



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2021 – 04:55 PM EDT

PDB ID : 3BB1  
Title : Crystal structure of Toc34 from Pisum sativum in complex with Mg<sup>2+</sup> and GMPPNP  
Authors : Koenig, P.; Sinning, I.; Schleiff, E.; Tews, I.  
Deposited on : 2007-11-09  
Resolution : 2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](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 following versions of software and data (see [references ⓘ](#)) 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 : 2.23.2  
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) : 2.23.2

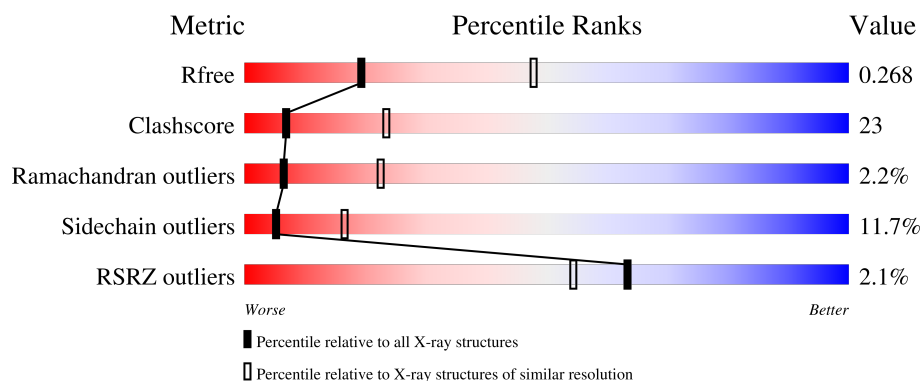
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	274	
1	B	274	
1	C	274	
1	D	274	
1	E	274	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	274	
1	G	274	
1	H	274	

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
5	GOL	C	291	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 15802 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 Translocase of chloroplast 34.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	252	Total	C	N	O	S	0	0	0
			1930	1226	328	372	4			
1	B	243	Total	C	N	O	S	0	0	0
			1853	1182	314	353	4			
1	C	246	Total	C	N	O	S	0	0	0
			1887	1202	320	361	4			
1	D	247	Total	C	N	O	S	0	0	0
			1882	1199	318	361	4			
1	E	248	Total	C	N	O	S	0	0	0
			1901	1211	323	363	4			
1	F	245	Total	C	N	O	S	0	0	0
			1882	1201	320	357	4			
1	G	251	Total	C	N	O	S	0	0	0
			1923	1221	327	371	4			
1	H	248	Total	C	N	O	S	0	0	0
			1897	1210	323	360	4			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	GLY	GLU	engineered mutation	UNP Q41009
A	267	LEU	-	expression tag	UNP Q41009
A	268	GLU	-	expression tag	UNP Q41009
A	269	HIS	-	expression tag	UNP Q41009
A	270	HIS	-	expression tag	UNP Q41009
A	271	HIS	-	expression tag	UNP Q41009
A	272	HIS	-	expression tag	UNP Q41009
A	273	HIS	-	expression tag	UNP Q41009
A	274	HIS	-	expression tag	UNP Q41009
B	10	GLY	GLU	engineered mutation	UNP Q41009
B	267	LEU	-	expression tag	UNP Q41009
B	268	GLU	-	expression tag	UNP Q41009
B	269	HIS	-	expression tag	UNP Q41009

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	270	HIS	-	expression tag	UNP Q41009
B	271	HIS	-	expression tag	UNP Q41009
B	272	HIS	-	expression tag	UNP Q41009
B	273	HIS	-	expression tag	UNP Q41009
B	274	HIS	-	expression tag	UNP Q41009
C	10	GLY	GLU	engineered mutation	UNP Q41009
C	267	LEU	-	expression tag	UNP Q41009
C	268	GLU	-	expression tag	UNP Q41009
C	269	HIS	-	expression tag	UNP Q41009
C	270	HIS	-	expression tag	UNP Q41009
C	271	HIS	-	expression tag	UNP Q41009
C	272	HIS	-	expression tag	UNP Q41009
C	273	HIS	-	expression tag	UNP Q41009
C	274	HIS	-	expression tag	UNP Q41009
D	10	GLY	GLU	engineered mutation	UNP Q41009
D	267	LEU	-	expression tag	UNP Q41009
D	268	GLU	-	expression tag	UNP Q41009
D	269	HIS	-	expression tag	UNP Q41009
D	270	HIS	-	expression tag	UNP Q41009
D	271	HIS	-	expression tag	UNP Q41009
D	272	HIS	-	expression tag	UNP Q41009
D	273	HIS	-	expression tag	UNP Q41009
D	274	HIS	-	expression tag	UNP Q41009
E	10	GLY	GLU	engineered mutation	UNP Q41009
E	267	LEU	-	expression tag	UNP Q41009
E	268	GLU	-	expression tag	UNP Q41009
E	269	HIS	-	expression tag	UNP Q41009
E	270	HIS	-	expression tag	UNP Q41009
E	271	HIS	-	expression tag	UNP Q41009
E	272	HIS	-	expression tag	UNP Q41009
E	273	HIS	-	expression tag	UNP Q41009
E	274	HIS	-	expression tag	UNP Q41009
F	10	GLY	GLU	engineered mutation	UNP Q41009
F	267	LEU	-	expression tag	UNP Q41009
F	268	GLU	-	expression tag	UNP Q41009
F	269	HIS	-	expression tag	UNP Q41009
F	270	HIS	-	expression tag	UNP Q41009
F	271	HIS	-	expression tag	UNP Q41009
F	272	HIS	-	expression tag	UNP Q41009
F	273	HIS	-	expression tag	UNP Q41009
F	274	HIS	-	expression tag	UNP Q41009
G	10	GLY	GLU	engineered mutation	UNP Q41009

*Continued on next page...*

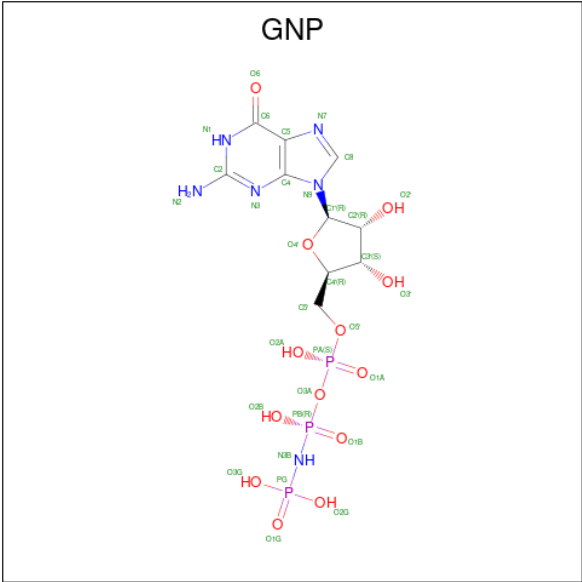
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	267	LEU	-	expression tag	UNP Q41009
G	268	GLU	-	expression tag	UNP Q41009
G	269	HIS	-	expression tag	UNP Q41009
G	270	HIS	-	expression tag	UNP Q41009
G	271	HIS	-	expression tag	UNP Q41009
G	272	HIS	-	expression tag	UNP Q41009
G	273	HIS	-	expression tag	UNP Q41009
G	274	HIS	-	expression tag	UNP Q41009
H	10	GLY	GLU	engineered mutation	UNP Q41009
H	267	LEU	-	expression tag	UNP Q41009
H	268	GLU	-	expression tag	UNP Q41009
H	269	HIS	-	expression tag	UNP Q41009
H	270	HIS	-	expression tag	UNP Q41009
H	271	HIS	-	expression tag	UNP Q41009
H	272	HIS	-	expression tag	UNP Q41009
H	273	HIS	-	expression tag	UNP Q41009
H	274	HIS	-	expression tag	UNP Q41009

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	B	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0
2	D	1	Total Mg 1 1	0	0
2	E	1	Total Mg 1 1	0	0
2	F	1	Total Mg 1 1	0	0
2	G	1	Total Mg 1 1	0	0
2	H	1	Total Mg 1 1	0	0

- Molecule 3 is PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER (three-letter code: GNP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>13</sub>P<sub>3</sub>).



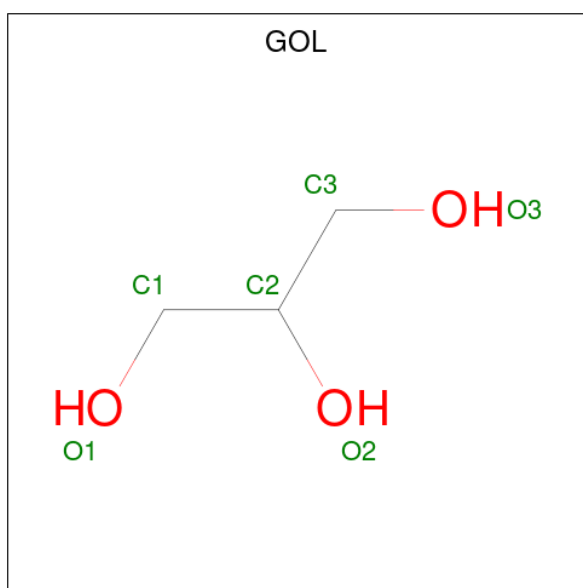
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
3	B	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
3	C	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
3	D	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
3	E	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
3	F	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
3	G	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
3	H	1	Total	C	N	O	P	0	0
			32	10	6	13	3		

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		
5	G	1	Total	C	O	0	0
			6	3	3		
5	G	1	Total	C	O	0	0
			6	3	3		

