



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 03:46 am BST

PDB ID : 4BZC  
Title : Crystal structure of the tetrameric dGTP-bound wild type SAMHD1 catalytic core  
Authors : Ji, X.; Yang, H.; Wu, Y.; Yan, J.; Mehrens, J.; DeLucia, M.; Hao, C.; Gronenborn, A.M.; Skowronski, J.; Ahn, J.; Xiong, Y.  
Deposited on : 2013-07-25  
Resolution : 2.88 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 : 2.11  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

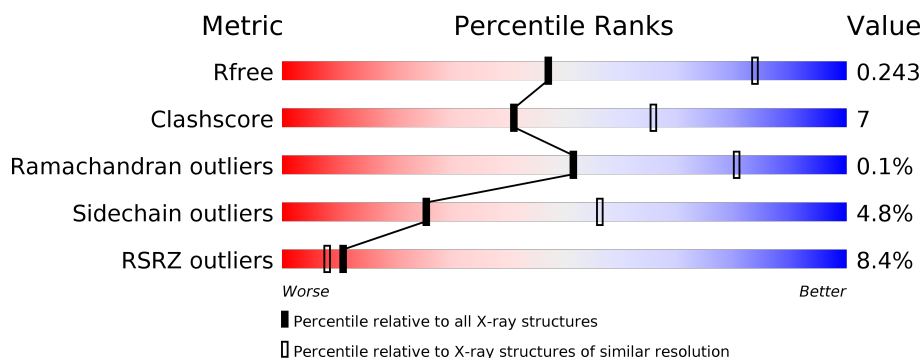
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.88 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	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	550	<div> <div>4%</div> <div> <div></div> <div>70%</div> <div>15%</div> <div>•</div> <div>13%</div> </div> </div>
1	B	550	<div> <div>11%</div> <div> <div></div> <div>73%</div> <div>11%</div> <div>•</div> <div>13%</div> </div> </div>
1	C	550	<div> <div>6%</div> <div> <div></div> <div>72%</div> <div>14%</div> <div>•</div> <div>13%</div> </div> </div>
1	D	550	<div> <div>8%</div> <div> <div></div> <div>70%</div> <div>13%</div> <div>•</div> <div>15%</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15954 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 DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHO-HYDROLASE SAMHD1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	476	Total	C	N	O	S	0	0	0
			3880	2483	672	705	20			
1	B	477	Total	C	N	O	S	0	0	0
			3887	2489	678	700	20			
1	C	481	Total	C	N	O	S	0	0	0
			3932	2517	683	712	20			
1	D	470	Total	C	N	O	S	0	0	0
			3841	2462	666	693	20			

There are 144 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	MET	-	expression tag	UNP Q9Y3Z3
A	78	GLY	-	expression tag	UNP Q9Y3Z3
A	79	SER	-	expression tag	UNP Q9Y3Z3
A	80	SER	-	expression tag	UNP Q9Y3Z3
A	81	HIS	-	expression tag	UNP Q9Y3Z3
A	82	HIS	-	expression tag	UNP Q9Y3Z3
A	83	HIS	-	expression tag	UNP Q9Y3Z3
A	84	HIS	-	expression tag	UNP Q9Y3Z3
A	85	HIS	-	expression tag	UNP Q9Y3Z3
A	86	HIS	-	expression tag	UNP Q9Y3Z3
A	87	SER	-	expression tag	UNP Q9Y3Z3
A	88	SER	-	expression tag	UNP Q9Y3Z3
A	89	GLY	-	expression tag	UNP Q9Y3Z3
A	90	LEU	-	expression tag	UNP Q9Y3Z3
A	91	VAL	-	expression tag	UNP Q9Y3Z3
A	92	PRO	-	expression tag	UNP Q9Y3Z3
A	93	ARG	-	expression tag	UNP Q9Y3Z3
A	94	GLY	-	expression tag	UNP Q9Y3Z3
A	95	SER	-	expression tag	UNP Q9Y3Z3
A	96	HIS	-	expression tag	UNP Q9Y3Z3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	97	MET	-	expression tag	UNP Q9Y3Z3
A	98	ALA	-	expression tag	UNP Q9Y3Z3
A	99	SER	-	expression tag	UNP Q9Y3Z3
A	100	MET	-	expression tag	UNP Q9Y3Z3
A	101	THR	-	expression tag	UNP Q9Y3Z3
A	102	GLY	-	expression tag	UNP Q9Y3Z3
A	103	GLY	-	expression tag	UNP Q9Y3Z3
A	104	GLN	-	expression tag	UNP Q9Y3Z3
A	105	GLN	-	expression tag	UNP Q9Y3Z3
A	106	MET	-	expression tag	UNP Q9Y3Z3
A	107	GLY	-	expression tag	UNP Q9Y3Z3
A	108	ARG	-	expression tag	UNP Q9Y3Z3
A	109	ASP	-	expression tag	UNP Q9Y3Z3
A	110	PRO	-	expression tag	UNP Q9Y3Z3
A	111	ASN	-	expression tag	UNP Q9Y3Z3
A	112	SER	-	expression tag	UNP Q9Y3Z3
B	77	MET	-	expression tag	UNP Q9Y3Z3
B	78	GLY	-	expression tag	UNP Q9Y3Z3
B	79	SER	-	expression tag	UNP Q9Y3Z3
B	80	SER	-	expression tag	UNP Q9Y3Z3
B	81	HIS	-	expression tag	UNP Q9Y3Z3
B	82	HIS	-	expression tag	UNP Q9Y3Z3
B	83	HIS	-	expression tag	UNP Q9Y3Z3
B	84	HIS	-	expression tag	UNP Q9Y3Z3
B	85	HIS	-	expression tag	UNP Q9Y3Z3
B	86	HIS	-	expression tag	UNP Q9Y3Z3
B	87	SER	-	expression tag	UNP Q9Y3Z3
B	88	SER	-	expression tag	UNP Q9Y3Z3
B	89	GLY	-	expression tag	UNP Q9Y3Z3
B	90	LEU	-	expression tag	UNP Q9Y3Z3
B	91	VAL	-	expression tag	UNP Q9Y3Z3
B	92	PRO	-	expression tag	UNP Q9Y3Z3
B	93	ARG	-	expression tag	UNP Q9Y3Z3
B	94	GLY	-	expression tag	UNP Q9Y3Z3
B	95	SER	-	expression tag	UNP Q9Y3Z3
B	96	HIS	-	expression tag	UNP Q9Y3Z3
B	97	MET	-	expression tag	UNP Q9Y3Z3
B	98	ALA	-	expression tag	UNP Q9Y3Z3
B	99	SER	-	expression tag	UNP Q9Y3Z3
B	100	MET	-	expression tag	UNP Q9Y3Z3
B	101	THR	-	expression tag	UNP Q9Y3Z3
B	102	GLY	-	expression tag	UNP Q9Y3Z3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	103	GLY	-	expression tag	UNP Q9Y3Z3
B	104	GLN	-	expression tag	UNP Q9Y3Z3
B	105	GLN	-	expression tag	UNP Q9Y3Z3
B	106	MET	-	expression tag	UNP Q9Y3Z3
B	107	GLY	-	expression tag	UNP Q9Y3Z3
B	108	ARG	-	expression tag	UNP Q9Y3Z3
B	109	ASP	-	expression tag	UNP Q9Y3Z3
B	110	PRO	-	expression tag	UNP Q9Y3Z3
B	111	ASN	-	expression tag	UNP Q9Y3Z3
B	112	SER	-	expression tag	UNP Q9Y3Z3
C	77	MET	-	expression tag	UNP Q9Y3Z3
C	78	GLY	-	expression tag	UNP Q9Y3Z3
C	79	SER	-	expression tag	UNP Q9Y3Z3
C	80	SER	-	expression tag	UNP Q9Y3Z3
C	81	HIS	-	expression tag	UNP Q9Y3Z3
C	82	HIS	-	expression tag	UNP Q9Y3Z3
C	83	HIS	-	expression tag	UNP Q9Y3Z3
C	84	HIS	-	expression tag	UNP Q9Y3Z3
C	85	HIS	-	expression tag	UNP Q9Y3Z3
C	86	HIS	-	expression tag	UNP Q9Y3Z3
C	87	SER	-	expression tag	UNP Q9Y3Z3
C	88	SER	-	expression tag	UNP Q9Y3Z3
C	89	GLY	-	expression tag	UNP Q9Y3Z3
C	90	LEU	-	expression tag	UNP Q9Y3Z3
C	91	VAL	-	expression tag	UNP Q9Y3Z3
C	92	PRO	-	expression tag	UNP Q9Y3Z3
C	93	ARG	-	expression tag	UNP Q9Y3Z3
C	94	GLY	-	expression tag	UNP Q9Y3Z3
C	95	SER	-	expression tag	UNP Q9Y3Z3
C	96	HIS	-	expression tag	UNP Q9Y3Z3
C	97	MET	-	expression tag	UNP Q9Y3Z3
C	98	ALA	-	expression tag	UNP Q9Y3Z3
C	99	SER	-	expression tag	UNP Q9Y3Z3
C	100	MET	-	expression tag	UNP Q9Y3Z3
C	101	THR	-	expression tag	UNP Q9Y3Z3
C	102	GLY	-	expression tag	UNP Q9Y3Z3
C	103	GLY	-	expression tag	UNP Q9Y3Z3
C	104	GLN	-	expression tag	UNP Q9Y3Z3
C	105	GLN	-	expression tag	UNP Q9Y3Z3
C	106	MET	-	expression tag	UNP Q9Y3Z3
C	107	GLY	-	expression tag	UNP Q9Y3Z3
C	108	ARG	-	expression tag	UNP Q9Y3Z3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	109	ASP	-	expression tag	UNP Q9Y3Z3
C	110	PRO	-	expression tag	UNP Q9Y3Z3
C	111	ASN	-	expression tag	UNP Q9Y3Z3
C	112	SER	-	expression tag	UNP Q9Y3Z3
D	77	MET	-	expression tag	UNP Q9Y3Z3
D	78	GLY	-	expression tag	UNP Q9Y3Z3
D	79	SER	-	expression tag	UNP Q9Y3Z3
D	80	SER	-	expression tag	UNP Q9Y3Z3
D	81	HIS	-	expression tag	UNP Q9Y3Z3
D	82	HIS	-	expression tag	UNP Q9Y3Z3
D	83	HIS	-	expression tag	UNP Q9Y3Z3
D	84	HIS	-	expression tag	UNP Q9Y3Z3
D	85	HIS	-	expression tag	UNP Q9Y3Z3
D	86	HIS	-	expression tag	UNP Q9Y3Z3
D	87	SER	-	expression tag	UNP Q9Y3Z3
D	88	SER	-	expression tag	UNP Q9Y3Z3
D	89	GLY	-	expression tag	UNP Q9Y3Z3
D	90	LEU	-	expression tag	UNP Q9Y3Z3
D	91	VAL	-	expression tag	UNP Q9Y3Z3
D	92	PRO	-	expression tag	UNP Q9Y3Z3
D	93	ARG	-	expression tag	UNP Q9Y3Z3
D	94	GLY	-	expression tag	UNP Q9Y3Z3
D	95	SER	-	expression tag	UNP Q9Y3Z3
D	96	HIS	-	expression tag	UNP Q9Y3Z3
D	97	MET	-	expression tag	UNP Q9Y3Z3
D	98	ALA	-	expression tag	UNP Q9Y3Z3
D	99	SER	-	expression tag	UNP Q9Y3Z3
D	100	MET	-	expression tag	UNP Q9Y3Z3
D	101	THR	-	expression tag	UNP Q9Y3Z3
D	102	GLY	-	expression tag	UNP Q9Y3Z3
D	103	GLY	-	expression tag	UNP Q9Y3Z3
D	104	GLN	-	expression tag	UNP Q9Y3Z3
D	105	GLN	-	expression tag	UNP Q9Y3Z3
D	106	MET	-	expression tag	UNP Q9Y3Z3
D	107	GLY	-	expression tag	UNP Q9Y3Z3
D	108	ARG	-	expression tag	UNP Q9Y3Z3
D	109	ASP	-	expression tag	UNP Q9Y3Z3
D	110	PRO	-	expression tag	UNP Q9Y3Z3
D	111	ASN	-	expression tag	UNP Q9Y3Z3
D	112	SER	-	expression tag	UNP Q9Y3Z3

