



# Full wwPDB X-ray Structure Validation Report ⓘ

May 19, 2020 – 01:26 am BST

PDB ID : 3E9P  
Title : Crystal Structure of Yeast Prp8, Residues 1827-2092  
Authors : Pena, V.; Rozov, A.; Wahl, M.C.  
Deposited on : 2008-08-22  
Resolution : 2.10 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 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 : 2.11  
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) : 2.11

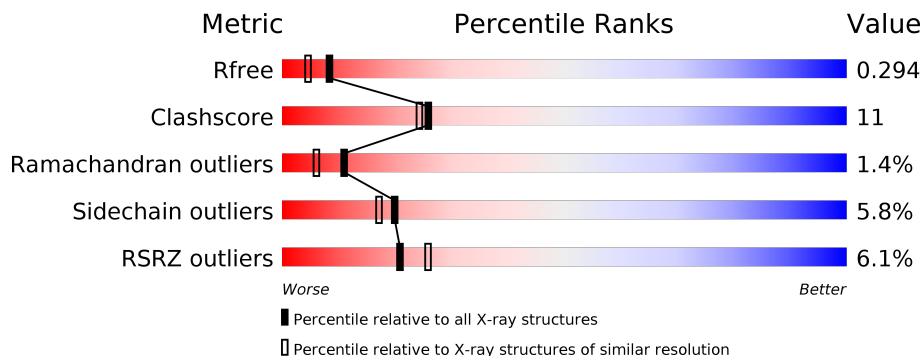
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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	255	
1	B	255	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4477 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 Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	255	Total 2110	C 1356	N 354	O 391	S 9	0	6	0
1	B	255	Total 2107	C 1354	N 354	O 391	S 8	0	5	0

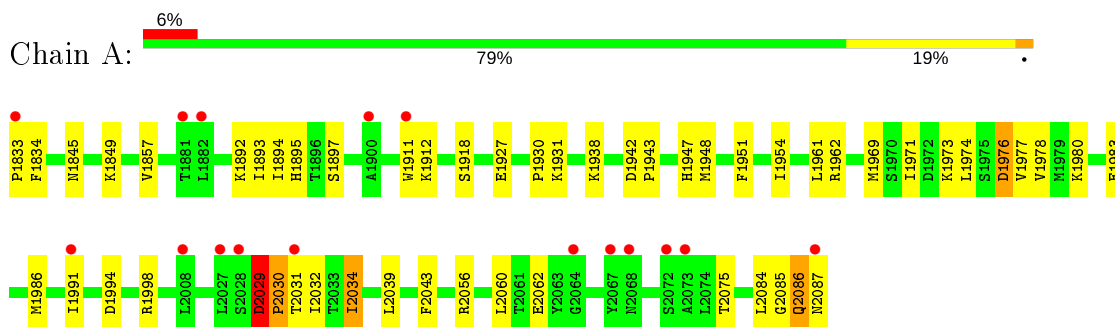
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	162	Total 163	O 163	0	2
2	B	97	Total 97	O 97	0	0

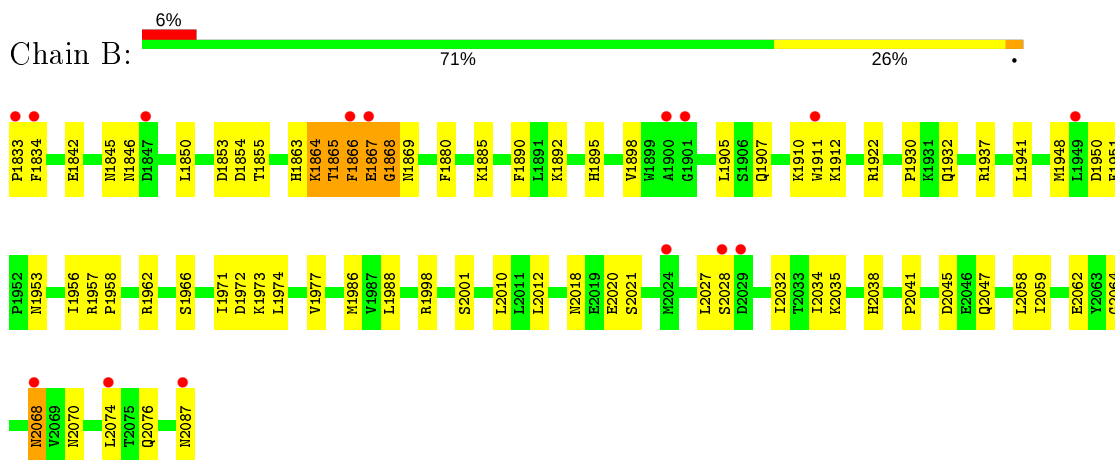
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Pre-mRNA-splicing factor 8



