Dynamics and Driving Forces of Macromolecular Complexes

Dissertation
for the award of the degree
"Doctor rerum naturalium"
of the Georg-August-Universität Göttingen

submitted by
Lars Bock

from Braunschweig
Göttingen 2012

Thesis Committee:

Prof. Dr. Helmut Grubmüller (Reviewer)
    Theoretical and Computational Biophysics,
    Max Planck Institute for Biophysical Chemistry
Prof. Dr. Ralf Ficner (Reviewer)
    Molecular Structural Biology,
    Georg-August-Universität Göttingen
Prof. Dr. Bernd Abel
    Molecular Physical and Biophysical Chemistry,
    Universität Leipzig

Date of the oral examination: 11.6.2012

Affidavit

Lars Bock
Eisenbahnstr. 7
37073 Göttingen

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertation mit dem Thema "Dynamics and Driving Forces of Macromolecular Complexes" selbstständig und ohne unerlaubte Hilfe verfasst habe.

Göttingen, 4.5.2012

Contents

Chapter 1

Introduction

Biological macromolecular complexes often perform sophisticated tasks in the cell, for which both efficient throughput and high accuracy are crucial. The spliceosome and the ribosome, both huge RNA-protein complexes of ~3 and ~2.6 MDa in size, as well as the nuclear pore complex (40–60 MDa) are responsible for key functions of transcriptional and translational processes [1–4]. The SNAP Receptor (SNARE) protein complex (~70 kDa) controls fast membrane fusion [5], e.g., facilitating neurotransmitter release from neurons. The membrane embedded F-ATPase, a protein complex (370–400 kDa), converts proton gradients across membranes into chemical energy stored in form of adenosine tri-phosphate (ATP) [6], providing energy for nearly all processes in the cell.

To elucidate the functional mechanisms of these machines, their motions and energetics need to be understood at an atomic level and at all relevant time scales. This is especially challenging for large complexes, due to the increased complexity. To meet this challenge, an interdisciplinary approach is necessary.

Nuclear magnetic resonance (NMR) spectroscopy allows to determine the atomistic structure of biomolecules in different states in solution [7], but it is currently not applicable to macromolecular complexes larger than 35 kDa. X-ray crystallography allows to determine the spatial arrangement of atoms within biomolecules at resolutions on the order of Å. To obtain such high resolutions, the molecules need to form rigid crystals, which is achieved under unphysiological conditions and at low temperatures. These conditions might affect the conformations of the molecules. Typically, for macromolecular complexes, only isolated components of the complex can be crystallized (e.g., spliceosome [8], F1-ATPase [9], nuclear pore complex [10, 11]). At best, if the complete structure is known, only few conformational states are captured by X-ray crystallography (e.g., ribosome [2, 12–15], SNARE complex

[16, 17]). Complete structures of macromolecular complexes, at resolutions between 6 and 30 Å were obtained by three dimensional reconstructions of cryo-electron microscopy (cryo-EM) images [18, 19] for the spliceosome [20, 21], the nuclear pore complex [22, 23], and the ribosome [24, 25]. Recent cryo-EM experiments resolved 50 different states of the ribosome with bound tRNAs undergoing spontaneous retro-translocation [26, 27]. Up to now, the resolution of cryo-EM reconstructions is low compared to resolutions that can be achieved with NMR and X-ray crystallography and is insufficient to construct atomistic models based on the reconstructions alone. To combine the information from partial or complete atomistic structures of a limited number of states with that provided by low resolution structures of multiple states, several methods to flexibly fit atomic structures into cryo-EM reconstructions have been developed [28–33].

In this thesis, to make use of the high spatial resolution achieved by X-ray crystallography and the temporal resolution achieved by cryo-EM, data obtained with both techniques were combined with molecular dynamics (MD) simulations. MD simulations allow to investigate the forces and dynamics of molecular systems at atomic resolution, employing the underlying basic laws of physics. To verify the validity of the results, they were checked against structural, biochemical, and mutagenesis data. Two macromolecular complexes were studied in detail, the SNARE complex and the ribosome. Here, a brief sketch of the two systems is given, a more detailed description is presented in the respective chapters.

The SNARE complex, in neurons, controls fusion between neurotransmitter containing vesicles and the cell membrane, inducing fast exocytosis and subsequent release of neurotransmitters into the synaptic cleft. The SNARE complex consists of three SNARE proteins, Synaptobrevin, Syntaxin 1A and SNAP-25B. Initially, Synaptobrevin is attached to the vesicle while Syntaxin 1A and SNAP-25B are attached to the cell membrane. Synaptobrevin and Syntaxin 1A each contain one and SNAP-25B contains two SNARE domains which, by forming a coiled-coil complex, pull the vesicle towards the membrane, thereby inducing fusion. The two SNARE domains of SNAP-25B are connected by a linker which has to be stretched out to allow formation of the SNARE complex.

Oxidative stress in the pre-synapse cell decreases neurotransmitter release [34] and has been linked to mitochondrial dysfunction in early stages of many neurodegenerative diseases [35]. Individual pre-treatment with a reactive oxygen species, such as H_2O_2, of the three SNARE proteins mentioned above showed that the oxidation of SNAP-25B had the largest influence on neurotransmitter release [34]. Since the SNAP-25B linker contains a cysteine-rich region, oxidation of cysteines forming one or two disulfide bonds might

lead to a shortening of the linker which in turn would not allow the SNAP-25B SNARE domains to form the full coiled-coil SNARE complex. As a consequence, the vesicle might not be pulled close enough to the cell membrane to induce fusion, impairing neurotransmitter release. This would lead to reduced synaptic transmission and possibly cognitive decline.

To test this hypothesis, first, a model of the SNARE complex based on a crystal structure [16] was built, including the SNAP-25B linker which was not resolved in the structure. Second, the disulfide bond resulting in the largest shortening of the linker was identified. Next, the time evolution of the SNARE complex in the presence and absence of this disulfide bond was calculated using molecular dynamics (MD) simulations. Finally, the simulations of the SNARE complex with and without disulfide bond were compared to see if the strain of the shortened linker leads to a destabilization of the complex itself.

The second macromolecular complex studied in this work is the ribosome, a protein-RNA complex consisting of two subunits. The ribosome synthesizes proteins in the cell. Protein synthesis needs to be sufficiently fast to ensure cell function and growth and at the same time highly accurate to avoid miscoded proteins that can be dysfunctional or even toxic. The tRNAs which deliver the amino acids to the growing peptide chain bind to the ribosome via three distinct binding sites A, P, and E. After peptide bond formation, the tRNA bound to the peptide chain occupies the A site and the empty (deacylated) tRNA is in the P site. Clearing the A site for the next tRNA to enter the ribosome, the two tRNAs translocate from A and P to P and E sites, respectively [36]. Translocation of tRNAs is accompanied by large-scale rotations of the small ribosomal subunit relative to the large subunit [27, 37]. Subsequently, the E-site tRNA dissociates from the ribosome.

Crystal structures of the ribosome bound to two tRNAs [12, 38–43] only describe states in which the tRNAs reside in the A, P, or E sites.

The cryo-EM reconstructions of spontaneous translocation were grouped into nine major intermediate states, representing the pathway of the tRNAs through the ribosome at resolutions between 9 and 30 Å [27]. However, an atomic description of translocation including the dynamics within and the barriers between these intermediate states has not been achieved yet. Aiming at such a description, crystal structures were refined against these cryo-EM reconstructions obtaining near-atomic structures for translocation intermediates and subsequently applied MD simulations to capture the dynamics and energetics of the tRNAs and the ribosome.

In particular, this thesis addresses the question of which motions of ribosomal components limit the transition rates between intermediate states and what mechanisms help to overcome these barriers, a necessary condition for

efficient translocation. A second prerequisite for translocation is the maintaining of subunit association during tranlocation. Here the question of how the affinity between the subunits is balanced despite the large-scale rotations of the subunits was addressed.

Chapter 2

Established Methods

Molecular dynamics (MD) simulations is the main method which was applied in the course of this thesis. This chapter outlines the principles and approximations of this simulation methodology as well as principal component analysis (PCA), a method that was used to extract main modes of motion from the simulations.

2.1 Principles of MD simulations

MD simulations is a computational method to calculate the motion of a many particle molecular systems by numerically integrating Newton's equation of motion of N interacting atoms.

2.1.1 Approximations of MD simulations

The conventional MD algorithm used in this work rests on three main approximations which reduce the computational effort for evaluating the time evolution of a macromolecular system: **(1)** the separation of nuclear and electronic degrees of freedom (Born-Oppenheimer approximation), **(2)** the application of a simplified potential energy function (force field), and **(3)** the classical description of nuclear motions.

(1) Born-Oppenheimer

The time evolution of a molecular system is described by the time-dependent Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, \mathbf{R}) = \hat{H}\Psi(\mathbf{r}, \mathbf{R}), \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator, i.e., the sum of the kinetic and potential energy operators, Ψ is the wave function of the system, and $\hbar$ is the reduced Planck constant. The wave function $\Psi = \Psi(\mathbf{R}, \mathbf{r})$ is a function of the positions of the m nuclei $\mathbf{R} = \{\mathbf{R}_1, ..., \mathbf{R}_m\}$ and the n electrons $\mathbf{r} = \{\mathbf{r}_1, ..., \mathbf{r}_n\}$ composing the system. The Born-Oppenheimer approximation [44] allows to separate the dynamics of the electrons from the dynamics of the nuclei. Due to the much lower mass of the electrons compared to the mass of the nuclei (mass ratio 10^4), the electronic wave function is assumed to follow the slow nuclei motion instantaneously. This assumption allows to separate the wave function into an electronic part Ψ_e and a nuclear part Ψ_n,

$$\Psi(\mathbf{r}, \mathbf{R}) = \Psi_n(\mathbf{R})\Psi_e(\mathbf{r}; \mathbf{R}), \tag{2.2}$$

where Ψ_e depends on the positions $\mathbf{R}$ of the nuclei only parametrically. Assuming fixed positions of the nuclei, the electronic wave function can be determined from the time-independent Schrödinger equation,

$$\hat{H}_e\Psi_e(\mathbf{r}; \mathbf{R}) = E_e(\mathbf{R})\Psi_e(\mathbf{r}; \mathbf{R}), \tag{2.3}$$

where the electronic Hamiltonian is $\hat{H}_e = \hat{H} - \hat{T}_n$ with the kinetic energy of the nuclei $\hat{T}_n$ being subtracted from the Hamiltonian of the complete system $\hat{H}$. Varying the positions $\mathbf{R}$, one obtains the ground state energy $E_e(\mathbf{R})$ as a function of $\mathbf{R}$, which is used as the potential energy surface (PES) of the system. As a result of the Born-Oppenheimer approximation, the dynamics of the nuclei on the PES is described by the time-dependent Schrödinger equation:

$$(\hat{T}_n + E_e(\mathbf{R}))\Psi_n(\mathbf{R}, t) = i\hbar\frac{\partial\Psi_n(\mathbf{R}, t)}{\partial t}. \tag{2.4}$$

The Born-Oppenheimer approximation holds as long as the PES of the ground state and of the first excited states do not approach each other too closely.

(2) Force field

Calculating the PES from the time-independent Schrödinger equation for the electronic wave function is still too expensive for macromolecular systems comprising several hundred thousands of atoms, as required in this thesis. Therefore, one further approximation is necessary to reduce the computational effort to a feasible amount. Accordingly, the PES $E_e(\mathbf{R})$ is approximated by the force field $V(\mathbf{R})$, a function which is set up to be computation-

ally relatively easy to calculate,

$$
\begin{aligned}
V(\mathbf{R}) =& V_b + V_a + V_d + V_{id} + V_{vdW} + V_{coul} \\
=& \underbrace{\sum_i^{\text{bonds}} \frac{k_i}{2}(b_i - b_{i,0})^2}_{\text{bond stretching}} + \underbrace{\sum_i^{\text{angles}} \frac{f_i}{2}(\theta_i - \theta_{i,0})^2}_{\text{angle bending}} \\
&+ \underbrace{\sum_i^{\text{dihedrals}} \frac{V_i}{2}\left[1 + \cos{(n\phi - \delta)}\right]}_{\text{dihedral angle}} + \underbrace{\sum_i^{\text{improper dihedrals}} \kappa_i(\xi_i - \xi_{i,0})^2}_{\text{improper dihedral angle}} \\
&+ \sum_{(i,j)}^{\text{pairs}} \left(\underbrace{4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right]}_{\text{van der Waals}} + \underbrace{\frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}}}_{\text{Coulomb}} \right) .
\end{aligned} \tag{2.5}
$$

Here, the bond stretching potential V_b, the bond angle potential V_a, and the improper dihedral potential V_{id} are described by harmonic potentials. Improper dihedral potentials restrict out-of-plane motions, e.g., to keep aromatic rings planar. Dihedral potentials are described by a cosine function with a periodicity n and the barrier height between different conformers is V_i. Non-bonded interactions, van-der-Waals V_{vdW} and coulombic V_{coul} interactions, are calculated pairwise using the atom-atom distance r_{ij}, thus neglecting many-particle effects. The van-der-Waals interactions are modeled by a Lennard-Jones potential with the parameters ϵ_{ij} and σ_{ij}. In the Coulomb potential, q_i is the partial charge of the respective atom.

The parameters for force fields are typically either derived from ab inito quantum chemical calculations or fitted to experimental data such as solvation free energies, NMR, or X-ray crystallography data. In this thesis, two force fields were used, OPLS/AA [45] for protein systems, and amber99sb [46] for systems comprising proteins and RNA.

(3) Classical description of nuclear motion

Even when using the Born-Oppenheimer approximation and a simplified potential function, solving the time-dependent Schrödinger equation for the nuclei is still rendered impossible by the large number of atoms in macromolecular systems. To reduce the computational effort even further, the motion of the nuclei is described classically by Newton's equation of motion,

$$
m_i \frac{\partial^2 \mathbf{R}_i}{\partial t^2} = m_i \mathbf{a}_i = \mathbf{F}_i, \quad i = 1...N, \tag{2.6}
$$

where m_i is the mass, $\mathbf{R}_i$ the position, and $\mathbf{a}_i$ is the acceleration of the ith atom. $\mathbf{F}_i$ denotes the force acting on atom i. These forces are obtained from the negative gradient of the force field $V(\mathbf{r}_1, ..., \mathbf{r}_N)$ which describes the interactions between all atoms,

$$\mathbf{F}_i = -\nabla_i V(\mathbf{R}_1, ..., \mathbf{R}_m). \tag{2.7}$$

The force $\mathbf{F}_i$ acting on atom i determines the velocity change of this atom leading to an updated position after a discrete time step Δt. All the afore-mentioned approximations enabled us to describe macromolecular systems at atomic detail in a computationally feasible manner. However, for any application the validity of these approximations should be checked, e.g., by comparing the obtained observables to experimental data. Further, chemical reactions, charge transfers and excited electronic states cannot be described by conventional MD simulations.

2.1.2 Integrating Newton's equation of motion

In MD simulations, Newton's equation of motion is numerically integrated to obtain the time evolution of the system. In this work, a Verlet-type algorithm [47], the leap-frog algorithm [48], was applied,

$$\mathbf{R}(t + \Delta t) = \mathbf{R}(t) + \mathbf{v}\left(t + \frac{\Delta t}{2}\right)\Delta t$$
$$\mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t - \frac{\Delta t}{2}\right) + \mathbf{a}(t)\Delta t. \tag{2.8}$$

Integration time step

The integration time step Δt needed for stable integration depends on the chosen propagation algorithm and on the fastest motions explicitly described. In biomolecular systems, these are the bond stretch vibrations of hydrogen atoms. To trade off avoiding numerical instabilities against sufficient phase space sampling, the time step should be five times smaller than the shortest period of motion when using the leap-frog algorithm [49]. Explicit simulation of bond vibrations requires an integration time step of 1 fs.

Bond constraints

Since bond vibrations hardly couple to the other degrees of freedom of macro-molecules, bond lengths were restrained, which allowed us to increase the integration time step to 2 fs. Furthermore, it has been argued that constrained

vibrations are a better representation of the quantum-mechanical nature of bond vibrations than the classical harmonic oscillator [50]. In this work, the LINCS algorithm [51] was used to constrain protein and RNA bond lengths and the SETTLE algorithm [52] was used for water molecules.

Virtual interaction sites

After constraining bond vibrations, the next shortest period is that of the bond-angle vibration of hydrogens with a period of 13 fs [53]. To increase the applicable time step to 4 fs, hydrogen atoms were defined as virtual interaction sites [53]. Using virtual interaction sites, the position of hydrogen is generated at every MD step based on the positions of three nearby heavy atoms, thus removing its high-frequency degrees of freedom. All the forces acting on the hydrogen are redistributed to these heavy atoms for the next integration step.

2.1.3   Thermodynamic ensembles

Integrating Newton's equation of motion in an isolated system generates a microcanonical ensemble (NVE) of conformations. Since the biomolecular systems studied in this thesis are usually in thermal and mechanical equilibrium with their environment, canonical (NVT) and isothermalisobaric (NPT) ensembles are the more natural choice.

In order to generate these ensembles, the temperature (NVT) or additionally the pressure (NPT) have to be coupled to given reservoirs. To account for short term temperature and pressure fluctuations in the simulated system, the coupling is not instantaneous, but rather, a coupling time constant is introduced. Several algorithms for temperature coupling, thermostats, and for pressure coupling, barostats have been developed. The Berendsen thermostat controls the temperature by rescaling the velocities of the particles and in this way approaches the desired temperature [54]. However, this weak coupling to an external bath was shown to not generate the correct canonical ensemble [55]. The velocity-rescaling thermostat which extended the Berendsen thermostat by a properly constructed random force, instead, generates a correct ensemble [56]. Therefore the velocity-rescaling thermostat was used for equilibrium simulations.

In addition, the Berendsen barostat has been used which couples the pressure of the system to the reference value by rescaling the simulation box and thus changes the volume of the simulation system [54].

2.2   Principal Component Analysis

MD trajectories represent a huge amount of high-dimensional data, i.e., the positions of all atoms of the system in phase space for each time frame. Principal component analysis (PCA) [57] is an orthogonal transformation that converts a set of coordinates, in which motions along different axes are possibly correlated, e.g., the Cartesian coordinates of the atoms, into a new coordinate system, the principal components. In the principal component system the motions along axes are uncorrelated. When sorted by descending fluctuations of the projection onto this component, the first components describe large correlated motions of the system. Since such motions are often related to the function of the macromolecular system and few principal components suffice to describe most of the structural fluctuations, PCA allows to reduce the dimensionality of the system to a much smaller number of essential coordinates.

The principal components are the eigenvectors of the covariance matrix C of the atomic coordinates x_i,

$$
\begin{aligned}
C =& \langle (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^T \rangle \\
\Rightarrow & C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle ,
\end{aligned}
\tag{2.9}
$$

where $i, j \in [1, 3N]$, and N is the number of considered atoms. As C is a symmetric matrix, it can be diagonalized by an orthogonal coordinate transformation T,

$$
\mathbf{x} - \langle \mathbf{x} \rangle = T\mathbf{q} \Rightarrow \mathbf{q} = T^T(\mathbf{x} - \langle \mathbf{x} \rangle)
$$

$$
\Rightarrow (T\mathbf{q})^T = \mathbf{q}^T T^T = (\mathbf{x} - \langle \mathbf{x} \rangle)^T .
$$

T transforms C into the diagonal matrix $\Lambda = \langle \mathbf{q}\mathbf{q}^T \rangle$ of eigenvalues λ_i,

$$
C = TT^T C TT^T = TT^T \langle (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^T \rangle TT^T
$$

$$
= T \langle T^T (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^T T \rangle T^T
$$

$$
= T \langle \mathbf{q}\mathbf{q}^T \rangle T^T .
$$

The columns of T are the eigenvectors of the covariance matrix C, also called principal or essential modes. The total positional fluctuation can be expressed by the eigenvalues λ_i,

$$
\sum_i \langle (x_i - \langle x_i \rangle)^2 \rangle = \langle (\mathbf{x} - \langle \mathbf{x} \rangle)^T (\mathbf{x} - \langle \mathbf{x} \rangle) \rangle
$$

$$
= \langle \mathbf{q}^T T^T T \mathbf{q} \rangle = \langle \mathbf{q}^T \mathbf{q} \rangle = \sum_i \langle q_i^2 \rangle = \sum_i \lambda_i .
$$

The eigenvalues describe the variance of the position along the eigenvectors. By projecting the trajectory on the principal modes the principal components $q_i(t)$ are obtained:

$$\mathbf{q}(t) = T^T(\mathbf{x}(t) - \langle \mathbf{x}(t) \rangle).$$

Chapter 3

Newly developed Methods

This chapter introduces the methods that have been developed in the course of this work. The method to obtain rates for transitions between different states of a macromolecular system from multiple simulations has been developed together with Christian Blau and Andrea Vaiana.

3.1 Transition rate estimates

Macromolecular complexes can adopt several meta-stable conformations (states) which are often linked to their function. Here, we develop a method to estimate barrier heights and transition rates for collective motions between these states. Transition rate estimates can be obtained for motions involving the whole complex or for motions involving components of the complex as done in this work. The rates were estimated in two steps. First, the heights of the free energy barriers $\Delta G_{\text{est}}^{\ddagger}$ were obtained from a fluctuation analysis of the trajectories of all different states. Second, these barrier estimates were gauged by comparing passage frequencies obtained from the initial barrier estimates to passage frequencies actually observed in the simulations to yield the free energy barriers $\Delta G^{\ddagger}$.

3.1.1 Free energy barrier estimates

Initial free energy barrier estimates for the collective motions of components of the complex were obtained as follows. Each trajectory (state) was projected onto d-dimensions spanned by d reaction coordinates, e.g., dominant PCA eigenvectors or rotation angles, describing the motion of the component (Fig 3.1A). The mean $\boldsymbol{\mu}$ and $d \times d$ covariance matrix $\boldsymbol{\Sigma}$ of each projected trajectory define a d-dimensional multivariate Gaussian probability distribution
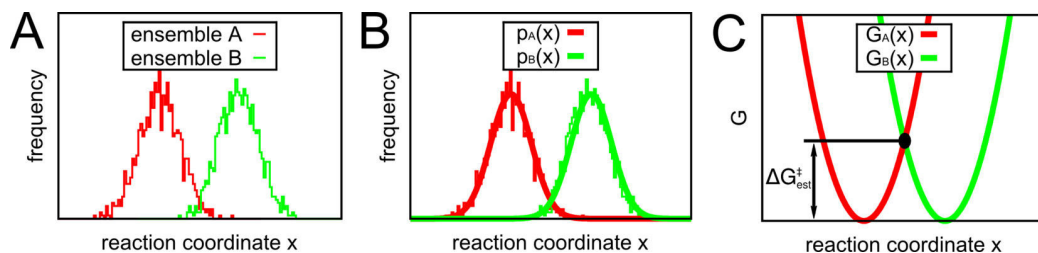
Figure 3.1: **Scheme of barrier height estimation for the transition between state A and B along an 1-dimensional reaction coordinate.** (**A**) Histogram of the trajectories projected onto the reaction coordinate (state A and state B). (**B**) Gaussian probability distribution functions for states A and B. (**C**) Free energy landscapes described by quasi-harmonic approximation, where the intersection yields the barrier height estimate $\Delta G^{\ddagger}$.

function,

$$p(\boldsymbol{x}) = \frac{A}{\sqrt{2\pi^3 \|\boldsymbol{\Sigma}\|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \tag{3.1}$$

for the corresponding state (Fig 3.1B). The free energy landscape

$$G(\boldsymbol{x}) = -k_b T \ln p(\boldsymbol{x}) \tag{3.2}$$

arising from such a distribution $p(\boldsymbol{x})$ is a d-dimensional quasi-harmonic approximation to the underlying free energy landscape (Fig 3.1C). For each pair of states, the intersection of the two quasi-harmonic approximations obtained for a given motion defines a hypersurface, the free energy minimum of which was used as an estimate for the barrier height $\Delta G^{\ddagger}_{\mathrm{est}}$ between the two states (Fig 3.1C). To test whether barrier estimates change with the number of dimensions used for the analysis, barrier heights were estimated using an increasing number of reaction coordinates. Even though the absolute values for the free energy barrier estimates increased with the number of dimensions, as expected, their values normalized by the mean value of all estimates did not change markedly in this work. Therefore, we assumed $\Delta G^{\ddagger} = c\Delta G^{\ddagger}_{\mathrm{est}}$ with a constant gauge factor c. In this work, we assumed that each barrier is scaled by the same gauge factor c for each transition by using the same factor for all motions and all transitions between states. Several transitions for low barriers are seen during the simulations, which specifies the gauge factor c. The same gauge factor can then be used for high barriers for which no transtions are found.

3.1.2 Attempt rate and gauge factor from passage frequencies in PCA and angle space

From the transitions between states actually observed in simulations with low free energy barriers, we estimated the gauge factor c for the free energy barrier estimates and determined the attempt rate of barrier crossing. To that aim, passage frequencies that describe how many transitions between states are seen in a given free energy barrier estimate interval were determined as follows. The trajectories were projected onto the reaction coordinates for each collective motion. The distance $d_{A,B}$ between two state ensembles A and B was defined as the minimum distance of all projections onto the reaction coordinates. A barrier between A and B states was considered to be crossed if the average distance within an ensemble was found to be larger than the distance $d_{A,B}$. For each motion, the frequency of barrier crossing $p_{A \to B}^{sim}$ in an energy interval $I = [\Delta G_{est}^{\ddagger} - \Delta G, \Delta G_{est}^{\ddagger} + \Delta G]$ was defined as

$$p_{A \to B}^{sim} \left(\Delta G_{est}^{\ddagger} \right) = \frac{n_{A \to B}}{n}, \tag{3.3}$$

where $n_{A \to B}$ is the number of observed barrier crossings from ensemble A to B with estimated free energy barriers in I, and n is the total number of barriers in the same energy interval. The probability of observing a transition from a state A to a state B in a time interval $[0, t]$ is also known from reaction kinetics to be

$$p_{A \to B} = 1 - \exp(-kt), \tag{3.4}$$

where k is the transition rate from A to B. Transition rates k are estimated by Arrhenius' law

$$k = A \exp\left(-\Delta G^{\ddagger}/k_b T\right), \tag{3.5}$$

where k_b is the Boltzmann constant, and T is the temperature. The combination of the latter,

$$p_{A \to B} = 1 - \exp\left[-A \exp\left(\frac{-c \Delta G_{est}^{\ddagger}}{k_b T}\right) t\right], \tag{3.6}$$

determines the gauge factor c and an attempt rate A by a least-square fit of $p_{A \to B}$ to $p_{A \to B}^{sim}$.

Transition time estimates obtained from Arrhenius' law

Refined transition times $\tau = 1/k$ between states determined from the half-times given by Arrhenius' law using the obtained estimated attempt rate A and the corrected free energy barrier estimate $\Delta G^{\ddagger}$.

3.2 Kinetic sequence of states

In this work, atomic structures were refined against several cryo-EM reconstructions representing different states [27]. These reconstructions have been grouped and put into sequence based on structual similarity. To address the question wether this sequence of states matches the kinetic sequence of states for the macromolecular complex, we determined the one which best matches the observed set of transition rates as follows. From all reaction sequences, the one which best matches the observed set of transition rates was determined: The best matching reaction sequence (i.e., one permutation of all states, {state A, state B, ...}) was assumed to be the one that yields the shortest overall half-time τ. This overall half-time is proportional to the sum of the half-times of the state transitions $\tau_{A \to B} \propto \exp\left(\Delta G^{\ddagger}_{A \to B}\right)$. From the barriers of the motions of the individual components, the highest transition barrier estimate dominates the overall half-time, so only this barrier which governs the transition for a given pair of states was taken into account. If we require all states to be visited, the sequence with the shortest overall half-time of states minimizes $\sum_{i}^{\text{transitions}} \exp\left(\Delta G^{\ddagger \text{max}}_{i}\right)$, where $\Delta G^{\ddagger \text{max}}_{i}$ is the highest barrier of all the barriers estimated for the individual component for transition i. To avoid that only very high barrier estimates with large errors dominate the calculation, in this work, barrier estimates higher than 80 kT were set to 80 kT.

3.3 Clustering of dynamic contacts between components of macromolecular complexes

The components of macromolecular complexes interact via contact interfaces. The interactions across these interfaces are crucial for the function of the complex, as they serve to balance the affinity between the components. The affinity needs to be high enough such that the components remain associated as long as necessary for accomplishing the function and low enough to allow the components to dissociate after accomplishment. In the case of large conformational changes of the components, this task is particularly challenging. Further, if information is exchanged between the components, e.g., during allosteric regulation, it has to be transmitted by the interactions between the components when passing the interface.

In this work, we studied the balancing of affinity for the 6000 Å^2 interface between the two subunits (50S and 30S) of the ribosome [58], which comprise hundreds of residue-residue contacts. To describe the dynamics
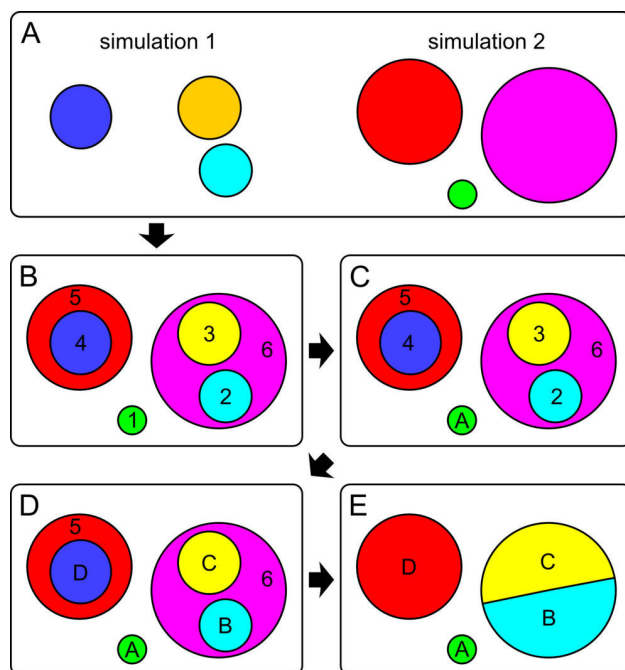
Figure 3.2: **Scheme of the clustering protocol applied to intersubunit contacts obtained from the simulations.** (**A**) Clusters of contacts are represented by colored circles (from simulation 1: blue, orange, cyan; from simulation 2: red, magenta, green). (**B**) Clusters of different simulations were combined and sorted by size (small to large). (**C**) The smallest cluster was labeled "unified" cluster A, and excluded from subsequent clustering. (**D**) With the remaining sorted clusters, this procedure was repeated until all clusters were assigned (clusters B, C, D). (**E**) Contacts from the remaining clusters were distributed to the closest "unified" cluster, where distance was defined as the number of contacts which connect the residue to the cluster.

and energetics of the interactions between subunits, residue-residue contacts were extracted and then clustered into groups of contacts. From the trajectory of each state, all pairs of atoms, one from the 50S and the other from the 30S subunit, whose distance was smaller than 3 Å for at least 1 % of time, were considered atom-atom contacts. An efficient method to extract these pairs of atoms, developed by Christian Blau, was used to cope with the huge amount of data in a reasonable time. These atom-atom contacts were assigned to residue-residue contacts. In the simulations of the different states, a different set of residues was involved in intersubunit contacts. To define localized groups of residue-residue contacts present in more than one state and to keep groups separate, which are not connected in at least one state, the contacts were clustered in the following four steps (see Fig. 3.2):

- For each state all residues connected by intersubunit contacts were put into one cluster (Fig. 3.2A).

