



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 04:43 PM EDT

PDB ID : 6CIL  
Title : PRE-REACTION COMPLEX, RAG1(E962Q)/2-INTACT/INTACT  
12/23RSS COMPLEX IN MN2+  
Authors : Chuenchor, W.; Chen, X.; Kim, M.S.; Gellert, M.; Yang, W.  
Deposited on : 2018-02-24  
Resolution : 4.15 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 : 2.35.1  
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

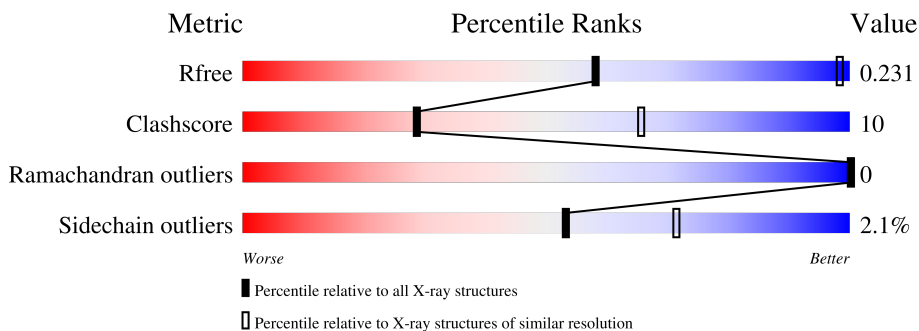
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.15 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	1020 (4.54-3.76)
Clashscore	141614	1028 (4.52-3.80)
Ramachandran outliers	138981	1005 (4.54-3.78)
Sidechain outliers	138945	1024 (4.54-3.76)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	625	
1	C	625	
2	B	359	
2	D	359	
3	N	163	
4	F	40	
5	I	40	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
6	G	55	 33% 62% . .
7	J	55	 27% 69% .

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 18260 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 V(D)J recombination-activating protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	585	Total	C	N	O	S	0	0	0
			4601	2903	802	863	33			
1	C	589	Total	C	N	O	S	0	0	0
			4547	2862	802	851	32			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	962	GLN	GLU	engineered mutation	UNP P15919
C	962	GLN	GLU	engineered mutation	UNP P15919

- Molecule 2 is a protein called V(D)J recombination-activating protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	337	Total	C	N	O	S	0	1	0
			2617	1676	443	481	17			
2	D	330	Total	C	N	O	S	0	1	0
			2528	1625	426	460	17			

- Molecule 3 is a protein called High mobility group protein B1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	N	55	Total	C	N	O	0	0	0
			273	162	55	56			

- Molecule 4 is a DNA chain called Intact 12RSS substrate reverse strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	F	37	Total	C	N	O	P	0	0	0
			768	366	138	227	37			

- Molecule 5 is a DNA chain called Intact 12RSS substrate forward strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	37	Total	C	N	O	P	0	0	0
			749	358	137	217	37			

- Molecule 6 is a DNA chain called Intact 23RSS substrate reverse strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	53	Total	C	N	O	P	0	0	0
			1086	519	189	325	53			

- Molecule 7 is a DNA chain called Intact 23RSS substrate forward strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	J	53	Total	C	N	O	P	0	0	0
			1087	517	206	311	53			

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	Zn	0	0
			1	1		
8	C	1	Total	Zn	0	0
			1	1		

