doi:10.1016/S0022-2836(02)00476-X available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2002) 320, 783–799

JMB

Structural Basis for Mobility in the 1.1 Å Crystal Structure of the NG Domain of *Thermus aquaticus* Ffh

Ursula D. Ramirez[1]†, George Minasov[1]†, Pamela J. Focia[1] Robert M. Stroud[2], Peter Walter[2], Peter Kuhn[3] and Douglas M. Freymann[1]*

[1]*Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Medical School Chicago, IL 60611, USA*

[2]*Department of Biochemistry and Biophysics, University of California, San Francisco San Francisco, CA 94143, USA*

[3]*Stanford Synchrotron Radiation Laboratory Stanford University Stanford, CA 94309, USA*

**Corresponding author*

The NG domain of the prokaryotic signal recognition protein Ffh is a two-domain GTPase that comprises part of the prokaryotic signal recognition particle (SRP) that functions in co-translational targeting of proteins to the membrane. The interface between the N and G domains includes two highly conserved sequence motifs and is adjacent in sequence and structure to one of the conserved GTPase signature motifs. Previous structural studies have shown that the relative orientation of the two domains is dynamic. The N domain of Ffh has been proposed to function in regulating the nucleotide-binding interactions of the G domain. However, biochemical studies suggest a more complex role for the domain in integrating communication between signal sequence recognition and interaction with receptor. Here, we report the structure of the apo NG GTPase of Ffh from *Thermus aquaticus* refined at 1.10 Å resolution. Although the G domain is very well ordered in this structure, the N domain is less well ordered, reflecting the dynamic relationship between the two domains previously inferred. We demonstrate that the anisotropic displacement parameters directly visualize the underlying mobility between the two domains, and present a detailed structural analysis of the packing of the residues, including the critical α4 helix, that comprise the interface. Our data allows us to propose a structural explanation for the functional significance of sequence elements conserved at the N/G interface.

Keywords: ultrahigh resolution; SRP; Ffh; GTPase; X-ray crystallography

Introduction

The signal recognition particle (SRP) is a phylogenetically conserved ribonucleoprotein that mediates co-translational targeting of nascent proteins to the membrane. SRP-mediated targeting is dependent on GTP binding and hydrolysis by two components of the pathway, SRP54 and SR, its membrane receptor. Homologs of these two proteins are found throughout evolution, and in prokaryotes they are termed Ffh and FtsY, respectively. Remarkably, the two proteins, the SRP GTPase and its receptor, each contains a structurally homologous two-domain module, the NG domain, which defines the SRP subfamily of GTPases. The structures of the NG domains from *Thermus aquaticus* and *Acidianus ambivalens* Ffh, and *Escherichia coli* FtsY, have been determined.[1–3] The Ffh NG comprises a four α-helix N domain and a G domain, which has similarity to other members of the GTPase superfamily of proteins (Figure 1). The N domain abuts the G domain across the α4 helix of that domain. (The α4 helix, so named with reference to the structure of small GTPases, is termed the α6 helix in other studies.[3]). The α4 helix is adjacent to one of the characteristic GTPase recognition motifs, termed motif IV.[4]

The structures of the complex of the *T. aquaticus* NG domain with GDP revealed that the relationship between the two domains is dynamic.[5] In the

† These authors contributed equally to this work.

Abbreviations used: ADP, anisotropic displacement parameter; SRP, signal recognition particle; TLS analysis, translation, liberation, and screw analysis.

E-mail address of the corresponding author: freymann@northwestern.edu

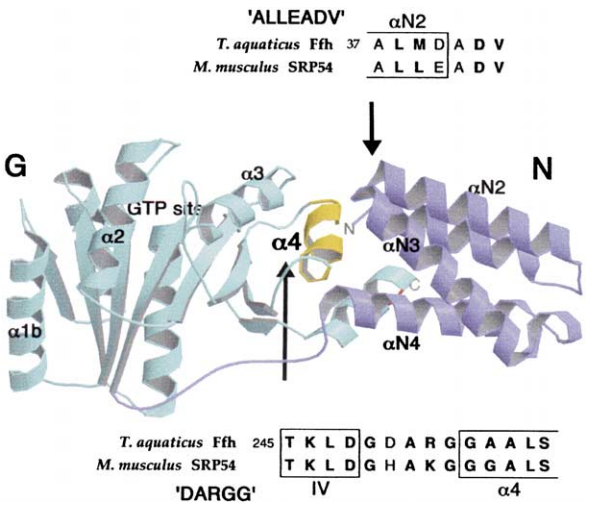

Figure 1. Two-domain structure of the SRP GTPase. The N domain, a four α-helix bundle, interacts with the G domain across the α4 helix of the G domain. The position of the GTP-binding site, on the far side of the plane of the image, is indicated. The α4 helix couples the position of the N domain to structural features of the nucleotide binding site. The highly conserved ALLEADV and DARGG sequence motifs contribute to the interaction across the N/G interface. The arrows point to the regions of the motifs from their respective sequence boxes. The ALLEADV motif includes the C-terminal end of αN2 and the turn between αN2 and αN3 (top arrow); the DARGG motif comprises the loop preceding the α4 helix and much of the helix itself (bottom arrow).

apo protein, the N domain is mobile, as evidenced by the fact that it can be trapped in different conformations under different crystallization conditions. The interface is closely associated with the position of the guanine recognition motif IV, so that its position is coupled to the position of that motif,[1] and, therefore, the relationship between the N domain and the G domain in the nucleotide-bound state is somewhat restricted. This coupling between the two domains, and the fact that the interface comprises sequence elements that are highly conserved between SRP GTPases (Figure 1), suggests that the N domain plays a role in the regulation of the nucleotide-binding state of the protein.

Although biochemical studies have elucidated, to some extent, the function of GTP binding and hydrolysis in SRP-mediated targeting, the role played by the N domain in this process remains unknown. The interface has been the focus of mutagenesis studies,[6,7] which suggest that this region plays a role in two different functionalities of the SRP GTPase; signal peptide recognition and interaction with the SRP receptor. In the first study, hydrophobic residues of the conserved sequence motif of the N domain part of the interface were changed to alanine, and it was found that the mutations caused defects in signal sequence recognition.[7] In the second study, a conserved glycine residue of the G domain motif was mutated, and it was found that this caused a defect in the interaction between SRP and receptor, but had little effect on GTP binding or hydrolysis.[6] These data are consistent with the N domain playing a central role in SRP-mediated targeting, as the NG interface is involved directly or indirectly in both the initial recognition of the signal sequence and the subsequent interaction with the membrane receptor.

Crystals of the apo NG domain of Ffh from *T. aquaticus* diffract to 1.0 Å resolution,[1] allowing us the opportunity to understand the protein structure in detail at ultrahigh resolution. The approximately tenfold increase in diffraction data relative to the previous report at 2.0 Å resolution[1] enables refinement of anisotropic displacement parameters (ADPs) and provides well-defined positions for most of the hydrogen atoms. Both allow us to define better the structural relationships that determine the function of the protein. The anisotropy gives information about the directionality of atomic and domain motions, and we find that it reflects directly the underlying mobility between the N and G domains. And, because tight protein packing is dependent on proper orientations of the small but numerous hydrogen atoms,[8] packing analyses of this structure reveals details that can explain the roles of the particular conserved hydrophobic side-chains at the N/G interface. Therefore, the increase in resolution between this structure and the previous one is significant, as it represents the difference between a model that can describe the protein packing only at the level of the heavy atoms of the structure, and one that defines the myriad of interactions between the hydrogen atoms of the structure that actually determine the fold. (In this context, it is worth noting that the difference is analogous to the transition that distinguishes the protein folding endgame.[9])

Our goal is to understand the structural basis for function in the SRP GTPases. Here, we focus on the interface between the N and G domains, taking advantage of the ultrahigh-resolution data to visualize directly the distribution of orientations of the two domains in the crystal. Packing analysis allows us to present a structural rationale for the conservation of residues central to the interface that is consistent with the functional roles suggested by biochemical studies. This work thus complements those studies. The α4 helix at the N/G interface is likely to play a central role in SRP GTPase function. In the *T. aquaticus* Ffh NG domain, the structural basis for its mobility is the persistence of tight packing interactions with the N domain, and loose packing with the underlying β-sheet that allow it to adjust to different conformations of the motif IV loop. Differences in the structures of the *T. aquaticus* Ffh NG and the available structures of other SRP GTPases can be understood in terms of this model.

Table 1. Data collection statistics

Space group	C_2	
Unit cell	$a = 99.73$, $b = 53.67$,	
a, b, c (Å), β (deg)	$c = 57.84$, β = 119.92	
Resolution (Å)	50.0–1.10	
Observations	694,235	
Unique reflections	102,368	
	50.0–1.10 Å	1.12–1.10 Å
R_{sym}[a]	0.037	0.324
Completeness (%)	95.4	89.9
Redundancy	4.27	2.96
Average I/σ	37.8	3.2

[a] $R_{sym} = \sum|I_h - \langle I_h\rangle|/\sum I_h$, where, $\langle I_h\rangle$ is the average intensity over symmetry equivalents.

Results and Discussion

Quality of the refined model

The structure of the Ffh NG domain was refined at 1.1 Å resolution using data measured from a single crystal at SSRL BL 9-1. Particular care was taken to ensure that the data were very complete, and the resulting dataset has excellent statistics (Table 1). Crystallographic refinement was carried out in several stages. The starting model for the refinement was the structure of the apo Ffh NG determined at 2.0 Å resolution (PDB ID 1ffh). Initial refinement was carried out using X-PLOR,[10] and the structure was subsequently refined using SHELXL-97[11] in order to introduce ADPs. The drop in R_{free} on adding anisotropy was ~4%, which is typical for structures at this resolution.[12,13] Following refinement with SHELXL, R_{cryst} was 13.3%; however, R_{free} remained relatively high, 18.7%, and a number of loops of the N domain were poorly ordered but could not be modeled in alternative conformations. Furthermore, although we had made a point of measuring complete data from the low to high-resolution limits, we were unable to use low-resolution data beyond 10 Å resolution because of inadequacies of the Babinet solvent model implemented in SHELXL. Therefore, the final stage of the refinement was carried out using the program REFMAC,[14] which implements a solvent model similar to that used in X-PLOR.[15] This stage resulted in little change in R_{cryst}, but yielded a significant 2% drop in R_{free} for all data over the full 50.0–1.10 Å resolution range (Table 2).

