Initiation and Inhibition of Protein Biosynthesis – Studies at High Resolution

Raz Zarivach[1], Anat Bashan[1], Frank Schluenzen[2], Joerg Harms[2], Marta Pioletti[3], François Franceschi[3] and Ada Yonath[1,2*]

[1] Dept. Structural Biology. Weizmann Inst. of Science, Rehovot, Israel, [2] Max-Planck-Res. Unit for Ribosomal Structure, Hamburg, Germany, [3] Max-Planck-Inst. for Molecular Genetics, Berlin, Germany

Abstract: Analysis of the high resolution structure of the small subunit from *Thermus thermophilus* shed light on its inherent conformational variability and indicated an interconnected network of features allowing concerted movements during translocation. It also showed that conformational rearrangements may be involved in subunit association and that a latch-like movement guarantees processivity and ensures maximized fidelity. Conformational mobility is associated with the binding and the anti association function of initiation factor 3, and antibiotics interfering with prevent the initiation of the biosynthetic process. Proteins stabilize the structure mainly by their long basic extensions that penetrate into the ribosomal RNA. When pointing into the solution, these extensions may have functional roles in binding of non-ribosomal factors participating in the process of protein biosynthesis. In addition, although the decoding center is formed of RNA, proteins seem to serve ancillary functions such as stabilizing ist required conformation and assisting the directionality of the translocation.

THE SMALL SUBUNIT AND ITS GLOBAL ORGANIZATION

Ribosomes are the universal cellular organelles catalyzing the sequential polymerization of amino acids according to the genetic blueprint, encoded in the mRNA. They are built of two subunits that associate for performing protein synthesis. The prokaryotic ribosomal subunit (called 30S) has a molecular weight of $8.5x10^5$ Dalton and contains one RNA chain of over 1500 nucleotides and 20 proteins. It has key roles in the initiation of the process; in decoding the genetic message; in discriminating against cognate, non- and near-cognate aminoacylated tRNA molecules; in controlling the fidelity of codon-anti-codon interactions and participates in mRNA/tRNA translocation.

The high-resolution structure of the small subunit from *Thermus thermophilus* has been determined by us (Fig. **1** and in [1, 2]) as well as by the group of V. Ramakrishnan [3]. The emerging particles from both electron density maps are very similar and contain the morphological features familiar from cryo-EM reconstructions. As seen in the secondary structure diagram [4], all main structural features of the small subunit radiate from a junction combining them. The traditional sub-divisions are a "head", a "neck" and a "body" that contains a "shoulder" and a "platform". All these features, as well as the overall shape and dimensions of the subunit, have been observed even in the early electron microscopy studies [5,6].

*Address correspondece to this author at the Dept. Structural Biology. Weizmann Inst. of Science, Rehovot, 76100, Israel; Tel: -972-8-9343028; Fax: -972-8-9344154; E-mail: ada.yonath@.weizmann.ac.il

Three long helices run parallel to the vertical axis of the body. Among them, two are located on the rather flat surface that faces the 50S subunit: H44, in the middle of the flat surface, and H16/H17 that form part of the shoulder, on the left. The longitudinal helical elements are likely to transmit structural re-arrangements, correlating events at the particle's far ends with the cycle of mRNA translocation at the decoding region, located at the middle of the meeting point between the body and the neck. Transverse features, placed like ladder rungs between them, link the three longitudinal helices. Principal among these transverse helices is an inclined line extending from the shoulder to the platform. The head contains most of the 3' region of the 16S RNA. It interacts with mRNA, elongation factor G, tRNA and several antibiotics. It consists mainly of short helices, in marked contrast to the long features of the body. It has a bi-lobal architecture, with a longer helix (H34) serving as the bridge between hemispheres. It joins the body through a slender connection, made of a single RNA helix, H28, which appears to act as a hinge.

The shoulder is built of four mutually perpendicular helices that create a compact fold. This highly structured region consists of a co-stack of over 120 Å long helical elements arranged around two key junctions that radiate towards the main elements involved in mRNA and tRNA binding. Although only half of it is well conserved, the shoulder forms the lower side of an elongated, curved channel, which we assigned as the entrance side of the path of the mRNA. A latch, which can be described as a non-covalent body-head connection, formed by the shoulder and the lower part of the head, is the feature that designates the