- All clusters of all states were sorted by size (Fig. 3.2B). A new set of "unified" clusters was defined and the smallest cluster was added to it (Fig. 3.2C). Iteratively, the next (larger) clusters which did not have overlaps with any of the "unified" clusters were added to the set of "unified" clusters (Fig. 3.2D).

- To each of the new clusters, residues from all remaining clusters, overlapping only with this "unified" cluster, were added (Fig. 3.2E).

- The residues of remaining clusters were distributed to the closest overlapping "unified" cluster, where distance was defined as the number of contacts which connect the residue to the cluster.

Chapter 4

Chemomechanical Regulation of SNARE Proteins[1]

4.1 Introduction

Biological membranes consisting of lipid bilayers enclose cells and separate compartments inside cells. Membrane fusion, the merging of two initially separated membranes into one membrane, is essential for communication across cell compartments and for communication of the cell with the surrounding [60]. Communication across membranes can be performed via the exchange of signaling agents, e.g., neurotransmitters and hormones. Signaling agents are released from cells via exocytosis, where a vesicle which contains signaling agents and is initially inside the cell, fuses with the limiting membrane of the cell. In this way, the content of the vesicle is expelled into the extracellular space (Fig. 4.1).

The exocytosis of synaptic vesicles, releasing neurotransmitters into the synaptic cleft, is mediated by three SNARE proteins [16], SNAP-25B, synaptobrevin, and syntaxin 1A, which form a macromolecular complex. Initially, synaptobrevin is anchored by its transmembrane helix to the synaptic vesicle, whereas the transmembrane helix of syntaxin 1A spans the plasma membrane and SNAP-25B is targeted to the plasma membrane by palmitoylation (Fig. 4.1A). The nearly identical isoform of SNAP-25B, SNAP-25A, is expressed early in development [61]. The key role of SNAP-25B is the formation of a protein complex leading to fusion of synaptic vesicles with the plasma membrane. The complex forms sequentially, with SNAP-25B first in-
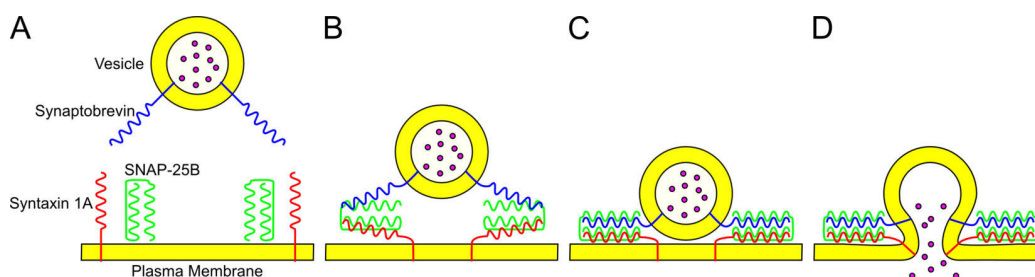

Figure 4.1: **Scheme of exocytosis mediated by SNARE proteins.** (**A**) Shown is the vesicle membrane (top) filled with neurotransmitters (magenta circles), the plasma membrane (bottom) and SNARE proteins (colored lines), before exocytosis. Synaptobrevin is anchored in the vesicle membrane and syntaxin 1A in the plasma membrane. SNAP-25B is targeted to the plasma membrane. (**B,C**) Upon formation of the SNARE complex by interwinding of the 4 helices, where synaptobrevin and syntaxin 1A each contribute one helix and SNAP-25B contributes two helices, the membranes are brought into close proximity. (**D**) Finally, the complete SNARE complex induces fusion, thereby releasing neurotransmitters.

teracting with syntaxin 1A and then with synaptobrevin within the synaptic vesicle membrane. The three proteins form a four helix bundle with SNAP-25B contributing two of the four alpha helices (helix 1 and helix 2). Each helix, known as a SNARE domain, is composed of ∼63 amino acids. The helices form a coiled-coil complex, with layers of highly conserved interacting residues in the center [62], such that the two SNARE domains of SNAP-25B have to stretch out a central linker region of 58 amino acids to the full length of the complex in order to align. The current model of SNARE protein function is that the two membranes are brought into close proximity during the formation of this SNARE complex (Fig. 4.1B,C), thereby inducing the fusion (Fig. 4.1D).

The SNARE complex has been crystallized showing the coiled-coil nature of the three protein complex [16, 17], but no structure has been determined so far for the SNAP-25B linker.

The two SNARE domains of SNAP-25B are connected by a linker region which has several unique features (Fig. 4.1). Firstly, removal of the linker slows down SNARE assembly [63] and decreases secretion from intact chromaffin cells [64]. Additionally, the linker contains a conserved region of amino acids that aids in targeting SNAP-25B to the membrane [65]. As shown in Fig. 4.2 A, this membrane binding domain (MBD) is close to the N-terminal of the linker and includes a cysteine-rich domain that contains four cysteines
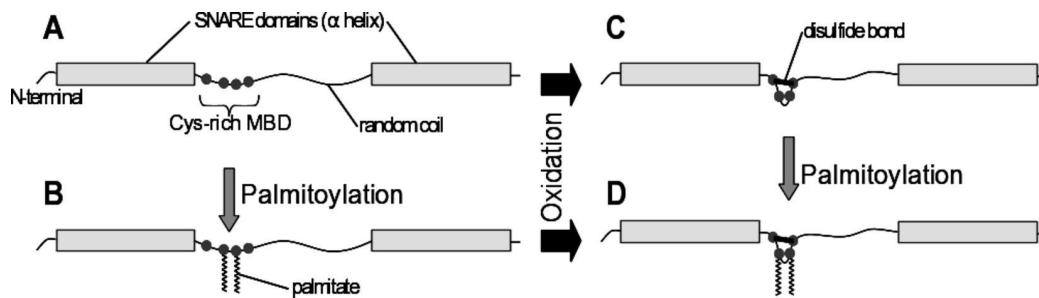
Figure 4.2: **Putative SNAP-25B response to oxidative stress.** (**A**) "naked" SNAP-25B with its two SNARE domains and four cysteines (C85, C88, C90, C92). (**B**) In vivo, SNAP-25B is posttranslationally modified by palmitoylation, which adds 1-4 palmitic acids to cysteines within the weak membrane-binding domain (MBD); palmitoylation of the middle two cysteines is shown. (**C**) Oxidation of SNAP-25B could result in disulfide bond formation between any of the free cysteines (one of six possibilities is shown). (**D**) One specific example of a SNAP-25B that has the inner two cysteines palmitoylated and the outer cysteines disulfide linked.

within an eight amino acid section. One or more palmitoylated cysteines contribute to the targeting SNAP-25B to the membrane [66, 67]. Binding to syntaxin may also help target SNAP-25B to membranes [68], but this is controversial [69].

It has generally been assumed that all four cysteines of SNAP-25B are palmitoylated and thus function solely in membrane targeting, but several other functions for these cysteines have been implied. Huang et al. [70] showed that the four cysteines of SNAP-25A can form an iron-sulfur cluster, although there is no data that such a cluster actually forms in vivo. Pallavi and Nagaraj [71] reported that peptides from the linkers of SNAP-25/23 can enhance membrane fusion depending on the extent of palmitoylation. Giniatullin et al. [34] showed that oxidation of SNAP-25B was the likely cause of the reduced secretion observed in neurons exposed to oxidative stress. Such stress is known to play a role in many neurodegenerative disorders (e.g. Alzheimer's and Parkinson's) [72, 73]. Although the site of oxidative modification was not identified, cysteines are the most susceptible residues and frequently are involved in redox control within cells through the formation of disulfide bonds [74].

Thus, the four cysteines of SNAP-25B may function as sites of palmitoylation, in membrane fusion, in heavy metal coordination, or in disulfide bond formation. These potential modifications need not be exclusive, since all three are reversible and one modification might have a physiological effect

simply by preventing a different modification. Additionally, cysteine oxidation and partial palmitoylation could occur simultaneously (Fig. 4.2 D). Only two cysteines are required to form a single disulfide bond and palmitoylation of two cysteines would still provide a strong membrane anchor for the protein. The extent of palmitoylation appears to depend on cell type and environment. In HIT (insulin-secreting) cells palmitoylation is not critical for function as the replacement of all four cysteines still allowed significant secretion [67], but in PC12 cells the same modification prevented exocytosis [75].

The possibility that cysteine oxidation in SNAP-25B may have functional significance is intriguing, but raises two questions. First, under which conditions might SNAP-25B be oxidized and how might oxidation appropriately alter the function of SNAP-25? The location of SNAP-25B at the synapse places it at a good location to sense oxidative stress from internal sources (mitochondria) and from the extracellular environment. Under such stresses, neuron survival would likely be enhanced if the energy-requiring steps of neurosecretion were temporarily reduced. Although the redox state of the intracellular environment is generally reduced, reactive oxygen species produced after intense neuronal activity [74, 76, 77] could oxidize the cysteines of SNAP-25. As previously noted by Giniatullin et al. [34], oxidation of SNAP-25B decreases formation of the SNARE complex and thus could be the direct cause of decreased secretion.

The second question is the focus of this chapter: How might oxidation of SNAP-25B prevent complex formation? We hypothesizes that since the linker is not much longer than the SNARE domain, formation of a disulfide bond between two cysteines could shorten the loop sufficiently to hinder or even block the proper alignment of the two SNARE domains in SNAP-25B and thus also the formation of the complete coiled-coil complex. To test this idea, we carried out multiple molecular dynamics (MD) simulations which indeed showed several marked conformational changes of SNAP-25B induced by formation of a single disulfide bond.

4.2   Methods

In this section, the methods for the estimation of linker shortening due to disulfide bond formation, which was done by Brian Hutchings in the lab of Dixon Woodbury and the methods for the extensive MD simulations of the SNARE complex in different setups and the subsequent analysis of the trajectories are described.

4.2.1 Estimation of linker shortening caused by disulfide bonds

In a first step, the length decreases due to the 10 different possibilities for disulfide bonds were estimated from short term simulations of an 18 amino acid peptide from the SNAP-25B linker. From these, the one with the largest effect was selected and studied in subsequent extended simulations of the whole SNARE complex.

Ten different configurations of the peptide from the SNAP-25B linker (G K F C(1) G L C(2) V C(3) P C(4) N K L K S S D) were constructed. The configurations differed in number and position of disulfide bonds: none, C1-C3, C2-C4, C2-C3, C1-C4, C1-C2, C3-C4, C1-C4 & C2-C3, C1-C3 & C2-C4, and C1-C2 & C3-C4. Each of the 10 configuration was then relaxed and stretched through 41 iterations using proportional atom translations followed by geometry optimization through energy minimization with the CHARMM22 force field [78]. After each iteration, the peptide was placed in a water box of size $12.5 \times 5 \times 5$ nm and K^+Cl^- at a concentration of 0.154 mol/l were added. In this way, a total of 410 configurations were retained.

Molecular dynamics were run on each of these systems using CHARMM version c35b1 [79] with the CHARMM22 force field and a time step of 2 fs. Periodic boundaries were applied using an orthogonal unit cell of $12.5 \times 5 \times 5$ nm. The N- and C-termini were restrained at their original separations during the dynamics run with an energy penalty of 1000 kcal·mol^{-1}nm^{-1}. Coordinates were stored every 2 ps. All bonds to hydrogen atoms were held rigid using the ShakeH algorithm as implemented in CHARMM. Langevin dynamics were used to maintain the pressure at 1 atm with a collision frequency of 20 ps^{-1}, a pressure piston mass of 400 amu, and a piston bath temperature of 300 K. Each system was heated from 48 K to 298 K for 50 ps and equilibrated for another 50 ps.

For each frame of the trajectories, the distances between the N-terminus nitrogen and the C-terminus carbon and the energy of the system without water and salt were computed and average values were obtained for each of the 410 systems.

4.2.2 Molecular dynamics simulation of the SNARE complex

To create the SNARE complex model used in this study, the structure reported by Sutton et al. [16] (pdb code: 1SFC) was used. It contains four

α-helix fragments, where the C-terminal of helix 1 and the N-terminal of helix 2 do not extend as far as the other helices and the linker between them and other parts of the proteins have been removed by proteolysis. A more recent crystal structure of the SNARE complex that includes the membrane spanning region of syntaxin and synaptobrevin, shows helices extending continuously through the SNARE domain and membrane spanning domain [80]. Therefore, the short ends of helices 1 and 2 were extended several coils such that they ended approximately even with the ends of the other helices, thus maximizing possible helix-helix interactions. The remaining residues between helix 1 and 2 (Cys88-Met127) were also added as an unstructured linker between the two helices. The resulting structure was then energy minimized.

This model of the complete SNARE complex was then solvated in a dodecahedron box with a minimum distance of 1.5 nm between the proteins and the box boundary. Ions were added at a physiological concentration of 0.154 mol l^{-1} with the program genion from the GROMACS simulation suite [81]. The total system size was $\sim$360,000 atoms. The simulation system was energy minimized, and the solvent was subsequently equilibrated for 1 ns, applying position restraints on the heavy atoms of the proteins using a force constant of k = 1000 kJ mol^{-1} nm^{-2}.

Five sets of simulations were carried out, termed "reduced" (without disulfide bond), "oxidized" (with disulfide bond), "cut linker uncharged", "cut linker charged" (with the linker cut between residues N107 and N108, using uncharged and charged termini for the linker, respectively), and "no linker" (removed linker residues, N93–Q126; charged termini). For the "reduced" simulations, the energy minimized system was equilibrated without any restraints. For the "oxidized" simulations, the sulfur atoms of the SNAP-25B residues C85 and C92 were pulled towards each other applying a harmonic potential with a linearly increasing force constant from zero to the value of the disulfide bond force constant during 1 ns of simulation. Subsequently, the disulfide bond was established and the simulation was continued without any restraints. For the "cut linker charged" simulations, the peptide bond between N107 and N108 was removed, charged termini were added, and the system was simulated for 1 ns with position restraints on the protein heavy atoms. Subsequently, the system was equilibrated without restraints. The same protocol was applied for "cut linker uncharged" simulations. Additionally, a simulation with removed linker (N93–Q126) was carried out ("no linker").

To create five independent trajectories for each of the five models, snapshots were taken at times 10, 20, 30, and 40 ns, respectively, new velocities were assigned to the atoms, and new simulations were started from these snapshots. After 100 ns of the "oxidized" simulations, snapshots from the

trajectory were taken, the disulfide bond was removed, the system was energy minimized, and subsequently equilibrated with position restraints on the backbone atom for 1 ns. Afterwords, new simulations without position restraints were started ("removed disulfide").

All simulations were carried out with the MD software package GRO-MACS 4 [81]. The OPLS all-atom force field [45] and the TIP4P [82] water model were used. The temperature was kept constant by velocity rescaling [56] at $T = 300$ K with a coupling time constant of $\tau_T = 0.1$ ps. The system pressure was kept at $p = 1$ atm by coupling the system to a Parrinello-Rahman barostat [83] with a coupling time constant of $\tau_p = 1$ ps and an isotropic compressibility of $4.5 \cdot 10^{-5}$ bar^{-1}. All bond lengths were constrained with the LINCS algorithm [51]. Electrostatic interactions were calculated pairwise for distances smaller than 0.9 nm. Long-range electrostatic interactions were calculated by particle-mesh Ewald summation [84] with a grid spacing of 0.12 nm. Lennard-Jones interactions were calculated for distances smaller than 1.4 nm. Coordinates were recorded every 1 ps.

4.2.3   Analysis

Root mean square deviations (rmsd values) were calculated after fitting to the backbone of the residues which are part of the helices in the crystal structure.

To monitor the angles between different parts of the SNAP-25B helix, directional vectors describing the orientation of each part were calculated for each frame. To calculate these vectors, coordinates of the backbone atoms of every four successive residues from one part were averaged, yielding positions close to the central axis of the helix. The eigenvector with the largest eigenvalue of the covariance matrix of these averaged coordinates was then used as the orientation vector of the respective part. For each frame, this vector was calculated for both parts, and the angle between these vectors was calculated.

Secondary structure content was determined with the program DSSP [85].

4.3   Results and Discussion

The motivation of this work was to test the hypothesis that the formation of a disulfide bond upon oxidation of two cysteines shortens the SNAP-25B linker, thus hindering SNARE complex formation, which would explain a depression of neurotransmitter release.

31
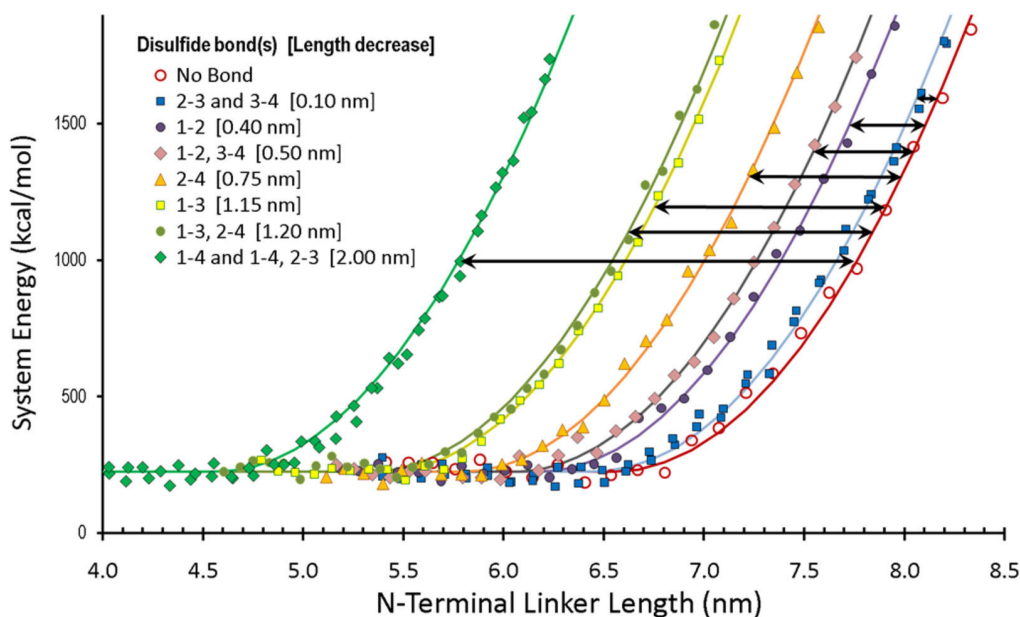
Figure 4.3: **Shortening of SNAP-25B linker due to formation of disulfide bonds.** N-terminal linker system energy as a function of end-to-end distance for different combinations of disulfide bonds in SNAP-25B. Semi-harmonic functions, $\frac{1}{2}k(x - \Delta x)^2$, were used to fit the data with a common spring coefficient k and adjustable offsets Δx. Length decrease for the nine disulfide linkages with respect to the reduced state is shown in brackets and by black arrows.

To test this hypothesis, we proceeded in two steps. First, we screened for the influence of all ten possible disulfide bonds on the linker length with geometry optimizations of the relevant part of SNAP-25B. Second, to investigate the effect of the most restrictive disulfide bond on the SNARE complex, we modeled the complete complex with extended SNAP-25B helix ends and the complete SNAP-25B linker, and compared simulations with ("oxidized") and without disulfide bond ("reduced"). Additionally, simulations with the cut linker and with removed linker were carried out to investigate the effects of the linker on the SNARE complex and to separate these effects from the effect of the mechanical strain generated by the shortening of the linker.

Name	Disulfide bond(s)	Length decrease [nm]	Shortening [#bonds]
1-2	C85-C88	0.38	4
1-3	C85-C90	1.16	10
1-4	C85-C92	1.98	16
2-3	C88-C90	0.092	1
2-4	C88-C92	0.76	7
3-4	C90-C92	0.109	1
1-2, 3-4	C85-C88 & C90-C92	0.50	5
1-3, 2-4	C85-C90 & C88-C92	1.22	10*
1-4, 2-3	C85-C92 & C88-C90	1.99	16

Table 4.1: **Length decrease for the nine disulfide linkages with respect to the reduced state.** Length decrease was estimated from the offset of the fitted semi-harmonic function for each data set. For each disulfide bond formed, the net number of bonds removed from the total segment is shown. As expected, there is a fairly linear relationship between the end-to-end distance and the number of bonds in the fully extended test segment. On average, the atom-atom distance is 0.128 nm or 0.38 nm per amino acid, which gives a simple way to estimate the maximum length of a peptide. *The shorting is not clear for this system because the 2-4 disulfide interacts with the loop of the 1-3 disulfide.

4.3.1 Influence of different disulfide bonds on linker length

To estimate how much the linker shortens after formation of different disulfide bonds, we generated all six possible single-disulfide linkers and the three double-disulfide linkers and compared the length of each to the linker without disulfides. Using the N-terminal 18 amino acids of the linker for each of the ten possibilities, molecular modeling was carried out for varying distance between the N- and C-terminal amino acids. Figure 4.3 shows the energy of each system as a function of end-to-end distance imposed by constraints on the termini followed by energy minimization as described in Methods. As expected for elastic stretching, each curve can be fitted well by a semi-harmonic function shifted by some offset. The offset change with respect to the offset of the reduced state was used to estimate the amount of shortening produced by the respective disulfide linkages (insert of Fig. 4.3 and Table 4.3.1). The largest shortening (2.0 nm) was observed for the C85-C92 single-disulfide and for the C85-C92 & C88-C90 double-disulfide.

4.3.2   Effects of the disulfide bond on the complete SNARE complex

We have chosen disulfide bond 1-4 (between cysteines C85 and C92), which showed the largest shortening of the linker, for the subsequent extended MD simulations of the whole complex. To this end, the sulfur atoms of these cysteines were pulled towards each other using a constraint during equilibration dynamics from an initial separation of 15 Å to 2.7 Å, then the disulfide bond was established. To investigate the effects of the shortening of the linker due to the disulfide bond between C85 and C92 on the structure of the SNARE complex, five 140 ns simulations were started from the model of the oxidized state. For comparison with the reduced state, five 180 ns simulations without disulfide bond were carried out. Figures 4.4B, C show a ribbon representation of the "oxidized" and "reduced" starting structures, respectively. If our hypothesis is correct, one would expect that the SNARE complex starts to unfold in the "oxidized" simulations, while remaining stable in the "reduced" simulations. To further check if conformational changes seen during the first 100 ns of the "reduced" simulations are reversible on short time scales, structures from each trajectory of the oxidized state at 100 ns were recorded, from which the disulfide bond was removed, and new simulations ("removed disulfide") of 100 ns were started.

Figure 4.5 shows the mean RMSD of each residue for each of the three setups with respect to the modeled structure, calculated from the final 40 ns of each trajectory. As can be seen, the residues which compose helices in the crystal structure ("core" helices; dark gray bars) show small RMSD values for all three proteins and for all setups. In contrast, the residues close to the termini exhibit larger RMSD values and also larger variances, reflecting their increased flexibility. The largest RMSD values are observed for the modeled SNAP-25B residues, including the newly established ends of the helices (light gray bar) and the linker.

Comparing the "oxidized" and "reduced" states, it can be seen that in the region of the two cysteines C85 and C92, the RMSD values are smaller for the "oxidized" than for the "reduced" simulations. In contrast, the RMSD values of the residues 106–124 and 126–132 are significantly larger for the "oxidized" compared to the "reduced" setup. Remarkably, these differences occur quite distant (between 4.8 and 11.2 nm in the model) from the cysteines C85 and C92. Comparing "removed disulfide" simulations with those of the two other setups, it can be seen that the RMSD values for the residues in the central part of the linker (residues 98–127) are more similar to those of the "oxidized" than to those of the "reduced" simulations. In contrast, the RMSD values for residues 127–140 fall between the values for the "oxidized"
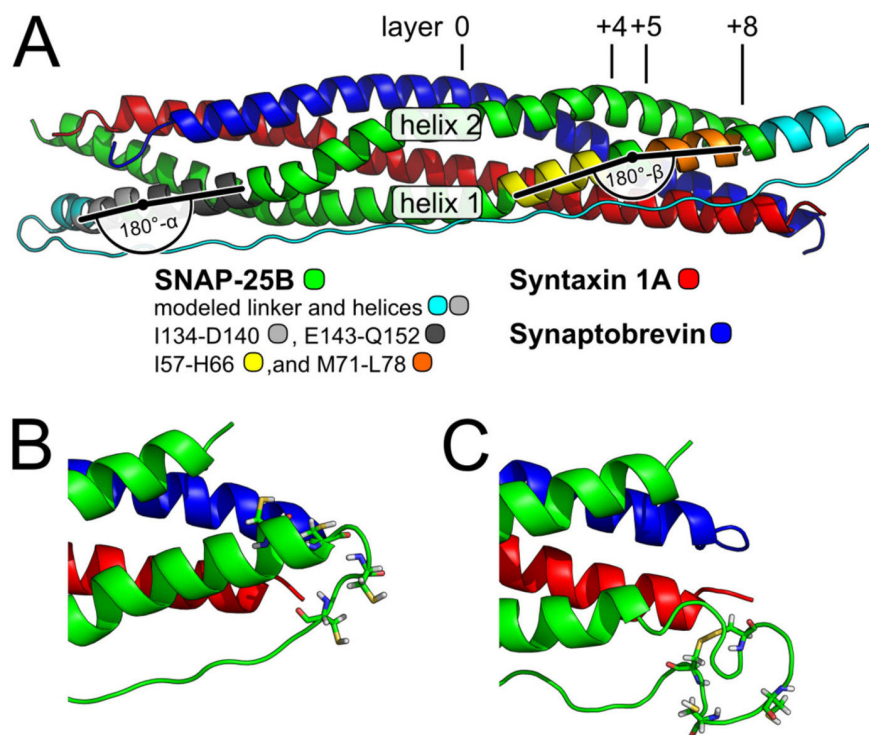
Figure 4.4: **Model of the SNARE complex including SNAP-25B linker.** Ribbon representation of the model of the SNARE complex. The complete model (**A**), close up on cysteines for the model (**B**) and for the established disulfide bond (**C**), respectively. The residues of SNAP-25B which stem from the crystal structure are drawn in green, dark gray, orange, and yellow, the modeled residues in cyan and light gray. Syntaxin 1A is drawn in red, and Synaptobrevin in blue. Regions of the SNAP-25B helices which define the angles α and β are colored in light and dark gray and in yellow and orange, respectively. Location of the helix-helix interaction layers 0, +4, +5, and +8 are shown in part **A**.
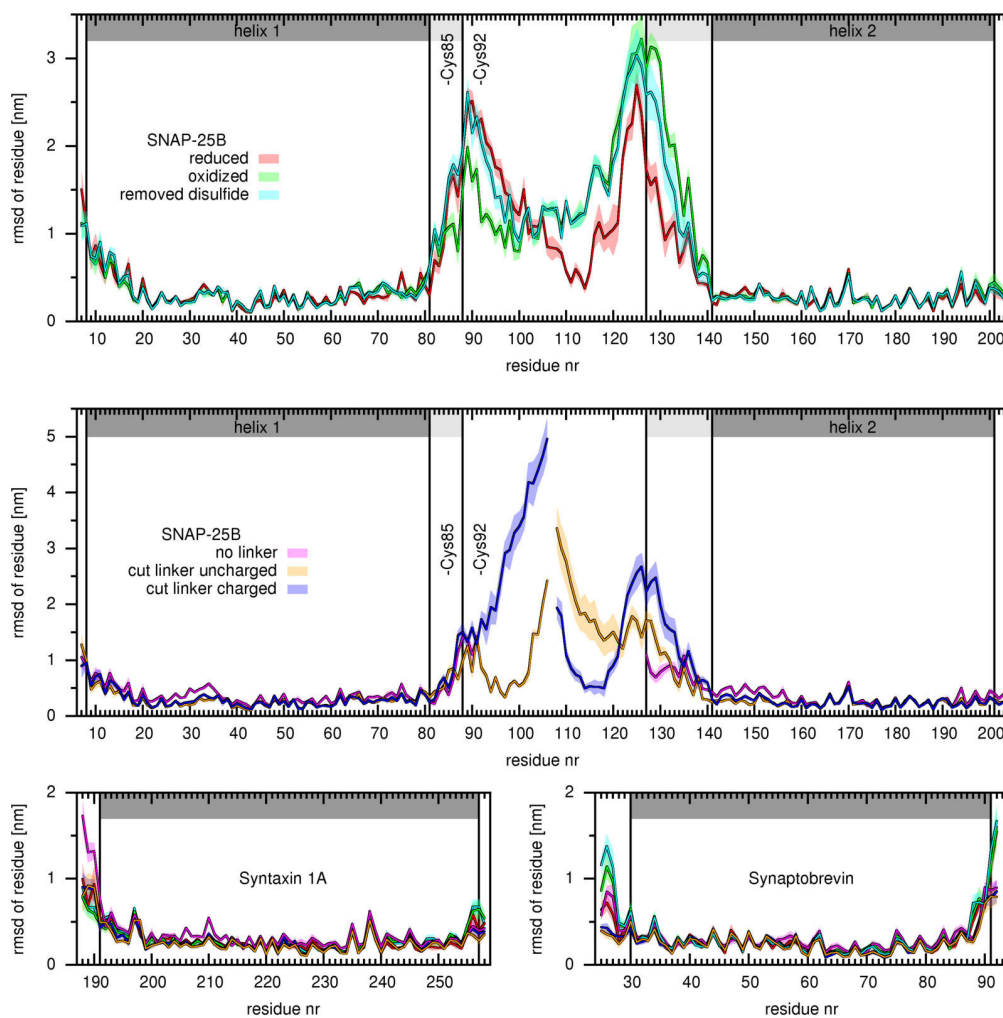
Figure 4.5: **Structural deviation for the different setups.** Shown are mean RMSD values (bold lines) and standard errors (shaded area) of each residue , determined from all 5 simulations of each setup (colors) and averaged over the final 40 ns. The helical part of the crystal structure and the modeled helix ends are indicated with gray and light gray background, respectively.
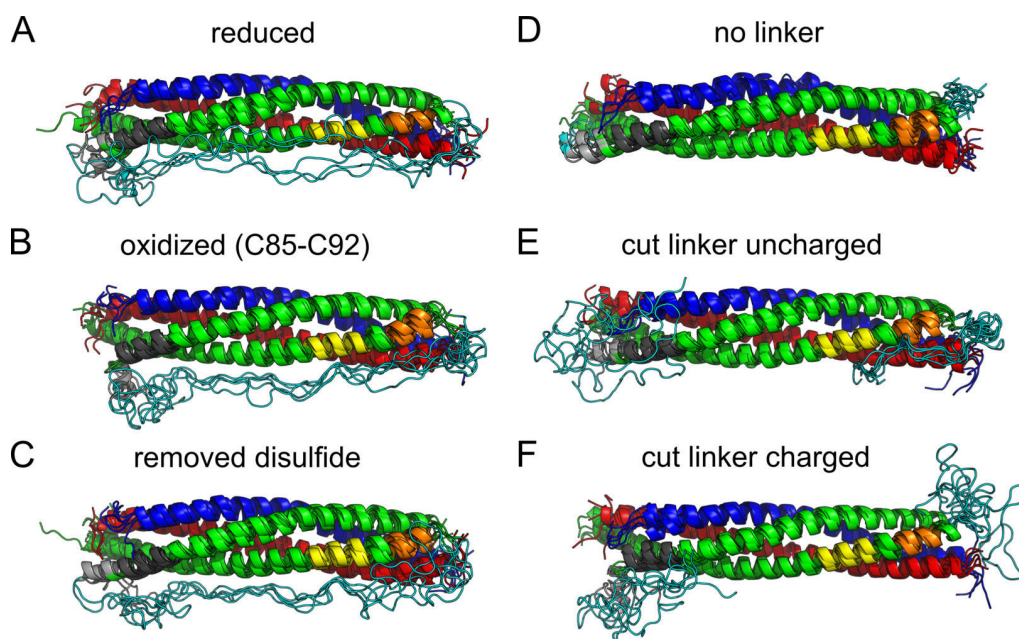
Figure 4.6: **Structural comparison of SNARE complex conformations.** Ribbon representation of the final frames of the simulations. For each setup the conformations at the end of the five trajectories are overlaid. Colors as in Fig. 4.4.

and "reduced" simulations.