Table 2. Refinement statistics

Resolution (Å)	50.0–1.10
No. reflections	92,430
R_{cryst}[a] (%)	13.5
R_{free} (%)	16.9
Protein atoms	2513
Alt. conformations	30
Solvent atoms	326
Average B_{iso} (Å^2)	
Protein	22.6
Solvent	39.9
rms bond (Å)	0.024
rms angle (deg.)	2.086

[a] $R_{cryst} = \sum|F_o - F_c|/\sum F_o$, R_{free} was calculated for a test set of reflections (8%) omitted from the refinement.

The current model comprises 289 amino acid residues and 324 solvent molecules, with a crystallographic R value of 13.5%, and an R_{free} of 16.9%. The same subset of reflections was used for R_{free} throughout the refinement. The geometry of the model is quite good (Table 2) and the electron density map in the well-ordered core of the G domain is remarkably clear (Figure 2(a)). All residues are within the allowed (94.4%) or additional allowed (5.6%) regions defined by PROCHECK.[16] There is only one striking deviation in the main-chain geometry, which occurs at the Phe102/Leu103 peptide bond in the interior of the G domain. There, the omega torsion is 167°, and the conformations of the two residues, which occur in the interior of the protein, are very well-defined in the electron density map.

Thirty residues exhibit clearly defined multiple conformations in the electron density map (Figure 2(b)). Most occur on the surface and are poorly constrained but, interestingly, a number occur in the interior of the protein. Nine leucine residues, including the well-buried Leu103, exhibit discrete disorder; both Leu118 and Leu202 exhibit χ1 rotamer flips, three others exhibit χ2 flips, and the remainder undergo small positional shifts. In the N domain, 13 residues are in alternative conformations, with over half of those (Leu45, Arg49, Arg35, Met39, Lys83, Thr77, and Glu80) clustered across a crystallographic 2-fold screw axis that relates adjacent N domains. The multiple conformations of these residues reflect accommodation of different conformational substates that arise as a consequence of mobility between the N and G domains (see below).

A number of loops are poorly ordered in the structure and cannot be modeled as discrete alternative conformations. The electron density of the distal loops of the N domain, in particular, is remarkably smeared, presumably due to convolution of intrinsic disorder of the loops with the distribution of orientations of the N domain in the crystal. Poorly defined electron density in the region of the DARGG motif loop (Figure 1) that links motif IV with the α4 helix[1] may be functionally significant. Although the overall conformation of the region is discerned readily and is similar to that in previous apo and GDP structures, in several recent structures of the nucleotide-bound protein, the loop (residues 250–252) has been observed to be displaced due to crystal packing interactions, in a position that uncouples the N and G domains (P.J.F., H. Alam & D.M.F., unpublished results). Finally, more than 10% of the solvent atoms have been built into well-defined but "peanut-like" density features. These are consistent with, and have been modeled as, discrete and mutually exclusive solvent sites; however, we cannot exclude the

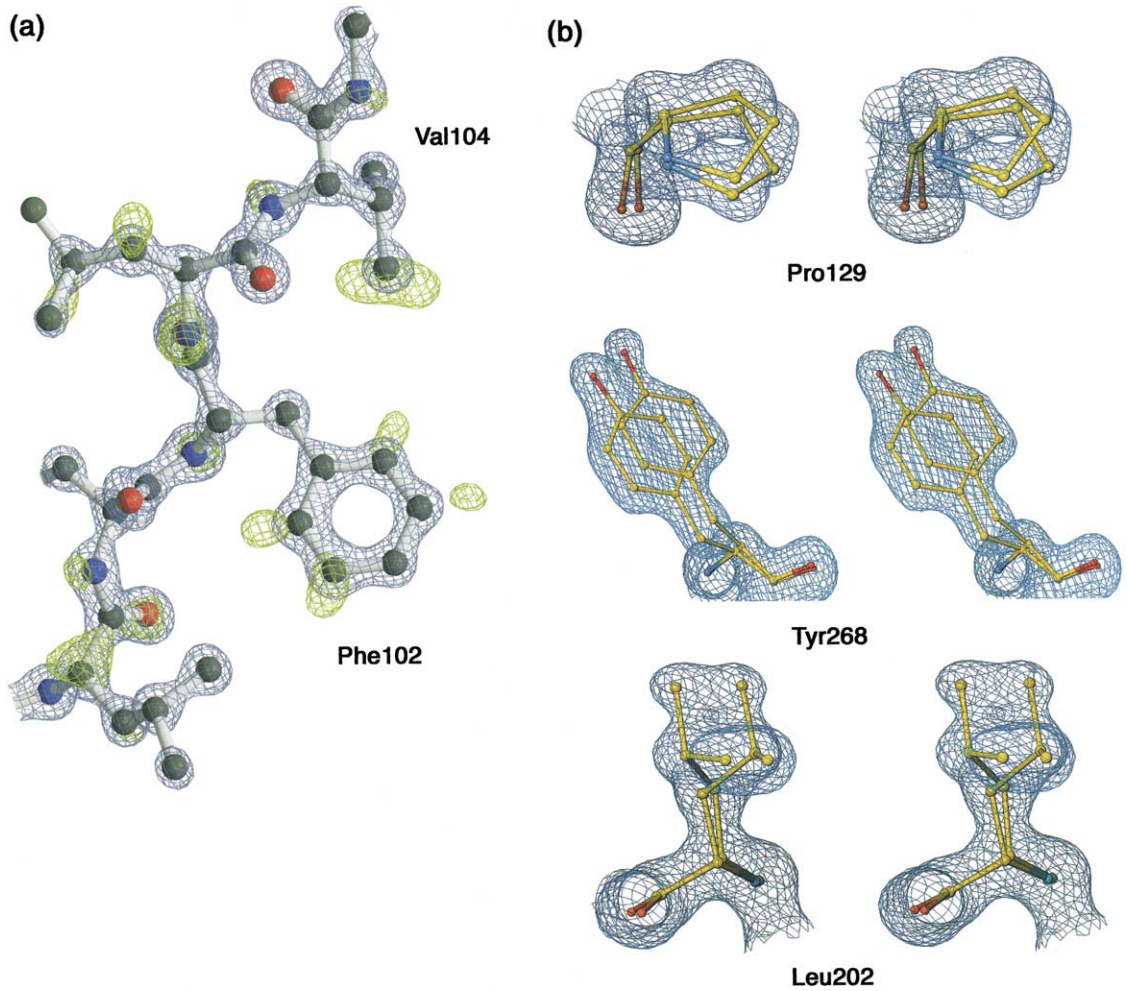

Figure 2. Electron density at 1.1 Å resolution. (a) Electron density in the well-ordered core of the G domain. The $2F_o - F_c$ map calculated using data from 50 to 1.1 Å resolution is contoured at 5σ. Superimposed is a hydrogen atom omit difference map, calculated using data from 5.0 to 1.1 Å resolution and contoured at 1.8σ. The positions of many hydrogen atoms are well defined in the well-ordered core of the G-domain (note the disordered methyl groups and backbone amide hydrogen atoms). (b) Residues in multiple conformations. In the first panel, the C^{γ} of proline occupies two positions; interestingly, at lower resolution in X-PLOR this residue refines as "flat". The final panel illustrates a typical leucine χ1 rotamer flip. The $2F_o - F_c$ electron density is contoured at 0.8σ.

possibility that they represent an uncharacterized solvent component.

Although the electron density map is excellent overall, its quality varies tremendously. Despite diffraction to 1.0 Å resolution, the equivalent isotropic *B* factors average ~25 Å^2, which is relatively high for a structure at this resolution.[17] This arises because two extensive crystal packing interactions stabilize the position of the G domain, but less well-defined interactions allow the N domain to occupy a range of conformational substates in the crystal. This is reflected as well in a skewed distribution of temperature factors between the N and G domains (see Figure 4), and it explains the relatively high crystallographic *R* and *B* factors in this structure. What is intriguing about the ultrahigh-resolution structure, however, is that it allows us to determine the directionality of the intrinsic intramolecular motion and propose a structural explanation for the functional mobility at the N/G domain junction.

Analysis of anisotropic displacement parameters

Following refinement of the ADPs, the average anisotropy (defined as the ratio of the minimum and maximum eigenvalue of the anisotropic displacement matrix) was found to be 0.55 ($\sigma = 0.13$), which is similar to the distribution seen in other proteins.[18] The mean value of the solvent anisotropy is ~0.40. We examined the distribution of anisotropy more closely and found that over much of the N domain, but not the remainder of the protein, the ellipsoids of motion of the alpha carbon atoms were clearly correlated (Figure 3). This suggested that there was a coherent motion of the domain, and was analyzed by applying a

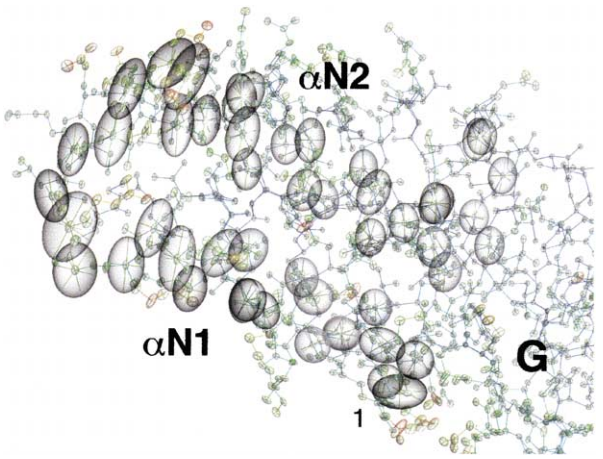

Figure 3. Anisotropic temperature factors reflect coherent motion of the N-domain. The anisotropic ellipsoids of motion at the 90% probability level for the α-carbon atoms of helices αN1 and αN2 highlight the mobility of the main-chain atoms of the N-domain. The principal axes of vibration along both of the helices are largely coherent (~up and down; compare with the ellipsoids near residue 1). The direction of this motion (presumably trapped substates in the crystal) is similar to that seen between the apo and GDP complex crystal structures.

rigid-body analysis of the domain movement in terms of a translation, liberation, and screw (TLS) model.[13] Because rigid-body motion is not considered explicitly during the refinement, an underlying anisotropic domain shift will be reflected in the atomic ADPs, and a description of the movement can be extracted from those ADPs.[19,20] There are two caveats to bear in mind: first, the diffraction data cannot themselves distinguish between correlated and uncorrelated motion,[20] and this inference has to be imposed during our interpretation of the data. However, as we discuss below, domain motion at the interface is well established, based on the comparison of the apo and GDP-bound structures, so this assumption is well justified in the structural analysis here. Second, the TLS parameters tend to over-fit the underlying motion, which may or may not be due to domain motions, and so provide only an upper limit on the rigid-body contribution.[13]

To carry out the analysis, we considered two domains of the protein as rigid bodies. The first comprised the helices of the N domain, residues 2–14, 26–59, and 71–88 (thus excluding the distal loops of the N domain) and helix α4 of the G domain, residues 253–262. The second comprised the rigid core of the G domain (residues 99–219). This is certainly a simplification of the dynamics of the protein; however, it works well for our purpose here, to characterize and explain the implicit motion of the N domain (which presumably arises as a series of conformational substates[21]) in the crystal. We used the program ANISOANL[22] to obtain the parameters of motion (translation, libration, and screw) for a rigid body comprising each domain from the distribution of atomic ADPs of the main-chain atoms. The derived TLS tensors were then analyzed using TLSANL.[23] That the TLS model accounts for much of the anisotropy observed in the structure is shown by the good fit between the model and the observed anisotropy and equivalent isotropic B values (Figure 4). The correlation coefficient over the equivalent B_{iso} values for both domains is ~0.85. The distal loops of the N domain, which are disordered, were not included in the TLS model and they degraded the fit when they were included.

