how these elements interact with each other to stabilize or produce the final native conformation. Some of the principles that control the relative arrangement of secondary structure elements were first put forth by Crick (1953), who proposed a "knobs into holes" model to describe helical coiled-coils. As the number of solved proteins grew in size and variety, other well-defined geometrical patterns began to emerge, and various models were advanced to rationalize α-helix/α-helix, β-sheet/β-sheet, and finally α-helix/β-sheet packing (Chothia et al., 1977, 1981; Chothia & Janin, 1981). It is this latter type of packing that is explored in the present study.

In the case of α/β proteins, Chothia and coworkers used the "ridges into groves" description of secondary structure packing, and proposed that adjacent rows of residues on the α-helix (i, $i+4$, $i+8\ldots$ and $i+1$, $i+5$, $i+9\ldots$) would pack against the smooth surface of the β-sheet, with the right-handed twists of the helix and the sheet efficiently complementing each other (Chothia et al., 1977; Janin & Chothia, 1980; Chothia, 1984; Chothia & Finkelstein, 1990). Contrasting with this smooth "twist-complementarity model," a specific "interdigitation pattern" of side chains was suggested (Cohen et al., 1982), where four residues on the α-helix ($i+1$, $i+4$, $i+5$, $i+8$) would enclose one of the amino acids on the β-sheet. Both models successfully explained the vertical ($\Omega \sim [-20,+20]$) and parallel ($\Theta \sim [0,+20]$) packing angles of α-helices onto β-sheets (Janin & Chothia, 1980), but presented opposite views on the specificity of side-chain packing at the helix/sheet interface. Another important aspect highlighted by the models was the relationship between the orientation of the helix and the

Reprint requests to: Marc Antoine Ceruso, Chemistry Department, University of Rome "La Sapienza," P. le Aldo Moro 5, Rome 00185, Italy; e-mail: mceruso@signac.chem.uniroma1.it.

Abbreviations: ANTIHB, antiparallel hydrogen bonding; MD, molecular dynamics; PARAHB, parallel hydrogen bonding; PGA, streptococcal protein G; R_G, radius of gyration; RMSD, RMS deviation; RMSF, RMS fluctuation; RMSIP, RMS inner product; SAS, solvent accessible surface; WT, wild-type; WT-PGA, wild-type streptococcal protein G.

concavity of the sheet (Chothia et al., 1977; Janin & Chothia, 1980; Cohen et al., 1982). These geometrical packing preferences were later quantified using conformational energy computations to confirm that the helix preferably lies parallel to the concave diagonal, maximizing its contact surface with the β-sheet (Chou et al., 1985).

Inspired by these approaches, we have chosen to study a simple α/β-sandwich protein, corresponding to one of the domains of protein G: entry 1pga (PGA) in the Brookhaven Protein Data Base (Gronenborn et al., 1991; Gallagher et al., 1994). The "architectural" nature of the biological function (as opposed to enzymatic), the small size, the remarkable stability, and the absence of prosthetic groups or disulfide bridges make such domains of protein G prime candidates for the study of secondary structure packing and protein folding. The tertiary structures of different IgG-binding domains of protein G have been resolved by NMR and X-ray crystallography, either as isolated domains (Gronenborn et al., 1991; Achari et al., 1992; Lian et al., 1992; Gallagher et al., 1994), or bound to an Fab fragment of mouse IgG (Derrick & Wigley, 1992; Lian et al., 1994), or bound to an Fc fragment of human IgG (Gronenborn & Clore, 1993; Lian et al., 1994). All IgG-binding domains of protein G are found to consist of a single α-helix packed against a four-stranded mixed β-sheet (Fig. 1). Among the reasons invoked for their remarkable stability, the tight and efficient packing of the core has received particular attention (Gronenborn et al., 1991; Achari et al., 1992; Alexander et al., 1992; Orban et al., 1995). However, reconstitution experiments showed that isolated fragments of the sequence, corresponding to the two hairpins and the helix, failed to interact with each other (Blanco & Serrano, 1995). In addition, a simple I6G mutation on the solvent exposed face of the sheet was reported to destabilize the protein dramatically (Smith et al., 1994).

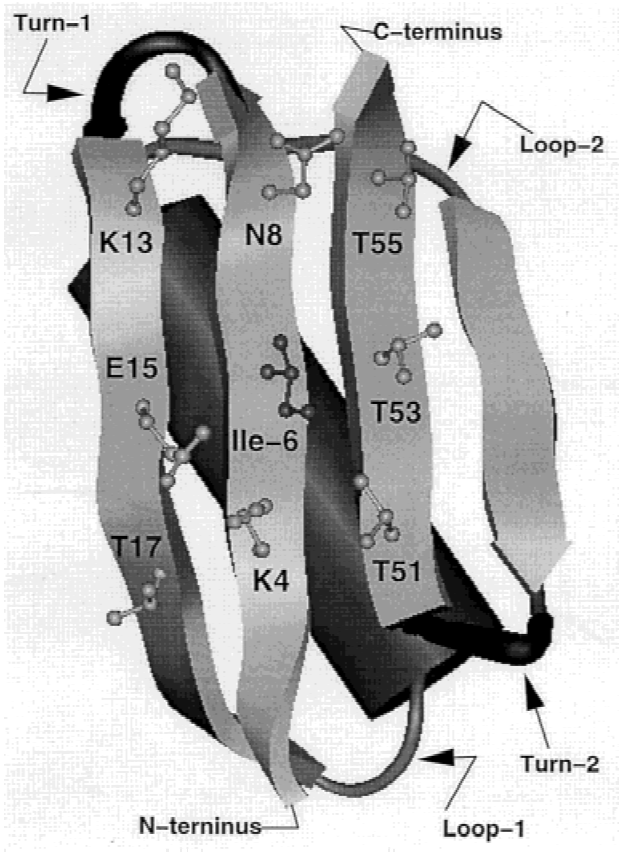

Fig. 1. Cartoon representation of the B1 domain of streptococcal protein G. The side chains of the residues that form the hydrophobic surface cluster around isoleucine-6 have been labeled.

In the current study, we have been interested in determining how well molecular dynamics (MD) simulations of PGA reflected the established packing principles of α/β folds, but also in extending such principles by characterizing the mechanical and dynamical relationships between the elementary building-blocks of PGA. In particular, we have sought to establish whether the building-blocks determined by MD corresponded to regular secondary structure elements, or whether such "rigid" sets reflected some local packing interactions or the overall fold's topology. Finally, we have attempted to interpret these mechanical and dynamical features in terms of protein stability and biological function. Consequently, we have studied the effects of a systematic series of mutations on the dynamics of wild-type PGA. Mutations were introduced on the solvent exposed side of the sheet, with the purpose of analyzing the mechanical and dynamical role of isoleucine-6 in the protein in general and the sheet in particular. This isoleucine was of particular interest since Smith et al. (1994) had reported that its mutation to a glycine lead to substantial destabilization of the protein, and Tisi and Evans (1995) had suggested a possible role in protein stability for such hydrophobic residues. The role of hydrophobic bulk was tested using I6G and I6F single amino acid mutations, while the effect of polarity was investigated using I6T and [I6T,T53G] single and double mutants. The combined use of mechanical response matrices (Wong et al., 1993; Chillemi et al., 1997) and principal component analysis of positional fluctuations (Garcia, 1992; Amadei et al., 1993) have afforded a description of the B1 domain of protein G as an assembly of elementary building-blocks, closely related to secondary structure elements and separated from each other by articulation points. Comparative analysis of native and "virtual" mutant simulations helped the refining of the three-dimensional organization of this α/β fold and emphasized its mechanical relationships. In particular, the results highlighted the importance of the surface residue Ile6, in controlling the integrity of the packing within the protein core, the flexibility of the first hairpin-turn and the sheets regularity (through surface interstrand interactions). The relative "rigid-body" motions of the helix and the sheet were consistent with literature data on α/β proteins in general, and protein G domains in particular. The specific role of packing in defining articulations and building-blocks was addressed by simulating isolated fragments of PGA's sequence, in the manner of Blanco and Serrano (1995). The results are discussed in relation to protein organization; details of our work could prove useful for the de novo design of α/β proteins.

Results

Isoleucine-6 and protein architecture

As a starting point, we chose to focus on the role of the hydrophobic residue Ile6, which is centrally located on the solvent exposed face of the sheet, and secluded in a protective cage formed by eight polar residues (Fig. 1). To assess the structural significance of this residue and/or specific location in the sequence, we have analyzed the molecular dynamics trajectories of wild-type PGA (WT-PGA) and three single point mutants at position 6 (P6):

I6G-PGA, I6F-PGA, and I6T-PGA. Finally, to clarify the effects seen in the I6T mutant, the double mutant [I6T,T53G]-PGA was studied as well.

Global structural stability

The overall structural stability of the proteins was confirmed by the fairly stable values in solvent accessible surface (SAS), radius of gyration (R_G, data not shown), and RMS deviation (RMSD) as a function of time (Fig. 2). In addition, for all systems, the mean solvent accessible surface area and the mean radius of gyration were comparable to the corresponding values in the X-ray structure (Table 1). Although further characterization is necessary, it is interesting to note that the behavior of the various mutants is consistent with their respective steric hindrance: the smaller and more flexible glycine having wider amplitude oscillations, the bulkier phenylalanine "paralyzing" the system, and the intermediate-sized but polar threonine deforming more profoundly the original system.