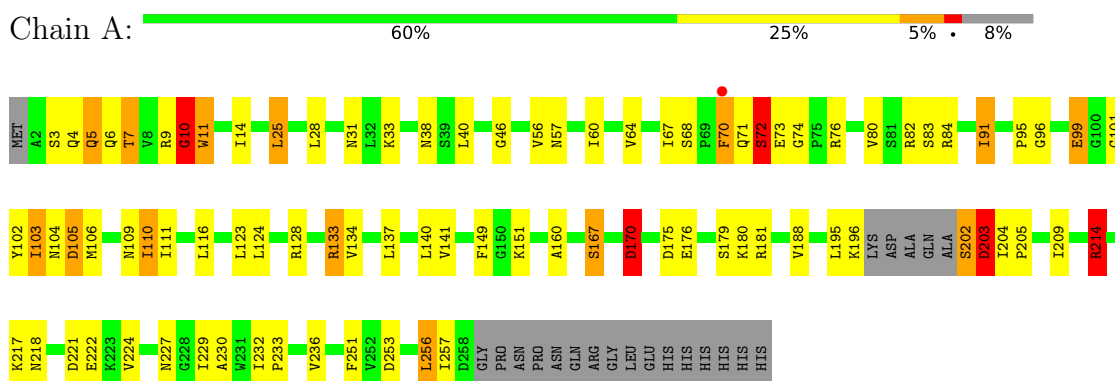
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	52	Total	O	0	0
			52	52		
6	B	24	Total	O	0	0
			24	24		
6	C	49	Total	O	0	0
			49	49		
6	D	39	Total	O	0	0
			39	39		
6	E	53	Total	O	0	0
			53	53		
6	F	35	Total	O	0	0
			35	35		
6	G	40	Total	O	0	0
			40	40		
6	H	35	Total	O	0	0
			35	35		

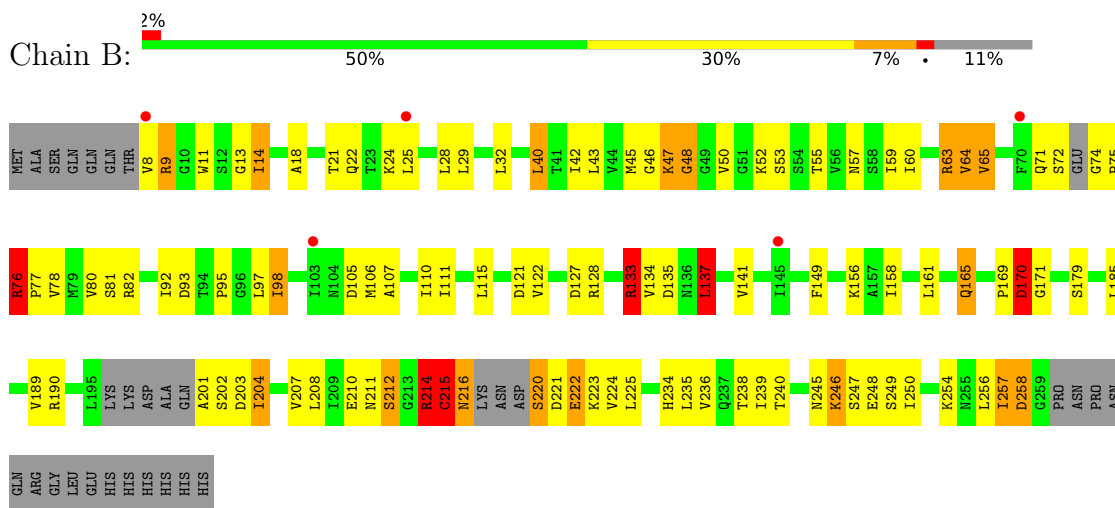
### 3 Residue-property plots

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.

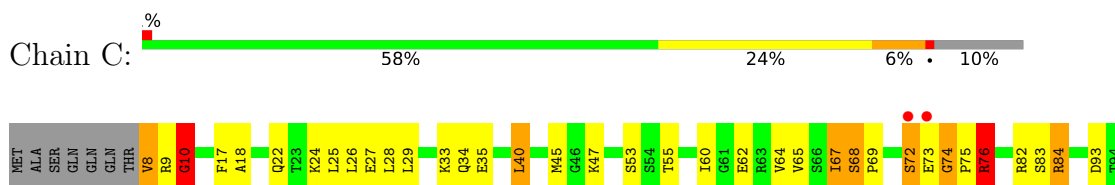
#### • Molecule 1: Translocase of chloroplast 34

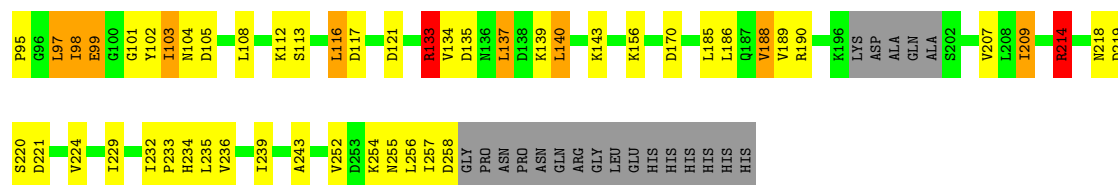


#### • Molecule 1: Translocase of chloroplast 34

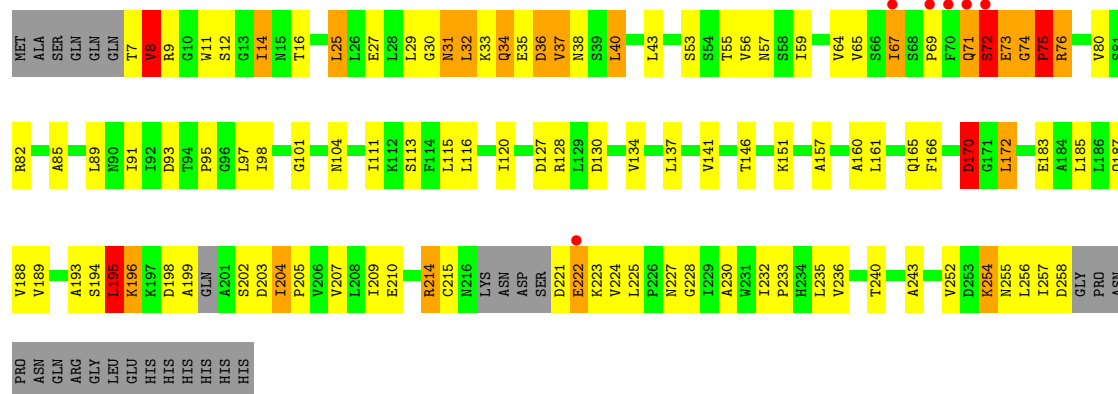


#### • Molecule 1: Translocase of chloroplast 34

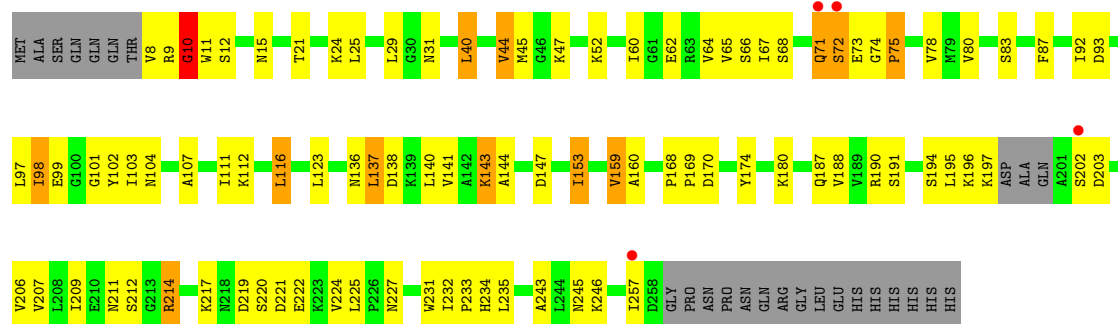




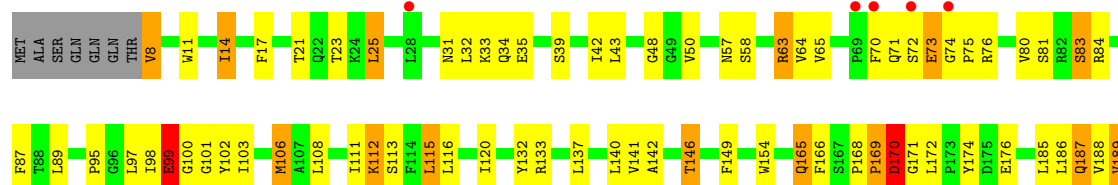
• Molecule 1: Translocase of chloroplast 34

