



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 06:41 am BST

PDB ID : 6QIC  
Title : Crystal structure of DEAH-box ATPase Prp22-S837A with bound ssRNA  
Authors : Hamann, F.; Ficner, R.; Enders, M.  
Deposited on : 2019-01-18  
Resolution : 2.70 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](mailto: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 ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

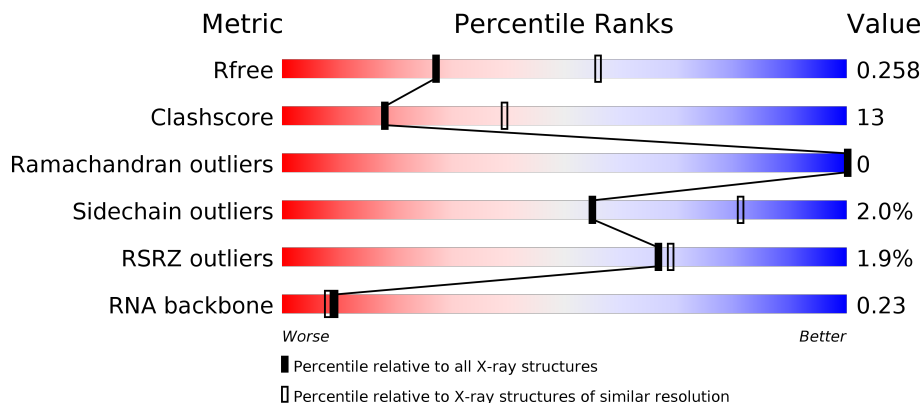
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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.






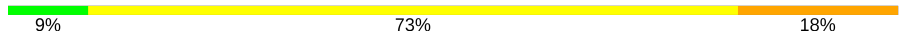
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)
RNA backbone	3102	1159 (3.00-2.40)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	677	 2% 71% 21% 7%
1	B	677	 2% 68% 24% 8%
1	C	677	 2% 74% 18% 7%
1	D	677	 2% 71% 21% 7%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	11	
2	F	11	
2	G	11	
2	H	11	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 20594 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 Putative pre-mRNA splicing factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	628	4932	3146	831	926	29	0	0	0
1	B	626	4934	3147	834	923	30	0	0	0
1	C	627	4942	3154	834	925	29	0	0	0
1	D	629	4953	3159	836	928	30	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	837	ALA	SER	engineered mutation	UNP G0S700
B	837	ALA	SER	engineered mutation	UNP G0S700
C	837	ALA	SER	engineered mutation	UNP G0S700
D	837	ALA	SER	engineered mutation	UNP G0S700

- Molecule 2 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	E	11	220	99	22	88	11	0	0	0
2	F	11	204	90	20	83	11	0	0	0
2	G	11	204	90	20	83	11	0	0	0
2	H	11	204	90	20	83	11	0	0	0

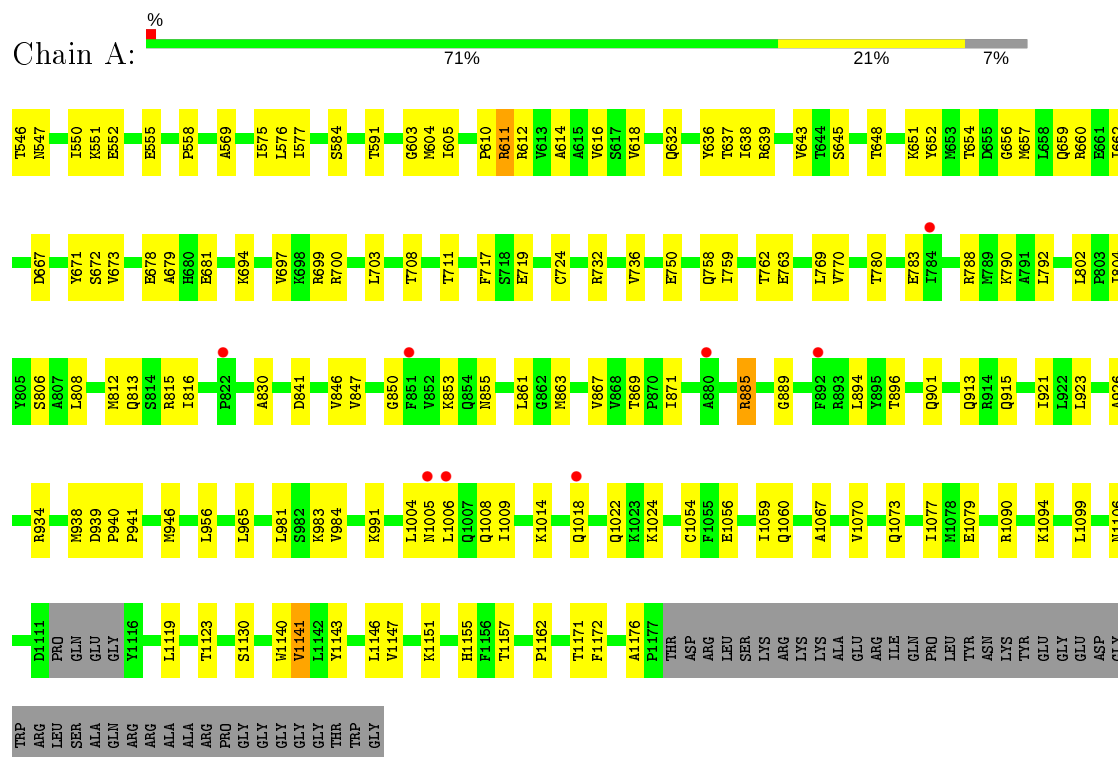
- Molecule 3 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	A	1	Total	O	0	0
			1	1		

