



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 11:21 AM UTC

PDB ID : 1U35 / pdb_00001u35
Title : Crystal structure of the nucleosome core particle containing the histone domain of macroH2A
Authors : Chakravarthy, S.; Gundimella, S.K.; Caron, C.; Perche, P.Y.; Pehrson, J.R.; Khochbin, S.; Luger, K.
Deposited on : 2004-07-20
Resolution : 3.00 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 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

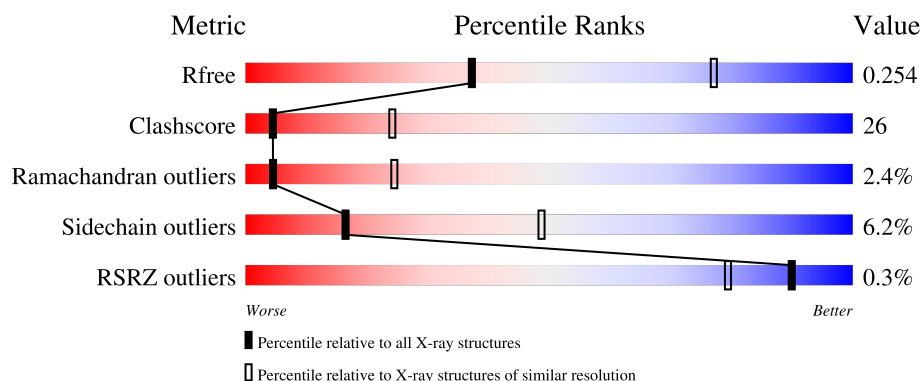
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	180053	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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 ≥ 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	32% 66% ..
1	J	146	37% 62% .
2	A	136	32% 37% . 28%
2	E	136	% 47% 23% .. 28%
3	B	103	42% 28% 6% . 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	103	% 54%23%19%
4	C	120	41%38%8%12%
4	G	120	46%33%8%12%
5	D	126	34%32%7%26%
5	H	126	% 36%38%24%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 12053 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 alpha-satellite DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	145	Total	C	N	O	P	0	0	0
			2970	1420	539	867	144			
1	J	145	Total	C	N	O	P	0	0	0
			2969	1420	536	869	144			

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

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

- Molecule 3 is a protein called Hist1h4i protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			
3	F	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			

- Molecule 4 is a protein called H2A histone family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	C	106	Total	C	N	O	S	0	0	0
			810	520	149	139	2			
4	G	106	Total	C	N	O	S	0	0	0
			810	520	149	139	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	867	VAL	GLY	engineered mutation	UNP O75367
G	1067	VAL	GLY	engineered mutation	UNP O75367

- Molecule 5 is a protein called histone 3, H2ba.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	93	Total	C	N	O	S	0	0	0
			731	459	131	139	2			
5	H	96	Total	C	N	O	S	0	0	0
			755	473	138	142	2			

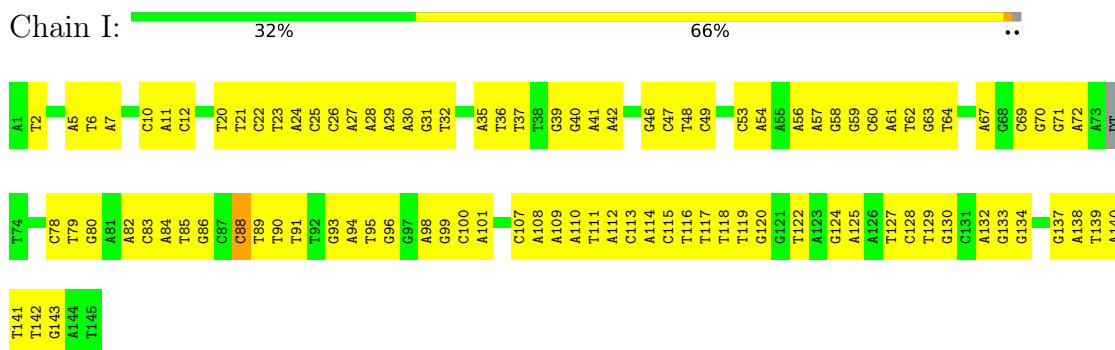
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	20	Total	O	0	0
			20	20		
6	J	20	Total	O	0	0
			20	20		
6	A	8	Total	O	0	0
			8	8		
6	B	6	Total	O	0	0
			6	6		
6	C	10	Total	O	0	0
			10	10		
6	D	9	Total	O	0	0
			9	9		
6	E	11	Total	O	0	0
			11	11		
6	F	13	Total	O	0	0
			13	13		
6	G	4	Total	O	0	0
			4	4		
6	H	4	Total	O	0	0
			4	4		

