Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 08:33 PM GMT

PDB ID	:	4TRH
Title	:	The Legionella effector SidC defines a unique family of ubiquitin ligases important for bacterial phagosomal remodeling
Authors	:	Hsu, F.S.; Luo, X.; Qiu, J.; Teng, Y.; Jin, J.; Smolka, M.B.; Luo, Z.Q.; Mao, Y.
Deposited on	:	2014-06-16
Resolution	:	2.03 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

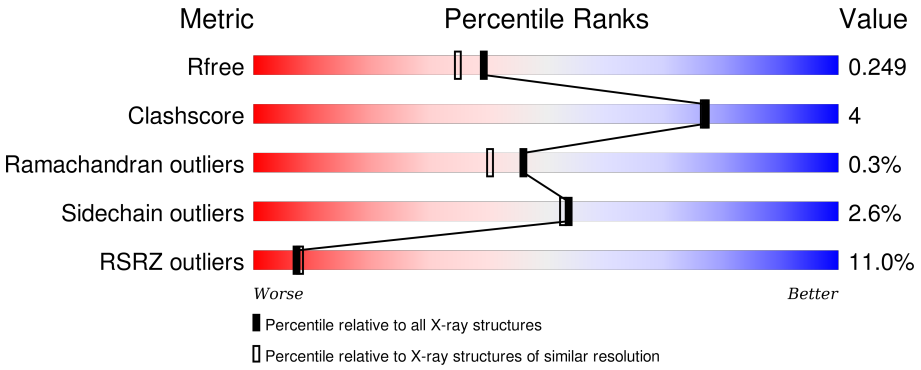
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.03 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}	91344	7858 (2.04-2.00)
Clashscore	102246	9060 (2.04-2.00)
Ramachandran outliers	100387	8952 (2.04-2.00)
Sidechain outliers	100360	8951 (2.04-2.00)
RSRZ outliers	91569	7873 (2.04-2.00)

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 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	542	10%, 85%, 8%, 6%
1	B	542	10%, 84%, 9%, 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8496 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 SidC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	509	Total	C	N	O	S	0	3	0
			4143	2618	704	815	6			
1	B	508	Total	C	N	O	S	0	1	0
			4120	2603	700	812	5			

- Molecule 2 is water.

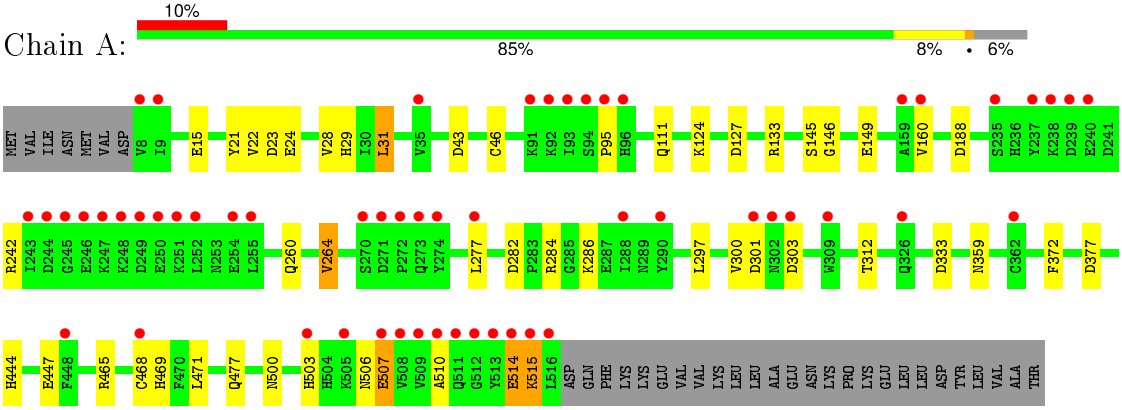
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	165	Total	O	0	0
			165	165		
2	B	68	Total	O	0	0
			68	68		

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 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.

