Quantitative Structural Analysis of Importin-β Flexibility: Paradigm for Solenoid Protein Structures

Jade K. Forwood, Allison Lange, Ulrich Zachariae, Mary Marfori, Callie Preast, Helmut Grubmüller, Murray Stewart, Anita H. Corbett, and Bostjan Kobe

SUMMARY

The structure of solenoid proteins facilitates a higher degree of flexibility than most folded proteins. In importin-β, a nuclear import factor built from 19 tandem HEAT repeats, flexibility plays a crucial role in allowing interactions with a range of different partners. We present a comprehensive analysis of importin-β flexibility based on a number of different approaches. We determined the crystal structure of unliganded *Saccharomyces cerevisiae* importin-β (Kap95) to allow a quantitative comparison with importin-β bound to different partners. Complementary mutagenesis, small angle X-ray scattering and molecular dynamics studies suggest that the protein samples several conformations in solution. The analyses suggest the flexibility of the solenoid is generated by cumulative small movements along its length. Importin-β illustrates how solenoid proteins can orchestrate protein interactions in many cellular pathways.

INTRODUCTION

Solenoid proteins, constructed from tandem structural repeats arranged in superhelical fashion, feature in many cellular processes (Kobe and Kajava, 2000). One such process is nucleocytoplastic transport, in which solenoid proteins constructed from HEAT repeats (the β-karyopherin superfamily) and armadillo (importin-α; Impα) repeats (Kobe et al., 1999; Peifer et al., 1994) constitute the principal transport receptors. A key structural property that differentiates solenoid proteins from other structured proteins is the lack of contacts between distal regions of protein sequence (sequence-distal contacts). For this reason, solenoid proteins are often more flexible than other structured proteins, and this flexibility is an important feature of their specific functions. In terms of flexibility, solenoid proteins are emerging as a structural class that falls between typical globular structured proteins and intrinsically unstructured proteins (Wright and Dyson, 1999).

Importin-β (karyopherin-β1; here abbreviated as Impβ1) is the most intensively studied member of the β-karyopherin superfamily. Its sequence consists of 19 tandem HEAT repeats, each repeat a structural unit comprised of two antiparallel (A- and B-) α helices (Andrade and Bork, 1995; Cingolani et al., 1999; Groves et al., 1999; Kobe et al., 1999). The units arrange into a solenoid, with the helices perpendicular to the solenoid axis (A helices on the concave face and B helices on the convex face). Impβ1 facilitates transport through the nuclear pore complex by transient interactions with FG-nucleoporins that line the central transport channel of the pore (Stewart, 2007). Impβ1 can carry into the nucleus a range of cargo proteins, which either bind to Impβ1 directly or through adaptor proteins such as Impα and snurportin1. The Impα/β-mediated nuclear import is considered the most widely used pathway (Lange et al., 2007). Cargo proteins bind to Impα through the classical basic nuclear localization sequence (cNLS); Impα in turn binds to Impβ through its Impβ1-binding (IBB) domain, and the trimeric complex enters the nucleus. In the nucleus, the complex is dissociated by binding of RanGTP to Impβ1. The direction of the transport is determined by the nucleotide state of Ran (nucleus: RanGTP; cytoplasm: RanGDP), which is in turn established by localization RanGAP (GTPase-activating protein) and RanGAP (guanine nucleotide exchange factor, RCC1) to the cytoplasm and the nucleus, respectively (Lange et al., 2007; Stewart, 2007).

The structure of Impβ1 has been studied extensively, using full-length and fragments of human (hImpβ1), mouse (mImpβ1) and *S. cerevisiae* (Kap95; here abbreviated as yImpβ1) proteins bound to different partners (Suel et al., 2006), such as Ran (Forwood et al., 2008; Lee et al., 2005; Vetter et al., 1999), nucleoporins (Bayliss et al., 2000; Liu and Stewart, 2005), and cargo proteins (Cingolani et al., 1999, 2002; Lee et al., 2003; Mitrousis et al., 2008; Wohlgend et al., 2007). Comparison of structures of Impβ1 bound to these different proteins shows large conformational differences (Cingolani et al., 2000; Conti et al., 2006; Bhardwaj and Cingolani, 2010). The solenoid structure plays a central role through providing a large surface area...
that can adjust to different binding partners. The use of different conformational states for different binding functions may finely tune the energies of binding events and ensure the appropriate gradation of affinities in different stages of the pathway (Cansizoglu and Chook, 2007; Conti et al., 2006; Forwood et al., 2008; Fukuhara et al., 2004; Zachariae and Grubmuller, 2008). We present a comprehensive analysis of importin-β flexibility based on a number of different approaches. As the structure of unliganded full-length Impβ has previously only been studied by small-angle X-ray scattering (SAXS) (Fukuhara et al., 2004), we determined the crystal structure of unliganded yImpβ to allow a quantitative comparison with importin-β bound to different partners. We used a number of complementary approaches including molecular dynamics (MD), SAXS, quantitative geometric analyses and TLS (translation/libration/screw) analyses to compare the conformations of the different Impβ structures and analyze the molecular basis of the molecule’s flexibility. We conclude that Impβ samples several conformations in solution, which result from cumulative small structural changes along the length of the solenoid. Such flexibility may be important to store internal energy in the structure to control the binding of different partners along the nuclear transport pathway.

