



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 05:42 am BST

PDB ID : 4OB1
Title : Crystal Structure of Nitrile Hydratase from *Pseudonocardia thermophila*
bound to Butaneboronic Acid via Co-crystallization
Authors : Rui, W.; Salette, M.; Ruslan, S.; Richard, H.; Dali, L.
Deposited on : 2014-01-06
Resolution : 1.63 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 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 : 2.11
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) : 2.11

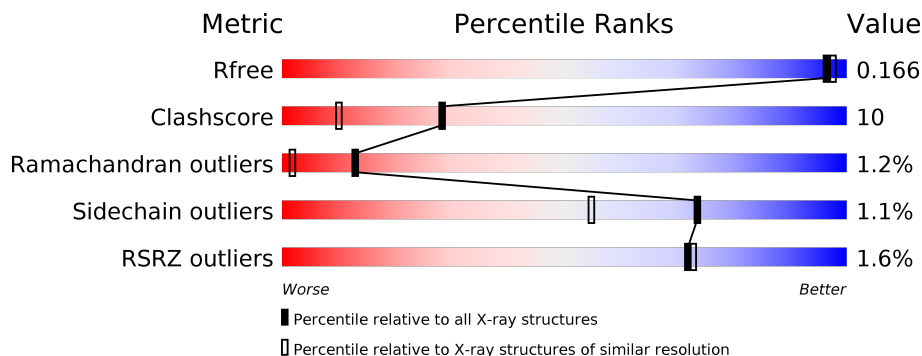
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.63 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}	130704	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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 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.

Mol	Chain	Length	Quality of chain
1	A	290	
2	B	233	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4310 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 Cobalt-containing nitrile hydratase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	203	1646	1049	271	315	11	0	7	0

