WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Aug 27, 2023 – 08:44 AM EDT

PDB ID	:	3HHT
Title	:	A mutant of the nitrile hydratase from Geobacillus pallidus having enhanced thermostability
Authors	:	Van Wyk, J.C.; Sewell, B.T.; Cowan, D.A.; Sayed, M.F.; Tsekoa, T.L.; Tastan Bishop, A.O.
Deposited on	:	2009-05-17
Resolution	:	1.16 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:

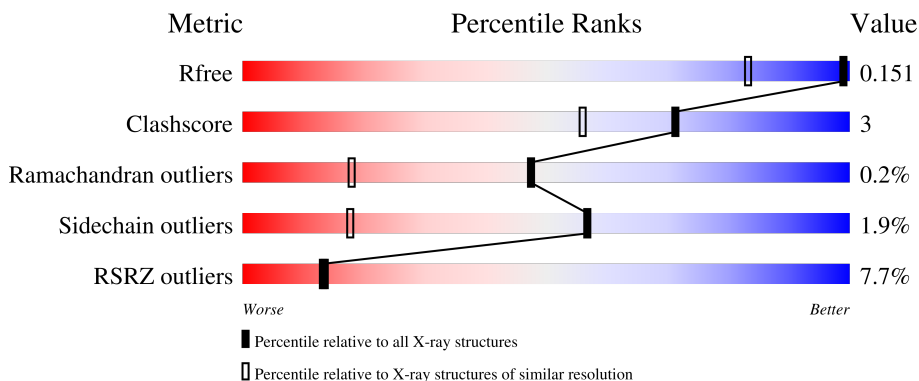|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.16 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		0.151
Clashscore		3
Ramachandran outliers		0.2%
Sidechain outliers		1.9%
RSRZ outliers		7.7%

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}	130704	1758 (1.20-1.12)
Clashscore	141614	1832 (1.20-1.12)
Ramachandran outliers	138981	1768 (1.20-1.12)
Sidechain outliers	138945	1768 (1.20-1.12)
RSRZ outliers	127900	1724 (1.20-1.12)

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	216	6% / 86% / 6% / 6%
2	B	229	8% / 87% / 12% / ·

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 4028 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 Nitrile hydratase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	202	1637	1043	284	302	8	0	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	4	GLY	ASP	engineered mutation	UNP Q84FS5

- Molecule 2 is a protein called Nitrile hydratase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	227	1869	1194	317	349	9	0	3	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	36	LEU	PHE	engineered mutation	UNP Q84FS6
B	103	SER	LEU	engineered mutation	UNP Q84FS6
B	127	ASN	TYR	engineered mutation	UNP Q84FS6

- Molecule 3 is COBALT (III) ION (three-letter code: 3CO) (formula: Co).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total 1	Cl 1	0	0

- Molecule 5 is water.

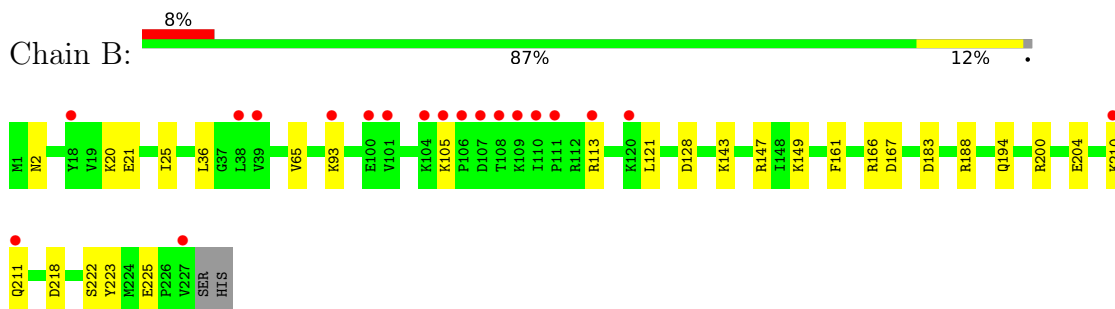| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	224	Total 224	O 224	0	0
5	B	295	Total 295	O 295	0	0

3 Residue-property plots ⓘ

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: Nitrile hydratase alpha subunit

Chain A:

6%

86%

6% 6%

MET THR ILE GLY GLN LYS ASN THR ASN I10 H16 H17 H18 E38 D45 E48 M61 T72 A75 Q78 R79 E82 H111 S120 C121 K145 D157 D160 R174 F175 M176 V177 G184 G187 M188 T189 E190 E191 I203 G204 P211