1389-2037/02 $35.00+.00

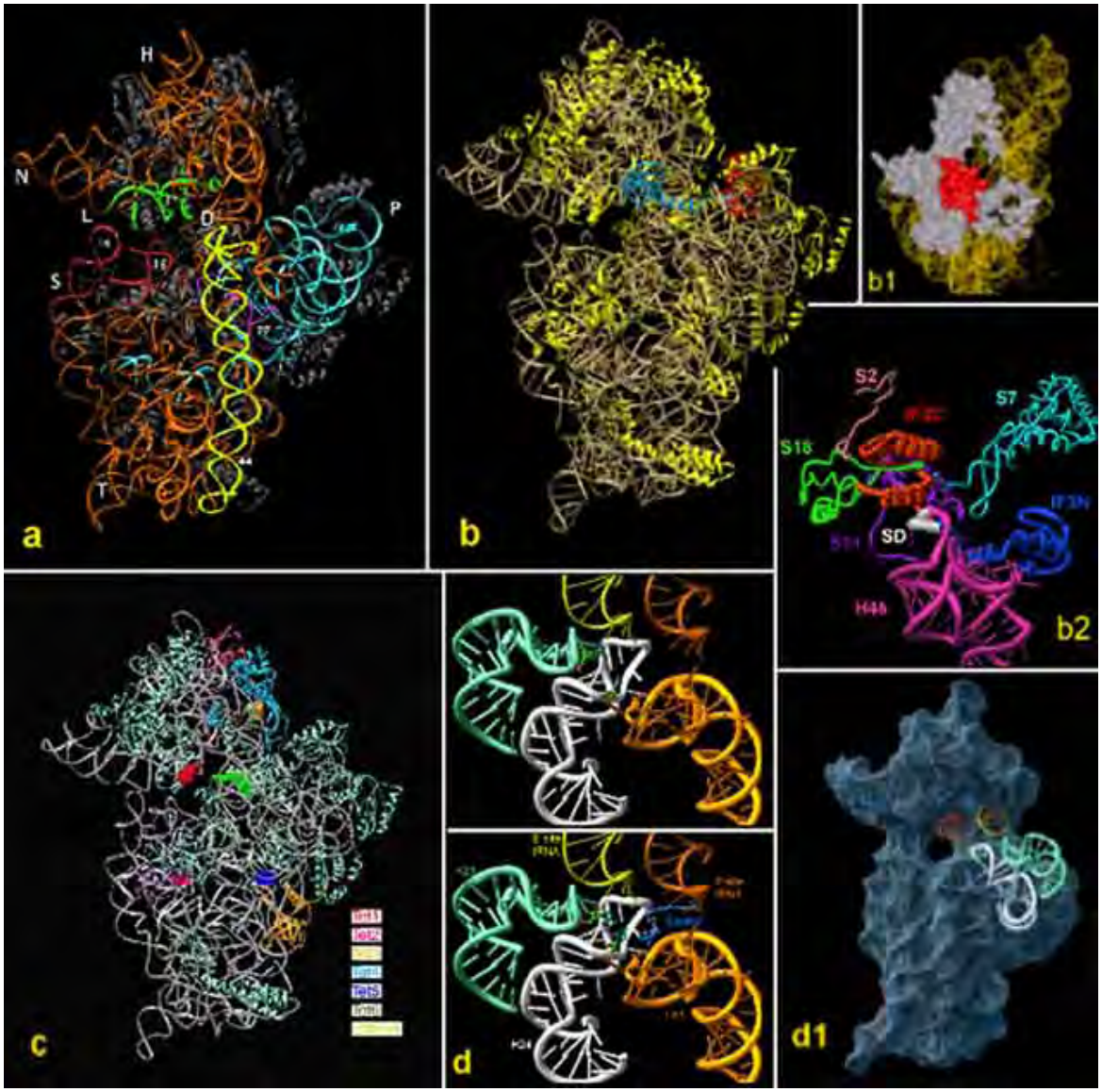

Fig. (1). (a) The three dimensional structure of the small ribosomal subunit from *Thermus thermophilus* (T30S), seen from the side that faces the large subunit within the ribosome- called also "front view" H=head, B=body, P=platform, S=shoulder. The RNA is shown in silver and the proteins in blue. The features participating in the decoding and the translocation are highlighted.

(b) The binding site of IF3C red and IF3N are marked on the 30S subunit shown with its RNA in brown and the proteins in yellow. (b1) a space filling representation showing the location of IF3C in the solvent side of the platform (in red) within the 30S proteins (in gray) and RNA in dark-yellow. Here the view is opposite to that shown in (a) and (b). (b2) The detailed environment of IF3.

(c) The binding site of edeine and tetracycline seen from the same direction as in (a).

(d) Detailed view of the edeine binding site of edeine (in blue). Top: before and bottom – after the binding. Note the newly formed base pair in green).

(d1) The small subunit shown at about 75 degrees rotation (around the vertical axis of the particle) compared to the view of Figure 1. The mRNA channel and the P-site tRNA (white) and E-site tRNA (cyan) are seen. Made with RIBBONS (62).

entrance to the mRNA channel. This latch facilitates mRNA threading and provides the special geometry that guarantees processivity and ensures maximized fidelity. It controls the entrance to the mRNA channel by creating a pore of varying diameter and its relative location may be dictated by the head twist.

The decoding region contains features from the upper part of the body and the lower part of the head. Mapping the conserved nucleotides in the 16S RNA on our structure showed remarkable conservation around this region, in accord with the universality of the decoding process. This region contains the upper part of H44, H45, the 3' and 5' ends of the 16S RNA, the switch helix [7], H27, the central pseudoknot, the stem loop of the shoulder helix, H18, and the main transverse head element, H34 (Fig. **1**). The most prominent feature in the decoding center is the upper portion of H44, which bends towards the neck and forms most of the intersubunit contacts in the assembled ribosome [8,9]. Its upper bulge forms the A- and P- tRNA sites. The switch helix, H27, which packs groove-to-groove with the upper end of H44, can undergo changes in its base-pairing scheme that may induce global conformational rearrangements [7].

We placed mRNA and tRNA in the decoding center by reference to the structures of the complex of the entire ribosome with three tRNA molecules that were determined at 7.8 Å and 5.5 Å resolution [8,9]. An important conclusion from this placement, which could be drawn even at the early stages of structure determination, is that the anticodon loops of the A- and P-site tRNAs and the codon segments of the mRNA do not contact any ribosomal proteins. The configuration of the 16S RNA therefore determines the arrangement of the interacting elements and enforces precision in codon-anticodon interactions. As seen below, proteins assist in the initiation and the translocation process by maintaining their accuracy and maximizing their efficiency.

ON THE CONFORMATIONAL MOBILITY OF THE SMALL RIBOSOMAL SUBUNIT

The ribosome is a precisely engineered molecular machine performing an intricate multi-step process that requires smooth and rapid switches between different conformations. Unlike the rather compact 50S subunit, the small subunit is built of loosely attached domains and contains structural elements that allow local rearrangements as well as the global motions required for its function. Its conformational variability has been detected by cryo electron microscopy [10,11], by surface RNA probing [12], by monitoring ribosomal activity, and by the analysis of the high-resolution structures of the small subunit complexes [1-3,13]. The conformational variability also explains why all the available cryo-EM reconstructions were not useful for extracting initial phase sets for the small subunit, whereas similar searches were performed successfully for the whole ribosome and for its large ribosomal subunit [14]. Our analysis of the 30S structure led us to suggest an interconnected network of features that could allow concerted movements during translocation [15]. The major contributors to the conformational mobility of the small subunit are described below.

The principal component of the subunit interface region is the long penultimate helix (H44). This feature consists of over a 100 nucleotides, of which the only evolutionarily conserved part comprising of less than two dozen nucleotides that are involved in decoding and in P-site tRNA binding. This part is situated at the upper end of H44. Examination of the temperature factors of the entire RNA fold of the small subunit detected a high degree of freedom in the central zone of H44. This indicates that inter-subunit association requires conformational rearrangements, similar to recent finding in the large subunit [16]. Head mobility was confirmed by Molecular Replacement studies**,** performed with AmoRe [17] with data collected from the crystals of T30S that diffract to low resolution together with the high resolution model, divided to three entities: the entire particle, the body and the head. These studies indicated that the low resolution crystals contain at least one conformation that differs from that of the crystals diffracting to high resolution. The pivotal point for this movement may be at the connection between the head and the neck. It is rather close to the binding site of the antibiotic spectinomycin that is known to hamper the head twist by trapping a particular conformation [18].

