WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 15, 2024 – 10:12 pm GMT

PDB ID	:	6YQM
Title	:	Human histidine triad nucleotide-binding protein 1 (hHINT1) complexed with dGMP and refined to 1.02 A
Authors	:	Dolot, R.D.; Seda, A.; Nawrot, B.C.
Deposited on	:	2020-04-17
Resolution	:	1.02 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

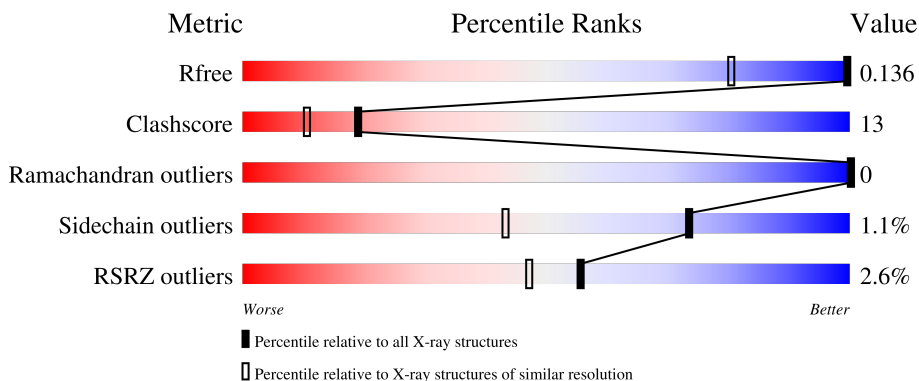|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.02 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		0.136
Clashscore		13
Ramachandran outliers		0
Sidechain outliers		1.1%
RSRZ outliers		2.6%

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	1188 (1.08-0.96)
Clashscore	141614	1253 (1.08-0.96)
Ramachandran outliers	138981	1178 (1.08-0.96)
Sidechain outliers	138945	1180 (1.08-0.96)
RSRZ outliers	127900	1158 (1.08-0.96)

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	AAA	126	2% / 76% 17% 7%
1	BBB	126	3% / 75% 10% • 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PEG	BBB	501	-	X	-	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2639 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 Histidine triad nucleotide-binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	117	Total	C	N	O	S	0	18	0
			1062	670	191	194	7			
1	BBB	112	Total	C	N	O	S	0	19	0
			1041	659	183	192	7			

- Molecule 2 is 2'-DEOXYGUANOSINE-5'-MONOPHOSPHATE (three-letter code: DGP) (formula: $C_{10}H_{14}N_5O_7P$) (labeled as "Ligand of Interest" by depositor).
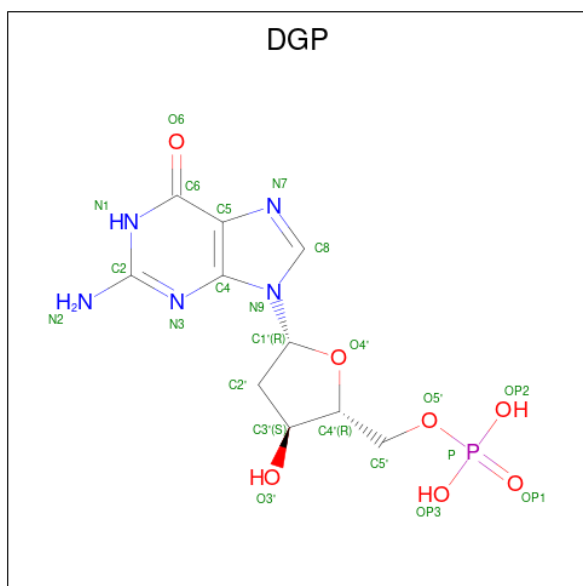
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	AAA	1	Total	C	N	O	P	0	1
			46	20	10	14	2		

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$) (labeled as "Ligand of Interest" by depositor).