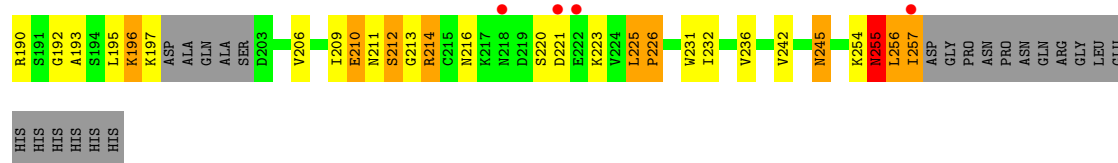


• Molecule 1: Translocase of chloroplast 34

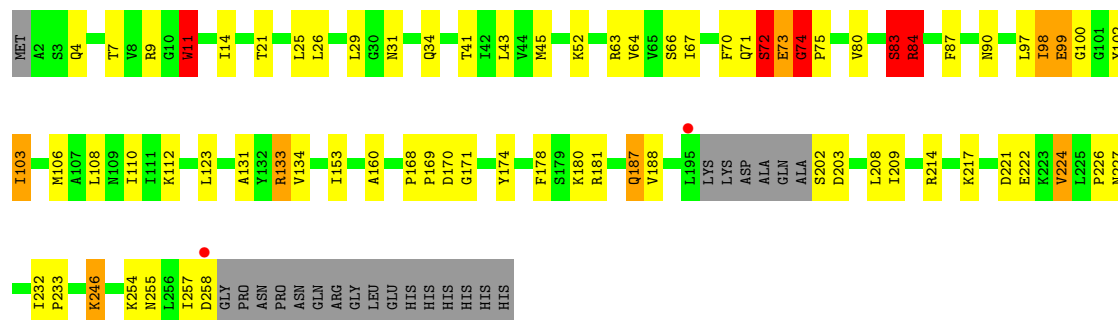


• Molecule 1: Translocase of chloroplast 34

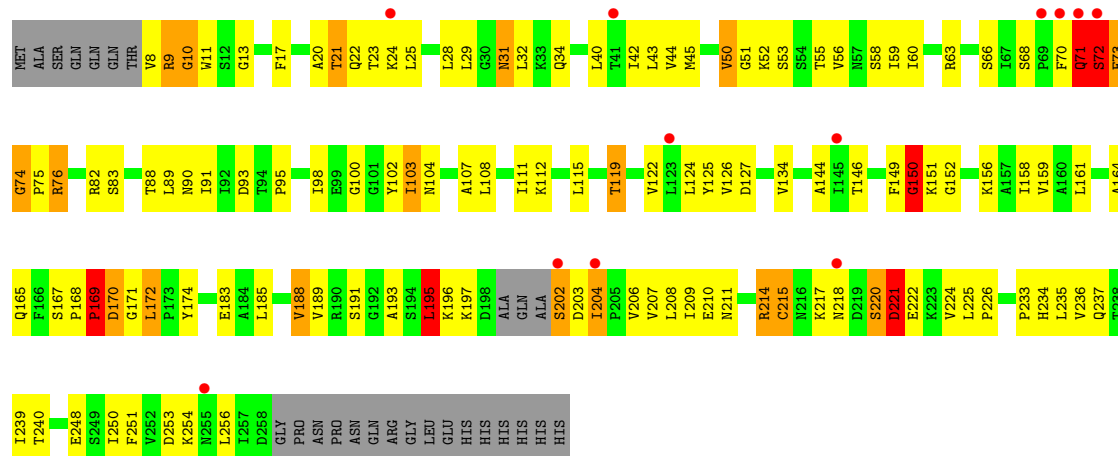
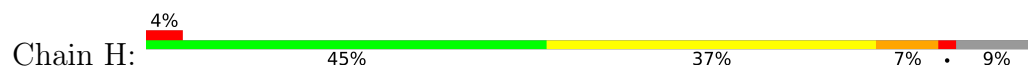




• Molecule 1: Translocase of chloroplast 34



• Molecule 1: Translocase of chloroplast 34



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	178.78Å 180.06Å 90.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.88 – 2.80 49.85 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.88-2.80) 97.5 (49.85-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.3.0008	Depositor
R, $R_{free}$	0.224 , 0.282 0.205 , 0.268	Depositor DCC
$R_{free}$ test set	3668 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.1	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 64.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.207 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	15802	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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: PGE, GOL, GNP, MG

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	0.85	1/1960 (0.1%)	0.95	4/2658 (0.2%)
1	B	0.77	0/1881	1.00	9/2549 (0.4%)
1	C	0.91	1/1917 (0.1%)	1.03	8/2599 (0.3%)
1	D	0.78	0/1911	1.04	6/2593 (0.2%)
1	E	0.88	1/1931 (0.1%)	0.93	4/2617 (0.2%)
1	F	0.78	0/1912	0.96	4/2591 (0.2%)
1	G	0.84	0/1953	0.93	5/2649 (0.2%)
1	H	0.72	0/1927	0.91	4/2612 (0.2%)
All	All	0.82	3/15392 (0.0%)	0.97	44/20868 (0.2%)

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	1	5
1	B	6	8
1	C	2	7
1	D	3	8
1	E	2	3
1	F	2	5
1	G	3	2
1	H	2	9
All	All	21	47

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	35	GLU	CD-OE1	5.87	1.32	1.25
1	E	214	ARG	CG-CD	5.63	1.66	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	196	LYS	C-O	5.54	1.33	1.23

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	10	GLY	N-CA-C	9.71	137.37	113.10
1	F	99	GLU	N-CA-C	8.64	134.33	111.00
1	C	214	ARG	NE-CZ-NH1	8.52	124.56	120.30
1	B	258	ASP	CB-CA-C	7.82	126.04	110.40
1	E	11	TRP	N-CA-C	-7.65	90.35	111.00

5 of 21 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	99	GLU	CA
1	B	47	LYS	CA
1	B	137	LEU	CA
1	B	203	ASP	CA
1	B	215	CYS	CA

5 of 47 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	10	GLY	Peptide
1	A	203	ASP	Peptide
1	A	70	PHE	Peptide
1	A	72	SER	Peptide
1	A	9	ARG	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	1930	0	1969	86	0
1	B	1853	0	1899	111	0
1	C	1887	0	1935	94	0
1	D	1882	0	1918	113	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1901	0	1953	70	0
1	F	1882	0	1939	99	0
1	G	1923	0	1963	67	0
1	H	1897	0	1945	115	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	32	0	13	0	0
3	B	32	0	13	1	0
3	C	32	0	13	1	0
3	D	32	0	13	5	0
3	E	32	0	13	0	0
3	F	32	0	13	0	0
3	G	32	0	13	1	0
3	H	32	0	13	2	0
4	A	10	0	14	0	0
4	F	10	0	14	4	0
5	A	6	0	8	1	0
5	C	6	0	8	4	0
5	E	6	0	8	0	0
5	F	6	0	8	0	0
5	G	12	0	16	3	0
6	A	52	0	0	9	0
6	B	24	0	0	8	0
6	C	49	0	0	10	0
6	D	39	0	0	8	0
6	E	53	0	0	2	0
6	F	35	0	0	10	0
6	G	40	0	0	12	0
6	H	35	0	0	16	0
All	All	15802	0	15701	721	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 23.

The worst 5 of 721 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:H:122:VAL:HG21	6:H:315:HOH:O	1.31	1.26
1:G:67:ILE:O	1:G:67:ILE:HD12	1.38	1.20
1:F:189:VAL:O	1:F:193:ALA:HB2	1.43	1.19
1:A:116:LEU:HD21	1:A:257:ILE:CD1	1.73	1.18
1:B:246:LYS:CB	1:B:247:SER:HA	1.63	1.14

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	248/274 (90%)	234 (94%)	13 (5%)	1 (0%)	34	66
1	B	235/274 (86%)	204 (87%)	25 (11%)	6 (3%)	5	18
1	C	242/274 (88%)	222 (92%)	17 (7%)	3 (1%)	13	39
1	D	241/274 (88%)	216 (90%)	17 (7%)	8 (3%)	4	13
1	E	244/274 (89%)	220 (90%)	20 (8%)	4 (2%)	9	31
1	F	241/274 (88%)	218 (90%)	20 (8%)	3 (1%)	13	39
1	G	247/274 (90%)	222 (90%)	18 (7%)	7 (3%)	5	17
1	H	244/274 (89%)	213 (87%)	20 (8%)	11 (4%)	2	8
All	All	1942/2192 (89%)	1749 (90%)	150 (8%)	43 (2%)	6	22

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	SER
1	B	77	PRO
1	B	170	ASP
1	B	215	CYS
1	D	9	ARG

### 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	217/236 (92%)	193 (89%)	24 (11%)	6	19
1	B	208/236 (88%)	183 (88%)	25 (12%)	5	15
1	C	213/236 (90%)	189 (89%)	24 (11%)	6	18
1	D	210/236 (89%)	183 (87%)	27 (13%)	4	13
1	E	214/236 (91%)	188 (88%)	26 (12%)	5	15
1	F	212/236 (90%)	178 (84%)	34 (16%)	2	7
1	G	217/236 (92%)	202 (93%)	15 (7%)	15	41
1	H	212/236 (90%)	187 (88%)	25 (12%)	5	16
All	All	1703/1888 (90%)	1503 (88%)	200 (12%)	5	16

5 of 200 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	153	ILE
1	F	146	THR
1	H	254	LYS
1	E	194	SER
1	F	33	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	31	ASN
1	H	71	GLN
1	H	211	ASN
1	D	187	GLN
1	D	155	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 24 ligands modelled in this entry, 8 are monoatomic - leaving 16 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	GOL	C	291	-	5,5,5	0.43	0	5,5,5	0.96	0
3	GNP	H	281	2	28,34,34	2.47	10 (35%)	30,54,54	2.89	9 (30%)
3	GNP	C	281	2	28,34,34	2.85	12 (42%)	30,54,54	2.69	6 (20%)
3	GNP	G	281	2	28,34,34	2.52	11 (39%)	30,54,54	3.03	10 (33%)
5	GOL	G	291	-	5,5,5	0.46	0	5,5,5	0.74	0
3	GNP	F	281	2	28,34,34	2.95	11 (39%)	30,54,54	2.22	7 (23%)
5	GOL	E	290	-	5,5,5	0.55	0	5,5,5	1.11	0
3	GNP	A	281	2	28,34,34	2.62	9 (32%)	30,54,54	2.93	10 (33%)
3	GNP	E	281	2	28,34,34	2.74	11 (39%)	30,54,54	2.94	11 (36%)
4	PGE	F	290	-	9,9,9	0.69	0	8,8,8	0.85	0
5	GOL	A	291	-	5,5,5	0.33	0	5,5,5	0.82	0
3	GNP	B	281	2	28,34,34	2.54	8 (28%)	30,54,54	2.33	6 (20%)
4	PGE	A	290	-	9,9,9	0.82	0	8,8,8	0.71	0
3	GNP	D	281	2	28,34,34	2.44	11 (39%)	30,54,54	2.42	9 (30%)
5	GOL	G	290	-	5,5,5	0.38	0	5,5,5	0.56	0
5	GOL	F	291	-	5,5,5	0.35	0	5,5,5	0.51	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	291	-	-	2/4/4/4	-
3	GNP	H	281	2	-	6/17/38/38	0/3/3/3
3	GNP	C	281	2	-	3/17/38/38	0/3/3/3
3	GNP	G	281	2	-	4/17/38/38	0/3/3/3
5	GOL	G	291	-	-	2/4/4/4	-
3	GNP	F	281	2	-	3/17/38/38	0/3/3/3
5	GOL	E	290	-	-	4/4/4/4	-
3	GNP	A	281	2	-	3/17/38/38	0/3/3/3
3	GNP	E	281	2	-	6/17/38/38	0/3/3/3
4	PGE	F	290	-	-	5/7/7/7	-
5	GOL	A	291	-	-	2/4/4/4	-
3	GNP	B	281	2	-	3/17/38/38	0/3/3/3
4	PGE	A	290	-	-	3/7/7/7	-
3	GNP	D	281	2	-	3/17/38/38	0/3/3/3
5	GOL	G	290	-	-	2/4/4/4	-
5	GOL	F	291	-	-	0/4/4/4	-