LYS VAL THR VAL GLY
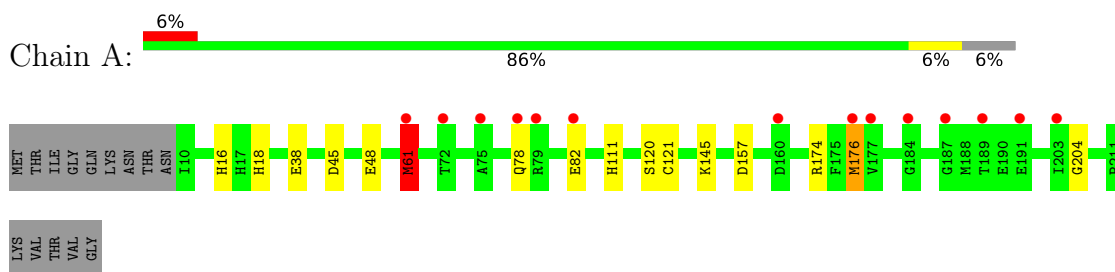
• Molecule 2: Nitrile hydratase beta subunit

Chain B:

8%

87%

12%

M1 N2 Y18 V19 K20 E21 I25 L36 G37 L38 V39 V65 K93 E100 V101 K104 K105 P106 D107 T108 K109 I110 P111 R112 R113 K120 L121 D128 K143 R147 I148 K149 F161 R166 D167 D183 R188 Q194 R200 E204 K210

Q211 D218 S222 Y223 M224 E225 P226 V227 SER HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	105.83Å 105.83Å 83.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.99 – 1.16 26.99 – 1.16	Depositor EDS
% Data completeness (in resolution range)	99.5 (26.99-1.16) 99.5 (26.99-1.16)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	4.02 (at 1.16Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.127 , 0.144 0.136 , 0.151	Depositor DCC
R_{free} test set	8144 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å²)	11.2	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.39 , 53.3	EDS
L-test for twinning [2]	$<\lvert L \rvert> = 0.53$, $<L^2> = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4028	wwPDB-VP
Average B, all atoms (Å²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<\lvert L \rvert>$, $<L^2>$ 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: CSD, CL, 3CO

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	0.77	2/1669 (0.1%)	0.89	5/2269 (0.2%)
2	B	0.84	4/1922 (0.2%)	0.93	6/2597 (0.2%)
All	All	0.81	6/3591 (0.2%)	0.91	11/4866 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	176	MET	SD-CE	-7.32	1.36	1.77
2	B	21	GLU	CD-OE1	5.88	1.32	1.25
2	B	25	ILE	CA-CB	5.58	1.67	1.54
2	B	65	VAL	CB-CG2	-5.56	1.41	1.52
2	B	204	GLU	CG-CD	-5.18	1.44	1.51
1	A	48	GLU	CG-CD	5.17	1.59	1.51

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	176	MET	CG-SD-CE	-16.51	73.79	100.20
2	B	188	ARG	NE-CZ-NH2	10.53	125.56	120.30
1	A	61[A]	MET	CG-SD-CE	7.39	112.03	100.20
1	A	61[B]	MET	CG-SD-CE	7.39	112.03	100.20
2	B	200	ARG	NE-CZ-NH1	5.92	123.26	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	218	ASP	CB-CG-OD1	-5.88	113.01	118.30
2	B	161	PHE	CB-CG-CD1	5.75	124.83	120.80
2	B	128	ASP	CB-CG-OD2	-5.69	113.18	118.30
1	A	157	ASP	CB-CG-OD1	5.62	123.35	118.30
1	A	45	ASP	CB-CG-OD1	5.56	123.31	118.30
2	B	223	TYR	CB-CG-CD1	5.30	124.18	121.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	210	LYS	Mainchain

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	1637	0	1619	12	0
2	B	1869	0	1814	12	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	224	0	0	5	0
5	B	295	0	0	5	0
All	All	4028	0	3433	24	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 3.

All (24) 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)
2:B:2:ASN:HD21	2:B:166:ARG:HH12	1.24	0.85
1:A:61[A]:MET:H	1:A:61[A]:MET:HE3	1.45	0.80
2:B:225[B]:GLU:HG2	5:B:377:HOH:O	1.89	0.71

