Full wwPDB X-ray Structure Validation Report (i)

Feb 27, 2014 – 03:57 PM GMT

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

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.33 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		0.268
Clashscore		8
Ramachandran outliers		0
Sidechain outliers		4.8%
RSRZ outliers		3.7%

Worse — Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
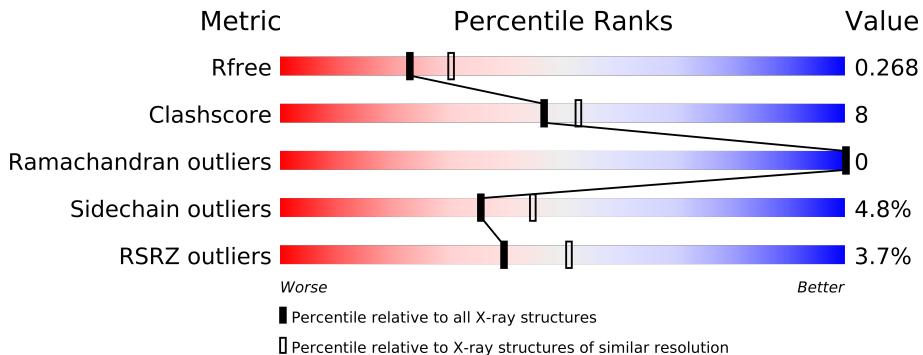
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	66092	4049 (2.38-2.30)
Clashscore	79885	1094 (2.36-2.32)
Ramachandran outliers	78287	1080 (2.36-2.32)
Sidechain outliers	78261	1081 (2.36-2.32)
RSRZ outliers	66119	4050 (2.38-2.30)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	263	
1	B	263	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4084 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Amino acid ABC transporter (ArtP).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	242	Total	C	N	O	S	0	0	0
			1896	1205	326	351	14			
1	B	242	Total	C	N	O	S	0	0	0
			1896	1205	326	351	14			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is PHOSPHOTHIOPHOSPHORICACID-ADENYLATE ESTER (three-letter code: AGS) (formula: $C_{10}H_{16}N_5O_{12}P_3S$).
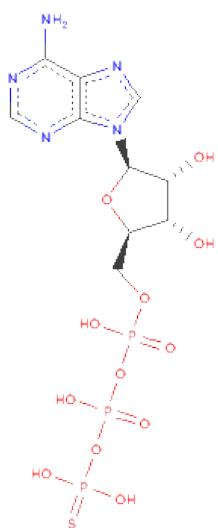
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	S	0	0
			31	10	5	12	3	1		
3	B	1	Total	C	N	O	P	S	0	0
			31	10	5	12	3	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	119	Total	O	0	0
			119	119		
4	B	109	Total	O	0	0
			109	109		

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: Amino acid ABC transporter (ArtP)

Chain A:
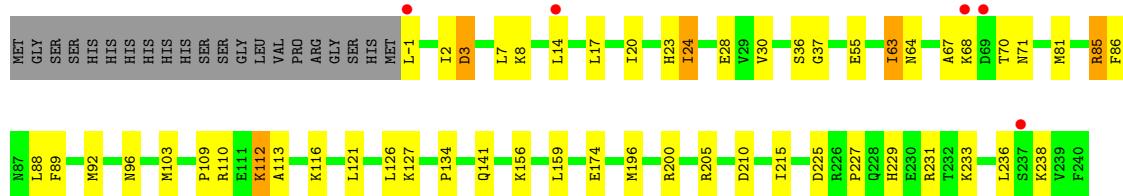
• Molecule 1: Amino acid ABC transporter (ArtP)