- Molecule 2 is 2'-deoxyguanosine-5'-O-(1-thiotriphosphate) (three-letter code: T8T) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>12</sub>P<sub>3</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mn 1 1	0	0
3	A	1	Total Mn 1 1	0	0
3	D	1	Total Mn 1 1	0	0
3	C	1	Total Mn 1 1	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Mg 1 1	0	0
4	A	1	Total Mg 1 1	0	0
4	D	1	Total Mg 1 1	0	0
4	C	1	Total Mg 1 1	0	0

- Molecule 5 is water.

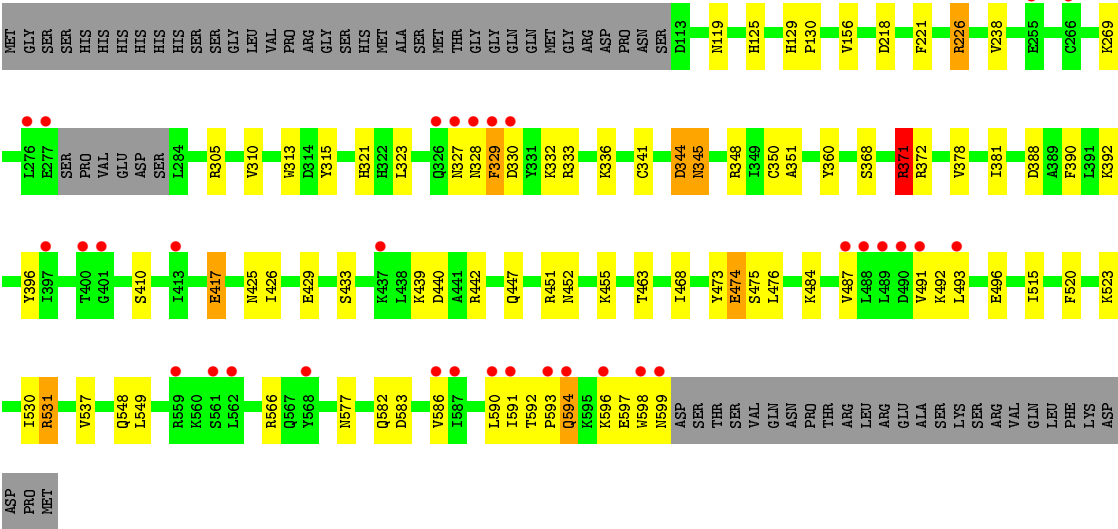
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	14	Total O 14 14	0	0
5	B	9	Total O 9 9	0	0
5	C	7	Total O 7 7	0	0
5	D	4	Total O 4 4	0	0



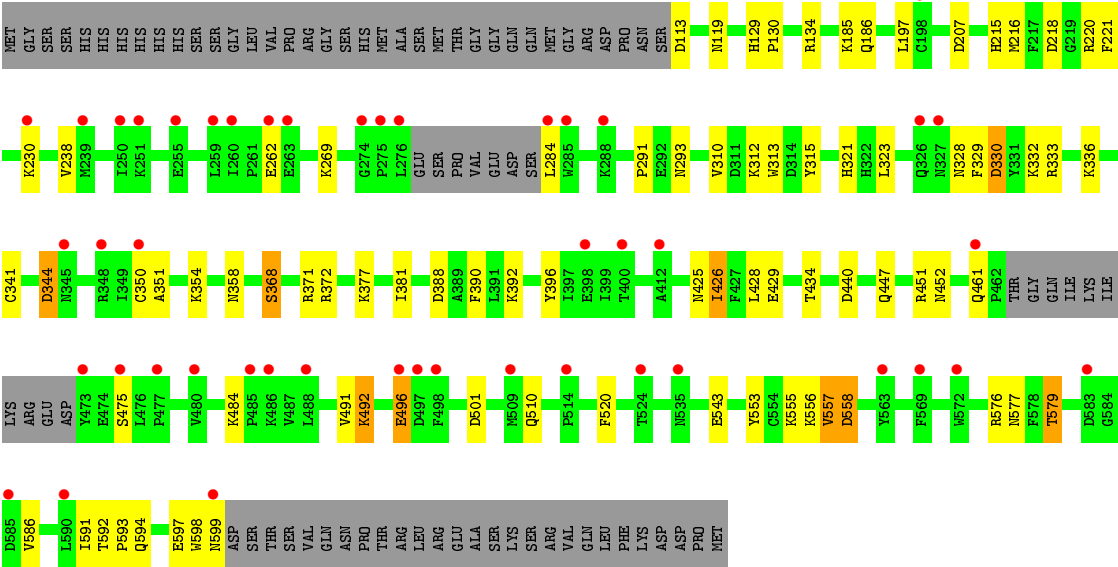


GLN  
ASN  
PRO  
THR  
ARG  
LEU  
ARG  
GLU  
ALA  
SER  
SER  
SER  
SER  
VAL  
GLN  
LEU  
PHE  
LYS  
ASP  
ASP  
PRO  
MET

● Molecule 1: DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHOHYDROLASE SAMHD1



● Molecule 1: DEOXYNUCLEOSIDE TRIPHOSPHATE TRIPHOSPHOHYDROLASE SAMHD1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.00Å 138.73Å 93.72Å 90.00° 113.56° 90.00°	Depositor
Resolution (Å)	49.71 – 2.88 49.66 – 2.88	Depositor EDS
% Data completeness (in resolution range)	77.5 (49.71-2.88) 77.5 (49.66-2.88)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.80 (at 2.86Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.190 , 0.245 0.193 , 0.243	Depositor DCC
$R_{free}$ test set	1668 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.8	Xtriage
Anisotropy	0.319	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 68.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.030 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	15954	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.33% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, T8T, MN

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	0.66	0/3972	0.84	10/5364 (0.2%)
1	B	0.61	0/3978	0.85	9/5369 (0.2%)
1	C	0.65	0/4025	0.82	5/5434 (0.1%)
1	D	0.62	0/3933	0.78	2/5311 (0.0%)
All	All	0.63	0/15908	0.82	26/21478 (0.1%)

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	D	0	1

There are no bond length outliers.

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	451	ARG	NE-CZ-NH1	12.89	126.74	120.30
1	B	451	ARG	NE-CZ-NH2	-12.80	113.90	120.30
1	C	226	ARG	CG-CD-NE	8.26	129.14	111.80
1	B	366	ARG	NE-CZ-NH2	-7.26	116.67	120.30
1	A	531	ARG	NE-CZ-NH1	7.26	123.93	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	330	ASP	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	3880	0	3844	76	0
1	B	3887	0	3872	54	0
1	C	3932	0	3913	59	2
1	D	3841	0	3817	69	2
2	A	93	0	39	3	0
2	B	93	0	39	4	0
2	C	93	0	39	12	0
2	D	93	0	39	6	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	14	0	0	1	0
5	B	9	0	0	1	0
5	C	7	0	0	0	0
5	D	4	0	0	0	0
All	All	15954	0	15602	225	2

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 7.

