

A Map of Protein-rRNA Distribution in the 70 S *Escherichia coli* Ribosome*

Received for publication, November 10, 1999, and in revised form, January 10, 2000

Dmitri I. Svergun†§¶ and Knud H. Nierhaus||

From the ‡European Molecular Biology Laboratory, EMBL c/o DESY, Notkestraße 85, D-22603 Hamburg, Germany, the §Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333 Moscow, Russia, and the ||Max Planck Institut für Molekulare Genetik, AG Ribosomen, Ihnestr. 73, 14195 Berlin, Germany

Neutron scattering exploits the enormous scattering difference between protons and deuterons. A set of 42 x-ray and neutron solution scattering curves from hybrid *Escherichia coli* ribosomes was obtained, where the proteins and rRNA moieties in the subunits were either protonated or deuterated in all possible combinations. This extensive data set is analyzed using a novel method. The volume defined by the cryoelectron microscopic model of Frank and co-workers (Frank, J., Zhu, J., Penczek, P., Li, Y. H., Srivastava, S., Verschoor, A., Rademacher, M., Grassucci, R., Lata, R. K., and Agrawal, R. K. (1995) *Nature* 376, 441–444) is divided into 7890 densely packed spheres of radius 0.5 nm. Simulated annealing is employed to assign each sphere to solvent, protein, or rRNA moieties to simultaneously fit all scattering curves. Twelve independent reconstructions starting from random approximations yielded reproducible results. The resulting model at a resolution of 3 nm represents the volumes occupied by rRNA and protein moieties at 95% probability threshold and displays 15 and 20 protein subvolumes in the 30 S and 50 S, respectively, connected by rRNA. 17 proteins with known atomic structure can be tentatively positioned into the protein subvolumes within the ribosome in agreement with the results from other methods. The protein-rRNA map enlarges the basis for the models of the rRNA folding and can further help to localize proteins in high-resolution crystallographic density maps.

Ribosomes are supramolecular complexes responsible for protein synthesis in all organisms. Each of the two unequal ribosomal subunits is a complicated assembly of proteins and nucleic acids. Thus, the prokaryotic 70 S ribosome from *Escherichia coli* has a total molecular mass of about 2.3×10^6 Da and consists of a 30 S and 50 S subunit (1). The 30 S subunit contains 21 individual proteins and a single 16 S rRNA molecule, and the 50 S subunit, 33 proteins and two rRNA molecules (5 S rRNA + 23 S rRNA). In both subunits the rRNA moieties account for about two-thirds of the mass. A better description of the interactions between proteins and rRNA in the ribosome is of paramount importance for the understanding of the structural mechanism of protein synthesis.

The structure of the ribosome has been extensively studied

for decades using various methods. In the last years, cryoelectron microscopy (cryo-EM)¹ yielded models of the overall shape of the ribosome at a resolution of 1.5 nm (2–5). Recent tremendous progress in the x-ray crystallography provided electron density maps down to 0.78 nm for ribosomal complexes containing tRNAs (6) and to 0.45 nm for the individual ribosomal subunits (7–10). Despite these remarkable achievements, limited information has so far been obtained about the spatial distribution of rRNA and proteins in the ribosome. Using the x-ray maps, six ribosomal proteins were positioned in the small (7, 10) and four in the large subunit (8). The low contrast between ribosomal proteins and rRNA for both electrons and x-rays makes it difficult to distinguish between the two components even at relatively high resolution. Moreover, disordered or flexible domains do not show up in the electron density maps obtained by the x-ray crystal analysis (6, 8).

Neutron scattering in solution, although yielding only low-resolution information (11), is a powerful tool to study complexes of proteins and nucleic acids. Scattering from a particle or its component is proportional to the squared contrast (difference between the scattering density of the particle/component and that of the solvent), and measurements at different contrasts provide additional information about the object (12). As the neutron scattering lengths of hydrogen and deuterium differ drastically, isotopic H/D substitution is widely employed for the contrast variation. In earlier studies using H₂O/D₂O mixtures (13, 14), integral parameters of the protein and rRNA moieties were established. Yet more information is provided by selective deuteration of the ribosomal components. Triangulation of labeled protein pairs in the ribosomal subunits so far provided the most comprehensive information about the protein positions (15, 16).

In a recent study (17, 18), 42 solution scattering curves from the hybrid *E. coli* ribosomes were measured, where the proteins and rRNA moieties in the subunits were selectively deuterated. This is probably the most extensive set of consistent x-ray and neutron contrast-variation data collected on a single object. The curves permitted to validate cryo-EM models of the ribosome (2, 4) and they were further interpreted in terms of a solid body model constructed from the envelopes of the subunits and those of the rRNA moieties. Here, we exploit the data set further using a new analysis method (19), where an overall shape of the object provided by cryo-EM is filled by densely packed small spheres. The algorithm assigns each sphere to the rRNA or protein moieties (or to the solvent) to simultaneously fit the available scattering data. Its efficiency was illustrated (19) by an *ab initio* restoration of a model structure represent-

* This work was supported by European Commission Grant BIO4-CT97-2143 (to D. I. S.), INTAS Grant 95-1272 (to D. I. S.), and Deutsche Forschungsgemeinschaft Grant Ni 174/8-2 (to K. H. N.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: EMBL c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany. Tel.: 49-40-89902-125; Fax: 49-40-89902-149; E-mail: Svergun@EMBL-Hamburg.DE.

¹ The abbreviations used are: cryo-EM, cryo-electron microscopy; TP30, TP50, RNA30, and RNA50, the total proteins and the rRNA moieties in the 30 S and 50 S subunits, respectively; DAM, dummy atoms model; SA, simulated annealing; TP, total proteins.

ing the envelope of the 30 S ribosomal subunit with several embedded proteins. The application of the new method significantly improves the resolution of the neutron-scattering maps of the ribosomal structure. A three-dimensional map of the protein-rRNA distribution is established which, in particular, reveals the likely positions of individual ribosomal proteins or protein complexes in the 70 S *E. coli* ribosome.

EXPERIMENTAL PROCEDURES

Experimental—The sample preparation and scattering experiments are described in detail elsewhere (17, 18). The 70 S samples of the *E. coli* ribosome were made by *in vitro* association of selectively deuterated ribosomal subunits and checked for structural integrity and biological activity. The neutron experiments were performed at the Risø Laboratory (Roskilde, Denmark (20)) and at the GKSS Research Center (Geesthacht, Germany (21, 22)). Using two sample-detector distances ("low angle" and "high angle" setting), the range of momentum transfer $0.09 < s < 2.3 \text{ nm}^{-1}$ was covered ($s = (4\pi/\lambda)\sin\theta$, where λ is the wavelength, and 2θ is the scattering angle). The x-ray synchrotron radiation scattering data were collected on the EMBL beamline (23–25) at HASYLAB (Hamburg, Germany) in the range $0.08 < s < 1.6 \text{ nm}^{-1}$. The full list of samples is presented in Table I.

