



wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2025 – 05:27 pm BST

PDB ID : 8QHN / pdb_00008qhn
Title : Streptococcus pyogenes GapN in complex with NADPH and erythrose-4-phosphate
Authors : Wirsing, R.; Schindelin, H.
Deposited on : 2023-09-08
Resolution : 1.99 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0rc1
Mogul : 1.8.4, CSD as541be (2020)
Xtrriage (Phenix) : 2.0rc1
EDS : 3.0
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.003 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.43.1

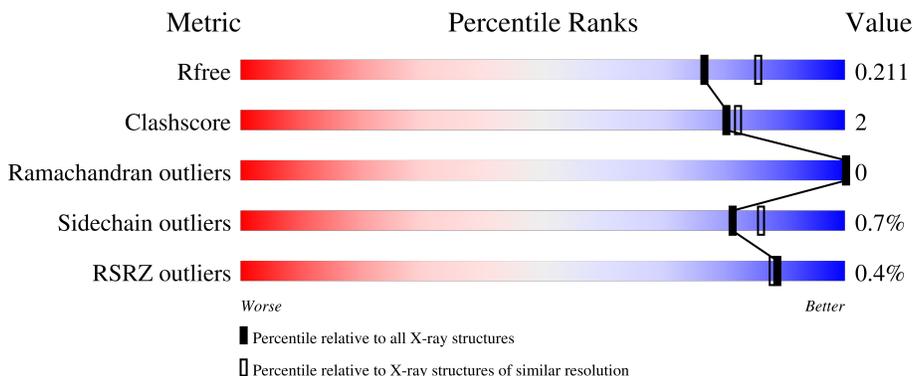
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 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.



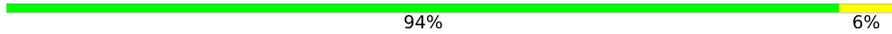
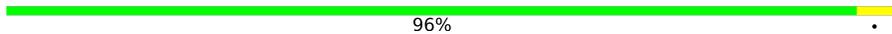
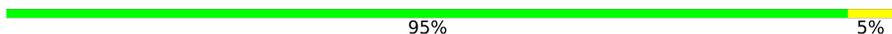
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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 $\leq 5\%$. 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	475	 94% 6%
1	B	475	 94% 5%
1	C	475	 96% .
1	E	475	 94% 5%
1	F	475	 93% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	475	 94% 6%
1	H	475	 96% .
2	D	480	 95% 5%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 61216 atoms, of which 29501 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 NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	474	7195	2268	3634	592	692	9	0	5	0
1	B	474	7281	2298	3680	593	702	8	0	11	0
1	C	474	7225	2277	3652	592	695	9	3	8	0
1	E	474	7216	2274	3647	594	692	9	3	5	0
1	F	474	7228	2285	3647	591	697	8	0	7	0
1	G	474	7225	2277	3650	592	698	8	3	8	0
1	H	474	7191	2267	3631	591	694	8	0	5	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	58	THR	ALA	conflict	UNP A0A4U9C786
A	284	SER	CYS	conflict	UNP A0A4U9C786
B	58	THR	ALA	conflict	UNP A0A4U9C786
B	284	SER	CYS	conflict	UNP A0A4U9C786
C	58	THR	ALA	conflict	UNP A0A4U9C786
C	284	SER	CYS	conflict	UNP A0A4U9C786
E	58	THR	ALA	conflict	UNP A0A4U9C786
E	284	SER	CYS	conflict	UNP A0A4U9C786
F	58	THR	ALA	conflict	UNP A0A4U9C786
F	284	SER	CYS	conflict	UNP A0A4U9C786
G	58	THR	ALA	conflict	UNP A0A4U9C786
G	284	SER	CYS	conflict	UNP A0A4U9C786
H	58	THR	ALA	conflict	UNP A0A4U9C786
H	284	SER	CYS	conflict	UNP A0A4U9C786

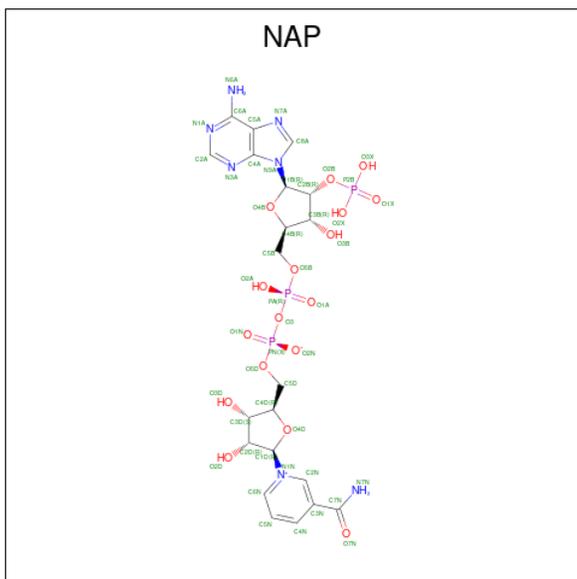
- Molecule 2 is a protein called NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
2	D	480	7364	2315	3724	607	709	9	1	11	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-10	ILE	-	expression tag	UNP A0A4U9C786
D	-9	GLU	-	expression tag	UNP A0A4U9C786
D	-8	GLY	-	expression tag	UNP A0A4U9C786
D	-7	ARG	-	expression tag	UNP A0A4U9C786
D	-6	ARG	-	expression tag	UNP A0A4U9C786
D	-5	ASP	-	expression tag	UNP A0A4U9C786
D	58	THR	ALA	conflict	UNP A0A4U9C786
D	284	SER	CYS	conflict	UNP A0A4U9C786

- Molecule 3 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (CCD ID: NAP) (formula: C₂₁H₂₈N₇O₁₇P₃) (labeled as "Ligand of Interest" by depositor).



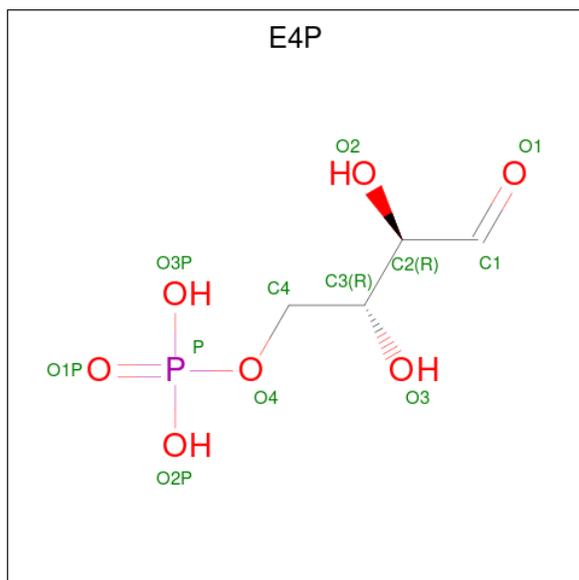
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			P
3	A	1	73	21	25	7	17	3	0	0
3	B	1	73	21	25	7	17	3	0	0
3	E	1	73	21	25	7	17	3	0	0

Continued on next page...

