Full wwPDB X-ray Structure Validation Report (i)

Aug 14, 2023 – 10:14 PM EDT

PDB ID	:	1S0M
Title	:	Crystal structure of a Benzo[a]pyrene Diol Epoxide adduct in a ternary complex with a DNA polymerase
Authors	:	Ling, H.; Sayer, J.M.; Boudsocq, F.; Plosky, B.S.; Woodgate, R.; Yang, W.
Deposited on	:	2003-12-31
Resolution	:	2.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (i)) were used in the production of this report:

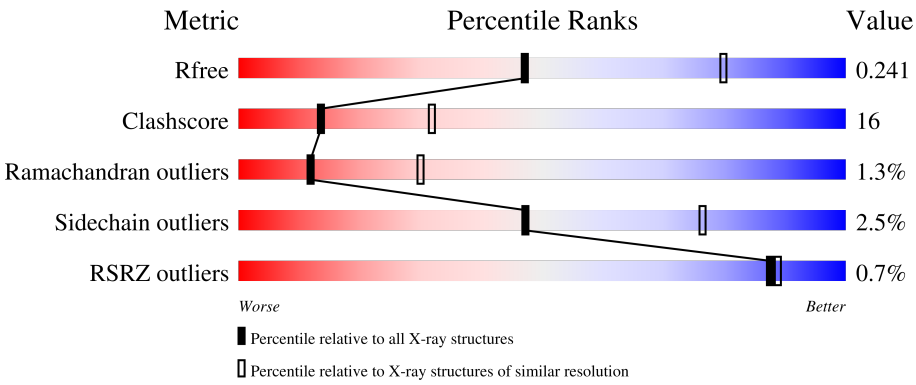
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 <=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	C	13	69% 31%
1	E	13	38% 62%
2	D	16	50% 44% 6%
2	F	16	56% 31% 6% 6%
3	A	352	70% 24% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	B	352	% 67% 28%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	BAP	F	1920	-	-	-	X
7	DTP	A	803	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 7121 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(*GP*GP*GP*GP*GP*AP*AP*GP*GP*AP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	13	274	130	56	76	12	0	0	0
1	E	13	274	130	56	76	12	0	0	0

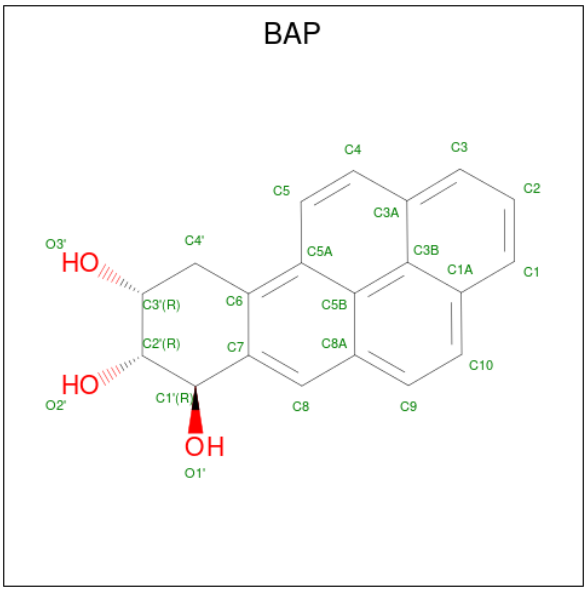
- Molecule 2 is a DNA chain called 5'-D(P*AP*TP*AP*AP*AP*TP*CP*CP*TP*TP*CP*CP*CP*CP*CP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	16	318	153	54	95	16	0	0	0
2	F	16	318	153	54	95	16	0	0	0

- Molecule 3 is a protein called DNA polymerase IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	341	2743	1760	472	504	7	0	0	0
3	B	341	2743	1760	472	504	7	0	0	0

- Molecule 4 is 1,2,3-TRIHYDROXY-1,2,3,4-TETRAHYDROBENZO[A]PYRENE (three-letter code: BAP) (formula: $C_{20}H_{16}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total C O 23 20 3	0	0
4	F	1	Total C O 23 20 3	0	0

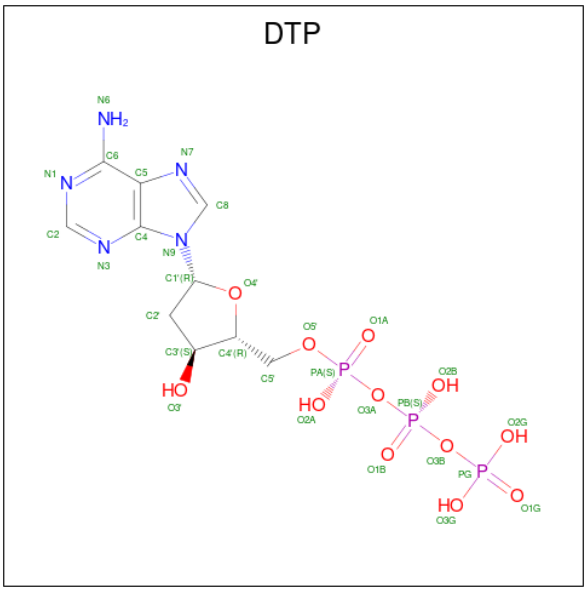
- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	1	Total Ca 1 1	0	0
5	A	3	Total Ca 3 3	0	0
5	B	3	Total Ca 3 3	0	0

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 7 is 2'-DEOXYADENOSINE 5'-TRIPHOSPHATE (three-letter code: DTP) (formula: $C_{10}H_{16}N_5O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
7	A	1	30	10	5	12	3	0	0
			Total	C	N	O	P		
7	B	1	30	10	5	12	3	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	C	26	26	26	0	0
			Total	O		
8	D	26	26	26	0	0
			Total	O		
8	E	32	32	32	0	0
			Total	O		
8	F	22	22	22	0	0
			Total	O		
8	A	138	138	138	0	0
			Total	O		
8	B	93	93	93	0	0

3 Residue-property plots (i)

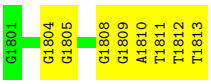
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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(*GP*GP*GP*GP*GP*AP*AP*GP*GP*AP*TP*TP*T)-3'

Chain C: 69% 31%

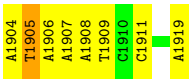
• Molecule 1: 5'-D(*GP*GP*GP*GP*GP*AP*AP*GP*GP*AP*TP*TP*T)-3'

Chain E: 38% 62%



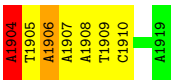
• Molecule 2: 5'-D(P*AP*TP*AP*AP*AP*TP*CP*CP*TP*TP*CP*CP*CP*CP*CP*A)-3'

Chain D: 50% 44% 6%



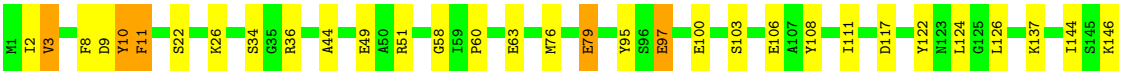
• Molecule 2: 5'-D(P*AP*TP*AP*AP*AP*TP*CP*CP*TP*TP*CP*CP*CP*CP*CP*A)-3'

