



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2018 – 07:17 am GMT

PDB ID : 4F1Q  
Title : Human Artd8 (Parp14, Bal2) - catalytic domain in complex with A16(E)  
Authors : Karlberg, T.; Andersson, C.D.; Lindgren, A.; Spjut, S.; Thorsell, A.G.; Ekblad, T.; Weigelt, J.; Elofsson, M.; Linusson, A.; Schuler, H.  
Deposited on : 2012-05-07  
Resolution : 2.80 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  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtrriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

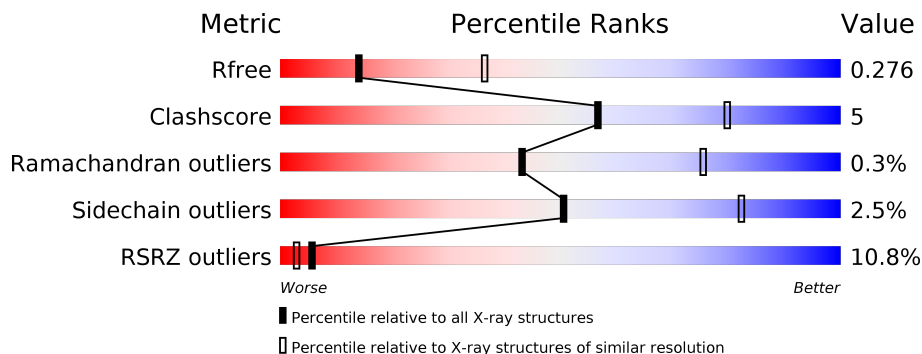
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

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	193	 9% 82% 10% •• 5%
1	B	193	 11% 83% 10% 7%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2959 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.

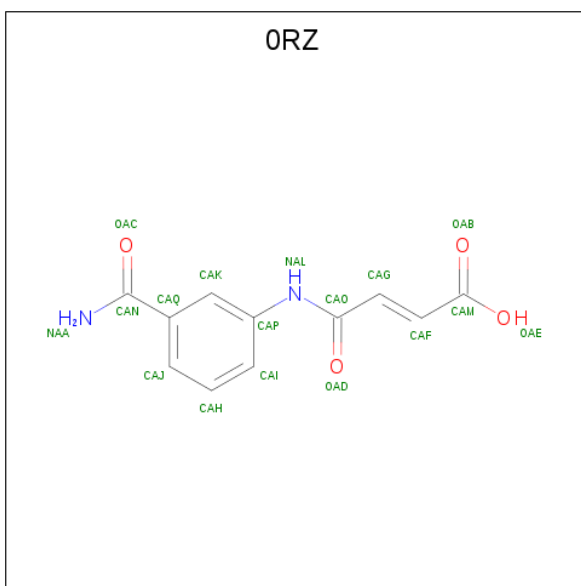
- Molecule 1 is a protein called Poly [ADP-ribose] polymerase 14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	183	1476	931	262	279	4	0	0	0
1	B	179	1445	914	255	272	4	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

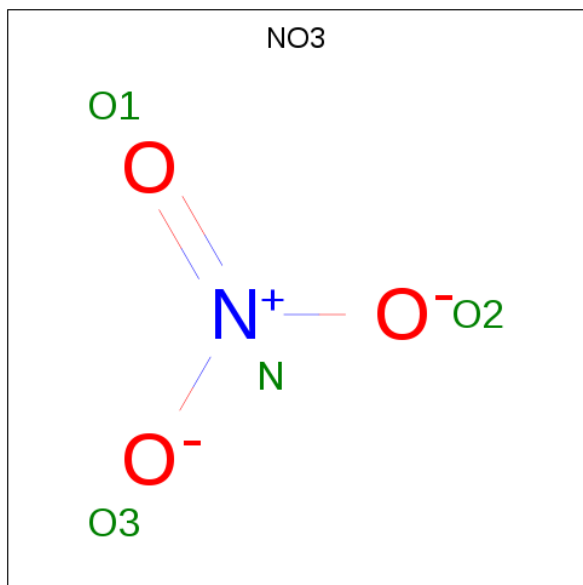
Chain	Residue	Modelled	Actual	Comment	Reference
A	1528	SER	-	EXPRESSION TAG	UNP Q460N5
A	1529	MET	-	EXPRESSION TAG	UNP Q460N5
B	1528	SER	-	EXPRESSION TAG	UNP Q460N5
B	1529	MET	-	EXPRESSION TAG	UNP Q460N5

