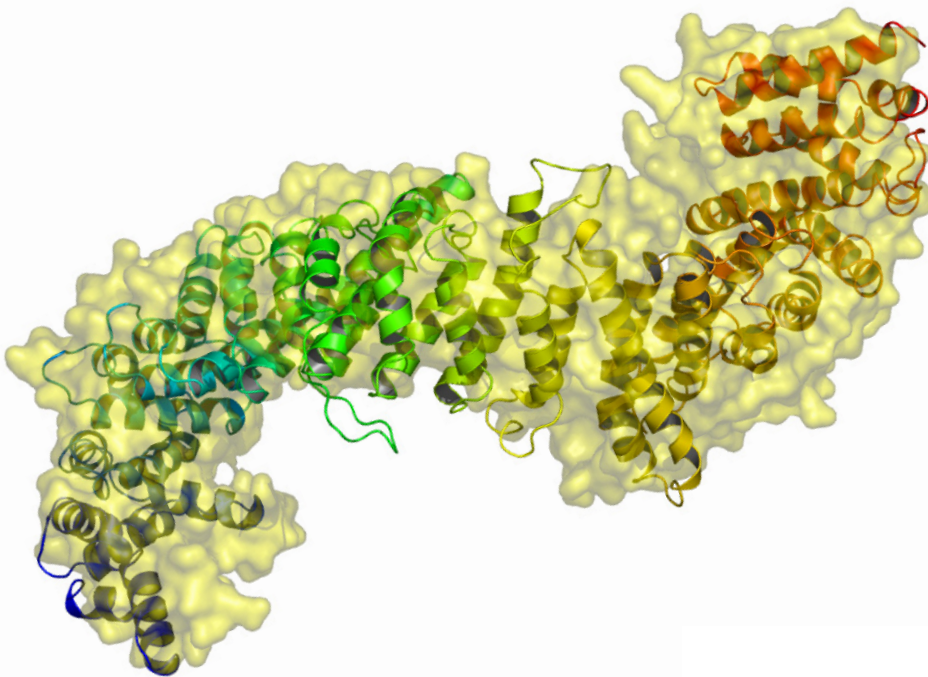


Berichte aus den Abteilungen

Prediction of a RanGTP-induced conformational switch in the exportin CAS/Cse1p by molecular dynamics simulations

Ulrich Zachariae, Helmut Grubmüller

Abt. Theoretische und computergestützte Biophysik



Introduction

The structure of water-soluble proteins forms around a cluster of hydrophobic residues which avoid contact with the solvent, thus exposing most charged and polar amino acids to the aqueous interface. This effect is known as 'hydrophobic collapse', and it is for this reason that the shape of many soluble proteins can very roughly be described as globular.

A class of soluble proteins that has recently caught increased attention among structural biologists and biophysicists are α -helical repeat proteins, which fall somewhat outside of this rule. These proteins

are formed from repeating arrangements of helix-loop-helix motifs, stacked to form large curved or superhelical structures (Fig. 1 A,B). Depending on their sequence and internal structure, the building blocks are called leucine-rich repeats (LRR), armadillo repeats, or HEAT-repeats; these building blocks form proteins such as nuclear transport receptors (karyopherins), clathrin, catenin, internalin, and ankyrin. As can be seen in Fig. 1A, their hydrophobic core is far from being globular, and due to the large hydrophilic surface area, these proteins are well suited for processes

involving the formation of multiple protein-protein contacts (Fig. 1C). Associated with the modular construction scheme, a high level of flexibility is ascribed to all of these proteins.

Nucleocytoplasmic transport in cells

An important cellular process, requiring the rapid formation of especially diverse protein-protein contacts, is the transport of proteins and RNA to and from the cell nucleus through the nuclear pore complex (NPC)¹. It is carried out by nuclear transport receptors. The members of their largest class, which are related to importin- β , are composed of ~20 HEAT repeats (their name is derived from the proteins in which they were first identified: huntingtin, elongation factor 3, 'A' subunit of protein phosphatase A, and TOR1 lipid kinase), which consist of approximately 40 amino acids that form helix-loop-helix structures (Fig. 1B).

