Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 07:47 PM GMT

PDB ID	:	1H7S
Title	:	N-TERMINAL 40KDA FRAGMENT OF HUMAN PMS2
Authors	:	Guarne, A.; Junop, M.S.; Yang, W.
Deposited on	:	2001-07-10
Resolution	:	1.95 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:

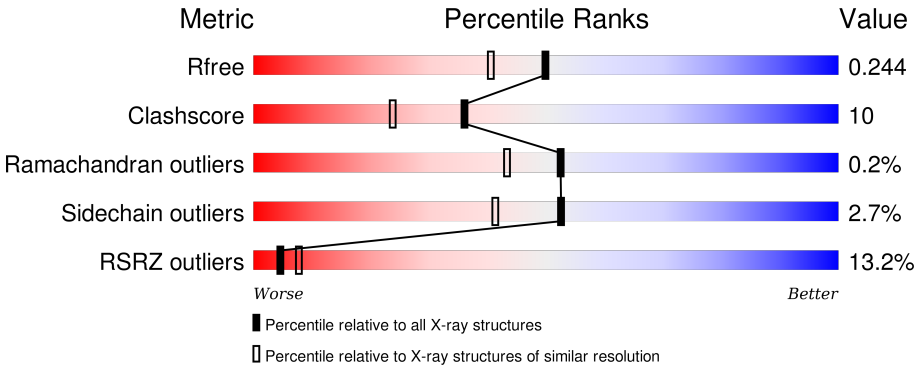
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.95 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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	365	12% 71% 16% • 12%
1	B	365	10% 64% 17% • 17%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5113 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 PMS1 PROTEIN HOMOLOG 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	321	Total	C	N	O	S	Se	0	0	0
			2486	1566	428	476	12	4			
1	B	302	Total	C	N	O	S	Se	0	0	0
			2337	1472	402	447	12	4			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	136	MSE	MET	MODIFIED RESIDUE	UNP P54278
A	184	MSE	MET	MODIFIED RESIDUE	UNP P54278
A	312	MSE	MET	MODIFIED RESIDUE	UNP P54278
A	362	MSE	MET	MODIFIED RESIDUE	UNP P54278
B	136	MSE	MET	MODIFIED RESIDUE	UNP P54278
B	184	MSE	MET	MODIFIED RESIDUE	UNP P54278
B	312	MSE	MET	MODIFIED RESIDUE	UNP P54278
B	362	MSE	MET	MODIFIED RESIDUE	UNP P54278

- Molecule 2 is water.

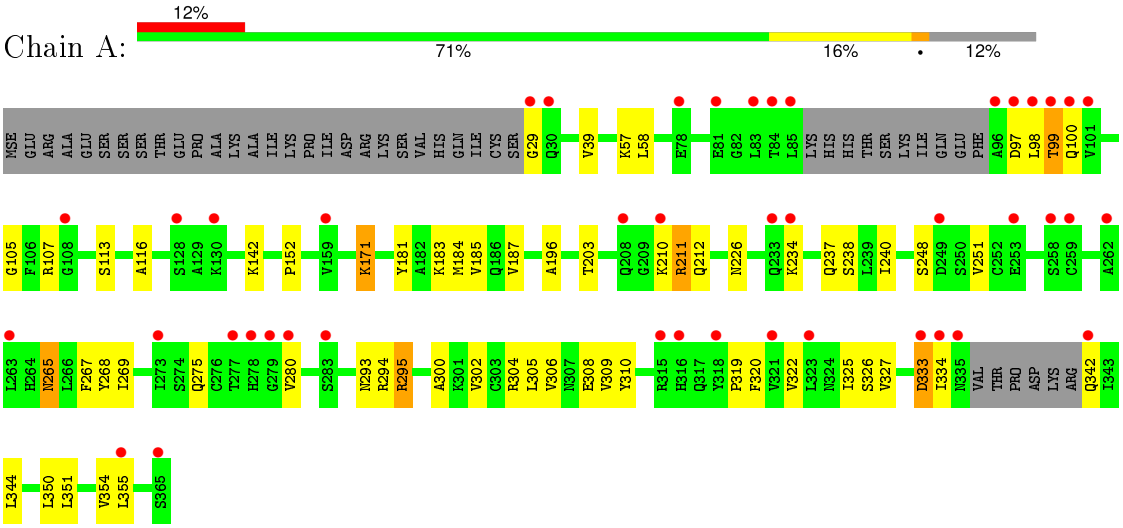
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	156	Total	O	0	0
			156	156		
2	B	134	Total	O	0	0
			134	134		

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.