Geometry and hydrogen bonding

To gain further insight into the behavior of the individual proteins, a number of geometrical properties were monitored in the 600–1,600 ps time range (the first 600 ps of simulation were discarded to ensure that calculated parameters reflect the intrinsic properties of each system). The total number of hydrogen bonds remained fairly constant in all simulations and it was, on average, slightly larger than that found in the X-ray structure, reflecting again the stability of the simulated systems and indicating that

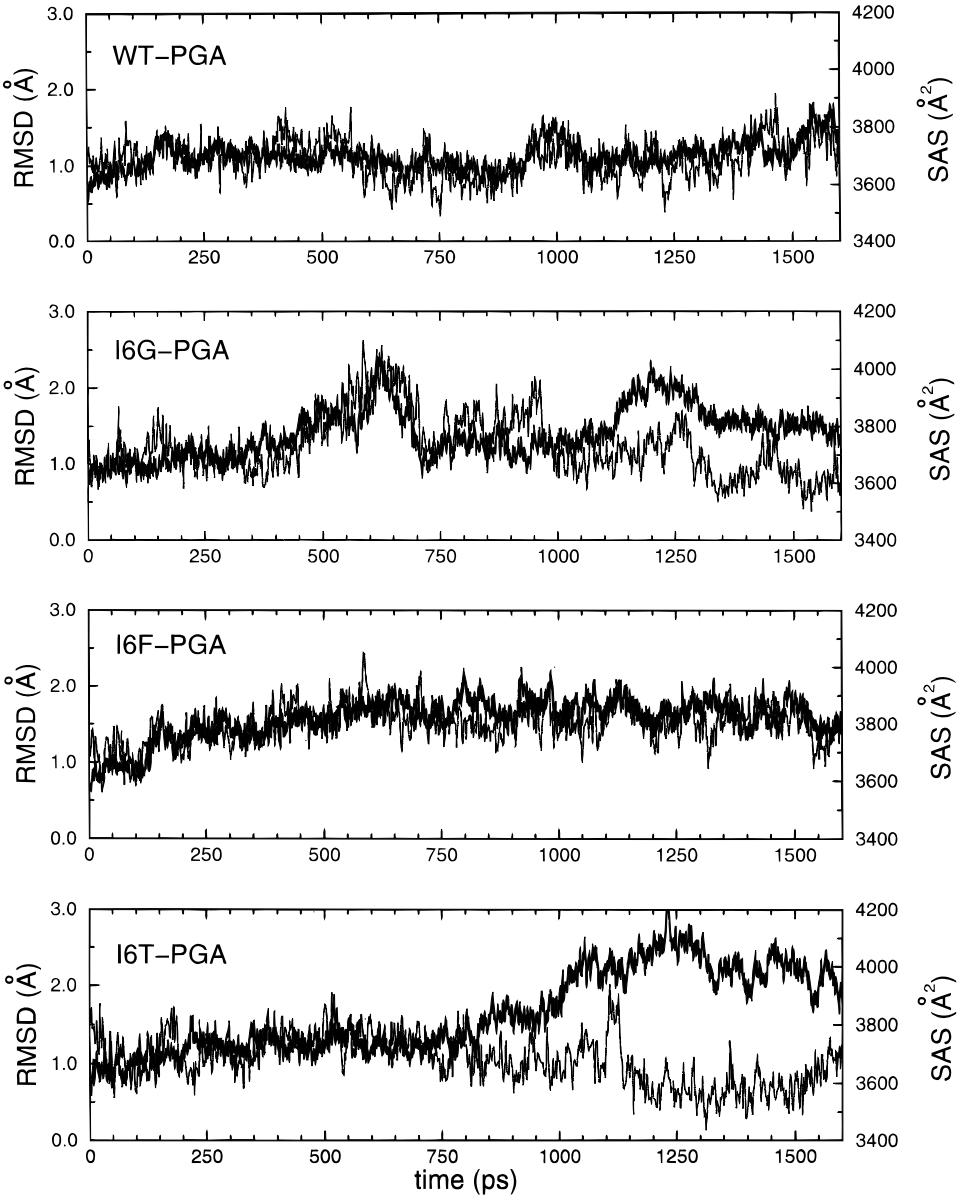

Fig. 2. Global conformational reporters. Left-hand side *y*-axis shows the time evolution of the RMSD of C-alpha atoms with respect to the crystal structure (black curve). Right-hand side *y*-axis shows time evolution of the total solvent accessible surface area (gray curve).

Table 1. *Structural statistics*[a] *(600.0–1,600.0 ps)*

System[b]	SAS (Å^2)	R_G (Å)	HBO	PARAHB	ANTIHB	HB4
X-ray	3,721	10.5	40	6	14	12
WT-PGA	3,677 (72)	10.4 (0.1)	44.3 (2.7)	5.8 (0.6)	12.0 (1.1)	13.7 (0.5)
I6G-PGA	3,733 (119)	10.4 (0.1)	44.3 (3.3)	5.8 (0.5)	12.4 (1.4)	12.9 (1.1)
I6F-PGA	3,812 (61)	10.5 (0.1)	44.8 (2.8)	5.9 (0.4)	12.2 (1.2)	12.7 (0.9)
I6T-PGA	3,649 (83)	10.4 (0.1)	41.7 (3.4)	5.3 (1.1)	9.1 (2.6)	12.3 (1.4)
I6T,T53G-PGA	3,695 (66)	10.4 (0.1)	42.4 (3.7)	5.7 (0.8)	12.6 (1.3)	12.1 (1.5)

[a] Properties were evaluated every 10 ps within the 600–1,600 ps time range. All statistics were obtained using DSSP (Kabsch & Sander, 1983). The radius of gyration was calculated using WHATIF (Vriend, 1990). All reported values are average values, numbers in parenthesis refer to standard deviations.
[b] SAS, GYR, HBO, PARAHB, ANTIHB, and HB4 designate, respectively, the protein's solvent accessible area, radius of gyration, the total number of hydrogen bonds in the protein, in parallel bridges, in antiparallel bridges, and finally the total number of hydrogen bonds of type $O_i \rightarrow HNi + 4$.

none of the systems underwent significant unfolding. For most systems (WT, I6G, and I6F), changes in secondary structure affected mostly antiparallel beta-strand interfaces (see ANTIHB values), lost BETA residues adopting either random-coil or turn and bend-like conformations (Table 2). In contrast, the number of hydrogen bonds across parallel strands (strands 1 and 4) remained close to that found in the X-ray reference conformation. These differences probably reflect the fact that antiparallel strands are located at the edge of the sheet and are consequently more solvent exposed than the central parallel strands. In addition, the hydrogen bonding partners for the edge strands are located on the protein termini, which are usually quite flexible. The time evolution of the secondary structure for each system (data not shown) confirmed that losses of structure (mostly extended beta-sheet conformation) occurred by fraying of the extremities of the first-β-hairpin and of the N-terminus of the second β-hairpin (around the 40th residue). The case of the I6T-mutant was again different from that of the other systems. For this mutant, both parallel (PARAHB) and antiparallel hydrogen bonding (ANTIHB) was affected, and the total loss in beta-sheet conformation was greater than for the rest of the systems (ANTIHB and BETA). The principal loss of secondary structure, for I6T-PGA, occurred in the third strand of the sheet and loop-2 (away from the site of mutation). Furthermore, important distortions at loop-2 and and strand-3 affected the C-terminus section of the helix, which extended into a turn-like conformation.

Distribution of strain

The total number of ("strained") residues lying outside favorable regions of the Ramachandran map (Ramachandran et al., 1963) was measured, using amino acid ϕ,ψ-propensities reported by Swindells et al. (1995). The number of "strained" dihedrals was low for all simulations (3.0 ± 2.3 in average, Table 2), and essentially the same five residues were concerned in all proteins: the four glycine residues (Gly9, Gly14, Gly38, and Gly41) and to a much lesser extent Lys50 in the second β-hairpin turn. Lys50 already lies on the right-hand side of the Ramachandran map in the crystal structure, and it is not uncommon to find glycines in "unfavorable" regions of the Ramachandran map since the absence of side chain confers greater flexibility to such residues. In comparison to wild-type all P6 mutants introduced an additional strain at Gly9. The frequency of deformation at this site increased with the size of side chain at P6: Phe $>$ Thr $>$ Gly. Interestingly, the marked distortion at Gly9 introduced by the I6F mutation was counterbalanced by a significant relief at Gly38.

Polarity and integrity of the sheet

At this stage, to evaluate a possible pathway for the effects of the I6T mutant, we analyzed the trajectory of the [I6T,T53G] double mutant construct. This double mutant was designed to attenuate interactions across the central strands (see Discussion), since in the I6T sequence, residue-6 faces three consecutive threonines on

Table 2. *Structural statistics*[a] *(600.0–1,600.0 ps)*

System[b]	ALPHA	BETA	TURN	COIL	STRAINED
X-ray	14	24	10	8	1
WT-PGA	15.0 (0.6)	20.6 (1.8)	9.8 (1.2)	10.6 (2.0)	2.6 (0.8)
I6G-PGA	14.9 (0.8)	21.9 (1.9)	10.1 (1.4)	9.2 (2.0)	2.3 (1.0)
I6F-PGA	14.1 (0.8)	21.6 (1.6)	11.0 (1.3)	9.2 (1.8)	3.3 (0.8)
I6T-PGA	14.0 (1.4)	17.8 (3.5)	10.4 (1.7)	13.7 (3.7)	3.2 (1.0)
I6T,T53G-PGA	13.4 (1.7)	21.5 (2.2)	11.5 (2.2)	9.5 (2.6)	3.5 (0.9)

[a] See footnote a of Table 1.
[b] ALPHA, BETA, TURN, and COIL designate, respectively, the total number of residues in alpha-helical, beta-strand, turn, and random-coil conformation. STRAINED designates the total number of residues in unfavorable regions of the Ramachandran map.

strand-4 (Thr51, Thr53, and Thr53). Global conformational reporters (RMSD, SAS, and R_G) for [I6T,T53G]-PGA resembled those of the single mutant, I6T-PGA, in their variation tendencies (data not shown), but the RMSD reached a flatter plateau, having a maximum of value of ~1.60 Å at the end of the simulation. In contrast, geometrical reporters (Table 1) indicated an increase in hydrogen bonding "stability" with respect to I6T, reaching levels comparable to those of the other systems. In particular, ANTIHB was re-established. Secondary structure reporters confirmed the more wild-type like behavior of this double-mutant (Table 2). But an increased deformation of the C-terminal section of the helix (residues 33–36) was observed in this case (data not shown).