The worst 5 of 225 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:D:556:LYS:O	1:D:557:VAL:HG22	1.44	1.17
1:D:215:HIS:NE2	2:D:700:T8T:O2A	1.84	1.09
1:A:325:ILE:C	1:C:328:ASN:HB2	1.77	1.04
1:B:402:ALA:O	1:B:405:LYS:HD2	1.60	0.99
1:D:368:SER:OG	1:D:372:ARG:NH1	1.97	0.97

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:493:LEU:O	1:D:396:TYR:OH[1_454]	2.10	0.10
1:C:396:TYR:OH	1:D:230:LYS:NZ[1_454]	2.10	0.10

## 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	470/550 (86%)	464 (99%)	6 (1%)	0	100	100
1	B	471/550 (86%)	466 (99%)	5 (1%)	0	100	100
1	C	477/550 (87%)	471 (99%)	6 (1%)	0	100	100
1	D	464/550 (84%)	458 (99%)	5 (1%)	1 (0%)	47	76
All	All	1882/2200 (86%)	1859 (99%)	22 (1%)	1 (0%)	51	80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	557	VAL

### 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	421/488 (86%)	402 (96%)	19 (4%)	27	59
1	B	421/488 (86%)	400 (95%)	21 (5%)	24	54

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	427/488 (88%)	405 (95%)	22 (5%)	23	53
1	D	417/488 (86%)	398 (95%)	19 (5%)	27	58
All	All	1686/1952 (86%)	1605 (95%)	81 (5%)	25	56

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	491	VAL
1	C	371	ARG
1	D	484	LYS
1	B	496	GLU
1	C	329	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	376	HIS
1	B	425	ASN
1	C	425	ASN
1	B	243	HIS
1	D	235	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 8 are monoatomic - leaving 12 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
2	T8T	C	900	4	24,33,33	1.13	2 (8%)	30,52,52	2.30	8 (26%)
2	T8T	B	800	4	24,33,33	1.39	3 (12%)	30,52,52	2.27	8 (26%)
2	T8T	A	900	4	24,33,33	0.96	1 (4%)	30,52,52	2.25	9 (30%)
2	T8T	D	800	4	24,33,33	1.00	1 (4%)	30,52,52	1.87	7 (23%)
2	T8T	C	800	4	24,33,33	1.22	3 (12%)	30,52,52	1.84	8 (26%)
2	T8T	A	700	-	24,33,33	1.06	1 (4%)	30,52,52	2.17	11 (36%)
2	T8T	B	700	-	24,33,33	1.58	2 (8%)	30,52,52	2.24	10 (33%)
2	T8T	C	700	-	24,33,33	0.97	1 (4%)	30,52,52	2.03	9 (30%)
2	T8T	D	700	-	24,33,33	1.20	3 (12%)	30,52,52	1.95	7 (23%)
2	T8T	A	800	4	24,33,33	0.98	1 (4%)	30,52,52	2.53	12 (40%)
2	T8T	D	900	4	24,33,33	1.10	2 (8%)	30,52,52	2.36	10 (33%)
2	T8T	B	900	4	24,33,33	0.89	1 (4%)	30,52,52	1.69	7 (23%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	T8T	C	900	4	-	5/14/34/34	0/3/3/3
2	T8T	B	800	4	-	3/14/34/34	0/3/3/3
2	T8T	A	900	4	-	3/14/34/34	0/3/3/3
2	T8T	D	800	4	-	1/14/34/34	0/3/3/3
2	T8T	C	800	4	-	2/14/34/34	0/3/3/3
2	T8T	A	700	-	-	5/14/34/34	0/3/3/3
2	T8T	B	700	-	-	6/14/34/34	0/3/3/3
2	T8T	C	700	-	-	3/14/34/34	0/3/3/3
2	T8T	D	700	-	-	4/14/34/34	0/3/3/3
2	T8T	A	800	4	-	4/14/34/34	0/3/3/3
2	T8T	D	900	4	-	2/14/34/34	0/3/3/3

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	T8T	B	900	4	-	1/14/34/34	0/3/3/3

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	700	T8T	C6-C5	5.06	1.50	1.41
2	B	700	T8T	PA-O5'	4.22	1.64	1.57
2	D	700	T8T	C6-C5	3.78	1.47	1.41
2	C	900	T8T	C6-C5	3.73	1.47	1.41
2	B	800	T8T	PA-O5'	3.52	1.63	1.57

The worst 5 of 106 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	900	T8T	C6-C5-C4	-6.41	114.68	120.80
2	A	800	T8T	N2-C2-N1	6.41	127.22	117.25
2	A	900	T8T	C2-N3-C4	6.24	122.48	115.36
2	D	900	T8T	C5-C6-N1	-5.82	115.47	123.43
2	A	800	T8T	C5-C6-N1	-5.45	115.98	123.43

There are no chirality outliers.

5 of 39 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	900	T8T	PB-O3B-PG-O1G
2	C	900	T8T	PB-O3B-PG-O2G
2	B	700	T8T	C3'-C4'-C5'-O5'
2	B	700	T8T	O4'-C4'-C5'-O5'
2	B	700	T8T	C4'-C5'-O5'-PA

There are no ring outliers.

11 monomers are involved in 25 short contacts:

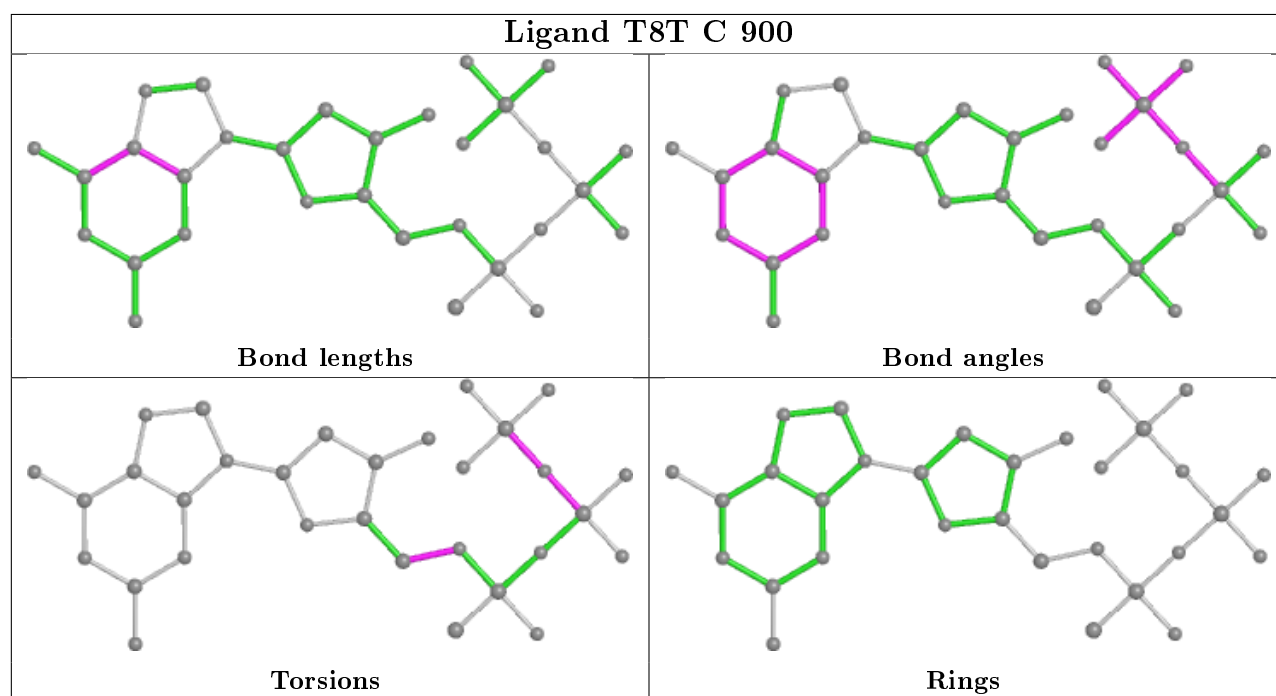
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	900	T8T	5	0
2	B	800	T8T	3	0
2	A	900	T8T	1	0
2	D	800	T8T	2	0
2	C	800	T8T	5	0
2	A	700	T8T	2	0
2	B	700	T8T	1	0