Inhalt

Abt. Theoretische und computergestützte Biophysik	1-6
Aktuelle Pressemeldungen	6-10
EU-Referat	8-9
Umzug Abt. Hell	10
„Deutschland - Land der Ideen“	11
Neueinstellungen	12
Abgänge; Gäste	13
Publikationen	13-15
GWDG Info	14-15
Promotionen und Auszeichnungen	15
BR-Wahl 2006	16
Radfahrer-Infos	16
PhDNet	17
Doktorandenseminar	18-20
KFB Lecture	20
Impressum	20



Depending on transport direction, karyopherins are classified into importins and exportins; some receptors mediate transfer in both ways. The prototypical and most widely studied import-export cycle is the importin- β /importin- α /CAS system (Fig. 2), for which X-ray structures of the major players became recently available. Importin- β carries proteins across the nuclear envelope which are destined to the nucleus by possessing a classical nuclear localisation sequence (NLS). Frequently, the adaptor protein importin- α is interposed between importin- β and the cargo protein. The protein-protein contacts must be readily switchable, because cargo is released in the nucleus when the small G-protein Ran in its GTP-bound form (RanGTP) binds to importin- β . The importin- β :RanGTP complex travels back to the cytoplasm, while the adaptor, importin- α , is recycled back to the cytoplasm by the specialized export receptor CAS, also in a complex with RanGTP. Hydrolysis of GTP to GDP, catalyzed by a cytoplasmic enzyme, disassembles both complexes and renders both importin- α and β ready for a new round of protein import. Disassembly of the export complex is accompanied by a drastic conformational change in CAS, which ensures that the surface of CAS is not open for reassociation with importin- α or Ran. All nucleocytoplasmic transport processes are driven by a steep RanGTP-RanGDP gradient across the nuclear envelope.

Karyopherins are inherently flexible structures

The crystal structures of karyopherins obtained thus far have shown that a high degree of intrinsic conformational flexibility seems to be characteristic for these structures. Chook and Blobel asked in a review article: "Are there discrete regions of flexibility within the proteins or are karyopherins capable of forming a continuum of conformations?"². Another important question is whether the varying conformations observed in complexes with different binding partners are inherent, i.e. pre-existing free energy minima given only by the structure, one of which is then favoured by a population shift upon binding, or, alternatively, if they are created by an induced protein-protein fit.

In order to shed some light on these questions and also to get information on

the involved time-scales, we have performed a number of multi-nanosecond molecular dynamics (MD) simulations on the exportin CAS, represented by its yeast equivalent, Cse1p³. In MD simulations, Newton's equations of motion are solved to produce a trajectory of the dynamics of a molecule, e.g., a protein, over a period of time, starting from an X-ray or NMR-derived structure. Cse1p exists in an open, nuclear structure (bound to RanGTP and importin- α) and a ring-shaped, closed cytoplasmic state.

Real time observation of the conformational transition in Cse1p

We used the crystal structure of the nuclear export complex⁴ (Fig. 1C), but removed both RanGTP and importin- α to examine the stability of the open state of Cse1p when it is not complexed with its binding partners. Starting from the nuclear X-ray structure, all simulations showed a pronounced decrease in the size of the protein. The ring-shaped, closed cytoplasmic form of CAS is more compact, therefore we compared the structures obtained from the MD trajectory with the crystal structure of

the closed protein⁵. Figure 3 shows the root mean square deviation (RMSD) of the C α positions in the first 16 HEAT repeats with respect to the cytoplasmic ring, which is formed between HEAT repeats 1-16 (white line). It can be seen that the nuclear form spontaneously evolved into a structure very close to the cytoplasmic crystal structure. In the simulation, the RMSD, initially at 0.9 nm, drops to ~0.3 nm over a time range of only 10 nanoseconds. In fact, this low value would typically have been expected for the deviation if we had started our simulation from the closed state already, because an RMSD of that size represents equilibrium fluctuations of large proteins in solution at a temperature of ~300K. In that sense, the simulations have predicted the closed structure without prior knowledge of this structure.

The conformational transition is visualized in Fig. 4, which shows the initial, open starting structure in panels A and D, as well as the final closed structure from our simulation in panels B and E, respectively. For comparison, the crystal structure of the closed state is shown in panels C and F. Within the simulated period of time, the