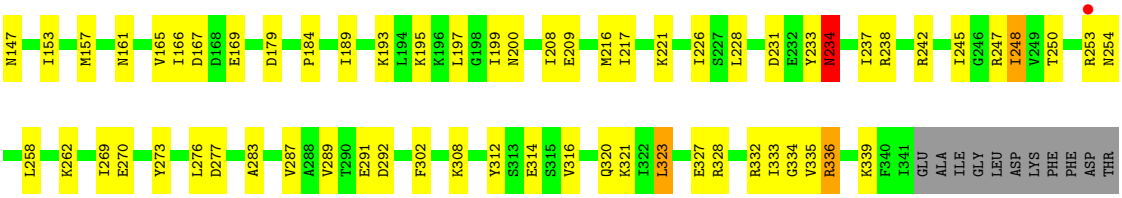
Chain F: 56% 31% 6% 6%



• Molecule 3: DNA polymerase IV

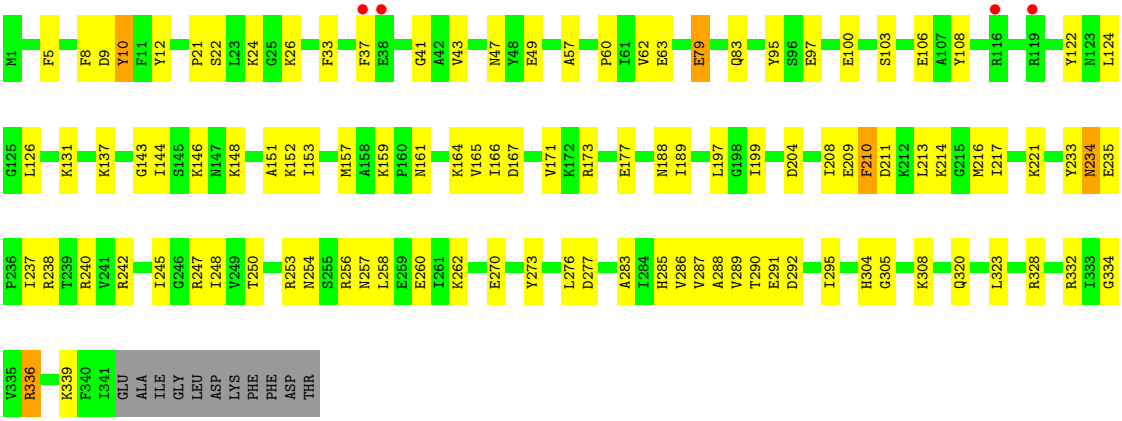
Chain A: 70% 24%





- Molecule 3: DNA polymerase IV

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	100.03Å 103.55Å 106.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.60 – 2.70 25.64 – 2.70	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.60-2.70) 89.6 (25.64-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.73 (at 2.72Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.207 , 0.246 0.203 , 0.241	Depositor DCC
R_{free} test set	703 reflections (2.54%)	wwPDB-VP
Wilson B-factor (Å^2)	50.7	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 39.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.013 for -h,l,k	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7121	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 42.57 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.0016e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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, CA, DTP, BAP

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	C	0.66	0/309	0.88	0/478
1	E	0.67	0/309	0.86	0/478
2	D	2.92	5/354 (1.4%)	1.43	6/541 (1.1%)
2	F	2.70	8/354 (2.3%)	2.23	8/541 (1.5%)
3	A	0.50	0/2782	0.63	0/3736
3	B	0.48	0/2782	0.62	0/3736
All	All	1.02	13/6890 (0.2%)	0.88	14/9510 (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
2	F	0	1

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1904	DA	P-O5'	-42.83	1.17	1.59
2	F	1904	DA	P-O5'	-41.49	1.18	1.59
2	D	1904	DA	O5'-C5'	-20.82	0.90	1.42
2	F	1904	DA	P-OP1	-15.73	1.22	1.49
2	D	1904	DA	C5'-C4'	14.78	1.67	1.51
2	D	1904	DA	P-OP1	-14.60	1.24	1.49
2	F	1904	DA	O3'-P	11.77	1.75	1.61
2	F	1904	DA	O5'-C5'	-8.62	1.20	1.42
2	F	1904	DA	C1'-N9	-8.62	1.35	1.47
2	D	1904	DA	P-OP2	7.32	1.61	1.49
2	F	1904	DA	C5'-C4'	-6.69	1.44	1.51
2	F	1906	DA	N3-C4	5.15	1.38	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1906	DA	C6-N6	5.02	1.38	1.33

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	1904	DA	O4'-C1'-N9	37.36	134.15	108.00
2	F	1904	DA	C5'-C4'-C3'	-17.34	82.88	114.10
2	D	1904	DA	P-O5'-C5'	15.84	146.25	120.90
2	F	1904	DA	O5'-C5'-C4'	12.67	142.69	111.00
2	D	1904	DA	O5'-P-OP1	12.65	125.89	110.70
2	D	1904	DA	OP1-P-OP2	-12.21	101.28	119.60
2	F	1904	DA	P-O3'-C3'	-10.09	107.59	119.70
2	F	1904	DA	OP1-P-O3'	8.76	124.47	105.20
2	F	1904	DA	O3'-P-O5'	-8.23	88.36	104.00
2	F	1904	DA	C8-N9-C1'	-7.71	113.82	127.70
2	F	1904	DA	N9-C1'-C2'	-7.35	98.64	112.60
2	D	1904	DA	C5'-C4'-O4'	-7.18	95.65	109.30
2	D	1905	DT	C6-N1-C1'	-5.52	112.12	120.40
2	D	1905	DT	C2-N1-C1'	5.46	126.93	118.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	1904	DA	Sidechain

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	C	274	0	148	4	0
1	E	274	0	148	8	0
2	D	318	0	180	15	0
2	F	318	0	180	13	0
3	A	2743	0	2889	83	0
3	B	2743	0	2889	98	0
4	D	23	0	15	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	23	0	15	4	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0
5	E	1	0	0	0	0
6	A	1	0	0	0	0
7	A	30	0	12	9	0
7	B	30	0	10	5	0
8	A	138	0	0	17	0
8	B	93	0	0	23	0
8	C	26	0	0	1	0
8	D	26	0	0	8	0
8	E	32	0	0	1	0
8	F	22	0	0	2	0
All	All	7121	0	6486	212	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 16.

