



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 02:10 PM UTC

PDB ID : 4NFT / pdb\_00004nft  
Title : Crystal structure of human lnkH2B-h2A.Z-Anp32e  
Authors : Shan, S.; Pan, L.; Mao, Z.; Wang, W.; Sun, J.; Dong, Q.; Liang, X.; Ding, X.;  
Chen, S.; Dai, L.; Zhang, Z.; Zhu, B.; Zhou, Z.  
Deposited on : 2013-11-01  
Resolution : 2.61 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

<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  
Xtrriage (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

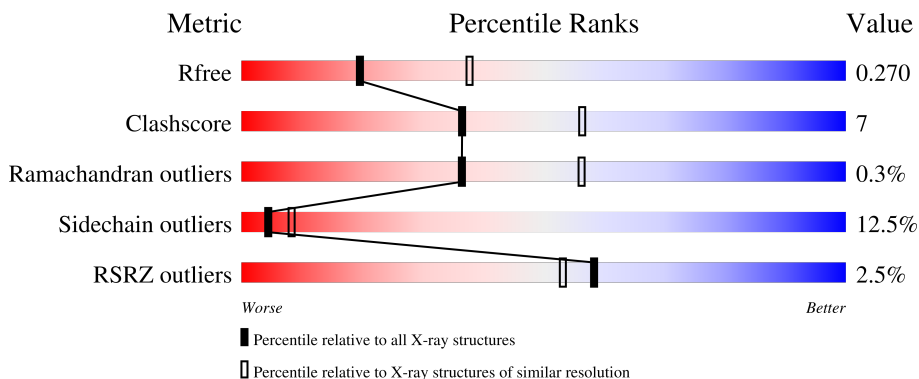
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.61 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	4951 (2.64-2.60)
Clashscore	190562	5303 (2.64-2.60)
Ramachandran outliers	187476	5217 (2.64-2.60)
Sidechain outliers	187428	5217 (2.64-2.60)
RSRZ outliers	180081	4950 (2.64-2.60)

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  $\geq 3$ , 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	A	209	 3% 69% 15% 12%
1	B	209	 3% 68% 16% 12%
1	C	209	 67% 14% 15%
1	D	209	 69% 10% 5% 15%
2	E	52	 2% 10% 6% 6% 79%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	52	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a small red segment (2%), a green segment (12%), a yellow segment (8%), and a large grey segment (79%). The percentages are labeled below the bar.</p>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5932 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 H2B type 2-E, Histone H2A.Z.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	184	1433	898	269	263	3	0	1	0
1	B	184	1433	898	269	263	3	0	1	0
1	C	178	1371	859	255	255	2	0	0	0
1	D	178	1371	859	255	255	2	0	0	0

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q16778
A	-1	SER	-	expression tag	UNP Q16778
A	0	HIS	-	expression tag	UNP Q16778
A	1	MET	-	expression tag	UNP Q16778
A	194	SER	-	expression tag	UNP P0C0S5
A	195	GLY	-	expression tag	UNP P0C0S5
A	196	SER	-	expression tag	UNP P0C0S5
A	197	GLY	-	expression tag	UNP P0C0S5
A	198	GLY	-	expression tag	UNP P0C0S5
A	199	SER	-	expression tag	UNP P0C0S5
A	200	GLY	-	expression tag	UNP P0C0S5
A	201	GLY	-	expression tag	UNP P0C0S5
A	202	GLY	-	expression tag	UNP P0C0S5
A	203	LEU	-	expression tag	UNP P0C0S5
A	204	VAL	-	expression tag	UNP P0C0S5
A	205	PRO	-	expression tag	UNP P0C0S5
A	206	ARG	-	expression tag	UNP P0C0S5
B	-2	GLY	-	expression tag	UNP Q16778
B	-1	SER	-	expression tag	UNP Q16778
B	0	HIS	-	expression tag	UNP Q16778
B	1	MET	-	expression tag	UNP Q16778

