



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2024 – 08:21 PM EDT

PDB ID : 5YRQ  
Title : Crystal structure of Rad5 and Rev1  
Authors : Chen, Z.C.; Zhou, C.Y.  
Deposited on : 2017-11-09  
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

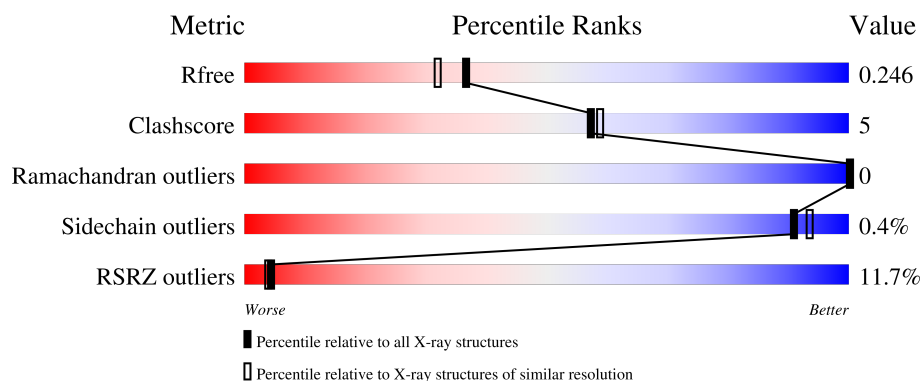
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	148	<div> <div>6%</div> <div> <div></div> <div>78%</div> <div>6%</div> <div>16%</div> </div> </div>
1	B	148	<div> <div>14%</div> <div> <div></div> <div>71%</div> <div>10%</div> <div>19%</div> </div> </div>
1	D	148	<div> <div>12%</div> <div> <div></div> <div>70%</div> <div>14%</div> <div>17%</div> </div> </div>
1	E	148	<div> <div>7%</div> <div> <div></div> <div>70%</div> <div>12%</div> <div>18%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4229 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 DNA repair protein RAD5,DNA repair protein REV1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	125	Total	C	N	O	S	0	0	0
			1039	664	187	184	4			
1	B	120	Total	C	N	O	S	0	0	0
			999	639	177	178	5			
1	D	123	Total	C	N	O	S	0	0	0
			1023	652	184	182	5			
1	E	121	Total	C	N	O	S	0	0	0
			1014	651	183	176	4			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	-	expression tag	UNP P32849
A	3	HIS	-	expression tag	UNP P32849
A	4	MET	-	expression tag	UNP P32849
A	857	GLY	-	linker	UNP P32849
A	858	GLY	-	linker	UNP P32849
A	859	SER	-	linker	UNP P32849
A	860	SER	-	linker	UNP P32849
A	861	SER	-	linker	UNP P32849
A	862	SER	-	linker	UNP P32849
A	863	LEU	-	linker	UNP P32849
A	864	VAL	-	linker	UNP P32849
A	865	PRO	-	linker	UNP P32849
A	866	ARG	-	linker	UNP P32849
A	867	GLY	-	linker	UNP P32849
A	868	SER	-	linker	UNP P32849
A	869	GLY	-	linker	UNP P32849
A	870	GLY	-	linker	UNP P32849
A	871	SER	-	linker	UNP P32849
A	872	GLY	-	linker	UNP P32849
A	873	GLY	-	linker	UNP P32849
A	874	SER	-	linker	UNP P32849

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	875	PRO	-	linker	UNP P32849
B	2	ALA	-	expression tag	UNP P32849
B	3	HIS	-	expression tag	UNP P32849
B	4	MET	-	expression tag	UNP P32849
B	857	GLY	-	linker	UNP P32849
B	858	GLY	-	linker	UNP P32849
B	859	SER	-	linker	UNP P32849
B	860	SER	-	linker	UNP P32849
B	861	SER	-	linker	UNP P32849
B	862	SER	-	linker	UNP P32849
B	863	LEU	-	linker	UNP P32849
B	864	VAL	-	linker	UNP P32849
B	865	PRO	-	linker	UNP P32849
B	866	ARG	-	linker	UNP P32849
B	867	GLY	-	linker	UNP P32849
B	868	SER	-	linker	UNP P32849
B	869	GLY	-	linker	UNP P32849
B	870	GLY	-	linker	UNP P32849
B	871	SER	-	linker	UNP P32849
B	872	GLY	-	linker	UNP P32849
B	873	GLY	-	linker	UNP P32849
B	874	SER	-	linker	UNP P32849
B	875	PRO	-	linker	UNP P32849
D	2	ALA	-	expression tag	UNP P32849
D	3	HIS	-	expression tag	UNP P32849
D	4	MET	-	expression tag	UNP P32849
D	857	GLY	-	linker	UNP P32849
D	858	GLY	-	linker	UNP P32849
D	859	SER	-	linker	UNP P32849
D	860	SER	-	linker	UNP P32849
D	861	SER	-	linker	UNP P32849
D	862	SER	-	linker	UNP P32849
D	863	LEU	-	linker	UNP P32849
D	864	VAL	-	linker	UNP P32849
D	865	PRO	-	linker	UNP P32849
D	866	ARG	-	linker	UNP P32849
D	867	GLY	-	linker	UNP P32849
D	868	SER	-	linker	UNP P32849
D	869	GLY	-	linker	UNP P32849
D	870	GLY	-	linker	UNP P32849
D	871	SER	-	linker	UNP P32849
D	872	GLY	-	linker	UNP P32849

