wwPDB X-ray Structure Validation Summary Report (i)

Nov 20, 2023 – 03:36 PM JST

PDB ID	:	7CTY
Title	:	Wild type plasmodium falciparum dihydrofolate reductase-thymidylate synthase (PfDHFR-TS), fragment 263, NADP+, dUMP
Authors	:	Vanichtanankul, J.; Vitsupakorn, D.
Deposited on	:	2020-08-20
Resolution	:	2.80 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:

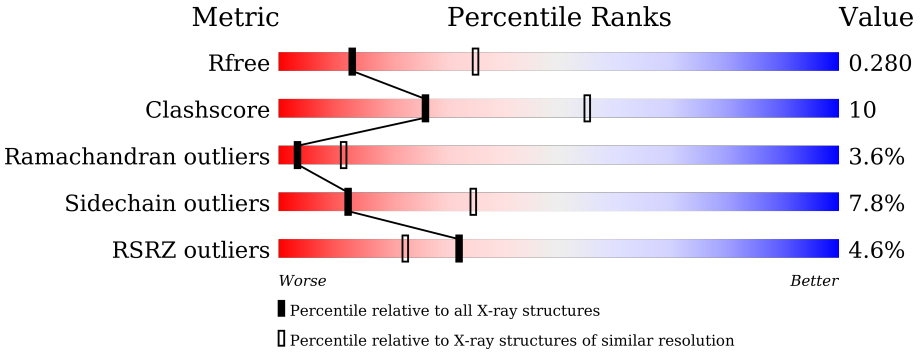
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	608	2% 63% 20% • 13%
1	B	608	6% 61% 19% • 16%

2 Entry composition (i)

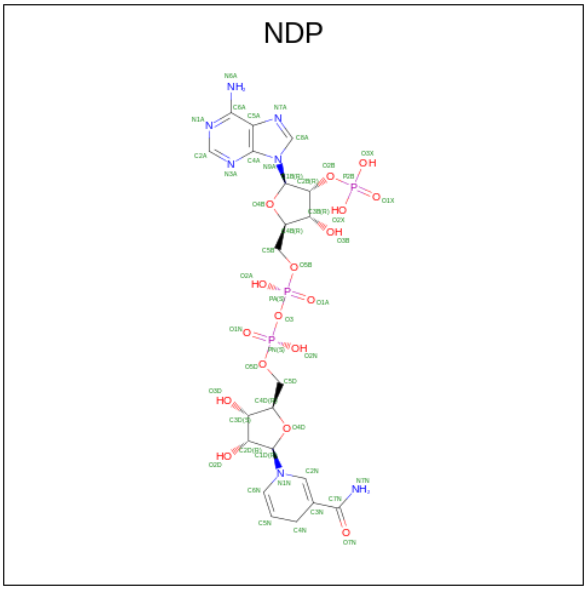
There are 5 unique types of molecules in this entry. The entry contains 17387 atoms, of which 8627 are hydrogens and 0 are deuteriums.

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 Bifunctional dihydrofolate reductase-thymidylate synthase.

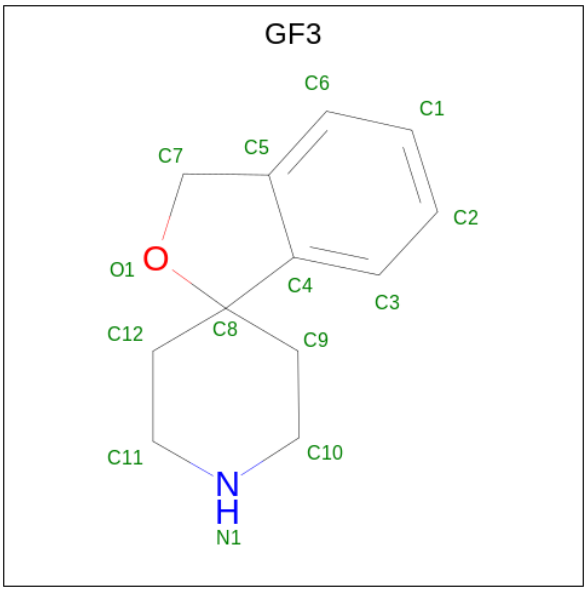
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	529	8756	2845	4355	725	806	25	256	0	0
1	B	512	8473	2757	4216	699	777	24	245	0	0

- Molecule 2 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: $C_{21}H_{30}N_7O_{17}P_3$).



