Supplementary Information

Energy barriers and driving forces of tRNA translocation through the ribosome

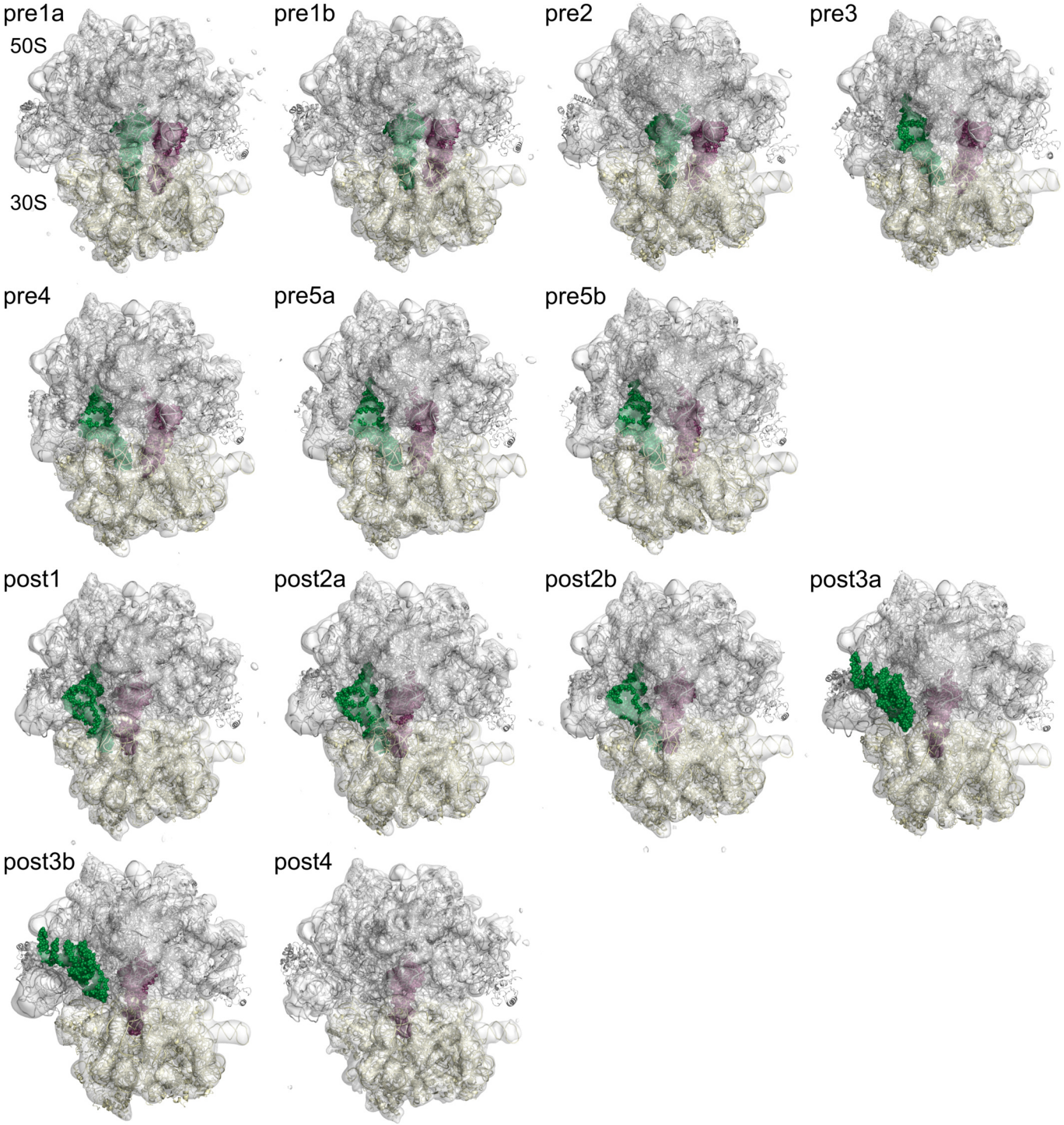
Lars V. Bock, Christian Blau, Gunnar F. Schröder, Iakov I. Davydov, Niels Fischer, Holger Stark, Marina V. Rodnina, Andrea C. Vaiana, Helmut Grubmüller

Contents

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

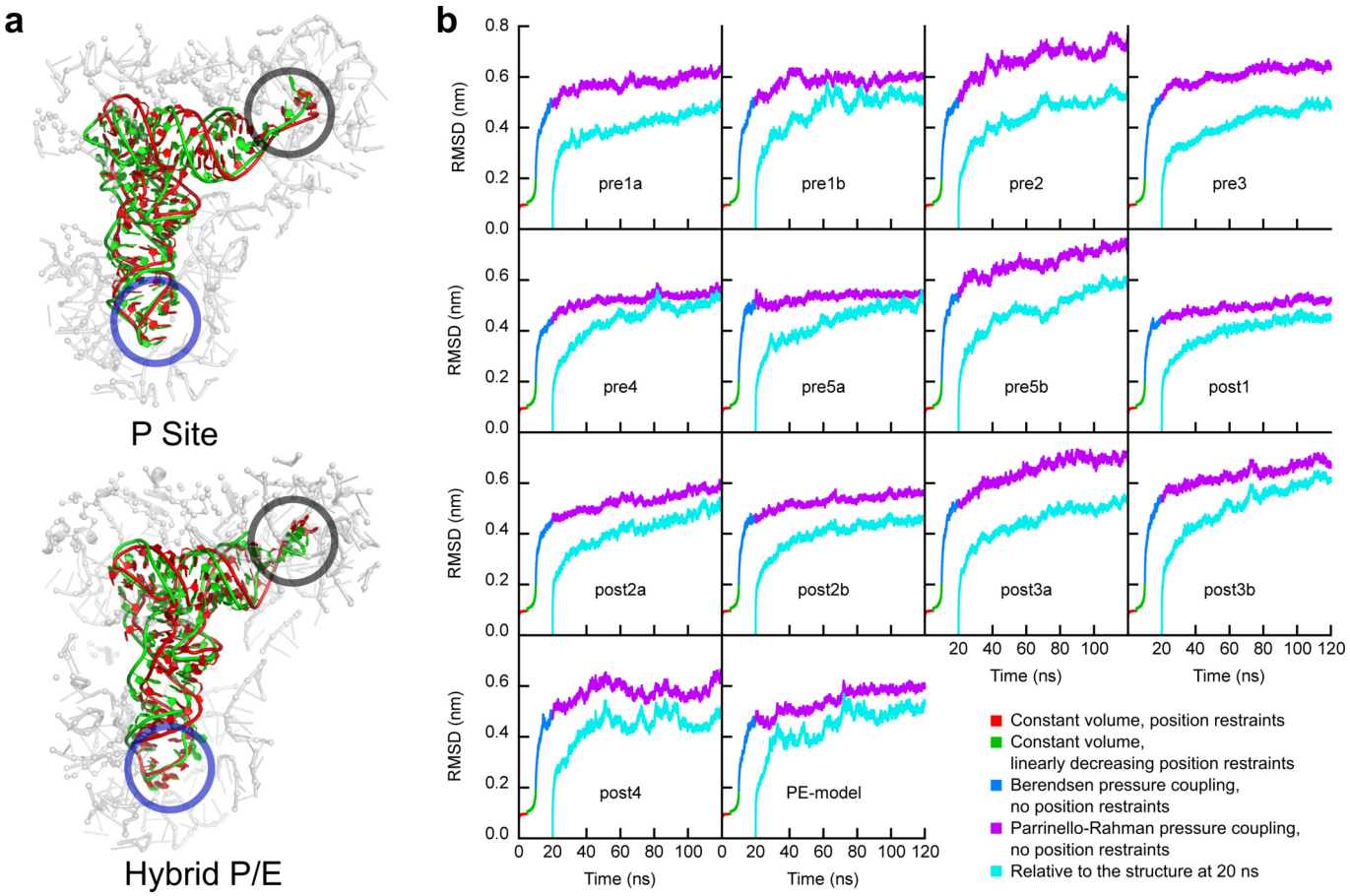
Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Figures



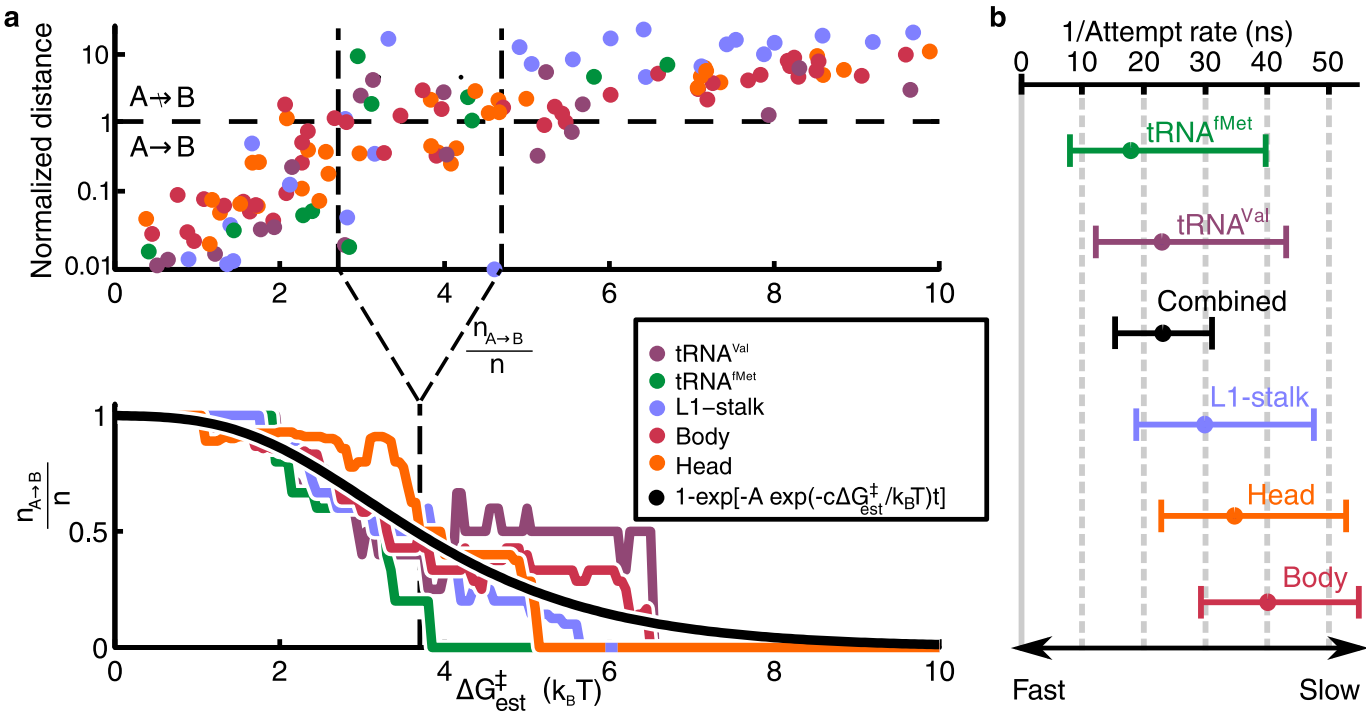
Supplementary Figure 1: All-atom models of pre1a–post4 states obtained from refinement of atomic models against cryo-EM maps[1]. For each state, the refined structure and an isosurface of the cryo-EM map (grey surface) are shown. The ribosomal subunits (50S and 30S) are shown in ribbon representation; $tRNA^{fMet}$ and $tRNA^{Val}$ atoms are depicted by magenta and green spheres, respectively.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690