Continued from previous page...

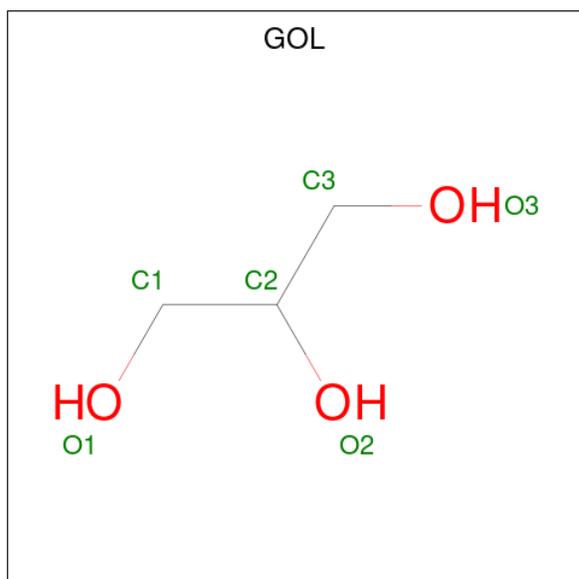
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			P
3	F	1	73	21	25	7	17	3	0	0

- Molecule 4 is ERYTHROSE-4-PHOSPHATE (CCD ID: E4P) (formula: $C_4H_9O_7P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
4	A	1	12	4	7	1	0	0

- Molecule 5 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			14	3	8	3		
5	B	1	Total	C	H	O	0	0
			14	3	8	3		
5	B	1	Total	C	H	O	0	0
			14	3	8	3		
5	C	1	Total	C	H	O	0	0
			14	3	8	3		
5	C	1	Total	C	H	O	0	0
			14	3	8	3		
5	C	1	Total	C	H	O	0	0
			14	3	8	3		
5	C	1	Total	C	H	O	0	0
			14	3	8	3		
5	D	1	Total	C	H	O	0	0
			14	3	8	3		
5	E	1	Total	C	H	O	0	0
			14	3	8	3		
5	F	1	Total	C	H	O	0	0
			14	3	8	3		
5	F	1	Total	C	H	O	0	0
			14	3	8	3		
5	G	1	Total	C	H	O	0	0
			14	3	8	3		
5	G	1	Total	C	H	O	0	0
			14	3	8	3		
5	G	1	Total	C	H	O	0	0
			14	3	8	3		
5	H	1	Total	C	H	O	0	0
			14	3	8	3		
5	H	1	Total	C	H	O	0	0
			14	3	8	3		
5	H	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	288	Total	O	0	0
			288	288		
6	B	331	Total	O	0	0
			331	331		
6	C	366	Total	O	0	0
			366	366		

Continued on next page...

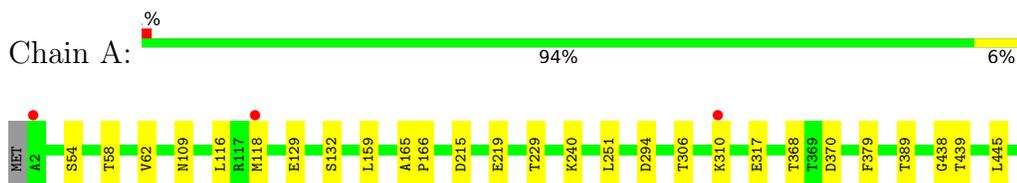
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	382	Total 382	O 382	0	0
6	E	264	Total 264	O 264	0	0
6	F	363	Total 363	O 363	0	0
6	G	371	Total 371	O 371	0	0
6	H	384	Total 384	O 384	0	0

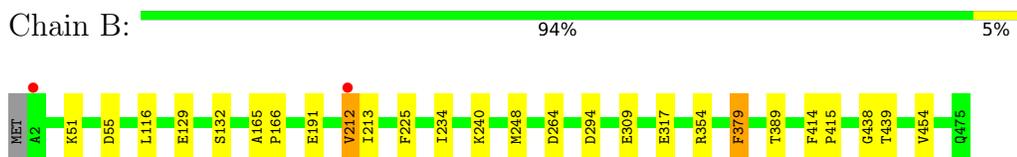
3 Residue-property plots i

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.

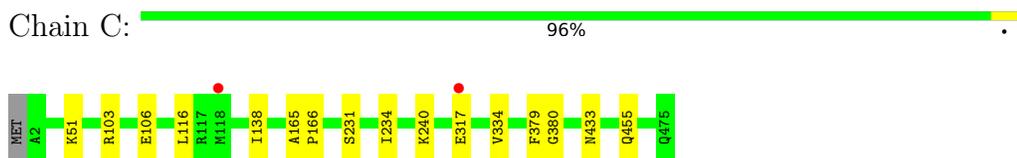
- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



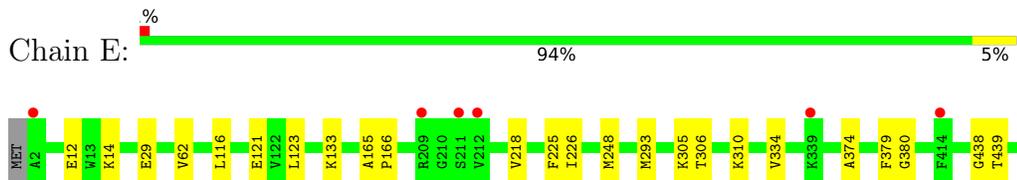
- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



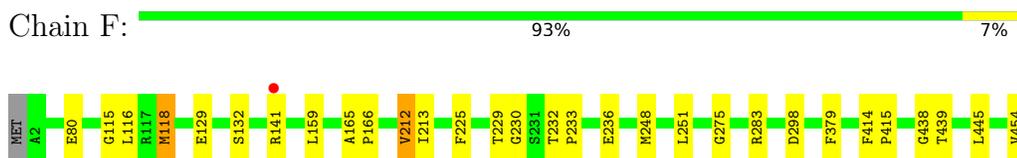
- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase





- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase

Chain G: 94% 6%



- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase

Chain H: 96%



- Molecule 2: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase

Chain D: 95% 5%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	97.52Å 98.48Å 102.98Å 77.77° 76.08° 67.79°	Depositor
Resolution (Å)	41.15 – 1.99 41.15 – 1.99	Depositor EDS
% Data completeness (in resolution range)	79.3 (41.15-1.99) 79.5 (41.15-1.99)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.38 (at 1.98Å)	Xtrriage
Refinement program	PHENIX 1.21.2_5419	Depositor
R, R_{free}	0.166 , 0.211 0.166 , 0.211	Depositor DCC
R_{free} test set	11310 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	20.7	Xtrriage
Anisotropy	0.010	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 41.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.000 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	61216	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.05% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: E4P, GOL, 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	0.13	0/3638	0.30	0/4933
1	B	0.11	0/3692	0.28	0/5006
1	C	0.14	0/3656	0.31	0/4959
1	E	0.12	0/3641	0.30	0/4936
1	F	0.19	1/3661 (0.0%)	0.29	0/4964
1	G	0.14	0/3656	0.30	0/4957
1	H	0.17	1/3632 (0.0%)	0.31	0/4925
2	D	0.14	0/3732	0.30	0/5056
All	All	0.14	2/29308 (0.0%)	0.30	0/39736

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	C	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	118	MET	CB-CG	-8.13	1.28	1.52
1	H	303	GLU	CD-OE2	-5.95	1.14	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	455	GLN	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	3561	3634	3619	18	0
1	B	3601	3680	3679	22	0
1	C	3573	3652	3639	13	0
1	E	3569	3647	3646	18	0
1	F	3581	3647	3646	25	0
1	G	3575	3650	3649	21	0
1	H	3560	3631	3630	13	0
2	D	3640	3724	3716	16	0
3	A	48	25	25	1	0
3	B	48	25	25	2	0
3	E	48	25	25	0	0
3	F	48	25	25	0	0
4	A	12	0	7	1	0
5	A	6	8	8	0	0
5	B	12	16	16	1	0
5	C	24	32	32	1	0
5	D	6	8	8	0	0
5	E	6	8	8	0	0
5	F	12	16	16	0	0
5	G	18	24	24	0	0
5	H	18	24	24	0	0
6	A	288	0	0	3	1
6	B	331	0	0	7	0
6	C	366	0	0	7	2
6	D	382	0	0	6	0
6	E	264	0	0	6	0
6	F	363	0	0	4	1
6	G	371	0	0	11	1
6	H	384	0	0	7	1
All	All	31715	29501	29467	143	3

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 2.

The worst 5 of 143 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:G:317:GLU:OE1	6:G:601:HOH:O	1.63	1.14
1:C:240:LYS:NZ	6:C:601:HOH:O	1.83	1.09
1:F:141:ARG:HH22	1:F:462:GLU:HA	1.32	0.94
1:F:118:MET:SD	1:F:141:ARG:NE	2.41	0.93
1:A:317:GLU:OE1	6:A:601:HOH:O	1.89	0.90

All (3) 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)
6:C:645:HOH:O	6:G:768:HOH:O[1_466]	1.89	0.31
6:C:602:HOH:O	6:F:814:HOH:O[1_466]	2.04	0.16
6:A:847:HOH:O	6:H:947:HOH:O[1_466]	2.17	0.03

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	477/475 (100%)	465 (98%)	12 (2%)	0	100	100
1	B	483/475 (102%)	472 (98%)	11 (2%)	0	100	100
1	C	480/475 (101%)	468 (98%)	12 (2%)	0	100	100
1	E	477/475 (100%)	464 (97%)	13 (3%)	0	100	100
1	F	479/475 (101%)	468 (98%)	11 (2%)	0	100	100
1	G	480/475 (101%)	468 (98%)	12 (2%)	0	100	100
1	H	477/475 (100%)	463 (97%)	14 (3%)	0	100	100
2	D	487/480 (102%)	474 (97%)	13 (3%)	0	100	100
All	All	3840/3805 (101%)	3742 (97%)	98 (3%)	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	374/370 (101%)	372 (100%)	2 (0%)	86	90
1	B	380/370 (103%)	377 (99%)	3 (1%)	79	84
1	C	377/370 (102%)	375 (100%)	2 (0%)	86	90
1	E	374/370 (101%)	369 (99%)	5 (1%)	65	71
1	F	376/370 (102%)	374 (100%)	2 (0%)	86	90
1	G	377/370 (102%)	375 (100%)	2 (0%)	86	90
1	H	374/370 (101%)	372 (100%)	2 (0%)	86	90
2	D	385/374 (103%)	380 (99%)	5 (1%)	65	71
All	All	3017/2964 (102%)	2994 (99%)	23 (1%)	81	84

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	293[A]	MET
1	F	212	VAL
1	E	379	PHE
1	F	379	PHE
1	C	433	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	423	GLN
1	G	7	ASN
1	H	433	ASN
1	H	353	ASN
1	H	423	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

22 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
5	GOL	C	504	-	5,5,5	0.33	0	5,5,5	0.28	0
3	NAP	E	501	-	45,52,52	1.58	3 (6%)	56,80,80	1.00	2 (3%)
3	NAP	B	501	-	45,52,52	1.60	3 (6%)	56,80,80	0.87	1 (1%)
4	E4P	A	502	-	9,11,11	1.09	0	12,15,15	0.66	0
5	GOL	H	502	-	5,5,5	0.34	0	5,5,5	0.39	0
5	GOL	G	503	-	5,5,5	0.32	0	5,5,5	0.38	0
3	NAP	F	502	-	45,52,52	1.57	3 (6%)	56,80,80	0.94	1 (1%)
5	GOL	B	503	-	5,5,5	0.28	0	5,5,5	0.27	0
5	GOL	C	503	-	5,5,5	0.33	0	5,5,5	0.32	0
5	GOL	H	501	-	5,5,5	0.37	0	5,5,5	0.44	0
3	NAP	A	501	-	45,52,52	1.60	3 (6%)	56,80,80	0.93	1 (1%)
5	GOL	F	501	-	5,5,5	0.36	0	5,5,5	0.35	0
5	GOL	E	502	-	5,5,5	0.32	0	5,5,5	0.36	0
5	GOL	C	502	-	5,5,5	0.30	0	5,5,5	0.41	0
5	GOL	G	502	-	5,5,5	0.37	0	5,5,5	0.40	0
5	GOL	H	503	-	5,5,5	0.36	0	5,5,5	0.34	0
5	GOL	A	503	-	5,5,5	0.30	0	5,5,5	0.37	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	GOL	C	501	-	5,5,5	0.36	0	5,5,5	0.42	0
5	GOL	F	503	-	5,5,5	0.38	0	5,5,5	0.50	0
5	GOL	G	501	-	5,5,5	0.44	0	5,5,5	0.53	0
5	GOL	B	502	-	5,5,5	0.39	0	5,5,5	0.42	0
5	GOL	D	501	-	5,5,5	0.35	0	5,5,5	0.40	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
5	GOL	C	504	-	-	0/4/4/4	-
3	NAP	E	501	-	-	1/31/67/67	0/5/5/5
3	NAP	B	501	-	-	2/31/67/67	0/5/5/5
4	E4P	A	502	-	-	4/10/12/12	-
5	GOL	H	502	-	-	2/4/4/4	-
5	GOL	G	503	-	-	2/4/4/4	-
3	NAP	F	502	-	-	1/31/67/67	0/5/5/5
5	GOL	B	503	-	-	4/4/4/4	-
5	GOL	C	503	-	-	2/4/4/4	-
5	GOL	H	501	-	-	0/4/4/4	-
3	NAP	A	501	-	-	1/31/67/67	0/5/5/5
5	GOL	F	501	-	-	2/4/4/4	-
5	GOL	E	502	-	-	2/4/4/4	-
5	GOL	C	502	-	-	0/4/4/4	-
5	GOL	G	502	-	-	3/4/4/4	-
5	GOL	H	503	-	-	2/4/4/4	-
5	GOL	A	503	-	-	2/4/4/4	-
5	GOL	C	501	-	-	4/4/4/4	-
5	GOL	F	503	-	-	2/4/4/4	-
5	GOL	G	501	-	-	1/4/4/4	-
5	GOL	B	502	-	-	2/4/4/4	-
5	GOL	D	501	-	-	4/4/4/4	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	NAP	P2B-O2B	8.51	1.75	1.59