Overall, the RMSD values reflect a marked influence of the disulfide bond on the linker conformation and in particular on the conformation of the newly established helix ends. In contrast, no influence on the "core" helices is seen. Further, the RMSD values calculated for the "removed disulfide" simulations show that parts of the structure start to approach the conformation of the "reduced" simulations.

Figures 4.6A, B, and C provide an overview of the structures which give rise to the observed RMSD values. Shown are the five final SNARE complex conformations obtained from the five trajectories of the "reduced", "oxidized", and "removed disulfide" simulations, respectively. As can be seen, the most pronounced difference between the conformations of the "oxidized" and "reduced" simulations is a kink of the partially unfolded extended helix end of helix 2, residues I134–D140 (light gray), Fig. 4.6B, which is not seen in Fig. 4.6A. Notably, the same part of helix 2 is straight in the final frames of two of the "removed disulfide" simulations (Fig. 4.6C). Also the C-terminal end of helix 1 becomes bent for three of five "oxidized" structures compared to the conformations obtained for the "reduced" structures.

37
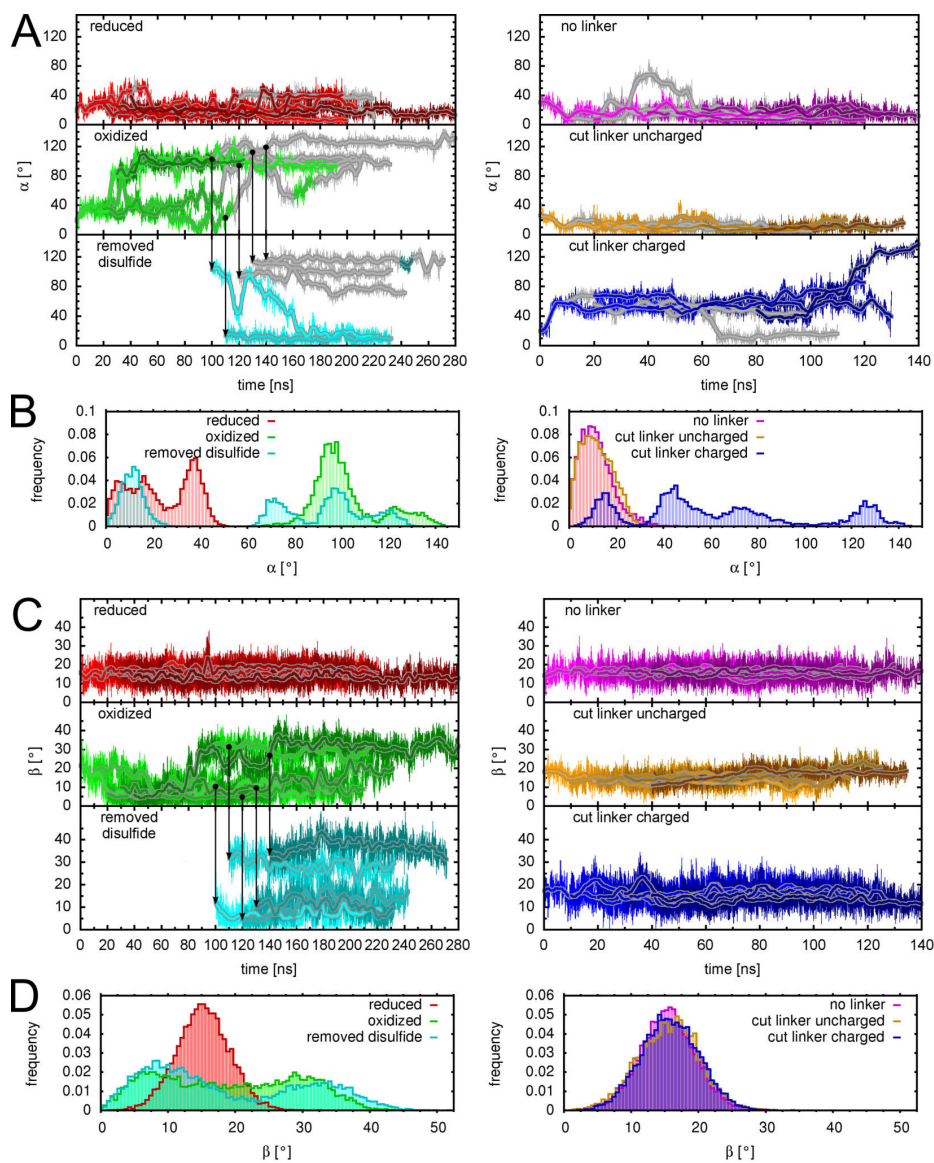
Figure 4.7: **Kinking of SNAP-25B helices.** Angle α between parts of helix 2, I134–D140 and E143–Q152, as a function of time (**A**), where gray lines denote parts of the trajectories in which the fraction of residues from I134–D140 in an α-helical secondary structure is below 0.5. Histogram of the angle α from the final 40 ns of each simulation (**B**). Angle β between parts of helix 1, L57–H66 and M71–L78, as a function of simulation time (**C**) and a histogram of the angle β from the final 40 ns of each simulation. The vertical black lines denote the points in time where the disulfide bond was removed from snapshots of the "oxidized" simulations and the "removed disulfide" simulations were started.

To analyze these drastic structural changes in more detail, Fig. 4.7A shows the angle α between two parts of helix 2, I134–D140 and E143–Q152 (colored light and dark gray in Fig. 4.4), as a function of time. Figure 4.7B shows a histogram of this angle α. As can be seen, the angle α fluctuates around a mean value of 30° for all ”reduced“ simulations, whereas it increases to over 80° for the five ”oxidized“ simulations. Remarkably, in one of the ”removed disulfide“ simulations α reverts to less than 30° within 60 ns. Additionally, in the ”removed disulfide“ simulation which started with a low value of α, the helix stays stretched during the simulation time. The three other trajectories apparently did not yet revert to low α values within the 100–130 ns simulation time, although one of them slightly approaches the stretched conformation. This result shows that reverting pathways exist.

Closer structural inspection of the trajectories reveals that, overall, the α-helical content of the helix extension decreases during the simulations compared to α-helical content of the initial model. Those parts of the trajectories, for which the fraction of the residues I134–D140 which are in an α-helical secondary structure is below 0.5, are shown as gray lines. In the ”reduced“ simulations, the trajectories alternate between states where the fraction is either smaller or larger than 0.5. In contrast, the fraction is lower than 0.5 at the end of 4 of 5 ”oxidized“ simulations. In two cases, the helix first kinked and then unfolded, while in two other cases the helix first unfolded and then α increased. This suggests that the equilibrium conformation of these residues is much closer to random coil in the ”oxidized“ state than it is for the ”reduced“ state.

In the ”removed disulfide“ simulation in which the helix was kinked but not unfolded, the angle decreased. Further, in the simulation which started with unkinked and helical residues I134–D140, these residues remained helical and the angle stayed low. In contrast, only slightly decreasing angles are observed for the simulations which started with a large angle and unfolded residues I134–D140. A short refolding of the extended helix is only observed in one of these cases.

The kinking observed in the C-terminal part of helix 1, which is at the opposite side of the linker, was also investigated by calculating the angle β of two parts of the helix, L57–H66 and M71–L78 (colored yellow and orange in Fig. 4.4). Figure 4.7C shows β during the course of the simulation, Fig. 4.7D shows a histogram of β. The angle β calculated from the ”reduced“ trajectories fluctuates around a mean value of 15°; however, in the ”oxidized“ simulations β increases to values fluctuating around 30° in two trajectories. Further, in the ”removed disulfide“ simulations which start with a large value of β, the angle does not decrease to the values of the ”reduced“ state.

To examine if these drastic structural differences between oxidized and
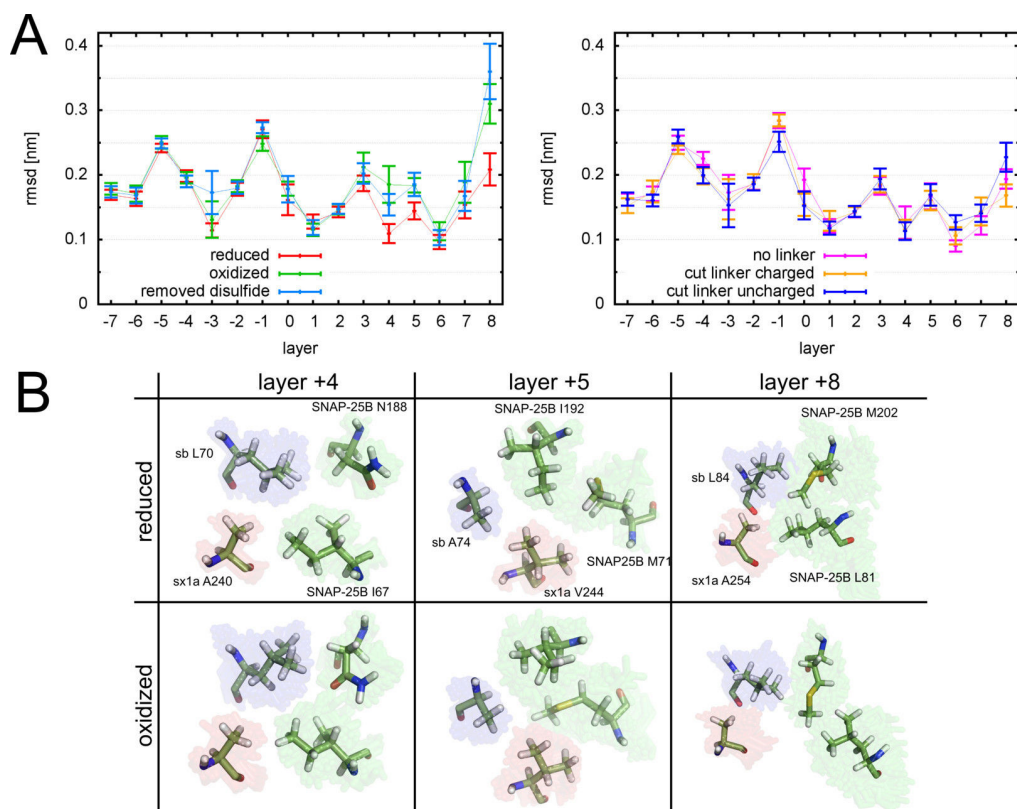
Figure 4.8: **Conformations of central hydrophobic layers.** (**A**) Structural deviations of the residues which contribute to the hydrophobic layers for the six simulations discussed in the text. Shown are the means and the standard errors for the RMSD values calculated from the final 40 ns of each simulation. The RMSD of each layer was calculated after fitting the positions of the Cα-atoms to those of the modeled structure. (**B**) For "reduced" and "oxidized" setups SNAP-25B, syntaxin 1A (sx1a), and synaptobrevin (sb) residues contributing to layers +4, +5, and +8 are shown in sticks representation. The solid sticks represent one conformation at the end of a trajectory; the transparent region depicts an ensemble from the final 40 ns of all trajectories.

reduced SNAP-25B also influence the central hydrophobic layers which are crucial for the association of the four SNARE helices [86], the RMSD for the residues of each layer was calculated with respect to the model structure after fitting the Cα atom coordinates to those of the model.

The 16 interacting layers are numbered −7 through +8. Figure 4.8A shows the mean RMSD values for each layer and each setup, averaged over the final 40 ns of the respective trajectories. The observed standard deviation was used to estimate the error of the mean (bars). As can be seen, the RMSD values of the layers are small compared, e.g., to the linker region, but also compared to the modeled helices (see Fig. 4.5) which implies that the layers are quite rigid. Further, the RMSD values for the different setups are similar for most of the layers. However, marked differences in RMSD values between the "oxidized" and "reduced" states are seen in layers +4, +5, and +8, where the RMSD values for the "oxidized" state are larger.

In particular, the largest difference between RMSD values for these two states is observed for layer +8, where the L81 of SNAP-25B, which contributes to this layer, is only four residues away from the C85 for which the disulfide bond was formed. Remarkably, the conformational differences expand further into the SNARE complex (layer +4 and +5). The RMSD values of the "removed disulfide" setup are very close to those of the setup with disulfide bond, thus the conformations of the layers did not change drastically after the removal of the disulfide bond during the relatively short simulations.

Figure 4.8B shows the conformations of the residues contributing to layers +4, +5, and +8 during the final 40 ns of each trajectory of the "reduced" and "oxidized" state, respectively. As can be seen, the residues of layer +4 are markedly more flexible in the oxidized state. In layer +5, the residues also show more conformational diversity in the "oxidized" ensemble as compared to the "reduced" ensemble. In particular, the methyl group of M71 also occupies the center of the layer in the "oxidized" simulations. The residues of layer +8 show drastic conformational differences between the "reduced" and "oxidized" simulations. L81 of SNAP-25B moves out of the center of the layer and looses contact to the residues L84 of synaptobrevin and A254 of syntaxin 1A. M202 of SNAP-25B also moves out of the center, maintaining contact with L81. In summary, oxidation of cysteines C85 and C92 induces significant rearrangements of many residues within these layers.

4.3.3   Role of the SNAP-25B linker

To further investigate the role of the SNAP-25B linker, simulations with the linker (N93–Q126) removed ("no linker") and with the linker cut between

residues N107 and N108 were performed. Comparison of these simulations with the previous simulations including the complete reduced linker should in particular enable us to distinguish between effects caused by the mechanical strain of the linker and those effects caused by the non-bonded interactions between the linker and the complex.

Cutting the linker of SNAP-25B between N107 and N108 removes the mechanical stress exerted by the linker, while retaining most non-bonded interactions between linker and helices. Two sets of simulations were performed, one set of simulations where the termini of the linker were charged ("cut linker charged") and a second set where the termini were left uncharged ("cut linker uncharged"). The first set mimics SNAP-25B cut chemically by a protease, while the second set is intended to remove the mechanical strain only, without further chemical or electrostatical modifications. Finally, complete removal of the linker removes the strain as well as non-bonded interactions to the complex.

To characterize the structural differences between the three setups, "no linker", "cut linker uncharged", and "cut linker charged", the RMSD of the residues with respect to the initial model of the complex was calculated and is shown in Fig. 4.5. The RMSD values of the residues in the two "cut linker" setups show significant differences in the loop region, which suggests that the charge of the termini strongly influences the conformation of the loop. Similar to the setups with the complete linker, the RMSD values of the "core" helices (dark gray bars) in the simulations with cut linker and without linker are small. Furthermore, the RMSD values calculated for residues 25–35 and 142–156 from the "no linker" trajectories are significantly larger than the respective values of the "cut linker" trajectories.

The obtained "no linker" conformations (Fig. 4.6D) show that the modeled helix end (I134–D140, light gray) is still α-helical and the helix is bent to a lesser extent as compared to the "reduced" setup (see Fig. 4.6A). The cut linker with uncharged termini, Fig. 4.6E, behaves very differently from the cut linker with charged termini, Fig. 4.6F. Whereas, the part of the cut linker carrying the N-terminal interacts with the helices when charged and the modeled helix end (I134–D140) is kinked and unfolds in three out of five simulations, it is very flexible when uncharged and interacting only weakly with the helices. Further, the SNAP-25 helix remains unkinked. In contrast, the C-terminal of the linker is less flexible in the simulations with the uncharged termini. For the uncharged linker, helix 2 of SNAP-25 is bent to a lesser extent as compared to the structures of the "reduced" setup, implying that the mechanical stress caused by the linker suffices to bend the helix.

The kink angle α of helix 2 (Figs. 4.7A and B) calculated from the "cut linker uncharged" and "no linker" simulations remains at values around 10°

during the simulation, whereas the five "cut linker charged" simulations show very diverse behavior.

These results show how sensitively the SNARE complex reacts to small changes despite its thermodynamic stability. In particular, as the "cut linker" simulations show, the outcome of an experiment which includes cutting of the linker must be expected to strongly depend on details of the chosen experimental setup.

The bending angle of helix 1 β for the "no linker", "cut linker uncharged", and "cut linker charged" simulations (Fig. 4.7C and D) is scattered around a mean value of 15°, similar to the distribution calculated from the "reduced" trajectories, suggesting that the influence of the reduced linker on the bending of helix 1 is small.

The RMSD values of the hydrophobic layers (Fig. 4.8A) from the three setups without linker or with a cut linker are similar to the RMSD values calculated for the "reduced" setup, except for layers -4, -3, and $+6$. In layer -4 and -3 the "no linker" simulations show a larger RMSD value than for the "reduced" simulations. Additionally, layer -3 for the "cut linker charged" simulation and layer $+6$ for the "cut linker uncharged" simulations shows slightly larger RMSD values as compared to the "reduced" simulations. This suggests that the different states, cut and removed linker, lead to conformational changes in some of the layers but the influence is not as pronounced as it is in the oxidized state.

4.3.4 Helicity of SNAP-25B

For the residues of the SNAP-25B helix "cores", (except for the first and last two residues), the fraction of residues contributing to an α-helix was calculated for each time step of the trajectories. Figure 4.9A shows a histogram of these fractions collected from the final 40 ns of each trajectory. As can be seen, the helical content for the "oxidized" state is markedly smaller than for the "reduced" state. The distribution for the "removed disulfide" simulations is comparable to that for the "reduced" simulations. Further, the simulations with a cut linker show large frequencies for large helical contents, whereas the distribution is flatter for the simulations with the removed linker.

These results suggest that presence of the disulfide bond causes a partial unfolding of the helices. After its removal the helices do not refold during the relatively short simulation times. The fact that the distribution is shifted to larger helical content for the cut linker simulations as compared to the "reduced" setup suggests that it is the mechanical stress induced by the linker and not interactions with the complex, which leads to the destabilization of the secondary structure. However, the fact that complete removal of the
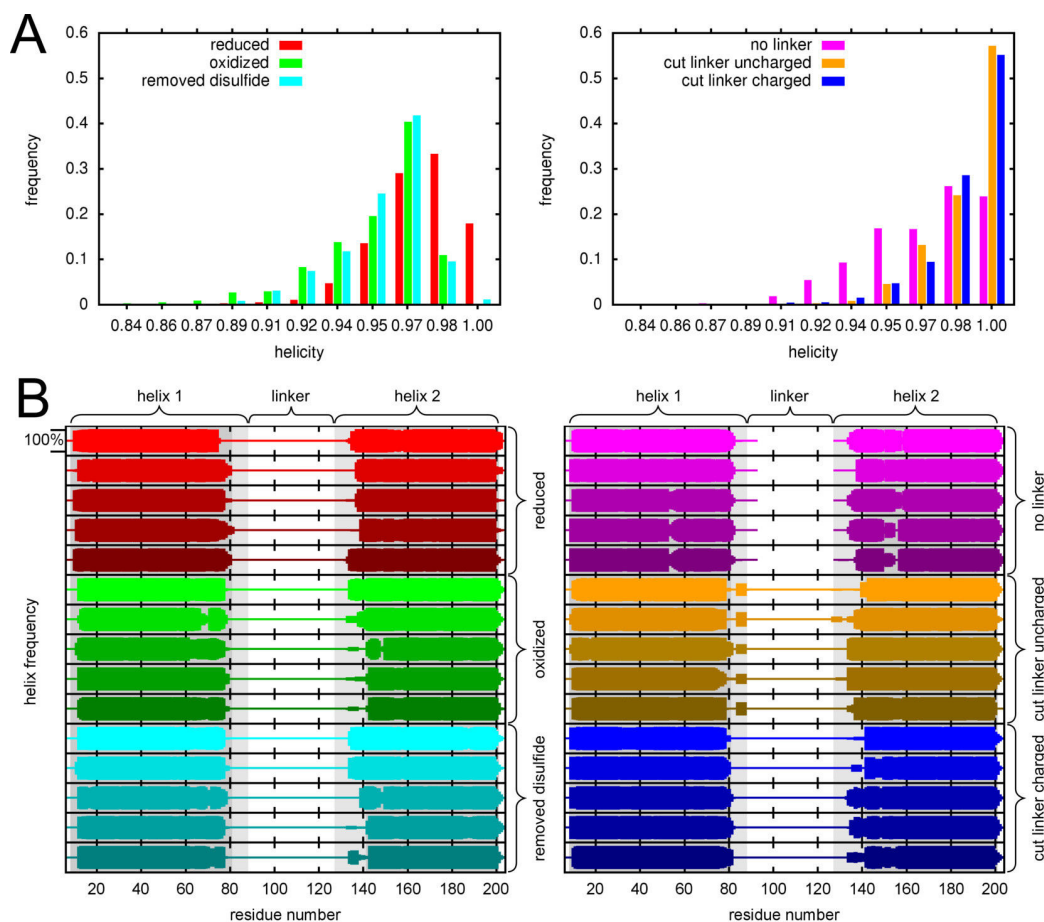
Figure 4.9: **Helicity of SNAP-25B.** (**A**) Remaining helix content of the SNAP-25B helix regions; from the residues which are helical in the crystal structure (except for the first and the last two residues), the fraction of residues which are in an α-helical secondary structure was calculated for each frame of the final 40 ns of the respective set of simulations. Shown is the histogram of these fractions for the six different setups, considered. (**B**) The frequency of each SNAP-25B residue of being in an α-helical structure calculated from the final 40 ns of each trajectory.

linker further enhances the frequencies of states with lower helical content suggests that the interactions of the cut linker with the helices stabilize the secondary structure of the helices.

To locate the origin of the differences in the overall helicity of SNAP-25B, Fig. 4.9B analyzes which SNAP-25B residues contribute to the loss of helical content. Shown is the frequency of α-helical secondary structure, calculated for each SNAP-25B residue from the final 40 ns of each trajectory. For the "no linker" simulations helicity is lower in three out of five cases around residues 53 and 155. None of the other data sets shows a loss of helicity at these residues, suggesting that the linker interacts specifically at these locations.

4.4 Conclusions

Protein	Amino Acid Sequence																	
				84	85			88		90		92						
SNAP-25B	D	L	G	K	F	C	G	L	C	V	C	P	C	N	K	L	K	S
SNAP-25A	D	L	G	K	C	C	G	L	F	I	C	P	C	N	K	L	K	S
SNAP-23A	E	L	N	K	C	C	G	L	C	V	C	P	C	N	R	T	K	N
SNAP-23B	E	L	N	K	C	C	G	L	C	V	C	P	C	N	S	I	T	N

Table 4.2: **Amino acid sequence of the cysteine-rich regions of several SNAP-25B homologs.** The numbering corresponds to both SNAP-25s, for SNAP-23, the first cysteine is at 79.

Molecular dynamics (MD) simulations were used to examine how the formation of a disulfide bond in the SNAP-25B linker region affects the conformation of the SNARE complex. The SNARE complex is a coiled-coil of proteins critical for membrane fusion, and in neurons is composed of three proteins, syntaxin 1A, synaptobrevin, and SNAP-25B. In SNAP-25B, there are four cysteines within a region of eight amino acids of the linker. Formation of a single disulfide bond (oxidation) between C85 and C92 shortens the linker by 2.0 nm (Fig. 4.3).

The linker was added to the crystal structure of the SNARE complex and the complete SNARE complex with both reduced and oxidized SNAP-25B linker was simulated. Indeed, alterations in conformational changes after oxidation were seen, as compared to both the "reduced" state as well as the oxidized and subsequently reduced ("removed disulfide") state. Not surprisingly, formation of the disulfide bond was found to reduce the flexibility of the linker in the region of the disulfide bond (residues 85-96), but unexpectedly

it induces conformational changes at the far end of the linker (residues 126–132) which is 10 nm apart. This shows that local chemical changes can lead to large-scale effects in macromolecular complexes. Additionally, removal of the linker was seen to destabilize the SNARE complex, testified by an RMSD increase in several regions of helix 1 and 2 for the "no linker" setup (e.g., close to residue 155 of Fig. 4.5) and by a decrease in helix frequency in some of these regions as compared to the "reduced" setup (Fig. S1B). However, when the individual residues at each layer of the complex were examined, as in Fig. 4.8A, there is no marked difference between "reduced" and "no linker" at the corresponding layers -7 and -6, marked differences are found nearby at layers -4 and -3. Since the data in Fig. 4.8A were obtained by fitting the coordinates of each layer separately, the increase in RMSD seen in Fig. 4.5 reflects a general bending or rotation of the whole complex after removal of the linker. Although our simulations, taken alone, do not establish functional significance of the observed bends of the SNARE complex upon addition of the linker, it is known that removal of the linker has profound effects in vivo [64, 87].

Closer examination of both ends of the linker showed a strong bend at the end of helix 2 as well as, somewhat weaker, within helix 1. These two bends, characterized by angles α and β respectively (Fig. 4.4A), show significantly different angles after formation of the disulfide bond (Fig. 4.7). The β kink is particularly interesting because it resides within helix 1 and forms part of the SNARE complex.

With respect to oxidation of SNAP25B, it is interesting to note that a key difference between the two SNAP-25 homologs is the relocation of the third cysteine of SNAP-25B at C88 upstream to C84 (Table 4.2) in SNAP-25A. The significance of this change in the two isoforms is not known, but we would predict that such a change would make SNAP-25A more sensitive than SNAP-25B to oxidative stress and iron exposure. This prediction is based on three different possible molecular mechanisms any of which would lower stability of the SNARE complex. (i) The formation of a C84-C92 disulfide bond (in SNAP-25A) would shorten the linker by one more amino acid ($\sim$2.4 nm instead of 2.0 nm) and thus would more strongly prevent complete zippering of the SNARE complex. (ii) The formation of a disulfide at C84 would more likely unwind the C-terminal of helix 1 leading to greater disruption of layer $+8$ and/or larger changes in angle β. (iii) The new cysteine configuration in SNAP-25A allows it to bind iron [70]; such binding would also shorten the linker suppressing complex formation.

If the two isoforms have similar sensitivity to oxidative stress, a reasonable alternate scenario might be that the alternate splicing endows SNAP-25A with a sensitivity to iron. Thus, the SNAP-25A linker would make the

complex sensitive to both oxidation and iron (or other heavy metal) exposure allowing the synapse to shut down under either environmentally harsh condition. Such a response could have developmental consequences, since SNAP-25A is expressed early in development [61].

An additional possibility for the alternate splicing between SNAP-25A and SNAP-25B is that the two different sequences may alter the effectiveness of cellular enzymes to modify each protein. For example, the extent of palmitoylation of each cysteine in each isoform will be a function of how effective of a target it is for the relevant palmitoylation (palmitoyl transferases). The stability of palmitoylation will depend on how good of a substrate it is for any de-palmitoylation (palmitoyl thioesterases) enzyme (for review, see reference [88]). Thus, rearrangement of cysteines could change enzyme binding and thereby explain the apparent inconsistencies reported on the stability of palmitoylation of SNAP-25. Specifically, Lane and Liu [87] reported that in PC12 cells, palmitoylation of SNAP-25 (primarily SNAP-25A) rapidly turned over with a half-life of 3 hours whereas Kang and associates [89, 90] observed no turnover in primary neuronal cultures (primarily SNAP-25B).

A report on SNAP-23 [91] is consistent with the interpretation that cellular palmitoylation of SNAP-25 is cysteine dependent and not simply a non-specific modification of all cysteines. SNAP-23 is a cellular homolog of the neuronal SNAP-25, and has yet another arrangement of cysteines in its linker (Table 4.2). Vogel and Roche [91] report that in transiently transfected HeLa (Human cervical carcinoma) cells, the extent of SNAP-23 palmitoylation was only 4% as much as SNAP-25. If the palmitoyl transferase non-specifically palmitoylated every cysteine in each linker, one would expect SNAP-23 to be palmitoylated by 25% more than SNAP-25, since it has five cysteines, compared to SNAP-25s four cysteines. Instead, the change in sequence drastically reduced palmitoylation of SNAP-23 (or enhanced de-palmitoylation). These data support a model consistent with Fig. 4.2 in which only certain cysteines are specifically palmitoylated/de-palmitoylated, potentially leaving other cysteines as targets for oxidation.

