wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:51 PM BST

PDB ID	:	3J46
EMDB ID:	:	EMD-5693
Title	:	Structure of the SecY protein translocation channel in action
Authors	:	Akey, C.W.; Park, E.; Menetret, J.F.; Gumbart, J.C.; Ludtke, S.J.; Li, W.; Whynot, A.; Rapoport, T.A.
Deposited on	:	2013-06-18
Resolution	:	10.10 Å(reported)
Based on PDB ID	:	2I2P, 3I8G, 3J01

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

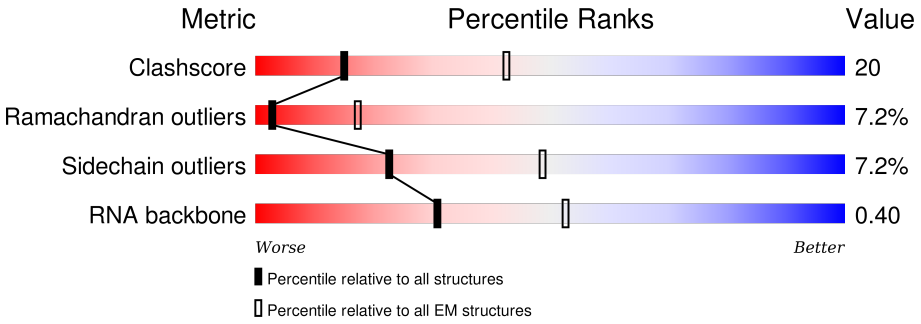
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 10.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



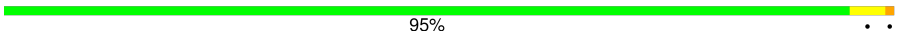
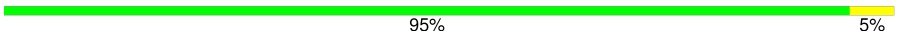
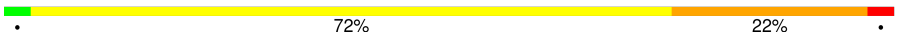
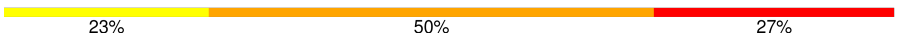
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	y	437	73% 20% 5% •
2	E	56	46% 48% 5%
3	G	67	58% 31% 7% •
4	n	101	59% 30% 6% 5%
5	p	76	45% 50% 5%
6	a	76	62% 34% •
7	5	234	75% 22% •
8	T	100	88% 10% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	U	103	
10	Y	63	
11	1	63	
12	2	36	
13	3	44	
14	4	109	

2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 17478 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Protein translocase subunit SecY.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	y	437	Total	C	N	O	S	0	1
			3361	2220	554	569	18		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
y	5	ACE	-	ACETYLATION	UNP P0AGA2
y	68	CYS	SER	ENGINEERED MUTATION	UNP P0AGA2
y	441	NH2	-	AMIDATION	UNP P0AGA2

- Molecule 2 is a protein called Preprotein translocase subunit SecE.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	E	56	Total	C	N	O	S	0	1
			433	283	76	73	1		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	73	ACE	-	ACETYLATION	UNP P0AG96
E	128	NH2	-	AMIDATION	UNP P0AG96

- Molecule 3 is a protein called Protein-export membrane protein SecG.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	G	67	Total	C	N	O	S	0	1
			461	301	74	82	4		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	8	ACE	-	ACETYLATION	UNP P0AG99
G	74	NH2	-	AMIDATION	UNP P0AG99

- Molecule 4 is a protein called NC100.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	n	101	Total	C	N	O	S	0	0
			760	480	132	146	2		

- Molecule 5 is a RNA chain called P-tRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	p	76	Total	C	N	O	P	0	0
			1621	722	287	536	76		

- Molecule 6 is a RNA chain called A-tRNA.

