wwPDB X-ray Structure Validation Summary Report (i)

Sep 6, 2023 – 08:57 AM EDT

PDB ID	:	4EBA
Title	:	Crystal structure of the Rna14-Rna15 complex
Authors	:	Paulson, A.R.; Tong, L.
Deposited on	:	2012-03-23
Resolution	:	3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

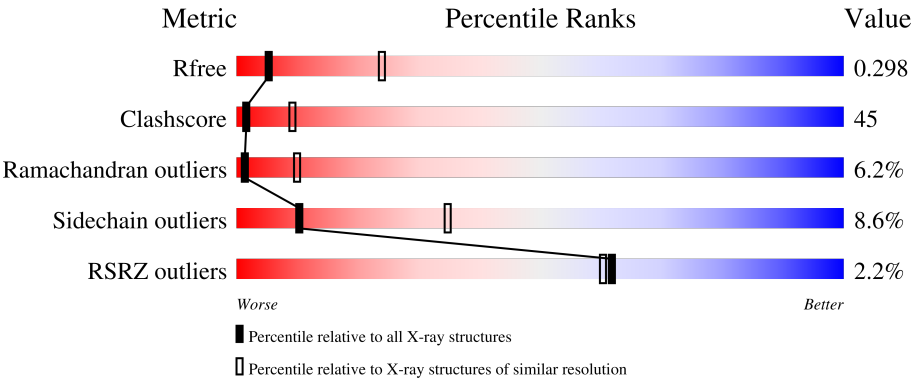
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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.



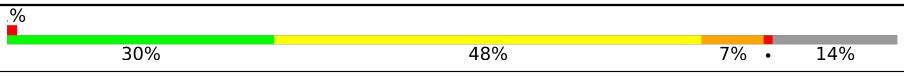
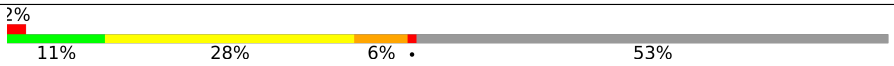
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	645	27% 53% 9% • 9%
1	B	645	29% 49% 8% • 13%
1	C	645	2% 35% 49% 6% • 9%
1	D	645	% 32% 47% 7% 15%
1	E	645	7% 27% 52% 10% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	645	
2	G	174	
2	H	174	
2	I	174	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 30268 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 mRNA 3'-end-processing protein RNA14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	585	Total	C	N	O	S	0	0	0
			4874	3149	803	897	25			
1	B	558	Total	C	N	O	S	0	0	0
			4646	2990	772	859	25			
1	C	585	Total	C	N	O	S	0	0	0
			4874	3149	803	897	25			
1	D	551	Total	C	N	O	S	0	0	0
			4589	2957	759	848	25			
1	E	578	Total	C	N	O	S	0	0	0
			4823	3118	795	885	25			
1	F	553	Total	C	N	O	S	0	0	0
			4608	2967	764	852	25			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	MET	-	expression tag	UNP Q6CII8
B	17	MET	-	expression tag	UNP Q6CII8
C	17	MET	-	expression tag	UNP Q6CII8
D	17	MET	-	expression tag	UNP Q6CII8
E	17	MET	-	expression tag	UNP Q6CII8
F	17	MET	-	expression tag	UNP Q6CII8

- Molecule 2 is a protein called Rna15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	81	Total	C	N	O	S	0	0	0
			618	395	98	119	6			
2	I	81	Total	C	N	O	S	0	0	0
			618	395	98	119	6			
2	H	81	Total	C	N	O	S	0	0	0
			618	395	98	119	6			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	85	MET	-	expression tag	UNP Q6CKN2
G	86	GLY	-	expression tag	UNP Q6CKN2
G	87	SER	-	expression tag	UNP Q6CKN2
G	88	SER	-	expression tag	UNP Q6CKN2
G	89	HIS	-	expression tag	UNP Q6CKN2
G	90	HIS	-	expression tag	UNP Q6CKN2
G	91	HIS	-	expression tag	UNP Q6CKN2
G	92	HIS	-	expression tag	UNP Q6CKN2
G	93	HIS	-	expression tag	UNP Q6CKN2
G	94	HIS	-	expression tag	UNP Q6CKN2
G	95	SER	-	expression tag	UNP Q6CKN2
G	96	SER	-	expression tag	UNP Q6CKN2
G	97	GLY	-	expression tag	UNP Q6CKN2
G	98	LEU	-	expression tag	UNP Q6CKN2
G	99	VAL	-	expression tag	UNP Q6CKN2
G	100	PRO	-	expression tag	UNP Q6CKN2
G	101	ARG	-	expression tag	UNP Q6CKN2
G	102	GLY	-	expression tag	UNP Q6CKN2
G	103	SER	-	expression tag	UNP Q6CKN2
G	104	HIS	-	expression tag	UNP Q6CKN2
I	85	MET	-	expression tag	UNP Q6CKN2
I	86	GLY	-	expression tag	UNP Q6CKN2
I	87	SER	-	expression tag	UNP Q6CKN2
I	88	SER	-	expression tag	UNP Q6CKN2
I	89	HIS	-	expression tag	UNP Q6CKN2
I	90	HIS	-	expression tag	UNP Q6CKN2
I	91	HIS	-	expression tag	UNP Q6CKN2
I	92	HIS	-	expression tag	UNP Q6CKN2
I	93	HIS	-	expression tag	UNP Q6CKN2
I	94	HIS	-	expression tag	UNP Q6CKN2
I	95	SER	-	expression tag	UNP Q6CKN2
I	96	SER	-	expression tag	UNP Q6CKN2
I	97	GLY	-	expression tag	UNP Q6CKN2
I	98	LEU	-	expression tag	UNP Q6CKN2
I	99	VAL	-	expression tag	UNP Q6CKN2
I	100	PRO	-	expression tag	UNP Q6CKN2
I	101	ARG	-	expression tag	UNP Q6CKN2
I	102	GLY	-	expression tag	UNP Q6CKN2
I	103	SER	-	expression tag	UNP Q6CKN2
I	104	HIS	-	expression tag	UNP Q6CKN2
H	85	MET	-	expression tag	UNP Q6CKN2
H	86	GLY	-	expression tag	UNP Q6CKN2