PEG

C4

OH O4

O2

O

C3

C1

C2

HO

O1

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	BBB	1	Total 7	C 4	O 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	274	Total 274	O 274	0	0
4	BBB	209	Total 209	O 209	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: Histidine triad nucleotide-binding protein 1

Chain AAA:
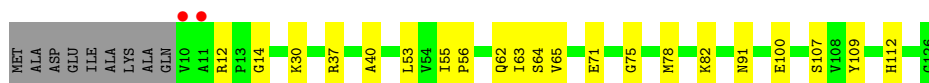
- Molecule 1: Histidine triad nucleotide-binding protein 1

Chain BBB:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	78.66Å 46.43Å 64.10Å 90.00° 94.24° 90.00°	Depositor
Resolution (Å)	39.96 – 1.02 39.96 – 1.02	Depositor EDS
% Data completeness (in resolution range)	99.1 (39.96-1.02) 99.1 (39.96-1.02)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.15 (at 1.02Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.112 , 0.136 0.113 , 0.136	Depositor DCC
R_{free} test set	6128 reflections (5.28%)	wwPDB-VP
Wilson B-factor (Å^2)	7.5	Xtriage
Anisotropy	0.641	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 45.0	EDS
L-test for twinning[2]	$< |L| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	2639	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.04% 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: PEG, DGP

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	AAA	0.81	1/1087 (0.1%)	1.02	4/1462 (0.3%)
1	BBB	1.48	11/1064 (1.0%)	1.15	9/1431 (0.6%)
All	All	1.19	12/2151 (0.6%)	1.09	13/2893 (0.4%)

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
1	BBB	0	1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	71[A]	GLU	CD-OE2	16.67	1.44	1.25
1	BBB	71[B]	GLU	CD-OE2	16.67	1.44	1.25
1	BBB	78[A]	MET	CG-SD	16.27	2.23	1.81
1	BBB	78[B]	MET	CG-SD	16.27	2.23	1.81
1	BBB	78[C]	MET	CG-SD	16.27	2.23	1.81

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	12	ARG	NE-CZ-NH2	-9.14	115.73	120.30
1	BBB	78[A]	MET	CB-CG-SD	-8.27	87.60	112.40
1	BBB	78[B]	MET	CB-CG-SD	-8.27	87.60	112.40
1	BBB	78[C]	MET	CB-CG-SD	-8.27	87.60	112.40
1	AAA	37[A]	ARG	NE-CZ-NH2	7.85	124.22	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	BBB	68[B]	ASP	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	AAA	1062	0	1052	30	0
1	BBB	1041	0	1028	37	0
2	AAA	46	0	24	2	0
3	BBB	7	0	8	1	0
4	AAA	274	0	0	6	2
4	BBB	209	0	0	4	0
All	All	2639	0	2112	56	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 13.

The worst 5 of 56 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:BBB:78[C]:MET:CG	1:BBB:78[C]:MET:SD	2.17	1.30
1:BBB:78[A]:MET:SD	1:BBB:78[A]:MET:CG	2.23	1.26
1:AAA:62:GLN:NE2	1:AAA:100[A]:GLU:OE2	1.78	1.16
1:AAA:63[B]:ILE:HG22	1:AAA:100[B]:GLU:OE1	1.44	1.15
1:BBB:78[A]:MET:SD	1:BBB:78[A]:MET:CB	2.62	0.88

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)
4:AAA:412:HOH:O	4:AAA:464:HOH:O[4_555]	2.01	0.19
4:AAA:437:HOH:O	4:AAA:545:HOH:O[4_555]	2.17	0.03

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	AAA	136/126 (108%)	136 (100%)	0	0	100	100
1	BBB	131/126 (104%)	130 (99%)	1 (1%)	0	100	100
All	All	267/252 (106%)	266 (100%)	1 (0%)	0	100	100

There are no Ramachandran outliers to report.

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	AAA	117/103 (114%)	116 (99%)	1 (1%)	78	51
1	BBB	115/103 (112%)	113 (98%)	2 (2%)	60	26
All	All	232/206 (113%)	229 (99%)	3 (1%)	73	35

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	82	LYS
1	BBB	82[A]	LYS
1	BBB	82[B]	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

3 ligands 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	$\#|Z| > 2$
2	DGP	AAA	201[A]	-	22,25,25	1.96	3 (13%)	26,38,38	1.19	1 (3%)
3	PEG	BBB	501	-	6,6,6	1.91	3 (50%)	5,5,5	3.73	5 (100%)
2	DGP	AAA	201[B]	-	22,25,25	1.59	6 (27%)	26,38,38	1.22	2 (7%)

