



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 10:13 AM UTC

PDB ID : 9RAZ / pdb\_00009raz  
Title : Streptococcus pyogenes GapN in complex with NADP and glyceraldehyde-3-phosphate  
Authors : Wirsing, R.; Schindelin, H.  
Deposited on : 2025-05-21  
Resolution : 1.38 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](mailto: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 Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Buster-report : wwPDB partial adaption of 1.1.7 (2018)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

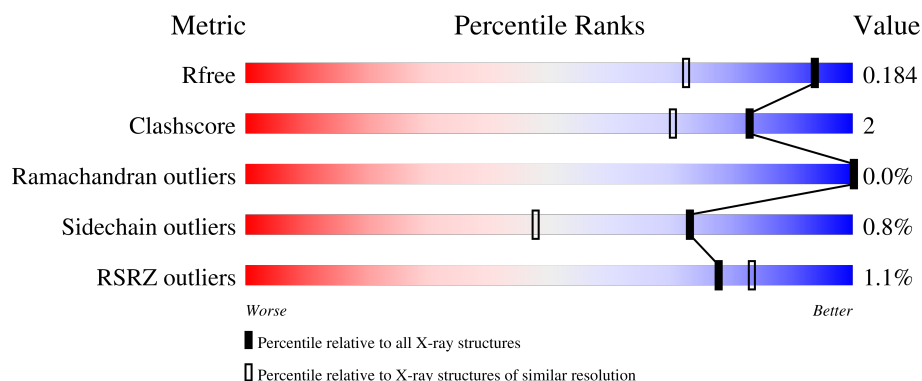
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.38 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.




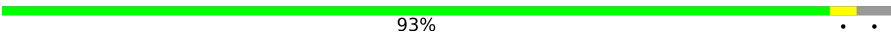

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	180053	4403 (1.40-1.36)
Clashscore	190562	4528 (1.40-1.36)
Ramachandran outliers	187476	4459 (1.40-1.36)
Sidechain outliers	187428	4458 (1.40-1.36)
RSRZ outliers	180081	4399 (1.40-1.36)

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  $\geq 3$ , 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	496	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>5%</div> <div>.</div> </div> </div>
1	B	496	<div> <div>2%</div> <div> <div></div> <div>90%</div> <div>5%</div> <div>.</div> </div> </div>
1	C	496	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div>.</div> </div> </div>
1	D	496	<div> <div>%</div> <div> <div></div> <div>91%</div> <div>5%</div> <div>.</div> </div> </div>
1	E	496	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div>.</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	496	 90% 6% .
1	G	496	 93% . .
1	H	496	 % 92% . .

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	G3H	F	501[B]	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 64525 atoms, of which 30616 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
1	A	474	Total	C	H	N	O	S	0	23	0
			7396	2332	3739	600	714	11			
1	B	475	Total	C	H	N	O	S	0	37	0
			7587	2384	3838	625	731	9			
1	C	475	Total	C	H	N	O	S	0	40	0
			7581	2377	3863	614	717	10			
1	D	476	Total	C	H	N	O	S	0	38	0
			7589	2393	3838	617	732	9			
1	E	474	Total	C	H	N	O	S	0	21	0
			7367	2321	3730	598	709	9			
1	F	475	Total	C	H	N	O	S	0	37	0
			7551	2368	3838	613	721	11			
1	G	475	Total	C	H	N	O	S	0	38	0
			7520	2367	3803	615	725	10			
1	H	476	Total	C	H	N	O	S	0	34	0
			7546	2376	3833	610	717	10			

There are 192 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	ALA	-	expression tag	UNP A0A4U9C786
A	-19	SER	-	expression tag	UNP A0A4U9C786
A	-18	TRP	-	expression tag	UNP A0A4U9C786
A	-17	SER	-	expression tag	UNP A0A4U9C786
A	-16	HIS	-	expression tag	UNP A0A4U9C786
A	-15	PRO	-	expression tag	UNP A0A4U9C786
A	-14	GLN	-	expression tag	UNP A0A4U9C786
A	-13	PHE	-	expression tag	UNP A0A4U9C786
A	-12	GLU	-	expression tag	UNP A0A4U9C786
A	-11	LYS	-	expression tag	UNP A0A4U9C786
A	-10	ILE	-	expression tag	UNP A0A4U9C786
A	-9	GLU	-	expression tag	UNP A0A4U9C786
A	-8	GLY	-	expression tag	UNP A0A4U9C786

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	ARG	-	expression tag	UNP A0A4U9C786
A	-6	ARG	-	expression tag	UNP A0A4U9C786
A	-5	ASP	-	expression tag	UNP A0A4U9C786
A	-4	ARG	-	expression tag	UNP A0A4U9C786
A	-3	GLY	-	expression tag	UNP A0A4U9C786
A	-2	PRO	-	expression tag	UNP A0A4U9C786
A	-1	GLU	-	expression tag	UNP A0A4U9C786
A	0	PHE	-	expression tag	UNP A0A4U9C786
A	1	LEU	MET	conflict	UNP A0A4U9C786
A	58	THR	ALA	conflict	UNP A0A4U9C786
A	284	SER	CYS	conflict	UNP A0A4U9C786
B	-20	ALA	-	expression tag	UNP A0A4U9C786
B	-19	SER	-	expression tag	UNP A0A4U9C786
B	-18	TRP	-	expression tag	UNP A0A4U9C786
B	-17	SER	-	expression tag	UNP A0A4U9C786
B	-16	HIS	-	expression tag	UNP A0A4U9C786
B	-15	PRO	-	expression tag	UNP A0A4U9C786
B	-14	GLN	-	expression tag	UNP A0A4U9C786
B	-13	PHE	-	expression tag	UNP A0A4U9C786
B	-12	GLU	-	expression tag	UNP A0A4U9C786
B	-11	LYS	-	expression tag	UNP A0A4U9C786
B	-10	ILE	-	expression tag	UNP A0A4U9C786
B	-9	GLU	-	expression tag	UNP A0A4U9C786
B	-8	GLY	-	expression tag	UNP A0A4U9C786
B	-7	ARG	-	expression tag	UNP A0A4U9C786
B	-6	ARG	-	expression tag	UNP A0A4U9C786
B	-5	ASP	-	expression tag	UNP A0A4U9C786
B	-4	ARG	-	expression tag	UNP A0A4U9C786
B	-3	GLY	-	expression tag	UNP A0A4U9C786
B	-2	PRO	-	expression tag	UNP A0A4U9C786
B	-1	GLU	-	expression tag	UNP A0A4U9C786
B	0	PHE	-	expression tag	UNP A0A4U9C786
B	1	LEU	MET	conflict	UNP A0A4U9C786
B	58	THR	ALA	conflict	UNP A0A4U9C786
B	284	SER	CYS	conflict	UNP A0A4U9C786
C	-20	ALA	-	expression tag	UNP A0A4U9C786
C	-19	SER	-	expression tag	UNP A0A4U9C786
C	-18	TRP	-	expression tag	UNP A0A4U9C786
C	-17	SER	-	expression tag	UNP A0A4U9C786
C	-16	HIS	-	expression tag	UNP A0A4U9C786
C	-15	PRO	-	expression tag	UNP A0A4U9C786
C	-14	GLN	-	expression tag	UNP A0A4U9C786