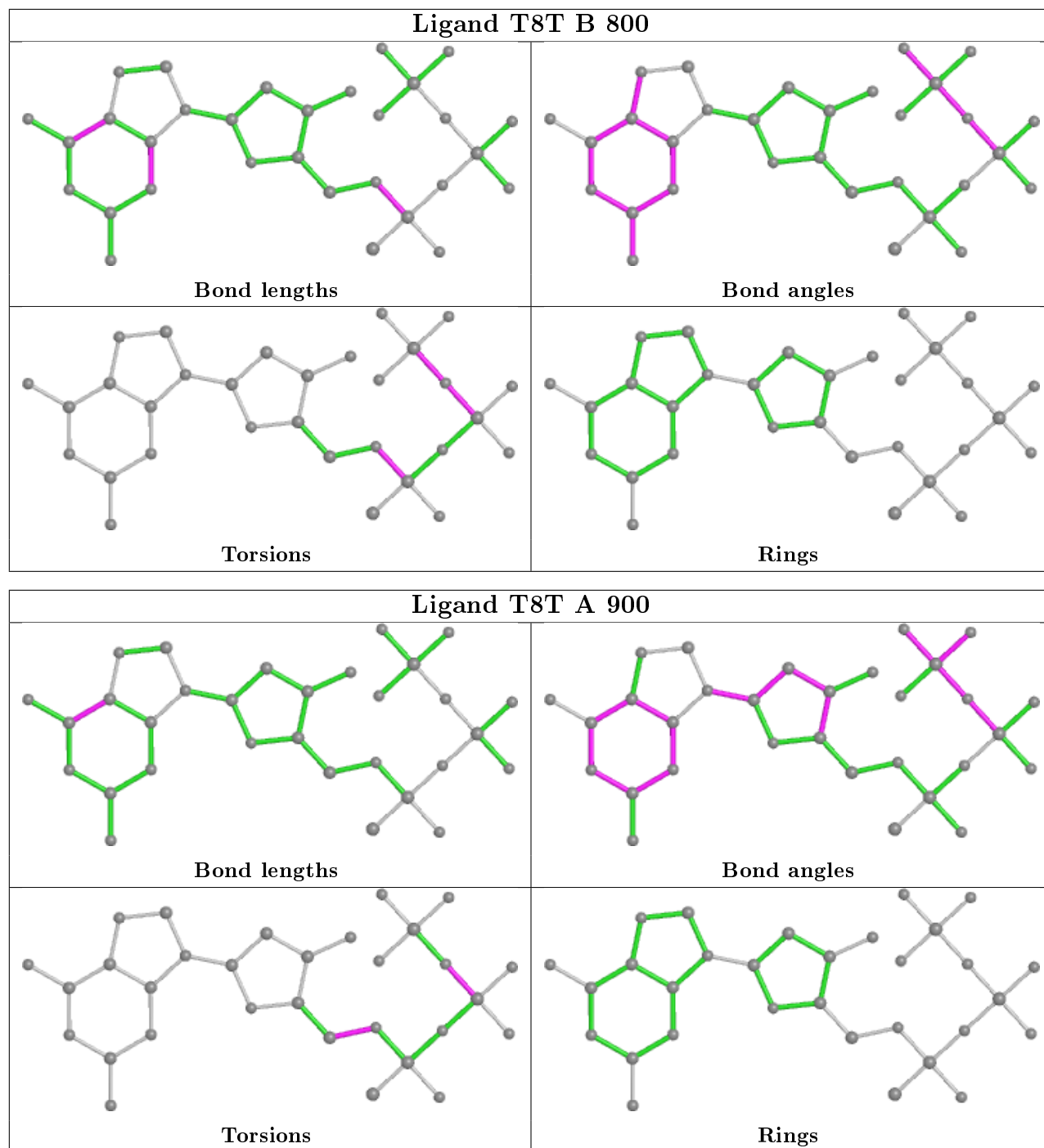
*Continued on next page...*

*Continued from previous page...*

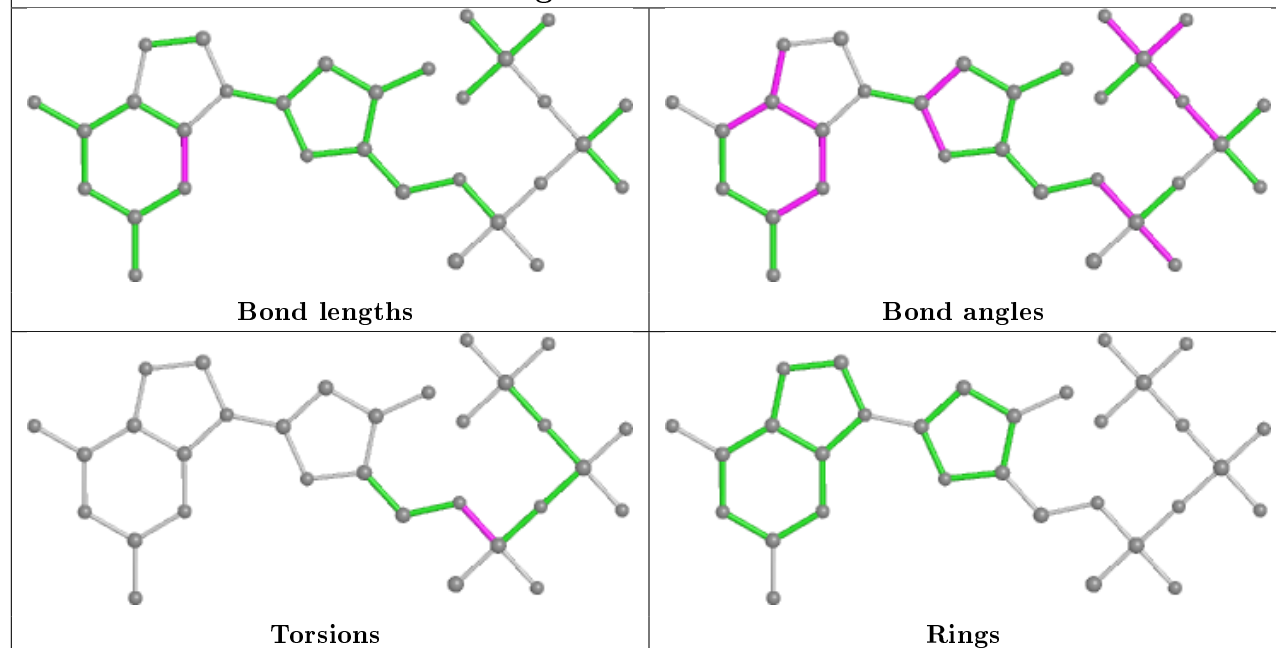
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	700	T8T	3	0
2	D	700	T8T	4	0
2	D	900	T8T	1	0
2	B	900	T8T	1	0

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.