Dummy Atoms Model—The 70 S ribosome consists of four components representing the total proteins (TP) and the rRNA moieties in the 30 S and 50 S subunits (denoted below as TP30, TP50, RNA30, and RNA50, respectively). The search volume was defined by drawing a regular grid of points in space corresponding to dense hexagonal packing of spheres of radius $r_0 = 0.5 \text{ nm}$ inside a parallelepiped determined by the gabarites of the cryo-EM reconstruction of the 70 S *E. coli* ribosome (2). If the distance between a point and the closest pixel in the EM reconstruction did not exceed $1.5r_0$, a dummy atom is placed at this point. This procedure expands the original model and thus reduces the bias and allows for deviations from the EM shape. The structure of the four component dummy atoms model (DAM) is defined by assigning a number (component index) $0 \leq X_j \leq 4$ to each dummy atom. As the subunits are well resolved in the cryo-EM model, the 30 S and 50 S subvolumes are defined where the dummy atoms may have the component index 0,1,2 (solvent, TP30 or RNA30) and 0,3,4 (solvent, TP50 or RNA50), respectively. The atoms at the subunit interface can belong to any component. The DAM in Fig. 1 contains 2644 atoms in the 30 S subvolume, 5020 atoms in the 50 S subvolume, 196 atoms at the interface, and encloses the volume of 5560 nm^3 (that of the original cryo-EM model is 3920 nm^3).

In keeping with the low resolution of the solution scattering data, the model must be constrained to have low resolution with respect to r_0 . For this, a list of neighbors (*i.e.* atoms at an offset $2r_0$) is defined for each dummy atom. Looseness (degree of isolation) of a non-solvent atom is calculated as $P(N_e) = \exp(-0.5N_e) - \exp(-0.5N_c)$, where N_e is the number of neighbors having the same index and $N_c = 12$ is the coordination number for hexagonal packing. Looseness of the configuration X (*i.e.* its non-compactness) is characterized by the average value $P(X) = \langle P(N_e) \rangle$ over all non-solvent atoms. Another condition requires connectivity, *i.e.* a possibility to connect two arbitrarily selected atoms belonging to a component by successively connecting neighboring atoms belonging to this component. The measure of connectivity of the k -th component is computed as $G_k(X) = \ln(N_k/M_k) \geq 0$, where N_k and M_k are the numbers of dummy atoms in the entire component and in the longest connected fragment, respectively.

Using the multipole expansion, the scattering intensity from a four component ribosomal DAM in solution is (17, 18, 26),

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_{k=1}^4 \{ [\Delta\rho_k A_{lm}^{(k)}(s)]^2 + 2 \sum_{n>k} \Delta\rho_k A_{lm}^{(k)}(s) \Delta\rho_n [A_{lm}^{(n)}(s)]^* \} \quad (\text{Eq. 1})$$

where $\Delta\rho_k$ and $A_{lm}^{(k)}(s)$ denote the contrast and the partial amplitudes of the k -th component, respectively. The contrasts of the "dry" protonated and deuterated ribosomal components in different solvents (Table I) were computed from their chemical composition as described elsewhere (17, 18). The partial amplitudes are expressed as (19),

$$A_{lm}^{(k)}(s) = i^l \sqrt{2/\pi} v_a \sum_j j_l(sr_j) Y_{lm}^*(\omega_j) \quad (\text{Eq. 2})$$

where the sum runs over the atoms of the k -th component, $(r_j, \omega_j) = r_j$ are their polar coordinates, $v_a = (4\pi r_0^3/3)/0.74$ is the displaced volume per dummy atom, $j_l(x)$ and $Y_{lm}(\omega)$ denote the spherical Bessel function and

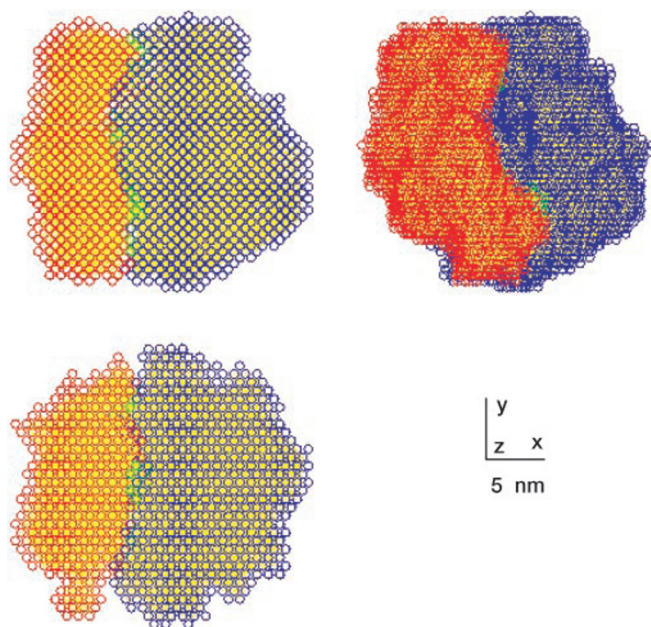


FIG. 1. Search-volume definition of the ribosomal DAM. Yellow pixels, cryo-EM model of (2) (pixel size corresponds to the original cryo-EM reconstruction); red and blue circles, dummy atoms belonging to the 30 S and 50 S subunits, respectively; green circles, dummy atoms that may belong to either subunit. The coordinates given in the inset correspond to the top left orientation. Top right and bottom orientations are rotated counterclockwise by 45° around y and by 90° around x , respectively. All three-dimensional images here and below are prepared on a SUN Workstation using the program ASSA (60).

the spherical harmonics, respectively. Equations 1 and 2 permit one to rapidly compute the scattering curves from the DAM for an arbitrary configuration X under the given contrasts of the components.

Minimization Procedure—To fit the data with a low resolution model one should find a configuration X minimizing $f(X) = \chi^2 + \alpha P(X)$, where the overall discrepancy is,

$$\chi^2 = \frac{1}{M} \sum_{i=1}^M \frac{1}{N(i)-1} \left\{ \sum_{j=1}^{N(i)} [I_{\text{exp}}^{(i)}(s_j) - I_{\text{calc}}^{(i)}(s_j)/\sigma(s_j)]^2 \right\} \quad (\text{Eq. 3})$$

$M = 42$ is the number of experimental curves, $N(i)$ is the number of points in the i -th curve, $I_{\text{exp}}(s)$, $I_{\text{calc}}(s)$, and $\sigma(s)$ are the experimental and calculated intensity and the experimental errors, respectively, and $\alpha > 0$ is the weight of the looseness penalty. The experimental data were scaled to the total dry volume of the ribosome 2350 nm^3 expected from its molecular weight and fitted in the range up to $s_{\text{max}} = 0.2 \text{ nm}^{-1}$. Series (1) over spherical harmonics was truncated at $l = 14$ and the calculated intensities were appropriately smeared to account for instrumental effects (17, 18).

The minimization was performed using simulated annealing (SA (27–29)). Starting from a random configuration X_0 , the assignment of a single atom is changed randomly (a move from X to X'). If $\Delta = f(X') - f(X) < 0$, the move is accepted, if not, it can be accepted with a probability $\exp(-\Delta/T)$, where T is annealing temperature. The latter is held constant for 70N moves or 7N accepted moves, whichever occurs first, then it is decreased ($T' = 0.9T$). The procedure starts at a sufficiently high temperature $T \approx 10^1$ and runs down to $T \approx 10^{-3}$ until no further reduction in $f(X)$ is observed. A complete run takes about 3 weeks CPU time on a 180 MHz SGI machine. After preliminary calculations, the penalty weight $\alpha \approx 20$ was selected so that the penalty term amounted to about 5–10% of $f(X)$ at the end of minimization (independent runs yielded discrepancy about $\chi \approx 4.1$ and looseness $P(X) \approx 0.08$).