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-13	PHE	-	expression tag	UNP A0A4U9C786
C	-12	GLU	-	expression tag	UNP A0A4U9C786
C	-11	LYS	-	expression tag	UNP A0A4U9C786
C	-10	ILE	-	expression tag	UNP A0A4U9C786
C	-9	GLU	-	expression tag	UNP A0A4U9C786
C	-8	GLY	-	expression tag	UNP A0A4U9C786
C	-7	ARG	-	expression tag	UNP A0A4U9C786
C	-6	ARG	-	expression tag	UNP A0A4U9C786
C	-5	ASP	-	expression tag	UNP A0A4U9C786
C	-4	ARG	-	expression tag	UNP A0A4U9C786
C	-3	GLY	-	expression tag	UNP A0A4U9C786
C	-2	PRO	-	expression tag	UNP A0A4U9C786
C	-1	GLU	-	expression tag	UNP A0A4U9C786
C	0	PHE	-	expression tag	UNP A0A4U9C786
C	1	LEU	MET	conflict	UNP A0A4U9C786
C	58	THR	ALA	conflict	UNP A0A4U9C786
C	284	SER	CYS	conflict	UNP A0A4U9C786
D	-20	ALA	-	expression tag	UNP A0A4U9C786
D	-19	SER	-	expression tag	UNP A0A4U9C786
D	-18	TRP	-	expression tag	UNP A0A4U9C786
D	-17	SER	-	expression tag	UNP A0A4U9C786
D	-16	HIS	-	expression tag	UNP A0A4U9C786
D	-15	PRO	-	expression tag	UNP A0A4U9C786
D	-14	GLN	-	expression tag	UNP A0A4U9C786
D	-13	PHE	-	expression tag	UNP A0A4U9C786
D	-12	GLU	-	expression tag	UNP A0A4U9C786
D	-11	LYS	-	expression tag	UNP A0A4U9C786
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	-4	ARG	-	expression tag	UNP A0A4U9C786
D	-3	GLY	-	expression tag	UNP A0A4U9C786
D	-2	PRO	-	expression tag	UNP A0A4U9C786
D	-1	GLU	-	expression tag	UNP A0A4U9C786
D	0	PHE	-	expression tag	UNP A0A4U9C786
D	1	LEU	MET	conflict	UNP A0A4U9C786
D	58	THR	ALA	conflict	UNP A0A4U9C786
D	284	SER	CYS	conflict	UNP A0A4U9C786
E	-20	ALA	-	expression tag	UNP A0A4U9C786

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-19	SER	-	expression tag	UNP A0A4U9C786
E	-18	TRP	-	expression tag	UNP A0A4U9C786
E	-17	SER	-	expression tag	UNP A0A4U9C786
E	-16	HIS	-	expression tag	UNP A0A4U9C786
E	-15	PRO	-	expression tag	UNP A0A4U9C786
E	-14	GLN	-	expression tag	UNP A0A4U9C786
E	-13	PHE	-	expression tag	UNP A0A4U9C786
E	-12	GLU	-	expression tag	UNP A0A4U9C786
E	-11	LYS	-	expression tag	UNP A0A4U9C786
E	-10	ILE	-	expression tag	UNP A0A4U9C786
E	-9	GLU	-	expression tag	UNP A0A4U9C786
E	-8	GLY	-	expression tag	UNP A0A4U9C786
E	-7	ARG	-	expression tag	UNP A0A4U9C786
E	-6	ARG	-	expression tag	UNP A0A4U9C786
E	-5	ASP	-	expression tag	UNP A0A4U9C786
E	-4	ARG	-	expression tag	UNP A0A4U9C786
E	-3	GLY	-	expression tag	UNP A0A4U9C786
E	-2	PRO	-	expression tag	UNP A0A4U9C786
E	-1	GLU	-	expression tag	UNP A0A4U9C786
E	0	PHE	-	expression tag	UNP A0A4U9C786
E	1	LEU	MET	conflict	UNP A0A4U9C786
E	58	THR	ALA	conflict	UNP A0A4U9C786
E	284	SER	CYS	conflict	UNP A0A4U9C786
F	-20	ALA	-	expression tag	UNP A0A4U9C786
F	-19	SER	-	expression tag	UNP A0A4U9C786
F	-18	TRP	-	expression tag	UNP A0A4U9C786
F	-17	SER	-	expression tag	UNP A0A4U9C786
F	-16	HIS	-	expression tag	UNP A0A4U9C786
F	-15	PRO	-	expression tag	UNP A0A4U9C786
F	-14	GLN	-	expression tag	UNP A0A4U9C786
F	-13	PHE	-	expression tag	UNP A0A4U9C786
F	-12	GLU	-	expression tag	UNP A0A4U9C786
F	-11	LYS	-	expression tag	UNP A0A4U9C786
F	-10	ILE	-	expression tag	UNP A0A4U9C786
F	-9	GLU	-	expression tag	UNP A0A4U9C786
F	-8	GLY	-	expression tag	UNP A0A4U9C786
F	-7	ARG	-	expression tag	UNP A0A4U9C786
F	-6	ARG	-	expression tag	UNP A0A4U9C786
F	-5	ASP	-	expression tag	UNP A0A4U9C786
F	-4	ARG	-	expression tag	UNP A0A4U9C786
F	-3	GLY	-	expression tag	UNP A0A4U9C786
F	-2	PRO	-	expression tag	UNP A0A4U9C786