In summary, with the exception of I6T-PGA, none of the parameters reflected large or global deformations in any of the systems with respect to the X-ray conformation. Only few residues appeared to "lose" secondary structure and hydrogen bonding. For I6T-PGA, notwithstanding the more pronounced loss of assigned secondary structure, global conformational reporters indicated the absence of disruption of the tertiary fold. In general, perturbation of the initial systems was inversely proportional to the bulk of the side chain at P6. Threonine-6, being a polar amino acid, introduced an additional perturbation. Interestingly mutations did not seem to affect their immediate surroundings. Effects were delocalized across the structure and distributed among distinct "strained" points. Finally, the substitution of Thr53 by a glycine in the double mutant [I6T,T53G]-PGA effectively attenuated the effects of the single mutation, by re-establishing the "integrity" of the sheet surface. In the following section, we will analyze the topographical distribution of the perturbations ensuing the various mutations, focusing on single mutant systems only.

Deviations and fluctuations

Deviations from the crystallographic conformation at a residue level were measured by the RMSD of individual C-alpha carbon atoms (Fig. 3). The first striking feature, when comparing mutant systems with WT-PGA, was the quasi absence of deformation within the site of mutation itself. For the rest, two types of perturbation were observed. The first one, present in all mutant systems, affected the regions surrounding the first β-hairpin turn, loop-2, and Val21 (this residue has one of the most elevated temperature factors in the 1pga crystal structure and the highest one in 1pgb (Gallagher et al., 1994)). The second type of perturbation was mutant specific, involving loop-1 and turn-2 for I6T, and the upper half of the helix for I6G and I6F. The absence of perturbation in the upper half of the helix (26–34) for I6T was intriguing since this mutation had shown to be more perturbative than I6F or I6G. In fact, I6T behaved very much like WT for this segment of the protein. The individual characteristics of the amino acid replacements could explain the specificity of the perturbation: both the bulkier (I6F, phenylalanine's side-chain volume: 175 Å^3) and the more flexible mutant (I6G) influenced helix conformation, on the contrary threonine, although smaller than isoleucine was capable of maintaining the packing organization seen in WT-PGA (threonine's side-chain volume is 102 Å^3 vs. 140 Å^3 for isoleucine).

Deformations were not necessarily associated with an increase in the mobility of the implicated domains (Fig. 3). Indeed, plots of the RMSF of individual C-alpha carbons showed that only I6G-PGA was more mobile in those regions that deviated from the crystallographic conformation (Fig. 3B). I6F-PGA was remarkably "stiff." I6T-PGA had a mixed behavior: the conformational changes in the first β-hairpin were not paralleled by a sizable increase in the fluctuations, but, stretching from the C-terminus of the helix to the end of the second β-hairpin-turn, an appreciable magnification of the mobility was observed (Fig. 3D). I6T-PGA displayed a general increase in flexibility in comparison to any of the systems.

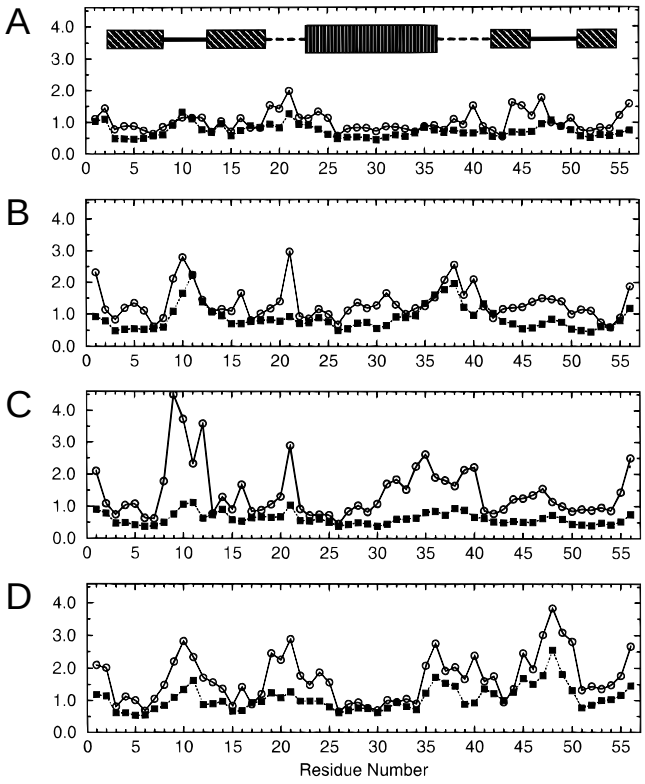

Fig. 3. Average RMSD (open circles) and RMSF (filled squares) of individual C-alpha atoms in Å. **A:** WT simulation with cartoon representation of the elements of secondary structure (from left to right: strand-1, strand-2, helix, strand-3, and strand-4). **B:** I6G. **C:** I6F. **D:** I6T.

Mechanical aspects of protein G's architecture

Mechanical response matrices

To investigate the mechanical correlations between the various protein domains, mechanical response matrices (see Methods) were constructed for each of the mutants (Fig. 4). The matrices are represented using identical gray-scales. The overall darkness of individual plots reflects what has been observed until now, e.g., I6F-mutation froze protein correlations in comparison to WT, while I6G and I6T intensified them, the latter being the most effective. The off-diagonal terms of the matrices link domains that would be deformed simultaneously following a small perturbation. In this way, a first mechanical decomposition of the protein was achieved. Principally, the topological organization of the protein resulted in the correlated interaction among the following six domains: D1, N-terminus of the protein (first two residues of the sequence); D2, first β-hairpin-turn (residues 9–12) and Gly14; D3, centered around Val21 extending from the end of the second strand to the first few residues of the helix (residues 17–26); D4, C-terminal end of the helix and loop-2 (residues 33–41); D5, second β-hairpin-turn (residues 46–51); D6, C-terminus of the protein (last two residues in

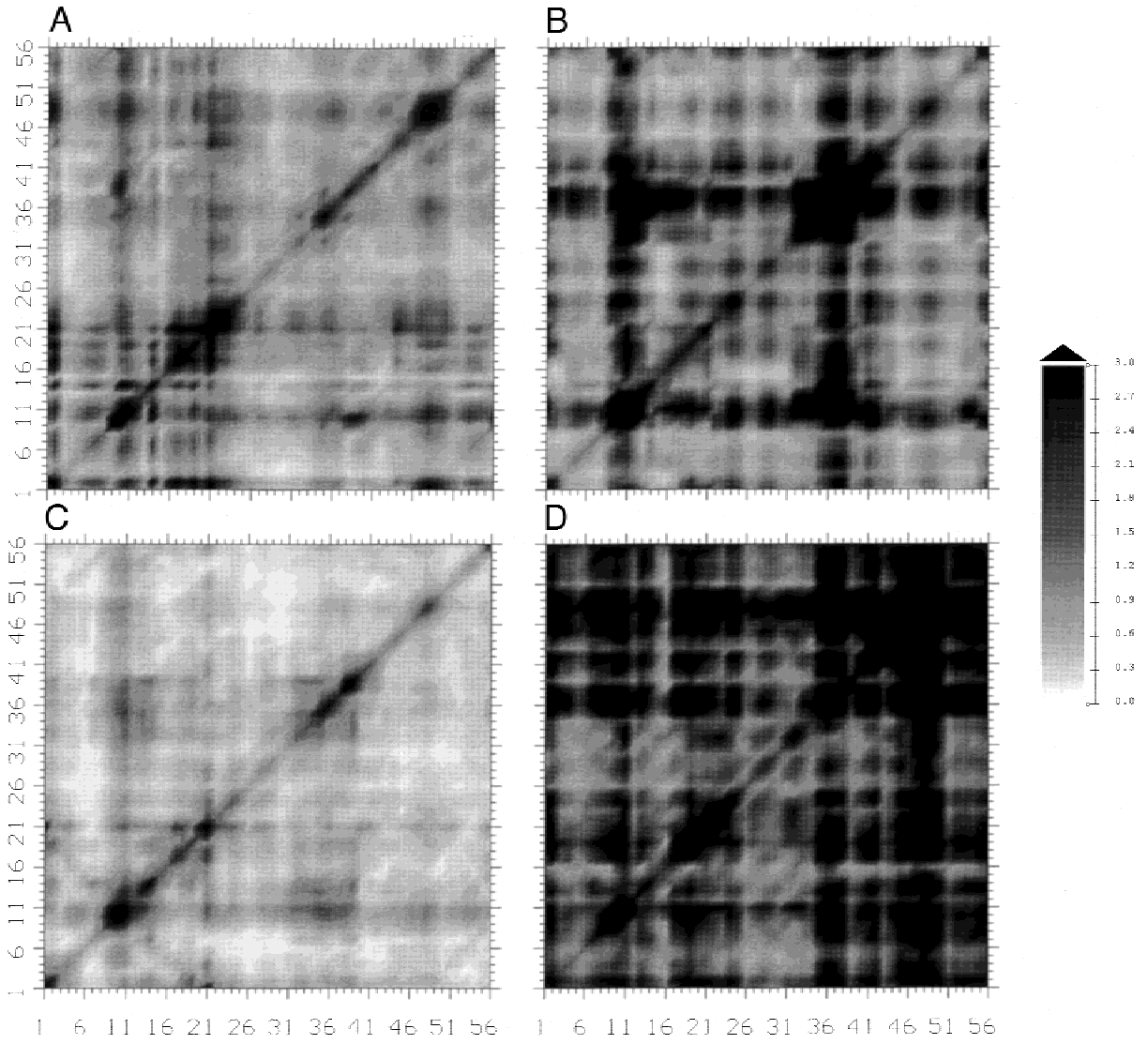

Fig. 4. Mechanical response matrices. Residue numbers are reported on *x*- and *y*-axes, intense correlations are darker. Scale units are in $10^{-2} \times$ Å. **A:** WT simulation. **B:** I6G-PGA. **C:** I6F-PGA. **D:** I6T-PGA.

the sequence). It should be noted that the residue boundaries in parenthesis are just indicative and should not be considered in a strict sense.

