



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 01:18 PM GMT

PDB ID : 1S32  
Title : Molecular Recognition of the Nucleosomal 'Supergroove'  
Authors : Edayathumangalam, R.S.; Weyermann, P.; Gottesfeld, J.M.; Dervan, P.B.;  
Luger, K.  
Deposited on : 2004-01-12  
Resolution : 2.05 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

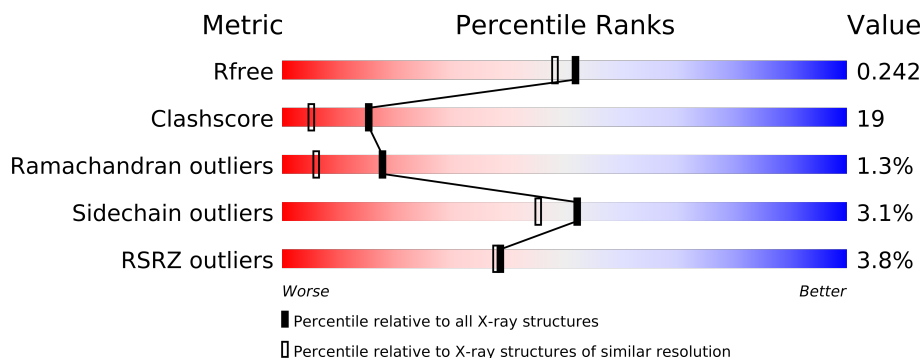
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.05 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))
$R_{free}$	66092	1380 (2.06-2.02)
Clashscore	79885	1577 (2.06-2.02)
Ramachandran outliers	78287	1565 (2.06-2.02)
Sidechain outliers	78261	1565 (2.06-2.02)
RSRZ outliers	66119	1381 (2.06-2.02)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	I	146	
1	J	146	
2	A	135	
2	E	135	
3	B	102	
3	F	102	
4	C	119	
4	G	119	
5	D	122	
5	H	122	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
6	MN	J	2008	-	X
6	MN	J	2011	-	X
8	IMT	I	1603	-	X
8	PYB	I	1604	-	X
8	ABU	I	1605	-	X
8	PYB	I	1606	-	X
8	PYB	I	1607	-	X
8	PYB	I	1608	-	X
8	DIB	I	1611	-	X
8	IMT	I	1621	-	X
8	PYB	I	1622	-	X
8	ABU	I	1625	-	X
8	PYB	I	1628	-	X
8	BAL	I	1630	-	X
8	DIB	I	1631	-	X
9	OGG	J	1700	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 13437 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA chain called palindromic alpha-satellite 146 bp DNA fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	146	Total	C	N	O	P	0	0	0
			2990	1430	541	874	145			
1	J	146	Total	C	N	O	P	0	0	0
			2990	1430	541	874	145			

- Molecule 2 is a protein called histone H3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	98	Total	C	N	O	S	0	1	0
			816	514	157	141	4			
2	E	99	Total	C	N	O	S	0	5	0
			867	547	170	146	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	MET	INTIATING METHIONINE	GB 30268544
E	?	-	MET	INTIATING METHIONINE	GB 30268544

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	81	Total	C	N	O	S	0	2	0
			666	418	132	115	1			
3	F	88	Total	C	N	O	S	0	2	0
			730	457	151	121	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	MET	INTIATING METHIONINE	UNP P02304
F	?	-	MET	INTIATING METHIONINE	UNP P02304

- Molecule 4 is a protein called histone H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	C	107	Total	C	N	O	0	1	0
			834	525	165	144			
4	G	107	Total	C	N	O	0	0	0
			828	522	162	144			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	MET	INITIATING METHIONINE	GB 30268540
G	?	-	MET	INITIATING METHIONINE	GB 30268540

- Molecule 5 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	101	Total	C	N	O	S	0	3	0
			828	520	154	152	2			
5	H	98	Total	C	N	O	S	0	2	0
			792	496	148	146	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	MET	INITIATING METHIONINE	GB 30268542
H	?	-	MET	INITIATING METHIONINE	GB 30268542