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	-1	GLU	-	expression tag	UNP A0A4U9C786
F	0	PHE	-	expression tag	UNP A0A4U9C786
F	1	LEU	MET	conflict	UNP A0A4U9C786
F	58	THR	ALA	conflict	UNP A0A4U9C786
F	284	SER	CYS	conflict	UNP A0A4U9C786
G	-20	ALA	-	expression tag	UNP A0A4U9C786
G	-19	SER	-	expression tag	UNP A0A4U9C786
G	-18	TRP	-	expression tag	UNP A0A4U9C786
G	-17	SER	-	expression tag	UNP A0A4U9C786
G	-16	HIS	-	expression tag	UNP A0A4U9C786
G	-15	PRO	-	expression tag	UNP A0A4U9C786
G	-14	GLN	-	expression tag	UNP A0A4U9C786
G	-13	PHE	-	expression tag	UNP A0A4U9C786
G	-12	GLU	-	expression tag	UNP A0A4U9C786
G	-11	LYS	-	expression tag	UNP A0A4U9C786
G	-10	ILE	-	expression tag	UNP A0A4U9C786
G	-9	GLU	-	expression tag	UNP A0A4U9C786
G	-8	GLY	-	expression tag	UNP A0A4U9C786
G	-7	ARG	-	expression tag	UNP A0A4U9C786
G	-6	ARG	-	expression tag	UNP A0A4U9C786
G	-5	ASP	-	expression tag	UNP A0A4U9C786
G	-4	ARG	-	expression tag	UNP A0A4U9C786
G	-3	GLY	-	expression tag	UNP A0A4U9C786
G	-2	PRO	-	expression tag	UNP A0A4U9C786
G	-1	GLU	-	expression tag	UNP A0A4U9C786
G	0	PHE	-	expression tag	UNP A0A4U9C786
G	1	LEU	MET	conflict	UNP A0A4U9C786
G	58	THR	ALA	conflict	UNP A0A4U9C786
G	284	SER	CYS	conflict	UNP A0A4U9C786
H	-20	ALA	-	expression tag	UNP A0A4U9C786
H	-19	SER	-	expression tag	UNP A0A4U9C786
H	-18	TRP	-	expression tag	UNP A0A4U9C786
H	-17	SER	-	expression tag	UNP A0A4U9C786
H	-16	HIS	-	expression tag	UNP A0A4U9C786
H	-15	PRO	-	expression tag	UNP A0A4U9C786
H	-14	GLN	-	expression tag	UNP A0A4U9C786
H	-13	PHE	-	expression tag	UNP A0A4U9C786
H	-12	GLU	-	expression tag	UNP A0A4U9C786
H	-11	LYS	-	expression tag	UNP A0A4U9C786
H	-10	ILE	-	expression tag	UNP A0A4U9C786
H	-9	GLU	-	expression tag	UNP A0A4U9C786
H	-8	GLY	-	expression tag	UNP A0A4U9C786

*Continued on next page...*



Chain	Residue	Modelled	Actual	Comment	Reference
H	-7	ARG	-	expression tag	UNP A0A4U9C786
H	-6	ARG	-	expression tag	UNP A0A4U9C786
H	-5	ASP	-	expression tag	UNP A0A4U9C786
H	-4	ARG	-	expression tag	UNP A0A4U9C786
H	-3	GLY	-	expression tag	UNP A0A4U9C786
H	-2	PRO	-	expression tag	UNP A0A4U9C786
H	-1	GLU	-	expression tag	UNP A0A4U9C786
H	0	PHE	-	expression tag	UNP A0A4U9C786
H	1	LEU	MET	conflict	UNP A0A4U9C786
H	58	THR	ALA	conflict	UNP A0A4U9C786
H	284	SER	CYS	conflict	UNP A0A4U9C786

- # NAP

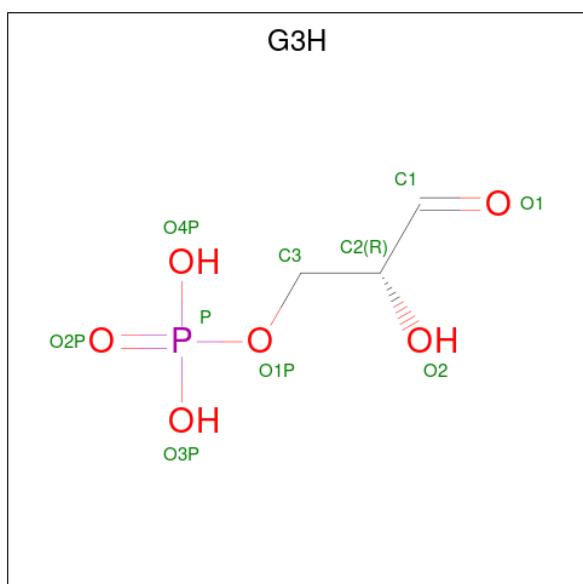
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total 73	C 21	H 25	N 7	O 17	P 3	0	0
2	B	1	Total 73	C 21	H 25	N 7	O 17	P 3	0	0
2	E	1	Total 73	C 21	H 25	N 7	O 17	P 3	0	0
2	F	1	Total 73	C 21	H 25	N 7	O 17	P 3	0	0

- 



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	D	1	Total	C	H	O	0	0
			14	3	8	3		
3	F	1	Total	C	H	O	0	0
			14	3	8	3		
3	F	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is GLYCERALDEHYDE-3-PHOSPHATE (CCD ID: G3H) (formula:  $C_3H_7O_6P$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	F	1	Total	C	H	O	P	0	1
			30	6	10	12	2		

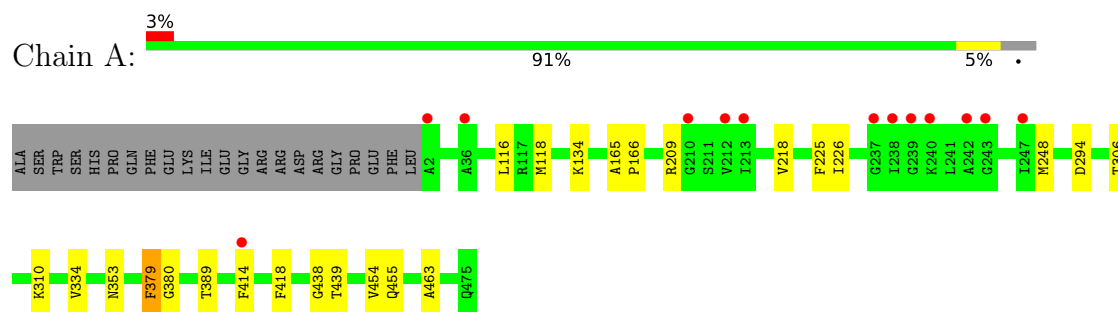
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	357	Total	O	0	3
			360	360		
5	B	382	Total	O	0	9
			391	391		
5	C	606	Total	O	0	4
			610	610		
5	D	552	Total	O	0	4
			556	556		
5	E	419	Total	O	0	3
			422	422		
5	F	588	Total	O	0	5
			593	593		
5	G	591	Total	O	0	13
			604	604		
5	H	485	Total	O	0	3
			488	488		

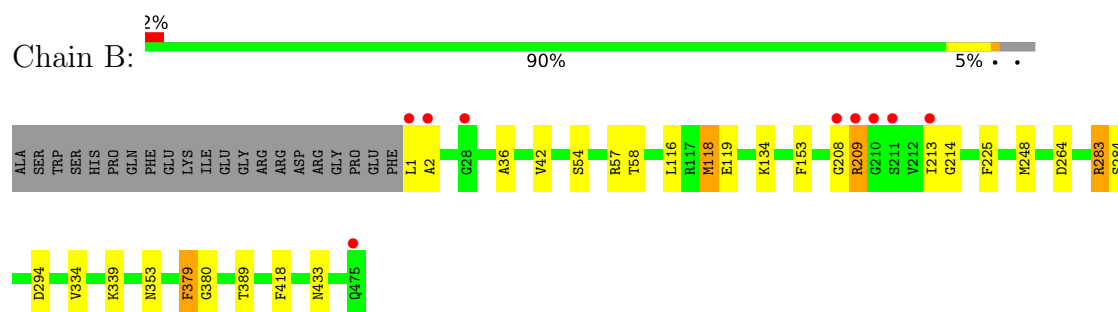
### 3 Residue-property plots [i](#)

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.