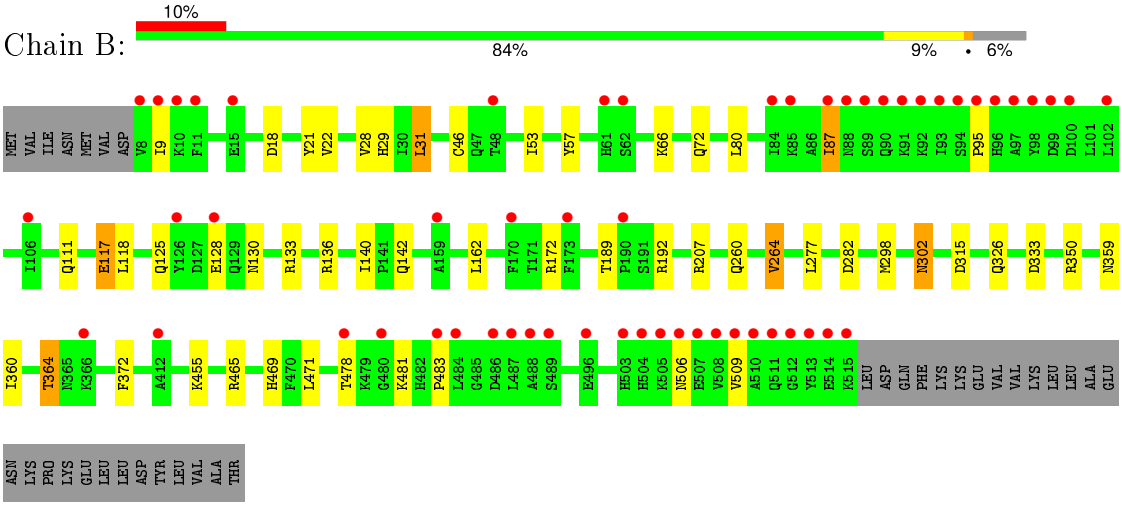
- Molecule 1: SidC

Chain A:



- Molecule 1: SidC

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.23Å 132.69Å 171.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.90 – 2.03 35.85 – 2.00	Depositor EDS
% Data completeness (in resolution range)	94.8 (35.90-2.03) 94.8 (35.85-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.74 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.213 , 0.249 0.219 , 0.249	Depositor DCC
R_{free} test set	4335 reflections (5.27%)	DCC
Wilson B-factor (Å^2)	35.2	Xtriage
Anisotropy	0.132	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 50.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 89783 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8496	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% 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

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	A	0.96	3/4230 (0.1%)	0.94	11/5709 (0.2%)
1	B	0.85	0/4207	0.88	7/5679 (0.1%)
All	All	0.91	3/8437 (0.0%)	0.91	18/11388 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	468	CYS	CB-SG	-7.82	1.69	1.82
1	A	149	GLU	CD-OE1	6.41	1.32	1.25
1	A	145	SER	CB-OG	-5.85	1.34	1.42

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	465	ARG	NE-CZ-NH2	-8.39	116.11	120.30
1	B	31	LEU	CA-CB-CG	-7.85	97.23	115.30
1	A	31	LEU	CA-CB-CG	-7.25	98.62	115.30
1	A	188	ASP	CB-CG-OD2	7.19	124.77	118.30
1	B	465	ARG	NE-CZ-NH2	-6.85	116.88	120.30
1	B	207	ARG	NE-CZ-NH1	6.83	123.72	120.30
1	B	298	MET	CG-SD-CE	-6.68	89.50	100.20
1	B	465	ARG	NE-CZ-NH1	6.67	123.63	120.30
1	A	465	ARG	NE-CZ-NH1	6.56	123.58	120.30
1	B	333	ASP	CB-CG-OD1	6.46	124.11	118.30
1	A	133	ARG	NE-CZ-NH1	6.42	123.51	120.30
1	A	43	ASP	CB-CG-OD2	5.74	123.47	118.30
1	A	377	ASP	CB-CG-OD2	-5.70	113.17	118.30
1	A	127	ASP	CB-CG-OD2	5.61	123.35	118.30
1	A	333	ASP	CB-CG-OD1	5.45	123.20	118.30
1	A	377	ASP	CB-CG-OD1	5.17	122.95	118.30
1	A	23	ASP	CB-CG-OD1	5.13	122.92	118.30
1	B	350	ARG	NE-CZ-NH1	5.06	122.83	120.30

