wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 06:08 pm BST

PDB ID	:	5IZD
Title	:	Wild-type glyceraldehyde dehydrogenase from Thermoplasma acidophilum in complex with NADP
Authors	:	Iermak, I.; Mesters, J.R.; Kuta Smatanova, I.
Deposited on	:	2016-03-25
Resolution	:	2.05 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 following versions of software and data (see references (1)) were used in the production of this report:

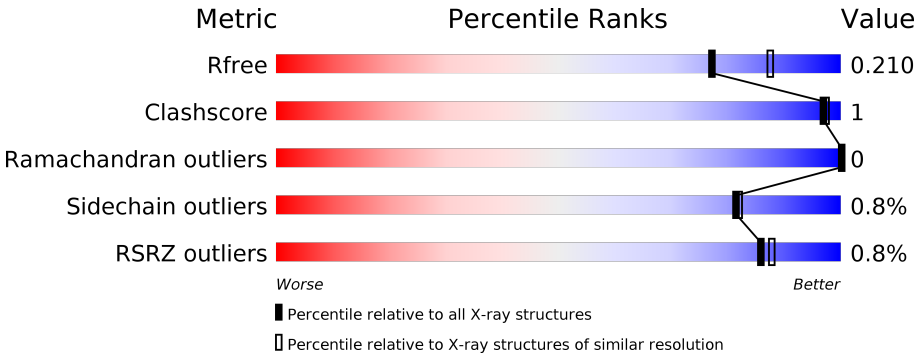
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.05 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria 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	508	94%
1	B	508	93%
1	C	508	92% 5%
1	D	508	91% 5%
1	E	508	92% 5%
1	F	508	93%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	508	
1	H	508	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 31616 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 D-glyceraldehyde dehydrogenase (NADP(+)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	494	Total	C	N	O	S	0	1	0
			3807	2430	634	730	13			
1	B	492	Total	C	N	O	S	0	1	0
			3776	2417	631	715	13			
1	C	484	Total	C	N	O	S	0	1	0
			3682	2355	618	696	13			
1	D	487	Total	C	N	O	S	0	0	0
			3682	2359	607	703	13			
1	E	492	Total	C	N	O	S	0	1	0
			3783	2420	626	724	13			
1	F	492	Total	C	N	O	S	0	2	0
			3790	2422	634	721	13			
1	G	492	Total	C	N	O	S	0	0	0
			3763	2409	624	717	13			
1	H	492	Total	C	N	O	S	0	3	0
			3794	2426	637	718	13			

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	494	SER	-	expression tag	UNP Q9HK01
A	495	GLY	-	expression tag	UNP Q9HK01
A	496	ARG	-	expression tag	UNP Q9HK01
A	497	PRO	-	expression tag	UNP Q9HK01
A	498	VAL	-	expression tag	UNP Q9HK01
A	499	LEU	-	expression tag	UNP Q9HK01
A	500	GLY	-	expression tag	UNP Q9HK01
A	501	SER	-	expression tag	UNP Q9HK01
A	502	SER	-	expression tag	UNP Q9HK01
A	503	HIS	-	expression tag	UNP Q9HK01
A	504	HIS	-	expression tag	UNP Q9HK01
A	505	HIS	-	expression tag	UNP Q9HK01
A	506	HIS	-	expression tag	UNP Q9HK01

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	507	HIS	-	expression tag	UNP Q9HK01
A	508	HIS	-	expression tag	UNP Q9HK01
B	494	SER	-	expression tag	UNP Q9HK01
B	495	GLY	-	expression tag	UNP Q9HK01
B	496	ARG	-	expression tag	UNP Q9HK01
B	497	PRO	-	expression tag	UNP Q9HK01
B	498	VAL	-	expression tag	UNP Q9HK01
B	499	LEU	-	expression tag	UNP Q9HK01
B	500	GLY	-	expression tag	UNP Q9HK01
B	501	SER	-	expression tag	UNP Q9HK01
B	502	SER	-	expression tag	UNP Q9HK01
B	503	HIS	-	expression tag	UNP Q9HK01
B	504	HIS	-	expression tag	UNP Q9HK01
B	505	HIS	-	expression tag	UNP Q9HK01
B	506	HIS	-	expression tag	UNP Q9HK01
B	507	HIS	-	expression tag	UNP Q9HK01
B	508	HIS	-	expression tag	UNP Q9HK01
C	494	SER	-	expression tag	UNP Q9HK01
C	495	GLY	-	expression tag	UNP Q9HK01
C	496	ARG	-	expression tag	UNP Q9HK01
C	497	PRO	-	expression tag	UNP Q9HK01
C	498	VAL	-	expression tag	UNP Q9HK01
C	499	LEU	-	expression tag	UNP Q9HK01
C	500	GLY	-	expression tag	UNP Q9HK01
C	501	SER	-	expression tag	UNP Q9HK01
C	502	SER	-	expression tag	UNP Q9HK01
C	503	HIS	-	expression tag	UNP Q9HK01
C	504	HIS	-	expression tag	UNP Q9HK01
C	505	HIS	-	expression tag	UNP Q9HK01
C	506	HIS	-	expression tag	UNP Q9HK01
C	507	HIS	-	expression tag	UNP Q9HK01
C	508	HIS	-	expression tag	UNP Q9HK01
D	494	SER	-	expression tag	UNP Q9HK01
D	495	GLY	-	expression tag	UNP Q9HK01
D	496	ARG	-	expression tag	UNP Q9HK01
D	497	PRO	-	expression tag	UNP Q9HK01
D	498	VAL	-	expression tag	UNP Q9HK01
D	499	LEU	-	expression tag	UNP Q9HK01
D	500	GLY	-	expression tag	UNP Q9HK01
D	501	SER	-	expression tag	UNP Q9HK01
D	502	SER	-	expression tag	UNP Q9HK01
D	503	HIS	-	expression tag	UNP Q9HK01

Continued on next page...

Continued from previous page...

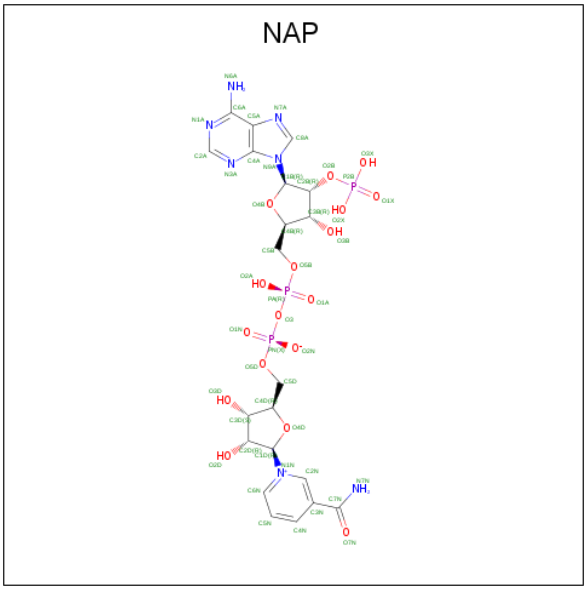
Chain	Residue	Modelled	Actual	Comment	Reference
D	504	HIS	-	expression tag	UNP Q9HK01
D	505	HIS	-	expression tag	UNP Q9HK01
D	506	HIS	-	expression tag	UNP Q9HK01
D	507	HIS	-	expression tag	UNP Q9HK01
D	508	HIS	-	expression tag	UNP Q9HK01
E	494	SER	-	expression tag	UNP Q9HK01
E	495	GLY	-	expression tag	UNP Q9HK01
E	496	ARG	-	expression tag	UNP Q9HK01
E	497	PRO	-	expression tag	UNP Q9HK01
E	498	VAL	-	expression tag	UNP Q9HK01
E	499	LEU	-	expression tag	UNP Q9HK01
E	500	GLY	-	expression tag	UNP Q9HK01
E	501	SER	-	expression tag	UNP Q9HK01
E	502	SER	-	expression tag	UNP Q9HK01
E	503	HIS	-	expression tag	UNP Q9HK01
E	504	HIS	-	expression tag	UNP Q9HK01
E	505	HIS	-	expression tag	UNP Q9HK01
E	506	HIS	-	expression tag	UNP Q9HK01
E	507	HIS	-	expression tag	UNP Q9HK01
E	508	HIS	-	expression tag	UNP Q9HK01
F	494	SER	-	expression tag	UNP Q9HK01
F	495	GLY	-	expression tag	UNP Q9HK01
F	496	ARG	-	expression tag	UNP Q9HK01
F	497	PRO	-	expression tag	UNP Q9HK01
F	498	VAL	-	expression tag	UNP Q9HK01
F	499	LEU	-	expression tag	UNP Q9HK01
F	500	GLY	-	expression tag	UNP Q9HK01
F	501	SER	-	expression tag	UNP Q9HK01
F	502	SER	-	expression tag	UNP Q9HK01
F	503	HIS	-	expression tag	UNP Q9HK01
F	504	HIS	-	expression tag	UNP Q9HK01
F	505	HIS	-	expression tag	UNP Q9HK01
F	506	HIS	-	expression tag	UNP Q9HK01
F	507	HIS	-	expression tag	UNP Q9HK01
F	508	HIS	-	expression tag	UNP Q9HK01
G	494	SER	-	expression tag	UNP Q9HK01
G	495	GLY	-	expression tag	UNP Q9HK01
G	496	ARG	-	expression tag	UNP Q9HK01
G	497	PRO	-	expression tag	UNP Q9HK01
G	498	VAL	-	expression tag	UNP Q9HK01
G	499	LEU	-	expression tag	UNP Q9HK01
G	500	GLY	-	expression tag	UNP Q9HK01