RESULTS AND DISCUSSION

Establishing a Model of Protein-rRNA Distribution—Each of the dummy atoms filling the search volume in Fig. 1 was assigned to a specific component (solvent, TP30, TP50, RNA30, or RNA50) by simultaneously fitting the available scattering curves in Table I. Already preliminary computations without using connectivity constrain yielded well separated volumes for

TABLE I
Samples measured in this study and the contrasts of ribosomal components

Sample abbreviations: H, protonated; D, deuterated; first letter, TP, second letter, rRNA (e.g. DH30 + HH50 denotes hybrid ribosome with the TP30 deuterated, the rest protonated). For neutron scattering experiments, D₂O concentration in the solvent Y (conventional scattering) or sample polarization P (spin dependent scattering) is indicated. The contrasts are given in units 10¹⁰ cm⁻². Contrasts for the spin dependent scattering and for the x-ray scattering are scaled to those of the conventional contrast variation data as described (17, 18).

	Sample	TP30	RNA30	TP50	RNA50
Conventional neutron scattering, hybrid 70 S particles					
1	HH30 + HH50, Y = 0% D ₂ O	2.309	4.245	2.309	4.245
2	HH30 + HH50, Y = 35% D ₂ O	.378	2.207	.378	2.207
3	HH30 + HH50, Y = 50% D ₂ O	-.446	1.328	-.446	1.328
4	HH30 + HH50, Y = 75% D ₂ O	-1.819	-.129	-1.819	-.129
5	HH30 + HH50, Y = 100% D ₂ O	-3.199	-1.586	-3.199	-1.586
6	HH30 + DD50, Y = 0% D ₂ O	2.309	4.245	7.185	7.143
7	HH30 + DD50, Y = 35% D ₂ O	.378	2.207	5.250	5.086
8	HH30 + DD50, Y = 50% D ₂ O	-.446	1.328	4.425	4.226
9	HH30 + DD50, Y = 75% D ₂ O	-1.819	-.129	3.057	2.768
10	DD30 + HH50, Y = 0% D ₂ O	7.121	7.143	2.348	4.245
11	DD30 + HH50, Y = 35% D ₂ O	5.199	5.086	.422	2.207
12	DD30 + HH50, Y = 50% D ₂ O	4.373	4.226	-.402	1.328
13	DD30 + HH50, Y = 75% D ₂ O	3.006	2.768	-1.780	-.129
14	DD30 + HH50, Y = 100% D ₂ O	1.625	1.309	-3.148	-1.586
15	DD30 + DD50, Y = 0% D ₂ O	7.121	7.143	7.185	7.143
16	DD30 + DD50, Y = 35% D ₂ O	5.199	5.086	5.250	5.086
17	DD30 + DD50, Y = 50% D ₂ O	4.373	4.226	4.425	4.226
18	DD30 + DD50, Y = 75% D ₂ O	3.006	2.768	3.057	2.768
19	DD30 + DD50, Y = 100% D ₂ O	1.625	1.309	1.677	1.309
20	DH30 + HH50, Y = 0% D ₂ O	7.121	4.245	2.348	4.245
21	DH30 + HH50, Y = 40% D ₂ O	4.928	1.907	.150	1.907
22	DH30 + HH50, Y = 60% D ₂ O	3.818	.744	-.949	.744
23	DH30 + HH50, Y = 100% D ₂ O	1.625	-1.586	-3.148	-1.586
24	HH30 + DH50, Y = 0% D ₂ O	2.309	4.245	7.185	4.245
25	HH30 + DH50, Y = 40% D ₂ O	.111	1.907	4.992	1.907
26	HH30 + DH50, Y = 60% D ₂ O	-.988	.744	3.883	.744
27	HH30 + DH50, Y = 100% D ₂ O	-3.199	-1.586	1.677	-1.586
Spin-dependent neutron scattering, hybrid 70 S particles					
28	HH30 + DD50, P = 0%	1.290	.748	-.116	-.037
29	HH30 + DD50, P = 100%	1.290	.729	.013	.000
30	DD30 + HH50, P = 0%	-.116	-.037	1.290	.767
31	DD30 + HH50, P = 100%	.013	.000	1.290	.729
32	HD30 + HD50, P = 0%	-.052	1.870	-.052	1.870
33	HD30 + HD50, P = 100%	.271	1.870	.271	1.870
Conventional neutron scattering, free subunits					
34	HH30, Y = 0% D ₂ O	2.309	4.245	.000	.000
35	HH30, Y = 100% D ₂ O	-3.199	-1.586	.000	.000
36	DD30, Y = 0% D ₂ O	7.121	7.143	.000	.000
37	HH50, Y = 0% D ₂ O	.000	.000	2.309	4.245
38	HH50, Y = 100% D ₂ O	.000	.000	-3.199	-1.586
39	DD50, Y = 0% D ₂ O	.000	.000	7.185	7.143
X-ray curves from all 70 S samples and from free subunits					
40	70 S	2.26	4.23	2.26	4.23
41	30 S	2.26	4.23	.000	.000
42	50 S	.000	.000	2.26	4.23

the protein and rRNA moieties. For all independent restorations starting from random initial approximations, proteins in both subunits converged to several isolated subvolumes, whereas the distribution of the dummy atoms belonging to the rRNAs was more uniform and featureless (see example in Fig. 2, top). To further constrain the model, the connectivity requirement was imposed on the rRNA moieties in both subunits and also on the entire subunit (see details under “Experimental Procedures”). Corresponding connectivity factors $G_k(X)$ were added to the looseness penalty $P(X)$ during minimization and 12 independent restorations yielded very consistent results. The positions of the individual protein subvolumes differ slightly in different restorations (typical deviations are illustrated in Fig. 2, bottom). Given that the restorations provide the same overall discrepancy $\chi \cong 4.1$ (typical fit to the experi-

mental data is presented in Fig. 3), this uncertainty has to be attributed to the low resolution of the solution scattering data. A similar ambiguity has been observed in model calculations on a ribosome-like structure (19).