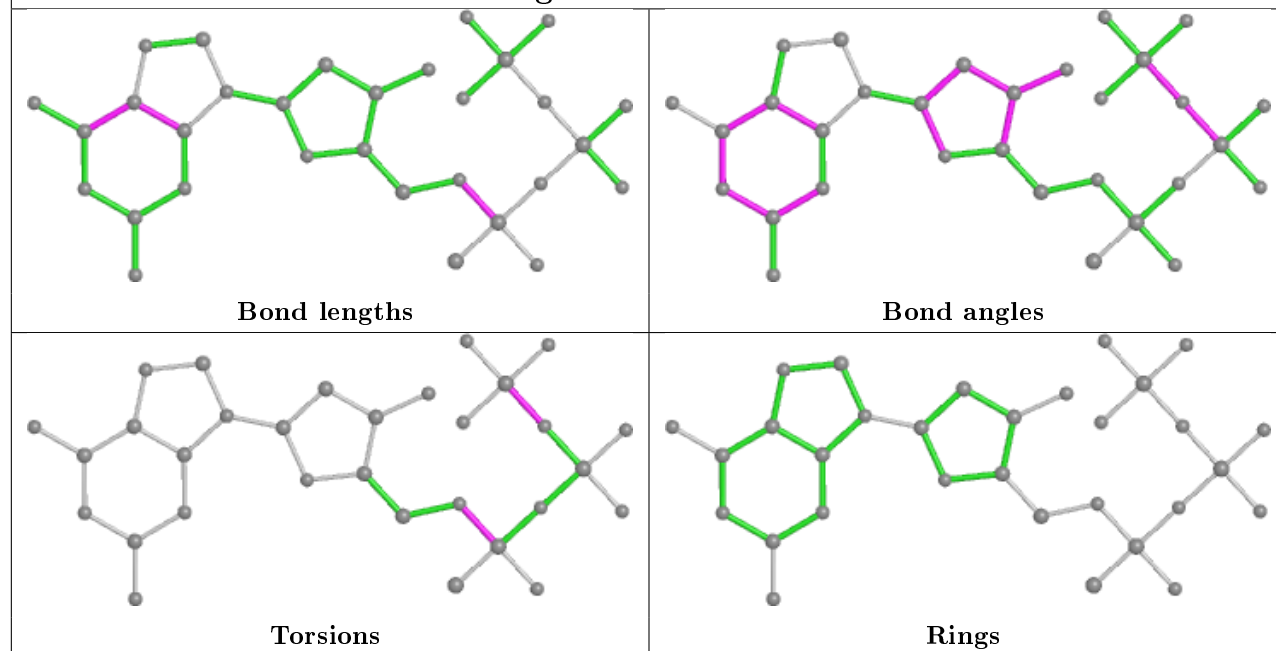


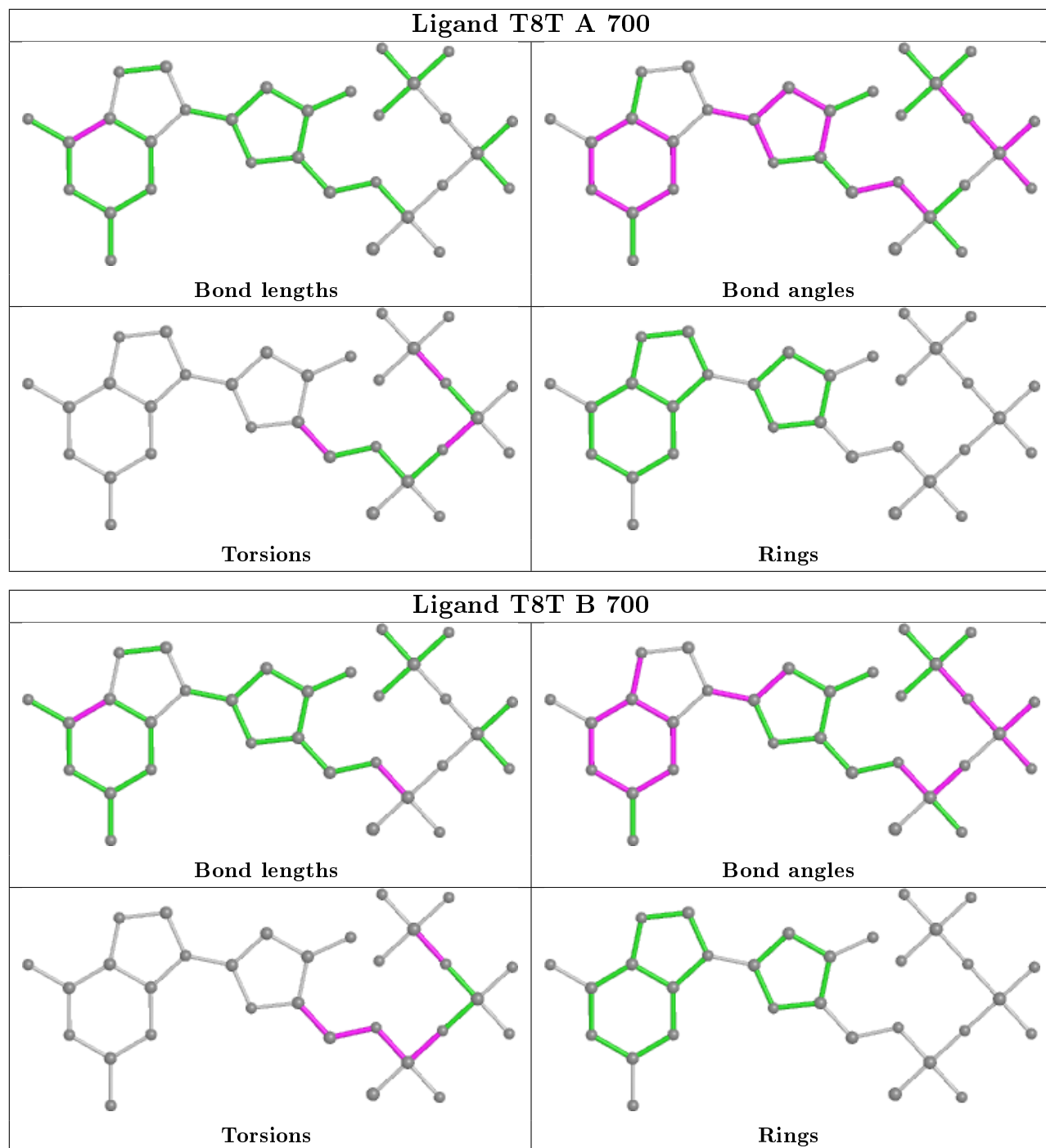


## Ligand T8T D 800

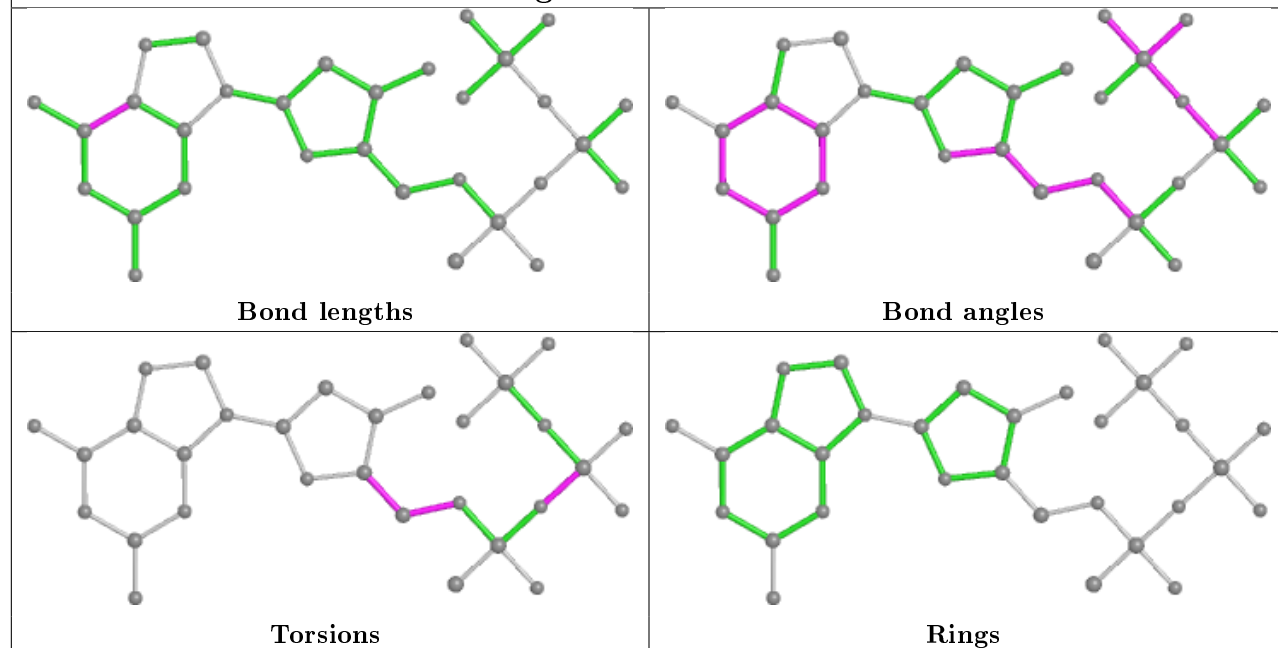


## Ligand T8T C 800

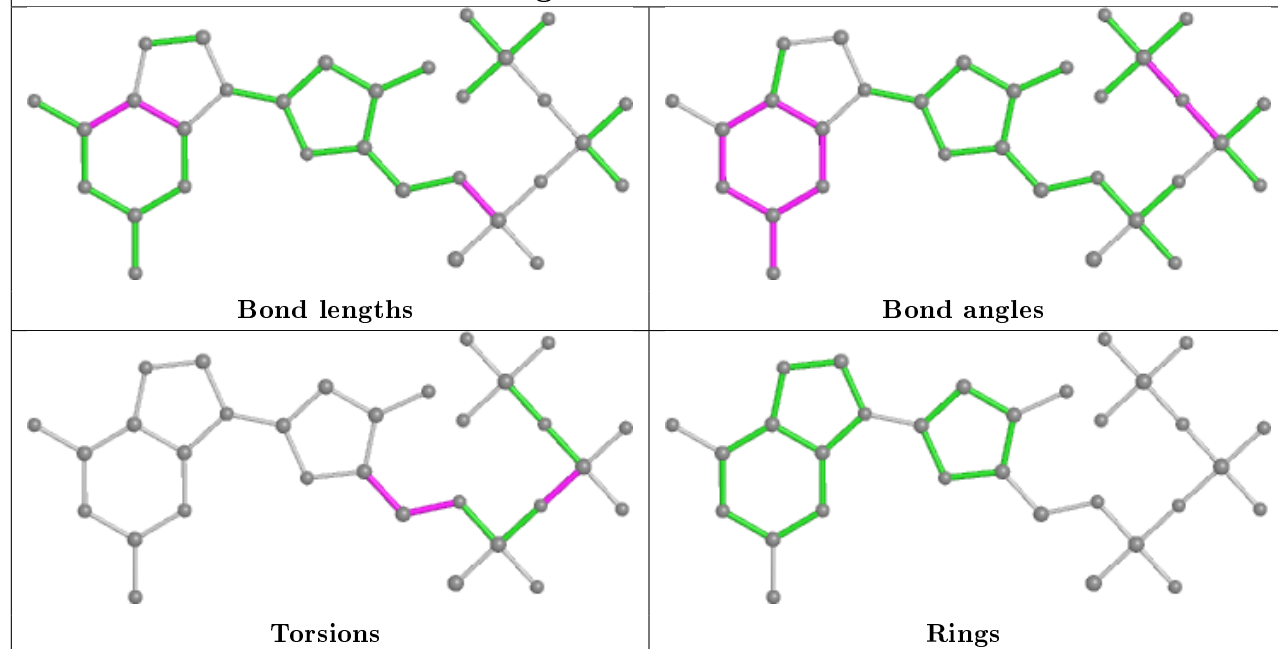




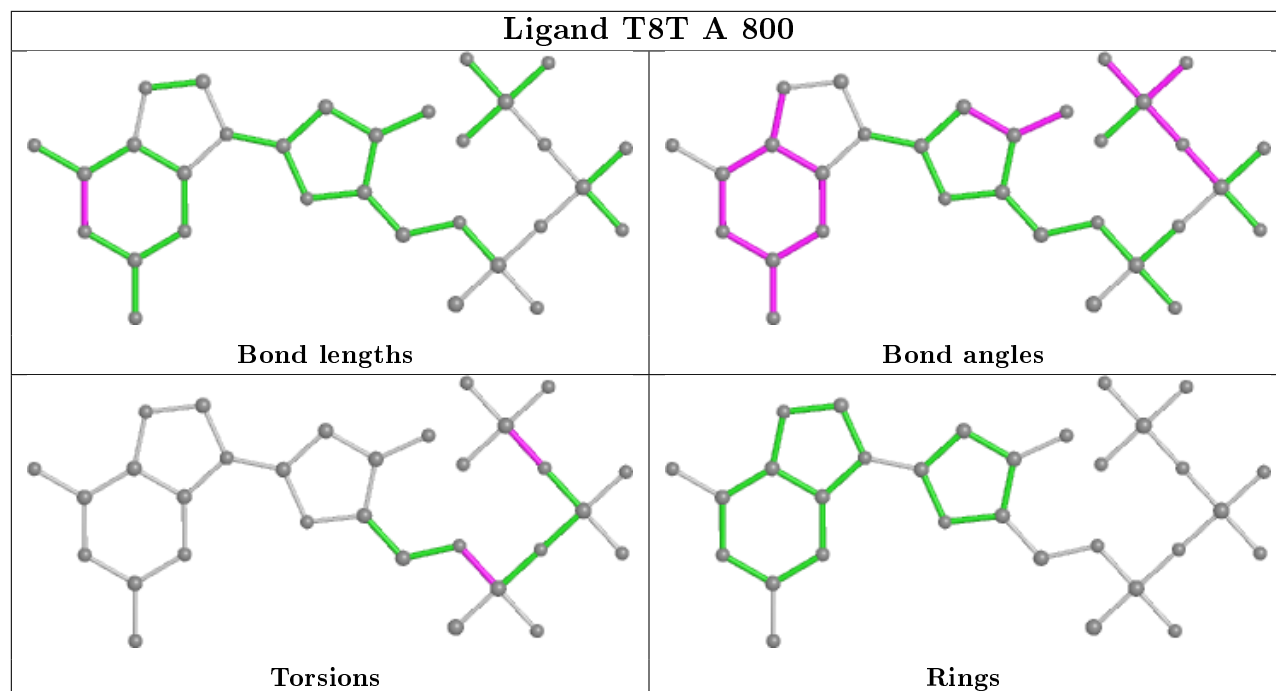
## Ligand T8T C 700



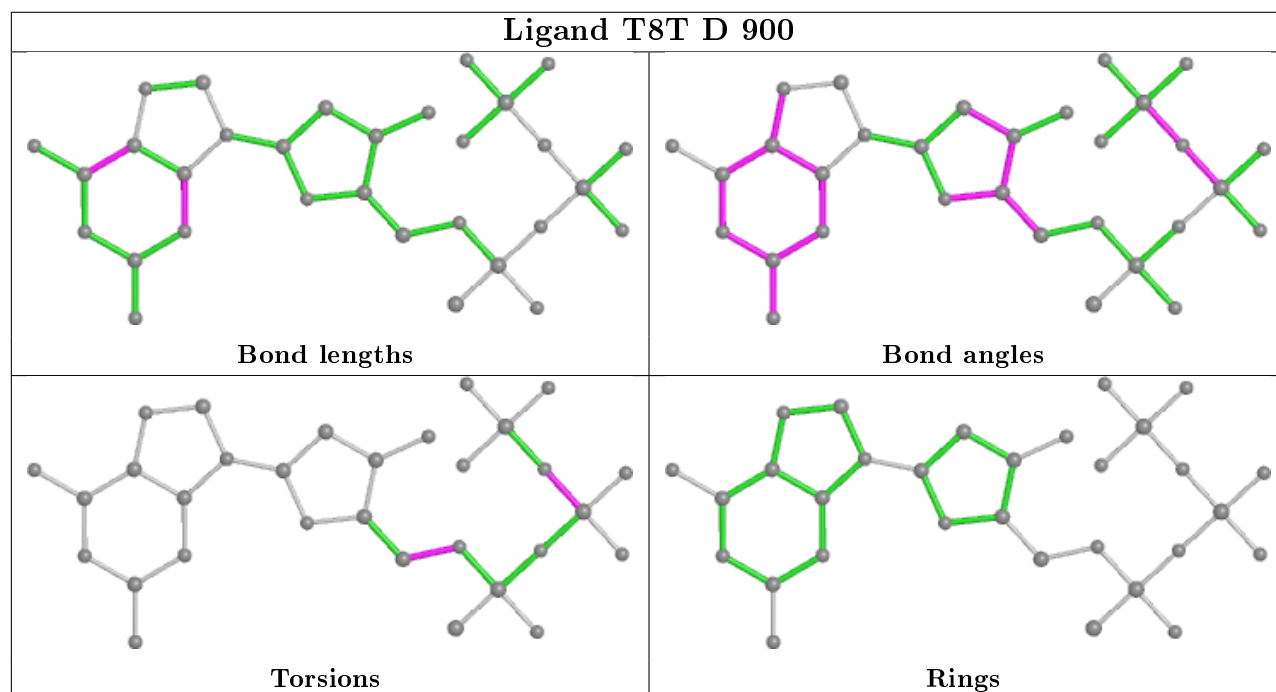
## Ligand T8T D 700

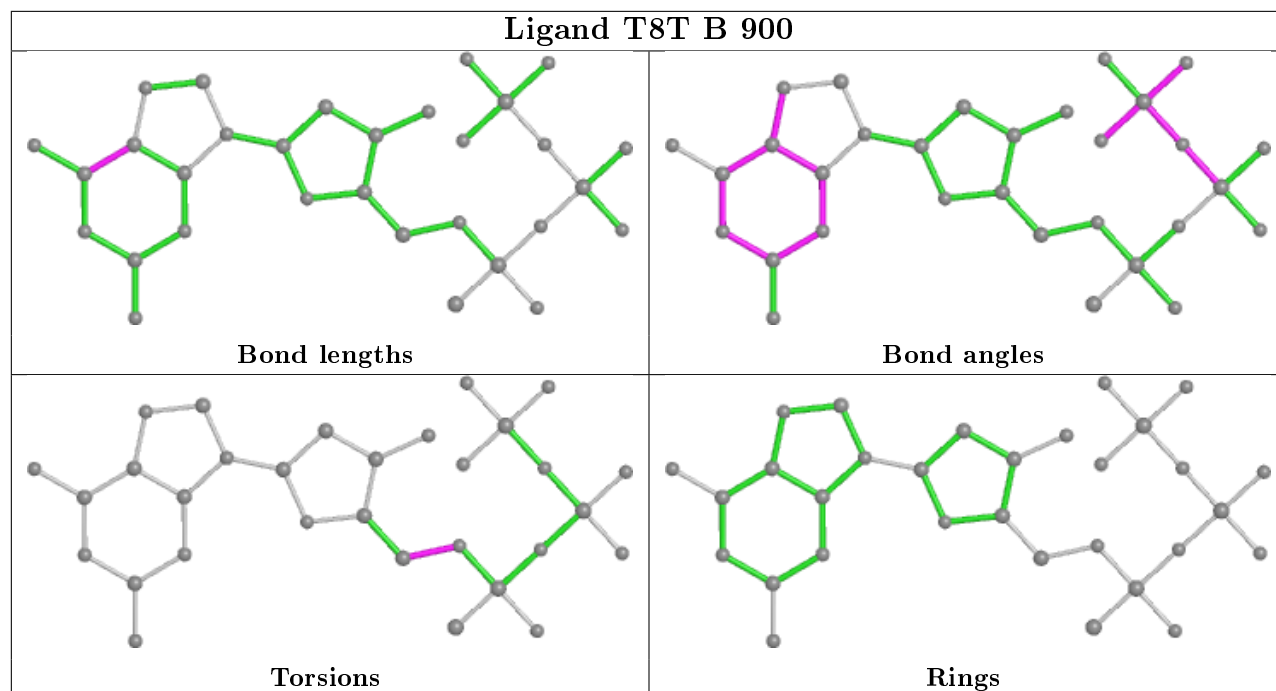


## Ligand T8T A 800



## Ligand T8T D 900





## 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	476/550 (86%)	0.26	22 (4%) 32 28	29, 53, 94, 127	0
1	B	477/550 (86%)	0.74	58 (12%) 4 2	40, 76, 117, 137	0
1	C	481/550 (87%)	0.34	33 (6%) 16 13	28, 56, 102, 140	0
1	D	470/550 (85%)	0.60	46 (9%) 7 5	39, 73, 108, 140	0
All	All	1904/2200 (86%)	0.48	159 (8%) 11 8	28, 65, 107, 140	0

The worst 5 of 159 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	490	ASP	9.2
1	B	488	LEU	6.9
1	B	491	VAL	6.4
1	A	345	ASN	5.7
1	D	412	ALA	5.5

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