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	501	SER	-	expression tag	UNP Q9HK01
G	502	SER	-	expression tag	UNP Q9HK01
G	503	HIS	-	expression tag	UNP Q9HK01
G	504	HIS	-	expression tag	UNP Q9HK01
G	505	HIS	-	expression tag	UNP Q9HK01
G	506	HIS	-	expression tag	UNP Q9HK01
G	507	HIS	-	expression tag	UNP Q9HK01
G	508	HIS	-	expression tag	UNP Q9HK01
H	494	SER	-	expression tag	UNP Q9HK01
H	495	GLY	-	expression tag	UNP Q9HK01
H	496	ARG	-	expression tag	UNP Q9HK01
H	497	PRO	-	expression tag	UNP Q9HK01
H	498	VAL	-	expression tag	UNP Q9HK01
H	499	LEU	-	expression tag	UNP Q9HK01
H	500	GLY	-	expression tag	UNP Q9HK01
H	501	SER	-	expression tag	UNP Q9HK01
H	502	SER	-	expression tag	UNP Q9HK01
H	503	HIS	-	expression tag	UNP Q9HK01
H	504	HIS	-	expression tag	UNP Q9HK01
H	505	HIS	-	expression tag	UNP Q9HK01
H	506	HIS	-	expression tag	UNP Q9HK01
H	507	HIS	-	expression tag	UNP Q9HK01
H	508	HIS	-	expression tag	UNP Q9HK01

- Molecule 2 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
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	G	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	H	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 3 is water.

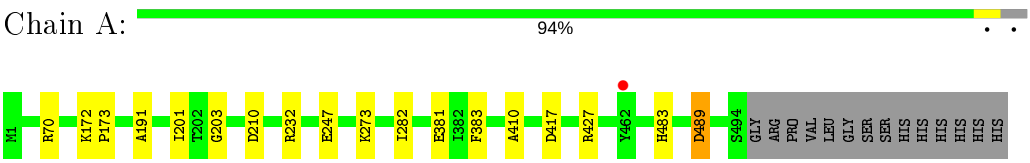
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	183	Total	O	0	0
			183	183		
3	B	164	Total	O	0	0
			164	164		
3	C	107	Total	O	0	0
			107	107		
3	D	106	Total	O	0	0
			106	106		
3	E	188	Total	O	0	0
			188	188		
3	F	195	Total	O	0	0
			195	195		
3	G	192	Total	O	0	0
			192	192		
3	H	218	Total	O	0	0
			218	218		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

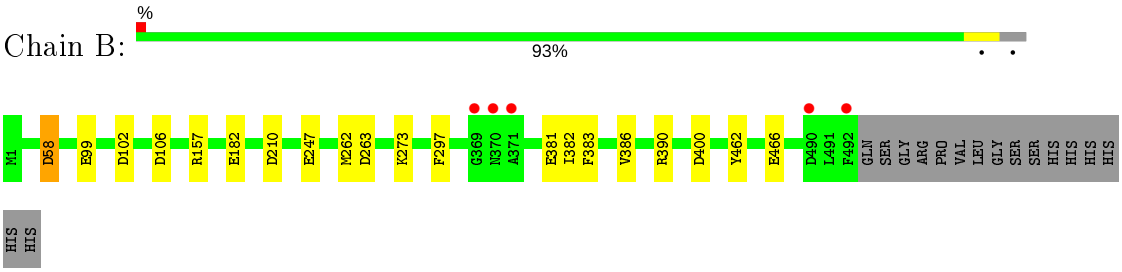
- Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain A:



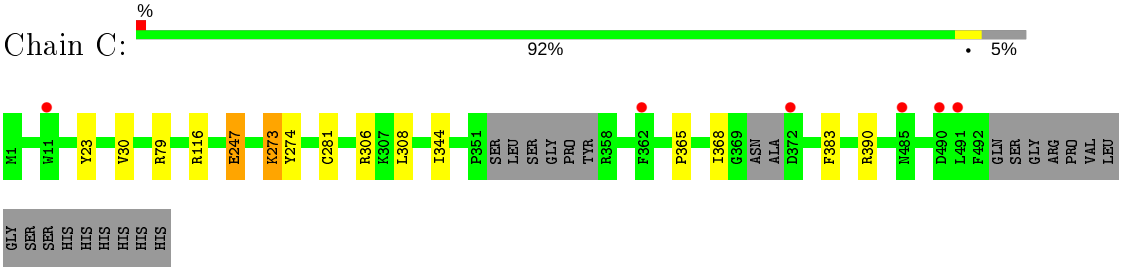
- Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain B:



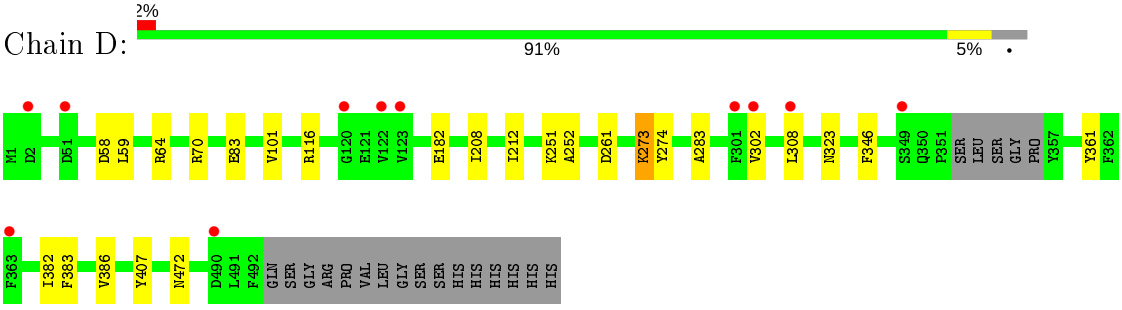
- Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain C:



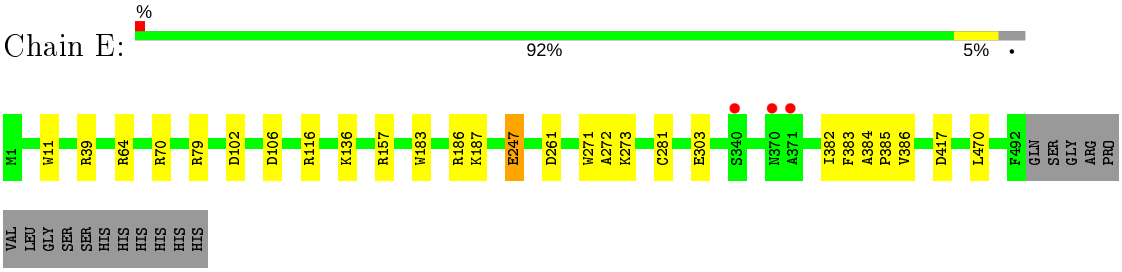
- Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain D:



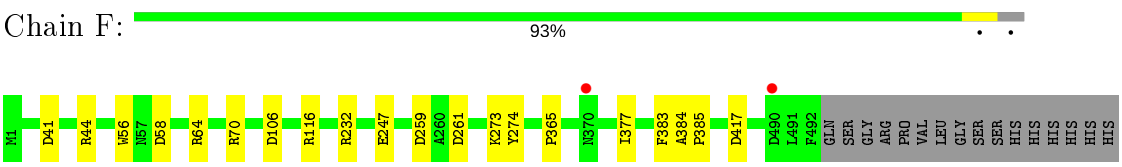
• Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain E:



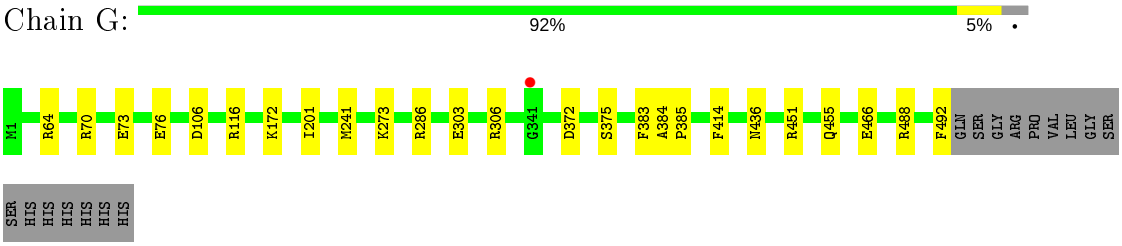
• Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain F:



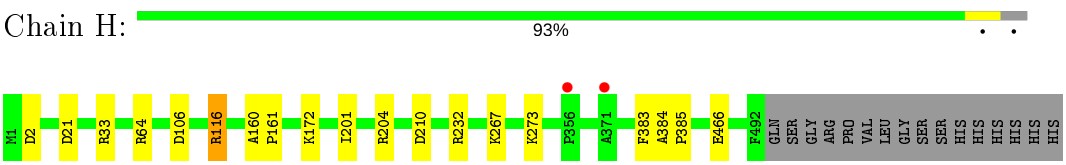
• Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain G:



• Molecule 1: D-glyceraldehyde dehydrogenase (NADP(+))

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	95.29Å 152.35Å 149.91Å 90.00° 92.19° 90.00°	Depositor
Resolution (Å)	48.10 – 2.05 48.10 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.10-2.05) 99.8 (48.10-1.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.59 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.168 , 0.205 0.177 , 0.210	Depositor DCC
R_{free} test set	15542 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	30.2	Xtriage
Anisotropy	0.317	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 45.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.018 for -h,-l,-k 0.015 for -h,l,k 0.030 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	31616	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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: CSD, NAP

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	0.86	2/3882 (0.1%)	0.88	2/5263 (0.0%)
1	B	0.80	2/3854 (0.1%)	0.88	12/5223 (0.2%)
1	C	0.82	0/3756	0.87	6/5090 (0.1%)
1	D	0.81	0/3754	0.84	6/5093 (0.1%)
1	E	0.85	0/3861	0.93	10/5233 (0.2%)
1	F	0.83	0/3871	0.88	9/5247 (0.2%)
1	G	0.87	3/3838 (0.1%)	0.93	9/5205 (0.2%)
1	H	0.88	1/3878 (0.0%)	0.91	7/5252 (0.1%)
All	All	0.84	8/30694 (0.0%)	0.89	61/41606 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	381	GLU	CD-OE2	6.67	1.32	1.25
1	G	76	GLU	CD-OE2	6.15	1.32	1.25
1	H	466	GLU	CD-OE2	5.74	1.31	1.25
1	G	466	GLU	CD-OE2	5.56	1.31	1.25
1	A	489	ASP	CB-CG	5.50	1.63	1.51

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	106	ASP	CB-CG-OD2	8.58	126.03	118.30
1	H	106	ASP	CB-CG-OD2	8.29	125.77	118.30
1	G	64	ARG	NE-CZ-NH1	7.76	124.18	120.30
1	C	390	ARG	NE-CZ-NH2	7.61	124.11	120.30
1	H	64	ARG	NE-CZ-NH1	7.18	123.89	120.30

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	A	3807	0	3687	8	0
1	B	3776	0	3666	4	0
1	C	3682	0	3544	6	0
1	D	3682	0	3528	10	0
1	E	3783	0	3668	9	1
1	F	3790	0	3686	5	0
1	G	3763	0	3650	6	1
1	H	3794	0	3694	5	0
2	A	31	0	11	0	0
2	B	31	0	11	0	0
2	C	31	0	11	0	0
2	D	31	0	11	0	0
2	G	31	0	11	0	0
2	H	31	0	11	0	0
3	A	183	0	0	0	0
3	B	164	0	0	0	0
3	C	107	0	0	0	0
3	D	106	0	0	0	0
3	E	188	0	0	0	0
3	F	195	0	0	0	0
3	G	192	0	0	1	0
3	H	218	0	0	0	0
All	All	31616	0	29189	51	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 1.

The worst 5 of 51 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:172:LYS:HE2	1:A:203:GLY:O	1.94	0.67
1:G:455:GLN:OE1	3:G:701:HOH:O	2.16	0.63
1:D:323:ASN:HA	1:D:361:TYR:CE2	2.36	0.61
1:A:172:LYS:NZ	1:A:173:PRO:O	2.33	0.61
1:D:382:ILE:HG21	1:D:386:VAL:HB	1.83	0.60

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:E:186:ARG:NH1	1:G:303:GLU:OE1[2_656]	2.16	0.04

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	492/508 (97%)	482 (98%)	10 (2%)	0	100	100
1	B	490/508 (96%)	482 (98%)	8 (2%)	0	100	100
1	C	478/508 (94%)	466 (98%)	12 (2%)	0	100	100
1	D	482/508 (95%)	467 (97%)	15 (3%)	0	100	100
1	E	490/508 (96%)	481 (98%)	9 (2%)	0	100	100
1	F	491/508 (97%)	481 (98%)	10 (2%)	0	100	100
1	G	489/508 (96%)	473 (97%)	16 (3%)	0	100	100
1	H	492/508 (97%)	482 (98%)	10 (2%)	0	100	100
All	All	3904/4064 (96%)	3814 (98%)	90 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	385/418 (92%)	382 (99%)	3 (1%)	81	82
1	B	377/418 (90%)	373 (99%)	4 (1%)	73	73
1	C	363/418 (87%)	360 (99%)	3 (1%)	81	82
1	D	362/418 (87%)	358 (99%)	4 (1%)	73	73
1	E	381/418 (91%)	378 (99%)	3 (1%)	81	82
1	F	383/418 (92%)	380 (99%)	3 (1%)	81	82
1	G	378/418 (90%)	376 (100%)	2 (0%)	88	89
1	H	381/418 (91%)	378 (99%)	3 (1%)	81	82
All	All	3010/3344 (90%)	2985 (99%)	25 (1%)	81	82

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	273	LYS
1	D	472	ASN
1	H	273	LYS
1	D	383	PHE
1	E	247	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	350	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)

8 non-standard protein/DNA/RNA residues 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	$\#\|Z\|>2$
1	CSD	B	281	1	3,7,8	0.32	0	1,8,10	1.78	0
1	CSD	A	281	1	3,7,8	0.81	0	1,8,10	5.04	1 (100%)
1	CSD	D	281	1	3,7,8	0.81	0	1,8,10	8.14	1 (100%)
1	CSD	C	281	1	3,7,8	0.58	0	1,8,10	0.26	0
1	CSD	F	281	1	3,7,8	0.63	0	1,8,10	3.82	1 (100%)
1	CSD	E	281	1	3,7,8	1.03	0	1,8,10	2.06	1 (100%)
1	CSD	H	281	1	3,7,8	0.63	0	1,8,10	2.19	1 (100%)
1	CSD	G	281	1	3,7,8	0.67	0	1,8,10	2.71	1 (100%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CSD	B	281	1	-	1/2/6/8	-
1	CSD	A	281	1	-	1/2/6/8	-
1	CSD	D	281	1	-	1/2/6/8	-
1	CSD	C	281	1	-	1/2/6/8	-
1	CSD	F	281	1	-	1/2/6/8	-
1	CSD	E	281	1	-	1/2/6/8	-
1	CSD	H	281	1	-	1/2/6/8	-
1	CSD	G	281	1	-	1/2/6/8	-

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	281	CSD	OD1-SG-CB	-8.14	90.04	105.54
1	A	281	CSD	OD1-SG-CB	-5.04	95.94	105.54
1	F	281	CSD	OD1-SG-CB	-3.82	98.26	105.54
1	G	281	CSD	OD1-SG-CB	-2.71	100.38	105.54
1	H	281	CSD	OD1-SG-CB	-2.19	101.38	105.54