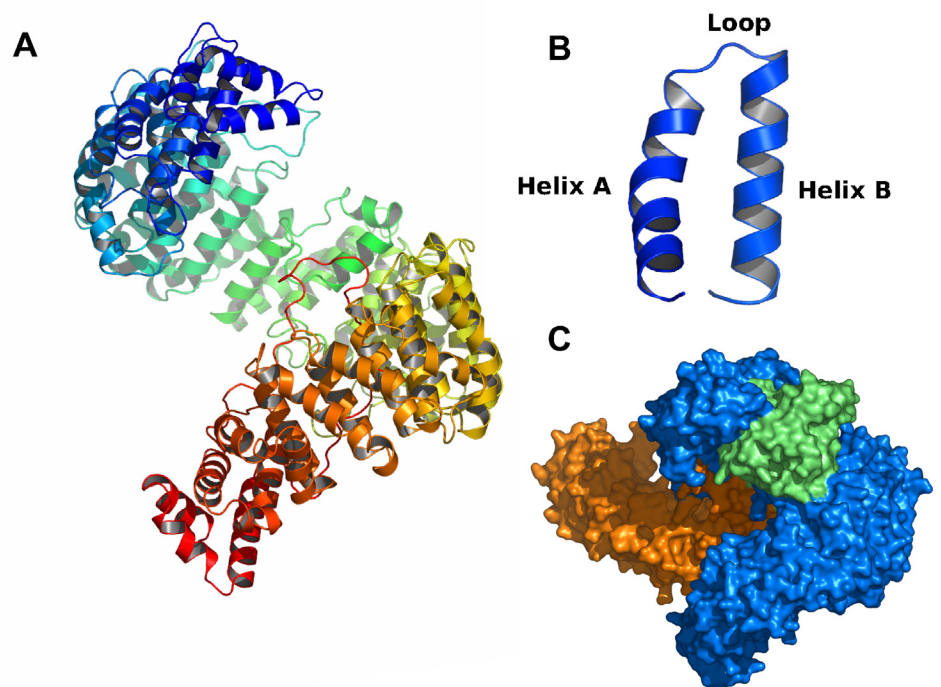


Fig. 1. (A) Superhelical structure of the exportin CAS/Cse1p. It consists of 20 HEAT repeats, a helix-loop-helix building block (B). (C) Cse1p (blue) wraps around RanGTP (light green) and importin- α (orange) for nuclear export.

Abb. 1: (A) Struktur des Exportins CAS/Cse1p in Form einer Superhelix. Cse1p besteht aus 20 HEAT-repeats, Bausteinen aus einem Helix-Loop-Helix-Motiv (B). (C) Cse1p (blau) umschließt RanGTP (hellgrün) und Importin- α (orange) während des Exports aus dem Zellkern.

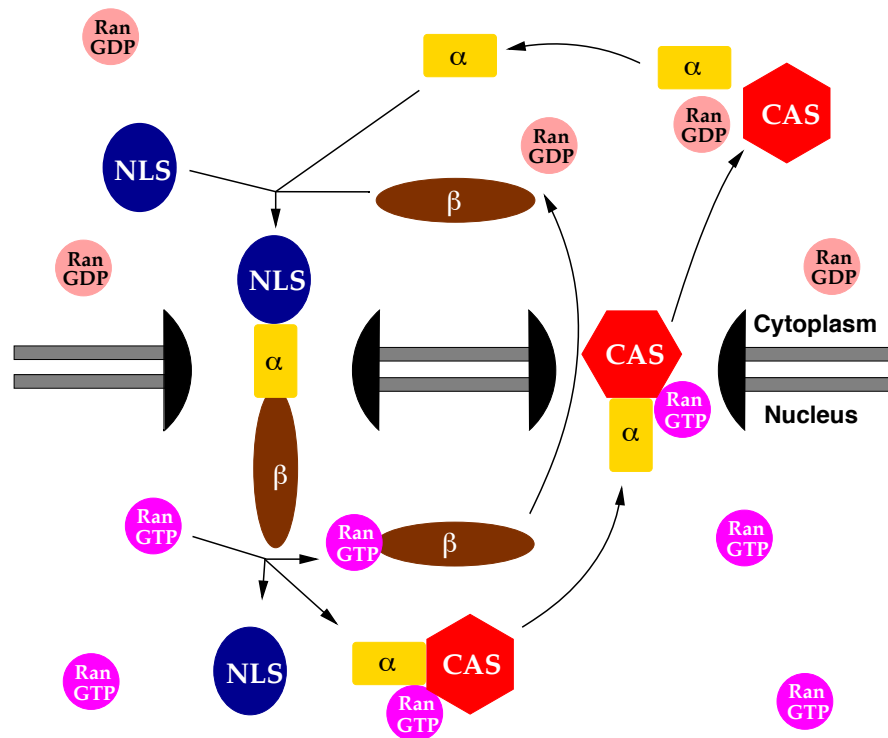


Fig. 2. Schematic diagram of nucleocytoplasmic transport processes involved in the nuclear import of cargo proteins which carry a classical nuclear localisation sequence (NLS).