The worst 5 of 83 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	281	GNP	C4-N9	-9.33	1.35	1.47
3	B	281	GNP	C4-N9	-8.57	1.36	1.47
3	E	281	GNP	C4-N9	-8.38	1.36	1.47
3	C	281	GNP	C5-C6	-8.34	1.38	1.52
3	H	281	GNP	C4-N9	-8.22	1.36	1.47

The worst 5 of 68 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	281	GNP	O1G-PG-N3B	-9.87	97.24	111.77
3	C	281	GNP	C4-C5-N7	9.32	114.81	102.46
3	E	281	GNP	C4-C5-N7	8.53	113.77	102.46
3	A	281	GNP	C4-C5-N7	7.94	112.98	102.46
3	A	281	GNP	C5-C6-N1	-7.48	108.96	118.19

There are no chirality outliers.

5 of 51 torsion outliers are listed below:

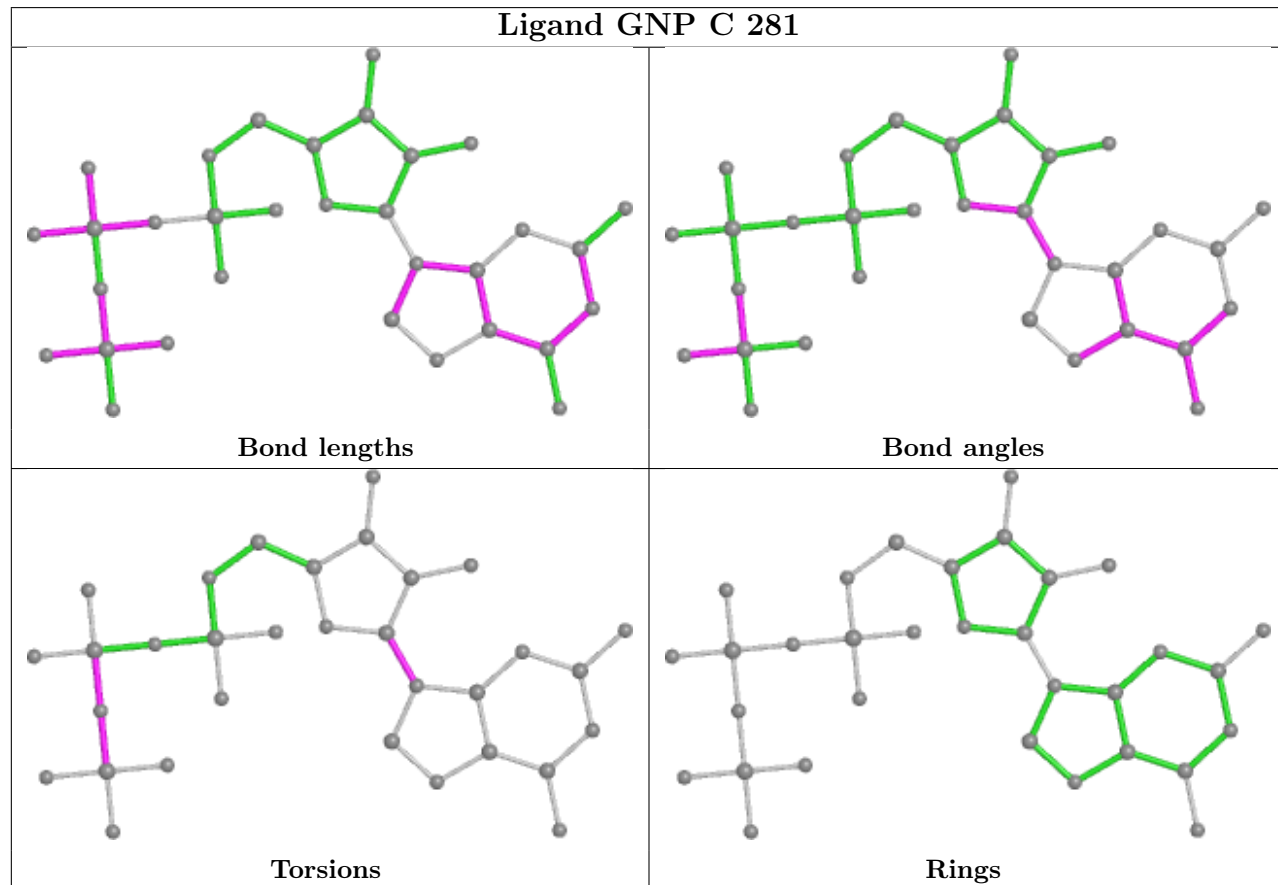
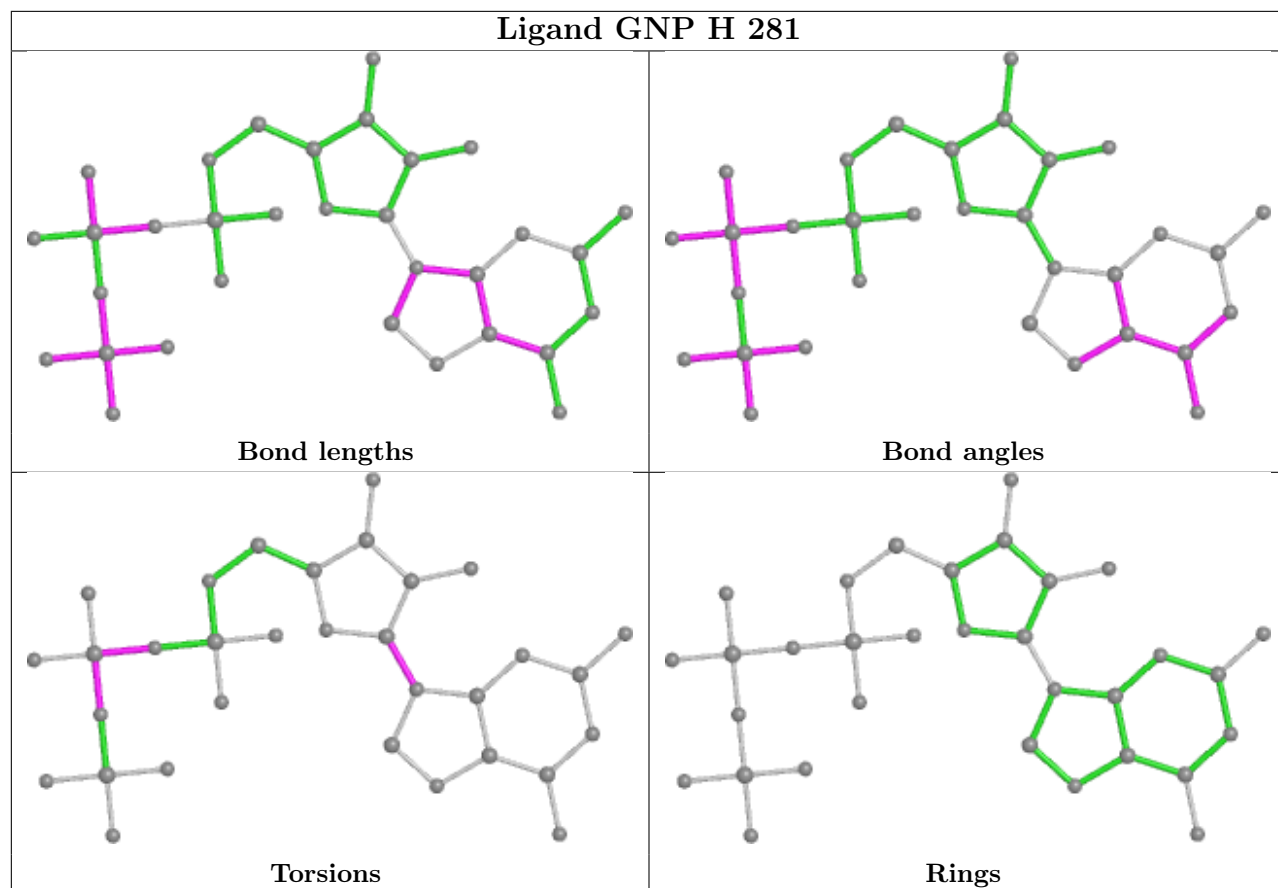
Mol	Chain	Res	Type	Atoms
3	A	281	GNP	PG-N3B-PB-O1B
3	A	281	GNP	PA-O3A-PB-O1B
3	A	281	GNP	C2'-C1'-N9-C4
3	B	281	GNP	PG-N3B-PB-O1B
3	B	281	GNP	PG-N3B-PB-O3A