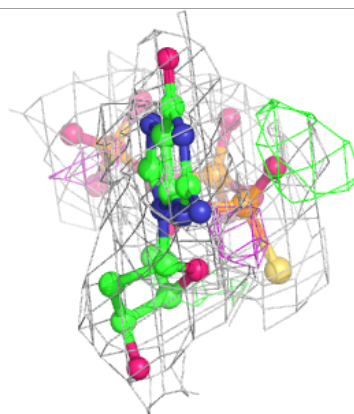
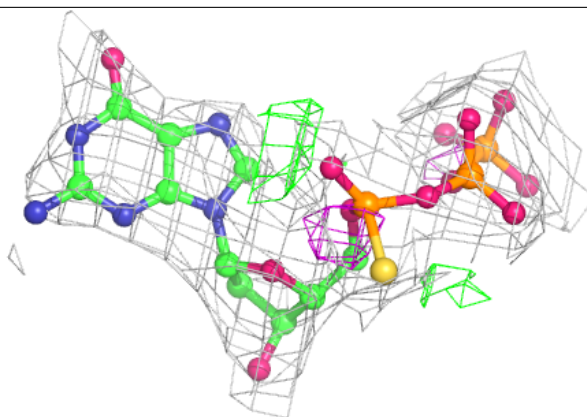
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 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
2	T8T	B	700	31/31	0.93	0.18	59,69,80,85	0
2	T8T	D	700	31/31	0.94	0.15	63,72,80,88	0
2	T8T	A	700	31/31	0.96	0.12	36,41,62,73	0
2	T8T	C	700	31/31	0.97	0.12	42,46,51,59	0
2	T8T	A	900	31/31	0.97	0.13	27,29,31,31	0
4	MG	B	950	1/1	0.97	0.11	36,36,36,36	0
2	T8T	C	800	31/31	0.98	0.11	30,32,36,37	0
2	T8T	D	800	31/31	0.98	0.13	39,43,49,55	0
2	T8T	B	900	31/31	0.98	0.11	36,40,46,47	0
2	T8T	C	900	31/31	0.99	0.10	36,39,41,42	0
2	T8T	B	800	31/31	0.99	0.11	37,40,48,48	0
4	MG	C	950	1/1	0.99	0.10	28,28,28,28	0
4	MG	A	950	1/1	0.99	0.10	25,25,25,25	0
2	T8T	A	800	31/31	0.99	0.13	23,24,25,26	0
3	MN	C	750	1/1	0.99	0.16	25,25,25,25	0
2	T8T	D	900	31/31	0.99	0.10	36,38,42,43	0
4	MG	D	950	1/1	0.99	0.09	34,34,34,34	0
3	MN	B	750	1/1	0.99	0.15	51,51,51,51	0
3	MN	D	750	1/1	1.00	0.15	43,43,43,43	0
3	MN	A	750	1/1	1.00	0.16	25,25,25,25	0