RESULTS AND DISCUSSION

Conformational Behavior of Unliganded S. cerevisiae Importin-β

Crystal Structure of Unliganded S. cerevisiae Importin-β

To enable a complete comparison to be made between different conformations of yImpβ, it was necessary to obtain a crystal structure of the unliganded molecule so that the effects of the binding of different partners could be assessed. We therefore crystallized and determined a 2.4 Å resolution crystal structure of yImpβ. The refined model (R = 18.6%, Rfree = 22.4%) (Table 1) includes all 861 residues of yImpβ and 168 water molecules. In this structure, the 19 HEAT repeats of yImpβ are arranged in a tightly coiled, compact conformation resembling the shape of a heart (Figure 1). The structure, including the crystal packing, resembles yImpβ in the yImpβ:Nup1 complex (Liu and Stewart, 2005) (rmsd 2.22 Å for 835 Cα atoms). Residues 142–861 (corresponding to HEAT repeats H4–19) superimpose particularly well, while HEAT repeats H1–3 diverge (Figure 1B). The similarity of these two structures suggests that Nup1 binding does not significantly alter the structure of Impβ. The unliganded yImpβ structure also resembles an unliganded mImpβ fragment (residues 1–454) closely (rmsd of 2.7 Å for 418 Cα atoms) (Lee et al., 2000; Figure 1C). The compact conformation appears to be mediated primarily through sequence-distal contacts involving HEAT repeats H2 and H4 interacting with H17 (the principal interactions involve Ser74 in Asp167 in H2 and Arg696, Glu737, Asn738 and Gly739 in H17; Table 2). These interactions bury 306 Å² of surface area, which is smaller than is normally seen in protein:protein interfaces and indicates that the amount of energy required to distort the flexible Impβ solenoid is relatively small. There are also interactions of H8 with H11 and H12, but similar interactions are found in other yImpβ structures (Forwood et al., 2008; Lee et al., 2005; Liu and Stewart, 2005) and so these are unlikely to contribute to the specific conformation of the unliganded state observed here.

The ring-like conformation found in the crystals of unliganded yImpβ is somewhat similar to exportins CRM1 (Dong et al., 2009a, 2009b; Monecke et al., 2009; Petosa et al., 2004) and unliganded Cse1 (Cook et al., 2005) (Figure 1A). The interactions that mediate ring formation are not conserved between Impβ and Cse1, which shows contacts between repeats H1–3 and H14–16. Because of the different roles in import and export, the physiological significance of these compacted structures is unlikely to be the same. In exportins, RanGTP and cargo binding in the nucleus cooperate to open up the structure and assemble the export complex (Cook et al., 2005; Dong et al., 2009a, 2009b; Petosa et al., 2004).

<table>
<thead>
<tr>
<th>Table 1. Structure Determination</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diffraction Data Statistics</strong></td>
</tr>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
</tr>
<tr>
<td>β (°)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>Unique reflections</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Multiplicity</td>
</tr>
<tr>
<td>Rmerge (%)</td>
</tr>
<tr>
<td>cRfinal</td>
</tr>
<tr>
<td><strong>Refinement Statistics</strong></td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>Number of unique reflections</td>
</tr>
<tr>
<td>Rcryst (%)</td>
</tr>
<tr>
<td>Rfree (%)</td>
</tr>
<tr>
<td>Number of Nonhydrogen Atoms</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Coordinate error (Å)</td>
</tr>
</tbody>
</table>

*Numbers in parentheses are for the highest resolution shell.

a Calculated with the program MOLPROBITY (Davis et al., 2007).

b Calculated with the program MOLPROBITY (Davis et al., 2007). The coordinate error is the maximum likelihood based coordinate error.
Small-angle X-Ray Scattering Studies of *S. cerevisiae* and Mouse Importin-β

Previous SAXS studies indicated that mlmpβ had an S-shaped conformation in solution (Fukuhara et al., 2004). We collected SAXS data to characterize the solution behavior of the yImpβ used in crystallization, and to compare it with mlmpβ. The scattering patterns of both yImpβ and mlmpβ display a shoulder at $q \approx 1 \text{nm}^{-1}$ (Figure 2A), as seen previously for unliganded mlmpβ, as well as β-karyopherins transportin and Xpo-1 (Fukuhara et al., 2004). The radius of gyration $R_g$ and the maximum dimension $D_{\text{max}}$ inferred from the pair distance distribution function (P(r), shown in Figure 2B) by the program GNOM (Svergun, 1992) were found to be 39 and 120 Å for yImpβ and 45 and 145 Å for mlmpβ, respectively. These results indicate that the average size of yImpβ in solution is larger than that seen in the crystals (calculated $R_g = 33$ Å, $D_{\text{max}} = 91$ Å), and smaller than for mlmpβ.

