From The Royal Institute of Technology,
School of Technology and Health

SINGLE-PARTICLE CRYO-
ELECTRON MICROSCOPY
OF MACROMOLECULAR
ASSEMBLIES

Kimberley Cheng

Stockholm 2009
To my family
ABSTRACT

In this thesis, single-particle cryo-electron microscopy (cryo-EM) was used to study the structure of three macromolecular assemblies: the two hemocyanin isoforms from *Rapana thomasiiana*, the *Pyrococcus furiosus* chaperonin, and the ribosome from *Escherichia coli*.

Hemocyanins are large respiratory proteins in arthropods and molluscs. Most molluscan hemocyanins exist as two distinct isoforms composed of related polypeptides. In most species the two isoforms differ in terms of their oligomeric stability, and thus we set out to investigate the two *Rapana thomasiiana* hemocyanins (RtH) in order to explain this behaviour. Our findings showed that the two RtH isoforms are identical at the experimental resolution. Furthermore, three previously unreported connections that most likely contribute to the oligomeric stability were identified.

Chaperonins are double-ring protein complexes that assist the folding process of nascent, non-native polypeptide chains. The chaperonin from the hyperthermophilic archaea *Pyrococcus furiosus* belongs to Group II chaperonins, and unlike most other group II chaperonins it appears to be homo-oligomeric. The 3D reconstruction of the *Pyrococcus furiosus* chaperonin revealed a di-octameric structure in a partially closed/open state, something in between the closed folding-active state and the open substrate-accepting state.

The ribosome is the molecular machine where protein synthesis takes place. In bacteria there is a unique RNA molecule called transfer-messenger RNA (tmRNA) that together with its helper protein SmpB rescues ribosomes trapped on defective messenger RNAs (mRNAs) through a process called trans-translation. tmRNA is about 4 times the size of a normal tRNA, and it is composed of a tRNA-like domain (TLD) that is connected to the mRNA-like domain (MLD) by several pseudoknots (PKs) and RNA helices. During trans-translation, tmRNA utilize its TLD to receive the incomplete polypeptide from the peptidyl-tRNA in the ribosomal P site of the stalled ribosome. Subsequently, its MLD is used to tag the incomplete polypeptide with a degradation signal. When tmRNA enters a stalled ribosome the MLD and pseudoknots form a highly structured arc that encircles the beak of the small ribosomal subunit. By utilizing maximum-likelihood based methods for heterogeneity analysis we could observe the *Escherichia coli* ribosome in a number of different tmRNA·SmpB-bound states. The cryo-EM map of the post-accommodated state revealed that the TLD·SmpB part of the tmRNA·SmpB complex mimics native tRNAs in the A site of stalled ribosomes. The density map also showed that the tmRNA arc remains well structured and that it is still attached to the beak of the small ribosomal subunit. The reconstructions of the double-translocation tmRNA-bound ribosome complex showed that the pseudoknots of tmRNA still form an arc, and that they are located at positions similar to the ones assigned for the pseudoknots in the post-accommodated state. In addition, the tmRNA arc exists in two states; one stable and highly structured and another more flexible and disorganized.
LIST OF PUBLICATIONS


## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Hemocyanin – An oxygen transporter</td>
<td>2</td>
</tr>
<tr>
<td>Structure of molluscan hemocyanins</td>
<td>2</td>
</tr>
<tr>
<td>Medical Significance</td>
<td>3</td>
</tr>
<tr>
<td><em>Rapana thomasiastana</em> hemocyanin</td>
<td>3</td>
</tr>
<tr>
<td>Chaperonin – protein folding assistant</td>
<td>6</td>
</tr>
<tr>
<td>Structure and mechanism</td>
<td>6</td>
</tr>
<tr>
<td>Ribosome – The workbench of translation</td>
<td>8</td>
</tr>
<tr>
<td>The message – mRNA</td>
<td>8</td>
</tr>
<tr>
<td>The adaptor molecule - tRNA</td>
<td>9</td>
</tr>
<tr>
<td>The protein-synthesizing machine – The ribosome</td>
<td>10</td>
</tr>
<tr>
<td>The translation process</td>
<td>12</td>
</tr>
<tr>
<td>Initiation</td>
<td>12</td>
</tr>
<tr>
<td>Elongation</td>
<td>12</td>
</tr>
<tr>
<td>Termination</td>
<td>14</td>
</tr>
<tr>
<td>Trans-Translation</td>
<td>14</td>
</tr>
<tr>
<td>tmRNA and SmpB: structure and function</td>
<td>15</td>
</tr>
<tr>
<td>The mechanism of trans-translation</td>
<td>17</td>
</tr>
<tr>
<td>Physiological significance of trans-translation</td>
<td>18</td>
</tr>
<tr>
<td>Structural studies of tmRNA·ribosome complexes</td>
<td>18</td>
</tr>
<tr>
<td>Single-Particle Cryo-Electron Microscopy</td>
<td>22</td>
</tr>
<tr>
<td>Specimen preparation</td>
<td>22</td>
</tr>
<tr>
<td>Negative Staining</td>
<td>22</td>
</tr>
<tr>
<td>Vitrification</td>
<td>22</td>
</tr>
<tr>
<td>Data Processing</td>
<td>23</td>
</tr>
<tr>
<td>Heterogeneity analysis</td>
<td>25</td>
</tr>
<tr>
<td>Conclusions</td>
<td>29</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>30</td>
</tr>
<tr>
<td>References</td>
<td>32</td>
</tr>
</tbody>
</table>
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D</td>
<td>Three-dimensional</td>
</tr>
<tr>
<td>5S RNA</td>
<td>Small RNA molecule in 50S</td>
</tr>
<tr>
<td>16S RNA</td>
<td>RNA molecule in 30S</td>
</tr>
<tr>
<td>23S RNA</td>
<td>Large RNA molecule in 50S</td>
</tr>
<tr>
<td>30S</td>
<td>Small subunit in 70S</td>
</tr>
<tr>
<td>50S</td>
<td>Large subunit in 70S</td>
</tr>
<tr>
<td>70S</td>
<td>Bacterial ribosome</td>
</tr>
<tr>
<td>aaRS</td>
<td>Aminoacyl-tRNA synthetase</td>
</tr>
<tr>
<td>Ala-tmRNA</td>
<td>tmRNA charged with alanine</td>
</tr>
<tr>
<td>ASL</td>
<td>Anticodon stem-loop</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>ATPase</td>
<td>Enzyme hydrolyzing ATP</td>
</tr>
<tr>
<td>C&lt;sub&gt;s&lt;/sub&gt;</td>
<td>Spherical aberration of the EM lens system</td>
</tr>
<tr>
<td>CCA</td>
<td>Nucleotide sequence at the 3'-end of all tRNA molecules</td>
</tr>
<tr>
<td>CCT</td>
<td>Eukaryotic chaperonin (Chaperonin containing TCP-1)</td>
</tr>
<tr>
<td>Cryo-EM</td>
<td>Cryo-electron microscopy</td>
</tr>
<tr>
<td>CTF</td>
<td>Contrast transfer function</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EF-G</td>
<td>Elongation factor G</td>
</tr>
<tr>
<td>EF-Tu</td>
<td>Elongation factor Tu</td>
</tr>
<tr>
<td>EM</td>
<td>Electron microscopy</td>
</tr>
<tr>
<td>FU</td>
<td>Functional unit</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanosine diphosphate</td>
</tr>
<tr>
<td>GroEL</td>
<td>Group I chaperonin in <em>E. coli</em></td>
</tr>
<tr>
<td>GroES</td>
<td>Co-protein to GroEL</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine triphosphate</td>
</tr>
<tr>
<td>GTPase</td>
<td>Enzyme hydrolyzing GTP</td>
</tr>
<tr>
<td><em>H. tuberculata</em></td>
<td><em>Haliotis tuberculata</em></td>
</tr>
<tr>
<td>HtH</td>
<td><em>Haliotis tuberculata</em> hemocyanin</td>
</tr>
<tr>
<td>IF</td>
<td>Initiation factor</td>
</tr>
<tr>
<td>KLH</td>
<td>Keyhole limpet hemocyanin</td>
</tr>
<tr>
<td><em>M. crenulata</em></td>
<td><em>Megathura crenulata</em></td>
</tr>
<tr>
<td>ML</td>
<td>Maximum-likelihood</td>
</tr>
<tr>
<td>MLD</td>
<td>mRNA-like domain</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NpH</td>
<td><em>Nautilus pompilius</em> hemocyanin</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td><em>P. furiosus</em></td>
<td><em>Pyrococcus furiosus</em></td>
</tr>
<tr>
<td>Pf/Cpn</td>
<td>Chaperonin from <em>P. furiosus</em></td>
</tr>
<tr>
<td>PK</td>
<td>Pseudoknot</td>
</tr>
<tr>
<td>PTC</td>
<td>Peptidyl transferase center</td>
</tr>
<tr>
<td>RF</td>
<td>Release factor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Enzyme degrading RNA</td>
</tr>
<tr>
<td>RRF</td>
<td>Recycling factor</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>RtH</td>
<td><em>Rapana thomasiana</em> hemocyanin</td>
</tr>
<tr>
<td>SmpB</td>
<td>Small protein B</td>
</tr>
<tr>
<td><em>smpB</em></td>
<td>Gene encoding SmpB</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td><em>ssrA</em></td>
<td>Gene encoding tmRNA</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscopy</td>
</tr>
<tr>
<td>TLD</td>
<td>tRNA-like domain</td>
</tr>
<tr>
<td>tmRNA</td>
<td>Transfer-messenger RNA</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
</tbody>
</table>
INTRODUCTION

Many tasks in the living cell are performed by proteins. Since structure and function are closely correlated, proteins must obtain a specific structure to function properly. While many proteins function by themselves, some are multiply used to form large macromolecular complexes resulting in an incredible large diversity of shapes and sizes of these assemblies [1]. Similar to individual proteins, macromolecules are not static in their native state. Instead they undergo functionally important conformational changes, often triggered by ligand binding, in order to perform their specific biological activity. Take hemoglobin for instance. Hemoglobin is a respiratory protein whose function is to transport oxygen in the cells. It is composed of four protein molecules (called subunits in the context of the complex). When hemoglobin go between the oxygenated and de-oxygenated state it “breathes”, i.e. the subunits move in relation to each other. This allows the oxygen-binding sites to communicate with each other to increase or decrease the affinity towards oxygen. Thus, by performing structural studies of a macromolecule complex in its different functional states, and characterizing the differences, a deeper insight into its biological process can be obtained.