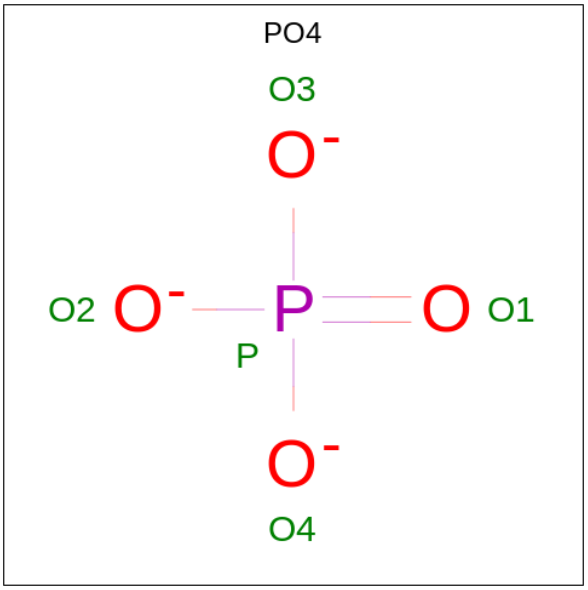
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	H	N	O	P		
2	A	1	74	21	26	7	17	3	7	0

- Molecule 3 is spiro[1H-2-benzofuran-3,4'-piperidine] (three-letter code: GF3) (formula: $C_{12}H_{15}NO$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
3	A	1	29	12	15	1	1	0	0
			Total	C	H	N	O		
3	B	1	29	12	15	1	1	0	0

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
4	A	1	5	4	1	0	0
			Total	O	P		
4	B	1	5	4	1	0	0

- Molecule 5 is water.

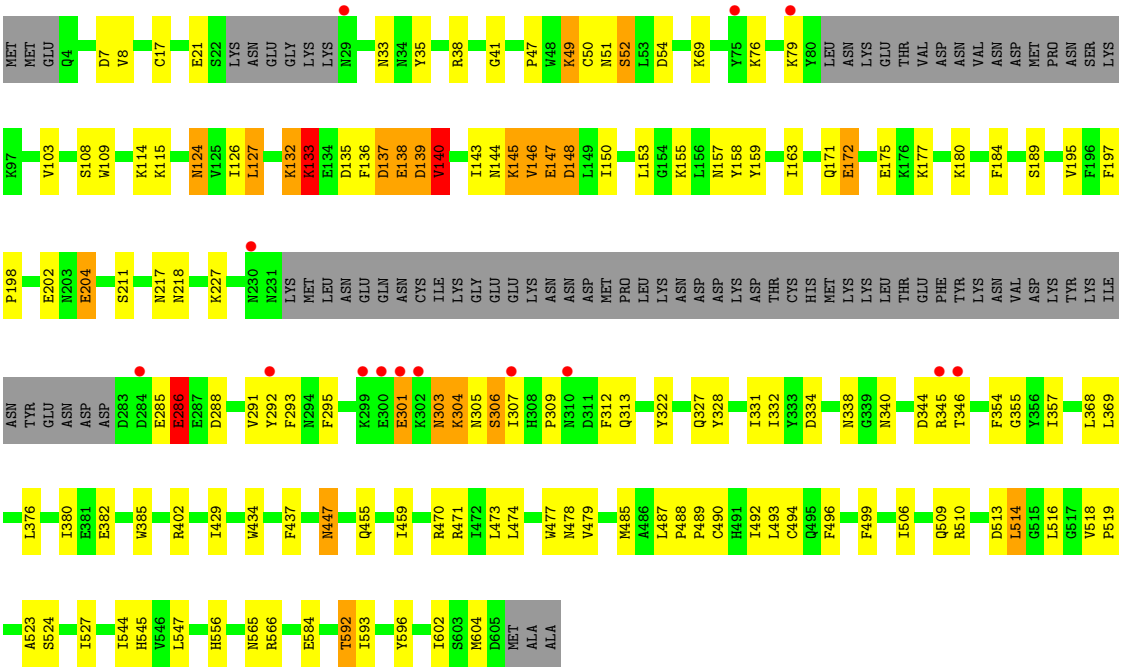
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	11	11	11	0	0
			Total	O		
5	B	5	5	5	0	0

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.

