A Highly Strained Nuclear Conformation of the Exportin Cse1p Revealed by Molecular Dynamics Simulations

Ulrich Zachariae[1] and Helmut Grubmüller[1,*]
[1] Department of Theoretical and Computational
  Biophysics
Max Planck Institute for Biophysical Chemistry
Am Faßberg 11
37077 Göttingen
Germany

Summary

To investigate the stability of the open nuclear state of the exportin Cse1p and its closing mechanism at the atomic level, we have performed multiple molecular dynamics simulations. The simulations revealed a strikingly fast transition of Cse1p from the open conformation to the closed cytoplasmic form, consistent with the proposal that Cse1p represents a "spring-loaded molecule." The structure of the ring-shaped state obtained in the simulations is remarkably close to the crystal structure of the cytoplasmic state, though the open nuclear structure was used as the only input. The conformational change is initially driven by release of strain due to RanGTP/importin-α binding. Subsequently, a stable closed state is formed, driven by attraction of electrostatically complementary interfaces. These results are consistent with and extend previous proposals. Reverse-charge and neutral mutants remained in an open state. The simulations predict a detailed reaction pathway and resolve the role of suggested hinge regions.

Introduction

Nuclear transport receptors selectively mediate the active transport of large macromolecules through nuclear pore complexes (NPCs). The members of their largest class, karyopherins, are all related to importin-β, and they interact intimately with both their cargo substrates and the NPC, which allows passage between the cell nucleus and the cytoplasm (Fahrenkrog and Aebi, 2003; Weis, 2003; Isgro and Schulten, 2005). To date, the superfamily of known karyopherins comprises more than 20 members in humans and 14 in yeast (Harel and Forbes, 2004).

Among these, the importin-α,β/CAS system is the most widely studied nuclear import-export cycle (Conti et al., 2006) (Figure 1). Importin-β (Kap95p in yeast) carries a wide range of cargo proteins bearing a classical nuclear localization signal (NLS) into the nucleus. Usually, the adaptor protein importin-α (Kap60p in yeast) is interposed between carrier and cargo. RanGTP disassembles the import complex in the nucleus, thereby releasing the cargo. Importin-α is returned to the cytoplasm by a specialized transport factor, the exportin CAS (Cse1p in yeast; Kutay et al., 1997; Künzler and Hurt, 1998; Solsbacher et al., 1998). CAS/Cse1p binds

importin-α with high affinity only in a heterotrimeric complex requiring the presence of RanGTP. Both the heterodimers of CAS/importin-α and CAS/RanGTP have a dramatically decreased stability (Kutay et al., 1997). The trimeric export complex is dissociated by catalyzed GTP hydrolysis in the cytoplasm. The directionality of all nucleocytoplasmic transport processes is introduced by a steep RanGTP/GDP gradient across the nuclear envelope (Izaurralde et al., 1997; Kalab et al., 2002).

Karyopherins are constructed from an array of ~20 HEAT repeats (Figure 2). Each HEAT repeat comprises about 40 residues and consists of a hairpin of 2 antiparallel α helices, designated A and B (Chook and Blobel, 2001; Harel and Forbes, 2004). The HEAT repeats are stacked to form an overall superhelical structure. Interrepeat linker loops and sharp intrarepeat turns connect the helices. Some of the turns are extended by inserts of up to ~50 residues.

The ability of karyopherins to wrap around cargo proteins of different sizes and shapes and to release or bind cargo upon association with the transport effector RanGTP is essential for their transport function and has been associated with an unusual degree of conformational flexibility (Stewart, 2003; Conti et al., 2006). Indeed, a number of crystallographic studies have demonstrated differences in the geometries of importin-β/Kap95p and Cse1p in complex with various cargoes and with RanGTP (Cingolani et al., 1999, 2002; Lee et al., 2000, 2003, 2005). Extensive conformational variability has also been inferred from small-angle X-ray scattering (SAXS) and dynamic force spectroscopy (DFS) experiments (Fukuhara et al., 2004; Nevo et al., 2003, 2004). Moreover, the asymmetric unit of most karyopherin crystal structures comprises two slightly different protein conformations, which also points toward enhanced conformational flexibility.

Recently, the X-ray structures of the export complex of Cse1p bound to its cargo Kap60p and RanGTP and of the free cytoplasmic state of Cse1p have been determined (Matsuura and Stewart, 2004; Cook et al., 2005; Protein Data Bank (PDB) entries 1WA5 and 1Z3H, respectively). Whereas the cargo bound form shows the usual superhelix and curls around RanGTP and the C-terminal section of Kap60p, the cytoplasmic form exhibits a closed ring formed by its N-terminal 16 HEAT repeats (see Figure 2). Ring closure leads to occlusion of the RanGTP binding site, which comprises the inner surface of HEAT repeats 1–3 and HEAT repeats 13 and 14 together with a stretch of insert 19, and to the distortion of the Kap60p binding site (outer surface of HEAT repeats 5–7, HEAT repeats 2–4, and insert 19). This likely prevents the functionally adverse reassociation of Cse1p with importin-α/Kap60p in the cytoplasm. The asymmetric unit of the 77K crystal structure of free Cse1p comprises two independent copies with a C_α root-mean-square deviation (rmsd) of 1.3 Å with respect to each other. Several loop regions are less well defined in form 2; therefore, Cook et al. (2005) focused on form 1 in their discussion. By comparing the rmsd of HEAT repeat pairs in their nuclear and the cytoplasmic states,
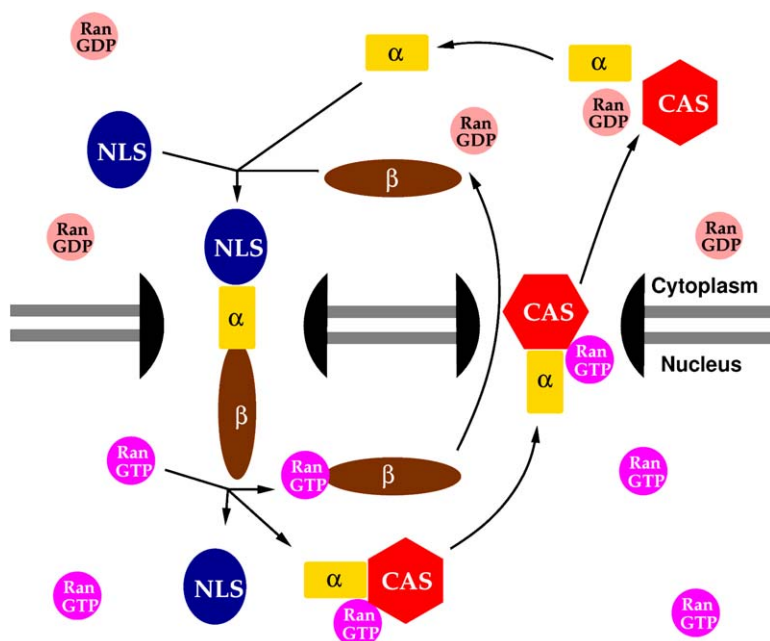
*Correspondence: hgrubmu@gwdg.de

Figure 1. Schematic Diagram of the Primary Processes Involved in the CAS/Importin-β/α Nucleocytoplasmic Transport Cycle

Importin-β (β), aided by the adaptor protein importin-α (α), transports NLS-containing cargo proteins into the nucleus. RanGTP binding disassembles the import complex in the nucleus. The export receptor CAS (Cse1p in yeast) is specialized to carry importin-α back to the cytoplasm in a complex with RanGTP; GTP hydrolysis to GDP then dissociates the export complex in the cytoplasm. All receptor-based nucleocytoplasmic transport processes are based on a steep RanGTP/GDP gradient across the nuclear envelope.