X-ray crystallography and Nuclear Magnetic Resonance (NMR) are the traditional methods used to determine the three-dimensional (3D) structure of proteins at atomic resolution. NMR has the advantage of analyzing proteins in solution, thus allowing the protein to be in all its functional states. However, this technique is limited to studying proteins smaller than 50kDa. X-ray crystallography on the other hand has been an extremely useful tool for determining the structure of large individual proteins, but it runs into difficulties with large multi-component complexes. In addition, to form crystals the protein complex needs to be in one conformational state only. In many cases this is not easy to achieve, especially when the complex has a flexible region. Thus, often formation of crystals only takes place when the flexible part has been removed. Furthermore, crystal packing may lead to steric constraints that limit the number of functional states the complex can assume and thereby restrict the functional states that can be studied.

A technique for structural studies that has attracted growing interest over the past decade is single-particle reconstruction. Single-particle reconstruction is an experimental/computational technique that utilizes transmission electron microscopy (TEM) to address problems which crystallographic methods are unable to cope with. Combined with cryo-electron microscopy (cryo-EM) it is often the choice of method for structural studies of macromolecular complexes.

In this thesis, single-particle cryo-EM was conducted on three different macromolecular assemblies. The first paper describes a comparison of the two hemocyanin isoforms from the marine snail *Rapana thomasiana*. Structural characterization of the chaperonin from the hyperthermophilic archaea *Pyrococcus furiosus* is carried out in paper II. In paper III and IV, two functional states of the ribosome from *Escherichia coli* stalled at a truncated messenger RNA (mRNA) are studied.
HEMOCYANIN – AN OXYGEN TRANSPORTER

Hemocyanins are giant copper-containing multi-subunit respiratory proteins, freely dissolved in the hemolymph of many arthropod and mollusc species [2]. The oxygen binding sites in the two phyla are very similar, both in the way they involve a pair of copper atoms to how the copper is coordinated via histidine ligands (Figure 1) [3, 4]. Apart from these similarities the molluscan and arthropod hemocyanins represent two individual classes of proteins with distinct tertiary and quaternary structure.

Figure 1. The oxygen-binding site in arthropod and molluscan hemocyanins. The copper atoms are coordinated via histidine ligands. Coppers are brown and oxygens are red. Upon oxygenation the copper-copper distance in molluscan hemocyanin increases [3, 5]. PDB code: 1JS8.

STRUCTURE OF MOLLUSCAN HEMOCYANINS

Molluscan hemocyanins are based on large polypeptide chains, each ranging from 350 kDa to 450 kDa, that are folded into 7 (cephalopods) or 8 (gastropods, chitons and bivalves) globular functional units (FUs, termed a-g/h). The FUs, each containing one oxygen-binding site, are connected by linker peptide strands thus yielding a subunit structure that resembles a pearl-necklace (Figure 2A). Two subunits are paired in an anti-parallel fashion as stable dimers (the repeating unit), and in the basic quaternary structure of molluscan hemocyanins five such oblique dimers assemble into a hollow cylindrical molecule (Figure 2B) with an internal collar complex. In cephalopods and chitons, the decamer is the only structure found, whereas in gastropods and bivalves these decamers form di-decamers or even higher oligomers [2, 6, 7].

Structural information of hemocyanins has come from EM (e.g. immunoelectron microscopy [8, 9], cryo-EM [10-13]), X-ray crystallography [3, 5] and Small-angle neutron scattering studies [14]. Several pathways for the subunits have been proposed based on the immunoelectron microscopy studies [9], and recently one of them was suggested by Gebauer and colleagues to be the most probable one [15].

It has been shown in some gastropods, e.g Megathura crenulata [16, 17] and Haliotis tuberculata [18], that the hemocyanin exists in two structurally distinct isoforms and that they are homo-oligomeric. In Helix pomatia, up to three different hemocyanin isoforms are found [19]. Although the different isoforms in these species are composed of closely related subunits, they differ in terms of their oligomeric stability [20-24], carbohydrate content [25] and oxygen-binding properties [23, 26].
The basic structure of the molluscan hemocyanin. (A) Schematic of the polypeptide chain (subunit) with its eight different functional units. (B) Schematic of two oblique anti-parallel subunits (the repeating unit) arranged about a molecular 5-fold symmetrical axis within the cylindrical decamer. Note, in order to simplify the description at this stage the two subunits in the dimer are displayed as linear. In reality, the pathway of the subunit has still not been firmly established. However, it is widely acknowledged that the two subunits within the dimer are anti-parallel.

**MEDICAL SIGNIFICANCE**

Over 30 years ago the keyhole limpet hemocyanin (KLH) from the marine gastropod *M. crenulata* was found to possess immunostimulatory properties [27-30]. Most hemocyanins are glycoproteins, and the immunological response of KLH has often been related to its specific carbohydrate content and composition. For instance, the success of using KLH for the treatment of bladder carcinoma (reviewed in [31]) is probably due to a carbohydrate epitope that is cross-reactive with an equivalent epitope on the bladder tumor cell surface [32].

KLH is also widely used as a hapten carrier [33]. Haptens are small molecules such as drugs, hormones and peptides that, in general, require the assistance of a larger carrier protein to stimulate the immune system for the production of antibodies. The efficiency of KLH as a carrier protein is due to its large molecular size and its abundance of lysine residues for coupling haptens. For other medical applications of KLH, see [33].

The successful use of KLH in both clinical and biotechnological applications have made it interesting to study hemocyanins from other species (mostly gastropods), especially their structure and carbohydrate content. Recently it was shown that the hemocyanin from *Concholepas concholepas* possess adjuvant immunostimulatory properties, as well as immunotherapeutic effects against bladder cancer [34, 35]. Moreover, the hemocyanin from *H. tuberculata* has been considered to be a possible substitute/complement for KLH in biotechnological purposes due to its higher oligomeric stability under certain ionic conditions [22].

**RAPANA THOMASIANA HEMOCYANIN**

*Rapana thomasiana* (prosobranch gastropod) is a marine snail that was originally found along the coast of Japan. However, fifty years ago it was discovered in relatively large amounts along the west coast of the Black Sea. Considered that the Black Sea is practically a closed sea with salinity less than half of what the Pacific Ocean contains, it might be expected that some physiological properties of this
Gastropod will have been affected to adapt to the new environmental conditions. Similar to *M. cremlata* and *H. tuberculata*, its hemocyanin (RhH) is also present in two independent homo-oligomeric isoforms [36]. RhH are composed of two structurally distinct subunits, RhH1 and RhH2, with molecular masses of 420 kDa and 450 kDa, respectively. Recently it was shown that RhH possess immunostimulatory properties similar to those of KLH [37, 38]. Previous EM studies of specimen prepared from native RhH containing both isoforms showed that the complexes were solitary di-decamers [36]. Apart from this, no additional structural studies were performed.

The different hemocyanin isoforms in most mollusc species, including *R. thomasiana*, exhibit differences in their re-association behaviour under different ionic conditions [20-24, 36]. In paper I the structure of the two RhH isoforms were studied in order to detect something that could, hopefully, provide an explanation to this behaviour. The cryo-EM maps showed that the quaternary structure of RhH1 and RhH2 (Figure 3) were identical at the resolution of the experiment. While the reconstructions did not give any clues to why the two isoforms differ in their oligomeric stability, they did reveal three previously unreported connections. Based on the subunit pathway proposed by Gebauer *et al.* [15], one connection was identified to be within the subunit dimer (Figure 3D, full arrow) while the other two were between two adjacent subunit dimers (Figures 3E and 3F, full arrows). The former connection is most likely involved in stabilizing the subunit dimer while the two latter connections are formed during the pentamerization process to stabilize the decameric ring. Molluscan hemocyanins require Ca²⁺ and/or Mg²⁺ ions to maintain their quaternary structure. A decrease of these ions results in dissociation of the oligomeric complex into decamers, subunit dimers or individual subunits depending on the ionic conditions [22, 36, 39]. It is therefore possible that the aforementioned non-covalent connections provide binding sites for Ca²⁺ and Mg²⁺.