- Molecule 1: Pre-mRNA-splicing factor 8



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.12Å 84.16Å 95.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.10 29.65 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.9 (30.00-2.10) 98.9 (29.65-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.50 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.237 , 0.298 0.233 , 0.294	Depositor DCC
$R_{free}$ test set	1814 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.6	Xtrriage
Anisotropy	0.241	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 49.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4477	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% 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	0.45	0/2153	0.63	0/2919
1	B	0.44	0/2150	0.59	0/2914
All	All	0.45	0/4303	0.61	0/5833

There are no bond length outliers.

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	2110	0	2153	43	0
1	B	2107	0	2150	47	0
2	A	163	0	0	5	0
2	B	97	0	0	2	0
All	All	4477	0	4303	90	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 11.

All (90) 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:1898:VAL:HB	1:B:1912:LYS:HE3	1.52	0.89

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2086:GLN:HG3	1:A:2087:ASN:H	1.36	0.88
1:B:1866:PHE:O	1:B:1867:GLU:HB2	1.73	0.85
1:B:1890:PHE:HB3	1:B:1986:MET:HE1	1.58	0.85
1:B:1892[B]:LYS:HG3	1:B:1986:MET:CE	2.11	0.80
1:B:1907:GLN:O	1:B:1911:TRP:HD1	1.65	0.79
1:A:1927:GLU:HG3	2:A:127:HOH:O	1.83	0.79
1:B:1892[B]:LYS:HG3	1:B:1986:MET:HE2	1.69	0.72
1:A:1951:PHE:HB3	1:A:1954:ILE:HD12	1.73	0.71
1:B:1890:PHE:HB3	1:B:1986:MET:CE	2.22	0.69
1:A:1973:LYS:NZ	1:A:2034:ILE:HD11	2.08	0.69
1:B:1892[B]:LYS:CG	1:B:1986:MET:HE3	2.23	0.68
1:A:2085:GLY:O	1:A:2086:GLN:HB3	1.94	0.68
1:A:2086:GLN:HG3	1:A:2087:ASN:N	2.09	0.66
1:B:1864:LYS:HG2	1:B:1868:GLY:HA2	1.77	0.65
1:A:2085:GLY:O	1:A:2086:GLN:CB	2.48	0.62
1:B:1842[B]:GLU:HG2	1:B:1845:ASN:ND2	2.20	0.56
1:A:1930:PRO:O	1:A:1954:ILE:HG12	2.07	0.55
1:A:1973:LYS:HZ2	1:A:2034:ILE:HD11	1.72	0.55
1:B:1892[B]:LYS:HD3	1:B:1986:MET:HE3	1.89	0.54
1:A:2034:ILE:O	1:A:2034:ILE:HG13	2.09	0.53
1:B:1962:ARG:NH2	1:B:2087:ASN:OD1	2.36	0.53
1:B:1892[B]:LYS:CD	1:B:1986:MET:HE3	2.39	0.52
1:B:1846:ASN:OD1	1:B:1885[B]:LYS:NZ	2.43	0.52
1:A:1834:PHE:HE1	1:A:1938:LYS:HZ2	1.58	0.51
1:B:1895:HIS:O	1:B:1898:VAL:HG22	2.10	0.51
1:A:1918:SER:HA	1:A:1948[B]:MET:HE3	1.92	0.51
1:A:2029:ASP:N	1:A:2030:PRO:HD2	2.26	0.51
1:A:2075:THR:HG23	2:A:241:HOH:O	2.10	0.51
