



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 04:12 am BST

PDB ID : 3X1U
Title : Crystal structure of nucleosome core particle in the presence of histone variants involved in reprogramming
Authors : Sivaraman, P.; Kumarevel, T.S.
Deposited on : 2014-11-28
Resolution : 3.25 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 ⓘ 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

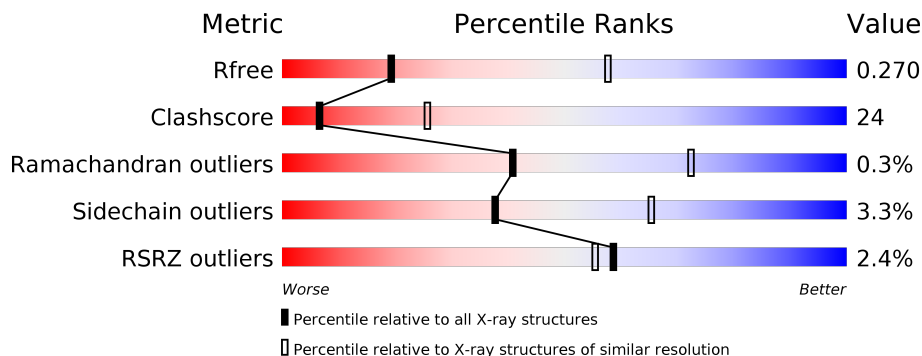
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.25 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	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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 ≥ 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	I	146	
1	J	146	
2	A	135	
2	E	135	
3	B	102	
3	F	102	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	C	128	
4	G	128	
5	D	124	
5	H	124	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	CL	E	202	-	-	X	-

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 12121 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 DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	I	146	2990	1431	540	874	145	0	0	0
1	J	146	2990	1431	540	874	145	0	0	0

- Molecule 2 is a protein called Histone H3.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	98	807	508	156	139	4	0	0	0
2	E	99	816	514	158	140	4	0	0	0

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	83	662	418	129	114	1	0	0	0
3	F	87	703	442	142	118	1	0	0	0

- Molecule 4 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	C	105	809	509	159	141	0	0	0
4	G	109	843	530	168	145	0	0	0

- Molecule 5 is a protein called Histone H2B type 1-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	94	Total	C	N	O	S	0	0	0
			736	462	132	140	2			
5	H	94	Total	C	N	O	S	0	0	0
			735	462	132	139	2			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	G	1	Total	Mn	0	0
			1	1		
6	I	2	Total	Mn	0	0
			2	2		
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	J	2	Total	Cl	0	0
			2	2		
7	I	3	Total	Cl	0	0
			3	3		
7	E	1	Total	Cl	0	0
			1	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	I	4	Total	O	0	0
			4	4		
8	J	6	Total	O	0	0
			6	6		
8	A	1	Total	O	0	0
			1	1		
8	B	1	Total	O	0	0
			1	1		
8	C	4	Total	O	0	0
			4	4		
8	E	1	Total	O	0	0
			1	1		
8	F	1	Total	O	0	0
			1	1		

Continued on next page...

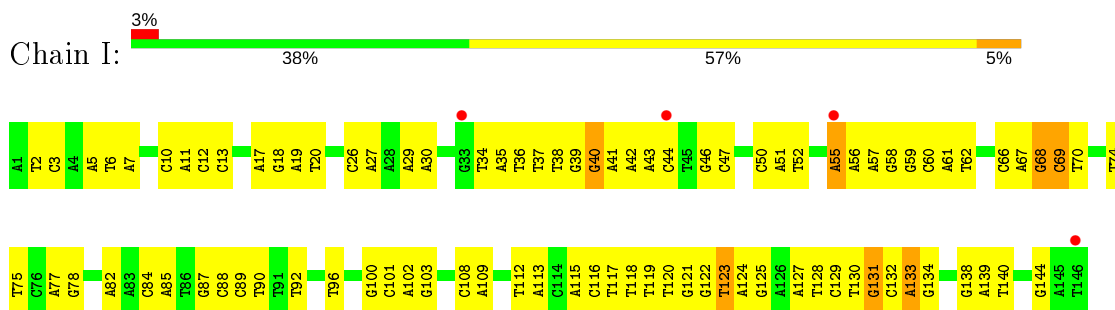
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	G	2	Total	O	0	0
			2	2		