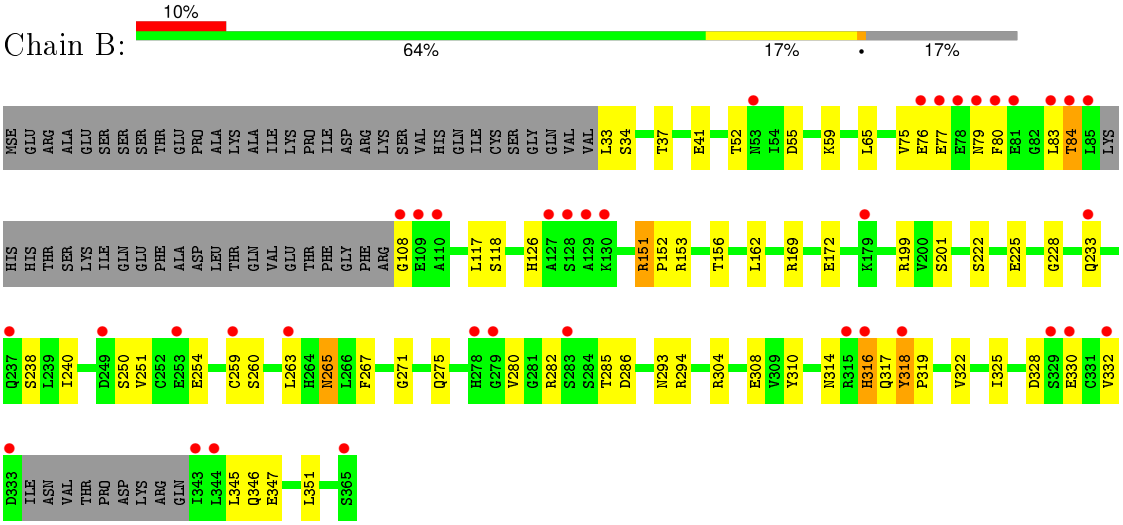
- Molecule 1: PMS1 PROTEIN HOMOLOG 2

Chain A:



- Molecule 1: PMS1 PROTEIN HOMOLOG 2

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	74.16Å 74.85Å 135.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 1.95 24.99 – 1.94	Depositor EDS
% Data completeness (in resolution range)	96.8 (19.96-1.95) 96.0 (24.99-1.94)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.26 (at 1.93Å)	Xtriage
Refinement program	CNS 0.9	Depositor
R, R_{free}	0.218 , 0.243 0.218 , 0.244	Depositor DCC
R_{free} test set	5470 reflections (10.17%)	DCC
Wilson B-factor (Å^2)	27.2	Xtriage
Anisotropy	0.306	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 50.1	EDS
Estimated twinning fraction	0.032 for k,h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 54553 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5113	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% 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	0.38	0/2523	0.66	0/3403
1	B	0.36	0/2372	0.63	0/3198
All	All	0.37	0/4895	0.64	0/6601

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	A	2486	0	2488	46	0
1	B	2337	0	2343	48	0
2	A	156	0	0	2	0
2	B	134	0	0	5	0
All	All	5113	0	4831	93	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 10.