The low-resolution shapes modeled based on their scattering profiles (Figure 2C) show a coiled structure for both proteins, comparable to the S-like conformation seen previously (Fukuhara et al., 2004). yImpβ is more compressed along the superhelical axis than mlmpβ.

The theoretical scattering profiles of previously solved yImpβ atomic models (Protein Data Bank [PDB] ID 3ea5, 2bku, 2bpt; binding partner coordinates removed) were calculated by the program CRYSOL (Svergun et al., 1995), and are shown in Figure 2D. These theoretical curves also exhibit a shoulder at $q \approx 1 \text{nm}^{-1}$; however, they do not readily superimpose onto the experimental scattering pattern, clearly evident by large χ² values ($\chi^2 > 4.9$). This suggests that the average shape of the molecule in solution does not correspond to any available crystallographic model. The program OLIGOMER (Konarev et al., 2003) was used to assess whether the scattering pattern for yImpβ could be better represented as a linear combination of multiple crystallographic structures. Using the calculated scattering intensities of the closed unliganded yImpβ and the more open yImpβ::RanGDP structure (PDB ID 3ea5, Ran coordinates removed), an improvement in the χ² value ($\chi^2 = 4.3$) was observed (using volume fractions of 3ea5 and unliganded yImpβ of 84.3% and 15.7%, respectively; Figure 2D). Including additional

<table>
<thead>
<tr>
<th>yImpβ Mutations</th>
<th>S74K</th>
<th>F514K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutation Predicted to Enhance Intramolecular Interactions</td>
<td>Ser74</td>
<td>Asn738</td>
</tr>
<tr>
<td>Interaction</td>
<td>Glu737</td>
<td>Asp167</td>
</tr>
<tr>
<td>Predicted to Disrupt Intramolecular Interactions</td>
<td>Arg696</td>
<td>D167K</td>
</tr>
</tbody>
</table>

Table 2. yImpβ Mutations

Figure 1. Comparison of yImpβ, Ligand-Bound Impβ and Cse1 Structures

(A) The structures are shown in cartoon representation in analogous orientation with cylinders representing α helices, yImpβ, crystal structure of unliganded yImpβ (gray). Residues targeted in mutagenesis studies are shown in stick representation and colored: red, Ser74 (HEAT repeat H2)/Glu737 (H17); blue, Asp167 (H4)/Asn738 (H17); green, Glu341 (H8)/Phe514 (H12) and Asp343 (H8)/Lys468 (H11). yImpβ::RanGTP, structure of yImpβ (cyan) bound to Ran (magenta) and GTP (blue in stick representation) (PDB ID 2bku) (Lee et al., 2005). yImpβ::Nup1, structure of yImpβ (cyan) bound to Nup1 (red) (2bpt; Liu and Stewart, 2005). hImpβ::snuportin1BB, structure of hImpβ (cyan) bound to the IBB domain of snuprt1 (green) (Cingolani et al., 1999). hImpβ::snuportin1BB, structure of hImpβ (cyan) bound to the IBB domain of snuprt1 (yellow) (2q5d) (Mitrousis et al., 2008). mlmpβ::SREBP-2, mlmpβ (cyan) bound to a dimer of SREBP-2 (orange; 1ukl) (Lee et al., 2003). Cse1, structure of unliganded Cse1 (magenta) (123h) (Cook et al., 2005).

(B) Superposition of the structures of unliganded yImpβ (gray) and yImpβ (cyan, HEAT repeats H1-H3 blue):Nup1p (red) complex (PDB ID 2bpt).

(C) Superposition of the structures of unliganded yImpβ (gray) and unliganded mlmpβ fragment comprising residues 1–454 (green) (PDB ID 1gcj). See also Figure S1.
structures of yImpβ (PDB ID 2bku and 2bpt) did not result in a decrease of $\chi^2$ value, presumably because of the high similarity of these conformations to those seen in 3ea5, and the unliganded yImpβ. However, a slight improvement in the $\chi^2$ value (to 4.0) was achieved through the addition of the scattering intensity calculated from the crystal structure of mImpβ (PDB ID 1ukl, coordinates of SREBP-2 removed) (using volume fractions of 3ea5, unliganded yImpβ and 1ukl of 56.5%, 7.3%, and 36.2% respectively). Nonetheless, only $\chi^2$ values $<2$ are considered meaningful using this analysis. Altogether, the data suggest that the yImpβ is most likely sampling a range of conformations in solution, which favors a more elongated state than seen in the crystal. However, the spectrum of conformations that yImpβ adopts in solution cannot simply be described as a combination of the currently available crystal structures.

