



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 29, 2025 – 02:01 AM EDT

PDB ID : 2PYJ / pdb_00002pyj
Title : Phi29 DNA polymerase complexed with primer-template DNA and incoming nucleotide substrates (ternary complex)
Authors : Berman, A.J.; Kamtekar, S.; Goodman, J.L.; Lazaro, J.M.; de Vega, M.; Blanco, L.; Salas, M.; Steitz, T.A.
Deposited on : 2007-05-16
Resolution : 2.03 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

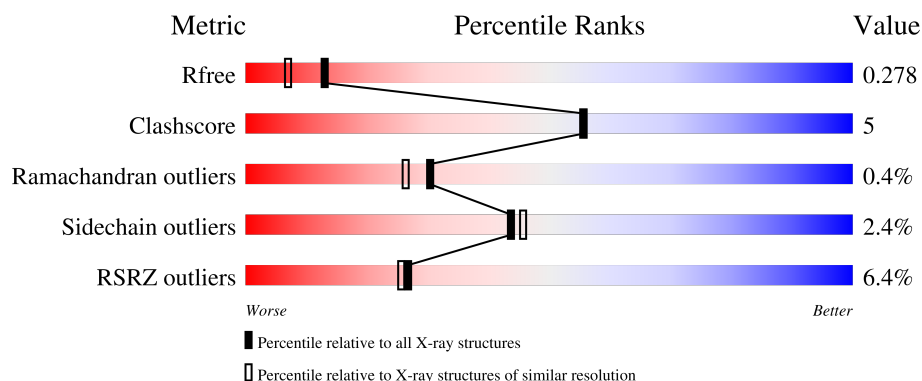
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.03 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	12358 (2.04-2.00)
Clashscore	180529	13897 (2.04-2.00)
Ramachandran outliers	177936	13770 (2.04-2.00)
Sidechain outliers	177891	13769 (2.04-2.00)
RSRZ outliers	164620	12358 (2.04-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	J	10	<div><div></div><div>50%</div><div>50%</div></div>
1	Q	10	<div><div></div><div>60%</div><div>40%</div></div>
1	X	10	<div><div></div><div>70%</div><div>30%</div></div>
2	K	14	<div><div></div><div>79%</div><div>21%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	R	14	<div><div></div><div>93%</div><div>7%</div></div>
2	Y	14	<div><div>7%</div><div>86%</div><div>14%</div></div>
3	A	575	<div><div>2%</div><div>89%</div><div>9%</div><div></div></div>
3	B	575	<div><div>11%</div><div>84%</div><div>13%</div><div></div></div>

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 11997 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 DNA chain called 5'-d(GACTGCTTA(DOC)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	10	Total	C	N	O	P	0	0	0
			199	97	35	58	9			
1	Q	10	Total	C	N	O	P	0	0	0
			199	97	35	58	9			
1	J	10	Total	C	N	O	P	0	0	0
			199	97	35	58	9			

- Molecule 2 is a DNA chain called 5'-d(ACACGTAAGCAGTC)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Y	14	Total	C	N	O	P	0	0	0
			284	136	56	79	13			
2	R	13	Total	C	N	O	P	0	0	0
			266	126	51	76	13			
2	K	14	Total	C	N	O	P	0	0	0
			284	136	56	79	13			

- Molecule 3 is a protein called DNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	567	Total	C	N	O	S	0	4	0
			4660	3041	751	847	21			
3	B	572	Total	C	N	O	S	0	5	0
			4707	3071	758	856	22			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	ALA	ASP	engineered mutation	UNP P03680
A	66	ALA	ASP	engineered mutation	UNP P03680
B	12	ALA	ASP	engineered mutation	UNP P03680
B	66	ALA	ASP	engineered mutation	UNP P03680

- Molecule 4 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	X	1	Total	C	O	0	0
			4	2	2		
4	X	1	Total	C	O	0	0
			4	2	2		
4	Y	1	Total	C	O	0	0
			4	2	2		
4	Y	1	Total	C	O	0	0
			4	2	2		
4	Q	1	Total	C	O	0	0
			4	2	2		
4	R	1	Total	C	O	0	0
			4	2	2		
4	K	1	Total	C	O	0	0
			4	2	2		
4	K	1	Total	C	O	0	0
			4	2	2		
4	K	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		

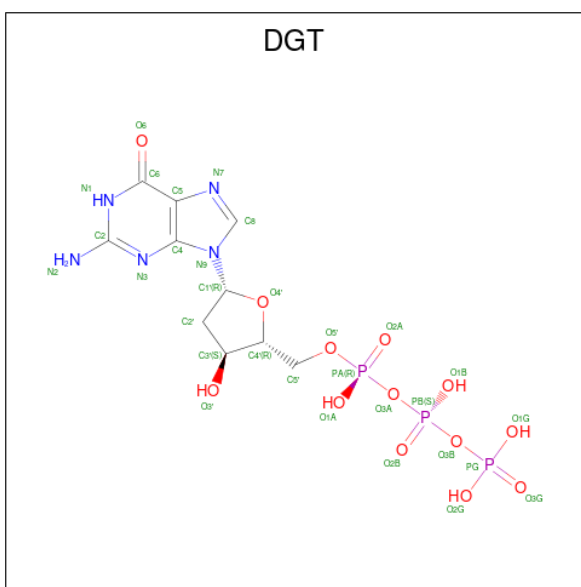
- Molecule 5 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mn	0	0
			1	1		
5	B	1	Total	Mn	0	0
			1	1		