Unpalmitoylated cysteines, although susceptible to oxidation, would normally not be oxidized due to the reducing environment maintained in healthy cells, however, during times of high energy output, reactive oxygen species (ROS) produced by the mitochondria could oxidize newly synthesized SNAP-25 [74, 76, 77]. Another possibility for cellular oxidation of SNAP-25 is by external ROS that cross the cell membrane. Since SNAP-25 binds membranes (even without palmitoylation), it would be in position to be oxidized by such entering ROS. This would also be true after partial palmitoylation of SNAP-25 by a membrane-bound palmitoyl transferase, because the palmitoylation of even one cysteine would bury the nearby free cysteines deeper

in the membrane where they become isolated from the reducing environment of the cell and exposed to external oxidative stress. Finally, it is possible that a disulfide bond-containing SNAP-25 is a better target for palmitoylation by the relevant palmitoyl transferase and that it is only after oxidation, palmitoylation, and reduction that the SNAP-25 becomes available for its normal role in exocytosis. All these scenarios are of course speculative, but experimentally testable.

In our simulations, formation of the disulfide bond also altered helix-helix interactions at layers +4, +5, and +8. This change was noted as an increase in the RMSD of residues in these layers (Fig. 4.8A) and a shift in the averaged location of the residues in helix 1 of SNAP-25 (Fig. 4.8B). Sørensen et al. [86] investigated the effects which mutations in the hydrophobic layers have on exocytosis. Mutations in the layer closest to the cysteines (+8) led to a slowdown of secretion. This effect was more severe with an additional mutation in layer +7. Further, mutations in layer +4 and +5 also led to a drastically decreased secretion. Therefore, our observations of changes in layers +4, +5, and +8 in the oxidized SNAP-25B are consistent with the hypothesis that an oxidation-induced conformational change may be a chemomechanical regulator of SNARE complex formation, leading to down regulation of neurotransmitter release during oxidative stress.

The presented possible mechanism of neurotransmitter release regulation represents a way of how small changes can lead to impairing the function of a macromolecular complex. Detailed structural information along with the time evolution calculated by MD simulations allow to investigate this mechanism at atomic level. In this case, the introduction of a disulfide bond was found to alter the conformation of one component, thereby changing the conformation of the whole complex and possibly leads to a dissociation, rendering the complex dysfunctional.

Chapter 5

tRNA Translocation through the Ribosome[1]

5.1 Introduction

To assemble proteins according to messenger RNA (mRNA) blueprints, the four-letter code of the mRNA nucleotides has to be translated to the 20-letter code of amino acids. In 1956, Francis Crick first proposed "adaptor molecules" which specifically combine an amino acid with the nucleotide template [92]. Transfer RNA (tRNA) molecules were found to act as this "adaptor molecule": For each amino acid, at least one tRNA species exists which specifically binds the amino acid and which contains a specific nucleotide triplet, the anticodon region. The anticodon nucleotides can form base pairs with the three nucleotides of one codon on the mRNA molecule, thus translating the sequence of mRNA nucleotides into a sequence of amino acids, the primary structure of proteins. The first tRNA crystal structure was solved by X-ray crystallography in 1973 [93].

  In 1974, Christian de Duve, Albert Claude, and George Emil Palade received the Nobel Prize in Physiology or Medicine for the discovery of the ribosomes, large protein and RNA complexes responsible for protein synthesis, which consist of two subunits. In prokaryotes, the large (50S) subunit consists of a 23S and a 5S ribosomal RNA (rRNA) strand and 31 proteins, whereas, the small (30S) subunit is composed of a 16S rRNA strand and of 21 proteins [12] (Fig. 5.1A). The ribosome catalyzes the peptide bond formation between subsequent amino acids by positioning the substrates [3] and it

---

[1]Lars V Bock, Christian Blau, Gunnar F Schröder, Niels Fischer, Holger Stark, Marina Rodnina, Andrea C Vaiana, Helmut Grubmüller. Energy barriers and driving forces of tRNA translocation through the ribosome. Submitted.
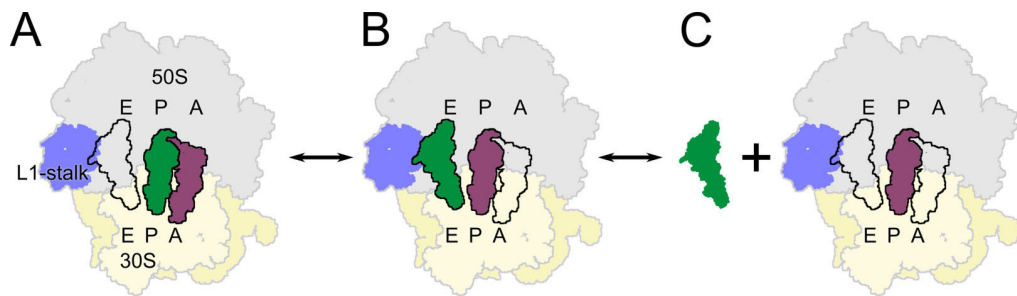
Figure 5.1: **Illustration of tRNA translocation.** (**A**) Pre-translocation state: tRNAs (green, magenta) occupy the A and P sites on the large (50S) and small (30S) subunit of the ribosome. The L1-stalk, composed of the L1 protein and helices 76-78 of the large subunit 23S ribosomal RNA, is depicted in blue. (**B**) Post-translocation state: tRNAs occupy P and E sites. (**C**) After exit of the E-site tRNA, only the tRNA in the P site remains.

ensures high fidelity of peptide chain elongation [94] which is crucial for cell viability. During elongation, the tRNAs bind to three major binding sites in the ribosome: the aminoacyl (A) site, the peptidyl (P) site, and the exit (E) site (Fig. 5.1A). At the beginning of an elongation cycle, the tRNA bound to the peptide chain (peptidyl-tRNA) occupies the P site, forming a base pair to the corresponding mRNA codon. In prokaryotes, elongation factor Tu (EF-Tu) mediates the entry of aminocacyl-tRNAs into the A site. EF-Tu functions by binding an aminoacyl-tRNA and a guanosine triphosphate (GTP) [95, 96]. This complex enters the ribosome and if the tRNA anticodon matches the mRNA codon presented in the A site, EF-Tu hydrolyzes GTP and dissociates from the ribosome, allowing the tRNA to fully accommodate into the A site [97]. Then, in the peptidyl transferase reaction, the α-amino group of the A-site tRNA attacks the carbonyl carbon of the peptidyl-tRNA in the P site [3]. The products of this reaction are a new peptidyl-tRNA, with a peptide that is one amino acid longer, in the A site and a deacylated tRNA in the P site.

After peptide bond formation, the two tRNAs bound to the A and P sites on the small (30S) and large (50S) ribosomal subunits translocate by more than 7 nm to the P and E sites, respectively, while moving the next mRNA codon into the A site (Fig. 5.1A,B). During translocation, tRNAs move first on the 50S subunit into the hybrid A/P and P/E positions [36, 98] with a concomitant large-scale rotation of the 30S relative to the 50S subunit, followed by the displacement of the codon-anticodon complexes on the 30S subunit and the reversal of the subunit rotation, yielding the post-translocation complex [27]. After the exit of the E-site tRNA, the elongation

cycle can start anew (Fig. 5.1C). Translocation of tRNAs is catalyzed by elongation factor G (EF-G) at the cost of hydrolyzing GTP [99, 100]. In the absence of the factor, spontaneous, thermally driven tRNA translocation can occur, and the directionality of uncatalyzed translocation is determined by the chemical specificity of each tRNA to its binding sites on the ribosome [26, 101].

Venkatraman Ramakrishnan, Thomas A. Steitz and Ada E. Yonath were awarded with the Nobel Prize in Chemistry 2009 for resolving structures of the ribosome at atomic resolution and the use of these structures for clarification of ribosomal functions and of the principles of antibiotics attacking ribosomes of bacterial pathogens. To date, now many structures of the ribosome with bound tRNAs and mRNA [12, 14, 38–43, 102] and different factors bound [13, 15, 96, 97, 103–113] have been resolved by X-ray crystallography, capturing the ribosome trapped in different states of translation. Previous simulation studies of the ribosome [114] have addressed tRNA accommodation into the A site [115–117], interactions of the nascent polypeptide chain with the ribosomal exit tunnel [118] and compared interactions of the L1-stalk with different P-site tRNA in a single state [119]. These structures and simulations do not completely describe the dynamics of the ribosome with bound tRNAs. In fact, recent cryo electron microscopy (cryo-EM) work by Fischer et al. [27] has demonstrated that, when tRNAs move through the ribosome, tRNAs and the ribosome sample a large number of conformational states. Fischer et. al resolved 50 different cryo-EM reconstructions representing 50 different states of translation. These states were grouped into 5 pre-translocation states (pre1–pre5) and four post-translocation states (post1–post4). However, precisely how the spontaneous thermal fluctuations of tRNAs and the various interacting parts of the ribosome cooperatively drive the tRNA movement remains unclear.

To aim at a near-atomic description of the mechanism of tRNA translocation, in this work, crystallography and cryo-EM data were combined with molecular dynamics (MD) simulations to analyze the interaction networks between the tRNAs and the ribosome during the movement and to study the dynamic fluctuations of the tRNAs and parts of the ribosome. Further, we estimated the interaction energies between individual components, and identified the forces that drive directed tRNA movements. In particular, two main questions regarding prerequisites for efficient translocation are addressed: First, which components of the ribosome are rate-limiting for transitions between different states and which mechanisms help overcome free energy barriers along the translocation pathway? Second, how does the network of interactions between the subunits adapt during the large-scale intersubunit rotations which accompany translocation? The interaction net-

work needs to balance the affinity between the subunits in the different states, such that the two subunits remain associated and that, at the same time, large free energy barriers for rotation are avoided.

5.2 Methods

In this section, all the methods used for the work presented in this chapter are described. The refinement of ribosome models based on crystal structures against cryo-EM reconstructions was done in close collaboration with Gunnar Schröder. The cryo-EM of analysis of the global ribosome conformation was entirely done by Niels Fischer in the group of Holger Stark.

5.2.1 General molecular dynamics setup

All molecular dynamics (MD) simulations were carried out with GROMACS 4 [81] using the amber99sb force field [46], and the SPC/E water model [120]. Each simulated model was first solvated in a dodecahedron box keeping a minimum distance of 1.5 nm between the model atoms and the box boundaries. K^+ and Cl^- force field parameters were taken from Joung and Cheatham [121]. Long-range electrostatic interactions beyond 0.9 nm were calculated by particle-mesh Ewald summation [84] with a grid spacing of 0.12 nm. Lennard-Jones interactions were calculated within a distance of 1.4 nm. Coordinates were recorded for analysis every 2 ps. Unless stated otherwise, an integration time step of 4 fs was used, applying virtual site constraints [53]. All bond lengths were constrained with the LINCS algorithm [122]. The system temperature was kept constant at $T = 300$ K using velocity rescaling [56] with a coupling time constant of $\tau_T = 0.1$ ps. Protonation states of amino acids were determined with WHATIF [123].

5.2.2 Models of the ribosome including tRNAs

Three initial atomic models of the *Escherichia coli* (*E.coli*) ribosome were built which were subsequently fitted into the cryo-EM maps provided by Fischer et al. [27]: First, a model of the ribosome with a P-site fMetVal-tRNA^{Val} (P-model), second, a model with a P-site fMetVal-tRNA^{Val} and an E-site tRNA^{fMet} (PE-model), and third, a model with an A-site fMetVal-tRNA^{Val} and a P-site tRNA^{fMet} (AP-model).

All models were constructed from the crystal structure of the intact *E.coli* ribosome by Zhang et al. (pdb ids: 3I1P, 3I1O) [14]. This was the best resolved (resolution: 3.19 Å) and most complete *E.coli* ribosome structure at

the time of modeling. All structural information, including the crystallo-graphic water molecules and ions, was used. In the following, I describe the modeling of parts of the L1-stalk and helix 38 of the 23S rRNA which were not resolved in this crystal structure.

Modeling of the L1-stalk

For the L1 protein and the parts of the L1-stalk rRNA which are not resolved in the structure of Zhang et al., the *Thermus thermophilus* (*T.thermophilus*) ribosome structure of Gao et al. [13] (pdb id: 2WRI) was used, where all residues are resolved and the sequence identity is sufficient to build a homology model.

First, a homology model of the L1 protein was built using the swiss-model server (http://swissmodel.expasy.org/) with the *E.coli* sequence and the *T.thermophilus* structure as a template (44% sequence identity).

Second, in the *E.coli* structure, 68 nucleotides (2111–2179) are not re-solved in the L1-stalk rRNA. These were also modeled using the *T.thermophilus* structure. The corresponding structurally aligned nucleotides [13] as well as ten nucleotides upstream and downstream, which form ten base pairs at the stem of the L1-stalk, were extracted from the *T.thermophilus* structure. The extracted nucleotides were mutated to match the *E.coli* sequence using WHATIF [123] (58% sequence identity).

The whole L1-stalk, comprising the mutated rRNA and the homology model of the L1 protein, was energy minimized in vacuum with position restraints on the P and C^α atoms (force constant: 1000 kJ mol^{-1} nm^{-2}). Next, water and K^+Cl^- ions at a concentration of 0.154 mol were added to mimic physiological conditions. This system was energy minimized and subsequently equilibrated for 1 ns, maintaining position restraints. After that, the stem base pairs were superimposed to the matching nucleotides in the *E.coli* structure. Finally, an MD simulation of the L1-stalk was carried out, restraining the positions of heavy atoms of the stem to those of the target *E.coli* structure. The force constant of the restraining potential was linearly increased from 0 to 500 kJ mol^{-1} nm^{-2} within 5 ns. The resulting L1-stalk structure was finally incorporated into the model.

Modeling of parts of helix 38

Helix 38 (H38) of the 23S rRNA of the large ribosomal subunit forms contacts to the small subunit via intersubunit bridge B1a [12]. Nucleotides G879–C897 of this helix are not resolved in the *E.coli* structure. Secondary structure prediction software S2S [124] predicted a pentaloop for the tip of the helix.

The nucleotides of a pentaloop from an NMR structure [125] (pdb id: 1NA2) were mutated to match H38 sequence using WHATIF. To fit this rRNA structural motif into the model of the ribosome, the same protocol as for the L1-stalk rRNA was used.

Modeling of tRNA$^{\text{Val}}$ and mRNA

The ribosomes used for the cryo-EM experiments contained an fMetVal-tRNA$^{\text{Val}}$. Since no high resolution structure was available, a solution structure of tRNA$^{\text{Val}}$ from a refinement of a homology model against residual dipolar coupling and SAXS data (pdb id: 2K4C) [126] was used for the models.

The *T.thermophilus* structure by Yusupova et al. [127] (pdb id: 2HGP) contains a P-site tRNA$^{\text{Phe}}$ and a 50 nucleotide long mRNA. This structure was rigid-body fitted to our model using structurally aligned nucleotides [127] from 16S rRNA of the small subunit. The fitted coordinates of the tRNA phosphates and of the mRNA were stored for later use. From this mRNA structure, the A-, P-, and E-site codons as well as three upstream and downstream nucleotides were extracted. These nucleotides were mutated with WHATIF to match the sequence of the mRNA used in the cryo-EM experiments. Appropriate tRNA modifications and the dipeptide were added to match experimental conditions [27]. Force field atom types for fMet were obtained with ANTECHAMBER [128], partial charges were determined using DFT-b3lyp with a 6-31/G* basis set. Side chain charges are the same as for Met and on the backbone, only the charges of the formylamino cap changed by more than 5 %. The modified fMetVal-tRNA$^{\text{Val}}$ structure was then fitted as a rigid body to the P atoms obtained from the fitted Yusupova structure. Next, a 1 ns simulation of the tRNA and the mRNA in solvent with position restraints on the P and C1' atoms was carried out. In the subsequent 5 ns simulation, positions of P and C1' atoms were restrained to those of the fitted Yusupova structure, thereby linearly increasing the force constant from 0 to 1000 kJ mol^{-1} nm^{-2} during the simulation. The resulting tRNA structure was then included into the ribosome model, yielding the P-model.

Modeling of tRNA$^{\text{fMet}}$

The *T.thermophilus* structure by Gao et al. [13] (pdb id: 2WRI) contains an E-site tRNA$^{\text{fMet}}$. The 30S subunit structure was rigid-body fitted to our model using structurally aligned nucleotides [13] from 16S rRNA. Nucleotide modifications were added to the tRNA to match the modifications of the tRNAs in the cryo-EM experiments and a 1 ns simulation of the tRNA in

solvent with position restraints on the P and C1' atoms was performed. The resulting tRNA structure was then included into the P-model, yielding the PE-model.

For the AP-model, the same tRNA structures were used as for the PE-model, but the tRNA$^{\text{Val}}$ was fitted into the A and the tRNA$^{\text{fMet}}$ into the P site. Almost the same protocol as for the addition of the P-site tRNA$^{\text{Val}}$ was used, except that the tRNA$^{\text{Val}}$ has 77 nucleotides and the tRNA$^{\text{Phe}}$ from the Yusupova structure which was used for fitting has 76 nucleotides. For the rigid-body fitting and the position restraints in the simulation all the nucleotides, except for 5 nucleotides upstream and 5 nucleotides downstream of the insertion were taken into account.

5.2.3 Flexible fitting of the atomic models into cryo-EM maps

High resolution structures of large macromolecular systems, such as the ribosome, are only available for a small set of functional states, most often for only one. Three-dimensional reconstructions obtained from a large number of cryo-electron microscopy (cryo-EM) images allow to determine density maps of many different conformational states [27]. The resolution of the 50 cryo-EM density maps used in this work range from higher than 9 Å to 30 Å [27]. In order to obtain near-atomic models of these conformational states, we have refined an initial model based on an atomistically resolved X-ray structure to fit into the map, using the real-space refinement program DireX [30, 31]. DireX employs CONCOORD [129], a geometrical conformational sampling method, to flexibly fit a model into a density map in an iterative procedure, while maintaining local stereochemistry and preventing steric clashes. A deformable elastic network (DEN) is used to incorporate prior structural knowledge into the refinement process.

The three ribosome models (P-, PE-, and AP-model) described in the previous paragraph were used as starting structures for subsequent structure refinements against the 25 different cryo-EM density maps, to provide a near-atomic interpretation of each individual conformational state. Initial placement of a starting model into a density map was done using the rigid-body fit module of the program Chimera [130]. DireX [30] was then used for all refinements. DireX computes a density map from an atomic model and refines the atomic coordinates to maximize the overlap between this model map and the cryo-EM density map. The model density maps were generated using a Gaussian kernel with a width adapted to the resolution of the corresponding cryo-EM density map.

AP-states		PE-states		P-state	
state of fitted structure	starting model	state of fitted structure	starting model	state of fitted structure	starting model
pre1b	AP-model	post1	PE-model	post4	P-model
pre1a	pre1b	post1b	post1		
pre2	pre1b	post1c	post1		
pre4	pre1b	post1d	post1		
pre4b	pre4	post1e	post1		
pre4c	pre4	post2a	post1		
pre4d	pre4	post2b	post2a		
pre4e	pre4	post3b	post2a		
pre3	pre4	post3a	post3b		
pre5a	pre1b				
pre5c	pre5a				
pre5d	pre5a				
pre5f	pre5a				
pre5e	pre5a				
pre5b	pre5e				

Table 5.1: **Sequence of structure refinements against cryo-EM maps.** For each state the starting model which was used for refinement against the corresponding map is shown.

Harmonic distance restraints were applied between randomly chosen pairs of atoms that are within 3 and 12 Å in the starting model using the DEN feature of DireX. It should, however, be noted that due to the low resolution of the density maps, the DEN restraints were chosen to be non-deformable, i.e. the γ-parameter was set to zero. The number of these restraints was chosen two times the number of atoms. Additional harmonic distance restraints (5562 in total) were used to preserve the hydrogen bonds between paired bases in the RNA. For each refinement, 2000 optimization steps were performed.

The initial AP-, PE-, and P-models were refined against the map which most closely resembled the state of the model (respectively: the pre1b map, the post1 map, and the post4 map). Subsequently, the refined structures were used as starting models for refinement against the remaining 22 maps in the sequence described in table 5.1.

5.2.4   Choice of models for MD simulations

For each of the major states of spontaneous retro-translocation (pre1 to post4), the structures refined against the one or two cryo-EM maps with the highest resolution were used as starting structures for MD simulations: pre1a (12 Å), pre1b (12 Å), pre2 (15 Å), pre3 (17 Å), pre4 (13 Å), pre5a (15 Å), post1 (12 Å), post2a (17 Å), post2b (17 Å), post3b (15 Å), and post4 (9 Å). Two additional structures were chosen for simulations, because of extreme intersubunit rotation angles (pre5b) and an extreme tRNA$^{\text{fMet}}$ conformation (post3a), to capture a large range of conformations accessible by the ribosome and the tRNAs.

5.2.5   MD simulations of the fitted models

Atomic models obtained from the flexible fitting to 13 cryo-EM maps were solvated, and the system was neutralized with K^+ ions before adding explicit salt to mimic the buffer conditions in the experiments [27] (7 mM $MgCl_2$ and 150 mM KCl) using the GENION program from the GROMACS suite [81]. The system was then equilibrated in four steps:

- 0–5 ns: Constant volume and position restraints on all ribosomal heavy atoms with a force constant of 1000 kJ mol^{-1} nm^{-2}, 2 fs time step.

- 5–10 ns: Constant volume and linearly decreasing the position restraints force constant to zero.

- 10–20 ns: The pressure was coupled to a Berendsen barostat [54] with coupling constant $\tau_p = 1$ ps and an isotropic compressibility of $4.5 \cdot 10^{-5}$ bar^{-1}.

- 20–120 ns: The pressure was coupled to a Parrinello-Rahman barostat [83] with coupling constant $\tau_p = 1$ ps and an isotropic compressibility of $4.5 \cdot 10^{-5}$ bar^{-1}.

The same protocol was used for simulations of the initial PE-model and of each tRNA in solution. For the simulations of the tRNAs in solution, the last step was prolonged to 1020 ns. All the production runs used in this study are summarized in Table 5.2.

5.2.6   Cryo-EM analysis of global ribosome conformation

Cryo-EM grids of vacant *E.coli* ribosomes were prepared at 18 °C under controlled environmental conditions [131] and imaged under cryo conditions with a Titan Krios electron microscope (FEI Company, the Netherlands) at 300 kV and defoci ranging from 1 μm to 2.5 μm on an Eagle 4k x 4k CCD camera (FEI Company, the Netherlands) using 2fold pixel binning, resulting in a final pixel size of 3.2 Å. Ribosome particles were selected semi-automically with Boxer [132], an interactive particle selection tool, and corrected locally for the contrast transfer function [133]. The resulting 9814 ribosome particles were coarsened twofold to a pixel size of 6.4 Å and classified according to 30S body rotation in 2.5° steps as described earlier [27]. Pre-translocation state ribosome particles (315108 in total) were obtained from an existing data set of *E.coli* ribosome complexes prepared for cryo-EM at different time-points of the retro-translocation reaction, using the same buffer and temperature of 18 °C as for the vacant ribosomes [26, 27]. Hierarchical classification resulted in 34 groups of ribosome particles representing structurally distinct pre-translocation states [27]. For each population of pre-translocation ribosomes, the 30S body rotation of the corresponding cryo-EM reconstruction was determined. In Fig. 5.14, the fraction of particles as a function of 30S body rotation was plotted in 2.5° steps for all pre-translocation state ribosomes and vacant ribosomes, respectively. Image processing was generally performed using IMAGIC-5 [134] and exhaustive alignment [135].

Nr of trajectories	system	state	trajectory length
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	PE-model	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	pre1a	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	pre1b	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	pre2	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	pre3	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	pre4	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	pre5a	100 ns
2	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	pre5b	100 ns
2	ribosome, mRNA, solvent	pre5b	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	post1	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	post2a	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	post2b	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	post3a	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, solvent	post3b	100 ns
1	ribosome, mRNA, tRNA$^{\text{Val}}$, solvent	post4	100 ns
1	tRNA$^{\text{Val}}$, solvent	-	1 μs
1	tRNA$^{\text{fMet}}$, solvent	-	1 μs

Table 5.2: **Summary of production simulations.** Trajectory length refers to the length of the production run after 20 ns equilibration.

5.3 Analysis

Here, the methods used for analysis of the conformations, motions, and energetics observed in the simulations are presented. The definition of the axes of for 30S head and body rotation and the transition rate estimates were done in close collaboration with Christian Blau.

5.3.1 Comparison to recent crystal structures

To compare our models to existing crystal structures for each state, an average structure was calculated from the last 10 ns of each trajectory. The root mean square deviations (RMSD) of these structures relative to three *E.coli* crystal structures (pdb ids: 3R8S, 3R8T [15], and 3I1R [14]) were calculated after rigid-body fitting using all resolved C^α and P atoms, except L9 protein atoms. L9 atoms were excluded, because in the crystal structures protein L9 forms a strong contact with the 30S subunit of the ribosome in the neighboring cell.

5.3.2 Definition of reaction coordinates for collective motions

Principal component analysis

To obtain reaction coordinates, principal component analyses (PCA) [136] were carried out for tRNA$^{\text{Val}}$, tRNA$^{\text{fMet}}$, and the L1-stalk using all the simulations. To define common sub-spaces, all trajectories were first superimposed by least-squares fit, using C^α and P atoms of the 50S subunit excluding the L1-stalk. Next, the C^α and P atoms of the tRNAs and the L1-stalk were extracted from the trajectories. For each of the three ribosomal parts the extracted trajectories of all states were concatenated, and the atomic displacement covariance matrix was calculated. The trajectories for each state were then projected on the first eigenvector of this matrix. The projections divided by $\sqrt{N}$, where N is the number of atoms used to construct the covariance matrix, yielded the reaction coordinates (r.c.). The minimum and maximum of this reaction coordinate for each state are shown in Fig. 5.5. The projections on the first three eigenvectors were used to estimate transition rates (see below).

In order to calculate the distance dependence of the binding enthalpies between tRNA$^{\text{fMet}}$ and the L1 protein, a distance coordinate was obtained from a PCA using their C^α and P atoms. The atomic displacement covariance matrix was constructed from the trajectories of those states in which the

tRNA and the L1 were in contact (pre3–pre2b). The projection on the first eigenvector of this matrix, divided by $\sqrt{N}$, where N is the number of atoms used to construct the covariance matrix, was used as the distance coordinate.

30S head and body rotation

Head and body rotations were quantified by comparing structures of each state, extracted from the respective trajectories at 200 ps intervals, to the post1a structure at 20 ns. The post1a structure was used to define zero degree body and head rotations. To define the axes of rotation and pivoting points, Christian Blau extended a non-linear least-squares fitting method [137] to also include the axis of minimal and median rotation. Final mean pivoting point and axes of rotation for head and body movement were determined by quaternion-based averaging [138] over all rotations obtained from all structures of all states. Rotation angles were then calculated relative to the mean axes of rotation and pivoting points.

5.3.3   Transition rate estimation

For the collective motions of the two tRNAs, the L1-stalk, 30S head and body rotation, the individual transition rates for all pairs of states were estimated as described in chapter 3.1. As reaction coordinates for the analysis the projections onto the three dominant eigenvectors obtained from PCA were used for the motions of tRNAs and L1-stalk, and the three rotation angles were used for 30S head and body rotations. The projected trajectories were analyzed in 200 ps intervals for rotations and in 10 ps intervals for projections on eigenvectors.

5.3.4   tRNA contacts

For each simulation, the distances between all pairs of atoms, respectively from the tRNAs and the ribosome, were calculated. A residue pair was considered to be in contact with each other if the distance between two atoms (one from each residue) was found to be below 3 Å in at least one frame of the 100 ns trajectory. Binding enthalpies between tRNA residues and contacting residues of L1, L5, or L16 proteins were calculated from potential energy function of the MD force field.

5.3.5 Short time diffusion constants

Translational short time diffusion constants of the tRNAs within the ribosome were calculated for the trajectories for which the tRNAs were bound to the A, P or E site (tRNA$^{\text{Val}}$: pre1–post4, tRNA$^{\text{fMet}}$: pre1–post2). For each state the trajectory was superimposed by least-squares fitting to the structure at 20 ns using C^α and P atoms of the 50S subunit except L1-stalk atoms. From 10 ns intervals of these trajectories the mean square deviation (MSD), $\langle x^2 \rangle$, of the center of mass of the tRNA atoms was calculated. For each interval i, a diffusion constant D_i was obtained from the slope of the MSD using the Einstein relation, $\langle x^2 \rangle = 6 D_i t$. To that aim, a linear function was least-squares fitted to the MSD for 1 ns $\leq t \leq 9$ ns of the respective interval. To estimate the uncertainty of the diffusion constant for each 10 ns interval, the absolute value of the difference ΔD_i between the diffusion constants obtained for 1 ns $\leq t \leq 5$ ns and for 5 ns $\leq t \leq 9$ ns, respectively, was calculated. The weighted mean $\bar{D}$ of the diffusion constants D_i and its error ΔD were calculated as follows,

$$\bar{D} = \frac{\sum_{i=1}^{N} \frac{D_i}{\Delta D_i^2}}{\sum_{i=1}^{N} \frac{1}{\Delta D_i^2}}, \qquad \Delta D = \sqrt{\frac{1}{\sum_{i=1}^{N} \frac{1}{\Delta D_i^2}}}. \tag{5.1}$$

To determine the diffusion constant of the tRNAs in solution, the same protocol, without the initial superimposition, was applied to the respective simulations of solvated tRNAs.