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	6	Total	Mn	0	0
			6	6		
6	I	7	Total	Mn	0	0
			7	7		
6	E	1	Total	Mn	0	0
			1	1		

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

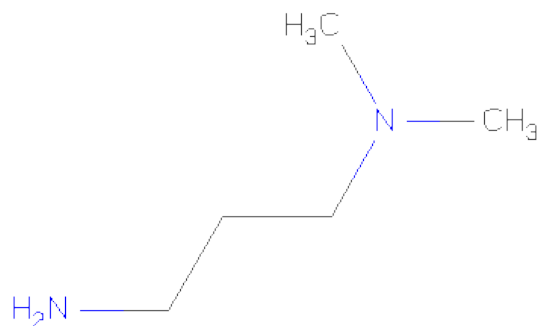
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	1	Total	Cl	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

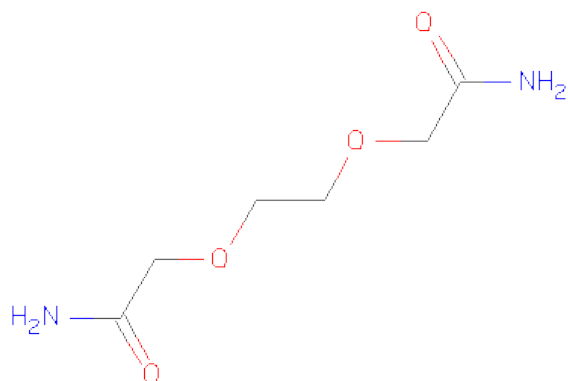
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Cl	0	0
			1	1		
7	D	1	Total	Cl	0	0
			1	1		
7	E	1	Total	Cl	0	0
			1	1		

- Molecule 8 is 4-AMINO-(1-METHYLIMIDAZOLE)-2-CARBOXYLICACID (three-letter code: DIB, PYB, IMT, BAL, ABU) (formula:  $C_5H_{14}N_2$ ,  $C_6H_8N_2O_2$ ,  $C_5H_7N_3O_2$ ,  $C_3H_7NO_2$ ,  $C_4H_9NO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	I	11	Total	C	N	O	0	0
			89	58	21	10		
8	I	11	Total	C	N	O	0	0
			89	58	21	10		

- Molecule 9 is 2-(2-CARBAMOYLMETHOXY-ETHOXY)-ACETAMIDE (three-letter code: OGG) (formula:  $C_6H_{12}N_2O_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	J	1	Total	C	N	O	0	0
			12	6	2	4		

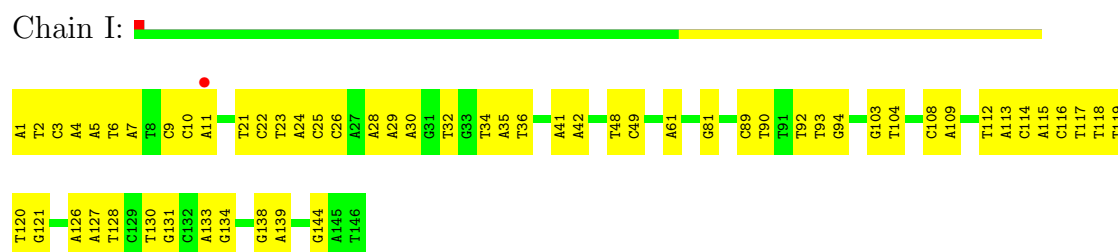
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	77	Total	O	0	0
			77	77		
10	B	53	Total	O	0	0
			53	53		
10	C	109	Total	O	0	0
			109	109		
10	D	62	Total	O	0	0
			62	62		
10	E	97	Total	O	0	0
			97	97		
10	F	85	Total	O	0	0
			85	85		
10	G	80	Total	O	0	0
			80	80		
10	H	29	Total	O	0	0
			29	29		
10	I	145	Total	O	1	0
			145	145		
10	J	151	Total	O	0	0
			151	151		

