WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 08:46 PM GMT

PDB ID	:	4U4J
Title	:	Crystal structure of Salmonella alpha-2-macroglobulin mutant Y1175G
Authors	:	Wong, S.G.; Dessen, A.
Deposited on	:	2014-07-23
Resolution	:	2.90 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865
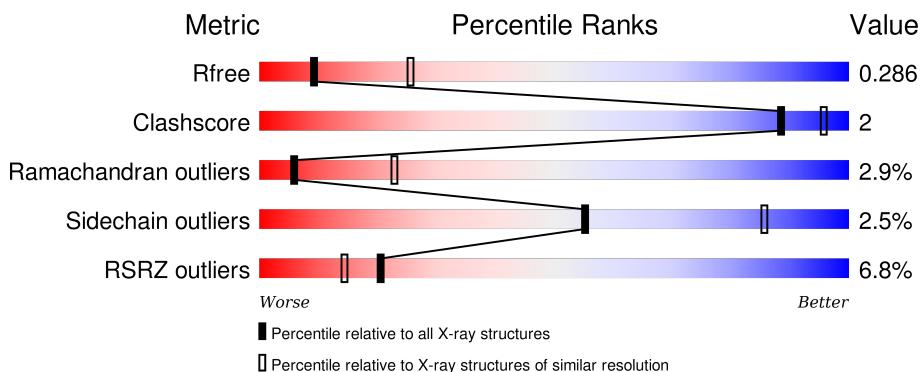
1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION
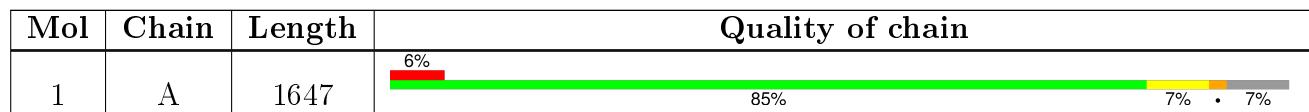
The reported resolution of this entry is 2.90 Å.

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	Percentile Ranks	Value
Rfree		0.286
Clashscore		2
Ramachandran outliers		2.9%
Sidechain outliers		2.5%
RSRZ outliers		6.8%

