Full wwPDB X-ray Structure Validation Report ⓘ

Oct 20, 2024 – 07:12 AM EDT

PDB ID	:	3W20
Title	:	Crystal Structure of a Novel N-Substituted L-Amino Acid Dioxygenase from Burkholderia ambifaria AMMD
Authors	:	Qin, H.M.; Miyakawa, T.; Jia, M.Z.; Nakamura, A.; Ohtsuka, J.; Xue, Y.L.; Kawashima, T.; Kasahara, T.; Hibi, M.; Ogawa, J.; Tanokura, M.
Deposited on	:	2012-11-26
Resolution	:	1.77 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:

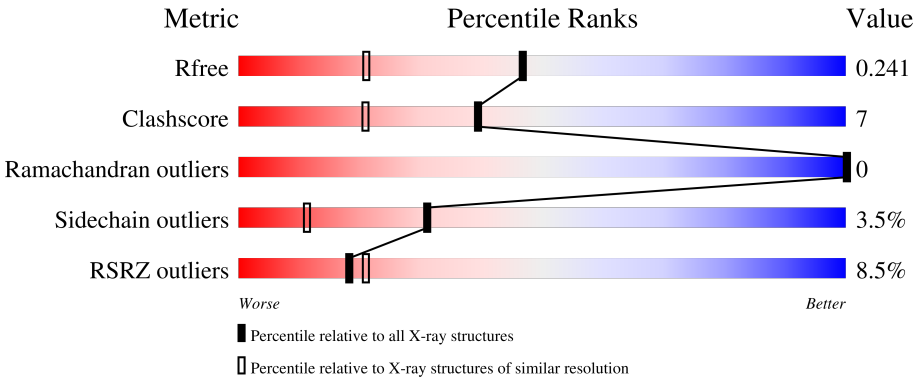
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.77 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	1191 (1.78-1.78)
Clashscore	180529	1282 (1.78-1.78)
Ramachandran outliers	177936	1270 (1.78-1.78)
Sidechain outliers	177891	1270 (1.78-1.78)
RSRZ outliers	164620	1191 (1.78-1.78)

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	273	10% 77% 14% •• 7%
1	B	273	5% 75% 13% • 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4086 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 Putative uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	254	Total	C	N	O	S	Se	0	1	0
			2004	1255	358	377	3	11			
1	B	250	Total	C	N	O	S	Se	0	0	0
			1955	1223	348	371	3	10			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is water.

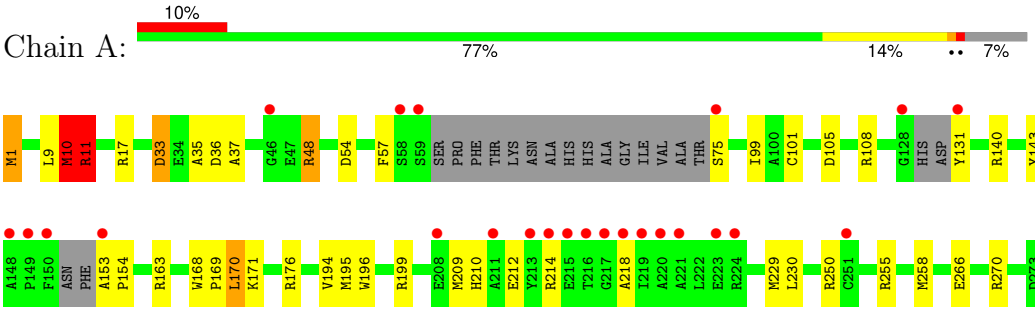
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	63	Total	O	0	0
			63	63		
3	B	62	Total	O	0	0
			62	62		

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.