The anisotropy observed in the structure appears to be fit particularly well by the TLS model (Figure 4(b)), with a correlation coefficient over both the N and G domains of 0.78. The mean translational displacement of the TLS component for the two domains was 0.20 Å and 0.16 Å, respectively, and the mean-square angles of libration[13] of the two domains were 1.78°^{2} and 1.28°^{2}, respectively. However, while the mean values were not dissimilar, the distribution of the displacements around the libration axes was quite different. Thus, for the G domain they range from 0.97 to 1.83°^{2}, while for the N domain, they range from 0.60 to 3.59°^{2}, a range of ~sixfold that is characteristic of anisotropic domain motion.[13,22] The major axis of libration of the N domain is approximately aligned with the αN4 helix (Figure 5(a)). This implies that the anisotropy of the domain arises predominantly from a longitudinal rotation of the N-domain helical bundle, and is remarkably reminiscent of the domain motion inferred from comparison of the apo and GDP-bound structures.[5]

To compare this explicitly, we carried out an analysis of the conformational change between the apo and ligand-bound states of Ffh NG in terms of a hinge rotation between the two domains using the model-independent analysis implemented in DYNDOM.[24] The resulting inter-domain rotation axis is consistent with and has similar directionality to the predominant libration axis of the N domain (Figure 5(b)). That they are not identical is, of course, expected, because both are idealizations of what is certainly a more complex structural change, and one represents a single hinge for rotation between the two conformational states, while the other represents a decomposition of the motion of the one domain. The scales of the two motions are also quite different (the mean rotation around the major libration axis in this structure is only 3.59°^{2}, while the hinge motion between the two domains that accompanies the conformational change on binding GDP is ~9°. Nevertheless, their approximate coincidence confirms that the directional preference for motion of the N domain is consistent with the conformational change between the apo and GDP-bound states, and this motion is, therefore, likely to be functionally significant.

It appears, therefore, that the relatively high temperature factors for the N domain are not due

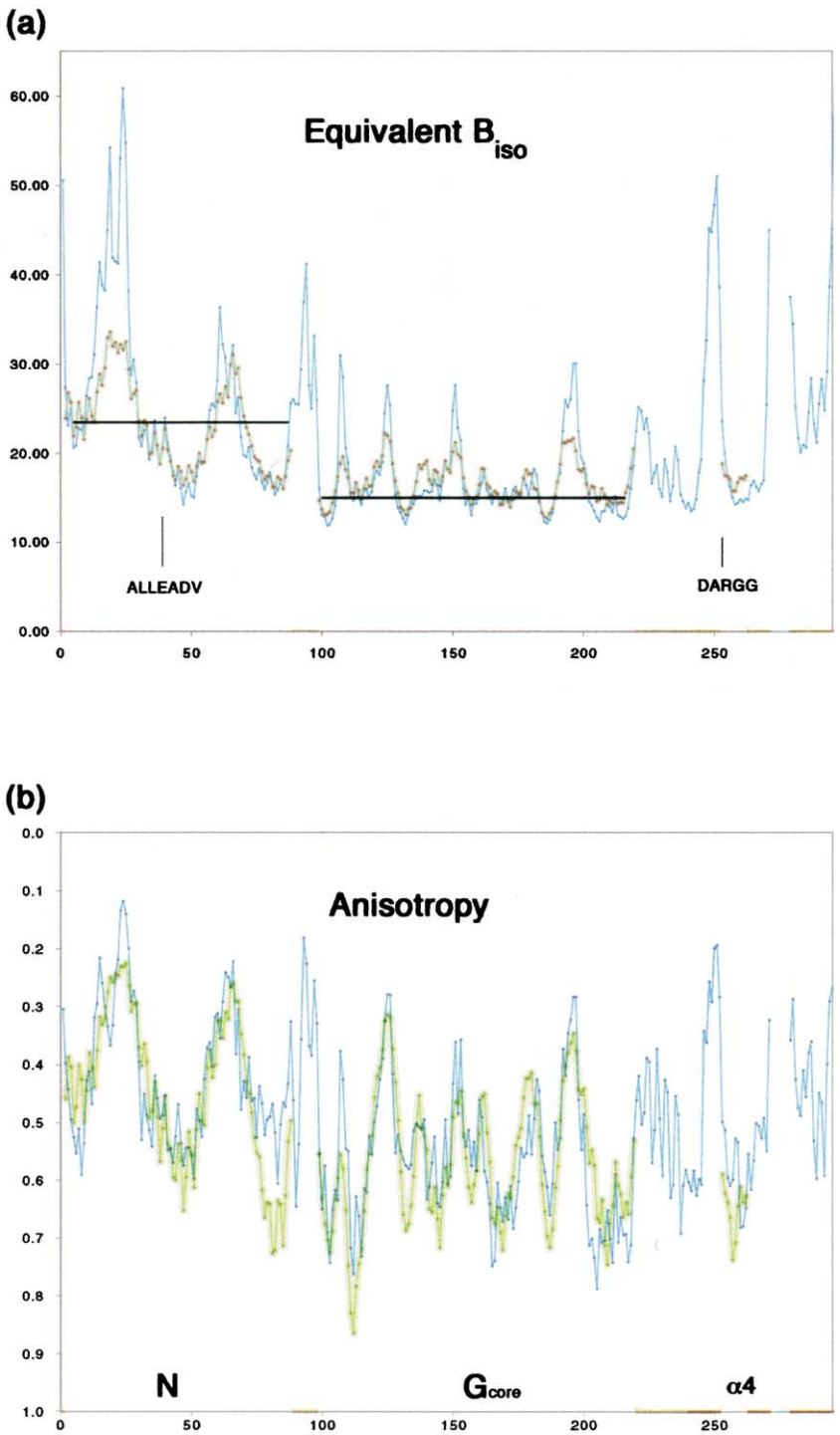

Figure 4. A rigid-body model accounts for much of the mobility of the N and G domains. TLS parameters were derived for rigid bodies taken as (i) the N domain and α4 helix (residues 2–88 and 253–262) and (ii) the core of the G domain (residues 99–219). In the upper plot, (a), the observed equivalent isotropic B values for each group are compared to the isotropic displacement parameters predicted from the rigid-body model (shown as the diffuse yellow-green line superimposed on the graph). The horizontal lines indicate the average B_{iso} for each of the domains. The loops of the N domain, which are poorly fit, exhibit very high temperature factors and are disordered in the electron density maps. In the lower plot, (b), the anisotropy in the refined structure is compared to the anisotropy predicted from the rigid-body model (shown as the diffuse yellow-green line). Note that anisotropy is defined such that isotropic 1.0 → 0.0 highly anisotropic.

to random displacements in the crystal, but instead arise from correlated motion of the N domain that is accommodated within the crystal packing (as evidenced, for example, by clusters of side-chains in multiple conformations) to yield a distribution of conformational substates. However, because it reveals directly a biologically-relevant structural variation, and because the protein model is very well refined with an R_{free} of ~17%, the structure affords the opportunity to examine in unprecedented detail the structural elements, primarily contributed by two highly conserved sequence motifs in the hinge region (see Figure 1), that enable motion across the N/G domain interface.

Analysis of the N/G domain interface

The N/G domain interface is unique to the SRP GTPase subfamily, and is characterized by a number of conserved sequence motifs (Figure 1). The interface between the two domains comprises the ALLEADV motif of the N domain,[7] elements

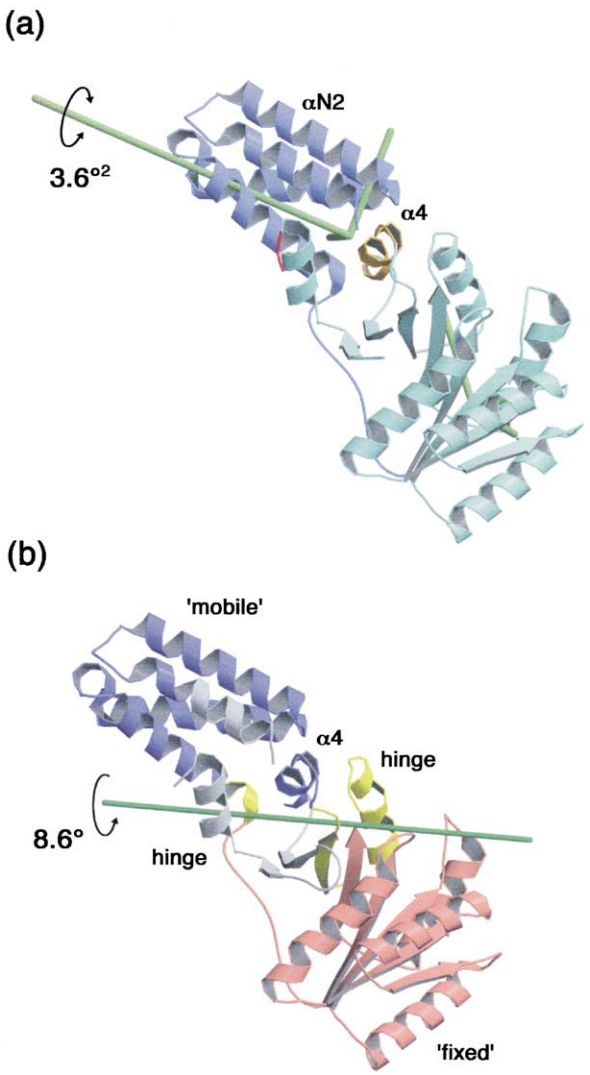

Figure 5. TLS analysis of the anisotropic displacement parameters reveals a prominent N-domain libration. (a) The libration axes of the TLS matrices derived separately for the N and G domains are plotted and superimposed on a ribbon representation of the structure. The length of each axis is proportional to its mean-square libration. The most striking feature is a predominant libration axis that lies approximately along the αN4 helix of the N domain. (b) The vector of rotation obtained following model-independent analysis of the conformational change between the structures of the apo and Mg^{2+}GDP-bound Ffh NG.[5] The vector generated by DYNDOM[24] is superimposed on the structure color-coded to distinguish a "hinge" region (in yellow) from the two domains (taken as mobile and fixed). Note that the α4 helix of the G domain is paired with the N domain in this model-independent analysis. Note also that the hinge axis is not coincident with, but is consistent with, the major axis of libration in the TLS model. However, the scales of the two motions are quite different.