Molecular Dynamics Simulations of S. cerevisiae Importin-β

Previous MD simulations indicated that unliganded Impβ was able to undergo rapid and large conformational changes in solution (Zachariae and Grubmuller, 2008). We used MD to test if the structure of unliganded yImpβ observed in the crystals was likely to persist in solution. Starting the simulation with the structure observed in the crystal, rapid conformational changes occurred during the early stages of the simulations that led to a marked...
Structural Analysis of Importin-β Flexibility

**Figure 3. Molecular Dynamics Simulations**
(A) Evolution of radius of gyration of during molecular dynamics simulation. Left, yImpβ. The starting condition of the simulation was the crystal structure. The observed tendency toward larger Rg values during the first part of the simulation indicates an evolution toward a more elongated structure in solution; however, another compact state is regained around t = 17 ns, together with a transient reformation of the original sequence-distal contacts. Middle, yImpβ S74K mutant. Right, GFP.

(B) Snapshots of structures during the MD simulation of yImpβ. Left, the starting compact conformation corresponding to the crystal structure. Middle, an extended conformation after 15 ns of simulation. Right, a compact conformation after 25 ns of simulation.

acid substitutions affect the essential function of yImpβ in vivo. jRSL1 cells, which lack the gene encoding yImpβ but contain a yImpβ maintenance plasmid, were transformed with vector alone or plasmid encoding wild-type or mutant yImpβ. Cells were plated on control plates or on plates containing the drug 5-FOA (Boeke et al., 1987), which removes the maintenance plasmid and leaves the yImpβ mutant as the only cellular copy of yImpβ (Figure 4A). None of the substitutions significantly impacted growth and therefore yImpβ function. Growth curve analysis also did not show significant differences in growth (not shown). All proteins were expressed at similar level as wild-type yImpβ-GFP (Figure 4B). To verify that amino acid substitutions within yImpβ can indeed disrupt protein function, a control variant of yImpβ (L329T, L330T, L332L, L333T) was assayed by plasmid shuffle (Figure 4C). Cells expressing this variant yImpβ as the only copy of yImpβ were unable to grow, consistent with previous studies, which show that a even single amino acid change within yImpβ can impair its function in vivo (Iovine and Wente, 1997).

We further examined whether any of the amino acid substitutions had an impact on transport of cargo proteins into the nucleus, by assessing the steady-state localization of the model classical NLS-containing cargo SV40TAgNLS-GFP-GFP or GFP-GFP alone as a control, in cells containing either wild-type or mutant yImpβ as the only copy of yImpβ (Figure 5A). Cells were cotained with Hoechst to verify the position of the nucleus (data not shown). The double GFP tag was used to minimize passive diffusion through nuclear pores, though some degree of nuclear localization of GFP-GFP was expected (Hodel et al., 2006). For each yImpβ mutant, GFP-GFP alone localized throughout the cell and SV40TAgNLS-GFP-GFP showed significant steady-state nuclear localization, indicating that the substitutions do not significantly impair yImpβ function in the classical nuclear import pathway. Furthermore, the interaction between yeast Impβ (Kap60/Srp1) and the yImpβ variants is not significantly altered, because the heterodimeric import receptor must form for classical import to occur. There may be a slight decrease in the nuclear localization of the NLS reporter in the K468D mutant, but attempts to confirm through counting cells

elongation of the protein (during the first 2.5 ns, Rg increased from 32.5 to 36 Å, and the Ca rmsd relative to the initial structure to 6 Å; Figure 3). The sequence-distal contacts between H2/4 and H17 were not preserved. After subsequent contraction (t=3 ns, Rg = 3.4 Å) and expansion (t = 15 ns, Rg = 36.5 Å, rmsd = 8 Å), yImpβ started to enter into another phase of contraction after 15 ns and reformed a compact state with a number of loose sequence-distal contacts between H2/4 and H17, as well as H7 and H18/19. The structure after t=22 ns resembled the crystal structure (rmsd ~3 Å), and appeared to be stable over the course of several ns.

