



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:28 PM GMT

PDB ID : 1P3I  
Title : Crystallographic Studies of Nucleosome Core Particles containing Histone 'Sin' Mutants  
Authors : Muthurajan, U.M.; Bao, Y.; Forsberg, L.J.; Edayathumangalam, R.S.; Dyer, P.N.; White, C.L.; Luger, K.  
Deposited on : 2003-04-17  
Resolution : 2.30 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  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

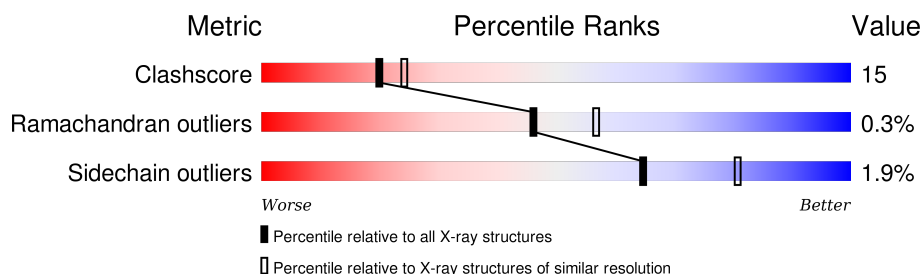
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)




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. 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.

Note EDS was not executed.

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	129	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	G	129	
5	D	125	
5	H	125	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12260 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 Palindromic 146bp Human Alpha-Satellite 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	0	0
			808	509	156	140	3			
2	E	104	Total	C	N	O	S	0	0	0
			852	535	164	150	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	434	GLU	GLY	CONFLICT	UNP Q7ZT64
A	435	SER	VAL	CONFLICT	UNP Q7ZT64
A	502	ALA	GLY	CONFLICT	UNP Q7ZT64
E	634	GLU	GLY	CONFLICT	UNP Q7ZT64
E	635	SER	VAL	CONFLICT	UNP Q7ZT64
E	702	ALA	GLY	CONFLICT	UNP Q7ZT64

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	79	Total	C	N	O	S	0	0	0
			626	395	120	110	1			
3	F	79	Total	C	N	O	S	0	0	0
			626	395	120	110	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	45	HIS	ARG	CONFLICT	UNP P62799
F	245	HIS	ARG	CONFLICT	UNP P62799

- Molecule 4 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	C	103	Total	C	N	O	0	0	0
			795	501	155	139			
4	G	104	Total	C	N	O	0	0	0
			804	507	157	140			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	814	ALA	SER	CONFLICT	UNP Q7ZT66
C	867	GLY	TRP	CONFLICT	UNP Q7ZT66
C	868	ASN	GLU	CONFLICT	UNP Q7ZT66
C	869	ALA	ARG	CONFLICT	UNP Q7ZT66
C	870	ALA	LEU	CONFLICT	UNP Q7ZT66
C	871	ARG	PRO	CONFLICT	UNP Q7ZT66
C	872	ASP	GLU	CONFLICT	UNP Q7ZT66
C	873	ASN	ILE	CONFLICT	UNP Q7ZT66
C	874	LYS	TRP	CONFLICT	UNP Q7ZT66
C	876	THR	ARG	CONFLICT	UNP Q7ZT66
C	877	ARG	PRO	CONFLICT	UNP Q7ZT66
C	878	ILE	VAL	CONFLICT	UNP Q7ZT66
C	879	ILE	LEU	CONFLICT	UNP Q7ZT66
C	880	PRO	SER	CONFLICT	UNP Q7ZT66
C	881	ARG	PRO	CONFLICT	UNP Q7ZT66
C	882	HIS	GLY	CONFLICT	UNP Q7ZT66
C	883	LEU	TRP	CONFLICT	UNP Q7ZT66
C	884	GLN	CYS	CONFLICT	UNP Q7ZT66
C	885	LEU	ASN	CONFLICT	UNP Q7ZT66
C	886	ALA	SER	CONFLICT	UNP Q7ZT66
C	887	VAL	LEU	CONFLICT	UNP Q7ZT66
C	888	ARG	CYS	CONFLICT	UNP Q7ZT66
C	923	ALA	SER	CONFLICT	UNP Q7ZT66
C	926	ALA	THR	CONFLICT	UNP Q7ZT66
G	1014	ALA	SER	CONFLICT	UNP Q7ZT66
G	1067	GLY	TRP	CONFLICT	UNP Q7ZT66
G	1068	ASN	GLU	CONFLICT	UNP Q7ZT66
G	1069	ALA	ARG	CONFLICT	UNP Q7ZT66
G	1070	ALA	LEU	CONFLICT	UNP Q7ZT66