Cook et al. (2005) suggested that the regions between HEAT repeats 7–9 and, less importantly, between HEAT repeats 3 and 4 function as main hinges for the overall conformational change.

In a recent review, Conti et al. (2006) pointed out that energy stored by distorting the karyopherins may compensate for the substantial binding energies involved in complex formation and may thereby facilitate disassembly of the complexes by the comparably low energy associated with GTP hydrolysis. This effect may contribute to the high transport rates measured for nucleocytoplasmic transport. Accordingly, Matsuura and Stewart (2004) proposed on the basis of the X-ray structure of

the nuclear Kap60p:RanGTP:Cse1p complex and mutational studies that the nuclear state of Cse1p represents a strained conformation. Such a "spring-loaded" molecule should be capable of undergoing spontaneous changes.

To test this hypothesis, we have carried out unbiased multinanosecond nonequilibrium molecular dynamics (MD) simulations of Cse1p in aqueous solution. Covering the complete ring-closing reaction of Cse1p, our simulations provide a time- and residue-resolved atomistic and causal picture of this pronounced structural transition as well as information on the timescales on which the transition occurs. In particular, our simulations enabled
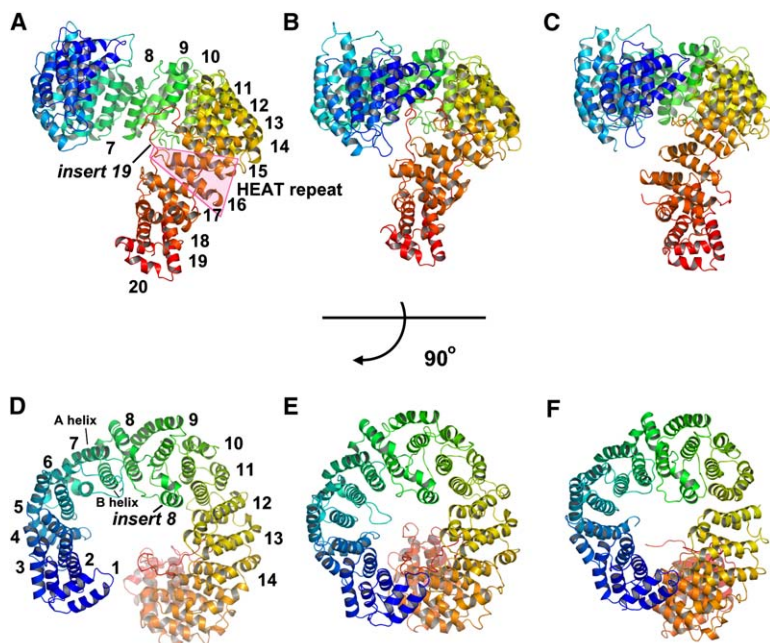

Figure 2. Transition of the Nuclear Structure of Cse1p into the Cytoplasmic Conformation

(A–C) Comparison of the (A) nuclear initial structure of Cse1p with its (B) structure obtained from a 10 ns molecular dynamics (MD) simulation and with the (C) crystal structure of the cytoplasmic state in form 2 (side views).

(D–F) (D), (E), and (F) are top views of (A), (B), and (C), respectively. HEAT repeat numbers and inserts 8 and 19 are indicated. Spontaneous ring formation is observed in MD simulations (from [A] to [B] or from [D] to [E]), during which HEAT repeats 1 and 2 are shifted by ~1.8 nm relative to HEAT repeats 14–16.

us to address the character of the most flexible regions as well as the underlying intramolecular forces that drive this functional conformational change.

Results and Discussion

All trajectories reported here were started from a "dissociated" X-ray structure of the nuclear state of Cse1p (entry 1WA5 in the PDB; Matsuura and Stewart, 2004). Accordingly, RanGTP and the Kap60p cargo were removed from the X-ray structure prior to the simulations. Four simulations in which this uncomplexed form of wild-type Cse1p spontaneously underwent significant conformational changes after removal of the position restraints on the protein heavy atoms were performed. In addition, control simulations were carried out on mutant Cse1p. Despite the fact that no biasing potentials were applied, the conformational transition led to a stable closed ring state in three of the wild-type simulations. By contrast, no conformational change was seen in control simulations of the complexed form of Cse1p, and mutant Cse1p remained open. The features shared by these three simulations are reported and analyzed here. Plots rest on snapshots taken from simulation 1.

Observation of Spontaneous Ring Closure

During the first ~2 ns of simulation 1, the radius of gyration dropped from 3.6 to 3.3 nm, indicating a substantial change in shape. To monitor this conformational motion, C_α rmsd fits to forms 1 and 2 of the crystal structure of cytoplasmic Cse1p (Cook et al., 2005) were calculated for the entire trajectory, and the N-terminal 16 HEAT repeats were chosen as reference because of their involvement in ring formation. It was found that form 2 of the crystallographic unit more adequately describes the conformational state of cytoplasmic Cse1p in solution at 310K (see below). Figure 2 reveals that the ring-like geometry of free form 2 is spontaneously adopted by Cse1p after removal of RanGTP and cargo within a few nanoseconds. The starting structure of Cse1p in its nuclear state is shown in Figure 2A (side view) and Figure 2D (top view); Figures 2B and 2E display the final structure obtained from the unbiased simulation. From a comparison with form 2 of PDB structure 1Z3H (Figures 2C and 2F), it can be seen that the final structure obtained from the MD simulation is in remarkably good agreement with the crystal structure of cytoplasmic Cse1p. This agreement is quantified by a C_α rmsd of only 0.3 nm, comparable to rmsd values obtained from equilibrium fluctuation of stable structures. After ring formation, HEAT repeats 1 and 2 are moved by ~2 nm relative to the interface at HEAT repeats 14–16. The C_α rmsd to the same reference section in crystallographic form 1 of the closed Cse1p state is slightly higher. During the same simulation, it decreases from 0.95 nm to 0.45 nm, and the other simulations also yield slightly larger values than the rmsd to form 2. Although form 2 is less well defined and has higher B factors in some regions, this result suggests that form 2 is in fact closer to the cytoplasmic conformation of Cse1p in aqueous solution at physiological temperature.

Figure 3A shows the rmsd (black line) of the Cse1p trajectory with respect to the cytoplasmic conformation, form 2. Starting from 0.9 nm, the low final rmsd of only 0.3 nm demonstrates that the structure has spontaneously evolved into a structure very close to the crystallographic form. In particular, for the equilibrium fluctuation of the complexed open form, an rmsd of 0.3 nm is obtained after 1–3 ns (dashed line), and in a simulation of the cytoplasmic form, the rmsd leveled off at ~0.3 nm after 0.5 ns; thus, no further drop of the rmsd in Figure 3A can be expected.