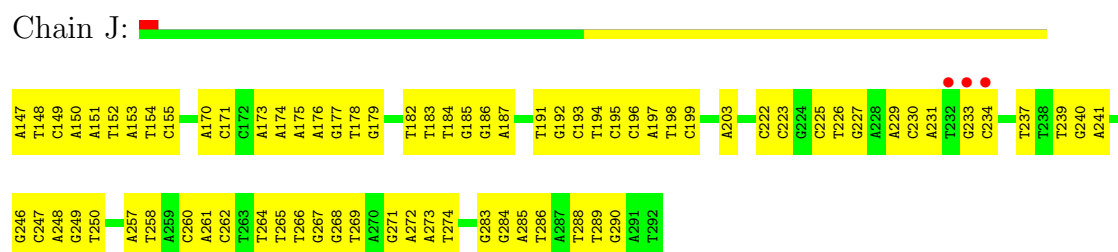
### 3 Residue-property plots

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 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.

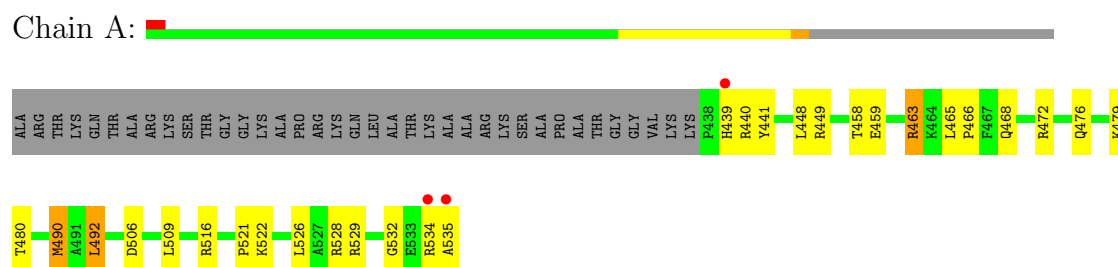
- Molecule 1: palindromic alpha-satellite 146 bp DNA fragment



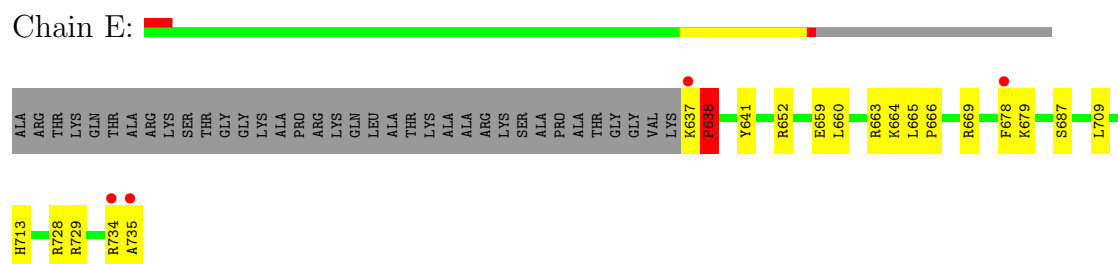
- Molecule 1: palindromic alpha-satellite 146 bp DNA fragment