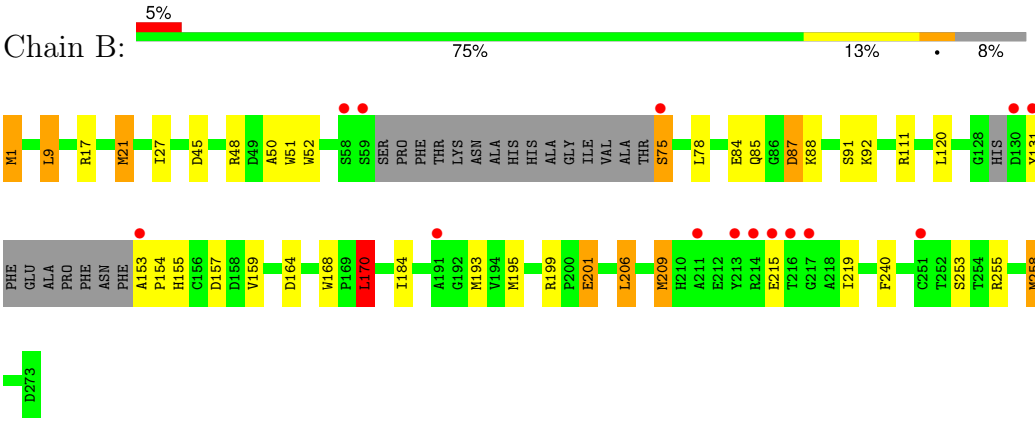
• Molecule 1: Putative uncharacterized protein

Chain A:



• Molecule 1: Putative uncharacterized protein

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	49.00Å 70.96Å 147.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.77 50.00 – 1.77	Depositor EDS
% Data completeness (in resolution range)	99.6 (50.00-1.77) 99.8 (50.00-1.77)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	8.10 (at 1.77Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.211 , 0.241 0.209 , 0.241	Depositor DCC
R_{free} test set	2596 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	22.2	Xtriage
Anisotropy	0.081	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 28.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4086	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% 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 (i)

5.1 Standard geometry (i)

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

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	1.05	3/2039 (0.1%)	1.19	16/2739 (0.6%)
1	B	1.07	3/1985 (0.2%)	1.21	16/2668 (0.6%)
All	All	1.06	6/4024 (0.1%)	1.20	32/5407 (0.6%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	168	TRP	CD2-CE2	7.05	1.49	1.41
1	B	51	TRP	CD2-CE2	6.16	1.48	1.41
1	A	196	TRP	CD2-CE2	6.14	1.48	1.41
1	B	52	TRP	CD2-CE2	5.51	1.48	1.41
1	B	168	TRP	CD2-CE2	5.40	1.47	1.41
1	A	199	ARG	CZ-NH2	-5.08	1.26	1.33

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	199	ARG	NE-CZ-NH2	-16.02	112.29	120.30
1	B	199	ARG	NE-CZ-NH2	-13.95	113.32	120.30
1	B	199	ARG	NE-CZ-NH1	11.95	126.27	120.30
1	A	199	ARG	NE-CZ-NH1	10.72	125.66	120.30
1	A	11	ARG	NE-CZ-NH2	-8.89	115.86	120.30
1	B	87	ASP	CB-CG-OD1	8.64	126.07	118.30
1	A	10	MSE	CG-SE-CE	-8.02	81.25	98.90
1	B	258	MSE	CG-SE-CE	-7.38	82.67	98.90
1	B	270	ARG	NE-CZ-NH2	-7.16	116.72	120.30
1	A	258	MSE	N-CA-CB	-7.10	97.82	110.60
1	A	163	ARG	NE-CZ-NH1	6.67	123.64	120.30
1	B	21	MSE	CB-CA-C	-6.34	97.73	110.40
1	A	11	ARG	NE-CZ-NH1	6.17	123.39	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	270	ARG	NE-CZ-NH1	5.94	123.27	120.30
1	A	199	ARG	CG-CD-NE	-5.82	99.58	111.80
1	A	195	MSE	CA-CB-CG	-5.78	103.48	113.30
1	A	176	ARG	NE-CZ-NH2	5.73	123.17	120.30
1	A	270	ARG	NE-CZ-NH1	5.65	123.12	120.30
1	B	21	MSE	CG-SE-CE	-5.64	86.49	98.90
1	B	209	MSE	CG-SE-CE	-5.64	86.49	98.90
1	B	1	MSE	CG-SE-CE	5.61	111.24	98.90
1	B	258	MSE	N-CA-CB	-5.54	100.62	110.60
1	A	209	MSE	CG-SE-CE	-5.47	86.87	98.90
1	B	170	LEU	CB-CG-CD2	-5.47	101.71	111.00
1	B	206	LEU	CB-CG-CD2	-5.46	101.72	111.00
1	A	48	ARG	NE-CZ-NH1	-5.31	117.64	120.30
1	B	157	ASP	CB-CG-OD1	5.25	123.03	118.30
1	B	199	ARG	CG-CD-NE	-5.23	100.81	111.80
1	A	33	ASP	CB-CG-OD1	5.23	123.01	118.30
1	A	270	ARG	NE-CZ-NH2	-5.22	117.69	120.30
1	B	9	LEU	CB-CG-CD2	-5.16	102.23	111.00
1	A	54	ASP	CB-CG-OD1	5.14	122.93	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	2004	0	1935	27	0
1	B	1955	0	1877	32	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	63	0	0	0	0
3	B	62	0	0	1	0
All	All	4086	0	3812	51	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 7.

