wwPDB X-ray Structure Validation Summary Report (i)

Aug 19, 2023 – 11:56 PM EDT

PDB ID	:	2GQ2
Title	:	Mycobacterium tuberculosis ThyX-NADP complex
Authors	:	Sampathkumar, P.; Turley, S.; Sibley, C.H.; Hol, W.G.
Deposited on	:	2006-04-19
Resolution	:	2.10 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:

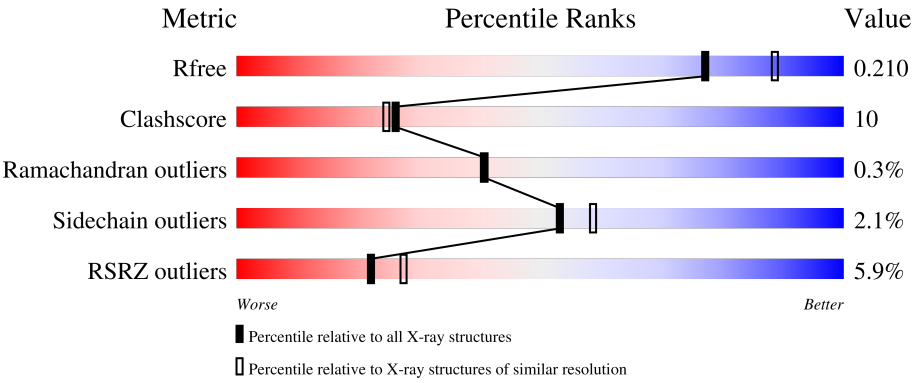
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	258	5% 75% 17% • 6%
1	B	258	4% 76% 12% 12%
1	C	258	7% 74% 13% 13%
1	D	258	5% 72% 15% 12%

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
2	IOD	A	515	-	-	X	-
2	IOD	B	510	-	-	X	-
2	IOD	C	513	-	-	X	-
2	IOD	D	507	-	-	X	-
6	GOL	A	401	-	-	X	-
6	GOL	B	402	-	-	X	-
6	GOL	C	403	-	-	X	-
6	GOL	D	404	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 7649 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 Thymidylate synthase thyX.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	242	Total	C	N	O	S	Se	0	1	0
			1851	1167	329	348	3	4			
1	B	226	Total	C	N	O	S	Se	0	1	0
			1722	1089	305	321	3	4			
1	C	225	Total	C	N	O	S	Se	0	1	0
			1721	1085	307	322	3	4			
1	D	226	Total	C	N	O	S	Se	0	1	0
			1705	1076	301	321	3	4			

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP P66930
A	65	MSE	ILE	engineered mutation	UNP P66930
A	122	MSE	MET	modified residue	UNP P66930
A	175	MSE	LEU	engineered mutation	UNP P66930
A	198	MSE	MET	modified residue	UNP P66930
A	251	LEU	-	cloning artifact	UNP P66930
A	252	GLU	-	cloning artifact	UNP P66930
A	253	HIS	-	expression tag	UNP P66930
A	254	HIS	-	expression tag	UNP P66930
A	255	HIS	-	expression tag	UNP P66930
A	256	HIS	-	expression tag	UNP P66930
A	257	HIS	-	expression tag	UNP P66930
A	258	HIS	-	expression tag	UNP P66930
B	1	MSE	MET	modified residue	UNP P66930
B	65	MSE	ILE	engineered mutation	UNP P66930
B	122	MSE	MET	modified residue	UNP P66930
B	175	MSE	LEU	engineered mutation	UNP P66930
B	198	MSE	MET	modified residue	UNP P66930
B	251	LEU	-	cloning artifact	UNP P66930
B	252	GLU	-	cloning artifact	UNP P66930
B	253	HIS	-	expression tag	UNP P66930

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	254	HIS	-	expression tag	UNP P66930
B	255	HIS	-	expression tag	UNP P66930
B	256	HIS	-	expression tag	UNP P66930
B	257	HIS	-	expression tag	UNP P66930
B	258	HIS	-	expression tag	UNP P66930
C	1	MSE	MET	modified residue	UNP P66930
C	65	MSE	ILE	engineered mutation	UNP P66930
C	122	MSE	MET	modified residue	UNP P66930
C	175	MSE	LEU	engineered mutation	UNP P66930
C	198	MSE	MET	modified residue	UNP P66930
C	251	LEU	-	cloning artifact	UNP P66930
C	252	GLU	-	cloning artifact	UNP P66930
C	253	HIS	-	expression tag	UNP P66930
C	254	HIS	-	expression tag	UNP P66930
C	255	HIS	-	expression tag	UNP P66930
C	256	HIS	-	expression tag	UNP P66930
C	257	HIS	-	expression tag	UNP P66930
C	258	HIS	-	expression tag	UNP P66930
D	1	MSE	MET	modified residue	UNP P66930
D	65	MSE	ILE	engineered mutation	UNP P66930
D	122	MSE	MET	modified residue	UNP P66930
D	175	MSE	LEU	engineered mutation	UNP P66930
D	198	MSE	MET	modified residue	UNP P66930
D	251	LEU	-	cloning artifact	UNP P66930
D	252	GLU	-	cloning artifact	UNP P66930
D	253	HIS	-	expression tag	UNP P66930
D	254	HIS	-	expression tag	UNP P66930
D	255	HIS	-	expression tag	UNP P66930
D	256	HIS	-	expression tag	UNP P66930
D	257	HIS	-	expression tag	UNP P66930
D	258	HIS	-	expression tag	UNP P66930

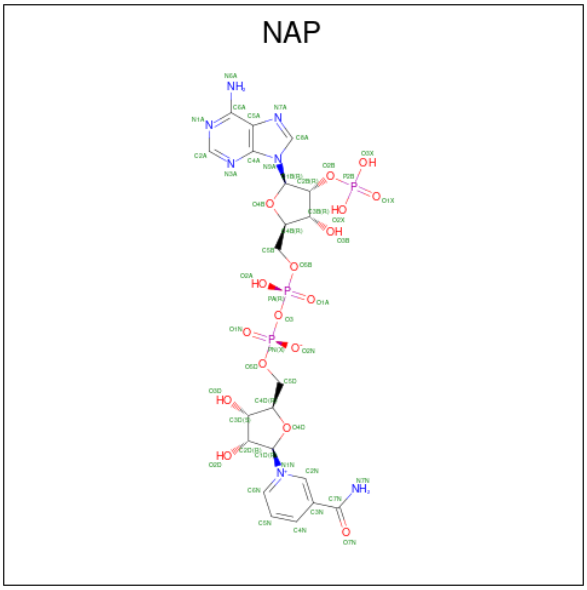
- Molecule 2 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	4	Total I 5 5	0	1
2	B	2	Total I 2 2	0	0
2	C	4	Total I 5 5	0	1
2	D	5	Total I 6 6	0	1