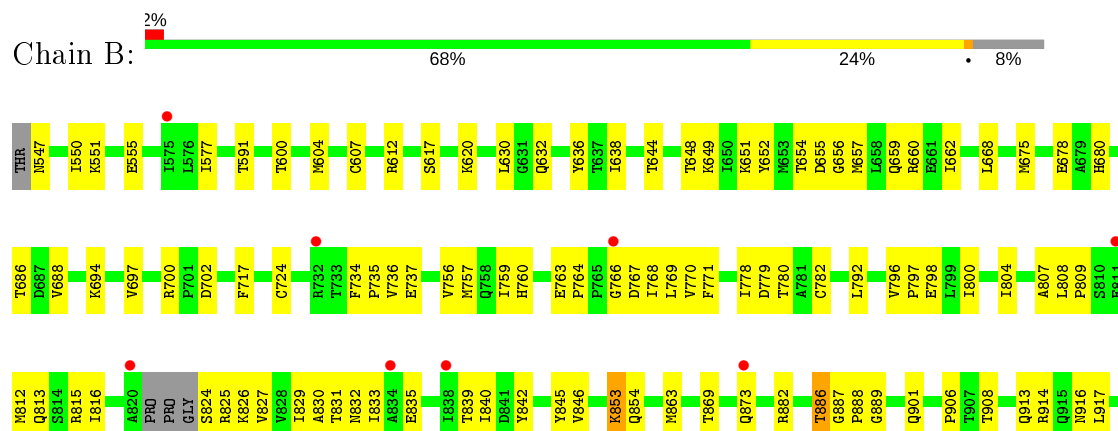
### 3 Residue-property plots i

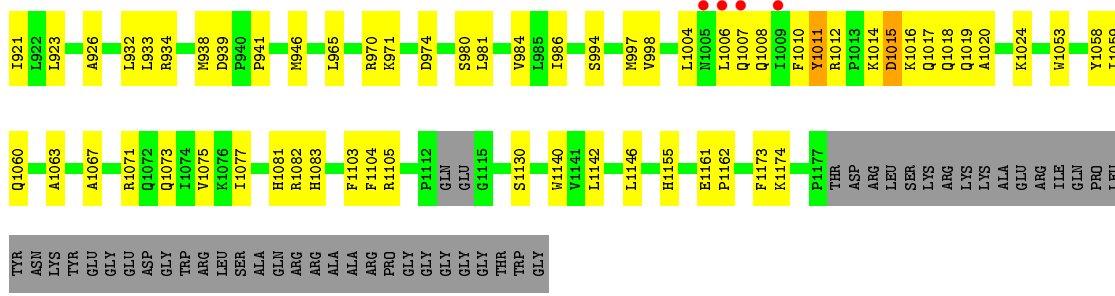
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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 ( $\text{RSRZ} > 2$ ). 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: Putative pre-mRNA splicing factor