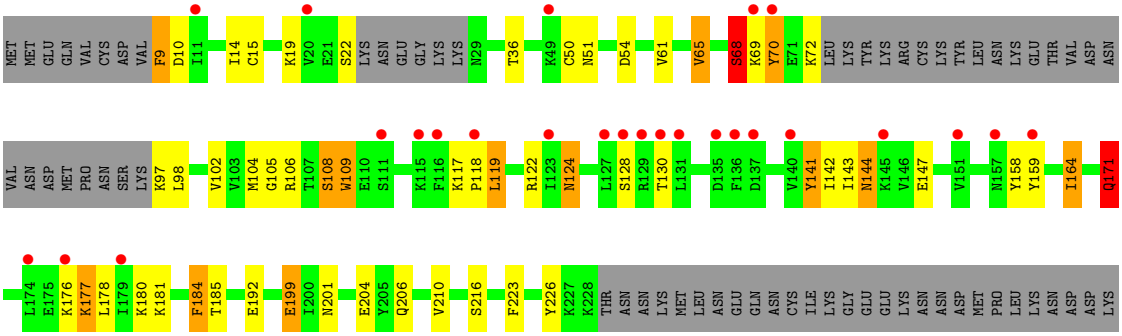
• Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase

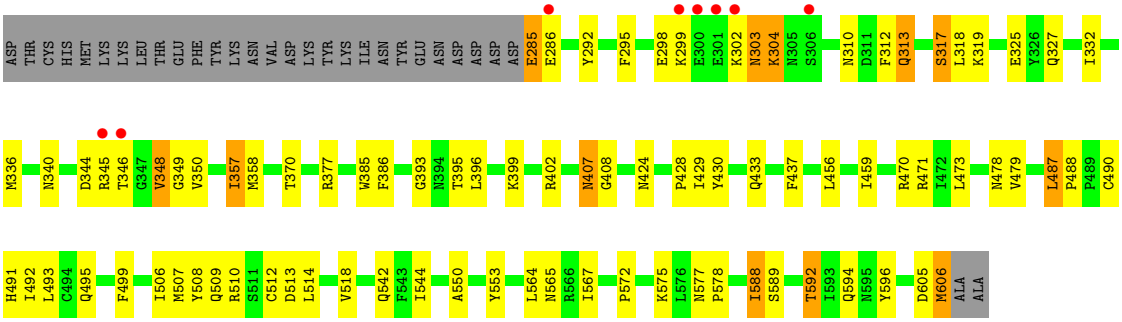
Chain A:



• Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase

Chain B:




ASP THR CYS HIS MET LYS LYS LEU THR GLU PHE TYR LYS ASN VAL ASP LYS TYR LYS ILE ASN TYR GLU ASN ASP ASP ASP ASP E285 E286 Y292 F295 E298 K299 E300 E301 K302 N303 K304 N305 S306 N310 D311 F312 Q313 S317 L318 K319 E325 Y326 Q327 I332
M336 N340 D344 R345 T346 G347 V348 G349 V350 I357 M358 T370 R377 W385 F386 G393 N394 T395 L396 K399 R402 N407 G408 N424 P428 I429 Y430 Q433 F437 L456 I459 R470 R471 I472 L473 N478 V479 L487 P488 P489 C490
H491 I492 L493 C494 Q495 F499 I506 M507 Y508 Q509 R510 S511 C512 D513 L514 V518 Q542 F543 I544 A550 Y553 L564 N565 R566 I567 P572 K575 L576 N577 P578 I588 S589 T592 I593 Q594 N595 Y596 D605 M606 ALA ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.14Å 154.09Å 163.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.70 – 2.80 24.08 – 2.80	Depositor EDS
% Data completeness (in resolution range)	81.2 (20.70-2.80) 92.8 (24.08-2.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.87 (at 2.80Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.234 , 0.303 0.207 , 0.280	Depositor DCC
R_{free} test set	1686 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	22.6	Xtriage
Anisotropy	0.305	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 6.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Reported twinning fraction	0.904 for H, K, L 0.096 for -H, -L, -K	Depositor
Outliers	0 of 34592 reflections	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	17387	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% 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: NDP, PO4, GF3

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	0.72	0/4505	0.92	0/6086
1	B	0.71	0/4359	0.88	0/5889
All	All	0.71	0/8864	0.90	0/11975