- Molecule 2 is (2E)-4-[(3-carbamoylphenyl)amino]-4-oxobut-2-enoic acid (three-letter code: ORZ) (formula: C<sub>11</sub>H<sub>10</sub>N<sub>2</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	17	11	2	4	0	0
2	B	1	17	11	2	4	0	0

- Molecule 3 is NITRATE ION (three-letter code: NO3) (formula: NO<sub>3</sub>).

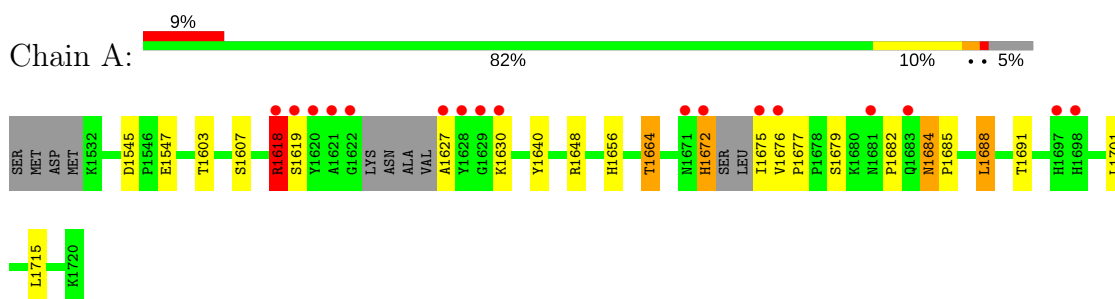


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	N	O		
3	B	1	4	1	3	0	0

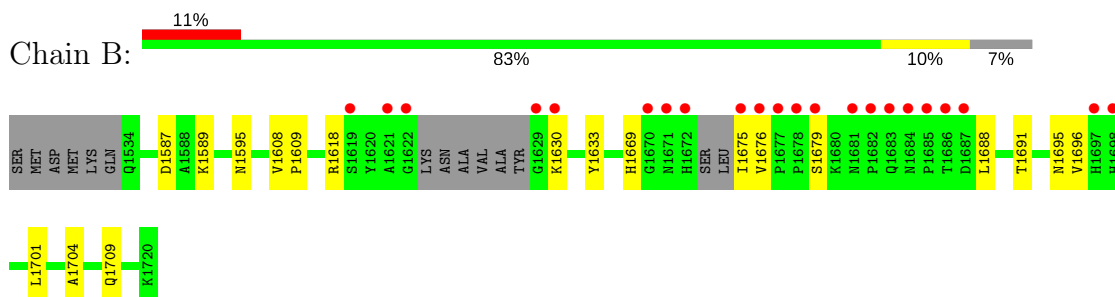
### 3 Residue-property plots [i](#)

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: Poly [ADP-ribose] polymerase 14



- Molecule 1: Poly [ADP-ribose] polymerase 14



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.68Å 144.99Å 83.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.24 – 2.80 33.24 – 2.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.24-2.80) 100.0 (33.24-2.80)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.99 (at 2.81Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.217 , 0.283 0.214 , 0.276	Depositor DCC
$R_{free}$ test set	1268 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.3	Xtrriage
Anisotropy	0.888	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 42.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.000 for 1/2*h+1/2*k,3/2*h-1/2*k,-l 0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	2959	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	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 37.06 % 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 4.5406e-04. 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NO3, ORZ

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	0.55	0/1517	0.64	0/2061
1	B	0.53	0/1486	0.63	0/2019
All	All	0.54	0/3003	0.64	0/4080

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	1476	0	1383	18	0
1	B	1445	0	1356	13	0
2	A	17	0	9	1	0
2	B	17	0	9	1	0
3	B	4	0	0	1	0
All	All	2959	0	2757	31	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 5.