Chain B:
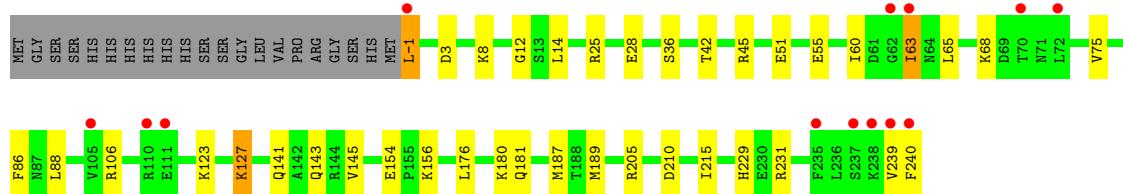
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.16Å 83.44Å 69.57Å 90.00° 107.50° 90.00°	Depositor
Resolution (Å)	41.70 – 2.33 40.87 – 2.33	Depositor EDS
% Data completeness (in resolution range)	87.4 (41.70-2.33) 87.6 (40.87-2.33)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.67 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.221 , 0.276 0.216 , 0.268	Depositor DCC
R_{free} test set	1045 reflections (5.50%)	DCC
Wilson B-factor (Å^2)	43.2	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 33.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.51, < L^2 > = 0.35$	Xtriage
Outliers	4 of 20049 reflections (0.020%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4084	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.24 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 8.2545e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0/1927	0.58	0/2589
1	B	0.36	0/1927	0.56	0/2589
All	All	0.37	0/3854	0.57	0/5178

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1896	0	1940	35	0
1	B	1896	0	1940	28	0
2	A	1	0	0	1	0
2	B	1	0	0	0	0
3	A	31	0	12	4	0
3	B	31	0	12	2	0
4	A	119	0	0	1	0
4	B	109	0	0	2	0
All	All	4084	0	3904	63	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

All (63) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:36:SER:HA	3:B:601:AGS:O2G	1.53	1.06
2:A:603:MG:MG	3:A:602:AGS:S1G	1.42	1.02
1:A:85:ARG:HH11	1:A:85:ARG:HG3	1.29	0.97
1:A:210:ASP:HB3	1:A:215:ILE:HD12	1.57	0.85
1:A:63:ILE:HD11	1:A:70:THR:HG23	1.55	0.85
1:A:85:ARG:NH1	1:A:85:ARG:HG3	2.03	0.70
1:A:229:HIS:HD2	1:A:231:ARG:H	1.45	0.64
1:B:45:ARG:NH1	4:B:614:HOH:O	2.31	0.63
1:B:210:ASP:HB3	1:B:215:ILE:HD12	1.81	0.62
1:B:229:HIS:HD2	1:B:231:ARG:H	1.48	0.62
1:A:103:MET:HE2	1:A:113:ALA:HB3	1.81	0.61
1:A:103:MET:HE1	1:A:110:ARG:HG3	1.83	0.60
1:A:2:ILE:HB	1:A:24:ILE:HG13	1.86	0.57
1:A:28:GLU:OE2	1:A:205:ARG:HD3	2.05	0.57
1:A:229:HIS:CD2	1:A:231:ARG:H	2.25	0.55
1:B:106:ARG:HH12	1:B:154:GLU:HG2	1.73	0.53
1:A:8:LYS:HB2	1:A:55:GLU:HB2	1.89	0.53
1:A:112:LYS:O	1:A:116:LYS:HG2	2.09	0.53
1:A:229:HIS:CD2	1:A:231:ARG:HB2	2.44	0.53
1:A:109:PRO:HG2	1:A:112:LYS:HB2	1.91	0.53
1:A:227:PRO:HG2	1:A:233:LYS:HG2	1.90	0.53
1:B:60:ILE:HD12	1:B:65:LEU:HD21	1.90	0.52
1:A:200:ARG:NH2	1:A:225:ASP:OD2	2.43	0.52
1:B:187:MET:HG3	1:B:189:MET:HG3	1.92	0.52
1:B:63:ILE:HD11	1:B:75:VAL:CG2	2.40	0.52
1:A:37:GLY:HA2	3:A:602:AGS:H5'2	1.93	0.50
1:A:233:LYS:NZ	4:A:664:HOH:O	2.44	0.50
1:A:81:MET:HG3	1:A:159:LEU:HB2	1.95	0.49
1:A:88:LEU:O	1:B:86:PHE:HB3	2.13	0.49
1:B:106:ARG:NH1	1:B:154:GLU:HG2	2.28	0.49
1:B:28:GLU:OE2	1:B:205:ARG:HD3	2.13	0.48
1:B:8:LYS:HB2	1:B:55:GLU:HG2	1.95	0.48
1:A:103:MET:HE1	1:A:110:ARG:HA	1.94	0.48
1:A:17:LEU:HD22	1:A:20:ILE:HD11	1.95	0.48
1:B:141:GLN:O	1:B:145:VAL:HG23	2.13	0.47
1:B:12:GLY:HA3	4:B:624:HOH:O	2.13	0.47
1:A:103:MET:CE	1:A:110:ARG:HA	2.45	0.46
1:B:8:LYS:HB2	1:B:55:GLU:CG	2.46	0.46
1:B:239:VAL:O	1:B:240:PHE:HB2	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:210:ASP:HB3	1:B:215:ILE:CD1	2.45	0.46
1:A:200:ARG:HH22	1:A:225:ASP:CG	2.18	0.46
3:A:602:AGS:S1G	3:A:602:AGS:O1B	2.73	0.46
1:B:25:ARG:H	1:B:25:ARG:HG2	1.62	0.46
1:A:196:MET:O	1:A:200:ARG:HG3	2.15	0.46
1:A:24:ILE:HD13	1:A:30:VAL:HG21	1.97	0.45
1:A:86:PHE:HB3	1:B:88:LEU:O	2.16	0.45
1:B:-1:LEU:N	1:B:25:ARG:HB3	2.32	0.45
1:A:3:ASP:OD1	1:A:23:HIS:HD2	2.00	0.45
1:A:36:SER:HA	3:A:602:AGS:O2G	2.17	0.44
1:B:176:LEU:O	1:B:180:LYS:HB2	2.17	0.43
1:A:64:ASN:HB3	1:A:67:ALA:HB2	2.01	0.42
1:B:156:LYS:HA	1:B:156:LYS:HD3	1.85	0.42
1:A:126:LEU:HD21	1:A:141:GLN:HG2	2.01	0.41
1:B:-1:LEU:H2	1:B:25:ARG:HB3	1.86	0.41
1:B:123:LYS:CG	1:B:181:GLN:HE22	2.33	0.41
1:B:127:LYS:H	1:B:127:LYS:HD3	1.85	0.41
1:A:89:PHE:O	1:A:134:PRO:HG2	2.20	0.41
1:A:121:LEU:HB3	1:A:127:LYS:HA	2.02	0.41
1:B:42:THR:OG1	3:B:601:AGS:O1A	2.28	0.41
1:B:45:ARG:HB3	1:B:51:GLU:HG3	2.02	0.41
1:B:143:GLN:OE1	1:B:143:GLN:HA	2.21	0.41
1:A:92:MET:CE	1:A:96:ASN:HB3	2.52	0.40
1:A:210:ASP:OD2	1:A:231:ARG:HD2	2.22	0.40