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	2
1	B	0	3
All	All	0	5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	140	VAL	Peptide
1	A	50	CYS	Peptide
1	B	304	LYS	Peptide
1	B	605	ASP	Peptide
1	B	68	SER	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	4401	4355	4340	107	1
1	B	4257	4216	4201	91	1
2	A	48	26	26	3	0
3	A	14	15	0	0	0
3	B	14	15	0	0	0
4	A	5	0	0	1	0
4	B	5	0	0	0	0
5	A	11	0	0	1	0
5	B	5	0	0	0	0
All	All	8760	8627	8567	180	1

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 180 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:139:ASP:O	1:A:140:VAL:HG13	1.57	1.02
1:A:171:GLN:NE2	1:A:175:GLU:OE1	2.03	0.92
1:B:606:MET:SD	1:B:606:MET:N	2.45	0.89
1:A:126:ILE:HG12	1:A:140:VAL:HG11	1.57	0.86
1:A:357:ILE:HD11	1:B:357:ILE:HB	1.66	0.78

All (1) 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)
1:A:138:GLU:H	1:B:377:ARG:HE[2_555]	1.27	0.33

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	521/608 (86%)	452 (87%)	50 (10%)	19 (4%)	3 11
1	B	504/608 (83%)	432 (86%)	54 (11%)	18 (4%)	3 11
All	All	1025/1216 (84%)	884 (86%)	104 (10%)	37 (4%)	3 11

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	49	LYS
1	A	133	LYS
1	A	139	ASP
1	A	147	GLU
1	A	148	ASP

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	495/570 (87%)	454 (92%)	41 (8%)	11 32
1	B	478/570 (84%)	443 (93%)	35 (7%)	14 38
All	All	973/1140 (85%)	897 (92%)	76 (8%)	12 35

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	199	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	490	CYS
1	B	210	VAL
1	B	336	MET
1	B	606	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	313	GLN
1	B	415	ASN
1	B	521	ASN
1	B	424	ASN
1	B	407	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)

5 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	$\#\|Z\| > 2$
3	GF3	A	702	-	16,16,16	0.51	0	18,23,23	1.06	1 (5%)
3	GF3	B	902	-	16,16,16	0.32	0	18,23,23	0.76	1 (5%)
4	PO4	A	703	-	4,4,4	0.69	0	6,6,6	0.65	0
4	PO4	B	901	-	4,4,4	0.95	0	6,6,6	0.40	0
2	NDP	A	701	-	45,52,52	0.69	0	53,80,80	0.77	1 (1%)

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	GF3	B	902	-	-	-	0/3/3/3
3	GF3	A	702	-	-	-	0/3/3/3
2	NDP	A	701	-	-	3/30/77/77	0/5/5/5

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	701	NDP	C5A-C6A-N6A	2.64	124.36	120.35
3	A	702	GF3	C7-C5-C4	-2.52	107.56	108.80
3	B	902	GF3	C11-C12-C8	2.14	113.52	111.82

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	NDP	C2B-O2B-P2B-O1X
2	A	701	NDP	C2B-O2B-P2B-O3X
2	A	701	NDP	O4D-C1D-N1N-C2N

There are no ring outliers.

