



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 3, 2023 – 12:59 PM EDT

PDB ID : 3SNP  
Title : Crystal structure analysis of iron regulatory protein 1 in complex with ferritin  
H IRE RNA  
Authors : Volz, K.; Selezneva, A.I.; Walden, W.E.  
Deposited on : 2011-06-29  
Resolution : 2.80 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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

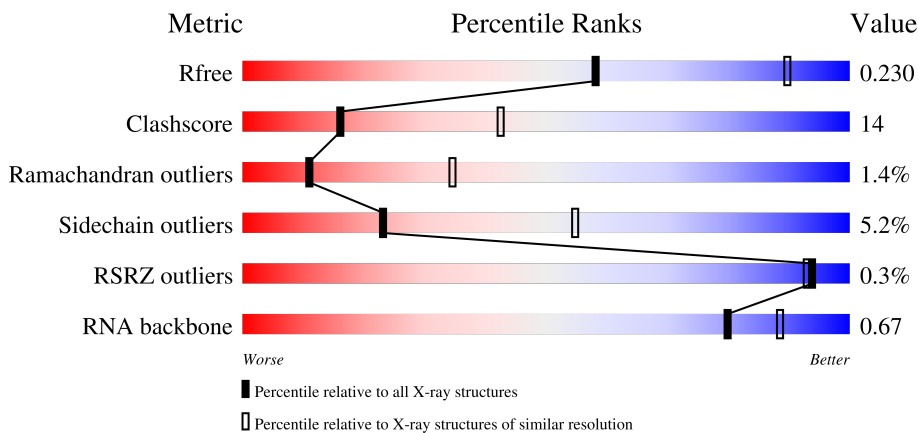
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)
RNA backbone	3102	1227 (3.10-2.50)

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  $\geq 3$ , 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	908	 65% 26% • 6%
1	B	908	 63% 28% • 6%
2	C	30	 60% 23% 17%
2	D	30	 50% 37% 7% 7%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 15113 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 Cytoplasmic aconitate hydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	850	6657	4259	1143	1231	24	0	0	0
1	B	850	6657	4259	1143	1231	24	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	MET	-	expression tag	UNP Q01059
A	-17	GLY	-	expression tag	UNP Q01059
A	-16	HIS	-	expression tag	UNP Q01059
A	-15	HIS	-	expression tag	UNP Q01059
A	-14	HIS	-	expression tag	UNP Q01059
A	-13	HIS	-	expression tag	UNP Q01059
A	-12	HIS	-	expression tag	UNP Q01059
A	-11	HIS	-	expression tag	UNP Q01059
A	-10	ALA	-	expression tag	UNP Q01059
A	-9	ASP	-	expression tag	UNP Q01059
A	-8	ASP	-	expression tag	UNP Q01059
A	-7	ASP	-	expression tag	UNP Q01059
A	-6	ASP	-	expression tag	UNP Q01059
A	-5	LYS	-	expression tag	UNP Q01059
A	-4	ASP	-	expression tag	UNP Q01059
A	-3	GLY	-	expression tag	UNP Q01059
A	-2	VAL	-	expression tag	UNP Q01059
A	-1	ASP	-	expression tag	UNP Q01059
A	0	LYS	-	expression tag	UNP Q01059
A	1	LEU	-	expression tag	UNP Q01059
A	283	PRO	LEU	SEE REMARK 999	UNP Q01059
A	437	SER	CYS	engineered mutation	UNP Q01059
A	503	SER	CYS	engineered mutation	UNP Q01059
A	874	PHE	LEU	SEE REMARK 999	UNP Q01059
B	-18	MET	-	expression tag	UNP Q01059

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-17	GLY	-	expression tag	UNP Q01059
B	-16	HIS	-	expression tag	UNP Q01059
B	-15	HIS	-	expression tag	UNP Q01059
B	-14	HIS	-	expression tag	UNP Q01059
B	-13	HIS	-	expression tag	UNP Q01059
B	-12	HIS	-	expression tag	UNP Q01059
B	-11	HIS	-	expression tag	UNP Q01059
B	-10	ALA	-	expression tag	UNP Q01059
B	-9	ASP	-	expression tag	UNP Q01059
B	-8	ASP	-	expression tag	UNP Q01059
B	-7	ASP	-	expression tag	UNP Q01059
B	-6	ASP	-	expression tag	UNP Q01059
B	-5	LYS	-	expression tag	UNP Q01059
B	-4	ASP	-	expression tag	UNP Q01059
B	-3	GLY	-	expression tag	UNP Q01059
B	-2	VAL	-	expression tag	UNP Q01059
B	-1	ASP	-	expression tag	UNP Q01059
B	0	LYS	-	expression tag	UNP Q01059
B	1	LEU	-	expression tag	UNP Q01059
B	283	PRO	LEU	SEE REMARK 999	UNP Q01059
B	437	SER	CYS	engineered mutation	UNP Q01059
B	503	SER	CYS	engineered mutation	UNP Q01059
B	874	PHE	LEU	SEE REMARK 999	UNP Q01059