There are no symmetry-related clashes.

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	240/263 (91%)	233 (97%)	7 (3%)	0	100	100
1	B	240/263 (91%)	232 (97%)	8 (3%)	0	100	100
All	All	480/526 (91%)	465 (97%)	15 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	207/225 (92%)	193 (93%)	14 (7%)	22	25
1	B	207/225 (92%)	201 (97%)	6 (3%)	55	69
All	All	414/450 (92%)	394 (95%)	20 (5%)	35	45

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-1	LEU
1	A	3	ASP
1	A	7	LEU
1	A	14	LEU
1	A	24	ILE
1	A	63	ILE
1	A	68	LYS
1	A	71	ASN
1	A	85	ARG
1	A	112	LYS
1	A	156	LYS
1	A	174	GLU
1	A	236	LEU
1	A	238	LYS
1	B	-1	LEU
1	B	3	ASP
1	B	14	LEU
1	B	63	ILE
1	B	68	LYS
1	B	127	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	48	ASN
1	A	71	ASN
1	A	91	HIS
1	A	96	ASN
1	A	229	HIS
1	B	23	HIS
1	B	48	ASN
1	B	84	GLN
1	B	96	ASN
1	B	181	GLN
1	B	228	GLN
1	B	229	HIS

5.3.3　RNA ⓘ

There are no RNA chains 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 ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	AGS	A	602	2	33,33,33	2.01	6 (18%)	52,52,52	2.34	10 (19%)
3	AGS	B	601	2	33,33,33	1.96	5 (15%)	52,52,52	2.32	13 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	AGS	A	602	2	-	0/21/38/38	0/1/3/3
3	AGS	B	601	2	-	0/21/38/38	0/1/3/3