The reliability of the results is assessed by the analysis of the 12 independently obtained configurations. For each dummy atom, the number of repetitions (*i.e.* the number of times a given position was attributed to the same specific component) was counted. The frequency histogram (numbers of atoms having the given number of repetitions) is presented in Fig. 4 along with the histogram computed from 12 randomly generated DAMs. The two distributions are principally different and that obtained from the SA has much higher proportion of atoms with high repetition rates. In particular, the random generations indicate that the probability for a dummy atom to be

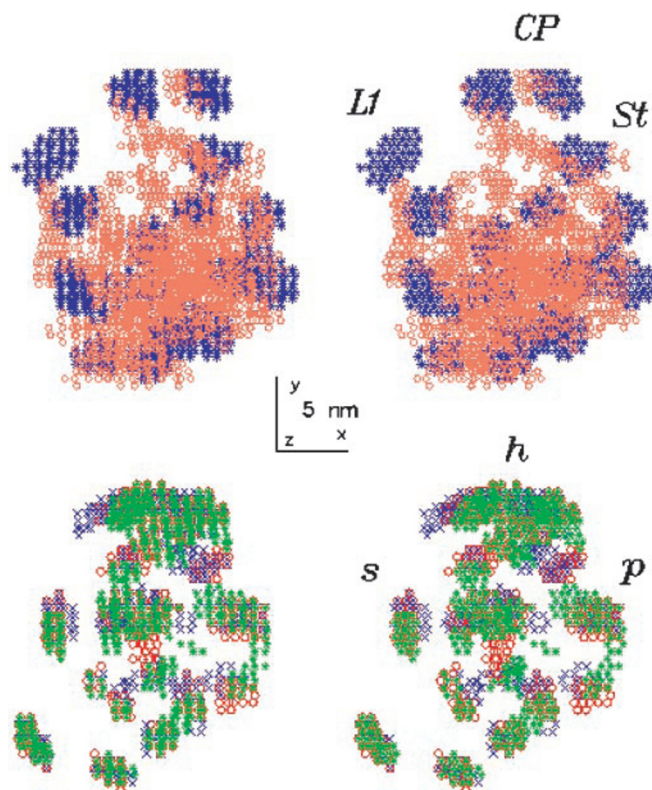


FIG. 2. Examples of the restoration of the ribosomal structure by SA. *Top*, a map of the 50 S subunit computed without connectivity requirements (blue asterisks, TP50; red circles, RNA50). Landmarks of the 50 S subunit: L1, L1 stalk; CP, central protuberance; St, L7/L12 stalk. *Bottom*, superposition of the TP30 for three independent runs (red circles, blue crosses, and green asterisks). Landmarks of the 30 S subunit: s, shoulder; h, head; p, platform. The stereo pairs are rotated by 45° around the y axis (30 S clockwise, 50 S counterclockwise) as compared with the upper left orientation in Fig. 1.

occasionally assigned to one and the same component 12 times out of 12 is about 2×10^{-6} . Along the fact that the frequency histogram contains 1755 such atoms underlines the statistical significance and reproducibility of the results.

The uncertainty illustrated in Fig. 2 (*bottom*), indicates that the assignment of the dummy atoms close to the protein-rRNA interface or to the subunit border may differ in different reconstructions. Given also that closely positioned proteins may be joined in single globules at low resolution, the method does not yield exact positions and shapes of all individual proteins. Nevertheless, the analysis of the independent reconstructions permits one to distinguish between the volumes occupied by proteins and rRNA in the ribosomal subunits with a reasonable probability. For each of the four ribosomal components (TP30, RNA30, TP50, and RNA50) we build a volume from the atoms that were ascribed to a specific component at least once in the 12 reconstructions. It is conceivable that these 4 volumes enclose the true volumes of the corresponding ribosomal components. In each of these volumes we now discard the atoms that appeared only once and keep those atoms ascribed to the given component twice and more. Simulations with randomly generated structures show that the remaining volumes would enclose the true volumes with the probability of 95%. The rRNA and protein moieties thus obtained are well separated in both subunits: the overlap (relative number of dummy atoms found in the two moieties simultaneously) is less than 17 and 20% in the 30 S and 50 S subunit, respectively. The overlapping positions are resolved by selecting the component with higher number of repetitions. The final map of the protein-rRNA distribu-

tion in the ribosome at a resolution of about $2\pi/s_{\max} \approx 3$ nm is presented in Fig. 5.

General Features of the Model—Given that no distinction is made between proteins and rRNA at the first stage (see Fig. 2, *top*), the mere fact that the proteins reproducibly converge to separated volumes connected by the rRNA moiety underlines the soundness of the restoration. The integral parameters of the ribosomal components in Table II are in a good agreement with those reported in (17, 18). It should be noted that the solid body model used in Refs. 17 and 18 could not fully account for intermingling of the ribosomal components and an artificial term from internal density fluctuations has to be added to the calculated scattering curves. The representation using the DAM is free from this limitation. The solutions are rather stable to the changes in the weight of the looseness penalty and also to the changes in the packing radius (restorations with $r_0 = 0.6$ and 0.7 nm yielded very similar results).

The cryo-EM model (2) has a nominal resolution of about 2.5 nm. As the construction of the search volume involved expansion of this model and (implicitly) smoothing of its surface, the use of the higher resolution reconstruction published later (3) would not result in significant difference. An attempt to use the search volume built from the cryo-EM model (4) yielded the spatial distribution of proteins similar to that presented in Fig. 5 but the agreement to the experimental data was worse ($\chi \approx 4.5$ for all independent runs). This corroborates previous conclusions (17, 18) that the model (4) is less consistent with the neutron scattering data.

In each of the independent SA reconstructions (*e.g.* those displayed in Fig. 2), the enclosed volumes of the four ribosomal components are equal to those computed from the chemical composition. As the final model in Fig. 5 is obtained by merging a dozen independent restorations, the enclosed volumes of the components should be higher, and the increase of the volume provides an estimate of the uncertainty in the spatial distribution of the component. As seen from Table II, the protein moieties in both subunits are better defined by the scattering data than the rRNA moieties. This should be expected as both TP30 and TP50 are organized in separate globules and are located closer to the periphery of the subunits. The map in Fig. 5 displays 15 and 20 protein islands in the 30 S and 50 S subunits, respectively, *i.e.* about 60 to 70% of the total number of ribosomal proteins.

Small Subunit—For the 30 S subunit, the comparison with the triangulation map of the protein centers (15) in Fig. 6 (*top row*) is of special interest. The radius of gyration of the TP30 evaluated from the triangulation map (6.36 nm) has been the object of a long standing controversy with significantly higher values obtained in numerous contrast variation studies (14, 17, 18, 30). Positioning of the triangulation map within the models of the 30 S (31–33) predicted no proteins in the bottom of the subunit. The TP30 in the SA model is indeed concentrated more at the top of the subunit, whereas the 16 S rRNA predominantly fills its bottom (Fig. 5; this is also in agreement with the recent crystallographic map (10)). The overlap in Fig. 6 displays a reasonable agreement between the triangulation map and the SA distribution: most of the protein globules in the latter could be assigned to a number (or numbers) from the triangulation map. However, the SA map displays at least two protein globules at the bottom of the subunit. One of them could be S20 that has been detected (34) at the bottom of the 30 S subunit (see also next paragraph). It follows that, in agreement with the results reported in Refs. 17 and 18, a complete absence of proteins from the bottom of the 30 S subunit is not compatible with the neutron scattering data. The separation between the centers of mass of TP30 and RNA30 was reported in the