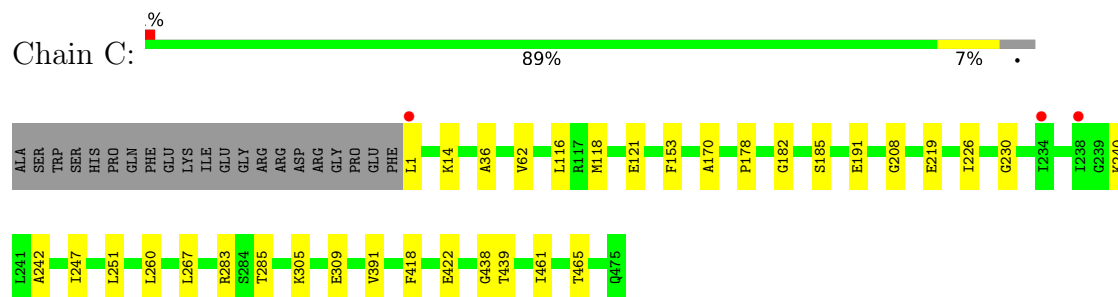
- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



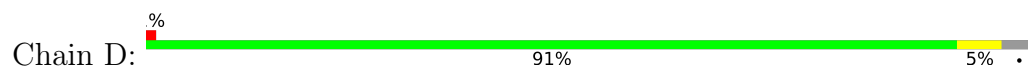
- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



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

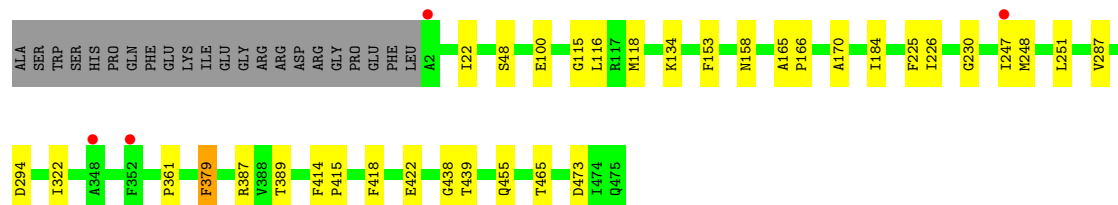
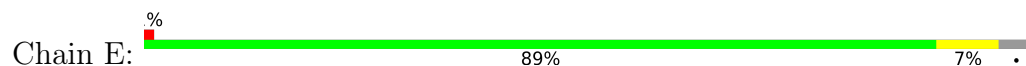


- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase

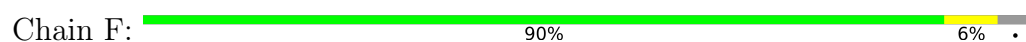




- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



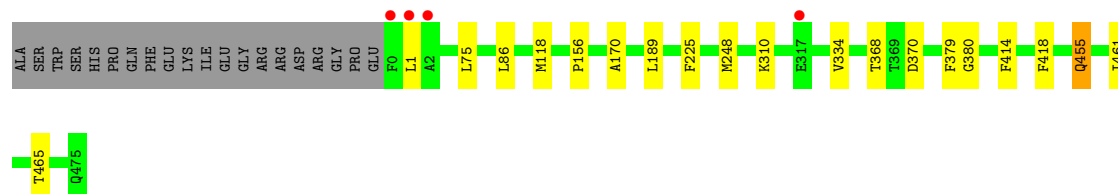
- Molecule 1: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.15Å 97.44Å 100.71Å 77.06° 78.19° 68.92°	Depositor
Resolution (Å)	48.62 – 1.38 48.62 – 1.38	Depositor EDS
% Data completeness (in resolution range)	76.4 (48.62-1.38) 77.2 (48.62-1.38)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.64 (at 1.37Å)	Xtriage
Refinement program	REFMAC 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.154 , 0.179 (Not available) , 0.184	Depositor DCC
$R_{free}$ test set	26722 reflections (3.27%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.9	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 39.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.009 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	64525	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: GOL, G3H, 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	0.25	0/3781	0.37	0/5123
1	B	0.26	0/3885	0.39	0/5261
1	C	0.34	0/3919	0.47	0/5310
1	D	0.31	0/3920	0.43	0/5307
1	E	0.28	0/3763	0.40	0/5100
1	F	0.33	0/3898	0.47	0/5278
1	G	0.34	0/3920	0.47	0/5309
1	H	0.29	0/3883	0.41	0/5260
All	All	0.30	0/30969	0.43	0/41948

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	E	0	1
1	F	0	1
1	H	0	1
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	455	GLN	Peptide
1	B	209[A]	ARG	Sidechain
1	E	455	GLN	Peptide
1	F	455	GLN	Peptide
1	H	455	GLN	Peptide

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3657	3739	3700	17	0
1	B	3749	3838	3789	27	0
1	C	3718	3863	3758	28	0
1	D	3751	3838	3762	20	0
1	E	3637	3730	3701	22	0
1	F	3713	3838	3728	21	0
1	G	3717	3803	3670	12	0
1	H	3713	3833	3757	12	0
2	A	48	25	25	1	0
2	B	48	25	25	2	0
2	E	48	25	25	1	0
2	F	48	25	25	1	0
3	D	6	8	8	0	0
3	F	12	16	16	0	0
4	F	20	10	10	5	0
5	A	360	0	0	3	1
5	B	391	0	0	9	0
5	C	610	0	0	7	1
5	D	556	0	0	1	1
5	E	422	0	0	2	1
5	F	593	0	0	4	0
5	G	604	0	0	3	3
5	H	488	0	0	4	1
All	All	33909	30616	29999	147	4

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 147 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:E:473:ASP:OD2	1:F:433[B]:ASN:ND2	2.15	0.79
1:D:236[A]:GLU:OE2	1:D:449:LYS:NZ	2.14	0.75
1:A:353:ASN:ND2	5:A:601:HOH:O	2.19	0.74
1:D:368:THR:OG1	1:D:370:ASP:OD1	2.08	0.70
1:A:294:ASP:OD2	1:A:389:THR:HG22	1.93	0.69

All (4) 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)
5:C:821:HOH:O	5:G:668:HOH:O[1_546]	2.13	0.07
5:D:1058:HOH:O	5:H:825:HOH:O[1_546]	2.17	0.03
5:A:604:HOH:O	5:G:527:HOH:O[1_645]	2.19	0.01
5:E:874:HOH:O	5:G:1029:HOH:O[1_556]	2.19	0.01

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	494/496 (100%)	481 (97%)	13 (3%)	0	100	100
1	B	509/496 (103%)	493 (97%)	15 (3%)	1 (0%)	43	19
1	C	513/496 (103%)	499 (97%)	14 (3%)	0	100	100
1	D	511/496 (103%)	499 (98%)	12 (2%)	0	100	100
1	E	493/496 (99%)	482 (98%)	11 (2%)	0	100	100
1	F	510/496 (103%)	500 (98%)	10 (2%)	0	100	100
1	G	511/496 (103%)	499 (98%)	12 (2%)	0	100	100
1	H	508/496 (102%)	496 (98%)	12 (2%)	0	100	100
All	All	4049/3968 (102%)	3949 (98%)	99 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	2	ALA