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	281	CSD	CA-CB-SG-OD1
1	A	281	CSD	CA-CB-SG-OD1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	D	281	CSD	CA-CB-SG-OD1
1	C	281	CSD	CA-CB-SG-OD1
1	F	281	CSD	CA-CB-SG-OD1

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	281	CSD	1	0
1	E	281	CSD	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

6 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	NAP	D	601	-	27,33,52	1.09	3 (11%)	35,52,80	0.87	1 (2%)
2	NAP	B	601	-	27,33,52	1.04	1 (3%)	35,52,80	0.80	1 (2%)
2	NAP	H	601	-	27,33,52	1.28	2 (7%)	35,52,80	0.83	0
2	NAP	A	601	-	27,33,52	1.42	4 (14%)	35,52,80	0.82	0
2	NAP	G	601	-	27,33,52	1.06	2 (7%)	35,52,80	0.78	1 (2%)
2	NAP	C	601	-	27,33,52	1.09	3 (11%)	35,52,80	0.79	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
2	NAP	D	601	-	-	5/17/37/67	0/3/3/5
2	NAP	B	601	-	-	1/17/37/67	0/3/3/5
2	NAP	H	601	-	-	4/17/37/67	0/3/3/5
2	NAP	A	601	-	-	4/17/37/67	0/3/3/5
2	NAP	G	601	-	-	4/17/37/67	0/3/3/5
2	NAP	C	601	-	-	3/17/37/67	0/3/3/5

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	601	NAP	O4B-C1B	3.83	1.46	1.41
2	A	601	NAP	O4B-C1B	3.81	1.46	1.41
2	A	601	NAP	P2B-O2B	3.71	1.66	1.59
2	B	601	NAP	O4B-C1B	3.57	1.46	1.41
2	C	601	NAP	O4B-C1B	3.57	1.46	1.41

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	601	NAP	C5A-C6A-N6A	2.30	123.85	120.35
2	G	601	NAP	C5A-C6A-N6A	2.18	123.66	120.35
2	B	601	NAP	C5A-C6A-N6A	2.15	123.62	120.35

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	601	NAP	C5B-O5B-PA-O1A
2	A	601	NAP	C5B-O5B-PA-O2A
2	G	601	NAP	C5B-O5B-PA-O1A
2	G	601	NAP	C5B-O5B-PA-O2A
2	G	601	NAP	C2B-O2B-P2B-O3X

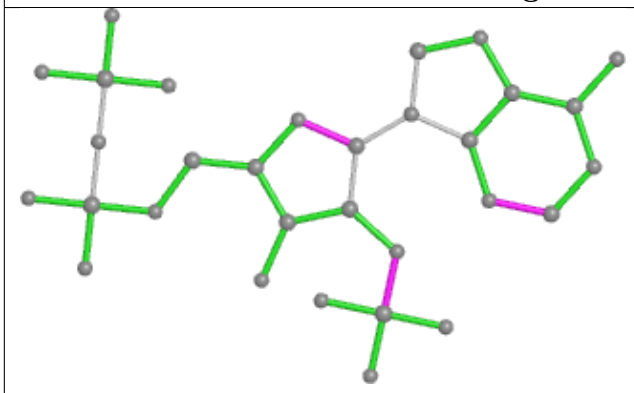
There are no ring outliers.

No monomer is involved in short contacts.

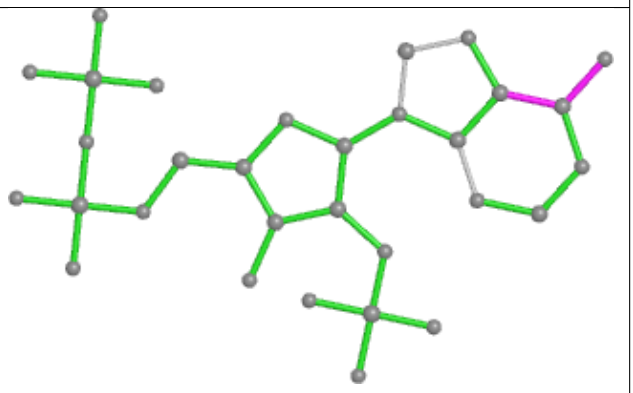
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