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	1071	ARG	PRO	CONFLICT	UNP Q7ZT66
G	1072	ASP	GLU	CONFLICT	UNP Q7ZT66
G	1073	ASN	ILE	CONFLICT	UNP Q7ZT66
G	1074	LYS	TRP	CONFLICT	UNP Q7ZT66
G	1076	THR	ARG	CONFLICT	UNP Q7ZT66
G	1077	ARG	PRO	CONFLICT	UNP Q7ZT66
G	1078	ILE	VAL	CONFLICT	UNP Q7ZT66
G	1079	ILE	LEU	CONFLICT	UNP Q7ZT66
G	1080	PRO	SER	CONFLICT	UNP Q7ZT66
G	1081	ARG	PRO	CONFLICT	UNP Q7ZT66
G	1082	HIS	GLY	CONFLICT	UNP Q7ZT66
G	1083	LEU	TRP	CONFLICT	UNP Q7ZT66
G	1084	GLN	CYS	CONFLICT	UNP Q7ZT66
G	1085	LEU	ASN	CONFLICT	UNP Q7ZT66
G	1086	ALA	SER	CONFLICT	UNP Q7ZT66
G	1087	VAL	LEU	CONFLICT	UNP Q7ZT66
G	1088	ARG	CYS	CONFLICT	UNP Q7ZT66
G	1123	ALA	SER	CONFLICT	UNP Q7ZT66
G	1126	ALA	THR	CONFLICT	UNP Q7ZT66

- Molecule 5 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	95	Total	C	N	O	S	0	0	0
			744	468	134	140	2			
5	H	94	Total	C	N	O	S	0	0	0
			735	462	132	139	2			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1219	GLN	PRO	CONFLICT	UNP P02281
D	1242	LEU	MET	CONFLICT	UNP P02281
D	1257	SER	GLY	CONFLICT	UNP P02281
D	1266	VAL	ILE	CONFLICT	UNP P02281
H	1419	GLN	PRO	CONFLICT	UNP P02281
H	1442	LEU	MET	CONFLICT	UNP P02281
H	1457	SER	GLY	CONFLICT	UNP P02281
H	1466	VAL	ILE	CONFLICT	UNP P02281

- Molecule 6 is water.

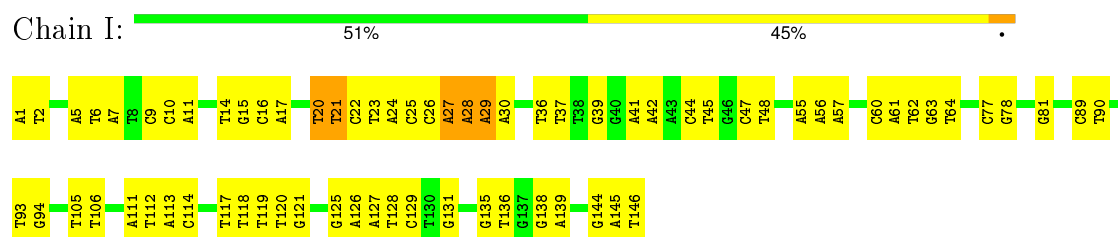
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	27	Total 27	O 27	0	0
6	B	17	Total 17	O 17	0	0
6	C	30	Total 30	O 30	0	0
6	D	21	Total 21	O 21	0	0
6	E	41	Total 41	O 41	0	0
6	F	30	Total 30	O 30	0	0
6	G	27	Total 27	O 27	0	0
6	H	9	Total 9	O 9	0	0
6	I	41	Total 41	O 41	0	0
6	J	47	Total 47	O 47	0	0

### 3 Residue-property plots

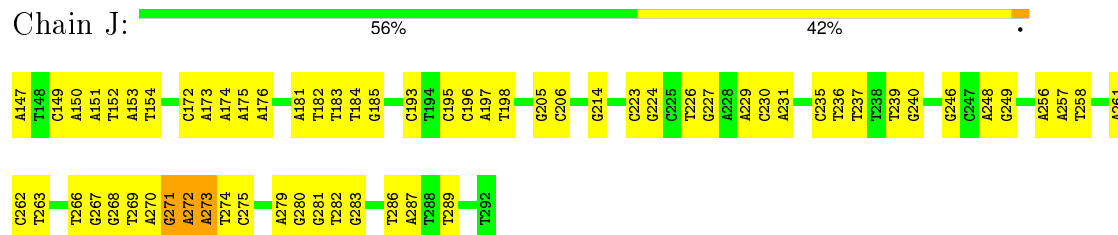
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 ( $\text{RSRZ} > 2$ ). 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.