After the publication of our work, the three non-covalent connections previously mentioned were also observed in HtH1 (one of the hemocyanin isoforms from *H. tuberculata*) [40], KLH1 [41] and also in the *Nautilus pompilius* (cephalopod) hemocyanin (NpH) [40, 42]. However, compared to our work the connections were designated differently due to the new subunit pathway proposed. The new pathway of the subunits in NpH was determined by fitting homology models of the FUs into a 9Å cryo-EM map [42]. Recently it was shown that this subunit arrangement is valid for KLH1 as well [41], and thus it was proposed that it also applies to other molluscan hemocyanins. The most remarkable with the new subunit pathway is the suggestion that the two FUs in the arch (FU-g) are from two adjacent subunit dimers [42]. If the FUs in our cryo-EM maps are re-assigned according to the new subunit pathway [42], the linker region between FU-fY (wall) and FU-gY (arch) will be absent. The covalent linker regions are very long, and are thus presumed to be very flexible. According to the new pathway, the FU-fY → FU-gY linker region spans a very long distance [42]. In addition, it is also highly exposed. It is therefore likely that it shows variability in conformation during the cooperative oxygen binding process, and as a result it is averaged out in our reconstructions. Thus, the absence of this linker region in our cryo-EM maps is most likely due to the presence of different hemocyanin conformers in the sample. In the recent study where a 9Å cryo-EM density map of KLH1 was obtained, the sample was kept in a chamber containing 25% oxygen prior to grid preparation for the EM work [41]. This procedure was carried out in order to ensure full oxygenation of the hemocyanin complexes, and is probably what enabled the authors to define this linker region in their reconstruction.
Figure 3. 3D reconstruction of RtH2. FUs labeled with yellow (FU$_Y$) and white (FU$_W$) colors correspond to FUs within the front subunit dimer, while FU$_L$ and FU$_R$ refer to FUs from its left and right neighboring subunit dimers, respectively. (A) Top view. (B) External side view. The three tiers of the wall and the major groove (broken line) within the subunit dimer are marked out. (C) Internal side view showing the arch (FU$_g$) and the collar elements (FU$_h$). (D) Fragment of the RtH2 map showing the non-covalent intra-subunit dimer connection between FU$_{gY}$ and the first tier of the wall (full arrow), and the covalent linker region (broken arrow). (E) Same fragment as in (D) rotated about 90° to reveal the linker region (broken arrow), and the non-covalent inter-subunit dimer connection (full arrow) to FU$_{gW}$. (F) The non-covalent collar-wall connection between individual subunit dimers (arrow).
**CHAPERONIN – PROTEIN FOLDING ASSISTANT**

Chaperonins are a specific class of molecular chaperones that assist the folding process of proteins in the crowded environment of the cell (reviewed in [43-45]). Their general structure of a multi-subunit double-ring assembly with two central cavities allows them to encapsulate protein substrates for ATP-driven protein folding (Figure 4). Chaperonins are found in all organisms studied to date (with the exception of a few Mycoplasmas), and they are essential for cell viability.

There are two main chaperonin groups (reviewed in [46, 47]): Group I (reviewed in [48]) consist of eubacterial chaperonins (exemplified by *E. coli* GroEL) and homologous proteins in mitochondria and chloroplasts, and Group II (reviewed in [49]) which the eukaryotic chaperonin CCT and chaperonins from archaea belong to. GroEL is composed of 14 identical subunits that form the 7-fold rotational symmetric double-ring complex [50]. The composition for group II chaperonins is more complex. CCT is composed of eight distinct subunits [51], whereas archaeal chaperonins consist of one, two or three different subunits [52, 53].

**STRUCTURE AND MECHANISM**

All chaperonin subunits have a similar overall structure consisting of three domains: apical, intermediate and equatorial [54, 55]. Within the double-ring assembly the subunits are arranged so that the equatorial domains form the inter-ring and inter-subunit contacts, and the apical domains, which contain the binding sites for the protein substrates, are located at each end of the cylinder (Figure 4). The intermediate domains connect the apical and equatorial domains. The major structural difference between the two chaperonin groups is that complexes belonging to Group II have an extra helical extension of the apical domain serving as a built-in lid that closes the central cavity [56]. GroEL, on the other hand, use the co-protein GroES for that function (Figure 4).

![Figure 4. Schematic of chaperonin structures.](image)

Figure 4. Schematic of chaperonin structures. Side view of Group I (left) and Group II (right) chaperonin in cross-section. The equatorial domains are represented by dark blue rectangles; the light blue ovals are the intermediate domains and the apical domains are represented by green ellipses. A co-protein (orange) is utilized by Group I to seal the folding chamber, whereas in Group II a helical protrusion (lighter green extension) serves as a built-in lid structure to close the cavity. Surfaces on the apical domain for substrate (grey, left only) binding are marked by brown patches. (Taken and modified from [57]).

GroEL is the most extensively studied Group I chaperonin, and a great deal is known about its reaction cycle (reviewed in [47, 58]). Briefly, the mechanism starts with

---

1 For simplicity, GroEL will be used to represent Group I chaperonins.
binding of the unfolded protein substrate to the apical domains in one ring (the \textit{cis} ring). Subsequent ATP binding to the equatorial domains increases the affinity for the co-protein GroES, which seals the folding chamber and induces the release of the substrate into the cavity. The substrate is encapsulated in the cavity until ATP binds to the opposite ring (the \textit{trans} ring), which cannot occur until ATP in the \textit{cis} ring has hydrolyzed. Substrate and ATP binding to the \textit{trans} ring triggers the release of GroES and the folded substrate from the \textit{cis} ring, and a new cycle of protein folding takes place on the \textit{trans} ring.

In contrast to the situation for GroEL, the functional cycle for Group II chaperonins is not well understood. X-ray crystallography and cryo-EM studies have identified several conformations of archaeal and eukaryotic chaperonins \cite{55, 59-62} that show resemblance to some functional states observed for GroEL \cite{54, 63-65}, but it is still unclear how they are coupled to the ATPase cycle. One factor that contributes to limit the amount of information in this matter is the hetero-oligomeric nature of Group II chaperonins, which impose difficulties to understand the properties of the distinct subunits and how these are related to the mechanism.

\textit{Pyrococcus furiosus} is a hyperthermophilic archaea that grows optimally at approximately 100°C. The \textit{P. furiosus} genome contain only one chaperonin gene, suggesting a homo-oligomeric chaperonin complex \cite{52, 66}. Due to its minimal oligomeric complexity, the \textit{P. furiosus} chaperonin is an attractive model system for studying the structure-function relationship of Group II chaperonins. The chaperonin from the closely related organism \textit{Pyrococcus horikoshii} has also been studied \cite{67}, but structural characterization is still unavailable for any \textit{Pyrococcus} chaperonins.

In paper II, structural characterization of the chaperonin from \textit{P. furiosus} (\textit{Pf} Cpn) was carried out. Expression and purification of \textit{Pf} Cpn was performed according to the protocol outlined in paper II, and electron micrographs of frozen-hydrated \textit{Pf} Cpn samples revealed ring shaped structures that are characteristic for chaperonins. Because the end-view was the strongly preferred orientation, not all projection images were selected. A total of 674 projection images, selected from 22 micrographs, were used in the 3D reconstruction process. The resulting cryo-EM map revealed a di-octameric structure (Figure 5), consistent with the number of subunits per ring predicted by Archibald \textit{et al.} \cite{52} for \textit{Pyrococcus} chaperonin. Moreover, the double-ring assembly had 8-fold rotational symmetry. \textit{Pyrococcus} species contain only one chaperonin gene per genome \cite{52}, with a high sequence identity between the chaperonin genes. Hence, it is very likely that all chaperonins within \textit{Pyrococcus} species exist as 8-fold symmetrical complexes. Comparison of our density map with previously determined functional states of Group II chaperonins suggest that the 16-mer assembly is in a partially closed/open state \cite{59, 60, 68}.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{Fig5.png}
\caption{3D reconstruction of \textit{Pf} Cpn. Top view (left) and side view (right).}
\end{figure}
RIBOSOME – THE WORKBENCH OF TRANSLATION

When a protein is synthesized the genetic information stored in DNA must be accurately transferred to a polypeptide chain. This flow of information passes through an RNA molecule (Figure 6). First, the information encoded in DNA is transcribed into messenger RNA (mRNA). Subsequently, mRNA is used as a template when the genetic message is translated into the 20 different amino acids. The code words in mRNA are deciphered by transfer RNAs (tRNAs), small adaptor molecules that understand the two languages nucleic acids and amino acids. The entire translation process takes place on the ribosome, a large molecular machine found in all cells. During translation, the ribosome interacts with mRNA, tRNA and other molecules that are necessary for the synthesis of the polypeptide.

Ribosomes and the other key components essential for translation have been extensively studied over the years, which in turn have resulted in a large number of publications. The purpose of this chapter is to give a brief background of protein synthesis and the components involved; hence, only a fraction of the publications will be covered. For a more detailed coverage the work of A. Liljas [69] is recommended. In the next few paragraphs his disposition will be followed.

Figure 6. Transfer of information from DNA to RNA to protein.

The message – mRNA

The mRNAs are a class of RNA molecules that carries genetic information, copied from DNA, to the ribosomes for translation into protein. Figure 7 shows the genetic code, which is the set of rules that dictates how the genetic message is translated. The code words are triplets of nucleotides called codons [70-72]. Since there are four different nucleotides in mRNA (A, U, C, G), 64 (or $4^3$) different triplet combinations are possible. Out of these 64 codons, 61 specify individual amino acids. Thus, the genetic code is degenerate, meaning that some amino acids are encoded by more than one codon. Only methionine and tryptophan have one codon. On the other hand, leucine, serine and arginine are each specified by up to six codons (Figure 7). Codons encoding one specific amino acid usually differ in their third base. For example, the codons for alanine are GCU, GCC, GCA, and GCG. In addition to encoding individual amino acids, some codons also signal the beginning and end of the translation process. Generally the start (initiation) codon is AUG, which is the same as the methionine codon. The three codons UAA, UGA, and UAG do not encode amino acids; instead they are stop (termination) codons that signal the end of the polypeptide chain synthesis. The sequence of codons in mRNA that is between the start and stop signal is termed a reading frame.
Figure 7. The genetic code. The codons (three-base code words) are translated into the 20 amino acids, which are given in their three letter codes. The three termination codons, as well as the initiation codon, are in bold.

The adaptor molecule - tRNA

The decoding process where the nucleotide sequence of mRNA is converted into the amino acid sequence of the protein is performed by tRNA molecules. Each amino acid has its own tRNA it is attached to. When a codon in the mRNA calls for a particular amino acid, the tRNA carrying that amino acid will recognize the codon and thus deliver its amino acid to the growing polypeptide chain.

tRNAs consist of a single stranded RNA molecule, about 75 nucleotides long, folded into a stem-loop arrangement that resembles a cloverleaf when drawn in two dimensions (Figure 8A). The structure contains four stems that are stabilized by Watson-Crick base pairing. Three of the stems have loops. The middle loop carries the anticodon, the nucleotide triplet that base-pair to the corresponding codon in mRNA. The other two loops, which contain several modified bases, are part of the D and T arms. The fourth stem is composed of the 3'- and 5'-ends of the chain. All tRNAs has a conserved CCA sequence at their 3'-end that is not base-paired. The free 2'-hydroxyl or 3'-hydroxyl of the adenosine is covalently attached to the proper amino acid, thereby yielding an aminoacyl-tRNA (aa-tRNA). Thus, this unlooped stem is referred to as the acceptor or aminoacyl stem.