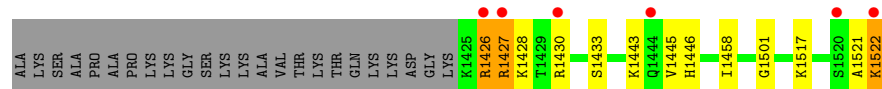
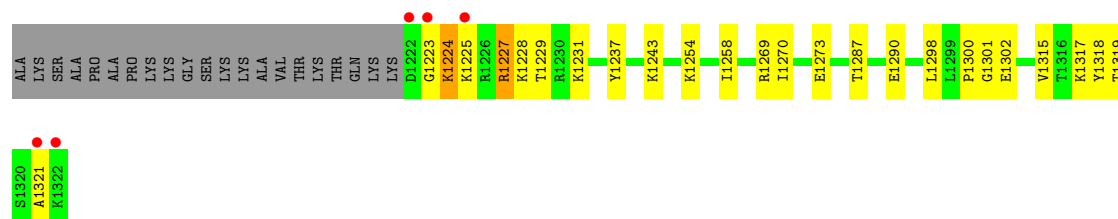
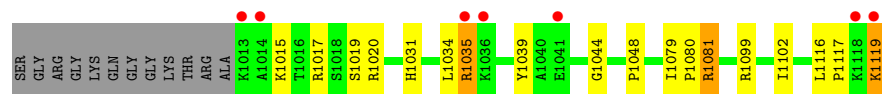
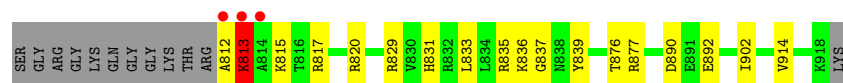
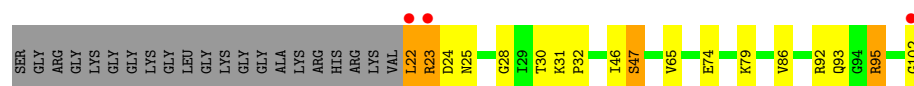
- Molecule 2: histone H3



- Molecule 2: histone H3







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.35Å 109.72Å 181.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	100.00 – 2.05 93.94 – 2.05	Depositor EDS
% Data completeness (in resolution range)	90.7 (100.00-2.05) 91.2 (93.94-2.05)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.89 (at 2.05Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.219 , 0.243 0.222 , 0.242	Depositor DCC
$R_{free}$ test set	6050 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	44.1	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 57.1	EDS
Estimated twinning fraction	0.016 for k,h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	2 of 124445 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13437	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL, MN, DIB, IMT, OGG, PYB, BAL, ABU

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	I	0.26	0/3354	0.68	0/5175
1	J	0.27	0/3354	0.68	0/5175
2	A	0.38	0/828	0.57	0/1109
2	E	0.51	0/880	0.64	1/1176 (0.1%)
3	B	0.37	0/673	0.61	0/899
3	F	0.42	0/738	0.71	1/983 (0.1%)
4	C	0.44	0/844	0.60	1/1138 (0.1%)
4	G	0.33	0/838	0.51	0/1128
5	D	0.41	0/840	0.62	0/1124
5	H	0.33	0/804	0.53	0/1077
All	All	0.34	0/13153	0.65	3/18984 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	218	HIS	N-CA-C	5.79	126.63	111.00
2	E	735	ALA	N-CA-C	5.11	124.78	111.00
4	C	813	LYS	N-CA-C	5.06	124.67	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	2990	0	1651	101	0
1	J	2990	0	1651	106	0
2	A	816	0	854	34	0
2	E	867	0	910	23	0
3	B	666	0	706	21	0
3	F	730	0	784	25	0
4	C	834	0	894	30	0
4	G	828	0	890	27	0
5	D	828	0	862	34	0
5	H	792	0	822	19	0
6	E	1	0	0	0	0
6	I	7	0	0	0	0
6	J	6	0	0	0	0
7	A	1	0	0	1	0
7	D	1	0	0	0	0
7	E	1	0	0	0	0
7	G	1	0	0	1	0
8	I	178	0	138	94	0
9	J	12	0	10	13	0
10	A	77	0	0	8	0
10	B	53	0	0	3	0
10	C	109	0	0	7	0
10	D	62	0	0	6	0
10	E	97	0	0	5	0
10	F	85	0	0	3	0
10	G	80	0	0	2	0
10	H	29	0	0	0	0
10	I	145	0	0	9	0
10	J	151	0	0	5	0
All	All	13437	0	10172	426	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 19.