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

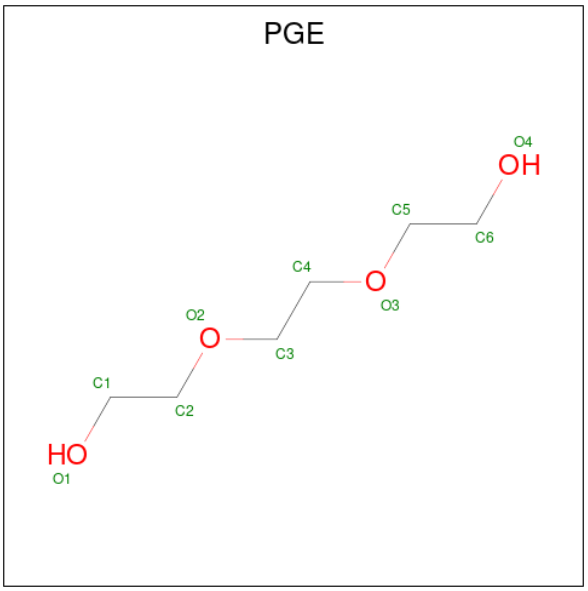
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total K 1 1	0	0
3	B	1	Total K 1 1	0	0
3	C	1	Total K 1 1	0	0
3	D	1	Total K 1 1	0	0

- Molecule 4 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: $C_{21}H_{28}N_7O_{17}P_3$).



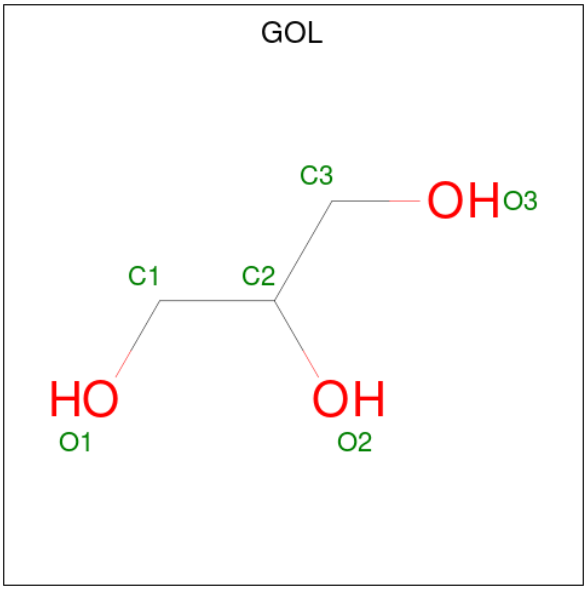
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O P 48 21 7 17 3	0	0
4	B	1	Total C N O P 48 21 7 17 3	0	0
4	C	1	Total C N O P 48 21 7 17 3	0	0
4	D	1	Total C N O P 48 21 7 17 3	0	0

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 10 6 4	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 6 3 3	0	0
6	B	1	Total C O 6 3 3	0	0
6	C	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0

- Molecule 7 is water.

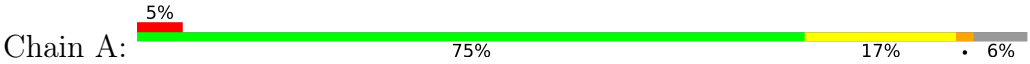
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	112	Total O 113 113	0	1
7	B	91	Total O 91 91	0	0
7	C	83	Total O 83 83	0	0
7	D	97	Total O 97 97	0	0

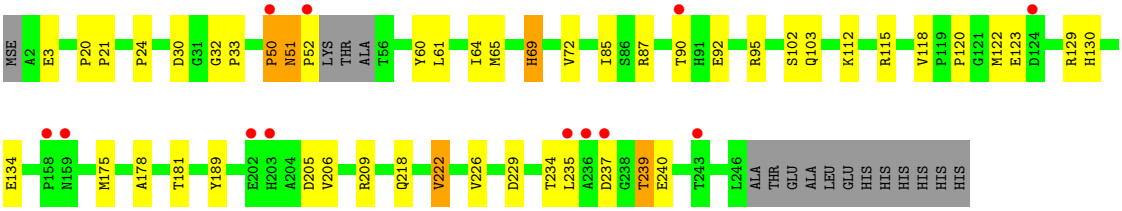
3 Residue-property plots (i)

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: Thymidylate synthase thyX

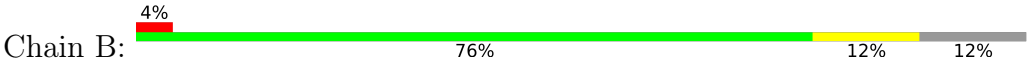
Chain A:

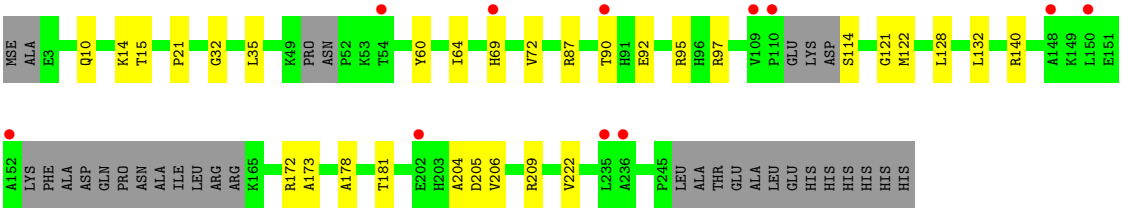




• Molecule 1: Thymidylate synthase thyX

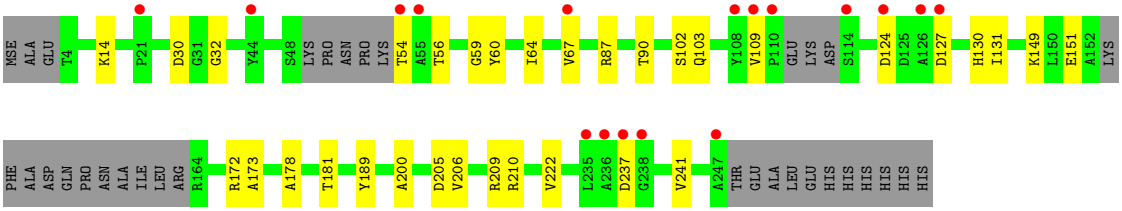
Chain B:





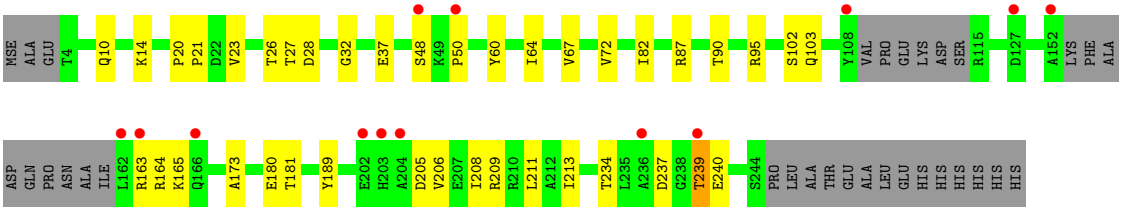
• Molecule 1: Thymidylate synthase thyX

Chain C:



• Molecule 1: Thymidylate synthase thyX

Chain D:


MSE ALA GLU T4 Q10 K14 P20 P21 D22 V23 T26 T27 D28 G32 E37 S48 K49 P50 Y60 I64 V67 V72 I82 R87 T90 R95 S102 Q103 Y108 VAL PRO GLU LYS ASP SER R115 D127 A152 LYS PHE ALA
ASP GLN PRO ASN ALA ILE L162 R163 R164 K165 Q166 A173 E180 T181 Y189 E202 H203 A204 D205 V206 E207 I208 R209 R210 L211 A212 I213 T234 L235 A236 D237 G238 T239 E240 S244 PRO LEU ALA THR GLU ALA LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	73.01Å 78.09Å 88.51Å 90.00° 95.67° 90.00°	Depositor
Resolution (Å)	35.20 – 2.10 35.20 – 2.10	Depositor EDS
% Data completeness (in resolution range)	96.7 (35.20-2.10) 96.7 (35.20-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	5.01 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.212 , 0.255 0.215 , 0.210	Depositor DCC
R_{free} test set	2824 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	23.9	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 66.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7649	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PGE, NAP, IOD, K

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	0.54	0/1893	0.61	0/2578
1	B	0.55	0/1759	0.60	0/2391
1	C	0.57	0/1757	0.64	0/2390
1	D	0.56	0/1742	0.64	0/2374
All	All	0.55	0/7151	0.62	0/9733

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	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	50	PRO	Peptide

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	1851	0	1786	50	0
1	B	1722	0	1673	32	0
1	C	1721	0	1668	26	0
1	D	1705	0	1627	39	0
2	A	5	0	0	5	0
2	B	2	0	0	3	0
2	C	5	0	0	3	0
2	D	6	0	0	6	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	48	0	25	7	0
4	B	48	0	25	0	0
4	C	48	0	25	6	0
4	D	48	0	25	5	0
5	A	10	0	14	1	0
6	A	6	0	8	7	0
6	B	6	0	8	5	0
6	C	12	0	16	9	0
6	D	18	0	22	8	0
7	A	113	0	0	5	0
7	B	91	0	0	5	0
7	C	83	0	0	4	0
7	D	97	0	0	3	0
All	All	7649	0	6922	148	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 10.

The worst 5 of 148 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:B:172:ARG:NH1	2:B:510:IOD:I	2.52	1.13
1:B:205:ASP:OD2	7:B:695:HOH:O	1.79	0.99
7:B:690:HOH:O	2:C:513:IOD:I	2.51	0.98
1:D:206:VAL:CG1	2:D:507:IOD:I	2.89	0.90
1:D:206:VAL:HG12	2:D:507:IOD:I	2.44	0.88

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	239/258 (93%)	231 (97%)	7 (3%)	1 (0%)	34	32
1	B	219/258 (85%)	214 (98%)	5 (2%)	0	100	100
1	C	218/258 (84%)	214 (98%)	3 (1%)	1 (0%)	29	26
1	D	221/258 (86%)	216 (98%)	4 (2%)	1 (0%)	29	26
All	All	897/1032 (87%)	875 (98%)	19 (2%)	3 (0%)	41	41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	51	ASN
1	C	237	ASP
1	D	164	ARG

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	188/204 (92%)	183 (97%)	5 (3%)	44	48
1	B	175/204 (86%)	174 (99%)	1 (1%)	86	90
1	C	175/204 (86%)	171 (98%)	4 (2%)	50	55
1	D	170/204 (83%)	165 (97%)	5 (3%)	42	46
All	All	708/816 (87%)	693 (98%)	15 (2%)	53	59

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	67	VAL
1	D	234	THR
1	C	124	ASP
1	D	239	THR
1	D	67	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	10	GLN
1	D	103	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)

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)