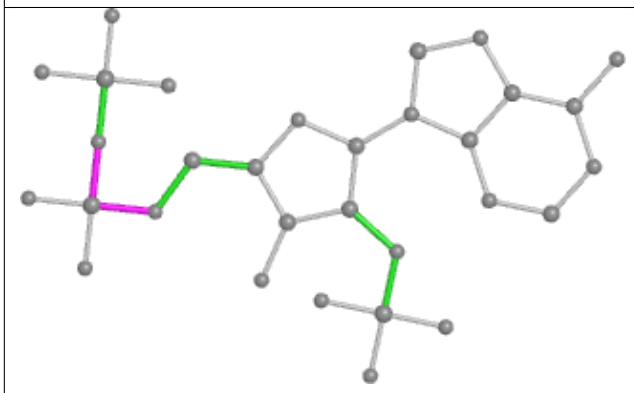
Ligand NAP D 601



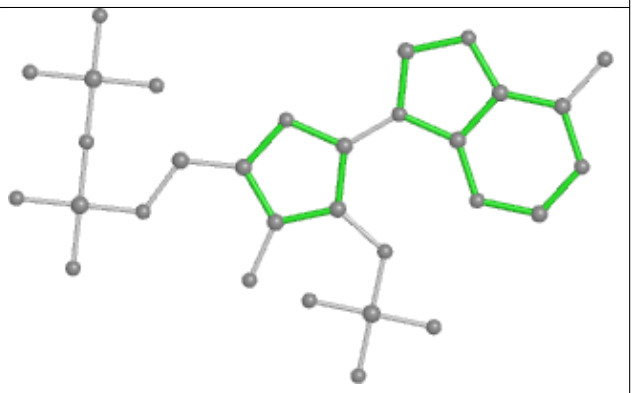
Bond lengths



Bond angles

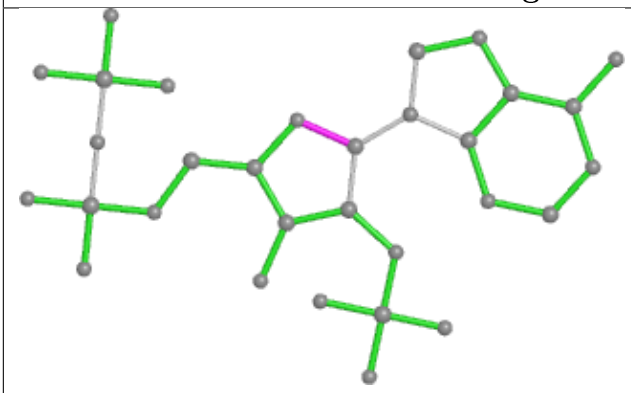


Torsions

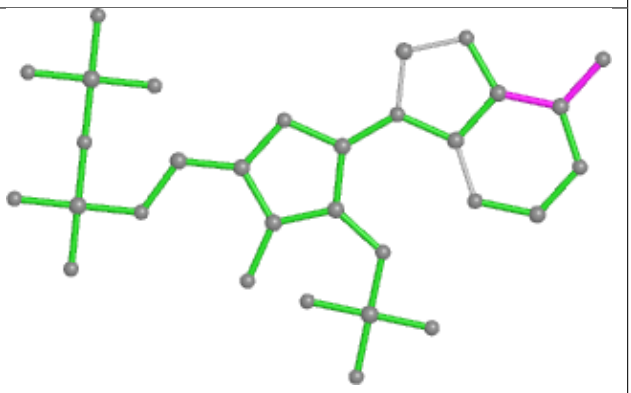


Rings

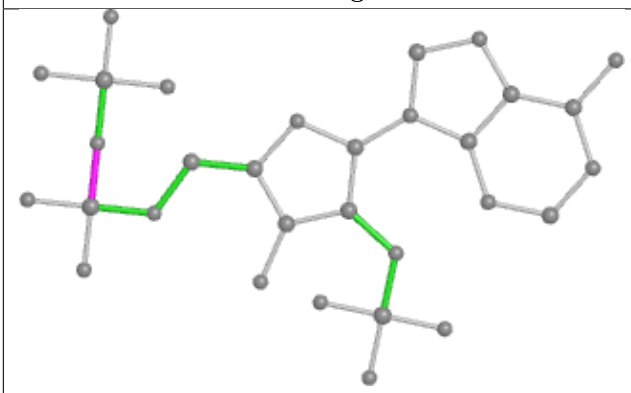
Ligand NAP B 601



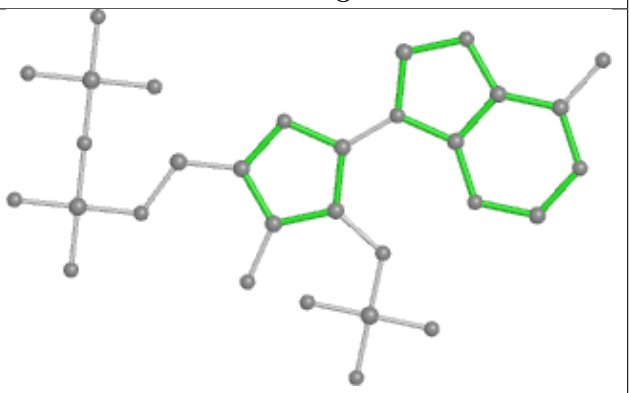
Bond lengths



Bond angles

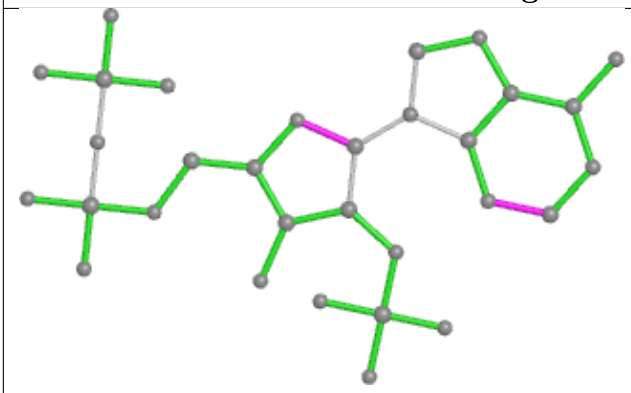


Torsions

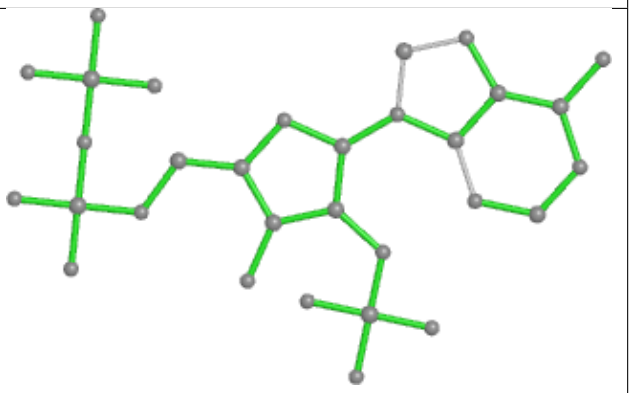


Rings

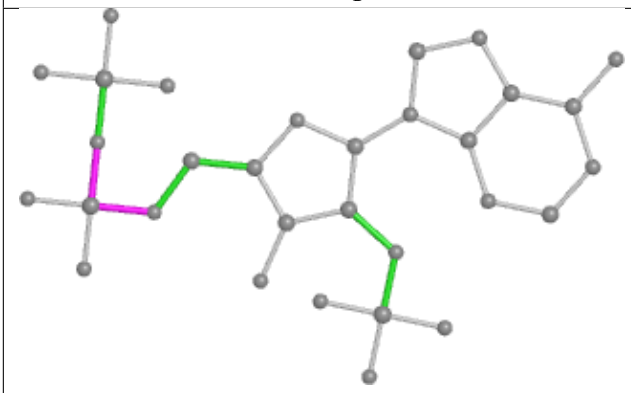
Ligand NAP H 601



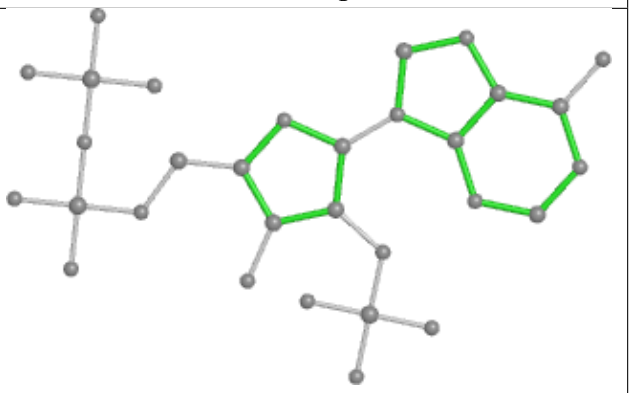
Bond lengths



Bond angles

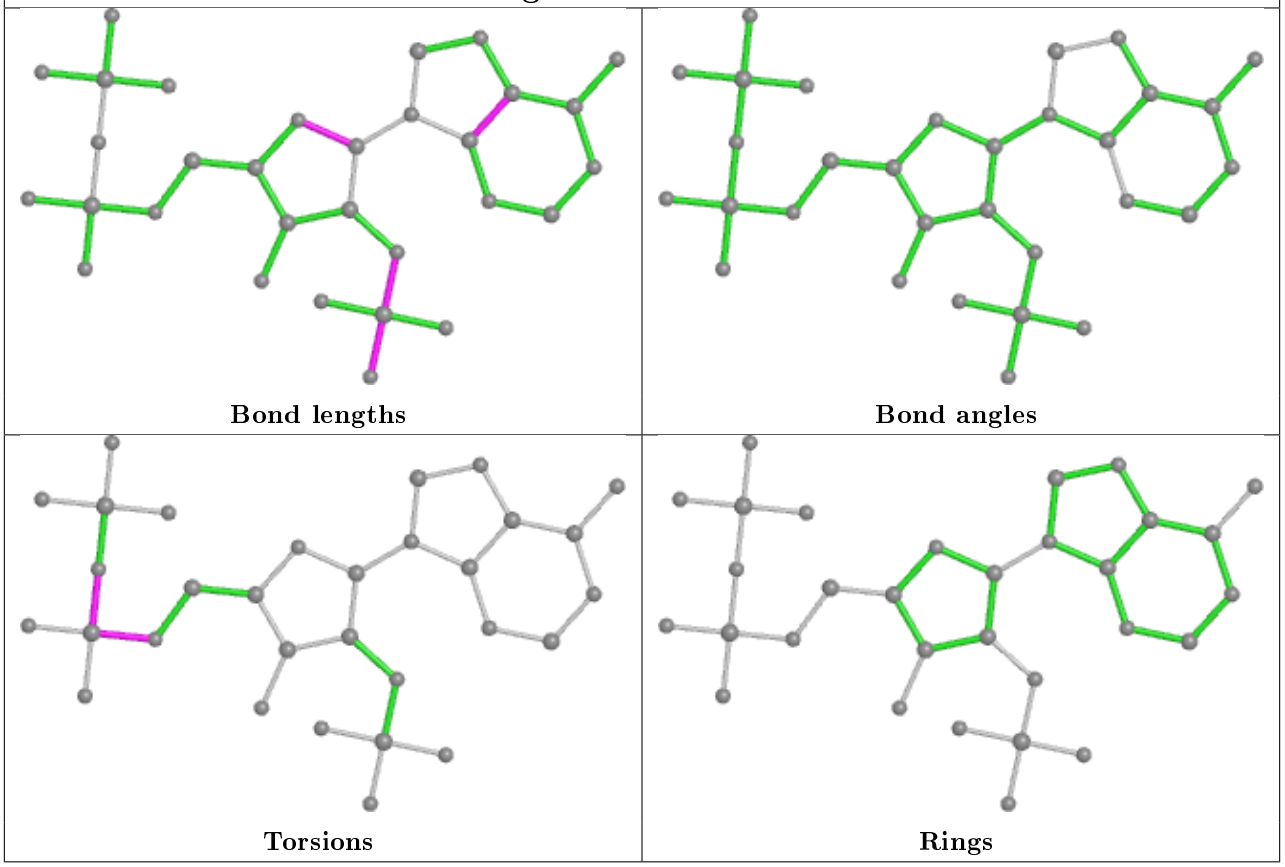


Torsions

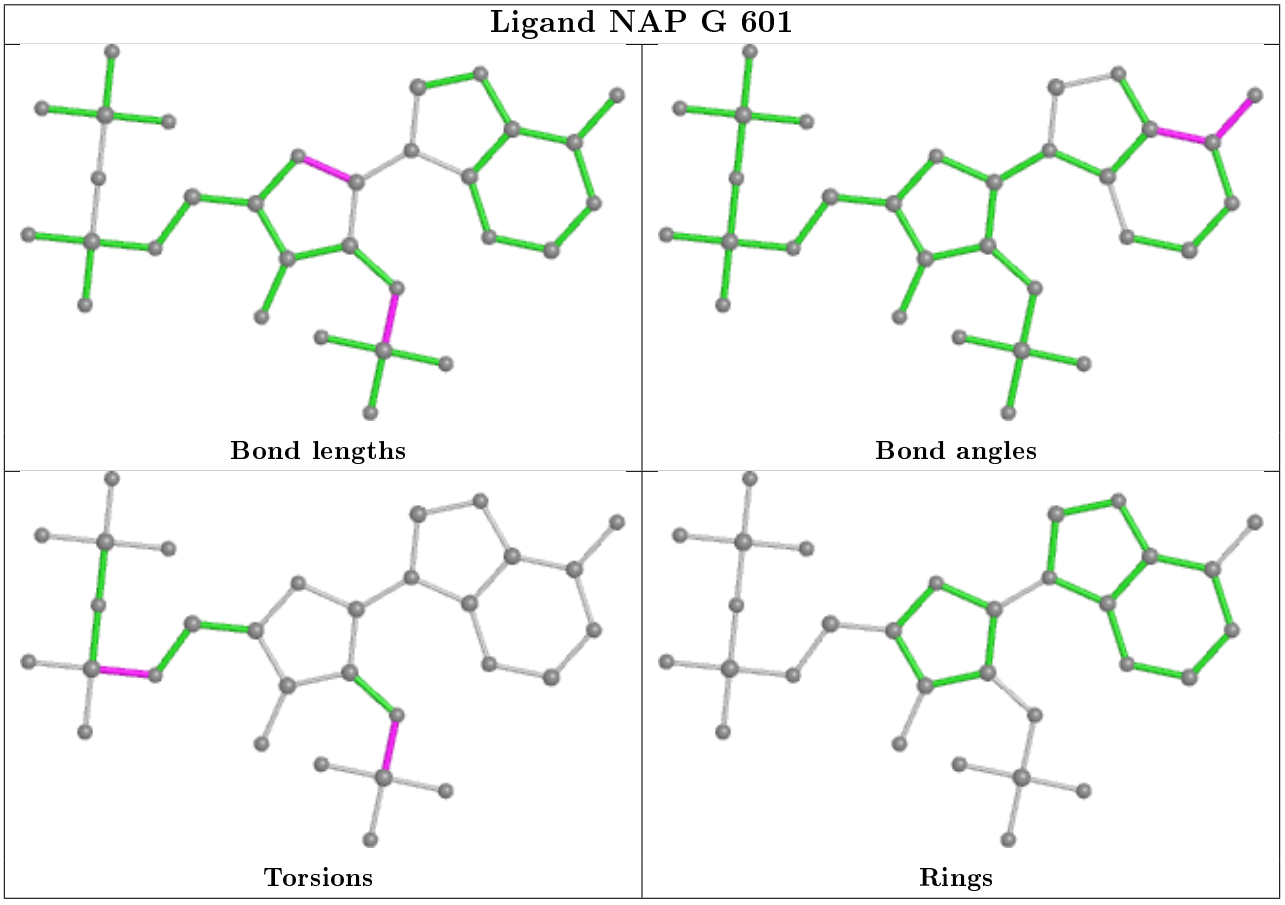


Rings

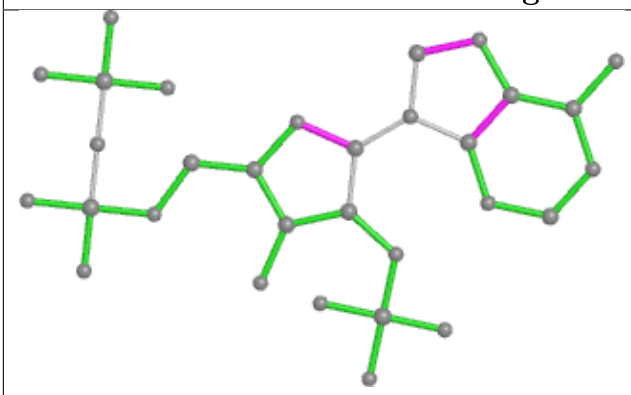
Ligand NAP A 601



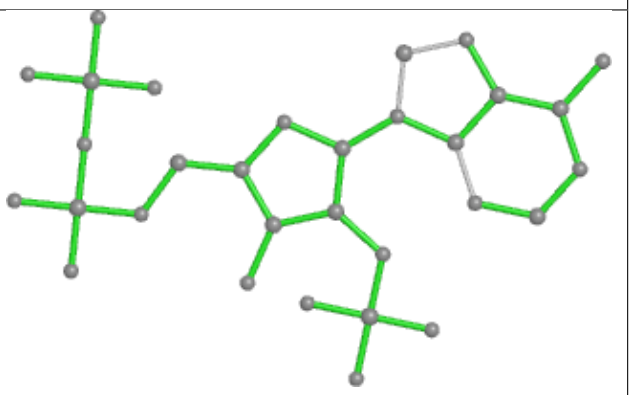
Ligand NAP G 601



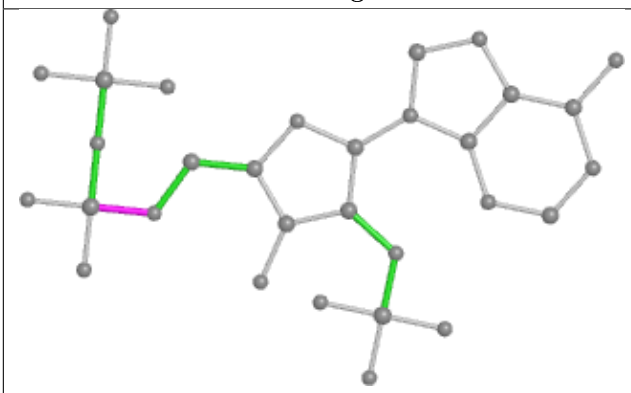