All (212) 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:A:10:TYR:HA	7:A:803:DTP:O2A	1.33	1.28
2:F:1905:DT:H1'	2:F:1906:DA:H5'	1.27	1.12
2:D:1919:DA:C4	3:B:21:PRO:HB2	2.08	0.89
2:D:1919:DA:N3	3:B:21:PRO:HB2	1.88	0.88
3:A:10:TYR:CA	7:A:803:DTP:O2A	2.22	0.87
3:A:97:GLU:H	3:A:97:GLU:CD	1.80	0.85
3:A:157:MET:HE2	3:A:166:ILE:HD11	1.58	0.85
3:A:226:ILE:HA	8:A:930:HOH:O	1.76	0.84
3:B:157:MET:HE2	3:B:166:ILE:HD11	1.61	0.82
3:A:60:PRO:HD2	3:A:63:GLU:HG3	1.61	0.82
1:E:1804:DG:H2"	1:E:1805:DG:OP2	1.78	0.81
3:B:214:LYS:HB2	8:B:864:HOH:O	1.80	0.80
2:D:1905:DT:H5"	8:D:44:HOH:O	1.83	0.78
3:A:179:ASP:HA	8:A:897:HOH:O	1.84	0.77
3:B:248:ILE:HG12	3:B:334:GLY:HA3	1.68	0.76
3:B:289:VAL:HA	8:B:860:HOH:O	1.86	0.76
2:F:1905:DT:C1'	2:F:1906:DA:H5'	2.12	0.75
3:B:97:GLU:H	3:B:97:GLU:CD	1.94	0.72
2:F:1907:DA:OP1	3:B:247:ARG:HD2	1.91	0.71
3:B:157:MET:CE	3:B:166:ILE:HD11	2.21	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:292:ASP:OD1	3:A:328:ARG:HD2	1.90	0.70
2:D:1911:DC:H5"	8:D:310:HOH:O	1.91	0.70
3:A:248:ILE:HG12	3:A:334:GLY:HA3	1.73	0.70
3:B:10:TYR:HA	7:B:804:DTP:PB	2.31	0.70
3:B:292:ASP:OD1	3:B:328:ARG:HD2	1.92	0.70
3:B:254:ASN:HD22	3:B:291:GLU:HB3	1.58	0.68
3:B:217:ILE:HD12	3:B:221:LYS:HB3	1.78	0.67
3:A:157:MET:CE	3:A:166:ILE:HD11	2.25	0.66
3:B:60:PRO:HD2	3:B:63:GLU:HG3	1.78	0.66
3:B:290:THR:HG23	8:B:824:HOH:O	1.95	0.65
3:B:336:ARG:HD2	8:B:811:HOH:O	1.98	0.64
3:B:100:GLU:HB2	3:B:237:ILE:HG23	1.78	0.64
7:A:803:DTP:O2A	8:A:841:HOH:O	2.15	0.64
3:B:295:ILE:HG12	8:B:860:HOH:O	1.98	0.63
2:F:1906:DA:N6	4:F:1920:BAP:H5	2.14	0.63
2:F:1905:DT:H5"	3:B:41:GLY:HA2	1.81	0.62
2:D:1919:DA:C2	3:B:21:PRO:HB2	2.34	0.62
8:D:114:HOH:O	3:B:137:LYS:HG3	1.99	0.62
3:B:214:LYS:HD3	8:B:893:HOH:O	1.98	0.62
3:A:167:ASP:OD1	3:A:169:GLU:HB3	2.00	0.61
2:D:1919:DA:O4'	3:B:21:PRO:HD2	2.00	0.61
3:B:144:ILE:HB	3:B:165:VAL:HG22	1.82	0.61
2:D:1907:DA:H2"	2:D:1908:DA:H5'	1.83	0.61
1:C:1801:DG:H5'	3:B:83:GLN:OE1	2.00	0.61
3:A:26:LYS:HG2	8:A:921:HOH:O	2.00	0.60
3:B:288:ALA:HB2	8:B:878:HOH:O	2.00	0.60
3:B:197:LEU:HD11	3:B:216:MET:HG2	1.83	0.60
3:B:12:TYR:CE2	7:B:804:DTP:H2'1	2.37	0.60
2:D:1906:DA:H4'	8:D:244:HOH:O	2.00	0.59
3:A:242:ARG:HD2	3:A:245:ILE:HD11	1.84	0.59
3:B:21:PRO:O	3:B:24:LYS:HB2	2.01	0.59
3:B:146:LYS:HG2	8:B:828:HOH:O	2.01	0.59
3:A:22:SER:O	3:A:26:LYS:HE3	2.02	0.59
3:B:304:HIS:HD2	3:B:305:GLY:O	1.86	0.59
3:B:257:ASN:OD1	3:B:260:GLU:HB2	2.03	0.59
3:B:100:GLU:OE2	3:B:240:ARG:NH2	2.35	0.58
3:B:286:VAL:HG12	8:B:878:HOH:O	2.02	0.58
8:E:167:HOH:O	4:F:1920:BAP:H1'	2.02	0.58
3:A:153:ILE:O	3:A:157:MET:HG3	2.03	0.58
3:A:144:ILE:HB	3:A:165:VAL:HG22	1.85	0.58
1:E:1812:DT:H2'	1:E:1813:DT:C6	2.40	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:1905:DT:C5'	3:B:41:GLY:HA2	2.35	0.57
3:A:122:TYR:CE2	3:A:126:LEU:HD11	2.40	0.57
2:F:1905:DT:H2"	2:F:1906:DA:O5'	2.06	0.56
3:B:47:ASN:HB2	8:B:825:HOH:O	2.06	0.56
3:B:199:ILE:HD11	3:B:208:ILE:HG21	1.88	0.56
3:A:283:ALA:HB2	3:A:339:LYS:HD2	1.89	0.55
3:B:10:TYR:HA	7:B:804:DTP:O1B	2.06	0.55
3:B:289:VAL:HG13	8:B:860:HOH:O	2.07	0.55
3:A:217:ILE:HD12	3:A:221:LYS:HB3	1.89	0.55
3:A:100:GLU:HG3	3:A:238:ARG:O	2.07	0.55
3:A:197:LEU:HD11	3:A:216:MET:HG2	1.88	0.55
3:B:43:VAL:O	3:B:57:ALA:HA	2.07	0.55
3:B:304:HIS:CD2	3:B:305:GLY:O	2.60	0.54
3:A:226:ILE:HG12	8:A:930:HOH:O	2.06	0.54
3:B:283:ALA:HB2	3:B:339:LYS:HD2	1.90	0.54
3:B:289:VAL:HB	3:B:332:ARG:HB2	1.89	0.54
8:F:320:HOH:O	3:B:242:ARG:HG3	2.07	0.54
3:A:100:GLU:HB2	3:A:237:ILE:HG23	1.89	0.54
3:A:234:ASN:CG	3:A:234:ASN:O	2.46	0.54
3:B:22:SER:O	3:B:26:LYS:HE3	2.08	0.53
3:A:199:ILE:HD11	3:A:208:ILE:HG21	1.91	0.53
3:B:258:LEU:HD22	3:B:320:GLN:NE2	2.24	0.52
3:B:210:PHE:O	3:B:213:LEU:N	2.41	0.52
3:A:97:GLU:CD	3:A:97:GLU:N	2.56	0.52
2:D:1908:DA:OP2	3:A:336:ARG:NH2	2.40	0.52
3:B:254:ASN:ND2	3:B:291:GLU:HB3	2.22	0.52
3:A:302:PHE:HZ	3:A:314:GLU:HG2	1.75	0.52
3:B:256:ARG:NH1	3:B:328:ARG:O	2.42	0.52
3:A:10:TYR:HA	7:A:803:DTP:PA	2.42	0.52
3:B:153:ILE:O	3:B:157:MET:HG3	2.09	0.52
3:B:221:LYS:HB2	8:B:896:HOH:O	2.10	0.52
3:B:37:PHE:HE1	8:B:876:HOH:O	1.92	0.51
3:A:3:VAL:HG11	3:A:147:ASN:C	2.31	0.51
3:A:234:ASN:O	3:A:234:ASN:OD1	2.29	0.51
3:A:242:ARG:CD	3:A:245:ILE:HG12	2.41	0.51
3:B:250:THR:HA	3:B:332:ARG:HG2	1.92	0.51
3:A:195:LYS:HD3	8:A:834:HOH:O	2.10	0.50
3:A:270:GLU:OE2	3:A:308:LYS:HD3	2.11	0.50
3:A:117:ASP:HB2	8:A:885:HOH:O	2.11	0.50
3:A:36:ARG:NH2	3:A:254:ASN:OD1	2.44	0.50
3:B:258:LEU:HD22	3:B:320:GLN:HE21	1.76	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:51:ARG:HD3	8:A:889:HOH:O	2.11	0.49
3:A:237:ILE:HD12	3:A:237:ILE:N	2.28	0.49