According to the rmsd, the conformational motion from the open to the closed form can be separated into three phases. The first phase is a very rapid drop, reducing the C_α rmsd of HEAT repeats 1–16 (residues 1–755) to the free form by 0.3 to ~0.6 nm within only 500 ps. During the second phase, the transition toward the cytosolic structure of Cse1p is markedly slowed down. After a subsequent short period of stagnation, relaxation continues. At about t = 2 ns, a nearly closed state is achieved, as defined by the formation of specific contacts between the N-terminal and the C-terminal junctures. During the third phase, further relaxation is seen on much slower timescales. An additional small step at t = 7.5 ns reduces the rmsd to ~0.3 nm. We note that the final closed structure was not used as an input for our simulations and that it is predicted in this sense.

In our additional independent simulations, the ring-closing reaction has been observed to occur at T = 310K and T = 300K within ~2 ns and ~10 ns, respectively (Figure 3D). Indeed, the C_α rmsd to the crystal structure was reduced from 0.9 to ~0.4–0.5 and ~0.5 nm, respectively, in these two simulations. Despite this structural heterogeneity, which is not unusual, in all cases the correct contact pattern (as compared to the crystal structure) and ring-shaped form were obtained. Only one simulation did not show the full closing transition of Cse1p within a 15 ns simulation, indicating that, due to the stochasticity of the conformational motions, the time required for completion of the closing process scatters and, in this case, exceeds the time window covered by the simulation. Also, this finding is not unexpected; for a single-exponential reaction, for example, individual transition times scatter by about one order of magnitude (Lange et al., 2006).

In addition, a principal components analysis (PCA) has been carried out on the full trajectory, and the projection of the motion onto the first eigenvector is plotted for comparison (Figure 3A, red line). The ring-closing transition can almost perfectly be described by motion on eigenvector 1, as illustrated by the exceptionally high correlation to the rmsd. The small differences near t = 0.5 ns are analyzed further below. The eigenvalue spectrum (inset) further stresses the dominance of the closing motion along the first eigenvector, which accounts for more than 50% of the total atomic motions of Cse1p described by the simulation.

Energetics and Driving Forces

Since the dynamics of the system are dominated by motion along the first eigenvector, we have analyzed the contributions of different regions of Cse1p to that motion, thus excluding all fluctuations that are less involved in the ring-closing process. A time-resolved plot of the structural deviations within 200 ps time windows along a trajectory projected upon the first eigenvector
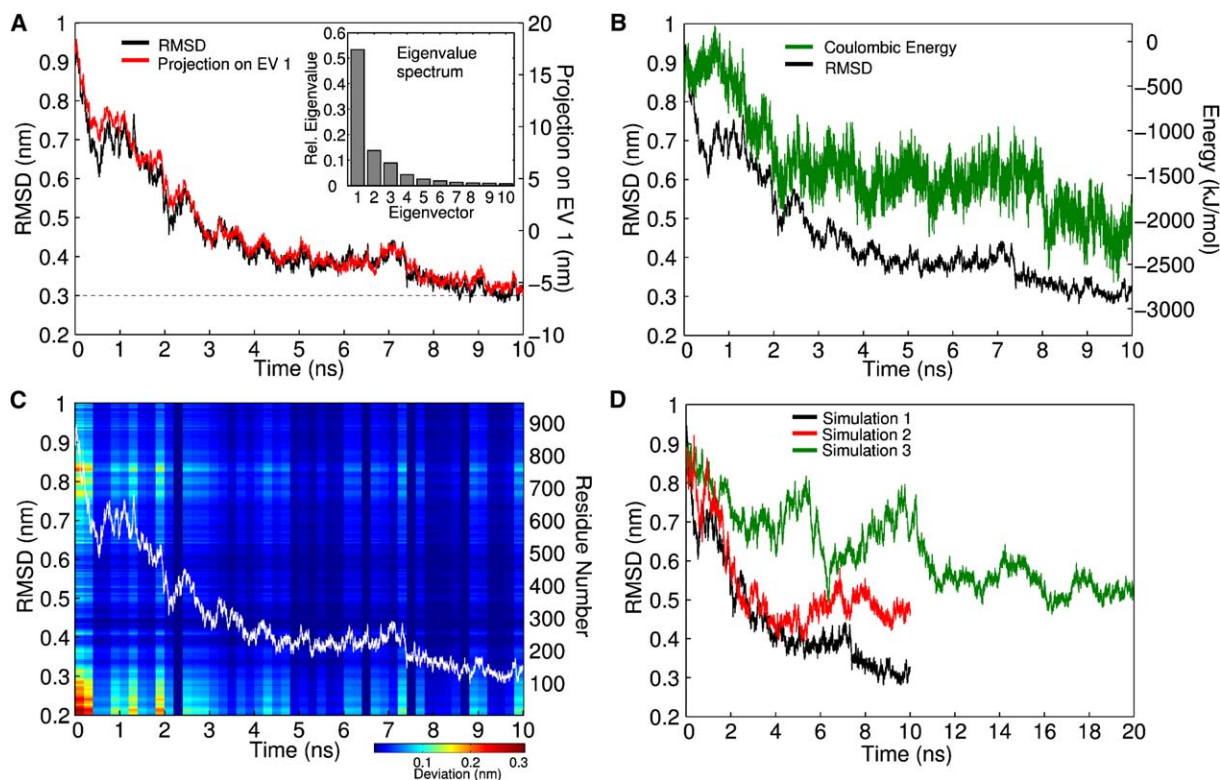
Figure 3. Analysis of the Trajectory Reflecting the Complete Structural Change

(A) C$_\alpha$ root-mean-square deviation (rmsd, black line) of HEAT repeats 1–16 of Cse1p with respect to form 2 of the crystal structure of the cytoplasmic state (1Z3H in the PDB) during closure, compared to projection of the trajectory upon the first eigenvector (red). Cse1p approaches and eventually reaches a conformation very similar to that of the cytoplasmic state from the nuclear initial conformation after removal of RanGTP and Kap60p. The dashed line indicates a typical rmsd value to the starting structure, seen in equilibration runs of proteins of comparable size (the cytoplasmic form of Cse1p reaches an rmsd of ~0.3 nm after 0.5 ns of equilibration).
(B) Coulombic energy between HEAT repeats 1–3 and residues 630–900 on the extended interface region (green line), and comparison with the drop in rmsd (black line).
(C) Time-resolved contribution of Cse1p residues to ring formation. The color range is from red (~0.3 nm deviation in 200 ps time windows) to blue (~0 nm). The overall rmsd to the crystal structure is overlayed in white. Areas of high variability include the N-terminal ~120 residues and helices 15A–17A. Frames showing a high structural change correlate well with significant drops in the rmsd curve.
(D) Comparison of the rmsd with respect to the cytoplasmic structure in the simulations that exhibited spontaneous ring formation.

(Figure 3B) shows that the regions forming the contact between HEAT repeats 1, 2, and 14–16 are already predominantly involved in the initial structural change. In agreement with the overall root-mean-square fluctuation (rmsf) distribution, it becomes evident that the N terminus especially as well as the region between HEAT repeats 14–17 change their positions most markedly and without any lagging period after the start of the simulation. The peaks in the latter region correspond to the A helices of HEAT repeats 15–17. Time windows showing a large difference between the initial and final positions of these areas correlate well with significant drops in the rmsd with respect to the cytoplasmic crystal structure.