The worst 5 of 426 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
9:J:1700:OGG:CG	9:J:1700:OGG:OH	1.65	1.44
9:J:1700:OGG:CI	9:J:1700:OGG:OH	1.70	1.39
9:J:1700:OGG:OE	9:J:1700:OGG:CD	1.80	1.29
8:I:1605:ABU:HB2	9:J:1700:OGG:CD	1.70	1.15

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:I:133:DA:H2"	1:I:134:DG:H5"	1.36	1.07

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 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	
2	A	97/135 (72%)	97 (100%)	0	0	100	100
2	E	102/135 (76%)	101 (99%)	0	1 (1%)	22	9
3	B	81/102 (79%)	78 (96%)	1 (1%)	2 (2%)	9	1
3	F	88/102 (86%)	82 (93%)	6 (7%)	0	100	100
4	C	106/119 (89%)	102 (96%)	4 (4%)	0	100	100
4	G	105/119 (88%)	103 (98%)	2 (2%)	0	100	100
5	D	102/122 (84%)	97 (95%)	1 (1%)	4 (4%)	5	0
5	H	98/122 (80%)	92 (94%)	3 (3%)	3 (3%)	7	1
All	All	779/956 (82%)	752 (96%)	17 (2%)	10 (1%)	18	6

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	23	ARG
3	B	25	ASN
5	D	1224	LYS
5	D	1227	ARG
5	D	1301	GLY

### 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 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	
2	A	86/110 (78%)	80 (93%)	6 (7%)	21	12
2	E	91/110 (83%)	90 (99%)	1 (1%)	84	83
3	B	68/78 (87%)	64 (94%)	4 (6%)	28	17
3	F	74/78 (95%)	73 (99%)	1 (1%)	78	77
4	C	85/92 (92%)	82 (96%)	3 (4%)	48	40
4	G	85/92 (92%)	82 (96%)	3 (4%)	48	40
5	D	89/102 (87%)	89 (100%)	0	100	100
5	H	86/102 (84%)	82 (95%)	4 (5%)	36	26
All	All	664/764 (87%)	642 (97%)	22 (3%)	52	42

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	95[B]	ARG
4	C	876	THR
5	H	1433	SER
4	C	813	LYS
4	C	829	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	439	HIS
4	C	831	HIS
5	D	1279	HIS
4	G	1031	HIS
5	H	1492	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

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

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 41 ligands modelled in this entry, 18 are monoatomic - leaving 23 for Mogul analysis.

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$
8	IMT	I	1601	-	6,8,10	2.23	2 (33%)	4,10,14	2.63	3 (75%)
8	PYB	I	1602	-	9,9,10	1.75	2 (22%)	8,12,14	2.64	3 (37%)
8	IMT	I	1603	-	9,9,10	2.25	3 (33%)	9,12,14	2.88	3 (33%)
8	PYB	I	1604	-	9,9,10	1.64	2 (22%)	8,12,14	2.66	3 (37%)
8	ABU	I	1605	-	3,5,6	1.97	1 (33%)	2,4,6	5.41	1 (50%)
8	PYB	I	1606	-	9,9,10	3.44	4 (44%)	8,12,14	2.72	4 (50%)
8	PYB	I	1607	-	9,9,10	1.97	1 (11%)	8,12,14	2.78	3 (37%)
8	PYB	I	1608	-	9,9,10	2.01	1 (11%)	8,12,14	2.89	3 (37%)
8	PYB	I	1609	-	9,9,10	1.57	1 (11%)	8,12,14	2.82	3 (37%)
8	BAL	I	1610	-	4,4,5	8.72	1 (25%)	1,3,5	0.03	0
8	DIB	I	1611	-	6,6,6	0.58	0	6,6,6	0.89	0
8	IMT	I	1621	-	6,8,10	2.23	2 (33%)	4,10,14	2.62	3 (75%)
8	PYB	I	1622	-	9,9,10	1.56	2 (22%)	8,12,14	2.55	3 (37%)
8	IMT	I	1623	-	9,9,10	2.06	3 (33%)	9,12,14	3.09	4 (44%)
8	PYB	I	1624	-	9,9,10	1.68	3 (33%)	8,12,14	2.48	3 (37%)
8	ABU	I	1625	-	3,5,6	0.77	0	2,4,6	4.86	1 (50%)
8	PYB	I	1626	-	9,9,10	2.34	4 (44%)	8,12,14	2.54	3 (37%)
8	PYB	I	1627	-	9,9,10	1.64	1 (11%)	8,12,14	2.62	3 (37%)
8	PYB	I	1628	-	9,9,10	1.76	1 (11%)	8,12,14	2.78	3 (37%)
8	PYB	I	1629	-	9,9,10	1.32	1 (11%)	8,12,14	2.61	3 (37%)
8	BAL	I	1630	-	4,4,5	8.75	1 (25%)	1,3,5	0.32	0
8	DIB	I	1631	-	6,6,6	0.64	0	6,6,6	0.84	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	OGG	J	1700	-	11,11,11	6.69	8 (72%)	12,12,12	2.98	5 (41%)