- Molecule 2 is a RNA chain called ferritin H IRE RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	30	Total	C	N	O	P	19	0	0
			621	275	106	210	30			
2	D	30	Total	C	N	O	P	19	0	0
			621	275	106	210	30			

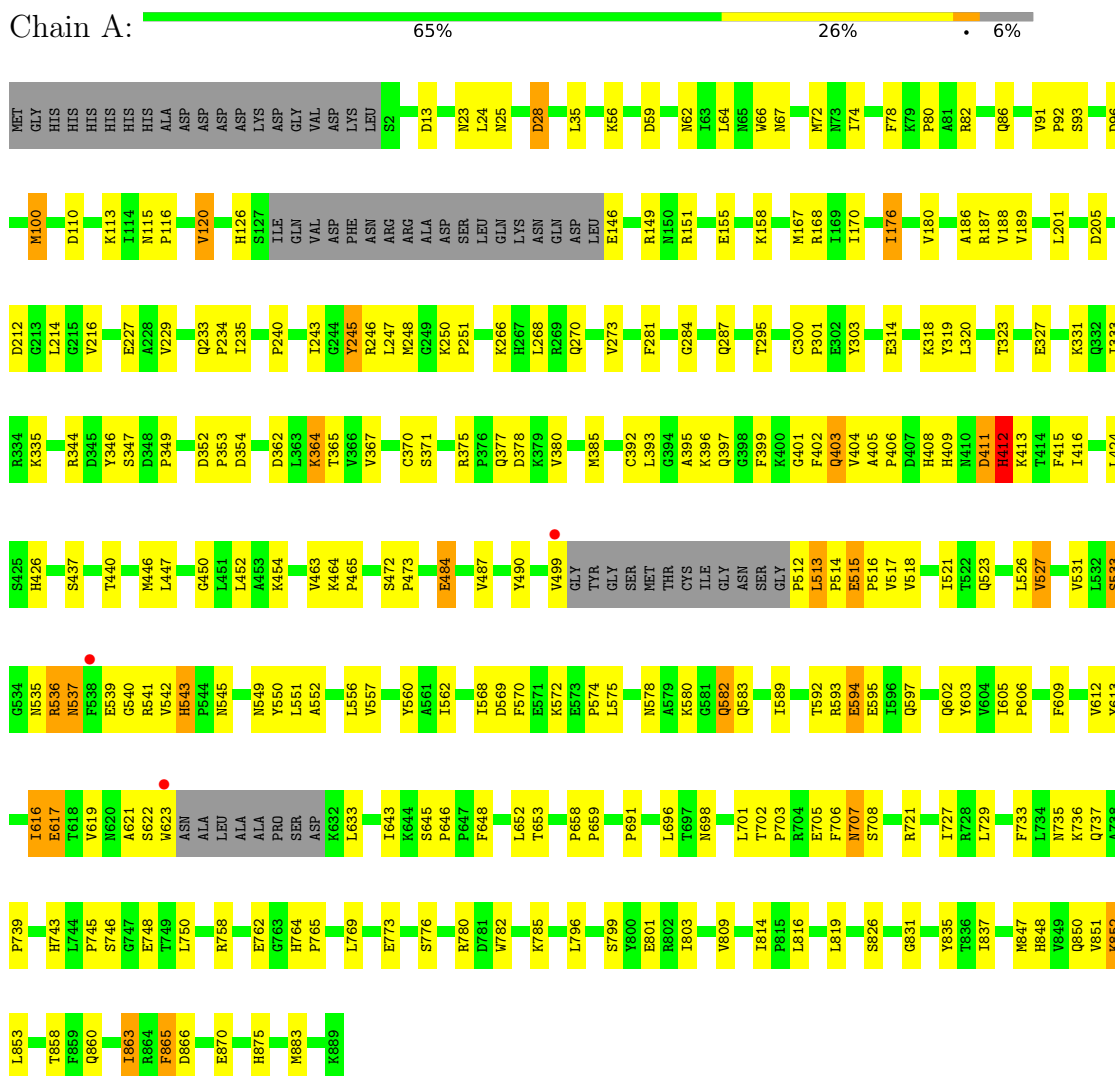
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	259	Total	O	0	0
			259	259		
3	B	226	Total	O	0	0
			226	226		
3	C	40	Total	O	0	0
			40	40		
3	D	32	Total	O	0	0
			32	32		