The latch is an additional feature that has more than one conformation. It is likely that the switch between the different latch conformations guarantees processivity and ensures maximized fidelity. The upper boundary of the mRNA channel is made by the head helix H34, and its relative location is dictated by the head twist. In addition, internal head axes may be utilized for facilitating global movements associated with protein biosynthesis. The platform is a large structural element that is known to shift when subunits associate and along the translation cycle [10]. Below we discuss the critical contribution of its conformational mobility to the initiation step. On its solvent side it contains the 3' end of the 16S RNA that was shown to be highly flexible, so that under specific conditions it can fold back on itself [3]. Part of the 3’ end is the region that pairs with the trigger (SD or Shine-Dalgarno) sequence in the mRNA. This interaction is critical for initiation of protein biosynthesis, a step of major importance in governing the accurate setting of the reading frame, as it facilitates the identification of the start codon of the mRNA.

SELECTION THE CORRECT INITIATOR TRNA, INITIATION OF PROTEIN BIOSYNTHESIS AND CONFORMATIONAL MOBILITY

The initiation complex contains, in prokaryotes, the small subunit, mRNA, three initiation factors (IF1, IF2-GTP and IF3) and initiator tRNA. IF3 plays multiple roles in the formation of the initiation complex. It influences the binding of the other ligands and acts as a fidelity factor by destabilizing non-canonical codon-anticodon interactions. It also selects the start mRNA codon [19] and the correct initiator tRNA to be positioned at the P-site (in prokaryotes, the f-met-tRNA). It stabilizes the binding of the fMet-

tRNA/IF2 complex to 30S and discriminates against leaderless mRNA chains [20]. IF3 also acts as an anti-association factor because it binds with a high affinity to the 30S subunit and shifts the dissociation equilibrium of the 70S ribosome towards free subunits, thus maintaining a pool of 30S [21]. It is a small basic protein of about 20kD, consisting of C- and N-terminus domains (IF3C and IF3N, respectively) connected by a rather long lysine-rich linker region. The structure of the entire protein has not been determined, but NMR [22,23] and X-ray structures of the N- and C-terminal domains have been reported [24,25]. Although the flexible inter-domain linker appears as a rigid alpha-helix in the crystals containing it and IF3N, NMR and biochemical results show clearly that even under physiological conditions, it is partially unfolded [26,27]. Furthermore, it has been suggested that the flexibility of the linker is related to the function of IF3 in the formation of the initiation complex, and thus has a functional importance.

The crystal structure of T30S in complex with IF3C, determined by us [2], shows that IF3C binds to the 30S particle at the upper end of the platform on the solvent side (Fig. **1**), near the anti-SD region. This location confirms the results of NMR and mutagenesis of the IF3 molecule [28,29] and almost all the crosslinks, footprints and protection patterns that were reported for the *Eschereichia coli* system [30-32]**.** The binding of IF3C at this site could affect the conformational mobility of the platform that is essential for the association of the two ribosomal subunits to form a productive ribosome, consistent with biochemical observations indicating that IF3C prevents subunit association or promotes dissociation by influencing the conformational dynamics of the subunit. The spatial proximity of IF3C binding site to the anti-SD region suggests a connection between them. Both the binding of IF3C and the hybridization of the anti-SD sequence may limit the combined head-platform-shoulder conformational changes. Upon the detachment of the SD anchor, required at the beginning of the translocation process, the platform may regain its conformational mobility.

Cryo EM reconstruction of a complex of the small ribosomal subunit from *Thermus thermophilus* with IF3 from *Thermotoga* localized IF3C at the subunit interface, suggesting that the anti-association activity of IF3 is the product of physical blockage at the interface between the two subunits [33]. On the other hand, EM studies on rat liver 40S in complex with the eukaryotic initiation factor 3 (eIF3) located eIF3 on the solvent side of the upper edge of the platform [34], in a region comparable to our findings. Although some aspects of the initiation process of protein biosyntheses were found to be different in eukaryotic and prokaryotic systems [35], neither of them indicates different locations of initiation factor 3. The consistency between our results and the location of the eukaryotic initiation factor may indicate that the main concepts underlying the initiation process and governing the anti-association properties of the initiation complex have been evolutionarily conserved.

The signal for the dislocation of the SD hybridized region may be linked to verification of codon-anticodon complementarity, which has been attributed mainly to IF3N [36]. We docked IF3N manually, according to its proposed function, the constraints posed by the position of IF3C and the existing biochemical data, as well as automatically, using MOLFIT [37]. The location found by both methods is in close proximity to the P-site. In this position IF3N interacts all helices known to be involved in the peptidyl-tRNA binding. The binding of IF3N leaves only a limited space for P-site tRNA, and requires small conformational changes for simultaneous binding of both IF3N and P-site tRNA. Hence, it seems that IF3N discriminates against non-cognate tRNA by space exclusion principles. Our results are, therefore, consistent with the proposal that the linker acts as a strap between the two domains.

EDEINE - AN ANTIBIOTIC INTERFERING WITH INITIATION BY LIMITING THE PLATFORM MOBILITY

Edeine is a universal antibiotic that has no clinical relevance since it inhibits the initiation of protein synthesis in all phylogenetic kingdoms [38,39]. It is a peptide-like compound, containing a spermidine-type moiety at its C-terminal end and a beta-tyrosine residue at its N-terminal end [40]. A subset of the 16S rRNA nucleotides protected by the P-site tRNA [31] and overlaps with those of kasugamycin and pactamycin [42,43]. The universal effect of edeine on initiation implies that structural elements important for the initiation process are universally conserved in prokaryotes and eukaryotes.