Ligand NAP C 601



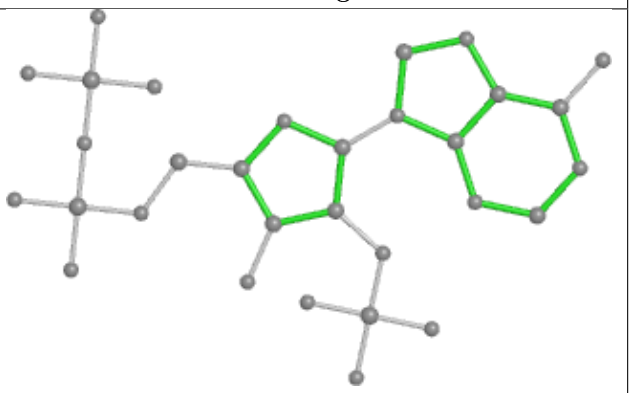
Bond lengths



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, 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	493/508 (97%)	-0.33	1 (0%)	95	95	27, 34, 48, 59	2 (0%)
1	B	491/508 (96%)	-0.37	5 (1%)	82	84	26, 36, 51, 65	0
1	C	483/508 (95%)	-0.01	6 (1%)	79	81	28, 43, 58, 70	2 (0%)
1	D	486/508 (95%)	0.08	11 (2%)	60	64	30, 44, 61, 73	2 (0%)
1	E	491/508 (96%)	-0.36	3 (0%)	89	91	22, 34, 51, 66	1 (0%)
1	F	491/508 (96%)	-0.38	2 (0%)	92	93	22, 33, 48, 59	0
1	G	491/508 (96%)	-0.40	1 (0%)	95	95	21, 33, 51, 64	1 (0%)
1	H	491/508 (96%)	-0.41	2 (0%)	92	93	21, 31, 50, 67	2 (0%)
All	All	3917/4064 (96%)	-0.27	31 (0%)	86	88	21, 36, 55, 73	10 (0%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	491	LEU	4.0
1	D	308	LEU	3.6
1	C	372	ASP	3.2
1	D	490	ASP	3.2
1	B	490	ASP	3.1