- Molecule 6 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Mg	0	0
			1	1		
6	B	1	Total	Mg	0	0
			1	1		

- Molecule 7 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (CCD ID: DGT) (formula: C₁₀H₁₆N₅O₁₃P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total 31	C 10	N 5	O 13	P 3	0	0
7	B	1	Total 31	C 10	N 5	O 13	P 3	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	X	44	Total O 44 44	0	0
8	Y	45	Total O 45 45	0	0
8	Q	15	Total O 15 15	0	0
8	R	26	Total O 26 26	0	0
8	K	28	Total O 28 28	0	0
8	J	5	Total O 5 5	0	0
8	A	463	Total O 463 463	0	0
8	B	343	Total O 343 343	0	0

3 Residue-property plots [i](#)

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

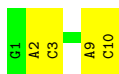
- Molecule 1: 5'-d(GACTGCTTA(DOC)-3'

Chain X: 



- Molecule 1: 5'-d(GACTGCTTA(DOC)-3'

Chain Q: 




- Molecule 1: 5'-d(GACTGCTTA(DOC)-3'

Chain J: 



- Molecule 2: 5'-d(ACACGTAAGCAGTC)-3'

Chain Y: 




- Molecule 2: 5'-d(ACACGTAAGCAGTC)-3'

Chain R: 




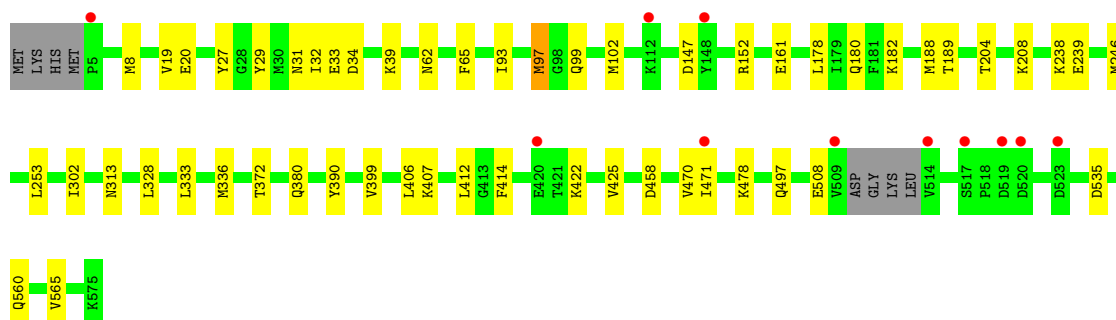
- Molecule 2: 5'-d(ACACGTAAGCAGTC)-3'