All (51) 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:210:HIS:HB3	1:A:214:ARG:HH12	1.11	1.09
1:A:210:HIS:HB3	1:A:214:ARG:NH1	1.84	0.91
1:A:131:TYR:OH	1:B:75:SER:HB2	1.71	0.90
1:A:10:MSE:HG2	1:A:37:ALA:O	1.84	0.76
1:A:171:LYS:HE2	1:B:87:ASP:OD2	1.88	0.74
1:B:85:GLN:HG2	1:B:164:ASP:HB2	1.71	0.72
1:A:194:VAL:HG22	1:A:229[A]:MSE:HG2	1.72	0.72
1:A:171:LYS:CE	1:B:87:ASP:OD2	2.40	0.70
1:A:75:SER:N	1:B:131:TYR:HH	1.89	0.70
1:B:9:LEU:HD22	1:B:17:ARG:NH1	2.13	0.64
1:B:209:MSE:HB3	1:B:219:ILE:HD13	1.82	0.62
1:B:84:GLU:HG3	1:B:92:LYS:HG3	1.83	0.61
1:A:131:TYR:OH	1:B:75:SER:CB	2.49	0.59
1:A:105:ASP:HA	1:A:108:ARG:NH1	2.17	0.59
1:A:11:ARG:HD2	1:A:36:ASP:OD2	2.03	0.58
1:A:11:ARG:HD3	1:A:35:ALA:HA	1.88	0.56
1:B:27:ILE:HD13	1:B:120:LEU:HD11	1.88	0.55
1:A:10:MSE:HE2	1:A:230:LEU:HD21	1.89	0.54
1:A:131:TYR:HH	1:B:75:SER:HB2	1.71	0.53
1:B:145:PRO:HG3	1:B:253:SER:HA	1.91	0.53
1:B:209:MSE:HE3	1:B:219:ILE:HG21	1.92	0.52
1:B:17:ARG:HH21	1:B:17:ARG:HG3	1.75	0.51
1:B:88:LYS:O	1:B:92:LYS:HG2	2.10	0.51
1:A:99:ILE:HD11	1:B:131:TYR:CG	2.45	0.50
1:B:50:ALA:HB1	1:B:111:ARG:HH22	1.76	0.50
1:A:10:MSE:HB3	1:A:10:MSE:HE3	1.57	0.47
1:A:9:LEU:HD22	1:A:17:ARG:HD2	1.97	0.47
1:A:105:ASP:HA	1:A:108:ARG:HH11	1.80	0.47
1:A:99:ILE:HD11	1:B:131:TYR:CB	2.44	0.47
1:A:1:MSE:HE2	1:A:1:MSE:HB2	1.96	0.46
1:B:21:MSE:HE2	1:B:21:MSE:HB3	1.68	0.46
1:B:153:ALA:HA	1:B:154:PRO:HD3	1.70	0.45
1:B:184:ILE:HG13	1:B:258:MSE:HE3	1.98	0.44
1:B:159:VAL:HG22	1:B:170:LEU:HD23	1.99	0.43
1:B:266:GLU:H	1:B:266:GLU:HG3	1.71	0.43
1:B:193:MSE:HE1	1:B:240:PHE:CD1	2.54	0.43
1:A:143:TYR:HB2	1:A:255:ARG:HB3	2.00	0.43
1:A:153:ALA:N	1:A:154:PRO:CD	2.81	0.43
1:B:45:ASP:HA	1:B:48:ARG:CZ	2.49	0.43
1:A:212:GLU:HG2	1:A:218:ALA:O	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:201:GLU:HB2	3:B:447:HOH:O	2.19	0.42
1:A:169:PRO:HG2	1:A:170:LEU:HD23	2.01	0.42
1:B:143:TYR:HB2	1:B:255:ARG:HB3	2.02	0.42
1:B:9:LEU:C	1:B:9:LEU:HD23	2.40	0.42
1:B:78:LEU:HD21	1:B:159:VAL:HG23	2.03	0.41
1:B:195:MSE:HB3	1:B:195:MSE:HE2	1.87	0.41
1:B:215:GLU:O	1:B:215:GLU:HG3	2.21	0.41
1:A:10:MSE:HE2	1:A:230:LEU:CD2	2.52	0.40
1:A:57:PHE:CD2	1:A:140:ARG:HD3	2.56	0.40
1:A:11:ARG:NH2	1:A:33:ASP:O	2.54	0.40
1:B:155:HIS:HA	1:B:206:LEU:HD11	2.03	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	247/273 (90%)	238 (96%)	9 (4%)	0	100	100
1	B	242/273 (89%)	239 (99%)	3 (1%)	0	100	100
All	All	489/546 (90%)	477 (98%)	12 (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	205/211 (97%)	197 (96%)	8 (4%)	27 8
1	B	199/211 (94%)	193 (97%)	6 (3%)	36 15
All	All	404/422 (96%)	390 (96%)	14 (4%)	31 11