Note EDS was not executed.

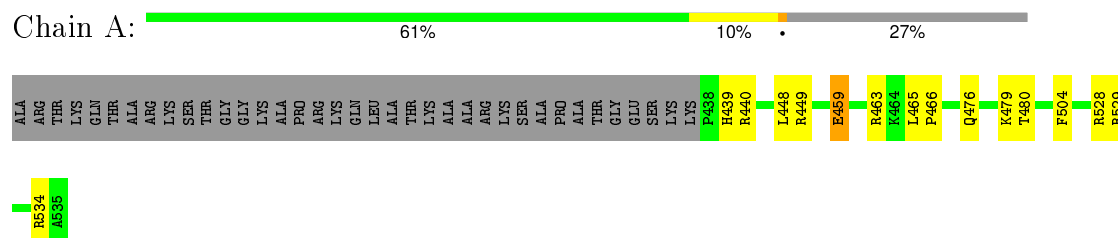
- Molecule 1: Palindromic 146bp Human Alpha-Satellite DNA fragment



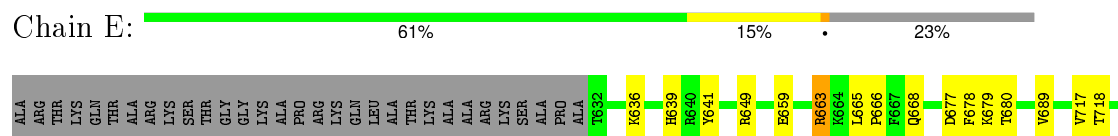
- Molecule 1: Palindromic 146bp Human Alpha-Satellite DNA fragment



- Molecule 2: Histone H3



- Molecule 2: Histone H3

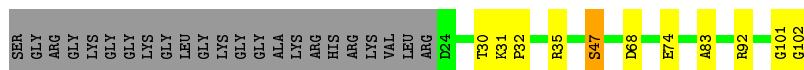






• Molecule 3: Histone H4

Chain B: 67% 10% 23%



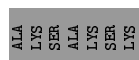
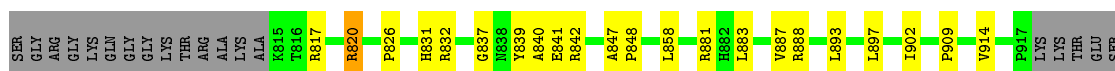
• Molecule 3: Histone H4

Chain F: 66% 11% 23%



• Molecule 4: Histone H2A

Chain C: 63% 16% 20%



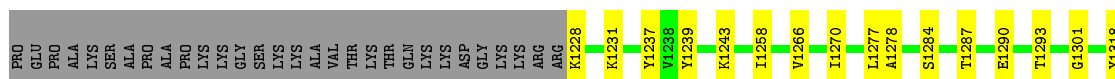
• Molecule 4: Histone H2A

Chain G: 65% 15% 19%



• Molecule 5: Histone H2B

Chain D: 62% 14% 24%



• Molecule 5: Histone H2B

Chain H: 64% 10% 25%

PRO	GLU	PRO	ALA	LYS	SER	ALA	PRO	ALA	PRO	LYS	LYS	GLY	SER	LYS	LYS	ALA	VAL	THR	LYS	THR	GLN	LYS	LYS	ASP	GLY	LYS	LYS	ARG	ARG	LYS	S1429	R1430	K1431	E1432	S1433	Q1444	T1449	I1458	E1473	L1477	R1489	L1498	L1499	P1500	G1501	E1502	L1503	E1510	K1522
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.61Å 109.75Å 182.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	100.00 – 2.30	Depositor
% Data completeness (in resolution range)	95.0 (100.00-2.30)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.242 , 0.275	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12260	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