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:225[B]:GLU:CD	5:B:377:HOH:O	2.38	0.61
1:A:16:HIS:H	1:A:16:HIS:CD2	2.20	0.60
1:A:111:HIS:HE1	1:A:204:GLY:O	1.85	0.59
2:B:225[B]:GLU:CG	5:B:377:HOH:O	2.50	0.55
1:A:61[A]:MET:H	1:A:61[A]:MET:CE	2.18	0.52
2:B:121:LEU:HD23	2:B:121:LEU:C	2.32	0.50
2:B:113:ARG:NH2	5:B:443:HOH:O	2.44	0.50
2:B:149:LYS:HE2	2:B:167:ASP:OD1	2.13	0.49
1:A:111:HIS:HD2	5:A:288:HOH:O	1.95	0.48
2:B:93:LYS:NZ	2:B:93:LYS:HB3	2.29	0.47
1:A:121:CSD:HB2	1:A:174:ARG:CZ	2.44	0.46
1:A:78:GLN:O	1:A:82:GLU:HG3	2.20	0.42
2:B:147:ARG:HH11	2:B:147:ARG:HD3	1.68	0.42
2:B:194:GLN:NE2	2:B:222:SER:H	2.18	0.42
1:A:82:GLU:HG3	5:A:305:HOH:O	2.20	0.41
2:B:93:LYS:NZ	2:B:93:LYS:CB	2.84	0.41
1:A:18:HIS:HE1	5:A:221:HOH:O	2.03	0.41
1:A:145:LYS:HE3	5:A:275:HOH:O	2.20	0.41
2:B:105:LYS:HD2	5:B:391:HOH:O	2.20	0.40
1:A:38:GLU:OE1	5:A:670:HOH:O	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	199/216 (92%)	194 (98%)	4 (2%)	1 (0%)	29	6
2	B	227/229 (99%)	223 (98%)	4 (2%)	0	100	100
All	All	426/445 (96%)	417 (98%)	8 (2%)	1 (0%)	47	18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	120	SER

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	178/189 (94%)	175 (98%)	3 (2%)	60	23
2	B	197/196 (100%)	191 (97%)	6 (3%)	41	6
All	All	375/385 (97%)	366 (98%)	9 (2%)	57	11

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	61[A]	MET
1	A	61[B]	MET
1	A	176	MET
2	B	20	LYS
2	B	36	LEU
2	B	143	LYS
2	B	183	ASP
2	B	211[A]	GLN
2	B	211[B]	GLN

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	16	HIS
1	A	18	HIS
1	A	56	HIS
1	A	78	GLN
1	A	111	HIS
2	B	2	ASN
2	B	152	ASN
2	B	194	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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	#$|Z| > 2$
1	CSD	A	121	3,1	3,7,8	0.75	0	1,8,10	0.20	0
1	CSD	A	119	1	3,7,8	0.94	0	1,8,10	0.73	0

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
1	CSD	A	121	3,1	-	0/2/6/8	-
1	CSD	A	119	1	-	0/2/6/8	-

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	121	CSD	1	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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 ⓘ

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(Å²)	Q<0.9
1	A	200/216 (92%)	0.54	14 (7%)	16	17	8, 13, 26, 35	3 (1%)
2	B	227/229 (99%)	0.63	19 (8%)	11	12	8, 12, 25, 36	7 (3%)
All	All	427/445 (95%)	0.58	33 (7%)	13	13	8, 13, 26, 36	10 (2%)

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	109	LYS	5.0
1	A	187	GLY	4.6
2	B	210	LYS	4.4
2	B	107	ASP	4.2
2	B	104	LYS	4.0
2	B	105	LYS	3.9
1	A	176	MET	3.5
2	B	227	VAL	3.5
1	A	75	ALA	3.5
2	B	106	PRO	3.1
2	B	18	TYR	3.0
1	A	189	THR	3.0
1	A	78	GLN	2.9
2	B	111	PRO	2.7
1	A	184	GLY	2.7
2	B	211[A]	GLN	2.6
1	A	160	ASP	2.5
1	A	79	ARG	2.5
2	B	110	ILE	2.5
2	B	39	VAL	2.4
2	B	38	LEU	2.4
2	B	100	GLU	2.4
1	A	82	GLU	2.4
1	A	72	THR	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	203	ILE	2.2
1	A	177	VAL	2.2
1	A	191	GLU	2.2
2	B	101	VAL	2.2
2	B	120	LYS	2.1
1	A	61[A]	MET	2.1
2	B	113	ARG	2.0
2	B	93	LYS	2.0
2	B	108	THR	2.0

6.2   Non-standard residues in protein, DNA, RNA chains ⓘ

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
1	CSD	A	119	8/9	0.99	0.07	7,8,9,10	0
1	CSD	A	121	8/9	0.99	0.06	8,9,10,11	1

6.3   Carbohydrates ⓘ

There are no monosaccharides 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. 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
4	CL	B	691	1/1	0.99	0.41	16,16,16,16	0
4	CL	A	218	1/1	1.00	0.04	12,12,12,12	0
3	3CO	A	217	1/1	1.00	0.03	7,7,7,7	0

6.5   Other polymers ⓘ

There are no such residues in this entry.