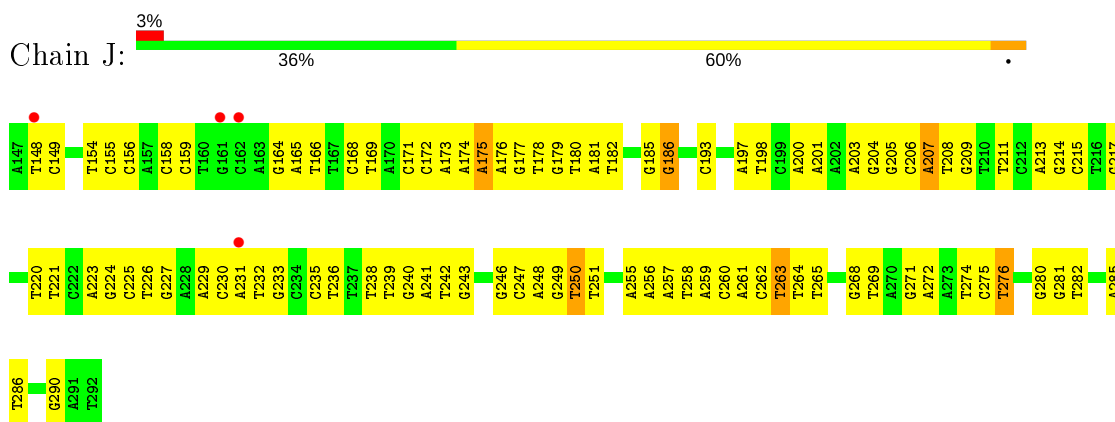
3 Residue-property plots [i](#)

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 ($RSRZ > 2$). 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.

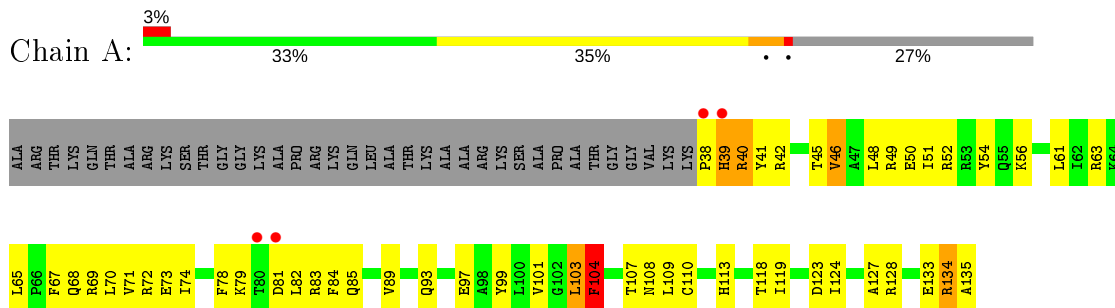
- Molecule 1: DNA (146-MER)



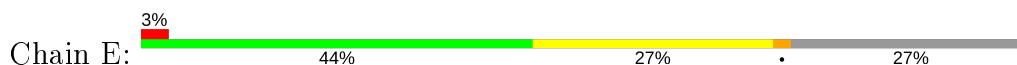
- Molecule 1: DNA (146-MER)

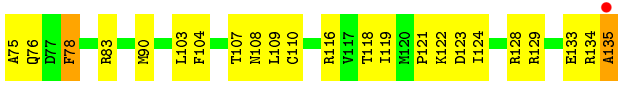
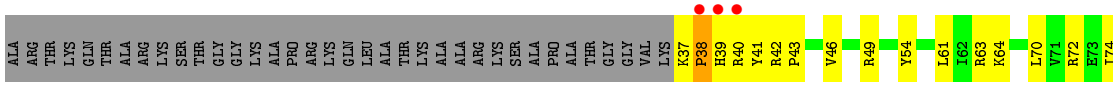


- Molecule 2: Histone H3.1

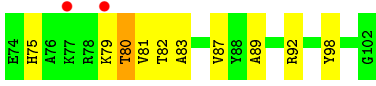


- Molecule 2: Histone H3.1





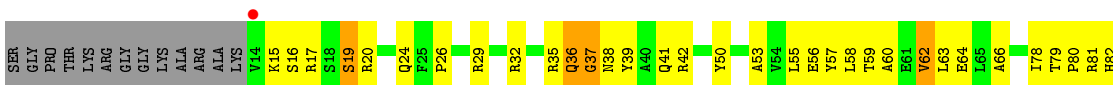
• Molecule 3: Histone H4



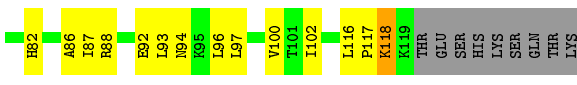
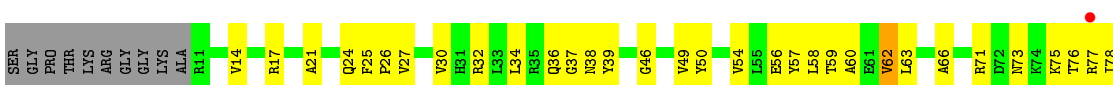
• Molecule 3: Histone H4



