Full wwPDB X-ray Structure Validation Report ⓘ

Jul 28, 2025 – 10:10 PM JST

PDB ID	:	9IYW / pdb_00009iyw
Title	:	Crystal structure of chimeric KSQ-AT didomain
Authors	:	Chisuga, T.; Miyanaga, A.
Deposited on	:	2024-07-31
Resolution	:	2.00 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:

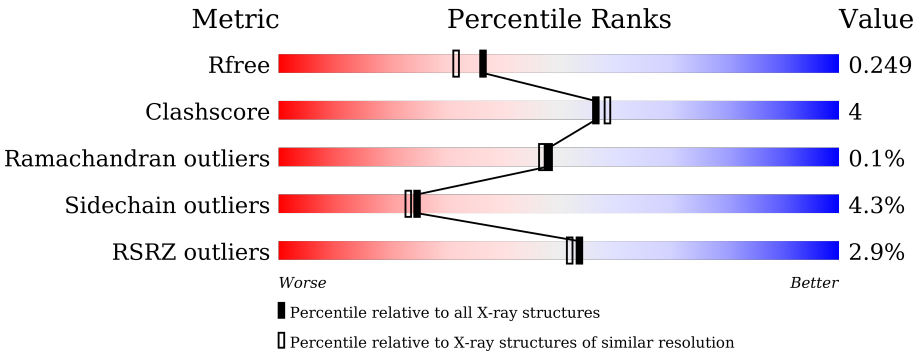
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

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 2.00 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	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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 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	A	922	3% 74% 15% • 10%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6371 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 Polyketide synthase,GfsA KSQ-AncAT chimeric protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	834	Total	C	N	O	S	0	1	0
			6193	3890	1126	1159	18			

- Molecule 2 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Na	0	0
			1	1		

- Molecule 3 is water.

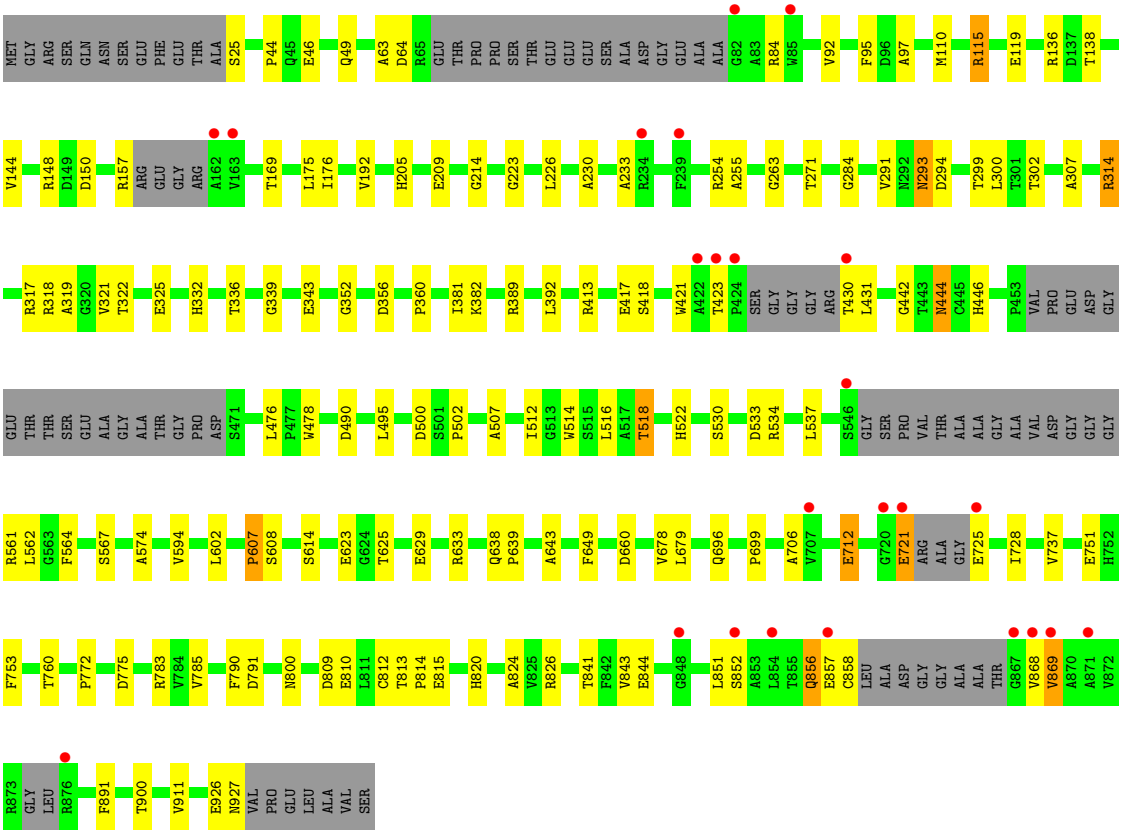
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	177	Total	O	0	0
			177	177		