The SAXS profiles from a number of the MD conformations were calculated using CRYOSOL (Svergun et al., 1995; Figure 2D). Additionally, the SAXS curves of more elongated structures from previous simulations were also calculated (Zachariae and Grubmuller, 2008). Although some of the MD conformations showed an improved fit to the solution scattering curve, indicated in a decrease in χ² (values ranging from 3.2 to 7.8) compared with the crystallographic models (see previous section; χ² values >4.9), no single conformation could adequately describe the measured scattering profile of yImpβ. These results further support the interpretation that Impβ samples a range of conformations in solution.

**Mutational Analysis of S. cerevisiae Importin-β**
To assess the in vivo requirement for the amino acids that may stabilize the circular arrangement of unliganded yImpβ found in the crystals, we engineered amino acid substitutions in the regions of H2, H4, and H17 (S74K, E737K, and D167K). As a control, we also engineered analogous mutations in H8, H11, and H12 (F514K, E341A, and K468D; Figure 1 and Table 2; see Figure S1A available online). The mutations were designed to either disrupt or stabilize the observed structure. Several of these residues are conserved in hImpβ (D167 is conserved and E737, E341 and K468 conserve the charge) (Figure S1B). A plasmid shuffle assay was used to determine if the amino acid substitutions affect the essential function of yImpβ in vivo. jRSL1 cells, which lack the gene encoding yImpβ but contain a yImpβ maintenance plasmid, were transformed with vector alone or
in a double-blind experiment showed no statistically significant difference.

Finally, a kinetic import assay was used to determine if the initial rate of NLS-cargo import was altered. Cells containing mutant or control yImp\(^b\) as the only copy and expressing either GFP-GFP or SV40TAg-NLS-GFP-GFP were incubated with azide and 2-deoxy-glucose, which deplete the cell of energy and cause redistribution of any nuclear cargo throughout the cell (Shulga et al., 1996). After washing out the inhibitors, import kinetics were measured by assessing the percentage of cells with nuclear accumulation of the reporter over time (Figure 5B). The rate of initial import of the NLS-reporter was similar in all cells analyzed (including the K468D mutant).

To obtain additional evidence of the effect the mutations on the structure of yImp\(^b\), we performed SAXS measurements and MD simulations on the S74K mutant using the same conditions as for wild-type yImp\(^b\). Consistent with the functional experiments, the mutant behaved very similar to wild-type (Figures 2 and 3).

Figure 4. Functional Analysis of the yImp\(^b\) Mutants In Vivo
(A) S. cerevisiae cells deleted for the endogenous ylm\(^p\) (DJSL1) but containing a ylm\(^p\) maintenance plasmid (ACY208) were transformed with vector alone (negative control) or a plasmid encoding wild-type (positive control) or mutant ylm\(^p\). Cells were plated on control plates or on plates containing the drug 5-FOA and grown at 25° C. 5-FOA removes the maintenance plasmid and leaves the ylm\(^p\) mutant as the only cellular copy of ylm\(^p\).
(B) The level of each of the ylm\(^p\)-GFP variants expressed in wild-type cells (ACY192) was detected by immunoblotting with an anti-GFP antibody. Levels of a control protein, Yrb1, were monitored with an anti-Yrb1 antibody (Schlenstedt et al., 1995) to ensure that an equal amount of total protein was loaded in each lane.
(C) S. cerevisiae cells deleted for the endogenous ylm\(^p\) but containing a ylm\(^p\) maintenance plasmid were transformed with a plasmid encoding wild-type ylm\(^p\), vector alone, or variant ylm\(^p\), serially diluted, and spotted on control or 5-FOA plates. See also Table S1.
free R factor during the TLS refinement. As the TLS model approximates the molecular disorder increasingly closely, the free R factor should decrease as the refined model becomes more consistent with the diffraction data, whereas overrefinement would be flagged by an increase in the free R factor. When the 2.4 Å resolution unliganded ylmpβ structure was refined in this way, using TLS regions generated by the http://skuld.bmsc.washington.edu/~tlsmd/ website (Painter and Merritt, 2006), the free R factor decreased progressively as the molecule was divided into progressively smaller regions, until six to eight regions were employed. After this, further subdivision into smaller regions did not result in a decrease in the free R factor and indeed, it began to increase slightly indicating that the structure was becoming overrefined (Figure 6B). Moreover, when six segments were employed, the model was approximating the B factor fluctuations along the molecule reasonably well. TLS models based in individual HEAT repeats did not improve the free R factor over that obtained using six segments.