Of 34 ligands modelled in this entry, 22 are monoatomic - leaving 12 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	NAP	D	300	3	45,52,52	1.90	5 (11%)	56,80,80	1.37	6 (10%)
4	NAP	A	300	3	45,52,52	1.86	4 (8%)	56,80,80	1.40	6 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAP	C	300	3	45,52,52	1.85	5 (11%)	56,80,80	1.33	5 (8%)
6	GOL	D	408	-	5,5,5	0.62	0	5,5,5	0.71	0
6	GOL	D	405	-	5,5,5	0.42	0	5,5,5	0.48	0
5	PGE	A	701	-	9,9,9	0.44	0	8,8,8	0.22	0
4	NAP	B	300	3	45,52,52	1.88	6 (13%)	56,80,80	1.38	5 (8%)
6	GOL	C	407	-	5,5,5	0.39	0	5,5,5	0.28	0
6	GOL	B	402	-	5,5,5	0.49	0	5,5,5	0.52	0
6	GOL	D	404	-	5,5,5	0.45	0	5,5,5	0.73	0
6	GOL	C	403	-	5,5,5	0.41	0	5,5,5	0.67	0
6	GOL	A	401	-	5,5,5	0.54	0	5,5,5	1.08	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	NAP	D	300	3	-	15/31/67/67	0/5/5/5
4	NAP	A	300	3	-	11/31/67/67	0/5/5/5
4	NAP	C	300	3	-	14/31/67/67	0/5/5/5
6	GOL	D	408	-	-	4/4/4/4	-
6	GOL	D	405	-	-	0/4/4/4	-
5	PGE	A	701	-	-	5/7/7/7	-
4	NAP	B	300	3	-	10/31/67/67	0/5/5/5
6	GOL	C	407	-	-	0/4/4/4	-
6	GOL	B	402	-	-	2/4/4/4	-
6	GOL	D	404	-	-	0/4/4/4	-
6	GOL	C	403	-	-	3/4/4/4	-
6	GOL	A	401	-	-	2/4/4/4	-

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	300	NAP	O7N-C7N	9.49	1.42	1.24
4	C	300	NAP	O7N-C7N	9.24	1.41	1.24
4	B	300	NAP	O7N-C7N	9.18	1.41	1.24
4	A	300	NAP	O7N-C7N	8.96	1.41	1.24
4	C	300	NAP	C2A-N3A	4.98	1.40	1.32

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	300	NAP	N3A-C2A-N1A	-5.93	119.41	128.68
4	B	300	NAP	N3A-C2A-N1A	-5.57	119.97	128.68
4	D	300	NAP	N3A-C2A-N1A	-5.55	120.00	128.68
4	C	300	NAP	N3A-C2A-N1A	-5.54	120.02	128.68
4	B	300	NAP	C6N-N1N-C2N	-4.27	118.08	121.97

There are no chirality outliers.

5 of 66 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	300	NAP	C5D-O5D-PN-O1N
4	A	300	NAP	O4D-C4D-C5D-O5D
4	A	300	NAP	C2D-C1D-N1N-C6N
4	B	300	NAP	C5D-O5D-PN-O1N
4	B	300	NAP	C5D-O5D-PN-O2N

There are no ring outliers.

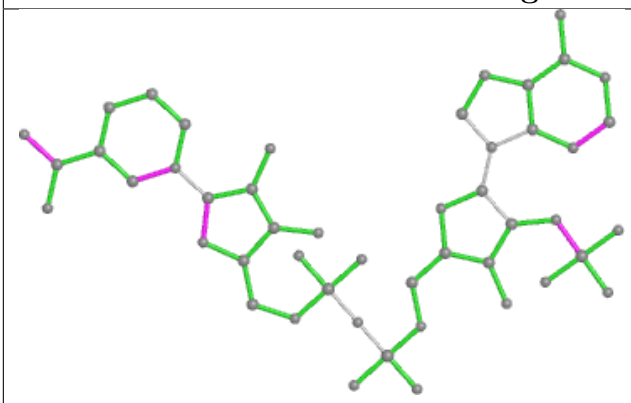
11 monomers are involved in 47 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	300	NAP	5	0
4	A	300	NAP	7	0
4	C	300	NAP	6	0
6	D	408	GOL	1	0
6	D	405	GOL	1	0
5	A	701	PGE	1	0
6	C	407	GOL	1	0
6	B	402	GOL	5	0
6	D	404	GOL	6	0
6	C	403	GOL	8	0
6	A	401	GOL	7	0

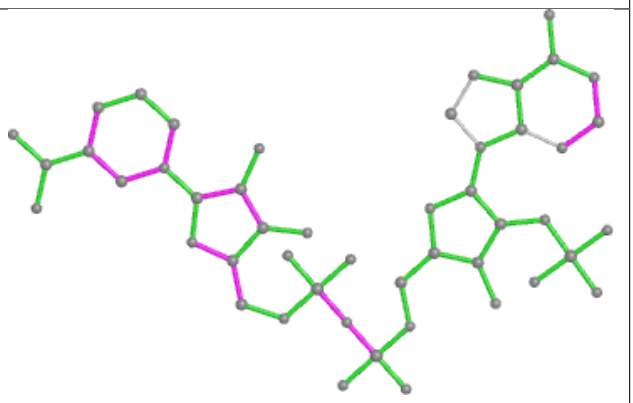
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

