



Full wwPDB X-ray Structure Validation Report i

Nov 16, 2024 – 06:44 AM EST

PDB ID : 3IA7
Title : Crystal Structure of CalG4, the Calicheamicin Glycosyltransferase
Authors : Chang, A.; Singh, S.; Bingman, C.A.; Thorson, J.S.; Phillips Jr., G.N.
Deposited on : 2009-07-13
Resolution : 1.91 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 i 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](#) i) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.20.1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.003 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

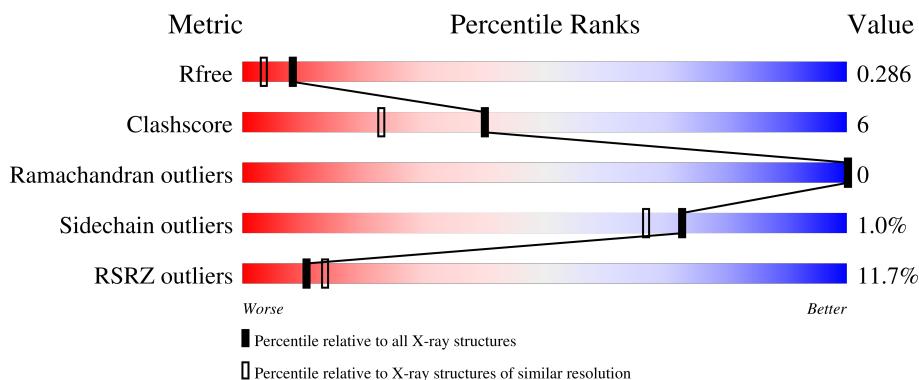
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

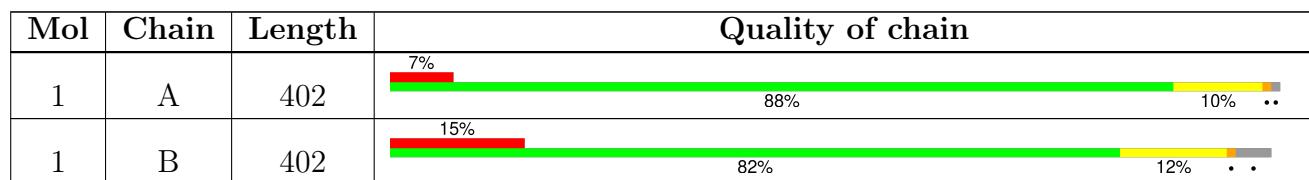
The reported resolution of this entry is 1.91 Å.

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}	164625	1028 (1.92-1.92)
Clashscore	180529	1100 (1.92-1.92)
Ramachandran outliers	177936	1087 (1.92-1.92)
Sidechain outliers	177891	1087 (1.92-1.92)
RSRZ outliers	164620	1028 (1.92-1.92)

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 $\leq 5\%$. 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.



2 Entry composition [\(i\)](#)

There are 4 unique types of molecules in this entry. The entry contains 6600 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 CalG4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	397	Total	C 3121	N 2002	O 552	S 562	Se 2	0	3	0
1	B	385	Total	C 3039	N 1947	O 542	S 545	Se 2	0	4	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	HIS	-	expression tag	UNP Q8KNC3
A	1	MSE	-	expression tag	UNP Q8KNC3
B	0	HIS	-	expression tag	UNP Q8KNC3
B	1	MSE	-	expression tag	UNP Q8KNC3

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	272	Total O 272 272	0	0

Continued on next page...