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
8	IMT	I	1601	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1602	-	-	0/0/2/4	0/1/1/1
8	IMT	I	1603	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1604	-	-	0/0/2/4	0/1/1/1
8	ABU	I	1605	-	-	0/2/3/4	0/0/0/0
8	PYB	I	1606	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1607	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1608	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1609	-	-	0/0/2/4	0/1/1/1
8	BAL	I	1610	-	-	0/1/2/3	0/0/0/0
8	DIB	I	1611	-	-	0/4/4/4	0/0/0/0
8	IMT	I	1621	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1622	-	-	0/0/2/4	0/1/1/1
8	IMT	I	1623	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1624	-	-	0/0/2/4	0/1/1/1
8	ABU	I	1625	-	-	0/2/3/4	0/0/0/0
8	PYB	I	1626	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1627	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1628	-	-	0/0/2/4	0/1/1/1
8	PYB	I	1629	-	-	0/0/2/4	0/1/1/1
8	BAL	I	1630	-	-	0/1/2/3	0/0/0/0
8	DIB	I	1631	-	-	0/4/4/4	0/0/0/0
9	OGG	J	1700	-	-	0/9/9/9	0/0/0/0

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	I	1630	BAL	O-C	17.45	1.23	1.11
8	I	1610	BAL	O-C	17.38	1.23	1.11
9	J	1700	OGG	OE-CD	15.67	1.80	1.42
9	J	1700	OGG	OH-CI	11.74	1.70	1.42
8	I	1606	PYB	CB-NG2	5.81	1.46	1.37

The worst 5 of 57 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	I	1605	ABU	CG-CB-CA	-7.64	99.20	112.71
9	J	1700	OGG	CD-CC-NB	7.51	132.24	116.59
8	I	1623	IMT	CG1-NB1-CA	7.51	108.63	102.87
8	I	1625	ABU	CG-CB-CA	-6.83	100.63	112.71
8	I	1603	IMT	CG1-NB1-CA	6.74	108.04	102.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

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 ⓘ

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	I	146/146 (100%)	0.14	1 (0%) 84 87	42, 97, 137, 157	0
1	J	146/146 (100%)	0.10	3 (2%) 60 61	46, 100, 145, 162	0
2	A	98/135 (72%)	0.40	3 (3%) 47 46	31, 44, 71, 93	0
2	E	99/135 (73%)	0.28	4 (4%) 36 36	26, 36, 61, 87	0
3	B	81/102 (79%)	0.37	3 (3%) 39 39	32, 42, 71, 115	0
3	F	88/102 (86%)	0.38	5 (5%) 23 21	26, 37, 81, 132	0
4	C	107/119 (89%)	0.41	3 (2%) 50 51	28, 41, 65, 93	0
4	G	107/119 (89%)	0.62	7 (6%) 18 16	35, 52, 77, 109	0
5	D	101/122 (82%)	0.56	5 (4%) 28 26	30, 43, 112, 138	0
5	H	98/122 (80%)	0.51	6 (6%) 21 18	34, 50, 105, 128	0
All	All	1071/1248 (85%)	0.36	40 (3%) 38 39	26, 50, 126, 162	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	C	812	ALA	13.2
4	G	1119	LYS	10.5
4	C	813	LYS	9.6
3	B	102	GLY	9.5
2	A	535	ALA	7.9