Worse ▮ *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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 >=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 <=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	1647	6% / 85% / 7% · 7%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 23627 atoms, of which 11754 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 Putative inner membrane lipoprotein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	1538	23627	7480	11754	2091	2279	23	0	0	0
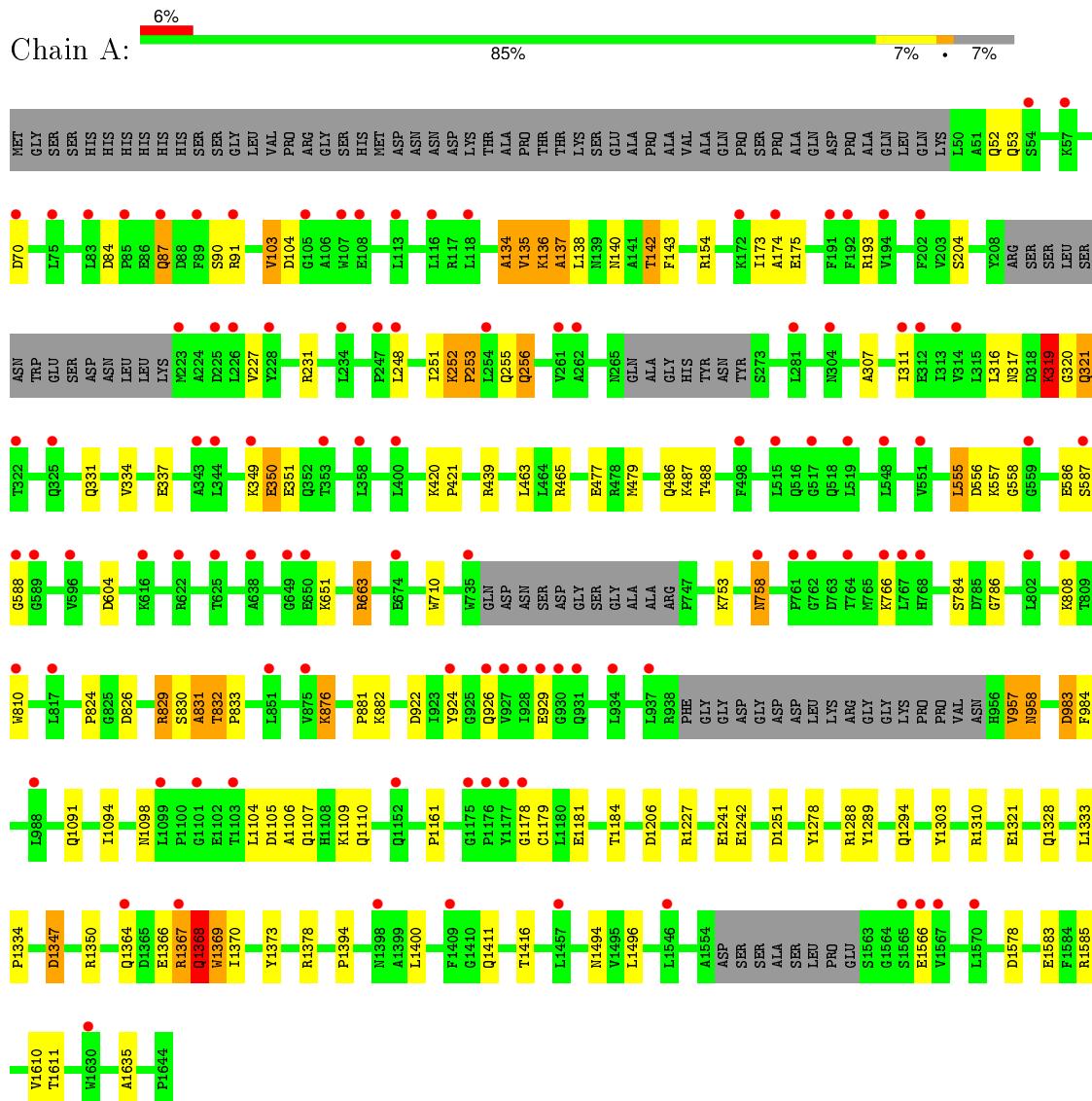
There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	initiating methionine	UNP Q8ZN46
A	-1	GLY	-	expression tag	UNP Q8ZN46
A	0	SER	-	expression tag	UNP Q8ZN46
A	1	SER	-	expression tag	UNP Q8ZN46
A	2	HIS	-	expression tag	UNP Q8ZN46
A	3	HIS	-	expression tag	UNP Q8ZN46
A	4	HIS	-	expression tag	UNP Q8ZN46
A	5	HIS	-	expression tag	UNP Q8ZN46
A	6	HIS	-	expression tag	UNP Q8ZN46
A	7	HIS	-	expression tag	UNP Q8ZN46
A	8	SER	-	expression tag	UNP Q8ZN46
A	9	SER	-	expression tag	UNP Q8ZN46
A	10	GLY	-	expression tag	UNP Q8ZN46
A	11	LEU	-	expression tag	UNP Q8ZN46
A	12	VAL	-	expression tag	UNP Q8ZN46
A	13	PRO	-	expression tag	UNP Q8ZN46
A	14	ARG	-	expression tag	UNP Q8ZN46
A	15	GLY	-	expression tag	UNP Q8ZN46
A	16	SER	-	expression tag	UNP Q8ZN46
A	17	HIS	-	expression tag	UNP Q8ZN46
A	18	MET	-	expression tag	UNP Q8ZN46
A	98	ALA	LYS	engineered mutation	UNP Q8ZN46
A	99	ALA	LYS	engineered mutation	UNP Q8ZN46
A	1175	GLY	TYR	engineered mutation	UNP Q8ZN46

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Putative inner membrane lipoprotein