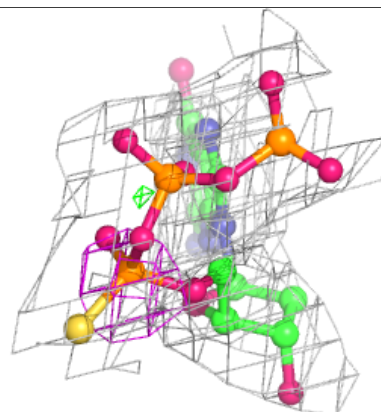
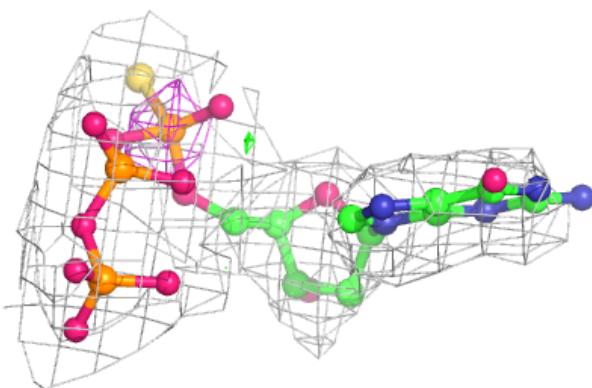
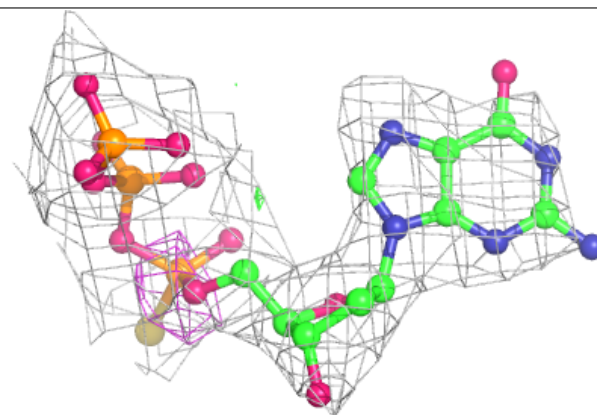
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around T8T B 700:**

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

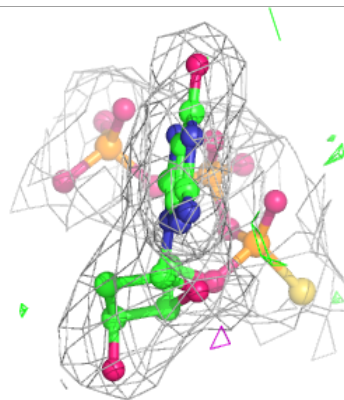
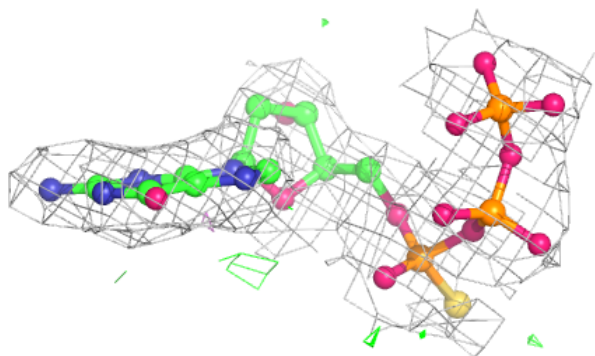
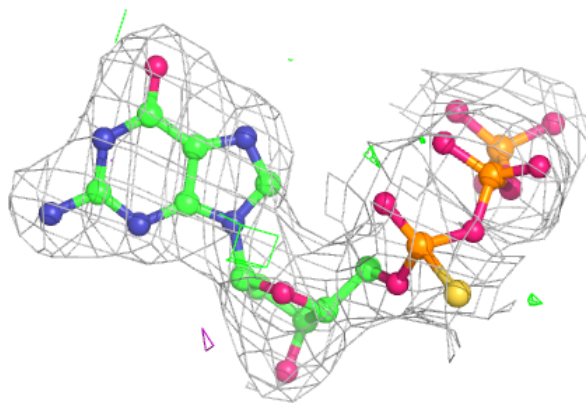
**Electron density around T8T D 700:**

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

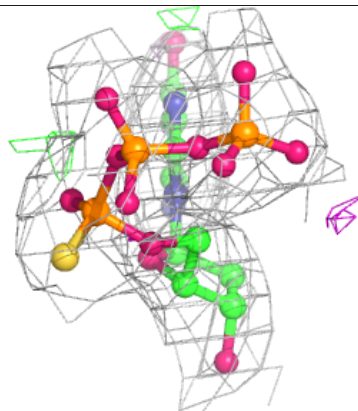
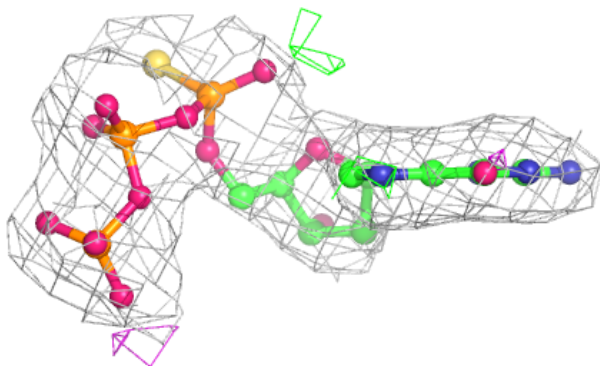
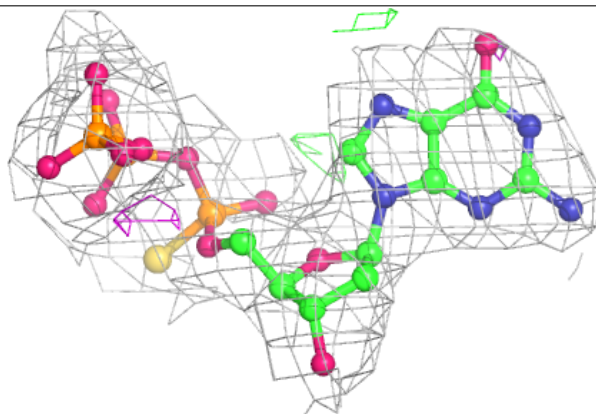


**Electron density around T8T A 700:**

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

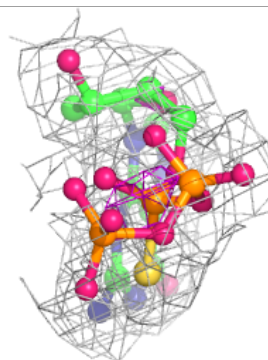
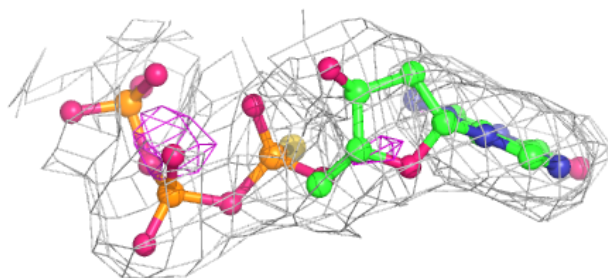
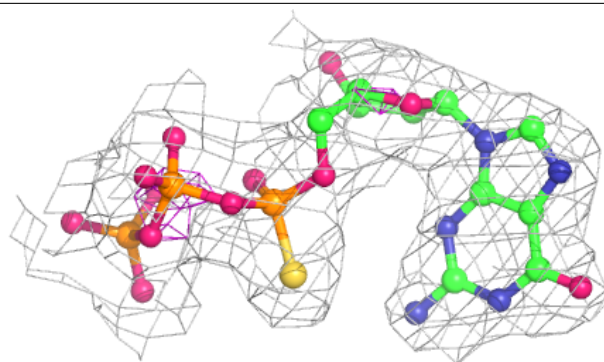
**Electron density around T8T C 700:**

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

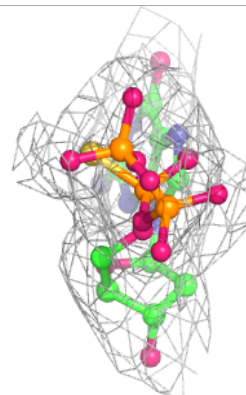
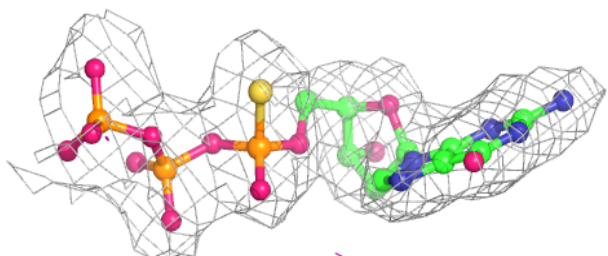
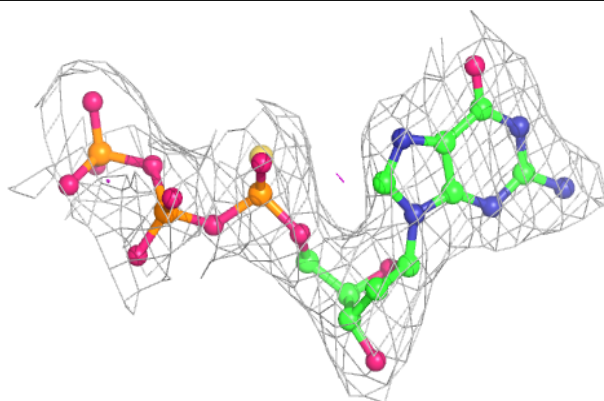


**Electron density around T8T A 900:**

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

**Electron density around T8T C 800:**

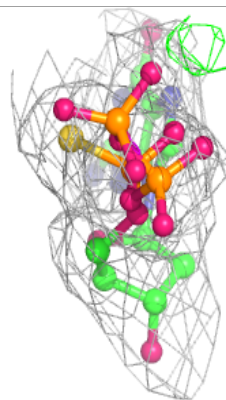
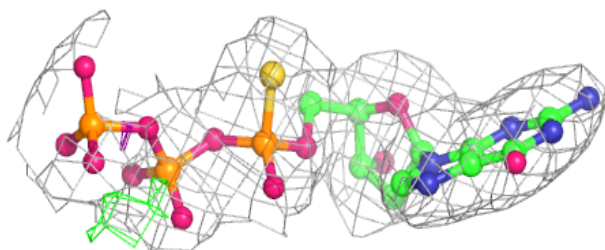
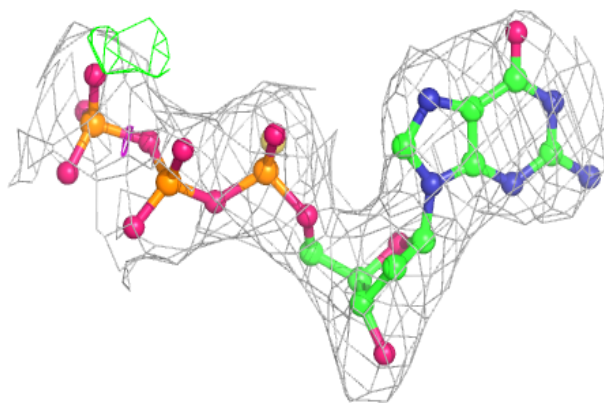
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