5.3.6 Conformational free energy of bound tRNAs

The tRNAs bound to the ribosome adopt conformations different from those in free solution. To estimate the conformational free energy stored in the tRNAs during translocation, we compared the ensemble of conformations adopted by each tRNA in solution, using the second half of the production trajectory (520-1020 ns), with the ensemble of conformations of the same tRNA bound to the ribosome in each state. From each trajectory, the positions of the tRNA backbone atoms, P, C4' and C3' were extracted. For each simulation of the tRNA in solution (tRNA$^{\text{Val}}$ and tRNA$^{\text{fMet}}$), each frame of the trajectory was least-squares fitted to the first frame and then the average structure was calculated. For each tRNA and each state (tRNA$^{\text{Val}}$: pre1–post4; tRNA$^{\text{fMet}}$: pre1–post3), the trajectory was also fitted on the first frame of the corresponding tRNA in solution trajectory, and subsequently the average structure was calculated. Next, for each tRNA and state, the fitted trajectories of the bound tRNA and of the tRNA in solution were projected

onto the difference vector between the two corresponding average structures. For each of the two projections, the range of the resulting values was divided into 100 equally spaced intervals and for each interval i , $P(x_i)$ the number of values in the interval divided by the total number of values was calculated, where x_i is defined as the center of the interval. The conformational free energy landscape for each tRNA in solution along this difference vector was estimated as

$$\Delta G^{\mathrm{sol}}(x_i) = -k_B T \ln \left(\frac{P^{\mathrm{sol}}(x_i)}{P^{\mathrm{sol}}_{max}} \right), \tag{5.2}$$

where T is the temperature, $P^{\mathrm{sol}}(x_i)$ is $P(x_i)$ for the tRNA in solution and P^{sol}_{max} is the maximum of $P^{\mathrm{sol}}(x_i)$. Since the projections of the two trajectories do not fully overlap, we assumed as the underlying probability distribution $p(x)$ for the tRNA in solution, a normalized sum of two Gaussian functions. This probability distribution was used to estimate the free energy in regions along the difference vector that are not visited by the trajectory of the tRNA in solution.

The function $F(x) = -k_B T \ln (p(x))$ was then least-squares fitted to $\Delta G^{\mathrm{sol}}(x_i)$. Finally, the conformational free energy ΔG_{conf} of the bound tRNA in comparison to the tRNA in solution was estimated by weighting $F(x)$ with the frequency for the bound tRNA $P^{\mathrm{bound}}(x_i)$,

$$\Delta G_{conf} = \sum_{i=1}^{100} P^{\mathrm{bound}}(x_i) \cdot F(x_i). \tag{5.3}$$

This free energy difference estimates the conformational free energy necessary to transfer the tRNA from solvent into the ribosome.

To investigate how many individual residues of the tRNAs contribute to the overall conformational free energy, an analysis similar to the one for the whole tRNA was applied. For each residue the center of geometry was calculated from the fitted trajectories. The trajectories of the center of geometry from the bound and the solution simulations were projected onto the difference vector between the corresponding average centers of geometry. Using these projections, we followed the same protocol as described above and derived an estimate for the conformational free energy of each tRNA residue in each state.

5.3.7 Clustering of intersubunit contacts

Residue-residue contacts between the two ribosomal subunits (50S and 30S) were extracted from the simulations of all the states, and clustered (see chapter 3.3). For all the resulting 52 clusters, force field binding enthalpies be-

tween the 50S and the 30S part of the cluster were calculated in each frame of the trajectory. For further analysis, only the 30 clusters were retained for which, in at least one state, the average binding enthalpy was larger than 1 % of the sum of all the enthalpies of the clusters in the given state.

5.3.8 Level of intersubunit contact change

Intersubunit residue-residue contacts which are only present in 3 or less states were defined as state-specific. The level of intersubunit contact change was defined as the ratio of the number of state-specific residue-residue contacts to the number of all residue-residue contacts for one cluster.

5.4 Results and Discussion

To obtain near-atomic structural models of translocation intermediates, crystal structures of *E.coli* ribosomes [14] were fitted into 25 cryo-EM density maps [27] which reflect the tRNA movement when ribosomes assembled in the post-translocation state, with P-site fMet-tRNA$^{\text{Val}}$ (P/P state) and deacylated E-site tRNA$^{\text{fMet}}$, were allowed to convert spontaneously to their A/A and P/P states, respectively [26] (Fig. 5.2). To study fast motions of the ribosome and to understand the underlying forces that drive tRNA translocation, I carried out all-atom explicit solvent MD simulations of the entire ribosome for 13 major pre- and post-translocational models selected from the 25 structural models (Table 5.2).

5.4.1 Structural deviation during MD simulations

For all 13 simulations of translocation intermediates and for the simulation starting from the PE-model, Fig. 5.3 shows the root mean square deviation (RMSD) relative to the respective starting structure or to the structure at 20 ns. In all simulations an RMSD of less than 8 Å was obtained, which is very low for a system of that size, indicating that the simulation systems are sufficiently stable. Values reported by other authors are ∼10 Å [139]. The PE-model simulation started from a crystal structure and thus can serve as a benchmark for the quality of the cryo-EM-fitted structures [140]. Notably, the RMSD values for the fitted structures are similar or only slightly larger, thus underscoring the quality of these structures.
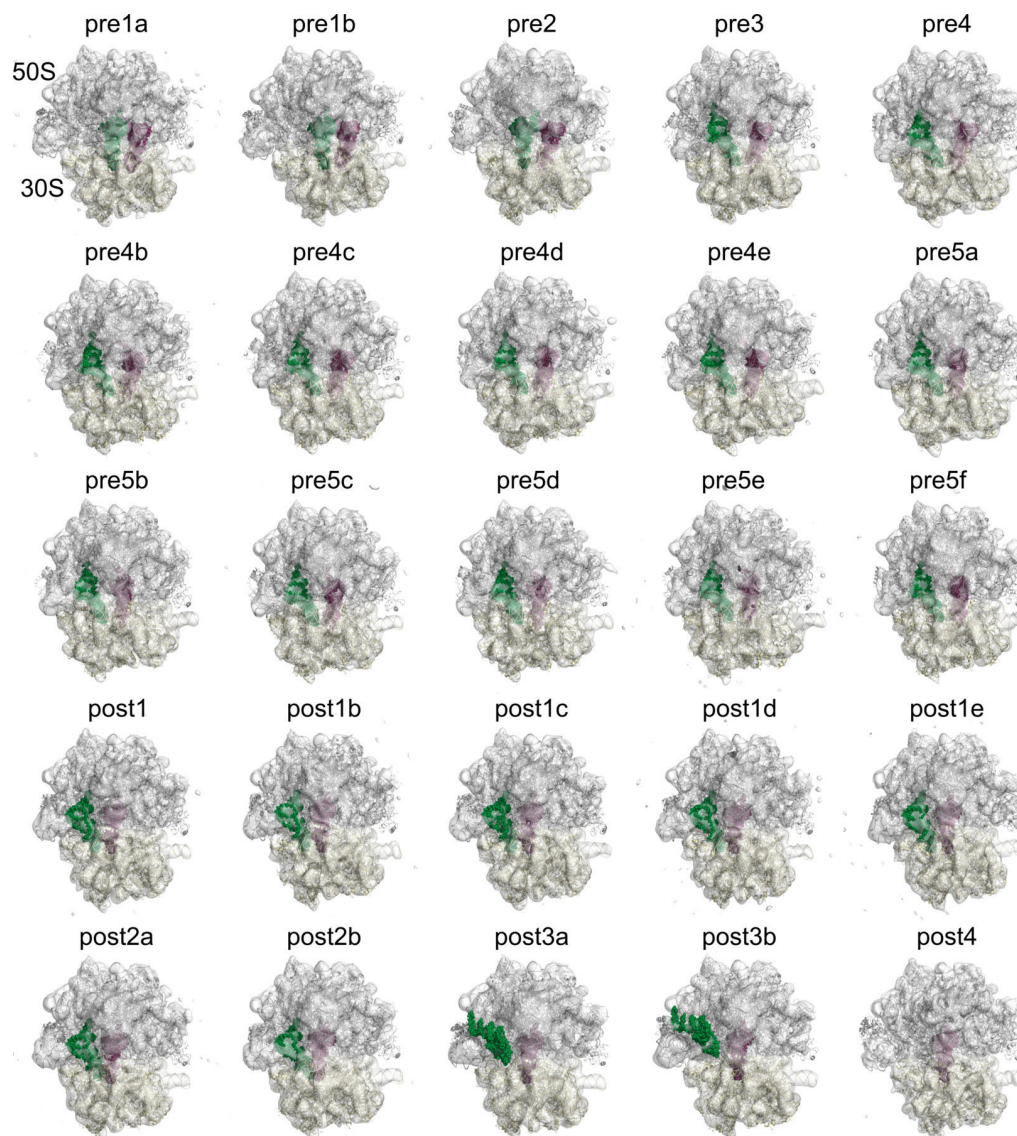
Figure 5.2: **Complete set of near-atomic models obtained from refinement against cryo-EM maps [27].** For the 25 structures, the fitted structure and an isosurface of the cryo-EM map (gray surface) is shown. The ribosomal subunits (50S and 30S) are shown in ribbon representation; tRNA$^{\text{fMet}}$ and tRNA$^{\text{Val}}$ are depicted by magenta and green spheres, respectively.
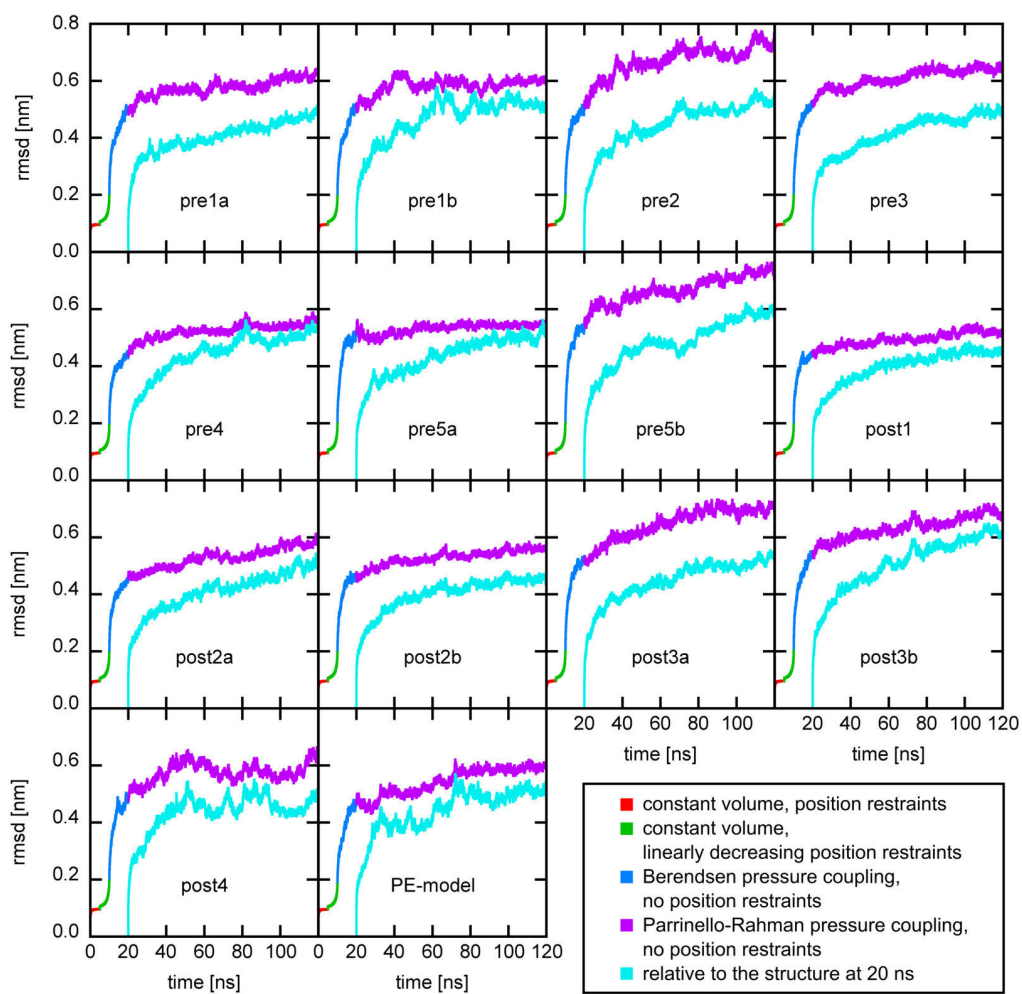
Figure 5.3: **Structural deviations during the simulations.** For each ribosome simulation, started from the model fitted to the cryo-EM map or from the PE-model, the RMSD relative to the starting structure is shown for the different steps of the simulation (red, green, blue, and magenta curves), and relative to the structure at 20 ns (cyan curve).
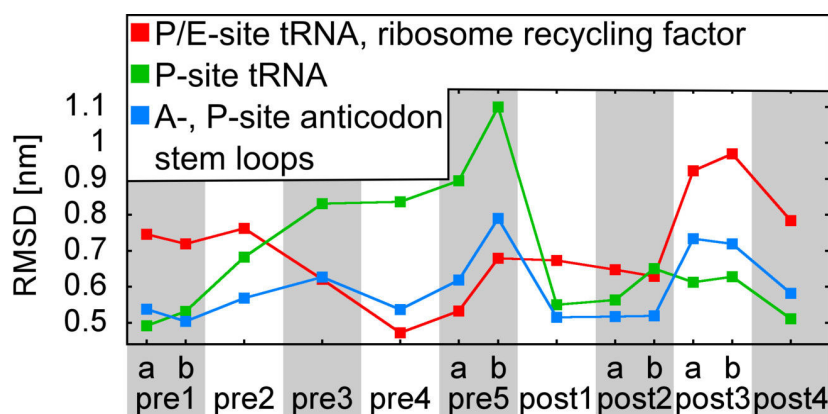
Figure 5.4: **Comparison of models to known crystal structures.** Backbone root mean square deviation (RMSD) between the MD structural models of the 13 translocation intermediates and three existing crystal structures (red and green curve [15], blue curve [14]).

5.4.2 Comparison of translocation intermediate models to crystal structures

Structures obtained from averaging the last 10 ns of the 13 translocation intermediate simulations were compared with crystal structures of *E.coli* ribosomes [14, 15] that have not been used for modeling (Fig. 5.4). The crystal structure of the ribosome trapped in the classical state [15] (green curve) closely resembles the model of the pre1a state, with a root mean square deviation (RMSD) of ∼5 Å. Other pre-translocation states markedly deviate from this classical state, showing RMSD values of up to 11 Å (pre5b). The structure by Zhang et al. (blue curve), comprises the ribosome and anticodon stem loops for A- and P-site tRNAs and shows 6° of 30S body rotation in reference to the 50S subunit [14]. The model that is most similar to this structure is the model representing the pre1b state with an RMSD value of below 5 Å. Recently, Dunkle et al. resolved the ribosome bound to a tRNA and the ribosome recycling factor [15] which is needed in the last step of translocation, dissociating the two subunits and rendering the mRNA available for the next round of translation [141, 142]. In this structure, the tRNA is found in a hybrid P/E conformation. The model for the late pre-translocation state pre4, with the P-site tRNA in the hybrid P/E position and a high degree of rotation between the ribosomal subunits, agrees well with this ribosome structure [15], with an RMSD value below 5 Å (red curve). The good agreement between the computational models of the pre1 and pre4 states and the respective crystal structures provides an independent quality check for the

structural fits and suggests similarly high accuracy for the models of other pre- and post-translocation states for which no crystal structures exist.

5.4.3 Choreography of ribosomal motions

To describe the complex motion of the ribosome during translocation, six dominant large-scale movements of ribosomal components were monitored: the motions of tRNA$^{\text{Val}}$ from A to P and of tRNA$^{\text{fMet}}$ from P to E site, the L1-stalk (comprising L1 protein bound to nucleotides 2084–2206 of 23S rRNA) from the open to the closed conformation, as well as swiveling and tilting of the 30S head and rotation of the 30S body (see schematic representation in Fig. 5.5A). To quantify these motions, Fig. 5.5B summarizes the choreography of motions for both tRNAs, and the L1-stalk. The reaction coordinates (r. c.) are obtained by the projections onto the first eigenvector of a PCA analysis (see chapter 5.3.2). When moving through the ribosome, the two tRNAs switch between binding sites: tRNA$^{\text{Val}}$ moves between the A (pre1–pre4), A/P (pre5), and P sites (post states), and tRNA$^{\text{fMet}}$ between the P (pre1–pre2), PE (pre3–pre5) and E (post1–post2), and L1 site (post3). Concomitantly, the position of the L1-stalk changes, moving from the open conformation (pre1–pre2) to the closed conformation (pre3–pre4) and then gradually into the open conformation (pre5–post4). The movements result in large alterations of the binding enthalpy between the tRNA and ribosomal proteins L1, L5, and L16 located on the 50S subunit (Fig. 5.5B, colored blocks). For each intermediate translocation state, black bars indicate the boundaries of rapid (10^7/s) fluctuations accessible to the simulations. Accordingly, lack of an overlap between adjacent states (e.g., between pre4 & pre5 or pre5 & post1 for tRNA$^{\text{Val}}$) indicates transitions between states that are much slower than 100 ns.

For the large-scale global 30S head and body rotations, pivot points, rotation axes, and angles (Fig. 5.5C) were determined, which allowed to disentangle swiveling and tilting motions of the 30S head [27, 58]. The pivot point for 30S head rotation was found to be close to 16S RNA residues G242 and U562, and for 30S body rotation it was found to be close to A923, U1194, G1386 (Fig. 5.6).

Upon transition from pre1 to pre5 states, 30S head tilting and body rotation angles gradually increase, before returning to the pre1 values upon translocation (pre5 to post1). In the post3 state, the rotation angles are even lower than in the pre1 or post1, and finally recover their initial value in post4, thus closing the conformational cycle of the 30S subunit after tRNA release from the E site. In comparison to the 30S head tilting and body rotation, the variations in the extent of the 30S head swiveling are much less pronounced,
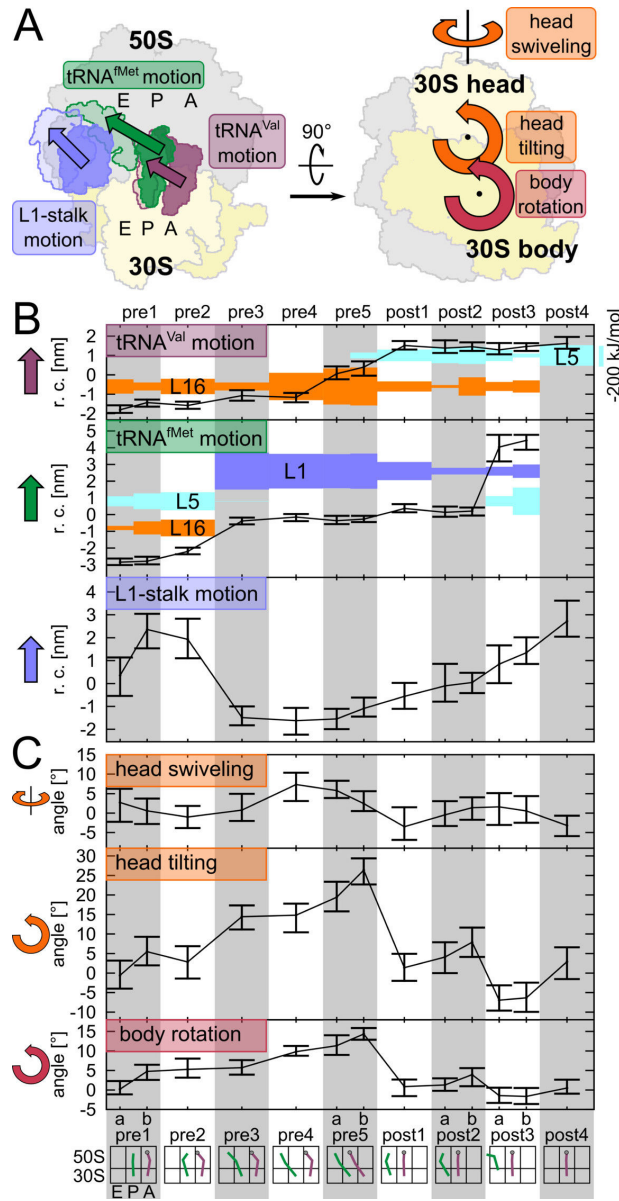
Figure 5.5: **Choreography of ribosome motions and tRNA-protein interactions during translocation.** (**A**) Schematic of tRNA^{Val}, tRNA^{fMet}, and L1-stalk motions (indicated by arrows, left), and of 30S head and body rotations (right). (**B**) Motions of tRNA^{Val}, tRNA^{fMet}, and the L1-stalk quantified by reaction coordinates (r.c.) for the 13 states. Arrows denote directions of motions as in (A); black bars indicate the fluctuation range covered by each simulation; colored bands denote interaction enthalpies between the respective tRNA and the 50S proteins L1, L5, and L16. (**C**) Changes in the angles of the 30S head tilting, swiveling, and 30S body rotation.
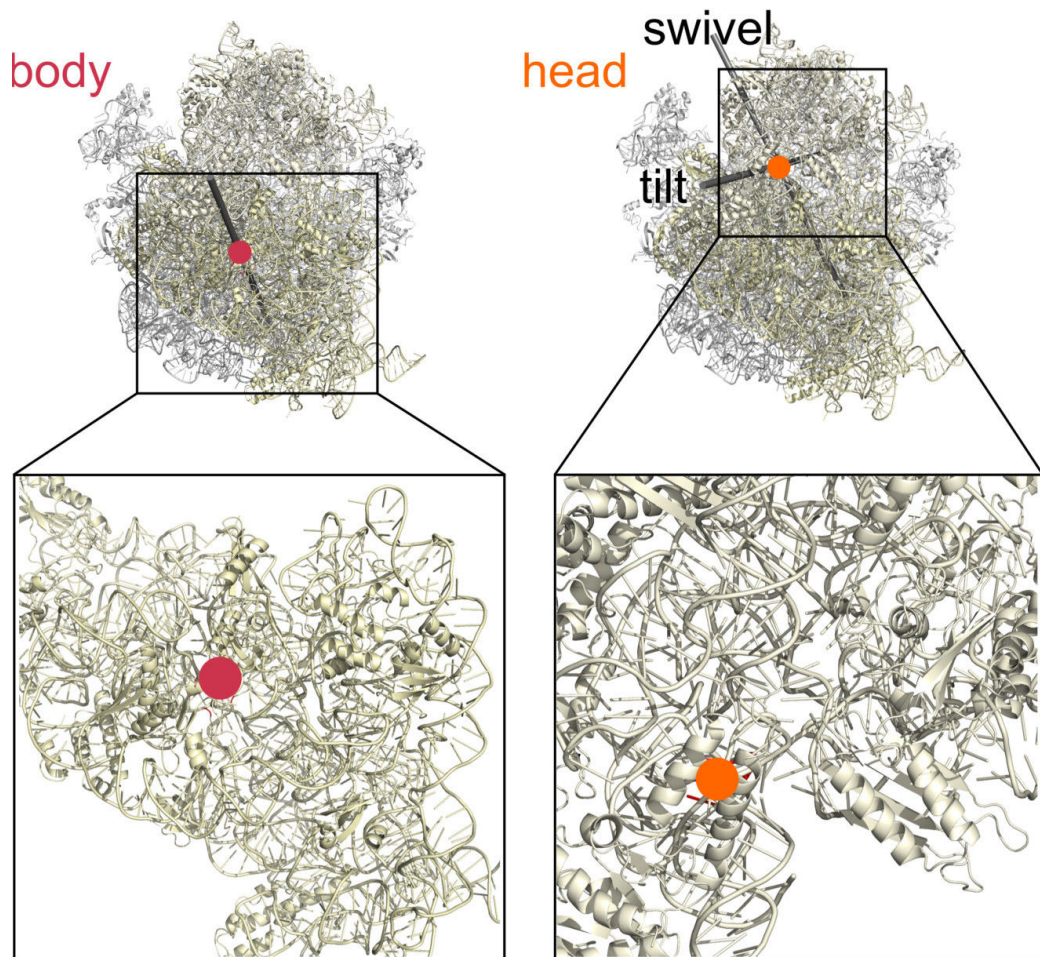
69

Figure 5.6: **Position of body and head rotation pivot points.** For 30S body rotation the main rotation axis (black cylinder) and the pivot point (circle) are shown (left). For head rotation, tilting and swiveling axes and the pivot point are shown (right). The ribosome structure, which was used as reference to determine the pivot points and axes of rotation, is shown in ribbon representation (30S light yellow; 50S light gray).

and consecutive states show larger overlaps, suggesting rapid head swiveling motions.

5.4.4 Individual transition rates of ribosomal components

To obtain rates for transitions between states for the above mentioned collective motions, the probability distributions along reaction coordinates extracted from the simulations were analyzed. First, initial free energy barrier heights $\Delta G^{\ddagger}_{\text{est}}$ for the transitions between all states for the motions of the L1-stalk, both tRNAs, and 30S head and body rotations were estimated from fluctuations extracted from the trajectories, as described in chapter 3.1. Next, the normalized distance between two ensembles representing two states in the space spanned by the reaction coodinates was calculated. Figure 5.7 (upper panel) shows this distance as a function of the barrier height estimate for each pair of state and for each motion considered. A distance smaller than one means that the average distance within the ensemble is smaller than the minimum distance between the ensembles. In this case the barrier was considered to be crossed in the simulation. From these distances, the transition frequency was calculated (Figure 5.7, lower panel).

To calculate transition rates using Arrhenius' law, the probability of barrier crossing $p_{\text{A}\rightarrow\text{B}}$ was least-squares fitted to the transition frequency $p^{\text{sim}}_{\text{A}\rightarrow\text{B}}$ obtained from the simulations. The fitting yielded an attempt rate of $A = (22.4\ \text{ns})^{-1}$ and a correction factor $c = 0.601$ at $t = 100$ ns, which allowed us to calculate the barrier height $\Delta G^{\ddagger} = c\Delta G^{\ddagger}_{\text{est}}$ and the respective transition rates k (Fig. 5.8).

A cluster of high free energy barriers with transition rates slower than $1/\mu$s is found at the transition from the pre to the post states, which provides an independent support for the notion that the pre-to-post transition is the rate-limiting step of spontaneous translocation [26, 27]. For the tRNA transitions (purple and green lines), low transition barriers for tRNA$^{\text{Val}}$ mostly correlate with high barriers for tRNA$^{\text{fMet}}$ and vice versa. This implies that, although intrinsically rapid, the excursions of tRNA$^{\text{fMet}}$ between the adjacent states (e.g., pre5b and post2a) are inhibited by the slow movement of tRNA$^{\text{Val}}$ due to coupling of the two tRNAs by base pairing to the mRNA. In comparison, most of the L1-stalk transitions are seen in the sub-μs range (blue lines), except for the slower movement from pre to post and from pre2 to pre3.

Notably, the analysis of the simulations reveals a large number of alternative transitions of the tRNAs and the L1-stalk which would take place in
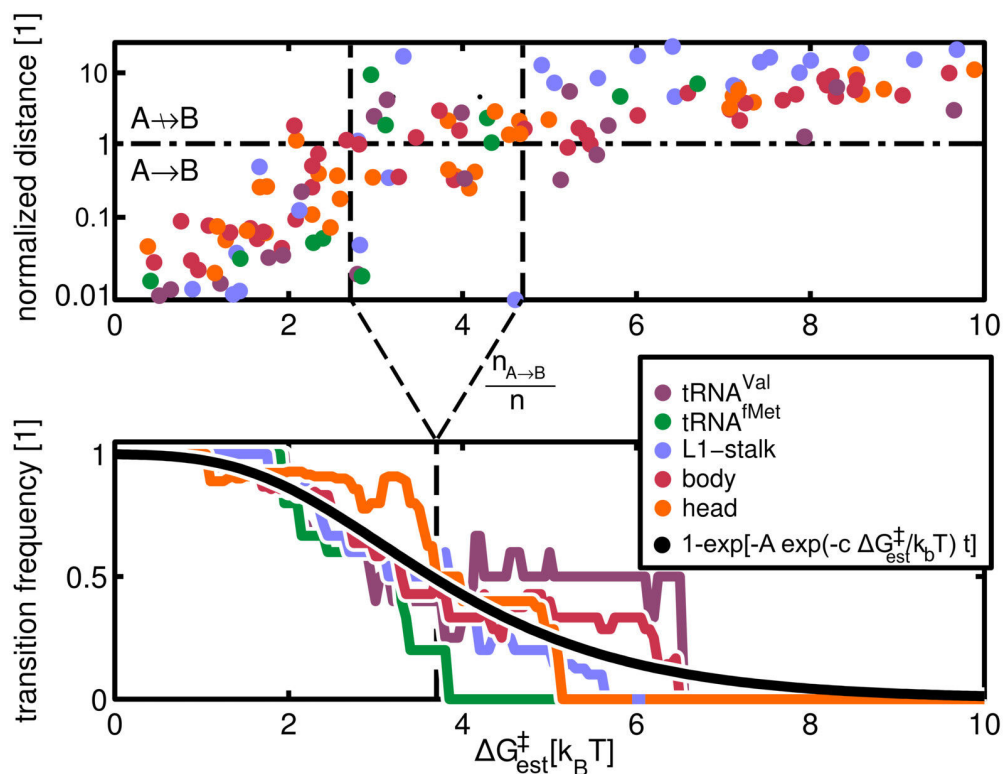
Figure 5.7: **Attempt rate and free energy gauge factor for free energy barrier estimation.** The upper panel shows the normalized distance between the ensembles for each pair of states versus the uncorrected free energy estimate $\Delta G_{est}^{\ddagger}$. This is done for each of the ribosome parts (colored circles). A barrier between two states is considered crossed if this distance is smaller than one. The lower panel shows the frequency of barrier crossings $p_{A \to B}^{sim} = (n_{A \to B})/n$ calculated for free energy intervals of 1 $k_b T$ (colored lines). The probability of barrier crossing $p_{A \to B}$ fitted to $p_{A \to B}^{sim}$ is shown as a black line.
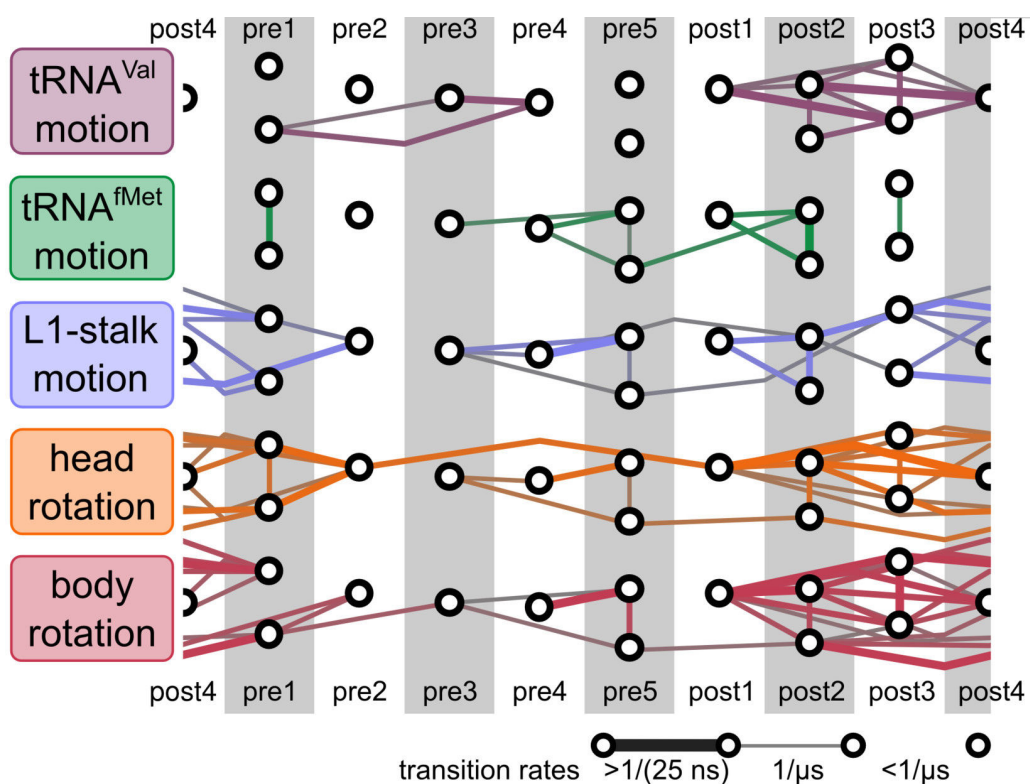
Figure 5.8: **Individual transition rates between states.** Estimated transition rates (encoded by line thickness) between states (circles) for tRNA and L1-stalk motions, and for 30S head and body rotations. Transition rate estimates slower than $1/\mu$s are not shown.

the μs and sub-μs range. However, due to coupling between the motions of the individual components, it is the slowest transition that determines the overall rates for transitions between states.