- Molecule 9 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

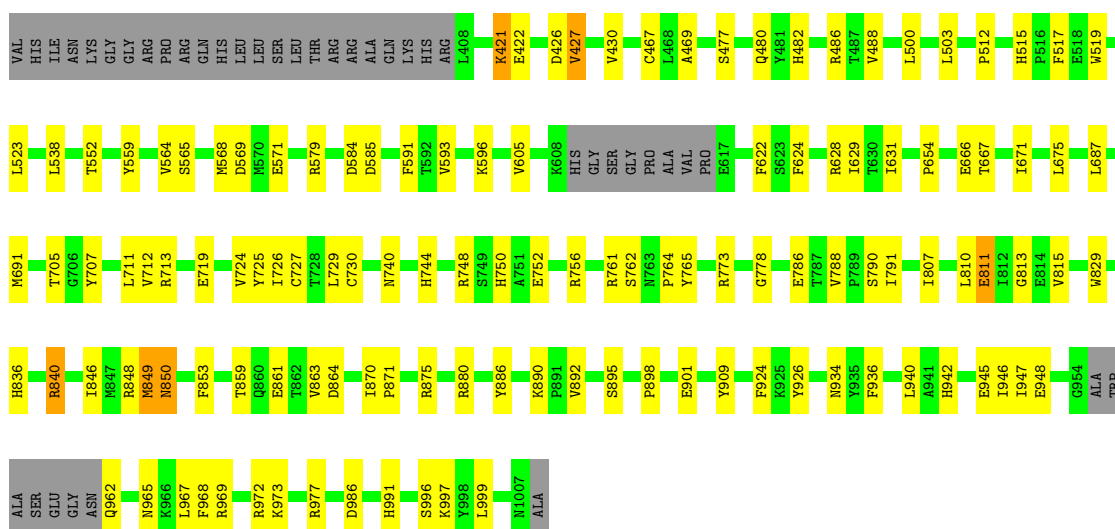
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Mn	0	0
			1	1		
9	C	1	Total	Mn	0	0
			1	1		

### 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. 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.

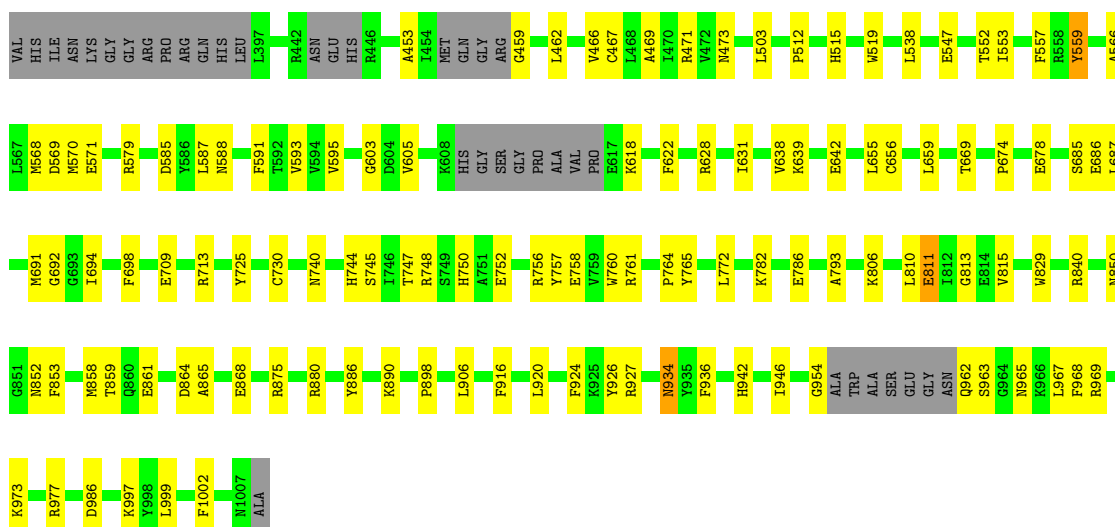
- Molecule 1: V(D)J recombination-activating protein 1

Chain A: 