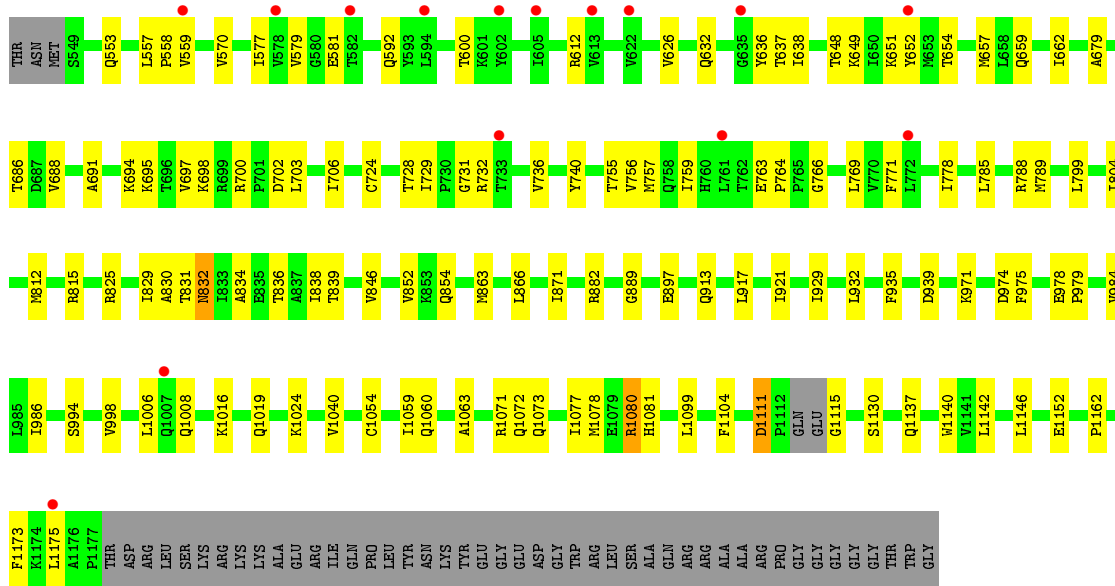
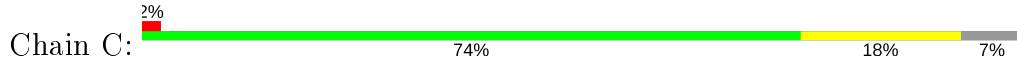


- Molecule 1: Putative pre-mRNA splicing factor

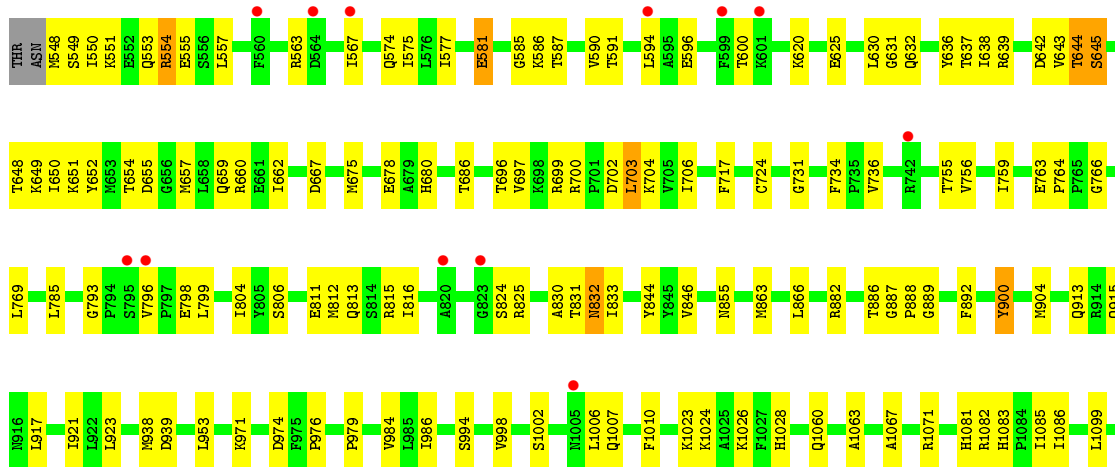


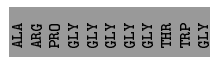
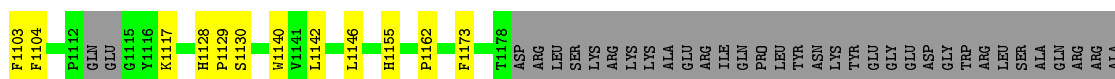


• Molecule 1: Putative pre-mRNA splicing factor



• Molecule 1: Putative pre-mRNA splicing factor

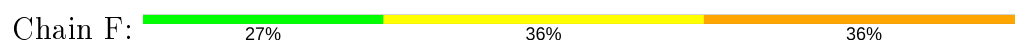




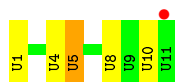
- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