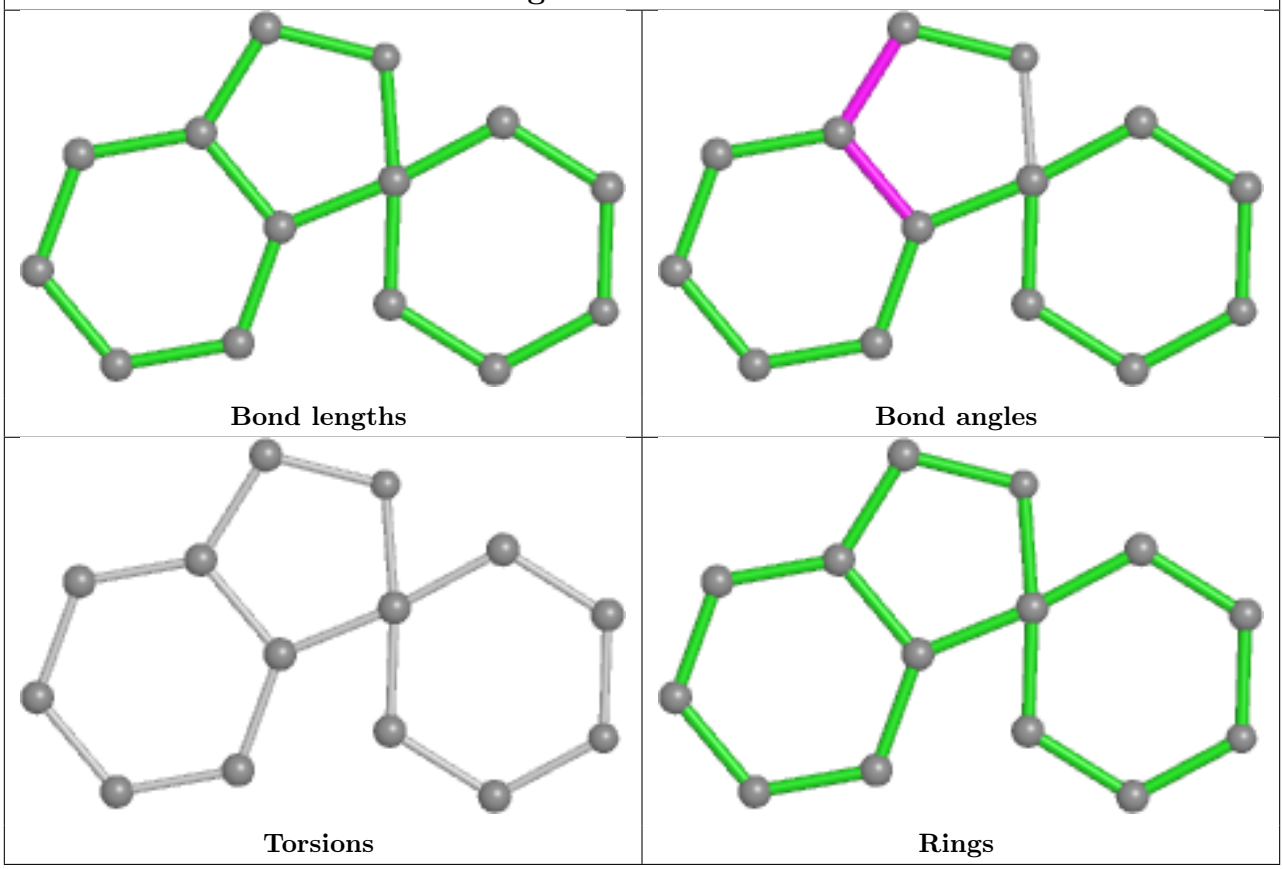
2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	703	PO4	1	0
2	A	701	NDP	3	0