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	390/388 (100%)	387 (99%)	3 (1%)	73	48
1	B	399/388 (103%)	390 (98%)	9 (2%)	44	13
1	C	408/388 (105%)	407 (100%)	1 (0%)	87	73
1	D	405/388 (104%)	403 (100%)	2 (0%)	81	62
1	E	388/388 (100%)	385 (99%)	3 (1%)	73	48
1	F	406/388 (105%)	402 (99%)	4 (1%)	68	40
1	G	405/388 (104%)	403 (100%)	2 (0%)	81	62
1	H	401/388 (103%)	397 (99%)	4 (1%)	68	40
All	All	3202/3104 (103%)	3174 (99%)	28 (1%)	73	43

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	379	PHE
1	H	418	PHE
1	E	379	PHE
1	H	1	LEU
1	E	48[B]	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	423	GLN
1	F	475	GLN
1	H	475	GLN
1	H	39	GLN
1	H	366	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

9 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	$\# Z  > 2$
4	G3H	F	501[A]	-	8,9,9	1.63	1 (12%)	7,12,12	1.60	2 (28%)
4	G3H	F	501[B]	-	8,9,9	1.66	2 (25%)	7,12,12	1.23	0
2	NAP	A	501	-	50,52,52	0.61	1 (2%)	71,80,80	0.64	1 (1%)
2	NAP	E	501	-	50,52,52	0.67	1 (2%)	71,80,80	0.63	1 (1%)
3	GOL	F	502	-	5,5,5	0.73	0	5,5,5	0.99	0
3	GOL	F	503	-	5,5,5	0.87	0	5,5,5	1.15	1 (20%)
2	NAP	B	501	-	50,52,52	0.64	1 (2%)	71,80,80	0.63	1 (1%)
3	GOL	D	501	-	5,5,5	0.79	0	5,5,5	0.99	0
2	NAP	F	504	-	50,52,52	0.65	1 (2%)	71,80,80	0.72	3 (4%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	G3H	F	501[A]	-	-	2/7/8/8	-
4	G3H	F	501[B]	-	-	6/7/8/8	-
2	NAP	A	501	-	-	4/35/67/67	0/5/5/5
2	NAP	E	501	-	-	4/35/67/67	0/5/5/5
3	GOL	F	502	-	-	1/4/4/4	-
3	GOL	F	503	-	-	0/4/4/4	-
2	NAP	B	501	-	-	6/35/67/67	0/5/5/5
3	GOL	D	501	-	-	0/4/4/4	-
2	NAP	F	504	-	-	4/35/67/67	0/5/5/5

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	501[A]	G3H	P-O1P	3.56	1.71	1.60
4	F	501[B]	G3H	P-O1P	3.38	1.71	1.60
2	E	501	NAP	P2B-O2B	2.64	1.64	1.59
2	F	504	NAP	P2B-O2B	2.52	1.63	1.59
2	B	501	NAP	P2B-O2B	2.47	1.63	1.59

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	NAP	P2B-O2B-C2B	-3.59	113.84	123.43
2	E	501	NAP	P2B-O2B-C2B	-3.14	115.06	123.43
2	B	501	NAP	P2B-O2B-C2B	-3.06	115.25	123.43
2	F	504	NAP	P2B-O2B-C2B	-2.94	115.58	123.43
2	F	504	NAP	O2B-P2B-O1X	-2.61	100.02	109.33

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

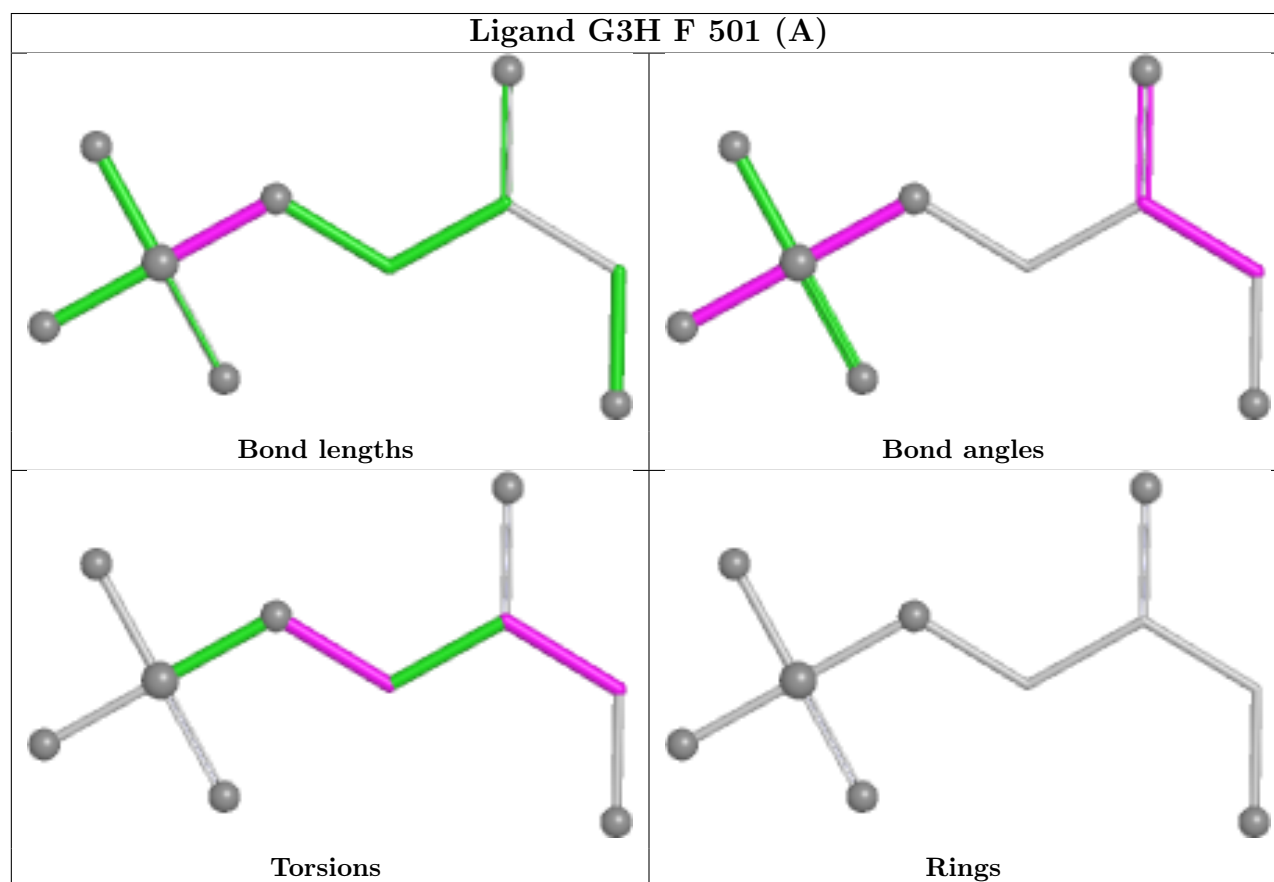
Mol	Chain	Res	Type	Atoms
4	F	501[A]	G3H	O1-C1-C2-C3
4	F	501[B]	G3H	C1-C2-C3-O1P
4	F	501[B]	G3H	O2-C2-C3-O1P
4	F	501[B]	G3H	C2-C3-O1P-P
4	F	501[B]	G3H	C3-O1P-P-O2P

There are no ring outliers.