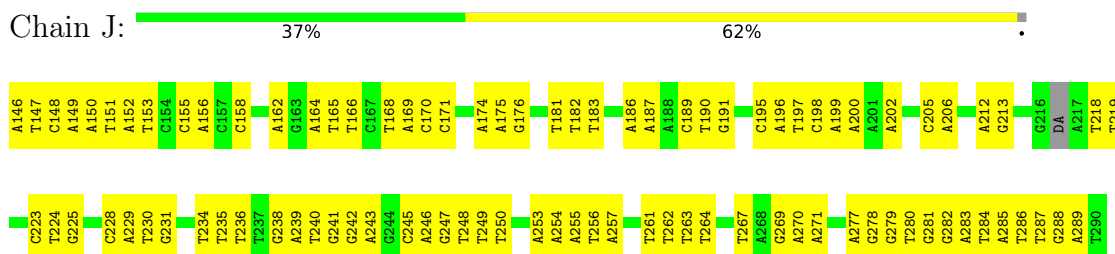
3 Residue-property plots

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.

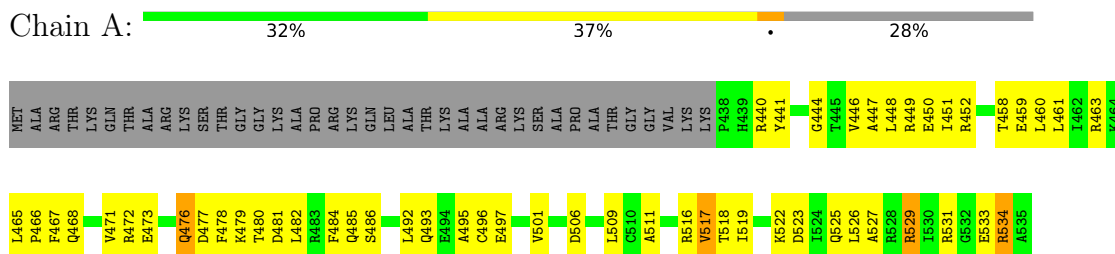
• Molecule 1: alpha-satellite DNA



• Molecule 1: alpha-satellite DNA

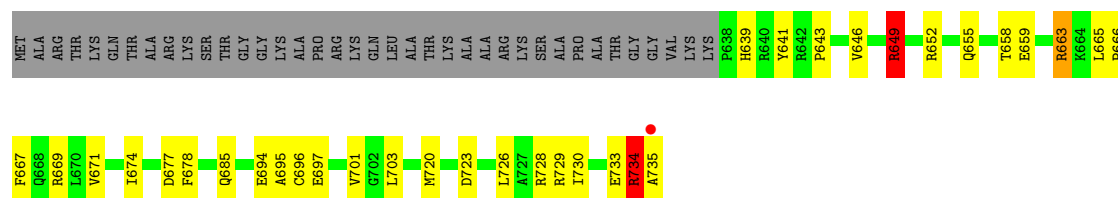


• Molecule 2: Histone H3.1

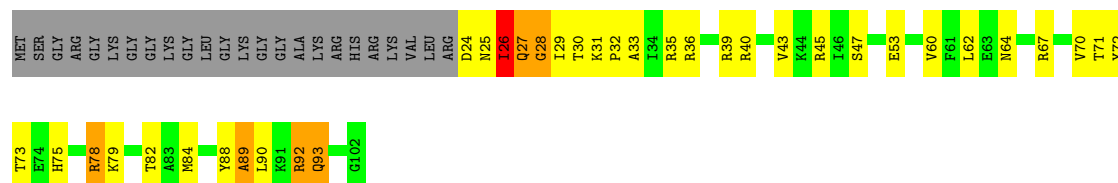


• Molecule 2: Histone H3.1

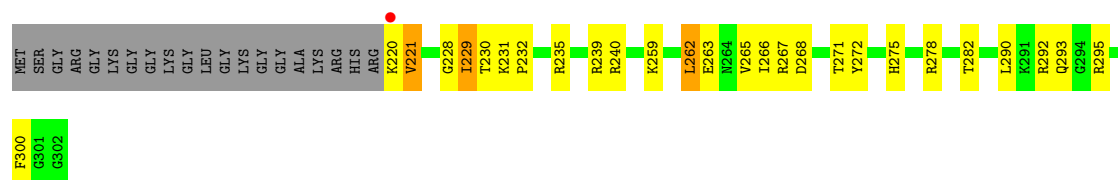




• Molecule 3: Hist1h4i protein



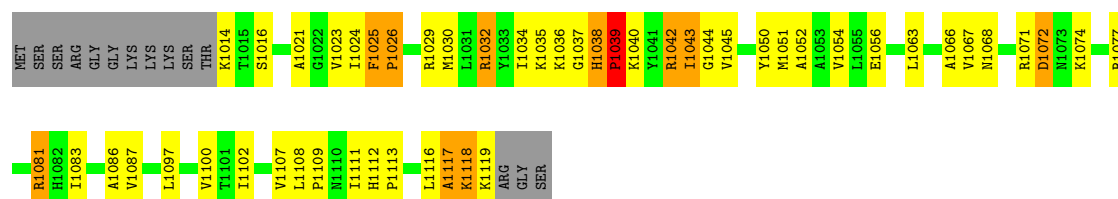
• Molecule 3: Hist1h4i protein



• Molecule 4: H2A histone family

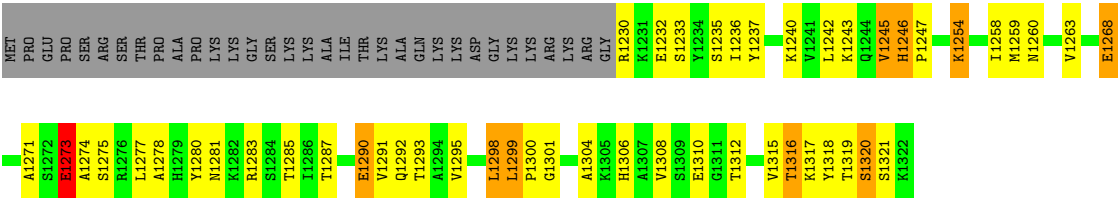


• Molecule 4: H2A histone family



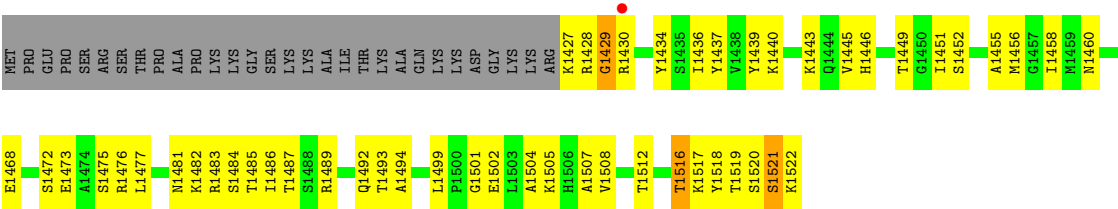
• Molecule 5: histone 3, H2ba

Chain D: 34% 32% 7% 26%



● Molecule 5: histone 3, H2ba

Chain H: 36% 38% 24%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.50Å 109.60Å 175.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.00 50.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	95.7 (50.00-3.00) 95.6 (50.00-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.43 (at 2.96Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.206 , 0.260 0.205 , 0.254	Depositor DCC
R_{free} test set	2169 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	78.6	Xtriage