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	194	SER	-	expression tag	UNP P0C0S5
B	195	GLY	-	expression tag	UNP P0C0S5
B	196	SER	-	expression tag	UNP P0C0S5
B	197	GLY	-	expression tag	UNP P0C0S5
B	198	GLY	-	expression tag	UNP P0C0S5
B	199	SER	-	expression tag	UNP P0C0S5
B	200	GLY	-	expression tag	UNP P0C0S5
B	201	GLY	-	expression tag	UNP P0C0S5
B	202	GLY	-	expression tag	UNP P0C0S5
B	203	LEU	-	expression tag	UNP P0C0S5
B	204	VAL	-	expression tag	UNP P0C0S5
B	205	PRO	-	expression tag	UNP P0C0S5
B	206	ARG	-	expression tag	UNP P0C0S5
C	-2	GLY	-	expression tag	UNP Q16778
C	-1	SER	-	expression tag	UNP Q16778
C	0	HIS	-	expression tag	UNP Q16778
C	1	MET	-	expression tag	UNP Q16778
C	194	SER	-	expression tag	UNP P0C0S5
C	195	GLY	-	expression tag	UNP P0C0S5
C	196	SER	-	expression tag	UNP P0C0S5
C	197	GLY	-	expression tag	UNP P0C0S5
C	198	GLY	-	expression tag	UNP P0C0S5
C	199	SER	-	expression tag	UNP P0C0S5
C	200	GLY	-	expression tag	UNP P0C0S5
C	201	GLY	-	expression tag	UNP P0C0S5
C	202	GLY	-	expression tag	UNP P0C0S5
C	203	LEU	-	expression tag	UNP P0C0S5
C	204	VAL	-	expression tag	UNP P0C0S5
C	205	PRO	-	expression tag	UNP P0C0S5
C	206	ARG	-	expression tag	UNP P0C0S5
D	-2	GLY	-	expression tag	UNP Q16778
D	-1	SER	-	expression tag	UNP Q16778
D	0	HIS	-	expression tag	UNP Q16778
D	1	MET	-	expression tag	UNP Q16778
D	194	SER	-	expression tag	UNP P0C0S5
D	195	GLY	-	expression tag	UNP P0C0S5
D	196	SER	-	expression tag	UNP P0C0S5
D	197	GLY	-	expression tag	UNP P0C0S5
D	198	GLY	-	expression tag	UNP P0C0S5
D	199	SER	-	expression tag	UNP P0C0S5
D	200	GLY	-	expression tag	UNP P0C0S5
D	201	GLY	-	expression tag	UNP P0C0S5

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	202	GLY	-	expression tag	UNP P0C0S5
D	203	LEU	-	expression tag	UNP P0C0S5
D	204	VAL	-	expression tag	UNP P0C0S5
D	205	PRO	-	expression tag	UNP P0C0S5
D	206	ARG	-	expression tag	UNP P0C0S5

- Molecule 2 is a protein called Acidic leucine-rich nuclear phosphoprotein 32 family member E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	11	Total	C	N	O	S	0	0	0
			89	57	12	19	1			
2	F	11	Total	C	N	O	S	0	0	0
			89	57	12	19	1			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-27	GLY	-	expression tag	UNP Q9BTT0
E	-26	SER	-	expression tag	UNP Q9BTT0
E	-25	HIS	-	expression tag	UNP Q9BTT0
E	-24	MET	-	expression tag	UNP Q9BTT0
F	-27	GLY	-	expression tag	UNP Q9BTT0
F	-26	SER	-	expression tag	UNP Q9BTT0
F	-25	HIS	-	expression tag	UNP Q9BTT0
F	-24	MET	-	expression tag	UNP Q9BTT0

- Molecule 3 is water.

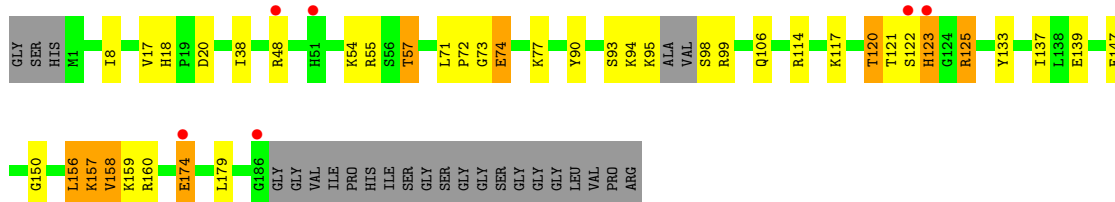
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	38	Total	O	0	0
			38	38		
3	B	36	Total	O	0	0
			36	36		
3	C	32	Total	O	0	0
			32	32		
3	D	34	Total	O	0	0
			34	34		
3	E	3	Total	O	0	0
			3	3		
3	F	3	Total	O	0	0
			3	3		

### 3 Residue-property plots [i](#)

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.