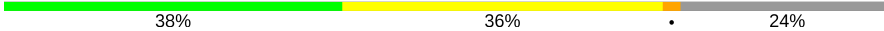
• Molecule 4: Histone H2A



• Molecule 4: Histone H2A




• Molecule 5: Histone H2B type 1-B

Chain D:  38% 36% 24%

GLU PRO SER LYS SER ALA PRO ALA PRO PRO LYS LYS GLY SER LYS LYS LYS ILE THR LYS ALA ALA GLN LYS LYS ASP GLY LYS LYS ARG ARG ARG ARG S32 R33 R34 E35 S36 Y37 Y38 I39 Y40 Y41 Y42 L45 K46 Q47 V48 H49 T52 G53 I54 I61 M62 F65 V66 M67 D68 I69

F70 E71 E76 A77 S78 F79 L80 M84 R85 R86 I94 Q95 T96 A97 V98 L101 L102 F103 LYS LYS G104 E105 A106 A107 K108 H109 A110 V111 S112 E113 G114 T115 K116 A117 V118 T119 S123 S124 K125

• Molecule 5: Histone H2B type 1-B

Chain H:  2% 43% 30% 24%

GLU PRO SER LYS SER ALA PRO ALA PRO PRO LYS LYS GLY SER LYS LYS LYS ILE THR LYS ALA ALA GLN LYS LYS ASP GLY LYS LYS ARG ARG ARG ARG S32 R33 R34 E35 S36 Y37 Y38 I39 Y40 Y41 Y42 K43 V44 L45 H49 T52 G53 I54 S55 I61 M62 V66 M67 D68 I69 F70

E71 R72 L80 Y83 S87 T88 E93 I94 Q95 T96 A97 R99 L100 L101 L102 E105 V111 S112 E113 G114 T115 V118 Y121 K125

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	104.43Å 109.18Å 175.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.24 – 3.25 37.24 – 3.21	Depositor EDS
% Data completeness (in resolution range)	99.6 (37.24-3.25) 92.7 (37.24-3.21)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.61 (at 3.18Å)	Xtriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, R_{free}	0.200 , 0.270 0.203 , 0.270	Depositor DCC
R_{free} test set	1920 reflections (5.85%)	wwPDB-VP
Wilson B-factor (Å ²)	83.9	Xtriage
Anisotropy	0.876	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 73.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.026 for k,h,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	12121	wwPDB-VP
Average B, all atoms (Å ²)	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.92% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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, CL

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	0.83	4/3354 (0.1%)	1.04	4/5175 (0.1%)
1	J	0.83	2/3354 (0.1%)	1.03	8/5175 (0.2%)
2	A	0.69	1/819 (0.1%)	1.03	5/1097 (0.5%)
2	E	0.92	2/828 (0.2%)	0.89	1/1109 (0.1%)
3	B	0.57	0/669	0.95	3/894 (0.3%)
3	F	0.74	0/711	0.97	2/948 (0.2%)
4	C	0.71	0/819	0.95	3/1107 (0.3%)
4	G	0.52	0/853	0.79	0/1150
5	D	0.71	0/747	0.98	1/1002 (0.1%)
5	H	0.59	0/746	0.79	1/1001 (0.1%)
All	All	0.77	9/12900 (0.1%)	0.99	28/18658 (0.2%)

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	A	0	2
3	B	0	1
5	D	0	1
All	All	0	4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	135	ALA	C-OXT	-15.88	0.93	1.23
1	I	68	DG	C3'-O3'	-6.06	1.36	1.44
2	E	110	CYS	CB-SG	-5.89	1.72	1.81
1	I	87	DG	C3'-O3'	-5.73	1.36	1.44
1	J	276	DT	C3'-O3'	-5.71	1.36	1.44

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	123	DT	O4'-C1'-N1	-9.85	101.11	108.00
2	A	39	HIS	N-CA-C	7.86	132.23	111.00
3	F	18	HIS	N-CA-C	7.54	131.34	111.00
2	A	104	PHE	N-CA-C	7.48	131.20	111.00
5	D	106	LEU	CA-CB-CG	7.32	132.13	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	103	LEU	Peptide
2	A	104	PHE	Peptide
3	B	35	ARG	Peptide
5	D	116	LYS	Peptide

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	I	2990	0	1652	107	0
1	J	2990	0	1652	104	0
2	A	807	0	844	80	0
2	E	816	0	856	60	0
3	B	662	0	709	53	0
3	F	703	0	755	46	0
4	C	809	0	861	64	0
4	G	843	0	905	49	0
5	D	736	0	758	53	0
5	H	735	0	755	49	0
6	E	1	0	0	1	0
6	G	1	0	0	0	0
6	I	2	0	0	0	0
7	E	1	0	0	2	0
7	I	3	0	0	0	0
7	J	2	0	0	1	0
8	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	B	1	0	0	0	0
8	C	4	0	0	1	0
8	E	1	0	0	0	0
8	F	1	0	0	0	0
8	G	2	0	0	1	0
8	I	4	0	0	0	0
8	J	6	0	0	1	0
All	All	12121	0	9747	512	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 24.