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	873	GLY	-	linker	UNP P32849
D	874	SER	-	linker	UNP P32849
D	875	PRO	-	linker	UNP P32849
E	2	ALA	-	expression tag	UNP P32849
E	3	HIS	-	expression tag	UNP P32849
E	4	MET	-	expression tag	UNP P32849
E	857	GLY	-	linker	UNP P32849
E	858	GLY	-	linker	UNP P32849
E	859	SER	-	linker	UNP P32849
E	860	SER	-	linker	UNP P32849
E	861	SER	-	linker	UNP P32849
E	862	SER	-	linker	UNP P32849
E	863	LEU	-	linker	UNP P32849
E	864	VAL	-	linker	UNP P32849
E	865	PRO	-	linker	UNP P32849
E	866	ARG	-	linker	UNP P32849
E	867	GLY	-	linker	UNP P32849
E	868	SER	-	linker	UNP P32849
E	869	GLY	-	linker	UNP P32849
E	870	GLY	-	linker	UNP P32849
E	871	SER	-	linker	UNP P32849
E	872	GLY	-	linker	UNP P32849
E	873	GLY	-	linker	UNP P32849
E	874	SER	-	linker	UNP P32849
E	875	PRO	-	linker	UNP P32849

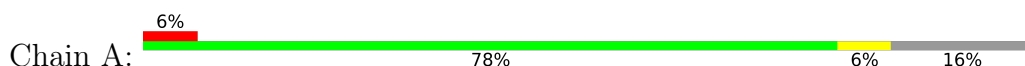
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	37	Total O 37 37	0	0
2	B	43	Total O 43 43	0	0
2	D	30	Total O 30 30	0	0
2	E	44	Total O 44 44	0	0

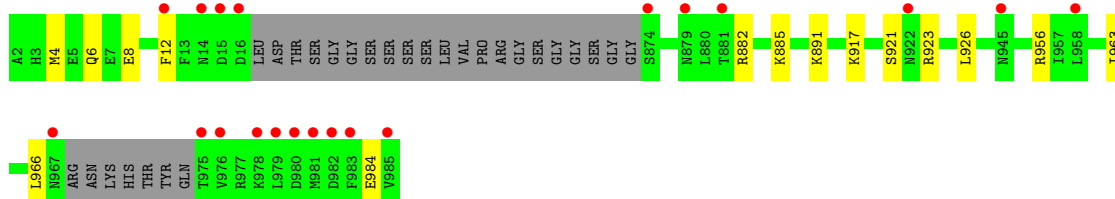
### 3 Residue-property plots [i](#)

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.

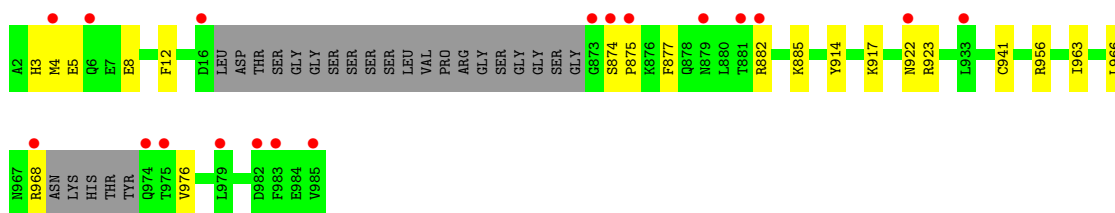
- Molecule 1: DNA repair protein RAD5,DNA repair protein REV1