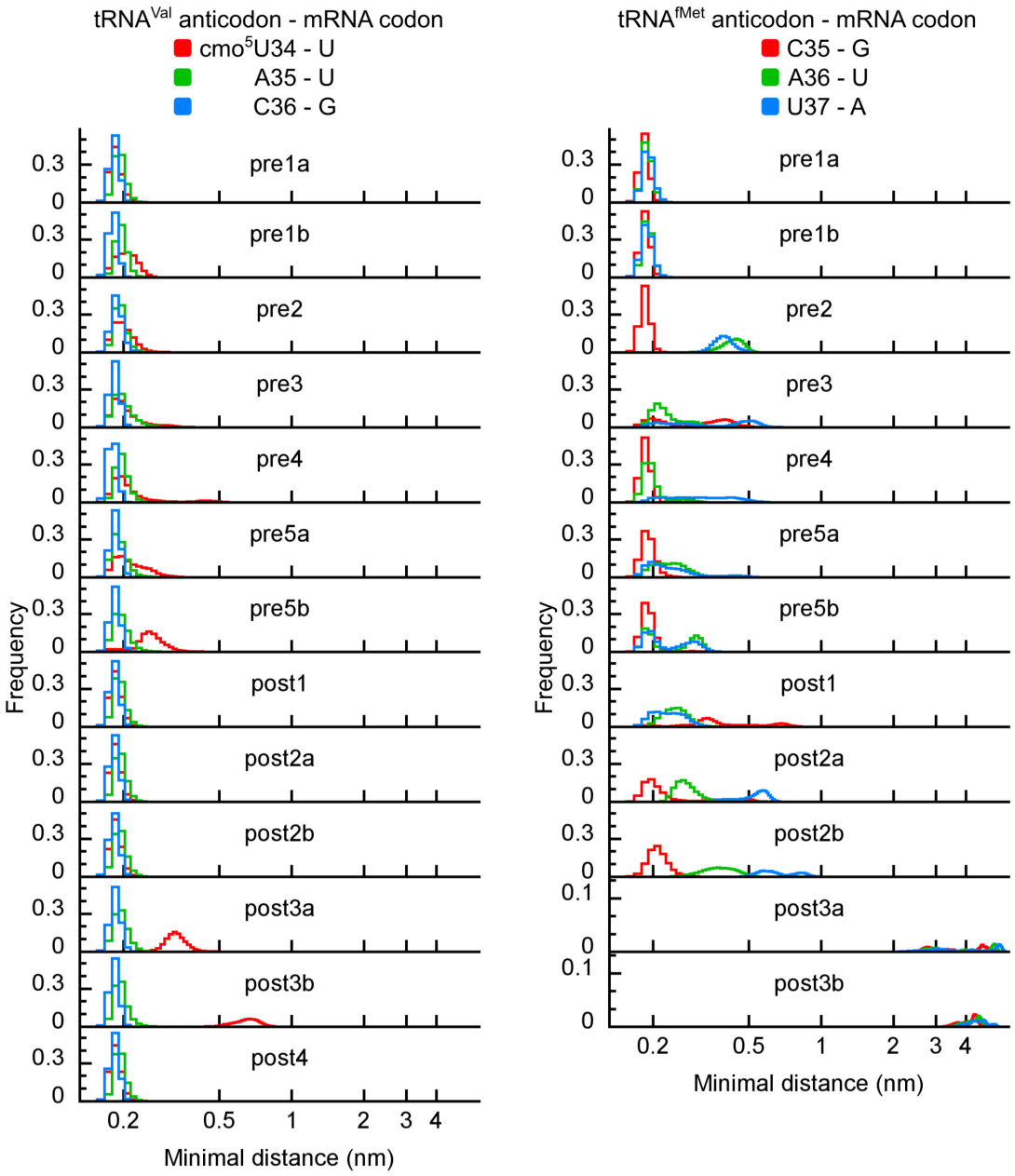
Supplementary Figure 2: Validation of models. (**a**) Comparison of tRNA positions between models and crystal structures in the P-site and P/E hybrid state. The tRNAs from crystal structures[2] and our models (left: pre1a, right: pre4) are shown as red and green ribbons, respectively, after rigid-body fitting of the binding region only (grey ribbons). C^{α} and P atoms used for fitting are depicted as grey spheres and CCA-tail and acceptor stem regions are indicated by black and blue circles, respectively. (**b**) Structural deviations during the simulations. For each ribosome simulation, started either from the model refined against the cryo-EM map or from the PE-model, the RMSD relative to the starting structure is shown for the different simulation steps (red, green, blue, and magenta curves), and relative to the structure at 20 ns (cyan curve).

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690



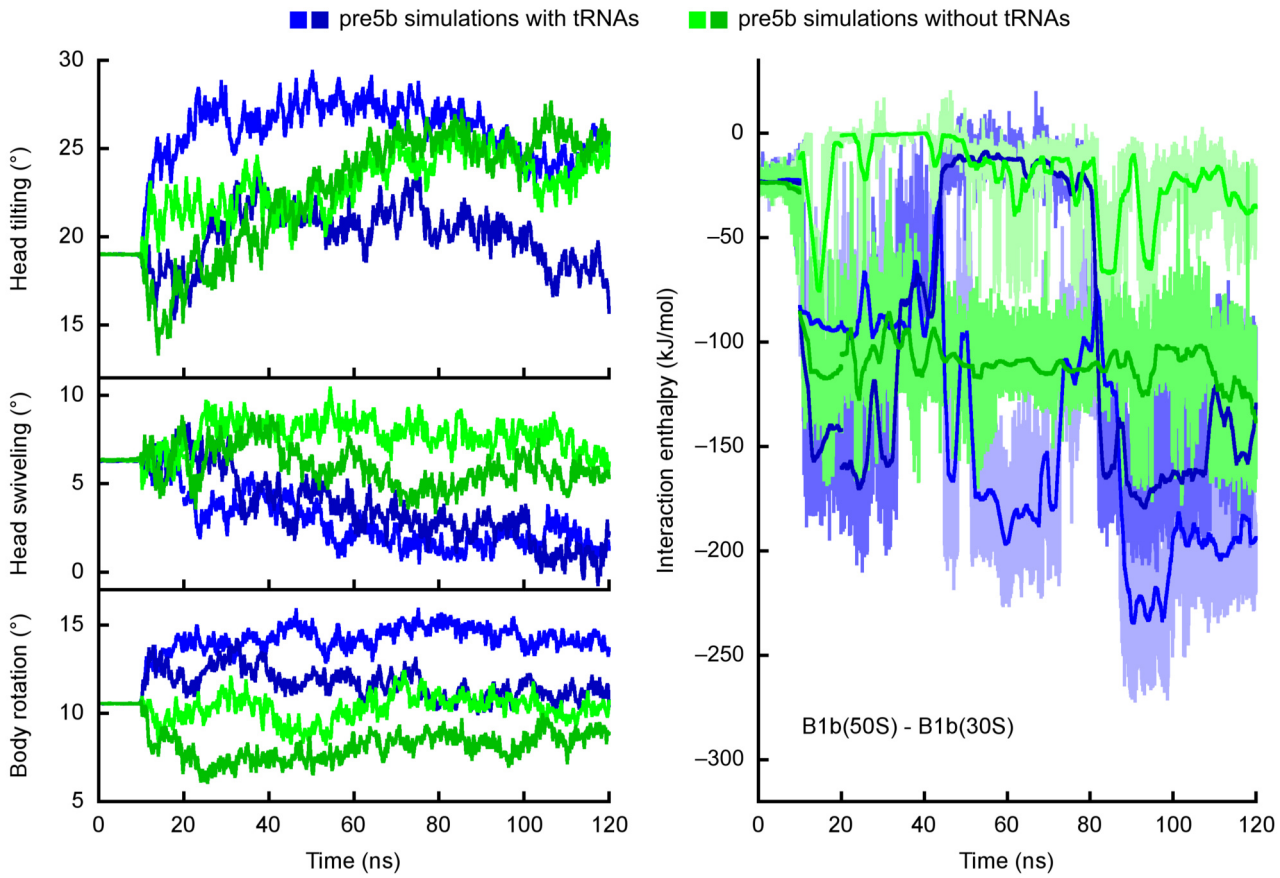
Supplementary Figure 3: Estimation of transition rates. (**a**) Attempt rate and free energy calibration factor. The upper panel shows an excerpt of the normalized distance between the ensembles for each pair of states versus the uncalibrated free energy estimate. This is done for each of the ribosome components (colored circles). A barrier between two states is considered crossed if this distance is smaller than one. The lower panel shows the frequency of barrier crossings $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}} = (n_{A\to B})/n$ calculated for free energy intervals of 1 k_bT (colored lines). The probability of barrier crossing $p_{\mathrm{A}\to\mathrm{B}}$ fitted to $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}$ is shown as a black line. (**b**) Statistical uncertainty of the attempt rate of the movement of individual ribosome components. Shown are the medium value of the distribution of the attempt rates A (circles) and standard deviation (bars). The overall attempt rate is shown as reference (black line).

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690



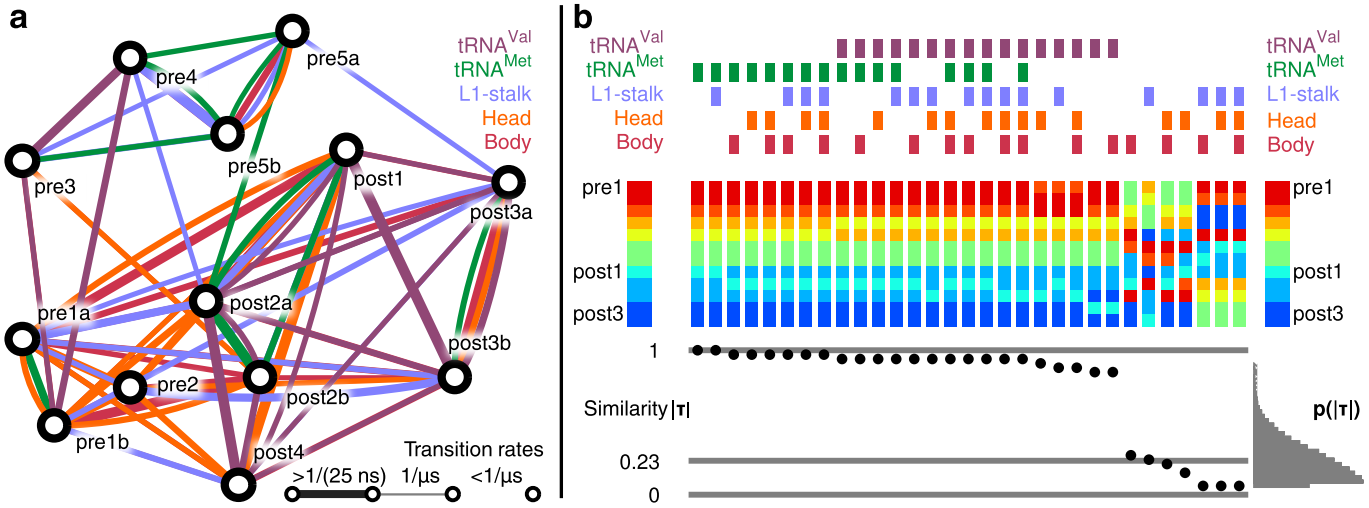
Supplementary Figure 4: Quality of tRNA-mRNA base-paring. For each state, histograms of the distances between codon residues of the mRNA and the corresponding anticodon residues of the two tRNAs are shown.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690



Supplementary Figure 5: Fast relaxation motions of the ribosome after tRNA removal during the simulations. Shown are time-traces of 30S head tilting, head swiveling, and body rotation angles (left panel), as well as of interaction enthalpies (right panel) for intersubunit bridge B1b, derived from four independent simulations. Blue curves refer to the two simulations started from the refined structure of the pre5b state with bound tRNAs, the green ones refer to simulations started from the same structure after removal of the tRNAs.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690



Supplementary Figure 6: Transiton Rates. (**a**) Schematic representation of the translocation intermediate states as a Markov model. Circles denote states, connecting lines encode the transition time estimates for L1-stalk, $tRNA^{fMet}$, $tRNA^{Val}$ motion as well as body and head rotation. We thank Benoit Roux for providing the idea. (**b**) Fastest progression sequences of translocation intermediate states ranked according to similarity to the sequence proposed by Fischer et al.[1] For all 31 possible combinations of ribosome components (top, color scheme as in Fig. 1a,d), the fastest progression sequence was determined as in 2.10. The similarity of each of the identified sequences (mid, columns) to the sequence given by Fischer et al.[1] was described using the absolute Kendall rank correlation coefficient τ (bottom). As a reference the mean τ value for random sequences (0.23) and their probability distribution $p(\tau)$ is shown.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Tables

Supplementary Table 1: Sequence of structure refinements against cryo-EM maps. For each state the starting model which was used for refinement against the corresponding map is shown. The pre5c structure was not used for simulations.

AP-states		PE-states		P-state	
state of refined structure	starting model	state of refined structure	starting model	state of refined structure	starting model
pre1b	AP-model	post1	PE-model	post4	P-model
pre1a	pre1b	post2a	post1		
pre2	pre1b	post2b	post2a		
pre4	pre1b	post3b	post2a		
pre3	pre4	post3a	post3b		
pre5a	pre1b				
pre5c	pre5a				
pre5b	pre5c				

Supplementary Table 2: Stereochemical parameters of our models compared to those of crystal structures. Shown are deviation of the distributions of key stereochemical parameters from distributions found in the protein data bank (PDB) for the 50S and 30S subunits. The upper part of each table shows deviations from the mean values in standard deviations σ. The relative width of the distribution compared to the standard distribution is shown in the lower part.