3:B:188:ASN:ND2	3:B:188:ASN:N	2.61	0.49
2:D:1907:DA:H1'	2:D:1908:DA:H5"	1.94	0.49
3:B:9:ASP:O	3:B:10:TYR:C	2.50	0.49
2:D:1906:DA:C6	4:D:1920:BAP:H5	2.48	0.49
3:A:11:PHE:N	8:A:841:HOH:O	2.45	0.48
2:F:1906:DA:OP2	3:B:332:ARG:NE	2.46	0.48
3:A:10:TYR:C	8:A:841:HOH:O	2.51	0.48
3:A:167:ASP:HB2	8:A:864:HOH:O	2.12	0.48
3:A:100:GLU:HB3	3:A:108:TYR:HB2	1.94	0.48
3:B:167:ASP:O	3:B:171:VAL:HG23	2.13	0.48
3:A:9:ASP:O	3:A:10:TYR:C	2.50	0.48
3:A:254:ASN:ND2	3:A:291:GLU:HB3	2.29	0.48
3:B:62:VAL:CG2	8:B:876:HOH:O	2.62	0.48
3:B:273:TYR:HA	3:B:276:LEU:HD12	1.94	0.48
8:D:124:HOH:O	3:A:242:ARG:HG2	2.14	0.48
3:A:258:LEU:O	3:A:262:LYS:HG3	2.14	0.48
1:E:1811:DT:H2"	1:E:1812:DT:O5'	2.14	0.47
3:B:5:PHE:CD2	3:B:152:LYS:HA	2.49	0.47
3:A:242:ARG:HD2	3:A:245:ILE:CD1	2.44	0.47
3:A:287:VAL:O	3:A:333:ILE:HD12	2.15	0.47
3:B:148:LYS:HG3	8:B:808:HOH:O	2.14	0.47
2:F:1904:DA:H1'	8:F:304:HOH:O	2.15	0.47
2:F:1909:DT:H2"	2:F:1910:DC:C6	2.50	0.47
3:B:100:GLU:HB3	3:B:108:TYR:HB2	1.96	0.47
3:B:173:ARG:NH1	3:B:177:GLU:OE1	2.43	0.47
3:B:188:ASN:N	3:B:188:ASN:HD22	2.10	0.47
3:A:8:PHE:CD1	3:A:8:PHE:N	2.81	0.47
3:A:9:ASP:O	3:A:11:PHE:N	2.48	0.47
3:B:95:TYR:HD1	3:B:124:LEU:HD11	1.79	0.47
2:D:1907:DA:H2"	2:D:1908:DA:C5'	2.44	0.47
3:A:269:ILE:HA	3:A:335:VAL:HG11	1.96	0.47
3:A:200:ASN:HB2	8:A:938:HOH:O	2.15	0.46
3:B:143:GLY:HA3	3:B:151:ALA:O	2.15	0.46
2:D:1905:DT:O4	7:A:803:DTP:N6	2.40	0.46
3:B:336:ARG:HB2	8:B:811:HOH:O	2.15	0.46
2:F:1908:DA:OP2	3:B:336:ARG:NH2	2.47	0.46
3:A:327:GLU:HG2	8:A:892:HOH:O	2.15	0.46
3:A:289:VAL:CG2	3:A:332:ARG:HB2	2.46	0.45
3:B:8:PHE:CD1	3:B:8:PHE:N	2.83	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:100:GLU:HG3	3:B:238:ARG:O	2.16	0.45
3:B:208:ILE:HG12	3:B:209:GLU:N	2.31	0.45
3:A:242:ARG:HD3	3:A:245:ILE:HG12	1.99	0.45
3:B:12:TYR:CD2	7:B:804:DTP:H2'1	2.51	0.45
3:B:242:ARG:HD2	8:B:815:HOH:O	2.16	0.45
3:A:2:ILE:HG22	3:A:111:ILE:CG1	2.48	0.44
3:A:60:PRO:HB2	3:A:63:GLU:HG2	1.99	0.44
3:B:289:VAL:CG2	3:B:332:ARG:HB2	2.46	0.44
8:D:44:HOH:O	3:A:34:SER:HB3	2.18	0.44
3:A:189:ILE:O	3:A:193:LYS:HG3	2.18	0.44
2:F:1906:DA:N6	4:F:1920:BAP:C5	2.81	0.44
3:B:33:PHE:HB3	8:B:854:HOH:O	2.18	0.44
3:B:131:LYS:HE2	8:B:894:HOH:O	2.18	0.44
2:D:1908:DA:C2	2:D:1909:DT:C2	3.06	0.43
3:B:122:TYR:CE2	3:B:126:LEU:HD11	2.53	0.43
3:A:63:GLU:HB3	8:A:909:HOH:O	2.18	0.43
3:A:302:PHE:CZ	3:A:314:GLU:HG2	2.52	0.43
1:E:1804:DG:H1'	1:E:1805:DG:H5'	1.99	0.43
3:B:285:HIS:HD2	8:B:883:HOH:O	2.00	0.43
3:A:95:TYR:HD1	3:A:124:LEU:HD11	1.84	0.43
1:E:1812:DT:H4'	1:E:1812:DT:OP1	2.18	0.43
3:A:242:ARG:HD2	3:A:245:ILE:CG1	2.49	0.43
3:B:171:VAL:CG2	8:B:828:HOH:O	2.67	0.43
8:D:124:HOH:O	3:A:242:ARG:CG	2.67	0.43
3:B:79:GLU:H	3:B:79:GLU:CD	2.20	0.43
7:A:803:DTP:H8	8:A:840:HOH:O	2.18	0.42
3:B:159:LYS:NZ	7:B:804:DTP:O1G	2.47	0.42
3:B:199:ILE:HG23	3:B:204:ASP:HB2	2.02	0.42
1:C:1812:DT:H71	8:C:224:HOH:O	2.18	0.42
2:F:1905:DT:H71	4:F:1920:BAP:H5	2.00	0.42
3:A:258:LEU:O	3:A:258:LEU:HD12	2.19	0.42
1:C:1808:DG:H2"	1:C:1809:DG:OP2	2.19	0.42
3:A:208:ILE:HG12	3:A:209:GLU:N	2.34	0.42
3:A:289:VAL:HB	3:A:332:ARG:HB2	2.01	0.42
7:A:803:DTP:O2G	7:A:803:DTP:H4'	2.19	0.42
3:B:287:VAL:C	8:B:878:HOH:O	2.58	0.42
3:B:97:GLU:CD	3:B:97:GLU:N	2.68	0.42
3:B:171:VAL:HG22	8:B:828:HOH:O	2.19	0.42
3:B:234:ASN:HD22	3:B:234:ASN:HA	1.70	0.42
3:A:137:LYS:HD2	8:A:928:HOH:O	2.19	0.42
3:A:312:TYR:O	3:A:316:VAL:HG23	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:1808:DG:H1'	1:E:1809:DG:H5"	2.02	0.42
3:A:323:LEU:HD12	3:A:323:LEU:HA	1.92	0.41
1:C:1801:DG:H2'	3:B:83:GLN:OE1	2.20	0.41
3:B:233:TYR:CZ	3:B:235:GLU:HB2	2.56	0.41
8:D:44:HOH:O	3:A:34:SER:CB	2.67	0.41
1:E:1813:DT:O3'	3:B:103:SER:HB2	2.21	0.41
3:A:321:LYS:HB2	3:A:321:LYS:HE3	1.79	0.41
3:A:228:LEU:HD23	3:A:233:TYR:HB3	2.03	0.41
3:A:58:GLY:CA	7:A:803:DTP:N6	2.84	0.41
3:A:250:THR:HA	3:A:332:ARG:HG2	2.03	0.41
3:A:316:VAL:O	3:A:320:GLN:HG3	2.21	0.41
3:A:146:LYS:NZ	3:A:231:ASP:OD1	2.54	0.41
3:B:103:SER:OG	3:B:106:GLU:HB2	2.20	0.41
3:B:270:GLU:OE2	3:B:308:LYS:HD3	2.21	0.41
1:E:1810:DA:H3'	3:B:189:ILE:HD12	2.03	0.41
3:B:164:LYS:HG3	3:B:165:VAL:N	2.36	0.41
3:A:103:SER:OG	3:A:106:GLU:HB2	2.21	0.41
3:A:79:GLU:CD	3:A:79:GLU:H	2.23	0.40
3:A:247:ARG:HD3	8:A:842:HOH:O	2.20	0.40
3:B:258:LEU:O	3:B:262:LYS:HG3	2.21	0.40
3:B:164:LYS:CG	3:B:165:VAL:N	2.84	0.40
7:A:803:DTP:H8	7:A:803:DTP:C3'	2.52	0.40
2:D:1905:DT:C2'	2:D:1906:DA:H8	2.35	0.40
3:A:44:ALA:HB2	3:A:76:MET:HE3	2.03	0.40
3:B:242:ARG:HD3	3:B:245:ILE:HD11	2.04	0.40
3:A:253:ARG:O	3:A:253:ARG:HG3	2.22	0.40
3:A:273:TYR:HA	3:A:276:LEU:HD12	2.04	0.40
3:B:9:ASP:CG	3:B:161:ASN:H	2.25	0.40