Continued on next page...

Continued from previous page...

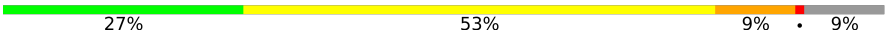
Chain	Residue	Modelled	Actual	Comment	Reference
H	87	SER	-	expression tag	UNP Q6CKN2
H	88	SER	-	expression tag	UNP Q6CKN2
H	89	HIS	-	expression tag	UNP Q6CKN2
H	90	HIS	-	expression tag	UNP Q6CKN2
H	91	HIS	-	expression tag	UNP Q6CKN2
H	92	HIS	-	expression tag	UNP Q6CKN2
H	93	HIS	-	expression tag	UNP Q6CKN2
H	94	HIS	-	expression tag	UNP Q6CKN2
H	95	SER	-	expression tag	UNP Q6CKN2
H	96	SER	-	expression tag	UNP Q6CKN2
H	97	GLY	-	expression tag	UNP Q6CKN2
H	98	LEU	-	expression tag	UNP Q6CKN2
H	99	VAL	-	expression tag	UNP Q6CKN2
H	100	PRO	-	expression tag	UNP Q6CKN2
H	101	ARG	-	expression tag	UNP Q6CKN2
H	102	GLY	-	expression tag	UNP Q6CKN2
H	103	SER	-	expression tag	UNP Q6CKN2
H	104	HIS	-	expression tag	UNP Q6CKN2

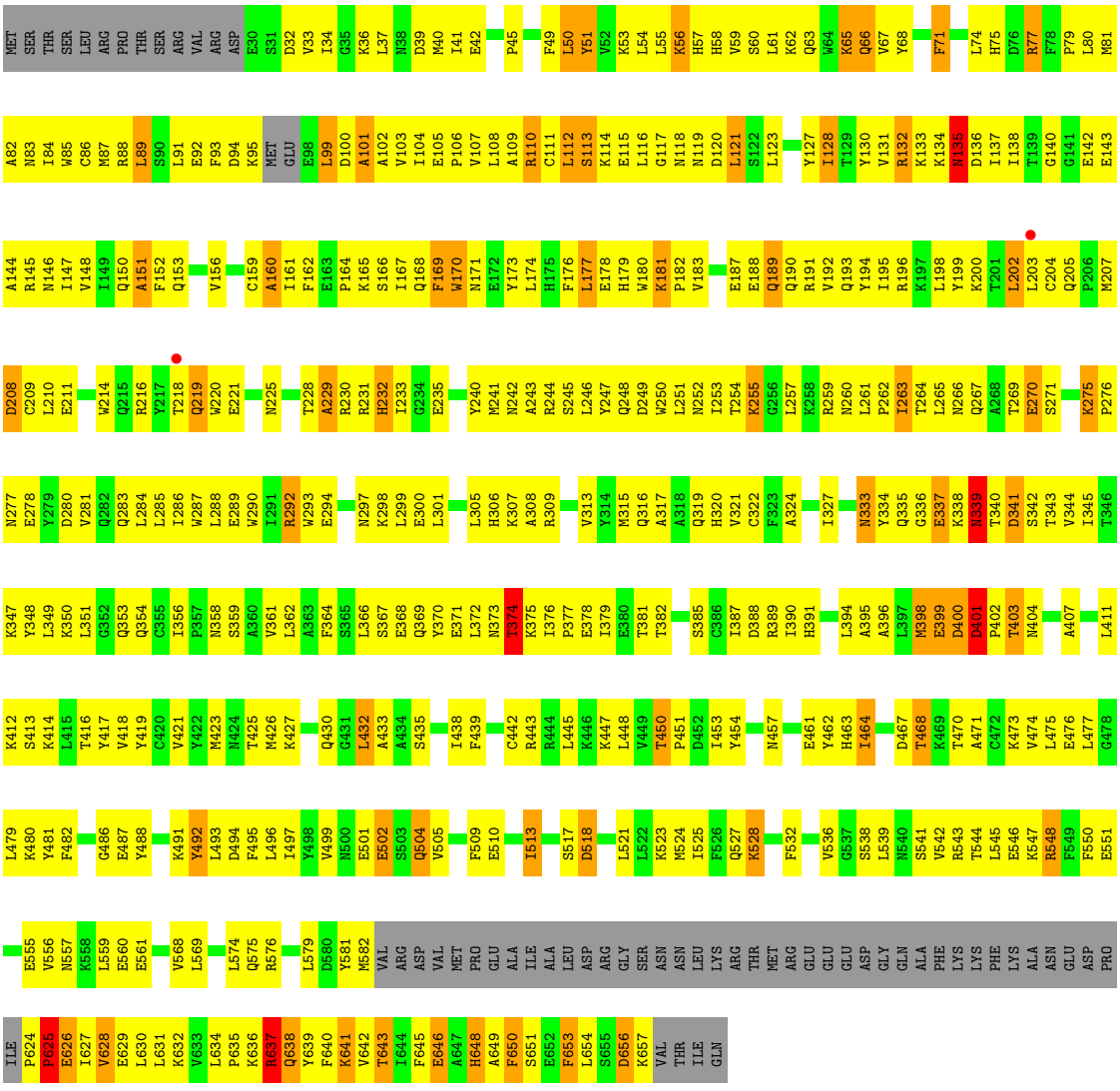
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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: mRNA 3'-end-processing protein RNA14