Over 35 years ago the first 3D structures of tRNA were determined [73, 74]. The structures revealed that the cloverleaf is folded into an L shape, with the anticodon loop and acceptor stem forming the ends of the two arms while the D- and T-loops form the elbow region (Figure 8B). In solution, the L shape is the basic structure for all tRNAs. On the ribosome, however, the tRNA can adopt several conformations.
Figure 8. Structure of tRNA. (A) The cloverleaf conformation of yeast alanine tRNA (tRNA\textsubscript{Ala}), the first nucleic acid sequence of tRNA to be determined [75]. Modified nucleotides: \textbf{D} = dihydrouridine, \textbf{I} = inosine, \textbf{T} = thymine, \textbf{\textupsilon} = pseudouridine and \textbf{m} = methyl group. The D arm nearly always contains a dihydrouridine. Likewise, the T or T\textupsilonCG arm almost always contains thymidylate and pseudouridylate. (B) The 3D structure of tRNA. The color coding is the same as in (A). The 3D structure of tRNA is reproduced with the kind permission of N.R. Voss.

The anticodon of a tRNA read the codon of mRNA by forming Watson-Crick base pairing. However, only the first and second bases of a codon form standard Watson-Crick base pairs with the third and second bases of the corresponding anticodon. The third base in the codon, the so-called “wobble” position, can form nonstandard pairing with the first base of the tRNA anticodon [76]. For this reason, a tRNA with a certain anticodon can recognize several codons.

Attachment of a particular amino acid to a tRNA is performed by the enzymes aminoacyl-tRNA synthetases (aaRS). Each of these 20 enzymes is specific for one amino acid, and they use different means to identify the correct tRNA. In general, the identity elements of the individual tRNAs are localized in the anticodon and acceptor stems [77]. For glutamine tRNA (tRNA\textsubscript{Gln}) the aaRS recognizes the correct anticodon [78], while the single G·U base pair in the acceptor arm of alanine tRNA (tRNA\textsubscript{Ala}) is sufficient as an identity element (reviewed in [79]). Once a tRNA has been charged with an amino acid, the cognate codon-anticodon interactions will guide the tRNA to deliver its amino acid to the growing polypeptide chain. For this reason, it is crucial for the fidelity of translation that the synthetases charge the tRNAs with their specific amino acid [80].

The protein-synthesizing machine – The ribosome

A ribosome consists of two subunits with unequal size, both composed of ribosomal RNA (rRNA) molecules and proteins. The size of the ribosome and its subunits, as well as some of the steps in the translation process, varies between different species. Thus, for simplicity only the bacterial ribosome will be covered.
In bacterial ribosomes (70S\(^2\)) the two subunits are called the 30S (small) and the 50S (large) subunit. These ribosomal subunits have irregular shape, but together they form a fairly globular complex with a cleft between them through which the mRNA and tRNA molecules transits (Figure 9). Various conformational changes takes place in the ribosome during translation (reviewed in [81]); where large-scale movements undergone by the 30S subunit is mainly in the head region whereas in the 50S subunit it is the side protuberances that show flexibility. Rotation of the two subunits with regard to each other is also observed.

Ribosomes contain about 65% rRNA and 35% protein, with the core of the ribosomal subunits formed by the rRNA molecules. The 30S subunit has one large rRNA molecule (16S RNA) while the 50S subunit has two rRNA molecules, one large (23S RNA) and one small (5S RNA). In both subunits it is the large rRNA molecule that provides the binding sites for the ribosomal proteins. Most of the ribosomal proteins are located at the external surface (solvent side) of the subunits [82-85]. The interface between the ribosomal subunits is mainly composed of rRNA. It is also here the functional sites of the ribosome are found; the decoding site on the 30S subunit [83, 86-89] and the peptidyl transferase center (PTC) on the 50S subunit [84, 85, 90, 91]. tRNA molecules are able to concurrently contact the decoding site on the 30S subunit and the PTC on the 50S subunit because of their L-shaped structure where the anticodon is at one side and the CCA tail is at the opposite side (Figure 8B), about 75 Å apart.

**Figure 9. The 70S ribosome.** L1: stalk of ribosomal protein L1; CP: Central protuberance; L7/L12: stalk of L7/L12; h: head; sh: shoulder; b: beak; sp: spur. Arrow (right) indicates the inter-subunit space.

There are three classical binding sites for tRNA molecules on the ribosome. These are the A site (for incoming aminoacyl-tRNA), the P site (for peptidyl-tRNA that carries the growing polypeptide chain) and the E site (for deacylated tRNA before it dissociates from the ribosome) [92-95]. The A site is located at the side of the L7/L12-stalk whereas the E site is at the L1-side of the ribosome. A fourth ribosomal binding site exists, which is where tRNA molecules are bound with EF-Tu prior to entry into the A site [96, 97]. This is called the A/T site and it is here the selection between cognate and non-cognate tRNA molecules takes place. tRNAs bound to this site are in a intermediate or hybrid state called the A/T state [97]. During their transit through the ribosome, tRNA molecules passes through two additional hybrid states, the A/P and the P/E state [98-100].

---

\(^2\) Naming of the ribosome and its subunits, as well as the rRNAs, is based on their sedimentation rates.
THE TRANSLATION PROCESS

Protein synthesis is divided into three stages; initiation, elongation and termination. Each step is catalyzed by specific translation factors. These factors, together with all the other components necessary for translation, are listed in Table 1.

Table 1. Components required for the translation process in *E. coli*.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Necessary components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>mRNA</td>
</tr>
<tr>
<td></td>
<td>fMet-tRNA\textsubscript{Met} (initiator tRNA)</td>
</tr>
<tr>
<td></td>
<td>Initiation/start codon (AUG) in mRNA</td>
</tr>
<tr>
<td></td>
<td>30S ribosomal subunit</td>
</tr>
<tr>
<td></td>
<td>50S ribosomal subunit</td>
</tr>
<tr>
<td></td>
<td>Initiation factors (IF1, IF2, IF3)</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
</tr>
<tr>
<td>Elongation</td>
<td>Initiation complex (70S mRNA· fMet-tRNA\textsubscript{Met} in P site)</td>
</tr>
<tr>
<td></td>
<td>Aminoacyl-tRNAs</td>
</tr>
<tr>
<td></td>
<td>Elongation factors (EF-Tu, EF-G)</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
</tr>
<tr>
<td>Termination and recycling</td>
<td>Termination codon in mRNA</td>
</tr>
<tr>
<td></td>
<td>Release factors (RF1, RF2, RF3)</td>
</tr>
<tr>
<td></td>
<td>Recycling factor (RRF)</td>
</tr>
</tbody>
</table>

*Initiation*

Protein synthesis is initiated when an mRNA molecule binds to the free 30S subunit. Correct initiation requires that the start codon (generally AUG) is positioned in the ribosomal P site. In bacteria, the 30S subunit identifies the start codon by forming interactions between its 16S RNA molecule and a nucleotide segment in mRNA that is located upstream of the AUG start codon. This mRNA sequence, called the *Shine-Dalgarno sequence* [101], is complementary to the 3'-end of the 16S RNA. After the correct placement of the start codon in the P site of the 30S subunit, the initiation factors IF1 and IF3 bind to the A and E sites, respectively [102-104], and guides the initiator tRNA (fMet-tRNA\textsubscript{Met}) into the P site. Subsequently, binding of IF2-GTP, release of IF3 and recruitment of the 50S subunit occurs. Lastly, GTP on IF2 is hydrolyzed and together with IF1 this factor leaves the ribosome. The result is the initiation complex composed of the 70S ribosome with fMet-tRNA\textsubscript{Met} in the P site that is ready for elongation.

*Elongation*

At the start of each elongation cycle, the 70S ribosome has a peptidyl-tRNA or the initiator tRNA (fMet-tRNA\textsubscript{Met}) bound in the P site. Elongation begins with aminoacyl-tRNA, elongation factor EF-Tu and GTP forming a ternary complex (aminoacyl-tRNA·EF-Tu·GTP) that binds to the A/T site on the ribosome. In the A/T site the
anticodon arm of the aminoacyl-tRNA molecule has a conformation that differs from those seen in normal L-shaped structures (Figure 10) [97, 100, 105, 106]. This modified tRNA structure allows codon-anticodon pairing in the decoding site on the 30S subunit while the aminoacyl end is bound to EF-Tu far from the PTC on the 50S subunit [100]. If the codon does not match the anticodon, the ternary complex leaves the ribosome and a new ternary complex is bound to the A/T site. If there is cognate codon recognition, EF-Tu will hydrolyze its bound GTP and undergo a conformational change. The EF-Tu·GDP complex has a low affinity for the aminoacyl-tRNA and the ribosome, thus it dissociates [107]. This will free the aminoacyl end of tRNA and allow it to swing into the PTC on the 50S subunit, which results in the tRNA molecule regaining its normal L-shaped structure. During this motion the position of the anticodon stem-loop (ASL) is maintained, i.e. the ASL in the A-site tRNA and A/T-site tRNA has a similar orientation (Figure 10) [100]. With a peptidyl-tRNA in P site and an aminoacyl-tRNA in A site, the peptide can be transferred. This leads to the pre-translocation ribosome complex with a peptidyl-tRNA in A site and a deacylated tRNA in P site.

Figure 10. A- and A/T-site tRNA. The A/T-site tRNA (blue) has a bent anticodon arm, compared to the A-site tRNA (red), due to a kink between the anticodon stem and the D stem. PDB id: 1QZA for A/T-site tRNA and 1QZB for A-site tRNA.