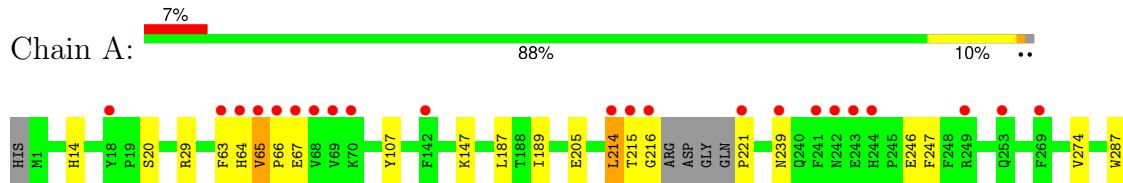
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	165	Total O 165 165	0	0

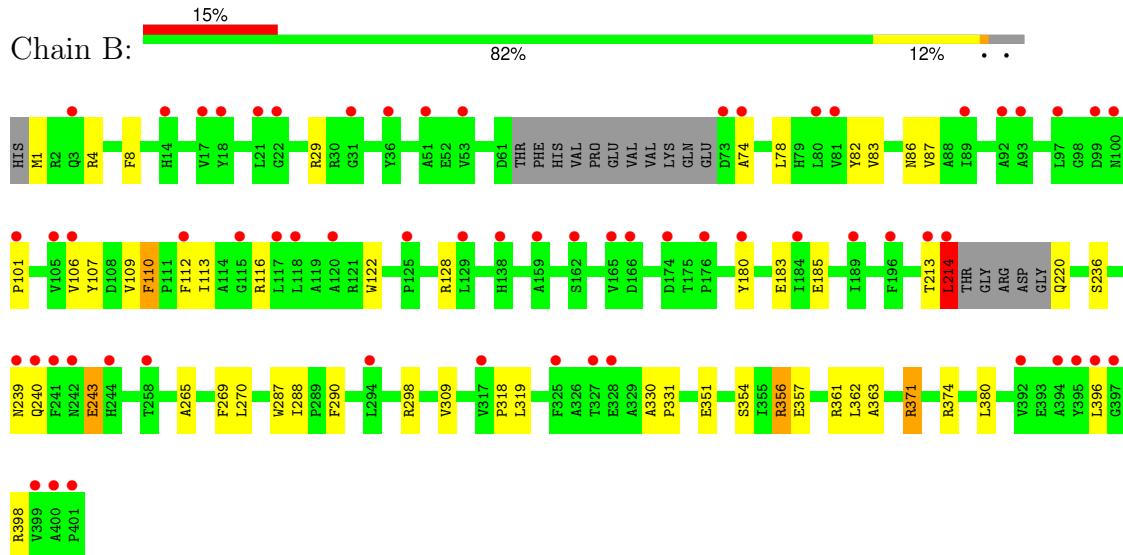
3 Residue-property plots [i](#)

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: CalG4



- Molecule 1: CalG4



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	84.75Å 66.97Å 86.62Å 90.00° 93.53° 90.00°	Depositor
Resolution (Å)	45.64 – 1.91 45.64 – 1.91	Depositor EDS
% Data completeness (in resolution range)	99.0 (45.64-1.91) 98.9 (45.64-1.91)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ ¹	2.32 (at 1.91Å)	Xtriage
Refinement program	REFMAC	Depositor
R , R_{free}	0.253 , 0.301 0.244 , 0.286	Depositor DCC
R_{free} test set	3745 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	30.1	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 43.1	EDS
L-test for twinning ²	$< L > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.013 for l,-k,h	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6600	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.33% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 i

5.1 Standard geometry i

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

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	0.89	2/3208 (0.1%)	0.78	2/4377 (0.0%)
1	B	0.85	2/3121 (0.1%)	0.77	4/4255 (0.1%)
All	All	0.87	4/6329 (0.1%)	0.77	6/8632 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	107	TYR	CD2-CE2	5.78	1.48	1.39
1	B	109	VAL	CB-CG2	-5.36	1.41	1.52
1	B	243	GLU	CD-OE2	5.31	1.31	1.25
1	A	20	SER	CB-OG	5.09	1.48	1.42

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	214	LEU	CA-CB-CG	5.87	128.80	115.30
1	A	362	LEU	CA-CB-CG	-5.84	101.87	115.30
1	B	356[A]	ARG	NE-CZ-NH2	-5.77	117.41	120.30
1	B	356[B]	ARG	NE-CZ-NH2	-5.77	117.41	120.30
1	B	371	ARG	NE-CZ-NH2	-5.45	117.58	120.30
1	A	348	ASP	CB-CG-OD1	5.38	123.14	118.30

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	3121	0	3058	46	0
1	B	3039	0	2982	42	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
4	A	272	0	0	3	0
4	B	165	0	0	4	0
All	All	6600	0	6040	79	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 6.