1:B:1865:THR:HB	1:B:1869:ASN:O	2.10	0.51
1:B:1855:THR:CG2	1:B:1937:ARG:HD2	2.41	0.51
1:A:2029:ASP:O	1:A:2031:THR:N	2.43	0.50
1:A:2029:ASP:N	1:A:2030:PRO:CD	2.76	0.49
1:B:1907:GLN:O	1:B:1911:TRP:CD1	2.56	0.49
1:A:1973:LYS:O	1:A:1977:VAL:HG13	2.11	0.49
1:A:1969:MET:HA	1:A:1969:MET:HE2	1.94	0.49
1:A:2029:ASP:C	1:A:2031:THR:H	2.16	0.49
1:A:1857:VAL:HG13	1:A:1894:ILE:HG13	1.95	0.48
1:A:1961:LEU:HD23	1:A:2086:GLN:HB2	1.94	0.48
1:B:1866:PHE:O	1:B:1867:GLU:CB	2.54	0.48
1:B:1998:ARG:HH21	1:B:2045:ASP:CG	2.17	0.48
1:A:1994:ASP:CG	1:A:1998:ARG:HH22	2.18	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1895:HIS:HD2	1:A:1897:SER:CB	2.28	0.47
1:A:1971:ILE:HG21	1:A:2039[A]:LEU:HD23	1.96	0.47
1:B:2032:ILE:HD11	1:B:2047:GLN:HE22	1.80	0.47
1:A:2075:THR:CG2	2:A:241:HOH:O	2.64	0.47
1:A:1895:HIS:CD2	1:A:1897:SER:H	2.33	0.46
1:A:1834:PHE:HE1	1:A:1938:LYS:NZ	2.14	0.46
1:A:1931:LYS:HB3	1:A:1931:LYS:HE2	1.81	0.46
1:A:1976:ASP:O	1:A:1980[B]:LYS:HG3	2.15	0.46
1:B:1833:PRO:HB2	1:B:1957:ARG:HH21	1.81	0.46
1:B:1850:LEU:HD12	1:B:1930:PRO:HG3	1.98	0.46
1:B:2010:LEU:HD12	1:B:2059:ILE:CD1	2.46	0.46
1:B:1892[B]:LYS:HG3	1:B:1986:MET:HE3	1.79	0.45
1:B:1853:ASP:HB3	1:B:1880:PHE:HB3	1.98	0.45
1:B:1973:LYS:O	1:B:1977:VAL:HG13	2.16	0.45
1:A:1845:ASN:OD1	1:A:1849:LYS:NZ	2.49	0.45
1:A:1911:TRP:HE3	1:A:1947:HIS:HE1	1.65	0.45
1:B:1863:HIS:HE1	2:B:97:HOH:O	1.99	0.45
1:B:1922[B]:ARG:HG2	1:B:1951:PHE:CE2	2.53	0.44
1:B:2064:GLY:O	1:B:2068:ASN:HA	2.16	0.44
1:A:1962:ARG:HD3	1:A:1962:ARG:HA	1.87	0.44
1:A:1961:LEU:HD22	1:A:2084:LEU:HB3	2.00	0.43
1:B:2035:LYS:HB2	1:B:2038:HIS:CG	2.53	0.43
1:B:1932:GLN:HE22	1:B:1957:ARG:HH11	1.66	0.43
1:A:1892:LYS:HE3	1:A:1893:ILE:O	2.19	0.43
1:B:1941:LEU:HD12	1:B:1956:ILE:HG22	2.01	0.43
2:A:22:HOH:O	1:B:1922[A]:ARG:HG2	2.19	0.43
1:B:1892[B]:LYS:CG	1:B:1986:MET:CE	2.82	0.42
1:B:1842[B]:GLU:HG2	1:B:1845:ASN:HD22	1.82	0.42
1:B:1833:PRO:O	1:B:1834:PHE:HB2	2.19	0.42
1:B:2018:ASN:HB3	1:B:2058:LEU:HD21	2.02	0.42
1:A:1977:VAL:HG23	1:A:1978:VAL:N	2.35	0.42
1:B:1972:ASP:HB2	2:B:130:HOH:O	2.20	0.42
1:A:1918:SER:HA	1:A:1948[B]:MET:CE	2.49	0.42
1:B:1885[A]:LYS:HD2	1:B:2001:SER:OG	2.20	0.42
1:A:2056[A]:ARG:NH1	2:A:79:HOH:O	2.52	0.42
1:A:1895:HIS:HD2	1:A:1897:SER:HB3	1.85	0.41
1:B:1867:GLU:HB3	1:B:1868:GLY:H	1.69	0.41
1:B:1854:ASP:OD1	1:B:1937:ARG:HD3	2.19	0.41
1:A:1991:ILE:HA	1:A:2039[B]:LEU:HD12	2.01	0.41
1:B:1971:ILE:HD11	1:B:2012:LEU:HD21	2.02	0.41
1:B:1941:LEU:HD22	1:B:1958:PRO:HG3	2.02	0.41