All individual transition barriers for 30S rotations turn out to be remarkably low (Fig. 5.8, orange and red lines), such that the intrinsic rotations of the 30S subunit, despite its large size, are in the sub-μs range, similar to the movements of the L1-stalk. The slowest 30S head transitions are found between pre2 and pre3, as well as between pre5 and post1 states, consistent with barriers inferred from cryo-EM and biochemical experiments [26, 27]. These slow barriers divide the states into two groups (group 1: pre3–pre5; group 2: post1–pre2) which, if they were uncoupled from other motions, could interconvert on fast sub-μs time scales. The body rotation angles for each group span a range of more than 10°, suggesting that these large-scale rotations are unexpectedly rapid. The abundance of low barriers for motions of individual components seems to be a general feature of the ribosome dynamics, which underscores the important role of the coupling between fluctuations of tRNAs and the ribosome.

Using all transition rate estimates determined from the simulations, one can predict which of the 479,001,600 possible linear kinetic sequences has the fastest progression rate. Remarkably, the most likely trajectory is identical to the one suggested by Fischer et al. [27], except for a swap of the two states post1 and post2, which are structurally very similar, suggesting that this sequence of states not only reflects the structural similarity, but also captures the kinetically favored pathway of tRNA movements.

Overall, a hierarchy of barriers is seen. The largest barriers separate pre from post states; smaller barriers and correspondingly faster transitions are seen between groups of pre-translocation states (pre1-pre2 and pre3-pre4-pre5). Sub-μs transitions dominate within these groups and between the post states. Of all the motions described above, barriers for tRNA transitions dominate, which suggests a possible explanation for the strong effect of tRNA species on the rate of spontaneous translocation [26, 101].

5.4.5 Mechanisms to overcome rate-limiting barriers

Rapid short-timescale tRNA diffusion

Analysis of the rapid dynamics of each tRNA revealed fluctuations of up to 4 Å within a given state. These large fluctuations, together with the observed attempt rate of one per 20 ns for barrier crossings, seem to be incompatible with the typically low diffusion constant of a tightly bound ligand. Indeed, averaged over all MD simulations, the short-timescale diffusion constant for
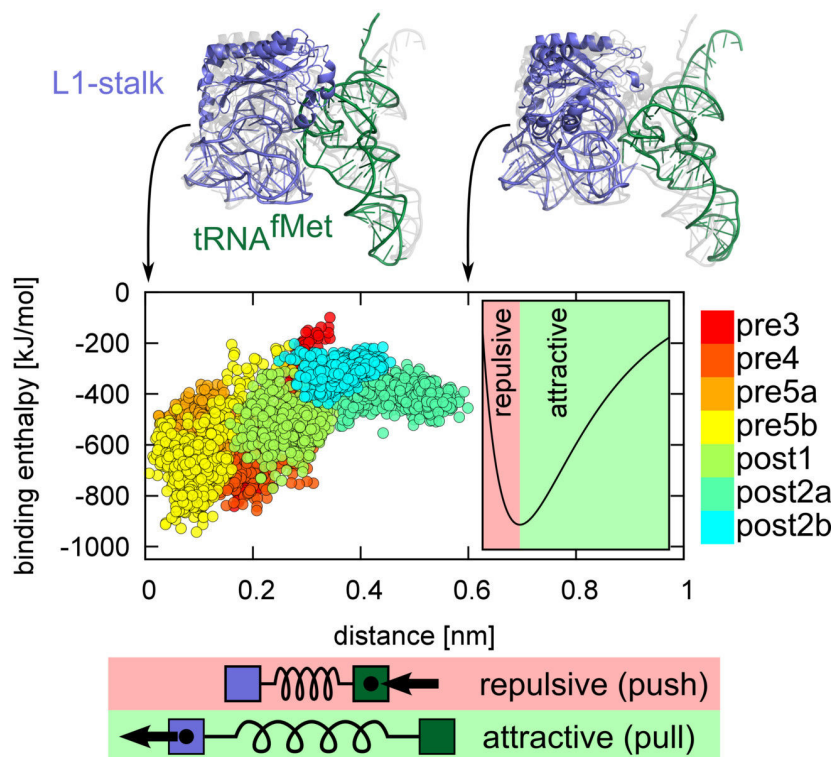
Figure 5.9: **Molecular driving forces between ribosomal protein L1 and tRNA.** Two structures representing the range of distances from the fully closed (top, left) to open (right) conformations. Distance dependence of binding enthalpies for different states (middle, left) compared to a typical interaction potential (right) and its mechanical analog (bottom).

tRNA motion is $\bar{D} = (2.2 \pm 0.2)10^{-9} \mathrm{cm}^2 \mathrm{s}^{-1}$, which is only two orders of magnitude smaller than in solution ($\bar{D} = (360 \pm 60)10^{-9} \mathrm{cm}^2 \mathrm{s}^{-1}$) and resembles free diffusion within an environment comparable to olive oil. Without the larger barriers which determine the ms dynamics and the overall translocation time, this diffusion constant would allow the tRNA to sample the full 10-nm pathway through the ribosome within only ∼175 μs. We suggest that the high inherent mobility of tRNA within the ribosome is a prerequisite for rapid tRNA translocation.

L1-stalk pulls tRNA from P to E site

The L1-stalk forms an important contact to the tRNA during translocation [37, 119, 143–147]. In the pre1 and pre2 states, the L1-stalk is found in an open conformation, and in the pre3 state it moves towards the E site where

it interacts with tRNA^{fMet} [27]. In the pre states the complex is stabilized by strong enthalpic interactions (Fig. 5.5B). The contact with the tRNA is gradually released in the post states, allowing the L1-stalk to rearrange into the open conformation. Analysis of how the binding enthalpy changes with distance between L1 and the tRNA^{fMet} (Fig. 5.9) enabled us to identify the driving agent of that process: Assuming a typical interaction potential (a mechanical analog is shown at the bottom), the agreement between the positive slope of the binding enthalpy as a function of distance between tRNA^{fMet} and the L1-stalk with the attractive branch of this potential (light green) suggests that it is the L1-stalk which pulls the tRNA upon transition from the pre3 to pre4/5 and to the post states, thus lowering the barrier for tRNA motion. A scenario (Fig. 5.9 light red) in which the tRNA pushes the L1-stalk is incompatible with the observed distance dependence. This finding explains both the lower translation rates [148] and the more stable classical tRNA states [149] observed for L1-depleted ribosomes.

Proteins L5 and L16 facilitate tRNA motion

To identify further elements of the ribosome that facilitate tRNA translocation, the main tRNA-50S subunit interactions were tracked along the complete tRNA pathway from the A to the E site. During translocation, the two tRNAs are handed over from ribosomal protein L16 to protein L5 and then to the L1-stalk (Fig. 5.10, right). These ribosome proteins remain tightly bound to the tRNAs by strong enthalpic contacts (Fig. 5.5B) despite considerable structural dynamics of all components. Closer analysis of the residue contacts of both tRNAs with the ribosome in each state (Fig. 5.10) suggests several different strategies for moving the tRNAs and maintaining the tight binding at the same time. During the whole A/A to P/P transition, tRNA^{Val} remains in contact with L16 (Fig. 5.10, top left). Along this transition, two adjacent interaction patches on L16 are utilized, each involving positively charged arginines (R51/R55/R50/R59 and R10/R81/R6) interacting with different parts of the negatively charged tRNA backbone. Upon hybrid state formation (pre4 to pre5) — when the acceptor stem of tRNA^{Val} moves into the P site of the 50S subunit while the L16 conformation remains unchanged — the contact region on tRNA^{Val} switches from the first to the second patch. Similar interactions with the second patch on L16 are seen for the tRNA^{fMet} in its P/P configuration (bottom left), suggesting that these contacts can stabilize any tRNA in the P site. During the tRNA handover from L16 to L5, the binding of tRNA^{Val} to L16 becomes weaker (Fig. 5.5B), whereas that of tRNA^{Val} to L5 (cyan bars) remains strong in the pre5b and throughout the post states. Strong contacts between the tRNA C56 and the P-site loop of L5
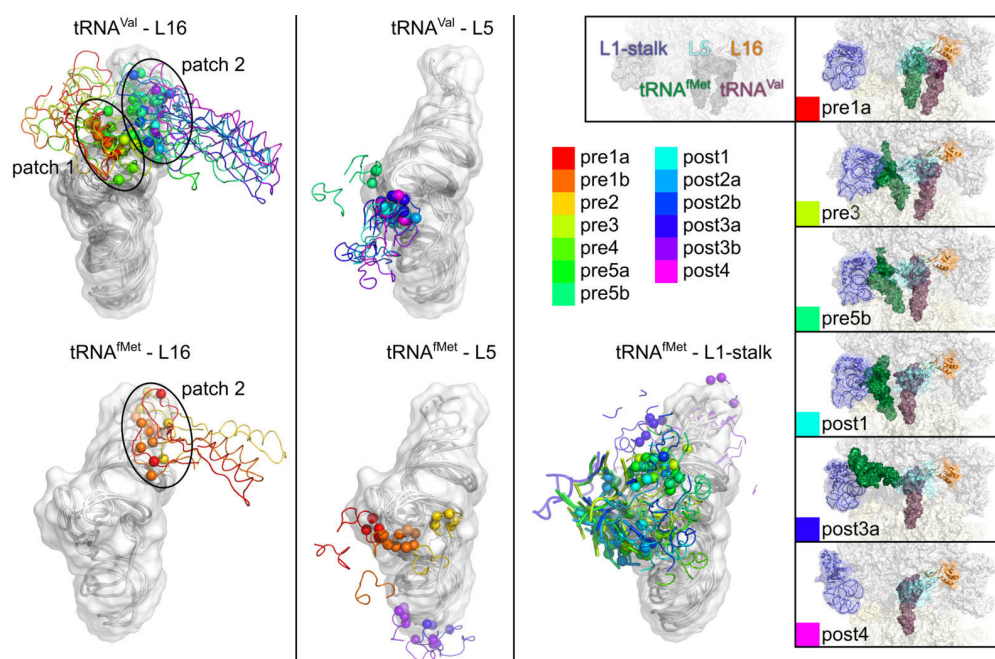
Figure 5.10: **Handover of tRNAs by ribosomal proteins during translocation.** Positions and key interacting residues (spheres) of the three main interaction partners, the ribosomal proteins L16, L5, and the L1-stalk (ribbons, colored according to the state), relative to tRNA$^{\text{Val}}$ (top row) and tRNA$^{\text{fMet}}$ (bottom). Selected conformations relative to the 50S subunit are shown for reference (right).

(A74–I78) [150] are also seen in the pre5b state. This finding, together with the large intersubunit rotation angle in the pre5b state, suggests a coupling of the tRNA movement via L5 to the intersubunit rotational motion, since L5 is involved in intersubunit contacts [12]. In the subsequent post states, the contacts of the L5 P-site loop shift down the D-loop of the tRNA$^{\text{Val}}$ (Fig. 5.10, top middle), now additionally involving nucleotide G19.

In contrast to the stepping motion of the rather rigid L16, the interactions between tRNA$^{\text{fMet}}$ and L5 (Fig. 5.10, bottom middle panel) are much less localized and more dynamic. In the pre1 and pre2 states, tRNA$^{\text{fMet}}$-L5 contacts shift from the D loop to the T loop of the tRNA. Upon handover to the L1-stalk in the pre3 state, the contact to L5 is lost, but new contacts between L5 and the anticodon stem-loop are formed just before the tRNA leaves the ribosome (in post4). Here, rather than stepping, the P-site loop of L5 slides smoothly over the tRNA, quite flexibly adapting to the changing tRNA position and orientation. Mutations of this loop in yeast ribosomes affect tRNA affinity to the P site [150], underscoring the importance of L5 as a guide for the P-site tRNA. The L1-stalk provides the final contact for tRNA$^{\text{fMet}}$ after its handover by L5 and L16. Interactions of L1 with tRNA$^{\text{fMet}}$ are established in the pre3 state and remain unchanged from pre3 to post2. The main interactions involve the tRNA backbone and positively charged L1 residues (R53, K54, R60, and R164). The abundance of non-specific contacts between charged residues and tRNA backbone suggests that this type of sequence-independent interactions may provide a general means for transport of different tRNA species through the ribosome.

The results of the above analysis suggest that tRNA binding and movement through the ribosome is supported by three different mechanisms, each adapted to the specific thermodynamic and functional requirements of the respective tRNA binding sites. First, the precise positioning of the tRNAs in the A and P sites is achieved by interactions with L16, which is rather rigid. To accommodate tRNA motion, L16 steps via discontinuous contact patches. Second, the flexible L5 P-site loop facilitates seamless tRNA sliding across the P site. Finally, L1, which exerts force to pull the tRNA out of the P site and hence requires internal stiffness, forms only one stable contact patch with tRNA, and moves together with the tRNA as a rather rigid body. In summary, interactions with L16, L5, and L1 accelerate tRNA translocation through the ribosome by reducing free energy barriers between adjacent binding states.
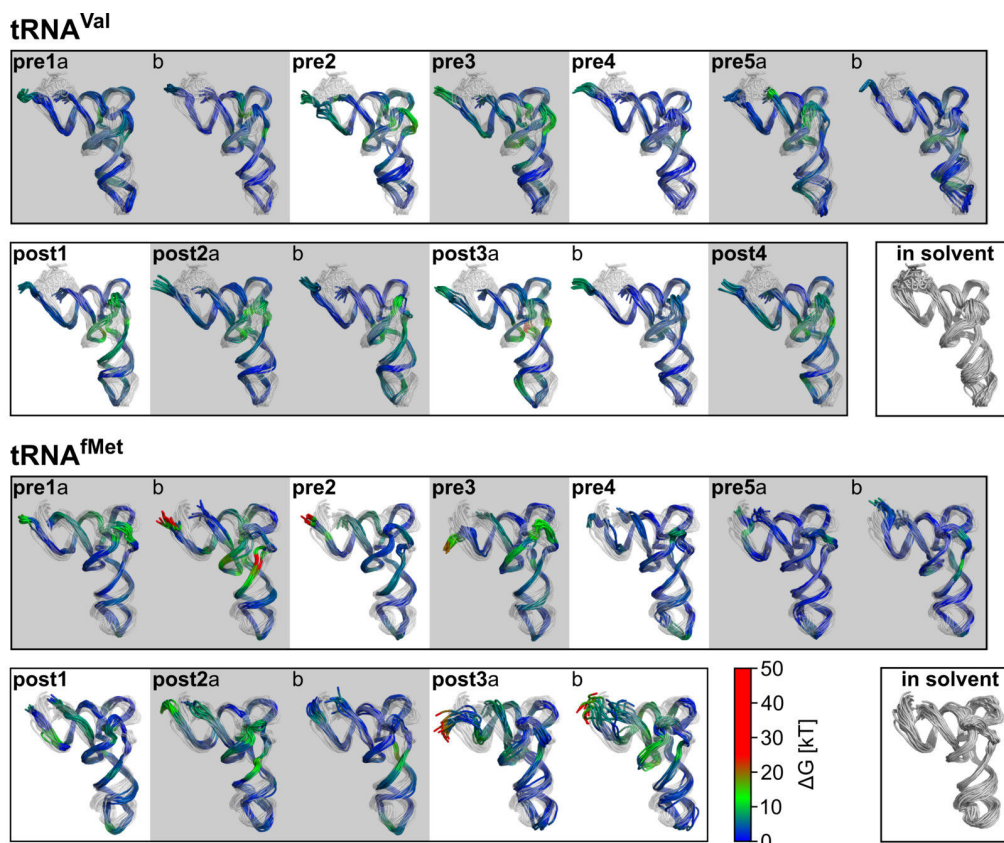
Figure 5.11: **tRNA conformational flexibility and entropic energy cost of binding.** For both tRNAs (tRNA$^{\text{Va}}$, top; tRNA$^{\text{fMet}}$, bottom), structural ensembles were extracted from simulations in solution (gray ribbons) and bound to the ribosome (states pre1–post4 colored; solution ensemble shown in light gray as reference). Estimated conformational free energy penalty (per residue) due to steric restrictions upon binding (color coded). Due to correlations, per residue entropies are not additive. Note that the solution ensembles cover larger time scales (1 μs each) than those bound to the ribosome (100 ns each).
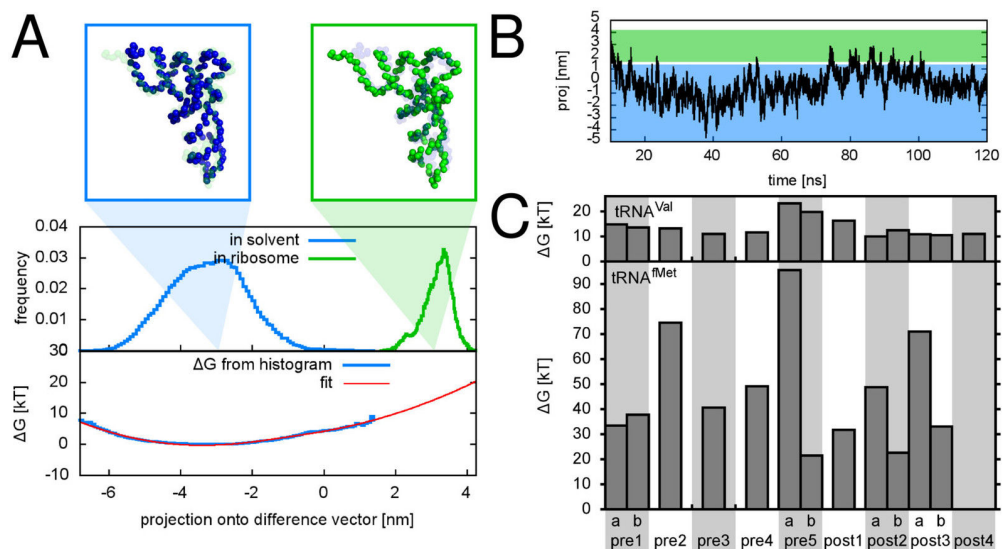
Figure 5.12: **Conformational free energy of the tRNAs bound to the ribosome.** (**A**) Histogram of the projections of the bound tRNA trajectory (tRNA[Val]) in the pre1a state, green curve, top panel) and of the tRNA in solution (blue curve) on the difference vector between the averaged structures (in solution, blue box; bound, green box). Conformational free energy ΔG (green curve) estimated from the histogram for the tRNA in solution and a fit of the sum of two quadratic functions (red curve). The conformational free energy was estimated by weighing the probability distribution from the bound simulation with the fitted function. (**B**) Projection on the difference vector of a tRNA simulation in solution, starting at the conformation when bound to the tRNA. The colored areas denote the range along the difference vector accessible to the bound and solution simulations as depicted in (A). (**C**) Conformational free energy estimated for tRNA[Val] (upper panel) and tRNA[fMet] (lower panel) for each state.

5.4.6　Conformational free energy of the tRNAs

Due to steric restrictions generated by the ribosomal environment during translocation, the tRNA dynamics inside the ribosome differ markedly from tRNA dynamics in solution (Fig. 5.11). To investigate the amount of conformational free energy needed to restrict tRNA dynamics to a certain region of phase space, the conformational free energy difference ΔG between each tRNA bound to the ribosome (pre1–post4) and each tRNA in solution was estimated. For each state and each tRNA, the trajectories of the tRNA bound to the ribosome and of the 1 μs simulation of the tRNA in solution were projected onto the difference vector between the average structures of the two trajectories (example for tRNA$^{\text{Val}}$, pre1, in Fig. 5.12A). From these projections, the conformational free energy ΔG was estimated (Fig. 5.12C). To examine if the conformations of the bound tRNAs represent stable conformations in solution which is not visited by our free tRNA simulations, the tRNA conformation was extracted from the ribosome simulation for each state, and simulated in solution (example for tRNA$^{\text{Val}}$, pre1, in Fig. 5.12B). In all cases, the tRNAs moved within 100 ns of simulations to the region explored by the 1 μs solution simulations, indicating that there are no large barriers between the conformations of the bound tRNAs and the conformations of the tRNAs in solution. This would mean that the conformations of the bound tRNAs do not represent additional minima of the free energy landscape and thus supports the hypothesis that free energy is stored in the tRNA conformations when bound to the ribosome.

The estimated conformational free energies of tRNA$^{\text{fMet}}$ are markedly larger than those of tRNA$^{\text{Val}}$. Note that tRNA$^{\text{Val}}$ in the post4 state and tRNA$^{\text{fMet}}$ in the pre1 states are bound to the P site. Apart from the sequence, the only difference between the two scenarios is that tRNA$^{\text{fMet}}$ has a neighboring A-site tRNA. The fact that, for each state, the conformational free energy of tRNA$^{\text{Val}}$ is smaller than that of tRNA$^{\text{fMet}}$ suggests that the difference is due to the tRNA sequence. The conformational free energy would then be an important contribution to the fine-tuning of tRNA affinity to the ribosomal binding sites. The free energy of a base pair is about $-8\,\text{k}_\text{B}\text{T}$ [151]. Therefore, the release of conformational free energy can be a significant contribution to the energetics of translocation, since the acceptance of a new tRNA by the ribosome depends on the presence of a single codon-anticodon base pair.

Figure 5.11 shows to which extent the individual nucleotides of the two tRNAs are affected by this binding free energy penalty. The tRNAs in the different states show diverse patterns of conformational free energy distributions by individual nucleotides. The parts of the tRNA which contribute most

are the CCA-tail, the D-loop and the D-stem (Fig. 5.11). The pronounced deformations of the CCA-tail required to place it at [115], maintain it at (all states for tRNA$^{\text{Val}}$, pre1–pre2 for tRNA$^{\text{fMet}}$), and subsequently remove it from the peptidyl-transferase center (PTC) during translocation, require the charging of the tRNAs with free energy. In the post3 states the CCA-tail of tRNA$^{\text{fMet}}$ strongly interacts with the L1 protein (Fig. 5.10), thus keeping a high free energy level. This free energy can then be used for dissociation from the ribosome compensating the loss of interaction energy between ribosome and tRNA (see Fig. 5.5).

5.4.7 Intersubunit contacts during rotation

Translocation entails large-scale intersubunit rotations. Therefore, the question arose of how the ribosome maintains and controls the affinity between the 50S and 30S subunits despite these rearrangements. To characterize the dynamics of bridges connecting the subunits, groups of residues of the 30S and 50S subunit that are in contact with each other at any time in the simulations (contact clusters) were identified (Fig. 5.13A; see Tab. 5.3 for the correspondence to previously described intersubunit bridges [12]), and their binding enthalpies during translocation (Fig. 5.13B) were calculated. The total binding enthalpy remains rather constant throughout translocation; differences between states are on the order of fluctuations within the 100 ns simulations, thus stabilizing the complex. Moreover, larger changes in binding enthalpy would imply free energy barriers which would slow down intersubunit rotation and thereby translocation.

Strong interactions are found for contact clusters 1–4; these contacts have already been suggested as key determinants for intersubunit binding free energy [152, 153]. Indeed, the contacts that form clusters 1, 3, and 4, which are located close to the axis of rotation, are maintained through the whole translocation process (Fig. 5.13C). In contrast, clusters 2 and 5–9 are situated at the periphery and change their interaction partners upon rotation (Fig. 5.13C). These clusters exhibit weaker and more variable binding enthalpies during rotation, which suggests that they may take part in the network of interactions that facilitate directed tRNA movement. One striking example is provided by the analysis of the interactions of cluster 7. The 50S subunit part of cluster 7 on the L1-stalk pulls the tRNA$^{\text{fMet}}$ towards the E site in the pre3–pre5 states. This motion correlates with that of the 30S subunit part of cluster 7, which in turn is moved by 30S subunit rotation, thereby connecting tRNA translocation to intersubunit rotation.

Notably, clusters 2 and 9 at the periphery of the intersubunit surface together maintain strong interactions throughout all states, despite a large-
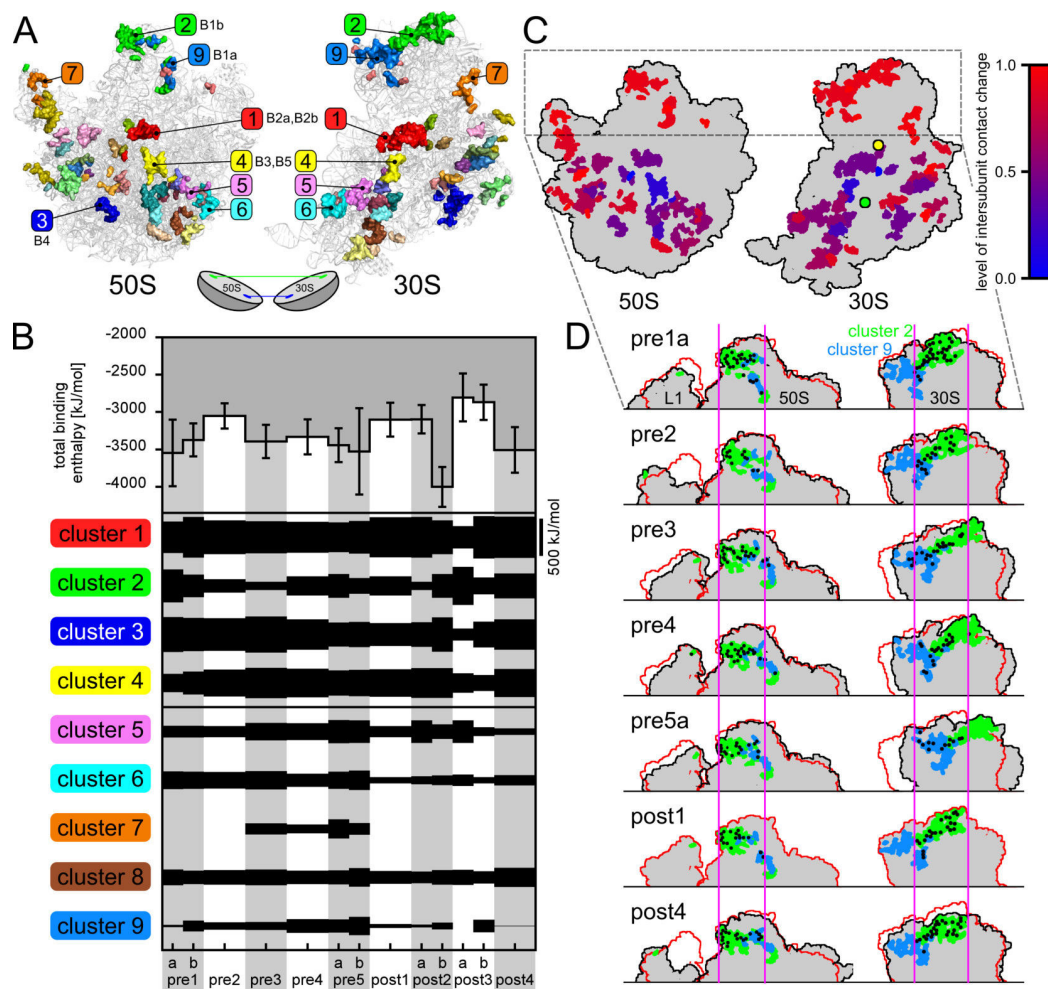
Figure 5.13: **Dynamics of intersubunit contacts.** (**A**) Clusters of residues contributing to intersubunit contacts (colored surfaces) between the 50S and 30S subunits (gray); key clusters are labeled by numbers and by the conventional bridge names [12]. (**B**) Binding enthalpies between 50S and 30S subunits (top) and contributions of individual clusters 1–9 during translocation. Note that due to entropic compensation and electrostatic shielding, the interaction *free* energies that determine affinity are weaker typically by 2–3 orders of magnitude. (**C**) Changes in intersubunit contacts during tranlocation; pivot points are indicated for head (yellow) and body rotation (green). (**D**) Clusters 2 (green) and 9 (blue) form a 'track glider'. Residues forming intersubunit contacts are marked as black circles. The outlines of the 50S and 30S subunits are shown in black, with post1 state as reference (red outline).