(a) 30S

	Models(pre1a–post4)		Dunkle et al.[2]		Zhang et al.[3]	
PDB id	100ns	refined	3R8N	3R8O	3I1Q	3I1Z
Ramachandran	-3.12	-7.62	-7.44	-7.52	-6.73	-6.96
χ_1-χ_2-rotamers	-2.04	-5.97	-6.97	-7.10	-6.26	-6.25
Backbone	-2.23	-4.66	-3.27	-3.47	-2.41	-2.39
Bond length	1.40	1.57	0.85	0.85	0.59	0.57
Bond angles	1.91	2.24	1.20	1.22	1.20	1.15
Omega angles	1.15	2.17	1.62	1.66	0.93	0.87
Side chain planarity	0.96	2.30	0.39	0.38	0.23	0.21
Improper dihedrals	1.04	1.90	0.75	0.72	0.53	0.46
Inside-Outside dist.	0.99	1.03	1.01	1.02	1.02	1.03

(b) 50S

	Models(pre1a–post4)		Dunkle et al.[2]		Zhang et al.[3]	
PDB id	100ns	refined	3R8S	3R8T	3I1R	3I20
Ramachandran	-3.05	-7.67	-5.75	-6.52	-6.01	-7.15
χ_1-χ_2-rotamers	-2.00	-5.95	-5.89	-6.23	-6.05	-5.78
Backbone	-1.98	-5.13	-1.92	-2.04	-2.78	-3.64
Bond length	1.33	1.48	1.49	1.04	0.77	0.61
Bond angles	1.98	2.34	1.53	1.38	1.40	1.21
Omega angles	1.14	2.22	1.52	1.47	1.20	0.87
Side chain planarity	1.00	2.32	0.50	0.41	0.31	0.20
Improper dihedrals	1.05	1.95	0.92	0.79	0.72	0.44
Inside-Outside dist.	0.99	1.02	0.99	0.99	1.00	1.03

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 3a: Contacting residues between $tRNA^{Val}$ and 23S.

$tRNA^{Val}$ residue	23S residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
C56	H38 A896													
G57	H38 A896													
Ψ55	H38 A896													
U17	H38 G882													
G19	H38 A896													
U17	H38 G881													
G18	H38 G882													
G19	H38 G882													
G19	H38 G883													
C56	H38 U895													
C56	H38 U894													
G20	H38 G883													
C56	H38 G882													
G18	H38 C897													
U17	H38 C898													
G18	H38 A896													
U17	H38 C897													
U17	H38 G880													
C56	H38 C897													
C56	H38 G881													
G20	H38 G882													
G19	H38 G881													
C56	H38 G880													
G57	H38 G880													
G57	H38 G881													
C25	H69 C1914													
U12	H69 U1915													
A38	H69 A1913													
G24	H69 C1914													
C11	H69 U1915													
U12	H69 A1916													
h^6m^1A37	H69 A1913													
C25	H69 A1913													
G39	H69 A1913													
C11	H69 C1914													
A26	H69 C1914													
G24	H69 U1915													
G10	H69 C1914													
G10	H69 U1915													
C27	H69 A1913													
C36	H69 A1913													
C27	H69 C1914													
U12	H69 G1910													
C13	H69 G1910													
C13	H69 C1909													
U12	H69 C1924													
U12	H69 U1923													
C11	H69 U1923													
C13	H69 C1924													
A69	H69 C1908													
A69	H69 G1907													
C70	H69 G1907													
C11	H69 C1909													
C72	H71 C1942													
A73	H71 U1943													
C71	H71 C1942													
C74	H71 U1943													
C74	H71 U1944													
C71	H71 C1941													
C72	H71 A1966													
C71	H71 A1966													
C71	H71 C1965													
V77	H74 A2451													
fMet78	H74 C2063													
fMet78	H74 A2439													
fMet78	H74 A2062													
fMet78	H74 A2451													
V77	H74 C2064													
V77	H74 A2450													
fMet78	H74 A2450													
fMet78	H74 C2064													
fMet78	H74 G2061													
C75	H74 A2451													
A76	H74 A2450													
A76	H74 C2064													
A76	H74 C2065													
C75	H74 A2450													
A76	H74 A2451													
V77	H74 C2063													
C75	H74 C2065													

Supplementary Table 3a: Continued: Contacting residues between $tRNA^{Val}$ and 23S.

$tRNA^{Val}$ residue	23S residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
A76	H74 C2063													
V77	H74 A2439													
fMet78	H74 U2441													
C75	H74 C2064													
V77	H74 A2062													
A76	H80 G2252													
A76	H80 G2253													
C75	H80 G2252													
A76	H80 C2254													
C75	H80 G2251													
A76	H80 G2251													
C74	H80 G2253													
C75	H80 G2253													
G3	H80 C2254													
G3	H80 G2255													
G2	H80 C2254													
C74	H80 G2252													
U4	H80 G2255													
A73	H80 G2251													
C72	H80 G2253													
A73	H80 G2252													
A73	H80 G2253													
G2	H80 G2255													
C74	H80 C2254													
G63	H81 G2279													
C56	H84 A2309													
C56	H84 G2308													
G19	H84 G2308													
G19	H84 A2309													
C56	H84 C2310													
G57	H84 A2309													
G57	H84 C2310													
A58	H84 A2309													
m^5U54	H85 A2327													
Ψ55	H85 C2326													
G52	H89 A2469													
U64	H89 C2483													
G52	H89 G2470													
U64	H89 A2482													
C51	H89 G2470													
G63	H89 C2483													
A76	H89 C2452													
G63	H89 A2482													
G52	H89 A2482													
G1	H89 C2462													
G63	H89 G2484													
fMet78	H89 C2452													
C51	H89 A2469													
V77	H89 C2452													
G53	H89 A2469													
V77	H89 U2506													
fMet78	H89 U2506													
A76	H89 U2506													
fMet78	H89 G2505													
C75	H89 C2452													
C74	H89 G2494													
C75	H90 C2507													
C74	H90 C2573													
A76	H90 C2507													
A76	H90 C2573													
C75	H92 G2553													
C74	H92 U2555													
C75	H92 U2554													
A73	H92 C2556													
C75	H92 U2555													
C74	H92 C2556													
A76	H92 G2553													
A76	H93 U2584													
A76	H93 A2602													
fMet78	H93 U2585													
A76	H93 G2583													
A76	H93 U2585													
C75	H93 G2583													
C75	H93 U2584													
V77	H93 U2585													
fMet78	H93 U2586													
C74	H93 A2602													
A73	H93 A2602													
fMet78	H93 U2584													
C75	H93 A2602													

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 3a: Continued: Contacting residues between $tRNA^{Val}$ and 23S.

$tRNA^{Val}$ residue	23S residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
A76	H93 U2604													
V77	H93 A2602													
V77	H93 U2584													
A76	H93 C2601													
A76	H93 A2600													
fMet78	H93 G2583													
A76	H93 U2586													
C74	H93 A2600													
C74	H93 C2601													
C72	H93 C2594													

Supplementary Table 3b: Contacting residues between $tRNA^{Val}$ and L5.

$tRNA^{Val}$ residue	L5 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
C56	A74													
C56	V73													
C56	G75													
Ψ55	A74													
C56	S72													
G20	K77													
C56	R79													
G19	A74													
G57	A74													
G19	R79													
G19	K77													
G57	V73													
G57	K77													
A58	K77													
C56	I78													
C56	K77													
C56	Q80													
G19	G75													
G19	V73													
G20	V73													
G19	F76													
G57	G75													
C56	K71													
G57	S72													
G20	F76													

Supplementary Table 3c: Contacting residues between $tRNA^{Val}$ and L16.

$tRNA^{Val}$ residue	L16 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
G53	R51													
m^5U54	R51													
G52	R55													
G53	R55													
G63	M1													
C62	M1													
m^5U54	R50													
C62	R44													
G63	R51													
m^5U54	T54													
G53	T54													
C62	R51													
G63	R44													
m^5U54	R59													
G53	R59													
U64	R44													
G53	R50													
A73	R81													
G1	N88													
G1	P77													
G1	R81													
G52	M1													
G63	Q3													
C62	R6													
m^5U54	K5													
G63	R10													
C72	R81													
A73	V80													
C74	V80													
U64	R10													
U64	Q3													
C75	R81													
G53	K5													
G63	R6													
G1	K76													
G2	N88													
G1	G87													
C51	M1													
C74	R81													
G1	E90													
G1	T74													
C65	Q3													
C65	R10													
G1	E75													
G2	G87													
A73	P77													
G52	R6													
G52	K5													
G53	R6													
C65	M1													
C51	K5													
A66	M1													
A66	Q3													
G2	K84													
G3	K84													
U64	R6													
U64	N88													
G53	K8													
G63	K8													
C65	R6													
G1	L78													
G52	K8													
m^5U54	K8													
C62	K8													
G1	G85													
G1	K86													
G2	K86													
G1	K84													
G1	V80													
G2	R81													
C65	K86													

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 3d: Contacting residues between $tRNA^{Val}$ and L27.

$tRNA^{Val}$ residue	L27 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
G3	ACE5													

Supplementary Table 3e: Contacting residues between $tRNA^{Val}$ and L33.

$tRNA^{Val}$ residue	L33 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
U17	ACE2													
U17	I4													
U17	G3													
U17	R27													
G19	R27													
C56	ACE2													
C56	R27													

Supplementary Table 3f: Contacting residues between $tRNA^{Val}$ and 16S.

$tRNA^{Val}$ residue	16S residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
A35	h18 G530													
cmo^5U34	h18 G530													
C36	h18 G530													
C36	h18 C518													
A35	h18 C518													
A38	h24 A790													
G39	h24 A790													
h^6m^1A37	h24 A790													
C31	h29 A1339													
G40	h29 A1339													
A41	h29 G1338													
C30	h29 A1339													
A41	h29 A1339													
C31	h29 A1340													
C32	h29 A1340													
C32	h29 U1341													
G39	h29 A1339													
G42	h29 G1338													
G40	h29 G1338													
C30	h29 G1338													
U29	h30 A1229													
C30	h30 C1230													
C30	h30 A1229													
C31	h30 C1230													
U29	h30 C1230													
C30	h30 G1231													
U29	h30 C1228													
cmo^5U34	h31 G966													
cmo^5U34	h34 C1054													
cmo^5U34	h34 A1196													
cmo^5U34	h34 A1197													
h^6m^1A37	h44 A1493													
A38	h44 A1493													
A35	h44 A1493													
C36	h44 A1493													
A38	h44 G1494													
cmo^5U34	h44 C1400													

Supplementary Table 3g: Contacting residues between $tRNA^{Val}$ and S9.

$tRNA^{Val}$ residue	S9 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
cmo^5U34	R129													
C32	K128													
C31	K128													
C31	R129													
U33	K128													
cmo^5U34	K128													
A35	K128													

Supplementary Table 3h: Contacting residues between $tRNA^{Val}$ and S13.

$tRNA^{Val}$ residue	S13 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
G44	K113													
G44	$NH_2$114													
G45	K113													
G43	K113													
G42	K113													
G44	P111													
G42	$NH_2$114													
G43	R112													
C28	K113													
G43	P111													
G43	$NH_2$114													

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 3i: Contacting residues between $tRNA^{fMet}$ and 23S.