Inspection of the "cross peaks" indicated that the response to a perturbation did not give rise to a generalized deformation of the protein, but that specific pathways were used to cope with external excitation. Distinct correlation pathways linked various protein domains, in a way that reflected the topological organization of the protein. Domains that were close to each other in three-dimensional space showed stronger coupling than distant ones. In particular, the data highlighted the importance of the first β-hairpin turn (D2) in the architectural organization of PGA. Comparison of the mutant matrices with that of WT-PGA suggested that the latter were the product of the amplification or attenuation of the correlations observed for wild-type protein G. The various systems differed by the specific "choices" of correlations that were emphasized. For I6G-PGA, the D2/D4 pathway was most intense, activating contemporaneously the associated secondary coherences seen in the WT protein. In I6F-PGA, most interactions were attenuated. On the contrary, for the I6T mutant, simultaneous activation of the pathways associated with D2 and D5 lead to a marked amplification of the protein mechanical response.

Dynamical correlations

In the previous section, a series of internal communication pathways were brought out, and comparison of the various mutants with WT-PGA seemed to indicate a certain similarity in the internal *mechanics* among the different systems. To assess the extent of overlap between the internal *dynamics* of the various systems, we proceeded using principal component analysis over the 600–1,600 ps time range (see Methods).

For each system, diagonalization of the covariance matrix of the positional fluctuations afforded a set of eigenvectors and corresponding eigenvalues. For all proteins ~80% of the total variance was represented by a few principal components (eigenvectors with largest eigenvalues) indicating that internal motions occurred mostly

along a limited set of directions (~10 eigenvectors) corresponding to a principal subspace (Amadei et al., 1993). The conformational flexibility of individual systems decreased in the order I6T (0.728) > I6G (0.514) > WT (0.327) > I6F (0.234), in agreement with the results of previous sections. In parenthesis, we have reported the total variance of each system. The similarity of the internal fluctuations between the various systems was evaluated by comparing the principal subspaces (first 10 eigenvectors) of each protein. Thus, for each pair of systems, we evaluated the RMS inner product (RMSIP) value between the two 10-eigenvector subsets (Table 3). The diagonal terms represent internal comparison parameters that allow to estimate the significance of the overlap (value of RMSIP) between two different protein systems. However, this sort of upper-limit value is just indicative, since segments of trajectory 500 ps long, are not necessarily sufficient to accurately define an essential subspace. The lower-end value of RMSIP depends on the dimensionality of the positional space, in this case: 3×56(C-alpha) $- 6$ (rotational and translational degrees of freedom), affording an estimated lower-end RMSIP at $\sqrt{10/162} = 0.25$. It can be shown that the density probability distribution for a casual overlap is not Gaussian and very narrow about the mean, and that the RMSIP values seen here are very unlikely the result of casual overlap (A. Amadei, unpubl. results).

Thus, the various systems showed similar and significant degree of overlap with WT and with one another. This suggested that the mutant systems oscillated within the framework of a conformational space closely similar to that of the WT protein. Thus, mutations resulted in amplification or attenuation of dynamical pathways already present in the WT protein, in the same manner that mutants shared common mechanical pathways.

The similitude of the internal dynamics of the various systems having been established, we will try to gain some geometrical insight into such pathways focusing on the WT-PGA system. The conformational evolution of WT-PGA along the first three principal components is represented using a gray-scale code, where each level of gray corresponds to a different conformation of the protein's C-alpha trace (Fig. 5). Such trajectory filtering as been described previously (Amadei et al., 1993).

Along the first eigenvector (Fig. 5A,B), a concerted inward (toward the protein interior) motion of the two hairpin turns, resulted in an overall backward tilt of the helix axis in relation to the plane of the sheet, the helix basis (its N-terminus) serving as pivot point. A slight bend was observed in the C-terminal section of the helix above Phe30. In addition, as the helix axis tilted away, loop-1 moved in toward the core, and the C-terminal section of the second strand, between residues 15 and 20, in concert with the first two residues (Met1 and Thr2), rotated laterally and away from the protein. As loop-1, strand-3 and turn-2 closed-in onto each other, the central portion of the sheet (residues 5–7 and 52–54) along with Gly14 arched over the helix. Loop-2 was almost rigidly carried along by strand-3 and the helix top.

Along the second eigenvector (Fig. 5C,D), the first hairpin-turn swung sideways on the sheet surface with an outward trajectory. Simultaneously, the second hairpin-turn had a similar outward lateral motion, in conjunction with an out-of-plane perpendicular bend, as the helix basis swung away. Consequently, the helix and the sheet were seen as rotating with respect to each other through an axis normal to the plane of the sheet and passing approximately through Phe30. The central section of the sheet and Gly14 remained tied together and quasi-motionless. And again, the N-terminus of protein G and the C-terminal section of strand-2 moved coherently, "pulled" by loop-1. Finally, the last two residues of protein G (55 and 56) accompanied the lateral motion of the first β-hairpin.

Along the third eigenvector (Fig. 5E,F), fluctuations were mainly concentrated on the first hairpin turn. As for the first eigenvector, this hairpin bent perpendicularly to the sheet-plane, but also had a slight lateral swing (as along eigenvector 2). The C-terminus of the helix was essentially rigid (above Phe30), while its N-terminus moved in toward the interior of the protein as turn-1 tilted away (note the bending of the helix axis). A slight torsional rotation of the helix as for the second eigenvector was also perceptible. The second hairpin of the sheet was essentially rigid, with a slight lateral motion at the hairpin turn similar to that seen along the second eigenvector. Loop-2 and turn-1 moved away from each other. The central section of the sheet, consisting of the two parallel strands and Gly14, formed a coherent set again, arching itself as the helix rotated toward the second hairpin. This time, the contorted motion of the N-terminus of the protein and the end of the second strand was less pronounced than along the first eigenvector. As along the second principal direction, the C-terminus of the protein G (residues 55 and 56) and turn-1 formed a single concerted domain.

In summary, the motions along individual eigenvectors pointed out the existence of a series of fairly constant groups of atoms that moved in a concerted fashion. β-Hairpin turns were essentially involved in out-of-plane motions with respect to the plane of the sheet, and these motions were more or less tilted with respect to the normal of the sheet surface. The helix behaved to a large extent as a single domain whose orientation, in relation to the sheet, changed either by tilting away from it at its C-terminal end (the N-terminus remaining in close contact with the plane of the sheet), or by rotating over the sheet's surface about an axis, passing through Phe30, and perpendicular to the sheet. But some local deformations were also observed at the C-terminal end of the helix (between residues 33 and 36). The central region of the sheet (strands 1 and 4), flanked by the flexible Gly14, formed clearly a single dynamical unit. Another dynamical domain consisted of loop-1 and the C-terminal section (beyond residue 15) of strand-2. In general, strand-2 presented marked contorted motions of the backbone, which contrasted with the more uniform motions of the other domains. The third strand, located at the edge of the sheet, was

Table 3. *Comparison of essential subspaces*[a]

	WT-PGA	I6G-PGA	I6F-PGA	I6T-PGA
WT-PGA	0.78[b]	0.74	0.72	0.73
I6G-PGA	0.74	0.74[b]	0.75	0.67
I6F-PGA	0.72	0.75	0.83[b]	0.68
I6T-PGA	0.73	0.67	0.68	0.66[b]

[a]The RMSIP value between the first 10 eigenvectors of two protein systems A and B is defined as

$$\mathrm{RMSIP} = \left(\frac{\sum_{i=1}^{10} \sum_{j=1}^{10} (\eta_i^{\mathrm{A}} \cdot \eta_j^{\mathrm{B}})^2}{10} \right)^{1/2}$$

where $\eta_i^{\mathrm{A(resp.B)}}$ is the ith eigenvector of set A (resp. B).