The final step of elongation is translocation. The end result of this procedure is the post-translocation ribosome complex where the next codon of mRNA is exposed in the A site and the peptidyl-tRNA in A site and the deacylated tRNA in P site has moved to the P and E site, respectively. This reaction is catalyzed by the elongation factor G (EF-G). During the course of translocation the two ribosomal subunits rotate with regard to each other in a ratchet-like motion [108, 109]. In addition, the peptidyl-tRNA and the deacylated tRNA passes through the hybrid A/P³ and P/E states, respectively [98]. The molecular mechanism for translocation is not well understood. However, current findings suggest the following model for the EF-G dependent translocation process (Figure 11) (reviewed in [110]). After peptidyl transfer, EF-G·GTP binds to the pre-translocation ribosome complex (a), with peptidyl-tRNA in A site and deacylated tRNA in P site, and hydrolyzes its GTP (b). Next, the two ribosomal subunits rotate with regard to each other (ratchet-like motion) and the P-site tRNA enters the hybrid P/E state (c). Subsequently, A-site tRNA shift to the A/P state (d). Then, structural re-arrangement of the ribosome allows the two tRNA molecules to enter their classical P and E sites (e). Finally, EF-G·GDP dissociates and the ribosome undergoes the second step of the ratchet-like movement and returns to its normal conformation where it is ready for a new cycle of elongation (f).

3 The first and second letter in A/P and P/E denotes the position of the ASL and the acceptor end, respectively. For instance, a tRNA in the A/P state means that its ASL is still bound in the 30S A site while its acceptor end has moved to the 50S P site.
Termination

Protein synthesis is terminated when one of the three stop codons UAA, UAG and UGA is exposed in the A site. In bacteria there are three release factors (RF1, RF2 and RF3) that work in concert to terminate the translation process. RF1 and RF2 recognize the stop codons UAG/UAA and UGA/UAA, respectively [111], and hydrolyze the polypeptide chain from the peptidyl-tRNA. RF3 is a GTPase, and its function is to remove RF1 and RF2 from the ribosome [112]. RF3 binds to the ribosome in complex with GDP [113]. Provided that RF1 or RF2 is bound to the A site, and the P site is occupied by a deacylated tRNA, RF3 exchange its GDP to GTP and subsequently releases RF1 or RF2 from the ribosome. Lastly, RF3 hydrolyze its GTP to GDP and dissociates from the ribosome.

After the termination step and the peptide have been released, the ribosome has to be prepared for a new round of translation. This task is performed by the ribosome-recycling factor (RRF) which, together with EF-G, promotes the separation of the two subunits and dissociation of the mRNA and the deacylated tRNA.

**TRANS-TRANSLATION**

If the ribosome reach the 3’-end of the mRNA molecule without encountering an in-frame stop codon, it leads to two undesirable consequences. First, it causes a significant
loss of translational efficiency since the ribosome stalled on the defective mRNA molecule will be unavailable for new rounds of translation. Secondly, the incomplete polypeptide produced may be toxic to the cell. In bacteria, both these problems are solved by the translation control system termed trans-translation (reviewed in [114-116]). The main components in trans-translation are an RNA molecule called transfer-messenger RNA (tmRNA) and its helper protein small protein B (SmpB). The SmpB·tmRNA system recognizes and rescues stalled ribosomes and simultaneously adds a C-terminal proteolysis tag to the incomplete polypeptide chain. Furthermore, the system also promotes degradation of the defective mRNA molecule, thus preventing any future ribosome stalling [117, 118].


tmRNA and SmpB: structure and function

As the name indicates, tmRNA is an RNA molecule that functions both as messenger and transfer RNA [119-122]. A diagram over the secondary structure of tmRNA is depicted in Figure 12. The 5'- and 3'-ends of the RNA molecule fold into the tRNA-like domain (TLD), which contains an acceptor arm, a T-arm and a D-loop. tmRNA contains a G·U wobble base pair in its acceptor arm, the identity element recognized by Alanyl-tRNA synthetase (Ala-RS). Hence, it is charged with Alanine by Ala-RS [121, 123, 124]. Instead of an anticodon arm, a long stem links the TLD to the other two domains in the tmRNA molecule, the mRNA-like domain (MLD) and the pseudoknot-rich domain consisting of four RNA pseudoknots (designated PK1 to PK4). The MLD contains the open reading frame (ORF) that encodes the proteolysis peptide tag (ANDENYLAA in E. coli).

Figure 12. Secondary structure diagram of tmRNA from Thermus thermophilus. The main three domains are the tRNA-like domain (TLD, boxed out), the mRNA-like domain (MLD) and the four pseudoknots (PK). Reprinted (and modified) with permission from [125]. Copyright © 2006 National Academy of Sciences, U.S.A.
The four pseudoknots are the largest features of tmRNA, encompassing about two-thirds of the tmRNA molecule (Figure 12). While the importance of the TLD and MLD for the tmRNA activity has been firmly established, the specific function of the pseudoknots is still unclear. PK2 to PK4 have been suggested to be important for the proper folding of tmRNA [126], while PK1 may have a functional role for the tmRNA activity [123, 127, 128]. PK1 is only 11 nucleotides upstream of the resume codon, and although the reading frame selection is determined by the nucleotides immediately upstream of the first codon [129, 130], this pseudoknot could be involved in positioning the tmRNA ORF into the A site.

The most significant protein partner of tmRNA is SmpB. This protein binds tmRNA with high affinity and specificity [131, 132], and it is essential for all known tmRNA functions. As previously described, during normal translation aminoacyl-tRNA binds the ribosome in a ternary complex with EF-Tu·GTP, and cognate codon-anticodon interactions trigger GTP hydrolysis on EF-Tu followed by the release of EF-Tu·GDP and accommodation of the tRNA acceptor stem into the PTC on the 50S subunit (see ‘elongation’ section). Likewise, Ala-tmRNA too is delivered to the ribosome in complex with EF-Tu and GTP. However, SmpB is also required since tmRNA-EF-Tu·GTP alone does not associate stably with stalled ribosomes [131, 133]. Since tmRNA lacks an ASL (Figure 12), accommodation of Ala-tmRNA into the ribosomal A site do not require codon-anticodon interactions. It is not clear how the SmpB-tmRNA mediated mechanism can bypass the decoding process and enter the stalled ribosome. However, numerous studies suggest that SmpB can substitute for the missing ASL, and that it does so with its highly conserved C-terminal tail [131, 134-138]. Sundermeier et al. [134] showed that SmpB proteins with a mutated or truncated C-terminal tail cannot support addition of the proteolysis tag, despite being able to bind tmRNA and promote association of the SmpB-tmRNA complex with stalled ribosomes. This indicates that the C-terminal tail of SmpB is required for a tmRNA function following the recruitment of the SmpB-tmRNA complex to the ribosome, but prior to trans-peptidation where the tmRNA-linked alanine is added to the incomplete polypeptide. Consistent with these findings, chemical probing experiments revealed that SmpB interact with the three key conserved nucleotides A1492, A1493 and G530 of the 16S RNA that form the 30S subunit decoding center, and that G530 in particular interact with the C-terminal end of SmpB [137]. Further support comes from X-ray crystallography studies of SmpB in complex with the TLD of tmRNA [135, 136]. SmpB binds to the elbow region of TLD (Figure 13) and it has been suggested that the protein substitutes for the anticodon arm and the D stem that is present in all native tRNAs but missing in the TLD of tmRNA. Furthermore, biochemical experiments demonstrated the necessity of SmpB on tmRNA for the ribosome-dependent GTPase activation of EF-Tu in the Ala-tmRNA-SmpB-EF-Tu·GTP quaternary complex [138]. Altogether, these and other related studies indicate that SmpB functions as an ASL mimic during the initial stages of trans-translation.
The mechanism of trans-translation

The SmpB-tmRNA mediated mechanism is depicted in Figure 14, and the process can be described as follows [114, 139]. tmRNA forms a complex with SmpB that is alanylated by Ala-RS. Subsequently, the Ala-tmRNA-SmpB complex is recognized by EF-Tu-GTP (box 1). The Ala-tmRNA-SmpB-EF-Tu-GTP quaternary complex binds ribosomes stalled at the 3’-end of an mRNA. The lack of an anticodon stem in tmRNA is compensated for by SmpB, which presumably interact with the ribosome near the decoding center and thus triggers GTP hydrolysis on EF-Tu.

After GTP hydrolysis, the TLD of Ala-tmRNA accommodates into the ribosomal A site and the incomplete polypeptide is transferred to Ala-tmRNA (trans-peptidation).
Next, the TLD moves to P site and the ribosome switch template from the defective mRNA to the ORF of tmRNA. Unlike normal translation, the resume codon is not specified by initiation factors. Instead, the reading frame of the tmRNA-encoded tag sequence is determined by bases upstream of the resume codon. After removal from the ribosome, the defective mRNA molecule is degraded by RNase R (box 2). Translation of the tmRNA ORF continues until a stop codon at the end of the reading frame is encountered. This promotes translation termination, ribosome recycling and polypeptide release. The released nascent polypeptide is recognized by several cellular proteases, due to the 11 amino acid degradation tag at its C-terminus, and is thus degraded (box 3). The end result is a rescued ribosome that is now free to engage in a new round of translation, while the defective mRNA and the incomplete polypeptide has been eliminated and will not be to any harm for the cell.

Physiological significance of trans-translation

To date, all sequenced bacterial genomes contain the genes encoding SmpB and tmRNA (smpB and ssrA, respectively) [114, 140, 141]. For most bacterial species, where the exceptions are some pathogenic bacteria (such as Neisseria gonorrhoeae, Mycoplasma genitalium and Mycoplasma pneumoniae [142, 143]), the tmRNA·SmpB system is non-essential under ideal growth conditions. However, tmRNA activity is frequently important when cells respond to stress, differentiate or during pathogenesis [144]. For instance, E. coli strains lacking ssrA and smpB exhibit slower growth at high temperature [145], and they also show increased sensitivity towards amino acid starvation [146] and to antibiotics that promote translational frameshifting, read-through and stalling [147]. Similarly, Bacillus subtilis deleted of ssrA and smpB show reduced growth rate at both high and low temperature and an increased level of tmRNA and SmpB during temperature stress [148, 149]. Furthermore, an eightfold increase of tmRNA levels has also been observed for Thermotoga maritime during antibiotic challenge [150].