There are 86 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-79	MET	-	initiating methionine	UNP Q7SID2
A	-78	THR	-	expression tag	UNP Q7SID2
A	-77	GLU	-	expression tag	UNP Q7SID2
A	-76	ASN	-	expression tag	UNP Q7SID2
A	-75	ILE	-	expression tag	UNP Q7SID2
A	-74	LEU	-	expression tag	UNP Q7SID2
A	-73	ARG	-	expression tag	UNP Q7SID2
A	-72	LYS	-	expression tag	UNP Q7SID2
A	-71	SER	-	expression tag	UNP Q7SID2
A	-70	ASP	-	expression tag	UNP Q7SID2
A	-69	GLU	-	expression tag	UNP Q7SID2
A	-68	GLU	-	expression tag	UNP Q7SID2
A	-67	ILE	-	expression tag	UNP Q7SID2
A	-66	GLN	-	expression tag	UNP Q7SID2
A	-65	LYS	-	expression tag	UNP Q7SID2
A	-64	GLU	-	expression tag	UNP Q7SID2
A	-63	ILE	-	expression tag	UNP Q7SID2
A	-62	THR	-	expression tag	UNP Q7SID2
A	-61	ALA	-	expression tag	UNP Q7SID2
A	-60	ARG	-	expression tag	UNP Q7SID2
A	-59	VAL	-	expression tag	UNP Q7SID2
A	-58	LYS	-	expression tag	UNP Q7SID2
A	-57	ALA	-	expression tag	UNP Q7SID2
A	-56	LEU	-	expression tag	UNP Q7SID2
A	-55	GLU	-	expression tag	UNP Q7SID2
A	-54	SER	-	expression tag	UNP Q7SID2
A	-53	MET	-	expression tag	UNP Q7SID2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-52	LEU	-	expression tag	UNP Q7SID2
A	-51	ILE	-	expression tag	UNP Q7SID2
A	-50	GLU	-	expression tag	UNP Q7SID2
A	-49	GLN	-	expression tag	UNP Q7SID2
A	-48	GLY	-	expression tag	UNP Q7SID2
A	-47	ILE	-	expression tag	UNP Q7SID2
A	-46	LEU	-	expression tag	UNP Q7SID2
A	-45	THR	-	expression tag	UNP Q7SID2
A	-44	THR	-	expression tag	UNP Q7SID2
A	-43	SER	-	expression tag	UNP Q7SID2
A	-42	MET	-	expression tag	UNP Q7SID2
A	-41	ILE	-	expression tag	UNP Q7SID2
A	-40	ASP	-	expression tag	UNP Q7SID2
A	-39	ARG	-	expression tag	UNP Q7SID2
A	-38	MET	-	expression tag	UNP Q7SID2
A	-37	ALA	-	expression tag	UNP Q7SID2
A	-36	GLU	-	expression tag	UNP Q7SID2
A	-35	ILE	-	expression tag	UNP Q7SID2
A	-34	TYR	-	expression tag	UNP Q7SID2
A	-33	GLU	-	expression tag	UNP Q7SID2
A	-32	ASN	-	expression tag	UNP Q7SID2
A	-31	GLU	-	expression tag	UNP Q7SID2
A	-30	VAL	-	expression tag	UNP Q7SID2
A	-29	GLY	-	expression tag	UNP Q7SID2
A	-28	PRO	-	expression tag	UNP Q7SID2
A	-27	HIS	-	expression tag	UNP Q7SID2
A	-26	LEU	-	expression tag	UNP Q7SID2
A	-25	GLY	-	expression tag	UNP Q7SID2
A	-24	ALA	-	expression tag	UNP Q7SID2
A	-23	LYS	-	expression tag	UNP Q7SID2
A	-22	VAL	-	expression tag	UNP Q7SID2
A	-21	VAL	-	expression tag	UNP Q7SID2
A	-20	VAL	-	expression tag	UNP Q7SID2
A	-19	LYS	-	expression tag	UNP Q7SID2
A	-18	ALA	-	expression tag	UNP Q7SID2
A	-17	TRP	-	expression tag	UNP Q7SID2
A	-16	THR	-	expression tag	UNP Q7SID2
A	-15	ASP	-	expression tag	UNP Q7SID2
A	-14	PRO	-	expression tag	UNP Q7SID2
A	-13	GLU	-	expression tag	UNP Q7SID2
A	-12	PHE	-	expression tag	UNP Q7SID2
A	-11	LYS	-	expression tag	UNP Q7SID2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	LYS	-	expression tag	UNP Q7SID2
A	-9	ARG	-	expression tag	UNP Q7SID2
A	-8	LEU	-	expression tag	UNP Q7SID2
A	-7	LEU	-	expression tag	UNP Q7SID2
A	-6	ALA	-	expression tag	UNP Q7SID2
A	-5	ASP	-	expression tag	UNP Q7SID2
A	-4	GLY	-	expression tag	UNP Q7SID2
A	-3	THR	-	expression tag	UNP Q7SID2
A	-2	GLU	-	expression tag	UNP Q7SID2
A	-1	ALA	-	expression tag	UNP Q7SID2
A	0	CYS	-	expression tag	UNP Q7SID2
A	205	HIS	-	expression tag	UNP Q7SID2
A	206	HIS	-	expression tag	UNP Q7SID2
A	207	HIS	-	expression tag	UNP Q7SID2
A	208	HIS	-	expression tag	UNP Q7SID2
A	209	HIS	-	expression tag	UNP Q7SID2
A	210	HIS	-	expression tag	UNP Q7SID2

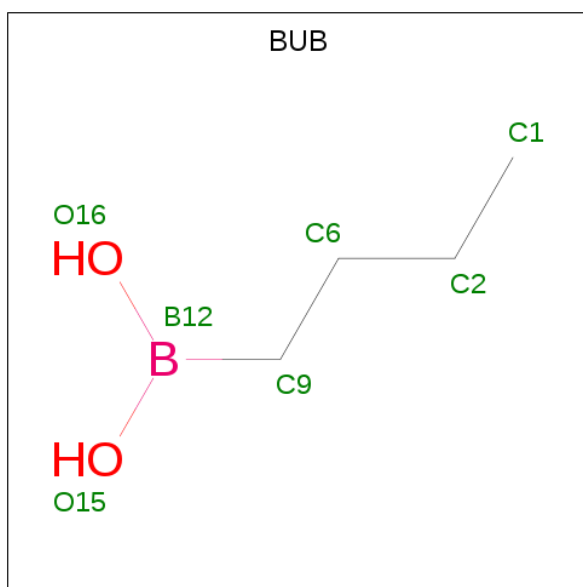
- Molecule 2 is a protein called Cobalt-containing nitrile hydratase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	228	1908	1225	329	347	7	0	11	0