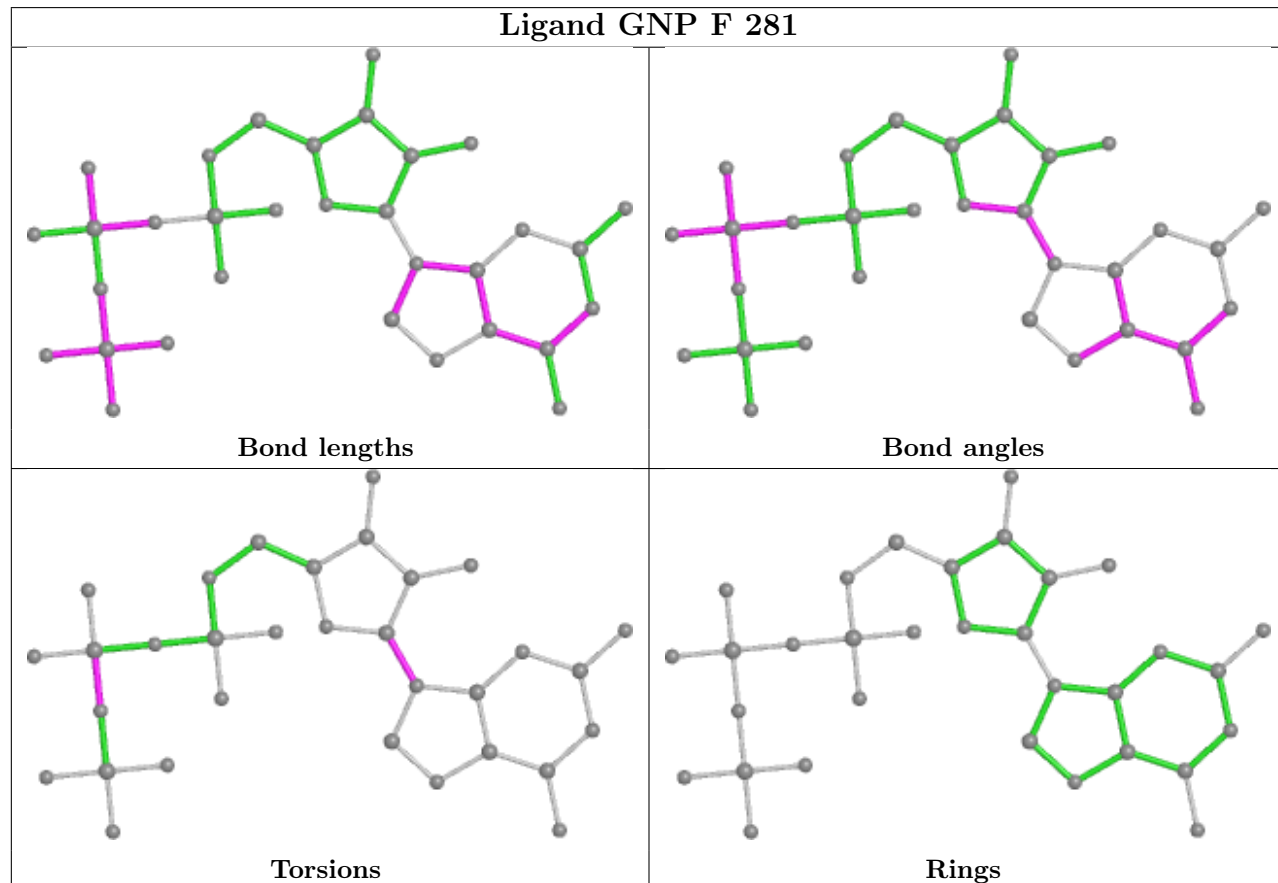
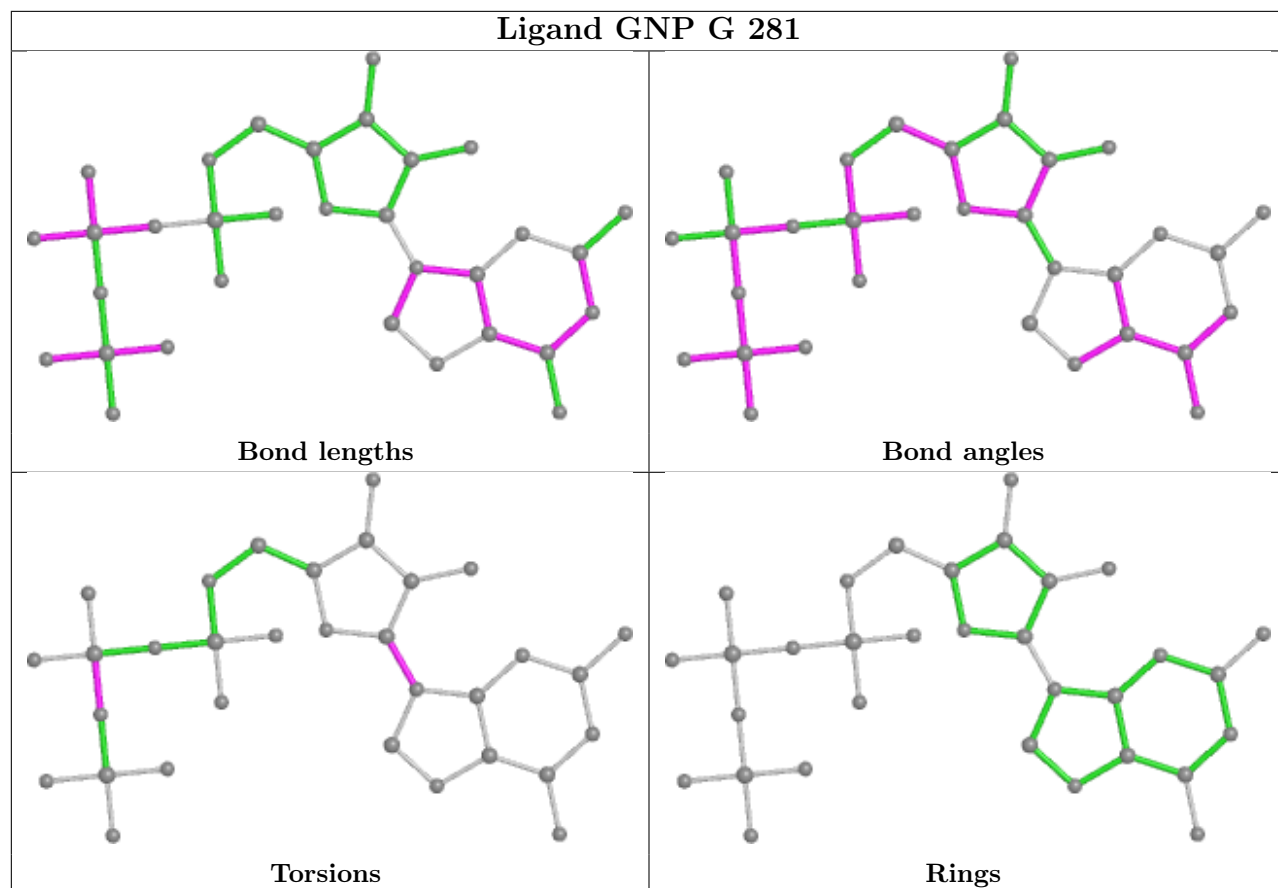
There are no ring outliers.

