wwPDB X-ray Structure Validation Summary Report (i)

Aug 23, 2023 – 09:20 AM EDT

PDB ID	:	3BEP
Title	:	Structure of a sliding clamp on DNA
Authors	:	Georgescu, R.E.; Kim, S.S.; Yurieva, O.; Kuriyan, J.; Kong, X.-P.; O'Donnell, M.
Deposited on	:	2007-11-19
Resolution	:	1.92 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 (i) 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 (i)) were used in the production of this report:

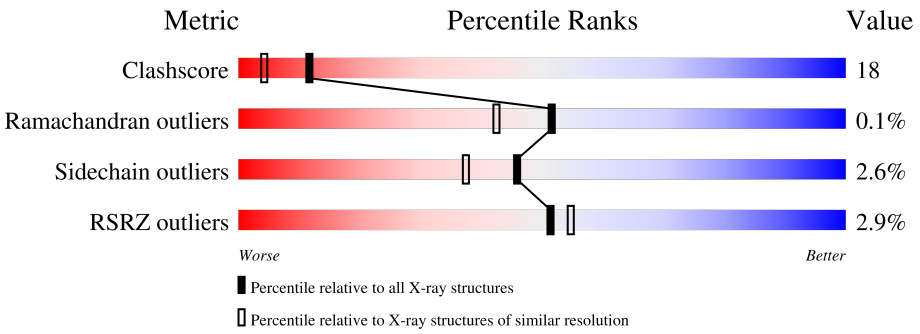
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.92 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	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	C	10	50%, 30%, 70%
2	D	14	36%, 43%, 57%
3	A	366	2%, 67%, 31%, •
3	B	366	2%, 73%, 26%, •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6797 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 DNA chain called DNA (5'-D(P*DCP*DCP*DCP*DAP*DTP*DCP*DGP*DTP*DAP*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	10	200	96	33	61	10	0	0	0

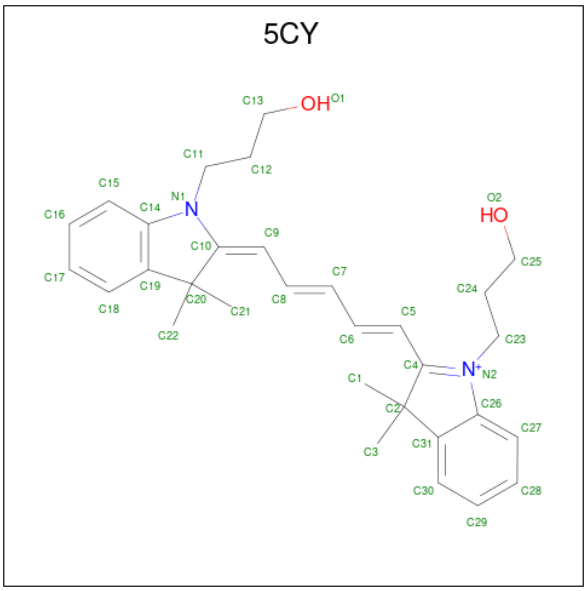
- Molecule 2 is a DNA chain called DNA (5'-D(*DTP*DTP*DTP*DTP*DAP*DTP*DAP*DCP*DGP*DAP*DTP*DGP*DGP*DG)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	14	287	139	50	85	13	0	0	0

- Molecule 3 is a protein called DNA polymerase III subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	366	2844	1786	498	541	19	0	0	0
			Total	C	N	O	S			
3	B	366	2844	1786	498	541	19	0	0	0

- Molecule 4 is 1-(3-hydroxypropyl)-2-{(1E,3E,5E)-5-[1-(3-hydroxypropyl)-3,3-dimethyl-1,3-dihydro-2H-indol-2-ylidene]penta-1,3-dien-1-y l}-3,3-dimethyl-3H-indolium (three-letter code: 5CY) (formula: $C_{31}H_{39}N_2O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	C	1	18	15	1	2	0	0

- Molecule 5 is water.

