Full wwPDB X-ray Structure Validation Report ⓘ

Nov 16, 2023 – 02:30 AM JST

PDB ID	:	6KK9
Title	:	A Crystal structure of OspA mutant
Authors	:	Shiga, S.; Makabe, K.
Deposited on	:	2019-07-24
Resolution	:	2.20 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*
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:

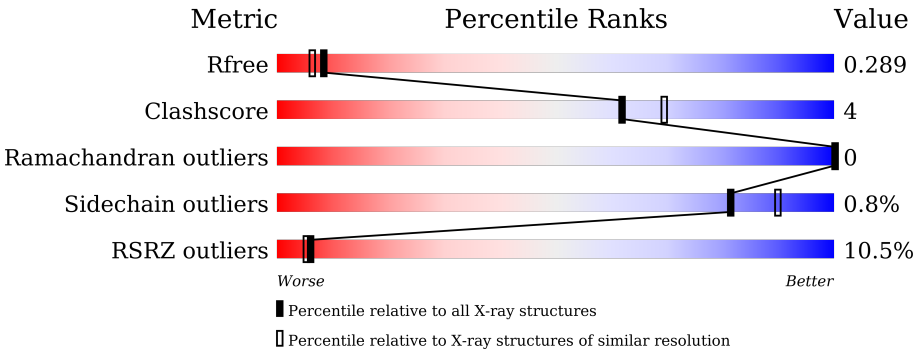
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 >=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 <=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	B	198	7% 87% 8% • •
1	C	198	12% 87% 7% • 5%
1	D	198	8% 86% 9% • 5%
1	O	198	14% 80% 14% 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5965 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 Outer Surface Protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	O	187	Total	C	N	O	S	0	0	0
			1447	894	238	314	1			
1	B	190	Total	C	N	O	S	0	0	0
			1475	912	241	321	1			
1	C	188	Total	C	N	O	S	0	0	0
			1456	899	239	317	1			
1	D	189	Total	C	N	O	S	0	0	0
			1468	908	240	319	1			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	125	LYS	LEU	engineered mutation	UNP D0VWU8
O	134	GLU	LEU	engineered mutation	UNP D0VWU8
O	148	LYS	LEU	engineered mutation	UNP D0VWU8
O	157	GLU	LEU	engineered mutation	UNP D0VWU8
O	171	LYS	LEU	engineered mutation	UNP D0VWU8
O	182	TYR	LYS	engineered mutation	UNP D0VWU8
O	194	LYS	LEU	engineered mutation	UNP D0VWU8
O	201	VAL	LEU	engineered mutation	UNP D0VWU8
O	203	GLU	LEU	engineered mutation	UNP D0VWU8
B	125	LYS	LEU	engineered mutation	UNP D0VWU8
B	134	GLU	LEU	engineered mutation	UNP D0VWU8
B	148	LYS	LEU	engineered mutation	UNP D0VWU8
B	157	GLU	LEU	engineered mutation	UNP D0VWU8
B	171	LYS	LEU	engineered mutation	UNP D0VWU8
B	182	TYR	LYS	engineered mutation	UNP D0VWU8
B	194	LYS	LEU	engineered mutation	UNP D0VWU8
B	201	VAL	LEU	engineered mutation	UNP D0VWU8
B	203	GLU	LEU	engineered mutation	UNP D0VWU8
C	125	LYS	LEU	engineered mutation	UNP D0VWU8
C	134	GLU	LEU	engineered mutation	UNP D0VWU8
C	148	LYS	LEU	engineered mutation	UNP D0VWU8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	157	GLU	LEU	engineered mutation	UNP D0VWU8
C	171	LYS	LEU	engineered mutation	UNP D0VWU8
C	182	TYR	LYS	engineered mutation	UNP D0VWU8
C	194	LYS	LEU	engineered mutation	UNP D0VWU8
C	201	VAL	LEU	engineered mutation	UNP D0VWU8
C	203	GLU	LEU	engineered mutation	UNP D0VWU8
D	125	LYS	LEU	engineered mutation	UNP D0VWU8
D	134	GLU	LEU	engineered mutation	UNP D0VWU8
D	148	LYS	LEU	engineered mutation	UNP D0VWU8
D	157	GLU	LEU	engineered mutation	UNP D0VWU8
D	171	LYS	LEU	engineered mutation	UNP D0VWU8
D	182	TYR	LYS	engineered mutation	UNP D0VWU8
D	194	LYS	LEU	engineered mutation	UNP D0VWU8
D	201	VAL	LEU	engineered mutation	UNP D0VWU8
D	203	GLU	LEU	engineered mutation	UNP D0VWU8