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	501	NAP	P2B-O2B	8.46	1.75	1.59
3	E	501	NAP	P2B-O2B	8.36	1.75	1.59
3	F	502	NAP	P2B-O2B	8.13	1.74	1.59
3	F	502	NAP	O4D-C1D	-3.24	1.36	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	501	NAP	C3B-C2B-C1B	-2.35	98.48	102.89
3	E	501	NAP	C3D-C2D-C1D	2.25	104.36	100.98
3	F	502	NAP	O3X-P2B-O2X	2.08	115.60	107.64
3	E	501	NAP	O3X-P2B-O2X	2.04	115.44	107.64
3	B	501	NAP	O3X-P2B-O2X	2.02	115.36	107.64

There are no chirality outliers.

5 of 43 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	502	E4P	C4-O4-P-O2P
4	A	502	E4P	C4-O4-P-O3P
5	B	502	GOL	C1-C2-C3-O3
5	B	502	GOL	O2-C2-C3-O3
5	B	503	GOL	O1-C1-C2-C3

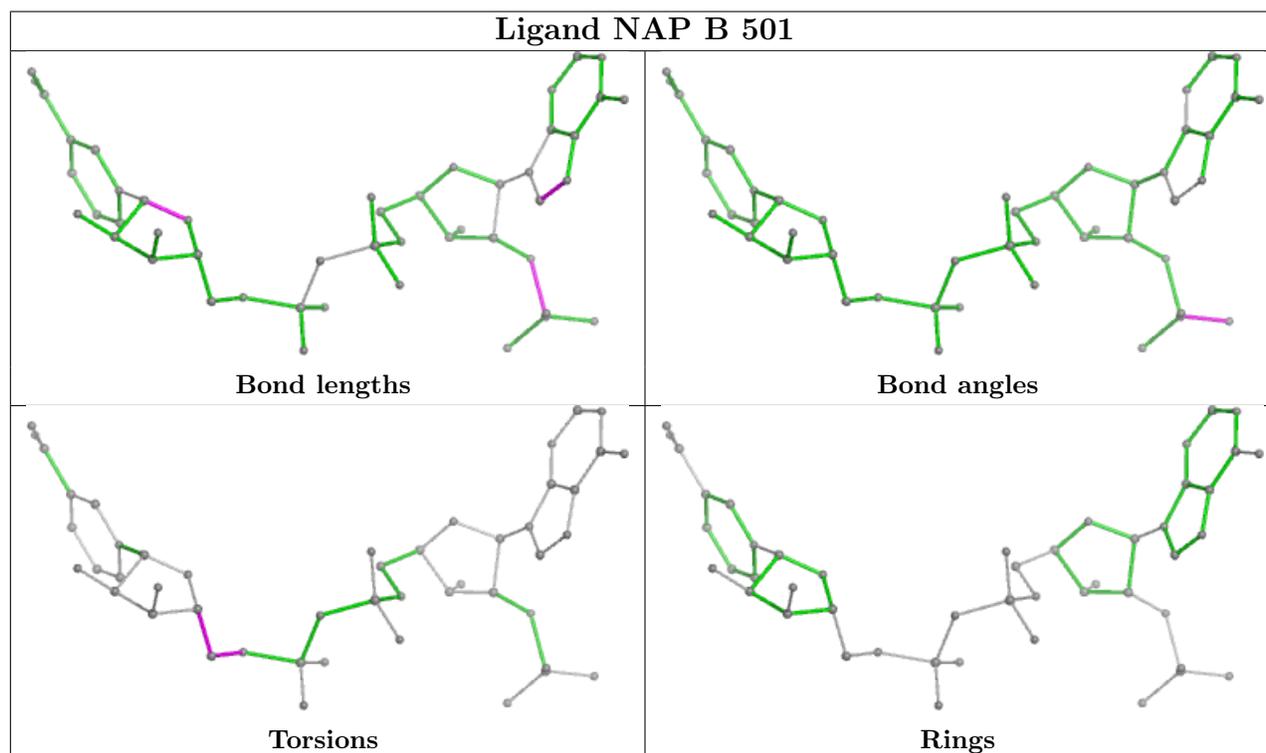
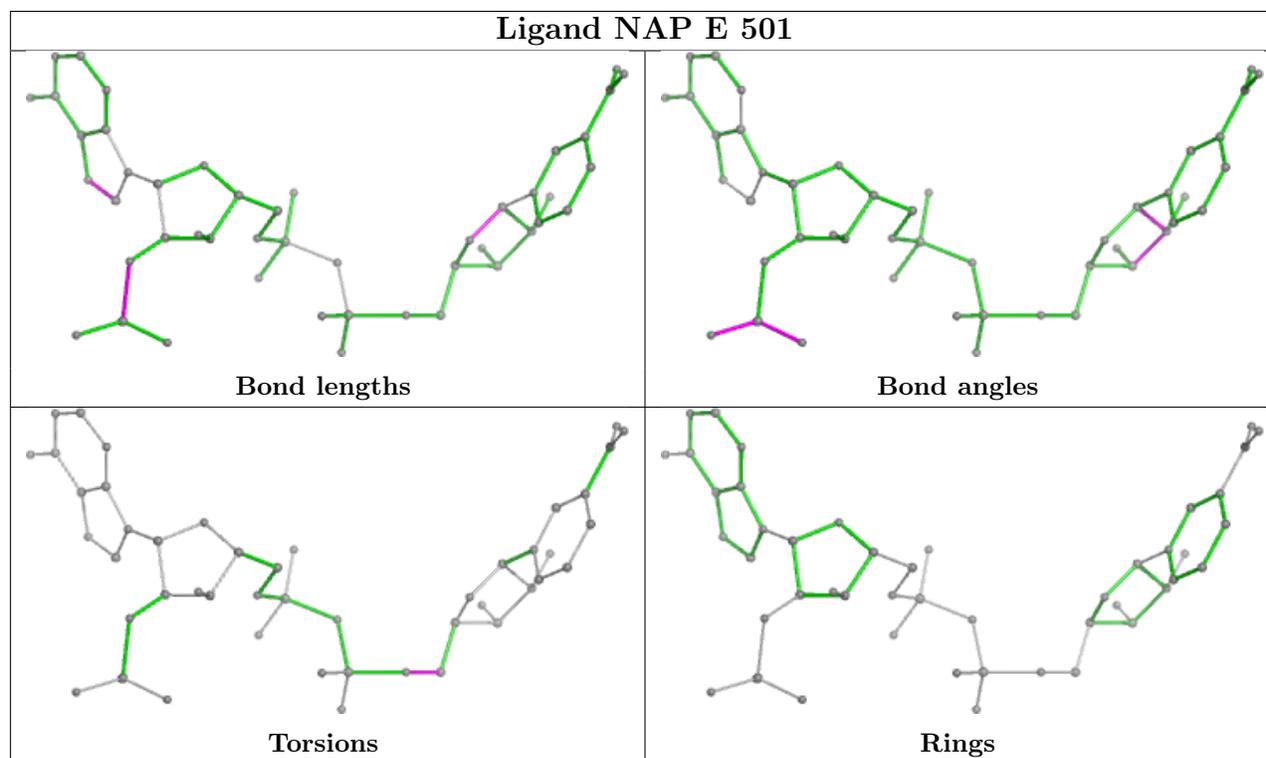
There are no ring outliers.