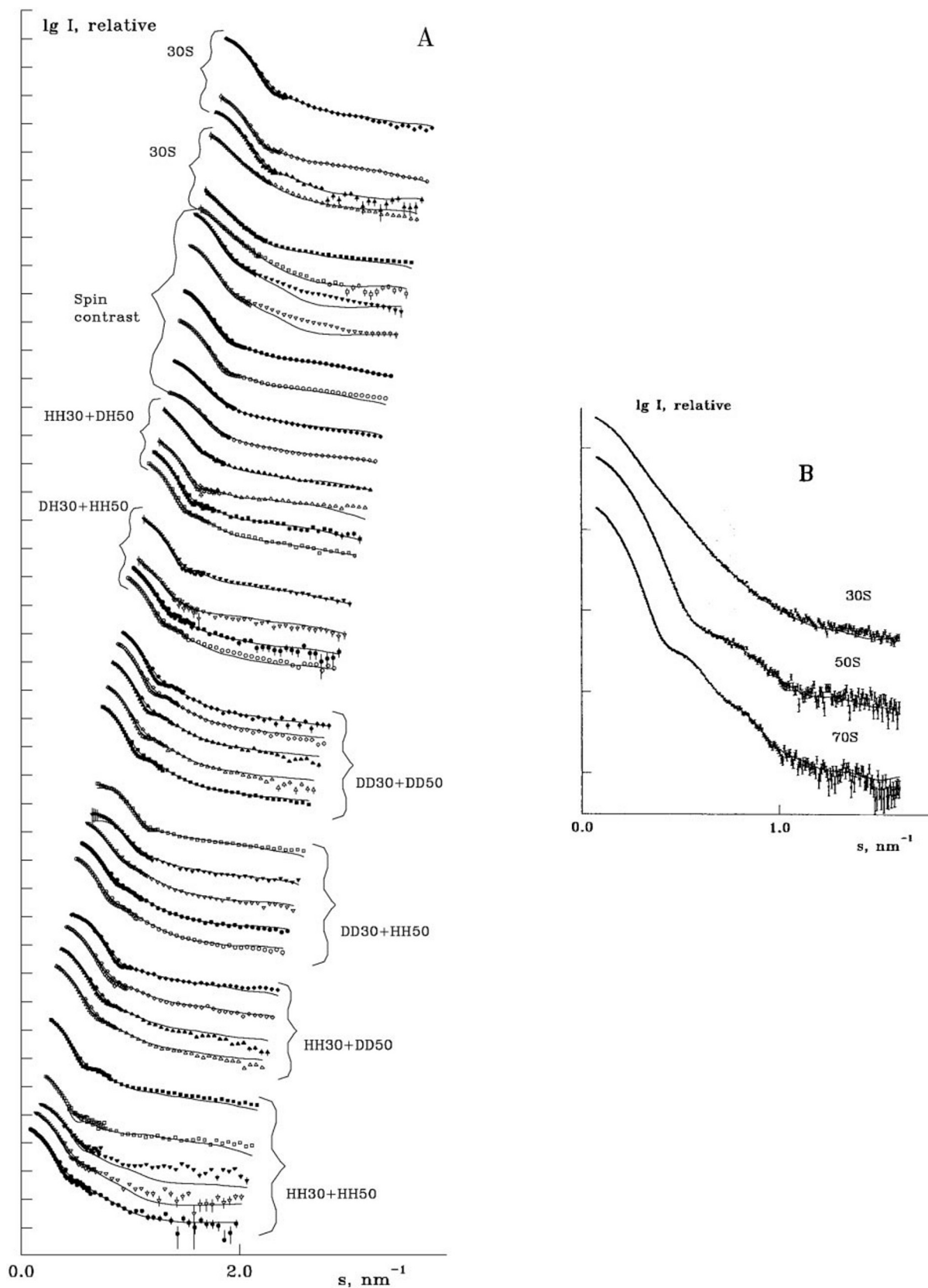


FIG. 3. Typical fit of the neutron (A) and x-ray (B) scattering data by the SA models. Successive curves are displaced up by one logarithmic unit corresponding to the distance between the ordinate tick marks (in panel A, also by $\Delta s = 0.05 \text{ nm}^{-1}$ along the abscissa) for better visualization. For the neutron scattering data in panel A, fits to the low angle and high angle settings are displayed separately, and the sequence of curves, from bottom to top, corresponds to that in Table I.

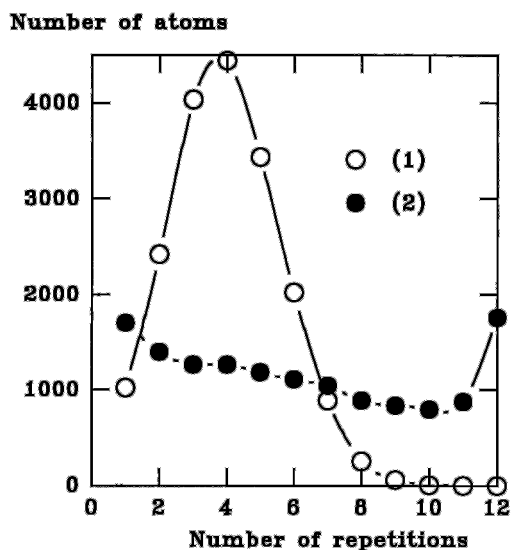


FIG. 4. Frequency histograms computed from 12 randomly generated DAMs (1) and from 12 independent SA solutions (2).

earlier contrast variation studies not to exceed 2 nm (in the present model a value of 5.5 nm is found, see Table II). It should be noted that most of the earlier studies were performed on protonated samples. The natural contrast between rRNA and proteins was demonstrated (35) to be insufficient to reliably estimate the separation. Experiments on specifically labeled samples are required as applied here.

The atomic models of several ribosomal proteins deposited in the Protein Data Bank (36) were interactively positioned in the SA map as shown in Fig. 6. From comparison with the triangulation map (15) in Fig. 6, *top row*, likely positions of S4 (37), S8 (PDB code 1SEI (38)), S15 (1AB3 (39)), and S17 (38) are easily established (Fig. 6, *middle row*). The globules accommodating S5 (1PKP (38)), S6 (1RIS (40)), and S7 (1RSS (41)) are selected in agreement with the relative positions and orientations of the proteins in the crystallographic map (10) and with the data (42). The form and orientation of the protein globule just below S17 is very similar to the tentative model of S20 as displayed in Ref. 10. Moreover, as in the latter map, the inter-subunit interface of the 30 S is relatively protein-free, and most of the proteins are concentrated in the head of the subunit.

Large Subunit—The prominent ribosomal features, *e.g.* the arms of the 50 S subunit containing the L1 and L7/L12 proteins, respectively, are adequately represented in the SA map.