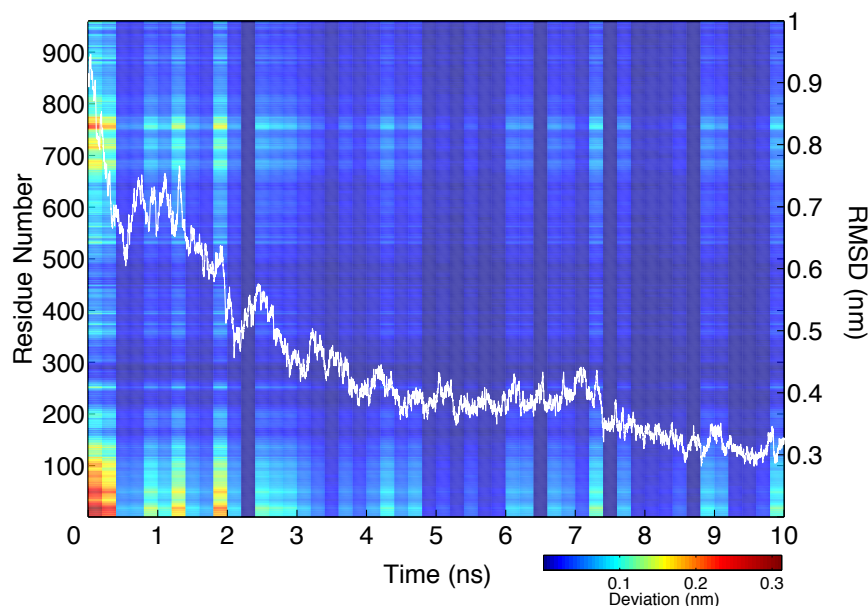
Abb. 2: Schema der Transportprozesse, die am Import von denjenigen Proteinen in den Zellkern beteiligt sind, die eine klassische Kern-Lokalisations-Sequenz (NLS) besitzen.

N-terminal HEAT repeat 1 has moved by ~2 nm relative to its interface at the opposing arch (HEAT repeats 14-16). Despite the small difference, we have thus seen the major part of the conformational change from the nuclear to the cytoplasmic state in the simulation.

The time required for the structural rearrangement is remarkably short. Furthermore, the conformational transition starts immediately after release of the atomic positions, taken from the 77K crystal structure in aqueous solution at 310K. Moreover, it was observed to take

place in three independent simulations at comparable time scales. This means that the open state of Cse1p is inherently unstable, which renders the population shift hypothesis rather unlikely. Instead, the open state seems to be stable only in complex with RanGTP and importin- α . Our simulations therefore suggest an induced fit mechanism for RanGTP and importin- α binding, without any significant intermediate state involved.

The colour plot underlying the RMSD curve in Fig. 3 depicts the structural deviation of Cse1p within time windows of



200 picoseconds on a trajectory filtered for the main mode of motion leading to ring closure. It becomes clear from this plot that the N-terminal three HEAT repeats (up to residue ~120), and, to a lesser degree, the interface-forming region around residues ~670-800 play an outstanding role in the structural change.

What is the driving force for the closure?

Complementing the static picture given by the crystal structures, our simulations provide a time-resolved and causal picture of the transition and the underlying driving forces. To reveal the origin of these forces, Fig. 5 shows the electrostatic potential resulting from the charged residues on the surface of Cse1p; red colour denotes negative, blue positive potential. The entire molecule has a large negative charge of -27 e, such that most parts of the surface have a strongly negative potential. A functionally important exception can be seen near the N-terminus, at HEAT repeats 1 and 2. Here, several positively charged amino acid residues lead to a stretch exhibiting a positive potential, depicted by its blue colour. The very strong correlation between the strength of the coulombic interaction between charges on the first three HEAT repeats and the opposing surface of the protein (Fig. 6, green line) and the progress on the reaction coordinate (RMSD, black line) shows that it is mainly the electrostatic attraction between the complementary charges on both interfaces which drives the conformational transition. This attrac-

Fig. 3. Spontaneous approach of the nuclear structure of Cse1p to its cytoplasmic state, as measured by RMSD (white curve), and residue-resolved regions of flexibility contributing most to this transition (increasing deviation from blue to red).

Abb. 3: Spontaner Übergang der im Kern vorliegenden, offenen Struktur von Cse1p in den Zustand (weiße Kurve) und flexible Regionen in der Sequenz von Cse1p, die zu diesem Übergang am meisten beitragen (Abweichung nimmt von blau nach rot zu).



tion also locks the molecule in its final closed ring form.

Whereas the electrostatics are the main driving force for closure, they are only part of the story. As the shape of the curve at the beginning of the simulation suggests (first ~ 1 ns), the electrostatic trap does not drive the early part of the rearrangement, which is particularly fast. Instead, an energy relaxation due to torsions around bonds is dominating during this regime. A closer look at the angles between adjacent helix pairs is revealing here. Whereas the angles do not differ much between the starting and the final structure, the angle

between HEAT repeats 3 and 4 shifts by over 20° - and back - in exactly this part of the simulation.