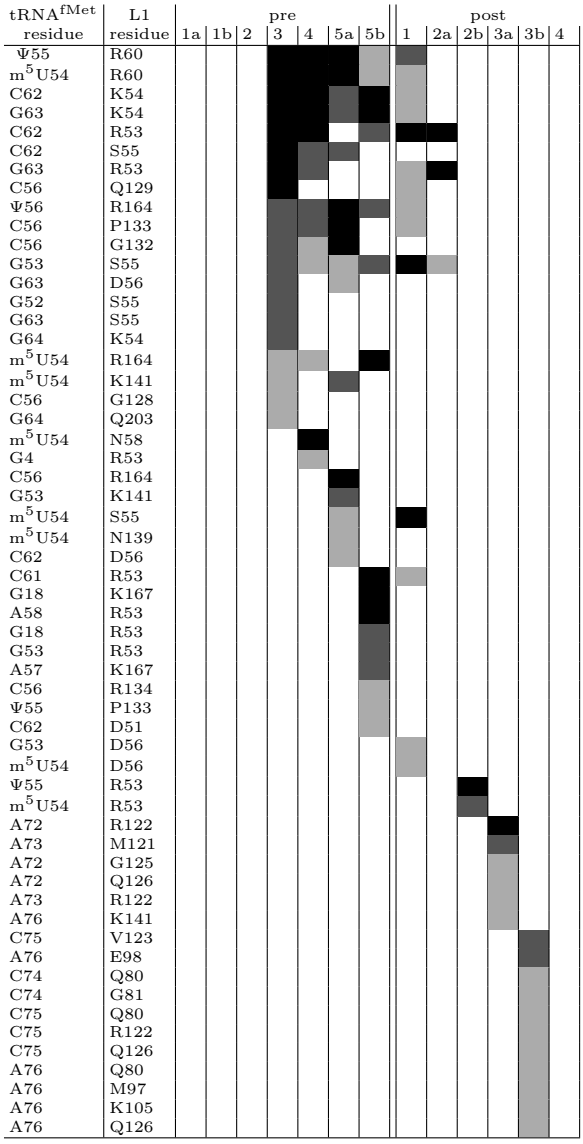
$tRNA^{fMet}$ residue	23S residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
A76	H13 C249													
G52	H38 G880													
G52	H38 G881													
G4	H68 G1850													
G4	H68 U1851													
C71	H68 C1893													
G70	H68 C1893													
C3	H68 U1851													
C71	H68 C1892													
C71	H68 U1851													
G70	H68 U1851													
C71	H68 U1852													
C71	H68 G1850													
A72	H68 G1891													
G70	H68 G1850													
G70	H68 C1892													
G4	H68 U1852													
G5	H68 U1851													
A76	H68 C1870													
A76	H68 A1871													
C75	H68 A1871													
C13	H69 C1924													
G12	H69 U1923													
G12	H69 C1924													
G12	H69 C1909													
C13	H69 C1909													
G12	H69 G1910													
U24	H69 U1923													
A11	H69 G1910													
U24	H69 G1922													
C25	H69 G1922													
A14	H69 C1925													
C13	H69 C1925													
C13	H69 G1907													
U36	H69 G1921													
U36	H69 G1922													
A37	H69 G1922													
A37	H69 U1923													
A76	H74 A2450													
C75	H74 C2064													
A76	H74 A2451													
A76	H74 C2064													
C75	H74 A2451													
A76	H74 C2065													
C75	H74 A2432													
A76	H74 A2432													
C75	H74 A2433													
C74	H74 A2433													
C74	H74 A2432													
U17A	H76 G2112													
G19	H76 G2112													
G19	H76 U2113													
G18	H76 G2112													
U17A	H76 U2111													
D20	H76 U2113													
D20	H76 G2112													
U17A	H76 U2180													
A21	H76 G2112													
U17A	H76 G2110													
A59	H76 G2112													
C17	H76 U2180													
C17	H76 U2181													
C56	H77 A2169													
G19	H77 A2169													
C56	H77 A2170													
C56	H77 U2122													
Ψ55	H77 A2169													
U17A	H78 C2145													
C3	H80 G2255													
C75	H80 G2251													
G2	H80 G2255													
G2	H80 C2254													
C1	H80 C2254													
C1	H80 G2253													
A76	H80 G2251													
C75	H80 G2252													
C3	H80 C2254													
C74	H80 G2252													
C74	H80 G2253													
C1	H80 G2255													

Supplementary Table 3i: Continued: Contacting residues between $tRNA^{fMet}$ and 23S.

$tRNA^{fMet}$ residue	23S residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
A76	H80 G2252													
C75	H80 G2253													
A76	H80 G2253													
C3	H80 G2256													
C74	H80 C2254													
C51	H84 A2309													
G52	H84 C2310													
G52	H84 A2309													
A76	H88 C2395													
A76	H88 C2394													
A76	H88 C2422													
C74	H88 C2422													
A76	H88 G2421													
C75	H88 C2422													
C75	H88 G2421													
A76	H88 G2396													
A76	H88 G2397													
C75	H88 G2397													
C75	H88 G2396													
C1	H88 G2397													
A76	H88 C2424													
A76	H88 U2423													
C74	H93 C2594													
A76	H93 A2602													
C74	H93 A2600													
C74	H93 A2602													
C75	H93 A2602													
A76	H93 C2601													

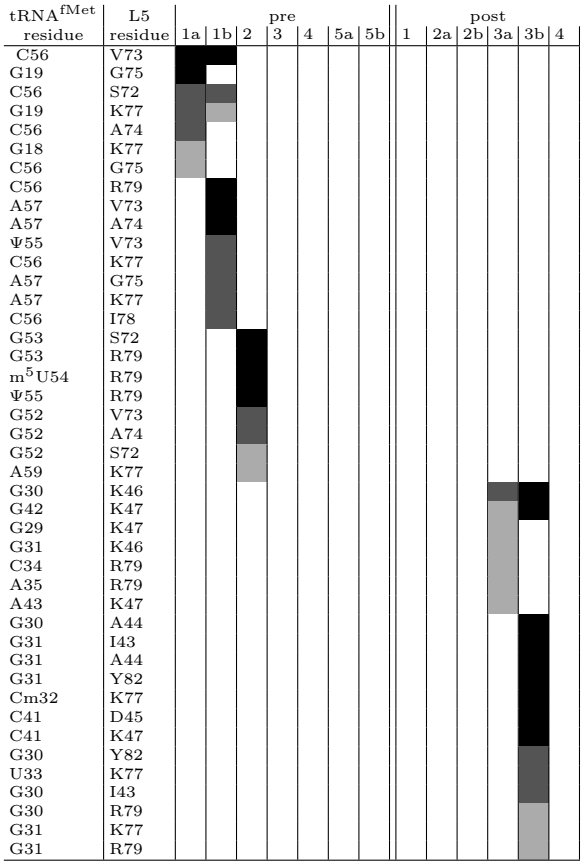
Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 3j: Contacting residues between $tRNA^{fMet}$ and L1.



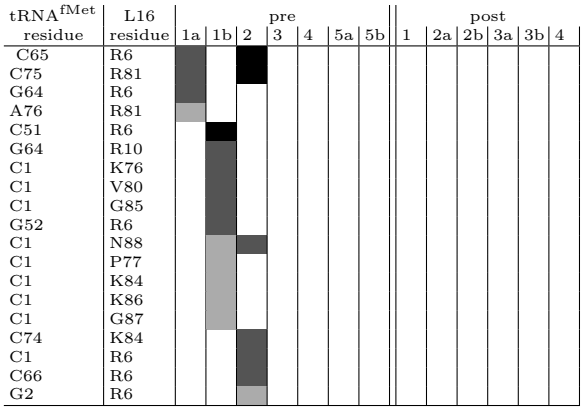
$tRNA^{fMet}$ residue	L1 residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
Ψ55	R60													
m^5U54	R60													
C62	K54													
G63	K54													
C62	R53													
C62	S55													
G63	R53													
C56	Q129													
Ψ56	R164													
C56	P133													
C56	G132													
G53	S55													
G63	D56													
G52	S55													
G63	S55													
G64	K54													
m^5U54	R164													
m^5U54	K141													
C56	G128													
G64	Q203													
m^5U54	N58													
G4	R53													
C56	R164													
G53	K141													
m^5U54	S55													
m^5U54	N139													
C62	D56													
C61	R53													
G18	K167													
A58	R53													
G18	R53													
G53	R53													
A57	K167													
C56	R134													
Ψ55	P133													
C62	D51													
G53	D56													
m^5U54	D56													
Ψ55	R53													
m^5U54	R53													
A72	R122													
A73	M121													
A72	G125													
A72	Q126													
A73	R122													
A76	K141													
C75	V123													
A76	E98													
C74	Q80													
C74	G81													
C75	Q80													
C75	R122													
C75	Q126													
A76	Q80													
A76	M97													
A76	K105													
A76	Q126													

Supplementary Table 3k: Contacting residues between $tRNA^{fMet}$ and L5.



$tRNA^{fMet}$ residue	L5 residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
C56	V73													
G19	G75													
C56	S72													
G19	K77													
C56	A74													
G18	K77													
C56	G75													
C56	R79													
A57	V73													
A57	A74													
Ψ55	V73													
C56	K77													
A57	G75													
A57	K77													
C56	I78													
G53	S72													
G53	R79													
m^5U54	R79													
Ψ55	R79													
G52	V73													
G52	A74													
G52	S72													
A59	K77													
G30	K46													
G42	K47													
G29	K47													
G31	K46													
C34	R79													
A35	R79													
A43	K47													
G30	A44													
G31	I43													
G31	A44													
G31	Y82													
Cm32	K77													
C41	D45													
C41	K47													
G30	Y82													
U33	K77													
G30	I43													
G30	R79													
G31	K77													
G31	R79													

Supplementary Table 3l: Contacting residues between $tRNA^{fMet}$ and L16.



$tRNA^{fMet}$ residue	L16 residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
C65	R6													
C75	R81													
G64	R6													
A76	R81													
C51	R6													
G64	R10													
C1	K76													
C1	V80													
C1	G85													
G52	R6													
C1	N88													
C1	P77													
C1	K84													
C1	K86													
C1	G87													
C74	K84													
C1	R6													
C66	R6													
G2	R6													

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 3m: Contacting residues between tRNA^{fMet} and L27.

tRNA^{fMet} residue	L27 residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
G2	G6													
G2	ACE5													
C1	ACE5													

Supplementary Table 3n: Contacting residues between tRNA^{fMet} and L28.

tRNA^{fMet} residue	L28 residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
C75	H19													

Supplementary Table 3o: Contacting residues between tRNA^{fMet} and L33.

tRNA^{fMet} residue	L33 residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
Ψ55	R27													
Ψ55	T28													
C56	R27													
C56	T28													
C65	R27													
C65	K26													
C66	K26													
C51	K29													
G52	K29													
C66	R27													
G52	P30													
G64	R27													
C65	I4													
C66	I4													
C1	I4													
C1	E6													
C74	G3													
A73	ACE2													
A73	G3													
C74	ACE2													
C74	I4													
C67	R27													
G64	K52													
C65	K52													
A37	P30													
A38	P30													
U36	P30													
G29	K29													
U36	T28													
U36	K29													
C28	K29													
C34	R27													
A35	R27													
A35	T28													

Supplementary Table 3p: Contacting residues between tRNA^{fMet} and 16S.

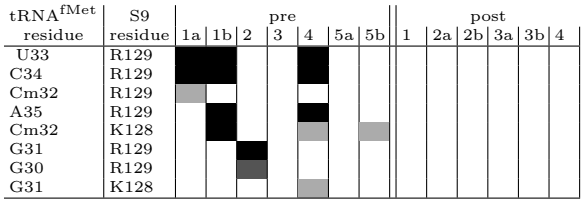
tRNA^{fMet} residue	16S residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
A37	h23 G693													
A38	h23 A694													
A37	h23 A694													
A38	h23 A695													
U36	h23 G693													
C39	h23 A694													
A38	h23 G693													
C39	h23 G693													
C39	h24 A790													
A38	h24 A790													
A37	h24 A790													
U36	h24 A790													
U36	h24 U789													
A37	h24 U789													
U33	h28 C1383													
C34	h28 C1383													
C41	h29 A1339													
G30	h29 A1339													
G31	h29 A1340													
C40	h29 A1339													
Cm32	h29 U1341													
G31	h29 A1339													
C41	h29 G1338													
Cm32	h29 A1340													
G29	h29 G1338													
G42	h29 G1338													
G31	h29 U1341													
G42	h29 A1339													
A43	h29 G1338													
G29	h29 A1339													
G30	h29 A1340													
C41	h29 A1340													
G30	h29 G1338													
U36	h29 A1339													
G29	h30 A1229													
G30	h30 C1230													
G42	h30 A1229													
A44	h30 C1228													
G30	h30 A1229													
G29	h30 C1230													
A43	h30 C1228													
A43	h30 A1229													
G31	h30 C1230													
G29	h30 C1228													
C28	h30 C1228													
C34	h31 G966													
A35	h31 G966													
C34	h44 C1400													

Supplementary Table 3q: Contacting residues between tRNA^{fMet} and S7.

tRNA^{fMet} residue	S7 residue	pre							post					
		1a	1b	2	3	4	5a	5b	1	2a	2b	3a	3b	4
Cm32	Q85													
U33	S82													
U33	T83													
A38	T83													
C40	A146													
C41	R142													
G42	R142													
Cm32	T83													
Cm32	Y84													
C40	M143													
C41	K135													
G29	K135													
U33	Q85													
C34	R78													
C40	N147													
C41	A146													
Cm32	R78													
U33	R78													

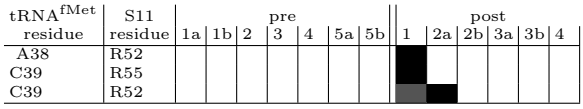
Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 3r: Contacting residues between $tRNA^{fMet}$ and S9.