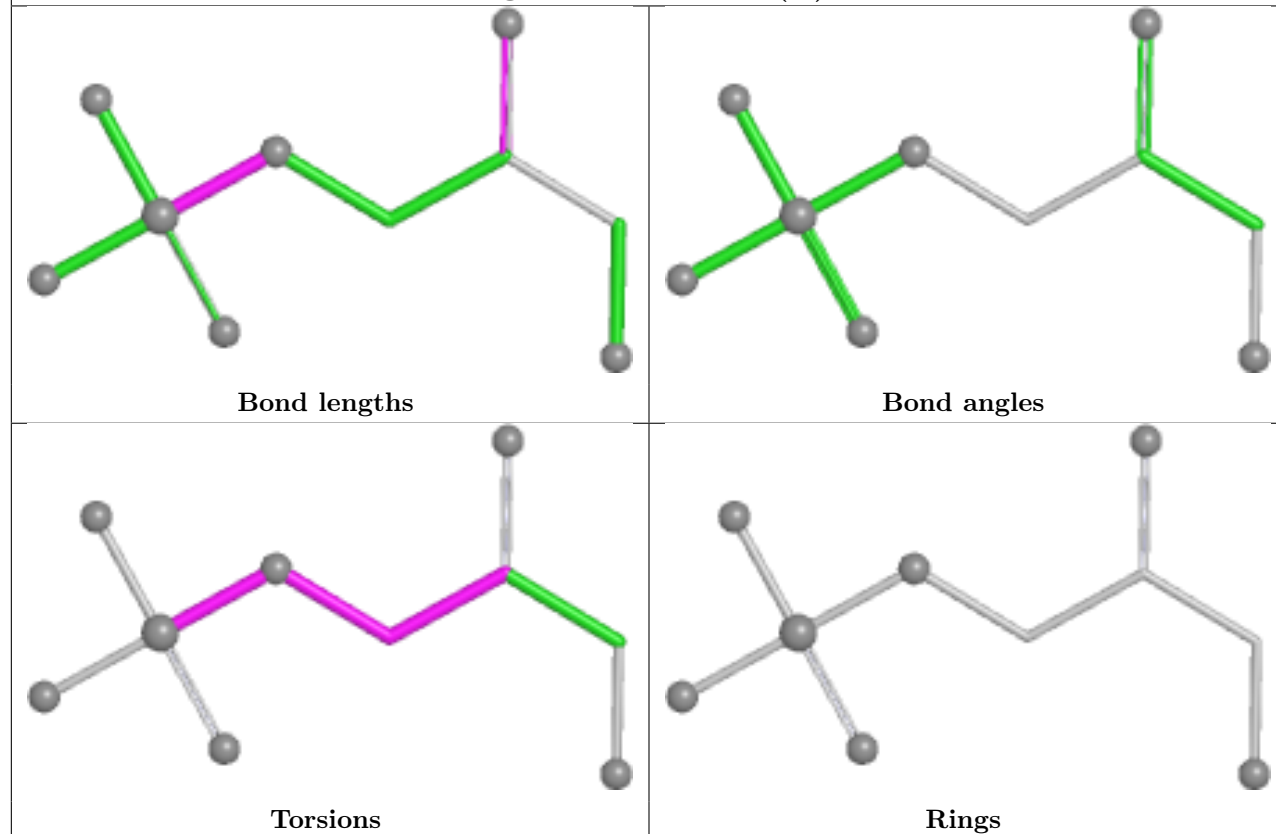
5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	501[B]	G3H	5	0
2	A	501	NAP	1	0
2	E	501	NAP	1	0
2	B	501	NAP	2	0
2	F	504	NAP	1	0

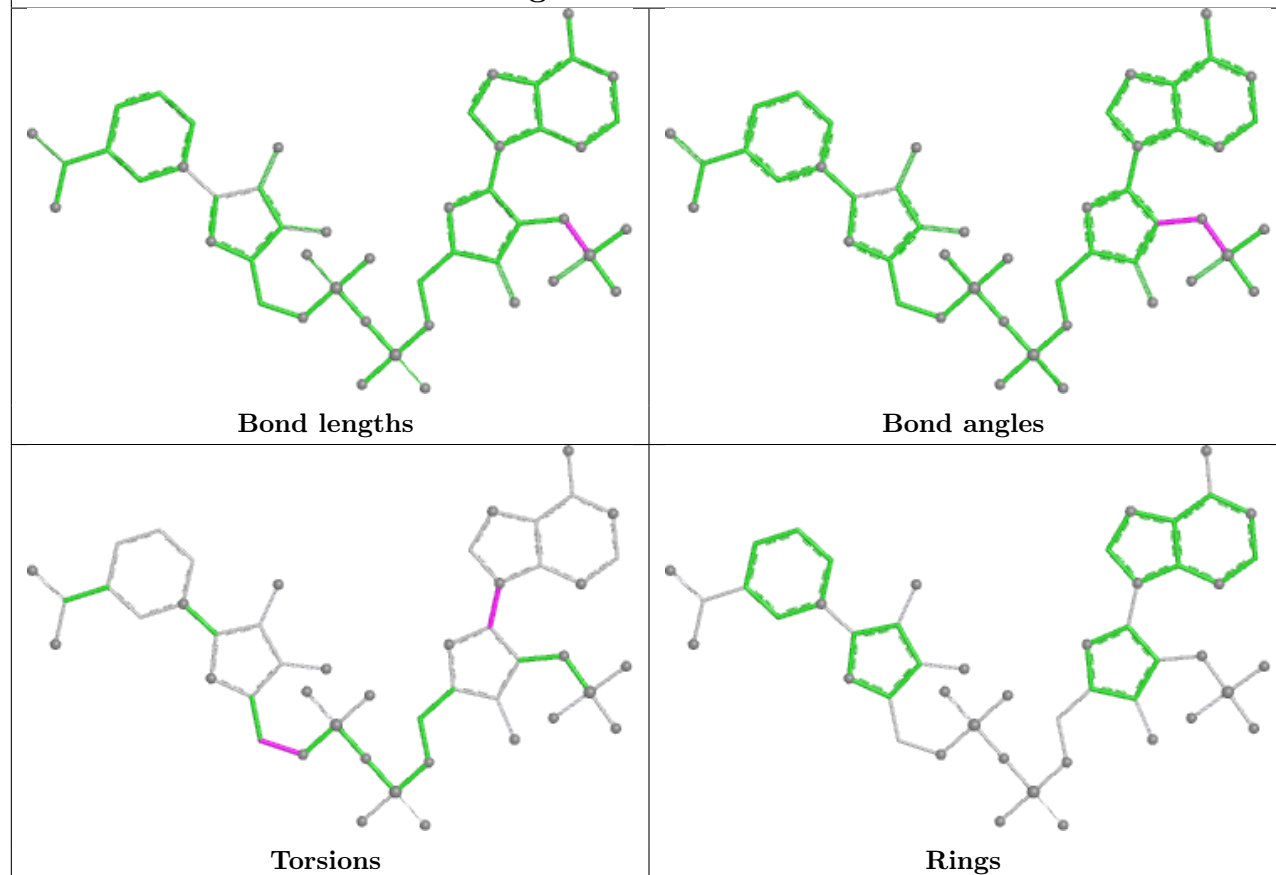
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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