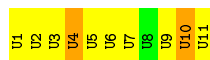
- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	139.61Å 139.96Å 164.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	106.49 – 2.70 106.49 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (106.49-2.70) 99.7 (106.49-2.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.212 , 0.261 0.213 , 0.258	Depositor DCC
$R_{free}$ test set	4442 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	77.2	Xtriage
Anisotropy	0.433	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 48.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.032 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	20594	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.67% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.69	0/5035	0.85	0/6835
1	B	0.71	0/5035	0.85	0/6825
1	C	0.70	0/5046	0.84	0/6845
1	D	0.72	0/5059	0.86	0/6863
2	E	0.52	0/241	0.70	0/370
2	F	0.51	0/223	0.69	0/343
2	G	0.49	0/223	0.70	0/343
2	H	0.52	0/223	0.69	0/343
All	All	0.70	0/21085	0.84	0/28767

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	A	4932	0	4932	135	0
1	B	4934	0	4954	163	0
1	C	4942	0	4969	111	0
1	D	4953	0	4980	137	0
2	E	220	0	111	14	0
2	F	204	0	100	15	0
2	G	204	0	100	10	0
2	H	204	0	100	15	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	1	0	0	0	0
All	All	20594	0	20246	547	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 13.

The worst 5 of 547 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:B:1011:TYR:HB2	1:B:1060:GLN:HE21	1.09	1.18
1:B:798:GLU:HB3	1:B:824:SER:HB3	1.21	1.15
1:C:975:PHE:CZ	1:C:1078:MET:CE	2.30	1.13
1:D:703:LEU:HD12	1:D:704:LYS:N	1.64	1.10
1:A:1006:LEU:HD11	1:A:1067:ALA:HB2	1.34	1.09

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	624/677 (92%)	607 (97%)	17 (3%)	0	100	100
1	B	620/677 (92%)	596 (96%)	24 (4%)	0	100	100
1	C	623/677 (92%)	598 (96%)	25 (4%)	0	100	100
1	D	626/677 (92%)	606 (97%)	20 (3%)	0	100	100
All	All	2493/2708 (92%)	2407 (97%)	86 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	535/580 (92%)	523 (98%)	12 (2%)	52	79
1	B	538/580 (93%)	522 (97%)	16 (3%)	41	70
1	C	539/580 (93%)	531 (98%)	8 (2%)	65	86
1	D	540/580 (93%)	532 (98%)	8 (2%)	65	86
All	All	2152/2320 (93%)	2108 (98%)	44 (2%)	55	81

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	901	GLN
1	B	1011	TYR
1	D	703	LEU
1	B	908	THR
1	B	934	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	573	ASN
1	C	680	HIS
1	D	1060	GLN
1	C	609	GLN
1	C	832	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	E	10/11 (90%)	5 (50%)	1 (10%)
2	F	9/11 (81%)	3 (33%)	1 (11%)
2	G	9/11 (81%)	2 (22%)	0
2	H	9/11 (81%)	5 (55%)	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
All	All	37/44 (84%)	15 (40%)	2 (5%)

5 of 15 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	E	3	U
2	E	4	U
2	E	5	U
2	E	7	U
2	E	10	U

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	E	3	U
2	F	3	U

#### 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 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.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	A	628/677 (92%)	0.08	8 (1%) 77 78	58, 86, 128, 159	0
1	B	626/677 (92%)	0.14	12 (1%) 66 69	60, 91, 131, 159	0
1	C	627/677 (92%)	0.16	15 (2%) 59 60	63, 94, 139, 169	0
1	D	629/677 (92%)	0.11	12 (1%) 66 69	61, 85, 121, 165	0
2	E	11/11 (100%)	-0.18	0 100 100	77, 87, 129, 178	0
2	F	11/11 (100%)	-0.29	0 100 100	78, 92, 142, 154	0
2	G	11/11 (100%)	0.02	1 (9%) 9 7	81, 94, 156, 166	0
2	H	11/11 (100%)	-0.11	0 100 100	70, 85, 134, 145	0
All	All	2554/2752 (92%)	0.12	48 (1%) 66 69	58, 89, 131, 178	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	820	ALA	4.1
1	B	766	GLY	4.0
1	C	602	TYR	3.9
1	B	834	ALA	3.9
1	D	795	SER	3.7

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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.