Anisotropy	0.412	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 75.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12053	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% 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

5.1 Standard geometry

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	0.30	0/3332	0.64	0/5141
1	J	0.30	0/3330	0.66	0/5138
2	A	0.50	0/819	0.99	2/1097 (0.2%)
2	E	0.62	0/819	1.01	1/1097 (0.1%)
3	B	0.56	0/634	1.23	7/848 (0.8%)
3	F	0.60	0/669	1.01	2/894 (0.2%)
4	C	0.70	1/823 (0.1%)	1.11	6/1111 (0.5%)
4	G	0.57	0/823	1.20	9/1111 (0.8%)
5	D	0.57	0/742	1.15	7/996 (0.7%)
5	H	0.50	0/766	0.90	0/1026
All	All	0.45	1/12757 (0.0%)	0.87	34/18459 (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
1	I	0	2
1	J	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	840	LYS	CE-NZ	6.11	1.67	1.49

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	27	GLN	OE1-CD-NE2	-9.22	113.38	122.60
3	B	28	GLY	N-CA-C	-8.72	105.08	114.67
4	G	1025	PHE	CA-C-N	8.31	130.23	119.84

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	G	1025	PHE	C-N-CA	8.31	130.23	119.84
5	D	1298	LEU	N-CA-C	8.05	121.16	111.82

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	67	DA	Sidechain
1	I	88	DC	Sidechain
1	J	212	DA	Sidechain