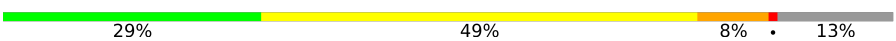
Chain A:

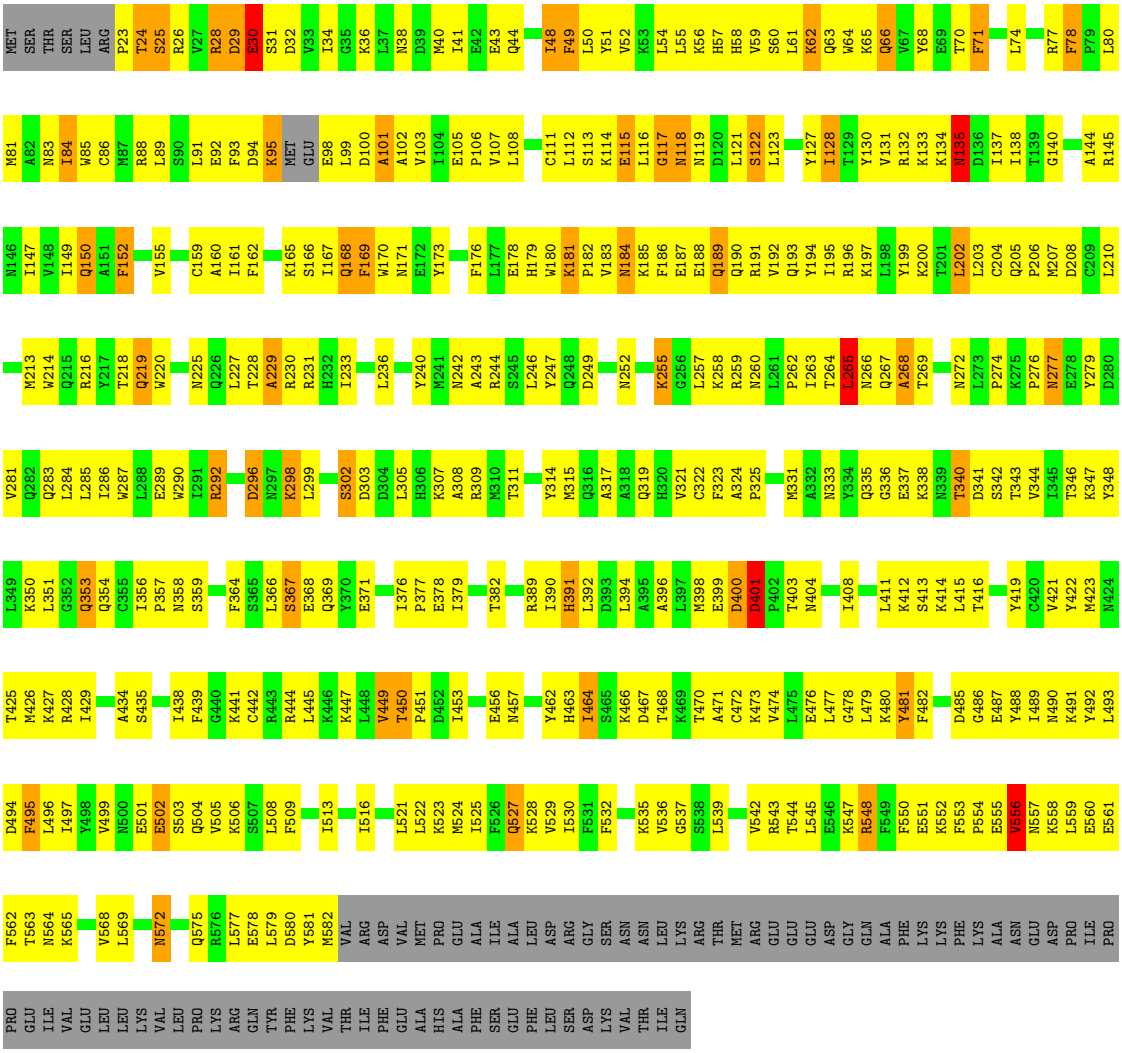




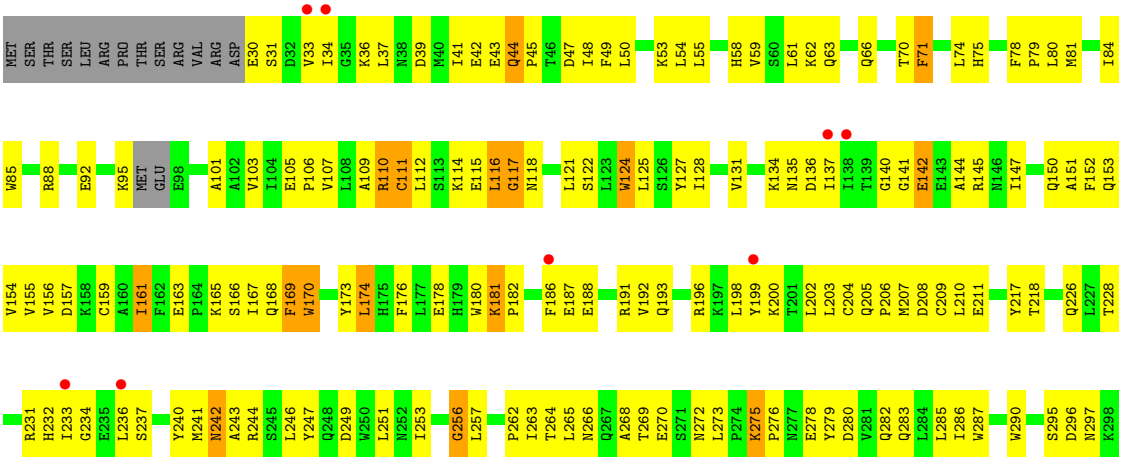
• Molecule 1: mRNA 3'-end-processing protein RNA14