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	339/352 (96%)	317 (94%)	17 (5%)	5 (2%)	10	26
3	B	339/352 (96%)	316 (93%)	19 (6%)	4 (1%)	13	32
All	All	678/704 (96%)	633 (93%)	36 (5%)	9 (1%)	12	30

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	277	ASP
3	B	277	ASP
3	A	10	TYR
3	B	10	TYR
3	A	234	ASN
3	B	210	PHE
3	B	211	ASP
3	A	11	PHE
3	A	161	ASN

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	
3	A	300/309 (97%)	291 (97%)	9 (3%)	41	70
3	B	300/309 (97%)	294 (98%)	6 (2%)	55	81
All	All	600/618 (97%)	585 (98%)	15 (2%)	47	76

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	3	VAL
3	A	49	GLU
3	A	79	GLU
3	A	97	GLU
3	A	184	PRO
3	A	234	ASN
3	A	248	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	323	LEU
3	A	336	ARG
3	B	49	GLU
3	B	79	GLU
3	B	234	ASN
3	B	253	ARG
3	B	323	LEU
3	B	336	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	188	ASN
3	A	320	GLN
3	B	188	ASN
3	B	234	ASN
3	B	304	HIS
3	B	320	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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	BAP	F	1920	2	27,27,27	2.45	11 (40%)	39,42,42	1.24	5 (12%)
4	BAP	D	1920	2	27,27,27	2.19	9 (33%)	39,42,42	0.75	0
7	DTP	B	804	5	26,32,32	2.15	5 (19%)	30,50,50	3.58	8 (26%)
7	DTP	A	803	6,5	26,32,32	2.35	7 (26%)	30,50,50	3.58	12 (40%)

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	BAP	F	1920	2	-	-	0/5/5/5
4	BAP	D	1920	2	-	-	0/5/5/5
7	DTP	B	804	5	-	3/18/34/34	0/3/3/3
7	DTP	A	803	6,5	-	4/18/34/34	0/3/3/3

All (32) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	804	DTP	C4-N3	6.03	1.44	1.35
7	A	803	DTP	C4-N3	5.62	1.43	1.35
7	A	803	DTP	C8-N7	-5.40	1.25	1.34
7	A	803	DTP	O4'-C1'	4.86	1.53	1.42
7	B	804	DTP	O5'-C5'	-4.69	1.26	1.44
4	F	1920	BAP	C6-C7	4.43	1.45	1.37
4	F	1920	BAP	C8-C7	4.40	1.44	1.36
4	D	1920	BAP	C4'-C6	4.28	1.56	1.51
4	F	1920	BAP	C4'-C6	4.05	1.56	1.51
4	F	1920	BAP	C6-C5A	3.99	1.49	1.43
7	A	803	DTP	PB-O2B	-3.97	1.36	1.55
7	B	804	DTP	PA-O5'	-3.96	1.43	1.59
4	D	1920	BAP	C8-C7	3.95	1.43	1.36
7	A	803	DTP	PA-O5'	3.92	1.75	1.59
4	D	1920	BAP	C6-C5A	3.72	1.49	1.43
4	D	1920	BAP	C6-C7	3.68	1.43	1.37
4	D	1920	BAP	C4'-C3'	3.64	1.58	1.52
4	F	1920	BAP	C2-C1	3.57	1.44	1.36