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	4143	0	4082	26	0
1	B	4120	0	4055	38	0
2	A	165	0	0	2	1
2	B	68	0	0	1	1
All	All	8496	0	8137	61	1

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 (61) 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:469:HIS:HD2	1:A:471:LEU:H	1.33	0.77
1:B:302:ASN:HD22	1:B:302:ASN:H	1.34	0.73
1:B:481:LYS:O	1:B:483:PRO:HD3	1.89	0.73
1:B:469:HIS:HD2	1:B:471:LEU:H	1.39	0.69
1:B:9:ILE:HD11	1:B:87:ILE:CD1	2.24	0.67
1:B:364:THR:HG23	1:B:455:LYS:HE2	1.76	0.66
1:A:146:GLY:H	1:A:477:GLN:HE22	1.44	0.64
1:B:359:ASN:ND2	1:B:372:PHE:H	1.97	0.63
1:A:301:ASP:OD1	1:A:303:ASP:OD1	2.18	0.61
1:A:15:GLU:OE1	1:A:510:ALA:HB2	2.01	0.61
1:B:125:GLN:HE21	1:B:125:GLN:HA	1.66	0.61
1:A:359:ASN:ND2	1:A:372:PHE:H	2.03	0.57
1:B:360:ILE:O	1:B:364:THR:HB	2.04	0.57
1:A:242:ARG:NH1	1:A:303:ASP:O	2.38	0.56
1:A:46[B]:CYS:SG	2:A:730:HOH:O	2.52	0.55
1:A:284:ARG:HD3	1:B:53:ILE:HD13	1.89	0.55
1:A:300:VAL:HG13	1:A:312:THR:HG21	1.88	0.54
1:A:146:GLY:H	1:A:477:GLN:NE2	2.05	0.54
1:B:302:ASN:H	1:B:302:ASN:ND2	2.06	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:125:GLN:HA	1:B:125:GLN:NE2	2.24	0.52
1:A:503:HIS:HE1	1:A:507:GLU:OE1	1.93	0.52
1:B:260:GLN:O	1:B:264:VAL:HG13	2.09	0.51
1:B:364:THR:HG23	1:B:455:LYS:CE	2.39	0.51
1:A:29:HIS:ND1	1:A:469:HIS:HE1	2.09	0.50
1:B:9:ILE:HD11	1:B:87:ILE:HD13	1.92	0.50
1:A:260:GLN:O	1:A:264:VAL:HG13	2.11	0.50
1:B:364:THR:HG23	1:B:455:LYS:CD	2.41	0.50
1:B:57:TYR:CD1	1:B:136:ARG:HD3	2.47	0.50
1:B:133:ARG:HH11	1:B:133:ARG:HG2	1.78	0.49
1:B:506:ASN:HB2	1:B:509:VAL:CG1	2.42	0.49
1:B:28:VAL:H	1:B:111:GLN:NE2	2.11	0.48
1:B:481:LYS:O	1:B:483:PRO:CD	2.61	0.48
1:B:66:LYS:HA	2:B:664:HOH:O	2.14	0.48
1:A:28:VAL:H	1:A:111:GLN:NE2	2.12	0.48
1:A:28:VAL:H	1:A:111:GLN:HE22	1.62	0.48
1:A:359:ASN:HD21	1:A:372:PHE:H	1.62	0.47
1:B:18:ASP:HB2	1:B:72:GLN:NE2	2.29	0.47
1:B:359:ASN:HD21	1:B:372:PHE:H	1.63	0.47
1:B:28:VAL:H	1:B:111:GLN:HE22	1.62	0.46
1:B:364:THR:HG23	1:B:455:LYS:HD2	1.98	0.46
1:B:130:ASN:ND2	1:B:133:ARG:HH21	2.13	0.45
1:A:444:HIS:HE1	1:B:315:ASP:OD2	2.00	0.45
1:B:29:HIS:ND1	1:B:469:HIS:HE1	2.14	0.45
1:A:444:HIS:HD2	2:A:704:HOH:O	2.00	0.45
1:A:297:LEU:HD22	1:B:140:ILE:HG13	1.98	0.45
1:B:18:ASP:HB2	1:B:72:GLN:HE22	1.83	0.44
1:A:469:HIS:CD2	1:A:471:LEU:H	2.24	0.43
1:B:302:ASN:ND2	1:B:302:ASN:N	2.66	0.43
1:B:21:TYR:HA	1:B:506:ASN:HD21	1.84	0.43
1:A:284:ARG:CZ	1:A:286:LYS:HE3	2.48	0.43
1:A:21:TYR:HA	1:A:506:ASN:HD21	1.85	0.42
1:B:22:VAL:H	1:B:506:ASN:ND2	2.18	0.42
1:A:515:LYS:N	1:A:515:LYS:CD	2.83	0.42
1:A:22:VAL:H	1:A:506:ASN:ND2	2.17	0.42
1:A:160:VAL:O	1:A:447:GLU:HA	2.20	0.42
1:B:364:THR:CG2	1:B:455:LYS:HE2	2.45	0.42
1:B:142:GLN:OE1	1:B:172:ARG:CZ	2.67	0.42
1:B:189:THR:O	1:B:192:ARG:NH1	2.52	0.41
1:B:117:GLU:HG3	1:B:118:LEU:N	2.36	0.40
1:A:515:LYS:HD2	1:A:515:LYS:N	2.36	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:87:ILE:HA	1:B:87:ILE:HD12	1.77	0.40