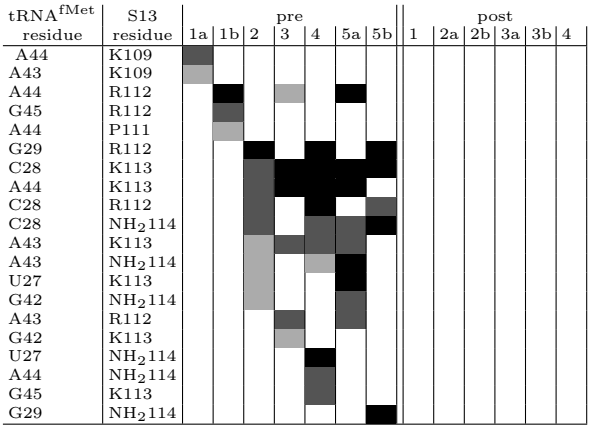
$tRNA^{fMet}$ residue	S9 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
U33	R129	■	■			■								
C34	R129	■	■			■								
Cm32	R129	▒												
A35	R129		■			■								
Cm32	K128		■			▒		▒						
G31	R129			■										
G30	R129			▓										
G31	K128					▒								

Supplementary Table 3s: Contacting residues between $tRNA^{fMet}$ and S11.



$tRNA^{fMet}$ residue	S11 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
A38	R52								■					
C39	R55													
C39	R52								▓	■				

Supplementary Table 3t: Contacting residues between $tRNA^{fMet}$ and S13.



$tRNA^{fMet}$ residue	S13 residue	pre 1a	1b	2	3	4	5a	5b	post 1	2a	2b	3a	3b	4
A44	K109	▓												
A43	K109	▒												
A44	R112		■		▒		■							
G45	R112		▓											
A44	P111		▒											
G29	R112			■		■		■						
C28	K113			▓	■	■	■	■						
A44	K113			▓	■	■	■							
C28	R112			▓		■		▓						
C28	$NH_2$114			▓		▓	▓	■						
A43	K113			▒	▓	▓	▓							
A43	$NH_2$114			▒		▒	■							
U27	K113			▒			■							
G42	$NH_2$114			▒			▓							
A43	R112				▓									
G42	K113				▒									
U27	$NH_2$114					■								
A44	$NH_2$114					▓								
G45	K113					▓								
G29	$NH_2$114							■						

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Table 4: Accession codes for models of spontaneous tRNA translocation. Shown are the EMDataBank accession codes for the cryo-EM densities used for refinement and the resolution of the respective cryo-EM maps. PDB-id codes are given for the 30S and 50S ribosomal subunit.

state	EMDB id	resolution [Å]	PDB id 30S	PDB id 50S
pre1a	1716	12	3J4V	3J52
pre1b	2472	12	3J4W	3J4X
pre2	1717	20	3J4Z	3J50
pre3	1718	17	3J4Y	3J51
pre4	1719	13	3J53	3J54
pre5a	2473	15	3J55	3J56
pre5b	1720	17	3J57	3J58
post1	1721	12	3J59	3J5A
post2a	1722	17	3J5B	3J5C
post2b	2474	17	3J5D	3J5E
post3a	2475	20	3J5F	3J5G
post3b	1723	15	3J5H	3J5J
post4	1724	9	3J5J	3J5K

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

Supplementary Note 1 (Methods)

1.1 General molecular dynamics setup

All molecular dynamics (MD) simulations were carried out with GROMACS 4[4] using the amber99sb force field[5], and the SPC/E water model[6]. Each simulated model was first solvated in a dodecahedron box keeping a minimum distance of 1.5 nm between the model atoms and the box boundaries. K^+ and Cl^- forcefield parameters were taken from Joung and Cheatham[7]. Long-range electrostatic interactions, beyond 0.9 nm were calculated by particle-mesh Ewald summation[8] with a grid spacing of 0.12 nm. Lennard-Jones interactions were calculated within a distance of 1.4 nm. Coordinates were recorded for analysis every 2 ps. Unless stated otherwise, an integration time step of 4 fs was used, applying virtual site constraints[9]. All bond lengths were constrained with the LINCS algorithm[10]. The system temperature was kept constant at $T = 300$ K using velocity rescaling[11] with a coupling time constant of $\tau_T = 0.1$ ps. Protonation states of amino acids were determined with WHATIF[12].

1.2 Models of the ribosome including tRNAs

Three initial atomic models of the *E.coli* ribosome were built, which were subsequently refined against the cryo-EM maps provided by Fischer et. al[1]: First, a model of the ribosome with a P-site fMetVal-tRNA$^{\mathrm{Val}}$ (P-model), second, a model with a P-site fMetVal-tRNA$^{\mathrm{Val}}$ and an E-site tRNA$^{\mathrm{fMet}}$ (PE-model), and third, a model with an A-site fMetVal-tRNA$^{\mathrm{Val}}$ and P-site tRNA$^{\mathrm{fMet}}$ (AP-model).

All models were constructed from the crystal structure by Zhang et al.[3]. This was the best resolved (resolution: 3.19 Å) and most complete *E.coli* ribosome structure at the time of modeling. All structural information (pdb ids: 3I1P, 3I1O), including the crystallographic water molecules and ions, was used.

For the L1 protein and the parts of the L1-stalk rRNA which are not resolved in the structure of Zhang et al., the *T.thermophilus* ribosome structure of Gao et al.[13] (pdb id: 2WRI) was used. A homology model of the L1 protein was built using the swissmodel server (swissmodel.expasy.org) with the *E.coli* sequence and the *T.thermophilus* structure as a template (44% sequence identity).

In the *E.coli* structure, 68 nucleotides (2111–2179) are not resolved in the L1-stalk rRNA. These were also modeled using the *T.thermophilus* structure. The corresponding structurally aligned nucleotides[13] as well as ten nucleotides upstream and downstream, which form ten base pairs at the stem of the L1-stalk, were extracted from the *T.thermophilus* structure. The extracted nucleotides were mutated to match the *E.coli* sequence using WHATIF[12] (58% sequence identity).

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

Helix 38 (H38) of the large ribosomal subunit forms contacts to the small subunit via intersubunit bridge B1a[14]. Nucleotides G879–C897 of this helix are not resolved in the *E.coli* structure. Secondary structure prediction software S2S[15] predicted a pentaloop for the tip of the helix. The nucleotides of a pentaloop from an NMR structure[16] (pdb id: 1NA2) were mutated to match H38 sequence using WHATIF. To fit this rRNA structural motif into the model of the ribosome, the same protocol as for the L1-stalk rRNA was used.

The ribosomes used for the cryo-EM experiments contained an fMetVal-tRNA$^{\mathrm{Val}}$. Since no high resolution structure was available, a solution structure of tRNA$^{\mathrm{Val}}$ from a refinement of a homology model against residual dipolar coupling and SAXS data (pdb id: 2K4C)[17] was used for the models.

The *T.thermophilus* structure by Yusupova et al.[18] (pdb id: 2HGP) contains a P-site tRNA$^{\mathrm{Phe}}$ and

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

a 50 nucleotide long mRNA. This structure was rigid-body fitted to our model using structurally aligned nucleotides[18] from 16S rRNA of the small subunit. The fitted coordinates of the tRNA phosphates and of the mRNA were stored for later use. From this mRNA structure, the A-, P-, and E-site codons as well as three upstream and downstream nucleotides were extracted. These nucleotides were mutated with WHATIF to match the sequence of the mRNA used in the cryo-EM experiments. Appropriate tRNA modifications and the dipeptide were added to match experimental conditions[1]. Atom types for fMet were obtained with ANTECHAMBER[19], partial charges were determined using DFT-b3lyp with a 6-31/G* basis set. Side chain charges are the same as in Met. On the backbone, only the charges of the formylamino cap changed more than 5 %. The modified fMetVal-tRNA$^{\mathrm{Val}}$ structure was then fitted as a rigid body to the P atoms obtained from the fitted Yusupova structure. Next, a 1 ns simulation of the tRNA and the mRNA in solvent with position restraints on the P and C1' atoms was carried out. In the subsequent 5 ns simulation, positions of P and C1' atoms were restrained to those of the fitted Yusupova structure, thereby linearly increasing the force constant from 0 to 1000 kJ mol^{-1} nm^{-2} during the simulation. The resulting tRNA structure was then included into the ribosome model, yielding the P-model.

The *T.thermophilus* structure by Gao et al.[13] (pdb id: 2WRI) contains an E-site tRNA$^{\mathrm{fMet}}$. The 30S subunit structure was rigid body fitted to our model using structurally aligned nucleotides[13] from 16S rRNA. Nucleotide modifications were added to the tRNA to match the modifications of the tRNAs in the cryo-EM experiments and a 1 ns simulation of the tRNA in solvent with position restraints on the P and C1' atoms was performed. The resulting tRNA structure was then included into the P-model, yielding the PE-model.

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

1.3 Refinement of the atomic models against cryo-EM maps

The three ribosome models obtained from crystal structures as described in the previous paragraph were used as starting structures for subsequent refinement against the 13 different cryo-EM density maps, to provide an all-atom interpretation of each individual conformational state. Initial placement of a starting model into a density map was done using the rigid-body fit feature of the program Chimera[20]. The real-space refinement program DireX[21] was then used for all refinements. DireX computes a density map from an atomic model and refines the atomic coordinates to maximize the overlap between this model map and the cryo-EM density map. The quality of the refinement by DireX has been found to be similar or better than that of other established methods in a comparative study[22]. The model density maps were generated using a Gaussian kernel with a width adapted to the resolution of the corresponding cryo-EM density map.

For each refinement, 2000 steps were performed which took 36 hours on average on one core of an Intel Core 2 Quad Processor Q9300 (2.5 GHz).

The initial AP-, PE-, and P-models were refined against the map which most closely resembled the state of the model (respectively: pre1b, post1, and post4). Subsequently, fitted structures were used as starting models for refinement against the remaining 10 maps in the sequence described in Supplementary Table 1.

Due to a program bug during the refinement process with DireX, in the 30S subunit, the tRNAs and the mRNA, several amino acids and nucleotides had wrong chiralities. In the structures extracted after 20 ns pre-equilibration, the errors were corrected by placing the chiral center atom on the opposite side of the plane defined by the three chiral neighbor heavy atoms. To that goal, the bond vector between the chiral center atom and the respective bound hydrogen atom was used for shifting the chiral center and the bound hydrogen atom. Subsequently, the bound hydrogen atom was flipped to the other side of the chiral center atom using the same bond vector. Alternatively, for the C2' in nucleotide sugars, chiral errors were corrected by swapping the positions of the O2' and H2' atoms. All corrected structures were again energy-

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

minimized. Since all the simulations were started from the structures containing these errors, we performed a 60 ns simulation from the energy-minimized corrected pre5b structure at 20 ns to make sure that these errors do not influence our results and conclusions. The pre5b state was chosen, because the refined structure contained the highest number of errors of all the structures. The root mean square deviation (RMSD) relative to the structure at 20 ns was calculated, a Principal Component Analysis (PCA) of $tRNA^{Val}$, $tRNA^{fMet}$, and L1-stalk motions was carried out and the intersubunit rotation angles were calculated as described in 2.3. The results were compared to the two independent pre5b simulations which started from different structures at 20 ns in order to see the effect of different starting structures (representing the same state) compared to the effect of changed chiralities in two simulations starting from the same structure. The PCAs of the tRNAs and the L1-stalk motions as well as the intersubunit rotation angles showed larger differences between the simulations starting from different structures than the difference due to the changed chiralities. The RMSD of the simulation started from the corrected structure was not markedly different from that of the other two simulations of the pre5b state. Hence, possible inaccuracies due to changed chiralities are found to be smaller than the statistical uncertainty due to limited sampling and, therefore, not significant. Next, we investigated the effect the changed chiralities have on our identification of residues involved in contacts between the tRNAs and proteins L1, L5 and L16. Each of these residues was characterized by the maximum of all its contact frequencies. To measure the similarity of contacting patterns in two simulations, we calculated the ratio of the number residues with the same level of contact frequency (12.5–25%, 25–50%, 50%–100%) in both simulations to the number of residues with different levels. The ratio extracted from the simulations with wrong chiralities and different starting structures was 0.56 and the ratio for the two simulations with the same starting structure but different chiralities was 1.16. Again, the possible inaccuracies due to changed chiralities are smaller than the estimated error range due to limited sampling. The energy-minimized corrected structures were submitted to the pdb-database under pdb-ids shown in Supplementary Table 4. Amino acids and nucleotides which were corrected for chirality are listed in the header of the pdb-files.