Virulence in some pathogenic bacteria is tmRNA·SmpB-dependent. Yersinia pseudotuberculosis and Salmonella enteric are two pathogens that exhibit defects in virulence when they lack tmRNA activity [151-153]. Studies have shown that Y. pseudotuberculosis cells lacking ssrA or smpB cannot cause lethal disease in a mouse infection model [151]. The mutant strains where also unable to proliferate in macrophages. Part of this virulence deficiency has been ascribed to defects in expression and delivery of the virulence effector proteins known as Yops. Production and secretion of Yops is regulated by the transcription factor VirF, and absence of tmRNA activity leads to misregulation of VirF and a delayed Yop secretion. In addition, Y. Pseudotuberculosis cells lacking the tmRNA·SmpB system are nonmotile, have increased sensitivity to antibiotics, and they are also temperature sensitive [151]. Similar to Y. Pseudotuberculosis, S. enteric also require tmRNA activity to proliferate in macrophages, and for virulence in mice [152, 153]. Altogether, these findings emphasize the importance of the SmpB-tmRNA system for survival under adverse conditions and in virulence of some pathogenic bacteria.

Structural studies of tmRNA·ribosome complexes

To date, the structural information available of tmRNA-bound ribosome complexes has come from cryo-EM studies, and the complex most extensively studied has been with tmRNA in the pre-accommodated state [125, 154]. This tmRNA-ribosome complex can be obtained with the antibiotic kirromycin (kir). Kirromycin permits GTP hydrolysis on EF-Tu but prevents EF-Tu-GDP from leaving the ribosome, thus allowing the
tmRNA-ribosome complex to be visualized in a state after GTP hydrolysis but prior to accommodation of tmRNA into the A site. The initial cryo-EM structure of the pre-accommodation tmRNA-ribosome complex revealed three major findings [154]; First, the TLD of tmRNA was found to bind EF-Tu in a manner comparable to an aminoacyl-tRNA. Second, the remaining density of tmRNA was highly structured and formed a large spiral that enclosed the beak of the 30S subunit in such a way that the ORF of the MLD was positioned close to the mRNA entrance channel. Third, one SmpB was found to bridge the TLD and the GTPase-associated center (GAC) on the 50S subunit. The latter was corrected for in a more recent cryo-EM study when the presence of two SmpB molecules in the pre-accommodation tmRNA-ribosome complex was observed [125]. In this density map one SmpB (SmpB-1) was found to bind near the site where it was previously observed [154], and the other (SmpB-2) was located close to the decoding site on the 30S ribosomal subunit. Both SmpB molecules interact with the TLD, where the complex formed between SmpB-2 and the TLD is similar to the TLD-SmpB complex observed by X-ray crystallography [135, 136].

One of the main questions regarding tmRNA is the fate of its structure during the course of trans-translation. As previously mentioned, the tmRNA molecule is highly structured when it enters the stalled ribosome [125, 154]. Thus, the question that arises is whether it stays structured or whether it unfolds. Wowér et al. [155] recently proposed that PK2 to PK4 of tmRNA unwinds as trans-translation progresses. However, other studies suggest that the pseudoknots of tmRNA, PK3 in particular, remain structured and that they stay outside the ribosome [156, 157].

In order to increase the structural information about trans-translation, we decided to study tmRNA-ribosome complexes captured at stages subsequent to the pre-accommodated state. The first is the post-accommodation complex (paper III) where EF-Tu-GDP has dissociated and tmRNA has accommodated into the ribosomal A site. The cryo-EM map shows the 70S ribosome with a density in the A site that is connected to the remaining tmRNA density (Figure 15A). Fitting of the recently published crystal structure of the T. thermophilus TLD-SmpB complex [135] into the A-site density revealed the presence of one SmpB molecule (corresponding to SmpB-2 in the pre-accommodated state [125]) in the density map. This is in agreement with previous biochemical studies [157, 158] where it was suggested that the SmpB molecule bound to the 50S subunit (SmpB-1 in [125]) dissociates upon accommodation of tmRNA into the A site. The fitting also revealed that the TLD-SmpB part of the tmRNA SmpB complex is positioned in the A site similarly to a canonical A-site tRNA (Figure 15B), in line with the suggestion that SmpB mimics the anticodon arm and the D-stem of native tRNAs [135-137]. Accordingly, during accommodation the SmpB molecule moves deeper into the A-site pocket on the 30S subunit while the acceptor stem of the TLD of the tmRNA-SmpB complex rotates away from its EF-Tu-bound position and moves into the A site of the 50S subunit (Figure 15C). The distance the acceptor arm of the TLD moves during accommodation is about 50 Å, which is very similar to the distance of 46 Å that the acceptor arm of a canonical aminoacyl-tRNA move during accommodation [159]. Altogether, our findings support the model in which the TLD-SmpB complex is suggested to functionally mimic tRNA during the first steps of trans-translation [135-137, 160, 161].
Figure 15. Reconstruction of the post-accommodation tmRNA·SmpB·ribosome complex. (A) tmRNA·SmpB·70S complex in the post-accommodated state. (B) Left, TLD·SmpB complex fitted into the A-site density. Right, the relative positions of the A-site tRNA and the A-site density. (C) Movement of the TLD·SmpB complex during accommodation. The ribosomal orientation for (B) and (C) is indicated by the corresponding thumbnails. tmRNA·SmpB is colored red, 30S and 50S ribosomal subunits is in yellow and blue, respectively. CP, central protuberance; L1, stalk of protein L1; L7/L12, stalk of proteins L7/L12; b, beak; sh, shoulder; sp, spur; dc, decoding center; post, post-accommodated state; pre, pre-accommodated state [125]. PDB codes: 2CZJ for TLD·SmpB, 1GIX for tRNA.

Similarly to previously reported cryo-EM maps of tmRNA·ribosome complexes [125, 154, 162], the tmRNA in our density map was highly structured and formed an arc around the beak of the 30S ribosomal subunit (Figure 15A). Comparison with the tmRNA structure in the pre-accommodation ribosomal complex [125] shows that a conformational change takes place in the arc when tmRNA goes from the pre-accommodated to the post-accommodated state. The most significant change is the downward flip undergone by PK1, which result in the MLD of tmRNA being positioned closer to the A-site entry. It is likely that the re-arrangement of the tmRNA structure may be caused by, or coordinated with, the rotation of the TLD·SmpB complex in order to facilitate the entry of the MLD into the decoding center of the 30S subunit.

These findings are in contradiction to the previously reported cryo-EM study of the post-accommodation complex [125] where no movement in TLD or SmpB was observed during tmRNA accommodation, and the arc density corresponding to PK2-PK4 was absent in the density map. We ascribe these contradictory findings to low occupancy of tmRNA in their post-accommodation complex [125].

In paper IV we studied the tmRNA·ribosome complex after two rounds of elongation, where the P and E site should be occupied by a peptidyl-tRNA\(^{Ala}\) and the deacylated tmRNA, respectively. The six reconstructions generated from the heterogeneity analysis showed strong density in all three ribosomal sites, consistent with the double translocation complex containing a deacylated tRNA\(^{Amp}\) in the A site. Density for the
tmRNA arc composed of the pseudoknots was also observed in all cryo-EM maps, and it was in a position somewhat similar to the arc in the post-accommodation ribosomal complex. However, the arc exists in two states, one stable and highly structured and another much more flexible and disorganized (Figure 16). The absence of density in the fragmented arc structures are mostly in the hinge regions between the individual pseudoknots, indicative of high flexibility in these parts of the structure that causes them to become averaged out in our reconstructions. Thus, our findings suggest that the pseudoknots in the double-translocation tmRNA-ribosome complex are still located outside the ribosome, and that the individual tmRNA elements are highly dynamic around the hinge regions.

Figure 16. Cryo-EM maps of double-translocation tmRNA-bound ribosome complexes. The reconstructions are displayed with the arc density in red. The top left side of each class shows the atomic coordinates of tmRNA in the post-accommodated state fitted into the isolated arc density for the corresponding cryo-EM maps. The resolution and the number of particles contributing to each class are indicated below each class. **Cyan**, PK1; **black**, linker between PK1 and ORF; **magenta**, ORF; **yellow**, helix 5; **purple**, PK2; **orange**, PK3; **gray**, PK4. PDB code: 2OB7 for tmRNA coordinates.
SINGLE-PARTICLE CRYO-ELECTRON MICROSCOPY

Transmission electron microscopy (TEM) is an extremely flexible imaging technique for a scientist. It can be utilized for 3D structural studies of protein molecules or macromolecular complexes, as well as for studying the organization of organelles within cells. Cryo-electron microscopy (cryo-EM) in combination with single-particle reconstruction is a widely used technique to study macromolecular complexes, as these assemblies are often either too large or too flexible for crystallographic methods [163]. The single-particle method is based on recording a large number of TEM images (projections) of the complex, determine the angular orientation of each image, and subsequently combine the images to recover the 3D structure of the complex. The main advantage with this technique over crystallographic methods is that the complexes are in a non-crystalline close-to-native state; hence, no steric constraints are present that will limit the functional states they can assume. Single-particle cryo-EM has enabled 3D structures of macromolecular complexes to be obtained at subnanometer resolution (reviewed in [164, 165]). Recently, near-atomic resolution has also been achieved ([166-169], reviewed in [170]).

SPECIMEN PREPARATION

Specimens observed in the EM are subjected to a high vacuum [171]; therefore they must be prepared in such a way that the structure of the initially hydrated molecules is preserved when the specimen is positioned in the vacuum. In addition, they must also be protected from the radiation damage caused by the electron beam. Two of the most commonly used specimen preparation methods will be covered in this section, namely negative staining and vitrification, with emphasize on the latter technique.