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	254.24Å　　82.16Å　　98.90Å 90.00°　　　90.00°　　　90.00°	Depositor
Resolution (Å)	58.99　–　2.90 58.99　–　2.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (58.99-2.90) 99.9 (58.99-2.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.86 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.244　,　0.284 0.246　,　0.286	Depositor DCC
R_{free} test set	2338 reflections (5.00%)	DCC
Wilson B-factor (Å^2)	87.1	Xtriage
Anisotropy	0.461	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 49.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$<|L|>=0.49, <L^2>=0.32$	Xtriage
Outliers	1 of 46790 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	23627	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.23	0/12113	0.43	1/16503 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	316	LEU	CA-CB-CG	7.46	132.47	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	252	LYS	Peptide
1	A	808	LYS	Peptide

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	11873	11754	11772	59	0
All	All	11873	11754	11772	59	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 2.

All (59) 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:A:87:GLN:N	1:A:87:GLN:OE1	2.19	0.76
1:A:1366:GLU:O	1:A:1368:GLN:N	2.19	0.75
1:A:1368:GLN:OE1	1:A:1369:TRP:N	2.30	0.65
1:A:463:LEU:O	1:A:465:ARG:NH1	2.34	0.60
1:A:317:ASN:O	1:A:319:LYS:N	2.33	0.60
1:A:1091:GLN:NE2	1:A:1110:GLN:OE1	2.36	0.58
1:A:337:GLU:N	1:A:337:GLU:OE1	2.37	0.57
1:A:1181:GLU:OE1	1:A:1416:THR:OG1	2.22	0.57
1:A:829:ARG:O	1:A:831:ALA:N	2.38	0.57
1:A:1294:GLN:OE1	1:A:1328:GLN:NE2	2.38	0.57
1:A:1566:GLU:N	1:A:1566:GLU:OE1	2.40	0.55
1:A:84:ASP:HB2	1:A:136:LYS:HE3	1.89	0.55
1:A:922:ASP:OD2	1:A:924:TYR:N	2.40	0.54
1:A:1241:GLU:OE2	1:A:1278:TYR:OH	2.19	0.54
1:A:349:LYS:O	1:A:351:GLU:N	2.41	0.53
1:A:663:ARG:NH2	1:A:710:TRP:O	2.40	0.53
1:A:784:SER:O	1:A:786:GLY:N	2.41	0.53
1:A:555:LEU:O	1:A:557:LYS:N	2.42	0.53
1:A:1105:ASP:OD1	1:A:1106:ALA:N	2.42	0.53
1:A:1181:GLU:O	1:A:1184:THR:N	2.42	0.52
1:A:1583:GLU:OE2	1:A:1585:ARG:NE	2.42	0.52
1:A:137:ALA:O	1:A:140:ASN:N	2.43	0.52
1:A:758:ASN:ND2	1:A:758:ASN:O	2.42	0.51
1:A:555:LEU:C	1:A:555:LEU:HD23	2.32	0.50
1:A:320:GLY:O	1:A:321:GLN:NE2	2.45	0.49
1:A:1227:ARG:NH1	1:A:1241:GLU:OE1	2.46	0.49
1:A:1347:ASP:OD2	1:A:1350:ARG:NH1	2.45	0.49
1:A:136:LYS:HE2	1:A:138:LEU:HA	1.96	0.48
1:A:1289:TYR:OH	1:A:1578:ASP:OD2	2.22	0.48
1:A:983:ASP:OD1	1:A:1109:LYS:NZ	2.47	0.47
1:A:1494:ASN:O	1:A:1496:LEU:N	2.47	0.47
1:A:204:SER:OG	1:A:319:LYS:O	2.27	0.46
1:A:103:VAL:HG22	1:A:104:ASP:H	1.80	0.46