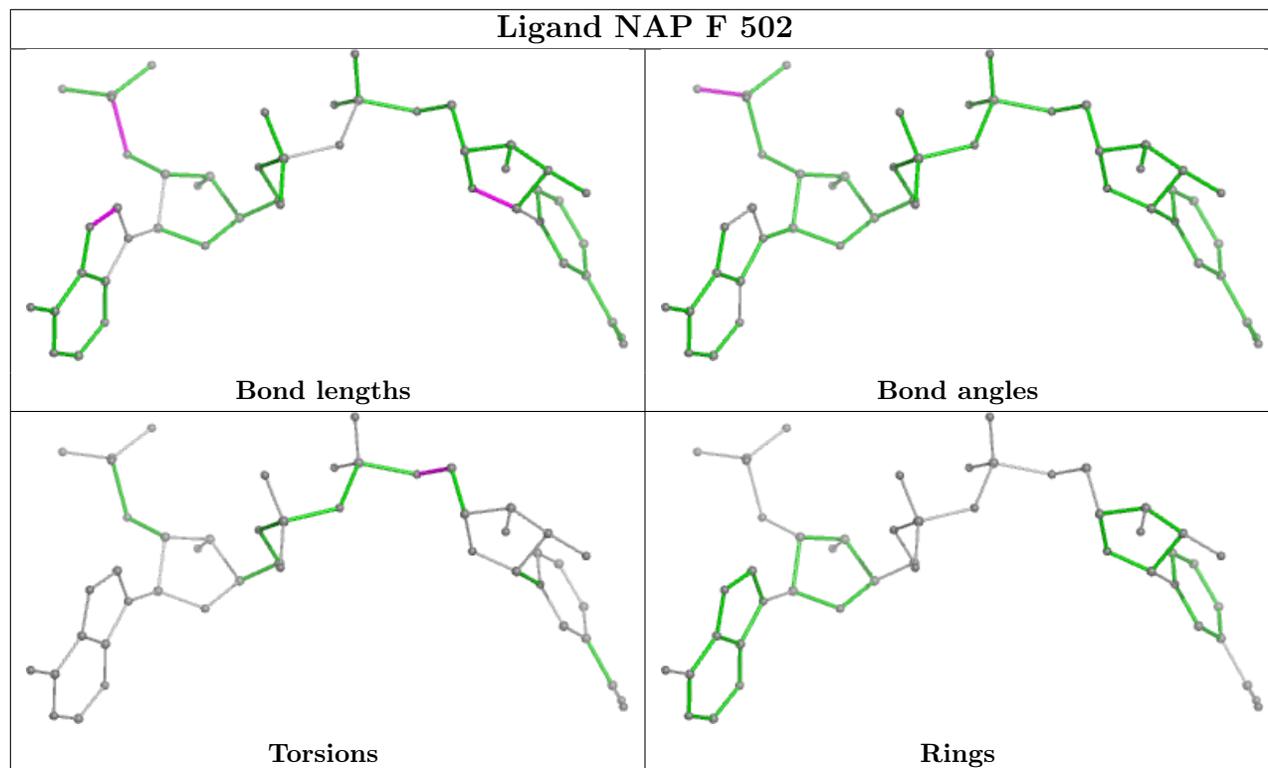
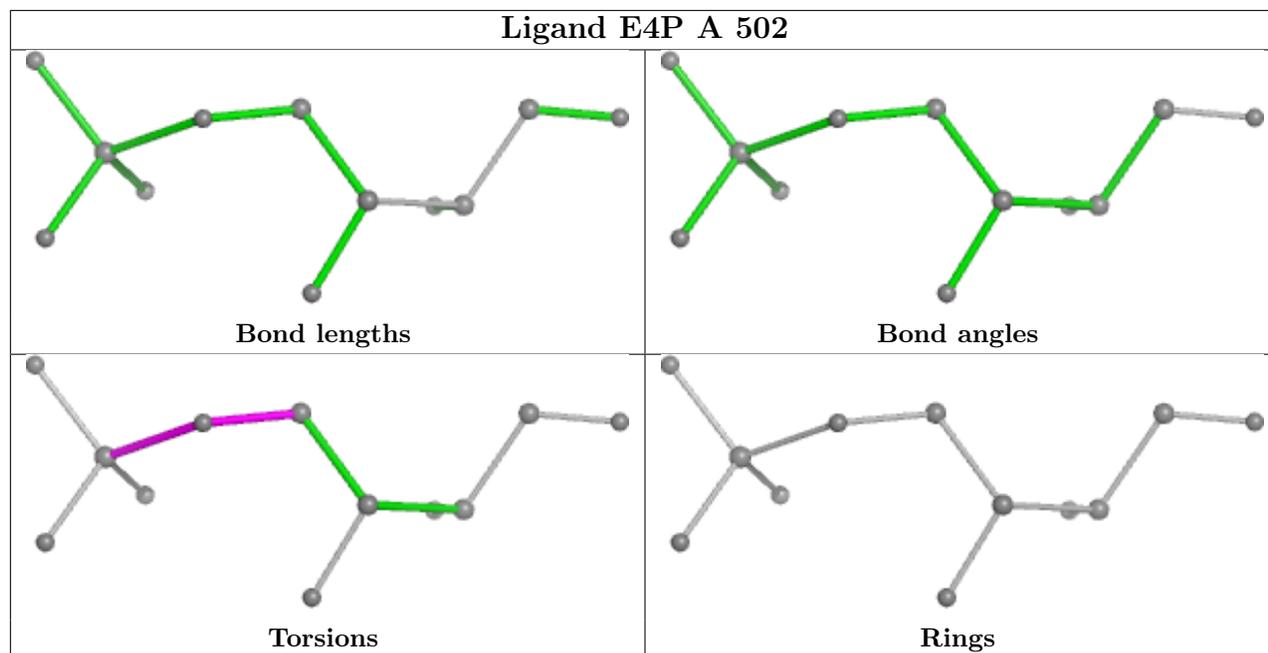
5 monomers are involved in 6 short contacts:

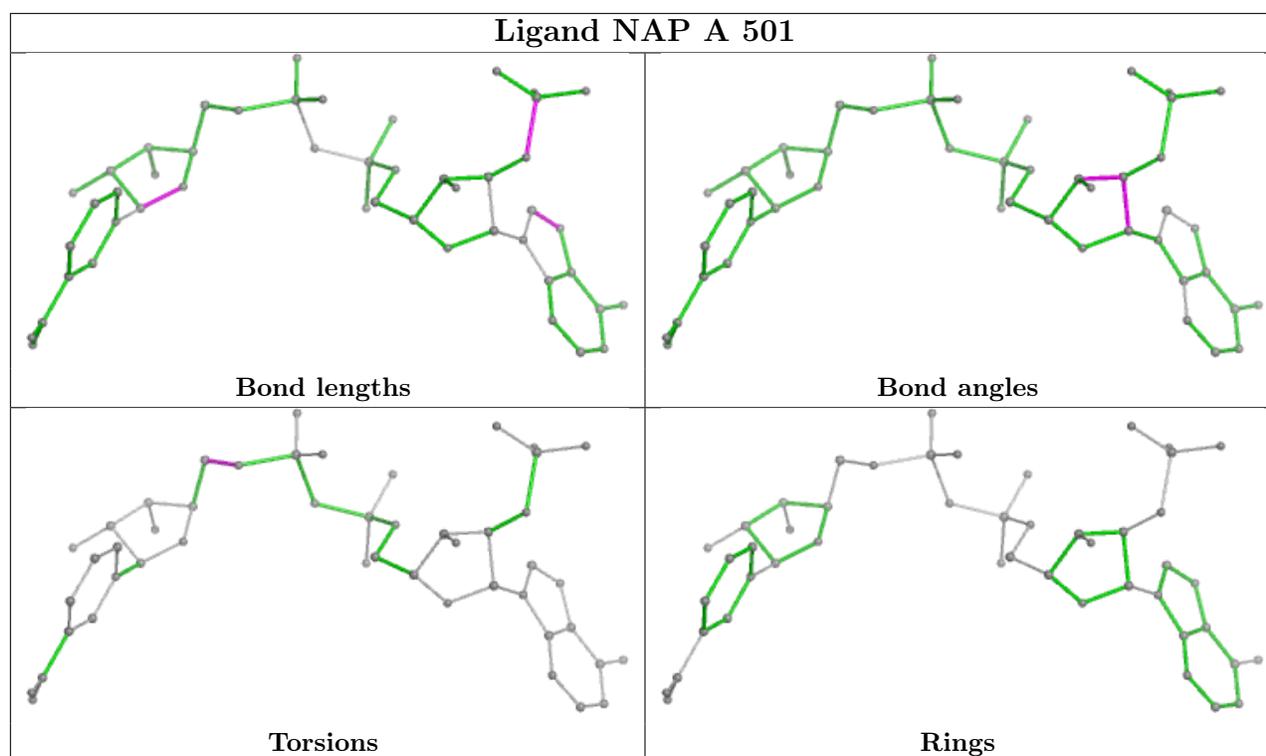
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	501	NAP	2	0
4	A	502	E4P	1	0
5	B	503	GOL	1	0
3	A	501	NAP	1	0
5	C	501	GOL	1	0

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







5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	-5:ASP	C	2:ALA	N	85.72

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	474/475 (99%)	-0.29	3 (0%) 85 85	12, 26, 44, 78	4 (0%)
1	B	474/475 (99%)	-0.50	2 (0%) 89 88	9, 21, 34, 68	11 (2%)
1	C	474/475 (99%)	-0.50	2 (0%) 89 88	7, 20, 32, 65	8 (1%)
1	E	474/475 (99%)	-0.15	6 (1%) 74 73	15, 30, 50, 83	7 (1%)
1	F	474/475 (99%)	-0.40	1 (0%) 92 91	9, 23, 40, 72	7 (1%)
1	G	474/475 (99%)	-0.55	0 100 100	9, 20, 29, 48	9 (1%)
1	H	474/475 (99%)	-0.51	0 100 100	9, 21, 33, 58	5 (1%)
2	D	480/480 (100%)	-0.47	2 (0%) 89 88	11, 21, 34, 72	11 (2%)
All	All	3798/3805 (99%)	-0.42	16 (0%) 89 88	7, 22, 40, 83	62 (1%)

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	ALA	4.6
1	F	141	ARG	4.4
1	C	118	MET	3.9
1	E	2	ALA	3.6
1	A	310	LYS	3.5

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](#)