1.4 Choice of models for simulation.

For each of the major states of spontaneous retro-translocation (pre1 to post4), the fit to the one or two cryo-EM maps with the highest resolution were used as starting structures for MD simulations: pre1a (12 Å), pre1b (12 Å), pre2 (15 Å), pre3 (17 Å), pre4 (13 Å), pre5a (15 Å), post1 (12 Å), post2a (17 Å), post2b (17 Å), post3b (15 Å), and post4 (9 Å). Two additional structures were chosen for simulations, because of extreme intersubunit rotation angles (pre5b) and an extreme $tRNA^{fMet}$ conformation (post3a), to capture a large range of conformations accessible by the ribosome and the tRNAs. To estimate the effect of refinement accuracy on our conclusions, the refined structure of the pre3 state was perturbed such that the conformation of $tRNA^{fMet}$ was closer to the pre2 state, but within the resolution limits set by the cryo-EM density.

1.5 MD simulations of the refined models

Atomic models, including the crystallographic resolved ions, obtained from the flexible fitting to 13 cryo-EM maps were solvated, and the system was neutralized with K^+ ions before adding additional explicit salt (7 mM $MgCl_2$ and 150 mM KCl) using the GENION program from the GROMACS suite[4] to mimic the conditions used for the cryo-EM experiments[1]. The system was then equilibrated in four steps:

- 0–5 ns: Constant volume and position restraints on all ribosomal heavy atoms with a force constant of 1000 kJ mol^{-1} nm^{-2}, 2 fs time step.
- 5–10 ns: Constant volume and linearly decreasing the position restraints force constant to zero.
- 10–20 ns: The pressure was coupled to a Berendsen barostat[23] with a coupling constant $\tau_p = 1$ ps and an isotropic compressibility of $4.5 \cdot 10^{-5}$ bar^{-1}.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

- 20–120 ns: The pressure was coupled to a Parrinello-Rahman barostat[24] with a coupling constant $\tau_p = 1$ ps and an isotropic compressibility of $4.5 \cdot 10^{-5}$ bar^{-1}.

For each state as well for the initial PE-model, one simulation was carried out except for state pre5b where two independent simulations were carried out. Additionally, two simulations of the pre5b state without tRNAs were completed. Here, the tRNAs were removed from the structure refined against the pre5b state and the ribosome structure was subsequently solvated and equilibrated in the same way as the other structures.

1.6 Cryo-EM analysis of global ribosome conformation

Cryo-EM grids of vacant *E.coli* ribosomes were prepared at 18°C under controlled environmental conditions[25] and imaged under cryo conditions with a Titan Krios electron microscope (FEI Company, the Netherlands) at 300 kV and defoci ranging from 1 μm to 2.5 μm on a Eagle 4k x 4k CCD camera (FEI Company, the Netherlands) using 2fold pixel binning, resulting in a final pixel size of 3.2 Å. Ribosome particles were selected semi-atomically with Boxer[26] and corrected locally for the CTF[27]. The resulting 9814 ribosome particles were coarsened twofold to a pixel size of 6.4 Å and classified according to 30S body rotation in 2.5° steps as described[1]. Pre-translocation state ribosome particles (315108 in total) were obtained from an existing dataset of *E.coli* ribosome complexes prepared for cryo-EM at different time-points of the retro-translocation reaction, using the same buffer and temperature of 18°C as for the vacant ribosomes[28;1]. Hierarchical classification resulted in 34 groups of ribosome particles representing structurally distinct pre-translocation states[1]. For each population of pre-translocation ribosomes, the 30S body rotation of the corresponding cryo-EM reconstruction was determined. In Fig. 2, the fraction of particles as a function of 30S body rotation was plotted in 2.5° steps for all pre-translocation state ribosomes and vacant ribosomes, respectively. Image processing was generally performed using IMAGIC-5[29] and exhaustive alignment[30].

Supplementary Note 2 (Analysis)

2.1 Comparison to recent crystal structures

To compare our models to existing crystal structures for each state, an average structure was calculated from the last 10 ns of each trajectory. The root mean square deviations (RMSD) of these structures relative to two *E.coli* ribosome crystal structures (pdb ids: 3R8S, 3R8T[2]) were calculated after rigid-body fitting using all resolved 70S C^α and P atoms, except L9 protein atoms (Fig. 1b).

Of particular importance for our analysis is the quality of our models in the tRNA binding region. To assess the accuracy of the models in this region, the RMSD of the C^α and P atoms which are within a 2-nm distance to the two tRNAs in any of the models was calculated after rigid-body fitting (Fig. 1b).

2.2 Independent MD based refinement

In order to provide an independent test of our refinement procedure, we performed an additional refinement of the AP-model against the pre1b cryo-EM map using an all-atom explicit solvent MD simulation with an additional biasing potential and no further restraints. This biasing potential maximizes correlation between the atomic model, using an adapted Gaussian kernel as described above, and the cryo-EM map[31;32]. The model density maps were calculated for each simulation time step, an effective potential constant[31] of $k = 10^6$ kJ/mol and a total simulation time of 2.7 ns were used. The AP-model was solvated in a cubic simulation box which matches the geometry of the cryo-EM map.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

2.3 Definition of reaction coordinates for collective motions

Principal component analysis

Principal component analysis (PCA)[33] was carried out for $tRNA^{Val}$, $tRNA^{fMet}$, and the L1-stalk using all the simulations of the whole ribosome. To define common sub-spaces, all trajectories were first superimposed by least square fit, using C^{α} and P atoms of the 50S subunit excluding the L1-stalk. Next, the C^{α} and P atoms of the tRNAs and the L1-stalk were extracted from the trajectories. For each of the three ribosomal components the extracted trajectories of all states were concatenated, and the atomic displacement covariance matrix was calculated. The trajectories of each state were then projected on the first eigenvector of this matrix. The projections divided by $\sqrt{N}$, where N is the number of atoms used to construct the covariance matrix, yielded the reaction coordinates (r.c.). The minimum and maximum of this reaction coordinate for each state are shown in Fig. 1c. The projections on the first three eigenvectors were used to estimate transition rates (see below).

In order to calculate the distance dependence of the interaction enthalpies, i.e. the sum of electrostatic and Lennard-Jones interactions, between $tRNA^{fMet}$ and the L1 protein, a distance coordinate was obtained from a PCA using C^{α} and P atoms of the tRNA and the L1-stalk. The atomic displacement covariance matrix was constructed from the trajectories of those states in which the tRNA and the L1 were in contact (pre3–pre2b). The projection on the first eigenvector of this matrix, divided by $\sqrt{N}$, where N is the number of atoms used to construct the covariance matrix, was used as the distance coordinate, where the smallest value observed in the simulations was set to zero.

30S head and body rotation

Head and body rotations were quantified by comparing structures of each state, extracted from the respective trajectories at 200 ps intervals, to the post1a structure at 20 ns. The post1a structure was used to define zero degree body and head rotations. To define the axes of rotation and pivoting points, we extended a non-linear least squares fitting method[34] to also include the axis of minimal and median rotation. Final mean pivoting point and axes of rotation for head and body movement were determined by quaternion-based averaging[35] over all rotations obtained from all structures of all states. Rotation angles were then calculated relative to the mean axes of rotation and pivoting points.

2.4 Transition rate estimates

Transition rates were estimated in two steps. Firstly, initial estimates for the free energy barrier heights $\Delta G^{\ddagger}_{\mathrm{est}}$ were obtained from a fluctuation analysis of the trajectories of all states. Secondly, these barrier estimates were calibrated by comparing passage frequencies obtained from the initial barrier estimates to passage frequencies actually observed in the simulations, to yield the free energy barriers $\Delta G^{\ddagger}$. The calibrated free energy barrier heights $\Delta G^{\ddagger}$ were then used to calculate the transition rates shown in Fig. 1d.

Free energy barrier estimates

The initial free energy barrier estimates $\Delta G^{\ddagger}_{\mathrm{est}}$ for the transitions between all states for the motions of the L1-stalk, both tRNAs, and 30S head and body rotations were obtained as follows. Each trajectory (state) was projected onto 3-dimensions spanned by the dominant PCA eigenvectors for the L1-stalk and tRNA motions and by the three rotation angles for the 30S head and body rotation. The mean $\boldsymbol{\mu}$ and the 3×3 covariance matrix $\boldsymbol{\Sigma}$ of each projected trajectory define a three-dimensional multivariate Gaussian probability distribution function,

$$p(\boldsymbol{x}) = \frac{w_{\mathrm{exp}}}{\sqrt{2\pi^3 \|\boldsymbol{\Sigma}\|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}, \quad (1)$$

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

for the corresponding state, where the weighting factor w_{exp} accounts for the experimentally measured population of the state[1]. The free energy landscape

$$G(\boldsymbol{x}) = -k_{\mathrm{B}}T \ln p(\boldsymbol{x}) \tag{2}$$

arising from such a distribution $p(\boldsymbol{x})$ describes a three-dimensional quasi-harmonic approximation to the underlying free energy landscape. For each pair of states, the intersection of the two quasi-harmonic approximations obtained for a given motion defines a hypersurface, the free energy minimum of which was used as an estimate for the barrier height $\Delta G^{\ddagger}_{\mathrm{est}}$ between the two states. To test whether barrier estimates change with the number of dimensions used for the analysis in PCA space, up to 20 dimensions were used for the L1-stalk and tRNA trajectory projections. Even though the absolute values for the free energy barrier estimates increased, their ratios to the respective mean value did not change markedly. Therefore, we assume $\Delta G^{\ddagger} = c\Delta G^{\ddagger}_{\mathrm{est}}$ with a constant calibration factor c.

To determine the statistical uncertainty of the free energy barrier estimate we randomly drew N points from distributions with given mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and recalculated mean $\boldsymbol{\mu}'$ and covariance matrix $\boldsymbol{\Sigma}'$ from the drawn samples. The number of stochastically independent data points that define the distributions was set to $N = N_{\mathrm{total}}\frac{t_{\mathrm{ac}}}{t}$, where t_{ac} denotes the autocorrelation time of the data points and t the total simulation time. This was repeated until the standard deviation of the free energy estimates for each transition with these newly defined harmonic potentials converged. The statistical uncertainty of the free energy barriers for each transition is the converged standard deviation of the free energy estimates.

Calibration of Arrhenius transition rates

Several conformational transitions between the 13 conformation states defined in Fig. 1c were actually observed during our 100 ns simulations. Comparison of the statistics of the observed transitions with transition rates obtained from the above barrier height estimates using Arrhenius' law therefore allowed to refine all barrier heights and transition rates by a common factor. To this aim, the barrier height estimates were grouped into height intervals of 2.494 kJ/mol width. For each interval, the fraction $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}$ of trajectories for which conformational transitions were observed was determined (cf. Supplementary Fig. 3a) as follows. The trajectories were projected onto the reaction coordinates for the six collective motions defined above and analyzed in 200 ps intervals for the rotational movements and 10 ps intervals for the projections in PCA space. The distance $d_{\mathrm{A,B}}$ between two state ensembles A and B was defined as the minimum distance of all projections onto the reaction coordinates. A barrier between A and B states was considered to be crossed if the average distance within an ensemble was found to be larger than the distance $d_{\mathrm{A,B}}$. For each motion, the frequency of barrier crossing $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}$ in an energy interval $I = [\Delta G^{\ddagger}_{\mathrm{est}} - \Delta G, \Delta G^{\ddagger}_{\mathrm{est}} + \Delta G]$ was defined as

$$p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}\left(\Delta G^{\ddagger}_{\mathrm{est}}\right) = \frac{n_{\mathrm{A}\to\mathrm{B}}}{n}, \tag{3}$$

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

$$p_{\mathrm{A}\to\mathrm{B}} = 1 - \exp\left(-kt\right), \tag{4}$$

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

$$k = A \exp\left(-\Delta G^{\ddagger}/k_{\mathrm{B}}T\right), \tag{5}$$

where k_{B} is the Boltzmann constant, and T is the temperature. Assuming the same linear calibration $\Delta G^{\ddagger} = a + c\Delta G^{\ddagger}_{\mathrm{est}}$, of all barrier heights and attempt rates, respectively, the calibration factors A and c were determined from a least square fit of the Arrhenius transition probability

$$p_{\mathrm{A}\to\mathrm{B}} = 1 - \exp\left[-A \exp\left(\frac{-c\Delta G^{\ddagger}_{\mathrm{est}}}{k_{\mathrm{B}}T}\right)t\right] \tag{6}$$

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

to the respective fraction $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}$ observed in the simulations. From the refined energy barrier heights, transition time estimates $\tau = 1/k$ were obtained from Arrhenius' law, Eq. (5).