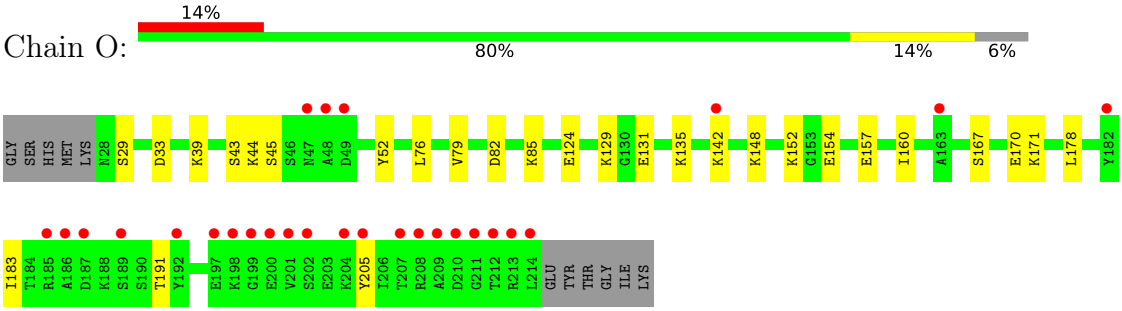
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	O	30	Total O 30 30	0	0
2	B	33	Total O 33 33	0	0
2	C	17	Total O 17 17	0	0
2	D	39	Total O 39 39	0	0

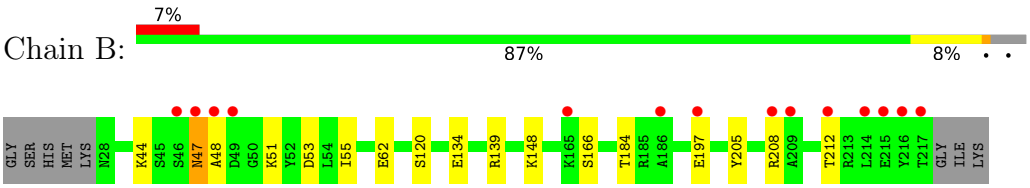
3 Residue-property plots (i)

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.

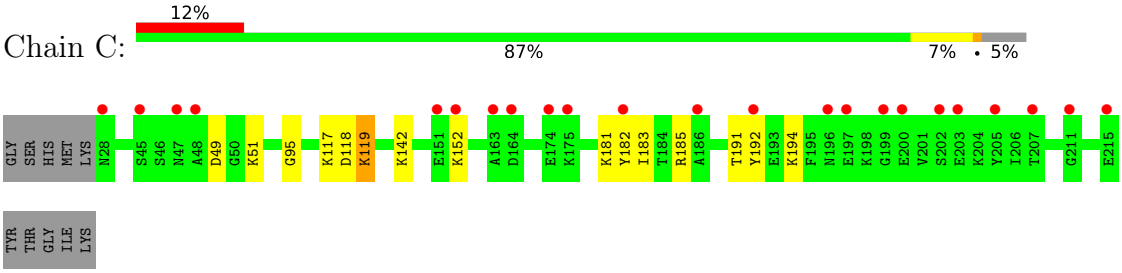
• Molecule 1: Outer Surface Protein A



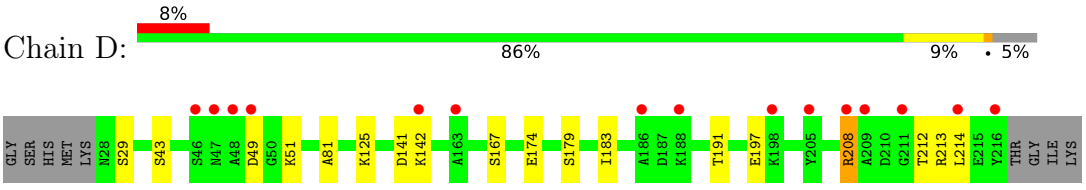
• Molecule 1: Outer Surface Protein A



• Molecule 1: Outer Surface Protein A



• Molecule 1: Outer Surface Protein A



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	102.54Å 138.64Å 62.51Å 90.00° 100.88° 90.00°	Depositor
Resolution (Å)	19.90 – 2.20 36.01 – 2.19	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.90-2.20) 99.2 (36.01-2.19)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.14 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.222 , 0.287 0.224 , 0.289	Depositor DCC
R_{free} test set	2178 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	33.1	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 42.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5965	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 37.52 % 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 4.2391e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	B	0.53	0/1490	0.70	0/1994
1	C	0.45	0/1470	0.64	0/1966
1	D	0.48	0/1483	0.66	0/1984
1	O	0.51	0/1461	0.63	0/1954
All	All	0.49	0/5904	0.66	0/7898

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

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	B	1475	0	1464	9	2
1	C	1456	0	1448	10	0
1	D	1468	0	1457	13	1
1	O	1447	0	1442	21	1
2	B	33	0	0	0	0
2	C	17	0	0	0	0
2	D	39	0	0	0	0
2	O	30	0	0	0	0
All	All	5965	0	5811	51	2

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 4.