Mol	Chain	Residues	Atoms						AltConf	Trace
6	a	76	Total	C	N	O	P	S	0	0
			1626	729	290	531	75	1		

- Molecule 7 is a protein called 50S ribosomal protein L1.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	5	234	Total	C	N	O	S	0	0
			1733	1081	315	330	7		

- Molecule 8 is a protein called 50S ribosomal protein L23P.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	T	100	Total	C	N	O	S	0	0
			787	496	146	143	2		

- Molecule 9 is a protein called 50S ribosomal protein L24P.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	U	103	Total	C	N	O	0	0
			789	498	148	143		

- Molecule 10 is a protein called 50S ribosomal protein L29P.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	Y	63	Total	C	N	O	S	0	0
			509	313	99	95	2		

- Molecule 11 is a RNA chain called 23S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	1	63	Total	C	N	O	P	0	0
			1350	603	245	439	63		

- Molecule 12 is a RNA chain called 23S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	2	36	Total	C	N	O	P	0	0
			775	345	142	252	36		

- Molecule 13 is a RNA chain called 23S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	3	44	Total	C	N	O	P	0	0
			948	423	180	301	44		

- Molecule 14 is a RNA chain called 23S ribosomal RNA.

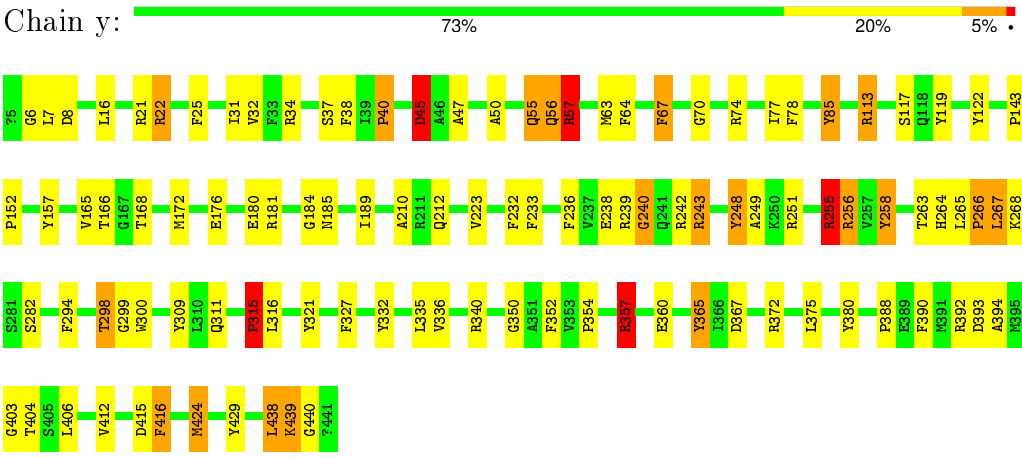
Mol	Chain	Residues	Atoms					AltConf	Trace
14	4	109	Total	C	N	O	P	0	0
			2325	1038	409	769	109		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

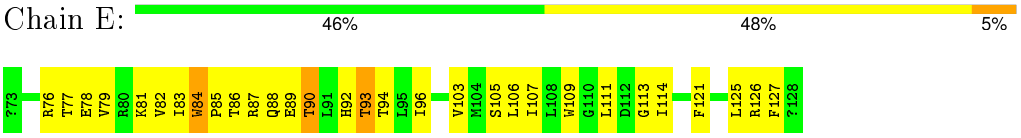
- Molecule 1: Protein translocase subunit SecY

Chain y:



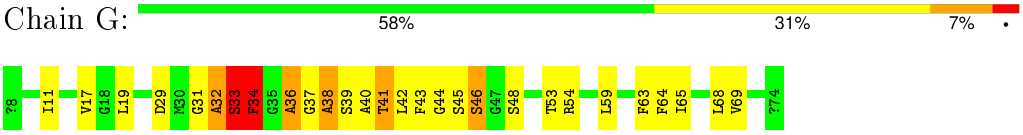
- Molecule 2: Preprotein translocase subunit SecE

Chain E:



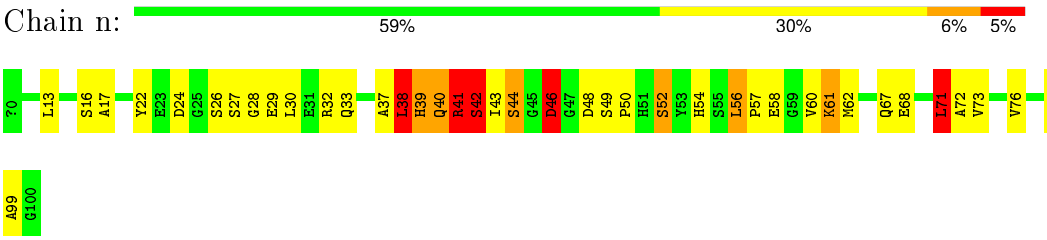
- Molecule 3: Protein-export membrane protein SecG