The statistical uncertainty of attempt rate A and calibration factor c were determined by reconstructing $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}$ from the free energy barrier estimate for a given transition. We randomly determined whether this transition would occur within a 100 ns simulation time using Eq. (6) with a random shift in $\Delta G^{\ddagger}_{\mathrm{est}}$, that accounts for the error estimated for $\Delta G^{\ddagger}_{\mathrm{est}}$. Fitting the reconstructed $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}$ to Eq. (6) yields a new calibration factor and attempt rate. When this procedure is repeated, the standard deviation of the redetermined calibration factors and attempt rates converges to the statistical uncertainty of the calibration factor and attempt rate.

2.5 tRNA contacts with the ribosome and mRNA

To assess the residue-residue contacts and the interaction enthalpy between the tRNAs and the ribosome, for each simulation, all pairs of atoms, respectively from the tRNAs and the ribosome, whose distances were below 3 Å were identified using g_contacts[36]. A residue pair was considered to be in contact if the distance between any two atoms (one from each residue) was found to be below 3 Å in at least one frame of the 100 ns trajectory. Interaction enthalpies between tRNA residues and contacting residues of L1, L5, or L16 proteins were calculated from the MD force field as the sum of electrostatic and Lennard-Jones interactions and averaged over the trajectory of each state (Fig. 1c).

In order to show the distance dependence of the L1-tRNA$^{\mathrm{fMet}}$ interaction enthalpy, the interaction enthalpy and the L1-tRNA$^{\mathrm{fMet}}$ distance (see section 2.3) were extracted from each simulation at intervals of 1 ns. For this, only simulations of the pre3–post2b states were taken into account (Fig. 1f).

To monitor tRNA-mRNA base-pairing, the minimal distance between the atoms from each codon nucleotide on the mRNA and the corresponding anticodon residue on the tRNA was calculated from each simulation.

Contacts between tRNA nucleosides and residues of the ribosomal proteins as well as the rRNA were found to have different levels of state-specifity, e.g some contacts are only present in a single state and others are present in several states. To asses the level of state-specifity, a contact entropy was calculated as follows: $\sum_{i=1}^{13} -f_i \ln(f_i)$, where f_i is the normalized contact frequency in state i. With this definition, contacts that are more state-specific have a smaller contact entropy than less specific ones.

2.6 Conservation of contact residues in L1, L5, and L16

Protein sequences of L1, L5 and L16 proteins were retrieved from the UniProt database[37]. 6,029 individual sequences of L1; 6,125 sequences of L5; and 6,031 sequences of L16 protein were used. To reduce computational complexity, sequences that had more than 90% identity were combined to a single cluster, each cluster represented by a single characteristic sequence. 1,174 characteristic sequences for L1; 1,106 sequences for L5; and 859 sequences for L16 protein were analyzed. Individual sequences were manually curated in order to exclude incomplete sequences leaving 1,153 sequences of L1, 1,079 sequences of L5 and 852 sequences of L16 protein.

Multiple sequence alignments were performed using Muscle software[38]. To construct a phylogenetic tree, incomplete positions of multiple sequence alignments were eliminated using Gblocks[39] software. Phylogenetic trees were constructed based on maximum-likelihood with JTT[40] model using PhyML[41]. Rate4Site software with JTT model was used to calculate conservation[42]. Calculation was performed using empirical Bayesian approach, which was shown to be superior to the maximum-likelihood method for site-specific conservation scores[42]. Conservation score was calculated for each individual position of the complete multiple sequence alignment. *E. coli* sequences were used as a reference. The conservation scores calculated by Rate4Site were inverted such that values higher than zero indicate conservation degree which is higher than for the protein in general, whereas values lower than zero indicate less than average conservation.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

In order to calculate the contact score, frequencies of contacts between tRNA and protein in every substate were used. For each pair of tRNA-protein contacts, the maximum contact frequency over all substates was determined. For every protein residue, the contact score was calculated as a sum of frequencies from all of its contact frequencies of different interactions. Residues that had a contact score of > 0.8 were considered contacting.

2.7 L1-stalk interaction with the 30S subunit

To monitor the interaction between the L1-stalk and the 30S subunit, the interaction enthalpy between the L1-stalk rRNA and proteins S7 and S11 was calculated. To that aim, the sum of electrostatic and Lennard-Jones interactions were averaged over the trajectory of each state (Fig. 1c).

2.8 L1-tRNA$^{\mathrm{fMet}}$ potential of mean force

The potential of mean force (PMF) between the L1-stalk and the tRNA$^{\mathrm{fMet}}$ was calculated using the extended umbrella sampling simulations[43]. The motion of the system was restricted at selected positions along the vector describing the distance between L1-stalk and tRNA (see section 2.3). First, 20 equally spaced positions x_i ($i = 1, \cdots, 20$) were chosen between the minimum and maximum value of the projection onto the distance vector observed in the simulations (pre3–post2b). For each i, the structure of the L1-stalk and the tRNA whose projection onto the vector was closest to x_i was then extracted from the trajectories. Next, the obtained structures were solvated (as described in section 1.1) and subsequently energy minimized. An additional harmonic umbrella potential with a force constant of $k_U = 100$ kJ/mol/nm^2 and centered at x_i was applied to the C$^\alpha$ and P atoms of the tRNA and the L1-stalk to restrain the movement along the distance vector.

Next, the solvent was equilibrated for 5 ns using position restraints on tRNA and the L1-stalk heavy atoms with a force constant of $k = 1000$ kJ/mol/nm^2. Subsequently, the system was simulated for 20 ns with the umbrella potential, but without position restraints.

The distances extracted from the 20 20-ns simulations were used to construct the free energy landscape (Fig. 1f) using the weighted histogram analysis method (WHAM)[44]. The WHAM implementation g_wham[45] was used and the statistical errors were calculated by bootstrapping new trajectories based on the umbrella histograms.

2.9 Bridge B1b interaction enthalpy

The 50S part of intersubunit bridge B1b[14] was defined as the set of residues of protein L5 that are in contact (see section 2.5) with the 30S residues in at least one of the states. The 30S part was defined as the set of all the 30S residues that are in contact with L5 in at least one of the states. The interaction enthalpy between 50S and 30S part of the intersubunit bridge was extracted as in section 2.5 from the pre5b simulation at intervals of 2 ps.

2.10 Kinetic sequence of states

From all reaction sequences (i.e. one permutation of all states, {pre1a, pre1b, pre2a, ..., post3b}), we determined the one which best matches the observed set of transition rates as follows. We assume the best matching reaction sequence to be the one that yields the shortest overall half-time τ. This overall half-time is proportional to the sum of the half-times of the state transitions $\tau_{\mathrm{A}\to\mathrm{B}} \propto \exp\left(\Delta G^{\ddagger}_{\mathrm{A}\to\mathrm{B}}\right)$. The highest transition barrier estimates dominate the overall half-time, so only the barrier for the ribosomal component governing the transition for a given pair of states was taken into account. The sequence with the shortest overall half-time of states minimizes $\sum_i^{\mathrm{transitions}} \exp\left(\Delta G_i^{\ddagger\mathrm{max}}\right)$, where $\Delta G_i^{\ddagger\mathrm{max}}$ is the highest barrier of all the barriers estimated for the individual ribosome components for transition i. To avoid that only very high

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

barrier estimates with large errors dominate the calculation, barrier estimates higher than 50 kT were set to 50 kT.

To check whether the omission of possible "off-track" states allows for a faster reaction sequence, we performed the above analysis for the fastest progression sequences with single states omitted. Two cases have been excluded here, which trivially enhance the progression rate. First, if end-states are omitted, the reaction sequence trivially becomes faster (e.g. choosing the simple sequence "pre3 pre4"). Second, if two barriers that include at least one capped barrier are replaced by just one capped barrier (e.g. replacing "pre2 pre3 pre4" by "pre2 pre4"). Excluding the two above trivial accelerations of the overall rate, we calculated the fastest progression rate for all left non-trivial combinations of omitting states from the fastest progression sequences.

2.11 Kinetic sequence of states for subsets of ribosome component movements

To assess which movements of individual ribosome components dominate the kinetic sequence of tRNA translocation, we calculated the fastest progression sequence using the movements of all ribosome components, of individual ribosome components, or of a combination of them. To that aim, only the maximum free energy barriers from the respective subsets of ribosome components were used for calculating the fastest progression sequences. To quantify the similarity of these sequences to the sequence introduced by Fischer et al.[1], the Kendall rank correlation coefficient τ was used, which reflects the minimum number of swaps of neighbouring states that are required to yield the desired sequence. Because a fully reversed sequence with negative τ leaves the progression rate unchanged, the absolute value of τ was used. Further, to account for the cases where the sequence presented by Fischer et al.[1] contains multiple substates for single states that have no specific order assigned (pre1a, pre1b, etc.) the maximum absolute τ for comparison to any permutation of substates was used.

After translocation, the tRNAs occupy other positions than before translocation, while body, head and L1-stalk return to their initial positions. To quantify how the derived sequences are affected by this fact, we determined a second set of reaction sequences with the first position fixed to a pre1 state.

Supplementary Note 3 (Results)

3.1 Refinement of atomic models

The AP-, PE-, and E-model were refined against 13 cryo-EM maps using DireX[21]. The complete set of all-atom ribosome models for the different states is depicted in Supplementary Fig. 1.

3.2 Comparison to recent crystal structures

Supplementary Figure 2a compares tRNA positions in our models with tRNA positions in two crystal structures[2] after aligning the binding site regions. For the comparison the pre1a and pre4 models were used which have the lowest RMSDs to these crystal structures (Fig. 1b). Even though the tRNAs used in our simulations and the tRNAs from the crystal structure represent different tRNA-species, they adopt almost the same position, especially in the functionally important anticodon and CCA-tail regions.

3.3 Structural deviation during the simulations

For each simulation, the root mean square deviations (RMSD) relative to the respective starting structure and relative to the structure at 20 ns are shown in Supplementary Fig. 2b. In all simulations an RMSD of less than 8 Å was obtained, which is very low for a system of this size. Typical values obtained by other authors are ∼10 Å[46]. Note that the PE-model simulation started from a crystal structure and thus can serve as a benchmark for the quality of the EM-fitted structures[47]. Notably, the RMSD values for the fitted structures are similar or only slightly larger, thus underscoring the quality of these structures.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

3.4 Independent MD based refinement

The independent MD based refinement against the pre1b cryo-EM map resulted in an RMSD of 3.26 Å relative to the corresponding structure obtained using DireX. Further, an RMSD of 4.54 Å was observed between the MD refined structure and the above mentioned crystal structure with a ribosome including a P-site tRNA[2]. A comparable RMSD of 5.31 Å was found for the DireX refined pre1b structure (cf. Fig. 1b, solid green line).

The fact that the two structures obtained by the two refinement methods are more similar to each other than to the reference crystal structure suggests that the refinement quality is independent of the choice of an elastic network based or an MD based refinement method.

3.5 Stereochemical parameters of the models

To further assess the quality of our models, we have carried out a statistical analysis of various stereochemical quantities. Supplementary Table 2 shows the deviation of the distributions of our models from those obtained from the protein data bank (PDB). To fully incorporate the additional effect of the simulations, the energy-minimzed refined structures and the energy-minimized structures after 100 ns of free MD simulation were used. Stereochemical parameters were calculated using WHATCHECK[48] and then averaged over all models (pre1a–post4). The ribosomal environment provides different conditions for protein and RNA folds than for most of the protein and RNA structures in the PDB. For this reason, we also checked ribosome crystal structures[2;3]. Supplementary Table 2 shows that upon initial refinement and energy-minimization through the force-field, the stereochemical parameters of our models are comparable to these of recent crystal strucutures. The stereochemical parameters of the models after 100 ns of free MD simulation, and subsequent energy-minimization further approached the parameters expected from an extensive analysis of the pdb-database[48]. Overall our results suggest that the stereochemical quality of our models is similar to that of crystal structures of comparable complexes.