[b]This number represents the RMSIP obtained when comparing the eigenvector sets derived from the first (600–1,100 ps) and second half (1,100–1,600 ps) of the trajectory of the given system.

dynamically linked to the second β-hairpin turn (first and third eigenvectors). Along the second eigenvector, however, the N-terminus of strand-3, moved independently of turn-2 and accompanied the motion of loop-2. Finally, the N- and C-terminal residues of the protein coupled their motions, respectively, to that of loop-1 and turn-1. Thus a certain degree of overlap existed between the mechanical domains of the previous section and the above dynamical domains, and just as for the mechanical domains, the exact boundaries were not strictly determined, since they were not conserved across eigenvectors. Nevertheless, for a given

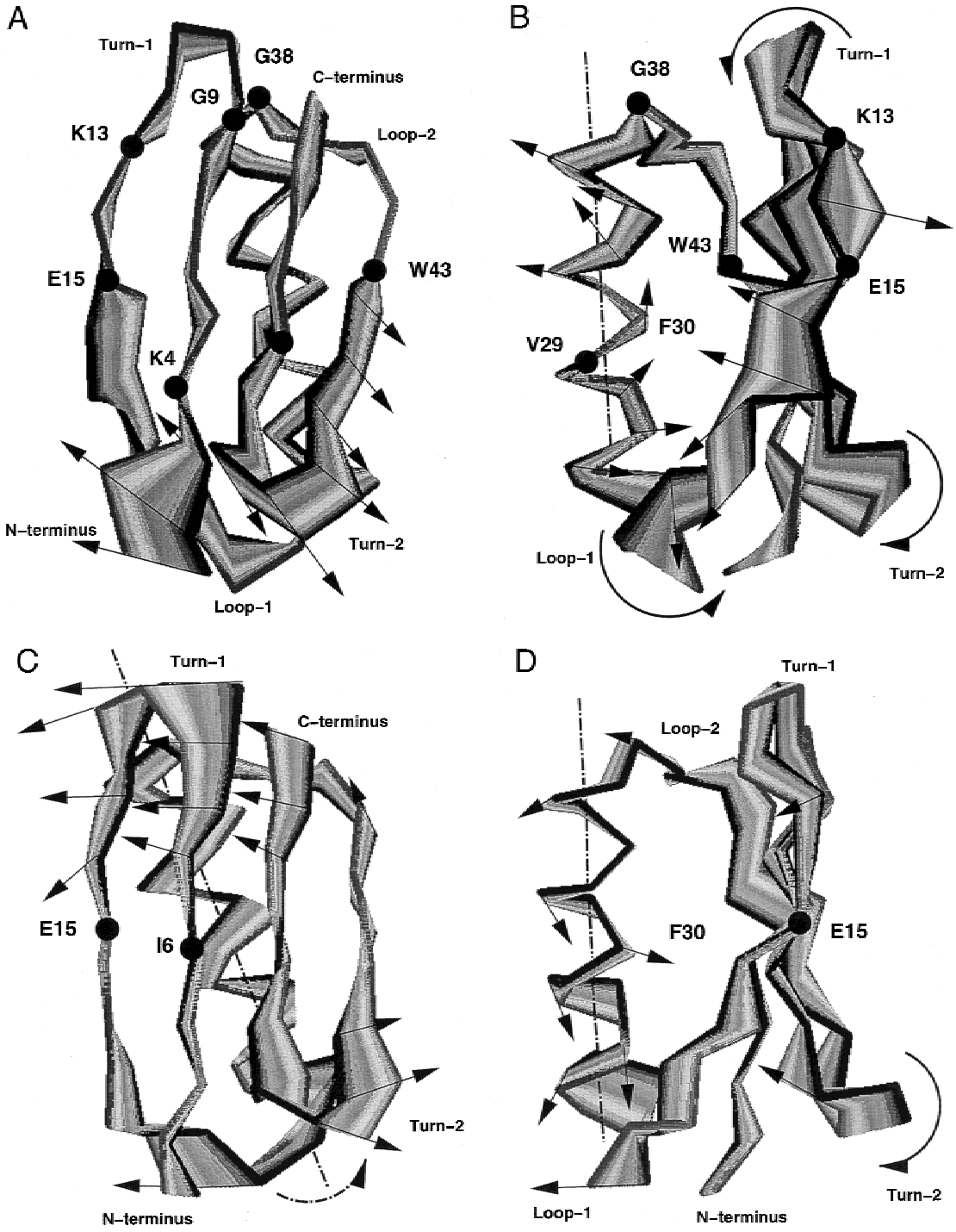

Fig. 5. Protein motions along principal eigenvectors. A gray-scale is used to represent motion in a film-like fashion. Vectors are used as qualitative indicators of the direction of motion of selected atom. Selected articulations are represented as a bullet. (**A**) Front and (**B**) lateral view of the C-alpha carbon trace of PGA as it "moves" along the first eigenvector direction, (**C**) and (**D**) for the second eigenvector, and (**E**) and (**F**) for the third eigenvector, respectively. (*Figure continues on facing page.*)

eigenvector, the atomic boundaries, or what we will call an articulation point (a group of one or more atoms across which atomic motion would occur in opposite directions), were fairly clear.

Effects of packing: Decomposition/reconstitution

To establish the extent to which the location and motion of these domains and articulation points reflected the three-dimensional packing organization of the protein, we have analyzed the molecular dynamics trajectories (1.6 ns) of individual segments of protein G (see Methods). To prevent a possible rapid unfolding (loss of topology) of the segments, simulations were carried out at 278 K (see Methods for setting of other parameters). Monitoring the evolution of the RMSD showed that all three systems moved sensibly away from their original conformation (data not shown), but both the R_G and the SAS of the individual segments did not increase significantly, indicating the absence of unraveling of the chains during the time of simulation. Structural reporters (Tables 4 and 5) confirmed that the topology of the individual chains remained mostly intact.

In general, atomic fluctuations and deviations increased considerably at the extremities of each segment (Fig. 6). Regions, such as strand-1 and the upper helix (residues 25–34), were strongly affected by the absence of packing partner. In both β-hairpins, fluctuations and deviations of the turn residues were decoupled from the rest of the peptide sequence by quasi-fixed residues at each end of the turns (8–13 for BH1 and 45–49 for BH2), in a manner reminiscent of the articulation points described previously. It should be noted that, as in the individual eigenvectors, these fixed points do not correspond exactly with the native endings, which are 9–12 and 46–50 for BH1 and BH2, respectively.

Table 4. *Structural statistics*[a] *(600.0–1,600.0 ps)*

System	SAS (Å^2)	R_G (Å)	HBO	PARAHB	HB4
BH1[1–20]	2,117 (54)	9.8 (0.2)	17.7 (2.8)	7.7 (1.0)	—
	2,117	—	11	8	—
HH1[19–41]	2,196 (45)	8.9 (0.2)	15.1 (1.2)	—	9.7 (0.8)
	2,196	—	15	—	12
BH2[41–56]	1,699 (40)	7.5 (0.4)	8.6 (2.2)	3.8(1.1)	0.6 (0.6)
	1,694	—	8	6	—

[a]For definitions, see footnote a of Table 1. For each structural reporter, values in the X-ray structure are reported on the second line. A dash appears where structural parameters are not applicable.

To further assess the parity between the dynamical fluctuations of the WT protein and those of the individual peptides, we compared the principal subspaces (first 10 eigenvectors) of BH1, HH1, and BH2 with the principal subspaces of their respective segments within WT-PGA. Thus, the covariance matrices of positional fluctuations for peptides BH1[2–19], HH1[23–36], and BH2[42–55], were built from the trajectories of the isolated peptides and that of the complete protein. Notice that these segments were shortened so as to reduce N- and C- termini effects. Diagonalization of these matrices afforded six sets of eigenvectors, two for each peptide segment, and for each sequence pair, "wild-type" and isolated, its RMSIP was evaluated. The calculated values were 0.81, 0.86, and

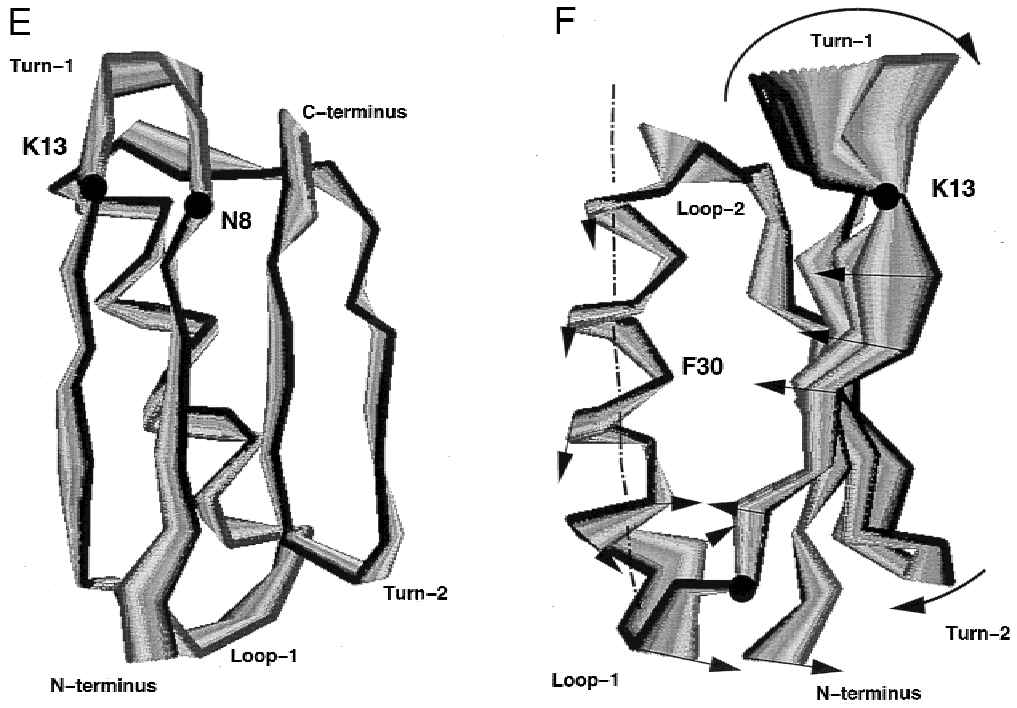

Fig. 5. *Continued.*

Table 5. *Structural statistics*[a] *(600.0–1,600.0 ps)*

System	Alpha	Beta	Turn	Coil	Strained
BH1[1–20]	—	13.1 (2.3)	2.5 (0.6)	4.3 (2.2)	0.5 (0.5)
	—	11	2	4	1 (Gly14)
HH1[19–41]	11.4 (1.3)	—	6.2 (1.8)	5.3 (1.2)	1.8 (0.9)
	14	—	4	4	0
BH2[41–56]	—	5.6 (2.1)	6.0 (1.9)	4.3 (1.9)	0.3 (0.5)
	—	10	4	2	0

[a]For definitions, see footnotes of Tables 2 and 4.