Toward a closer structural analysis of the closing motion, we will consider the three main regimes of structural rearrangement and their respective driving forces. Figure 3C shows that ring formation is predominantly based on the strong electrostatic attraction of the interface regions, owing to the complementarity of their electrostatic surface potential. From t = 1 ns to the end of the simulation, the electrostatic energy between HEAT repeats 1–3 and residues 630–900 on the opposite surface shows a significant drop of ~2000 kJ/mol. Figure 4 dis-

plays the electrostatic surface potential of Cse1p, taking into account all charged residues. The strongly negative surface potential distributed over most parts of the molecule is due to a large number of charged groups, located not only at RanGTP and importin-α binding sites, but also in many other parts of the protein. This gives rise to a surplus of 27 negative charges.

However, the interface on HEAT repeats 1 and 2 is an exception in that basic residues form a continuous patch of positive charges here. This stretch of structurally adjacent bases on helices 1B and 2B (Lys21, Arg25, Arg28, Lys66, and Lys67), also reported to make contacts with an acidic cluster on HEAT repeat 14 (Glu652, Asp653, and Glu656) in the crystal structure of the cytoplasmic state (Cook et al., 2005), is strongly attracted by the negatively charged surface on the opposite interface. Together with the strong attraction between Asp71 and Glu72, both situated on a loop connecting helices 2B and 3A, and the conserved Lys733 on HEAT repeat 16, the electrostatic interaction pulls the N-terminal section of Cse1p toward the interface between HEAT repeats 14 and 16 (for mutant studies, see below). This leads to ring closure after ~2 ns and can be clearly
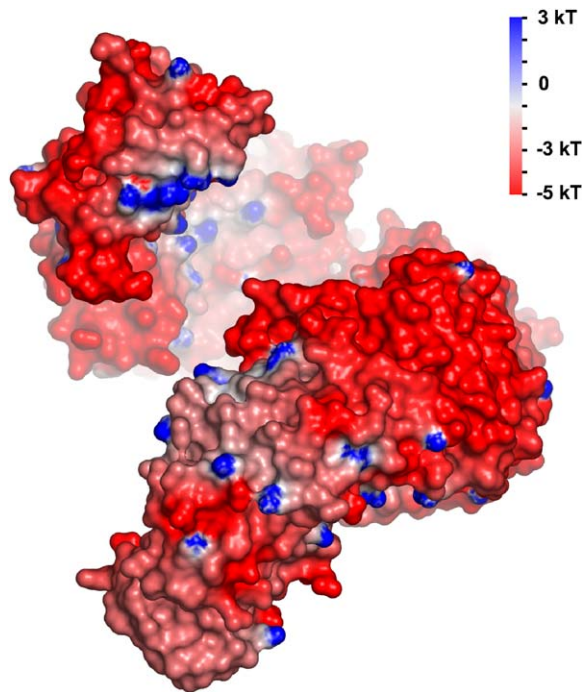
Figure 4. Electrostatic Potential Representation of the Surface of Cse1p in the Nuclear State

Colors range from red (−5 kT, negative surface potential) to blue (+3 kT, positive surface potential).

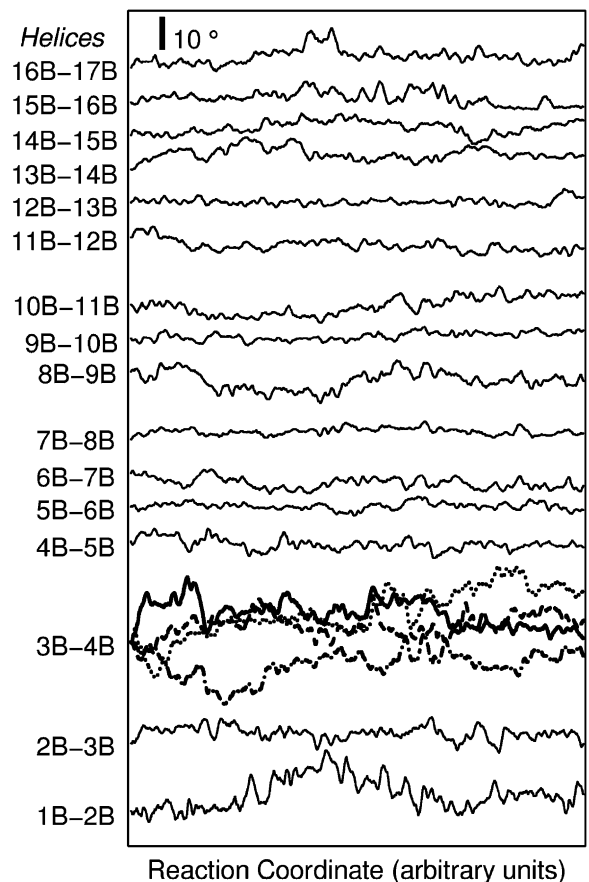Reaction Coordinate (arbitrary units)

Figure 5. Variations of the Angles between Adjacent HEAT Repeats during the Course of the Simulation

The largest variations are seen for the angle between HEAT repeats 3 and 4. A comparison of the evolution of this angle is plotted for all four simulations. Simulation 3 covered 20 ns; therefore, the time was scaled by a factor of two for this simulation (1, solid line; 2, dashed line; 3, dotted line; 4, dash-dotted line).

seen from the drop in Coulombic energy between the interfaces. In the complex, Asp653 and Glu656 form salt bridges to Lys37 and Lys152, respectively, of Ran.

The electrostatic nature of ring formation in the simulation agrees well with previous proposals by Cook et al. (2005). Whereas the interaction between Lys21 and Asp653, seen in the crystal structure, is reproduced in the simulation after closing, Glu72 makes contact with Arg728 rather than Lys733 in the crystal structure. The electrostatic surface (Figure 4), however, suggests that the complementarity between the N terminus and HEAT repeats 14–16 is not restricted to single salt bridges, but rather to groups of charged residues distributed over the entire surface, allowing substitutions of specific interactions with neighboring side chains while preserving the overall functional electrostatic pattern. This might also serve as an explanation for the paradox that although RanGTP uses the same residues to bind to different karyopherins, the binding sites on the karyopherins themselves have a low level of conserved residues; instead, a larger-scale electrostatic complementarity is conserved.

The Role of Proposed Hinge Regions

The initial rapid relaxation phase (t < 1 ns) is not based on electrostatic attraction between the interface regions. In fact, the Coulombic energy between residues 1–120 and 630–900 shows a slight increase during the first 0.7 ns (Figure 3C). This suggests that the first stage of approach toward the cytoplasmic structure is rather based on relaxation from mechanical stress, originally introduced by clamping RanGTP and importin-α in the nuclear complexed form. Initial relaxation primarily in-

volves HEAT repeats 1–3 (Figure 3B), which constitute the main binding site for Ran in the complex and a binding site for importin-α (Matsuura and Stewart, 2004), but it also involves HEAT repeats 14–17, which are close to the second RanGTP binding site.