We also applied this analysis to the refinement of the 2.0 Å resolution structure of the ylmpβ:Nup1 complex (PDB ID 2bpt) (Liu and Stewart, 2005). For this higher resolution structure, the minimum Rfree was obtained with 13–14 TLS segments.
consistent with the greater number of observations enabling a more precise description of the anisotropic disorder to be made. Thus, for refinement of the structures of both the unliganded yImp\(\beta\) and its complex with Nup1, a progressively larger number of TLS blocks gave a progressively closer description of the motion of the molecule until a resolution-dependent limit was reached, after which the structure started to become overrefined. The overrefinement most probably resulted from the number of parameters (in this case, the number of segments) exceeding the limit set by the number of observations in the structure factor data set. This suggests that an even greater number of blocks would be appropriate, were it possible to obtain a higher resolution data set that contained a greater number of observations. Thus, this analysis tends to support approximating the motion of yImp\(\beta\) by a segmented chain model with a large number of segments, which, in the limit, approaches a continuously flexible chain or at least one in which the individual HEAT repeats or their constituent helices formed the segments. Overall, therefore, the TLS analysis was consistent with the yImp\(\beta\) molecule having a relatively continuously flexible conformation rather than being constructed from simply two or three rigid domains separated by very flexible hinges.

**Quantitative Geometric Analysis**

To perform a systematic analysis of HEAT repeats in different Imp\(\beta\) structures, we used an approach similar to that used recently for leucine-rich repeat (LRR) proteins (Bublitz et al., 2008). The orientation of each repeat was defined with respect to the preceding repeat using three angles: curvature, twist, and lateral bending. In the LRR study, each repeat was represented by a triangle based on three chosen C\(\alpha\) positions (Bublitz et al., 2008). To increase the general applicability to different repeats and the ability to automate the calculations, we instead defined three principal axes for each repeat based on the C\(\alpha\) positions of all residues in a repeat, and calculated the interrepeat angles based on the principal axes of two adjacent repeats (Figure 7).

Analyses indicate that there is a remarkable conservation of the interrepeat angles between human, mouse, and yeast Imp\(\beta\), which can therefore be compared directly. This conservation suggests that the particular interrepeat geometry is functionally critical and is consistent with an extensive interface holding adjacent repeats together. The results will be illustrated on six representative structures, which cluster qualitatively in four distinct patterns of interrepeat angles: (i) yImp\(\beta\) and Nup1-p:yImp\(\beta\); (ii) Imp\(\alpha\) and snurportin1 IBB:hImp\(\beta\); (iii) RanGT-P:yImp\(\beta\); and (iv) SREBP-2:mImp\(\beta\) (Figure 7C). For comparison, a similar analysis performed on structures simulated by MD calculations is shown in Figure S2. There is a pronounced irregularity of the distribution of interrepeat angles along the chain, as compared with LRR proteins (Bublitz et al., 2008). This may reflect a higher sequence variation between HEAT repeats, and the presence of a \(\beta\) sheet in the LRR structure that may limit the structural variation as compared to helix-helix contacts as progressively more TLS groups are added until 13–14 groups are employed (these correspond to residues 1–77; 78–166; 167–172; 173–274; 275–391; 392–438; 439–546; 547–592; 593–670; 671–690; 691–765; 766–831; and 832–861 in the case of 13 groups). Adding further groups results in a small increase in \(R_{\text{free}}\) indicating that the structure may be becoming overrefined.
fre quency of unliganded and RanGTP-bound yImpβ show substantial differences in curvature. A movement is accompanied by smaller relative changes in large structured proteins are usually assumed to involve movement of rigid domains around a small number of hinges (Qi et al., 2005), our results indicate that this is not the case in Impβ.

In summary, our results are consistent with Impβ being an inherently flexible molecule, which performs its functions by taking advantage of conformational transitions resulting from cumulative small structural changes along the length of the solenoid. Such continuous flexing can store internal energy much more efficiently than movements of rigid domains around flexible hinges. Our results may have general implications for solenoid proteins, which use their structural properties to carry out their cellular functions, their inherent flexibility allowing the specific binding of a single protein to a number of binding partners at different stages of a pathway. This flexibility combined with the storage of internal mechanical energy can generate a gradation of affinities to orchestrate the appropriate sequence of binding reactions in the cell.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

Full-length GST-yImpβ and GST-mImpβ were expressed in Escherichia coli strain BL21(DE3)pLysS and purified by glutathione affinity chromatography (Forwood et al., 2008; Roman et al., 2009). The GST fusion-tag was cleaved using thrombin and removed by size exclusion chromatography. Fractions containing the purified recombinant yImpβ were passed through glutathione-Sepharose affinity matrix to remove residual traces of GST and uncleaved GST-Impβ. The protein was concentrated to 40 mg/ml and stored in 20 mM Tris pH 7.4, 50 mM NaCl at −80 °C.