All (31) 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:1675:ILE:HD12	1:B:1676:VAL:HG13	1.61	0.82
1:A:1679:SER:HA	1:A:1688:LEU:HD12	1.65	0.79
1:B:1669:HIS:CD2	1:B:1696:VAL:HG23	2.20	0.76
1:A:1627:ALA:HB1	1:A:1672:HIS:HB3	1.81	0.62
1:A:1672:HIS:ND1	1:A:1672:HIS:C	2.53	0.61
1:A:1545:ASP:OD1	1:A:1547:GLU:HB2	2.02	0.60
1:B:1704:ALA:HB1	1:B:1709:GLN:OE1	2.07	0.54
1:B:1679:SER:HA	1:B:1688:LEU:HD23	1.91	0.52
1:B:1675:ILE:CD1	1:B:1676:VAL:HG13	2.38	0.51
1:A:1656:HIS:HB3	1:A:1715:LEU:HD11	1.93	0.50
1:A:1701:LEU:HD12	1:A:1701:LEU:C	2.33	0.49
1:B:1587:ASP:OD1	1:B:1595:ASN:ND2	2.40	0.48
1:B:1701:LEU:C	1:B:1701:LEU:HD12	2.34	0.48
1:B:1609:PRO:HA	3:B:1801:NO3:O2	2.13	0.48
1:B:1630:LYS:HG3	1:B:1630:LYS:O	2.14	0.48
1:A:1640:TYR:CE2	2:A:1801:ORZ:CAH	2.97	0.47
1:A:1684:ASN:O	1:A:1684:ASN:ND2	2.43	0.47
1:A:1618:ARG:HG2	1:A:1619:SER:N	2.29	0.46
1:A:1603:THR:OG1	1:A:1607:SER:HB2	2.15	0.46
1:A:1684:ASN:C	1:A:1684:ASN:HD22	2.18	0.46
1:A:1701:LEU:HD12	1:A:1701:LEU:O	2.15	0.46
1:A:1664:THR:HA	1:A:1691:THR:HG21	1.99	0.44
1:B:1608:VAL:N	1:B:1609:PRO:CD	2.80	0.44
1:A:1676:VAL:CG2	1:A:1677:PRO:HD2	2.47	0.44
1:A:1630:LYS:HD3	1:A:1675:ILE:HG23	1.98	0.44
1:B:1669:HIS:CD2	1:B:1695:ASN:HA	2.54	0.43
1:A:1630:LYS:HB3	1:A:1675:ILE:HA	2.01	0.42
1:A:1682:PRO:O	1:A:1685:PRO:HG3	2.18	0.42
1:B:1691:THR:HG22	1:B:1704:ALA:HA	2.00	0.42
1:B:1633:TYR:HB3	2:B:1802:ORZ:CAI	2.50	0.41
1:A:1618:ARG:CG	1:A:1619:SER:N	2.83	0.41

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	177/193 (92%)	170 (96%)	6 (3%)	1 (1%)	27	60
1	B	173/193 (90%)	164 (95%)	9 (5%)	0	100	100
All	All	350/386 (91%)	334 (95%)	15 (4%)	1 (0%)	43	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1618	ARG

### 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	159/170 (94%)	153 (96%)	6 (4%)	36	70
1	B	156/170 (92%)	154 (99%)	2 (1%)	71	92
All	All	315/340 (93%)	307 (98%)	8 (2%)	50	82

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1618	ARG
1	A	1648	ARG
1	A	1664	THR
1	A	1672	HIS
1	A	1684	ASN
1	A	1688	LEU
1	B	1589	LYS
1	B	1618	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1534	GLN
1	B	1614	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](#)