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	1920	BAP	C10-C9	3.56	1.45	1.35
4	F	1920	BAP	C2-C3	3.33	1.44	1.36
4	D	1920	BAP	C2-C3	3.22	1.44	1.36
4	D	1920	BAP	C2-C1	3.19	1.44	1.36
4	F	1920	BAP	C7-C1'	3.12	1.56	1.51
7	B	804	DTP	C8-N7	-3.12	1.29	1.34
4	D	1920	BAP	C10-C9	3.06	1.43	1.35
7	A	803	DTP	C2-N3	2.83	1.36	1.32
4	F	1920	BAP	C4'-C3'	2.78	1.56	1.52
4	D	1920	BAP	C5-C4	2.72	1.42	1.35
4	F	1920	BAP	C2'-C1'	2.51	1.58	1.52
7	B	804	DTP	PB-O1B	-2.51	1.42	1.50
4	F	1920	BAP	C5-C4	2.45	1.41	1.35
7	A	803	DTP	PG-O2G	-2.13	1.46	1.54

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	B	804	DTP	O5'-C5'-C4'	15.94	163.85	108.99
7	A	803	DTP	O5'-C5'-C4'	9.93	143.19	108.99
7	A	803	DTP	PA-O5'-C5'	8.72	172.84	121.68
7	A	803	DTP	C2'-C1'-N9	-6.72	98.76	114.27
7	B	804	DTP	O5'-PA-O1A	-6.43	83.94	109.07
7	A	803	DTP	C2'-C3'-C4'	-6.11	90.04	102.76
7	A	803	DTP	O4'-C4'-C3'	-5.79	92.15	105.67
7	B	804	DTP	C5'-C4'-C3'	-4.38	89.02	114.74
4	F	1920	BAP	C4'-C3'-C2'	-4.28	105.15	110.30
7	B	804	DTP	PA-O5'-C5'	4.25	146.58	121.68
7	A	803	DTP	C3'-C2'-C1'	-4.23	91.95	102.54
7	A	803	DTP	C4-C5-N7	4.05	113.62	109.40
7	A	803	DTP	O2G-PG-O3B	3.50	116.39	104.64
7	A	803	DTP	O3'-C3'-C4'	3.23	122.46	110.10
4	F	1920	BAP	C4'-C6-C7	-3.13	115.38	121.09
7	B	804	DTP	O4'-C4'-C5'	-3.08	99.24	109.37
7	A	803	DTP	C4'-O4'-C1'	-3.02	102.17	109.45
7	B	804	DTP	C4-C5-N7	3.00	112.53	109.40
7	A	803	DTP	O4'-C1'-C2'	-2.52	101.49	106.25
7	B	804	DTP	O4'-C4'-C3'	2.49	111.48	105.67
4	F	1920	BAP	C4'-C6-C5A	2.46	125.39	120.24
4	F	1920	BAP	O3'-C3'-C4'	2.41	115.11	109.49
7	B	804	DTP	PA-O3A-PB	2.32	140.80	132.83
4	F	1920	BAP	C6-C4'-C3'	2.27	116.59	111.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	803	DTP	O2A-PA-O1A	2.10	122.62	112.24

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	803	DTP	C4'-C5'-O5'-PA
7	A	803	DTP	O4'-C4'-C5'-O5'
7	B	804	DTP	C4'-C5'-O5'-PA
7	A	803	DTP	PA-O3A-PB-O2B
7	B	804	DTP	PG-O3B-PB-O2B
7	A	803	DTP	PA-O3A-PB-O1B
7	B	804	DTP	PB-O3A-PA-O1A

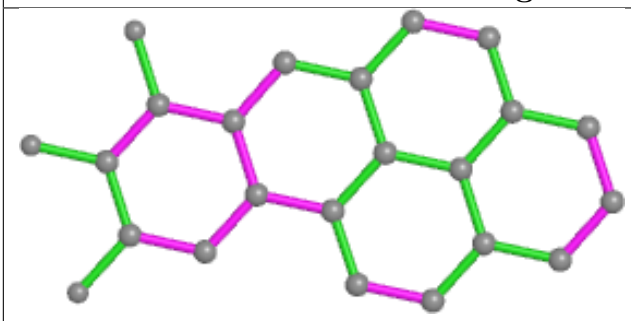
There are no ring outliers.

4 monomers are involved in 19 short contacts:

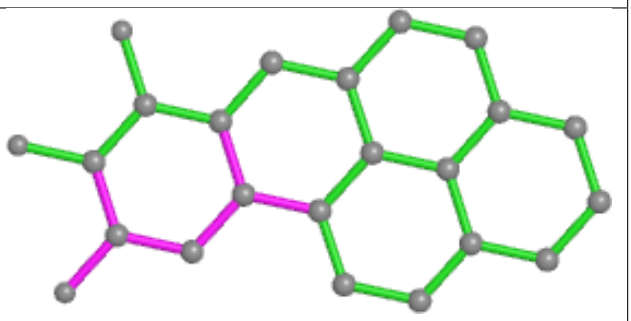
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	1920	BAP	4	0
4	D	1920	BAP	1	0
7	B	804	DTP	5	0
7	A	803	DTP	9	0

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 then 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.

