



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 08:47 PM GMT

PDB ID : 4Z66  
Title : Nucleosome disassembly by RSC and SWI/SNF is enhanced by H3 acetylation near the nucleosome dyad axis  
Authors : Dechassa, M.L.; Luger, K.; Chatterjee, N.; North, J.A.; Manohar, M.; Prasad, R.; Ottessen, J.J.; Poirier, M.G.; Bartholomew, B.  
Deposited on : 2015-04-03  
Resolution : 2.50 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20026982

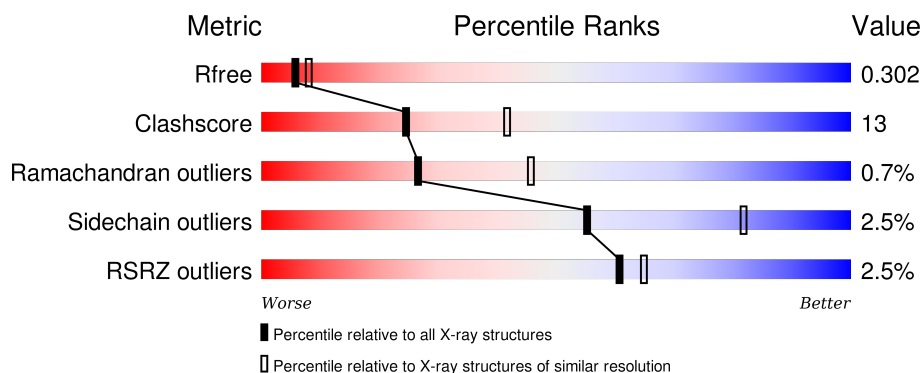
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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.

Mol	Chain	Length	Quality of chain
1	A	98	<div> <div></div> <div>78%22%</div> </div>
1	E	98	<div> <div>2%</div> <div>78%21%</div> <div>.</div> </div>
2	B	82	<div> <div>2%</div> <div>80%16%</div> <div>.</div> </div>
2	F	82	<div> <div>%</div> <div>70%30%</div> </div>
3	C	107	<div> <div>3%</div> <div>78%17%5%</div> <div>.</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	107	
4	D	94	
4	H	94	
5	I	147	
6	J	147	

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
1	ALY	A	522	-	-	X	-
1	ALY	E	722	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12214 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 Histone H3.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	S	0	0	0
			811	511	156	141	3			
1	E	98	Total	C	N	O	S	0	0	0
			811	511	156	141	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	502	ALA	GLY	conflict	UNP P84233
E	702	ALA	GLY	conflict	UNP P84233

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			
2	F	82	Total	C	N	O	S	0	0	0
			653	412	127	113	1			

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	107	Total	C	N	O		0	0	0
			826	520	161	145				
3	G	106	Total	C	N	O		0	0	0
			818	516	160	142				

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	94	Total	C	N	O	S	0	0	0
			736	463	132	139	2			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	94	Total	C	N	O	S	0	0	0
			736	463	132	139	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1229	THR	SER	conflict	UNP P02281
H	1429	THR	SER	conflict	UNP P02281

- Molecule 5 is a DNA chain called DNA (147-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	147	Total	C	N	O	P	0	0	0
			3011	1440	546	879	146			

- Molecule 6 is a DNA chain called DNA (147-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	147	Total	C	N	O	P	0	0	0
			3010	1440	543	881	146			

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	16	Total	O	0	0
			16	16		
7	B	12	Total	O	0	0
			12	12		
7	C	26	Total	O	0	0
			26	26		
7	D	13	Total	O	0	0
			13	13		
7	E	35	Total	O	0	0
			35	35		
7	F	20	Total	O	0	0
			20	20		
7	G	14	Total	O	0	0
			14	14		
7	H	6	Total	O	0	0
			6	6		
7	I	17	Total	O	0	0
			17	17		

*Continued on next page...*


*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	J	16	Total	O	0	0
			16	16		

### 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 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: Histone H3.2

Chain A: 




#### • Molecule 1: Histone H3.2

Chain E: 



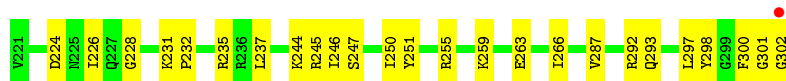
#### • Molecule 2: Histone H4

Chain B: 




#### • Molecule 2: Histone H4

Chain F: 




#### • Molecule 3: Histone H2A

Chain C: 