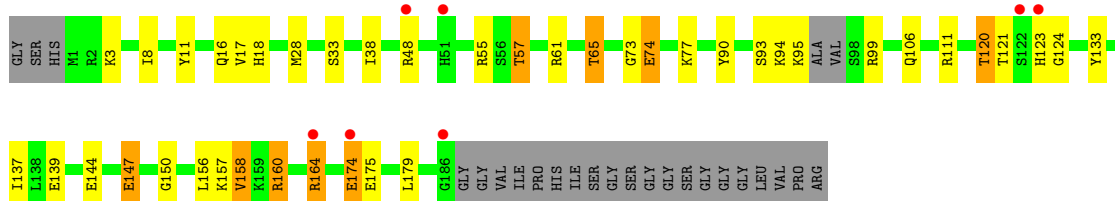
- Molecule 1: Histone H2B type 2-E, Histone H2A.Z

Chain A: 3% 69% 15% 12%



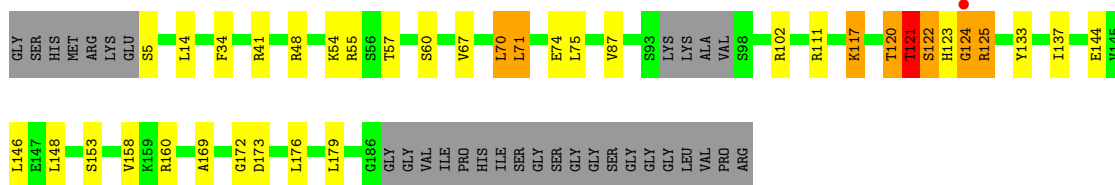
- Molecule 1: Histone H2B type 2-E, Histone H2A.Z

Chain B: 3% 68% 16% 12%



- Molecule 1: Histone H2B type 2-E, Histone H2A.Z

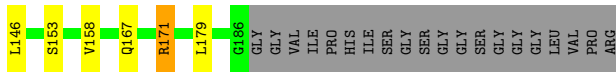
Chain C: 67% 14% 15%



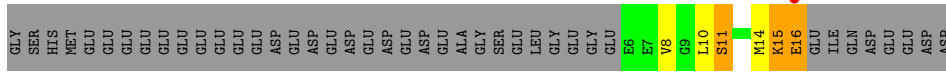
- Molecule 1: Histone H2B type 2-E, Histone H2A.Z