Chain K:  79% 21%




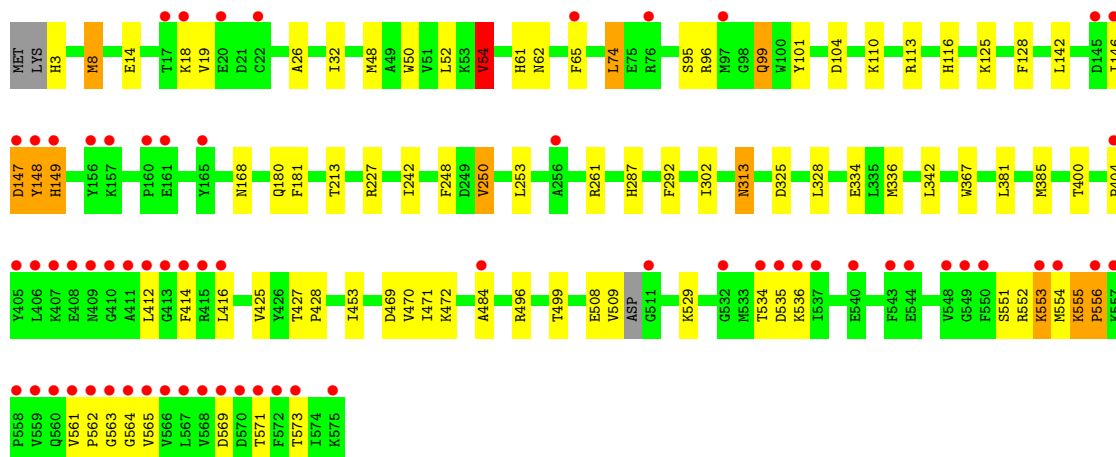
• Molecule 3: DNA polymerase

Chain A:  2% 89% 9%



• Molecule 3: DNA polymerase

Chain B:  11% 84% 13%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	72.83Å 114.67Å 104.76Å 90.00° 94.07° 90.00°	Depositor
Resolution (Å)	41.07 – 2.03 41.07 – 2.03	Depositor EDS
% Data completeness (in resolution range)	97.7 (41.07-2.03) 97.6 (41.07-2.03)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.38 (at 2.03Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, R_{free}	0.189 , 0.234 0.242 , 0.278	Depositor DCC
R_{free} test set	5354 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	29.8	Xtriage
Anisotropy	0.204	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 34.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11997	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.88% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DGT, EDO, DOC, MN, 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	J	0.54	0/202	0.98	2/310 (0.6%)
1	Q	0.59	0/202	0.79	0/310
1	X	0.57	0/202	1.02	1/310 (0.3%)
2	K	0.40	0/319	0.86	1/490 (0.2%)
2	R	0.46	0/298	0.83	0/457
2	Y	0.54	0/319	0.96	1/490 (0.2%)
3	A	0.68	0/4791	0.72	1/6463 (0.0%)
3	B	0.71	2/4843 (0.0%)	0.76	4/6532 (0.1%)
All	All	0.67	2/11176 (0.0%)	0.77	10/15362 (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
3	B	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	250	VAL	C-O	5.78	1.30	1.24
3	B	553	LYS	N-CA	5.70	1.49	1.46

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	553	LYS	N-CA-C	8.01	120.12	108.34
2	Y	3	DA	C5'-C4'-O4'	6.88	119.72	109.40
2	K	4	DA	C5'-C4'-O4'	5.85	118.18	109.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	9	DA	C4'-C3'-O3'	-5.61	101.59	110.00
3	B	556	PRO	N-CA-C	5.52	119.29	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	552	ARG	Peptide
3	B	555	LYS	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	J	199	0	115	1	0
1	Q	199	0	115	2	0
1	X	199	0	115	1	0
2	K	284	0	158	1	0
2	R	266	0	146	0	0
2	Y	284	0	158	1	0
3	A	4660	0	4685	32	0
3	B	4707	0	4722	60	0
4	A	68	0	102	1	0
4	B	60	0	90	2	0
4	K	12	0	18	0	0
4	Q	4	0	6	1	0
4	R	4	0	6	0	0
4	X	8	0	12	4	0
4	Y	8	0	12	1	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	31	0	12	0	0
7	B	31	0	12	0	0
8	A	463	0	0	3	0
8	B	343	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	J	5	0	0	0	0
8	K	28	0	0	0	0
8	Q	15	0	0	0	0
8	R	26	0	0	0	0
8	X	44	0	0	1	0
8	Y	45	0	0	0	0
All	All	11997	0	10484	98	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 5.

The worst 5 of 98 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)
3:B:302:ILE:HD11	3:B:336:MET:HE2	1.38	1.05
3:B:554:MET:HB2	3:B:569:ASP:HB3	1.60	0.83
3:B:328:LEU:HD11	3:B:336:MET:HE1	1.64	0.79
3:B:302:ILE:CD1	3:B:336:MET:HE2	2.13	0.76
3:B:328:LEU:CD1	3:B:336:MET:HE1	2.19	0.71

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	
3	A	567/575 (99%)	547 (96%)	18 (3%)	2 (0%)	30	26
3	B	573/575 (100%)	551 (96%)	20 (4%)	2 (0%)	37	34
All	All	1140/1150 (99%)	1098 (96%)	38 (3%)	4 (0%)	30	26

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	62	ASN
3	B	62	ASN
3	A	425	VAL
3	B	425	VAL

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	
3	A	503/506 (99%)	493 (98%)	10 (2%)	50	53
3	B	508/506 (100%)	494 (97%)	14 (3%)	38	39
All	All	1011/1012 (100%)	987 (98%)	24 (2%)	44	46

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	74	LEU
3	B	147	ASP
3	B	125	LYS
3	B	148	TYR
3	A	406	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	168	ASN
3	B	171	GLN
3	B	497	GLN
3	B	183	GLN
3	A	497	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
1	DOC	Q	10	2,1	16,19,20	1.13	1 (6%)	20,26,29	1.13	1 (5%)
1	DOC	J	10	2,1	16,19,20	1.17	1 (6%)	20,26,29	1.14	1 (5%)
1	DOC	X	10	2,1	16,19,20	0.90	1 (6%)	20,26,29	0.97	1 (5%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	DOC	Q	10	2,1	-	2/7/18/19	0/2/2/2
1	DOC	J	10	2,1	-	0/7/18/19	0/2/2/2
1	DOC	X	10	2,1	-	2/7/18/19	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	Q	10	DOC	C6-N1	-2.38	1.32	1.38
1	X	10	DOC	C6-C5	2.34	1.40	1.35
1	J	10	DOC	C6-N1	-2.22	1.32	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	10	DOC	C4'-O4'-C1'	-2.35	107.59	109.81
1	Q	10	DOC	O4'-C1'-C2'	2.14	108.95	106.41
1	X	10	DOC	C3'-C2'-C1'	2.08	105.27	102.87