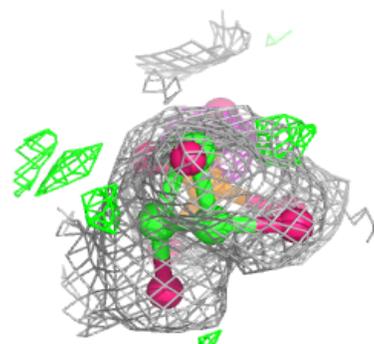
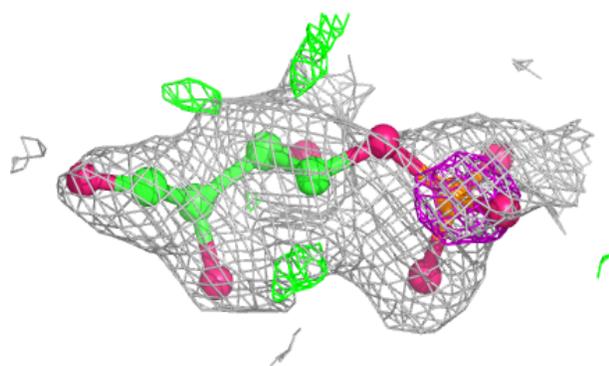
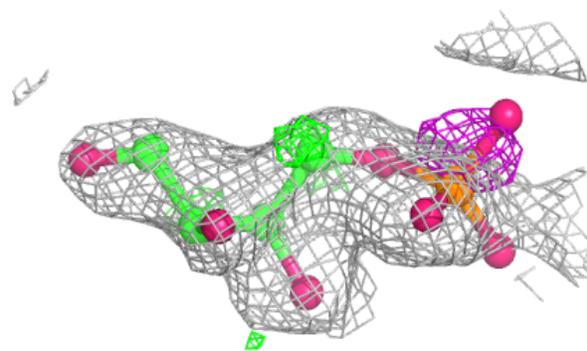
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	GOL	C	502	6/6	0.74	0.14	30,39,47,47	0
5	GOL	G	503	6/6	0.74	0.12	37,44,50,50	0
5	GOL	G	502	6/6	0.75	0.15	29,41,49,53	0
5	GOL	H	502	6/6	0.75	0.16	33,41,51,51	0
5	GOL	C	503	6/6	0.80	0.18	35,42,47,49	0
5	GOL	E	502	6/6	0.81	0.11	43,52,65,65	0
4	E4P	A	502	12/12	0.81	0.15	24,31,46,47	12
3	NAP	E	501	48/48	0.83	0.14	24,36,43,48	73
5	GOL	B	502	6/6	0.84	0.14	24,34,41,47	0
5	GOL	H	503	6/6	0.84	0.11	32,38,43,48	0
5	GOL	B	503	6/6	0.85	0.09	25,32,45,49	0
5	GOL	C	501	6/6	0.85	0.17	26,34,41,44	0
5	GOL	A	503	6/6	0.85	0.11	32,41,50,51	0
5	GOL	C	504	6/6	0.86	0.13	26,33,39,39	0
5	GOL	F	501	6/6	0.86	0.17	30,37,42,46	0
5	GOL	D	501	6/6	0.86	0.19	25,33,37,44	0
5	GOL	F	503	6/6	0.87	0.09	32,40,48,50	0
3	NAP	B	501	48/48	0.88	0.11	17,27,33,39	73
5	GOL	H	501	6/6	0.88	0.13	40,49,61,63	0
5	GOL	G	501	6/6	0.89	0.14	21,28,31,37	0
3	NAP	F	502	48/48	0.90	0.10	22,28,35,38	73
3	NAP	A	501	48/48	0.93	0.09	15,26,33,34	73

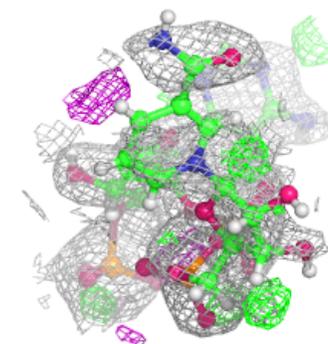
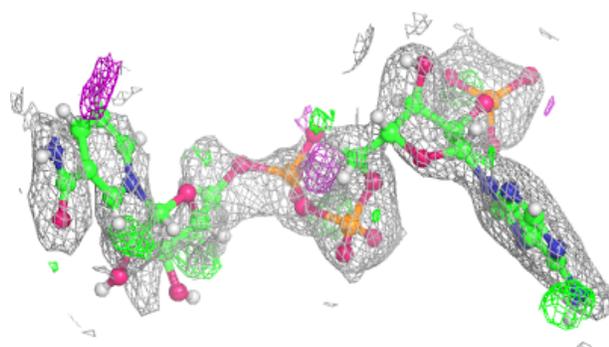
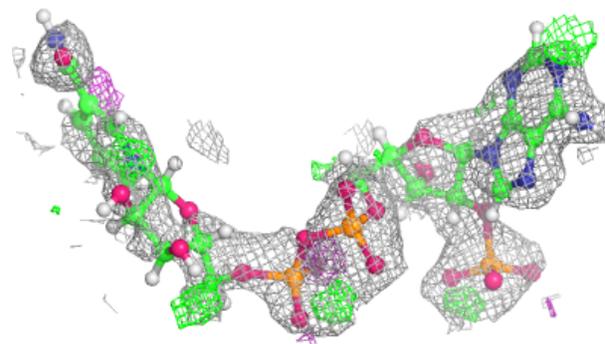
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around E4P A 502:

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

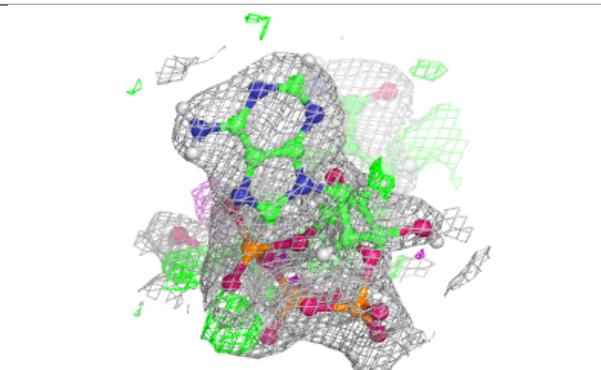
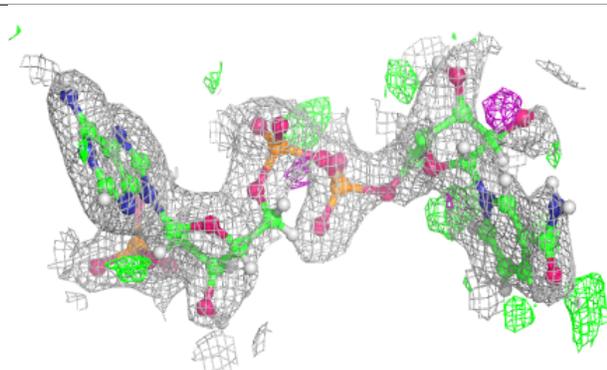
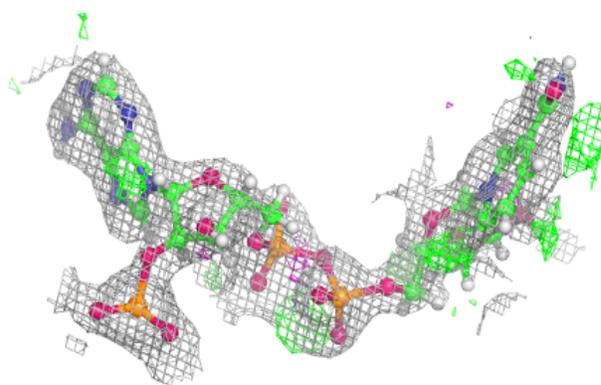
**Electron density around NAP E 501:**