of helix α3 of the G domain, the DARGG motif and the α4 helix that follows it,[1] and a number of water molecules. The functions of the solvent-exposed residues of the interface motifs remain largely unknown, as only a few interactions are defined in the available crystal structures[2,5] and there is little biochemical evidence for their function. Instead, it is the residues that contribute to the hydrophobic core of the interface that have been shown to be sensitive to mutation.[6,7] Mutagenesis of the leucine, valine and glycine residues of the ALLEADV and DARGG motifs produce readily detectable phenotypes, while mutations of the hydrophilic residues do not. Furthermore, the effects of these mutations are functionally far-reaching. Although they do not affect the intrinsic GTP-binding and hydrolysis activity of the proteins, different mutations cause defects in the recognition of the signal sequence of the nascent polypeptide, which is mediated primarily by the C-terminal M-domain,[25,26] and in the formation of the SRP/SR complex. The data imply that the junction between the domains and, presumably, their relative orientation is of central significance to the function of the SRP GTPase.

As a first step towards understanding the molecular architecture of the interface in atomic detail, we carried out packing analysis using a number of computational tools that quantify and visualize the molecular goodness-of-fit.[8,27,28] These tools, implemented in the program PROBE, explicitly evaluate hydrogen atom contacts, and so are particularly effective when used with structures refined at ultrahigh resolution.[27,29] The hydrophobic nature of a number of the functionally significant residues of the N/G interface makes the interface particularly accessible to this type of analysis. Thus, by taking advantage of the well-refined structure of the Ffh NG domain, we can clearly understand the packing interactions of these residues, and this, because their interactions are primarily hydrophobic, provides us with information about their functional roles.

The analysis carried out using PROBE allowed a direct visualization of the packing interactions of each atom as well as a quantitative measure of packing density in terms of a "score" that reflected the fraction of the available atomic surface area participating in packing interactions. A qualitative indication of the extent of atomic packing interactions over the entire structure is provided by a contour density plot,[28] and in this representation a number of features of the protein packing are highlighted (Figure 6). Of course, the core of the protein exhibits higher packing density than the surface, but a number of other features, such as intrahelical packing, stand out as having high "density" as well. In the structure of the Ffh NG domain, however, this representation reveals an additional and unexpected feature—a packing density hole that occurs adjacent to the N/G domain interface (Figure 6(a)). This region of low packing density does not occur between the two domains, but instead occurs between the α4 helix of the G-domain and the underlying β-sheet of the G-domain. The feature can be understood as arising from a particular pattern of residue

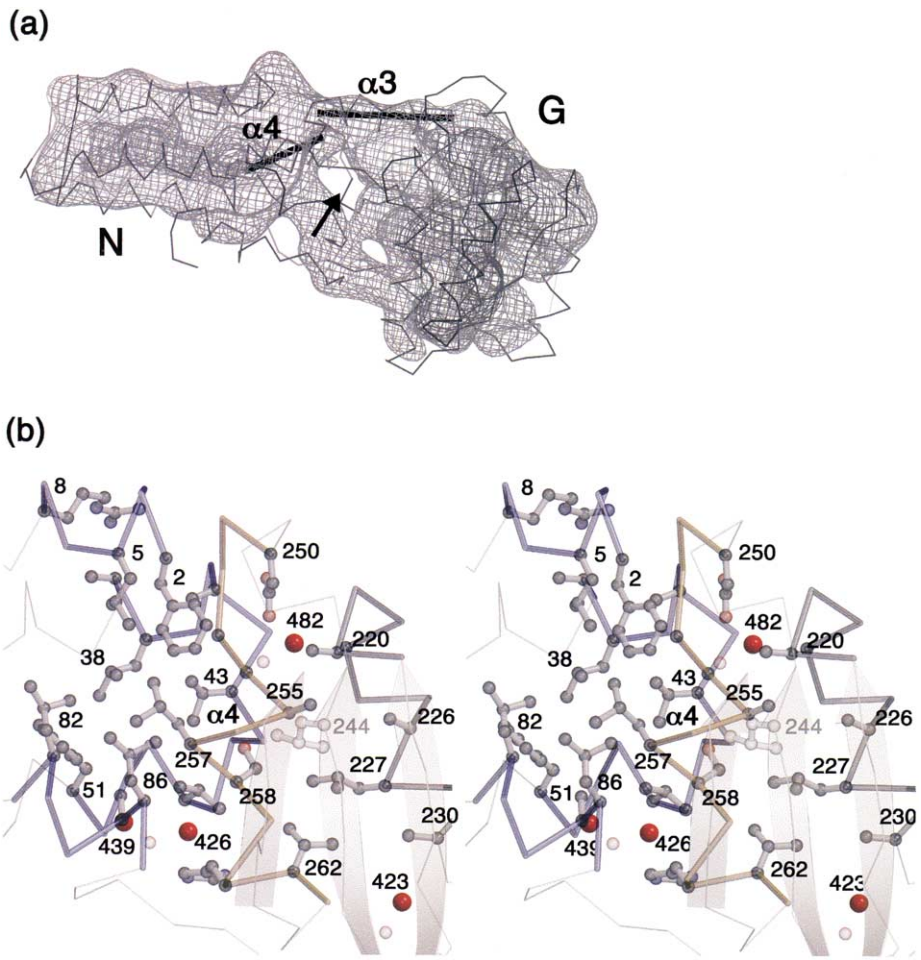

Figure 6. Contact density analysis reveals a loosely packed hinge region. (a) A contour representation of the packing density is superimposed on the α-carbon backbone of the NG domain. The large hole between the two domains represents a region of low packing density between the α4 helix of the G domain and the underlying β-sheet that may activate it to shift across the β-sheet during binding and release of nucleotide. The contour level corresponds to an atomic packing density of approximately 20% of the total available. (b) A stereo overview of the residues at the N/G interface. In this view, the ALLEADV motif (residues 37–43) is on the far side the α4 helix (orange). The N domain is on the left, and the G domain on the right. Residues discussed in the text are labeled, and the water molecules discussed in the text are highlighted. Three strands of the β-sheet of the G domain are indicated, in the right foreground.

conservation (see below). It occurs just above Wat423, which can, therefore, be considered to plug the hole (Figure 6(b)).

A packing cavity at the N/G interface is intriguing, of course, as it suggests an underlying design that facilitates movement of the N domain and helix α4 against the G domain consistent with that inferred previously from the comparison of the apo and GDP-bound structures (see Figure 5(b)).[5] Indeed, visualization of the packing interactions between the α4 helix and the N and G domains separately (Figure 7) reveals a clear clustering such that the α4 helix is packed tightly (i.e. has many interactions) with the conserved ALLEADV motif of the N domain. It interacts also with the α3 helix of the G domain, which has been shown to move in concert with it.[5] In contrast, the helix has remarkably few packing interactions with the underlying β-sheet of the G domain (Figure 7(b)). The helix, therefore, can be considered to be packed with the N domain rather than the G domain from which it arises, and we propose that this "unpacking" from the G domain occurs in the SRP GTPase so that the helix can facilitate the relative motion of the two domains.

Packing interactions of conserved residues at the interface

The residues of the ALLEADV motif occur at the turn between helices αN2 and αN3 of the N domain (Figures 1 and 6(b)). The motif shows some variation in sequence between different species (in the *T. aquaticus* Ffh, the sequence of the motif is ALMDADV); however, the conformation of the motif is well-conserved.[1,2] The solvent-exposed residues Met39 and Asp40 have no obvious interactions at the interface, and it has been speculated that they, along with conserved residues of the DARGG motif and the α3 helix of

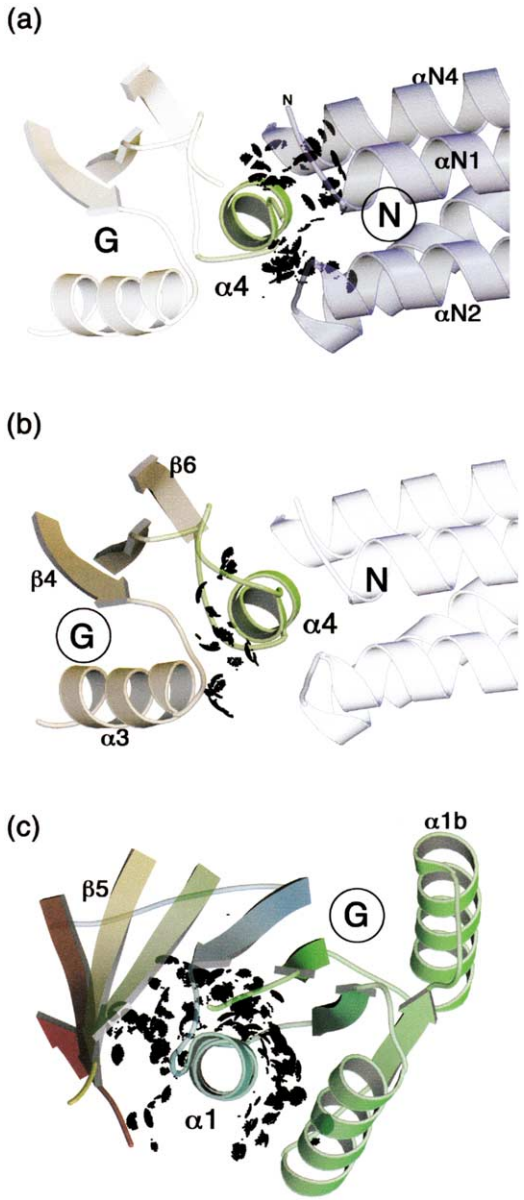

Figure 7. PROBE contact dots on the N domain side, but not the G domain side, of helix α4 indicate tight packing. Small probe contact dots[8] are shown in black between the α4 helix and surrounding areas of the N and G domains. (a) Contact dots from residues of the α4 helix to only the N domain. The N terminus (residues 1–4), the loop between helices αN2 and αN3 (ALLEADV motif), and the C-terminal end of αN4 exhibit numerous interactions with the α4 helix. (b) Contact dots from residues of the α4 helix to the G domain. Interactions are primarily with residues 220–224 of helix α-3. (c) A similar analysis of the packing of helix α1 of the G domain. This helix, which contributes the phosphate binding P-loop, motif I, is tightly packed relative to the other helices of the NG domain (see Table 3).