Interestingly, the region between HEAT repeats 1 and 3 is one of the main binding sites for RanGTP in the complex. Obviously, this part of the molecule undergoes a rapid mechanical change when the strain induced by association with RanGTP is released. Closer analysis of this motion shows that it actually is a quasi-rigid body motion of the first three HEAT repeats relative to the rest of the molecule, and that it occurs around a hinge given by a conserved motif of three amino acids, Asp¹¹⁶,

Phe¹¹⁷, and Pro¹¹⁸, in the turn between helix 3B and helix 4A.

After ring closure, relaxation continues on a slower time scale. This phase is characterized by slight shifts of the angles between HEAT repeats, by rearrangements of amino acid side chains, and by changes in the geometry of some helices. The A helices are longer and more curved than the B helices, and especially helices 9A and 12A contribute to the adaptation of the molecular structure to the newly formed ring of HEAT repeats by changing their bending angle or developing a new kink in the centre, respectively.

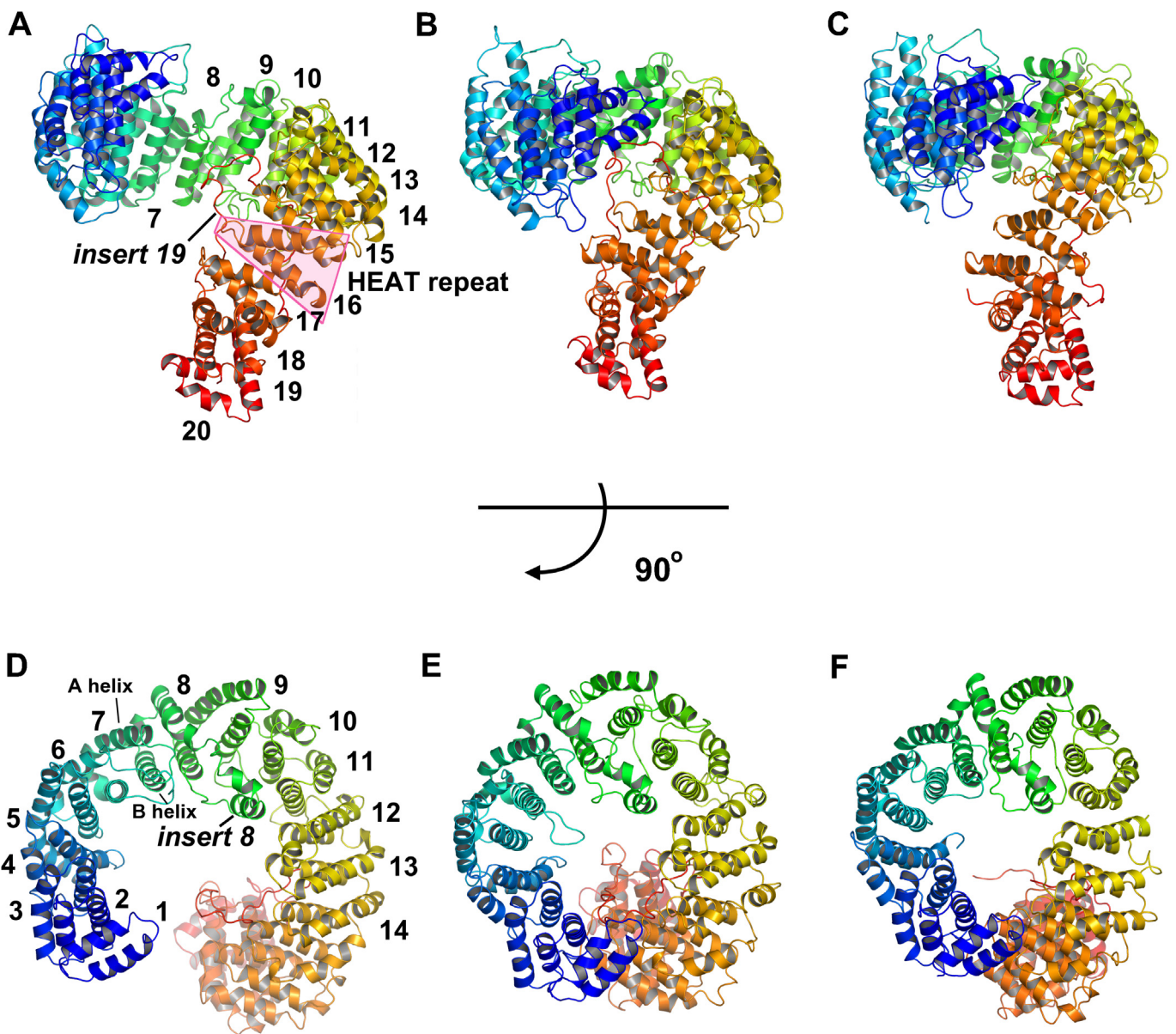


Fig. 4. (A) Initial nuclear starting structure without cargo. (B) Final structure obtained from the MD simulation. (C) Crystal structure of the cytoplasmic state. (D), (E), and (F) are top views of (A), (B), and (C), respectively.