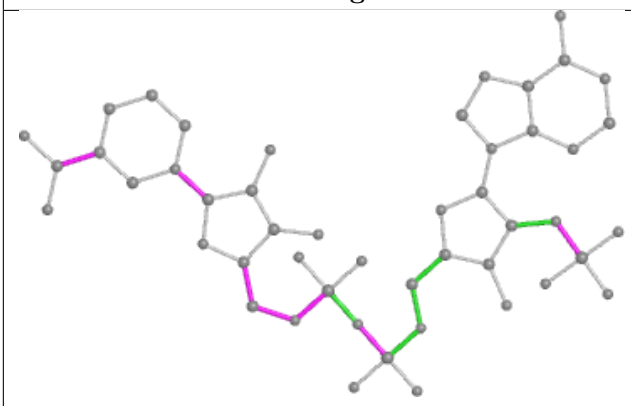
Ligand NAP D 300



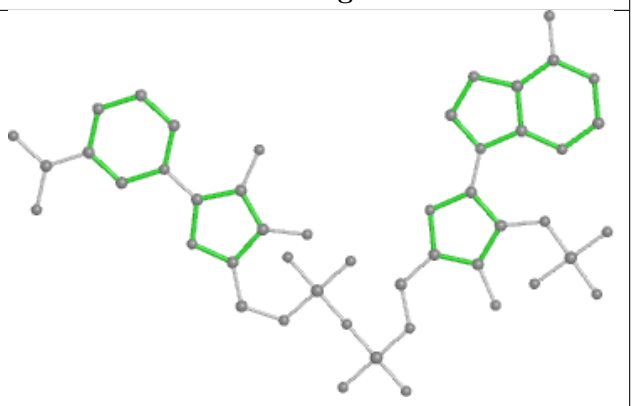
Bond lengths



Bond angles

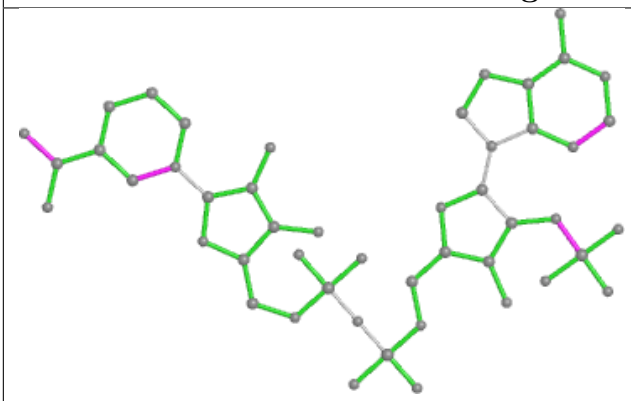


Torsions

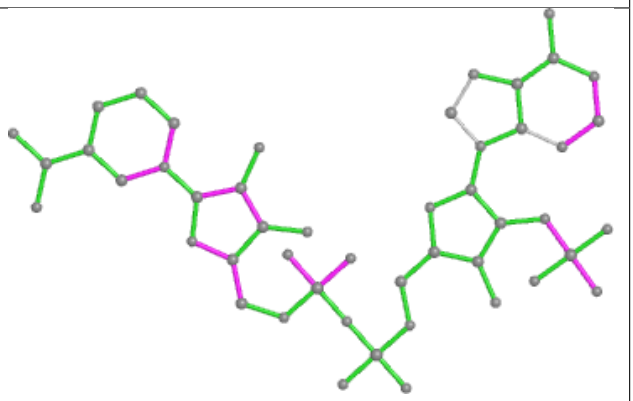


Rings

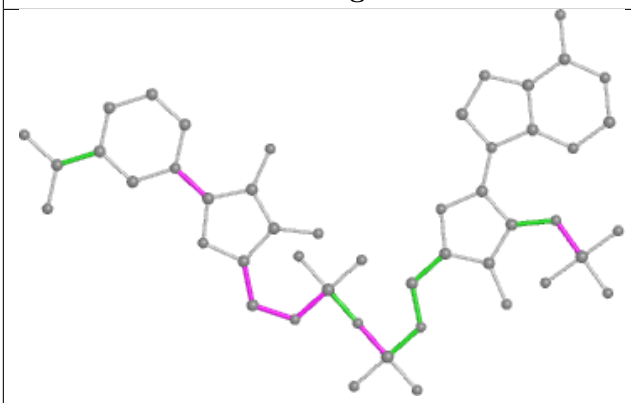
Ligand NAP A 300



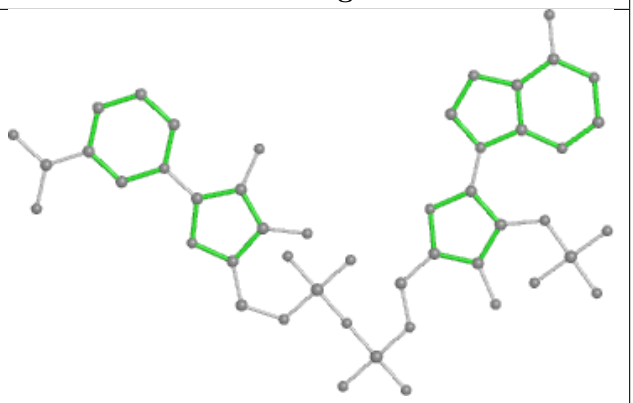
Bond lengths



Bond angles

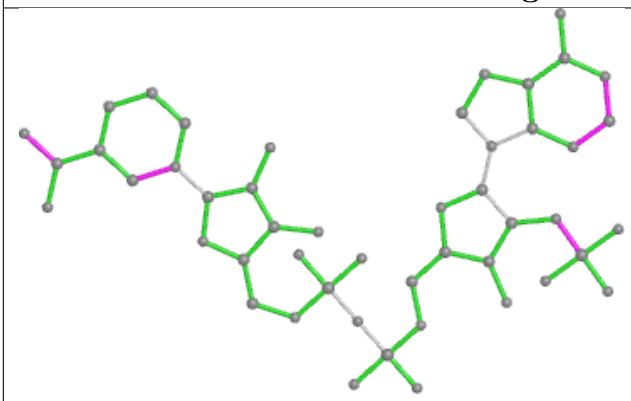


Torsions

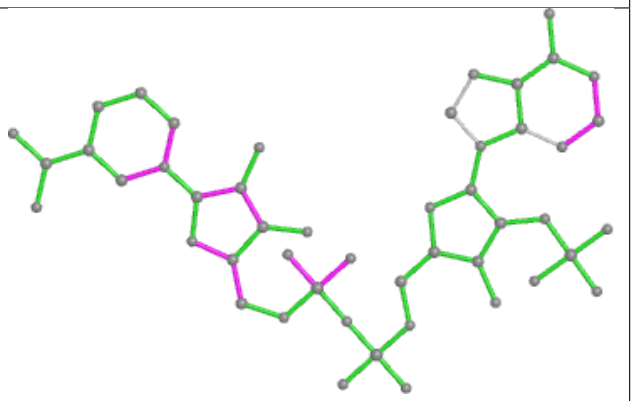


Rings

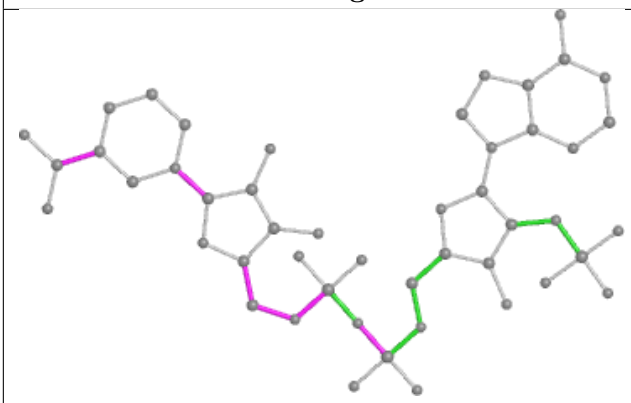
Ligand NAP C 300



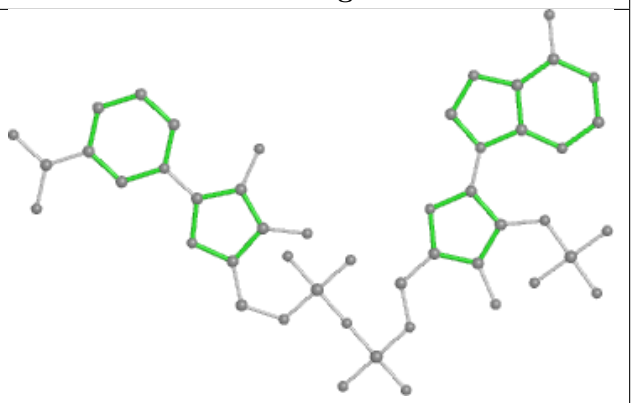
Bond lengths



Bond angles

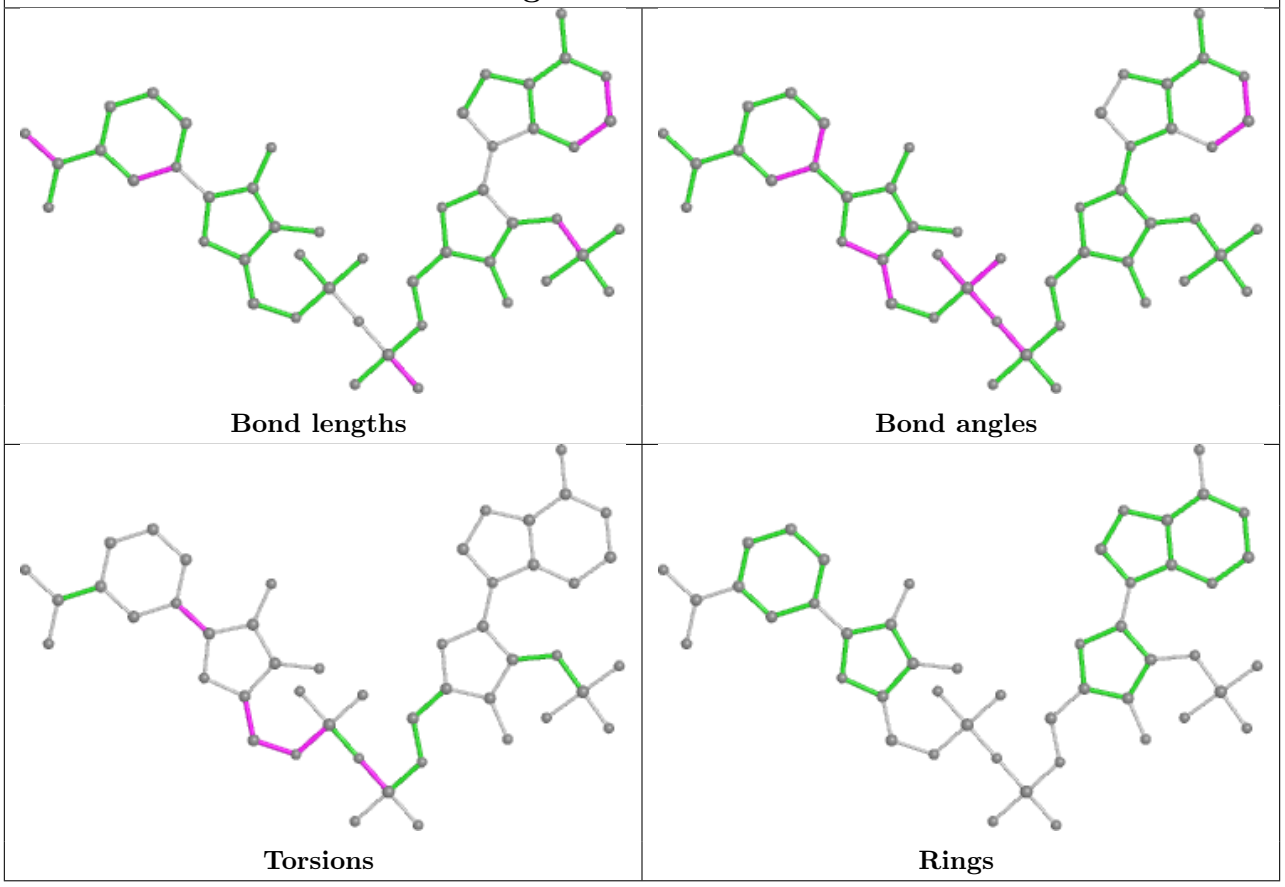


Torsions



Rings

Ligand NAP B 300



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

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, 95^{th} 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	238/258 (92%)	0.50	12 (5%)	28	34	11, 25, 41, 47	1 (0%)
1	B	222/258 (86%)	0.45	11 (4%)	28	34	11, 25, 41, 48	1 (0%)
1	C	221/258 (85%)	0.51	17 (7%)	13	17	11, 25, 42, 49	0
1	D	222/258 (86%)	0.44	13 (5%)	22	27	11, 25, 41, 47	1 (0%)
All	All	903/1032 (87%)	0.48	53 (5%)	22	27	11, 25, 42, 49	3 (0%)

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	203	HIS	6.4
1	C	236	ALA	4.7
1	C	235	LEU	4.5
1	D	204	ALA	4.4
1	C	110	PRO	4.4