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
2	DGP	AAA	201[A]	-	-	2/6/22/22	0/3/3/3
3	PEG	BBB	501	-	-	2/4/4/4	-
2	DGP	AAA	201[B]	-	-	2/6/22/22	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	201[A]	DGP	O3'-C3'	7.09	1.58	1.43
2	AAA	201[A]	DGP	C8-N7	-3.83	1.28	1.35
2	AAA	201[B]	DGP	O3'-C3'	3.67	1.51	1.43
3	BBB	501	PEG	C2-C1	-3.21	1.32	1.49
2	AAA	201[B]	DGP	O4'-C4'	-2.71	1.38	1.45

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	BBB	501	PEG	C3-O2-C2	-5.59	89.07	113.29
3	BBB	501	PEG	O4-C4-C3	3.58	132.56	111.81
2	AAA	201[A]	DGP	C8-N7-C5	3.38	109.43	102.99
2	AAA	201[B]	DGP	C8-N7-C5	3.34	109.35	102.99
3	BBB	501	PEG	O1-C1-C2	3.16	130.13	111.81

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	201[B]	DGP	O4'-C4'-C5'-O5'
2	AAA	201[A]	DGP	C4'-C5'-O5'-P
3	BBB	501	PEG	O2-C3-C4-O4
2	AAA	201[B]	DGP	C3'-C4'-C5'-O5'
2	AAA	201[A]	DGP	O4'-C4'-C5'-O5'

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	201[A]	DGP	1	0
3	BBB	501	PEG	1	0
2	AAA	201[B]	DGP	1	0

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 DGP AAA 201 (A)

Bond lengths

Bond angles

Torsions

Rings

Ligand PEG BBB 501

Bond lengths

Bond angles

Torsions

Rings

Ligand DGP AAA 201 (B)

