



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 05:21 am GMT

PDB ID : 1C0U
Title : CRYSTAL STRUCTURE OF HIV-1 REVERSE TRANSCRIPTASE IN
COMPLEX WITH BM+50.0934
Authors : Ren, J.; Esnouf, R.M.; Hopkins, A.L.; Stuart, D.I.; Stammers, D.K.
Deposited on : 1999-07-19
Resolution : 2.52 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 : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

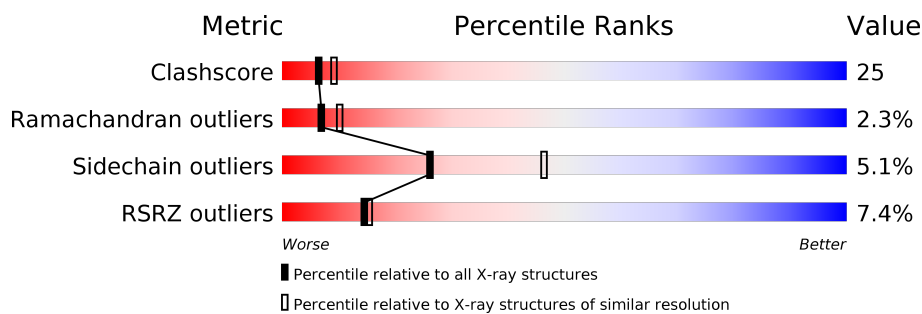
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.52 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))
Clashscore	122126	5751 (2.54-2.50)
Ramachandran outliers	120053	5650 (2.54-2.50)
Sidechain outliers	120020	5652 (2.54-2.50)
RSRZ outliers	108989	4938 (2.54-2.50)

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. 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	560	
2	B	440	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7938 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.

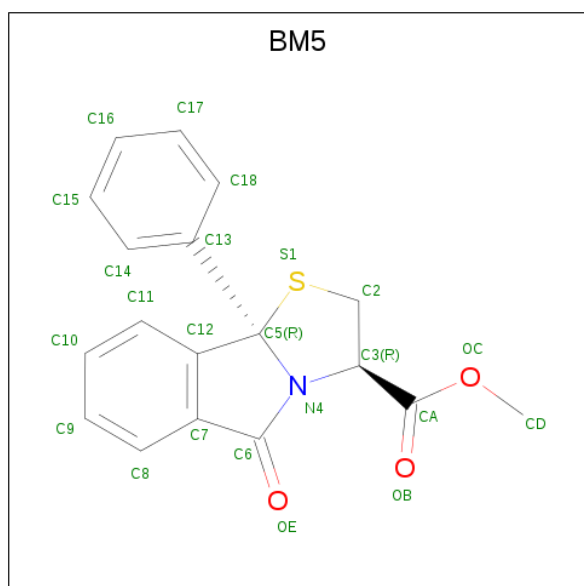
- Molecule 1 is a protein called HIV-1 REVERSE TRANSCRIPTASE (A-CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	539	Total	C	N	O	S	0	0	0
			4410	2853	734	815	8			

- Molecule 2 is a protein called HIV-1 REVERSE TRANSCRIPTASE (B-CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	414	Total	C	N	O	S	0	0	0
			3415	2219	566	623	7			

- Molecule 3 is (R)-(+)-5(9BH)-OXO-9B-PHENYL-2,3-DIHYDROTHIAZOLO[2,3-A]IS OINDOL-3-CARBOXYLIC ACID METHYL ESTER (three-letter code: BM5) (formula: C₁₈H₁₅NO₃S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			23	18	1	3	1		