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

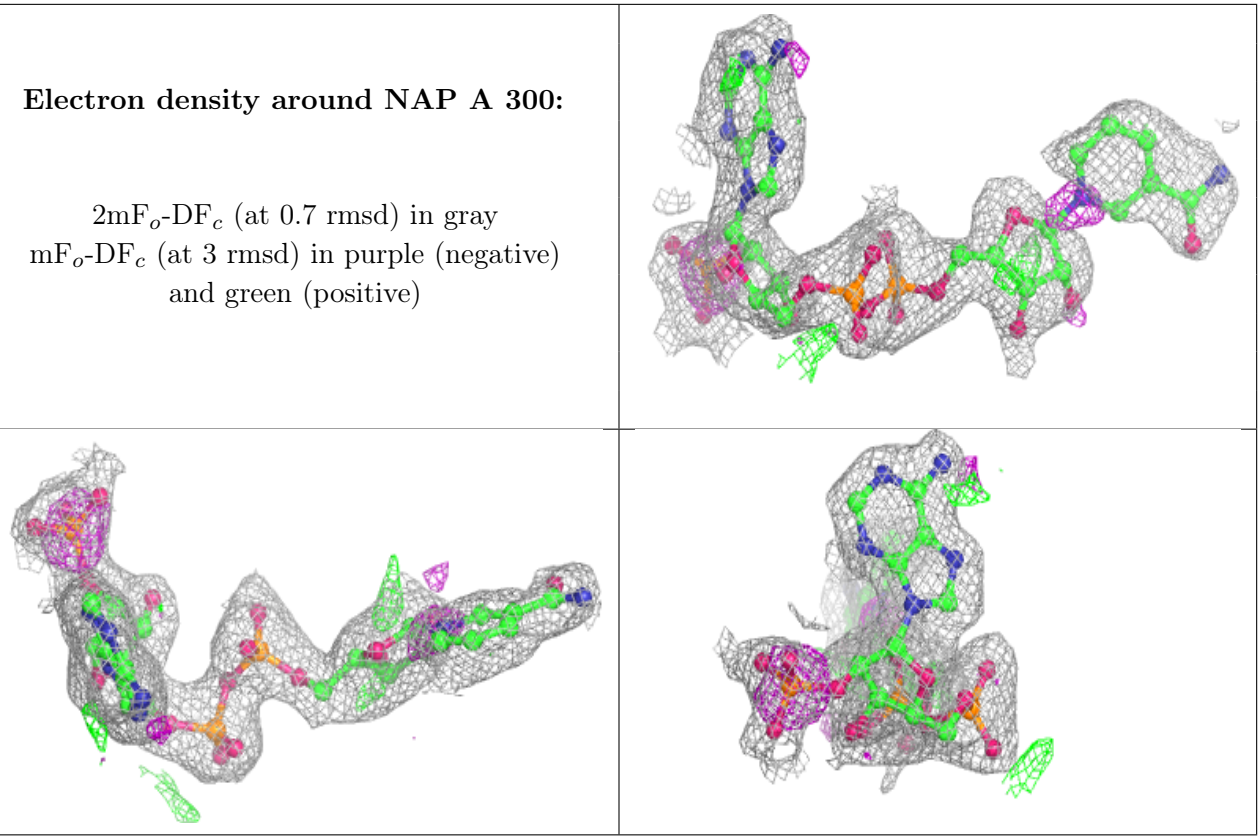
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
6	GOL	B	402	6/6	0.79	0.25	32,35,38,40	0
6	GOL	A	401	6/6	0.81	0.23	22,26,28,29	0
5	PGE	A	701	10/10	0.83	0.26	47,51,54,56	0
6	GOL	D	405	6/6	0.83	0.21	36,38,40,42	0
6	GOL	D	404	6/6	0.86	0.19	25,27,28,29	0
6	GOL	C	407	6/6	0.88	0.17	33,37,38,39	0
2	IOD	D	514	1/1	0.89	0.08	55,55,55,55	1
4	NAP	A	300	48/48	0.89	0.19	22,29,44,45	0
4	NAP	B	300	48/48	0.89	0.19	19,29,51,52	0
6	GOL	C	403	6/6	0.90	0.16	17,23,26,34	0
2	IOD	D	512[B]	1/1	0.92	0.09	52,52,52,52	1
4	NAP	C	300	48/48	0.92	0.17	17,26,47,48	0
4	NAP	D	300	48/48	0.92	0.16	19,27,52,52	0
2	IOD	C	506	1/1	0.92	0.12	39,39,39,39	1
2	IOD	D	512[A]	1/1	0.92	0.09	39,39,39,39	1
6	GOL	D	408	6/6	0.94	0.14	8,9,10,10	6
2	IOD	C	511[A]	1/1	0.95	0.09	40,40,40,40	1
2	IOD	C	511[B]	1/1	0.95	0.09	47,47,47,47	1
2	IOD	D	507	1/1	0.95	0.07	37,37,37,37	1
2	IOD	B	504	1/1	0.96	0.20	22,22,22,22	1
2	IOD	A	516	1/1	0.97	0.20	30,30,30,30	1
2	IOD	C	513	1/1	0.97	0.15	36,36,36,36	1
2	IOD	D	503	1/1	0.98	0.10	29,29,29,29	1
2	IOD	D	505	1/1	0.98	0.07	38,38,38,38	1
3	K	A	603	1/1	0.98	0.14	25,25,25,25	0
2	IOD	A	502[A]	1/1	0.98	0.06	28,28,28,28	1
2	IOD	A	502[B]	1/1	0.98	0.06	26,26,26,26	1
2	IOD	B	510	1/1	0.99	0.07	46,46,46,46	1
2	IOD	A	501	1/1	0.99	0.09	22,22,22,22	1
2	IOD	A	515	1/1	0.99	0.04	49,49,49,49	1
3	K	B	604	1/1	0.99	0.13	18,18,18,18	0
3	K	C	601	1/1	0.99	0.12	13,13,13,13	0
3	K	D	602	1/1	0.99	0.08	18,18,18,18	0
2	IOD	C	509	1/1	1.00	0.08	24,24,24,24	1

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.