All (1) 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:608:HOH:O	2:B:603:HOH:O[4_555]	2.06	0.14

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	510/542 (94%)	503 (99%)	5 (1%)	2 (0%)	39	32
1	B	507/542 (94%)	498 (98%)	8 (2%)	1 (0%)	52	47
All	All	1017/1084 (94%)	1001 (98%)	13 (1%)	3 (0%)	46	40

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	95	PRO
1	A	514	GLU
1	B	95	PRO

5.3.2 Protein sidechains

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	463/491 (94%)	453 (98%)	10 (2%)	60	60
1	B	460/491 (94%)	446 (97%)	14 (3%)	48	46
All	All	923/982 (94%)	899 (97%)	24 (3%)	54	53

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	GLU
1	A	31	LEU
1	A	124	LYS
1	A	264	VAL
1	A	277	LEU
1	A	282	ASP
1	A	500	ASN
1	A	507	GLU
1	A	514	GLU
1	A	515	LYS
1	B	31	LEU
1	B	46	CYS
1	B	80	LEU
1	B	87	ILE
1	B	117	GLU
1	B	128	GLU
1	B	162	LEU
1	B	264	VAL
1	B	277	LEU
1	B	282	ASP
1	B	302	ASN
1	B	326	GLN
1	B	364	THR
1	B	478	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (37) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	71	HIS
1	A	72	GLN
1	A	111	GLN
1	A	120	GLN
1	A	221	ASN
1	A	260	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	273	GLN
1	A	359	ASN
1	A	365	ASN
1	A	381	HIS
1	A	410	ASN
1	A	444	HIS
1	A	469	HIS
1	A	477	GLN
1	A	491	GLN
1	A	495	GLN
1	A	500	ASN
1	A	503	HIS
1	A	506	ASN
1	B	71	HIS
1	B	72	GLN
1	B	88	ASN
1	B	111	GLN
1	B	125	GLN
1	B	130	ASN
1	B	142	GLN
1	B	221	ASN
1	B	273	GLN
1	B	302	ASN
1	B	359	ASN
1	B	365	ASN
1	B	381	HIS
1	B	411	HIS
1	B	469	HIS
1	B	491	GLN
1	B	500	ASN
1	B	506	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

There are no carbohydrates 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

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	A	509/542 (93%)	0.60	56 (11%) 7 8	21, 38, 103, 149	3 (0%)
1	B	508/542 (93%)	0.75	56 (11%) 7 8	27, 56, 110, 157	1 (0%)
All	All	1017/1084 (93%)	0.68	112 (11%) 7 8	21, 49, 109, 157	4 (0%)