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	601	AGS	PG-S1G	8.99	2.07	1.90
3	A	602	AGS	PG-S1G	8.78	2.06	1.90
3	B	601	AGS	C5-C4	3.41	1.48	1.40
3	A	602	AGS	C5-C4	3.36	1.48	1.40
3	B	601	AGS	PG-O2G	-2.58	1.51	1.56
3	B	601	AGS	C4-N9	-2.49	1.34	1.37
3	A	602	AGS	C4-N9	-2.45	1.34	1.37
3	A	602	AGS	PG-O2G	-2.37	1.51	1.56
3	A	602	AGS	PG-O3G	-2.22	1.52	1.56
3	A	602	AGS	PG-O3B	2.19	1.62	1.60
3	B	601	AGS	PG-O3G	-2.17	1.52	1.56

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	602	AGS	O3B-PG-S1G	-9.22	110.44	114.53
3	B	601	AGS	O2G-PG-S1G	-8.67	103.59	112.73
3	A	602	AGS	N3-C2-N1	-6.46	123.30	128.71
3	B	601	AGS	N3-C2-N1	-6.39	123.37	128.71
3	A	602	AGS	O2G-PG-S1G	-6.22	106.17	112.73
3	B	601	AGS	N3-C4-N9	5.42	135.21	125.43
3	A	602	AGS	N3-C4-N9	5.41	135.19	125.43
3	A	602	AGS	O3G-PG-S1G	-4.76	107.71	112.73
3	B	601	AGS	O3G-PG-S1G	-4.39	108.11	112.73
3	B	601	AGS	PB-O3B-PG	-4.05	117.96	131.81
3	B	601	AGS	O3B-PG-S1G	-4.02	112.74	114.53
3	B	601	AGS	C3'-C2'-C1'	3.24	105.98	100.91
3	A	602	AGS	C5-C4-N3	-3.11	118.93	125.70
3	B	601	AGS	C5-C4-N3	-3.10	118.95	125.70
3	A	602	AGS	C4-C5-N7	-2.85	107.08	109.52
3	B	601	AGS	O3A-PB-O3B	2.73	107.22	101.66
3	B	601	AGS	C4-C5-N7	-2.65	107.25	109.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	602	AGS	PA-O3A-PB	-2.55	124.22	131.68
3	B	601	AGS	C2-N3-C4	2.43	120.92	114.01
3	A	602	AGS	C2-N3-C4	2.37	120.75	114.01
3	A	602	AGS	C8-N9-C4	2.06	108.47	106.90
3	B	601	AGS	O4'-C1'-N9	2.03	110.33	108.44
3	B	601	AGS	O3G-PG-O2G	2.01	120.86	105.74

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

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, 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	242/263 (92%)	0.40	5 (2%)	60	71	27, 39, 55, 61	0
1	B	242/263 (92%)	0.48	13 (5%)	25	35	33, 47, 62, 70	0
All	All	484/526 (92%)	0.44	18 (3%)	39	51	27, 42, 59, 70	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	240	PHE	8.3
1	A	-1	LEU	8.0
1	B	-1	LEU	6.4
1	B	239	VAL	5.8
1	A	68	LYS	4.4
1	B	72	LEU	3.4
1	B	62	GLY	3.2
1	B	235	PHE	3.0
1	B	110	ARG	2.8
1	B	238	LYS	2.8
1	A	14	LEU	2.8
1	B	63	ILE	2.6
1	A	69	ASP	2.4
1	B	70	THR	2.1
1	B	105	VAL	2.1
1	B	111	GLU	2.0
1	A	237	SER	2.0
1	B	237	SER	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 ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	AGS	B	601	31/31	0.20	1.35	47,61,64,65	4
3	AGS	A	602	31/31	0.16	0.25	40,51,55,55	4
2	MG	B	604	1/1	0.11	-0.87	37,37,37,37	0
2	MG	A	603	1/1	0.08	-6.18	41,41,41,41	0

6.5 Other polymers ⓘ

There are no such residues in this entry.