Figure 5 shows that most interrepeat angles remain rather stable throughout all phases of the simulation; angles associated with the interfaces (between HEAT repeats 1 and 2, between HEAT repeats 13–17) are variable within limits of ~10°. However, the angle made by HEAT repeats 3 and 4 varies in an interval of ~20°. Whereas this angle increased shortly after the beginning of simulation 1, yielding the fastest and most complete relaxation (solid line), it increased more slowly in the two additional simulations that exhibited a slower ring formation reaction (broken and dotted lines), and decreased in the simulation that did not show ring closure. This tilting motion of HEAT repeats 1–3 suggests that a hinge is situated between HEAT repeats 3 and 4.

In this connection, a small deviation between the drop in rmsd and motion along eigenvector 1 can be seen in Figure 3A, especially around t = 0.5 ns, and this deviation stems from motion along eigenvector 3 (in contrast, eigenvector 2 covers conformational changes in HEAT
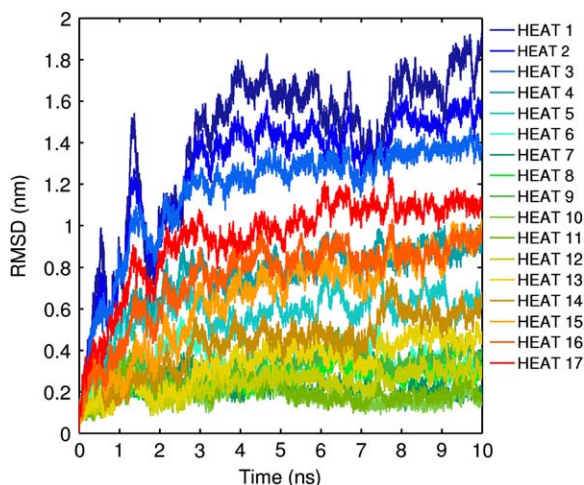
Figure 6. C_α Rmsd Plot of the First 17 HEAT Repeats with Respect to the Initial Nuclear Structure

HEAT repeats 1–3 exhibit the largest deviation and can be clearly separated from the rest of the protein. HEAT repeat 17 is also shifted to a higher extent than most of the other HEAT repeats. Notably, internal repeats 7–13 are almost invariant. Colors correspond to the coding used in Figure 2.
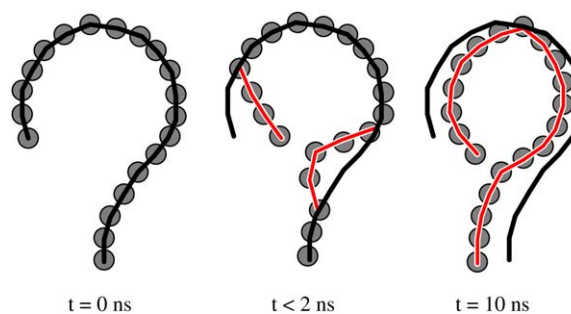

Figure 7. Schematic Diagram of Major Steps in Ring Formation

The electrostatic properties and the flexibility of the N-terminal three HEAT repeats as well as HEAT repeats 14–17 are the basis for ring closure. Subsequent domino-like relaxation of the HEAT repeats between both interfaces leads to the occurrence of an apparent hinge near HEAT repeat 8.

repeats 18–20 without any major contribution to ring formation). By DynDom analysis (Hayward and Berendsen, 1998) of the trajectory projected onto eigenvector 3, the quasirigid body motion of HEAT repeats 1–3 can be traced back to the loop connecting helices 3B and 4A. A conserved DFP motif is present here in CAS/Cse1p, and a proline residue is located at the same position in importin-β. In Crm1, it is replaced by EWP. This suggests a general functional relevance of this region in karyopherins. The importance of proline residues serving as hinges of conformational transitions is well known (Barlow and Thornton, 1988; Visiers et al., 2000). Figure 6 shows that the rmsd of HEAT repeats 1–3 with respect to the starting structure exhibits two cusps below t = 2 ns (Figure 6), and that it lies clearly above all other HEAT repeats in the simulation, which corroborates the importance of the interface between HEAT repeats 3 and 4.

Ring formation is followed by a slower relaxation motion toward the final closed structure that also involves the remaining regions of the protein. This relaxation constitutes an adaptation of the regions between the interfaces to the drastic structural changes at HEAT repeats 1–3 and 14–17. Although these rearrangements are seen as a combination of many small contributions from most parts of the molecule, we identified the flexibility of the A helices, which are often longer and more curved than the B helices, as an important factor. Apart from the interface-forming helices 15A–17A, which undergo some changes in their geometry already in the first and second, faster regimes, the extension of the curved helix 9A, which is situated in a more variable region (Cook et al., 2005), plays a major role in this relaxation process. Its flexibility is illustrated by a C_α rmsd of >0.15 nm (helix-forming atoms), whereas most other helices exhibit an rmsd of ∼0.05 nm. Similarly, helix 12A shows an increased rmsd of ∼0.15 nm, which can be attributed to the development of a kink in this helix at residue Leu549.

The combined structural rearrangements throughout the stacked helices lead to a final structure for which DynDom (Hayward and Berendsen, 1998) analysis with respect to the starting structure yields a hinge motion around an extended region centered at HEAT repeat 8. From our simulation, it can be seen that this "virtual" hinge motion was brought about by the strong attraction and movement of the interfaces. Adaptation of the chain of HEAT repeats between the interfaces leads to an apparent hinge half way between the interfaces (HEAT repeat 8), as schematically shown in Figure 7. The region around HEAT repeat 8, including insert 8, is structurally particularly well suited for accommodating structural changes originating from both directions, and this is due to the fact that the regular interrepeat arrangement of small right-handed twists is disrupted by a large left-handed twist between HEAT repeats 8 and 9 as well as to helix 9A being flexible (see above). According to our simulations, the separation of Cse1p into two moving arches is only a secondary effect, whereas the functional "hot spots" are situated in the rather peripheral sections. This observation has implications for the general understanding of other karyopherin structures, which are also considered to consist of two "independent" arches.

Influence of Insert 19

In our simulations, the high density of negatively charged residues on insert 19 contributes significantly to the attraction of the basic stretch on HEAT repeats 1 and 2. In fact, two additional salt bridges are formed between Lys66 and Arg67 from HEAT repeat 2 and Glu886 and Glu887 on the insert, respectively. These interactions are not present in the cytoplasmic crystal structure because insert 19 is disordered there; therefore, they are discussed here.

The slight structural heterogeneity in the ring geometry that has formed spontaneously in the three simulation runs is mainly due to different degrees of contact between HEAT repeats 1–3 and this loop. This observation might serve as an explanation for the fact that crystallization of Cse1p was not achieved unless loop 19 was cleaved at residue 875. Our results might thus point to a certain role of insert 19 in opening the ring in order to allow association with RanGTP and Kap60p. It should be noted that insert 19 is the only part of Cse1p that

interacts both with Ran and Kap60p in the nuclear state. The negatively charged groups on insert 19 interact with basic groups on HEAT repeats 15–17 in the nuclear complex (e.g., Lys695, Arg728) and thereby shield charges that make contacts with HEAT repeats 1–3 in the ring-shaped form, where insert 19 is displaced. Thus, this particular conformation of insert 19 should stabilize the open state in favor of the closed state.