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

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
1	CSD	C	281	8/9	0.90	0.14	43,45,50,51	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	CSD	F	281	8/9	0.93	0.11	32,33,41,42	0
1	CSD	B	281	8/9	0.94	0.08	32,34,41,41	0
1	CSD	H	281	8/9	0.94	0.10	28,30,37,37	0
1	CSD	E	281	8/9	0.96	0.08	32,33,40,42	0
1	CSD	D	281	8/9	0.96	0.11	42,44,47,48	0
1	CSD	A	281	8/9	0.97	0.08	30,32,40,41	0
1	CSD	G	281	8/9	0.97	0.08	29,31,35,36	0

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

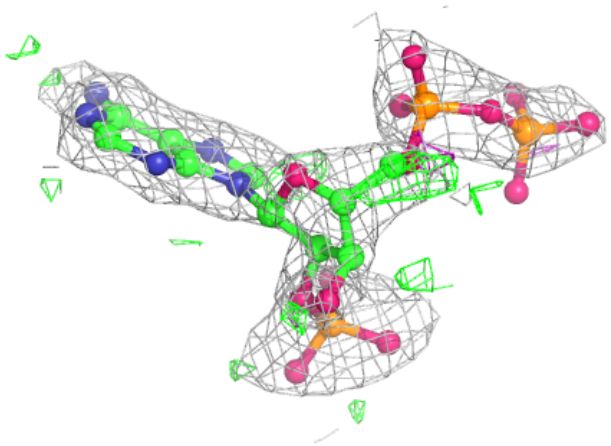
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

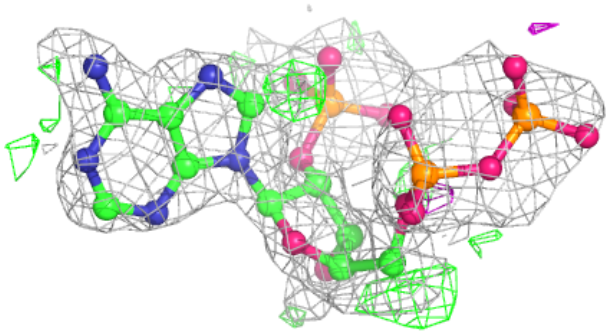
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAP	D	601	31/48	0.81	0.20	41,51,63,67	31
2	NAP	B	601	31/48	0.81	0.20	41,46,59,63	31
2	NAP	G	601	31/48	0.87	0.16	30,40,54,59	31
2	NAP	C	601	31/48	0.93	0.10	27,39,60,63	0
2	NAP	A	601	31/48	0.94	0.11	26,31,67,70	0
2	NAP	H	601	31/48	0.96	0.09	26,32,54,55	0

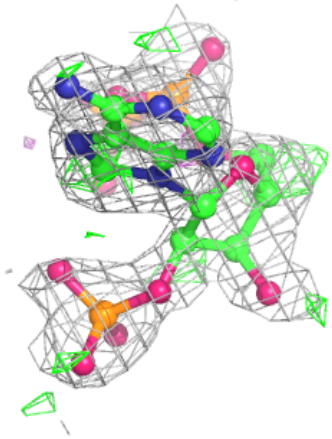
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around NAP D 601:

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

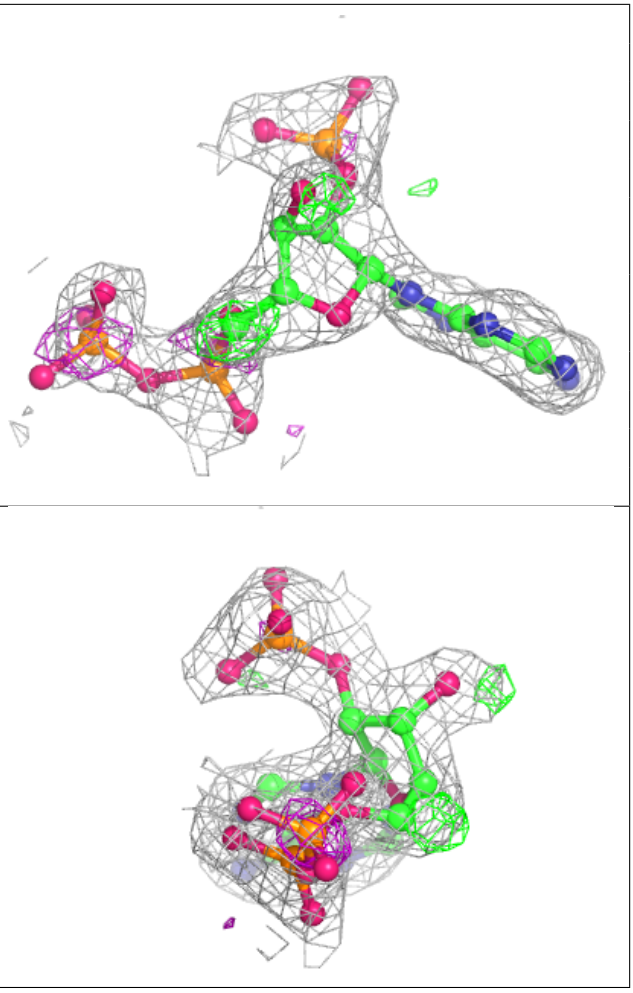






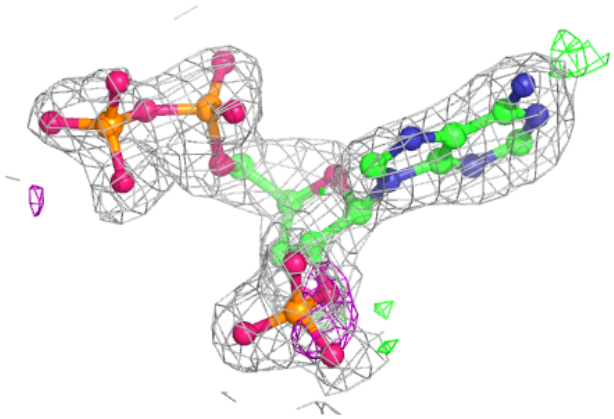
Electron density around NAP B 601:

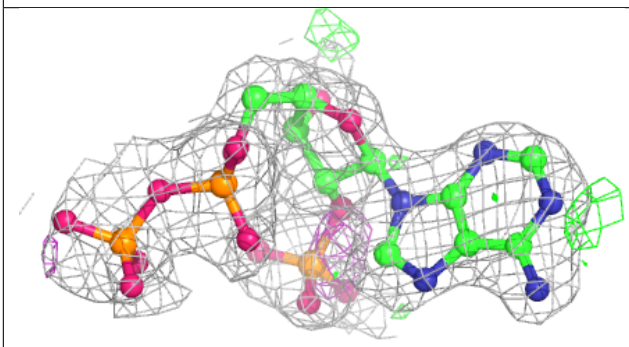
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

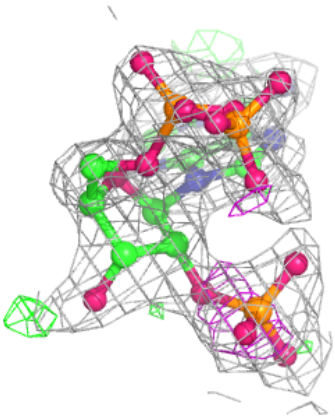


Electron density around NAP G 601:

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

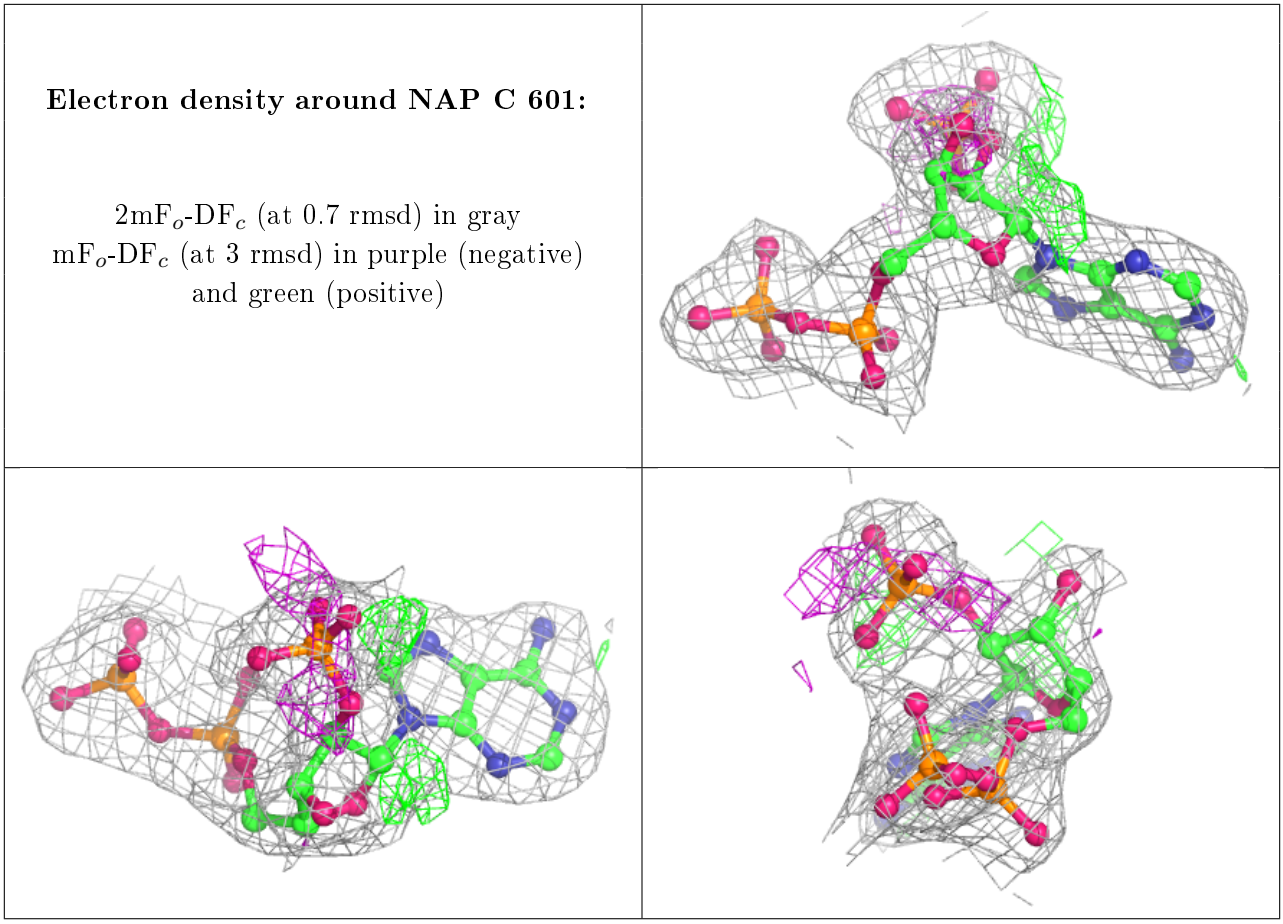






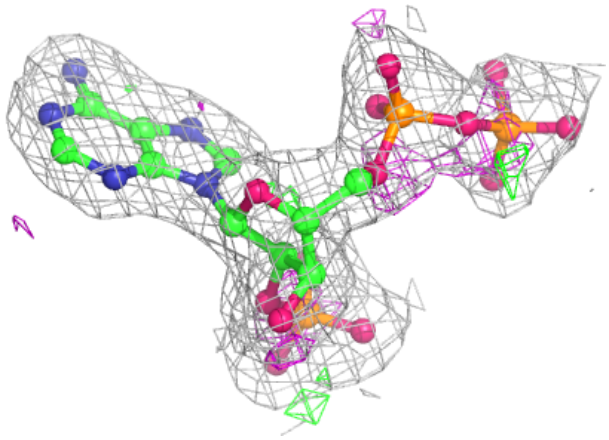
Electron density around NAP C 601:

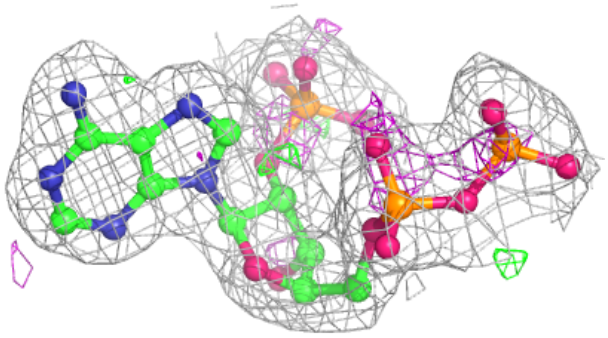
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

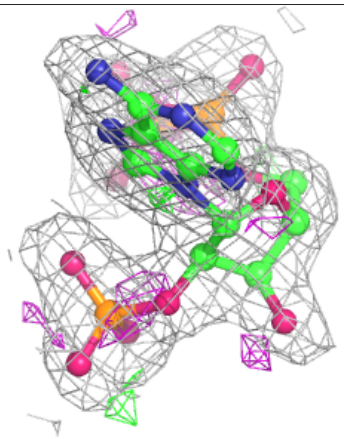


Electron density around NAP A 601:

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

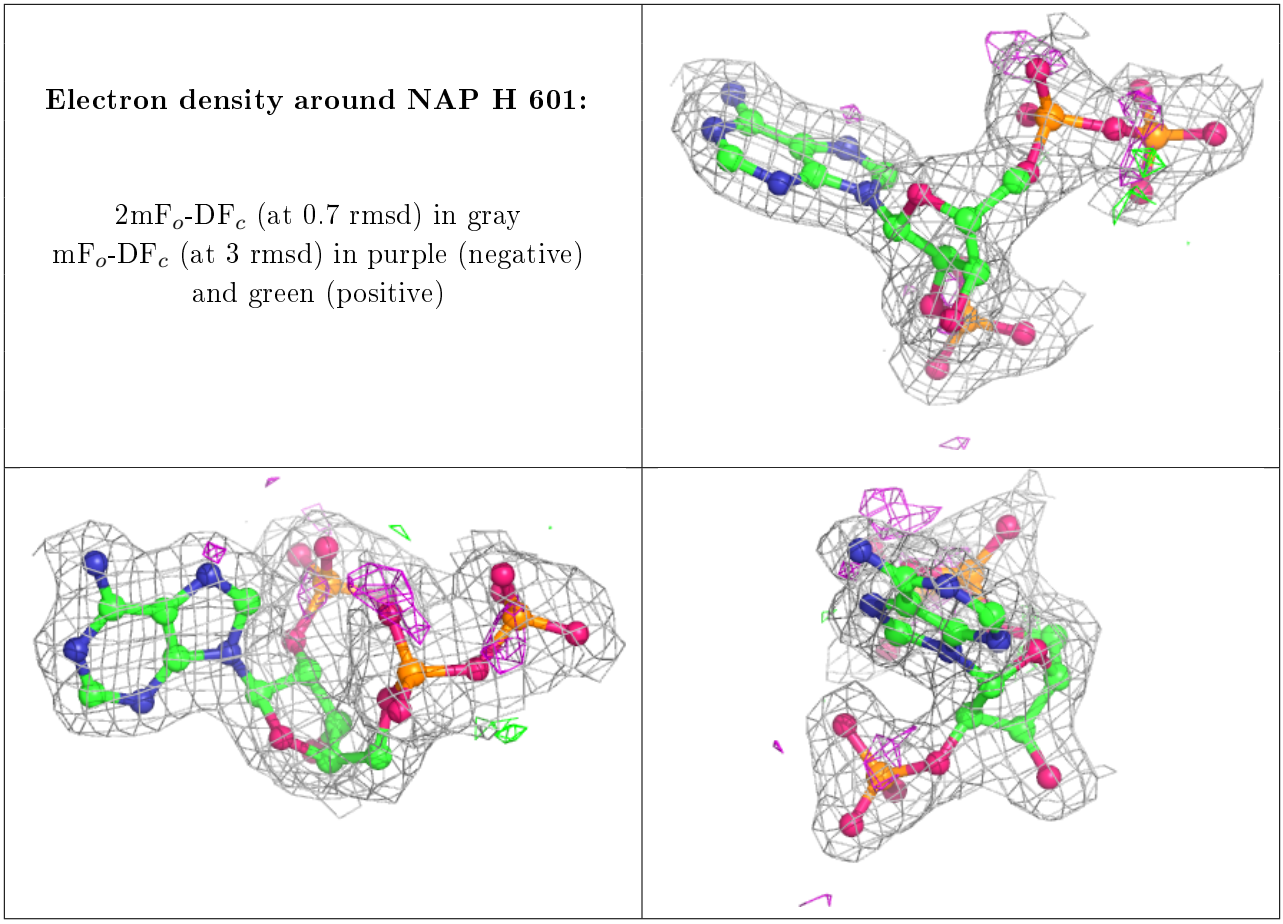






Electron density around NAP H 601:

$2mF_o$-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.