All (79) 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:107:TYR:OH	1:B:185:GLU:HG3	1.53	1.05
1:A:214:LEU:HD23	1:B:29:ARG:HH21	1.40	0.85
1:B:107:TYR:OH	1:B:185:GLU:CG	2.24	0.84
1:A:221:PRO:HB3	1:B:239:ASN:HB2	1.60	0.82
1:A:65:VAL:N	1:A:66:PRO:HD2	2.00	0.75
1:A:288:ILE:HD12	1:A:289:PRO:CD	2.20	0.72
1:A:63:PHE:HA	1:A:66:PRO:HG3	1.74	0.70
1:A:288:ILE:HD12	1:A:289:PRO:HD3	1.76	0.66
1:A:215:THR:HB	1:A:311:GLU:OE2	1.95	0.65
1:B:298[B]:ARG:NH2	4:B:781:HOH:O	2.31	0.64
1:A:63:PHE:HA	1:A:66:PRO:CG	2.28	0.63
1:B:213:THR:O	1:B:214:LEU:HB3	1.97	0.63
1:B:374:ARG:HG3	4:B:640:HOH:O	1.98	0.63
1:A:306:THR:HG22	1:A:332:SER:HB3	1.79	0.62
1:A:214:LEU:HD23	1:B:29:ARG:NH2	2.13	0.61
1:A:65:VAL:N	1:A:66:PRO:CD	2.63	0.61
1:A:215:THR:HG21	4:A:608:HOH:O	2.00	0.61
1:B:128[B]:ARG:NH2	1:B:185:GLU:H	1.98	0.60
1:B:330:ALA:HB3	1:B:331:PRO:HD3	1.84	0.59
1:A:64:HIS:N	1:A:66:PRO:HD2	2.18	0.58
1:A:288:ILE:CD1	1:A:289:PRO:HD2	2.33	0.58
1:A:215:THR:OG1	1:A:291[B]:HIS:HB3	2.05	0.57
1:A:290:PHE:CE2	1:B:220:GLN:HG2	2.41	0.56
1:A:64:HIS:H	1:A:66:PRO:HD2	1.69	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:265:ALA:HB2	1:B:288:ILE:HG12	1.90	0.53
1:A:221:PRO:HB3	1:B:239:ASN:CB	2.36	0.53
1:A:318:PRO:HB2	1:A:362:LEU:HD21	1.90	0.53
1:A:312:ALA:CB	1:A:319:LEU:HD11	2.39	0.53
1:A:67:GLU:HA	1:A:67:GLU:OE1	2.08	0.53
1:A:287:TRP:HD1	1:B:220:GLN:NE2	2.06	0.52
1:A:205:GLU:H	1:A:205:GLU:CD	2.13	0.52
1:A:288:ILE:CD1	1:A:289:PRO:CD	2.88	0.52
1:B:398:ARG:NH1	4:B:825:HOH:O	2.43	0.52
1:B:180:TYR:HA	1:B:183:GLU:HG3	1.92	0.51
1:A:313:PHE:CE1	1:A:319:LEU:HD13	2.46	0.50
1:B:236:SER:OG	1:B:290:PHE:HZ	1.95	0.50
1:B:4:ARG:HG3	1:B:396:LEU:HD13	1.94	0.49
1:B:101:PRO:HB3	1:B:122:TRP:CD1	2.47	0.48
1:A:380:LEU:HB2	4:A:666:HOH:O	2.14	0.48
1:B:128[B]:ARG:HH22	1:B:185:GLU:H	1.61	0.48
1:A:290:PHE:CZ	1:B:220:GLN:HG2	2.49	0.48
1:B:357:GLU:HG2	1:B:361:ARG:HH12	1.78	0.48
1:A:216:GLY:HA2	1:A:291[B]:HIS:CD2	2.49	0.47