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	X	10	DOC	C3'-C4'-C5'-O5'
1	X	10	DOC	O4'-C4'-C5'-O5'
1	Q	10	DOC	C3'-C4'-C5'-O5'
1	Q	10	DOC	O4'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 47 ligands modelled in this entry, 4 are monoatomic - leaving 43 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	$\# Z > 2$
4	EDO	B	2777	-	3,3,3	0.42	0	2,2,2	0.30	0
4	EDO	Q	2785	-	3,3,3	0.28	0	2,2,2	0.42	0
4	EDO	B	2791	-	3,3,3	0.51	0	2,2,2	0.27	0
4	EDO	A	2786	-	3,3,3	0.42	0	2,2,2	0.33	0
4	EDO	X	2782	-	3,3,3	0.44	0	2,2,2	0.41	0
4	EDO	A	2797	-	3,3,3	0.42	0	2,2,2	0.38	0
4	EDO	B	2789	-	3,3,3	0.40	0	2,2,2	0.48	0
4	EDO	A	2771	-	3,3,3	0.46	0	2,2,2	0.16	0
4	EDO	B	2775	-	3,3,3	0.62	0	2,2,2	0.04	0
4	EDO	A	2794	-	3,3,3	0.51	0	2,2,2	0.27	0
4	EDO	A	2800	-	3,3,3	0.57	0	2,2,2	0.23	0
4	EDO	B	2802	-	3,3,3	0.40	0	2,2,2	0.23	0
4	EDO	X	2798	-	3,3,3	0.25	0	2,2,2	0.22	0
4	EDO	A	2773	-	3,3,3	0.54	0	2,2,2	0.11	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
7	DGT	A	1588	5,6	29,33,33	1.30	3 (10%)	37,52,52	1.36	5 (13%)
4	EDO	A	2804	-	3,3,3	0.65	0	2,2,2	0.05	0
4	EDO	A	2781	-	3,3,3	0.51	0	2,2,2	0.25	0
4	EDO	B	2769	-	3,3,3	0.27	0	2,2,2	0.41	0
4	EDO	Y	2788	-	3,3,3	0.52	0	2,2,2	0.23	0
4	EDO	B	2779	-	3,3,3	0.52	0	2,2,2	0.22	0
4	EDO	A	2801	-	3,3,3	0.45	0	2,2,2	0.34	0
4	EDO	B	2787	-	3,3,3	0.54	0	2,2,2	0.11	0
4	EDO	B	2776	-	3,3,3	0.50	0	2,2,2	0.20	0
4	EDO	A	2770	-	3,3,3	0.43	0	2,2,2	0.31	0
4	EDO	A	2792	-	3,3,3	0.46	0	2,2,2	0.25	0
4	EDO	B	2763	-	3,3,3	0.43	0	2,2,2	0.31	0
4	EDO	B	2778	-	3,3,3	0.63	0	2,2,2	0.17	0
4	EDO	K	2766	-	3,3,3	0.41	0	2,2,2	0.29	0
4	EDO	Y	2772	-	3,3,3	0.51	0	2,2,2	0.16	0
4	EDO	K	2767	-	3,3,3	0.60	0	2,2,2	0.07	0
4	EDO	B	2762	-	3,3,3	0.49	0	2,2,2	0.21	0
4	EDO	B	2795	-	3,3,3	0.49	0	2,2,2	0.26	0
4	EDO	A	2784	-	3,3,3	0.33	0	2,2,2	0.23	0
7	DGT	B	1589	5,6	29,33,33	1.21	3 (10%)	37,52,52	1.27	3 (8%)
4	EDO	R	2796	-	3,3,3	0.57	0	2,2,2	0.27	0
4	EDO	B	2805	-	3,3,3	0.40	0	2,2,2	0.26	0
4	EDO	A	2780	-	3,3,3	0.41	0	2,2,2	0.32	0
4	EDO	A	2803	-	3,3,3	0.31	0	2,2,2	0.66	0
4	EDO	A	2774	-	3,3,3	0.39	0	2,2,2	0.38	0
4	EDO	A	2793	-	3,3,3	0.54	0	2,2,2	0.24	0
4	EDO	A	2790	-	3,3,3	0.42	0	2,2,2	0.36	0
4	EDO	K	2765	-	3,3,3	0.49	0	2,2,2	0.22	0
4	EDO	B	2764	-	3,3,3	0.40	0	2,2,2	0.33	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
4	EDO	B	2777	-	-	0/1/1/1	-
4	EDO	Q	2785	-	-	0/1/1/1	-
4	EDO	B	2791	-	-	1/1/1/1	-
4	EDO	A	2786	-	-	0/1/1/1	-
4	EDO	X	2782	-	-	0/1/1/1	-
4	EDO	A	2797	-	-	0/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	B	2789	-	-	0/1/1/1	-
4	EDO	A	2771	-	-	1/1/1/1	-
4	EDO	B	2775	-	-	1/1/1/1	-
4	EDO	A	2794	-	-	0/1/1/1	-
4	EDO	A	2800	-	-	1/1/1/1	-
4	EDO	B	2802	-	-	0/1/1/1	-
4	EDO	X	2798	-	-	0/1/1/1	-
4	EDO	A	2773	-	-	1/1/1/1	-
7	DGT	A	1588	5,6	-	2/18/34/34	0/3/3/3
4	EDO	A	2804	-	-	1/1/1/1	-
4	EDO	A	2781	-	-	1/1/1/1	-
4	EDO	B	2769	-	-	0/1/1/1	-
4	EDO	Y	2788	-	-	0/1/1/1	-
4	EDO	B	2779	-	-	0/1/1/1	-
4	EDO	A	2801	-	-	1/1/1/1	-
4	EDO	B	2787	-	-	1/1/1/1	-
4	EDO	B	2776	-	-	0/1/1/1	-
4	EDO	A	2770	-	-	0/1/1/1	-
4	EDO	A	2792	-	-	1/1/1/1	-
4	EDO	B	2763	-	-	0/1/1/1	-
4	EDO	B	2778	-	-	1/1/1/1	-
4	EDO	K	2766	-	-	0/1/1/1	-
4	EDO	Y	2772	-	-	0/1/1/1	-
4	EDO	K	2767	-	-	1/1/1/1	-
4	EDO	B	2762	-	-	1/1/1/1	-
4	EDO	B	2795	-	-	1/1/1/1	-
4	EDO	A	2784	-	-	0/1/1/1	-
7	DGT	B	1589	5,6	-	5/18/34/34	0/3/3/3
4	EDO	R	2796	-	-	1/1/1/1	-
4	EDO	B	2805	-	-	1/1/1/1	-
4	EDO	A	2780	-	-	1/1/1/1	-
4	EDO	A	2803	-	-	1/1/1/1	-
4	EDO	A	2774	-	-	0/1/1/1	-
4	EDO	A	2793	-	-	1/1/1/1	-
4	EDO	A	2790	-	-	0/1/1/1	-
4	EDO	K	2765	-	-	0/1/1/1	-
4	EDO	B	2764	-	-	0/1/1/1	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1588	DGT	PB-O3A	3.16	1.62	1.59