All (93) 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:83:LEU:O	1:B:84:THR:HG23	1.70	0.91

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:151:ARG:HD3	1:B:152:PRO:O	1.75	0.87
1:A:238:SER:HB2	1:A:280:VAL:HG13	1.64	0.78
1:B:75:VAL:HG11	1:B:83:LEU:HD11	1.71	0.72
1:B:325:ILE:HD13	1:B:351:LEU:HD21	1.71	0.71
1:B:172:GLU:HG2	2:B:2069:HOH:O	1.88	0.71
1:B:77:GLU:HA	1:B:80:PHE:CE1	2.27	0.70
1:A:238:SER:HB2	1:A:280:VAL:CG1	2.23	0.68
1:A:58:LEU:HD13	1:A:184:MSE:HE1	1.76	0.67
1:A:58:LEU:CD1	1:A:184:MSE:HE1	2.24	0.67
1:B:304:ARG:O	1:B:308:GLU:HG3	1.96	0.66
1:B:238:SER:HB2	1:B:280:VAL:CG1	2.28	0.64
1:B:83:LEU:HA	1:B:108:GLY:HA2	1.81	0.63
1:A:234:LYS:HA	1:A:237:GLN:HE21	1.64	0.62
1:B:304:ARG:HD2	2:B:2130:HOH:O	2.02	0.60
1:B:238:SER:HB2	1:B:280:VAL:HG13	1.83	0.60
1:B:77:GLU:HA	1:B:80:PHE:CZ	2.37	0.59
1:A:99:THR:HG23	1:A:99:THR:O	2.03	0.59
1:A:105:GLY:HA3	1:A:113:SER:HB3	1.84	0.58
1:B:318:TYR:N	1:B:318:TYR:CD1	2.73	0.56
1:A:142:LYS:HD2	2:A:2057:HOH:O	2.05	0.56
1:B:314:ASN:OD1	1:B:317:GLN:HB2	2.06	0.56
1:A:248:SER:OG	1:A:251:VAL:HG23	2.06	0.56
1:A:183:LYS:HE3	1:A:187:VAL:HG23	1.88	0.56
1:B:83:LEU:O	1:B:84:THR:CG2	2.51	0.55
1:B:228:GLY:HA3	1:B:233:GLN:NE2	2.21	0.55
1:A:57:LYS:HE2	2:A:2097:HOH:O	2.05	0.55
1:A:203:THR:HB	1:A:211:ARG:HG3	1.89	0.55
1:A:99:THR:CG2	1:A:99:THR:O	2.55	0.54
1:B:59:LYS:HB2	1:B:65:LEU:HB3	1.90	0.53
1:A:107:ARG:CZ	1:A:116:ALA:HB2	2.38	0.53
1:A:300:ALA:O	1:A:304:ARG:HG3	2.09	0.53
1:B:260:SER:O	1:B:263:LEU:HB2	2.08	0.53
1:A:305:LEU:O	1:A:309:VAL:HG23	2.09	0.53
1:A:295:ARG:HH11	1:A:295:ARG:CG	2.21	0.53
1:A:302:VAL:O	1:A:306:VAL:HG23	2.09	0.53
1:A:240:ILE:HG23	1:A:275:GLN:HG2	1.91	0.52
1:A:350:LEU:O	1:A:354:VAL:HG23	2.10	0.52
1:A:234:LYS:HA	1:A:237:GLN:NE2	2.25	0.52
1:A:39:VAL:HG21	1:A:184:MSE:HE2	1.92	0.52
1:B:285:THR:HG23	1:B:318:TYR:CD2	2.45	0.52
1:A:334:ILE:HA	1:A:342:GLN:O	2.10	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:294:ARG:HD2	2:B:2083:HOH:O	2.10	0.51
1:B:33:LEU:HD23	1:B:117:LEU:HD12	1.93	0.51
1:B:75:VAL:HG21	1:B:83:LEU:HD11	1.93	0.51
1:B:151:ARG:NH2	1:B:156:THR:OG1	2.44	0.50
1:B:153:ARG:HB3	1:B:153:ARG:NH1	2.27	0.50
1:B:293:ASN:O	1:B:294:ARG:HB2	2.12	0.50