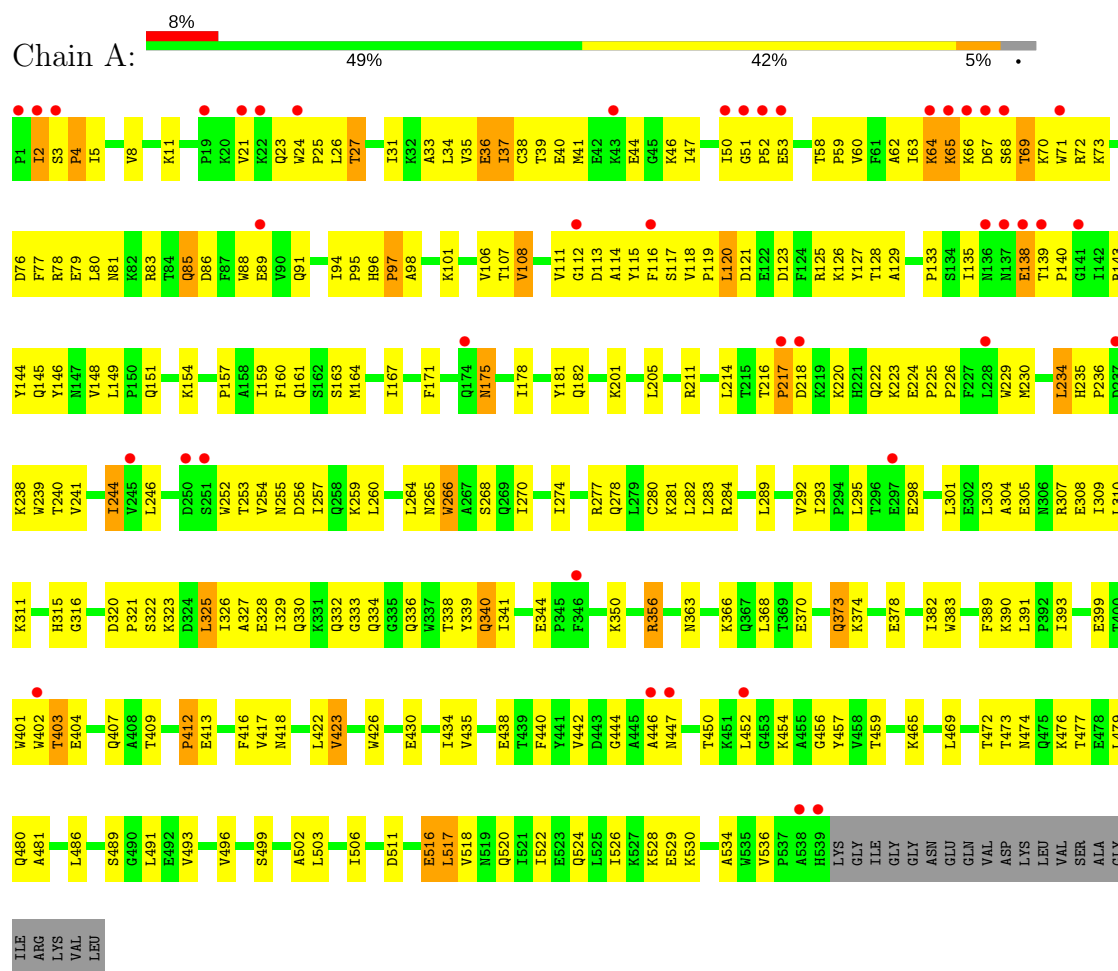
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	57	Total 57	O 57	0	0
4	B	33	Total 33	O 33	0	0

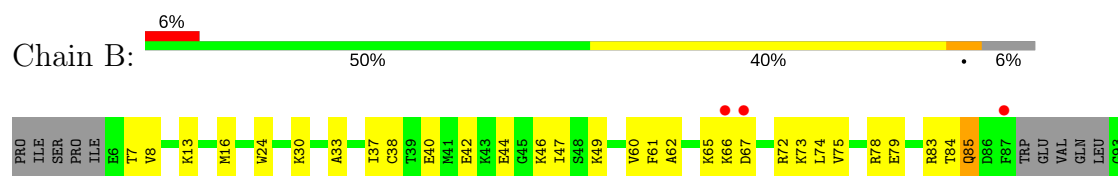
3 Residue-property plots

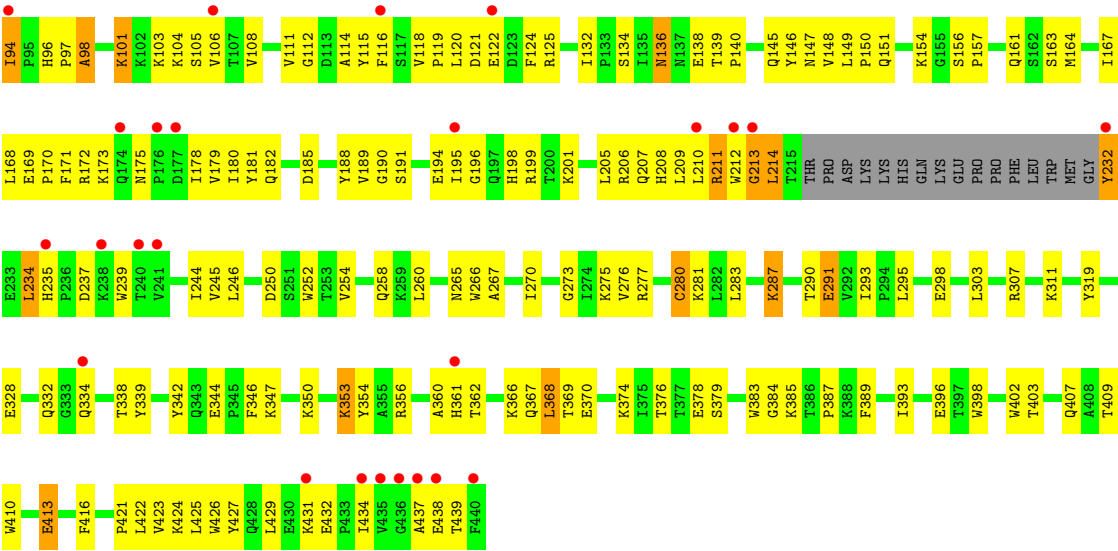
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: HIV-1 REVERSE TRANSCRIPTASE (A-CHAIN)