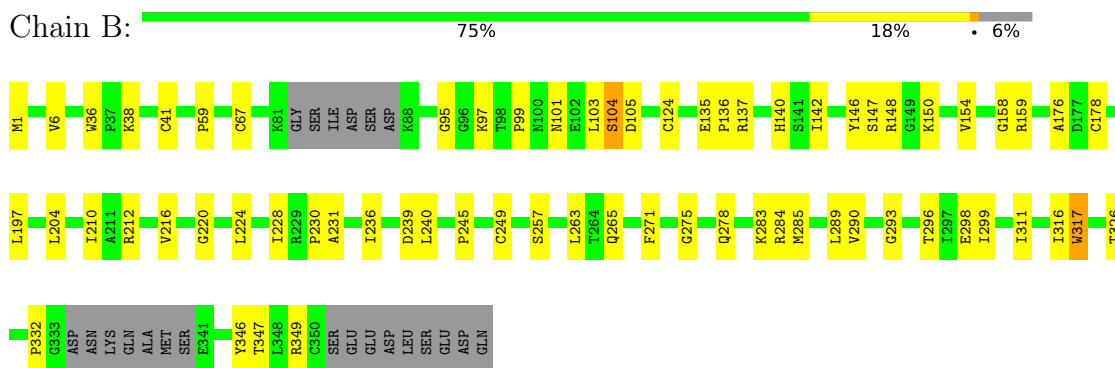


- Molecule 1: V(D)J recombination-activating protein 1

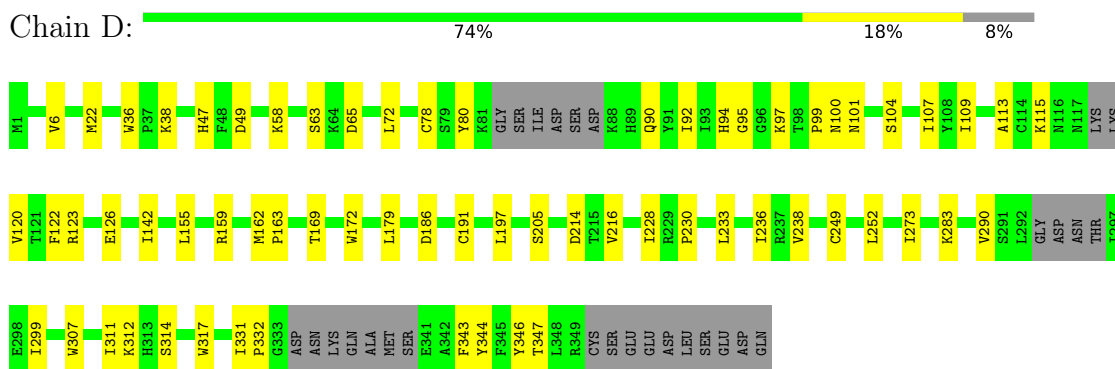
Chain C: 



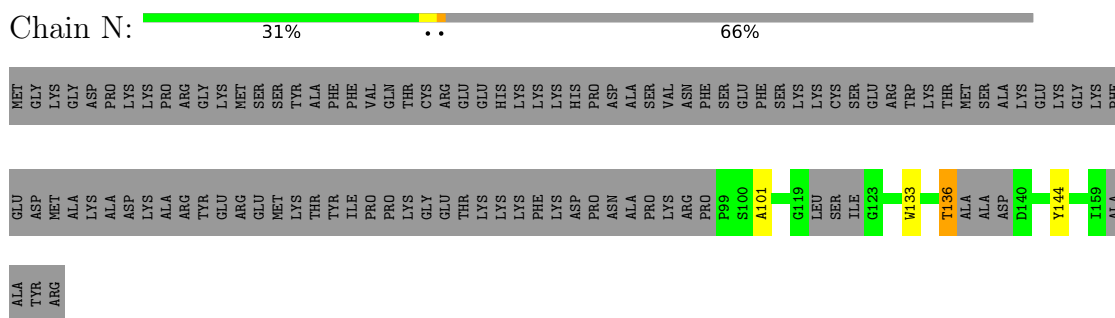
- Molecule 2: V(D)J recombination-activating protein 2



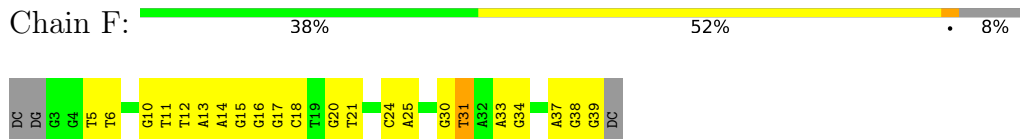
- Molecule 2: V(D)J recombination-activating protein 2



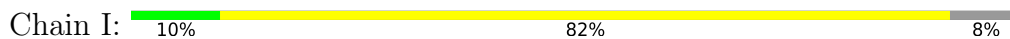
- Molecule 3: High mobility group protein B1

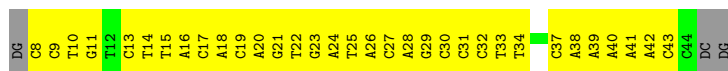


- Molecule 4: Intact 12RSS substrate reverse strand



- Molecule 5: Intact 12RSS substrate forward strand

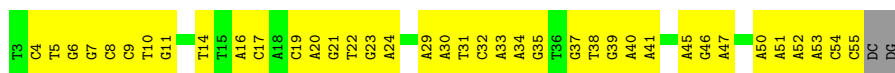
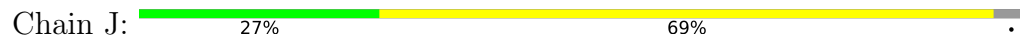