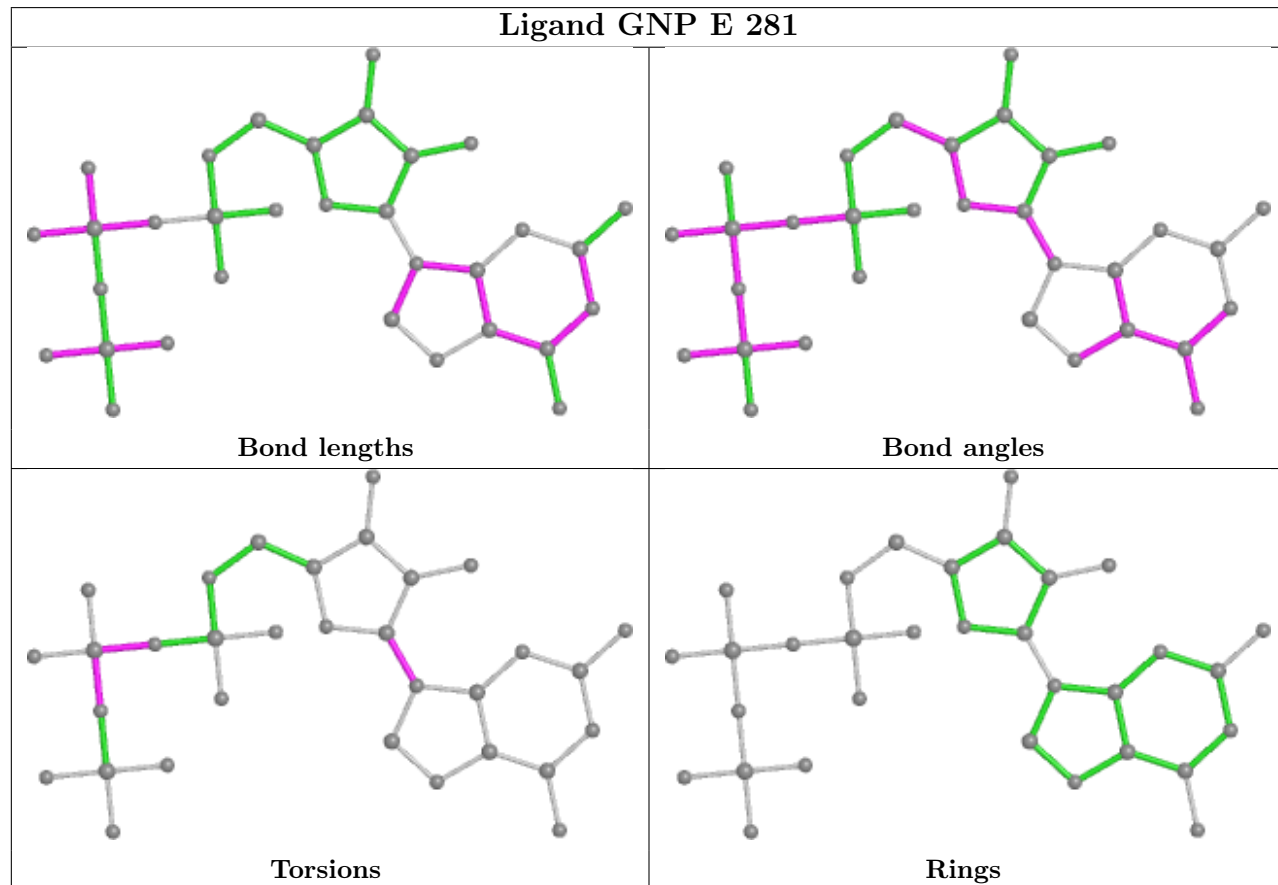
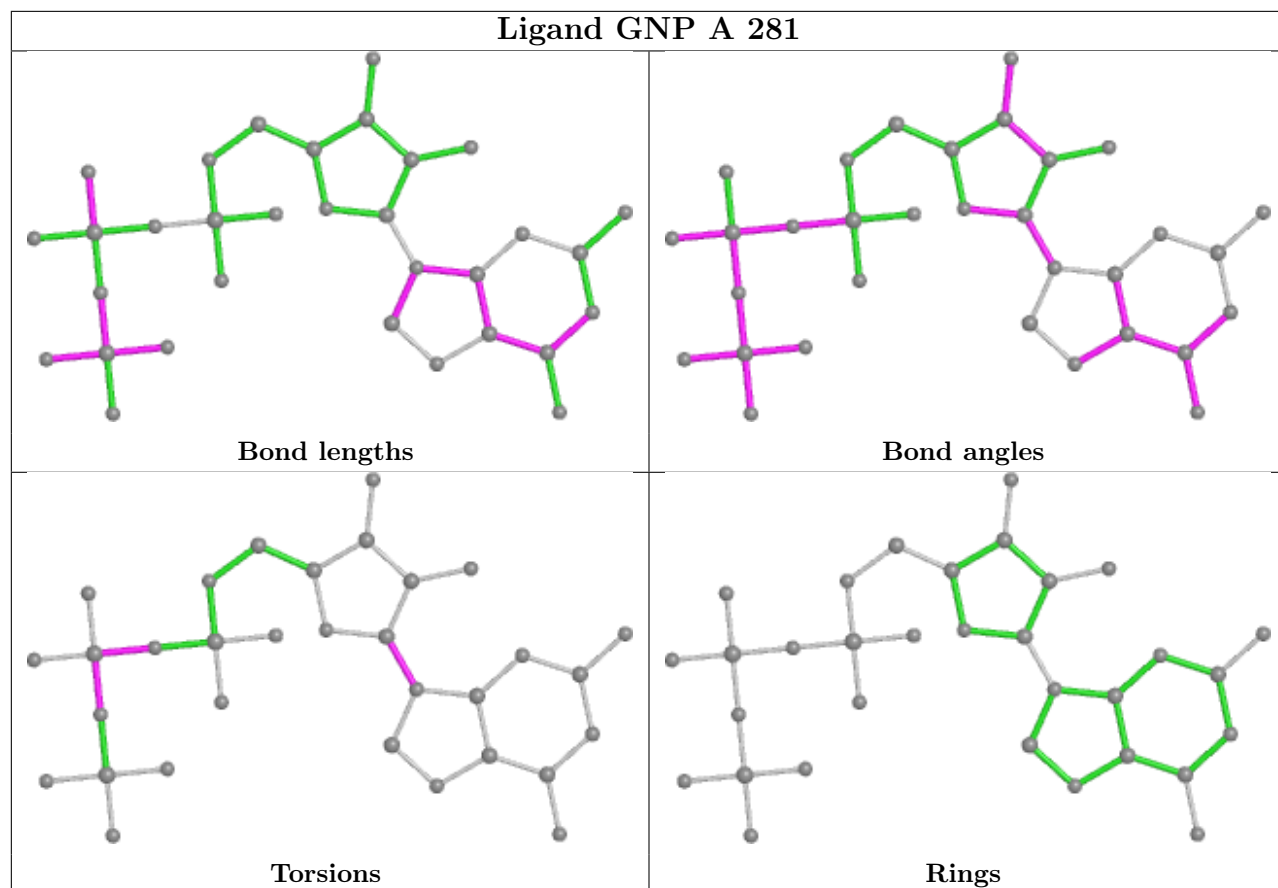
10 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	291	GOL	4	0
3	H	281	GNP	2	0
3	C	281	GNP	1	0
3	G	281	GNP	1	0
5	G	291	GOL	1	0
4	F	290	PGE	4	0
5	A	291	GOL	1	0
3	B	281	GNP	1	0
3	D	281	GNP	5	0
5	G	290	GOL	2	0

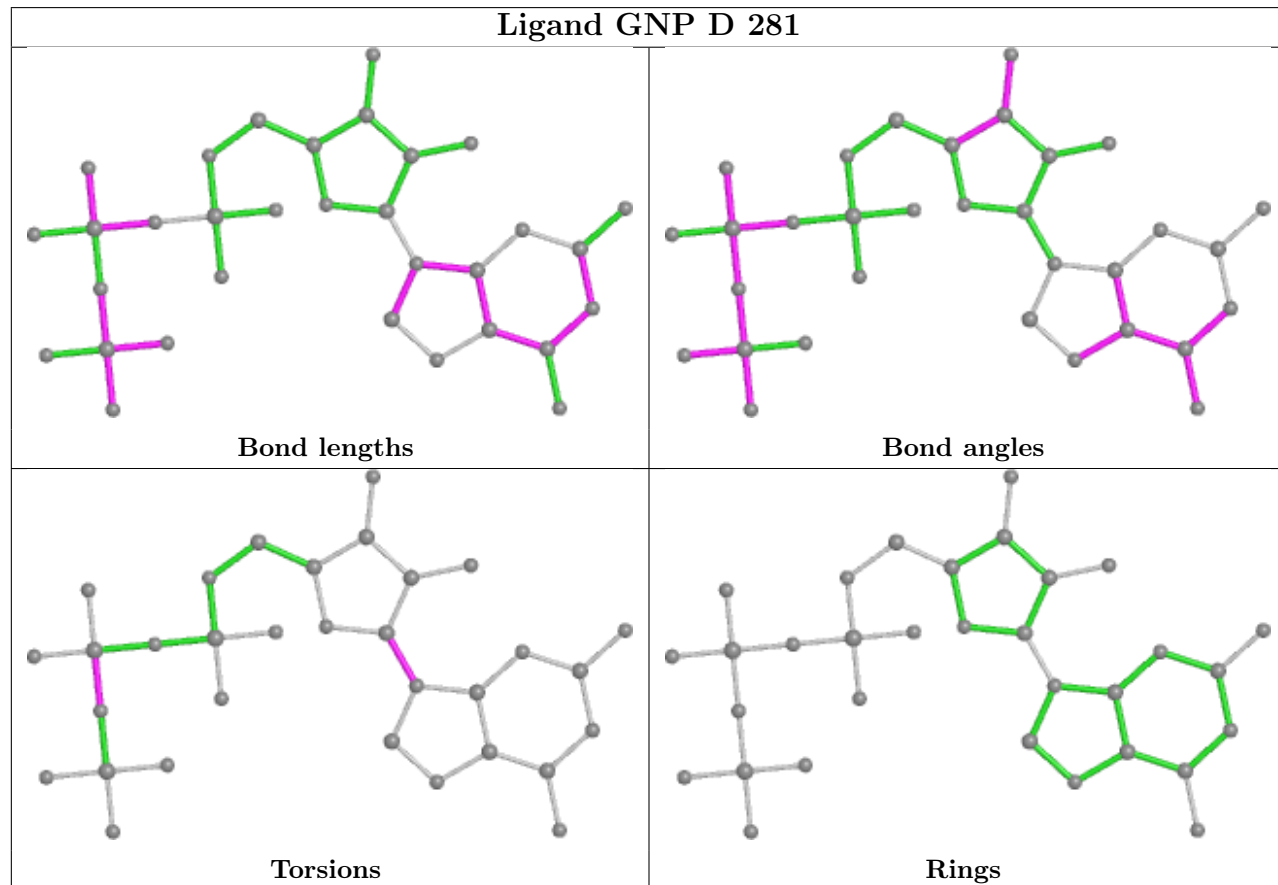
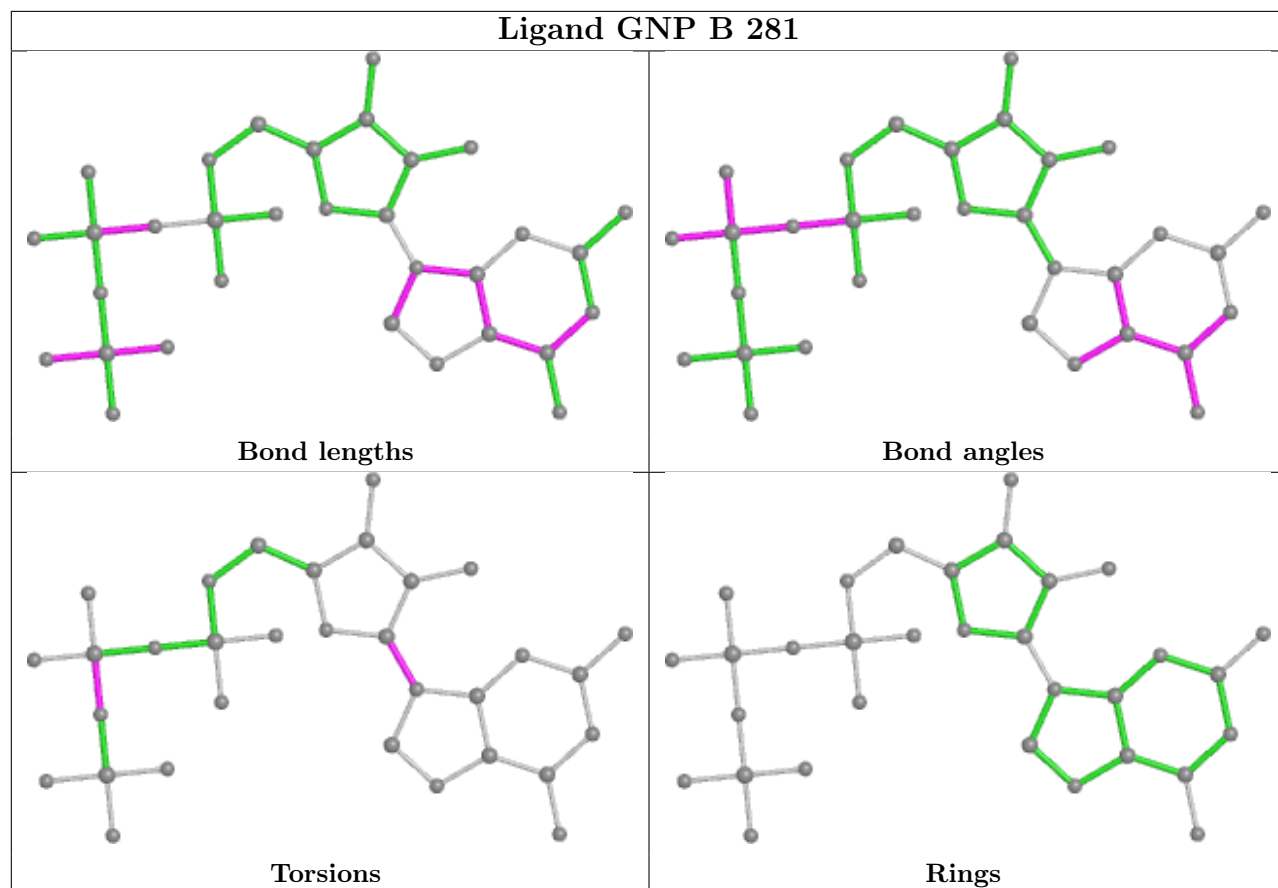
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