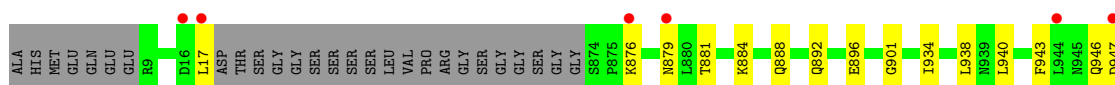
- Molecule 1: DNA repair protein RAD5,DNA repair protein REV1

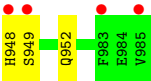


- Molecule 1: DNA repair protein RAD5,DNA repair protein REV1



- Molecule 1: DNA repair protein RAD5,DNA repair protein REV1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.59Å 49.28Å 103.07Å 90.00° 92.20° 90.00°	Depositor
Resolution (Å)	29.13 – 2.00 29.13 – 2.00	Depositor EDS
% Data completeness (in resolution range)	95.5 (29.13-2.00) 95.5 (29.13-2.00)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.03 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.205 , 0.246 0.206 , 0.246	Depositor DCC
$R_{free}$ test set	1911 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.7	Xtriage
Anisotropy	0.713	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 48.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.169 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4229	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.99% 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, 0.333 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	0.39	0/1059	0.53	0/1425
1	B	0.36	0/1017	0.50	0/1366
1	D	0.37	1/1041 (0.1%)	0.48	0/1397
1	E	0.36	0/1034	0.52	0/1391
All	All	0.37	1/4151 (0.0%)	0.51	0/5579

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	941	CYS	CB-SG	-5.34	1.73	1.81

There are no bond angle outliers.

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	1039	0	1051	5	0
1	B	999	0	1004	12	0
1	D	1023	0	1028	15	0
1	E	1014	0	1032	14	0
2	A	37	0	0	1	1
2	B	43	0	0	1	2
2	D	30	0	0	1	0
2	E	44	0	0	2	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4229	0	4115	40	4

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 5.

All (40) 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:956:ARG:HE	1:D:956:ARG:HE	1.16	0.93
1:B:891:LYS:NZ	2:B:1001:HOH:O	2.05	0.89
1:B:921:SER:OG	1:B:923:ARG:NH1	2.20	0.74
1:B:923:ARG:HB3	1:B:926:LEU:HD13	1.70	0.74
1:D:963:ILE:HA	1:D:966:LEU:HD23	1.71	0.70
1:B:956:ARG:HE	1:D:956:ARG:NE	1.89	0.69
1:D:968:ARG:NH2	2:D:1001:HOH:O	2.26	0.66
1:D:874:SER:HA	1:D:923:ARG:HH12	1.61	0.65
1:E:876:LYS:HG2	1:E:881:THR:HG22	1.83	0.60
1:E:940:LEU:HA	2:E:1021:HOH:O	2.03	0.59
1:E:17:LEU:HD12	1:E:879:ASN:HD21	1.69	0.57
1:B:4:MET:HG3	1:D:8:GLU:HG2	1.87	0.57
1:B:8:GLU:HG2	1:D:4:MET:HG3	1.89	0.55
1:A:876:LYS:HG2	1:A:881:THR:HG22	1.89	0.55
1:E:943:PHE:HD2	2:E:1021:HOH:O	1.89	0.54
1:E:901:GLY:HA3	1:E:948:HIS:HB2	1.89	0.54
1:D:923:ARG:HE	1:D:976:VAL:HG21	1.74	0.52
1:B:963:ILE:HA	1:B:966:LEU:HD23	1.92	0.52
1:B:882:ARG:HE	1:B:885:LYS:NZ	2.08	0.52
1:A:892:GLN:O	1:A:896:GLU:HG3	2.12	0.50
1:E:901:GLY:HA2	1:E:947:ASP:O	2.13	0.49
1:A:963:ILE:HA	1:A:966:LEU:HD23	1.94	0.49
1:D:922:ASN:HB2	1:D:976:VAL:HG23	1.96	0.48
1:B:984:GLU:HG2	1:E:946:GLN:NE2	2.29	0.47
1:D:874:SER:CA	1:D:923:ARG:HH12	2.26	0.46
1:E:892:GLN:O	1:E:896:GLU:HG3	2.16	0.46
2:A:1019:HOH:O	1:E:884:LYS:HB3	2.16	0.45
1:E:876:LYS:HZ2	1:E:881:THR:HG22	1.83	0.44
1:D:3:HIS:NE2	1:D:5:GLU:OE1	2.42	0.44
1:D:875:PRO:HB3	1:D:917:LYS:HE2	1.98	0.44
1:E:17:LEU:CD1	1:E:879:ASN:HD21	2.31	0.44
1:B:6:GLN:HG2	1:D:8:GLU:HG3	2.00	0.43
1:E:884:LYS:O	1:E:888:GLN:HG3	2.19	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:934:ILE:O	1:E:938:LEU:HG	2.18	0.43
1:D:882:ARG:HH21	1:D:885:LYS:HE3	1.83	0.43
1:A:901:GLY:HA3	1:A:948:HIS:HB2	2.00	0.42
1:E:949:SER:HB3	1:E:952:GLN:CD	2.40	0.41
1:D:877:PHE:HB2	1:D:914:TYR:CD1	2.55	0.41
1:B:917:LYS:HB3	1:B:917:LYS:HE2	1.66	0.41
1:A:901:GLY:HA2	1:A:947:ASP:O	2.21	0.40