The sequence of loop 19 is only moderately conserved in CAS/Cse1p, but it invariably exhibits a high density of negatively charged residues (the overall charge of the solvent-exposed part is −5 to −8). Loop 19 is the C-terminal part of Cse1p that is closest to the basic stretch on HEAT repeats 1 and 2, and its shape is complementary to the interface (see Figure 4). Therefore, the assumption that interaction occurs appears to be straightforward. One might argue, of course, that the formation of contacts between HEAT repeats 1 and 2 and insert 19 in the simulation is artificial and results from the usage of the nuclear conformation of this loop in the starting structure. However, because insert 19 is the only part of Cse1p that is involved in both RanGTP and importin-α binding, our results do suggest an essential role in switching between the closed and open states of Cse1p. Further experiments and simulation studies will be required to substantiate this idea.

Behavior of Mutants

Cook et al. (2005) found that both mutants K21E/R25E/R28E and D653R favor the open state to a higher extent than the wild-type protein. In order to test the hypothesis that ring closure is based on electrostatic forces, we performed a number of control simulations on Cse1p in silico mutants. Since D653R shows a moderate export defect (Cook et al., 2005), which might point to a more significant structural effect on Cse1p, we did not study this mutant. Instead, we focused on the effect of electrostatic changes on the basic stretch of the N-terminal interface.

Accordingly, the reverse-charge mutant K21E/R25E/R28E was simulated. In addition, we studied the conformational changes of four mutants in which all residues that formed specific salt bridges to the C-terminal interface were exchanged, i.e., K21, R25, R28, D71, and E72. Two of these mutants also carried reversed charges (K21E/R25E/R28E/D71K/E72K), whereas in the other two cases, the charged groups were substituted by neutral residues (polar, K21Q/R25Q/R28Q/D71N/E72Q, or alanine, K21A/R25A/R28A/D71A/E72A, respectively).

Figure 8A shows the rmsd of the reverse-charge mutants with respect to the cytoplasmic form (green line, K21E/R25E/R28E; cyan and red lines, K21E/R25E/R28E/D71K/E72K; black line, comparison with wild-type). Figure 8B displays the rmsd of the neutral mutants (red line, polar mutant; green line, alanine mutant). As can be seen, all of the mutants remain in a highly flexible open state. Figures 8C and 8D show the electrostatic interaction between the interface regions of the respective mutants and the wild-type. In contrast to the wild-type, the interfaces slightly repel each other. The apparent correlation between increases in rmsd, i.e., deviation from the closed form, and decreases in electrostatic repulsion points toward electrostatic driving forces. A similar correlation was seen for the wild-type.

Conclusions

By extended MD simulations, we have observed the complete conformational transition of the exportin Cse1p from the open state, present in the cell nucleus, to the ring-shaped closed state of the cytoplasm. Our simulations strongly suggest that the free open state of Cse1p is in fact a "spring-loaded" molecule. The almost immediate contraction toward the ring-shaped form in our simulations shows that the open form of Cse1p is highly unstable in the absence of RanGTP and Kap60p. Thus, the complex of Cse1p with Ran and Kap60p constitutes a clear example of an induced protein-protein fit. It puts into question the view that karyopherins constantly go through extensive defined structural changes in equilibrium and that, accordingly, binding to their partners merely leads to a dynamic population shift in which one of the conformations is preferred (Nevo et al., 2003).

The closure reaction can be separated into three regimes, each characterized by different driving forces. As suggested by Cook et al. (2005), the driving force for the formation of a stable ring is the strong electrostatic attraction between a stretch of positively charged residues on HEAT repeats 1 and 2 and the highly negative surface of HEAT repeats 14–16 (Figure 3C). In the complex, this attraction is outweighed by the binding of Ran and importin-α between these interfaces, and by the interaction of insert 19 with additional positive charges on HEAT repeats 15–17, which form salt bridges to HEAT repeats 1 and 2 in the closed state. The electrostatic nature of this conformational switch differs from the one inferred from a simulation study on a related multihelical repeat protein, ankyrin, in which packing effects were suggested to dominate (Sotomayor et al., 2005).

In contrast, the initial phase of especially fast contraction is found to be driven by the release of mechanical stress, induced by the association with RanGTP and importin-α, and involves the N-terminal section of Cse1p especially. In particular, our simulations revealed the importance of a hinge between HEAT repeats 3 and 4, and they highlight the pronounced dynamics of HEAT repeats 1–3 in this regime (see Figures 3B and 6). The electrostatic attraction and the initial mechanical relaxation involve similar regions because the ring-forming interfaces are identical to the main binding sites for RanGTP.

Further evidence for the existence of two different driving forces during the first two regimes comes from the simulation of the neutral mutants (Figures 8B and 8D). Here, the two phases can be clearly separated, because the strong electrostatic attraction in the wild-type is replaced by a weak repulsion. The fact that, nevertheless, a significant initial drop in rmsd can be observed strongly points to some relaxation from purely mechanical strain by RanGTP and importin-α binding. Accordingly, further relaxation and formation of the closed state are not seen in the mutants, because the second phase of complete ring closure in the wild-type must be attributed to the electrostatic complementarity of the interface regions.

The initial and main ring formation motion is followed by a slower relaxation regime. Here, a domino-like effect adjusts the HEAT repeats between the already shifted electrostatic switches at HEAT repeats 1–3 and 14–16 to the new ring-shaped geometry. This step involves bending and kinking of several A helices and a large number of minor rearrangements in inter-HEAT repeat
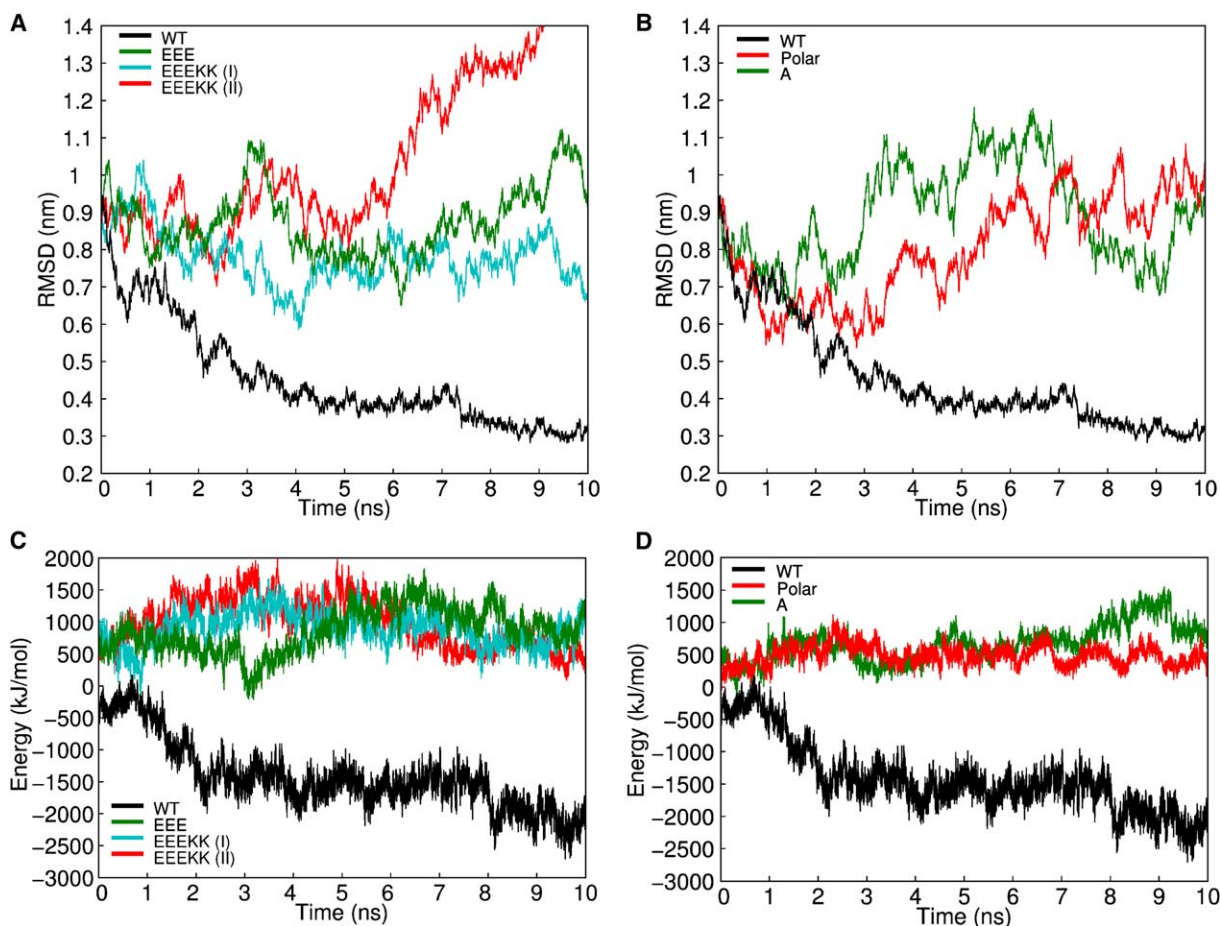
Figure 8. Simulations of Mutated Cse1p