the G domain, contribute to a functionally important intermolecular interaction surface.[2,5] Asp42 reaches across the interface toward the G domain, where it contributes a hydrogen bond at the amino terminus of helix α3.[5] The remaining residues of the motif, all hydrophobic, contribute to the core packing of the N domain and the N/G interface. Significantly, many of their packing interactions are with other highly conserved residues (Figure 6(b), and see Figure 8). Thus, Ala37 packs against the aliphatic chain of Arg8 and against Leu5, and Ala41, completely buried, contributes extensive interactions to the interface between the aliphatic chain of Arg252, the N domain, and Leu5. Val43 seals off the hydrophobic core of the N domain, as it rests against α4 of the G domain, and is situated adjacent to Leu257 of the α4 helix. Mutation of the two hydrophobic residues of the motif corresponding to Leu38 and Met 39 in *T. aquaticus* in the eukaryotic SRP54 to either valine or alanine causes defects in signal sequence recognition in *in vitro* translation/targeting assays.[7] Leu5, which is involved in packing interactions with a number of residues of the motif, is very highly conserved in the SRP GTPases.[30] Although its function has previously been obscure, the interactions identified above suggest that Leu5 contributes important packing interactions to the residues that form the interface between the two domains, and so can be considered part of the N/G interface of Ffh.

The conserved DARGG motif of the G-domain comprises the residues Asp250 through Ser258, and includes the loop that precedes helix α4 and most of the helix itself (Figures 1 and 6(b)). The first three residues are solvent-exposed, and their function is defined only partly. Asp250 hydrogen bonds to Wat481, which helps initiate the α4 helix and both Ala251 and Arg252 are somewhat disordered, with the position of the arginine side-chain very poorly defined at the surface. The remainder of the motif is buried, however, and packing analysis illuminates a number of functionally significant features. Both the glycine residues, in particular, are very tightly packed at the N/G interface (see Table 3). Gly253, highly evolutionarily conserved, is situated between the side-chains of Phe2 and Asp250 such that the alpha carbon hydrogen atoms are in extensive van der Waals contact with the phenyl group, on the one side, and with Asp250 and Wat481 at the amino terminus of helix α4, on the other (Figure 8(a)). Gly254, completely conserved in the SRP GTPases, plays a central role in the interface between the two domains. Its alpha carbon atom is nestled against the 42/43 peptide bond, with extensive packing interactions against the main-chain and side-chain atoms of both residues. The carbonyl oxygen atom of residue 41 is adjacent, and is in position to form a long hydrogen bond to the backbone amide group of the glycine residue. Not surprisingly, the packing density scores for Gly253 and Gly254 are approximately double that of average glycine residues in this structure, with Gly253 having contacts over 42% of its possible contact surface area and Gly254 having contacts over 34% of its possible contact surface area (Table 3). The two glycine residues clearly function here to allow both tight

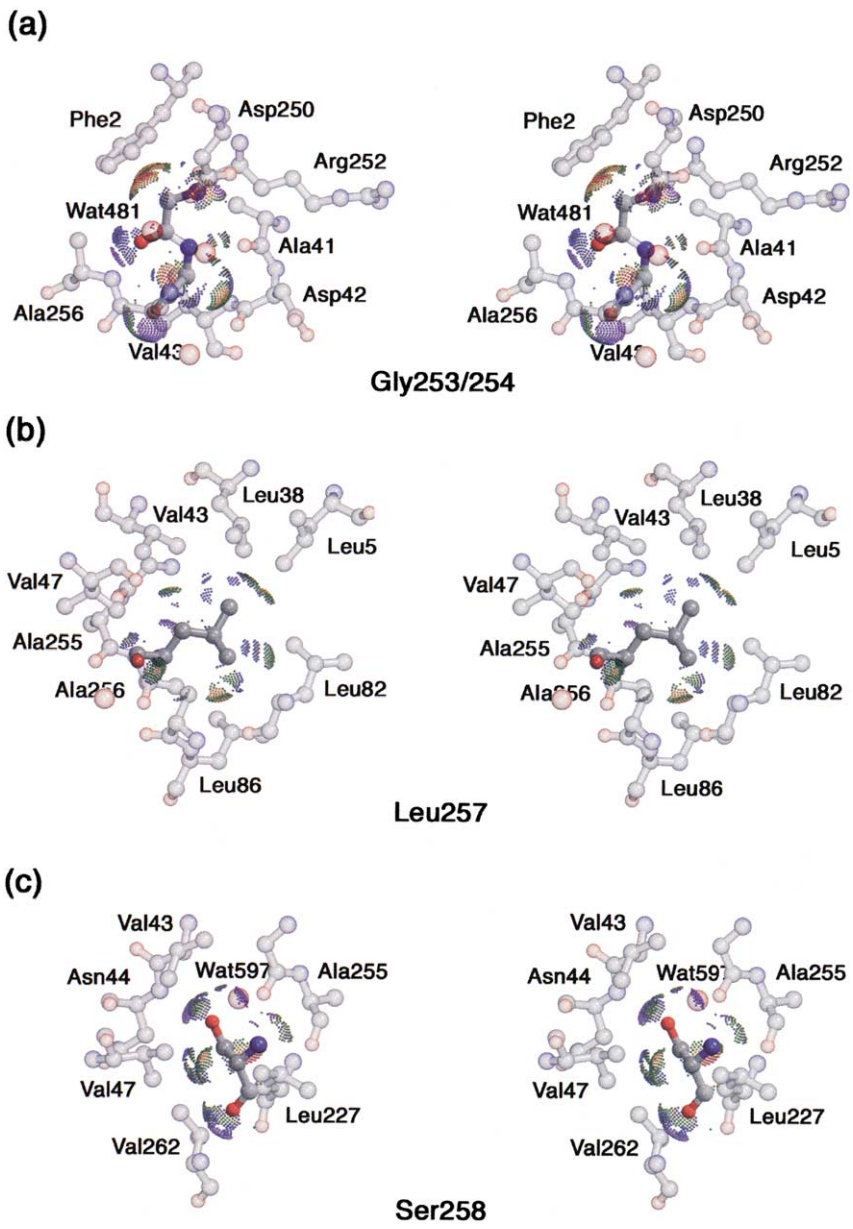

Figure 8. Conserved residues at the NG interface contribute to a well-packed interface. Small probe contact dots[8] are shown at a dot density of 50 dots/Å^2. The color of the dots indicates the types of interactions, with blue dots indicating wide contacts, greater than 0.25 Å, green indicating close contacts between 0 Å and 0.25 Å, yellow indicating small overlaps of less than 0.2 Å, and purple indicating hydrogen bonds. Hydrogen atoms are included in the analysis, but are not shown in the Figure. (a) The packing density around Gly253 and Gly254 (at the center) is especially tight for glycine residues and thus explains conservation of these two residues. (b) Leu257 appears to have relatively few interactions; however, those are with highly conserved residues, including Leu5, Val43, Leu38, and Leu82. Note the interaction between Leu5 and Leu257, which involves methyl hydrogen atoms of the two residues. (c) Interactions with Ser258, which is a completely conserved residue in the SRP GTPases. Both the Gly253/Gly254 and Ser258 interactions serve as the "glue" across the N/G interaction surface.

packing of the α4 helix and the close apposition of the N and G domains. It is mutation of the second of these glycine residues that has been shown to cause a functional defect in the *E. coli* Ffh.[6]

In contrast to the well-packed glycine residues of the interface, the conserved Ala255 that follows them appears to be much less constrained. The side-chain of the alanine residue is directed towards the β5-α3 loop of the G domain and, while the residue is completely buried (Table 3), it has remarkably few interactions with the remainder of the polypeptide. Its packing interactions are limited to Wat481 and Wat482, which initiate the helix α4, and to the carbonyl group of Gly223. Significantly, the alanine residue is directed towards the packing hole between the α4 helix and the β-sheet of the G-domain. In Ffh and SRP54, the residue at the corresponding position occurs only as alanine in prokaryotes and only as glycine in eukaryotes. Ala255 therefore plays a central, and presumably functionally significant, role in creating the packing hole below the α4 helix.

The side-chain of Leu257 is directed from the α4 helix into the center of the N-domain helix bundle. There, it is situated by packing interactions with surprisingly few residues: Leu5, Leu38, Val43, Leu82 and Leu86 (Figure 8(b)). Indeed, Leu257 has an average packing density score compared to other leucine residues (Table 3), substantially less than the maximum observed in this structure (41%). However, of the five residues that the leucine packs against, two, Val43 and Leu82, are conserved completely, and two, Leu5 and Leu38, are highly conserved in the SRP GTPase subfamily. This implies that although the packing interactions are relatively sparse, they maintain a precise structural relationship; one that would be disrupted,

Table 3. Packing scores and solvent accessibilities

	Density score[a]	Deviation[b]	Accessibility (Å^2)
A. *ALLEADV 37–43*			
Ala37	28	0.7	0.3
Leu38	35	+1.2	1.6
Met39	14	0.1	77.1
Asp40	13	−0.5	77.2
Ala41	35	+1.4	0.1
Asp42	18	0.1	51.6
Val43	31	0.4	2.1
B. *DARGG 250–258*			
Asp250	18	0.1	11.7
Ala251	10	−1.1	40.8
Arg252	16	0.1	124.3
Gly253	42	+3.1	0.0
Gly254	34	+2.4	1.0
Ala255	16	−0.5	0.1
Ala256	24	0.3	0.1
Leu257	24	0.1	2.3
Ser258	43	+1.3	0.0
C. *G-domain helices*[c]			
	Density score[a]	Deviation[b]	Context[d]
α1	23	1.4	P-loop
α1a	11	−0.9	IBD
α1b	12	−0.7	IBD
α2	12	−0.7	
α3	16	0.2	
α4[e]	8	−1.5	N/G interface

[a] Density score: % of total possible van der Waals contact surface, evaluated using at 1.0 Å PROBE sphere.
[b] Deviation from mean in standard deviations for that residue type (residues) or all helices.
[c] Scores are evaluated for the helix packing against the remainder of the G-domain.
[d] Context: identifying context for that helix.
[e] The α4 helix when considered in the context of both the N and G domains has a density score of 22.

for example, by the substitution of leucine for valine or *vice versa*. Their highly conserved interactions, therefore, may be functionally significant, and we speculate that they serve to facilitate or accommodate the mobility of the N/G domain interface.