1:A:325:ILE:HD13	1:A:351:LEU:HD21	1.94	0.49
1:B:83:LEU:HA	1:B:108:GLY:CA	2.43	0.49
1:A:152:PRO:HB3	1:B:52:THR:HB	1.95	0.49
1:A:295:ARG:HG3	1:A:295:ARG:NH1	2.27	0.48
1:A:333:ASP:HB3	1:A:344:LEU:HB2	1.96	0.47
1:A:295:ARG:HH11	1:A:295:ARG:HG3	1.81	0.46
1:B:76:GLU:HG2	1:B:126:HIS:CD2	2.50	0.46
1:A:310:TYR:CD2	1:A:319:PRO:HD3	2.50	0.46
1:B:318:TYR:N	1:B:318:TYR:HD1	2.13	0.46
1:B:251:VAL:O	1:B:254:GLU:HB2	2.16	0.46
1:B:250:SER:O	1:B:254:GLU:HG3	2.17	0.45
1:A:320:PHE:CE1	1:A:322:VAL:HG23	2.51	0.45
1:B:328:ASP:OD1	1:B:330:GLU:HG2	2.16	0.45
1:B:282:ARG:NH2	1:B:286:ASP:OD2	2.36	0.45
1:A:98:LEU:HA	1:A:98:LEU:HD23	1.73	0.44
1:A:171:LYS:HA	1:A:171:LYS:HD3	1.76	0.44
1:B:233:GLN:NE2	2:B:2106:HOH:O	2.49	0.44
1:B:316:HIS:O	1:B:317:GLN:HG3	2.18	0.44
1:A:295:ARG:CG	1:A:295:ARG:NH1	2.81	0.44
1:B:41:GLU:HG3	2:B:2004:HOH:O	2.18	0.44
1:B:346:GLN:O	1:B:347:GLU:HB2	2.17	0.43
1:A:267:PHE:HD1	1:A:327:VAL:CG1	2.31	0.43
1:A:183:LYS:HA	1:A:183:LYS:HD2	1.76	0.43
1:B:271:GLY:HA3	1:B:322:VAL:O	2.18	0.43
1:B:118:SER:HB3	1:B:162:LEU:HA	2.01	0.43
1:B:199:ARG:NH1	1:B:199:ARG:HG2	2.34	0.43
1:A:265:ASN:C	1:A:265:ASN:HD22	2.23	0.42
1:B:240:ILE:HG23	1:B:275:GLN:HG2	2.01	0.42
1:A:293:ASN:O	1:A:294:ARG:HB2	2.18	0.42
1:B:332:VAL:HG22	1:B:345:LEU:CD2	2.50	0.42
1:B:34:SER:H	1:B:37:THR:HG1	1.68	0.42
1:A:238:SER:CB	1:A:280:VAL:HG13	2.41	0.42
1:B:55:ASP:OD1	1:B:201:SER:OG	2.38	0.42
1:A:269:ILE:HD13	1:A:355:LEU:CD2	2.50	0.42
1:B:265:ASN:ND2	1:B:267:PHE:H	2.18	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:199:ARG:HH11	1:B:199:ARG:HG2	1.84	0.41
1:A:29:GLY:HA3	1:A:100:GLN:O	2.20	0.41
1:A:268:TYR:HB2	1:A:326:SER:HB3	2.02	0.41
1:A:196:ALA:HA	1:A:226:ASN:HD21	1.86	0.41
1:A:181:TYR:O	1:A:185:VAL:HG23	2.21	0.41
1:A:210:LYS:CE	1:A:212:GLN:HE22	2.34	0.40
1:A:58:LEU:HD11	1:A:184:MSE:HE1	1.99	0.40
1:B:222:SER:HB3	1:B:225:GLU:HG3	2.03	0.40
1:A:210:LYS:HE3	1:A:212:GLN:HE22	1.87	0.40
1:B:310:TYR:CD2	1:B:319:PRO:HD3	2.56	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	315/365 (86%)	307 (98%)	8 (2%)	0	100 100
1	B	296/365 (81%)	281 (95%)	14 (5%)	1 (0%)	46 35
All	All	611/730 (84%)	588 (96%)	22 (4%)	1 (0%)	52 43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	84	THR