1:A:215:THR:OG1	1:A:291[A]:HIS:HB2	2.14	0.47
1:B:1:MSE:N	1:B:1:MSE:HE2	2.29	0.46
1:B:82:TYR:HH	1:B:112:PHE:HE2	1.61	0.46
1:B:243:GLU:HB3	1:B:270:LEU:HD13	1.98	0.45
1:A:29:ARG:HG2	1:B:380:LEU:O	2.17	0.45
1:A:187:LEU:HD21	1:A:189:ILE:HD11	1.99	0.45
1:A:239:ASN:ND2	4:A:604:HOH:O	2.33	0.45
1:A:312:ALA:HB3	1:A:319:LEU:HD11	1.99	0.45
1:A:318:PRO:C	1:A:319:LEU:HD12	2.37	0.45
1:B:351:GLU:HG2	1:B:354:SER:HB3	1.99	0.44
1:A:346:ARG:HB3	1:A:347:PRO:HD2	1.98	0.44
1:A:303:HIS:CD2	1:A:305:THR:HG23	2.52	0.44
1:B:86:ASN:OD1	1:B:113:ILE:HB	2.17	0.44
1:B:318:PRO:HB2	1:B:362:LEU:HD21	2.00	0.43
1:B:309:VAL:HG22	1:B:319:LEU:HD13	2.00	0.43
1:A:288:ILE:HG13	1:A:289:PRO:HD2	2.01	0.43
1:B:110:PHE:CD2	1:B:110:PHE:C	2.92	0.43
1:A:319:LEU:CD1	1:A:319:LEU:N	2.82	0.43
1:A:366:SER:O	1:A:369[A]:ARG:HB3	2.19	0.42
1:B:371:ARG:O	1:B:374:ARG:HG2	2.19	0.42
1:A:377:ARG:HD3	1:A:377:ARG:HA	1.79	0.42
1:B:240:GLN:O	1:B:269:PHE:CD2	2.72	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:14:HIS:HE1	4:B:766:HOH:O	2.03	0.41
1:A:246:GLU:HG2	1:A:247:PHE:N	2.35	0.41
1:B:1:MSE:HE2	1:B:1:MSE:H1	1.85	0.41
1:B:101:PRO:HB3	1:B:122:TRP:CG	2.55	0.41
1:A:313:PHE:HE1	1:A:319:LEU:HD13	1.85	0.41
1:A:147:LYS:HE3	1:A:147:LYS:HB2	1.85	0.41
1:A:216:GLY:HA2	1:A:291[B]:HIS:CG	2.56	0.41
1:B:288:ILE:O	1:B:288:ILE:HG13	2.19	0.41
1:B:298[B]:ARG:HG2	1:B:363:ALA:HB1	2.02	0.41
1:B:8:PHE:HA	1:B:106:VAL:O	2.20	0.41
1:A:288:ILE:CD1	1:B:287:TRP:CD1	3.04	0.41
1:B:74:ALA:O	1:B:78:LEU:HG	2.21	0.41
1:B:112:PHE:O	1:B:116:ARG:HG3	2.21	0.40
1:B:83:VAL:O	1:B:87:VAL:HG23	2.21	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	396/402 (98%)	388 (98%)	8 (2%)	0	100 100
1	B	383/402 (95%)	377 (98%)	6 (2%)	0	100 100
All	All	779/804 (97%)	765 (98%)	14 (2%)	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	A	322/320 (101%)	319 (99%)	3 (1%)	75 70
1	B	312/320 (98%)	308 (99%)	4 (1%)	65 55
All	All	634/640 (99%)	627 (99%)	7 (1%)	73 63