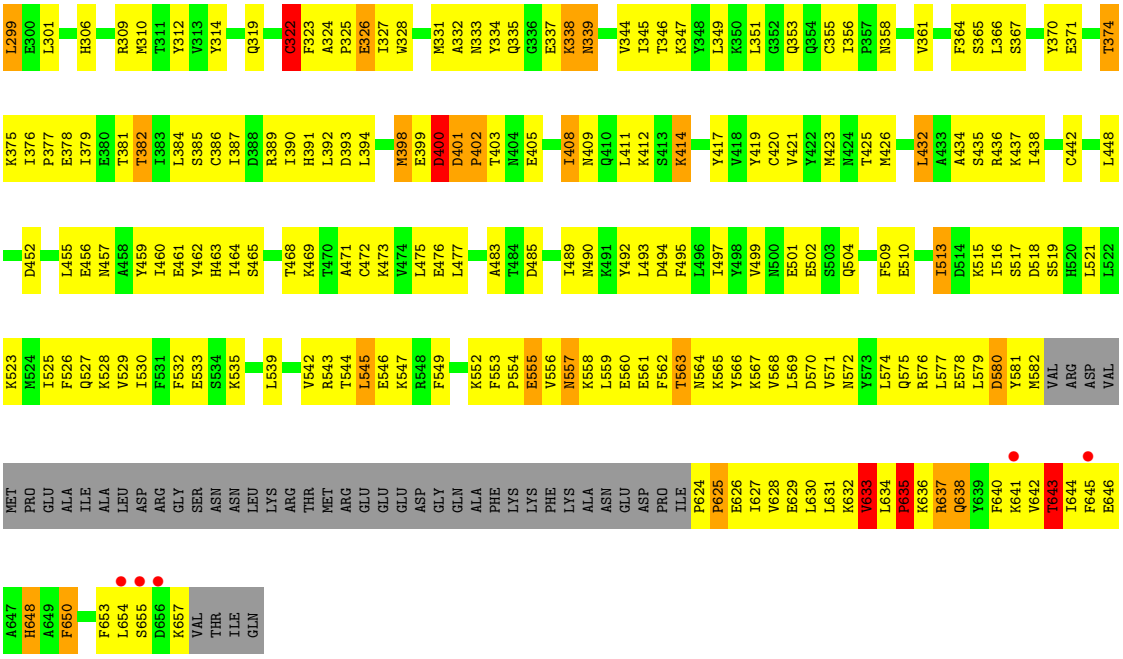
Chain B:





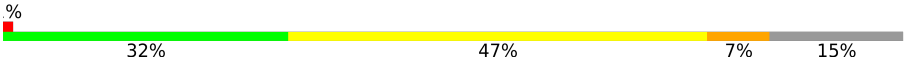
• Molecule 1: mRNA 3'-end-processing protein RNA14