Chain G:



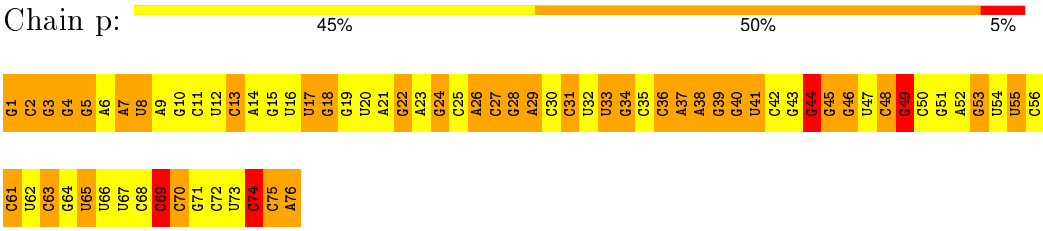
- Molecule 4: NC100

Chain n:



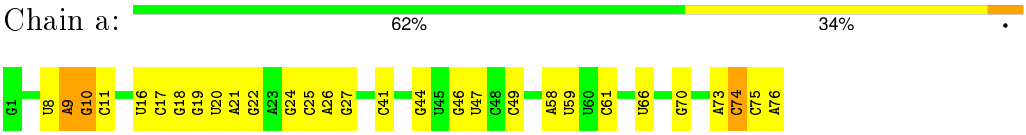
- Molecule 5: P-tRNA

Chain p:



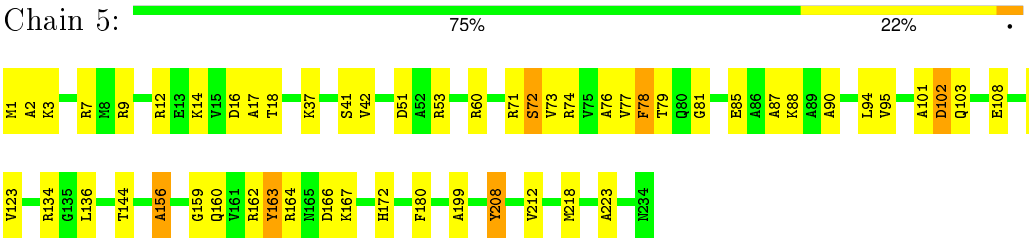
- Molecule 6: A-tRNA

Chain a:



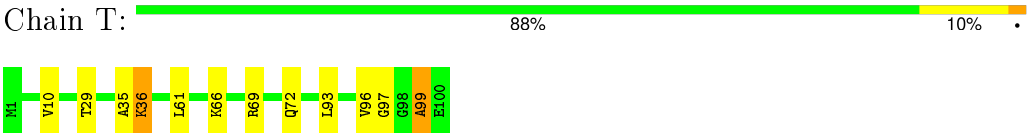
- Molecule 7: 50S ribosomal protein L1

Chain 5:



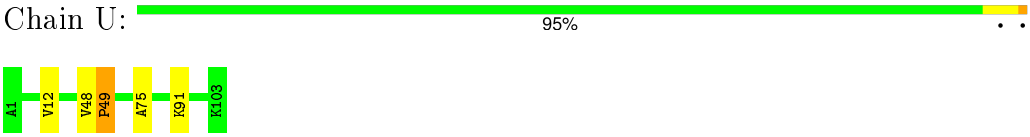
- Molecule 8: 50S ribosomal protein L23P

Chain T:



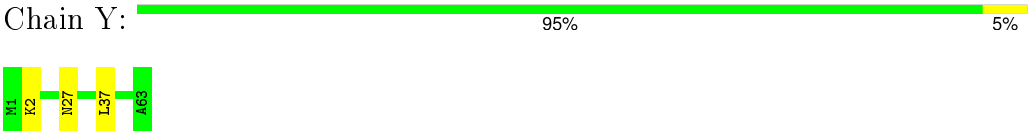
- Molecule 9: 50S ribosomal protein L24P

Chain U:



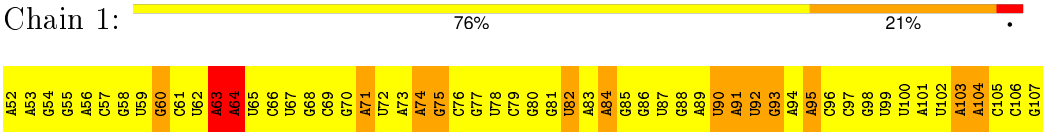
- Molecule 10: 50S ribosomal protein L29P

Chain Y:



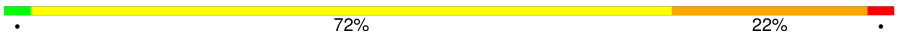
- Molecule 11: 23S ribosomal RNA

Chain 1:



• Molecule 12: 23S ribosomal RNA

Chain 2:

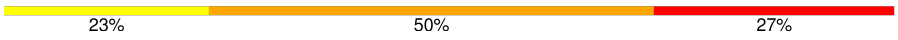


• Molecule 13: 23S ribosomal RNA

Chain 3:

• Molecule 14: 23S ribosomal RNA

Chain 4:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	53000	Depositor
Resolution determination method	FSC at 0.5 cut-off for a comparison between the full experimental 3D density map and a calculated map of the docked E. coli ribosome model (this map was calculated to 7 Angstrom resolution with EMAN).	Depositor
CTF correction method	per micrograph	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	160	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	42000	Depositor
Image detector	Not provided	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MIA, ACE, NH2

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	y	1.71	27/3434 (0.8%)	2.07	97/4657 (2.1%)
10	Y	1.00	0/510	0.90	0/677
11	1	1.62	1/1511 (0.1%)	2.50	178/2354 (7.6%)
12	2	1.63	1/867 (0.1%)	2.45	97/1351 (7.2%)
13	3	3.35	127/1062 (12.0%)	3.85	256/1655 (15.5%)
14	4	3.46	373/2599 (14.4%)	3.88	636/4049 (15.7%)
2	E	1.70	3/437 (0.7%)	2.26	24/596 (4.0%)
3	G	1.65	3/463 (0.6%)	2.12	17/622 (2.7%)
4	n	1.63	2/774 (0.3%)	1.60	14/1048 (1.3%)
5	p	3.30	222/1810 (12.3%)	3.82	397/2820 (14.1%)
6	a	0.35	0/1783	0.77	4/2776 (0.1%)
7	5	1.66	11/1748 (0.6%)	1.92	30/2355 (1.3%)
8	T	0.98	0/794	1.09	1/1060 (0.1%)
9	U	0.96	0/797	1.04	0/1062
All	All	2.21	770/18589 (4.1%)	2.66	1751/27082 (6.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	y	2	18
11	1	0	2
12	2	0	2
13	3	0	19
14	4	0	63
2	E	0	2
3	G	0	1
4	n	3	0
5	p	0	29

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
6	a	0	1
7	5	0	3
9	U	0	1
All	All	5	141

The worst 5 of 770 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	n	99	ALA	C-N	31.42	1.89	1.33
13	3	1516	G	O3'-P	25.56	1.91	1.61
4	n	22	TYR	C-N	22.00	1.84	1.34
13	3	1551	A	O3'-P	17.96	1.82	1.61
14	4	2098	U	C2-N3	14.56	1.48	1.37

The worst 5 of 1751 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	p	19	G	N1-C6-O6	24.02	134.31	119.90
14	4	2108	A	N1-C6-N6	22.36	132.01	118.60
5	p	18	G	N1-C6-O6	20.82	132.39	119.90
5	p	37	A	N1-C6-N6	20.71	131.02	118.60
13	3	1535	A	N1-C6-N6	20.48	130.89	118.60

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	y	55	GLN	CA
1	y	56	GLN	CA
4	n	38	LEU	CA
4	n	39	HIS	CA
4	n	40	GLN	CA

5 of 141 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	y	22	ARG	Sidechain
1	y	34	ARG	Sidechain
1	y	38	PHE	Sidechain
1	y	45	ASP	Sidechain
1	y	57	ARG	Sidechain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	y	3361	0	3514	0	0
2	E	433	0	466	22	0
3	G	461	0	485	7	0
4	n	760	0	742	0	0
5	p	1621	0	818	0	0
6	a	1626	0	833	0	0
7	5	1733	0	1824	8	0
8	T	787	0	844	20	0
9	U	789	0	847	0	0
10	Y	509	0	543	0	0
11	1	1350	0	678	2	0
12	2	775	0	389	0	0
13	3	948	0	478	0	0
14	4	2325	0	1168	8	0
All	All	17478	0	13629	45	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 20.