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	65	VAL
1	A	214	LEU
1	A	274	VAL
1	B	110	PHE
1	B	214	LEU
1	B	356[A]	ARG
1	B	356[B]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 oligosaccharides in this entry.

5.6 Ligand geometry [\(i\)](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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(Å ²)	Q<0.9
1	A	394/402 (98%)	0.44	29 (7%) 22 28	5, 16, 35, 46	4 (1%)
1	B	382/402 (95%)	1.03	62 (16%) 5 6	7, 21, 32, 49	4 (1%)
All	All	776/804 (96%)	0.73	91 (11%) 10 14	5, 18, 33, 49	8 (1%)

All (91) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	68	VAL	7.4
1	A	241	PHE	7.0
1	A	69	VAL	6.5
1	A	65	VAL	5.8
1	B	241	PHE	5.4
1	B	401	PRO	5.2
1	A	239	ASN	4.8
1	B	99	ASP	4.6
1	B	115	GLY	4.5
1	A	64	HIS	4.1
1	B	214	LEU	4.0
1	B	159	ALA	3.9
1	A	66	PRO	3.7
1	B	399	VAL	3.7
1	B	392	VAL	3.5
1	A	243	GLU	3.5
1	A	18	TYR	3.5
1	A	329	ALA	3.4
1	B	18	TYR	3.4
1	B	184	ILE	3.4
1	B	239	ASN	3.4
1	B	162	SER	3.4
1	A	63	PHE	3.3
1	B	97	LEU	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	327	THR	3.2
1	A	67	GLU	3.2
1	B	93	ALA	3.0
1	A	221	PRO	3.0
1	B	125	PRO	3.0
1	A	216	GLY	2.9
1	B	400	ALA	2.9
1	B	394	ALA	2.9
1	B	317	VAL	2.9
1	B	106	VAL	2.9
1	B	120	ALA	2.9
1	A	330	ALA	2.8
1	B	80	LEU	2.7
1	B	14	HIS	2.7
1	B	22	GLY	2.7
1	A	351	GLU	2.7
1	B	105	VAL	2.7
1	A	215	THR	2.6
1	B	242	ASN	2.6
1	B	165	VAL	2.6
1	B	176	PRO	2.6
1	A	242	ASN	2.6
1	B	100	ASN	2.6
1	B	327	THR	2.6
1	B	112	PHE	2.5
1	B	166	ASP	2.5
1	B	180	TYR	2.5
1	B	395	TYR	2.5
1	A	244	HIS	2.5
1	B	17	VAL	2.5
1	B	31	GLY	2.5
1	B	240	GLN	2.4
1	B	325	PHE	2.4
1	B	118	LEU	2.4
1	B	138	HIS	2.3
1	B	397	GLY	2.3
1	A	253	GLN	2.3
1	B	101	PRO	2.3
1	B	213	THR	2.3
1	B	21	LEU	2.3
1	B	73	ASP	2.3
1	B	117	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	174	ASP	2.3
1	B	328	GLU	2.3
1	B	81	VAL	2.3
1	A	142	PHE	2.3
1	B	244	HIS	2.3
1	A	326	ALA	2.2
1	B	74	ALA	2.2
1	B	92	ALA	2.2
1	B	89	ILE	2.2
1	B	3	GLN	2.2
1	A	348	ASP	2.2
1	A	269	PHE	2.2
1	B	53	VAL	2.1
1	A	290	PHE	2.1
1	B	196	PHE	2.1
1	B	396	LEU	2.1
1	A	249	ARG	2.1
1	B	294	LEU	2.1
1	B	189	ILE	2.1
1	A	214	LEU	2.0
1	B	258	THR	2.0
1	B	36	TYR	2.0
1	A	70	LYS	2.0
1	B	51	ALA	2.0
1	B	129	LEU	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 [\(i\)](#)

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, 95th 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(Å ²)	Q<0.9
3	CL	A	403	1/1	0.94	0.17	34,34,34,34	0
2	CA	B	402	1/1	0.99	0.04	31,31,31,31	0
2	CA	A	402	1/1	0.99	0.15	20,20,20,20	0

6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.