Using co-crystals of edeine and T30S, we found that edeine binds in the platform between the loop of Helix 24 and Helix 45 (Fig. **1**, and in [2]). In this position it might affect the binding of the linker and the N-terminal domain of IF3, as well as the 30S mobility, the interaction of the 3' end with IF3C and the interaction of the 30S and 50S subunits. By physically linking these four helices, critical points for tRNA, IF3, and mRNA binding, edeine could lock the small subunit into a fixed configuration and hinder the conformational changes that accompany the translation process [10,44]. The binding of edeine to the 30S subunit induces the formation of a new base pair between, connecting H24 and H23. H23 plays an important role in the binding of the C-terminal domain of IF3, and H24 are directly involved in 30S-50S subunit association). The newly induced base pair alters the mRNA path and would impose constraints on the mobility of the platform. Hence, it is conceivable that the formation of this base pair interferes substantially with the initiation. Thus, our data suggest that the initiation process is the main target of this universal antibiotic, and that edeine induces an allosteric change by the formation of a new base pair - an important new principle of antibiotic action, that fits nicely with our hypothesis for the mechanism of the initiation step.

Independent studies show that pactamycin, an antibiotic agent that shares a protection pattern with edeine, bridges H23b and H24a [45], the same helices that are linked by the new base-pair that is induced by edeine. This agent is known to interfere with the initiation process, and it is likely that besides reducing the mobility of the platform by locking these two helices, it also alters the mRNA path at the E-site. Like edeine, pactamycin interacts with the extended loop of

S7 – the upper border of the path of the exiting mRNA/tRNA complex, and its mode of interaction suggests that it may interfere with the pairing of the SD sequence or prevent it. The universal effect of edeine on initiation implies that the main structural elements important for the initiation process are conserved in all kingdoms [35]. Analysis of our results show that the rRNA bases defining the edeine binding site are conserved in chloroplasts, mitochondria, and the three phylogenetic domains. Among those are two conserved nucleotides along the path of the messenger. Hence, an additional effect of edeine may be preventing hybridization of the incoming mRNA.

TETRACYCLINE – A MULTI-SITE ANTIBIOTIC THAT BLOCKS A-SITE BINDING

Tetracyclines inhibit protein synthesis by interfering with the binding of aminoacylated tRNA (aa-tRNA) to the A-site of the 30S subunit [46]. They are products of the aromatic polyketide biosynthetic pathways and belong to a family of bacteriostatic antibiotics that act against a wide variety of bacteria. We identified six tetracycline-binding sites (numbered Tet-1, tet-2, etc., reciprocally to their relative occupancies) on the 30S subunit (Fig. **1**) with no common structural trait, all located in the head and the upper half of the body. The observation of multiple tetracycline binding sites was not unexpected, since several binding sites for tetracycline had already been suggested by biochemical experiments [2]. The six positions revealed by us can well explain the sometimes contradictory reported biochemical and functional data for tetracycline binding to the 30S subunits.

The Tet-1 binding site is within in a clamp-like pocket formed by the head helices H34 and H31, at the A-site tRNA binding site, consistent with toeprinting experiments in the presence of tetracycline [47]. Upon binding, the gap between these two bases becomes slightly wider. Aside from this local conformational change we did not observe any significant changes of the 30S subunit. Tet-1 interacts with the sugar-phosphate backbone of H34 through a magnesium ion, in a fashion similar to the interaction between the Mg-tetracycline complex and the tetracycline repressor tet(R) [48]. Tet-2 is located in a hydrophobic pocket of S4 and is the only tetracycline binding site not involved in interactions with the 16S rRNA. Tet-3 is buried inside H40, between its stem loop and its tetraloop. Tet-4 is located in a cavity at the bottom of the head. Tet-5 lies in a rather tight pocket confined by H11, H20, H27 and protein S17, with its hydrophilic side interacting with the phosphate-sugar backbone of the switch region of H27, with the bulged H11. Tet-6 is located in the vicinity of the E-site, interacting with the N-terminal end of S7A, via a Mg^{2+} ion in a similar fashion to that observed for the Tet-1 site.

The inhibitory action of tetracycline is mainly assigned to Tet-1, the site that interferes with the location where the A-site tRNA [8,9] was docked onto the 30S structure [1]. Thus, tetracycline can physically prevent the binding of the tRNA to the A-site. This mode of interaction is consistent with the classical model of tetracycline as an inhibitor of A-site occupation, and hence offers a clear explanation for the bacteriostatic effect of tetracycline. There are two ribosome-related mechanisms of tetracycline resistance in bacteria. Both relate to the Tet-1 site. In one, the resistance is mediated by ribosomal protection proteins [49] and in the other by the mutation on 16S rRNA [50]. Ribosomal protection proteins, such as TetM, TetO, and TetS, confer resistance only at low concentrations of tetracycline, and show some sequence and structural homology with the elongation factors G and EF-Tu [51]. It has been proposed that TetM binds to the A-site and upon GTP hydrolysis actively releases the tetracycline bound to it. The mutation could hamper the local base pairing system and lead to a conformational change that results in closing the Tet-1 binding pocket. These two tetracycline resistance mechanisms reflect the importance of the Tet-1 binding site in the antibiotic action of tetracycline and indicate that the physical blockage of the A-site tRNA binding by tetracycline bound at Tet-1, can account for the inhibitory action of tetracycline.

The presence of five additional binding sites, the biochemical evidence for different locations of tetracycline, accounting for Tet-1, Tet-4, Tet-5 and Tet-6 [2], and the low level of resistance conferred by the ribosomal protection proteins, demand more complex explanations about the possible functional relevance of the five additional sites. Although it is not certain that any of the five minor sites is involved in tetracycline action, we hypothesize that, with the exception of the tetracycline at the Tet-3 site, these sites could act synergistically to contribute to the bacteriostatic effect of tetracycline. The four proteins that come in contact with tetracycline, S4 (Tet-2), S7 (Tet-6), S9 (Tet-4 and Tet-6), and S17 (Tet-5) are primary rRNA binding proteins [52]. S4 and S7 are the two proteins that initiate the assembly of the 30S subunit [53]. Therefore, tetracycline binding at the Tet-2, Tet-4, Tet-5, and Tet-6 sites, may not influence the decoding process, but could disturb the early assembly steps of new 30S particles, contributing to the overall inhibitory effect of tetracycline. Tet-5 site, that was also revealed as a minor site by [42] is located between H11 and H27, the switch helix [7]. Tetracycline bound to Tet-5 position could limit the mobility needed by H27 to switch between conformations, and in this way hamper the processivity of the translation.