The residue that follows, Ser258, is, like the glycine residues of the motif, particularly well-packed in the *T. aquaticus* Ffh NG, with a packing density 1.3 σ above the average for serine residues (Table 3). The serine hydroxyl group accepts a hydrogen bond from the amide nitrogen atom of Asn44 (which follows the ALLEADV loop) and interacts with weakly bound Wat597. It also donates a hydrogen bond back to the α4 helix main chain. The Cβ atom is positioned by packing interactions with Val47, Leu227 and Val262, which function, therefore, to bridge the N and G domains near the C-terminal end of the α4 helix.

The key conclusion from these observations is that highly conserved residues of the two sequence motifs contribute to specific packing interactions at the N/G interface that incorporate a number of additional highly conserved residues to assemble a specific structural relationship between the two domains. There is tight packing at the interface between the N domain and helix α4, and there is loose packing between the helix and the underlying β-sheet. This supports the notion that the α4 helix is designed to move in concert with the N domain instead of being a fixed part of the G domain, and is consistent with the observation of intrinsic anisotropy of the N domain arising from the loose fit at the junction.

Three well-ordered water molecules also contribute to the N/G interface; their positions are well-defined in the 1.1 Å resolution structure, and their interactions appear to be preserved between the different structures of the apo, Mg^{2+}GDP-bound, and GMPPNP-bound Ffh NG domain.[1,5,31] Two of the water molecules (Wat426 and Wat439) are associated with the packing between α4 and the N domain, and one (Wat423) occurs between the helices α3 and α4 and the underlying β-sheet (Figure 6(b)). Wat426 provides an important additional contact bridging the α4 helix of the G domain and αN4 of the N domain by hydrogen bonding the carbonyl oxygen atom of Ala85 and Leu257 and the amide nitrogen atom of His261. The water is particularly well-ordered in the different structures of the NG domain; in the 1.1 Å structure its equivalent B_{iso} value is 16.8 Å^2. Wat439, also well ordered, is linked to Wat426 by hydrogen bonding the intervening Wat440. Wat439 hydrogen bonds the carbonyl group of Val43 of the ALLEADV motif, and packs tightly against the side-chain of Val47 and the main chain of Phe51. Finally, Wat423 is positioned in a hydrophobic pocket between the C-terminal ends of the α3 and α4 helices by hydrogen bonds to the carbonyl group of Ala230 and to an adjacent Wat424. This water molecule effectively closes the divergent end of the packing hole that occurs between the α4 helix and the β-sheet. The conserved positions of these water molecules in this variety of structures suggests that they have a structural role; in particular, all can be considered to fill gaps in the fold by both hydrogen bonding and packing interactions, and they may function to maintain the ability of the two domains to move relative to one another, thus acting as "lubricants" for the polypeptide interface.

The design of the N/G interface

We identify the following features of the interface between the N and G domains of Ffh: first, on one side of the interface a well-defined network of side-chain and main-chain packing interactions contributes to the interface, and these interactions in the interior of the N-domain provide the first explanation for why specific hydrophobic residues are sparsely distributed but highly conserved in the N domain sequence. Second, on the other side of the interface, a packing defect is maintained by conservation of a number of small hydrophobic residues that enable mobility of the α4 helix relative to the remainder of the G domain. The fact that many of the residues involved are

evolutionarily conserved implies that these structural features are functionally important. Because it has been shown that the hydrophobic residues of the N/G interface can critically affect function,[6,7] the interactions we identify are, therefore, likely to be key to understanding the biological function of the protein. Thus, for example, the glycine residues allow close approach of the backbone of the ALLEADV motif to the loops of the G-domain because both are highly packed. The role of the glycine residues here is reminiscent of a similar role for glycine residues in the tight self-association of the glycophorin transmembrane helix.[32] Leu257, in contrast, although completely solvent-inaccessible, is relatively loosely packed. Nevertheless, its interactions are functionally significant, as they are largely with other highly conserved residues of the N domain. Thus, instead of providing a "glue" for association, as apparently in the case of the two glycine residues, the leucine residue may provide a loosely constructed joint that contributes to the conformational freedom of the domain interface. Finally, the alanine residues of the α4 motif are directed towards the underlying β-sheet, and have relatively few interactions. These define one edge of the functionally significant packing hole discussed above.

The α4 helix shifts over the underlying β-sheet on the transition between the apo and Mg^{2+}GDP complexes of the NG domain.[5] The movement is tangential to the β-sheet and results in displacement of the helix by ~1.5 Å along its axis.[5] The precisely located interactions we identify and the preferential directionality of those packing interactions provide a means to understand and predict, at the structural level, the dynamics of this interface. As we do not yet have structures of the NG nucleotide complexes beyond ~2.0 Å resolution, we cannot provide a definitive analysis of the shift between the two structural elements that accompanies nucleotide binding. What is clear, however, is that there is no large conformational change or structural reorganization. There are no residues in alternate conformations in the interior of the interface; only one residue bordering the packing hole, Leu244, undergoes a rotamer shift between the apo and Mg^{2+}GDP complex structures. This shift, a relatively minor χ2 side-chain reorientation, is relatively isosteric, and does not appear to change the packing requirements significantly. Leu244 occurs as valine, isoleucine or leucine in the SRP GTPases, and is directed towards the hole so that it is packed relatively loosely. Thus, it appears that the side-chains buried in the interface are accommodated by relatively non-specific restraints on their position along the packing defect and the mobility of the α4 helix relative to the G domain is maintained.

Interestingly, the packing density of only two of the helices of the G domain varies significantly from the average (Table 3). Helix α4 is the most loosely associated, with only 8% of the available surface contributing to packing interactions with the G-domain (compared to an average of ~15%). In contrast, helix α1 of the G domain is exceptionally well-packed against the core protein structure (Figure 7(c)), with 23% of its surface in packing interactions (2σ above average). It can, therefore, be considered to be relatively rigid, consistent with its central position providing the immobile platform of the phosphate-binding P-loop in the GTPase active site. The remaining helices of the G domain (α1a, α1b, α2, and α3) exhibit lower packing densities than helix α1, but are more tightly associated than helix α4. It is intriguing to propose that, perhaps in contrast to the mobile glide surface provided by helix α4,[33] these helices switch between specific conformational states. It is still not known whether the helices of the IBD (α1a, α1b) undergo conformational change, but conformational change of the α3 helix has been observed directly in the comparison of the apo and GDP-bound structures, and it can be inferred for helix α2 by analogy to the behavior of the helix in other GTPases during GTP binding and hydrolysis.

Comparison with other NG domain structures

Unfortunately, two structures of the NG domains of Ffh and FtsY from different species reveal a number of conformational differences that limit the generalization of our analysis. First, the structure of the *E. coli* FtsY NG domain reveals an irregular conformation of the region corresponding to αN1 in Ffh such that the N terminus extends beyond the helical bundle of the N domain.[3] More significantly, the published structure of the *A. ambivalens* Ffh NG domain[2] differs remarkably from structure of the *T. aquaticus* Ffh NG domain in precisely the region of the N/G interface discussed above. However, the constructs used for the latter two structures differ in significant ways. The *A. ambivalens* construct is a C-terminal His_6-tagged protein in which the tag is inserted just before a completely conserved leucine residue[30] that precedes a poorly ordered linker peptide to the C-terminal M-domain. In the structure of the *T. aquaticus* Ffh NG (and in the structure of *E. coli* FtsY NG), two hydrophobic residues of the final turn of the C-terminal helix of the G domain pack against the open face of the N-terminal domain hydrophobic core.[1] In the *A. ambivalens* Ffh NG, only one hydrophobic residue, Leu293, is available, and this may have consequences for both the stability of the N terminus of the protein (the first two residues of the αN1 helix are disordered) and for the position of the N domain. Thus, relative to the G domain, the N domain is shifted by ~3 to 4 Å from its position in the NG domain of *T. aquaticus* Ffh. Interestingly, the conformation of the ALLEADV motif is similar in the two structures (despite sequence differences[2]), and the loops and beta-strands of the G domain that precede and follow helix α4 are in similar positions. Remarkably however, helix α4 itself is found to be

out of phase by one residue between the two structures. Superimposition of the G domains of the two proteins overlays (within 1 Å on alpha carbon atoms) Arg252 with Lys254 (of *A. ambivalens*) at the N terminal end of the helix, and Thr263 with Thr265 (of *A. ambivalens*) at the C terminal end. Different loop conformations, however, displace the ten intervening residues such that, although the helix is clearly in a similar overall position (with corresponding alpha carbon atoms 0.7–2.3 Å apart), the sequence is shifted so that each residue at position N in *T. aquaticus* is at the position of the residue $N+1$ in the *A. ambivalens* structure. Thus the glycine residues discussed above no longer pack against the ALLEADV motif in the *A. ambivalens* structure, and the leucine residue (Leu257 in *T. aquaticus*) is not directed to the hydrophobic core of the N domain. Indeed, the serine residue is shifted so that it is directed instead in that direction.

Given the observation that the available *A. ambivalens* Ffh and *E. coli* FtsY structures differ from the structure of the *T. aquaticus* Ffh, what confidence can we have in our analysis? First, the residues of the interface are highly conserved in the SRP GTPase family, and the structure of the *T. aquaticus* Ffh, but not that of the *A. ambivalens* Ffh, allows us to rationalize this sequence conservation as reflecting a role in the functionally important interface between the two domains. The residues have been demonstrated in biochemical studies to be important for function,[6,7] and are seen in our structure to be involved in a large number of specific packing interactions at the interface. The conformation observed in the *A. ambivalens* structure is less consistent with the biochemical studies,[6] and, we argue, makes less structural sense than does the conformation observed in the structures of the *T. aquaticus* NG domain. For example, in the *A. ambivalens* structure, the glycine residue corresponding to Gly254 is directed to a large space between the N terminus and motif IV that would easily accommodate a larger side-chain. Further structural studies will be required to resolve this discrepancy.