3 ligands 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
2	ORZ	A	1801	-	14,17,17	3.48	4 (28%)	18,22,22	1.22	2 (11%)
3	NO3	B	1801	-	1,3,3	3.34	1 (100%)	0,3,3	0.00	-
2	ORZ	B	1802	-	14,17,17	3.25	4 (28%)	18,22,22	1.21	2 (11%)

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
2	ORZ	A	1801	-	-	0/11/13/13	0/1/1/1
3	NO3	B	1801	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ORZ	B	1802	-	-	0/11/13/13	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1801	ORZ	CAQ-CAN	-8.09	1.37	1.50
2	B	1802	ORZ	CAQ-CAN	-7.17	1.39	1.50
2	A	1801	ORZ	CAP-NAL	-4.33	1.33	1.41
2	B	1802	ORZ	CAP-NAL	-3.72	1.34	1.41
3	B	1801	NO3	O1-N	3.34	1.40	1.24
2	A	1801	ORZ	CAG-CAO	3.50	1.54	1.48
2	B	1802	ORZ	CAG-CAO	3.70	1.55	1.48
2	B	1802	ORZ	CAG-CAF	8.20	1.54	1.33
2	A	1801	ORZ	CAG-CAF	8.43	1.55	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1802	ORZ	CAP-NAL-CAO	-3.59	122.71	128.25
2	A	1801	ORZ	CAP-NAL-CAO	-2.63	124.19	128.25
2	B	1802	ORZ	OAD-CAO-NAL	2.24	125.57	122.95
2	A	1801	ORZ	OAD-CAO-NAL	2.74	126.15	122.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1801	ORZ	1	0
3	B	1801	NO3	1	0
2	B	1802	ORZ	1	0

## 5.7 Other polymers [i](#)

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

### 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	183/193 (94%)	0.05	17 (9%) <b>8</b> <b>4</b>	6, 21, 63, 74	0
1	B	179/193 (92%)	0.23	22 (12%) <b>4</b> <b>2</b>	8, 24, 90, 99	0
All	All	362/386 (93%)	0.14	39 (10%) <b>6</b> <b>3</b>	6, 22, 67, 99	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1683	GLN	7.8
1	A	1675	ILE	6.2
1	B	1681	ASN	6.1
1	B	1684	ASN	6.1
1	B	1619	SER	6.0
1	B	1675	ILE	5.6
1	B	1686	THR	5.6
1	A	1619	SER	4.5
1	A	1622	GLY	4.5
1	B	1629	GLY	4.3
1	B	1687	ASP	4.2
1	B	1630	LYS	3.8
1	B	1682	PRO	3.6
1	A	1628	TYR	3.6
1	A	1621	ALA	3.5
1	A	1620	TYR	3.4
1	A	1629	GLY	3.4
1	B	1698	HIS	3.2
1	B	1679	SER	3.2
1	B	1685	PRO	3.2
1	B	1621	ALA	3.2
1	B	1672	HIS	3.1
1	B	1622	GLY	3.1
1	A	1672	HIS	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	1678	PRO	2.9
1	A	1676	VAL	2.9
1	B	1677	PRO	2.8
1	B	1676	VAL	2.8
1	B	1671	ASN	2.7
1	A	1627	ALA	2.7
1	A	1681	ASN	2.6
1	B	1697	HIS	2.5
1	A	1671	ASN	2.3
1	A	1683	GLN	2.3
1	A	1697	HIS	2.2
1	A	1630	LYS	2.1
1	A	1698	HIS	2.1
1	B	1670	GLY	2.1
1	A	1618	ARG	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](#)

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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	0RZ	B	1802	17/17	0.75	0.38	47,56,67,68	0
2	0RZ	A	1801	17/17	0.81	0.32	36,44,55,56	0
3	NO3	B	1801	4/4	0.97	0.18	27,28,29,29	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.