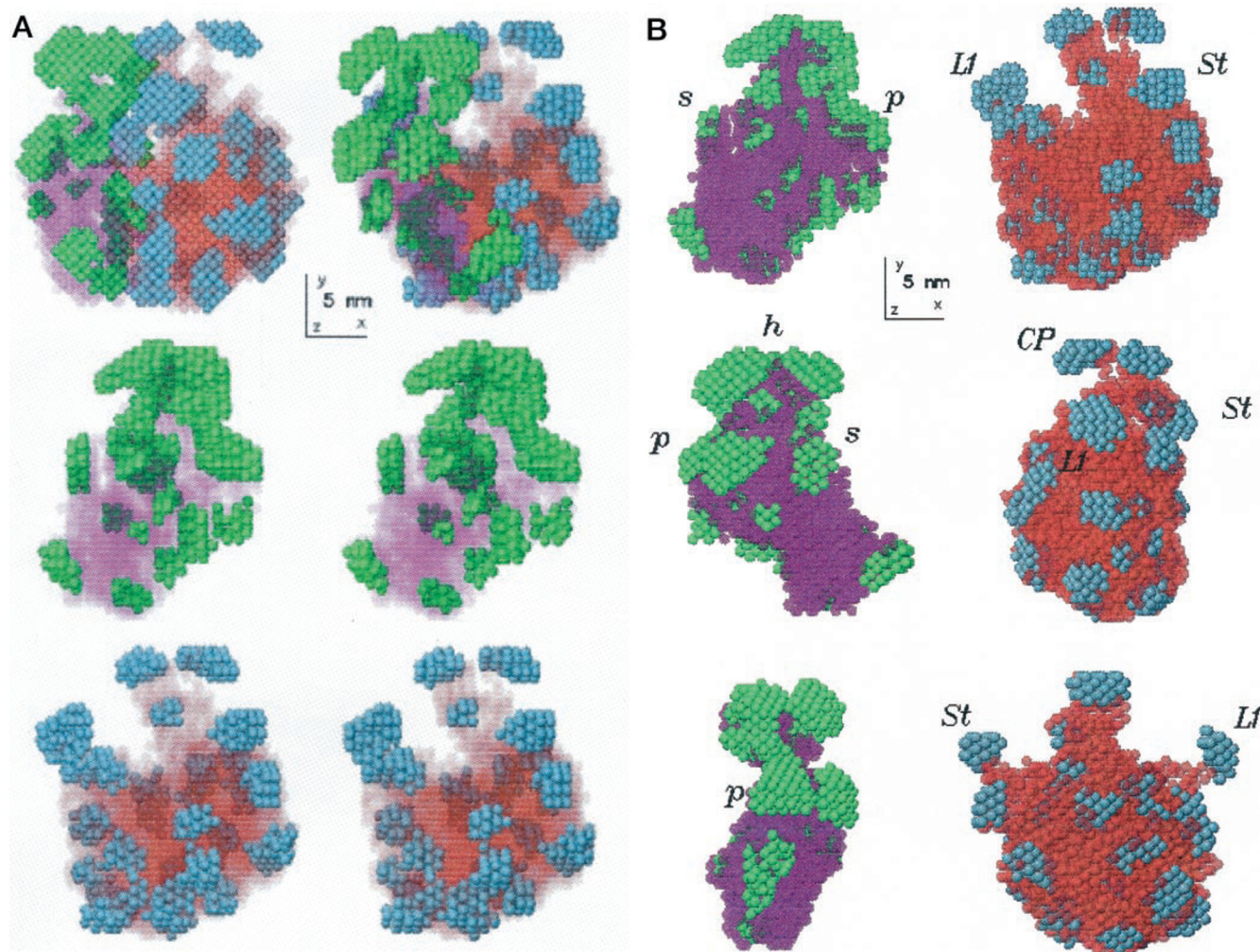


FIG. 5. The final SA model of the 70 S ribosome. Green, TP30; magenta, RNA30; blue, TP50; red, RNA50. The spheres representing proteins are displayed as solids, those representing rRNA as semi-transparent bodies. The sphere radius is $r_0 = 0.5$ nm. A, *top row* corresponds to the orientations of the 70 S ribosomes in the *top row* in Fig. 1, *middle* and *bottom row* display separately stereo pairs of the 30 S and 50 S subunits, respectively, in the orientations as in Fig. 2. Protein islands are green in the 30 S subunits and blue in the 50 S subunits. B, views of the subunits represented with less transparent rRNA. *Top row* has the orientations as that in the stereo pairs in A, *bottom* and *middle rows* are rotated counterclockwise around y axis by 120° and 240° , respectively. Landmarks are as in Fig. 2.

The appearance of the central protuberance of the 50 S subunit containing three protein globules and an rRNA fragment loosely connected with the rest of the rRNA moiety is worth noting. The shape of this rRNA fragment resembles that of the models of the free 5 S rRNA (43, 44). It is also well known that the 5 S rRNA forms stable complexes with the three ribosomal proteins L18, L25, and L5 (45). This suggests that the fragment in the head of the 50 S subunit represents the 5 S rRNA (possibly in a more compact form than the free molecule), whereas the main rRNA moiety in the body of the subunit is the 23 S rRNA molecule. A possible tunnel in the body of the 50 S subunit (reported *e.g.* by Refs. 3, 8, 9, and 46) would just be a higher resolution feature in the central part of the 23 S rRNA moiety. This feature has little impact on the scattering pattern

without specific labeling, and it would thus be unrealistic to search for it in Fig. 5.

A comprehensive immunoelectron microscopy model of the protein distribution in the large subunit (47) was used to place the available atomic models (Fig. 6, *bottom row*). The most obvious position is that of L1 (1AD2 (48)). The semi-concave shape of the globule formed by dummy atoms yields an orientation of the protein similar to that reported (3). A tentative tetramer of the C-terminal of the L7/L12 protein (1CTF (49)) is well accommodated within the upper globule in the L7/L12 stalk. The lower globule should in this case contain the N-terminal domains of L7/L12, and the distance between the two globules corresponds approximately to the length of the flexible hinge region between the N- and C-domains of L7/L12 (50, 51). The position of the RNA-binding domain of L2 (1RL2 (52)) near the peptidyl transferase center, as well as those of L6 (1RL6 (53)), L11 (1ACI (54)), and L14 (1WHI (55)) proximal to the GTPase associated center were selected following the x-ray map (8). The highly anisometric proteins, L9 (1DIV (56)), L11 (1ACI (54)), and L30 (1BXY (57)), can be appropriately positioned to agree with the map (47).

The three proteins bound to the 5 S rRNA in the head of the subunit are of special interest. The atomic model of L25 (58) is neatly fitted into the lower globule in the head in agreement with Ref. 47. The two upper globules should correspond to L5 and L18; the larger size of the globule on the cytoplasmic side suggests that the latter is L5, whereas the L18 contacts the 30