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	1589	DGT	PA-O3A	3.09	1.62	1.59
7	A	1588	DGT	PB-O3B	2.94	1.62	1.59
7	A	1588	DGT	C6-N1	-2.68	1.33	1.37
7	B	1589	DGT	PA-O2A	2.48	1.59	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	1589	DGT	C2'-C1'-N9	-4.57	104.01	114.61
7	A	1588	DGT	C2'-C1'-N9	-4.46	104.25	114.61
7	A	1588	DGT	C8-N7-C5	2.91	107.51	102.55
7	A	1588	DGT	O1B-PB-O2B	2.41	123.65	112.44
7	A	1588	DGT	C5-C6-N1	2.26	118.38	114.07

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	1588	DGT	PB-O3B-PG-O2G
4	K	2767	EDO	O1-C1-C2-O2
4	A	2781	EDO	O1-C1-C2-O2
4	B	2791	EDO	O1-C1-C2-O2
4	A	2792	EDO	O1-C1-C2-O2

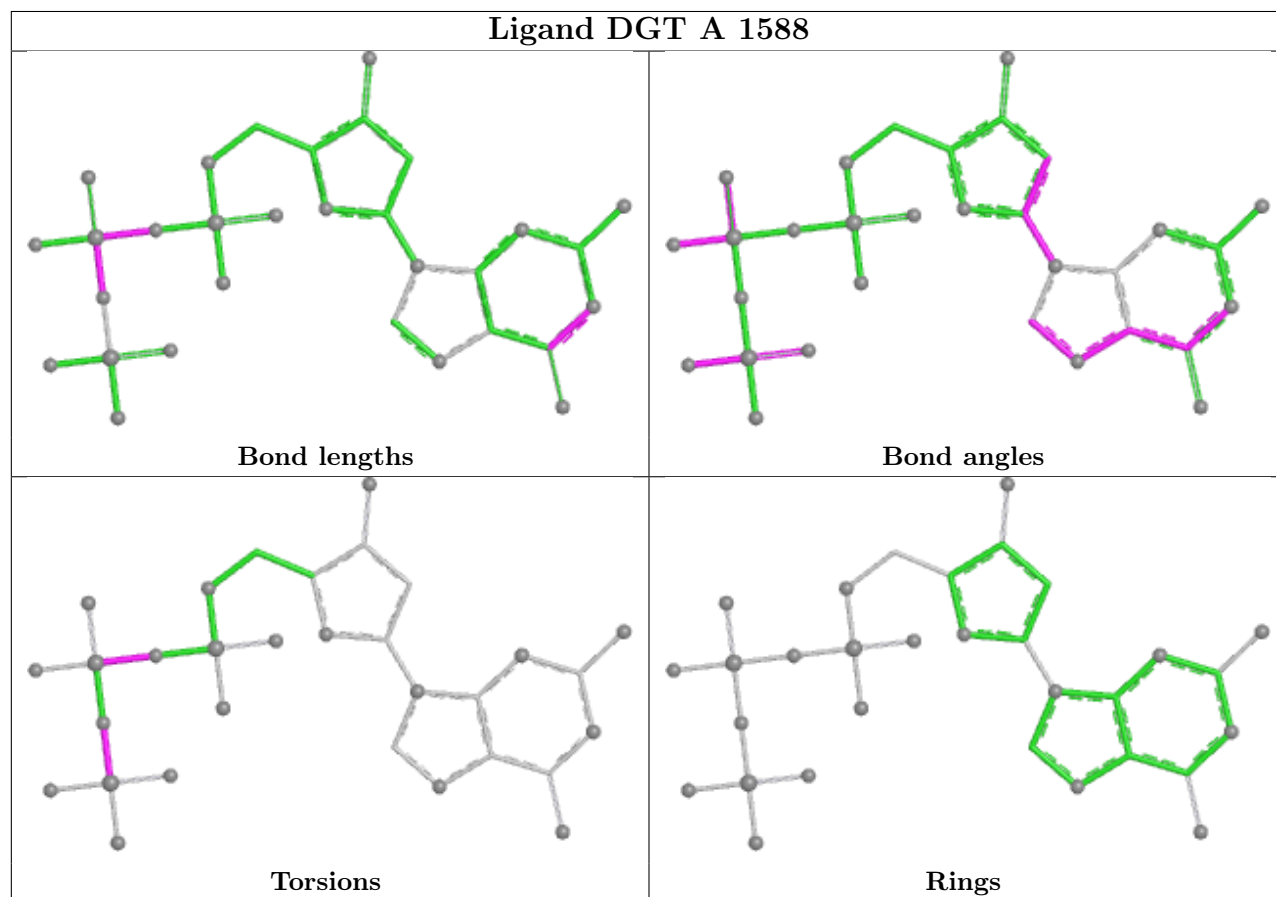
There are no ring outliers.

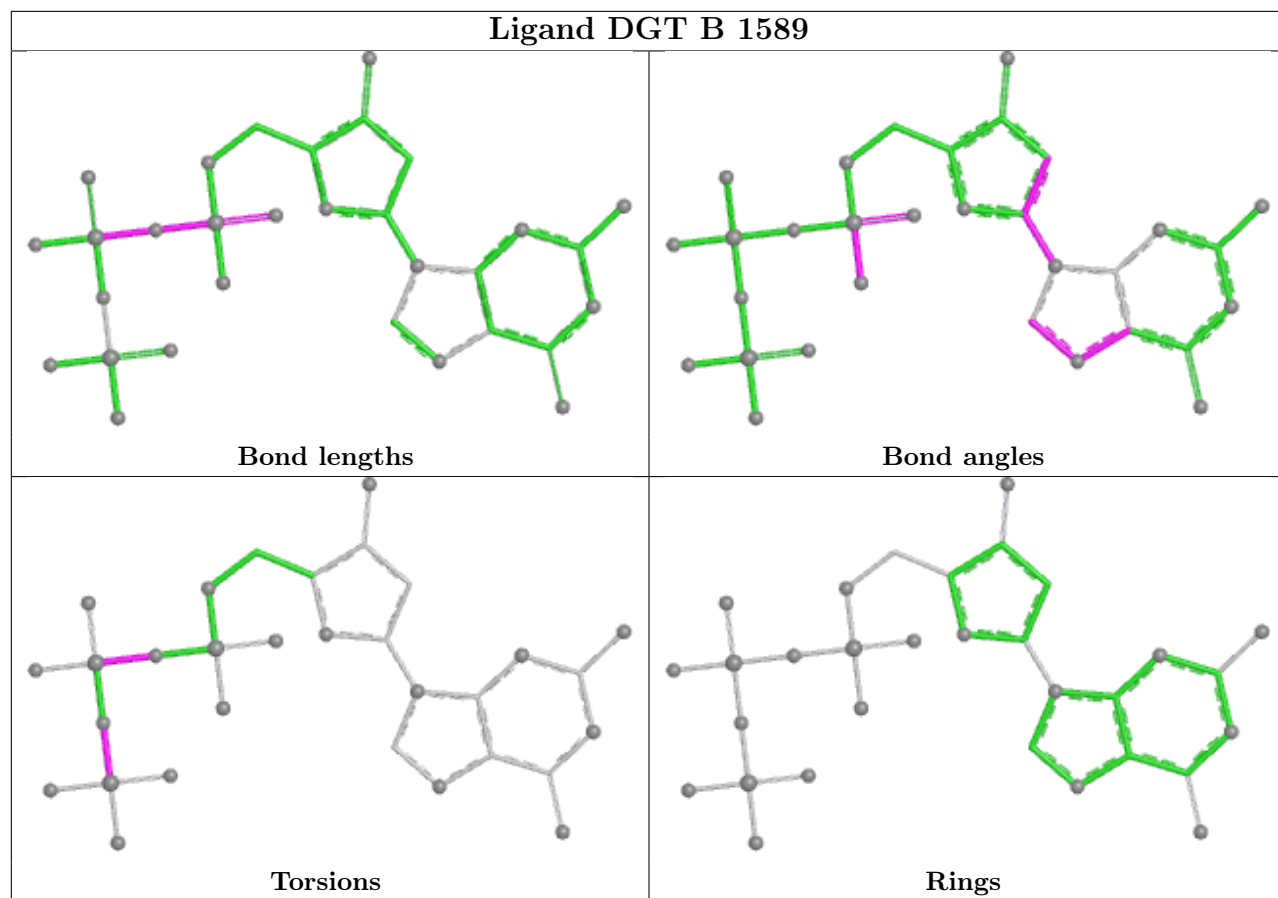
7 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	2777	EDO	1	0
4	Q	2785	EDO	1	0
4	X	2782	EDO	2	0
4	A	2771	EDO	1	0
4	X	2798	EDO	2	0
4	Y	2772	EDO	1	0
4	B	2805	EDO	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier.

Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å ²)	Q<0.9
1	J	9/10 (90%)	0.65	0	100	100	14, 16, 20, 20	0
1	Q	9/10 (90%)	0.73	0	100	100	14, 16, 23, 24	0
1	X	9/10 (90%)	-0.46	0	100	100	11, 15, 17, 18	0
2	K	14/14 (100%)	0.17	0	100	100	12, 15, 20, 21	0
2	R	13/14 (92%)	0.56	0	100	100	12, 15, 23, 24	0
2	Y	14/14 (100%)	-0.02	1 (7%)	23	22	11, 14, 26, 48	0
3	A	567/575 (98%)	0.05	11 (1%)	66	65	7, 14, 24, 38	4 (0%)
3	B	572/575 (99%)	0.62	65 (11%)	11	10	3, 14, 26, 43	5 (0%)
All	All	1207/1222 (98%)	0.33	77 (6%)	27	25	3, 14, 25, 48	9 (0%)

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	563	GLY	6.7
3	B	562	PRO	6.5
3	B	556	PRO	6.2
3	B	564	GLY	6.0
3	B	565	VAL	5.5

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

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
1	DOC	J	10	18/19	0.88	0.10	5,16,35,35	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
1	DOC	Q	10	18/19	0.96	0.11	10,12,14,15	0
1	DOC	X	10	18/19	0.97	0.08	10,14,15,16	0

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	EDO	B	2778	4/4	0.53	0.22	29,29,29,30	0
4	EDO	A	2801	4/4	0.56	0.20	35,35,36,38	0
4	EDO	B	2779	4/4	0.61	0.23	34,36,38,39	0
4	EDO	B	2791	4/4	0.61	0.20	44,45,45,46	0
4	EDO	A	2773	4/4	0.62	0.16	39,39,39,39	0
4	EDO	A	2793	4/4	0.63	0.17	40,41,42,42	0
4	EDO	K	2767	4/4	0.67	0.29	27,27,28,29	0
4	EDO	B	2787	4/4	0.68	0.18	38,38,38,39	0
4	EDO	B	2762	4/4	0.70	0.16	25,26,26,27	0
4	EDO	K	2765	4/4	0.72	0.24	36,37,38,39	0
4	EDO	A	2790	4/4	0.74	0.18	35,35,36,36	0
4	EDO	B	2776	4/4	0.74	0.15	23,24,24,24	0
4	EDO	R	2796	4/4	0.76	0.17	32,33,33,34	0
4	EDO	B	2777	4/4	0.76	0.16	24,26,27,30	0
4	EDO	A	2786	4/4	0.77	0.21	24,26,27,27	0
4	EDO	A	2781	4/4	0.78	0.13	32,32,32,33	0
4	EDO	B	2763	4/4	0.78	0.16	48,48,48,49	0
4	EDO	Y	2772	4/4	0.79	0.24	20,21,21,22	0
4	EDO	B	2795	4/4	0.82	0.12	35,35,35,35	0
4	EDO	X	2782	4/4	0.83	0.12	15,18,19,19	0
4	EDO	B	2802	4/4	0.83	0.12	24,24,25,25	0
4	EDO	A	2803	4/4	0.84	0.11	7,8,9,9	0
4	EDO	Y	2788	4/4	0.85	0.11	31,31,32,32	0
4	EDO	A	2792	4/4	0.85	0.14	38,38,39,39	0
4	EDO	K	2766	4/4	0.85	0.16	27,27,28,30	0
4	EDO	B	2764	4/4	0.85	0.12	37,37,38,39	0

Continued on next page...