3.6 30S head and body rotation

The 30S body rotation pivot point was found to be close to the 16S RNA residues G242 and U562, and the 30S head rotation pivot point is close to the 16S RNA residues A923, U1194, and G1386.

3.7 Transition rates

The fit of the probability of barrier crossing $p_{\mathrm{A}\to\mathrm{B}}$ to the frequency of barrier crossings $p^{\mathrm{sim}}_{\mathrm{A}\to\mathrm{B}}$ obtained from the simulations (see Supplementary Fig. 3a), yielded an attempt rate of $A = (22.4\mathrm{ns})^{-1}$ with an error interval from $(15.95\mathrm{ns})^{-1}$ to $(30.72\mathrm{ns})^{-1}$ and a calibration factor $c = 0.601 \pm 0.069$ at $t = 100$ ns. The relative statistical uncertainty of the free energy barrier estimates is 57%. Together these errors affect the statistical uncertainty of the transition rates by

$$\frac{\Delta k}{k} = c\sqrt{\left(\frac{\Delta A}{A}\frac{1}{c\Delta G^{\ddagger}}\right)^2 + \left(\frac{\Delta\Delta G^{\ddagger}}{\Delta G^{\ddagger}}\right)^2 + \left(\frac{\Delta c}{c}\right)^2} \tag{7}$$

$$\frac{\Delta k}{k} = 0.601\sqrt{0.429\left(\frac{1}{c\Delta G^{\ddagger}}\right)^2 + 0.338}. \tag{8}$$

The relative statistical uncertainty of the transition rate estimates ranges from 52% for fast transitions to 35% for slow transitions. The attempt rates and error estimates for the individual ribosome components are shown in Supplementary Fig. 3b. Notably, the attempt rates for the individual ribosome components do not differ markedly from the overall attempt rate, which was determined by combining the data from all transitions of all ribosome components.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

3.8 tRNA-mRNA base-pairing

For $tRNA^{Val}$ the base-pairing to the anticodon is maintained in all the simulations with minimal distances around 0.19 nm (Supplementary Fig. 4), except for the wobble base pair cmo^5U35-U which shows larger minimal distances in states pre5b, post3a and post3b. For $tRNA^{fMet}$ the base-pairing is partially disrupted in the simulations of the pre2 and pre3 states and, as expected, when the tRNA is moving out of the E site (post1–post4).

3.9 Rapid angular rearrangement after tRNA removal

The rotation angle between the 50S and 30S subunits changes quickly after tRNA removal (Supplementary Fig. 5). Compared to the simulations with bound tRNA (red curves), consistently lower body rotation angles are observed for the two simulations of the tRNA depleted ribosome structure (green). The interaction enthalpies of intersubunit bridge B1b between residues of the 50S protein L5 (109–111, 114, 135, 141, 143, 145–146, 177) and 30S proteins S13 (2, 6, 8, 56-57, 60, 63–65, 69–70) and S19 (26) were calculated. The interaction enthalpies were found to be weaker for simulations without tRNAs, indicating that the presence of tRNAs leads to a stabilization of large rotation angles by bridge B1b in the pre5b state. Figure 2b,c show histograms of B1b interaction enthalpies and 30S body rotation angles extracted from the last 50 ns of the simulations.

3.10 Influence of refinement accuracy on our conclusions

To estimate if and to which extent our conclusions might be affected by possible structural inaccuracies resulting from our refinement of an X-ray structure against the 13 cryo-EM maps has on our conclusions, we repeated a simulation and our analyses using the pre3 structure with the $tRNA^{fMet}$ in an artificially perturbed conformation. This perturbed structure also obeys all stereochemical and energetic constraints set by the force field and is within the resolution limits set by the cryo-EM map of the pre3 state, with a tRNA conformation closer to that of the pre2 state. From a simulation starting from this perturbed structure, for the reaction coordinates (r.c.) shown Figs. 1c,e, one would expect the changes to mostly affect $tRNA^{fMet}$ motion, such that it is similar to the motion in the pre2 state. Indeed, the r.c. values for this motion are lower in the perturbed simulation, close to those observed in the pre2 state, and the $tRNA^{fMet}$-L1 interaction enthalpy is weakened (data not shown). The $tRNA^{Val}$-L16 interaction enthalpies are weaker in the pre3 state than in the neighbouring states. Any change in the $tRNA^{Val}$ conformation in this state is expected to result in a stronger interaction enthalpy, as is the case for the perturbed simulation. The changes observed, however, would not have changed the qualitative picture of the interaction enthalpies offered by Fig. 1c. Interestingly, the changes in body rotation angle observed in the perturbed pre3 state reflect the coupling between tRNA conformation and body rotation reported in the main text (compare Fig. 2). The influence of the perturbation on the other motions is small.

Upon replacement of the pre3 state by the perturbed pre3 state in the the transition rate estimation, the barrier heights for head rotation and $tRNA^{Val}$ motion did not change significantly. In particular, only transition barriers larger than $\sim$4 kT are affected corresponding to rates slower than $1/\mu$s, thus leaving Fig. 1d unchanged.

3.11 Markov-state like representation of states and transition barriers

Supplementary Fig. 6a shows a schematic representation of the translocation intermediate states as a Markov model.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

3.12 Kinetic sequence of states

The sequences of translocation intermediate states with the shortest overall half-time within the range of the stochastical uncertainties of the correction factor and barrier heights are {pre1a, pre1b, pre2, pre4, pre3, pre5b, pre5a, post2a, post2b, post1, post3a, post3b} and {pre1a, pre1b, pre2, pre3, pre4, pre5b, pre5a, post2a, post1, post2b, post3a, post3b}. They closely resemble the sequence derived by Fischer et al.[1] based on structural similarity.

Omitting states in the proposed sequences did not yield a faster overall half-time, with the exception of omitting the pre5b state. According to this analysis only pre5b is a possible "off-track" state, suggesting that all other states are kinetically relevant.

3.13 Kinetic sequence of states for subsets of ribosome component movements

As shown in 3.12, sequences based on structural similarity reflect kinetic sequences, i.e. the ones with the fastest progression rate for the whole ribosome. Since sequence of translocation intermediate states presented by Fischer et al.[1] is based on structural similarity of the tRNAs, the fastest progression sequence based only on the barriers heights impeding tRNA movements should also match this structure based sequence. To test this idea, we determined the sequences, including all 12 states with two tRNAs, with the fastest progression rate for all 31 possible subsets of ribosome components. Indeed, Supplementary Fig. 6b shows that all fastest progression sequences including only tRNA movements match the structure based sequence. Further, the inclusion of movements of other ribosome components leaves sequence of states essentially unchanged, suggesting that the movement of the tRNAs dominates the movements of the other ribosome components. As a negative control, we looked at the fastest progression sequences from body, head and L1-stalk movements without tRNA movements. Indeed, completely different sequences were obtained, whose τ values are comparable to those of randomly drawn sequences (Supplementary Fig. 6b). Some of these sequence contain multiple pre to post transtions.

Since the fastest progression sequences with the first position fixed to a pre1 state are almost as similar to the structure based sequence as the ones with an unfixed first state (data not shown), we conclude that the fact that the tRNAs occupy different positions before and after translocation, while body, head and L1-stalk return to their initial positions does not markedly influence the determined sequences.

3.14 Contacts between the tRNAs and the ribosome and conservation of involved protein residues

The frequencies of contacts between the two tRNAs and the ribosome were determined for all states from the respective trajectories as described above. The contacting residues of L1, L5, and L16 are generally found to be clustered, e.g., D51–R60, R122–K141 and R164–K167 in L1, I43–K47, S72–Y82 in L5, M1–R10, R44–R59 and K76–E90 in L16 protein (Fig. 3b). All these contacting fragments have a high conservation degree and appear more conserved than the protein on average.

To test whether contacting residues had in fact a higher degree of conservation than the rest of the protein, we used a one-sided permutation test[49]. Differences in means were calculated for the contacting and non-contacting surface residues. Then, the same difference was calculated for every possible permutation of contacting and non-contacting groups. The p-values were calculated as a proportion of sampled mean differences larger than the observed value. The significance level was set at 0.05 (or 5%). Indeed, tRNA-contacting residues have a higher conservation degree for L1 (p-value=0.0019), L5 (p-value=0.027) and L16 (p-value=1.05×10^{-5}) proteins, with an overall p-value=6.62×10^{-8}. The high degree of evolutionary conservation of the protein residues which were identified to contact tRNA during translocation provides an independent evidence for their potential functional importance.

Overall, the contact entropy of tRNA-rRNA contacts is larger than that of tRNA-protein contacts: 64 % of the tRNA-rRNA contacts, but only 38 % of the tRNA-protein contacts have a contact entropy above 0.5.

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

This means that there are more state-specific contacts between the tRNAs and the proteins, which is an interesting finding in itself.

Supplementary Tables 3a–3h list the frequencies of contacts between $tRNA^{Val}$ and the ribosome for each state. Supplementary Tables 3i–3t list the contacts between $tRNA^{fMet}$ and the ribosome. The gray-scale level of the cells indicates the frequency of atom-atom contacts corresponding to the residue pairs, white (0–12.5%), light gray (12.5–25%), dark gray (25–50%), and black (50–100%).

Supplementary References

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

[2] Dunkle, J. A. *et al.* Structures of the bacterial ribosome in classical and hybrid states of trna binding. *Science* **332**, 981 (2011).

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

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

[5] Hornak, V. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).

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

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

[8] Essmann, U. *et al.* A smooth particle mesh ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).

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

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

[11] Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).

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

[13] Gao, Y. *et al.* The structure of the ribosome with elongation factor g trapped in the posttranslocational state. *Science* **326**, 694–699 (2009).

[14] Yusupov, M. *et al.* Crystal structure of the ribosome at 5.5 angstrom resolution. *Science* **292**, 883–896 (2001).

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

[16] Theimer, C. A., Finger, L. D., Trantirek, L. & Feigon, J. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase rna. *P. Natl. Acad. Sci. USA* **100**, 449–454 (2003).

Nature Structural and Molecular Biology: doi:10.1038/nsmb.2690

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

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

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

[20] Pettersen, E. F. *et al.* Ucsf chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

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

[22] Ahmed, A. & Tama, F. Consensus among multiple approaches as a reliability measure for flexible fitting into cryo-em data. *Journal of Structural Biology* **182**, 67–77 (2013).

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

[24] Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamic-smethod. *J. Appl. Phys.* **52**, 7182 (1981).

[25] Bellare, J. R., Davis, H. T., Scriven, L. E. & Talmon, Y. Controlled environment vitrification system: an improved sample preparation technique. *J. Electron Micr. Tech.* **10**, 87–111 (1988).

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

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

[28] Konevega, A. L. *et al.* Spontaneous reverse movement of mrna-bound trna through the ribosome. *Nat. Struct. Mol. Biol.* **14**, 318–324 (2007).

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

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

[31] Orzechowski, M. & Tama, F. Flexible fitting of high-resolution x-ray structures into cryoelectron microscopy maps using biased molecular dynamics simulations. *Biophysical Journal* **95**, 5692–5705 (2008).

[32] Ratje, A. H. *et al.* Head swivel on the ribosome facilitates translocation by means of intra-subunit trna hybrid sites. *Nature* **468**, 713–716 (2010).

[33] Amadei, A., Linssen, A. B. & Berendsen, H. J. C. Essential dynamics of proteins. *Proteins* **17**, 412–425 (1993).

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

[35] Gramkow, C. On averaging rotations. *Int. J. Comput. Vision* **42**, 7–16 (2001).

[36] Blau, C. & Grubmuller, H. g_contacts: Fast contact search in bio-molecular ensemble data. *Computer Physics Communications* – (2013). Doi:10.1016/j.cpc.2013.07.018.

[37] Apweiler, R. *et al.* Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res.* **39**, D214–219 (2011).

[38] Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).

[39] Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).

[40] Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 (1992).

[41] Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).

[42] Mayrose, I., Graur, D., Ben-Tal, N. & Pupko, T. Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. *Mol. Biol. Evol.* **21**, 1781–1791 (2004).

[43] Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **91**, 275–282 (1995).

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

[45] Hub, J. S., de Groot, B. L. & van der Spoel, D. g_wham - a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010).

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

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

[48] Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. Errors in protein structures. *Nature* **381**, 272–272 (1996).

[49] Röhmel, J. Precision intervals for estimates of the difference in success rates for binary random variables based on the permutation principle. *Biometrical journal* **38**, 977–993 (2007).