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: Polyketide synthase,GfsA KSQ-AncAT chimeric protein

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	76.06Å 76.06Å 352.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.21 – 2.00 48.21 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.21-2.00) 99.8 (48.21-2.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.201 , 0.246 0.208 , 0.249	Depositor DCC
R_{free} test set	4063 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	45.7	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 36.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.048 for -h,-k,l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	6371	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	1.27	19/6318 (0.3%)	1.54	39/8593 (0.5%)

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	300	LEU	C-O	7.47	1.32	1.24
1	A	205	HIS	CE1-NE2	7.24	1.39	1.32
1	A	446	HIS	C-O	7.05	1.31	1.23
1	A	512	ILE	C-O	6.94	1.31	1.24
1	A	343	GLU	C-O	6.52	1.31	1.24
1	A	214	GLY	C-O	-6.41	1.15	1.24
1	A	205	HIS	C-O	5.93	1.30	1.24
1	A	138	THR	C-O	5.80	1.32	1.23
1	A	263	GLY	C-O	5.67	1.30	1.23
1	A	649	PHE	C-O	5.59	1.30	1.24
1	A	841	THR	C-O	5.51	1.30	1.24
1	A	339	GLY	C-O	5.43	1.31	1.23
1	A	284	GLY	C-O	5.31	1.29	1.24
1	A	293	ASN	C-O	-5.20	1.17	1.23
1	A	223	GLY	C-O	5.06	1.29	1.23
1	A	444	ASN	C-O	5.04	1.30	1.23
1	A	522	HIS	CE1-NE2	5.04	1.37	1.32
1	A	226	LEU	C-O	5.02	1.30	1.24
1	A	518	THR	C-O	-5.02	1.17	1.23

All (39) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	97	ALA	CA-C-N	7.84	128.85	119.99
1	A	97	ALA	C-N-CA	7.84	128.85	119.99

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	607	PRO	CB-CA-C	7.56	118.47	110.00
1	A	271	THR	CA-CB-OG1	-7.17	98.85	109.60
1	A	712	GLU	CB-CA-C	7.15	122.10	110.88
1	A	46	GLU	CB-CG-CD	-6.94	100.80	112.60
1	A	209	GLU	CB-CG-CD	6.69	123.97	112.60
1	A	44	PRO	CB-CA-C	-6.20	103.07	112.11
1	A	332	HIS	N-CA-C	-6.10	104.32	110.97
1	A	533	ASP	CA-C-N	5.67	128.20	120.54
1	A	533	ASP	C-N-CA	5.67	128.20	120.54
1	A	721	GLU	CB-CG-CD	5.66	122.23	112.60
1	A	820	HIS	CA-C-N	5.61	128.44	120.53
1	A	820	HIS	C-N-CA	5.61	128.44	120.53
1	A	753	PHE	CB-CA-C	5.59	120.36	110.85
1	A	381	ILE	CA-C-O	-5.59	114.85	121.05
1	A	783	ARG	CG-CD-NE	-5.57	99.74	112.00
1	A	800	ASN	CB-CA-C	5.49	119.88	110.01
1	A	92	VAL	N-CA-C	-5.47	107.79	113.10
1	A	84	ARG	N-CA-C	-5.46	106.20	112.92
1	A	352	GLY	CA-C-N	5.42	127.86	120.54
1	A	352	GLY	C-N-CA	5.42	127.86	120.54
1	A	800	ASN	N-CA-C	-5.40	106.47	113.16
1	A	119	GLU	CA-C-O	-5.39	115.16	120.82
1	A	751	GLU	CB-CA-C	5.35	120.94	110.67
1	A	574	ALA	CA-C-O	-5.32	114.86	120.92
1	A	95	PHE	CA-CB-CG	5.23	119.03	113.80
1	A	699	PRO	CA-C-N	5.18	129.63	120.87
1	A	699	PRO	C-N-CA	5.18	129.63	120.87
1	A	891	PHE	CA-C-O	-5.18	115.37	121.07
1	A	813	THR	CA-CB-OG1	-5.17	101.84	109.60
1	A	44	PRO	N-CA-C	5.17	120.34	113.57
1	A	679	LEU	CA-C-O	-5.17	115.12	120.70
1	A	336	THR	CA-CB-OG1	-5.16	101.86	109.60
1	A	314	ARG	N-CA-C	-5.13	105.77	111.36
1	A	824	ALA	CA-C-N	5.07	129.20	122.01
1	A	824	ALA	C-N-CA	5.07	129.20	122.01
1	A	869	VAL	CB-CA-C	5.06	118.17	110.62
1	A	115	ARG	CB-CG-CD	-5.06	99.66	111.30