ON THE FUNCTIONAL AND STRUCTURAL RELEVANCE OF RIBOSOMAL PROTEINS

The striking architecture of the small subunit allows for substantial domain mobility. Yet, the individual structural elements are rather stable. The features that contribute to the local stability include specific RNA folds, by a high G-C content at the rims of strategically located junctions and by the ribosomal proteins. Almost all ribosomal proteins are peripheral, located on the particle's surface, at its solvent side, indicating that they may have evolved at later stages to fine tune the complex tasks and intricate recognitions required for the decoding process. The involvement of proteins in the stabilization of the structure is mainly through their long extensions that penetrate into rRNA regions and serve as molecular linkers, struts and supports. Protein extensions were seen in several protein/nucleic acid complexes, such as viruses [54], nucleosome [55].

Table 1. The Frequency of Selected Amino Acids Within the 30S Proteins

Frequency within all 30S proteins (Total 2289)

TRP	HIS	ASN	GLN	ARG	LYS	
15	40	61	86	279	195	No.
0.7	1.8	2.7	3.8	12.3	8.5	frequency

Frequency within the extensions (Total 323)

TRP	HIS	ASN	GLN	ARG	LYS	
1	8	9	16	55	40	No.
0. 1	2.5	2.8	4.9	17.0	12.4	frequency

Frequency within the parts of the extensions that bind rRNA

TRP	HIS	ASN	GLN	ARG	LYS	
1	3	2	2	19	12	No.
1.9	5.7	3.8	3.8	35.8	22.6	frequency

Frequency within globular domains (Total 1966)

TRP	HIS	ASN	GLN	ARG	LYS	
14	32	52	70	225	156	No.
0.7	1.6	2.6	3.6	11.4	7.9	frequency

Frequency within typical non-ribosomal proteins

TRP	HIS	ASN	GLN	ARG	LYS	
1.3	2.2	4.4	4	5.7	5.7	frequency

Summary

	HIS	ASN	GLN	ARG	LYS
ribosomal/non-ribosomal	0.79	0.61	0.94	2.14	1.49
glob/all	0.93	0.99	0.95	0.94	0.93
Exten/all	1.42	1.05	1.32	1.40	1.45
Exten/glob	1.52	1.05	1.39	1.49	1.56
% in exten, bound to RNA	37.5	22.2	12.5	34.5	30

(*) Ribos=ribosomal, glob=globular, exten=extensions

We defined a protein extension as a part of a protein that does not make contacts with the globular parts of it. Extended loops may have interactions within their own members. Most of the ribosomal proteins are basic. Table **1** shows some comparisons between all the 30S proteins and typical non-ribosomal proteins. The total amino acid content

of the small subunit ribosomal proteins has 1.5- and 2.13-fold excess of lysines and arginines, respectively. Higher ratios were obtained for the protein extensions. As expected, a large number of the basic amino acids of the extensions are bound to the ribosomal RNA (Table **1**). The tails of a few proteins are pointing into the solution [2]. Some of these may make crucial contributions to the efficient binding of non-ribosomal factors participating in the process of protein biosynthesis.

Protein S17 may shed light, at the molecular level, on one aspect of the intriguing question of how thermophilic organisms manage to live at high temperature. This protein possesses a long C-terminal tail, consisting of over 22 amino acids only in thermophilic bacteria (*T. thermophilus T. aquaticus* and *Thermotoga maritima)*. Such a long tail has not been detected in any mesophile. This protein is located near the lower side of the platform. Its long alpha helical tail is wrapped along the surface of the body, within a narrow groove that separates the platform and the body, curling around and following the contour of the platform, one of the ribosomal components that moves during translocation. The long tail of protein S17 is situated in a position that can act as a physical stop to block uncontrolled sliding that may occur at elevated temperature. This hypothesis is supported by the temperature sensitivity of a mutant of *T. thermophilus* lacking protein S17 [56].

Although the major functional features of the small subunit consist mainly of RNA elements, some proteins have functional relevance. Proteins S5, S3 and S8 are part of the walls of the mRNA channel. These, together with S4 and S12 may contribute to the fidelity and the directionality of the translocation.

Protein S12 is located at the RNA-rich surface that interacts with the large subunit. Protein S7, located at the rims of the subunit interface region, is part of the mRNA/tRNA exit path. The extensions of protein S13 may have both functional and structural roles, as they create an intersubunit bridge. Flexible tails that are not involved in the stabilization of the structure, but point towards the solution or lie loosely on the ribosomal surface, may be involved in binding factors and substrates [16]. An example is protein S18 of which the N- and C-terminal extensions are rather flexible and appear to act as tentacles, which enhance the binding and the placement of IF3C.

Several globular proteins or their parts may assist in the folding of the rRNA and shaping the structure, as it is likely that the proteins and the RNA co-assemble, or even dictate the ribosome fold. Over three decades ago Nomura and coworkers showed that the 30S ribosomal subunit from *E coli* can be reconstituted *in vitro* from individually purified ribosomal proteins and 16S ribosomal RNA, and that for efficient assembly, sequential addition of the ribosomal proteins, is preferred. Consequently the cooperative nature of the assembly process was determined, and an "assembly map" was constructed [52,57]. Recent studies indicate that defined ribosomal constructs, such as the central domain or the 5' domains, are capable of independent assembly [58,59] and analysis of the high resolution structure of the small subunit showed that proteins are often bound at the junctions in a fashion that may dictate or assist achieving the locale conformation.

A STRUCTURAL APPROACH FOR THE IN-VITRO ASSEMBLY

A structural approach for the in-vitro assembly that presents a plausible scheme for the events that may take place during the assembly of the head is described in Fig. [**2**]. Our scheme deviates slightly from the traditional map, in which the assembly of S20, a protein that was found at the bottom end of the particle, was required for the incorporation of S13, which is located at the head. It also incorporates protein Thx [60], a small polypeptide that, so far was detected only in T30S.

This scheme is based on the assumption that the assembly progresses in discrete steps; that the 16S RNA is not expected to fold by itself; that the ribosomal proteins play an important role in controlling the proper folding of the rRNA; that structured and unfolded regions of RNA and proteins exist simultaneously during the assembly process; that individual proteins interact sequentially with different regions of the RNA; that the proteins fold while being assembled, so that their long extensions acquire their conformation while interacting with the RNA; and that it is likely that the globular domains of the proteins are formed before the long extensions fold together with the neighboring partially folded RNA regions, and that the extensions can attach to both partially and fully formed domains.