83

cluster	bridge	30S residue	50S residues
1	B2a, B2b	16S (1406-1411, 1491-1497, 1517), S12 (37-38, 40, 43, 48)	23S (1911-1920, 1932)
2	B1b	S13 (1-11, 22, 24, 46-49, 52-54, 56-67, 69-71, 73-78)	L5 (107-110, 112-113, 115-116, 132-148, 177), 23S (882-884, 890, 2306), L1 (98)
3	B4	S15 (35, 39, 42-44, 46, 49, 52-53, 55-56, 59, 62-63, 85-88), 16S (762, 764)	23S (710-717)
4	B5, B3	16S (1418-1420, 1483-1485)	23S (1768, 1947-1950, 1958-1962), L14 (44, 54)
5		16S (50, 52-53, 317, 337-340, 1481)	L14 (13-14, 49-51, 97-98, 100, 114)
6	B8	16S (160, 341, 344-349)	L14 (105, 108, 110, 113, 118-119, 121-123), L19 (33-36, 38, 40-41)
7		S7 (128-133, 135-136, 138, 141-142, 148)	23S (2113-2117, 2147-2148, 2167)
8		16S (1431-1433, 1463-1468)	L19 (62, 64-65, 103-108, 110)
9	B1a	S13 (72, 80-83, 88, 90-92), S19 (23, 27, 29, 45, 47, 58-60, 62-65, 68, 73, 78-81)	23S (885, 888-889, 891), L5 (111, 114, 178)
10		S6 (14, 24, 27, 77-83)	L2 (119-123, 125, 129, 133-139, 166, 168-169, 191-192)
11		16S (1440-1444, 1461-1462)	L19 (24, 43, 84-88, 111-114)
12	B7a	16S (702-703)	23S (1846-1849, 1894-1895, 2097, 2099-2100, 2191)
13	B6	16S (1426, 1474-1476)	23S (1688-1690, 1700-1702, 1766, 1987-1988)
14		S11 (11-12, 37, 74-75, 79), 16S (709)	23S (2139-2145, 2182-2185), L9 (75, 122-126, 128-129, 143)
15		16S (1470-1473)	L14 (17-18), 23S (1989), L19 (66)
16		16S (711-714)	L2 (162, 164-165, 174, 180, 270)
17		16S (679-681, 710, 777)	L2 (172, 267-269, 271-272)
18	B2b	16S (783-784)	23S (1835-1837)
19		16S (775-776)	L2 (176)
20	B6	16S (1428-1430)	23S (1703-1705)
21	B7b	16S (771-773)	L2 (4, 14-16, 160, 201), 23S (1793)
22		16S (810, 899)	23S (1693, 1830-1831), L2 (10, 13)
23		S6 (72-73, 76)	L2 (124)
24	B5	16S (1421-1423, 1480)	L14 (45-48, 53)
25		16S (270-271, 1438-1439)	23S (1747-1751)
26		16S (342-343)	L14 (116-117, 120)
27		16S (682-683)	23S (2192-2194), L9 (91)
28		16S (1516)	23S (1931)
29	B1a	S13 (79, 93, 113), S19 (26, 55)	23S (887), L11 (91, 94)
30		16S (790-791)	23S (1921-1922)

Table 5.3: **Clusters of intersubunit contacts.** Intersubunit contact clusters with the 30S and 50S residues, compared to the bridges previously described by Yusupov et al.[12].

scale rearrangement of the ribosome which would move the centers of these bridges by more than 35 . This is achieved by a strategy different from both the stepping and the sliding mechanisms discussed above for tRNA transport. These clusters, mainly consisting of proteins L5 and S13, also differ from the others in their shape, with the long stretch of residues located tangentially on the periphery of the 30S subunit, and a compact contact area on the 50S subunit. During the transition from the pre1 through post3 states (Fig. 5.13D), the contacts of the 50S subunit shift along the residues of the extended cluster on the 30S subunit synchronously to the 30S body rotation, whereas on the 50S subunit these interactions remain local, involving the same residues throughout the cycle, thereby resembling a glider (on the 50S subunit) moving on a track (composed of the 30S subunit residues). By combining the rather rigid components of the above stepping mechanism with the continuity of the sliding mechanism, this track glider strategy not only ensures the stability of the complex throughout the whole translocation cycle, but also explains why removal of S13 results in defective subunit association [154].

5.4.8 tRNA-50S-30S interactions stabilize high rotation states

The maximum head tilting and body rotation angles at $\sim 20°$ and $\sim 15°$, respectively (Fig. 5.5, substate pre5b) correlate with strong interactions involving L1-stalk (cluster 7) and L5 (cluster 2, 9), both between the subunits (Fig. 5.13B) and with the tRNAs (Fig. 5.5, blue and light blue bars). This result suggests that removal of the tRNAs would weaken also these intersubunit interactions and therefore reduce the stability of highly rotated states. This notion is further supported by MD simulations of ribosomes from which the two tRNAs were removed, indicating that in the absence of tRNAs the intersubunit interactions at clusters 2 and 9 are weakened and the 30S subunit body rotation angles are decreased (Fig. 5.14A,B,C). In agreement with the MD simulations, cryo-EM analysis of ribosome conformations in the absence of tRNAs indicates markedly reduced frequencies of high body rotation angles (Fig. 5.14D). A similar effect should be observed upon weakening of the intersubunit interactions, e.g., by mutation of charged L5 residues R109, R111, D143, D146 to uncharged residues. Indeed, mutations of corresponding residues in yeast affect translational fidelity [155].
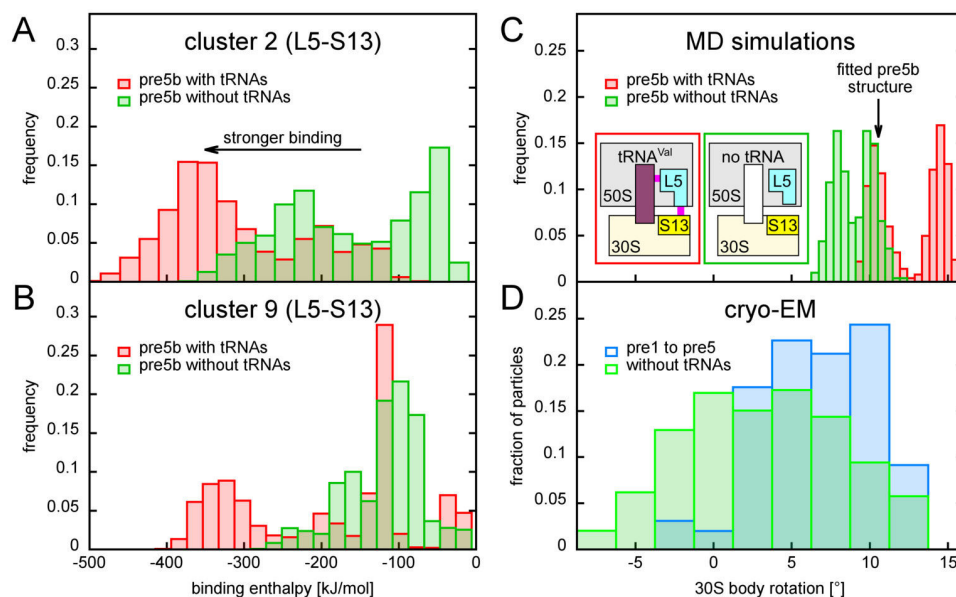
Figure 5.14: **Effect of the tRNAs on the large-angle 30S body rotations.** (**A**, **B**) Histograms of intersubunit binding enthalpy for clusters 2 and 9 calculated from simulations of the pre5b structural model with (red bars) and without tRNA (green bars). (**C**) Histogram of respective rotation angles. The arrow marks the angle of the starting structure as reference. Inset: Schematic of the tRNA-L5-S13 interaction network. In the presence of the tRNA, proteins L5 and S13 are in contact via clusters 2 and 9, stabilizing the rotated state; in absence of tRNAs the intersubunit interaction is weaker leading to a destabilizing the rotated state. (**D**) Histogram of 30S body rotation angles derived from cryo-EM analysis of ribosomes in the absence of tRNAs (green bars) compared to those of states pre1 to pre5 (blue bars).

5.5 Summary and Conclusions

By combining X-ray crystallography and cryo-EM data with the flexible fitting protocol DireX and extensive MD simulations, 25 structures of translocation intermediates at near-atomic resolution were determined and 100 ns dynamics for 13 states were calculated using subsequent MD simulations. This hybrid approach was validated by comparison to recent crystal structures. From an analysis of the fluctuations observed in the simulations, rates for transitions between states were estimated for the large-scale motions of the two tRNAs, the L1-stalk and rotations of 30S head and body in reference to the 50S subunit (Figs. 5.5 and 5.8).

The estimated rates for 30S head and body rotations and for the L1-stalk motion are faster than $1/\mu$s, except for the pre2 to pre4 transition and the pre5 to post1 transition which are known to be rate-limiting for translocation [27]. This suggests that these large-scale motions are not rate-limiting for the remaining transitions. In contrast, for the motions of the tRNAs, we found an abundance of high barriers rendering the tRNAs as rate-limiting for these transitions.

The observation of tRNA motions as rate-limiting for most transitions during translocation, suggested that one way to accelerate this process would be to specifically lower barriers for tRNA motions. The question addressed here is: Does the ribosome employ mechanisms to lower these barriers? The stable contacts between tRNA$^{\text{fMet}}$ and the L1 protein (Figs. 5.5 and 5.10) when the tRNA moves first from the P/E hybrid state (pre3–pre5) to the E site (post1–post2) and then out of the E site (post3) as well as the large-scale L1-stalk motion, render the L1-stalk a promising candidate for actively lowering tRNA motion barriers. Indeed, the dependence of the binding enthalpy between the L1-stalk and the tRNA on their mutual distance leads to the conclusion that the L1-stalk actively pulls the tRNA instead of being pushed by the tRNA (Fig. 5.9). Additionally, the tRNAs were found to have a fast short-timescale diffusion constant compared to free diffusion in olive oil, leading to a large barrier crossing attempt rate of $1/(20 \text{ ns})$ which is, besides lowering barriers another prerequisite for rapid transitions between states.

Further, two ribosomal proteins guide the movement of tRNAs through the ribosome by different mechanisms. First, protein L16 presents two distinct binding patches to the tRNAs which allows for precise tRNA positioning in the A and P sites (Fig. 5.10). Second, the P-site loop of protein L5 flexibly adapts to the tRNA moving through the P site (Fig. 5.10), thereby keeping the binding enthalpy to the tRNA rather constant and, as a consequence, lowering the barriers for tRNA motion.

Finally, the tRNAs adopt different conformations inside the ribosome as compared to tRNAs in solution (Fig. 5.11). By this restriction to a certain region in phase space, free energy is stored in the tRNA conformations (Fig. 5.12), which might be used to overcome barriers between different translocation states and to finally dissociate the tRNA from the ribosome.

Next, the question of how the observed fast transition rates for the large-scale rotations of the 30S head and body are achieved was addressed. To that aim, residue-residue contacts between subunits were extracted from the simulations and subsequently clustered into groups. This clustering captured the intersubunit interaction sites (bridges) known from previous structural work [12] (Fig. 5.13A), and additionally allowed to describe the dynamics of formation and disruption of contacts during the intersubunit rotations. The overall binding enthalpy between the two subunits remains rather constant (Fig. 5.13B). If we assume that, due to entropic compensation and electrostatic shielding, the binding free energies that determine the affinity are proportional to the binding enthalpy, this would mean that the binding free energy is similar in different states of rotation.

The rather constant binding free energy allows the subunits to remain associated despite large-scale rotations and implies low barriers between the states, leading to the observed fast transition rates. The similarity of binding enthalpy for different rotation angles, despite the large mutual displacement of contact clusters which are located far away from the rotation axis, is achieved by a change in contact partners (Fig. 5.13C). In particular, the localized 50S part of clusters 2 and 9 consists of few residues involved in contacts in all states of rotation which glide along the extended 30S part of the clusters thereby switching contact partners (Fig. 5.13D).

The largest rotation angles for 30S head tilting and body rotation were observed in the pre5b state (Fig 5.5). Further, clusters 2 and 9 both strongly interact in this state, thereby stabilizing the large rotation angles (Fig 5.13B). In the same state, tRNA$^{\text{Val}}$ is in contact with protein L5, whose residues form the 50S part of clusters 2 and 9, suggesting that the tRNA-L5-30S interaction network stabilizes large rotations. Indeed, after removal of the tRNAs in additional MD simulations, the binding of the clusters weakened and the angle values decreased. Subsequent cryo-EM experiments showed a lower probability for large body rotation angles for empty ribosomes compared to tRNA-ribosome complexes in the pre states, which is in agreement with the simulations.

The presented approach which combines atomistic crystal structures of the ribosome with cryo-EM data using MD simulations has contributed to the understanding of ribosomal translocation at near-atomic resolution on various timescales: The 100 ns dynamics of ribosome within the individual states,

the estimated rates for transitions between the states in the sub-microsecond time scales, and finally structures of translocation intermediates describing the whole pathway of tRNA translocation through the ribosome. The simulations allowed to characterize driving forces and mechanisms to lower barriers for tRNA translocation and to fine tune the affinity between the two subunits during large-scale rotations. Besides having rationalized the effects of several existing mutations affecting translation and translocation rates, the detailed knowledge of residue-residue contacts will enable directed mutagenesis studies stabilizing and destabilizing interactions in specific states, thus altering translocation kinetics, which will lead to a detailed understanding of the process of tRNA translocation.

Chapter 6

Conclusions and Outlook

6.1 Conclusions

In this thesis, data obtained by X-ray crystallography and cryo-EM were combined with MD simulations to elucidate the dynamics and driving forces of macromolecular complexes. The functions of two complexes, the downregulation of neurotransmitter release by the SNARE complex under oxidative stress and the translocation of tRNAs through the ribosome, were investigated.

In the first part of this thesis, I tested the hypothesis that the oxidation of two cysteines on the SNAP-25B linker and consequent disulfide bond formation shortens the linker sufficiently such that SNARE complex formation is hindered. For this purpose, MD simulations of the SNARE complex with and without the disulfide bond were compared.

In the simulations, the formation of the disulfide bond led to a reduced flexibility of the linker and induced conformational changes at the end of the linker, more than 10 nm away from the disulfide bond. The four SNARE domains comprising the SNARE complex are associated by the formation of central hydrophobic layers. In the simulations including the disulfide bond, kinking of the two SNAP-25B SNARE domains and marked conformational changes in three specific layers were observed. Mutations of residues contributing to these layers have been shown to reduce neurotransmitter release [86], which suggests that the stability of these layers is crucial for SNARE complex formation. These results support the hypothesis that a small chemical change, the disulfide bond formation, leads to a destabilization of the whole macromolecular SNARE complex and thus renders it dysfunctional. The presented mechanism might be a chemomechanical regulation to shut down neurotransmitter release under oxidative stress. Oxida-

tive stress and reduced levels of synaptic transmission have been linked to early stages of neurodegenerative diseases [35], rendering the understanding of this mechanism important for the elucidation of the molecular basis of neurodegenerative diseases.

The second focus of this thesis is the process of spontaneous tRNA translocation through the ribosome which takes place on the order of minutes [26]. Here, initially, structures of the ribosome bound to tRNAs in translocation intermediate states at near-atomic resolution were obtained from refinement of a model based on crystal structures [13, 14] against 3d cryo-EM reconstructions representing pre- and post-translocation states [27]. MD simulations, starting from the refined structures, capture the 100 ns time dynamics and the energetics of the ribosome within the respective states. The small deviation (RMSD< 5 Å) of our late pre-translocation model from a recent crystal structure of a ribosome with highly rotated subunits [15] suggests a similarly high accuracy for the other models of translocation intermediates.

Individual rates for transitions between the different states were estimated from the fluctuations within the 100 ns simulations for the most prominent motions of ribosomal components during translocation. Notably, the large-scale intersubunit rotations and the movement of the L1-stalk were found to be intrinsically rapid. Slow transition rates were only seen for two transitions giving independent support for barriers identified earlier [27]. In contrast to these rapid motions, the motion of the tRNAs was found to be rate-limiting for most transitions between states, suggesting that efficient translocation relies on mechanisms which lower barriers for tRNA motion.

The results suggested several novel mechanisms facilitating translocation by lowering barriers affecting tRNA motion. The L1-stalk moves into the ribosome and forms contacts with the P-site tRNA and subsequently pulls the tRNA to the E site. The ribosomal proteins L5 and L16 of the large subunit interact with the tRNAs according to specific thermodynamic and functional requirements. To precisely position the tRNAs in the A and P site, which is needed for the peptide transferase activity, L16 presents two charged contact patches to the tRNAs. The P-site loop of protein L5 flexibly adapts to the tRNA, allowing the tRNA to "slide" across the P site. These interactions involve the charged backbone of the tRNA, rendering them sequence-independent which allows the ribosome to similarly process different tRNA species. Further, rapid short-timescale diffusion of the bound tRNAs leads to a high attempt frequency increasing the tRNA transition rates. Finally, free energy was found to be stored in the tRNA conformations by restrictions of the phase space accessible to the tRNAs due to ribosomal environment. The release of this conformational free energy could be used as a mechanism to overcome barriers and facilitate tRNA dissociation from the ribosome.

During translocation, the ribosome undergoes large-scale intersubunit rotations. The affinity between the subunits needs to be finely balanced in order to keep the subunits associated and to allow dissociation after termination. Additionally, the affinity needs to be sufficiently constant for different rotational states to avoid the imposition of large free energy barriers for rotation, which would slow down translocation. To address the question of how the affinity is balanced, clusters of residue–residue contacts between the subunits were obtained from the simulation. These independently derived clusters include the previously described groups of contacts (intersubunit bridges) [12]. A complex interplay of the clusters, including formation and rupture of contacts, ensures that the overall binding enthalpy shows only small fluctuations across the different states. In particular, for clusters far away from the pivot points, a high level of contact partner change was observed. The 50S part of clusters 2 and 9 was found to move like a glider on a track, formed by the extended 30S part of these clusters. In this way, strong, balanced interactions are maintained throughout translocation.

A network of interactions between the A-site tRNA, the 50S protein L5 and the 30S protein S13 was seen for the state with the highest 30S body rotation. The hypothesis that this network stabilizes the highly rotated state was supported by MD simulations of the same state after removal of the tRNAs, where the interactions between L5 and S13 weakened and the body rotation angles decreased. In agreement with this hypothesis, subsequent cryo-EM analysis of the empty ribosome showed a decreased probability for high rotation states.

The approach of combining data from X-ray crystallography and cryo-EM with extensive MD simulations applied in this thesis has allowed to elucidate an abundance of phenomena concerning the function of macromolecular complexes.

It has been shown, that small changes, e.g., disulfide bond formation, can affect the stability and consequently the function of macromolecular complexes. This allows the regulation of the activity as a reaction to changes of the (chemical) environment. This regulatory mechanism serves as an example of the necessity to study these systems at atomic detail.

Going beyond simulations based on single structures, we have here for the first time addressed a molecular machine of the size of the ribosome by combining data from X-ray crystallography and cryo-EM with extensive MD simulations. This combination enabled us to obtain a unified picture of the functional dynamics of this molecular machine at very different time scales: The time scale of the entire process described by different states obtained from experiment, here, on the order of seconds and the time scale of dynamics within states which are addressed by MD simulations. The

combined information allows to describe dynamics in terms of Markov states and rates even for large macromolecular complexes.

Using the obtained transition rates, it was shown that the sequence of states purely derived by structural similarity, even when obtained from non-atomic resolution structures, gives a good indication for the most probable pathway during translocation. The good agreement between structural and kinetic sequence can be investigated for other macromolecules applying the presented techniques. This allows to directly conclude a probable kinetic sequence of states from single cryo-EM experiments.

The motions of components of the complex which limit the rate for a certain transition between states were identified on the basis of the obtained individual transition rates. Driving forces and key interactions, which are determinants for efficient function of the complex, were revealed, knowing which motions limit the rates. The mechanisms for lowering free energy barriers and for fine-tuning affinity across the intersubunit surface, identified in this thesis, might have developed similarly in other macromolecular complexes.

Finally, our results suggest a number of experimental tests and, further, suggest new experimental directions. The effects of the possible key interactions between the tRNAs and the ribosome can be evaluated, e.g., with mutagenesis experiments. Mutations of residues involved in contacts which were found to be specific for certain states can be used to either stabilize or destabilize the ribosome in these states. Finally, the obtained near-atomic structures of translocation intermediates allow to (re)interpret FRET efficiency measurements.

6.2 Outlook

Despite the wealth of information obtained in this thesis, the sheer size of the complexes, combined with the slow time scales at which functional processes occur, posed — and still poses — a number of challenges and raises several caveats, the most pressing of which are summarized here.

The first caveat is the limited simulation times of ∼500 ns and 100 ns for each state of the SNARE complex and the ribosome, respectively, is not sufficient to capture the full dynamics within all of the states. Therefore, extended simulations will help to improve the accuracy of the results presented in this thesis. This is especially true for the rough first estimates of the transition rates and the conformational free energy.

Secondly, the 100 ns dynamics were obtained for 13 out of 50 translocation intermediate states resolved by cryo-EM [27]. These states were chosen

94

according to high population (low free energy) and structural diversity to capture as much of the translocation dynamics as feasible. However, including the remaining states will improve the sampling of the free energy landscape for translocation. Further, transition states cannot be resolved by cryo-EM due to small number of images capturing the ribosome in these states. Therefore, other methods are necessary to investigate complete transition pathways.

Further, electrostatic interactions were found to be crucial in the simulations which is in line with the known sensitivity of ribosomal structure and function to the ion concentration of the surrounding medium, most notably of magnesium ions [156–159]. The magnesium ions resolved in the ribosome crystal structure [14] were included in the simulations. The distribution of the remaining ions is not fully equilibrated, due to the long residence time of magnesium ions near RNA molecules which is on the order of μs [160]. Further, their direct effects on the mechanisms identified in the thesis were not analyzed. E.g., since most of the intersubunit contact clusters include nucleotides, these clusters could be additionally stabilized by magnesium binding. Future work will address the role of ions, first, by identifying binding sites near the ribosome structure and, second, by calculating their contribution to the binding enthalpy of tRNAs and ribosomal components as well as of the two ribosomal subunits.

6.2.1 Free energy landscape of translocation

As a first step, the estimation of rates for transitions between translocation intermediates was based on a quasi-harmonic analysis of the fluctuations observed in the simulations. The next step to improve the estimations and to sample transition pathways as well as the free energy landscape along these pathways, two independent approaches are being developed together with Andrea Vaiana and Carsten Kutzner: Gutter dynamics and a method to carry out MD simulations targeting cryo-EM maps. Applying these methods to all pairs of states would allow us to go one step further towards elucidating the complete free energy landscape of spontaneous translocation.

In gutter dynamics, the system is driven from state A towards state B by adding a harmonic potential to the force field, restraining the dynamics of the system along the difference vector between the structures representing states A and B. This allows the system to move freely in all other directions. In the course of the simulation the minimum of the additional potential is shifted along the difference vector from state A to state B. Since states A and B are local free energy minima the trajectory is expected to end up close to the structure representing state B.

A second method to sample a transition pathway from state A to state B with targeted MD would be to drive the system from the cryo-EM map representing state A to the map representing state B. To that aim, a potential which depends on the correlation coefficients of a map generated from the coordinates of the system with the cryo-EM maps of states A and B would be added to the force field.

Since the time scales that can be addressed with MD simulations for large macromolecular complexes are short in comparison to typical transition times, a careful check for convergence of the targeted simulations is necessary. For both methods, the convergence can be checked by moving the system from B towards A which should lead to a similar pathway. Another check is to start an unbiased simulation from the final coordinates to see if they represent a stable state or if the simulation quickly drifts away.

The potential of mean force can be estimated by combining the umbrella sampling simulation method [161] along the obtained transition pathway with weighted histogram analysis methods (WHAM) [162]. The potential of mean force would further improve the estimations of the transition rates and the structural pathways would lead to a more complete picture of translocation.

6.2.2 Conformational free energy of the complete tRNA cycle

In this thesis, a first estimate of the conformational free energy of tRNAs bound to the ribosome is presented. Since the ensembles of conformations adopted inside the ribosome show little overlap with the ensembles of the tRNAs in solution, the estimate based on the fluctuations in the simulations is rough.

Umbrella sampling simulations along the difference vector between the averaged conformation of the tRNA in solution and the bound conformation, will markedly improve the estimate of the stored conformational free energy.

Further, the analysis of the conformational free energy of the tRNAs during translocation will be extended to the complete cycle tRNAs undergo. To this aim, the conformations of the tRNAs within the ternary complex, i.e. bound to EF-Tu and GTP [95], and the conformations of the tRNAs within the ternary complex bound to the ribosome [97] will be taken into account.

The increased accuracy and the extended set of bound conformations will allow us to identify where in the tRNA cycle the tRNAs are loaded with conformational free energy and where this energy is released. This will drastically improve the current knowledge of tRNA kinetics.

6.2.3 Ribosome-selB complex

The combination of approaches here exemplified for the *E.coli* ribosome should of course be applicable to other molecular machines. Although systems that have been characterized both structurally and biochemically to a comparable detail are rare, Marina Rodnina and Holger Stark have recently succeeded in obtaining cryo-EM reconstructions of the ribosome-selB complex.

SelB is a specific translational elongation factor that helps to incorporate the non-standard amino acid selenocysteine which occurs in dozens of proteins called selenoproteins [163]. Selenocysteine is not directly encoded by an mRNA codon. It is encoded by a UGA codon and a characteristic downstream mRNA nucleotide sequence (SECIS element). The selenocysteine is delivered to the ribosome by a specialized tRNA (Sec-tRNA$^{\text{Sec}}$) [164], which is bound selB and GTP, forming a ternary complex.

From MD simulations of the refined structure in different states, key interactions, correlated motions and kinetics will be determined. In particular, this study will allow us to investigate how the interactions of selB with the ribosome differ from those of elongation factor EF-Tu and how the signal of SECIS element recognition by selB domain IV is transfered to the selB GTPase center.

Acknowledgements

the office as well as lots of thoughts, chocolate, and laughs from the first to the last day of PhD studies I want to thank Frank Wiederschein very much. Furthermore, I would like to thank all the current and former members of the department for the fun and productive atmosphere.

Thanks to Evi Heinemann, Ansgar Esztermann, and Martin Fechner for the continuous administrative support and for providing an excellent working environment. Carsten Kutzner helped a lot setting up GROMACS on the supercomputers. Further, I would like to acknowledge the people at GGNB graduate school, especially Antje Erdmann, Michaela Böttcher, and Frauke Bergmann, for their continuous support.

For carefully proofreading the thesis I would like to thank Petra Kellers, Vera Bock, Andrea Vaiana, Dirk Matthes, and Christian Blau.

I want to thank my parents for their constant support, especially in providing child care during long working days. Finally, my deep gratitude goes to Lisa and Mika for making my day, every day.

Bibliography

[1] R. Reed. Mechanisms of fidelity in pre-mrna splicing. *Curr. Opin. Cell. Biol.*, 12(3):340–345, Jun 2000.

[2] V. Ramakrishnan. Ribosome structure and the mechanism of translation. *Cell*, 108(4):557–572, Feb 2002.

[3] M.V. Rodnina, M. Beringer, and W. Wintermeyer. How ribosomes make peptide bonds. *Trends Biochem. Sci.*, 32(1):20–26, 2007.

[4] K.J. Ryan and S.R. Wente. The nuclear pore complex: a protein machine bridging the nucleus and cytoplasm. *Curr. Opin. Cell. Biol.*, 12(3):361–371, Jun 2000.

[5] R. Jahn and R.H. Scheller. Snares–engines for membrane fusion. *Nat. Rev. Mol. Cell Biol.*, 7(9):631–643, Sep 2006.

[6] J.A. Leyva, M.A Bianchet, and L.M. Amzel. Understanding atp synthesis: structure and mechanism of the f1-atpase (review). *Mol. Membr. Biol.*, 20(1):27–33, 2003.

[7] K. Wüthrich. Protein structure determination in solution by nmr spectroscopy. *J. Biol. Chem.*, 265(36):22059–22062, Dec 1990.

[8] I. Vidovic, S. Nottrott, K. Hartmuth, R. Lührmann, and R. Ficner. Crystal structure of the spliceosomal 15.5 kd protein bound to a u4 snrna fragment. *Mol. Cell.*, 6(6):1331–1342, 2000.

[9] C. Gibbons, M. G. Montgomery, A. G. Leslie, and J. E. Walker. The structure of the central stalk in bovine f(1)-atpase at 2.4 a resolution. *Nat. Struct. Biol.*, 7(11):1055–1061, 2000.

[10] S.R. Solmaz, R. Chauhan, G. Blobel, and I. Melcak. Molecular architecture of the transport channel of the nuclear pore complex. *Cell*, 147(3):590–602, 2011.

[11] P. Sampathkumar, T. Gheyi, S.A. Miller, K.T. Bain, M. Dickey, J.B. Bonanno, S.J. Kim, J. Phillips, U. Pieper, J. Fernandez-Martinez, J.D. Franke, A. Martel, H. Tsuruta, S. Atwell, D.A. Thompson, J.S. Emtage, S.R. Wasserman, M.P. Rout, A. Sali, J. M. Sauder, and S.K. Burley. Structure of the c-terminal domain of saccharomyces cerevisiae nup133, a component of the nuclear pore complex. *Proteins*, 79(5):1672–1677, 2011.

[12] M.M. Yusupov, G.Zh. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H.D. Cate, and H.F. Noller. Crystal structure of the ribosome at 5.5 angstrom resolution. *Science*, 292(5518):883–896, 2001.

[13] Y. Gao, M. Selmer, C. Dunham, A. Weixlbaumer, A. Kelley, and V. Ramakrishnan. The structure of the ribosome with elongation factor g trapped in the posttranslocational state. *Science*, 326(5953):694–699, 2009.

[14] W. Zhang, J.A. Dunkle, and J.H.D. Cate. Structures of the ribosome in intermediate states of ratcheting. *Science*, 325(5943):1014, 2009.

[15] J.A. Dunkle, L. Wang, M.B. Feldman, A. Pulk, V.B. Chen, G.J. Kapral, J. Noeske, J.S. Richardson, S.C. Blanchard, and J.H.D. Cate. Structures of the bacterial ribosome in classical and hybrid states of trna binding. *Science*, 332(6032):981, 2011.

[16] R.B. Sutton, D. Fasshauer, R. Jahn, and A.T. Brunger. Crystal structure of a snare complex involved in synaptic exocytosis at 2.4 å resolution. *Nature*, 395(6700):347–354, 1998.