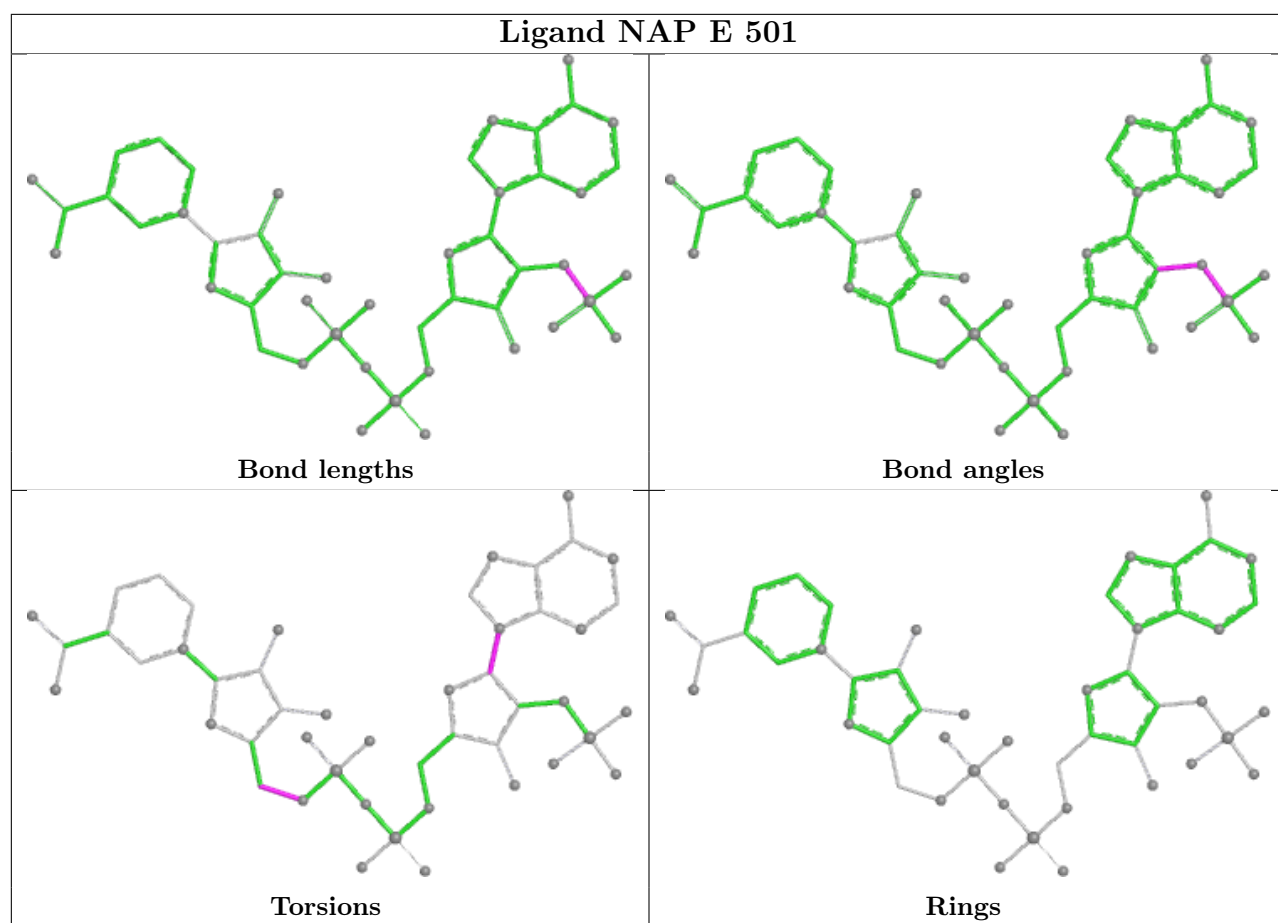
## Ligand G3H F 501 (B)