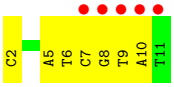
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	C	10	10	10	0	0
			Total	O		
5	D	13	13	13	0	0
			Total	O		
5	A	256	256	256	0	0
			Total	O		
5	B	325	325	325	0	0

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: DNA (5'-D(P*DCP*DCP*DCP*DAP*DTP*DCP*DGP*DTP*DAP*DT)-3')

Chain C:

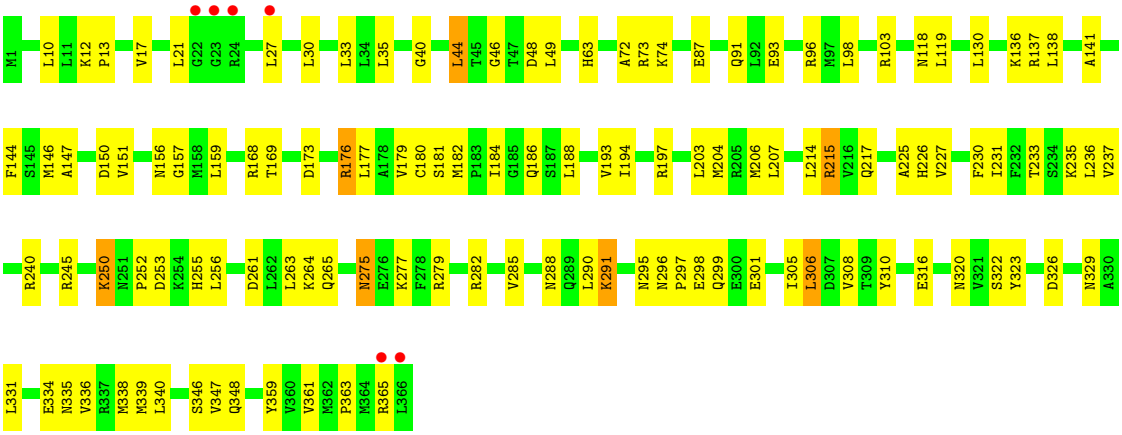


• Molecule 2: DNA (5'-D(*DTP*DTP*DTP*DTP*DAP*DTP*DAP*DCP*DGP*DAP*DTP*DGP*DGP*DG)-3')

Chain D:

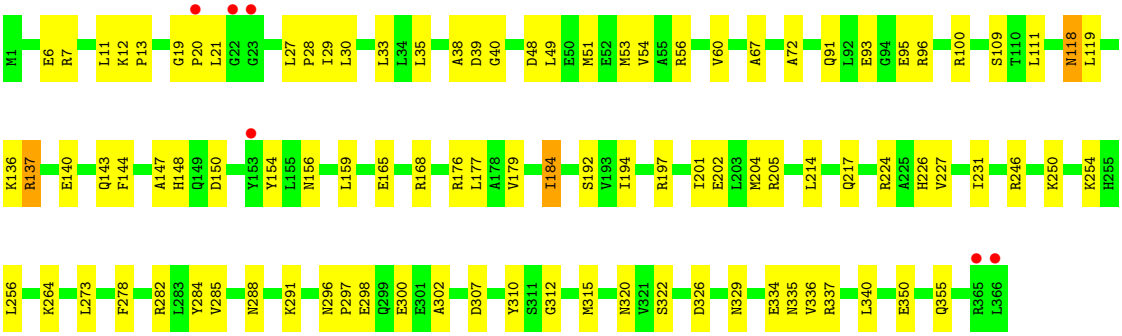
• Molecule 3: DNA polymerase III subunit beta

Chain A:



• Molecule 3: DNA polymerase III subunit beta

Chain B:


M1
E6 R7
L11 K12 P13
G19 P20 L21 G22 G23
L27 P28 I29 L30
L33 L34 L35
A38 D39 G40
D48 L49 E50 M51 E52 M53 V54 A55 R56
V60
A67
A72
Q91 L92 E93 G94 E95 R96
R100
S109 T110 L111
N118 L119
K136 R137 E140
Q143 F144
A147 H148 Q149 D150
Y153 Y154 L155 N156
L159
E165
R168
R176 L177 A178 V179
I184
S192 V193 I194
R197
I201 E202 L203 M204 R205
L214
Q217
R224 A225 H226 V227
I231
R246
K250
K254 H255
L256
K264
L273
F278
R282 L283 Y284 V285
N288
K291
N296 P297 E298 Q299 E300 E301 A302
D307
Y310 S311 G312
M315
N320 V321 S322
D326
N329
E334 N335 V336 R337
L340
E350
Q355
R365 L366