0.85 for BH1[2–19], HH1[23–36], and BH2[42–55], respectively. However, the lower dimensionality of the systems increases the probability of a casual overlap between any two subsets of eigenvectors. In the case of the peptide fragments of about 16 residues, this lower-end limit can be calculated at ~0.50, resulting in a ratio between RMSIP and lower-end limit of about 1.6. This ratio is smaller than that observed for the full sequence systems (mutant proteins and WT-PGA), but remains nonnegligible, and is far above that expected from a casual overlap (vide supra). Thus, the fluctuations of individual fragments were in close relationship with those of the corresponding fragments in WT, notwithstanding the absence of packing partners, the principal difference being the amplitude of the fluctuations rather than their nature.

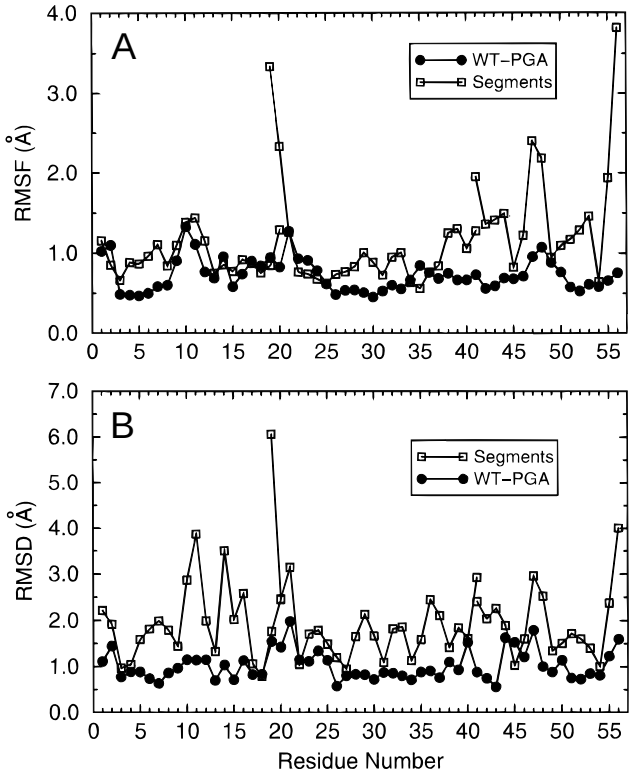

Fig. 6. Average RMSF and RMSD of C-alpha atoms in angstrom. (**A**) Average RMSF and (**B**) RMSD values are reported for every C-alpha atom. Results from peptidic segments (BH1, HH1, and BH2) are represented by open squares, and compared to WT filled circles.

In summary, conformational reporters indicated that individual peptide fragments had maintained enough topology so as to proceed with the analysis of packing effects. Partial removal of packing interactions, resulted in local changes in conformation and mobility, highlighting those amino acids and peptide segments involved in core interactions. Deformations and fluctuations reflected the topology of the fragments and the articulated building-blocks of the fragments resembled closely those within the full sequence protein. The packing interactions controlled the amplitude of the motions, while the directions of fluctuations reflected the specific topology of the chain.

Discussion

The three-dimensional organization of the B1 domain of protein G confers a remarkable thermal stability to this small protein, T_m = 87 °C (Gronenborn et al., 1991). The unusual topology of the chain, [−1, +3*x*, −1] (Richardson, 1977), the tight packing of the hydrophobic core, and the great number of hydrogen bonds in the protein have all been proposed to be determinant factors in the stability of WT-PGA (Gronenborn et al., 1991; Achari et al., 1992; Orban et al., 1995). However, the simple mutation of the solvent exposed isoleucine-6 to a glycine was reported to thermally destabilize the protein by "an unprecedented" 25 °C (Smith et al., 1994). The following discussion pertains to the description of WT-PGA as an assembly of elementary building-blocks, and to the analysis of the mechanisms that the protein uses in response to mutations of Ile6 and sequence deletions.

Response mechanisms

How consistent were the responses of the native fold with the physical and conformational properties of the substituting amino acids? Three different types of amino acid were used to replace the isoleucine at P6, and globally the level of perturbation decreased in the following order I6T > I6G > I6F. The large and hydrophobic phenylalanine, in I6F-PGA, conserved and rigidified the native protein configuration most (Figs. 2, 3; Tables 1, 2). This is not surprising since phenylalanine and other aromatic or β-branched residues have highest β-sheet forming propensities (Kim & Berg, 1993; Minor & Kim, 1994; Smith et al., 1994). One of the reasons advanced for such tendency lies in the size of their side chain that acting as a solvent-blocking agent would strengthen backbone hydrogen bonding in β-sheet structures in particular (Bai & Englander, 1994). But the number of hydrogen bonds across β-strands was not better conserved in I6F-PGA than in I6G-PGA (Table 1), where the poor β-sheet forming glycine is used. Another explanation for the effect of the I6F mutation may lie in the hydrophobic cluster that surrounds P6. Such clusters contribute to the organization and thermal stability of surface layers (Van Den Burg et al., 1994; Tisi & Evans, 1995; Frigerio et al., 1996), and β-sheets in particular could benefit from the associative effect that would drive the formation of a local solvent-exposed core (Mayo et al., 1996). These factors can also explain the more disruptive character of the flexible glycine mutant, especially since glycine is one of the amino acids with least β-sheet propensity (Kim & Berg, 1993; Minor & Kim, 1994; Smith et al., 1994). The I6T mutation is not as clear. On one hand, threonine is a good β-sheet forming residue and its smaller side chain, in comparison to isoleucine, should not create new steric clashes. On the other hand, its polarity could have a negative effect on the integrity of the sheet's surface. First,

because the hydrophobic drive for compactness (vide supra) would be compromised. Second, in I6T-PGA, Thr6 faces a row of polar threonines on strand-4 and it is known that pairing of threonines across antiparallel strands is not favored energetically (Smith et al., 1995). However, despite their unfavorable interaction energy, threonine–threonine pairs are often found in β-sheets (Smith et al., 1995), e.g., Thr53/Thr44 in PGA, and they have been found to promote β-hairpin formation when located on the hydrogen-bonded face of antiparallel strands (de Alba et al., 1997). But here, Thr6 and its paired neighbor Thr53 lay across parallel strands. The most plausible explanation for the strong perturbative effect of the I6T mutation remains the disruption of the native hydrophobic cluster, and this effect can be reduced by removing the side chain at position 53, as in the [I6T,T53G]-PGA double mutant.

The distribution of mutational effects across the chain was also of interest. Conformational changes were not limited to the surroundings of the mutation site. Instead specific and distant protein domains were affected (Figs. 3, 4). Propagation of the perturbations took place through *surface interactions*, as seen for I6T in the preceding paragraph, but *surface/core interplay*, between the helix and the sheet, was also present (Fig. 3). Interaction between the sheet and the helix is controlled principally by Leu7 on strand-1 (Achari et al., 1992; Gallagher et al., 1994). However, we did not see large variations in RMSD or RMSF values at P7 (Fig. 3). In contrast, the effect of P6 mutants on the helix core residues was clear (see Results and Fig. 3 between residues 25–34). This surface/core interaction was directly related to the size of the side chain at P6. In this respect, it is interesting to note that the B2 domain, which contains a smaller valine residue at P6 (and Leu at P7), is less stable thermally than the B1 domain (T_m = 79.4 °C vs. T_m = 87.5 °C), which contains the Ile6/Leu7 pair (Alexander et al., 1992; Orban et al., 1992).

Another remarkable apsect was the delocalization of stress across the sequence. Mostly glycine residues were affected, as though these residues were acting as stress-relief points. These data suggest that strategic positioning of glycine residues across a sequence might be determinant for the thermostability of a protein. This is in close connection with a recent report that relates protein thermostability and the apparition of strain in residues at the interface between conformationally regular regions (Karplus, 1996).

It should be noted that no straightforward connection exists between flexibility and stability (Lazaradis et al., 1997). Thus, it is not possible to ascertain that I6F-PGA will be more stable than WT-PGA, or that I6T-PGA will be less stable, on the basis of flexibility alone. However, mutational effects were shown to be consistent with amino acid types (supporting the validity of the simulations), and resulted not only in flexibility changes but also in conformational reorganizations consistent with I6G-PGA being less thermally stable than WT-PGA: disruption of hydrophobic cluster on the sheet's surface and of helix/sheet packing interaction as controlled by P6/P7 conformational coupling (vide supra). In addition, the loss of hydrophobic surface burial within the core of I6G-PGA, as a result of the increase in the fluctuations of the first hairpin-turn (see below and Results), should contribute significantly to its destabilization.

Architectural decomposition

The evolution of all systems happened by amplitude modulation of the large concerted motions that characterize the native protein fold (Figs. 2–4, 6; Table 3). Both mechanical and dynamical analyses suggested a decomposition of the native fold into a series of interacting building-blocks (Figs. 4, 5). These articulated objects corresponded mostly to elements of secondary structure (helix, strands, turns, and loops), but they did not consist exclusively of sequential stretches of residues (see Results). Among those building-blocks that extended across space were the two central and parallel strands, which formed the backbone of the β-sheet, and the protein N- and C- termini, which interacted coherently with loop-1 and turn-1, respectively. In this respect, the dynamic coupling of Gly14 with the central strands is remarkable. The side chain of Gly14 should point toward the core of the protein and face the helix if strand-2 were regular; however, as a flexible glycine residue, it can accomplish the double task of not interfering sterically with Tyr33 on the helix and participating to the surface hydrophobic cluster that surrounds Ile6 (vide supra).