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2032:ILE:HG21	1:A:2043:PHE:CD1	2.56	0.41
1:A:2086:GLN:CG	1:A:2087:ASN:H	2.21	0.41
1:B:1950:ASP:OD2	1:B:1950:ASP:N	2.53	0.41
1:A:1976:ASP:HB3	1:A:1980[B]:LYS:NZ	2.36	0.40
1:B:2034:ILE:HG12	1:B:2041:PRO:HA	2.04	0.40
1:A:1942:ASP:HB3	1:A:1943:PRO:HD3	2.03	0.40
1:B:1855:THR:HG23	1:B:1937:ARG:HD2	2.03	0.40

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	259/255 (102%)	243 (94%)	13 (5%)	3 (1%)	13	8
1	B	258/255 (101%)	245 (95%)	9 (4%)	4 (2%)	9	5
All	All	517/510 (101%)	488 (94%)	22 (4%)	7 (1%)	11	6

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2030	PRO
1	B	1867	GLU
1	B	1868	GLY
1	A	2086	GLN
1	A	2029	ASP
1	B	1866	PHE
1	B	2068	ASN

### 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	239/233 (103%)	229 (96%)	10 (4%)	30	30
1	B	238/233 (102%)	221 (93%)	17 (7%)	14	11
All	All	477/466 (102%)	450 (94%)	27 (6%)	20	18

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1833	PRO
1	A	1912	LYS
1	A	1974	LEU
1	A	1976	ASP
1	A	1983	GLU
1	A	1986	MET
1	A	2029	ASP
1	A	2034	ILE
1	A	2060	LEU
1	A	2062	GLU
1	B	1864	LYS
1	B	1865	THR
1	B	1905	LEU
1	B	1910	LYS
1	B	1948	MET
1	B	1953	ASN
1	B	1966	SER
1	B	1974	LEU
1	B	1988	LEU
1	B	2020	GLU
1	B	2021	SER
1	B	2027	LEU
1	B	2028	SER
1	B	2062	GLU
1	B	2070	ASN
1	B	2074	LEU
1	B	2076	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1895	HIS
1	A	1907	GLN
1	A	2054	GLN
1	A	2068	ASN
1	B	1863	HIS
1	B	1932	GLN
1	B	2054	GLN
1	B	2070	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

### 6.1 Protein, DNA and RNA chains

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	255/255 (100%)	0.35	16 (6%) 20 24	23, 34, 48, 56	0
1	B	255/255 (100%)	0.37	15 (5%) 22 27	26, 38, 48, 52	0
All	All	510/510 (100%)	0.36	31 (6%) 21 26	23, 35, 48, 56	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2073	ALA	5.3
1	B	1833	PRO	4.5
1	B	1834	PHE	4.2
1	A	1833	PRO	4.1
1	A	2087	ASN	4.0
1	A	2067	TYR	3.8
1	B	1911	TRP	3.7
1	A	2031	THR	3.5
1	B	1900	ALA	3.1
1	A	2064	GLY	3.1
1	A	2068	ASN	2.9
1	B	2087	ASN	2.8
1	B	1867	GLU	2.8
1	B	1847	ASP	2.7
1	A	1991	ILE	2.7
1	A	1882	LEU	2.4
1	A	2027	LEU	2.3
1	B	1866	PHE	2.3
1	B	2068	ASN	2.2
1	B	1901	GLY	2.2
1	A	1881	THR	2.2
1	B	2074	LEU	2.2
1	B	2028	SER	2.2
1	A	2028	SER	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	2072	SER	2.1
1	B	2029	ASP	2.1
1	A	1900	ALA	2.1
1	A	1911	TRP	2.1
1	B	2024	MET	2.0
1	A	2008	LEU	2.0
1	B	1949	LEU	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 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.