## 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	0.56	3/3354 (0.1%)	1.09	24/5175 (0.5%)
1	J	0.55	2/3354 (0.1%)	0.86	8/5175 (0.2%)
2	A	0.71	4/820 (0.5%)	0.72	2/1099 (0.2%)
2	E	0.68	0/864	0.73	0/1157
3	B	0.60	0/634	0.70	0/849
3	F	0.80	2/634 (0.3%)	0.97	4/849 (0.5%)
4	C	0.62	0/805	0.71	0/1088
4	G	0.52	0/814	0.64	0/1099
5	D	0.63	0/755	0.67	0/1013
5	H	0.53	0/746	0.65	0/1002
All	All	0.60	11/12780 (0.1%)	0.88	38/18506 (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	1
1	J	0	1
3	F	0	1
All	All	0	3

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	272	DA	C5-C6	-12.14	1.30	1.41
1	I	21	DT	C2-O2	-9.22	1.15	1.22
2	A	504	PHE	CD1-CE1	-6.83	1.25	1.39
1	I	21	DT	N1-C2	-6.59	1.32	1.38
2	A	504	PHE	CB-CG	5.68	1.61	1.51

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	28	DA	O5'-P-OP1	-26.42	78.99	110.70
1	J	273	DA	O5'-P-OP2	-23.07	83.01	110.70
1	I	28	DA	OP2-P-O3'	-22.68	55.30	105.20
1	I	28	DA	OP1-P-O3'	20.47	150.24	105.20
1	I	28	DA	O5'-P-OP2	-16.13	91.18	105.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	F	275	HIS	Sidechain
1	I	131	DG	Sidechain
1	J	214	DG	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	2990	0	1651	111	0
1	J	2990	0	1651	93	0
2	A	808	0	846	17	0
2	E	852	0	892	24	0
3	B	626	0	657	9	0
3	F	626	0	657	16	0
4	C	795	0	846	36	0
4	G	804	0	859	26	0
5	D	744	0	771	22	0
5	H	735	0	758	15	0
6	A	27	0	0	0	0
6	B	17	0	0	1	0
6	C	30	0	0	6	0
6	D	21	0	0	1	0
6	E	41	0	0	8	0
6	F	30	0	0	1	0
6	G	27	0	0	4	0
6	H	9	0	0	3	0
6	I	41	0	0	4	0
6	J	47	0	0	8	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	12260	0	9588	316	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 15.

The worst 5 of 316 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:I:21:DT:O2	1:J:272:DA:N1	1.58	1.36
1:I:21:DT:H3	1:J:272:DA:N6	1.22	1.32
4:C:820:ARG:HB3	4:C:820:ARG:NH1	1.53	1.21
1:J:261:DA:H2"	1:J:262:DC:H5"	1.25	1.17
4:C:820:ARG:HH11	4:C:820:ARG:CB	1.62	1.13

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%)	94 (98%)	2 (2%)	0	100	100
2	E	102/135 (76%)	98 (96%)	4 (4%)	0	100	100
3	B	77/102 (76%)	77 (100%)	0	0	100	100
3	F	77/102 (76%)	76 (99%)	1 (1%)	0	100	100
4	C	101/129 (78%)	99 (98%)	2 (2%)	0	100	100
4	G	102/129 (79%)	99 (97%)	3 (3%)	0	100	100
5	D	93/125 (74%)	91 (98%)	1 (1%)	1 (1%)	17	18
5	H	92/125 (74%)	88 (96%)	3 (3%)	1 (1%)	17	18
All	All	740/982 (75%)	722 (98%)	16 (2%)	2 (0%)	46	57

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	H	1501	GLY
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	85/111 (77%)	84 (99%)	1 (1%)	78	89
2	E	90/111 (81%)	88 (98%)	2 (2%)	60	77
3	B	64/78 (82%)	63 (98%)	1 (2%)	70	84
3	F	64/78 (82%)	64 (100%)	0	100	100
4	C	82/100 (82%)	79 (96%)	3 (4%)	41	55
4	G	83/100 (83%)	82 (99%)	1 (1%)	78	89
5	D	81/105 (77%)	81 (100%)	0	100	100
5	H	80/105 (76%)	76 (95%)	4 (5%)	30	41
All	All	629/788 (80%)	617 (98%)	12 (2%)	65	81

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	659	GLU
2	E	663	ARG
5	H	1433	SER
4	C	909	PRO
5	H	1431	LYS

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	1292	GLN
4	G	1031	HIS
4	G	1073	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
5	D	1279	HIS
4	G	1038	ASN

### 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](#)

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](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.