Crystal Structure Analysis

Crystalization and diffraction data collection are described by Roman et al. (2009). Raw data were autoindexed, integrated, and scaled using the HKL2000 package (HKL Research, Inc.). The crystals displayed P21 symmetry, with unit cell dimensions of a = 58.17 Å, b = 127.25 Å, c = 68.52 Å; β = 102.23°. The crystals are similar to the crystals of the yImpβ:Nup1 complex (Liu and Stewart, 2005), which displayed the same symmetry and unit cell dimensions of a = 58.08 Å, b = 125.97 Å, c = 69.35 Å; β = 110.2°. The asymmetric unit contains one yImpβ molecule. Initial phases were obtained by molecular replacement using Phaser (Storoni et al., 2004) and yImpβ from the yImpβ:Nup1p structure (Liu and Stewart, 2005) as a search model. Local rebuilding using COOT (Emsley and Cowtan, 2004) and refinement with REFMAC (Murshudov et al., 1997) from the CCP4 program suite (CCP4, 1994) followed by Phenix (Zwart et al., 2008) and the use of TLS displacement model (Painter and Merritt, 2008) yielded a final model with good overall stereochemistry (Table 1). All molecular structure figures were prepared using PyMol (DeLano Scientific LLC).

Small Angle X-ray Scattering

Small angle X-ray scattering (SAXS) data were collected using an Anton Paar/ Panalytical SAXSess system, with CuKα radiation from a sealed X-ray tube, line collimation and a CCD detector (Princeton Instruments). yImpβ and mImpβ were dialyzed overnight in 20 mM Tris 7.4, 125 mM NaCl, 10 mM DTT at 4 °C. SAXS measurements were collected at 20 °C at a concentration of 2–5 mg/ml, in a 1 mm silica capillary. To monitor for radiation damage, two to six successive 15 min exposures were compared, and no differences in scattering intensity were seen for either protein, suggesting that no structural changes had occurred. The scattering data were collected from q = 0.05 to 2.2 nm⁻¹ and were reduced to remove the contributions from the dark current of the detector, scattering of the dialysis buffer and empty capillary. The data were further normalized to absolute intensity using pure water as a calibration standard. The indirect transform program GNOM (Svergun, 1992) was used desmear the scattering profiles and evaluate the pair distance distribution function p(r) and the maximum size Dmax. The radius of gyration, Rg, and the forward scatter, I(0), was determined from the second and zeroth moments of p(r), respectively, as well as the maximum linear dimension of the particle, Dmax. Ten dummy atom models were generated by the program DAMMIN.
Figure 7. Variation of Interrepeat Angles between Successive HEAT Repeats
(A) Principal axes in successive repeats, shown in the structure of mImp (shown as ribbon diagram in gray) from the SREBP-2 complex as an example. Axis $P_1$ is the axis perpendicular to the helices in the plane of the repeat (blue); axis $P_2$ is the axis parallel to the helices in the plane of the repeat in individual repeats (cyan); and axis $P_3$ is the axis perpendicular to axes $P_1$ and $P_2$ (red).
**Quantitative Geometric Analysis**

The analysis of angles describing curvature, twist, and lateral bending between successive pairs of HEAT repeats was performed using an assignment of principal axes in individual repeats. The axes were obtained by a principal component analysis of the positions of Cx atoms in helical environments of a respective helix pair. Curvature is the angle between eigenvectors (ev) 2 of HEAT repeats N and N-1, projected onto a plane spanned by ev 3 and ev 2 of HEAT repeat N+1; twist is the angle made by ev 1 of HEAT repeats N and N+1, projected onto a plane spanned by ev 1 and ev 2 of repeat N+1; and lateral bending is the angle between ev 3 of HEAT repeats N and N+1, projected onto a plane spanned by ev 1 and ev 3 of helix pair N+1 (Figure 7). Our method is generally applicable to solenoid proteins, as the principal axes of repetitive motifs can be defined without the need for specifically conserved sequence positions, and also independent of secondary structure elements such as $\alpha$ helices. The program used for our geometric analysis (CuTLat: curvature, twist, lateral bending) is available for use on the web server [http://www.mpibpc.mpg.de/home/grubmueller/downloads/cutlat/index.html](http://www.mpibpc.mpg.de/home/grubmueller/downloads/cutlat/index.html).

**Molecular Dynamics Simulations**

The crystal structure of yImp was immersed in a dodecahedral box with box vectors of 14.9 nm, filled with TIP4P water molecules (Jorgensen et al., 1983), and Na$^+$/Cl$^-$ ions according to an ionic strength of 0.15 M. A production run of 23 ns length followed, in which the temperature and pressure were kept constant by weak coupling to a temperature bath of 310 K and a pressure of 1 bar, respectively. Long-range electrostatic interactions were calculated using the particle-mesh-Ewald method (Darden et al., 1993); short-range electrostatic interactions and van der Waals interactions were calculated explicitly up to a cutoff distance of 1 nm. The time-step used was 2 fs, and all bond lengths were constrained using the LINCS algorithm (Hess et al., 1997). The energy minimizations and force field (Jorgensen et al., 1996) in Gromacs version 3.3 (Lindahl et al., 2001) were then further refined using the generated damstart.pdb file as an input into DAMMIN (Konarev et al., 2006). Theoretical scattering curves of crystallographic atomic models were computed using CRYSOUL (Svergun et al., 1995). The program OLIGOMER (Konarev et al., 2003) was subsequently employed to calculate the volume fractions of a mixture of several distinct crystallographic conformations, to give the best fit to the experimental curve.