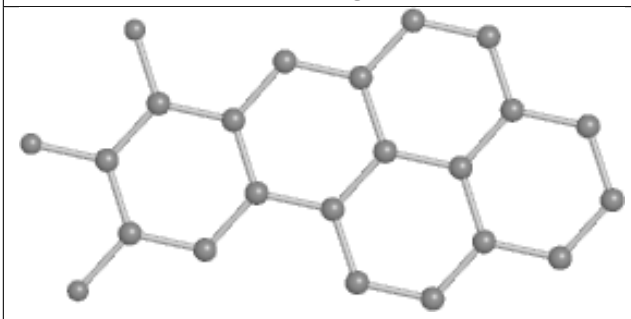
Ligand BAP F 1920



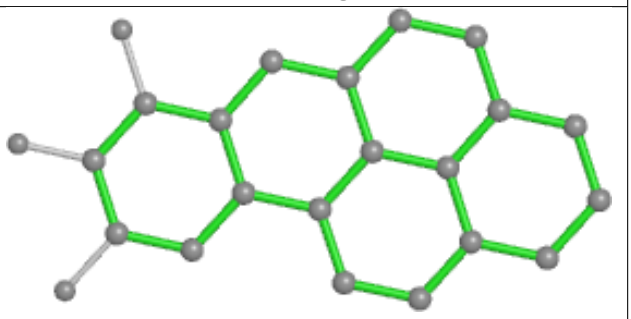
Bond lengths



Bond angles

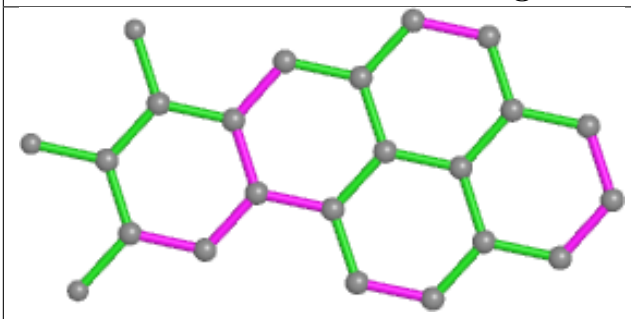


Torsions

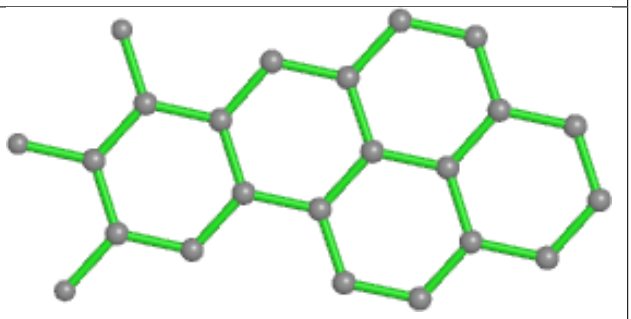


Rings

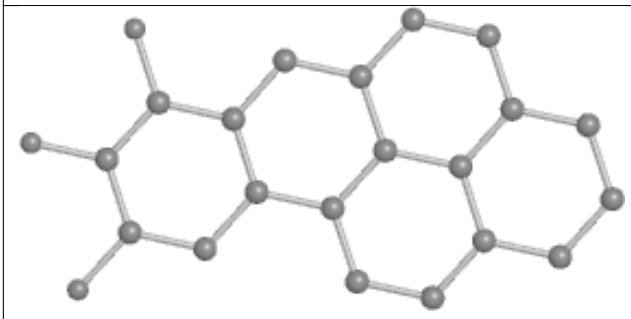
Ligand BAP D 1920



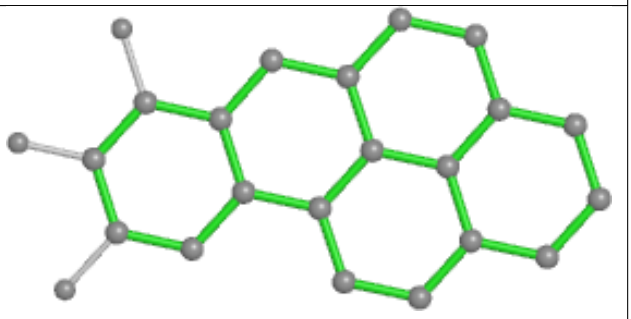
Bond lengths



Bond angles

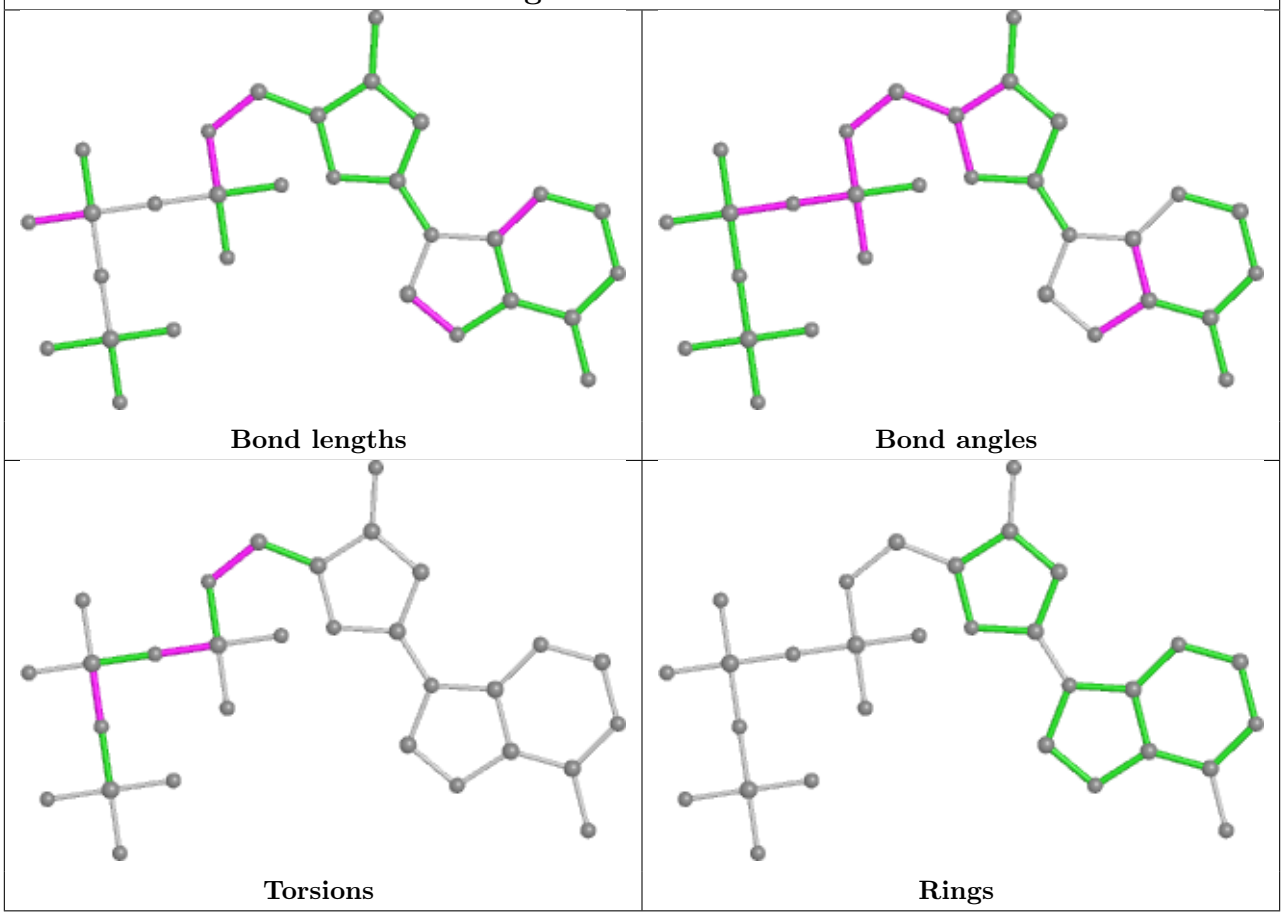


Torsions

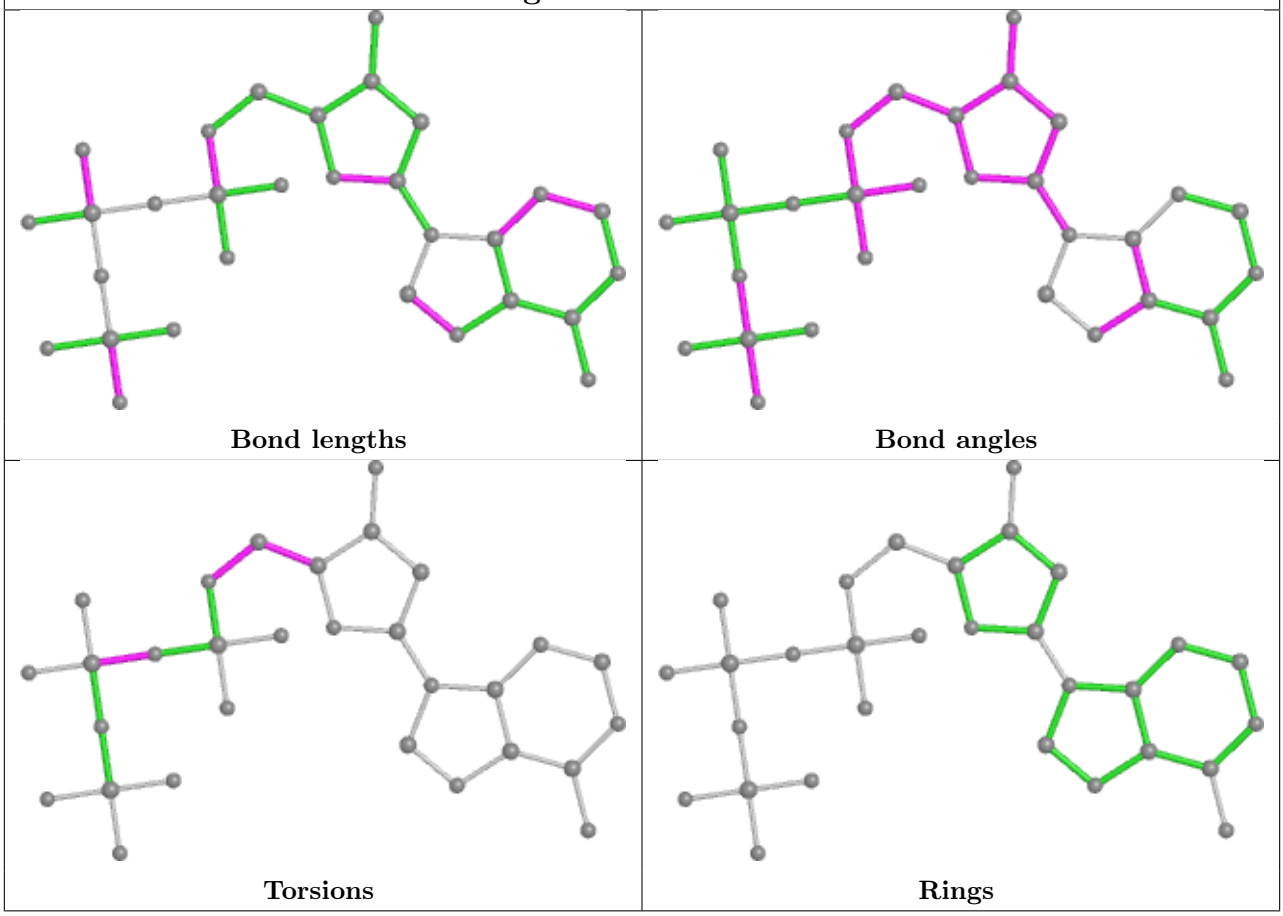


Rings

Ligand DTP B 804



Ligand DTP A 803



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, 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	C	13/13 (100%)	-0.61	0	100	100	29, 43, 53, 58	0
1	E	13/13 (100%)	-0.50	0	100	100	34, 40, 53, 88	0
2	D	16/16 (100%)	-0.44	0	100	100	40, 54, 65, 86	0
2	F	16/16 (100%)	-0.51	0	100	100	36, 43, 66, 111	0
3	A	341/352 (96%)	-0.40	1 (0%)	94	95	21, 40, 59, 69	0