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1251:ASP:OD1	1:A:1310:ARG:NH1	2.50	0.45
1:A:486:GLN:O	1:A:488:THR:N	2.50	0.45
1:A:134:ALA:O	1:A:136:LYS:N	2.51	0.44
1:A:557:LYS:HG2	1:A:558:GLY:H	1.82	0.44
1:A:1610:VAL:HG23	1:A:1611:THR:N	2.33	0.44
1:A:251:ILE:HG23	1:A:253:PRO:HD2	2.00	0.44
1:A:134:ALA:C	1:A:136:LYS:H	2.20	0.43
1:A:84:ASP:HB2	1:A:136:LYS:HD3	2.01	0.43
1:A:1364:GLN:OE1	1:A:1364:GLN:N	2.49	0.42
1:A:604:ASP:OD1	1:A:604:ASP:N	2.52	0.42
1:A:420:LYS:HB2	1:A:421:PRO:HD2	2.01	0.42
1:A:136:LYS:HZ3	1:A:137:ALA:CA	2.33	0.42
1:A:957:VAL:O	1:A:958:ASN:CB	2.68	0.42
1:A:193:ARG:NH2	1:A:256:GLN:OE1	2.53	0.41
1:A:252:LYS:HG3	1:A:255:GLN:HB2	2.01	0.41
1:A:1094:ILE:HB	1:A:1107:GLN:HB2	2.01	0.41
1:A:1288:ARG:NH2	1:A:1289:TYR:OH	2.54	0.41
1:A:1242:GLU:OE1	1:A:1373:TYR:OH	2.31	0.41
1:A:810:TRP:CD1	1:A:810:TRP:N	2.89	0.41
1:A:1368:GLN:N	1:A:1368:GLN:CD	2.73	0.41
1:A:173:ILE:CD1	1:A:175:GLU:HB2	2.51	0.41
1:A:477:GLU:O	1:A:479:MET:N	2.54	0.41
1:A:832:THR:HA	1:A:833:PRO:HD2	2.00	0.40
1:A:1333:LEU:HB3	1:A:1334:PRO:HD3	2.03	0.40
1:A:136:LYS:HE2	1:A:137:ALA:C	2.42	0.40
1:A:135:VAL:CG2	1:A:143:PHE:HB2	2.51	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	1526/1647 (93%)	1375 (90%)	106 (7%)	45 (3%)	6	23

All (45) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	321	GLN
1	A	556	ASP
1	A	587	SER
1	A	651	LYS
1	A	758	ASN
1	A	882	LYS
1	A	957	VAL
1	A	1098	ASN
1	A	1367	ARG
1	A	1368	GLN
1	A	1370	ILE
1	A	134	ALA
1	A	137	ALA
1	A	142	THR
1	A	231	ARG
1	A	829	ARG
1	A	830	SER
1	A	958	ASN
1	A	983	ASP
1	A	1347	ASP
1	A	1635	ALA
1	A	135	VAL
1	A	487	LYS
1	A	588	GLY
1	A	876	LYS
1	A	1369	TRP
1	A	1411	GLN
1	A	90	SER
1	A	307	ALA
1	A	319	LYS
1	A	350	GLU
1	A	832	THR
1	A	1104	LEU
1	A	174	ALA
1	A	253	PRO
1	A	831	ALA
1	A	881	PRO
1	A	984	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1178	GLY
1	A	1206	ASP
1	A	1394	PRO
1	A	929	GLU
1	A	1161	PRO
1	A	103	VAL
1	A	824	PRO

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	1259/1347 (94%)	1227 (98%)	32 (2%)	55	85

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	52	GLN
1	A	53	GLN
1	A	70	ASP
1	A	87	GLN
1	A	91	ARG
1	A	136	LYS
1	A	142	THR
1	A	154	ARG
1	A	227	VAL
1	A	248	LEU
1	A	256	GLN
1	A	311	ILE
1	A	319	LYS
1	A	331	GLN
1	A	334	VAL
1	A	350	GLU
1	A	439	ARG
1	A	555	LEU
1	A	586	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	663	ARG
1	A	753	LYS
1	A	766	LYS
1	A	826	ASP
1	A	876	LYS
1	A	926	GLN
1	A	1179	CYS
1	A	1303	TYR
1	A	1321	GLU
1	A	1367	ARG
1	A	1368	GLN
1	A	1378	ARG
1	A	1400	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	758	ASN
1	A	956	HIS
1	A	1091	GLN
1	A	1328	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 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 (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, 95th 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(Å²)	Q<0.9
1	A	1538/1647 (93%)	0.64	105 (6%)	20	14	51, 90, 137, 213	0