Abb. 4: (A) Startstruktur, wie sie im Zellkern vorliegt, ohne Bindungspartner. (B) Endstruktur aus der MD-Simulation. (C) Kristallstruktur des zytoplasmatischen Zustands. (D), (E) und (F) sind um 90° gedrehte Ansichten von (A), (B), bzw. (C).

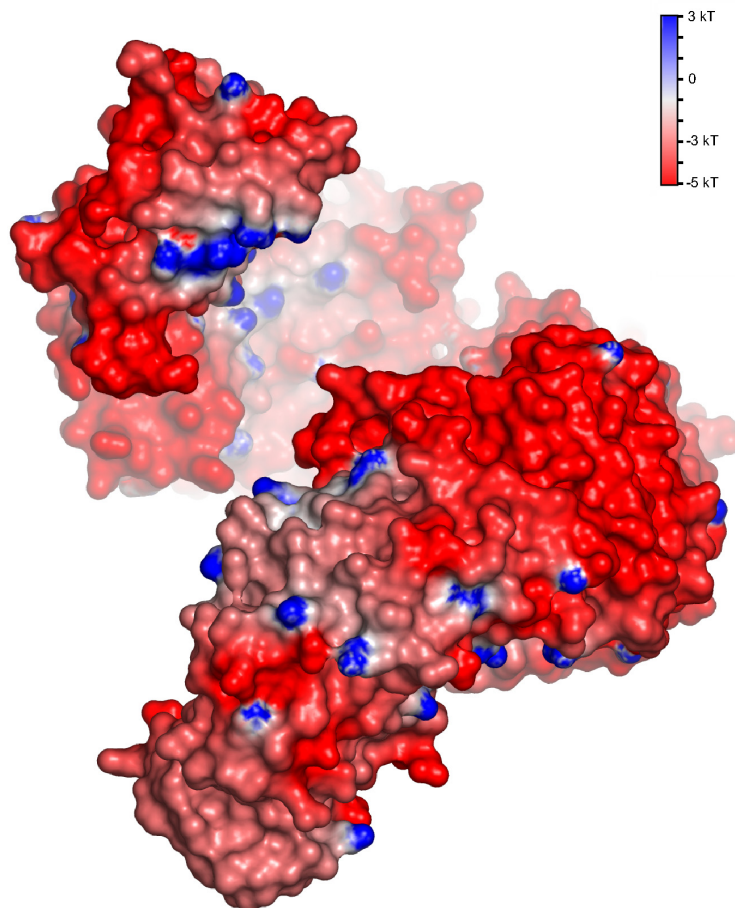


Fig. 5. Electrostatic potential on the surface of Cse1p.

Abb. 5: Elektrostatishes Oberflächenpotential von Cse1p.

connecting helices 3B and 4A appears to be an important factor in initiating the reaction which is then executed by the electrostatic trap.

Outlook

In close collaboration with the Görlich group, we will now turn our attention also towards Cse1p's "big sister", importin- β , in different complexes and the free state and its conformational flexibility in its different states. In addition, force probe simulations can be directly compared to data derived from experiments using the atomic force microscope (AFM) to stretch the complex formed by importin- β and RanGTP. Preliminary results indicate that the lack of a compact hydrophobic core not only enables changes in the tertiary structure of these extended α -helical proteins, but also destabilizes secondary structure elements under applied stress such as in the AFM. It is at present unclear to which extent this effect has biological relevance among karyopherins in general. For ankyrin, however, the capability of quick unfolding and refolding seems to be of functional relevance, also because it might be important for its function in cell adhesion.