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

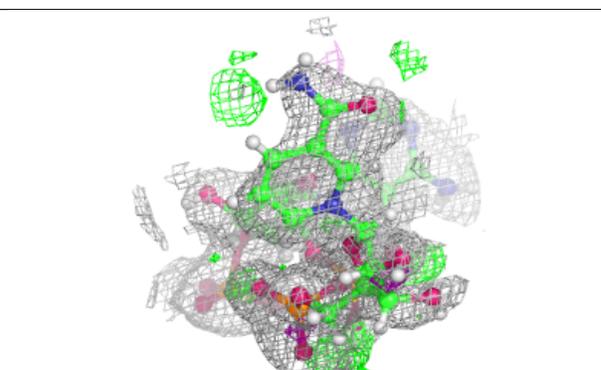
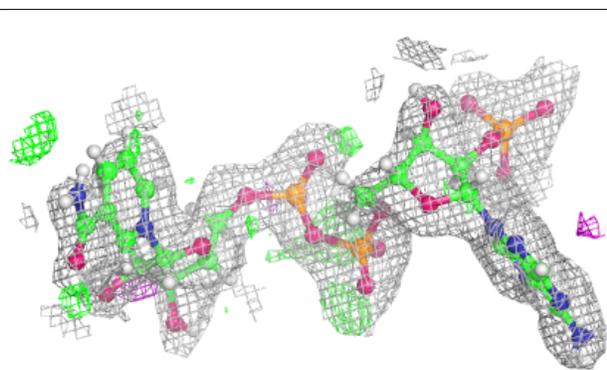
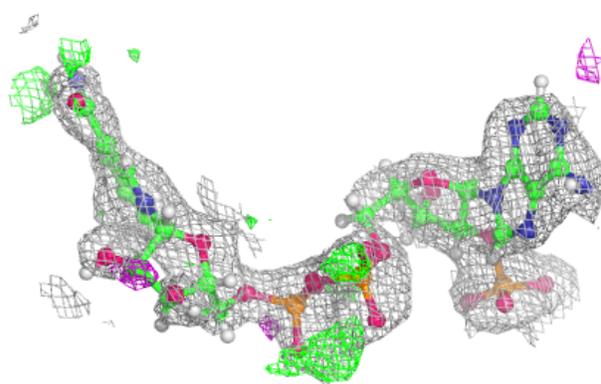


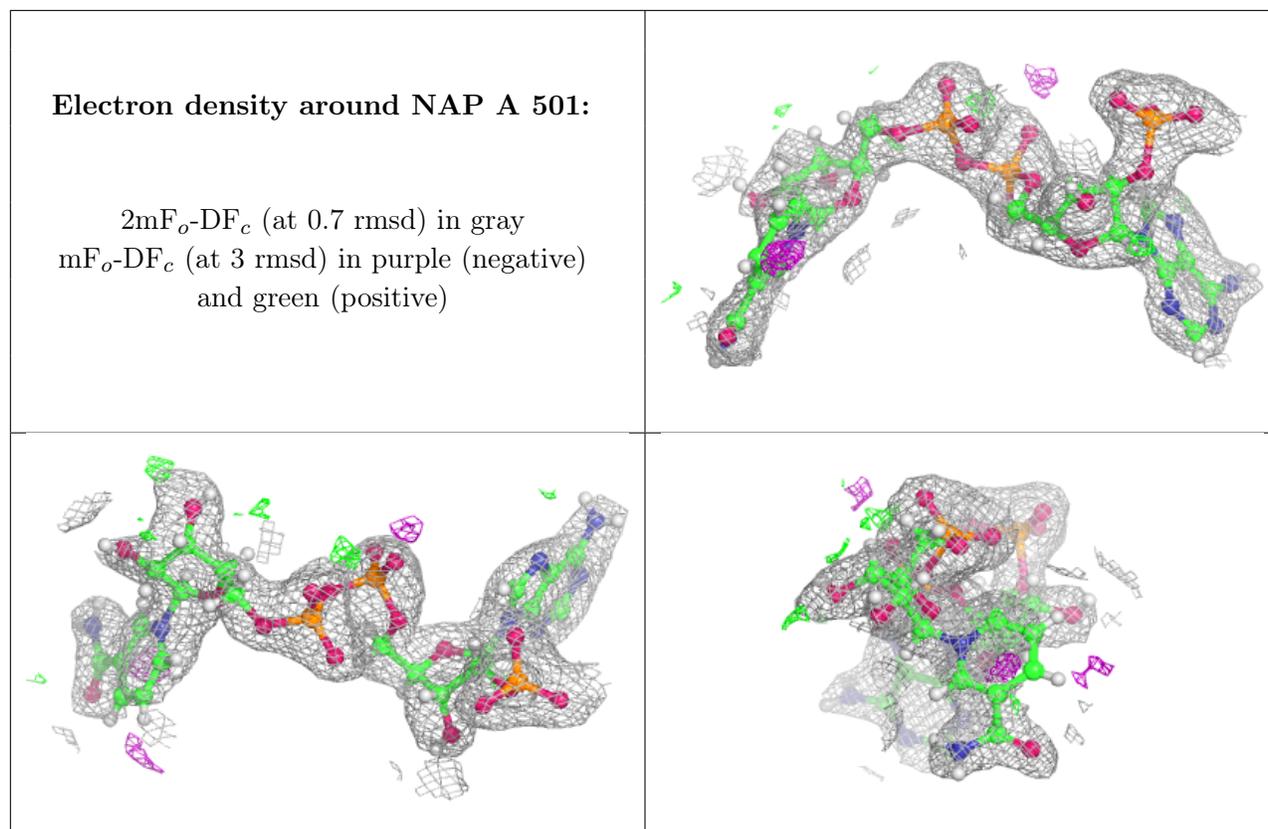
Electron density around NAP B 501:

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

**Electron density around NAP F 502:**

$2mF_o-DF_c$ (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.