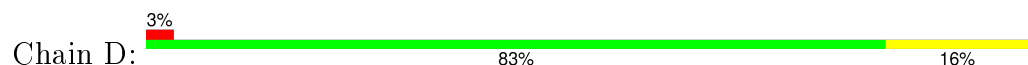


#### • Molecule 3: Histone H2A

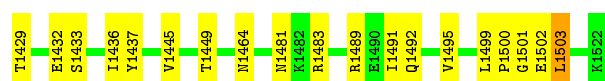
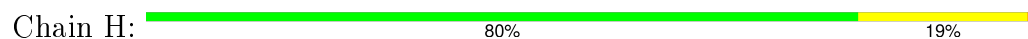
Chain G: 



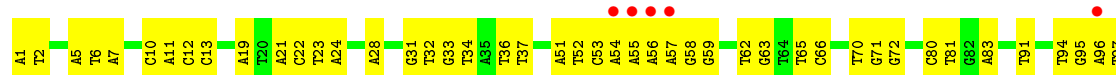
• Molecule 4: Histone H2B 1.1



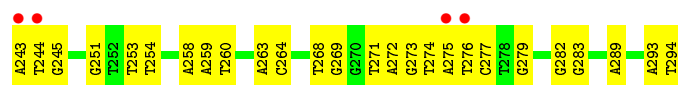
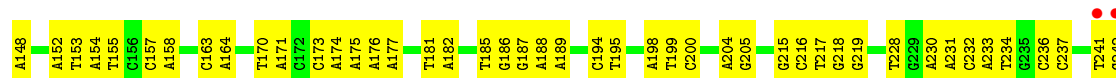
• Molecule 4: Histone H2B 1.1



• Molecule 5: DNA (147-MER)



• Molecule 6: DNA (147-MER)





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.73Å 109.63Å 181.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.87 – 2.50 42.87 – 2.40	Depositor EDS
% Data completeness (in resolution range)	89.7 (42.87-2.50) 94.7 (42.87-2.40)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 2.39Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.254 , 0.296 0.259 , 0.302	Depositor DCC
$R_{free}$ test set	2037 reflections (2.94%)	DCC
Wilson B-factor (Å <sup>2</sup> )	47.7	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 53.5	EDS
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	2 of 148178 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12214	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: ALY

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	0.38	0/810	0.61	0/1085
1	E	0.44	0/810	0.65	0/1085
2	B	0.42	0/634	0.65	0/848
2	F	0.48	0/660	0.75	1/883 (0.1%)
3	C	0.40	0/836	0.63	0/1127
3	G	0.36	0/828	0.61	0/1117
4	D	0.42	0/747	0.64	0/1004
4	H	0.38	0/747	0.58	0/1004
5	I	0.35	0/3378	0.69	0/5212
6	J	0.34	0/3376	0.71	0/5209
All	All	0.38	0/12826	0.67	1/18574 (0.0%)

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
5	I	0	1
6	J	0	2
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	301	GLY	N-CA-C	-5.41	99.56	113.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	I	134	DA	Sidechain
6	J	148	DA	Sidechain
6	J	215	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	A	811	0	847	20	0
1	E	811	0	847	30	0
2	B	627	0	663	15	0
2	F	653	0	696	26	0
3	C	826	0	884	34	0
3	G	818	0	877	26	0
4	D	736	0	760	13	0
4	H	736	0	760	14	0
5	I	3011	0	1662	73	0
6	J	3010	0	1663	74	0
7	A	16	0	0	4	0
7	B	12	0	0	2	0
7	C	26	0	0	3	0
7	D	13	0	0	0	0
7	E	35	0	0	13	0
7	F	20	0	0	3	0
7	G	14	0	0	0	0
7	H	6	0	0	3	0
7	I	17	0	0	1	1
7	J	16	0	0	3	1
All	All	12214	0	9659	276	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 13.