### 3 Residue-property plots [i](#)

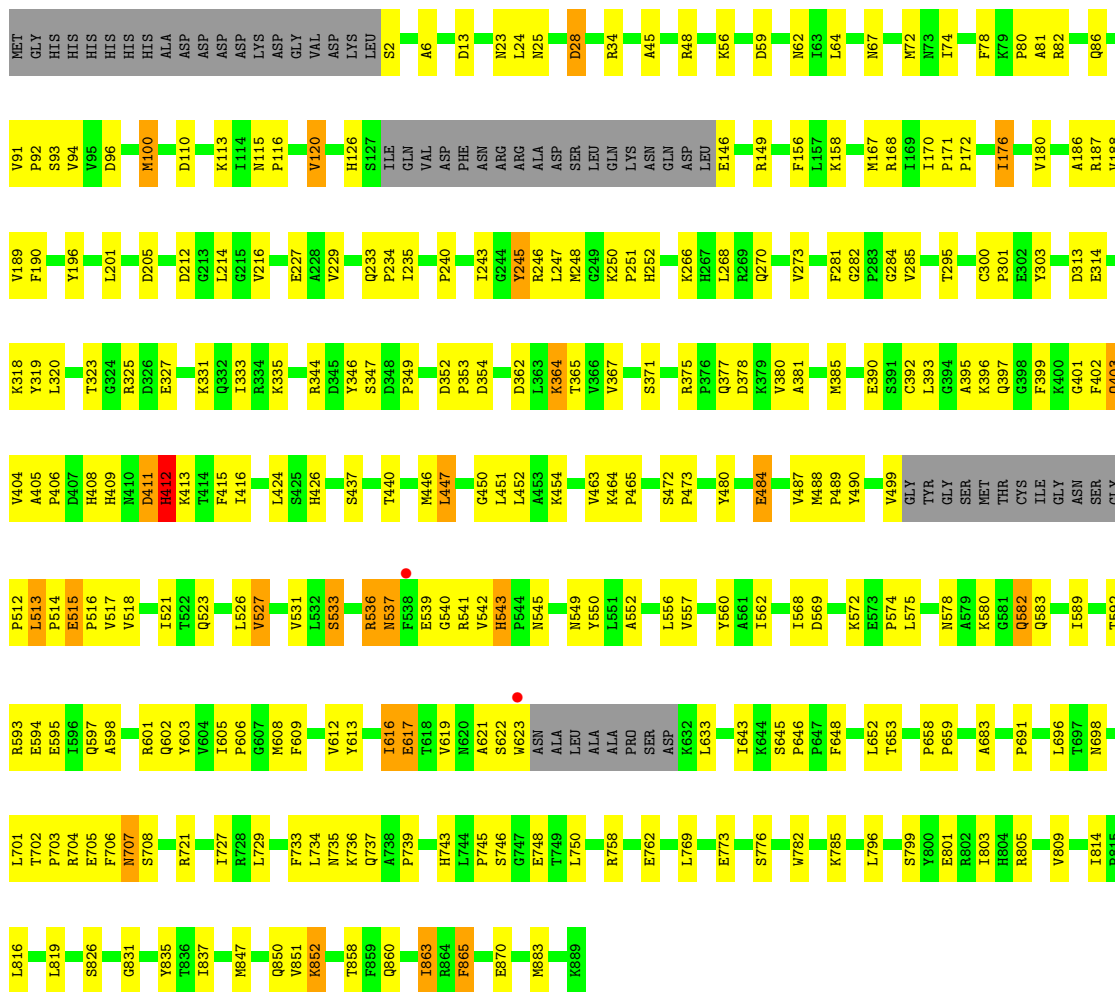
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Cytoplasmic aconitate hydratase