The head emerges and terminates at the helix 28 (called the "neck"), which is the only covalent connection between the two main domains of the small ribosomal subunit and is highly conserved. In principle, the RNA segment that comprises the fold of the head, a chain of more than 340 nucleotides, could fold in all directions. However, in reality, most of the head is situated above the shoulder, to enable the creation of the tRNA sites and the upper part of the mRNA path. Only a small part of the head is built above the platform, composed mainly of the ribosomal protein S7 and part of its "minor domain" are involved in E-site. Protein S7, which is an assembly initiator protein [53], serves as a physical barrier (Fig. **2b**), enforcing constraints to limit structural build-up toward the platform's direction (to the right, looking from the subunit-interface side as in Fig. **1**). This significant influence of S7 on the architecture of the head may result from the tight interaction between its N-terminal extension with the major groove of the upper part of helix 28. Interestingly, protein S7 was suggested to be a "structure organizer" based on its fold in isolation, which was found to be similar to a protein acting as a DNA architectural factor [61]**.** The neck (H28) ends at the junction with H29 and H43, and H29 runs into a junction with H30, H41, and H42. Protein S7 interacts with both junctions, and it is conceivable that it influences their fold. Once a favorable local structure arrangement has been created, protein S9 and S19 may be attached in parallel. Protein S19 (Fig. **2c**) can provide a "molecular clamp" between the features that are later positioned in the two hemispheres of the head. It attaches the loops of helices 42 and 30 to helix

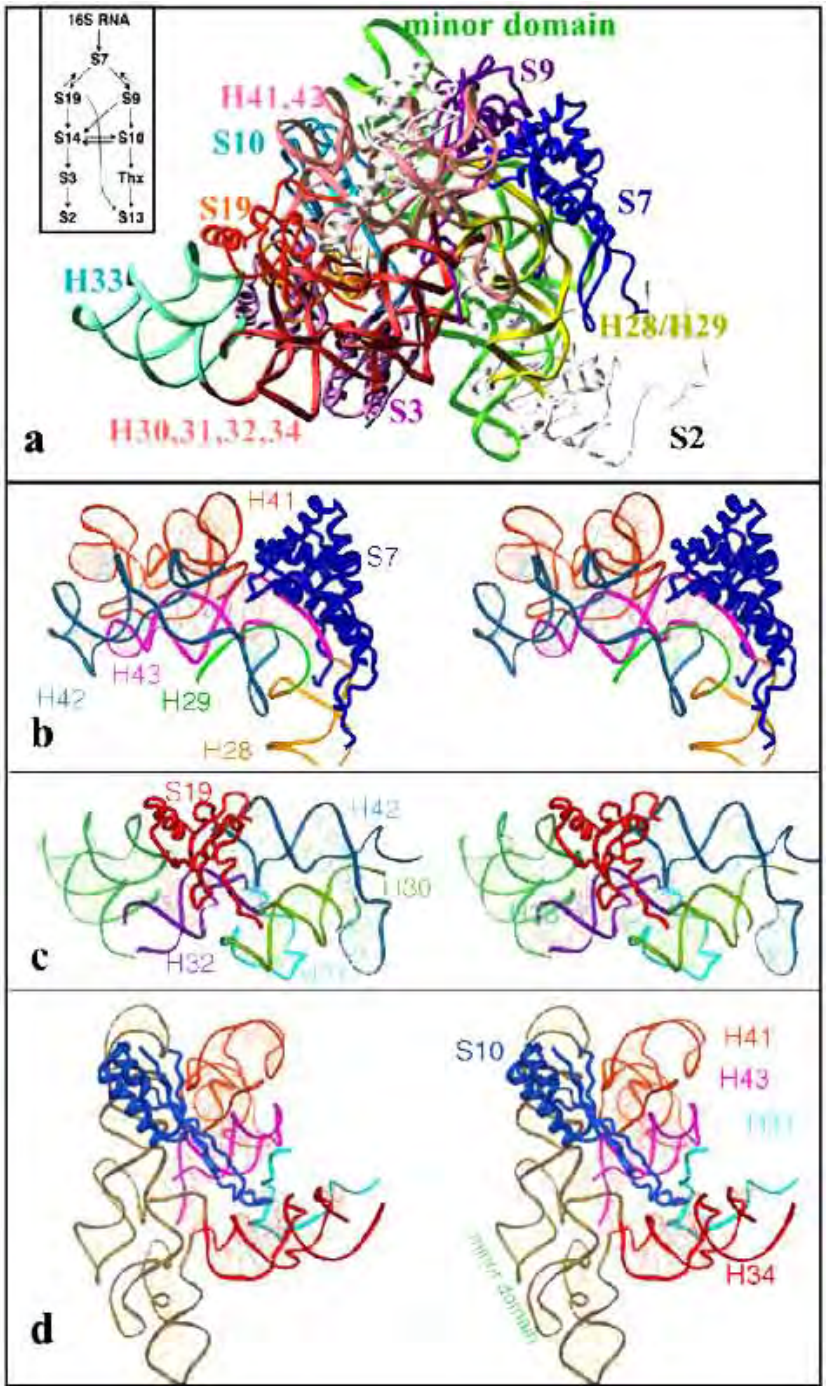

Fig. (2). Selected steps in the assembly of the head.

The entire head is shown in (a). Insert: the proposed diagram of the head assembly process. (b-d) The proposed progression of the assembly process. The components participating in the various steps are marked. Made with DINO [63].

33, one of the main components of the "nose", which in the mature 30S particle is positioned above the shoulder and its lower end comprises the upper side of the mRNA channel. S19 also interacts with H32 and H30 via their minor grooves, keeping them sandwiched to the upper layer, created by H42 and H33.

In the mature 30S particle, the globular part of protein S9 resides at the top of the head, and its long extension interacts with H30 and H34 and reaches the P-site tRNA position, at the lower part of the head. It also provides a contact from the main head components to the minor head domain. The latter, composed of helices H35-H40, is a rather isolated rod like feature, situated on the right side of the head (looking from the interface direction). Protein S10 can now join and complete the attachment of the minor and main head domains utilizing its extension to clips H31 and H34 together, thereby creating the lower head boundary (Fig. **2d**). In this way it enhances the contacts between the two parts of the head and leads to a definition of a structural pocket that is later utilized for binding of additional proteins. S10 also stabilizes the bend of H41, which interacts with protein Thx, which we suggest to assemble right after S10.