**Functional Analysis of yImp Mutants**

All chemicals were obtained from US Biological or Sigma unless noted otherwise. All media were prepared and all DNA manipulations were performed as described previously (Adams et al., 1997; Sambrook and Russel, 2001). All yeast strains and plasmids used in this study are described in Table S1.

RSL1 is the gene that encodes yImp (Kap95) in S. cerevisiae.

GFP-fusion proteins were localized in live S. cerevisiae cells using direct fluorescence microscopy on an Olympus BX60 epifluorescence microscope equipped with a Photometrics Quantix digital camera. To localize the yImp variants, wild-type cells (ACY192) expressing wild-type or variant yImp-GFP under its own promoter were grown overnight in selective media, diluted in fresh media, and grown for 3 hr to log phase prior to localization. To assess the functional impact of amino acid substitutions in yImp variants, yeast strains expressing wild-type RSL1 (ACY208) or each yImp variant as the only cellular copy of yImp. Transformants were plated on plates containing 5-fluoroorotic acid.
acids (S-FOA) (Boeke et al., 1987) to remove the maintenance plasmid and were transformed with a plasmid encoding either SV40TagNLS-GFP-GFP (Hodel et al., 2006) (pAC1065) or a GFP-GFP control (pAC1069) under the control of the MET25 promoter. Cells were grown overnight in selective media, washed once in dH2O, resuspended in media lacking methionine to induce expression of the reporter proteins, and incubated overnight prior to localization studies.

Immunoblot analyses were performed essentially as previously described (Towbin et al., 1979). Wild-type cells (ACY192) transformed with plasmids encoding the yImpβ-GFP variants were grown to log phase, collected by centrifugation, and washed once with dH2O and twice with cold PBSMT (PBS, 5 mM MgCl2, 0.5% Triton X-100). Glass bead lysis was conducted in PBSMT in the presence of the protease inhibitors phenylmethylsulfonyl fluoride (PMSF; 0.5 mM) and PLAC (pepstatin A, leupeptin, aprotinin, chymostatin; 3 μg/ml each). Lysates were cleared by centrifugation and total protein concentration was assessed by the Bradford assay. Thirty micrograms of protein was resolved by SDS-PAGE, transferred to a nitrocellulose membrane, and probed with an anti-GFP antibody (1:3,000, rabbit) to detect the yImpβ-GFP proteins and an anti-Yrb1 antibody (1:50,000, rabbit; Schlenstedt et al., 1995) as a loading control.

The in vivo function of each of the yImpβ mutants was assessed using a plasmid shuffle technique (Boeke et al., 1987). For the plasmid shuffle, vector alone or plasmids encoding the wild-type or variant yImpβ proteins were transformed into JRS11 cells containing a wild-type RSL1 URA3 maintenance plasmid (ACY208). Single transformants were streaked onto control ura-leu-2-glucose plates or on selective leu-glu plates containing 5-FOA. Plates were incubated at 25°C. A variant of yImpβ containing four leucine (L329, L330, L332, L333)-to-threonine changes was employed as a control to impact yImpβ protein function.

The cNLS import assay was performed essentially as previously described (Hodel et al., 2006; Shulga et al., 1996). Cells containing the yImpβ variants as the only copy of yImpβ and expressing the SV40TagNLS-GFP-GFP reporter protein were grown and induced. After collection by centrifugation, cells were pelleted, resuspended in glucose-free media containing sodium azide and 2-deoxy-D-glucose, and incubated at 30°C. Protein was resolved by SDS-PAGE, transferred to a nitrocellulose membrane, and probed with an anti-GFP antibody (1:3,000, rabbit) to detect the yImpβ-GFP variants and an anti-Yrb1 antibody (1:50,000, rabbit; Schlenstedt et al., 1995) as a loading control.

Immunoblot analyses were performed essentially as previously described (Hodel et al., 2006) (pAC1065) or a GFP-GFP control (pAC1069) under the control of the MET25 promoter. Cells were grown overnight in selective media, washed once in dH2O, resuspended in media lacking methionine to induce expression of the reporter proteins, and incubated overnight prior to localization studies.

REFERENCES


Structure

Structural Analysis of Importin-β Flexibility