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	6193	0	6110	55	0
2	A	1	0	0	0	0
3	A	177	0	0	4	0
All	All	6371	0	6110	55	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 4.

All (55) 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:856:GLN:HE21	1:A:868:VAL:HB	1.48	0.78
1:A:843:VAL:HG12	1:A:869:VAL:HG12	1.73	0.71
1:A:843:VAL:HA	1:A:869:VAL:O	1.99	0.62
1:A:608:SER:HA	1:A:633:ARG:NH2	2.16	0.61
1:A:607:PRO:O	1:A:696:GLN:NE2	2.34	0.61
1:A:926:GLU:O	1:A:927:ASN:ND2	2.33	0.61
1:A:507:ALA:O	1:A:534:ARG:NH2	2.35	0.59
1:A:518:THR:HG22	1:A:900:THR:OG1	2.04	0.57
1:A:851:LEU:HD13	3:A:1194:HOH:O	2.04	0.57
1:A:514:TRP:O	1:A:518:THR:HG23	2.05	0.57
1:A:594:VAL:HG11	1:A:643:ALA:HA	1.87	0.57
1:A:254:ARG:O	1:A:255:ALA:C	2.51	0.53
1:A:856:GLN:HE21	1:A:868:VAL:CB	2.19	0.52
1:A:790:PHE:HB3	1:A:812:CYS:HA	1.92	0.52
1:A:856:GLN:NE2	1:A:868:VAL:HB	2.21	0.52
1:A:561:ARG:HA	1:A:660:ASP:OD2	2.10	0.52
1:A:356:ASP:N	1:A:356:ASP:OD1	2.43	0.51
1:A:476:LEU:O	1:A:530:SER:HA	2.13	0.49
1:A:144:VAL:O	1:A:192:VAL:HA	2.13	0.48
1:A:785:VAL:HG23	1:A:814:PRO:HB3	1.95	0.48
1:A:478:TRP:CE2	1:A:537:LEU:HD13	2.49	0.48
1:A:360:PRO:HA	1:A:413:ARG:O	2.14	0.47
1:A:169:THR:HB	1:A:175:LEU:CD2	2.44	0.47
1:A:826:ARG:HD2	3:A:1197:HOH:O	2.15	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:638:GLN:HB2	1:A:639:PRO:HD3	1.98	0.46
1:A:564:PHE:CD1	1:A:843:VAL:HG22	2.51	0.46
1:A:299:THR:OG1	1:A:302:THR:OG1	2.27	0.46
1:A:725:GLU:OE2	1:A:725:GLU:HA	2.15	0.45
1:A:319:ALA:O	1:A:321:VAL:HG23	2.17	0.45
1:A:322:THR:OG1	1:A:325:GLU:HG3	2.17	0.45
1:A:110:MET:HE2	1:A:115:ARG:CG	2.47	0.44
1:A:678:VAL:HA	1:A:791:ASP:O	2.17	0.44
1:A:728:ILE:HG23	1:A:737:VAL:HG13	1.99	0.44
1:A:392:LEU:HG	1:A:421:TRP:HB2	1.99	0.44
1:A:625:THR:O	1:A:629:GLU:HG3	2.18	0.44
1:A:148:ARG:HD2	1:A:150:ASP:OD1	2.17	0.44
1:A:790:PHE:CB	1:A:812:CYS:HA	2.47	0.44
1:A:230:ALA:O	1:A:233:ALA:HB3	2.18	0.43
1:A:564:PHE:CD1	1:A:843:VAL:CG2	3.01	0.43
1:A:418:SER:HB3	3:A:1112:HOH:O	2.18	0.43
1:A:25:SER:HB3	1:A:318:ARG:O	2.17	0.43
1:A:567:SER:HA	1:A:851:LEU:HD22	2.01	0.42
1:A:382:LYS:NZ	3:A:1123:HOH:O	2.52	0.42
1:A:291:VAL:HA	1:A:444:ASN:O	2.19	0.42
1:A:421:TRP:NE1	1:A:431:LEU:HD22	2.34	0.42
1:A:495:LEU:HD21	1:A:516:LEU:HD11	2.02	0.42
1:A:169:THR:HB	1:A:175:LEU:HD23	2.01	0.42
1:A:169:THR:O	1:A:175:LEU:HD23	2.20	0.41
1:A:293:ASN:HA	1:A:442:GLY:O	2.21	0.41
1:A:706:ALA:O	1:A:760:THR:HA	2.20	0.41
1:A:844:GLU:OE1	1:A:851:LEU:HB2	2.21	0.41
1:A:294:ASP:HA	1:A:307:ALA:CB	2.51	0.40
1:A:144:VAL:HG11	1:A:176:ILE:CG2	2.51	0.40
1:A:421:TRP:CD1	1:A:431:LEU:HD22	2.56	0.40
1:A:772:PRO:O	1:A:775:ASP:HB2	2.22	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	817/922 (89%)	788 (96%)	28 (3%)	1 (0%)	48 47