Bond lengths

Bond angles

Torsions

Rings
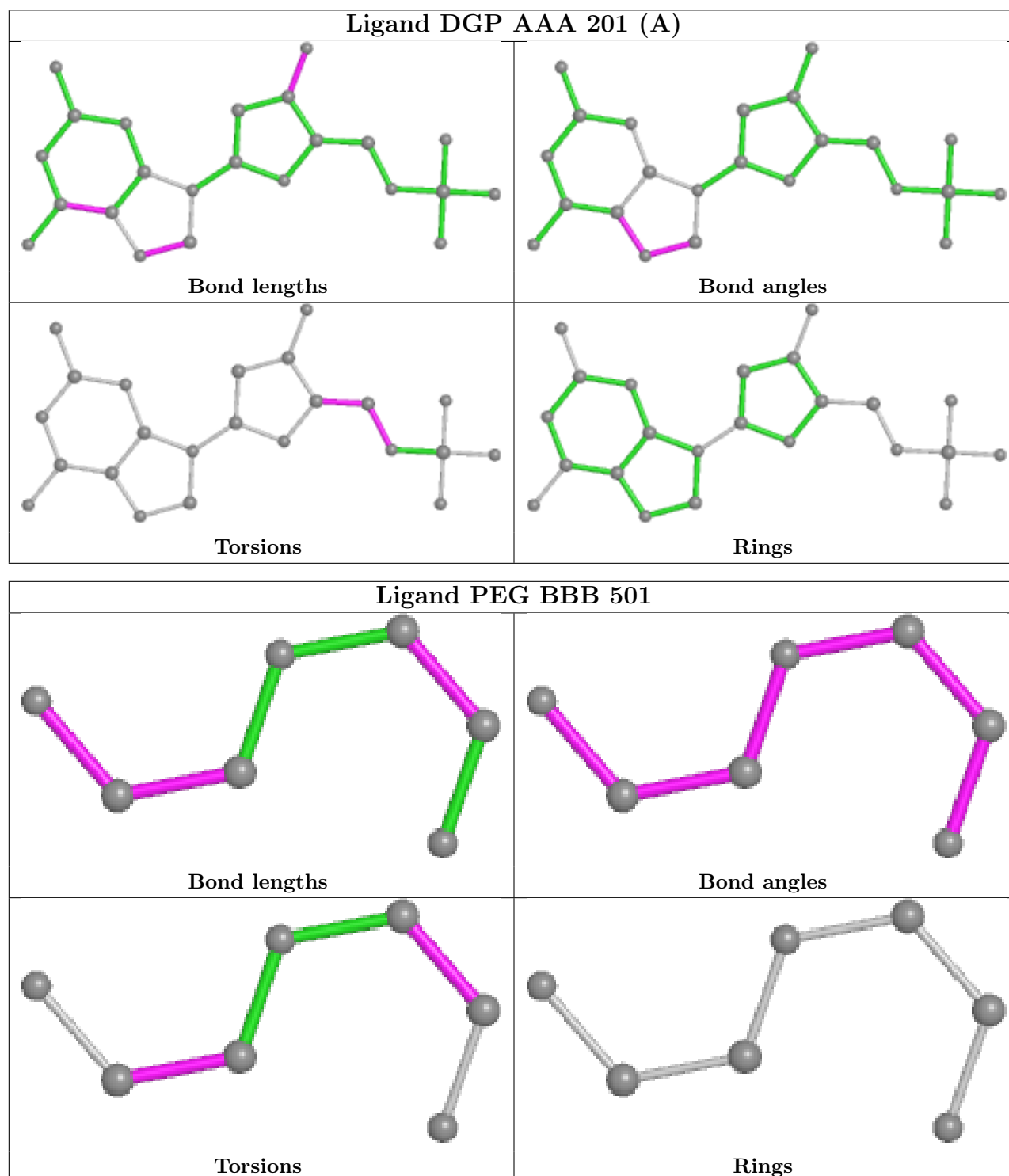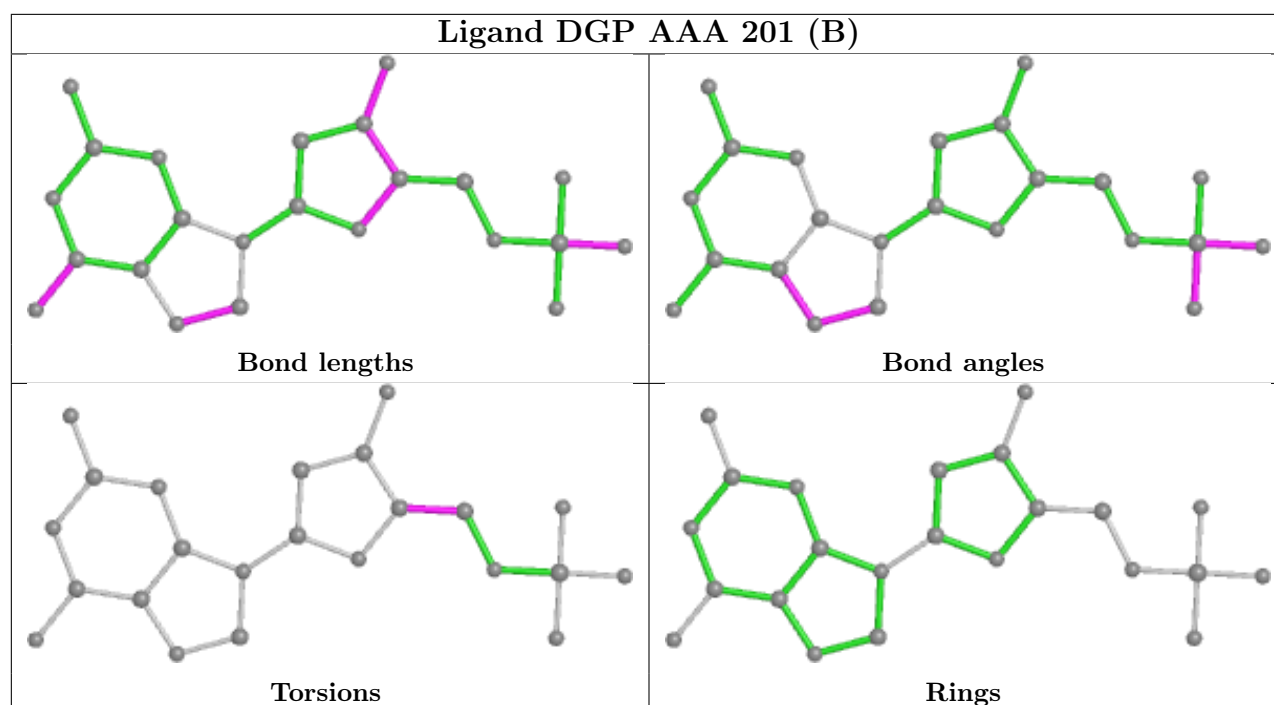
5.7 Other polymers ⓘ