All (51) 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:O:129:LYS:HE2	1:O:131:GLU:CG	1.78	1.14
1:O:129:LYS:HE2	1:O:131:GLU:HG2	1.48	0.92
1:D:167:SER:O	1:D:183:ILE:HD12	1.76	0.85
1:C:49:ASP:HB3	1:C:51:LYS:HG3	1.58	0.84
1:D:208:ARG:HB2	1:D:212:THR:HG22	1.61	0.81
1:C:181:LYS:HE2	1:C:183:ILE:HD11	1.67	0.75
1:O:129:LYS:HE2	1:O:131:GLU:HG3	1.68	0.73
1:C:192:TYR:HE2	1:C:194:LYS:HG3	1.61	0.66
1:O:152:LYS:HD2	1:O:154:GLU:HB3	1.77	0.66
1:B:134:GLU:HG2	1:B:148:LYS:HD2	1.78	0.66
1:O:183:ILE:HG22	1:O:191:THR:HB	1.78	0.65
1:D:141:ASP:C	1:D:142:LYS:HD2	2.19	0.63
1:O:183:ILE:CG2	1:O:191:THR:HB	2.29	0.62
1:O:129:LYS:HE2	1:O:131:GLU:CD	2.20	0.61
1:O:33:ASP:OD2	1:O:39:LYS:HE3	2.02	0.60
1:O:124:GLU:HG2	1:O:135:LYS:HG3	1.82	0.59
1:D:174:GLU:CD	1:D:174:GLU:H	2.04	0.59
1:O:79:VAL:HG22	1:O:85:LYS:HD3	1.86	0.57
1:D:49:ASP:HB3	1:D:51:LYS:HG2	1.87	0.57
1:B:47:ASN:HD22	1:B:47:ASN:C	2.07	0.57
1:O:205:TYR:OH	1:D:125:LYS:HD2	2.05	0.56
1:D:197:GLU:H	1:D:197:GLU:CD	2.10	0.55
1:C:192:TYR:CE2	1:C:194:LYS:HG3	2.40	0.55
1:D:183:ILE:CG2	1:D:191:THR:HB	2.40	0.52
1:O:157:GLU:HG2	1:O:171:LYS:HD2	1.96	0.48
1:C:142:LYS:HA	1:C:142:LYS:HD3	1.60	0.47
1:B:197:GLU:N	1:B:197:GLU:OE2	2.47	0.47
1:C:182:TYR:HE2	1:C:192:TYR:CE1	2.32	0.47
1:B:48:ALA:HB3	1:B:51:LYS:HB2	1.95	0.47
1:O:160:ILE:O	1:O:167:SER:HA	2.16	0.45
1:B:120:SER:HB2	1:B:139:ARG:HA	1.99	0.45
1:O:142:LYS:HA	1:O:142:LYS:HD3	1.64	0.44
1:B:166:SER:HB2	1:B:184:THR:O	2.19	0.43
1:B:208:ARG:NH1	1:B:212:THR:OG1	2.52	0.43
1:D:213:ARG:O	1:D:214:LEU:HG	2.18	0.43
1:O:29:SER:HB3	1:O:43:SER:HB3	2.01	0.42
1:C:118:ASP:O	1:C:119:LYS:HB2	2.19	0.42
1:B:55:ILE:HD11	1:B:62:GLU:CD	2.40	0.42
1:O:129:LYS:CE	1:O:131:GLU:HG2	2.34	0.42
1:O:152:LYS:HD2	1:O:154:GLU:CB	2.47	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:185:ARG:HH22	1:C:191:THR:HG21	1.85	0.42
1:D:29:SER:HB3	1:D:43:SER:HB2	2.01	0.42
1:O:148:LYS:HD3	1:D:81:ALA:HB2	2.03	0.41
1:D:183:ILE:HG23	1:D:191:THR:HB	2.01	0.41
1:O:76:LEU:HD12	1:O:76:LEU:N	2.36	0.41
1:B:44:LYS:HD2	1:B:51:LYS:O	2.20	0.41
1:C:95:GLY:HA2	1:C:117:LYS:HE2	2.03	0.41
1:C:152:LYS:HB2	1:C:152:LYS:HE2	1.87	0.40
1:O:170:GLU:HG2	1:O:178:LEU:HD11	2.03	0.40
1:O:44:LYS:HD3	1:O:52:TYR:CE2	2.57	0.40
1:D:214:LEU:HD23	1:D:214:LEU:HA	1.91	0.40