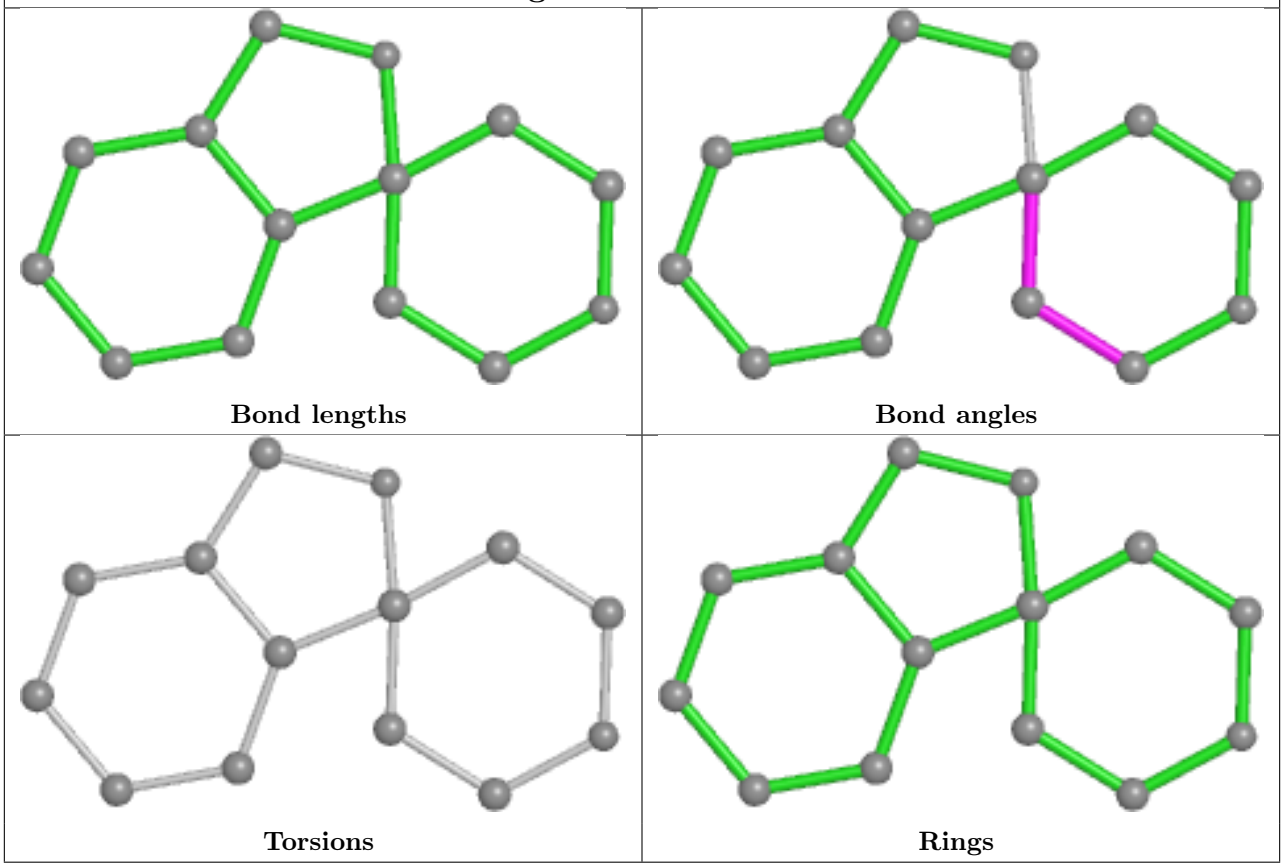
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

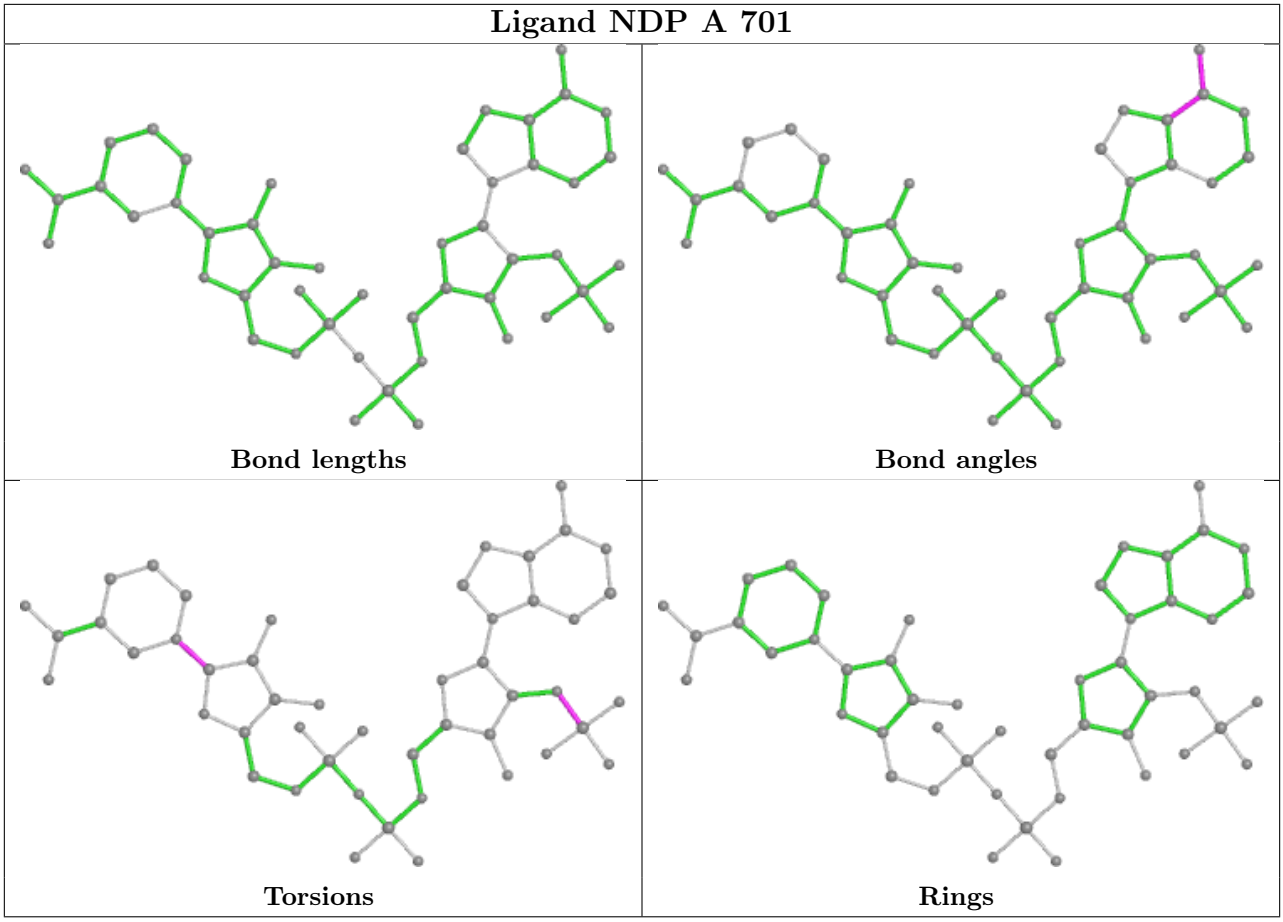
bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 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 GF3 A 702