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	40.20Å 70.08Å 73.87Å 113.28° 92.07° 99.36°	Depositor
Resolution (Å)	36.56 – 1.92 36.56 – 1.91	Depositor EDS
% Data completeness (in resolution range)	91.0 (36.56-1.92) 90.4 (36.56-1.91)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 1.91Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.210 , 0.268 0.210 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	21.8	Xtriage
Anisotropy	0.495	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6797	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.84% 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: 5CY

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	C	0.20	0/222	0.64	0/339
2	D	0.21	0/321	0.60	0/495
3	A	0.32	0/2893	0.59	0/3915
3	B	0.33	0/2893	0.61	0/3915
All	All	0.32	0/6329	0.60	0/8664

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 (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	C	200	0	114	5	0
2	D	287	0	162	10	0
3	A	2844	0	2861	116	0
3	B	2844	0	2861	88	0
4	C	18	0	18	0	0
5	A	256	0	0	5	0
5	B	325	0	0	10	0
5	C	10	0	0	1	0
5	D	13	0	0	0	0
All	All	6797	0	6016	214	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 18.

The worst 5 of 214 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)
3:A:184:ILE:HD11	3:A:188:LEU:HD11	1.51	0.93
3:A:35:LEU:HD22	3:A:44:LEU:HD12	1.58	0.85
3:A:150:ASP:H	3:A:156:ASN:HD21	1.27	0.82
3:B:137:ARG:HG2	3:B:137:ARG:HH11	1.43	0.82
3:B:264:LYS:NZ	3:B:329:ASN:HD21	1.80	0.79

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	
3	A	364/366 (100%)	354 (97%)	10 (3%)	0	100	100
3	B	364/366 (100%)	350 (96%)	13 (4%)	1 (0%)	41	31
All	All	728/732 (100%)	704 (97%)	23 (3%)	1 (0%)	51	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	312	GLY

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	
3	A	313/313 (100%)	303 (97%)	10 (3%)	39	29
3	B	313/313 (100%)	307 (98%)	6 (2%)	57	51
All	All	626/626 (100%)	610 (97%)	16 (3%)	46	37

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	184	ILE
3	B	176	ARG
3	A	291	LYS
3	B	137	ARG
3	A	275	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	156	ASN
3	B	329	ASN
3	A	221	ASN
3	A	226	HIS
3	A	275	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 monosaccharides 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	$\#\|Z\| > 2$
4	5CY	C	111	1	18,18,38	4.52	6 (33%)	23,26,55	6.66	9 (39%)

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	5CY	C	111	1	-	1/5/26/56	0/2/2/4

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	111	5CY	C9-C10	16.72	1.53	1.33
4	C	111	5CY	C14-N1	5.52	1.50	1.39
4	C	111	5CY	C8-C9	-3.85	1.35	1.49
4	C	111	5CY	C18-C19	3.11	1.43	1.39
4	C	111	5CY	C15-C14	2.30	1.43	1.39

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	111	5CY	C22-C20-C21	-19.83	81.78	111.16
4	C	111	5CY	C22-C20-C19	-13.73	89.66	110.53
4	C	111	5CY	C22-C20-C10	-12.25	92.00	111.28
4	C	111	5CY	C21-C20-C10	10.18	127.30	111.28
4	C	111	5CY	C21-C20-C19	9.79	125.41	110.53

There are no chirality outliers.

All (1) torsion outliers are listed below:

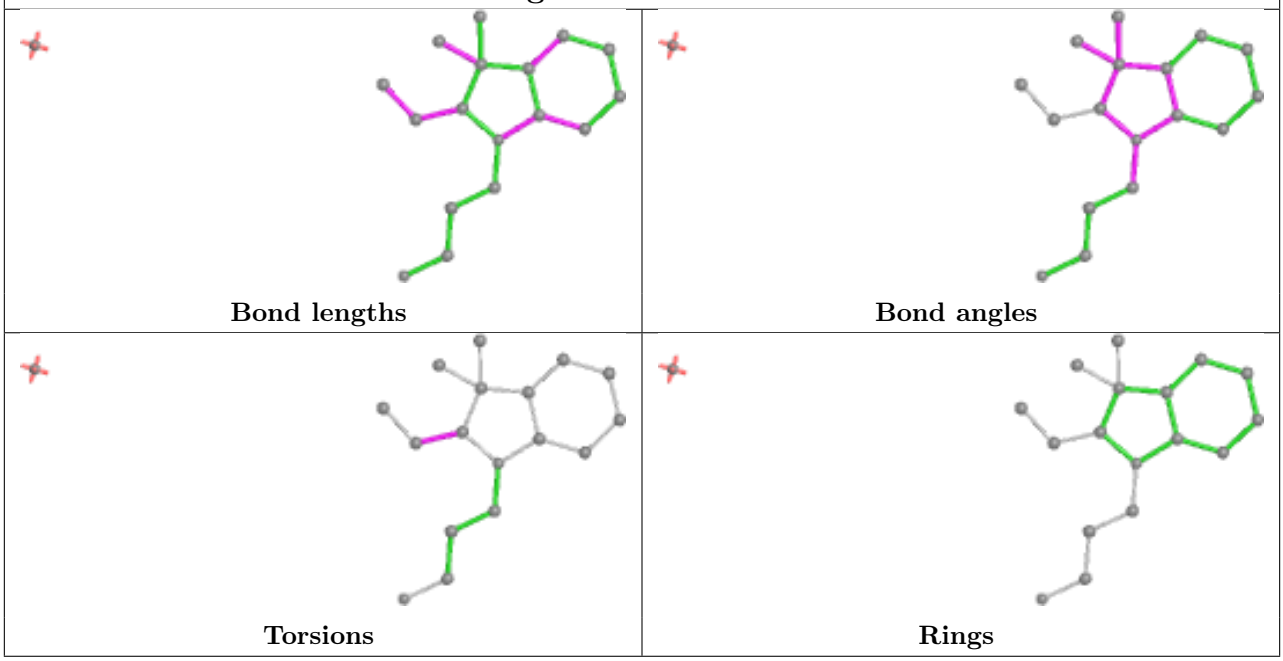
Mol	Chain	Res	Type	Atoms
4	C	111	5CY	C20-C10-C9-C8

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand 5CY C 111



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	C	10/10 (100%)	2.22	5 (50%)	0	0	59, 85, 108, 111	0
2	D	14/14 (100%)	1.64	5 (35%)	0	0	51, 87, 95, 96	0
3	A	366/366 (100%)	0.00	6 (1%)	72	74	16, 28, 44, 67	0
3	B	366/366 (100%)	-0.08	6 (1%)	72	74	12, 25, 41, 62	0
All	All	756/756 (100%)	0.02	22 (2%)	51	55	12, 27, 48, 111	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	366	LEU	7.5
1	C	9	DT	4.8
3	A	22	GLY	4.0
3	B	22	GLY	3.9
1	C	11	DT	3.9

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)

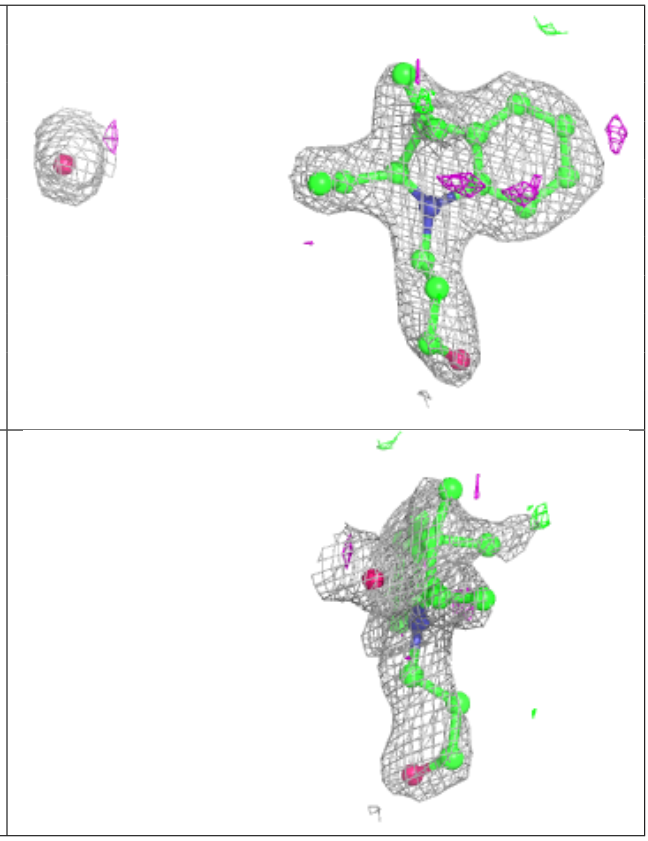
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	5CY	C	111	18/35	0.87	0.15	36,38,45,75	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 5CY C 111:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.