All (4) 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:A:1019:HOH:O	2:E:1021:HOH:O[2_646]	2.00	0.20
2:B:1002:HOH:O	2:E:1013:HOH:O[2_646]	2.07	0.13
2:E:1024:HOH:O	2:E:1040:HOH:O[2_646]	2.13	0.07
2:B:1013:HOH:O	2:E:1028:HOH:O[2_646]	2.14	0.06

## 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	121/148 (82%)	120 (99%)	1 (1%)	0	100	100
1	B	114/148 (77%)	113 (99%)	1 (1%)	0	100	100
1	D	117/148 (79%)	116 (99%)	1 (1%)	0	100	100
1	E	117/148 (79%)	116 (99%)	1 (1%)	0	100	100
All	All	469/592 (79%)	465 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	118/134 (88%)	118 (100%)	0	100	100
1	B	113/134 (84%)	112 (99%)	1 (1%)	78	83
1	D	115/134 (86%)	114 (99%)	1 (1%)	78	83
1	E	115/134 (86%)	115 (100%)	0	100	100
All	All	461/536 (86%)	459 (100%)	2 (0%)	91	93

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	12	PHE
1	D	12	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	948	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 monosaccharides in this entry.

## 5.6 Ligand geometry

There are no ligands in this entry.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	125/148 (84%)	0.70	9 (7%) 15 14	12, 23, 42, 61	0
1	B	120/148 (81%)	1.16	20 (16%) 1 1	15, 30, 59, 86	0
1	D	123/148 (83%)	1.07	18 (14%) 2 2	11, 30, 57, 77	0
1	E	121/148 (81%)	0.72	10 (8%) 11 10	12, 24, 43, 65	0
All	All	489/592 (82%)	0.91	57 (11%) 4 4	11, 26, 54, 86	0

All (57) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	948	HIS	9.4
1	E	948	HIS	8.4
1	D	873	GLY	7.0
1	E	17	LEU	6.6
1	B	985	VAL	6.5
1	B	983	PHE	6.0
1	B	979	LEU	5.9
1	D	974	GLN	5.5
1	D	16	ASP	5.1
1	B	922	ASN	5.0
1	B	16	ASP	5.0
1	D	985	VAL	4.7
1	B	975	THR	4.3
1	A	947	ASP	3.9
1	D	979	LEU	3.9
1	D	983	PHE	3.9
1	B	14	ASN	3.8
1	B	879	ASN	3.8
1	D	922	ASN	3.8
1	B	874	SER	3.7
1	D	968	ARG	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	981	MET	3.5
1	A	983	PHE	3.4
1	D	874	SER	3.3
1	B	982	ASP	3.3
1	B	978	LYS	3.2
1	B	15	ASP	3.2
1	D	879	ASN	3.2
1	E	16	ASP	3.1
1	B	12	PHE	3.0
1	D	982	ASP	2.8
1	E	947	ASP	2.8
1	A	949	SER	2.7
1	D	975	THR	2.7
1	D	4	MET	2.7
1	E	985	VAL	2.6
1	B	967	ASN	2.6
1	E	983	PHE	2.6
1	B	980	ASP	2.5
1	D	882	ARG	2.4
1	A	946	GLN	2.4
1	B	945	ASN	2.4
1	B	976	VAL	2.3
1	B	881	THR	2.3
1	E	944	LEU	2.3
1	E	876	LYS	2.3
1	D	6	GLN	2.3
1	D	881	THR	2.3
1	A	9	ARG	2.3
1	D	875	PRO	2.2
1	A	956	ARG	2.2
1	E	879	ASN	2.1
1	A	985	VAL	2.1
1	E	949	SER	2.1
1	B	958	LEU	2.1
1	A	879	ASN	2.0
1	D	933	LEU	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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.