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	252/274 (91%)	0.03	1 (0%) 92 91	58, 66, 84, 99	0
1	B	243/274 (88%)	0.28	5 (2%) 63 54	66, 82, 104, 116	0
1	C	246/274 (89%)	0.11	2 (0%) 86 81	55, 64, 84, 100	0
1	D	247/274 (90%)	0.23	6 (2%) 59 49	62, 76, 99, 114	0
1	E	248/274 (90%)	0.06	4 (1%) 72 66	54, 63, 88, 94	0
1	F	245/274 (89%)	0.21	9 (3%) 41 31	62, 74, 97, 109	0
1	G	251/274 (91%)	0.01	2 (0%) 86 81	56, 67, 86, 100	0
1	H	248/274 (90%)	0.33	12 (4%) 30 21	70, 85, 102, 111	0
All	All	1980/2192 (90%)	0.16	41 (2%) 63 54	54, 72, 98, 116	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	72	SER	4.8
1	C	73	GLU	4.3
1	H	202	SER	4.2
1	D	69	PRO	3.9
1	F	70	PHE	3.6

### 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 ⓘ

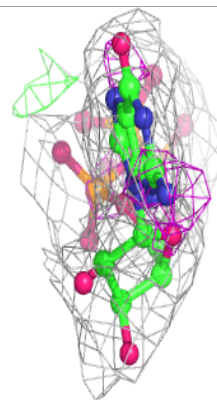
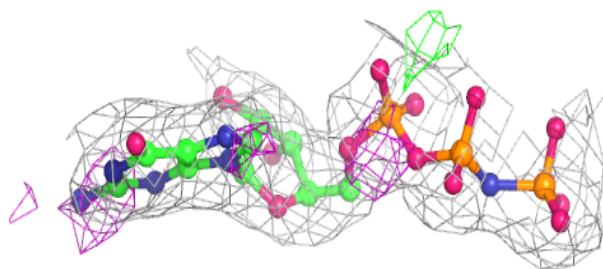
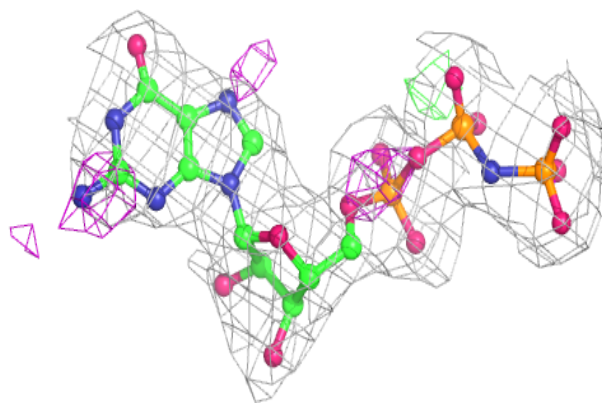
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	GOL	G	290	6/6	0.78	0.27	110,110,111,111	0
4	PGE	F	290	10/10	0.82	0.22	99,103,104,104	0
5	GOL	G	291	6/6	0.86	0.42	87,90,91,93	0
5	GOL	A	291	6/6	0.88	0.23	96,98,99,99	0
4	PGE	A	290	10/10	0.89	0.50	75,81,82,82	0
5	GOL	F	291	6/6	0.89	0.17	88,89,89,90	0
5	GOL	C	291	6/6	0.90	0.29	89,91,91,91	0
5	GOL	E	290	6/6	0.95	0.20	55,59,61,62	0
2	MG	H	282	1/1	0.95	0.10	71,71,71,71	0
3	GNP	H	281	32/32	0.97	0.15	67,71,75,76	0
2	MG	B	282	1/1	0.97	0.08	66,66,66,66	0
3	GNP	G	281	32/32	0.97	0.15	59,62,65,66	0
3	GNP	F	281	32/32	0.98	0.17	56,64,65,66	0
2	MG	C	282	1/1	0.98	0.07	63,63,63,63	0
2	MG	D	282	1/1	0.98	0.07	58,58,58,58	0
2	MG	E	282	1/1	0.98	0.04	64,64,64,64	0
2	MG	F	282	1/1	0.98	0.06	63,63,63,63	0
2	MG	A	282	1/1	0.98	0.10	65,65,65,65	0
3	GNP	A	281	32/32	0.98	0.17	58,63,65,66	0
3	GNP	B	281	32/32	0.98	0.15	67,70,73,77	0
3	GNP	C	281	32/32	0.98	0.13	60,64,66,67	0
3	GNP	D	281	32/32	0.98	0.15	60,65,68,69	0
3	GNP	E	281	32/32	0.98	0.13	59,62,67,72	0
2	MG	G	282	1/1	0.99	0.12	63,63,63,63	0

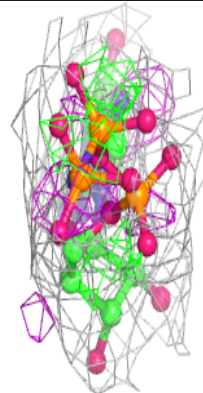
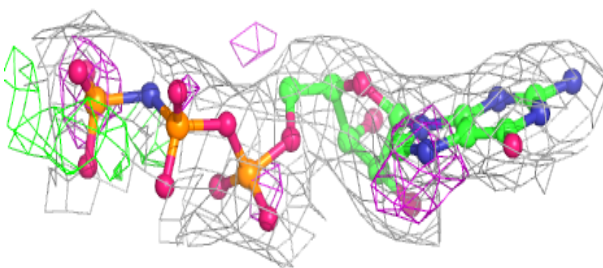
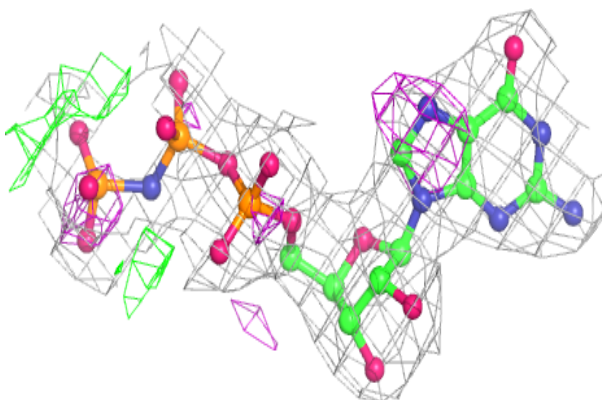
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around GNP H 281:**