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

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

- Molecule 4 is 1-BUTANE BORONIC ACID (three-letter code: BUB) (formula: C₄H₁₁BO₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	B	C	O		
4	A	1	7	1	4	2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	316	316	316	0	0
5	B	432	432	432	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	65.87Å 65.87Å 186.15Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.06 – 1.63 36.06 – 1.63	Depositor EDS
% Data completeness (in resolution range)	100.0 (36.06-1.63) 100.0 (36.06-1.63)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.97 (at 1.63Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.3_1479)	Depositor
R, R_{free}	0.142 , 0.166 0.143 , 0.166	Depositor DCC
R_{free} test set	2998 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	11.8	Xtrriage
Anisotropy	0.082	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 50.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	4310	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: BUB, CSD, CO

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	0.81	0/1694	0.82	0/2298
2	B	0.87	4/2003 (0.2%)	0.96	11/2722 (0.4%)
All	All	0.84	4/3697 (0.1%)	0.90	11/5020 (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	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	180	ASP	C-O	7.04	1.36	1.23
2	B	226	VAL	CA-C	-6.34	1.36	1.52
2	B	225	LEU	CA-C	6.05	1.68	1.52
2	B	226	VAL	CB-CG2	-5.83	1.40	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	226	VAL	CG1-CB-CG2	-10.37	94.31	110.90
2	B	226	VAL	CB-CA-C	-7.17	97.77	111.40
2	B	226	VAL	N-CA-C	-6.64	93.06	111.00
2	B	225	LEU	C-N-CA	-6.64	105.10	121.70
2	B	226	VAL	N-CA-CB	6.06	124.83	111.50

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	180	ASP	Mainchain
2	B	225	LEU	Peptide

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	1646	0	1655	17	0
2	B	1908	0	1830	53	0
3	A	1	0	0	0	0
4	A	7	0	10	0	0
5	A	316	0	0	7	0
5	B	432	0	0	15	2
All	All	4310	0	3495	69	2

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 10.

The worst 5 of 69 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:146[A]:ARG:HB2	2:B:226:VAL:HG23	1.34	1.10
2:B:167:GLY:N	5:B:707:HOH:O	1.86	1.08
2:B:146[B]:ARG:HB2	2:B:226:VAL:HG23	1.38	1.02
1:A:8:LYS:NZ	5:A:536:HOH:O	1.98	0.94
2:B:17[A]:ASN:ND2	5:B:694:HOH:O	2.01	0.93

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:B:676:HOH:O	5:B:687:HOH:O[4_555]	1.60	0.60
5:B:668:HOH:O	5:B:668:HOH:O[4_555]	2.13	0.07

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	207/290 (71%)	203 (98%)	3 (1%)	1 (0%)	29	11
2	B	237/233 (102%)	225 (95%)	8 (3%)	4 (2%)	9	0
All	All	444/523 (85%)	428 (96%)	11 (2%)	5 (1%)	13	2

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	173	HIS
2	B	225	LEU
2	B	166	THR
2	B	171	LYS
1	A	112	SER

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	185/254 (73%)	185 (100%)	0	100	100
2	B	199/191 (104%)	193 (97%)	6 (3%)	41	14
All	All	384/445 (86%)	378 (98%)	6 (2%)	73	39

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	13[B]	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	180	ASP
2	B	46[A]	MET
2	B	13[A]	LEU
2	B	46[B]	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	141	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

1 non-standard protein/DNA/RNA residue is 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	111	1	3,7,8	0.82	0	1,8,10	3.02	1 (100%)

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	111	1	-	1/2/6/8	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	111	CSD	OD1-SG-CB	-3.02	99.78	105.54

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	111	CSD	CA-CB-SG-OD1

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	$\# Z > 2$
4	BUB	A	302	3	4,6,6	4.50	1 (25%)	3,6,6	0.25	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
4	BUB	A	302	3	-	1/2/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	302	BUB	B12-C9	8.99	1.66	1.56

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	302	BUB	C2-C6-C9-B12

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	202/290 (69%)	-0.63	1 (0%) 91 91	6, 11, 24, 41	0
2	B	228/233 (97%)	-0.46	6 (2%) 56 55	6, 10, 24, 127	0
All	All	430/523 (82%)	-0.54	7 (1%) 72 73	6, 11, 24, 127	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	226	VAL	21.4
2	B	227	ASP	6.2
2	B	228	THR	4.1
2	B	170	VAL	3.7
2	B	206	PRO	2.5

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	CSD	A	111	8/9	0.99	0.08	5,8,10,11	2

6.3 Carbohydrates [i](#)

There are no carbohydrates 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
4	BUB	A	302	7/7	0.94	0.17	22,25,28,29	0
3	CO	A	301	1/1	1.00	0.05	6,6,6,6	1

6.5 Other polymers [i](#)

There are no such residues in this entry.