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	513	TYR	23.3
1	B	8	VAL	11.8
1	B	510	ALA	10.8
1	A	513	TYR	10.6
1	B	509	VAL	9.4
1	A	512	GLY	8.8
1	B	514	GLU	8.6
1	B	9	ILE	8.1
1	B	508	VAL	8.0
1	A	510	ALA	7.8
1	A	303	ASP	7.3
1	B	95	PRO	7.3
1	A	516	LEU	7.2
1	B	94	SER	7.1
1	A	243	ILE	6.9
1	A	245	GLY	6.7
1	A	514	GLU	6.7
1	B	96	HIS	6.3
1	B	515	LYS	6.2
1	A	511	GLN	6.1
1	B	480	GLY	5.8
1	B	91	LYS	5.7
1	A	302	ASN	5.6
1	A	244	ASP	5.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	89	SER	5.2
1	B	102	LEU	5.1
1	B	484	LEU	5.1
1	A	93	ILE	5.1
1	B	98	TYR	4.9
1	B	99	ASP	4.9
1	B	93	ILE	4.8
1	A	8	VAL	4.6
1	A	246	GLU	4.6
1	B	511	GLN	4.5
1	B	92	LYS	4.4
1	A	240	GLU	4.4
1	B	512	GLY	4.1
1	A	250	GLU	4.0
1	B	488	ALA	4.0
1	A	272	PRO	4.0
1	A	515	LYS	4.0
1	B	97	ALA	3.9
1	A	95	PRO	3.9
1	B	503	HIS	3.9
1	A	247	LYS	3.9
1	A	251	LYS	3.9
1	A	273	GLN	3.9
1	A	468	CYS	3.8
1	B	11	PHE	3.8
1	A	507	GLU	3.8
1	A	270	SER	3.8
1	A	508	VAL	3.8
1	A	290	TYR	3.7
1	B	87	ILE	3.6
1	B	90	GLN	3.6
1	B	506	ASN	3.6
1	B	106	ILE	3.5
1	B	507	GLU	3.5
1	B	504	HIS	3.4
1	B	61	HIS	3.3
1	A	160	VAL	3.2
1	A	509	VAL	3.2
1	A	248	LYS	3.2
1	A	288	ILE	3.2
1	A	254	GLU	3.1
1	A	274	TYR	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	362	CYS	3.0
1	B	100	ASP	3.0
1	B	10	LYS	3.0
1	A	271	ASP	3.0
1	B	478	THR	2.9
1	A	91	LYS	2.9
1	A	301	ASP	2.9
1	B	126	TYR	2.9
1	B	85	LYS	2.9
1	B	486	ASP	2.7
1	B	128	GLU	2.7
1	A	159	ALA	2.7
1	B	505	LYS	2.6
1	B	483	PRO	2.6
1	A	503	HIS	2.6
1	B	15	GLU	2.6
1	B	62	SER	2.5
1	A	96	HIS	2.5
1	A	249	ASP	2.5
1	B	84	ILE	2.5
1	A	505	LYS	2.5
1	A	9	ILE	2.5
1	B	412	ALA	2.4
1	A	255	LEU	2.4
1	B	487	LEU	2.4
1	B	159	ALA	2.4
1	A	277	LEU	2.4
1	A	235	SER	2.4
1	A	238	LYS	2.4
1	B	190	PRO	2.3
1	A	94	SER	2.3
1	A	35	VAL	2.2
1	B	366	LYS	2.2
1	A	309	TRP	2.2
1	A	237	TYR	2.2
1	A	252	LEU	2.1
1	A	326	GLN	2.1
1	A	448	PHE	2.1
1	B	489	SER	2.1
1	B	48	THR	2.1
1	B	170	PHE	2.1
1	B	173	PHE	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	239	ASP	2.1
1	B	88	ASN	2.0
1	B	496	GLU	2.0
1	A	92	LYS	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.