The worst 5 of 45 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:81:LYS:HD2	8:T:93:LEU:CD1	1.30	1.62
2:E:81:LYS:CD	8:T:93:LEU:HD12	1.04	1.51
2:E:81:LYS:CD	8:T:93:LEU:CD1	1.94	1.21
2:E:81:LYS:HZ3	8:T:93:LEU:CG	1.57	1.17
2:E:81:LYS:CE	8:T:93:LEU:HD12	1.77	1.13

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	y	435/437 (100%)	392 (90%)	18 (4%)	25 (6%)	2 28
2	E	54/56 (96%)	49 (91%)	4 (7%)	1 (2%)	10 52
3	G	65/67 (97%)	57 (88%)	4 (6%)	4 (6%)	2 26
4	n	99/101 (98%)	42 (42%)	26 (26%)	31 (31%)	0 0
7	5	232/234 (99%)	198 (85%)	25 (11%)	9 (4%)	4 36
8	T	98/100 (98%)	71 (72%)	20 (20%)	7 (7%)	1 22
9	U	101/103 (98%)	84 (83%)	14 (14%)	3 (3%)	5 42
10	Y	61/63 (97%)	48 (79%)	11 (18%)	2 (3%)	5 40
All	All	1145/1161 (99%)	941 (82%)	122 (11%)	82 (7%)	3 22

5 of 82 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	y	45	ASP
1	y	56	GLN
1	y	258	TYR
1	y	266	PRO
1	y	298	THR

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	y	353/353 (100%)	330 (94%)	23 (6%)	21 58

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	E	47/47 (100%)	45 (96%)	2 (4%)	35 70
3	G	46/46 (100%)	39 (85%)	7 (15%)	3 22
4	n	77/77 (100%)	61 (79%)	16 (21%)	1 10
7	5	181/181 (100%)	170 (94%)	11 (6%)	23 60
8	T	84/84 (100%)	79 (94%)	5 (6%)	24 60
9	U	84/84 (100%)	82 (98%)	2 (2%)	57 82
10	Y	55/55 (100%)	54 (98%)	1 (2%)	66 87
All	All	927/927 (100%)	860 (93%)	67 (7%)	23 55

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	59	LEU
4	n	44	SER
8	T	66	LYS
3	G	65	ILE
4	n	38	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
4	n	39	HIS
4	n	40	GLN
7	5	168	ASN
4	n	33	GLN
7	5	165	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
11	1	62/63 (98%)	13 (20%)	1 (1%)
12	2	35/36 (97%)	8 (22%)	1 (2%)
13	3	43/44 (97%)	7 (16%)	2 (4%)
14	4	108/109 (99%)	39 (36%)	8 (7%)
5	p	75/76 (98%)	17 (22%)	0
6	a	74/76 (97%)	28 (37%)	0
All	All	397/404 (98%)	112 (28%)	12 (3%)

5 of 112 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	p	2	C
5	p	4	G
5	p	5	G
5	p	7	A
5	p	8	U

5 of 12 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
14	4	2144	G
14	4	2145	C
14	4	2157	G
14	4	2132	U
14	4	2152	G

5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	MIA	a	37	6	22,31,32	1.41	2 (9%)	26,44,47	3.46	6 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	MIA	a	37	6	-	0/11/33/34	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	a	37	MIA	C13-C12	-3.99	1.35	1.52
6	a	37	MIA	C2-S10	4.17	1.79	1.75

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	a	37	MIA	C12-N6-C6	-12.12	109.44	123.46
6	a	37	MIA	C5-C6-N1	-2.80	117.74	120.58
6	a	37	MIA	C16-C14-C15	2.01	120.33	110.52
6	a	37	MIA	C2-N1-C6	2.18	119.11	113.13
6	a	37	MIA	C1'-N9-C4	2.75	129.88	126.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.