The worst 5 of 276 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:125:DA:H2"	5:I:126:DG:H5"	1.30	1.13

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:I:146:DA:H2''	5:I:147:DT:H5''	1.33	1.08
6:J:273:DG:H2''	6:J:274:DT:H5''	1.37	1.06
5:I:58:DG:H4'	5:I:59:DG:H5'	1.40	1.02
1:A:522:ALY:HG2	7:A:601:HOH:O	1.62	0.98

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)
7:I:208:HOH:O	7:J:305:HOH:O[3_744]	2.00	0.20

## 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	
1	A	95/98 (97%)	94 (99%)	1 (1%)	0	100	100
1	E	95/98 (97%)	93 (98%)	2 (2%)	0	100	100
2	B	77/82 (94%)	76 (99%)	1 (1%)	0	100	100
2	F	80/82 (98%)	80 (100%)	0	0	100	100
3	C	105/107 (98%)	99 (94%)	5 (5%)	1 (1%)	19	34
3	G	104/107 (97%)	99 (95%)	4 (4%)	1 (1%)	19	34
4	D	92/94 (98%)	90 (98%)	0	2 (2%)	8	13
4	H	92/94 (98%)	89 (97%)	2 (2%)	1 (1%)	17	31
All	All	740/762 (97%)	720 (97%)	15 (2%)	5 (1%)	26	46

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	918	LYS

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
4	D	1230	ARG
4	D	1301	GLY
4	H	1501	GLY
3	G	1072	ASP

### 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	
1	A	84/84 (100%)	83 (99%)	1 (1%)	78	93
1	E	84/84 (100%)	83 (99%)	1 (1%)	78	93
2	B	64/67 (96%)	64 (100%)	0	100	100
2	F	67/67 (100%)	64 (96%)	3 (4%)	34	59
3	C	85/85 (100%)	79 (93%)	6 (7%)	18	34
3	G	84/85 (99%)	82 (98%)	2 (2%)	57	82
4	D	80/80 (100%)	80 (100%)	0	100	100
4	H	80/80 (100%)	77 (96%)	3 (4%)	40	67
All	All	628/632 (99%)	612 (98%)	16 (2%)	55	82

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	663	ARG
2	F	224	ASP
3	G	1088	ARG
3	C	919	LYS
4	H	1433	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	1292	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	H	1506	HIS
3	G	1031	HIS
1	A	476	GLN
3	G	1112	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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
1	ALY	A	522	1	9,11,12	2.06	2 (22%)	10,12,14	2.16	4 (40%)
1	ALY	E	722	1	9,11,12	1.75	2 (22%)	10,12,14	1.88	2 (20%)

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
1	ALY	A	522	1	-	0/8/10/12	0/0/0/0
1	ALY	E	722	1	-	0/8/10/12	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	522	ALY	CH-NZ	2.59	1.40	1.33
1	E	722	ALY	CE-NZ	3.12	1.53	1.46
1	E	722	ALY	CH-NZ	3.13	1.41	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	522	ALY	CE-NZ	4.33	1.56	1.46

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	722	ALY	CE-NZ-CH	-5.49	112.96	122.38
1	A	522	ALY	CD-CE-NZ	-4.51	99.52	112.29
1	A	522	ALY	CE-NZ-CH	-3.47	116.41	122.38
1	A	522	ALY	CG-CD-CE	-2.60	102.55	113.69
1	A	522	ALY	O-C-CA	-2.12	120.03	125.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	522	ALY	6	0
1	E	722	ALY	10	0

## 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](#)

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	A	97/98 (98%)	0.25	0 <b>100</b> <b>100</b>	25, 49, 72, 118	0
1	E	97/98 (98%)	0.35	2 (2%) 67 71	25, 40, 67, 145	0
2	B	79/82 (96%)	0.20	2 (2%) 61 65	35, 45, 69, 129	0
2	F	82/82 (100%)	0.31	1 (1%) 81 83	22, 39, 55, 115	0
3	C	107/107 (100%)	0.25	3 (2%) 56 61	25, 43, 74, 168	0
3	G	106/107 (99%)	0.18	3 (2%) 56 61	35, 51, 88, 163	0
4	D	94/94 (100%)	0.18	3 (3%) 51 56	22, 46, 69, 134	0
4	H	94/94 (100%)	0.24	0 <b>100</b> <b>100</b>	22, 51, 80, 144	0
5	I	147/147 (100%)	0.22	6 (4%) 41 46	47, 82, 119, 152	0
6	J	147/147 (100%)	0.12	6 (4%) 41 46	45, 83, 118, 145	0
All	All	1050/1056 (99%)	0.22	26 (2%) 61 65	22, 51, 105, 168	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	920	THR	10.5
5	I	56	DA	6.6
5	I	55	DA	5.1
6	J	242	DG	5.1
3	G	1119	LYS	4.8

### 6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	ALY	A	522	12/13	0.79	0.21	-	21,24,26,30	0
1	ALY	E	722	12/13	0.79	0.28	-	21,24,26,30	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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.