- Molecule 1: Cytoplasmic aconitate hydratase

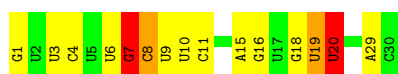




● Molecule 2: ferritin H IRE RNA



● Molecule 2: ferritin H IRE RNA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.56Å 80.85Å 142.87Å 90.00° 92.03° 90.00°	Depositor
Resolution (Å)	99.00 – 2.80 43.09 – 2.70	Depositor EDS
% Data completeness (in resolution range)	91.1 (99.00-2.80) 98.0 (43.09-2.70)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.26 (at 2.69Å)	Xtrriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.198 , 0.218 0.217 , 0.230	Depositor DCC
$R_{free}$ test set	3442 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	55.4	Xtrriage
Anisotropy	0.083	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 39.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.169 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15113	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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	0.38	0/6813	0.64	1/9249 (0.0%)
1	B	0.38	0/6813	0.64	1/9249 (0.0%)
2	C	0.60	2/691 (0.3%)	0.94	3/1072 (0.3%)
2	D	0.59	2/691 (0.3%)	0.94	4/1072 (0.4%)
All	All	0.41	4/15008 (0.0%)	0.67	9/20642 (0.0%)

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
2	C	0	1
2	D	0	3
All	All	0	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	G	OP3-P	-7.27	1.52	1.61
2	D	1	G	OP3-P	-7.20	1.52	1.61
2	C	6	U	O3'-P	-7.14	1.52	1.61
2	D	6	U	O3'-P	-6.96	1.52	1.61

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	19	U	C2'-C3'-O3'	7.28	125.51	109.50
2	C	19	U	C2'-C3'-O3'	6.23	123.67	113.70
2	D	7	G	N9-C1'-C2'	5.83	121.58	114.00
2	C	7	G	N9-C1'-C2'	5.68	121.39	114.00
1	A	865	PHE	N-CA-C	-5.50	96.14	111.00



There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	16	G	Sidechain
2	D	16	G	Sidechain
2	D	18	G	Sidechain
2	D	20	U	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	A	6657	0	6662	198	0
1	B	6657	0	6662	207	0
2	C	621	0	311	8	0
2	D	621	0	311	9	0
3	A	259	0	0	11	0
3	B	226	0	0	17	0
3	C	40	0	0	2	0
3	D	32	0	0	0	0
All	All	15113	0	13946	411	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 14.

The worst 5 of 411 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:A:450:GLY:HA3	1:A:487:VAL:HG11	1.44	1.00
1:B:450:GLY:HA3	1:B:487:VAL:HG11	1.44	0.99
1:B:537:ASN:HD21	1:B:540:GLY:HA3	1.39	0.88
1:A:537:ASN:HD21	1:A:540:GLY:HA3	1.40	0.87
1:A:158:LYS:HD2	1:A:623:TRP:CE2	2.12	0.84

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	842/908 (93%)	781 (93%)	49 (6%)	12 (1%)	11	34
1	B	842/908 (93%)	782 (93%)	48 (6%)	12 (1%)	11	34
All	All	1684/1816 (93%)	1563 (93%)	97 (6%)	24 (1%)	11	34

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	395	ALA
1	A	403	GLN
1	A	536	ARG
1	B	395	ALA
1	B	403	GLN

### 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	725/772 (94%)	687 (95%)	38 (5%)	23	55
1	B	725/772 (94%)	688 (95%)	37 (5%)	24	55
All	All	1450/1544 (94%)	1375 (95%)	75 (5%)	23	55

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	452	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	826	SER
1	B	513	LEU
1	B	582	GLN
1	A	533	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	698	ASN
1	B	707	ASN
1	B	850	GLN
1	A	707	ASN
1	A	698	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	28/30 (93%)	3 (10%)	1 (3%)
2	D	28/30 (93%)	5 (17%)	1 (3%)
All	All	56/60 (93%)	8 (14%)	2 (3%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	9	U
2	C	19	U
2	C	20	U
2	D	7	G
2	D	8	C

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	C	19	U
2	D	19	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 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 [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	850/908 (93%)	-0.50	3 (0%) 92   91	22, 50, 90, 124	0
1	B	850/908 (93%)	-0.52	2 (0%) 95   94	23, 50, 91, 124	0
2	C	30/30 (100%)	-0.70	0 100   100	40, 57, 85, 93	2 (6%)
2	D	30/30 (100%)	-0.71	0 100   100	40, 58, 83, 94	2 (6%)
All	All	1760/1876 (93%)	-0.52	5 (0%) 94   93	22, 50, 91, 124	4 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	623	TRP	2.5
1	B	623	TRP	2.4
1	A	499	VAL	2.2
1	B	538	PHE	2.2
1	A	538	PHE	2.1

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.