Ligand GF3 B 902





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	529/608 (87%)	-0.22	14 (2%)	56	46	5, 19, 62, 97	0
1	B	512/608 (84%)	0.13	34 (6%)	18	11	5, 20, 90, 132	0
All	All	1041/1216 (85%)	-0.05	48 (4%)	32	22	5, 19, 81, 132	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	301	GLU	6.4
1	B	130	THR	5.5
1	B	136	PHE	4.9
1	B	70	TYR	4.5
1	B	300	GLU	4.4

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)

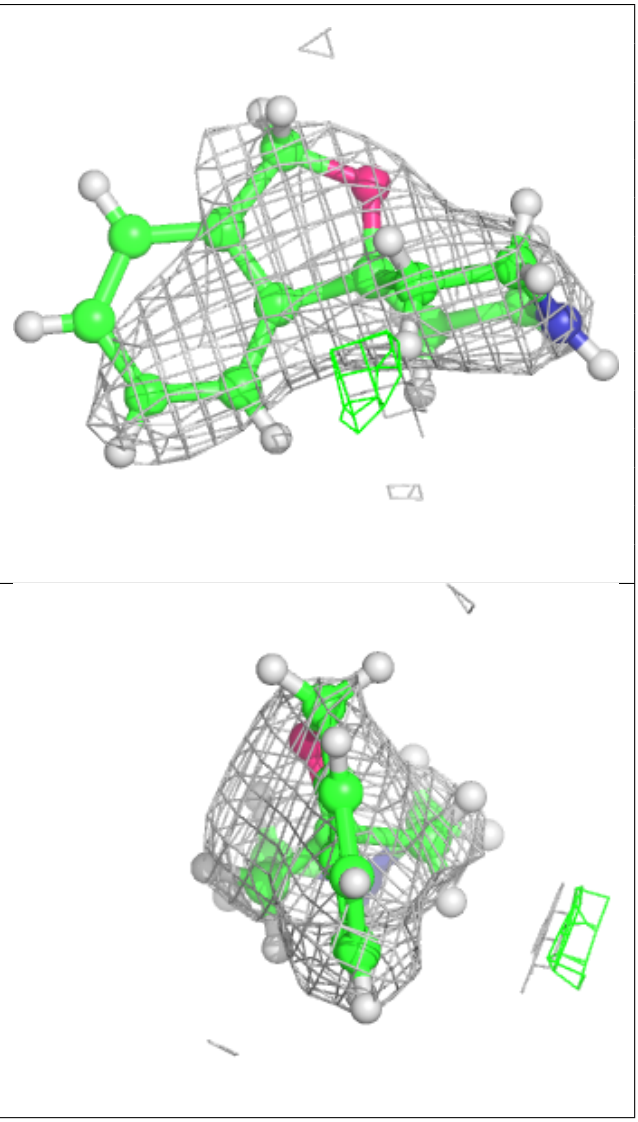
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GF3	B	902	14/14	0.88	0.29	0,47,59,60	0
4	PO4	B	901	5/5	0.93	0.18	43,47,50,57	0
3	GF3	A	702	14/14	0.94	0.17	0,20,21,21	0
4	PO4	A	703	5/5	0.95	0.16	35,36,38,40	0
2	NDP	A	701	48/48	0.97	0.14	14,19,23,26	7

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.