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	588	GLY	7.9
1	A	928	ILE	7.4
1	A	70	ASP	6.1
1	A	57	LYS	5.8
1	A	85	PRO	5.5
1	A	929	GLU	5.4
1	A	559	GLY	5.0
1	A	937	LEU	4.8
1	A	83	LEU	4.6
1	A	589	GLY	4.4
1	A	649	GLY	4.3
1	A	930	GLY	4.1
1	A	587	SER	4.1
1	A	1176	PRO	4.0
1	A	194	VAL	4.0
1	A	118	LEU	3.8
1	A	223	MET	3.8
1	A	1570	LEU	3.7
1	A	325	GLN	3.6
1	A	343	ALA	3.6
1	A	650	GLU	3.5
1	A	758	ASN	3.5
1	A	768	HIS	3.5
1	A	314	VAL	3.4
1	A	87	GLN	3.4
1	A	851	LEU	3.4
1	A	761	PRO	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	225	ASP	3.4
1	A	519	LEU	3.3
1	A	622	ARG	3.3
1	A	108	GLU	3.2
1	A	75	LEU	3.2
1	A	1101	GLY	3.2
1	A	1178	GLY	3.1
1	A	228	TYR	3.1
1	A	764	THR	3.1
1	A	551	VAL	3.0
1	A	1175	GLY	3.0
1	A	762	GLY	3.0
1	A	596	VAL	3.0
1	A	1567	VAL	2.9
1	A	548	LEU	2.8
1	A	1364	GLN	2.8
1	A	247	PRO	2.6
1	A	498	PHE	2.6
1	A	1099	LEU	2.6
1	A	311	ILE	2.6
1	A	234	LEU	2.5
1	A	802	LEU	2.5
1	A	202	PHE	2.5
1	A	988	LEU	2.5
1	A	926	GLN	2.5
1	A	349	LYS	2.5
1	A	248	LEU	2.5
1	A	172	LYS	2.5
1	A	262	ALA	2.5
1	A	808	LYS	2.5
1	A	517	GLY	2.4
1	A	353	THR	2.4
1	A	515	LEU	2.4
1	A	89	PHE	2.4
1	A	344	LEU	2.4
1	A	767	LEU	2.4
1	A	174	ALA	2.4
1	A	1152	GLN	2.3
1	A	281	LEU	2.3
1	A	625	THR	2.3
1	A	927	VAL	2.3
1	A	817	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	924	TYR	2.3
1	A	674	GLU	2.3
1	A	1630	TRP	2.3
1	A	91	ARG	2.3
1	A	358	LEU	2.3
1	A	638	ALA	2.2
1	A	116	LEU	2.2
1	A	312	GLU	2.2
1	A	1457	LEU	2.2
1	A	735	TRP	2.2
1	A	304	ASN	2.2
1	A	1177	TYR	2.2
1	A	934	LEU	2.2
1	A	191	PHE	2.2
1	A	810	TRP	2.2
1	A	1103	THR	2.2
1	A	226	LEU	2.2
1	A	54	SER	2.2
1	A	322	THR	2.1
1	A	875	VAL	2.1
1	A	113	LEU	2.1
1	A	254	LEU	2.1
1	A	261	VAL	2.1
1	A	931	GLN	2.1
1	A	400	LEU	2.1
1	A	616	LYS	2.1
1	A	766	LYS	2.1
1	A	107	TRP	2.1
1	A	1566	GLU	2.1
1	A	1546	LEU	2.1
1	A	1398	ASN	2.0
1	A	1367	ARG	2.0
1	A	192	PHE	2.0
1	A	1565	SER	2.0
1	A	105	GLY	2.0
1	A	1409	PHE	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 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.