All (2) 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)
1:B:53:ASP:OD2	1:D:208:ARG:NH2[4_556]	1.99	0.21
1:O:82:ASP:OD2	1:B:205:TYR:OH[4_556]	2.03	0.17

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	B	188/198 (95%)	179 (95%)	9 (5%)	0	100	100
1	C	186/198 (94%)	175 (94%)	11 (6%)	0	100	100
1	D	187/198 (94%)	177 (95%)	10 (5%)	0	100	100
1	O	185/198 (93%)	171 (92%)	14 (8%)	0	100	100
All	All	746/792 (94%)	702 (94%)	44 (6%)	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	B	168/174 (97%)	167 (99%)	1 (1%)	86 93
1	C	166/174 (95%)	165 (99%)	1 (1%)	86 93
1	D	167/174 (96%)	165 (99%)	2 (1%)	71 83
1	O	165/174 (95%)	164 (99%)	1 (1%)	86 93
All	All	666/696 (96%)	661 (99%)	5 (1%)	81 90

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	O	45	SER
1	B	47	ASN
1	C	119	LYS
1	D	179	SER
1	D	208	ARG

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

Mol	Chain	Res	Type
1	B	47	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 monosaccharides 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^{th} 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^2)	Q<0.9
1	B	190/198 (95%)	0.45	14 (7%) 14 13	19, 33, 61, 91	0
1	C	188/198 (94%)	0.58	23 (12%) 4 3	24, 41, 77, 94	0
1	D	189/198 (95%)	0.53	15 (7%) 12 11	22, 35, 71, 93	0
1	O	187/198 (94%)	0.80	27 (14%) 2 2	19, 35, 88, 115	0
All	All	754/792 (95%)	0.59	79 (10%) 6 5	19, 36, 80, 115	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	O	210	ASP	10.6
1	O	209	ALA	9.5
1	O	211	GLY	7.4
1	D	216	TYR	6.7
1	D	208	ARG	6.4
1	C	205	TYR	6.0
1	O	186	ALA	5.7
1	O	199	GLY	5.7
1	O	212	THR	5.4
1	B	217	THR	5.1
1	O	197	GLU	5.0
1	D	46	SER	4.7
1	D	48	ALA	4.6
1	O	201	VAL	4.5
1	C	174	GLU	4.5
1	D	209	ALA	4.4
1	D	47	ASN	4.3
1	B	48	ALA	4.2
1	D	205	TYR	4.1
1	D	163	ALA	4.0
1	B	214	LEU	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	208	ARG	3.9
1	B	216	TYR	3.9
1	D	186	ALA	3.9
1	O	205	TYR	3.8
1	C	48	ALA	3.8
1	O	200	GLU	3.6
1	O	214	LEU	3.6
1	C	196	ASN	3.6
1	O	189	SER	3.5
1	C	215	GLU	3.4
1	O	208	ARG	3.4
1	B	215	GLU	3.3
1	O	187	ASP	3.3
1	C	202	SER	3.2
1	C	182	TYR	3.1
1	O	185	ARG	3.0
1	C	200	GLU	3.0
1	O	207	THR	3.0
1	O	198	LYS	2.9
1	C	211	GLY	2.9
1	D	211	GLY	2.8
1	C	203	GLU	2.8
1	B	49	ASP	2.7
1	D	188	LYS	2.7
1	D	198	LYS	2.7
1	D	49	ASP	2.7
1	C	151	GLU	2.7
1	C	175	LYS	2.7
1	O	204	LYS	2.7
1	B	212	THR	2.6
1	O	49	ASP	2.6
1	O	47	ASN	2.6
1	B	47	ASN	2.6
1	O	213	ARG	2.6
1	C	45	SER	2.6
1	B	197	GLU	2.6
1	C	199	GLY	2.5
1	C	47	ASN	2.5
1	O	163	ALA	2.5
1	O	48	ALA	2.5
1	C	28	ASN	2.5
1	B	209	ALA	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	164	ASP	2.4
1	C	163	ALA	2.4
1	B	165	LYS	2.3
1	B	46	SER	2.3
1	O	192	TYR	2.3
1	D	142	LYS	2.3
1	O	142	LYS	2.3
1	C	197	GLU	2.2
1	D	214	LEU	2.2
1	C	152	LYS	2.1
1	O	202	SER	2.1
1	C	192	TYR	2.0
1	B	186	ALA	2.0
1	O	182	TYR	2.0
1	C	207	THR	2.0
1	C	186	ALA	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 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.