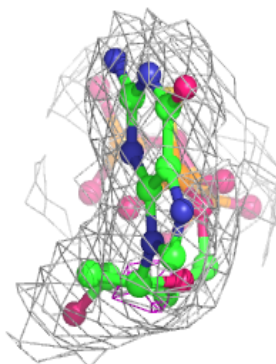
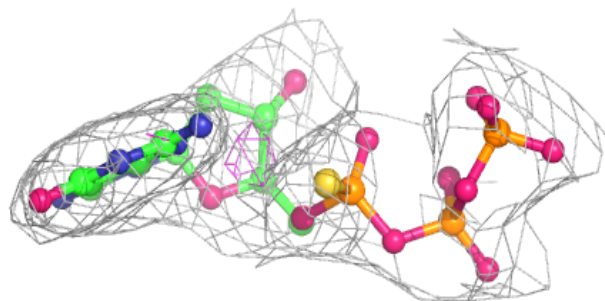
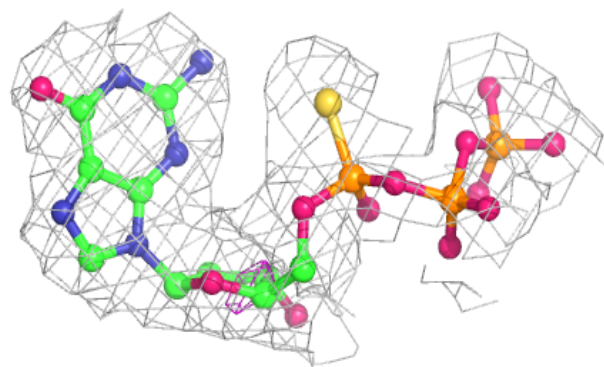


**Electron density around T8T D 800:**

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

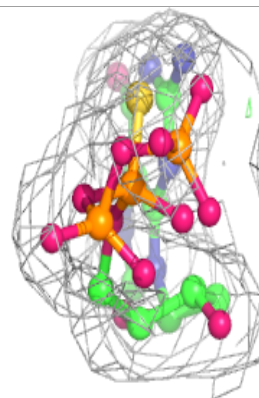
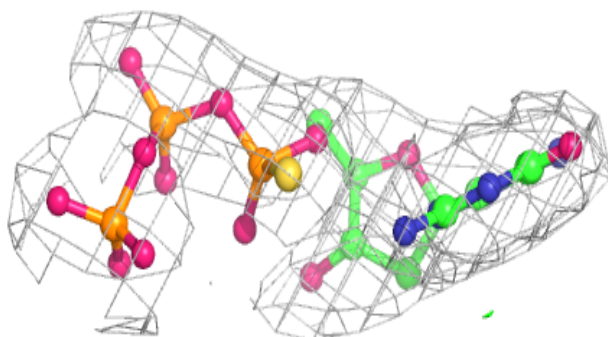
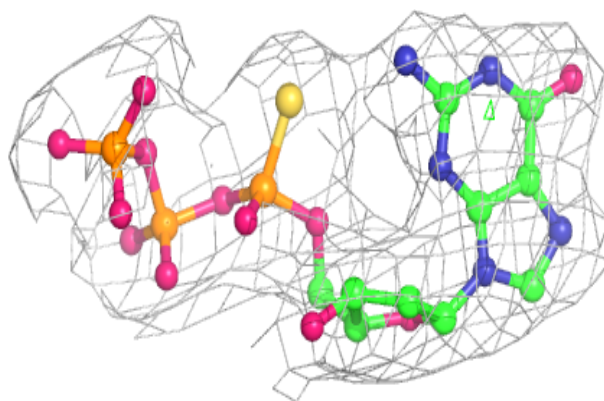
**Electron density around T8T B 900:**

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

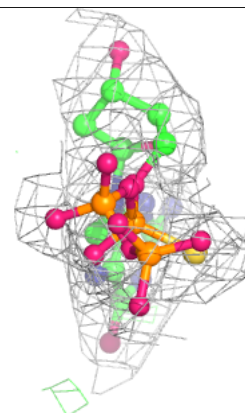
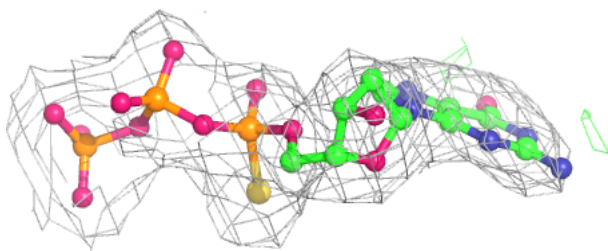
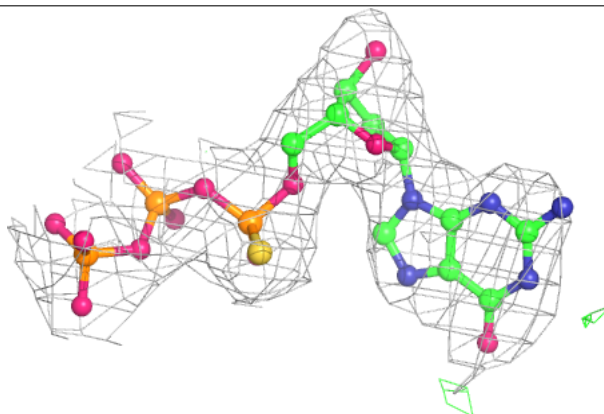


**Electron density around T8T C 900:**

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

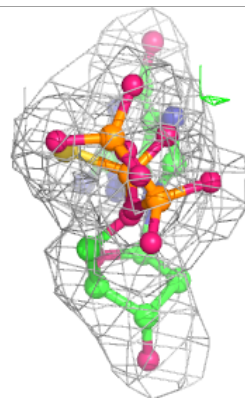
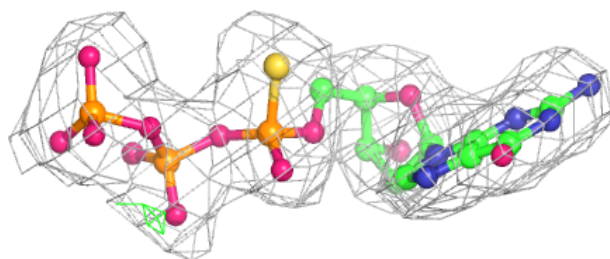
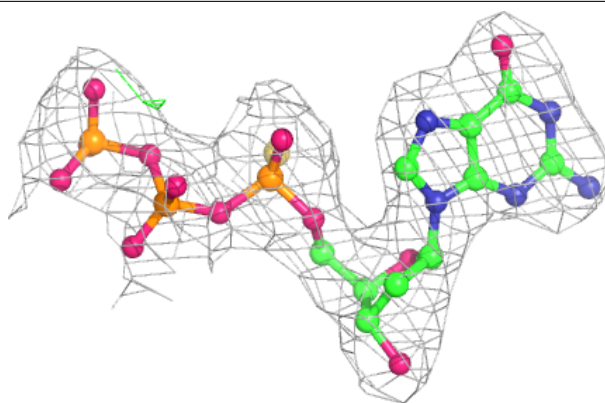
**Electron density around T8T B 800:**

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

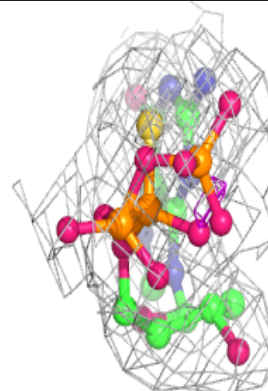
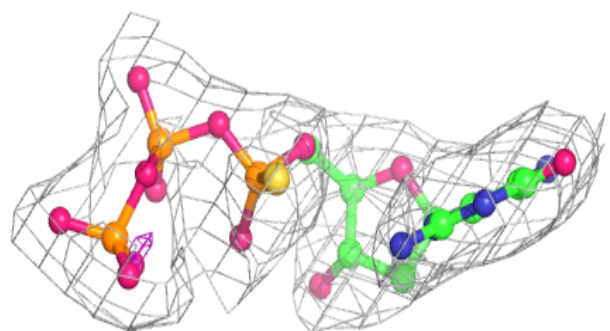
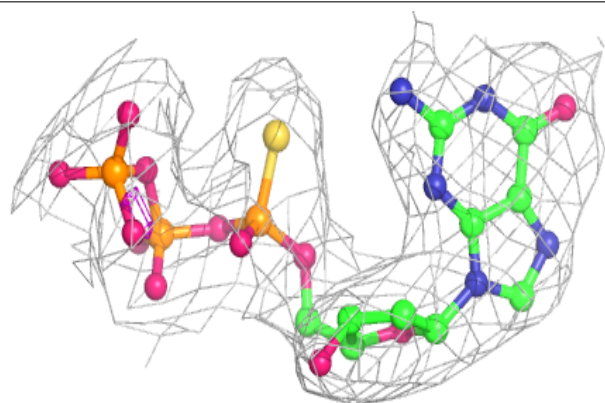


**Electron density around T8T A 800:**

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

**Electron density around T8T D 900:**

$2mF_o-DF_c$  (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.