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	ALA

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	622/677 (92%)	594 (96%)	28 (4%)	23 21

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	GLN
1	A	64	ASP
1	A	136[A]	ARG
1	A	136[B]	ARG
1	A	157	ARG
1	A	314	ARG
1	A	317	ARG
1	A	389	ARG
1	A	417	GLU
1	A	423	THR
1	A	430	THR
1	A	490	ASP
1	A	500	ASP
1	A	502	PRO
1	A	562	LEU
1	A	602	LEU
1	A	614	SER
1	A	623	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	712	GLU
1	A	721	GLU
1	A	809	ASP
1	A	810	GLU
1	A	815	GLU
1	A	852	SER
1	A	856	GLN
1	A	857	GLU
1	A	858	CYS
1	A	911	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	416	GLN
1	A	444	ASN
1	A	569	GLN
1	A	717	HIS
1	A	752	HIS
1	A	856	GLN
1	A	923	HIS
1	A	927	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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

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	834/922 (90%)	0.13	24 (2%) 54 52	25, 55, 83, 115	1 (0%)

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	82	GLY	4.8
1	A	424	PRO	4.5
1	A	162	ALA	4.5
1	A	848	GLY	3.9
1	A	430	THR	3.3
1	A	720	GLY	3.1
1	A	85	TRP	2.9
1	A	163	VAL	2.9
1	A	854	LEU	2.8
1	A	868	VAL	2.8
1	A	871	ALA	2.8
1	A	852	SER	2.7
1	A	423	THR	2.6
1	A	869	VAL	2.4
1	A	546	SER	2.3
1	A	725	GLU	2.3
1	A	867	GLY	2.2
1	A	239	PHE	2.2
1	A	721	GLU	2.2
1	A	707	VAL	2.2
1	A	422	ALA	2.2
1	A	857	GLU	2.1
1	A	876	ARG	2.0
1	A	234	ARG	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 oligosaccharides 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, 95^{th} 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
2	NA	A	1001	1/1	0.99	0.03	52,52,52,52	0

6.5 Other polymers (i)

There are no such residues in this entry.