Alternatively, that the other structures differ may suggest a conformational complexity to the interface we have yet to understand. Perhaps the differences hint that the N domain structure changes during the functional cycle, perhaps by unfolding so that it does not function as a four-helix bundle. Indeed, recent studies of the *T. aquaticus* Ffh and FtsY that suggest a specific conformational change of the N domain of FtsY occurs during nucleotide-dependent formation of the targeting complex.[34]

Conclusion

GTPases undergo conformational changes on binding and hydrolyzing GTP, and an understanding of their molecular mechanisms is important. The structures of many nucleotide complexes of signal transduction and translation GTPases have revealed a number of large conformational changes that accompany binding. These are generally localized to a set of structurally homologous "switch" regions that surround the active site and exhibit structural properties that enable them to undergo conformational changes during binding and hydrolysis of GTP. The α2 helix of the canonical GTPase fold, for example, can be considered to have, at least, two different stable packing arrangements that function during the GTPase cycle.[35]

The SRP GTPases, while members of the GTPase superfamily, are distinct. As yet, the nature of the conformational changes that accompany GTP activation and hydrolysis have not been determined.[31] The N domain, unique to the SRP GTPases, is distant from the switch regions that surround the active site; however, it now appears that it plays a role in integrating the distinct signal transduction events that occur during co-translational protein targeting.[6] One feature of its function is that it is in dynamic relation to the G domain. Here, we take advantage of X-ray diffraction data to 1.1 Å resolution to demonstrate that under the crystallization conditions the position of the N domain varies in a way consistent with the motion inferred from comparison of the structures of the apo and GDP bound complexes. Further, we show by packing analysis that the α4 helix of the G domain is a key player in this movement, as, on the one hand, it is coupled tightly to the N domain through interactions across a conserved interface and, on the other, it has few packing constraints against the underlying β-sheet of the G domain. Thus the α4 helix couples the N domain to the G domain, and facilitates its motion. The high-resolution packing analysis further explains cryptic patterns of sequence conservation in the SRP GTPases, most notably that of hydrophobic residues at the proximal core of the N domain that pack against the conserved leucine side-chain of the α4 helix.

What our data do not explain, however, is why the different structures of SRP GTPases from other species have, so far, not revealed a consistent arrangement of the two domains. The possibility is raised that the large deviations between members of the SRP GTPase family actually reflect exploration of a large conformational space accessible to the N domain sequence during SRP targeting. That is, the structures hint that the four α-helix bundle of the N domain is not static, a possibility that may itself be consistent with loose packing of the domain relative to other four-helix bundles. However, the well-defined structure of the domain interface in the ultrahigh-resolution structure of the *T. aquaticus* Ffh NG is consistent with the results of mutational studies, and therefore, if it defines only one endpoint of a range of mobility, it defines a significant one.

The 1.1 Å resolution structure of *T. aquaticus* Ffh NG domain expands the basis for future studies of the structural mechanism of the SRP GTPases.

The structure reported here directly visualizes a functionally significant motion that occurs at the interface of the N and G domains of the SRP GTPase, Ffh. It will be of interest to see, as the structures of the nucleotide bound complexes are determined at ultrahigh resolution, whether we can develop predictive algorithms that, based on packing analysis and an understanding of the anisotropic motion in each structure, can explain the intrinsic motions of this family of proteins. Mutations that affect the packing (and thus, presumably, the relative orientation) of the two domains produces striking defects, implying that the interface between the two domains is exquisitely designed for specific function. Because we identify a "neighborhood" of conserved residues that contribute to packing relationships on the tight side and on the hole side of the interface, our data help explain why the mutations of hydrophobic residues have such a significant effect. The intimate relationship between the N domain and the α4 helix of the G domain, and the packing interactions that facilitate mobility between the N and G domains, are therefore, likely to be important in understanding the structural mechanisms by which Ffh and SRP54 integrate signal sequence recognition, GTP binding and hydrolysis, and targeting of nascent polypeptides to their membrane-associated receptor.

Materials and Methods

Crystallization and data collection

The NG domain of *T. aquaticus* Ffh was expressed in *E. coli*, and purified as described.[1] Protein was concentrated to 33 mg/ml in water, and was crystallized by sitting-drop, vapor-diffusion over 30% PEG MME 550, 200 mM $MgCl_2$, 50 mM *N*-tris[hydroxymethyl]methyl-3-aminopropanesulfonic acid (Taps) (pH 9.0) at room temperature. Large bipyramidal crystals grew over several months to ~800 μm maximum dimension. The crystal was mounted on the surface of a ~700 μm nylon loop, so that it was surrounded by relatively little residual mother liquor and frozen immediately in the −170 °C N_2 cryostream.[36]

Data were measured at SSRL BL 9-1 on a MAR 300 image plate detector using a wavelength of 0.98 Å. Initial ultrahigh-resolution data collection experiments carried out on SSRL BL 7-1 established that these crystals exhibited a decay in the ultrahigh-resolution diffraction signal over a relatively short period of time. Our goal was to obtain complete, well-measured data from a single crystal; therefore, data were measured in three resolution segments, and the high-resolution data measured at four different translations of the crystal in the beam. Low-resolution datasets were obtained at the start and end of data collection to monitor overall decay. Low and medium-resolution data were collected with constant exposure time, the high-resolution data with constant dose. Data were integrated using DENZO and the seven individual datasets were scaled together using SCALEPACK[37] with a -3σ cutoff. The overall R_{sym} was 3.7% over the resolution range 50–1.10 Å, and was 32.4% in the high-resolution shell (1.12–1.10 Å). The data quality statistics are summarized in Table 1.

Refinement

The crystallographic refinement was begun using the model of the apo NG domain (PDB ID 1ffh) determined at 2.0 Å resolution.[1] The unit cells of the crystals used for the 2.0 Å and 1.10 Å datasets were similar (space group C_2, cell parameters $a = 99.90$ Å, $b = 53.91$ Å, $c = 57.36$ Å, $\beta = 119.8°$ and $a = 99.73$ Å, $b = 53.67$ Å, $c = 57.84$ Å, $\beta = 119.9°$, respectively) so that it was straightforward to begin several cycles of conventional positional and simulated-annealing refinement with X-PLOR 3.851.[10] $2F_o - F_c$ and $F_o - F_c$ electron density maps were used for manual corrections of the refined model using the graphics program O.[38] The refinement using isotropic temperature factors converged at an R_{cryst} value of 23.3% (with an R_{free} value of 27.2%) and at this stage the model included three residues with alternate conformations, 120 water molecules and one hydrated magnesium ion at a crystal lattice contact. The crystallographic R at this stage was somewhat higher than expected, although not unprecedented.[39,40] The position of the disordered closing loop noted in the 2.0 Å resolution structure was not resolved in the higher-resolution electron density map.

Subsequent introduction of ADPs in SHELXL resulted in a relatively large decrease of over 4% in both R_{cryst} (to 18%) and R_{free} (to 21%). At this point, the ratio of observations to parameters was about 4.2. Test refinements were carried out to optimize the stereochemical weighting factors by monitoring R_{cryst} and R_{free}, and the anisotropic restraints were optimized by validating the distribution of ADPs with PARVATI.[18] Occupancies of alternative conformations were refined using the FVAR instruction. After each manual rebuilding, the corrected parts of the model were refined with isotropic displacement parameters for five to ten cycles before the introduction of individual ADPs. Introduction of explicit hydrogen atoms in the latter stages of refinement caused a further drop of ~1% in R_{cryst}. The program REFMAC[14,41] was used during the final stages of the refinement in order to apply an empirical solvent model, which allowed us to include all low-resolution data to 50 Å resolution. The R_{cryst} value changed little at this stage, but we observed a significant ~2% drop in R_{free} (from 18.7 to 16.9%). The resulting electron density maps were somewhat better defined than those produced after SHELXL refinement (which included data from 10–1.1 Å only). Statistics from the final REFMAC refinement are summarized in Table 2. Multiconformer refinement[42] was not carried out.

Analysis

The ADPs were validated using PARVATI,[18] as noted above. We found that as restraints towards isotropy were released, the average anisotropy moved towards 0.5; further weakening of the restraints increased the number of highly anisotropic outliers but did not shift the overall average significantly. The atomic ADPs were then analyzed in terms of a TLS model for domain motion[19,20,43] using the programs ANISOANL[22] and TLSANL.[23] The first TLS domain was defined to include the four α-helices of the N domain and the α4 helix of the G domain; the second comprised residues 99–217 of the G domain. Following least-squares refinement of the

TLS parameters, the *R* values for the fit were 0.23 and 0.24, and the goodness of fit parameters were 18.18 and 17.72, respectively, indicating that the two TLS domains were modeled equally well. The correlation coefficient over the equivalent *B* values for the N domain was 0.88, and for the G domain it was 0.84.