• Molecule 2: HIV-1 REVERSE TRANSCRIPTASE (B-CHAIN)





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	137.30Å 108.50Å 72.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.52 14.91 – 2.52	Depositor EDS
% Data completeness (in resolution range)	88.8 (15.00-2.52) 88.7 (14.91-2.52)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.78 (at 2.51Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.232 , 0.298 0.243 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	46.8	Xtriage
Anisotropy	0.387	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 58.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7938	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CSD, BM5

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	0/4518	0.65	0/6140
2	B	0.40	0/3509	0.63	0/4763
All	All	0.39	0/8027	0.64	0/10903

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	4410	0	4452	239	0
2	B	3415	0	3445	171	0
3	A	23	0	14	1	0
4	A	57	0	0	4	0
4	B	33	0	0	4	0
All	All	7938	0	7911	401	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 25.

The worst 5 of 401 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:503:LEU:HA	1:A:506:ILE:HD12	1.43	0.99
2:B:125:ARG:HE	2:B:147:ASN:HA	1.33	0.94
1:A:244:ILE:HD13	1:A:244:ILE:H	1.32	0.93
1:A:244:ILE:HD11	1:A:310:LEU:HD22	1.50	0.93
2:B:111:VAL:HB	2:B:214:LEU:HD23	1.50	0.91

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	535/560 (96%)	474 (89%)	48 (9%)	13 (2%)	6	9
2	B	408/440 (93%)	360 (88%)	39 (10%)	9 (2%)	7	11
All	All	943/1000 (94%)	834 (88%)	87 (9%)	22 (2%)	7	10

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ILE
1	A	138	GLU
2	B	360	ALA
1	A	36	GLU
1	A	37	ILE

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	483/499 (97%)	457 (95%)	26 (5%)	24	43
2	B	375/400 (94%)	357 (95%)	18 (5%)	28	49
All	All	858/899 (95%)	814 (95%)	44 (5%)	26	46

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	373	GLN
1	A	516	GLU
2	B	361	HIS
1	A	403	THR
1	A	474	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	407	GLN
1	A	500	GLN
2	B	332	GLN
1	A	480	GLN
1	A	487	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	280	1	4,7,8	0.79	0	2,8,10	1.74	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	280	1	-	1/2/6/8	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	280	CSD	CA-CB-SG-OD1

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	280	CSD	2	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand 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	$\# Z > 2$
3	BM5	A	999	-	24,26,26	2.39	5 (20%)	28,39,39	2.25	3 (10%)

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	BM5	A	999	-	-	0/12/42/42	0/4/4/4

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	999	BM5	C2-C3	-8.15	1.38	1.52
3	A	999	BM5	C2-S1	-2.92	1.76	1.80
3	A	999	BM5	C14-C13	2.12	1.42	1.39
3	A	999	BM5	C5-C13	3.12	1.58	1.53
3	A	999	BM5	C3-N4	4.58	1.52	1.46

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	999	BM5	C2-C3-N4	-2.72	103.02	106.37
3	A	999	BM5	CA-C3-N4	2.38	117.39	111.57
3	A	999	BM5	C2-C3-CA	10.69	128.54	111.37

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
3	A	999	BM5	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	280:CSD	C	281:LYS	N	2.65

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	538/560 (96%)	0.23	42 (7%) 13 14	19, 51, 97, 150	0
2	B	414/440 (94%)	0.20	28 (6%) 17 18	20, 47, 102, 148	0
All	All	952/1000 (95%)	0.22	70 (7%) 14 15	19, 49, 101, 150	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	67	ASP	8.7
2	B	440	PHE	7.6
1	A	68	SER	7.4
2	B	436	GLY	7.3
2	B	437	ALA	6.6

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	280	8/9	0.94	0.15	40,46,64,64	0

6.3 Carbohydrates [i](#)

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. 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
3	BM5	A	999	23/23	0.88	0.25	25,47,57,61	0

6.5 Other polymers

There are no such residues in this entry.