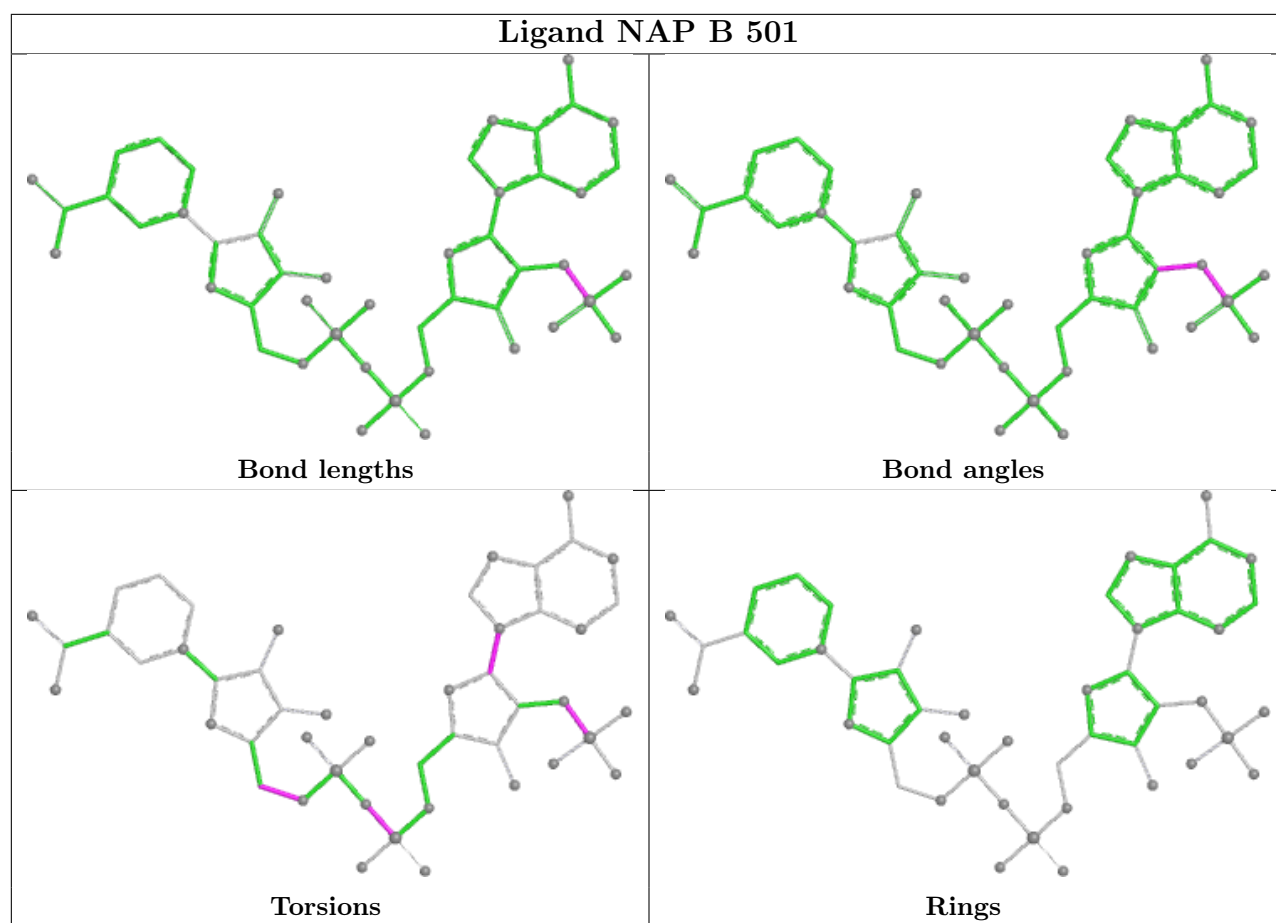


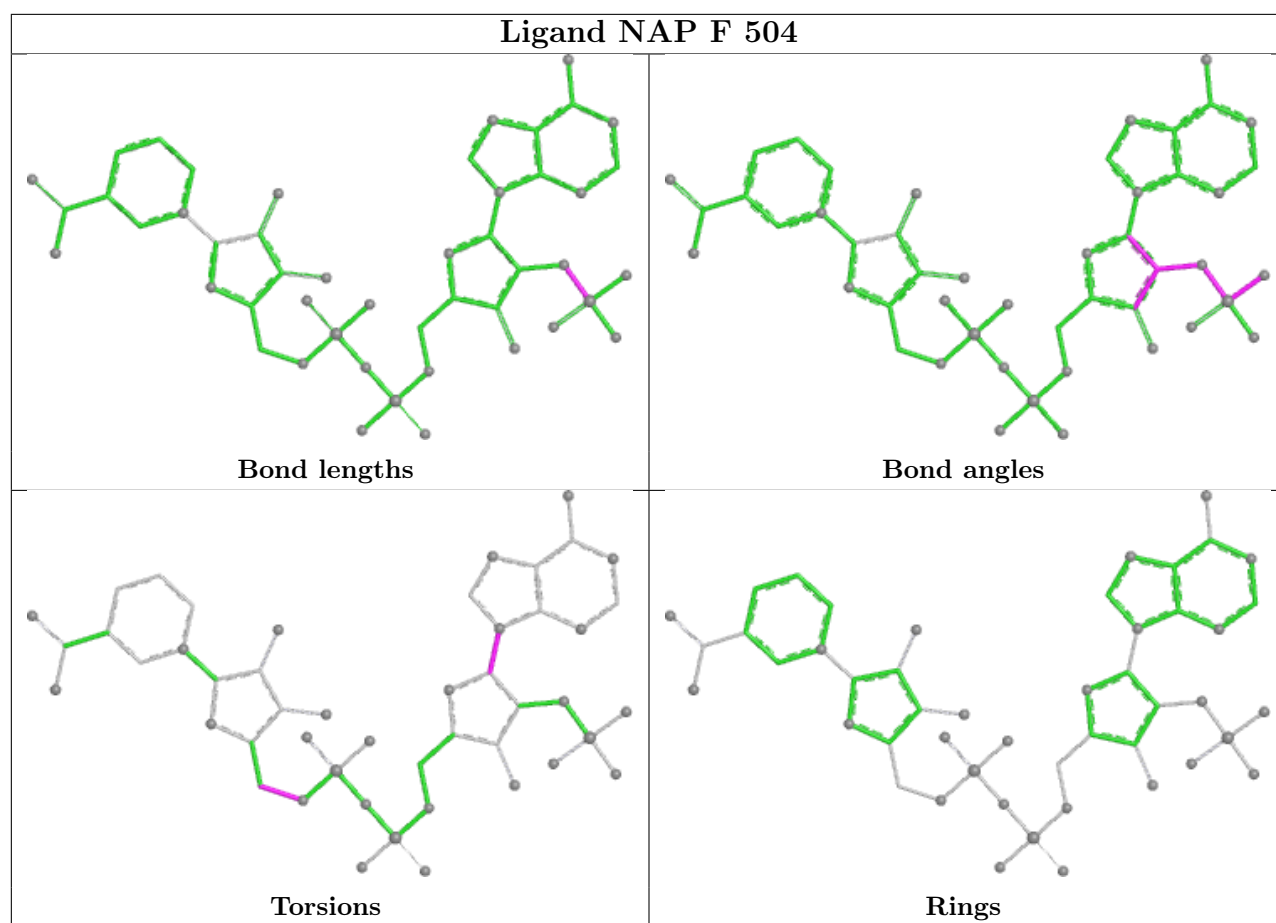
## Ligand NAP A 501











## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	474/496 (95%)	0.08	13 (2%) 56 61	9, 27, 53, 82	18 (3%)
1	B	475/496 (95%)	0.14	9 (1%) 66 72	9, 28, 44, 112	30 (6%)
1	C	475/496 (95%)	-0.49	3 (0%) 85 89	8, 17, 29, 71	29 (6%)
1	D	476/496 (95%)	-0.32	4 (0%) 82 87	8, 21, 33, 82	29 (6%)
1	E	474/496 (95%)	-0.04	4 (0%) 82 87	8, 25, 40, 93	17 (3%)
1	F	475/496 (95%)	-0.53	1 (0%) 91 93	7, 18, 28, 73	26 (5%)
1	G	475/496 (95%)	-0.49	2 (0%) 88 92	8, 17, 30, 86	26 (5%)
1	H	476/496 (95%)	-0.26	4 (0%) 82 87	8, 22, 36, 69	25 (5%)
All	All	3800/3968 (95%)	-0.24	40 (1%) 78 83	7, 21, 40, 112	200 (5%)

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	1	LEU	6.7
1	G	2	ALA	5.2
1	E	2	ALA	5.1
1	B	211[A]	SER	5.0
1	B	213[A]	ILE	4.8

### 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 oligosaccharides 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	GOL	D	501	6/6	0.81	0.11	45,63,79,79	0
3	GOL	F	502	6/6	0.83	0.14	35,57,67,69	0
4	G3H	F	501[A]	10/10	0.83	0.14	14,22,40,48	15
4	G3H	F	501[B]	10/10	0.83	0.14	5,30,45,54	15
3	GOL	F	503	6/6	0.84	0.10	35,52,62,71	0
2	NAP	A	501	48/48	0.85	0.11	29,45,62,67	0
2	NAP	B	501	48/48	0.88	0.10	22,39,51,56	0
2	NAP	E	501	48/48	0.92	0.09	22,30,42,44	0
2	NAP	F	504	48/48	0.97	0.05	14,18,23,28	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.