- Molecule 6: Intact 23RSS substrate reverse strand



- Molecule 7: Intact 23RSS substrate forward strand





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	157.33Å 118.83Å 198.96Å 90.00° 100.59° 90.00°	Depositor
Resolution (Å)	37.98 – 4.15 37.98 – 4.15	Depositor EDS
% Data completeness (in resolution range)	99.4 (37.98-4.15) 99.4 (37.98-4.15)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 4.13Å)	Xtrriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.206 , 0.231 0.206 , 0.231	Depositor DCC
$R_{free}$ test set	1396 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	190.4	Xtrriage
Anisotropy	0.262	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.22 , 134.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	18260	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	217.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.99% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MN, ZN

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	0.45	0/4694	0.63	1/6336 (0.0%)
1	C	0.44	0/4635	0.61	0/6260
2	B	0.41	0/2688	0.61	0/3650
2	D	0.39	0/2593	0.59	0/3519
3	N	0.36	0/270	0.50	0/371
4	F	0.94	1/861 (0.1%)	1.08	0/1330
5	I	0.85	0/839	0.98	0/1289
6	G	0.88	0/1215	1.11	4/1874 (0.2%)
7	J	0.87	0/1221	1.05	2/1881 (0.1%)
All	All	0.56	1/19016 (0.0%)	0.75	7/26510 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	31	DT	C3'-O3'	-5.18	1.37	1.44

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	25	DT	O5'-P-OP1	-9.26	97.37	105.70
7	J	14	DT	O4'-C1'-N1	6.93	112.85	108.00
6	G	25	DT	O5'-P-OP2	6.80	118.86	110.70
1	A	523	LEU	CB-CG-CD1	-5.55	101.56	111.00
6	G	17	DT	N3-C4-O4	5.28	123.06	119.90

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	4601	0	4467	78	0
1	C	4547	0	4340	71	0
2	B	2617	0	2539	39	0
2	D	2528	0	2421	38	0
3	N	273	0	125	2	0
4	F	768	0	421	24	0
5	I	749	0	416	35	0
6	G	1086	0	602	29	0
7	J	1087	0	595	36	0
8	A	1	0	0	0	0
8	C	1	0	0	0	0
9	A	1	0	0	0	0
9	C	1	0	0	0	0
All	All	18260	0	15926	329	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 10.

The worst 5 of 329 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)
5:I:21:DG:H2''	5:I:22:DT:H71	1.67	0.75
5:I:8:DC:H2''	5:I:9:DC:H5''	1.71	0.73
7:J:50:DA:H2''	7:J:51:DA:N7	2.03	0.73
2:B:105:ASP:HB3	2:B:136:PRO:HG3	1.69	0.72
4:F:17:DG:H2''	4:F:18:DC:C5	2.26	0.70

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	579/625 (93%)	561 (97%)	18 (3%)	0	100	100
1	C	579/625 (93%)	563 (97%)	16 (3%)	0	100	100
2	B	332/359 (92%)	323 (97%)	9 (3%)	0	100	100
2	D	321/359 (89%)	309 (96%)	12 (4%)	0	100	100
3	N	49/163 (30%)	45 (92%)	4 (8%)	0	100	100
All	All	1860/2131 (87%)	1801 (97%)	59 (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	490/549 (89%)	477 (97%)	13 (3%)	44	66
1	C	468/549 (85%)	457 (98%)	11 (2%)	49	68
2	B	287/316 (91%)	283 (99%)	4 (1%)	67	80
2	D	270/316 (85%)	267 (99%)	3 (1%)	73	84
3	N	1/139 (1%)	0	1 (100%)	0	0
All	All	1516/1869 (81%)	1484 (98%)	32 (2%)	53	71

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	101	ASN
2	D	273	ILE
1	A	967	LEU
1	A	936	PHE
2	D	317	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	990	HIS
2	B	4	GLN
2	D	27	GLN
1	C	633	HIS
1	A	809	GLN

### 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](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.