Atomic packing analysis including explicit hydrogen atoms was carried out using the programs PROBE and REDUCE.[8] The PROBE algorithm simulates rolling a 0.25 Å radius sphere along the van der Waals surfaces; where the probe sphere contacts two surfaces, each is marked with an indication (a dot) that classifies whether the surfaces are in wide contact (>0.25 Å apart), close contact (0.25–0.0 Å), overlapped (interpenetrating to −0.25 Å), or clashing (> −0.25 Å). Bad contacts, or clashes, can be considered to indicate inaccuracies in the model and are found to increase with decreasing crystallographic resolution.[8] Therefore, to normalize for the lower effective resolution of the N domain, packing density scores were calculated throughout using an average of the score including bad overlap indicators and the score without them. The explicit hydrogen atoms used were placed (and in some cases refined) by SHELXL and/or REFMAC; their positions were essentially the same as those generated by REDUCE. The distribution of contact dots was contoured and displayed using ScoreDotsAtAtom and Kin3Dcont.[44] For Table 3, a packing density score was calculated for each residue by first determining the van der Waals surface area accessible to a 1.0 Å radius probe in the context of the tripeptide in which it occurred in the structure. Since hydrogen is the smallest atom involved in packing contacts, the 1.0 Å probe, corresponding to the van der Waals sphere for hydrogen used in PROBE,[8] provided an estimate of the maximal possible contact surface area. The packing interactions of each residue were then evaluated using a 0.25 Å probe, and the fraction of the van der Waals area involved in packing interactions was tabulated. The ratio of the two was taken as a packing density score. Averages were calculated for each residue type in the structure; for glycine (23 residues) the average was 9.6%, and for leucine (32 residues) the average was 23.3%, with a maximum of 41.2% and a minimum of 1.4%. For scoring the packing of individual helices, the van der Waals surface of the helix as a whole was evaluated similarly in the context of the remainder of the structure. The contour level that illustrates the packing hole between helix α4 and the β-sheet of the G domain in Figure 6 corresponds to an atomic packing density of approximately 20% of the maximum. Surface accessibilities were calculated using AREAIMOL.[45]

The domain motion observed between the structures of the apo and Mg^{2+}GDP-bound NG domain (PDB ID 1ng1) was analyzed using DYNDOM.[46] The model-free method implemented in that program was found to be relatively insensitive to the input parameters, although the limits of the domains so identified varied somewhat. The α4 helix of the G domain was always observed to be grouped with the helices of the N domain, consistent with the model presented here. Least-squares superposition of structures was carried out using LSQMAN.[47] Figures were generated using MOLSCRIPT,[48] Raster3D and RASTEP,[49] MAGE[50] and O.[38]

Protein Data Bank accession codes

The atomic coordinates and structure factors have been deposited with the RCSB PDB, with accession code 1LS1.

Acknowledgments

We thank J. Brunzelle for interesting and helpful discussions, and we thank Garib Murshudov for discussions and assistance in implementing our refinement strategy. This work was supported by grant GM-58500 from the NIH. The data presented here were measured while D.M.F. was a postdoctoral fellow in the laboratories of P.W. and R.M.S. (UCSF). SSRL is supported by the DOE Office of Basic Energy Sciences. The SSRL Biotechnology Program is supported by the NIH, National Center for Research Resources, Biomedical Technology Program, and by the DOE Office of Biological and Environmental Research.

References

1. Freymann, D. M., Keenan, R. J., Stroud, R. M. & Walter, P. (1997). Structure of the conserved GTPase domain of the signal recognition particle. *Nature*, **385**, 361–364.
2. Montoya, G., Kaat, K. t., Moll, R., Schäfer, G. & Sinning, I. (2000). The crystal structure of the conserved GTPase of SRP54 from the archeon *Acidianus ambivalens* and its comparison with related structures suggests a model for the SRP–SRP receptor complex. *Structure*, **8**, 515–525.
3. Montoya, G., Svensson, C., Luirink, J. & Sinning, I. (1997). Crystal structure of the NG domain from the signal-recognition particle receptor FtsY. *Nature*, **385**, 365–369.
4. Bourne, H. R., Sanders, D. A. & McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. *Nature*, **349**, 117–127.
5. Freymann, D. M., Keenan, R. J., Stroud, R. M. & Walter, P. (1999). Functional changes in the structure of the SRP GTPase on binding GDP and Mg^{2+}GDP. *Nature Struct. Biol.* **6**, 793–801.
6. Lu, Y., Qi, H.-Y., Hyndman, J. B., Ulbrandt, N. D., Teplyakov, A., Tomasevic, N. & Bernstein, H. D. (2001). Evidence for a novel GTPase priming step in the SRP protein targeting pathway. *EMBO J.* **20**, 6724–6734.
7. Newitt, J. A. & Bernstein, H. D. (1997). The N-domain of the signal recognition particle 54 kDa subunit promotes efficient signal sequence binding. *Eur. J. Biochem.* **245**, 720–729.
8. Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K. *et al.* (1999). Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* **285**, 1711–1733.
9. Levitt, M., Gerstein, M., Huang, E., Subbiah, S. & Tsai, J. (1997). Protein folding: the endgame. *Annu. Rev. Biochem.* **66**, 549–579.
10. Brünger, A. T. (1992). *X-PLOR: A System for X-Ray Crystallography and NMR*, Yale University Press, New Haven, CT.
11. Sheldrick, G. & Schneider, T. (1997). SHELXL: higher resolution refinement. *Methods Enzymol.* **277**, 319–343.
12. Walsh, M. A., Schneider, T. R., Sieker, L. C., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1998). Refinement of triclinic hen egg-white lysozyme at atomic resolution. *Acta Crystallog. sect. D*, **54**, 522–546.
13. Schneider, T. R. (1996). What can we learn from anisotropic temperature factors? *Proceedings of the CCP4 Study Weekend* (Dodson, E., Moore, M., Ralph,

A. & Bailey, S., eds), pp. 133–144, SERC Daresbury Laboratory, Daresbury.
14. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
15. Jiang, J.-S. & Brünger, A. T. (1994). Protein hydration observed by X-ray diffraction. *J. Mol. Biol.* **243**, 100–115.
16. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
17. Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1995). Proteins at atomic resolution. *Curr. Opin. Struct. Biol.* **5**, 784–790.
18. Merritt, E. A. (1999). Expanding the model: anisotropic displacement parameters in protein structure refinement. *Acta Crystallog. sect. D*, **55**, 1109–1117.
19. Schomaker, V. & Trueblood, K. N. (1968). On the rigid-body motion of molecules in crystals. *Acta Crystallog. sect. B*, **24**, 63–76.
20. Dunitz, J. D., Schomaker, V. & Trueblood, K. N. (1988). Interpretation of atomic displacement parameters from diffraction studies of crystals. *J. Phys. Chem.* **92**, 856–867.
21. Rader, S. D. & Agard, D. A. (1997). Conformational substates in enzyme mechanism: the 120 K structure of alpha-lytic protease at 1.5 Å resolution. *Protein Sci.* **6**, 1375–1386.
22. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallog. sect. D*, **57**, 122–133.
23. Howlin, B., Butler, S. A., Moss, D. S., Harris, G. W. & Driessen, H. P. C. (1993). TLSANL: TLS parameter analysis program for segmented anisotropic refinement of macromolecular structures. *J. Appl. Crystallog.* **26**, 622–624.
24. Hayward, S. & Berendsen, H. J. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins: Struct. Funct. Genet.* **30**, 144–154.
25. Lütcke, H., High, S., Römisch, K., Ashford, A. J. & Dobberstein, B. (1992). The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences. *EMBO J.* **11**, 1543–1551.
26. Zopf, D., Bernstein, H. D., Johnson, A. E. & Walter, P. (1990). The methionine-rich domain of the 54 kDa protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. *EMBO J.* **9**, 4511–4517.
27. Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (1999). Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **285**, 1735–1747.
28. Word, J. M., Bateman, R. C., Jr, Presley, B. K., Lovell, S. C. & Richardson, D. C. (2000). Exploring steric constraints on protein mutations using MAGE/PROBE. *Protein Sci.* **9**, 2251–2259.
29. Lovell, S. C., Word, J. M., Richardson, J. S. & Richardson, D. C. (1999). Asparagine and glutamine rotamers: *B*-factor cutoff and correction of amide flips yield distinct clustering. *Proc. Natl Acad. Sci. USA*, **96**, 400–405.
30. Gorodkin, J., Knudsen, B., Zwieb, C. & Samuelsson, T. (2001). SRPDB (signal recognition particle database). *Nucl. Acids Res.* **29**, 169–170.
31. Padmanabhan, S. & Freymann, D. M. (2001). The conformation of bound GMPPNP suggests a mechanism for gating the active site of the SRP GTPase. *Structure*, **9**, 859–867.
32. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. (1997). A transmembrane helix dimer: structure and implications. *Science*, **276**, 131–133.
33. Gerstein, M., Lesk, A. M. & Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry*, **33**, 6739–6749.
34. Shepotinovskaya, I. V. & Freymann, D. M. (2002). Conformational change of the N-domain on formation of the complex between the GTPase domains of *Thermus aquaticus* Ffh and FtsY. *Biochim. Biophys. Acta*, **1597**, 92–99.
35. Sprang, S. R. (1997). G protein mechanisms: insights from structural analysis. *Annu. Rev. Biochem.* **66**, 639–678.
36. Bellamy, H. D., Phizackerly, R. P., Soltis, S. M. & Hope, H. (1994). An open-flow cryogenic cooler for single-crystal diffraction experiments. *J. Appl. Crystallog.* **27**, 967–970.
37. Otwinowski, Z. (1993). Oscillation data reduction program. In *Data Collection and Processing* (Sawyer, L., Isaacs, N. W. & Bailey, S., eds), pp. 55–62, SERC Daresbury Laboratory, Warrington.
38. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
39. Deacon, A., Gleichmann, T., Kalb, A. J., Price, H., Raftery, J., Bradbrook, G. *et al.* (1997). The structure of concanavalin A and its bound solvent determined with small-molecule accuracy at 0.94 Å resolution. *J. Chem. Soc., Faraday Trans.* **93**, 4305–4312.
40. Tong, L., Warren, T. C., King, J., Betageri, R., Rose, J. & Jakes, S. (1996). Crystal structures of the human $p56_{lck}$ SH2 domain in complex with two short phosphotyrosyl peptides at 1.0 Å and 1.8 Å resolution. *J. Mol. Biol.* **256**, 601–610.
41. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallog. sect. D*, **55**, 247–255.
42. Wilson, M. A. & Brunger, A. T. (2000). The 1.0 Å crystal structure of Ca(2 +)-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. *J. Mol. Biol.* **301**, 1237–1256.
43. Schomaker, V. & Trueblood, K. N. (1998). Correlation of internal torsional motion with overall molecular motion in crystals. *Acta Crystallog. sect. B*, **54**, 507–514.
44. Word, J. (2000). All-atom small-probe contact surface analysis: an information-rich description of molecular goodness-of-fit. PhD Thesis, Duke University
45. Collaborative Computational Project, No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
46. Hayward, S., Kitao, A. & Berendsen, H. J. (1997). Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins: Struct. Funct. Genet.* **27**, 425–437.

47. Kleywegt, G. J. & Jones, T. A. (1997). Detecting folding motifs and similarities in protein structures. *Methods Enzymol.* **277**, 525–545.
48. Kraulis, P. J. (1991). MOLSCRIPT—a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
49. Merritt, E. A. & Bacon, D. J. (1997). Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.
50. Richardson, D. C. & Richardson, J. S. (1992). The kinemage: a tool for scientific communication. *Protein Sci.* **1**, 3–9.

Edited by D. Rees

(Received 22 January 2002; received in revised form 8 May 2002; accepted 14 May 2002)