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(Å^2)	Q<0.9
1	AAA	117/126 (92%)	-0.32	2 (1%)	70	61	5, 8, 16, 113	0
1	BBB	112/126 (88%)	-0.22	4 (3%)	42	34	5, 8, 20, 93	0
All	All	229/252 (90%)	-0.27	6 (2%)	56	47	5, 8, 19, 113	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	68[A]	ASP	7.7
1	AAA	10	VAL	4.2
1	BBB	69[A]	ASP	4.1
1	BBB	15	GLY	3.2
1	AAA	11	ALA	3.1

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

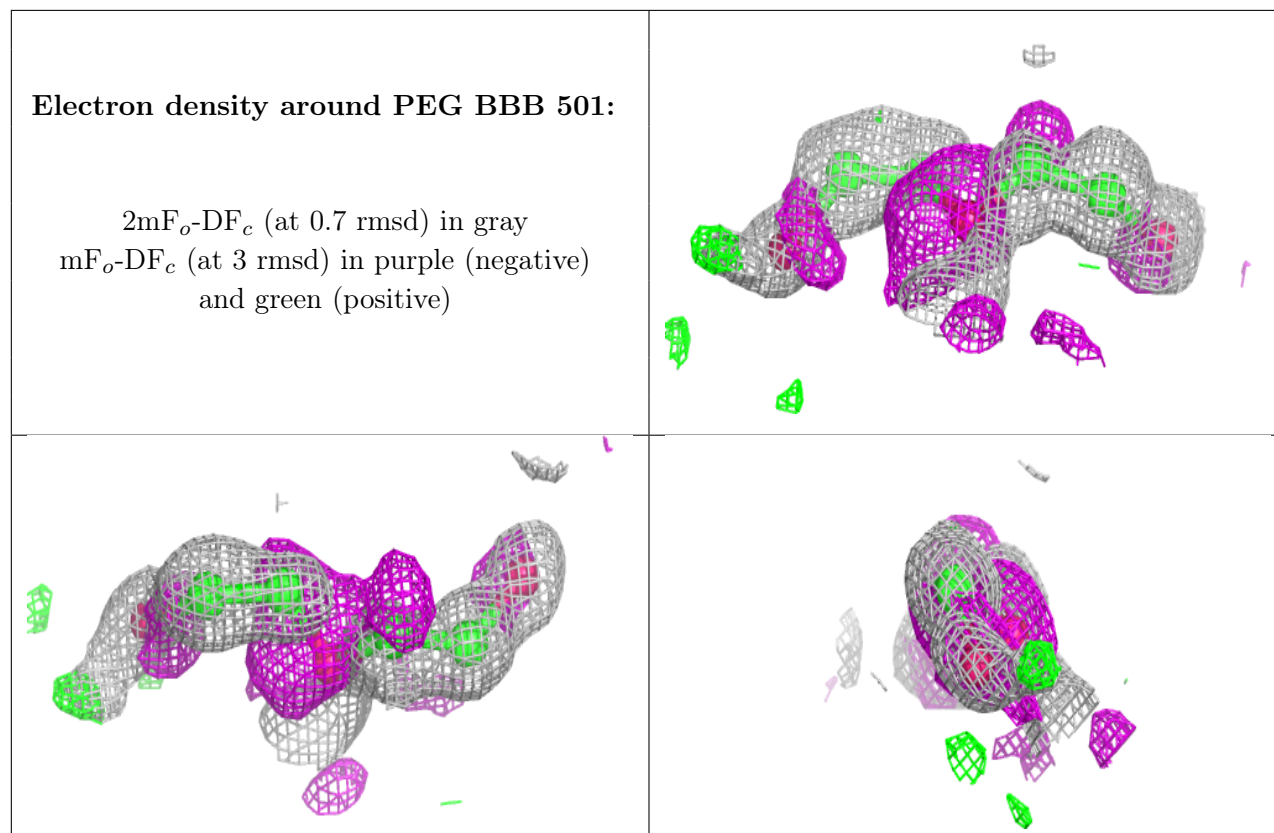There are no non-standard protein/DNA/RNA residues in this entry.