Taken together, as sketched in Fig. 7, the slow relaxation smoothes the curvature of the section between the interfaces at HEAT repeats 1-3 and 13-16, and thus leads to the appearance of a „virtual“ hinge in the centre halfway between them, at HEAT repeat 8, after fast ring closure by the electrostatic lock.

as the basis for its flexibility. Indeed, the Lennard-Jones energies quantifying the hydrophobic interactions related to packing, e.g. in the core of a protein, do not show a significant change during the course of the simulation. By contrast, the release of tension centered at a hinge in the turn

Conclusion

From the crystallographic studies, it was proposed that the nuclear state of Cse1p constitutes a „spring-loaded“ molecule. Our simulations confirm and quantify this picture, as the open form readily collapses into the ring-shaped state usually found in the cytoplasm⁶. Generally, α -helical repeat proteins are found to be inherently flexible because of their lack of a globular hydrophobic core. Their extended, curved shape allows for multiple conformations and seamless and smooth transitions between these. What is surprising in this first free MD analysis of a conformational change in one of these proteins is that the change is primarily driven by electrostatic attractions between two functionally important parts of the molecule. For ankyrin, reconstitution of ideal packing conditions in the extended hydrophobic core was suggested

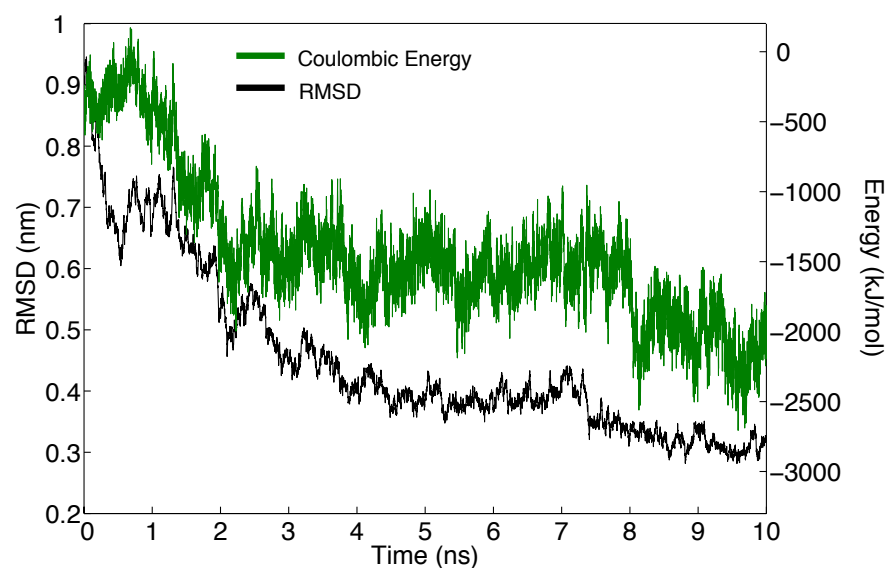


Fig. 6. Comparison of the electrostatic energy between HEAT repeats 1-3 and the interface on the opposite side of Cse1p with the decrease in RMSD.

Abb. 6: Vergleich der elektrostatischen Energie zwischen HEAT repeats 1-3 und der gegenüber liegenden Kontaktzone von Cse1p mit der Reduktion des RMSD.



References

1. Weis, K. Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell* **2003**, 112, 441-451.
2. Chook, Y. M.; Blobel, G. Karyopherins and nuclear import. *Curr. Opin. Struct. Biol.* **2001**, 11, 703-715.
3. Kutay, U.; Bischoff, F. R.; Kostka, S.; Kraft, R., Görlich, D. Export of importin- α from the nucleus is mediated by a specific nuclear transport factor. *Cell* **1997**, 90, 1061-1071.
4. Matsuura, Y.; Stewart, M. Structural basis for the assembly of a nuclear export complex. *Nature* **2004**, 432, 872-877.
5. Cook, A.; Fernandez, E.; Lindner, D.; Ebert, J.; Schlenstedt, G.; Conti, E. The structure of the nuclear export receptor Cse1 in its cytosolic state reveals a closed conformation incompatible with cargo binding. *Mol. Cell* **2005**, 18, 355-367.
6. Zachariae, U.; Grubmüller, H., *submitted*.



Ulrich Zachariae

studied chemistry in Munich and Guildford and did his Ph.D. work at the Max Planck Institute of Biochemistry in Martinsried. After postdoctoral work at Nestlé Research Center he is now working as a Postdoc in the Department of Theoretical and Computational Biophysics since 2005.