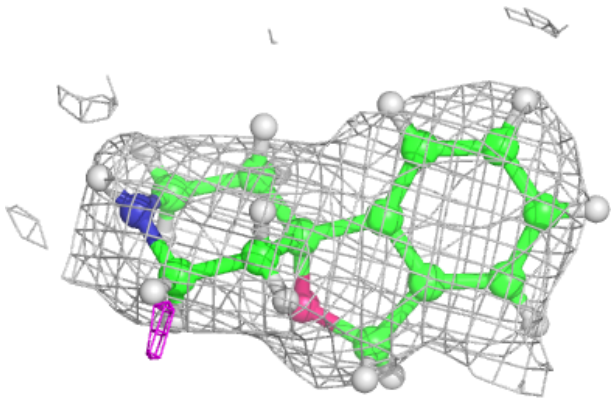
Electron density around GF3 B 902:

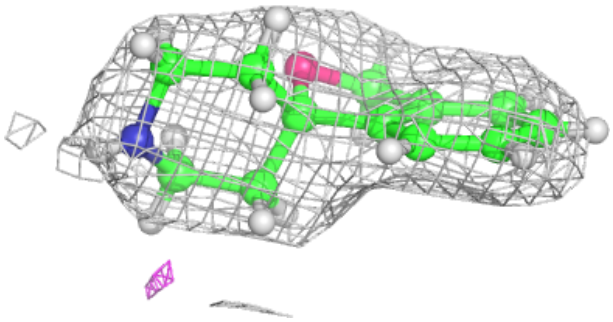
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

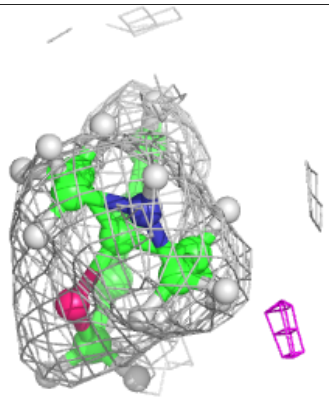


Electron density around GF3 A 702:

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

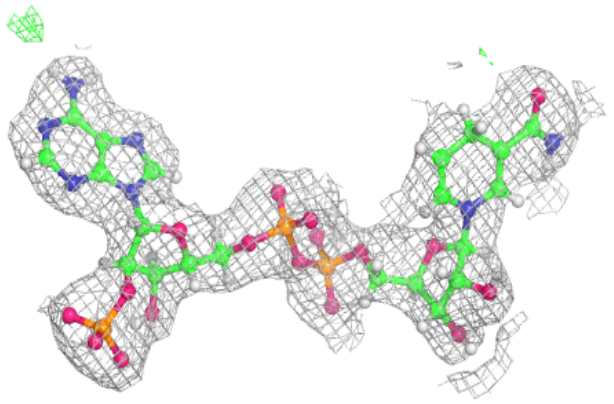


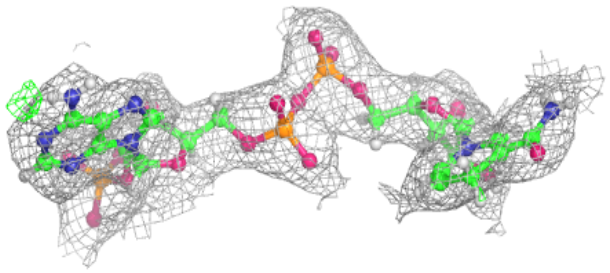


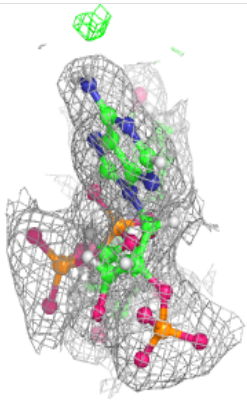


Electron density around NDP A 701:

$2mF_o$-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.