(A) Behavior of reverse-charge mutants. The rmsd to the cytoplasmic structure is plotted for the mutants K21E/R25E/R28E (EEE, green line) and K21E/R25E/R28E/D71K/E72K (two simulations, EEEKK, cyan and red lines) and is compared to the behavior of the wild-type (black line).
(B) Simulations of neutral mutants. The rmsd to the cytoplasmic conformation is shown for the polar neutral mutant K21Q/R25Q/R28Q/D71N/E72Q (red line) and the alanine mutant K21A/R25A/R28A/D71A/E72A (green line) and is compared with the wild-type (black line).
(C and D) (C) and (D) display the Coulombic energy between the N terminus and the C-terminal interface region for the reverse-charge and the neutral mutants, respectively.

angles. Relaxations from both interfaces proceed toward the center dissecting this extended region, which remains largely in place. The structure around HEAT repeat 8 is particularly well suited to accommodate divergent structural changes; thus, an ostensive hinge appears near HEAT repeat 8. Our results, however, suggest that the structural change is driven by the movement of HEAT repeats 1–3 and 14–17.

The final structures from our wild-type simulations and the crystal structure (form 2) of the cytoplasmic state are very similar. In our simulation, the initially nucleoplasmic structure approaches the crystal structure of the cytoplasmic state to ∼0.3 nm in rmsd, as measured over the ring-forming section, within a time of ∼9 ns at T = 310K. This rmsd is equal to values that are typically obtained from routine equilibrations of large proteins. In fact, simulations of the complexed open form yielded an rmsd of 0.3 nm with respect to the starting structure after 1–3 ns, a value that we also obtained from a simulation of the cytoplasmic structure of Cse1p after 0.5 ns. An rmsd on this order, therefore, reflects equilibrium fluctuations rather than conformational motions; thus, further reduction of the rmsd is not expected. We note that the only part of the molecule that differs more significantly from the crystal structure is the far C-terminal section. We attribute this to the combined influence of crystal contacts and the specific crystallization process. The structure of the cytoplasmic form of Cse1p was obtained after cleaving the peptide bond between residues Leu875 and Lys876, which leads to a separation of the far C-terminal part of Cse1p from the rest of the protein. Uncleaved Cse1p did not crystallize (Cook et al., 2005). The ring shape of the closed conformation was correctly predicted by the MD simulation, and even specific and crucial side chain contacts reported in the crystal structure of the ring-like state were seen and, thus, correctly predicted.

Experimental Procedures

All simulations were carried out by using the MD software packages GROMACS 3.2.1 and GROMACS 3.3 (Lindahl et al., 2001). The OPLS-all atom force field (Jorgensen et al., 1996) was used for the protein, and TIP4P was used as a water model (Jorgensen et al.,

1983). All simulations were performed in the NPT ensemble. The temperature was kept constant at T = 300K or T = 310K by coupling to an isotropic Berendsen thermostat (Berendsen et al., 1984) with a coupling time of τ_t = 0.1 ps. The pressure was coupled to a Berendsen barostat (Berendsen et al., 1984) with τ_p = 1.0 ps and an isotropic compressibility of 4.5×10^{-5} bar^{-1} in the x, y, and z directions. All bonds were constrained by using the LINCS algorithm (Hess et al., 1997). An integration time step of 2 fs was used. Lennard-Jones interactions were calculated with a cut-off of 10 Å. Electrostatic interactions were calculated explicitly at a distance smaller than 10 Å; long-range electrostatic interactions were calculated by particle-mesh Ewald summation (Darden et al., 1993) with a grid spacing of 0.19 nm and sixth order B-spline interpolation.

The simulation system was set up as follows. The crystal structure of the nuclear and the cargo/RanGTP bound states of Cse1p was used (Matsuura and Stewart, 2004; PDB entry 1WA5), and cargo and RanGTP atoms were removed. Short missing loop sections were modeled by using Whatif (Vriend, 1990). Only 5 residues of insert 19 were disordered in the 2.1 Å resolution crystal structure of the nuclear state of Cse1p (residues 881–885), which was used as our starting geometry, whereas 20 or 19 residues are missing from this ∼50 residue stretch in forms 1 and 2 of the cytoplasmic crystal structure, respectively. The 5 residues missing from Cse1p in PDB entry 1WA5 were inserted into the loop prior to the simulation. Notably, both additional salt bridges to HEAT repeats 1 and 2 seen in the simulation are formed with a section of insert 19 that was originally resolved in the nuclear crystal structure. Protonation states of titratable groups of Cse1p were determined by using Whatif (Vriend, 1990) and its interface to DelPhi (Nicholls and Honig, 1991). In silico mutants were made and optimized by using Whatif (Vriend, 1990). The proteins were solvated in a rhombic dodecahedral box with box vectors with lengths of 137.6 Å . The system comprised ∼238,000 atoms. A total of 164 Na$^+$ and 137 Cl$^-$ ions were added to neutralize the system and to obtain a physiological ionic strength of ∼150 mM. Energy minimization by 200 steps by using the steepest descent algorithm was followed by a 400 ps MD simulation at the target temperature using harmonic position restraints on the heavy atoms of the protein with a force constant of k = 1000 kJ mol^{-1} nm^{-2} to equilibrate water and ions. Subsequently, trajectories of 7–20 ns length were produced by free (unbiased) MD simulations. No explicit equilibration phase for the entire system was required for these intrinsically nonequilibrium simulations.