Chain D: 69% 10% 5% 15%

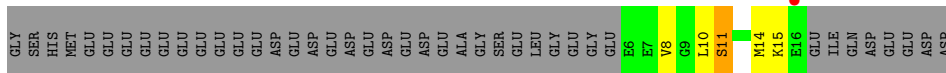




- Molecule 2: Acidic leucine-rich nuclear phosphoprotein 32 family member E



- Molecule 2: Acidic leucine-rich nuclear phosphoprotein 32 family member E



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.41Å 104.30Å 124.74Å 90.00° 99.15° 90.00°	Depositor
Resolution (Å)	39.80 – 2.61 39.80 – 2.61	Depositor EDS
% Data completeness (in resolution range)	98.0 (39.80-2.61) 99.0 (39.80-2.61)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.72 (at 2.61Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, $R_{free}$	0.220 , 0.273 0.223 , 0.270	Depositor DCC
$R_{free}$ test set	1493 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.8	Xtrriage
Anisotropy	0.657	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 40.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5932	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 95.87 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.4087e-10. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.74	1/1456 (0.1%)	0.95	4/1956 (0.2%)
1	B	0.76	1/1456 (0.1%)	0.95	1/1956 (0.1%)
1	C	0.65	2/1390 (0.1%)	0.89	1/1872 (0.1%)
1	D	0.69	0/1390	0.90	3/1872 (0.2%)
2	E	1.03	0/89	1.05	0/117
2	F	0.87	0/89	1.01	0/117
All	All	0.72	4/5870 (0.1%)	0.93	9/7890 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	124	GLY	C-O	-5.56	1.17	1.23
1	A	125	ARG	C-O	-5.24	1.17	1.23
1	B	144	GLU	C-O	-5.14	1.18	1.24
1	C	125	ARG	CA-C	-5.04	1.46	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	123	HIS	N-CA-C	-10.02	97.30	110.43
1	B	124	GLY	N-CA-C	7.91	122.34	112.14
1	D	121	THR	N-CA-C	-7.27	101.15	110.53
1	D	92	SER	CB-CA-C	-6.42	100.16	112.43
1	D	123	HIS	N-CA-C	-5.81	105.81	114.64

There are no chirality outliers.

There are no planarity outliers.

### 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	A	1433	0	1497	28	0
1	B	1433	0	1497	22	0
1	C	1371	0	1420	20	0
1	D	1371	0	1420	17	0
2	E	89	0	87	4	0
2	F	89	0	87	2	0
3	A	38	0	0	1	0
3	B	36	0	0	0	0
3	C	32	0	0	1	0
3	D	34	0	0	0	0
3	E	3	0	0	0	0
3	F	3	0	0	0	0
All	All	5932	0	6008	85	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 7.

The worst 5 of 85 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:B:174:GLU:OE1	1:B:175:GLU:HG2	1.09	1.26
1:B:94:LYS:O	1:B:95:LYS:HB3	1.50	1.06
1:B:174:GLU:OE1	1:B:175:GLU:CG	2.05	1.04
1:A:121:THR:O	1:A:123:HIS:O	1.80	0.99
1:A:94:LYS:O	1:A:95:LYS:HG3	1.71	0.89

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	
1	A	181/209 (87%)	178 (98%)	2 (1%)	1 (1%)	21	39
1	B	181/209 (87%)	177 (98%)	3 (2%)	1 (1%)	21	39
1	C	174/209 (83%)	168 (97%)	6 (3%)	0	100	100
1	D	174/209 (83%)	169 (97%)	5 (3%)	0	100	100
2	E	9/52 (17%)	9 (100%)	0	0	100	100
2	F	9/52 (17%)	9 (100%)	0	0	100	100
All	All	728/940 (77%)	710 (98%)	16 (2%)	2 (0%)	36	56

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	73	GLY
1	B	73	GLY

### 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	
1	A	153/167 (92%)	139 (91%)	14 (9%)	8	18
1	B	153/167 (92%)	136 (89%)	17 (11%)	6	11
1	C	146/167 (87%)	129 (88%)	17 (12%)	5	10
1	D	146/167 (87%)	124 (85%)	22 (15%)	3	5
2	E	10/46 (22%)	6 (60%)	4 (40%)	0	0
2	F	10/46 (22%)	7 (70%)	3 (30%)	0	0
All	All	618/760 (81%)	541 (88%)	77 (12%)	4	8

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	74	GLU
2	E	11	SER
1	D	102	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	123	HIS
2	F	11	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	106	GLN
1	B	106	GLN

### 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, 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	184/209 (88%)	-0.09	6 (3%) 49 44	18, 29, 51, 76	1 (0%)
1	B	184/209 (88%)	-0.02	7 (3%) 44 39	18, 29, 51, 76	1 (0%)
1	C	178/209 (85%)	-0.21	1 (0%) 85 83	16, 28, 48, 64	0
1	D	178/209 (85%)	-0.19	3 (1%) 69 65	16, 28, 48, 64	0
2	E	11/52 (21%)	0.13	1 (9%) 15 11	30, 38, 57, 60	0
2	F	11/52 (21%)	0.17	1 (9%) 15 11	29, 39, 58, 60	0
All	All	746/940 (79%)	-0.12	19 (2%) 58 53	16, 29, 51, 76	2 (0%)

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	174	GLU	6.2
1	A	174	GLU	6.1
1	B	186	GLY	5.8
1	A	123	HIS	3.9
1	B	164	ARG	3.6

### 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.