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
3	PEG	BBB	501	7/7	0.87	0.24	10,19,25,38	0
2	DGP	AAA	201[B]	23/23	0.95	0.14	4,9,12,14	23
2	DGP	AAA	201[A]	23/23	0.95	0.14	4,9,11,13	23

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.
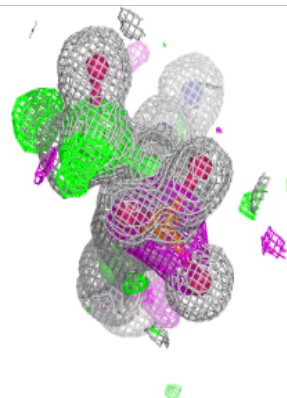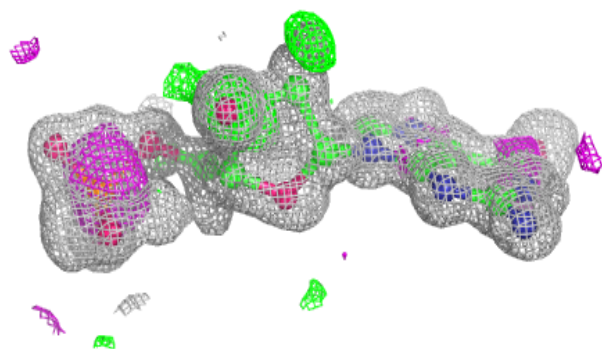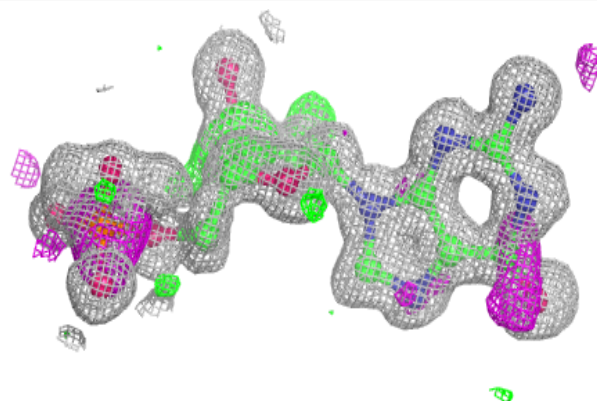
Electron density around PEG BBB 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative) and green (positive)

Electron density around DGP AAA 201 (B):

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
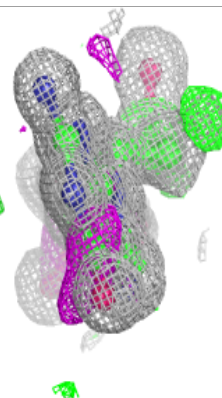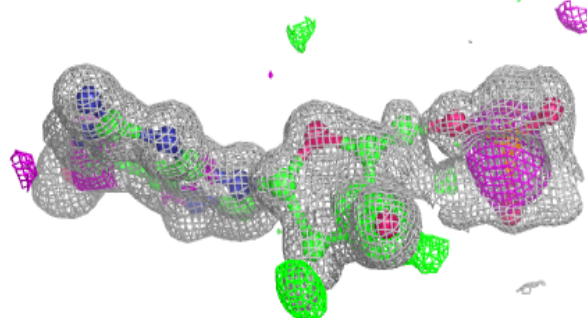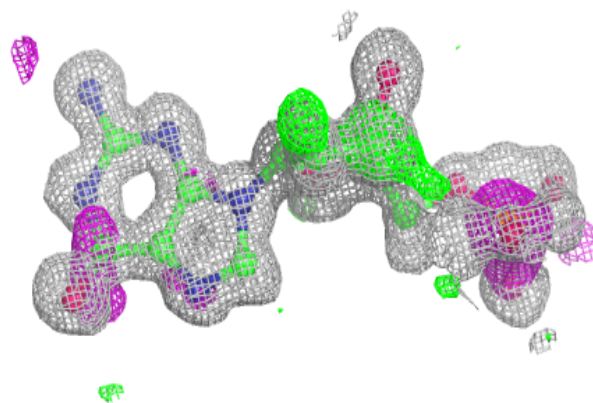
Electron density around DGP AAA 201 (A):

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers (i)

There are no such residues in this entry.