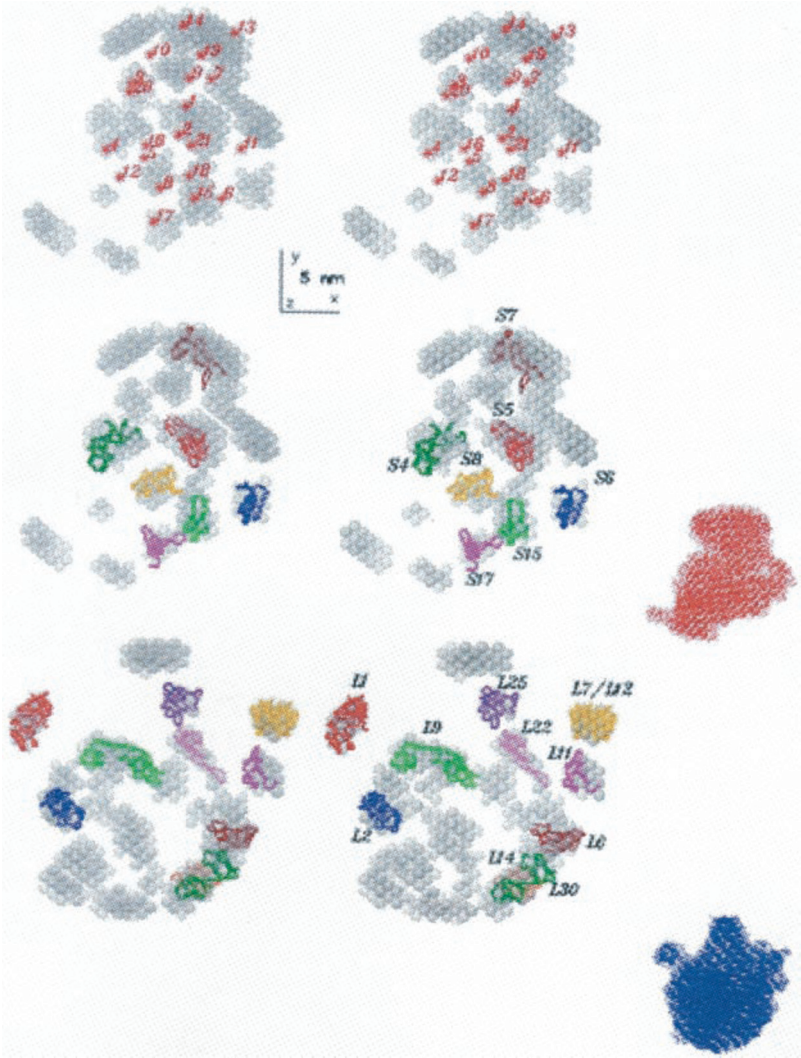
TABLE II

Integral parameters of the model of the 70 S ribosome

V_{SA}/V_{dry} is the ratio of the enclosed volume of the component in the SA model to its dry volume computed from the chemical composition; Δ is the distance between the centers of mass of TP and rRNA in the given subunit.

Component	Enclosed volume, V_{SA}, nm^3	Volume ratio, V_{SA}/V_{dry}	R_g, nm	Δ
TP30	454	1.29	7.61	5.5
RNA30	855	1.68	6.31	
TP50	525	1.16	9.48	
RNA50	1900	1.90	6.80	2.0

FIG. 6. Stereo pair of the positions of proteins in the ribosomal subunits. Top, protein map in the 30 S subunit (semi-transparent spheres) superimposed with the triangulation map of (15) (numbered asterisks). Middle and bottom, tentative positions of the atomic models of proteins presented as colored C_α traces. Dummy atoms representing protein components are shown as semi-transparent spheres. Middle, 30 S subunit; bottom, 50 S subunit. For both subunits, interface orientation is displayed. For clarity, the shapes of the subunits in this orientation taken from the cryo-EM reconstruction of 2, are shown (red, 30 S; blue, 50 S).



S subunit. Among the proteins positioned in the large subunit, L2, L9, L25, and L30 are on the subunit interface, while L6, L14, and L22 are closer to the cytoplasmic side.

Conclusions—The new approach presented here significantly expands the potential of structural analysis of biological macromolecules by neutron scattering techniques. The method is not limited to ribosomes but can be used to elucidate low-resolution *in situ* structures of various complexes in solution, especially when the components can be selectively labeled. Its sensitivity is enhanced if the components are organized in separated globules like ribosomal proteins.

What is the significance of the results presented in view of the recent high resolution x-ray maps (6–10)? First, the SA map yields an independently obtained model of the entire biologically active 70 S ribosome in solution at nearly physiological conditions. Second, one can expect to find extensive flexible or disordered domains in the ribosome (in particular, on its periphery) that do not contribute to the crystal diffraction, whereas neutron scattering is especially sensitive to the peripheral structural elements and provides their average structure. This is clearly seen in the rather flexible L7/L12 stalk of the 50 S subunit. This prominent ribosomal feature yields no density in the x-ray maps of both isolated 50 S subunit (8) and programmed 70 S ribosomes containing tRNAs (6), but is unequivocally identified as a proteins globule in all 12 SA reconstructions. Furthermore, neutron scattering can be applied to ribosomes from virtually all species, in particular, to *E. coli*, where most of the genetic, biochemical, and structural data have been accumulated. In contrast, the crystallographic studies are obviously limited to those organisms that provide good crystals (the crystals of the *E. coli* ribosomes are not sufficiently good). Finally, the established protein islands inside the ribosome can further help to localize proteins in the crystallographic maps (examples are the two protein globules at the bottom of the 30 S subunit and the three globules in the central protuberance of the 50 S subunit).

Acknowledgments—We are grateful to the authors of Refs. 2 and 4 for kindly providing the coordinates of their cryo-EM reconstructions of the 70 S ribosome in vitreous ice and to the authors of Refs. 59 and 58 for providing the atomic coordinates of S4 and L25, respectively. We thank M. H. J. Koch, EMBL, Hamburg, and S. Connell, MPI, Berlin, for help and advice during the preparation of the manuscript.