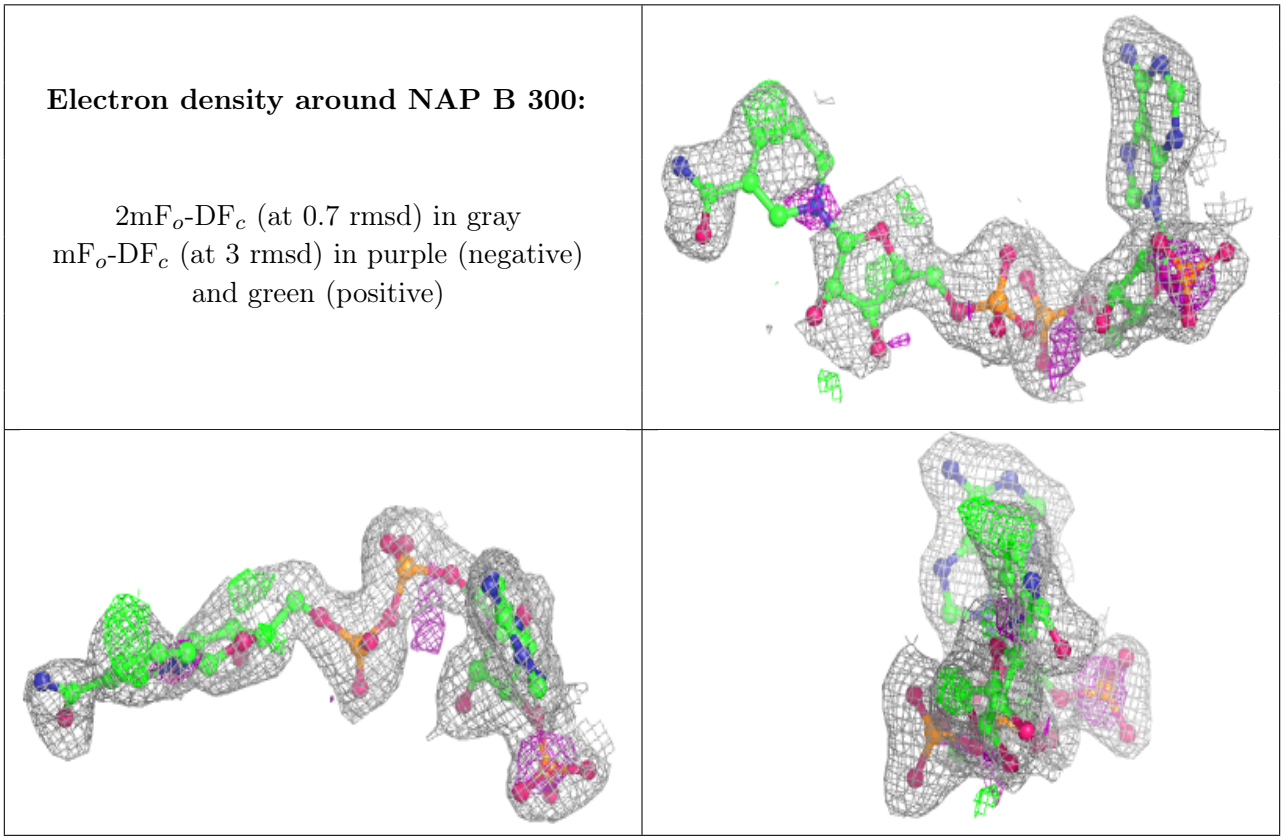
Electron density around NAP A 300:

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



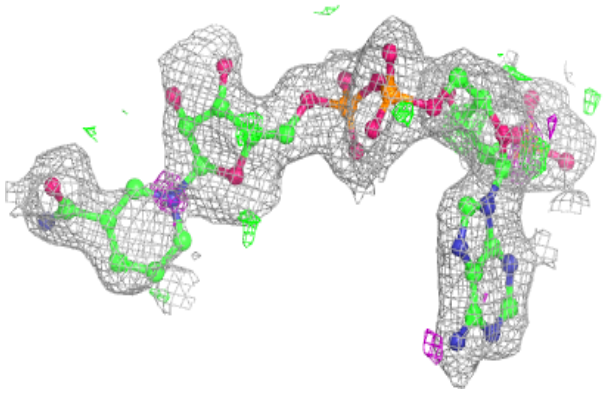
Electron density around NAP B 300:

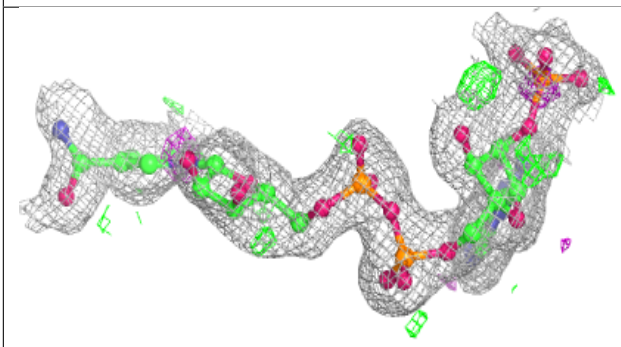
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

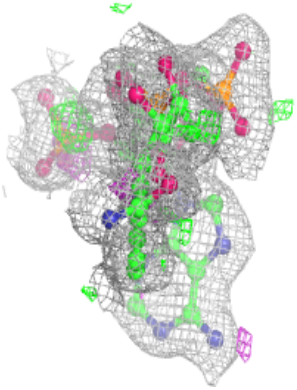


Electron density around NAP C 300:

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

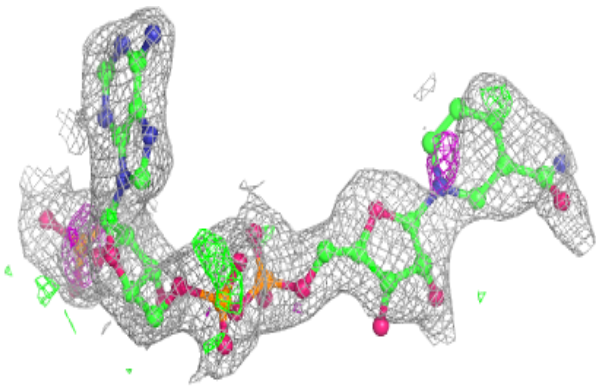


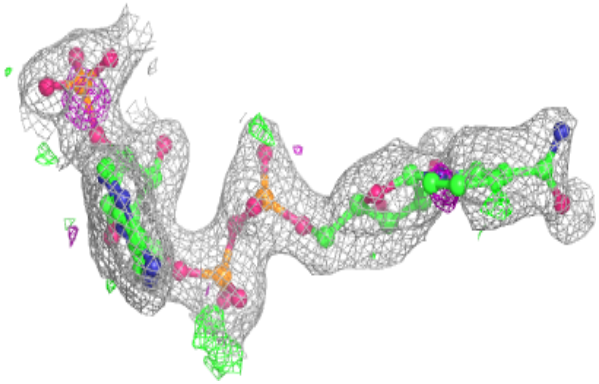


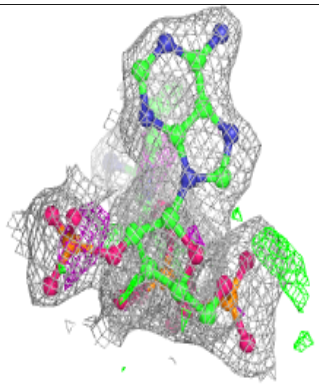


Electron density around NAP D 300:

$2mF_o$-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.