[17] A. Stein, G. Weber, M.C. Wahl, and R. Jahn. Helical extension of the neuronal snare complex into the membrane. *Nature*, 460(7254):525–528, Jul 2009.

[18] J. Frank. Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu. Rev. Biophys. Biomol. Struct.*, 31:303–319, 2002.

[19] B. Sander, M. M. Golas, R. Lührmann, and H. Stark. An approach for de novo structure determination of dynamic molecular assemblies by electron cryomicroscopy. *Structure*, 18(6):667–676, Jun 2010.

[20] D. Boehringer, E.M. Makarov, B. Sander, O.V. Makarova, B. Kastner, R. Lührmann, and H. Stark. Three-dimensional structure of a pre-catalytic human spliceosomal complex b. *Nat. Struct. Mol. Biol.*, 11(5):463–468, May 2004.

[21] M. Azubel, S.G. Wolf, J. Sperling, and R. Sperling. Three-dimensional structure of the native spliceosome by cryo-electron microscopy. *Mol. Cell*, 15(5):833–839, Sep 2004.

[22] D. Stoffler, B. Feja, B. Fahrenkrog, J. Walz, D. Typke, and U. Aebi. Cryo-electron tomography provides novel insights into nuclear pore architecture: implications for nucleocytoplasmic transport. *J. Mol. Biol.*, 328(1):119–130, Apr 2003.

[23] N. Elad, T. Maimon, D. Frenkiel-Krispin, R. Y. H. Lim, and O. Medalia. Structural analysis of the nuclear pore complex by integrated approaches. *Curr. Opin. Struct. Biol.*, 19(2):226–232, Apr 2009.

[24] J. Frank, J. Zhu, P. Penczek, Y. Li, S. Srivastava, A. Verschoor, M. Radermacher, R. Grassucci, R.K. Lata, and R.K. Agrawal. A model of protein synthesis based on cryo-electron microscopy of the e. coli ribosome. *Nature*, 376(6539):441–444, 1995.

[25] J. Frank. Cryo-electron microscopy as an investigative tool: the ribosome as an example. *Bioessays*, 23(8):725–732, Aug 2001.

[26] A.L. Konevega, N. Fischer, Y.P. Semenkov, H. Stark, W. Wintermeyer, and M.V. Rodnina. Spontaneous reverse movement of mrna-bound trna through the ribosome. *Nat. Struct. Mol. Biol.*, 14(4):318–324, 2007.

[27] N. Fischer, A. Konevega, W. Wintermeyer, M. V. Rodnina, and H. Stark. Ribosome dynamics and trna movement by time-resolved electron cryomicroscopy. *Nature*, 466(7304):329–333, 2010.

[28] M. Delarue and P. Dumas. On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models. *Proc. Natl. Acad. Sci. U. S. A.*, 101(18):6957–6962, May 2004.

[29] F. Tama, O. Miyashita, and C.L. Brooks. Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. *J. Mol. Biol.*, 337(4):985–999, Apr 2004.

[30] G.F. Schröder, A.T. Brunger, and M. Levitt. Combining efficient conformational sampling with a deformable elastic network model facilitates structure refinement at low resolution. *Structure*, 15(12):1630 – 1641, 2007.

[31] G.F. Schröder, M. Levitt, and A.T. Brunger. Super-resolution biomolecular crystallography with low-resolution data. *Nature*, 464(7292):1218–1222, 2010.

[32] L.G. Trabuco, E. Villa, K. Mitra, J. Frank, and K. Schulten. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure*, 16(5):673–683, May 2008.

[33] A. Ahmed, P.C. Whitford, K.Y. Sanbonmatsu, and F. Tama. Consensus among flexible fitting approaches improves the interpretation of cryo-em data. *J. Struct. Biol.*, 177(2):561–570, Feb 2012.

[34] A.R. Giniatullin, F. Darios, A. Shakirzyanova, B. Davletov, and R. Giniatullin. Snap25 is a pre-synaptic target for the depressant action of reactive oxygen species on transmitter release. *J. Neurochem.*, 98(6):1789, 2006.

[35] D.J. Keating. Mitochondrial dysfunction, oxidative stress, regulation of exocytosis and their relevance to neurodegenerative diseases. *J. Neurochem.*, 104(2):298–305, Jan 2008.

[36] S. C. Blanchard, H. D. Kim, R. L. Gonzalez, J. D. Puglisi, and S. Chu. trna dynamics on the ribosome during translation. *Proc. Natl. Acad. Sci. U. S. A.*, 101(35):12893–12898, Aug 2004.

[37] M. Valle, A. Zavialov, J. Sengupta, U. Rawa, M. Ehrenberg, and J. Frank. Locking and unlocking of ribosomal motions. *Cell*, 114(1):123–134, 2003.

[38] G.Z. Yusupova, M.M. Yusupov, J.H.D. Cate, H.F. Noller, et al. The path of messenger rna through the ribosome. *Cell*, 106(2):233–242, 2001.

[39] M. Selmer, C. M. Dunham, F.V. Murphy, A. Weixlbaumer, S. Petry, A.C. Kelley, J.R. Weir, and V. Ramakrishnan. Structure of the 70s ribosome complexed with mrna and trna. *Science*, 313(5795):1935–1942, 2006.

[40] A. Korostelev, S. Trakhanov, M. Laurberg, and H.F. Noller. Crystal structure of a 70s ribosome-trna complex reveals functional interactions and rearrangements. *Cell*, 126(6):1065–1077, 2006.

[41] A. Korostelev, S. Trakhanov, H. Asahara, M. Laurberg, L. Lancaster, and H. F. Noller. Interactions and dynamics of the shine–dalgarno helix

in the 70s ribosome. *Proc. Natl. Acad. Sci. U. S. A.*, 104(43):16840, 2007.

[42] R.M. Voorhees, A. Weixlbaumer, D. Loakes, A.C. Kelley, and V. Ramakrishnan. Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70s ribosome. *Nat. Struct. Mol. Biol.*, 16(5):528–533, 2009.

[43] L.B. Jenner, N. Demeshkina, G. Yusupova, and M. Yusupov. Structural aspects of messenger rna reading frame maintenance by the ribosome. *Nat. Struct. Mol. Biol.*, 17(5):555–560, 2010.

[44] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[45] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, and W.L. Jorgensen. Evaluation and reparametrization of the opls-aa force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B*, 105(28):6474–6487, 2001.

[46] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins*, 65(3):712–725, 2006.

[47] L. Verlet. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Phys. Rev.*, 159:98–103, 1967.

[48] R.W. Hockney, S.P. Goel, and J.W. Eastwood. Quiet high-resolution computer models of a plasma. *J. Comput. Phys.*, 14(2):148 – 158, 1974.

[49] A.K. Mazur. Common molecular dynamics algorithms revisited: Accuracy and optimal time steps of störmer-leapfrog integrators. *J. Comput. Phys.*, 136(2):354 – 365, 1997.

[50] M. Christen, C.D. Christ, and W.F. van Gunsteren. Free energy calculations using flexible-constrained, hard-constrained and non-constrained molecular dynamics simulations. *ChemPhysChem*, 8:1557–1564, 2007.

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

[52] S. Miyamoto and P. A. Kollman. Settle: An analytical version of the shake and rattle algorithm for rigid water models. *J. Comput. Chem.*, 13:952–962, 1992.

[53] K.A. Feenstra, B. Hess, and H.J.C. Berendsen. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.*, 20:786–798, 1999.

[54] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, and J.R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81:3684–3690, October 1984.

[55] T. Morishita. Fluctuation formulas in molecular-dynamics simulations with the weak coupling heat bath. *J. Chem. Phys.*, 113:2976–2982, 2000.

[56] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126:014101, 2007.

[57] A. Kitao, F. Hirata, and N. Go. The effects of solvent on the conformation and the collective motions of proteins - normal mode analysis and molecular-dynamics simulations of melittin in water and vacuum. *Chem. Phys.*, 158:447–472, 1991.

[58] B.S. Schuwirth, M.A. Borovinskaya, C.W. Hau, W. Zhang, A. Vila-Sanjurjo, J.M. Holton, and J.H.D. Cate. Structures of the bacterial ribosome at 3.5 a resolution. *Science*, 310(5749):827–834, Nov 2005.

[59] L. V. Bock, B. H., H. Grubmüller, and D. J. Woodbury. Chemomechanical regulation of snare proteins studied with molecular dynamics simulations. *Biophys. J.*, 99(4):1221–1230, Aug 2010.

[60] S. Martens and H.T. McMahon. Mechanisms of membrane fusion: disparate players and common principles. *Nat. Rev. Mol. Cell. Biol.*, 9(7):543–556, Jul 2008.

[61] C. Bark, F.P. Bellinger, A. Kaushal, J.R. Mathews, L.D. Partridge, and M.C. Wilson. Developmentally regulated switch in alternatively spliced snap-25 isoforms alters facilitation of synaptic transmission. *J. Neurosci.*, 24(40):8796–8805, 2004.

[62] D. Fasshauer, R.B. Sutton, A.T. Brunger, and R. Jahn. Conserved structural features of the synaptic fusion complex: Snare proteins reclassified as q-and r-snares. *Proc. Natl. Acad. Sci. U. S. A.*, 95(26):15781, 1998.

[63] K.M.S. Misura, L.C. Gonzalez, A.P. May, R.H. Scheller, and W.I. Weis. Crystal structure and biophysical properties of a complex between the n-terminal snare region of snap25 and syntaxin 1a. *J. Biol. Chem.*, 276(44):41301–41309, 2001.

[64] L. Wang, M.A. Bittner, D. Axelrod, and R.W. Holz. The structural and functional implications of linked snare motifs in snap25. *Mol. Biol. Cell*, 19(9):3944, 2008.

[65] S. Gonzalo, W.K. Greentree, and M.E. Linder. Snap-25 is targeted to the plasma membrane through a novel membrane-binding domain. *J. Biol. Chem.*, 274(30):21313–21318, 1999.

[66] M. Veit, T.H. Söllner, and J.E. Rothman. Multiple palmitoylation of synaptotagmin and the t-snare snap-25. *FEBS Lett.*, 385(1-2):119–123, 1996.

[67] C. Gonelle-Gispert, M. Molinete, P.A. Halban, and K. Sadoul. Membrane localization and biological activity of snap-25 cysteine mutants in insulin-secreting cells. *J. Cell. Sci.*, 113(18):3197–3205, 2000.

[68] K. Vogel, J.P. Cabaniols, and P.A. Roche. Targeting of snap-25 to membranes is mediated by its association with the target snare syntaxin. *J. Biol. Chem.*, 275(4):2959–2965, 2000.

[69] S.S. Loranger and M.E. Linder. Snap-25 traffics to the plasma membrane by a syntaxin-independent mechanism. *J. Biol. Chem.*, 277(37):34303–34309, 2002.

[70] Q. Huang, X. Hong, and Q. Hao. Snap-25 is also an iron–sulfur protein. *FEBS Lett.*, 582(10):1431–1436, 2008.

[71] B. Pallavi and R. Nagaraj. Palmitoylated peptides from the cysteine-rich domain of snap-23 cause membrane fusion depending on peptide length, position of cysteines, and extent of palmitoylation. *J. Biol. Chem.*, 278(15):12737–12744, 2003.

[72] D.A. Butterfield, M. Perluigi, and R. Sultana. Oxidative stress in alzheimer's disease brain: new insights from redox proteomics. *Eur. J. Pharmacol.*, 545(1):39–50, 2006.

[73] M. Facheris, S. Beretta, and C. Ferrarese. Peripheral markers of oxidative stress and excitotoxicity in neurodegenerative disorders: tools for diagnosis and therapy? *J. Alzheimers Dis.*, 6(2):177–184, 2004.

[74] Y.M.W. Janssen-Heininger, B.T. Mossman, N.H. Heintz, H.J. Forman, B. Kalyanaraman, T. Finkel, J.S. Stamler, S.G. Rhee, and A. van der Vliet. Redox-based regulation of signal transduction: principles, pitfalls, and promises. *Free Radic. Biol. Med.*, 45(1):1–17, 2008.

[75] P. Washbourne, V. Cansino, J.R. Mathews, M. Graham, R.D. Burgoyne, and M.C. Wilson. Cysteine residues of snap-25 are required for snare disassembly and exocytosis, but not for membrane targeting. *Biochem. J.*, 357(Pt 3):625, 2001.

[76] T. Finkel. Oxygen radicals and signaling. *Curr. Opin. Cell Biol.*, 10(2):248–253, 1998.

[77] A.R. Giniatullin, S.N. Grishin, E.R. Sharifullina, A.M. Petrov, A.L. Zefirov, and R.A. Giniatullin. Reactive oxygen species contribute to the presynaptic action of extracellular atp at the frog neuromuscular junction. *J. Physiol.*, 565(1):229–242, 2005.

[78] B. Brooks, R. Bruccoleri, B. Olafson, D States, S. Swaminathan, and M. Karplus. Charmm: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4(2):187–217, 1983.

[79] B.R. Brooks, C.L. Brooks III, A.D. Mackerell Jr, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, et al. Charmm: the biomolecular simulation program. *J. Comput. Chem.*, 30(10):1545–1614, 2009.

[80] A. Stein, G. Weber, M.C. Wahl, and R. Jahn. Helical extension of the neuronal snare complex into the membrane. *Nature*, 460(7254):525–528, 2009.

[81] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory. Comput.*, 4(3):435–447, 2008.

[82] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926, 1983.

[83] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamicsmethod. *J. Appl. Phys.*, 52:7182, 1981.

[84] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen. A smooth particle mesh ewald method. *J. Chem. Phys.*, 103:8577–8593, 1995.

[85] W. Kabsch and C. Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12), 1983.

[86] J.B. Sørensen, K. Wiederhold, E.M. Müller, I. Milosevic, G. Nagy, B.L. de Groot, H. Grubmüller, and D. Fasshauer. Sequential n- to c-terminal snare complex assembly drives priming and fusion of secretory vesicles. *EMBO J.*, 25(5):955–966, 2006.

[87] S.R. Lane and Y. Liu. Characterization of the palmitoylation domain of snap-25. *J. Neurochem.*, 69(5):1864, 1997.

[88] G.R. Prescott, O.A. Gorleku, J. Greaves, and L.H. Chamberlain. Palmitoylation of the synaptic vesicle fusion machinery. *J. Neurochem.*, 110(4):1135–1149, 2009.

[89] R. Kang, R. Swayze, M.F. Lise, K. Gerrow, A. Mullard, W.G. Honer, and A. El-Husseini. Presynaptic trafficking of synaptotagmin i is regulated by protein palmitoylation. *J, Biol. Chem.*, 279(48):50524, 2004.

[90] R. Kang, J. Wan, P. Arstikaitis, H. Takahashi, K. Huang, A.O. Bailey, J.X. Thompson, A.F. Roth, R.C. Drisdel, R. Mastro, et al. Neural palmitoyl-proteomics reveals dynamic synaptic palmitoylation. *Nature*, 456(7224):904–909, 2008.

[91] K. Vogel and P.A. Roche. Snap-23 and snap-25 are palmitoylated in vivo. *Biochem. Biophys. Res. Commun.*, 258(2):407–410, 1999.

[92] F Crick. On degenerate templates and the adaptor hypothesis: A note for the rna tie club, 1956.

[93] S.H. Kim, G.J. Quigley, F.L. Suddath, A. McPherson, D. Sneden, J.J. Kim, J. Weinzierl, and A. Rich. Three-dimensional structure of yeast phenylalanine transfer rna: folding of the polynucleotide chain. *Science*, 179(4070):285, 1973.

[94] H.S. Zaher and R. Green. Fidelity at the molecular level: lessons from protein synthesis. *Cell*, 136(4):746–762, 2009.

[95] P. Nissen, M. Kjeldgaard, S. Thirup, G. Polekhina, L. Reshetnikova, B.F. Clark, and J. Nyborg. Crystal structure of the ternary complex of phe-trnaphe, ef-tu, and a gtp analog. *Science*, 270(5241):1464–1472, Dec 1995.

[96] R.M. Voorhees, T.M. Schmeing, A.C. Kelley, and V. Ramakrishnan. The mechanism for activation of gtp hydrolysis on the ribosome. *Science*, 330(6005):835–838, 2010.

[97] T.M. Schmeing, R.M. Voorhees, A.C. Kelley, Y. Gao, F.V. Murphy, .R. Weir, and V. Ramakrishnan. The crystal structure of the ribosome bound to ef-tu and aminoacyl-trna. *Science*, 326(5953):688–694, Oct 2009.

[98] D. Moazed and H. F. Noller. Intermediate states in the movement of transfer rna in the ribosome. *Nature*, 342(6246):142–148, Nov 1989.

[99] M.V. Rodnina, A. Savelsbergh, V.I. Katunin, and W. Wintermeyer. Hydrolysis of gtp by elongation factor g drives trna movement on the ribosome. *Nature*, 385(6611):37–41, Jan 1997.

[100] R.K. Agrawal, A.B. Heagle, P. Penczek, R.A. Grassucci, and J. Frank. Ef-g-dependent gtp hydrolysis induces translocation accompanied by large conformational changes in the 70s ribosome. *Nat. Struct. Biol.*, 6(7):643–647, Jul 1999.

[101] S. Shoji, S. Walker, and K. Fredrick. Reverse translocation of trna in the ribosome. *Mol. Cell*, 24(6):931–942, 2006.

[102] V. Berk, W. Zhang, R.D. Pai, and J.H.D. Cate. Structural basis for mrna and trna positioning on the ribosome. *Proc. Natl. Acad. Sci. U. S. A.*, 103(43):15830–15834, 2006.

[103] A. Vila-Sanjurjo, B.S. Schuwirth, C.W. Hau, and J.H.D. Cate. Structural basis for the control of translation initiation during stress. *Nat. Struct. Mol. Biol.*, 11(11):1054–1059, 2004.

[104] S. Petry, D.E. Brodersen, F.V. Murphy IV, C.M. Dunham, M. Selmer, M.J. Tarry, A.C. Kelley, and V. Ramakrishnan. Crystal structures of the ribosome in complex with release factors rf1 and rf2 bound to a cognate stop codon. *Cell*, 123(7):1255–1266, 2005.

[105] A. Weixlbaumer, S. Petry, C.M. Dunham, M. Selmer, A.C. Kelley, and V. Ramakrishnan. Crystal structure of the ribosome recycling factor bound to the ribosome. *Nat. Struct. Mol. Biol.*, 14(8):733–737, 2007.

[106] M. Laurberg, H. Asahara, A. Korostelev, J. Zhu, S. Trakhanov, and H.F. Noller. Structural basis for translation termination on the 70s ribosome. *Nature*, 454(7206):852–857, 2008.

[107] R. Bingel-Erlenmeyer, R. Kohler, G. Kramer, S. Arzu Sandikci, T. Maier, C. Schaffitzel, B. Wiedmann, and N.B. Bernd Bukau. A peptide deformylase–ribosome complex reveals mechanism of nascent chain processing. *Nature*, 452(7183):108–111, 2008.

[108] A. Weixlbaumer, H. Jin, C. Neubauer, R.M. Voorhees, S. Petry, A.C. Kelley, and V. Ramakrishnan. Insights into translational termination from the structure of rf2 bound to the ribosome. *Science*, 322(5903):953–956, 2008.

[109] G. Blaha, R.E. Stanley, and T.A. Steitz. Formation of the first peptide bond: the structure of ef-p bound to the 70s ribosome. *Science*, 325(5943):966–970, 2009.

[110] A. Korostelev, J. Zhu, H. Asahara, and H.F. Noller. Recognition of the amber uag stop codon by release factor rf1. *EMBO Journal*, 29(15):2577–2585, 2010.

[111] H. Jin, A.C. Kelley, D. Loakes, and V. Ramakrishnan. Structure of the 70s ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. *Proc. Natl. Acad. Sci. U. S. A.*, 107(19):8593, 2010.

[112] H. Jin, A.C. Kelley, and V. Ramakrishnan. Crystal structure of the hybrid state of ribosome in complex with the guanosine triphosphatase release factor 3. *Proc. Natl. Acad. Sci. U. S. A.*, 108(38):15798–15803, 2011.

[113] J. Zhou, L. Lancaster, S. Trakhanov, and H.F. Noller. Crystal structure of release factor rf3 trapped in the gtp state on a rotated conformation of the ribosome. *RNA*, 18(2):230–240, 2012.

[114] K.Y. Sanbonmatsu. Computational studies of molecular machines: the ribosome. *Curr. Opin. Struct. Biol.*, 22(2):168–174, Apr 2012.

[115] K.Y. Sanbonmatsu, S. Joseph, and C.S.s Tung. Simulating movement of trna into the ribosome during decoding. *Proc. Natl. Acad. Sci. U. S. A.*, 102(44):15854–15859, Nov 1 2005.

[116] P.C. Whitford, P. Geggier, R. Altman, S.C. Blanchard, J. Onuchic, and K. Sanbonmatsu. Accommodation of aminoacyl-trna into the ribosome involves reversible excursions along multiple pathways. *RNA*, 16(6):1196–1204, 2010.

[117] P.C. Whitford, J.N. Onuchic, and K.Y. Sanbonmatsu. Connecting energy landscapes with experimental rates for aminoacyl-trna accommodation in the ribosome. *J. Am. Chem. Soc.*, 132(38):13170–13171, Sep 2010.

[118] L.G. Trabuco, C.B. Harrison, E. Schreiner, and K. Schulten. Recognition of the regulatory nascent chain tnac by the ribosome. *Structure*, 18(5):627–637, May 2010.

[119] L.G. Trabuco, E. Schreiner, J. Eargle, P. Cornish, T. Ha, Z. Luthey-Schulten, and K. Schulten. The role of l1 stalk-trna interaction in the ribosome elongation cycle. *J. Mol. Biol.*, 402(4):741–760, 2010.

[120] H.J.C. Berendsen, J.R. Grigera, and T.P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91(24):6269–6271, November 1987.

[121] I.S. Joung and T.E. Cheatham. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B*, 112(30):9020–9041, July 2008.

[122] B. Hess. P-lincs: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.*, 4:116–122, 2008.

[123] G. Vriend. What if: a molecular modeling and drug design program. *J. Mol. Graphics*, 8(1):52–56, 1990.

[124] F. Jossinet and E. Westhof. Sequence to structure (s2s): display, manipulate and interconnect rna data from sequence to structure. *Bioinformatics*, 21(15):3320, 2005.

[125] C.A. Theimer, L.D. Finger, L. Trantirek, and J. Feigon. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase rna. *Proc. Natl. Acad. Sci. U. S. A.*, 100(2):449–454, Jan 2003.

[126] A. Grishaev, J. Ying, M.D. Canny, A. Pardi, and A. Bax. Solution structure of trna val from refinement of homology model against residual dipolar coupling and saxs data. *J. Biomolecular NMR*, 42(2):99–109, 2008.

[127] G. Yusupova, L. Jenner, B. Rees, D. Moras, and M. Yusupov. Structural basis for messenger rna movement on the ribosome. *Nature*, 444(7117):391–394, 2006.

[128] J. Wang, W. Wang, P.A. Kollman, and D.A. Case. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model*, 25(2):247–260, 2006.

[129] B.L. de Groot, D.M.F. van Aalten, R.M. Scheek, A. Amadei, G. Vriend, and H.J.C. Berendsen. Prediction of protein conformational freedom from distance constraints. *Proteins: Struct. Funct. Genet.*, 29:240–251, 1997.

[130] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, and T.E. Ferrin. Ucsf chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25(13):1605–1612, 2004.

[131] J.R. Bellare, HT Davis, L.E. Scriven, and Y. Talmon. Controlled environment vitrification system: an improved sample preparation technique. *J. Elect. Micr. Tech.*, 10(1):87–111, 1988.

[132] S.J. Ludtke, P.R. Baldwin, and W. Chiu. Eman: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.*, 128(1):82–97, 1999.

[133] B. Sander, M.M. Golas, and H. Stark. Automatic ctf correction for single particles based upon multivariate statistical analysis of individual power spectra. *J. Struct. Biol.*, 142(3):392–401, 2003.

[134] M. van Heel, G. Harauz, E.V. Orlova, R. Schmidt, M. Schatz, et al. A new generation of the imagic image processing system. *J. Struct. Biol.*, 116:17–24, 1996.

[135] B. Sander, M.M. Golas, and H. Stark. Corrim-based alignment for improved speed in single-particle image processing. *J. Struct. Biol.*, 143(3):219–228, 2003.

[136] A. Amadei, A.B. Linssen, and H.J.C. Berendsen. Essential dynamics of proteins. *Proteins*, 17(4):412–425, Dec 1993.

[137] K. Halvorsen, M. Lesser, and A. Lundberg. A new method for estimating the axis of rotation and the center of rotation. *J. Biomech,*, 32(11):1221 – 1227, 1999.

[138] C. Gramkow. On averaging rotations. *Int. J. Comput. Vision*, 42:7–16, 2001. 10.1023/A:1011129215388.

[139] R. Brandman, Y. Brandman, and V.S. Pande. A-site residues move independently from p-site residues in all-atom molecular dynamics simulations of the 70s bacterial ribosome. *PLoS One*, 7(1):e29377, 2012.

[140] B.L. de Groot, A. Engel, and H. Grubmüller. The structure of the aquaporin-1 water channel: a comparison between cryo-electron microscopy and x-ray crystallography. *J. Mol. Biol.*, 325(3):485–493, 2003.

[141] A. Hirashima and A. Kaji. Factor-dependent release of ribosomes from messenger rna. requirement for two heat-stable factors. *J. Mol. Biol.*, 65(1):43–58, Mar 1972.

[142] G. Hirokawa, N. Demeshkina, N. Iwakura, H. Kaji, and A. Kaji. The ribosome-recycling step: consensus or controversy? *Trends Biochem. Sci.*, 31(3):143–149, Mar 2006.

[143] X. Agirrezabala, J. Lei, J. Brunelle, R. Ortiz-Meoz, R. Green, and J. Frank. Visualization of the hybrid state of trna binding promoted by spontaneous ratcheting of the ribosome. *Mol. Cell*, 32(2):190–197, 2008.

[144] P. Julian, A. Konevega, S. Scheres, M. Lazaro, D. Gil, and W. Wintermeyer. Structure of ratcheted ribosomes with trnas in hybrid states. *Proc. Natl. Acad. Sci. U. S. A.*, 105(44):16924–16927, 2008.

[145] P. Cornish, D. Ermolenko, D. Staple, L. Hoang, R. Hickerson, and H.F. Noller. Following movement of the l1 stalk between three functional states in single ribosomes. *Proc. Natl. Acad. Sci. U. S. A.*, 106(8):2571–2576, 2009.

[146] J. Fei, P. Kosuri, D. MacDougall, and R. Gonzalez. Coupling of ribosomal l1 stalk and trna dynamics during translation elongation. *Mol. Cell*, 30(3):348–359, 2008.

[147] J.B. Munro, R.B. Altman, C.S. Tung, J.H.D. Cate, K.Y. Sanbonmatsu, and S.C. Blanchard. Spontaneous formation of the unlocked state of the ribosome is a multistep process. *Proc. Natl. Acad. Sci. U. S. A.*, 107(2):709–714, 2010.

[148] A.R. Subramanian and E.R. Dabbs. Functional studies on ribosomes lacking protein l1 from mutant escherichia coli. *Eur. J. Biochem.*, 112(2):425–430, Nov 1980.

[149] J.B. Munro, R.B. Altman, N. O'Connor, and S.C. Blanchard. Identification of two distinct hybrid state intermediates on the ribosome. *Mol. Cell*, 25(4):505–517, 2007.

[150] M.H.J. Rhodin and J.D. Dinman. A flexible loop in yeast ribosomal protein l11 coordinates p-site trna binding. *Nucleic Acids Research*, 38(22):8377, 2010.

[151] F. Vendeix, A. Munoz, and P. Agris. Free energy calculation of modified base-pair formation in explicit solvent: A predictive model. *RNA*, 15(12):2278–2287, 2009.

[152] H. Gao, J. Sengupta, M. Valle, A. Korostelev, N. Eswar, S.M. Stagg, P. Van Roey, R.K. Agrawal, S.C. Harvey, A. Sali, M.S. Chapman, and J. Frank. Study of the structural dynamics of the e. coli 70s ribosome using real-space refinement. *Cell*, 113(6):789 – 801, 2003.

[153] K. Kipper, C. Hetényi, S. Sild, J. Remme, and A. Liiv. Ribosomal intersubunit bridge b2a is involved in factor-dependent translation initiation and translational processivity. *J. Mol. Biol.*, 385(2):405–422, Jan 2009.

[154] A.R. Cukras and R. Green. Multiple effects of s13 in modulating the strength of intersubunit interactions in the ribosome during translation. *J. Mol. Biol.*, 349(1):47–59, May 2005.

[155] M.H.J. Rhodin and J.D. Dinman. An extensive network of information flow through the b1b/c intersubunit bridge of the yeast ribosome. *PLoS One*, 6(5):e20048, 2011.

[156] F.C. Chao. Dissociation of macromolecular ribonucleoprotein of yeast. *Arch. Biochem. Biophys.*, 70(2):426–431, Aug 1957.

[157] B.J. McCarthy. The effects of magnesium starvation on the ribosome content of *Escherichia coli*. *Biochim. Biophys. Acta*, 55(6):880–889, 1962.

[158] B.W. Kimes and D.R. Morris. Cations and ribosome structure. ii. effects on the 50s subunit of substituting polyamines for magnesium ion. *Biochemistry*, 12(3):442–449, Jan 1973.

[159] D.J. Klein, P.B. Moore, and T.A. Steitz. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA*, 10(9):1366–1379, Sep 2004.

[160] Y. Hashem and P. Auffinger. A short guide for molecular dynamics simulations of rna systems. *Methods*, 47(3):187–197, Mar 2009.

[161] G.M. Torrie and J.P. Valleau. Monte carlo free energy estimates using non-boltzmann sampling: Application to the sub-critical lennard-jones fluid. *Chem. Phy. Lett.*, 28(4):578–581, 1974.

[162] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, and P.A. Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comput. Chem.*, 13(8):1011–1021, 1992.

[163] G.V. Kryukov and V.N. Gladyshev. The prokaryotic selenoproteome. *EMBO Rep.*, 5(5):538–543, May 2004.

[164] C. Baron and A. Böck. The length of the aminoacyl-acceptor stem of the selenocysteine-specific trna(sec) of escherichia coli is the determinant for binding to elongation factors selb or tu. *J. Biol. Chem.*, 266(30):20375–20379, Oct 1991.