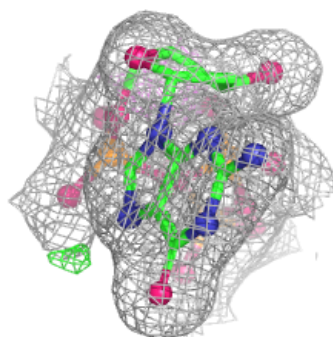
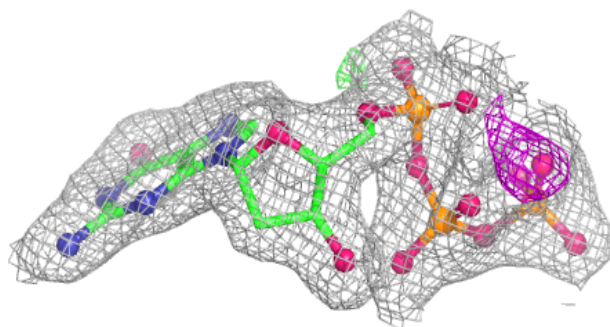
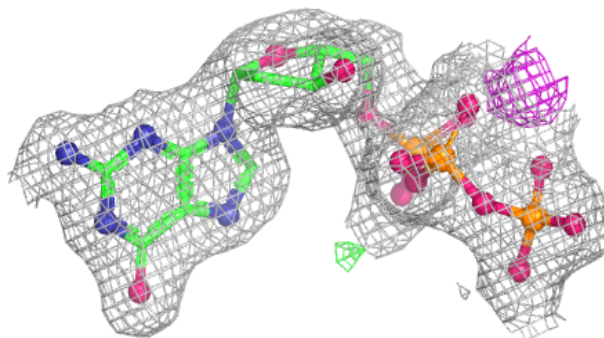
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	EDO	A	2794	4/4	0.85	0.14	31,32,32,32	0
4	EDO	X	2798	4/4	0.85	0.22	16,17,18,19	0
4	EDO	B	2805	4/4	0.85	0.14	18,20,21,24	0
4	EDO	A	2771	4/4	0.86	0.14	15,17,18,19	0
4	EDO	A	2770	4/4	0.87	0.10	31,31,31,31	0
4	EDO	A	2774	4/4	0.87	0.11	31,31,32,33	0
4	EDO	Q	2785	4/4	0.88	0.10	10,11,11,13	0
4	EDO	B	2775	4/4	0.90	0.10	3,8,8,11	0
4	EDO	A	2780	4/4	0.90	0.08	15,18,19,20	0
4	EDO	A	2800	4/4	0.91	0.09	13,14,14,15	0
4	EDO	A	2804	4/4	0.91	0.10	4,11,14,14	0
4	EDO	A	2784	4/4	0.91	0.09	17,20,21,22	0
6	MG	A	9004	1/1	0.92	0.14	16,16,16,16	0
4	EDO	B	2789	4/4	0.93	0.16	15,16,16,17	0
4	EDO	B	2769	4/4	0.94	0.12	10,11,11,14	0
4	EDO	A	2797	4/4	0.96	0.09	3,3,3,4	0
5	MN	A	9003	1/1	0.97	0.09	5,5,5,5	0
6	MG	B	9002	1/1	0.97	0.04	6,6,6,6	0
7	DGT	A	1588	31/31	0.97	0.04	2,2,6,8	0
7	DGT	B	1589	31/31	0.97	0.05	2,2,2,2	0
5	MN	B	9001	1/1	0.98	0.04	2,2,2,2	0

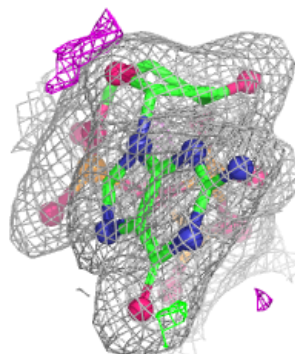
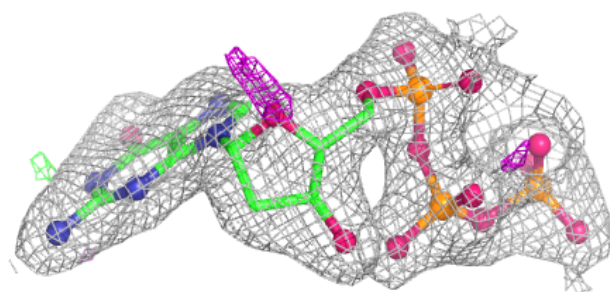
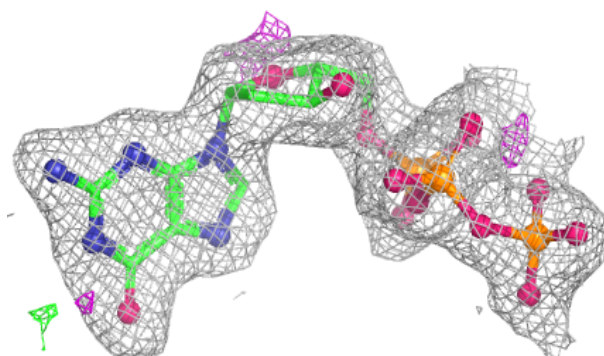
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DGT A 1588:

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

**Electron density around DGT B 1589:**

$2mF_o-DF_c$ (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.