REFERENCES

- Wittmann, H. G. (1982) *Annu. Rev. Biochem.* **51**, 155–183
- Frank, J., Zhu, J., Penczek, P., Li, Y. H., Srivastava, S., Verschoor, A., Rademacher, M., Grassucci, R., Lata, R. K., and Agrawal, R. K. (1995) *Nature* **376**, 441–444
- Malhotra, A., Penczek, P., Agrawal, R. K., Gabashvili, I. S., Grassucci, R. A., Jünemann, R., Burkhardt, N., Nierhaus, K. H., and Frank, J. (1998) *J. Mol. Biol.* **280**, 103–116
- Stark, H., Mueller, F., Orlova, E. V., Schatz, M., Dube, P., Erdemir, T., Zemlin, F., Brimacombe, R., and Heel, M. V. (1995) *Structure* **3**, 815–821
- Stark, H., Rodnina, M. V., Rinke-Appel, J., Brimacombe, R., Wintermeyer, W., and Heel, M. V. (1997) *Nature* **389**, 403–406
- Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. E., and Noller, H. F. (1999) *Science* **285**, 2095–2104
- Tocij, A., Schlünzen, F., Janell, D., Glühmann, M., Hansen, H. A. S., Harms, J., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 14252–14257
- Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B., and Steitz, T. A. (1999) *Nature* **400**, 841–847
- Ban, N., Freeborn, B., Nissen, P., Penczek, P., Grassucci, R. A., Sweet, R. A., Frank, J., Moore, P. B., and Steitz, T. A. (1998) *Cell* **93**, 1105–1115
- Clemons, W. M. J., Jr., May, J. L., Wimberly, B. T., McCutcheon, J. P., Capel, M. S., and Ramakrishnan, V. (1999) *Nature* **400**, 833–840
- Feigin, L. A., and Svergun, D. I. (1987) *Structure Analysis by Small-angle X-ray and Neutron Scattering*, Plenum Press, New York
- Stuhrmann, H. B., and Kirste, R. G. (1965) *Z. Phys. Chem. Neue Folge.* **46**, 247–250
- Stuhrmann, H. B., Koch, M. H. J., Parfait, R., Haas, J., Ibel, K., and Crichton, R. R. (1978) *J. Mol. Biol.* **119**, 203–212
- Koch, M. H. J., and Stuhrmann, H. B. (1979) *Methods Enzymol.* **59**, 670–706
- Capel, M. S., Engelmann, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., Sillers, I.-Y., Yabuki, S., and Moore, P. B. (1987) *Science* **238**, 1403–1406
- May, R. P., Nowotny, V., Nowotny, P., Voss, H., and Nierhaus, K. H. (1992) *EMBO J.* **11**, 373–378
- Svergun, D. I., Burkhardt, N., Pedersen, J. S., Koch, M. H. J., Volkov, V. V., Kozin, M. B., Meerwinck, W., Stuhrmann, H. B., Diedrich, G., and Nierhaus, K. H. (1997) *J. Mol. Biol.* **271**, 588–601
- Svergun, D. I., Burkhardt, N., Pedersen, J. S., Koch, M. H. J., Volkov, V. V., Kozin, M. B., Meerwinck, W., Stuhrmann, H. B., Diedrich, G., and Nierhaus, K. H. (1997) *J. Mol. Biol.* **271**, 602–618
- Svergun, D. I. (1999) *Biophys. J.* **76**, 2879–2886
- Lebech, B. (1990) *Neutron News* **1**, 7–13
- Stuhrmann, H. B. (1991) *Biochimie (Paris)* **73**, 899–910
- Stuhrmann, H. B., Burkhardt, N., Diedrich, G., Jünemann, R., Meerwinck, W., Schmitt, M., Wadzack, J., Willumeit, R., Zhao, J., and Nierhaus, K. H. (1995) *Nucl. Instr. Methods Phys. Res.* **356**, 124–132
- Koch, M. H. J., and Bordsas, J. (1983) *Nucl. Instrum. Methods* **208**, 461–469
- Boulin, C., Kempf, R., Koch, M. H. J., and McLaughlin, S. M. (1986) *Nucl. Instrum. Methods Sect. A* **249**, 399–407
- Boulin, C. J., Kempf, R., Gabriel, A., and Koch, M. H. J. (1988) *Nucl. Instrum. Methods Sect. A* **269**, 312–320
- Svergun, D. I. (1994) *Acta Crystallogr. Sec. A* **50**, 391–402
- Kirkpatrick, S., Gelatt, C. D., Jr., and Vecchi, M. P. (1983) *Science* **220**, 671–680
- Press, W. H., Teukolsky, S. A., Wetherling, W. T., and Flannery, B. P. (1992) *Numerical Recipes*, University Press, Cambridge, MA
- Ingber, L. (1993) *Math. Comp. Model.* **18**, 29–57
- Serdyuk, I. N. (1979) *Methods Enzymol.* **59**, 750–775
- Stern, S., Weiser, B., and Noller, H. F. (1988) *J. Mol. Biol.* **204**, 447–481
- Malhotra, A., and Harvey, S. C. (1994) *J. Mol. Biol.* **240**, 308–340
- Brimacombe, R. (1995) *Eur. J. Biochem.* **230**, 365–383
- Schwedler, G., Albrecht-Ehrlich, R., and Rak, K. H. (1993) *Eur. J. Biochem.* **217**, 361–369
- Ramakrishnan, V. (1986) *Science* **231**, 1562–1564
- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol.* **112**, 535–542
- Davies, C., Gerstner, R. B., Draper, D. E., White, S. W., and Ramakrishnan, V. (1998) *EMBO J.* **17**, 4545–4558
- Davies, C., Ramakrishnan, V., and White, S. W. (1996) *Structure* **4**, 1093–1104
- Berglund, H., Rak, A., Serganov, A., Garber, M., and Hard, T. (1997) *Nature Struct. Biol.* **4**, 20–23
- Lindahl, M., Svensson, L. A., Lijas, A., Sedelnikova, S. E., Eliseikina, I. A., Fomenkova, N. P., Nevskaya, N., Nikonov, S. V., Garber, M. B., Muranova, T. A., Rykono, A. I., and Amons, R. (1994) *EMBO J.* **13**, 1249–1254
- Wimberly, B. T., White, S. W., and Ramakrishnan, V. (1997) *Structure* **5**, 1187–1198
- Tanaka, I., Nakagawa, A., Hosaka, H., Wakatsuki, S., Mueller, F., and Brimacombe, R. (1998) *RNA* **4**, 542–550
- Müller, J. J., Zalkova, T. N., Zirwer, D., Misselwitz, R., Gast, K., Serdyuk, I. N., Welfle, H., and Damaschun, G. (1986) *Eur. Biophys. J.* **13**, 301–307
- Westhof, E., Romby, P., Romaniuk, P. J., Ebel, J.-P., Ehresmann, C., and Ehresmann, B. (1989) *J. Mol. Biol.* **207**, 417–431
- Moore, P. B. (1996) in *Ribosomal RNA-Structure, Evolution, Processing and Function in Protein Biosynthesis* (Zimmermann, R. A., and Dahlberg, A. E., eds) pp. 199–236, CRC Press, Boca Raton, FL
- Yonath, A., and Berkowitch-Yellin, Z. (1993) *Curr. Opin. Struct. Biol.* **3**, 175–181
- Stöffler-Meilicke, M., and Stöffler, G. (1990) in *The Ribosome: Structure, Function and Evolution* (Hill, W. E., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, D., and Warner, J. R., eds) pp. 123–133, American Society for Microbiology, Washington, D. C.
- Unge, J., AlKaradaghi, S., Liljas, A., Jonsson, B. H., Eliseikina, I., Ossina, N., Nevskaya, N., Fomenkova, N., Garber, M., and Nikonov, S. (1997) *FEBS Lett.* **411**, 53–59
- Leijonmarck, M., and Liljas, A. (1987) *J. Mol. Biol.* **195**, 555–580
- Bocharov, E. V., Gudkov, A. T., and Arseniev, A. S. (1996) *FEBS Lett.* **379**, 291–294
- Bocharov, E. V., Gudkov, A. T., Budovskaya, E. V., and Arseniev, A. S. (1998) *FEBS Lett.* **423**, 347–350
- Nakagawa, A., Nakashima, T., Taniguchi, M., Hosaka, H., Kimura, M., and Tanaka, I. (1999) *EMBO J.* **18**, 1459–1467
- Golden, B. L., Ramakrishnan, V., and White, S. W. (1993) *EMBO J.* **12**, 4901–4908
- Xing, Y. Y., and Draper, D. E. (1996) *Biochemistry* **35**, 1581–1588
- Davies, C., White, S. W., and Ramakrishnan, V. (1996) *Structure* **4**, 55–66
- Hoffman, D. W., Davies, C., Gerchman, S. E., Kycia, J. H., Porter, S. J., White, S. W., and Ramakrishnan, V. (1994) *EMBO J.* **13**, 205–212
- Fedorov, R., Nevskaya, N., Khairullina, A., Tishchenko, S., Mikhailov, A., Garber, M., and Nikonov, S. (1999) *Acta Crystallogr. Sec. D* **55**, 1827–1833
- Stoldt, M., Wöhnert, J., Görlach, M., and Brown, L. R. (1998) *EMBO J.* **17**, 6377–6384
- Davies, C., Bussiére, D. E., Golden, B. L., Porter, S. J., Ramakrishnan, V., and White, S. W. (1998) *J. Mol. Biol.* **279**, 873–888
- Kozin, M. B., Volkov, V. V., and Svergun, D. I. (1997) *J. Appl. Crystallogr.* **30**, 811–815