How consistent, then, is the native building-block decomposition with known dynamical features? Globally, very small amplitude fluctuations were present; turns, loops, and protein termini being the most mobile sections, in close agreement with a recent MD report (Sheinerman & Brooks, 1997). The details of the building-blocks (see Results) were remarkably consistent with the decomposition derived from ^{15}N relaxation measurements (Barchi et al., 1994): the central strands were essentially rigid, and strand-2, that manifests chemical exchange line-broadening (Barchi et al., 1994), was found to have very marked contorted motions in comparison to the other strands (Figs. 3, 5). As mentioned before, building-blocks and secondary structure elements were not directly superimposable. This was particularly interesting in the case of the helix, where the mechanics and dynamics of the C-terminus disrupted helical continuity, often linking this section to the first residues of loop-2. This discontinuity of the helix at its C-terminus has also been observed in hydrogen-exchange rate measurements of main-chain amide protons (Orban et al., 1995) and NMR-determined backone dynamics (Barchi et al., 1994), the C-terminal section of the helix being more exposed to solvent. Furthermore, in structural studies (Gronenborn et al., 1991), those few last residues of the helix have been reported to adopt a more extended 3_{10} helical conformation.

The dynamical relationships among the various building-blocks are in agreement with the principles of α/β-packing (see Introduction). The principal motions pertain to the relative orientation of the helix and the sheet. It was found that the helix could both rotate about the sheet surface and tilt away from it, and that these motions were brought about by changes in the concavity of the sheet's plane (see Dynamical correlations in results). The latter observation has also been advanced in the course of ^{15}N relaxation measurement studies (Barchi et al., 1994). In addition, it is important to note that differences in the orientation of the helix in relation to the sheet have been observed between the different IgG-binding domains of protein G (Achari et al., 1992; Gallagher et al., 1994) and that these helical rotations were related to conformational changes in the first hairpin-turn and to hydrophobic core packing differences at Leu7. In this respect, it is worth mentioning that the B2 domain binds IgG more efficiently than the B1 domain (Alexander et al., 1992), suggesting that the orientation of the helix as controlled by the P6/P7 pair (vide supra), could play a determining role in the biological function of these protein G domains. This is consistent with our finding that hairpin turns form independent mechanical and dynamical units, and that the amplitude of their motions (lateral or perpendicular out of plane swinging) is controlled by the size of side chain at P6.

Role of packing

What is the role of specific packing interactions in the definition of the building-blocks and their articulated relationships? Models of α/β-packing vary in their interpretation of specific packing events (see Introduction). In the case of PGA, the type of interdigitation pattern proposed by Cohen et al. (1982) is difficult to identify. However, Phe30, on the helix, points directly toward the sheet interior surface, and it is this helix residue which appears encircled by other residues on the sheet (Phe52, Tyr3, Thr18, and Leu5). Furthermore, the small Ala23 and Ala34 are the only other helix residues that directly point their side chain to the sheet, all other helix core residues interacting laterally rather than directly. This could be particularly relevant because the small side chain of alanine has been shown to provide greatest flexibility in packing orientation of helices (Walther et al., 1996). In terms of dynamics and building-blocks, the analysis of the simulations of isolated fragments shows that most of the structural decomposition and dynamical information is maintained in the absence of packing. These data suggest that there is no synergistic effect on the dynamics of the protein, which arises from specific packing interactions, and consequently, packing of the core might not be as specific as expected, in agreement with recent findings (Schultz Beardsley & Kauzmann, 1996; Bowie, 1997). Just as in the case of the mutants, packing modulated the amplitude of motions, acting as a sterical constraint that limits and coordinates the fluctuations of elementary building-blocks that are already defined by the local topology of the chain. As a consequence, perturbations, such as core mutations, would only lead to changes in the relative orientation of these elementary blocks and their levels of fluctuation. This is mostly what has been observed until now in core repacking experiments (Lim & Sauer, 1989; Baldwin et al., 1993). This phenomenon has also been observed and exploited in protein design experiments (Harbury et al., 1995; Dahiyat & Mayo, 1996, 1997; Su & Mayo, 1997). In those experiments concerning the design of PGA variants (Dahiyat & Mayo, 1997; Su & Mayo, 1997), a rigid body description of whole secondary structure elements was used to derive novel residue packings in the core of the protein. We hypothesize that the use of more detailed and flexible descriptors in combination with the knowledge of the specific interplay between these elementary building-blocks could lead to a more efficient protein design approach.

In summary, the present work has emphasized the dominant role of P6 in orchestrating interactions across the sheet's surface through the formation of a surface hydrophobic cluster (Tisi & Evans, 1995) and in controlling the packing of the helix onto the sheet through surface/core interactions. In addition, P6 was also shown to modulate the amplitude of the fluctuations at the first β-hairpin, and consequently, affect other sections of the protein as well. Mutational effects were shown to be distributed across the sequence within specific building-blocks units, the nature of the residue at P6 controlling the amplitude of the motions of such articulated entities. The structural decomposition, thus achieved, was consistent with known dynamical features of protein G and with the principles of α/β packing. Elementary building-blocks were closely related to elements of secondary structure, but did not consist exclusively of sequential segments. The nature of the building-blocks was shown to be mostly determined by the chain's topology rather than arising through specific packing interactions. Finally, it was suggested that knowledge of such mechanical and dynamical information could be incorporated into a de novo designed protein to control stability and biological function.

Methods

All simulations were performed with the GROMACS simulation package (Van der Spoel et al., 1995). A modification (Van Buuren et al., 1993) of the GROMOS87 (Van Gunsteren & Berendsen, 1987) force field was used with additional terms for aromatic hydrogens (Van Gunsteren et al., 1996) and improved carbon–oxygen interaction parameters (Van Buuren et al., 1993). The SHAKE algorithm (Ryckaert et al., 1977) was used to constrain bond lengths, allowing a time step of 2 fs. The initial wild-type protein configuration (WT-PGA) was taken from entry 1pga (Gallagher et al., 1994) of the Protein Data Bank. Mutants (I6G-PGA, I6F-PGA, I6T-PGA, and [I6T,T53G]-PGA) were constructed within Insight II 97.0 (Biosym/MSI, San Diego, California) using the crystallographic coordinates as a template and replacing target residues with the desired amino acid. Initial peptide configurations were obtained by cleaving the desired segments off the WT coordinates and adjusting the newly obtained termini with free ammonium and carboxylate groups. Peptide sequences corresponding to the first (BH1[1–20]) and second (BH2[41–56]) β-hairpin as well as the helix (HH1[19–41]) were built in this manner. The N- and C-terminus of each peptide is reported in brackets. Each system, WT, mutant, or peptide, was immersed in a pre-equilibrated box of SPC water (Berendsen et al., 1981), and water molecules with highest electrostatic potential were replaced by counter ions (Na or Cl), to give an electrically neutral cubic box. Care was taken that all crystallographic water molecules be conserved in each case.

To prepare each solvated system for molecular dynamics, a three step procedure was followed. Using a restraining harmonic potential, all heavy atoms of the protein and the crystallographic water oxygens were constrained to their initial positions, while surrounding SPC water molecules were first minimized and then submitted to 5 ps of constant volume MD at 300 K. The resulting system was then minimized, without any constraints, before starting constant temperature and constant volume molecular dynamics. A nonbonded cutoff of 1.2 nm was used for both Lennard–Jones and Coulomb potentials. The pair lists were updated every 10 steps. A constant temperature of 300 K was maintained by coupling to an external bath (Berendsen et al., 1984) using a coupling constant (τ = 0.002 ps) equal to the integration time step. Peptide simulations were carried out at 278 K and a nonbonded cutoff of 1.0 nm was used instead, all other parameters were unchanged.

Analysis of MD runs

For each system, 1.6 ns of simulation were produced in this manner, of which only the last 1.0 ns of trajectory was used for comparative analyses. For principal component and mechanical response analysis, all configurations were fitted to the same reference structure, e.g., the crystallographic configuration (PGA), by first translating all center of masses to the origin of coordinates and then superimposing the configurations using a least-squares fitting procedure. For the peptides, the corresponding subsequence in PGA was used as reference configuration for fitting purposes. The covariance matrix of the positional fluctuations was constructed using the coordinates of C-alpha atoms (Amadei et al., 1993), and diagonalization of this matrix afforded the eigenvectors and eigenvalues used in principal component analysis. Mechanical response

matrices were built from the above covariance matrix as described by Chillemi et al. (1997).

Acknowledgments

We would like to thank CASPUR (Consorzio per le Applicazione di Supercalcolo per Universitá e Ricerca) for access to computational facilities. This work was supported by EEC grant ERBFMRX-CT96-0013 and by the Ministerio dell' Università e della Ricerca Scientifica, project "Structural Biology."

References

Achari A, Hale SP, Howard AJ, Clore GM, Gronenborn AM, Hardman KD, Whitlow M. 1992. 1.67-Å X-ray structure of the B2 immunoglobulin-binding domain of streptococcal protein G and comparison to the NMR structure of the B1 domain. *Biochemistry 31*:10449–10457.

Alexander P, Fahnestock S, Lee T, Orban J, Bryan P. 1992. Thermodynamic analysis of the folding of streptococcal protein G IgG-binding domains B1 and B2: Why small proteins tend to have high denaturation temperatures. *Biochemistry 31*:3597–3603.

Amadei A, Linssen ABM, Berendsen HJC. 1993. Essential dynamics of proteins. *Proteins Struct Funct Genet 17*:412–425.

Bai Y, Englander SW. 1994. Hydrogen bond strength and β-sheet propensities: The role of a side chain blocking effect. *Proteins Struct Funct Genet 8*:262–266.

Baldwin EP, Hajiseyedjavadi O, Baase WA, Matthews BW. 1993. The role of backbone flexibility in the accommodation of variants that repack the core of T4 lysozyme. *Science 262*:1715–1717.