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	284/320 (89%)	276 (97%)	8 (3%)	51 39
1	B	268/320 (84%)	261 (97%)	7 (3%)	54 43
All	All	552/640 (86%)	537 (97%)	15 (3%)	52 41

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	97	ASP
1	A	99	THR
1	A	171	LYS
1	A	211	ARG
1	A	265	ASN
1	A	295	ARG
1	A	308	GLU
1	A	333	ASP
1	B	79	ASN
1	B	151	ARG
1	B	169	ARG
1	B	259	CYS
1	B	265	ASN
1	B	316	HIS
1	B	318	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	140	ASN
1	A	212	GLN
1	A	237	GLN
1	A	265	ASN
1	A	342	GLN
1	B	140	ASN
1	B	160	GLN
1	B	170	HIS
1	B	233	GLN
1	B	265	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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	317/365 (86%)	0.78	44 (13%)	4	6	19, 31, 59, 65	0
1	B	298/365 (81%)	0.82	37 (12%)	5	9	19, 32, 63, 73	0
All	All	615/730 (84%)	0.80	81 (13%)	4	8	19, 32, 61, 73	0

All (81) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	83	LEU	12.1
1	A	334	ILE	8.4
1	B	343	ILE	7.7
1	A	365	SER	7.4
1	B	332	VAL	7.3
1	A	96	ALA	6.9
1	B	84	THR	6.3
1	A	278	HIS	6.3
1	A	335	ASN	6.3
1	A	316	HIS	6.2
1	A	100	GLN	6.1
1	B	344	LEU	6.1
1	B	278	HIS	6.0
1	B	128	SER	5.9
1	B	85	LEU	5.7
1	A	29	GLY	5.5
1	A	130	LYS	5.4
1	B	316	HIS	5.3
1	B	130	LYS	5.2
1	A	85	LEU	4.9
1	A	99	THR	4.7
1	A	97	ASP	4.7
1	B	109	GLU	4.6
1	A	315	ARG	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	365	SER	4.3
1	A	101	VAL	4.2
1	B	315	ARG	4.2
1	A	233	GLN	4.0
1	B	259	CYS	3.9
1	A	78	GLU	3.9
1	B	78	GLU	3.9
1	B	279	GLY	3.9
1	A	98	LEU	3.9
1	B	333	ASP	3.9
1	B	80	PHE	3.7
1	B	318	TYR	3.5
1	B	253	GLU	3.5
1	A	342	GLN	3.5
1	A	279	GLY	3.5
1	A	283	SER	3.4
1	B	108	GLY	3.3
1	B	283	SER	3.3
1	B	76	GLU	3.2
1	B	129	ALA	3.2
1	A	108	GLY	3.2
1	A	259	CYS	3.2
1	B	79	ASN	3.2
1	B	81	GLU	3.1
1	A	318	TYR	3.1
1	B	127	ALA	3.0
1	A	208	GLN	2.9
1	A	333	ASP	2.8
1	A	159	VAL	2.8
1	A	128	SER	2.7
1	A	234	LYS	2.6
1	A	253	GLU	2.6
1	A	30	GLN	2.6
1	A	323	LEU	2.6
1	A	84	THR	2.5
1	A	210	LYS	2.5
1	A	280	VAL	2.4
1	B	77	GLU	2.4
1	A	277	THR	2.4
1	A	81	GLU	2.3
1	A	249	ASP	2.3
1	A	83	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	330	GLU	2.3
1	A	262	ALA	2.3
1	B	233	GLN	2.2
1	B	110	ALA	2.1
1	B	237	GLN	2.1
1	B	179	LYS	2.1
1	A	258	SER	2.1
1	A	355	LEU	2.1
1	A	321	VAL	2.1
1	B	263	LEU	2.0
1	B	329	SER	2.0
1	B	249	ASP	2.0
1	B	53	ASN	2.0
1	A	263	LEU	2.0
1	A	273	ILE	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.