5.2 Too-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 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	2970	0	1639	132	0
1	J	2969	0	1640	107	0
2	A	807	0	844	62	0
2	E	807	0	844	43	1
3	B	627	0	663	53	0
3	F	662	0	709	30	0
4	C	810	0	873	99	0
4	G	810	0	873	89	0
5	D	731	0	753	53	0
5	H	755	0	782	53	0
6	A	8	0	0	1	0
6	B	6	0	0	1	0
6	C	10	0	0	2	0
6	D	9	0	0	1	1
6	E	11	0	0	2	0
6	F	13	0	0	1	0
6	G	4	0	0	1	0
6	H	4	0	0	1	0
6	I	20	0	0	1	0
6	J	20	0	0	2	0
All	All	12053	0	9620	564	1

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 26.

The worst 5 of 564 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)
4:C:840:LYS:CE	4:C:840:LYS:NZ	1.67	1.57
3:B:26:ILE:CG2	3:B:27:GLN:H	1.25	1.41
4:C:840:LYS:HG3	4:G:1039:PRO:CB	1.64	1.25
3:B:26:ILE:HG23	3:B:27:GLN:N	1.34	1.17
4:C:835:LYS:O	4:C:839:PRO:HD3	1.43	1.15

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:677:ASP:OD2	6:D:301:HOH:O[3_745]	1.96	0.24

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	96/136 (71%)	92 (96%)	1 (1%)	3 (3%)	3	19
2	E	96/136 (71%)	92 (96%)	3 (3%)	1 (1%)	12	45
3	B	77/103 (75%)	72 (94%)	4 (5%)	1 (1%)	9	38
3	F	81/103 (79%)	78 (96%)	2 (2%)	1 (1%)	10	40
4	C	104/120 (87%)	93 (89%)	8 (8%)	3 (3%)	3	20
4	G	104/120 (87%)	94 (90%)	7 (7%)	3 (3%)	3	20
5	D	91/126 (72%)	79 (87%)	10 (11%)	2 (2%)	5	26
5	H	94/126 (75%)	79 (84%)	11 (12%)	4 (4%)	2	12
All	All	743/970 (77%)	679 (91%)	46 (6%)	18 (2%)	4	24

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	534	ARG
3	B	26	ILE
4	C	840	LYS
5	D	1320	SER
2	E	734	ARG

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	85/111 (77%)	81 (95%)	4 (5%)	23	58
2	E	85/111 (77%)	82 (96%)	3 (4%)	32	65
3	B	64/79 (81%)	62 (97%)	2 (3%)	35	68
3	F	68/79 (86%)	65 (96%)	3 (4%)	25	60
4	C	84/95 (88%)	78 (93%)	6 (7%)	13	43
4	G	84/95 (88%)	77 (92%)	7 (8%)	10	37
5	D	81/108 (75%)	72 (89%)	9 (11%)	6	25
5	H	83/108 (77%)	78 (94%)	5 (6%)	17	50
All	All	634/786 (81%)	595 (94%)	39 (6%)	16	49

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	G	1039	PRO
5	H	1484	SER
4	G	1045	VAL
4	G	1074	LYS
5	H	1512	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
5	H	1492	GLN
5	H	1460	ASN
3	F	293	GLN
5	H	1446	HIS
3	F	275	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 oligosaccharides 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, 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	145/146 (99%)	0.17	0	100 100	57, 128, 181, 200	0
1	J	145/146 (99%)	0.11	0	100 100	60, 127, 168, 198	0
2	A	98/136 (72%)	-0.58	0	100 100	39, 66, 124, 151	0
2	E	98/136 (72%)	-0.67	1 (1%)	79 59	36, 52, 80, 169	0
3	B	79/103 (76%)	-0.50	0	100 100	46, 68, 104, 156	0
3	F	83/103 (80%)	-0.69	1 (1%)	76 55	34, 51, 75, 151	0
4	C	106/120 (88%)	-0.57	0	100 100	38, 60, 124, 195	0
4	G	106/120 (88%)	-0.46	0	100 100	50, 70, 119, 162	0
5	D	93/126 (73%)	-0.67	0	100 100	44, 64, 104, 150	0
5	H	96/126 (76%)	-0.54	1 (1%)	79 59	43, 72, 131, 159	0
All	All	1049/1262 (83%)	-0.38	3 (0%)	90 79	34, 71, 156, 200	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	H	1430	ARG	3.7
2	E	735	ALA	2.9
3	F	220	LYS	2.0

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 oligosaccharides 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.