At this stage of the assembly process, most of the structural elements of the head have been formed and are ripe for final three-dimensional organization. Protein S13 binds H41 and H42, once the main structural elements have been assembled. It stabilizes the junction of helices H41-H42 at the solvent side, and its long tail may wrap around the features that are located mostly in the inter-subunit interface side. In the final stage S14 fills the head pocket, keeping together four helical features: the loop of H43, H34 and the connection between H31 and H32. Protein S3 is added, so that it covers S14 and S10, and while binding to H34, stabilizes the bottom connection of the minor domain to the main head. At this point the head is nearly fully assembled and protein S2 can bind to the lower part of the minor head domain.

CONCLUSIONS

The picture of the ribosomal small subunit that emerges from our analysis is of a precisely engineered machine, with shifts coupled over long as well as short distances to create a defined sequence of events at the decoding center. Thus, mutational changes at the foot of the subunit can affect the accuracy of translation via the long helices link the foot with the center of the subunit. The motions of the head may open and close the mRNA latch, these, together with the platform movements are likely to be correlated by the state of the switch helix, H27.

The decoding center itself is formed of RNA, and proteins seem to serve ancillary functions such as stabilizing required conformation and assisting the directionality of the translocation. The analysis of the structure of the complex of T30S with IF3C sheds light on the nature of the binding and action of this factor and indicates that the exterior protein-tails have important functional tasks that benefit from their significant flexibility. These tasks are quite different from those assigned to the protein extensions in the interior of the particle, which are assumed to be involved mainly in the stabilization of the ribosome structure.

ACKNOWLEDGEMENTS

We thank J.M. Lehn for indispensable advice, M. Pope for the heavy atom compounds, M. Wilchek, A. Tocilj and W. Traub for critical discussions, M. Eisensteinand E. Ben-Zeev for assisting with modelling and R. Albrecht, W.S. Bennett, H. Burmeister, C. Brune, C. Glotz, G. Goeltz, M. Kessler, M. Laschever, S. Meier, J. Muessig, M. Peretz, C. Radzwill, B. Schmidt, A. Sitka and A. Vieweger for contributing to different stages of these studies. These studies could not be performed without the cooperation of the staff of the synchrotron radiation facilities at EMBL & MPG beam lines at DESY; ID14/2 & ID14/4 at EMBL and ESRF and ID19/APS/ANL. The Max-Planck Society, the US National Institute of Health (GM34360), the German Ministry for Science and Technology (Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie Grant 05-641EA) and the Kimmelman Center for Macromolecular Assembly at the Weizmann Institute, provided support. AY holds the Martin S. Kimmel Professorial Chair.

REFERENCES

[1] Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. and Yonath, A. (**2000**) *Cell, 102,* 615-623.

[2] Pioletti, M., Schlünzen, F., Harms, J., Zarivach, R., Glühmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A. and Franceschi, F. (**2001**) *EMBO J., 20,* 1829-39.

[3] Wimberly, B., Brodersen, D., Clemons, W., Morgan-Warren, R., Carter, A., Vonrhein, C., Hartsch, T. and Ramakrishnan, V. (**2000**) *Nature, 407,* 327-339.

[4] Gutell, R.R., Subashchandran, S., Schnare, M., Du, Y., Lin, N., Madabusi, L., Muller, K., Pande, N., Yu, N., Shang, Z., Date, S., Konings, D., Schweiker, V., Weiser, B. and Cannone, J.J. (**1999**) http://www.rna.icmb.utexas.edu/.

[5] Stoeffler, G. & Stoeffler-Meilicke, M. (**1984**) *Annu. Rev. Biophys. Bioeng., 13,* 303-330.

[6] Lake, J.A. (**1985**) *Annu. Rev. Biochem., 54,* 507–30.

[7] Lodmell, J.S. and Dahlberg, A.E. (**1997**) *Science, 277,* 1262-1267.

[8] Cate, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N. and Noller, H.F. (**1999**) *Science, 285,* 2095-2104.

[9] Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., *et al.* (**2001**) *Science, 292,* 883-96.

[10] Gabashvili, I.S., Agrawal, R.K., Grassucci, R., Squires, C.L., Dahlberg, A.E. and Frank, J. (**1999**) *EMBO Journal, 18,* 6501-6507.

[11] Stark, H., Orlova, E.V., RinkeAppel, J., Junke, N., Mueller, F., Rodnina, M., Wintermeyer, W., Brimacombe, R. and vanHeel, M. (**1997**) *Cell, 88,* 19-28.

[12] Alexander, R.W., Muralikrishna, P. and Cooperman, B.S. (**1994**) *Biochemistry, 33,* 12109-12118.

[13] Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr. Tarry, M.J., Carter, A.P. and Ramakrishnan, V. (**2001**) *Science, 292,* 897-902.

[14] Harms, J., Tocilj, A., Levin, I., Agmon, I., Stark, H., Kölln, I., van Heel, M., Cuff, M., Schlünzen, F., Bashan, A., Franceschi, F. and Yonath, A. (**1999**) *Structure Fold. Des., 7,* 931-941.

[15] Ramakrishnan, V. and Moore, P.B. (**2001**) *Current Opinion in Structural Biology, 11,* 144–154.

[16] Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Bartels, H., Agmon, I. Franceschi, F. and Yonath, A. (**2001**) *Cell,* in the press.

[17] Navaza, J. (**1994**) *Acta Crystallogr. Sect. A, 50,* 157-163.

[18] Carter, A., Clemons, W., Brodersen, D., Morgan-Warren, R., Wimberly, B. and Ramakrishnan, V. (**2000**) *Nature, 407,* 340-348.

[19] Sussman, J.K., Simons, E.L., Simons, R.W. (**1996**) *Molecular Microbiology, 21,* 347-60.

[20] Tedin, K., Moll, I., Grill, S., Resch, A., Graschopf, A., *et al*. (**1999**) *Molecular Microbiology, 31,* 67-77.

[21] Grunberg-Manago, M., Dessen, P., Pantaloni, D., Godefroy-Colburn, A.D, Wolfe, T., *et al*. (**1975**) *J. Mol. Biol., 94,* 461-78.

[22] Garcia, C., Fortier, P., Blanquet, S., Lallemand, J.Y., Dardel, F. (**1995**) *Europ. J. of Bioch., 228,* 395-402.