### 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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	DIB	I	1611	7/7	0.38	59.14	142,142,142,142	0
8	IMT	I	1621	8/10	0.33	19.45	147,147,148,148	0
8	PYB	I	1622	9/10	0.30	15.54	147,147,147,148	0
8	DIB	I	1631	7/7	0.81	12.72	142,143,143,143	0
8	PYB	I	1628	9/10	0.33	9.60	144,144,145,145	0
6	MN	J	2011	1/1	0.16	8.49	93,93,93,93	1
8	PYB	I	1608	9/10	0.22	8.18	144,144,145,145	0
8	ABU	I	1625	6/7	0.41	4.61	137,139,141,143	0
8	PYB	I	1604	9/10	0.23	4.47	146,148,148,148	0
8	PYB	I	1606	9/10	0.19	3.89	140,143,145,146	0
8	ABU	I	1605	6/7	0.20	3.21	136,138,141,144	0
6	MN	J	2008	1/1	0.18	2.97	109,109,109,109	1
8	PYB	I	1607	9/10	0.16	2.43	145,145,145,145	0
9	OGG	J	1700	12/12	0.25	2.39	111,119,129,130	0
8	IMT	I	1603	9/10	0.20	2.36	148,148,148,148	0
8	BAL	I	1630	5/6	0.39	2.28	143,143,143,143	0
8	BAL	I	1610	5/6	0.25	1.85	141,142,142,143	0
8	PYB	I	1624	9/10	0.27	1.82	144,146,146,146	0
6	MN	I	2003	1/1	0.15	1.32	83,83,83,83	0
8	IMT	I	1623	9/10	0.17	1.14	146,146,147,147	0
6	MN	I	2009	1/1	0.16	1.09	106,106,106,106	1
8	PYB	I	1626	9/10	0.25	0.75	141,144,146,146	0
6	MN	I	2004	1/1	0.13	0.16	77,77,77,77	1
8	PYB	I	1602	9/10	0.17	0.16	148,148,149,149	0
8	IMT	I	1601	8/10	0.31	-0.03	149,149,149,149	0
8	PYB	I	1629	9/10	0.25	-0.18	143,144,144,144	0
6	MN	J	2010	1/1	0.12	-0.21	104,104,104,104	1
8	PYB	I	1609	9/10	0.21	-0.25	143,144,144,144	0
7	CL	D	2018	1/1	0.13	-0.38	50,50,50,50	0
8	PYB	I	1627	9/10	0.17	-0.66	145,145,145,145	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	MN	E	2001	1/1	0.11	-0.98	34,34,34,34	0
6	MN	J	2013	1/1	0.12	-1.12	90,90,90,90	1
7	CL	G	2016	1/1	0.12	-1.22	54,54,54,54	0
6	MN	I	2014	1/1	0.11	-1.30	107,107,107,107	1
6	MN	J	2002	1/1	0.10	-1.38	72,72,72,72	1
6	MN	J	2005	1/1	0.10	-1.57	77,77,77,77	0
6	MN	I	2007	1/1	0.10	-1.94	89,89,89,89	1
6	MN	I	2006	1/1	0.10	-2.41	101,101,101,101	1
7	CL	A	2017	1/1	0.09	-3.20	56,56,56,56	0
7	CL	E	2015	1/1	0.11	-3.62	48,48,48,48	1
6	MN	I	2012	1/1	0.06	-8.60	60,60,60,60	1

## 6.5 Other polymers ⓘ

There are no such residues in this entry.