Negative Staining

The negative staining method developed by Brenner and Horne [172] is based on utilizing a heavy metal salt (e.g., uranyl acetate) to reduce the radiation damage and protect the molecules from collapsing, while at the same time it also leads to an increase in the contrast. However, since it is the heavy metal atoms distributed around the molecules that are visualized, the structural information is limited to the shape of the molecule. Furthermore, the structure of the molecules is only preserved to some extent, deformation of the molecules occurs as a result of air drying [173]. Therefore, negative stained specimens are limited in resolution [174, 175]. Nevertheless, negative staining is technically easy to perform and is often used to evaluate whether a molecule is suitable or not for high-resolution EM studies.

Vitrification

The technique of vitrifying samples for cryo-EM studies was developed more than 30 years ago (reviewed in [176], [177, 178]), and it revolutionized the field of biological EM as it became possible to observe macromolecules under close-to-native conditions. The trick with this technique is the amorphous ice (or vitreous ice) the sample is embedded in, which enables the molecules to be fully hydrated despite the requirement for vacuum conditions inside the EM. Thus, the molecule is preserved in a biologically relevant conformation. Amorphous ice is produced when water is cooled extremely fast
(cooling rate $\sim 10^5$ K/s) [171], thus preventing formation of ice crystals that could damage the specimen. This is done by plunging the grid into a cryogen, often liquid ethane cooled by liquid nitrogen. The preparation steps are as follows: the sample is applied to a grid. The grid is then blotted with filter paper to remove excess liquid, leaving only a thin layer of the sample.Immediately after blotting, the grid is plunged into liquid ethane.

The advantage with this technique is the reduction of radiation damage at low temperatures [179] and that specimen collapse is avoided. Another advantage is that the contrast observed is related to the biological molecule itself rather than some contrasting agent, such as the heavy metal atoms in negative staining. However, the latter case is also what leads to the low contrast in EM images of vitrified specimens, which causes difficulties for the subsequent image processing procedures. In addition, the electron dose needs to be limited in order to reduce the radiation damage; the projection images produced are therefore very noisy with a low signal-to-noise ratio (SNR), thus further imposing problems when the images are analyzed. Despite these drawbacks, cryo-EM has still enabled structures of macromolecular complexes to be determined with high accuracy. A significant factor contributing to this success is the progress that has been made in developing more powerful software platforms, as well as the increased computer power, to deal with the noisy EM images.

During imaging, the electron beam can eject electrons (called ‘secondary electrons’) from the specimen. This leads to a buildup of positive charge that can distort the recorded image through an effect termed charging. Charging can be reduced by the carbon film that coats the EM grids, an effect that increases if the carbon support grids have been pre-irradiated under the electron beam prior to specimen preparation [180]. Pre-irradiated carbon support grids are used mostly for studies of 2D crystals. However, due to severe charging problems with the ribosome specimens (paper III and IV), these grids were tested. The result was a significant reduction of the charging effect.

The pre-irradiated grids used were ‘holey carbon grids’ coated with an extra (continuous) carbon film. Holey carbon grids are basically what the name implies, i.e. grids coated with holey carbon. When the specimen is vitrified, in cases where regular holey carbon grids are used, the macromolecules are dispersed in the holes away from the carbon film. This is to avoid any orientation preference that can occur when the molecules adsorb to the carbon film. However, it has been shown that the carbon actually has an opposite effect on the ribosome [163]. Hence, not only did the extra carbon film applied to the holey carbon grids reduce the charging effect, but it also generated a more random orientation distribution of the ribosome complexes.

**DATA PROCESSING**

Once a set of electron micrographs has been collected, the individual projection images of the molecules needs to be selected for the subsequent image processing procedure. This can be done manually or automatically. The automatic particle selection methods available are usually based on locating the particles by cross-correlating rotationally averaged reference images with the entire micrograph. The user manually selects several particles, which are then rotationally averaged and used as templates. A combination of the two methods was used for the hemocyanin and the ribosome projects, where boxer from the EMAN package [181] was first used to select all
particles in the right size range as the molecule, and then the gallery of particle candidates was visually checked. Due to the strong preference of the chaperonin complex to be in the end-orientation, particles for the chaperonin dataset were selected manually. First all side-views were selected, and subsequently a reasonable amount of images representing the end-orientation of the molecule was added to the dataset.

The next step is to correct the contrast transfer function (CTF) of the projection images [163]. The rationale for this correction is that a number of factors (predominantly the amount of defocus and the spherical aberration constant of the lens system (Cₐ)) contribute to affect the acquired images. The effect of the CTF is best described in Fourier space, where it will act as a multiplicative function. It has the shape of a sine wave with the period compressed at high resolution. The CTF for a certain micrograph, and accordingly also for all the particles in the micrograph, can be obtained by calculating Fourier transforms of images selected from this micrograph that contain carbon film. By fitting the theoretical expression of CTF to the CTF of the carbon film, defocus and CTF parameters for each micrograph can be determined. The Cₛ value is taken to be a known constant for a specific microscope. With these values known, correction of the sign changes (and possibly damping) caused by the CTF can be performed on the individual projection images that originate from a certain micrograph. CTFit from the EMAN package [181] was used for defocus determination, as well as the CTF phase correction of the datasets from all projects. CTF amplitude correction was only carried out on the hemocyanin (paper I) and the chaperonin (paper II) datasets, and it was performed as part of the reconstruction procedure (see below).

Once the projection images have been selected and CTF correction (in our case, only the phases) has been performed, the angular orientation of each projection needs to be determined (angular assignment) so the dataset can be combined to recover the 3D structure of the molecule (3D reconstruction). There are several software packages available that provide the tools to process the projection images, and they all use different approaches. For instance, SPIDER [182], FREALIGN [183] and STRUL [184] define the orientation angles by aligning the experimental projections against 3D reference volumes. For cases where no initial 3D volumes are available, 2D class averages are generated from the individual projections and taken to represent SNR enhanced projections. These can then be assigned Euler angles by maximizing the common lines correlation coefficient, a technique employed by IMAGIC [185] and RAD [186].

For the hemocyanin and chaperonin projects, the EMAN package [181] was used. EMAN use projection matching to determine the orientation angles of the projection images. However, instead of combining the individual projections to recover the 3D volumes, ‘refined’ class averages are used. The reconstruction procedure utilized by EMAN (figure 17) is as follows: First, the individual particles are selected from the electron micrographs. Subsequently, the CTF parameters are determined and the phases are corrected. 2D reference projections (so called re-projections) are then generated from the preliminary 3D volume and subsequently used to classify the projection images. Once the projections have been grouped into individual classes, ‘refined’ class averages representing the different projection views are generated. In parallel, CTF amplitude correction is also carried out. Re-assignment of Euler angles for the class
averages are then performed with projection matching. Finally, a new 3D model for the next round of refinement is generated from the class averages.

EMAN [181] was also used to generate the initial models of hemocyanin and chaperonin, which utilize a technique based on the rotational symmetry of the complex. First, a search for projections with the best $n$-fold rotational symmetry (‘top’ views), and projections with the best mirror or pseudo-mirror symmetry and the poorest $n$-fold rotational symmetry (‘side’ views) is performed. Class averages of the top and side views of the complex are then generated and subsequently used to construct a preliminary 3D model. D5 and D8 symmetry was applied for the hemocyanin and chaperonin complexes, respectively.

Figure 17. Schematics of the reconstruction process in EMAN. (Taken and modified from [181])

**HETEROGENEITY ANALYSIS**

In single-particle reconstruction, the underlying assumption is that the projection images represent randomly oriented views of the same structure. If the sample contains a mixture of conformational states, e.g. different conformers or different ligand binding states, the dataset collected will be structurally heterogeneous. The 3D volume recovered from such a dataset does not reflect any of the co-existing structures. Consequently, the only way to generate an interpretable structure is to divide the heterogeneous dataset into homogeneous subsets, which can then be used to recover the individual 3D volumes. There are two approaches that can, in principle, perform this task. These are known as *supervised* and *unsupervised* classification. The former approach is a reference based classification method, where the heterogeneous dataset is divided according to resemblance between the projection images and certain 3D template volumes that represent different conformational states of the molecule. These 3D templates can be obtained from other experiments [187] or created artificially [188].
Two main unsupervised 3D classification methods have been described. The first one, introduced by Penczek et al. [189], is based on focusing the classification to regions that show high variance. These high variance regions are identified and localized using 3D variance maps. The procedure for focused classification begins with the calculation of a 3D variance map [190]. A 3D spherical mask is then placed in the site corresponding to the high variance region, and subsequently projected into a set of 2D masks. Next, the experimental projections are sorted into angular sets according to the closest projection direction as the 2D masks. Classification of each set is then performed based on the regions outlined by the respective masks. In the case the structures differs with regard to ligand or no ligand bound, high pixel values indicate the presence of ligand and low pixel values indicate the absence of ligand. Two groups will therefore be generated from the classification, one with classes containing images with high pixel values inside the masked regions and another with low pixel values inside the masked regions. These two groups are used to reconstruct two 3D volumes that are subsequently used in a multi-reference 3D projection alignment procedure (i.e. supervised classification) against all projection images in order to refine the angles.