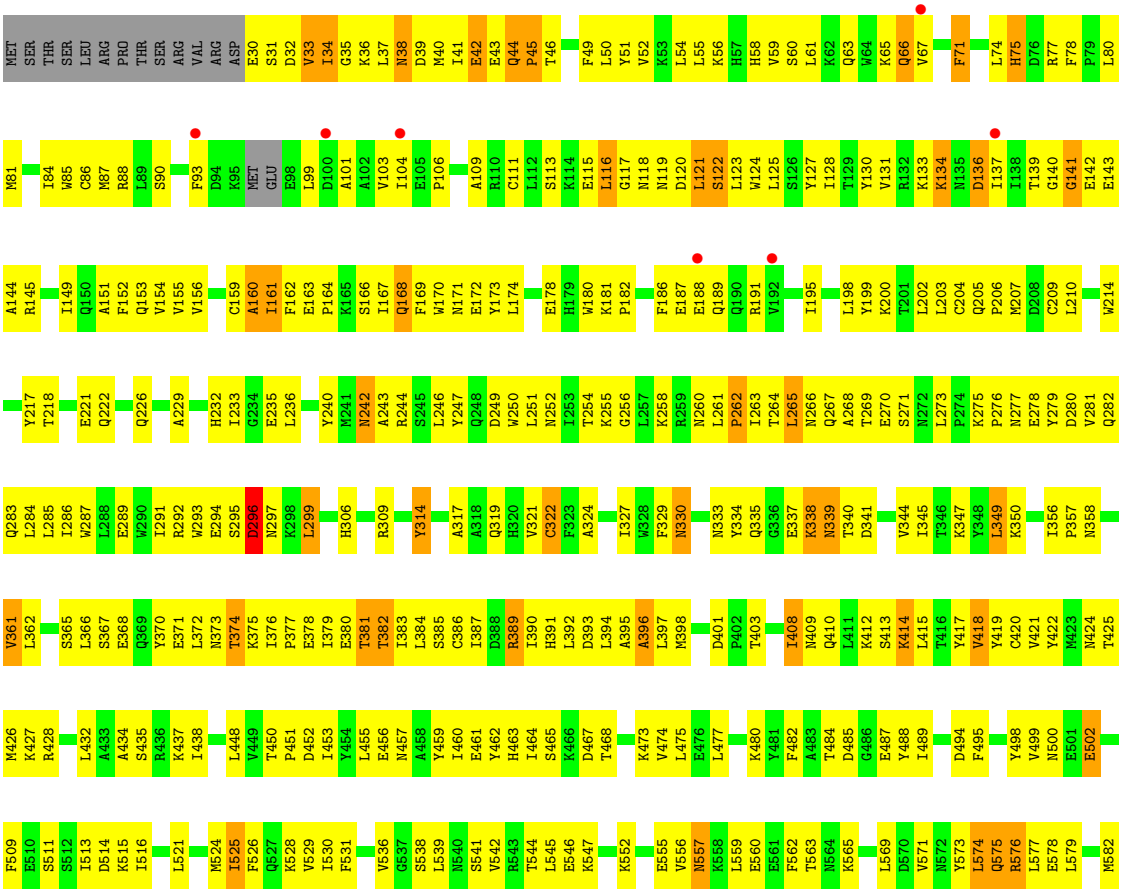


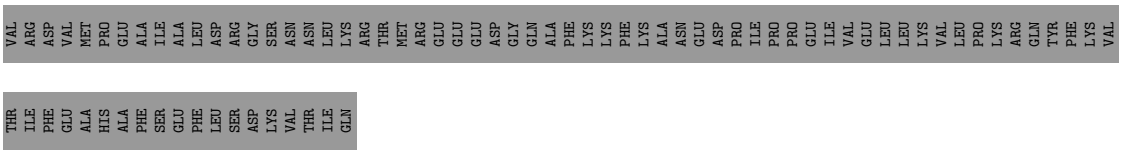


• Molecule 1: mRNA 3'-end-processing protein RNA14

Chain D:

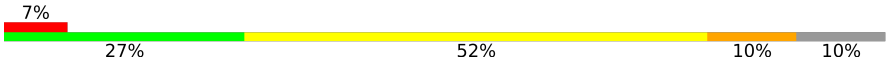


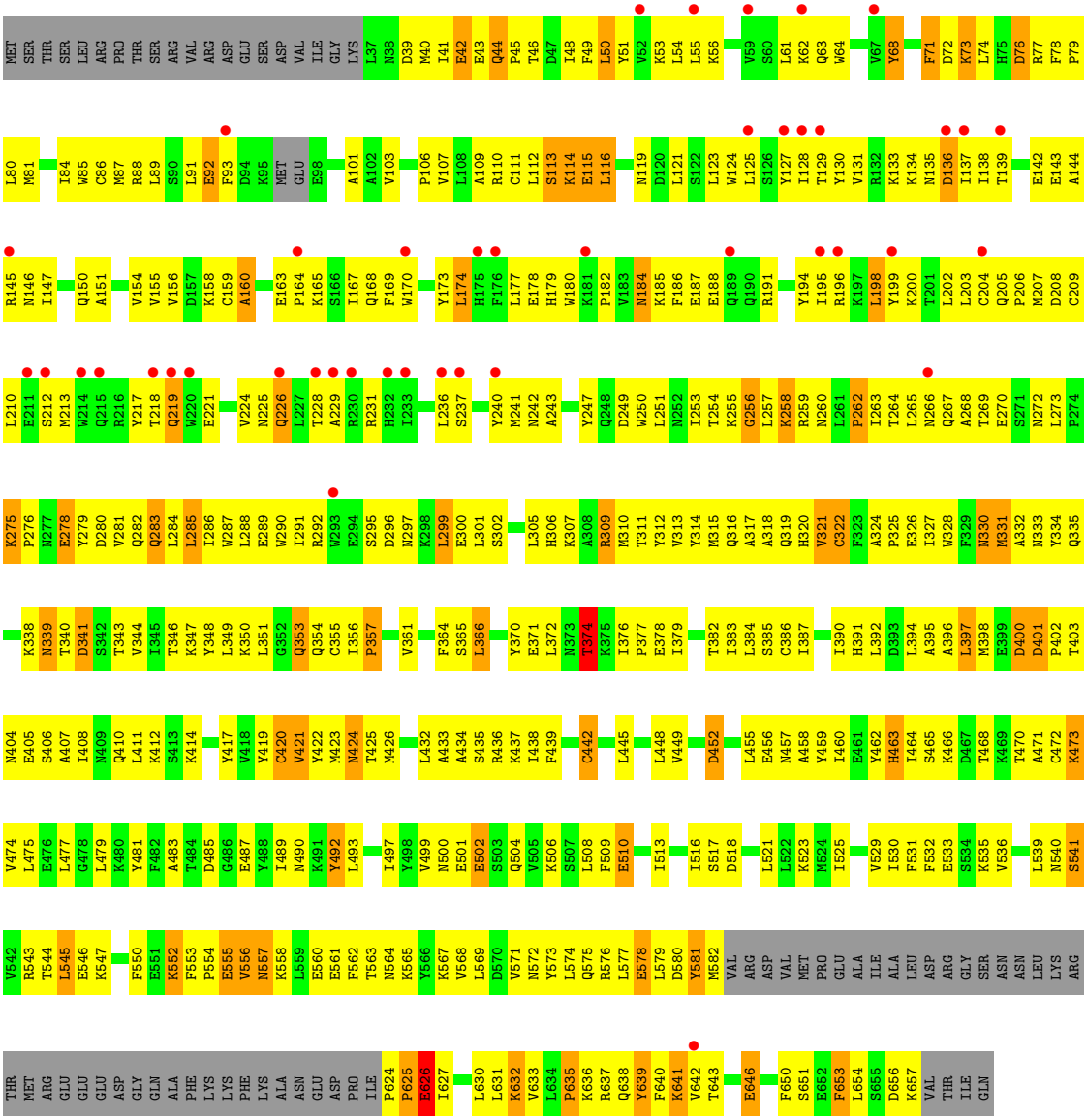




• Molecule 1: mRNA 3'-end-processing protein RNA14

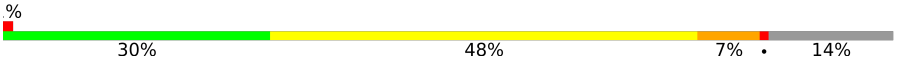
Chain E:

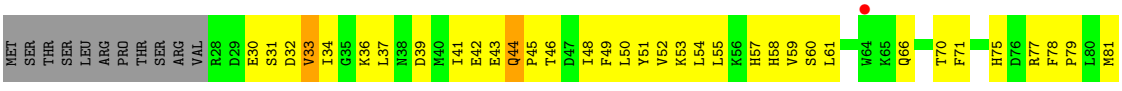


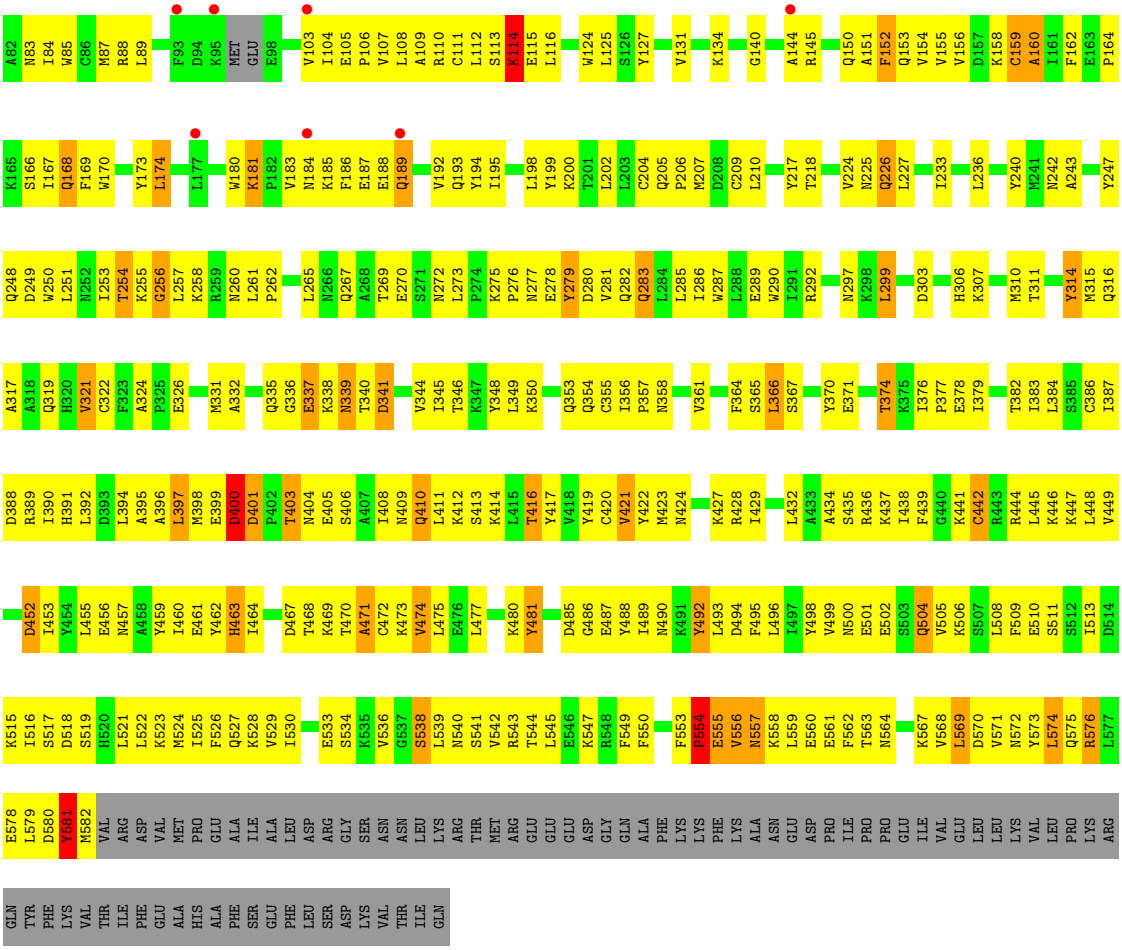


• Molecule 1: mRNA 3'-end-processing protein RNA14

Chain F:

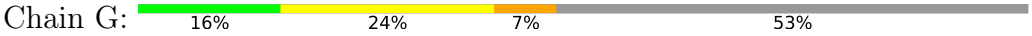


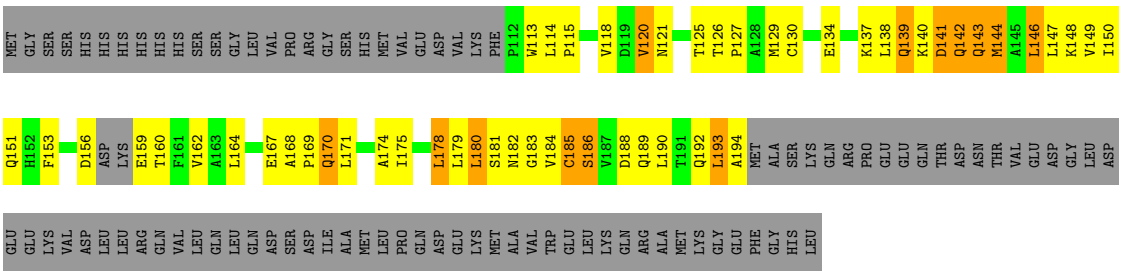




• Molecule 2: Rna15

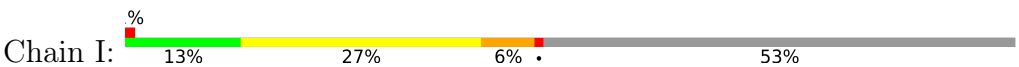
Chain G:

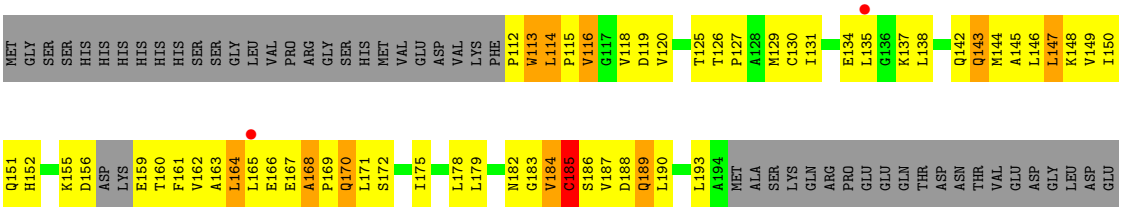




• Molecule 2: Rna15

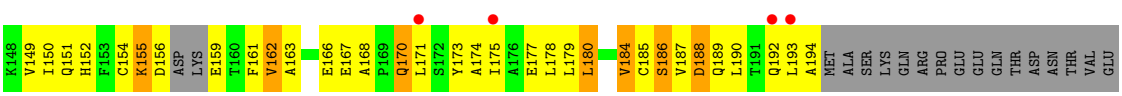
Chain I:





- Molecule 2: Rna15

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	162.04Å 162.04Å 177.54Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.79 – 3.30 47.79 – 3.30	Depositor EDS
% Data completeness (in resolution range)	86.4 (47.79-3.30) 86.3 (47.79-3.30)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 3.33Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.226 , 0.298 0.226 , 0.298	Depositor DCC
R_{free} test set	3646 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	84.8	Xtriage
Anisotropy	0.158	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 76.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.018 for -h,-k,l 0.038 for h,-h-k,-l 0.025 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	30268	wwPDB-VP
Average B, all atoms (Å^2)	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.09% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.46	0/4982	0.64	0/6732
1	B	0.44	0/4747	0.67	1/6416 (0.0%)
1	C	0.40	0/4982	0.60	0/6732
1	D	0.40	0/4689	0.59	0/6338
1	E	0.46	0/4931	0.66	1/6664 (0.0%)
1	F	0.42	0/4708	0.61	0/6363
2	G	0.39	0/626	0.71	0/849
2	H	0.37	0/626	0.67	0/849
2	I	0.38	0/626	0.63	0/849
All	All	0.43	0/30917	0.63	2/41792 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	31	SER	N-CA-C	-5.59	95.91	111.00
1	E	112	LEU	CA-CB-CG	5.05	126.91	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	A	4874	0	4885	509	0
1	B	4646	0	4650	399	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	4874	0	4885	391	0
1	D	4589	0	4591	385	0
1	E	4823	0	4834	511	0
1	F	4608	0	4608	398	0
2	G	618	0	632	81	0
2	H	618	0	632	93	0
2	I	618	0	632	92	0
All	All	30268	0	30349	2735	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 45.

The worst 5 of 2735 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)
1:D:563:THR:HG22	1:D:574:LEU:HD12	1.23	1.17
1:C:557:ASN:HD22	1:C:560:GLU:HB2	1.11	1.16
1:D:468:THR:HG21	1:D:499:VAL:HG11	1.31	1.07
1:A:345:ILE:H	1:A:345:ILE:HD12	1.19	1.03
1:A:225:ASN:HB2	1:A:229:ALA:HB2	1.37	1.02

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	A	579/645 (90%)	433 (75%)	110 (19%)	36 (6%)	1 10
1	B	554/645 (86%)	414 (75%)	108 (20%)	32 (6%)	1 11
1	C	579/645 (90%)	435 (75%)	110 (19%)	34 (6%)	1 10
1	D	547/645 (85%)	398 (73%)	117 (21%)	32 (6%)	1 10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	E	572/645 (89%)	429 (75%)	106 (18%)	37 (6%)	1 9
1	F	549/645 (85%)	422 (77%)	95 (17%)	32 (6%)	1 11
2	G	77/174 (44%)	57 (74%)	12 (16%)	8 (10%)	0 3
2	H	77/174 (44%)	56 (73%)	12 (16%)	9 (12%)	0 2
2	I	77/174 (44%)	58 (75%)	14 (18%)	5 (6%)	1 9
All	All	3611/4392 (82%)	2702 (75%)	684 (19%)	225 (6%)	1 10

5 of 225 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	LYS
1	A	110	ARG
1	A	135	ASN
1	A	255	LYS
1	A	337	GLU

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	A	544/598 (91%)	483 (89%)	61 (11%)	6 23
1	B	519/598 (87%)	472 (91%)	47 (9%)	9 31
1	C	544/598 (91%)	506 (93%)	38 (7%)	15 43
1	D	512/598 (86%)	476 (93%)	36 (7%)	15 43
1	E	538/598 (90%)	487 (90%)	51 (10%)	8 29
1	F	514/598 (86%)	479 (93%)	35 (7%)	16 44
2	G	71/154 (46%)	62 (87%)	9 (13%)	4 19
2	H	71/154 (46%)	65 (92%)	6 (8%)	10 35
2	I	71/154 (46%)	64 (90%)	7 (10%)	8 28
All	All	3384/4050 (84%)	3094 (91%)	290 (9%)	10 35

5 of 290 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	641	LYS
2	H	155	LYS
1	F	152	PHE
1	F	504	GLN
1	B	449	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 144 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	189	GLN
2	H	192	GLN
1	F	242	ASN
1	F	575	GLN
1	C	57	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides 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.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	585/645 (90%)	-0.29	2 (0%) 94 94	39, 76, 136, 149	0
1	B	558/645 (86%)	-0.29	0 100 100	36, 73, 136, 148	0
1	C	585/645 (90%)	-0.10	13 (2%) 62 60	60, 106, 149, 151	0
1	D	551/645 (85%)	-0.13	7 (1%) 77 77	58, 104, 149, 151	0
1	E	578/645 (89%)	0.20	43 (7%) 14 14	57, 106, 150, 151	0
1	F	553/645 (85%)	-0.08	8 (1%) 75 75	55, 104, 148, 151	0
2	G	81/174 (46%)	-0.30	0 100 100	63, 102, 138, 151	0
2	H	81/174 (46%)	0.06	4 (4%) 29 27	90, 128, 148, 151	0
2	I	81/174 (46%)	-0.03	2 (2%) 57 54	96, 125, 147, 151	0
All	All	3653/4392 (83%)	-0.11	79 (2%) 62 60	36, 99, 148, 151	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	125	LEU	4.9
1	E	214	TRP	4.3
1	E	215	GLN	4.3
1	E	218	THR	4.1
1	C	655	SER	4.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.