[23] Garcia, C., Fortier, P., Blanquet, S., Lallemand, J-Y and Dardel F. (**1995**) *J. Mol. Biol., 254,* 247-59.

[24] Kycia, J.H., Biou, V., Shu, F., Gerchman, S.E., Graziano, V., *et al*. (**1995**) *Biochemistry, 34,* 6183-6187.

[25] Biou, V., Shu, F., Ramakrishnan, V. (**1995**) *EMBO J.,* **14**, 4056-64.

[26] de Cock, E., Springer, M. and Dardel, F. (**1999**) *Molecular Microbiology, 32,* 193-202.

[27] Moreau, M., De, C.E., Fortier, P.L., Garcia, C., Albaret, C., *et al*. (**1997**) *J. Mol. Biol., 266,* 15-22.

[28] Sette, M., Spurio, R., Van Tilborg, P., Gualerzi, C.O., Boelens, R. (**1999**) *RNA, 5,* 82-92.

[29] Firpo, M.A., Connelly, M.B, Goss, D.J, Dahlberg, A.E. (**1996**) *J. Biol. Chem., 271,* 4693-8.

[30] Mackeen, L.A., Kahan, L., Wahba, A.J., Schwartz, I. (**1980**) *J. of Biol. Chem., 255,* 526-31.

[31] Moazed, D., Samaha, R.R., Gualerzi, C., Noller, H.F. (**1995**) *J. Mol. Biol., 248,* 207-10.

[32] Sacerdot, C., de Cock, E., Engst, K., Graffe, M., Dardel, F., Springer, M. (**1999**) *J. Mol. Biol., 288,* 803-10.

[33] McCutcheon, J.P., Agrawal, R.K., Philips, S.M., Grassucci, R.A., Gerchman, S.E. *et al*. (**1999**) *Proc. Nat. Acad. Sci. USA, 96,* 4301-06.

[34] Srivastava, S., Verschoor, A. and Frank, J. (**1992**) *J. Mol. Biol., 226,* 301-304.

[35] Hershey, J.W., Asano, K., Naranda, T., Vornlocher, H.P., Hanachi, P., Merrick, W.C. (**1996**) *Biochimie, 78,* 903-7.

[36] Bruhns, J. and Gualerzi, C.O. (**1980**) *Biochemistry, 19,* 1670-6.

[37] Eisenstein, M., Shariv, I., Koren, G., Friesem, A.A. and Katchalski-Katzir, E. (**1997**) *J. Mol. Biol., 266,* 135-43.

[38] Odon, O.W., Kramer, G., Henderson, A.B., Pinphanichakarn, P. and Hardesty, B. (**1978**) *J. Biol. Chem., 253*(6), 1807-16.

[39] Altamura, S., Sanz, J.L, Amils, R., Cammarano, P., Londei, P. (**1988**) *Systematic & Applied Microbiology, 10,* 218-25.

[40] Kurylo-Borowska, Z. (**1975**) *Biochim. Biophys. Acta, 399,* 31-41.

[41] Moazed, D. and Noller H.F. (**1987**) *Nature, 327,* 389-94.

[42] Mankin, A.S. (**1997**) *J. Mol. Biol., 274,* 8-15.

[43] Woodcock, J., Moazed, D., Cannon, M., Davies, J. and Noller, H.F. (**1991**) *EMBO J., 10,* 3099-3104.

[44] VanLoock, M.S., Agrawal, R.K., Gabashvili, I., Frank, J., Harvey S.C. (**2000**) *J. Mol. Biol., 304,* 507-515.

[45] Brodersen, D.E., Clemons, W.M., Carter, A.P., Morgan-Warren, R.J., Wimberly, B.T. *et al*. (**2000**) *Cell, 103,* 1143-54.

[46] Spahn, C.M.T. and Prescott, C.D. (**1996**) *J. Mol. Med., 74,* 423-439.

[47] Jerinic, O. and Joseph, S. (**2000**) Conformational changes in the ribosome induced by translational miscoding agents. *J. Mol. Biol., 304,* 707-713.

[48] Orth, P., Saenger, W. and Hinrichs, W. (**1999**) Tetracycline-chelated Mg2+ ion initiates helix unwinding in Tet repressor induction. *Biochemistry, 38,* 191-8.

[49] Roberts, M.C. (**1996**) *FEMS Microbiol. Rev., 19,* 1-24.

[50] Ross, J.I., Eady, E.A., Cove, J.H. and Cunliffe, W.J. (**1998**) *Antimicrob. Agents Chemother., 42,* 1702-1705.

[51] Dantley, K.A., Dannelly, H.K. and Burdett, V. (**1998**) *J. Bacteriol., 180,* 4089-92.

[52] Held, W.A., Ballou, B., Mizushim, S. and Nomura, M. (**1974**) *J. Biol. Chem., 249,* 3103-3111.

[53] Nowotny, V. and Nierhaus, K.H. (**1988**) *Biochemistry, 27,* 7051-7055.

[54] Huang, H., Chopra, R., Verdine, G.L. and Harrison, S.C. (**1998**) *Science, 282,* 1669-1675.

[55] Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (**1997**) *Nature, 389,* 251-260.

[56] Simitsopoulou, M., Avila, H. and Franceschi, F. (**1999**) *Eur. J. Biochem., 266,* 524-532.

[57] Mizushima, S. and Nomura, M. (**1970**) *Nature, 226,* 1214-1218.

[58] Nikulin, A., Serganov, A., Ennifar, E., Tishchenko, S., Nevskaya, N., Shepard, W., Portier, C., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S. and Dumas, P. (**2000**) *Nat. Struct. Biol., 7*, 273-277.

[59] Agalarov, S.C., Sridhar Prasad, G., Funke, P.M., Stout, C.D. and Williamson, J.R. (**2000**) *Science, 288*, 107-113.

[60] Choli, T., Franceschi, F., Yonath, A. and Wittmann-Liebold, B. (**1993**) *Biological Chemistry, 346*, 377-383.

[61] Hosaka, H., Nakagawa, A., Tanaka, I., Harada, N., Sano, K., Kimura, M., Yao, M. and Wakatsuki, S. (**1997**) *Structure, 5*, 1199-1208.

[62] Carson, M. (**1997**) *Methods Enzymol.*, **277**, 493-505.

[63] Philippsen, A. (**2000**) http://www.bioz.unibas.ch/~xray/dino.