The unsupervised 3D classification method used for the ribosome project (paper III and IV) is based on the maximum-likelihood (ML) principle. The underlying idea in the ML method is to find the statistical model that maximizes the probability to observe the experimental data. Mathematically this is achieved by optimizing the following log-likelihood function [191]

$$L(\Theta) = \sum_{j=1}^{I} \ln \left( \sum_{k=1}^{K} \int P(X_j | k, \varphi, \Theta) P(k, \varphi | \Theta) d\varphi \right)$$

where the model $\Theta$ consists of $K$ number of 3D structures and some parameters that describe the probability to observe different projections through these 3D structures. The first factor $P(X_j | k, \varphi, \Theta)$ represents the probability of observing image $X_j$ when viewing class $k$ in orientation $\varphi$. The second factor $P(k, \varphi | \Theta)$ is the probability of observing class $k$ in projection direction $\varphi$. The optimization of the log-likelihood function is performed iteratively by introducing a weighting factor, which describes the probability of viewing class $k$ in projection direction $\varphi$ when presented to image $X_j$. This weighting factor, which is updated for each image in every iteration, determines the contribution each projection image makes to every class and projection direction. The XMIPP software package [192, 193] was used for the ML classification, and the procedure is as follows: Projections are generated from the $K$ different reference structures. Subsequently, all the experimental images are compared with these reference projections. Based on the resemblance, probabilities are calculated for each image in all the different projection directions. These probabilities are then used to generate new probability-weighted 2D averages. Finally the 2D averages are combined into $K$ different 3D reconstruction, which will be used in the next iteration. In addition, the standard deviation for e.g. $P(X_j | k, \varphi, \Theta)$, as well as the probabilities for the different projection directions, are also updated for the next iteration.

To generate the initial reference maps, a single iteration of the ML classification for $K$ randomly split subsets of the experimental data is performed with a low-resolution volume that show the general features of the average structure as a starting model. For
paper III and IV, a ribosome structure low-pass filtered to 80 Å was used as the starting model [191].

Figure 18 shows the first two ML classifications performed on the ribosome dataset from paper III. In the first ML3D run [194], the 50S subunit particles and the intact 70S ribosome particles were separated. Subsequently, a second classification was performed using only the intact 70S ribosome particles. Likewise, in paper IV the 50S and 70S particles were also separated before continuing with the intact 70S ribosome particles, but in this case the first classification was carried out in Fourier space (MLF3D) [195]. The classes from paper III were refined separately by projection matching. Due to the limited number of projections in the individual classes, in combination with poor SNR of the projection images, no refinement was performed in paper IV.
Figure 18. Flow-chart showing the first two ML classifications of the tmRNA-SmpB-70S dataset from paper III. The number of particles in each class is indicated below the corresponding reconstructions.
CONCLUSIONS

Paper I – A comparison of the two *R. thomasi*ana hemocyanin isoforms revealed that up to a resolution of 19 Å, the two assemblies show no structural differences. Three previously unreported connections were identified in both complexes, one within the subunit dimer and two between adjacent subunit dimers, which could contain binding sites for the ions Ca$^{2+}$ and Mg$^{2+}$.

**My contributions:** Design of research (together with PJBK), specimen preparation for electron microscopy, data collection (together with TT), analysis of results (together with PJBK and HE), writing of paper.

Paper II – Electron micrographs of vitrified *P. furiosus* chaperonin samples revealed ring-shaped particles, and the density map obtained from subsequent image processing suggest that the 16-mer assembly with an 8-fold symmetry is in a partially closed/open state.

**My contributions:** Design of research (together with PJBK), protein purification (together with MK), specimen preparation for electron microscopy (together with PP), data collection (together with PP), analysis of results (together with PJBK and MK), writing of paper (together with MK and PJBK).

Paper III – The structure of the tmRNA·SmpB-bound ribosome complex in the post-accommodated state showed that the TLD·SmpB part of the tmRNA·SmpB complex mimics native tRNAs in the A site of stalled ribosomes. In addition, the tmRNA arc formed by the pseudoknots remains highly structured after the structural re-arrangements that occur during accommodation.

**My contributions:** Design of research (together with ML), specimen preparation for electron microscopy, data collection, analysis of results (together with SHWS, ML and HH), writing of paper (together with MP, ME, SHWS, ML, HH).

Paper IV – The pseudoknots of tmRNA in the double-translocation tmRNA·SmpB·70S complex is located outside the ribosome in positions similar to the ones assigned for the pseudoknots in the post-accommodated state. Furthermore, the maps show indications of dynamic behavior around the hinge regions connecting the individual pseudoknots.

**My contributions:** Design of research (together with ML), specimen preparation for electron microscopy, data collection, analysis of results (together with SHWS, MP, HH and ML), writing paper (together with MP).
ACKNOWLEDGEMENTS

At last I have made it to the end of my PhD studies. It has not been easy and I could not have pulled it off without all the people who supported and believed in me. So now I would like to thank them, in special:

Hans Hebert for making me part of your lab these years. I would also like to thank you for your encouragement and support, especially at times when I doubted myself.

Martin Lindahl for sharing the ribosome project with me, and for supervising me during the second part of my PhD studies. I am also thankful for your patience with all my questions. Not only were they never-ending, but they were also repeated several times.

Philip Koeck, also for supervising me and for your patience with my questions. But mostly I am grateful that you asked me to join the EM-group.

Past and present members of the EM-group, especially; Anna-Karin - thank you so much for being such a good friend and colleague during these years! I could not have asked for a better companion along this winding road. Caroline for being the glue in the group that keeps things together. Pasi for all the help with the microscope. Hans E and Dominika for your pep talk at the end when things were a bit rocky. Urban, for always helping me more than you needed to. Ulrika, Priya, Karina, Peter, Tomas, Björn, Romny, Supra and Qie.

My collaborators: Krassimira Idakieva for providing me with hemocyanin samples. Rudolf Ladenstein and Mikael Karlström for letting me be part of the chaperonin project which gave me the chance to do some protein purification work. Måns Ehrenberg, Misha Pavlov and Natalia Ivanova for a nice collaboration on the ribosome project. José María Carazo and Sjors Scheres, also for the nice collaboration on the ribosome project and for introducing me to your awesome software package XMIPP!

Yoshi – thank you so much for welcoming me to your lab! That month really was one of the best times of my life! I also want to thank Mie for all the help with the paperwork and accommodation, and for the nice morning chats. Chris for preventing many lost in translation scenarios. Tomohiro for teaching me all there is to know about how to pre-irradiate grids and also about mountain climbing.

Kurt Berndt for your great lectures at Södertörns Högskola. Thank you for opening my eyes to structural biology.

Other colleagues at NOVUM; Tobias for being a good friend that is sweet, warm, funny and smart - all at the same time! Sofia, Boel, Katarina and Linda for brightening up our corner of level 7. Sunny, for our nice talks and for attending all my dance classes. Patrick, also for our nice talks and for teaching me that stress is just a waste of time and energy. Merita and Birgitta for always being happy and cheerful whenever we
run into each other in the hallways. Johan, Jianxin, Xiaofeng, Ivan, Joelle, Luca, Magnus.

My students at Södertörns Högskola, best wishes to all of you!

All the staff at NOVUM, especially; Inger M for teaching me that sometimes it’s not me, it’s ‘them’. Anders and Erik for coming to my rescue every time my computer was giving me a hard time. Kristina for taking such good care of us PhD students!

Karl-Erik, tusen tack för all hjälp med det administrativa inför disputationen! Det hade inte gått lika smidigt utan dig.


Agneta, för att du fick mig att inse hur jag vill att mitt liv ska vara.

Pia för att du är en fantastisk vän samt den bästa kusinen man kan önska sig! Din självständighet, glädje samt intellekt är något jag strävar efter att uppnå. Din humor har jag redan…. ☺

Min syster Grace - tack för att du inte gav upp hoppet om mig! Också en stor kram för att du alltid fanns där och delade alla upp- och nedgångar i mitt liv. Lovar att inte klona en människa!

Min bror Alex - du är en outtömlig källa av positiv energi! Ditt sätt att se allt från den ljusa sidan är verkligen inspirerande.

Mamma och pappa, för ert stöd samt alla visdomsord ni delar med er.

Min sväger Magnus, min svägerska Veronika samt syskonbarnen Kevin, Viggo och Axel, ni är verkligen ett härligt tillskott till min smått galna och vilda familj.

Anders – gigantisk puss och kram för ditt stöd, den har varit ovärderlig! Utan dig hade jag aldrig klarat av att slutföra det hela. Tack för att du finns i mitt liv!
REFERENCES

[17] Lieb, B., Altenhein, B., Lehner, R., Gebauer, W., Markl, J. Subunit organization of the abalone Haliotis tuberculata hemocyanin type 2 and the cDNA sequence coding for its functional units d, e, f, g and h, European Journal of biochemistry. 265 (1999) 134-144.
[22] Harris, J.R., Scheffler, D., Gebauer, W., Lehner, R., Markl, J. Haliotis tuberculata hemocyanin (HtH): analysis of oligomeric stability of HtH1 and HtH2, and comparison with keyhole limpet hemocyanin KLH1 and KLH2, Micron. 31 (2000) 613-22.


Okochi, M., Matsuzaki, H., Nomura, T., Ishii, N., Yohda, M. Molecular characterization of the group II chaperonin from the hyperthermophilic archaea Pyrococcus horikoshii OT3, Extremophiles. 9 (2005) 127-34.


Lee, S., Ishii, M., Tadaki, T., Muto, A., Himeno, H. Determinants on tmRNA for initiating efficient and precise trans-translation: some mutations upstream of the tag-encoding sequence of Escherichia coli tmRNA shift the initiation point of trans-translation in vitro, Rna. 7 (2001) 999-1012.


on the bootstrap 3D variance analysis, and its application to EF-G-dependent

[190] Penczek, P.A., Yang, C., Frank, J., Spahn, C.M. Estimation of variance in
single-particle reconstruction using the bootstrap technique, J Struct Biol. 154

[191] Scheres, S.H., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P., Frank, J.,
Carazo, J.M. Disentangling conformational states of macromolecules in 3D-EM

la Fraga, L.G., Vaquerrizo, C., Carazo, J.M. Xmipp: An Image Processing

Scheres, S.H., Carazo, J.M., Pascual-Montano, A. XMIPP: a new generation of
an open-source image processing package for electron microscopy, J Struct

Image processing for electron microscopy single-particle analysis using

[195] Scheres, S.H., Nunez-Ramirez, R., Gomez-Llorente, Y., San Martin, C.,
Eggermont, P.P., Carazo, J.M. Modeling experimental image formation for
1167-77.