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

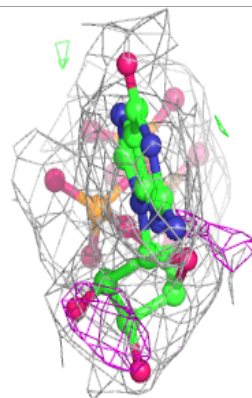
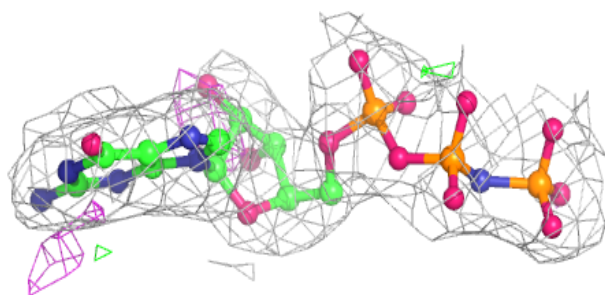
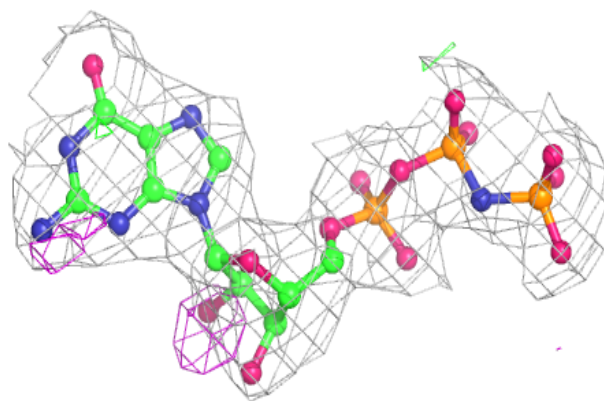
**Electron density around GNP G 281:**

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

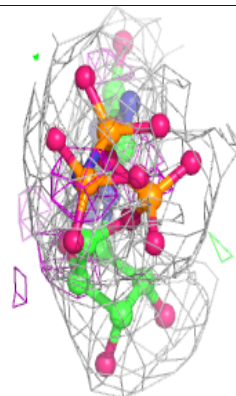
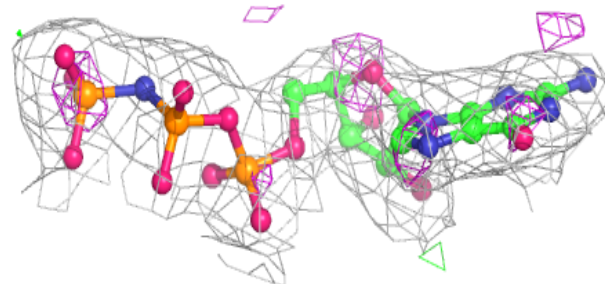
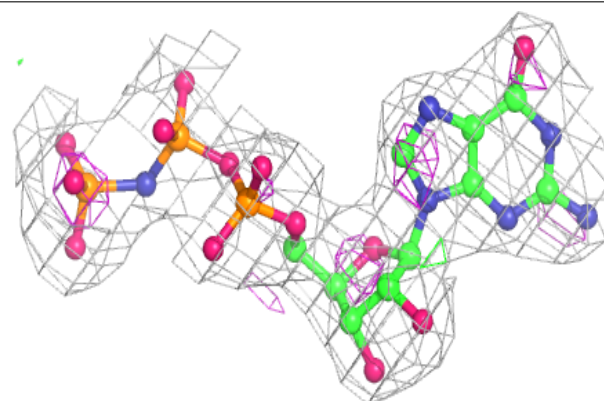


**Electron density around GNP F 281:**

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

**Electron density around GNP A 281:**

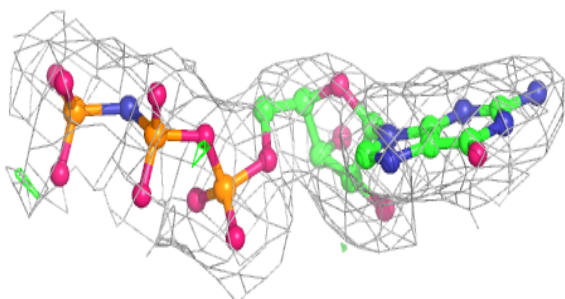
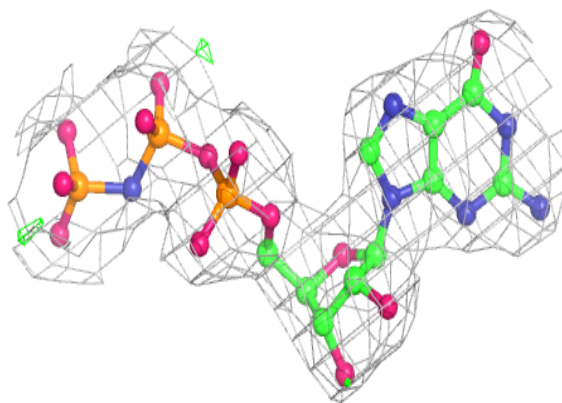
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



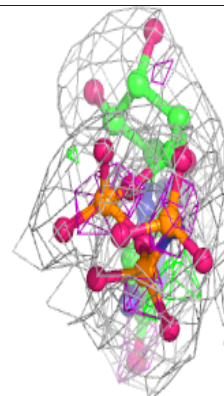
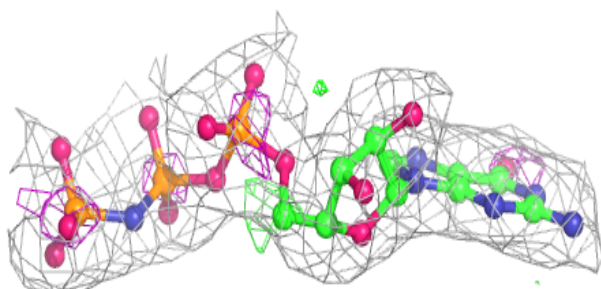
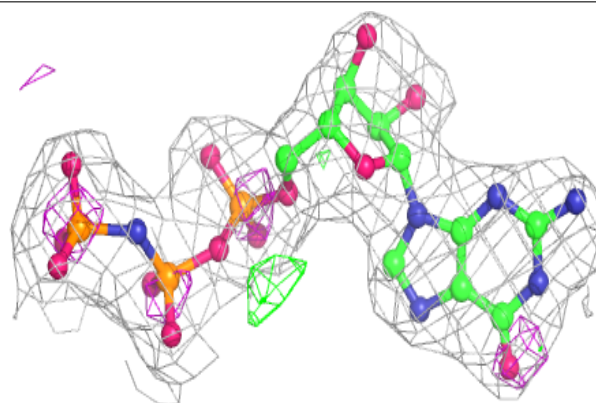


**Electron density around GNP B 281:**

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

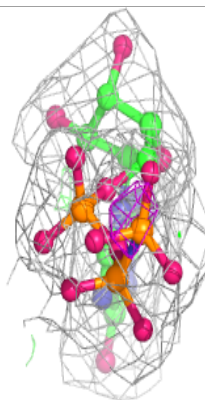
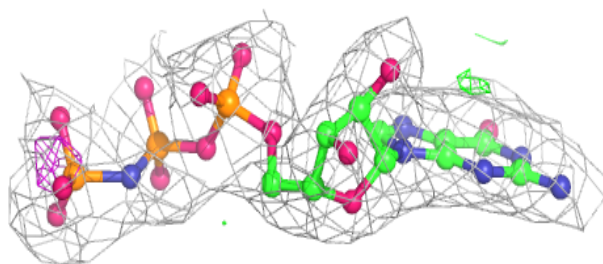
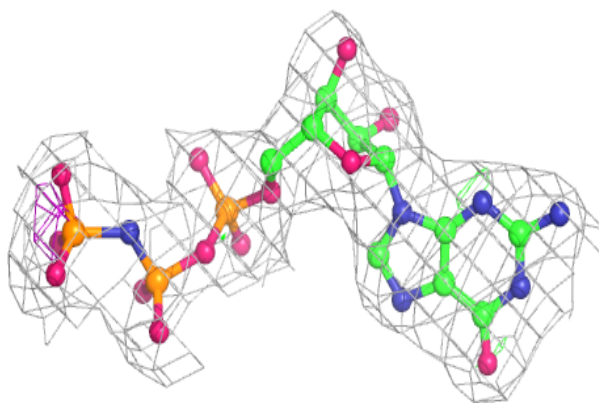
**Electron density around GNP C 281:**

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

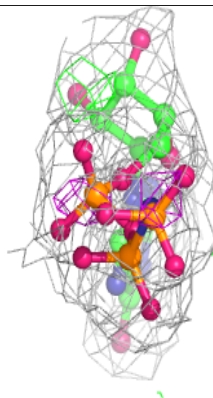
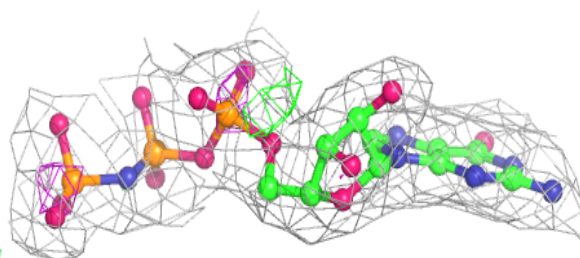
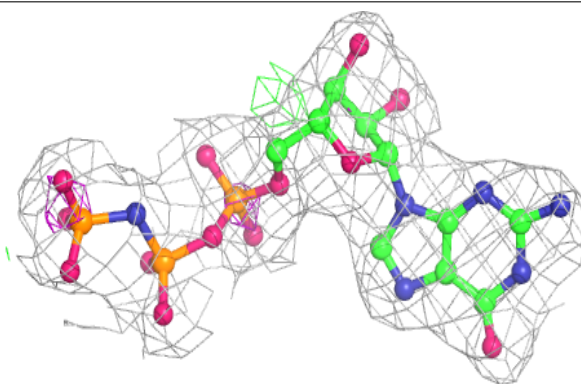


**Electron density around GNP D 281:**

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

**Electron density around GNP E 281:**

$2mF_o-DF_c$  (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.