Helmut Grubmüller

1994: PhD in Theoretical Physics, Technical University of Munich. 1994-1998: Postdoctoral assistant at the Theoretical Biophysics Group, University of Munich. Research visits: Laboratoire de Biophysique Moléculaire et Cellulaire, CENG, Grenoble, France; Theoretical Biophysics Group, Beckman Institute, University of Illinois at Urbana/Champaign, U.S.A.; Institute for Molecular Biology and Biophysics, ETH Zürich, Switzerland. 1998-2003: Head of the Theoretical Molecular Biophysics Group at the Max-Planck Institute for Biophysical Chemistry. 2002: Habilitation and *venia legendi* in Physics at the Univ. of Göttingen. 2003: Associate Professor at the EPFL Lausanne. Since 2003: Director of the Theoretical and Computational Biophysics Department at the Max-Planck Institute for Biophysical Chemistry, Göttingen, and Scientific Member of the Max-Planck Society. Since 2005: Honorary Professor for Physics at the Univ. of Göttingen.

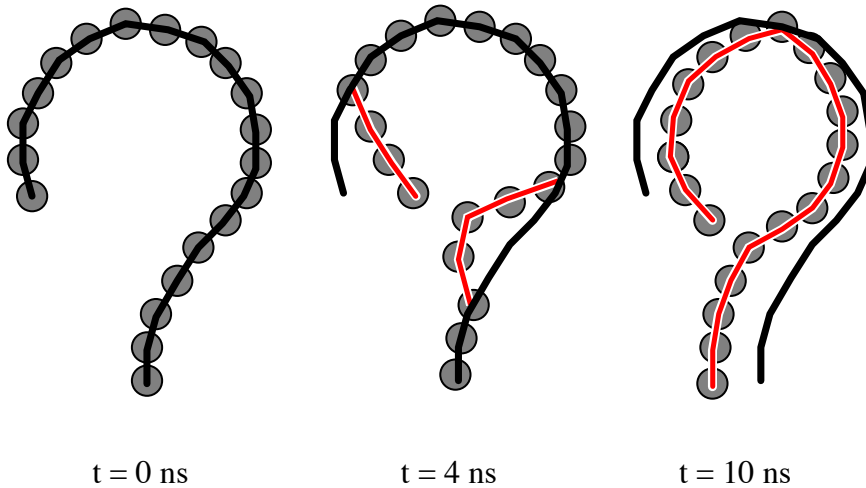


Fig.7. Schematic diagram of the closing mechanism.

Abb.7: Schematische Zeichnung des Schließmechanismus.

Zusammenfassung

Die Zellen aller höheren Lebewesen besitzen einen Zellkern, in dem die genetische Information in Form von DNA gespeichert ist. Eine Hülle, die aus einer Doppelmembranschicht besteht, grenzt diesen Zellkern vom restlichen Volumen der Zelle ab. Da DNA nur von Proteinen abgelesen werden kann, und die abgelesene Information in Form von RNA aus dem Zellkern gebracht werden muss, existieren in dieser Hülle Kanalstrukturen, sogenannte Kernporen-Komplexe. Die Zelle hat nun einen einzigartigen Mechanismus entwickelt, nur diejenigen Proteine bzw. RNA durch diese Poren passieren zu lassen, die auch im Zellkern benötigt werden. Eigens dafür spezialisierte Proteine greifen diese Fracht auf und transportieren sie durch den Kanal, während alle anderen Proteine nicht durch

den Kanal wandern können. Darunter gibt es sogenannte Importine und Exportine, je nach Transportrichtung. Diese Transport-Proteine entstammen einer Klasse von Proteinen, deren Struktur eine hohe Flexibilität vermuten lässt, so dass ihre große Oberfläche zahlreiche Bindungen mit anderen Proteinen (bzw. RNA) ermöglicht.

Wir sind dieser Vermutung mit Molekulardynamik-Simulationen nachgegangen und haben die Flexibilität eines Exportins (Cse1p) untersucht; dabei wurden im Computer die Newtonschen Bewegungsgleichungen gelöst, um die Dynamik des Proteins zu charakterisieren.

Das Exportin bindet seine Fracht (Importin- α) im Zellkern zusammen mit einem weiteren Protein (RanGTP) und transportiert beide durch die Kernpore. Außerhalb

des Kerns bricht dieser Komplex aus drei Proteinen auseinander und es ergibt sich eine große strukturelle Veränderung im Exportin (Ringschluss). Wir konnten mit unserer Arbeit zeigen, dass die im Zellkern vorkommende Form von Cse1p stark vorgespannt ist und daher nur durch die Wechselwirkung mit seinen zwei Partnern stabil bleibt. Die Simulationen zeigten ferner, dass die Strukturänderung im Wesentlichen durch elektrostatische Kräfte des Proteins angetrieben wurde. Schließlich konnte eine mechanische Schnittstelle im Protein identifiziert werden, durch die die Umlagerung eingeleitet wird. Der Ringschluss wurde durch die Computersimulation ohne Kenntnis der Endstruktur korrekt vorhergesagt.