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MSE
1	A	10	MSE
1	A	11	ARG
1	A	48	ARG
1	A	101	CYS
1	A	170	LEU
1	A	250	ARG
1	A	266	GLU
1	B	1	MSE
1	B	75	SER
1	B	91	SER
1	B	170	LEU
1	B	201	GLU
1	B	266	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	233	GLN
1	B	20	HIS
1	B	233	GLN
1	B	241	GLN
1	B	244	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 2 ligands modelled in this entry, 2 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

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	244/273 (89%)	0.36	26 (10%)	12	15	12, 22, 48, 62	1 (0%)
1	B	240/273 (87%)	0.25	15 (6%)	27	33	13, 22, 46, 74	1 (0%)
All	All	484/546 (88%)	0.31	41 (8%)	18	21	12, 22, 48, 74	2 (0%)

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	150	PHE	5.7
1	A	218	ALA	5.3
1	A	153	ALA	4.8
1	A	213	TYR	4.6
1	B	213	TYR	4.1
1	B	58	SER	4.0
1	B	131	TYR	3.8
1	B	216	THR	3.8
1	B	251	CYS	3.5
1	A	131	TYR	3.5
1	B	59	SER	3.5
1	A	220	ALA	3.5
1	A	219	ILE	3.5
1	A	59	SER	3.4
1	A	217	GLY	3.3
1	A	128	GLY	3.3
1	A	146	PHE	3.1
1	A	224	ARG	3.1
1	A	208	GLU	3.1
1	A	211	ALA	2.9
1	A	75	SER	2.8
1	B	214	ARG	2.8
1	B	130	ASP	2.7
1	A	216	THR	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	214	ARG	2.6
1	B	145	PRO	2.6
1	A	149	PRO	2.6
1	B	211	ALA	2.5
1	A	251	CYS	2.4
1	B	191	ALA	2.4
1	B	215	GLU	2.4
1	A	46	GLY	2.3
1	A	223	GLU	2.2
1	A	148	ALA	2.1
1	B	75	SER	2.1
1	A	144	GLU	2.1
1	B	153	ALA	2.1
1	A	215	GLU	2.1
1	B	217	GLY	2.1
1	A	58	SER	2.1
1	A	221	ALA	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, 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	ZN	B	301	1/1	0.86	0.13	50,50,50,50	1
2	ZN	A	301	1/1	0.98	0.04	18,18,18,18	1

6.5 Other polymers (i)

There are no such residues in this entry.