The worst 5 of 512 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)
2:E:128:ARG:NH2	2:E:135:ALA:HA	1.13	1.40
2:E:128:ARG:HH21	2:E:135:ALA:CA	1.36	1.39
5:D:32:SER:O	5:D:33:ARG:CG	1.72	1.36
4:C:35:ARG:NH2	4:C:36:GLN:NE2	1.74	1.35
3:F:18:HIS:O	3:F:19:ARG:HD3	1.43	1.15

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	
2	A	96/135 (71%)	93 (97%)	3 (3%)	0	100	100
2	E	97/135 (72%)	93 (96%)	3 (3%)	1 (1%)	15	47
3	B	81/102 (79%)	76 (94%)	5 (6%)	0	100	100
3	F	85/102 (83%)	82 (96%)	3 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	C	103/128 (80%)	100 (97%)	3 (3%)	0	100	100
4	G	107/128 (84%)	105 (98%)	2 (2%)	0	100	100
5	D	92/124 (74%)	87 (95%)	5 (5%)	0	100	100
5	H	92/124 (74%)	84 (91%)	7 (8%)	1 (1%)	14	46
All	All	753/978 (77%)	720 (96%)	31 (4%)	2 (0%)	41	72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	38	PRO
5	H	54	ILE

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	
2	A	85/110 (77%)	83 (98%)	2 (2%)	49	72
2	E	86/110 (78%)	85 (99%)	1 (1%)	71	83
3	B	68/78 (87%)	64 (94%)	4 (6%)	19	50
3	F	72/78 (92%)	70 (97%)	2 (3%)	43	69
4	C	83/101 (82%)	79 (95%)	4 (5%)	25	56
4	G	86/101 (85%)	83 (96%)	3 (4%)	36	64
5	D	81/105 (77%)	80 (99%)	1 (1%)	71	83
5	H	80/105 (76%)	76 (95%)	4 (5%)	24	55
All	All	641/788 (81%)	620 (97%)	21 (3%)	38	65

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	C	99	ARG
2	E	78	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	H	55	SER
4	C	62	VAL
5	H	71	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
5	D	84	ASN
5	D	109	HIS
2	E	76	GLN
4	C	84	GLN
2	E	39	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 10 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 [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, 95th 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	I	146/146 (100%)	0.01	4 (2%) 54 51	91, 143, 176, 183	0
1	J	146/146 (100%)	-0.08	4 (2%) 54 51	97, 143, 175, 182	0
2	A	98/135 (72%)	0.26	4 (4%) 37 34	68, 94, 128, 162	0
2	E	99/135 (73%)	0.05	4 (4%) 38 35	52, 70, 111, 145	0
3	B	83/102 (81%)	0.11	2 (2%) 59 55	70, 89, 119, 162	0
3	F	87/102 (85%)	-0.10	1 (1%) 80 80	53, 67, 91, 135	0
4	C	105/128 (82%)	-0.14	2 (1%) 66 64	51, 69, 98, 127	0
4	G	109/128 (85%)	-0.09	1 (0%) 84 84	68, 86, 128, 162	0
5	D	94/124 (75%)	0.01	0 100 100	43, 74, 108, 128	0
5	H	94/124 (75%)	0.21	3 (3%) 47 45	47, 86, 121, 152	0
All	All	1061/1270 (83%)	0.02	25 (2%) 59 55	43, 88, 162, 183	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	146	DT	7.1
2	A	38	PRO	4.6
2	A	81	ASP	4.2
2	E	135	ALA	4.0
4	C	14	VAL	3.9

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 [i](#)

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. The B-factors column lists the minimum, median, 95th 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	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	MN	G	201	1/1	0.48	0.24	111,111,111,111	0
7	CL	I	204	1/1	0.51	0.22	122,122,122,122	0
7	CL	J	301	1/1	0.64	0.24	142,142,142,142	0
6	MN	I	201	1/1	0.66	0.20	143,143,143,143	0
6	MN	I	202	1/1	0.72	0.10	177,177,177,177	0
7	CL	I	203	1/1	0.81	0.34	133,133,133,133	0
7	CL	J	302	1/1	0.84	0.11	100,100,100,100	0
7	CL	I	205	1/1	0.92	0.14	130,130,130,130	0
6	MN	E	201	1/1	0.96	0.17	80,80,80,80	0
7	CL	E	202	1/1	0.99	0.20	78,78,78,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.