Structures were written out every 1 ps for subsequent analysis. The trajectory was analyzed in terms of rmsd with respect to form 2 of the crystal structure of Cse1p (Cook et al., 2005; 1Z3H in the PDB) by fitting the C$_\alpha$ atoms of HEAT repeats 1–16 onto the crystal structure and calculating the rmsd for the same range of atoms. A principal components analysis (PCA) was performed on the full trajectory of the simulation, taking into account C$_\alpha$ atoms. The sequence-resolved rmsd plot was obtained by comparing C$_\alpha$ positions from snapshots taken every 200 ps on the trajectory filtered for motion on eigenvector 1 after fitting the structures on each other. For analysis of the rigid-body domain motion of Cse1p, the program DynDom (Hayward and Berendsen, 1998) was used. Electrostatic surface representations were obtained with DelPhi (Nicholls and Honig, 1991) together with Pymol (DeLano, 2002), which was also used for all other molecular representations. Graphs and color surface plots were created by using Matlab 7.0.1 (The Mathworks, Inc.). Interhelical angles were determined between the principal axes of the helices, derived from a PCA of the helix C$_\alpha$ positions. The B helices were taken as reference, since they are usually straighter than the A helices, resulting in a more accurate definition of the helix axes.

References

Barlow, D.J., and Thornton, J.M. (1988). Helix geometry in proteins. J. Mol. Biol. *201*, 601–619.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Di Nola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684–3690.

Chook, Y.M., and Blobel, G. (2001). Karyopherins and nuclear import. Curr. Opin. Struct. Biol. *11*, 703–715.

Cingolani, G., Petosa, C., Weis, K., and Müller, C.W. (1999). Structure of importin-β bound to the IBB domain of importin-α. Nature *399*, 221–229.

Cingolani, G., Bednenko, J., Gillespie, M.T., and Gerace, L. (2002). Molecular basis for the recognition of a nonclassical nuclear localization signal by importin β. Mol. Cell *10*, 1345–1353.

Conti, E., Müller, C.W., and Stewart, M. (2006). Karyopherin flexibility in nucleocytoplasmic transport. Curr. Opin. Struct. Biol. *16*, 237–244.

Cook, A., Fernandez, E., Lindner, D., Ebert, J., Schlenstedt, G., and Conti, E. (2005). The structure of the nuclear export receptor Cse1 in its cytosolic state reveals a closed conformation incompatible with cargo binding. Mol. Cell *18*, 355–367.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald – an *Nlog(N)* method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (http://www.pymol.org).

Fahrenkrog, B., and Aebi, U. (2003). The nuclear pore complex: nucleocytoplasmic transport and beyond. Nat. Rev. Mol. Cell Biol. *4*, 757–766.

Fukuhara, N., Fernandez, E., Ebert, J., Conti, E., and Svergun, D. (2004). Conformational variability of nucleo-cytoplasmic transport factors. J. Biol. Chem. *279*, 2176–2181.

Harel, A., and Forbes, D.J. (2004). Importin β: Conducting a much larger cellular symphony. Mol. Cell *16*, 319–330.

Hayward, S., and Berendsen, H.J. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. Proteins *30*, 144–154.

Hess, B., Bekker, H., Berendsen, H.J.C., and Fraaije, J.G.E.M. (1997). LINCS: a linear constraint solver for molecular simulations. J. Comput. Chem. *18*, 1463–1472.

Isgro, T.A., and Schulten, K. (2005). Binding dynamics of isolated nucleoporin repeat regions to importin-β. Structure *13*, 1869–1879.

Izaurralde, E., Kutay, U., von Kobbe, C., Mattaj, I., and Görlich, D. (1997). The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. EMBO J. *16*, 6535–6547.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935.

Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. (1996). Development and testing of the OPLS-AA force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. *118*, 11225–11236.

Kalab, P., Weis, K., and Heald, R. (2002). Visualization of a Ran-GTP gradient in interphase and mitotic *Xenopus* egg extracts. Science *295*, 2452–2456.

Künzler, M., and Hurt, E.C. (1998). Cse1p functions as the nuclear export receptor for importin α in yeast. FEBS Lett. *433*, 185–190.

Kutay, U., Bischoff, F.R., Kostka, S., Kraft, R., and Görlich, D. (1997). Export of importin α from the nucleus is mediated by a specific nuclear transport factor. Cell *90*, 1061–1071.

Lange, O.F., Schäfer, L.V., and Grubmüller, H. (2006). Flooding in GROMACS: accelerated barrier crossings in molecular dynamics. J. Comput. Chem., in press. Published online August 9, 2006. 10.1002/jcc.20473.

Lee, S.J., Imamoto, N., Sakai, H., Nakagawa, A., Kose, S., Koike, M., Yamamoto, M., Kumasaka, T., Yoneda, Y., and Tsukihara, T. (2000). The adoption of a twisted structure of importin-β is essential for the

protein-protein interaction required for nuclear transport. J. Mol. Biol. *302*, 251–264.

Lee, S.J., Sekimoto, T., Yamashita, E., Nagoshi, E., Nakagawa, A., Imamoto, N., Yoshimura, M., Sakai, H., Chong, K.T., Tsukihara, T., and Yoneda, Y. (2003). The structure of importin-β bound to SREBP-2: nuclear import of a transcription factor. Science *302*, 1571–1575.

Lee, S.J., Matsuura, Y., Liu, S.M., and Stewart, M. (2005). Structural basis for nuclear import complex dissociation by RanGTP. Nature *435*, 693–696.

Lindahl, E., Hess, B., and van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. J. Mol. Model. (Online) *7*, 306–317.

Matsuura, Y., and Stewart, M. (2004). Structural basis for the assembly of a nuclear export complex. Nature *432*, 872–877.

Nevo, R., Stroh, C., Kienberger, F., Kaftan, D., Brumfeld, V., Elbaum, M., Reich, Z., and Hinterdorfer, P. (2003). A molecular switch between alternative conformational states in the complex of Ran and importin β1. Nat. Struct. Biol. *10*, 553–557.

Nevo, R., Brumfeld, V., Kapon, R., Hinterdorfer, P., and Reich, Z. (2004). Direct discrimination between models of protein activation by single-molecule force measurements. Biophys. J. *87*, 2630–2634.

Nicholls, A., and Honig, B. (1991). A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. J. Comput. Chem. *12*, 435–445.

Solsbacher, J., Maurer, P., Bischoff, F.R., and Schlenstedt, G. (1998). Cse1p is involved in export of yeast importin α from the nucleus. Mol. Cell. Biol. *18*, 6805–6815.

Sotomayor, M., Corey, D.P., and Schulten, K. (2005). In search of the hair-cell gating spring elastic properties of ankyrin and cadherin repeats. Structure *13*, 669–682.

Stewart, M. (2003). Nuclear trafficking. Science *302*, 1513–1514.

Visiers, I., Braunheim, B.B., and Weinstein, H. (2000). Prokink: a protocol for numerical evaluation of helix distortions by proline. Protein Eng. *9*, 603–606.

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

Weis, K. (2003). Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. Cell *112*, 441–451.