Barchi JJ, Grasberger B, Gronenborn AM, Glore GM. 1994. Investigation of the backbone dynamics of the IgG-binding domain of the streptococcal protein G by heteronuclear two-dimensional ^{1}H-^{15}N nuclear magnetic resonance spectroscopy. *Protein Sci 3*:15–21.

Berendsen HJC, Postma JPM, DiNola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. *J Chem Phys 81*:3684–3690.

Berendsen HJC, Postma JPM, Van Gunsteren WF, Hermans J. 1981. Interaction models for water in relation to protein hydration. In: Pullman B, ed. *Intermolecular forces*. Dordrecht: D. Reidel Publishing Company. pp 331–342.

Blanco FJ, Serrano L. 1995. Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. *Eur J Biochem 230*:634–649.

Bowie JU. 1997. Helix packing angle preferences. *Nature Struct Biol 4*:915–917.

Brenner SE, Chothia C, Hubbard TJP. 1997. Population statistics of protein structures: Lessons from structural classifications. *Curr Opin Struct Biol 7*:369–376.

Chillemi G, Falconi M, Amadei A, Zimatore G, Desideri A, Di Nola A. 1997. The essential dynamics of Cu,Zn superoxide dismutase: Suggestion of intersubunit communication. *Biophys J 73*:1007–1018.

Chothia C. 1984. Principles that determine the structure of proteins. *Ann Rev Biochem 53*:537–572.

Chothia C, Finkelstein AV. 1990. The classification and origins of protein folding patterns. *Annu Rev Biochem 59*:1007–1039.

Chothia C, Janin J. 1981. Relative orientation of close packed β-pleated sheets in proteins. *Proc Natl Acad Sci USA 78*(7):4146–4150.

Chothia C, Levitt M, Richardson D. 1977. Structure of proteins: Packing of α-helices and pleated sheets. *Proc Natl Acad Sci USA 74*(10):4130–4134.

Chothia C, Levitt M, Richardson D. 1981. Helix to helix packing in proteins. *J Mol Biol 145*:215–250.

Chou KC, Némethy G, Rumsey S, Tuttle RW, Scheraga HA. 1985. Interactions between an α-helix and a β-sheet. Energetics of α/β packing in proteins. *J Mol Biol 186*:591–609.

Cohen FE, Sternberg MJ, Taylor WR. 1982. Analysis and prediction of the packing of α-helices against a β-sheet in the tertiary structure of globular proteins. *J Mol Biol 156*:821–862.

Crick FHC. 1953. The packing of alpha-helices: Simple coiled-coils. *Acta Crystallogr 6*:689–697.

Dahiyat BI, Mayo SL. 1996. Protein design automation. *Protein Sci 5*:895–903.

Dahiyat BI, Mayo SL. 1997. Probing the role of packing specificity in protein design. *Proc Natl Acad Sci USA 94*:10172–10177.

de Alba E, Rico M, Jiménez MA. 1997. Cross-strand side-chain interactions versus turn conformation in beta-hairpins. *Protein Sci 6*:2548–2560.

Derrick JP, Wigley DB. 1992. Crystal structure of a streptococcal protein G domain bound to an Fab fragment. *Nature 359*:752–754.

Doolittle R. 1995. The multiplicity of protein domains. *Ann Rev Biochem 64*:287–314.

Efimov AV. 1994. Common structural motifs in small proteins and domains. *FEBS Lett 355*:213–219.

Finkelstein AV, Ptitsyn OB. 1987. Why do globular proteins fit the limited set of folding patterns. *Prog Biophys Molec Biol 50*:171–190.

Frigerio F, Margarit I, Nogarotto R, de Filippis V, Grandi G. 1996. Cumulative stabilizing effects of hydrophobic interactions on the surface of the neutral protease from *Bacillus subtilis*. *Protein Eng 9*:439–445.

Gallagher T, Alexander P, Bryan P, Gilliand GL. 1994. Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. *Biochemistry 33*:4721–4729.

Garcia AE. 1992. Large-amplitude nonlinear motions in proteins. *Phys Rev Lett 68*:2696–2699.

Gronenborn AM, Clore GM. 1993. Identification of the contact surface of a streptococcal protein G domain complexed with a human Fc fragment. *J Mol Biol 233*:331–335.

Gronenborn AM, Filipula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. 1991. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science 253*:657–661.

Harbury PB, Tidor B, Kim PS. 1995. Repacking protein cores with backbone freedom: Structure prediction for coiled coils. *Proc Natl Acad Sci USA 92*:8408–8412.

Janin J, Chothia C. 1980. Packing of alpha-helices onto beta-pleated sheets and the anatomy of alpha/beta proteins. *J Mol Biol 143*:95–128.

Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*:2577–2637.

Karplus PA. 1996. Experimentally observed conformation-dependent geometry and hidden strain in proteins. *Protein Sci 5*:1406–1420.

Kim CA, Berg JM. 1993. Thermodynamic β-sheet propensities measured unsing a zinc-finger host peptide. *Nature 362*:267–270.

Lazaradis T, Lee I, Karplus M. 1997. Dynamics and unfolding pathways of hyperthermophilic and a mesophilic rubredoxin. *Protein Sci 6*:2589–2605.

Lian LY, Barsukov IL, Derrick JP, Roberts GCK. 1994. Mapping the interactions between streptococcal protein G and the Fab fragment of IgG in solution. *Nature Struct Biol 1*:355–357.

Lian LY, Derrick JP, Sutcliffe MJ, Yang JC, Roberts GCK. 1992. Determination of the solution structures of domains II and III of protein G from *streptococcus* by ^{1}H nuclear magnetic resonance. *J Mol Biol 228*:1219–1234.

Lim WA, Sauer RT. 1989. Alternative packing arrangements in the hydrophobic core of λ repressor. *Nature 339*:31–36.

Mayo KH, Iliyina E, Park H. 1996. A recipe for designing water-soluble β-sheet-forming peptides. *Protein Sci 5*:1301–1315.

Minor DL, Kim PS. 1994. Measurement of the β-sheet-forming propensities of amino acids. *Nature 367*:660–663.

Orban J, Alexander P, Bryan P. 1992. Sequence-specific ^{1}H NMR assignments and secondary structure of the streptococcal protein G B2-domain. *Biochemistry 31*:3604–3611.

Orban J, Alexander P, Bryan P, Khare D. 1995. Assessment of stability differences in the protein G B1 and B2 domains from hydrogen-deuterium exchange: Comparison with calorimetric data. *Biochemistry 34*:15921–15300.

Ramachandran GN, Ramakrishnan C, Sasisekharan V. 1963. Stereochemistry of polypeptide chain configurations. *J Mol Biol 7*:95–99.

Richardson JS. 1977. β-sheet topology and the relatedness of proteins. *Nature 268*:495–500.

Richardson JS. 1985. Describing patterns of protein tertiary structure. *Methods Enzymol 115*:341–380.

Ryckaert JP, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J Comp Phys 23*:327–341.

Schultz Beardsley D, Kauzmann WJ. 1996. Local densities orthogonal to β-sheet amide planes: Patterns of packing in globular proteins. *Proc Natl Acad Sci USA 93*:4448–4453.

Sheinerman FB, Brooks CL III. 1997. A molecular dynamics simulation study of segment B1 of protein G. *Proteins Struct Funct Genet 29*:193–202.

Smith CK, Withka JM, Regan L. 1994. A thermodynamic scale for the β-sheet forming tendencies of the amino acids. *Biochemistry 33*:5510–5517.

Smith LJ, Mark AE, Dobson CM, Van Gunsteren WF. 1995. Comparison of MD simulations and NMR experiments for hen lysozyme. Analysis of local fluctuations, cooperative motions, and global changes. *Biochemistry 34*(34):10918–10931.

Su A, Mayo SL. 1997. Coupling backbone flexibility and amino acid sequence selection in protein design. *Protein Sci 6*:1701–1707.

Swindells MB, MacArthur MW, Thornton JM. 1995. Intrinsic Φ,Ψ propensities of amino acids, derived from the coil regions of known structures. *Nature Struct Biol 2*:596–603.

Tisi LC, Evans PA. 1995. Conserved structural features on protein surfaces. Small exterior hydrophobic clusters. *J Mol Biol 249*:251–258.

Van Buuren AR, Marrink SJ, Berendsen HJC. 1993. A molecular dynamics study of the decane/water interface. *J Phys Chem 97*(36):9206–9212.

Van Den Burg B, Dijkstra BW, Vriend G, Van Der Vinne B, Venema G. 1994. Protein stabilization by hydrophobic interactions at the surface. *Eur J Biochem 220*:981–985.

Van der Spoel D, Berendsen HJC, Van Buuren AR, Apol E, Meulenhoff PJ, Sijbers ALTM, Van Drunen R. 1995. *Gromacs user manual*. Nijenborgh 4, 9747 AG Groningen, The Netherlands. Internet: http://rugmd0.chem.rug.nl/~gmx.

Van Gunsteren W, Billeter S, Eising A, Hünenberger P, Krüger P, Mark A, Scott W, Tironi I. 1996. *Biomolecular simulation: The GROMOS96 manual and user guide*. Zürich, Groningen: Biomos b.v.

Van Gunsteren WF, Berendsen HJC. 1987. *Gromos manual*. BIOMOS, Biomolecular Software, Laboratory of Physical Chemistry, University of Groningen, The Netherlands.

Vriend G. 1990. WHAT IF: A molecular modeling and drug design program. *J Mol Graph 8*:52–56.

Walther D, Eisenhaber F, Argos P. 1996. Principles of helix-helix packing in proteins: The helical lattice superposition model. *J Mol Biol 255*:536–553.

Wong CF, Zheng C, Shen J, McCammon A, Wolynes PG. 1993. Cytochrome C: A molecular proving ground for computer simulations. *J Phys Chem 97*:3100–3110.