3	B	341/352 (96%)	-0.23	4 (1%)	79	80	29, 49, 69, 90	0
All	All	740/762 (97%)	-0.33	5 (0%)	87	89	21, 44, 66, 111	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	37	PHE	2.9
3	B	38	GLU	2.4
3	B	116	ARG	2.4
3	A	253	ARG	2.3
3	B	119	ARG	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

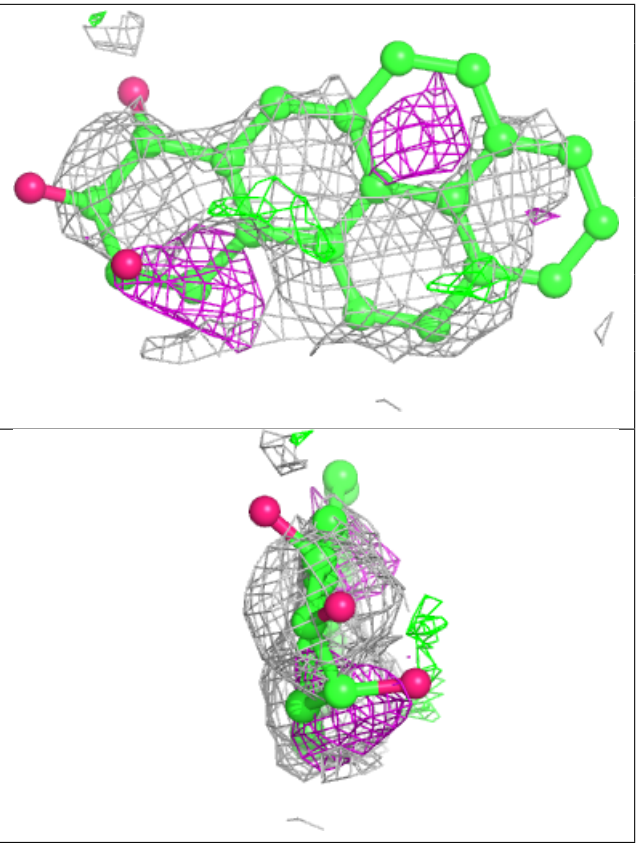
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	BAP	F	1920	23/23	0.64	0.53	80,82,84,84	0
4	BAP	D	1920	23/23	0.86	0.32	57,59,63,66	0
7	DTP	A	803	30/30	0.86	0.20	78,85,91,92	0
5	CA	B	403	1/1	0.89	0.09	61,61,61,61	0
6	MG	A	401	1/1	0.91	0.13	52,52,52,52	0
5	CA	A	402	1/1	0.94	0.06	36,36,36,36	0
7	DTP	B	804	30/30	0.94	0.15	54,56,66,67	0
5	CA	E	408	1/1	0.95	0.27	97,97,97,97	0
5	CA	A	406	1/1	0.96	0.05	61,61,61,61	0
5	CA	B	404	1/1	0.97	0.09	39,39,39,39	0
5	CA	A	407	1/1	0.98	0.03	41,41,41,41	0
5	CA	B	405	1/1	0.99	0.03	54,54,54,54	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.

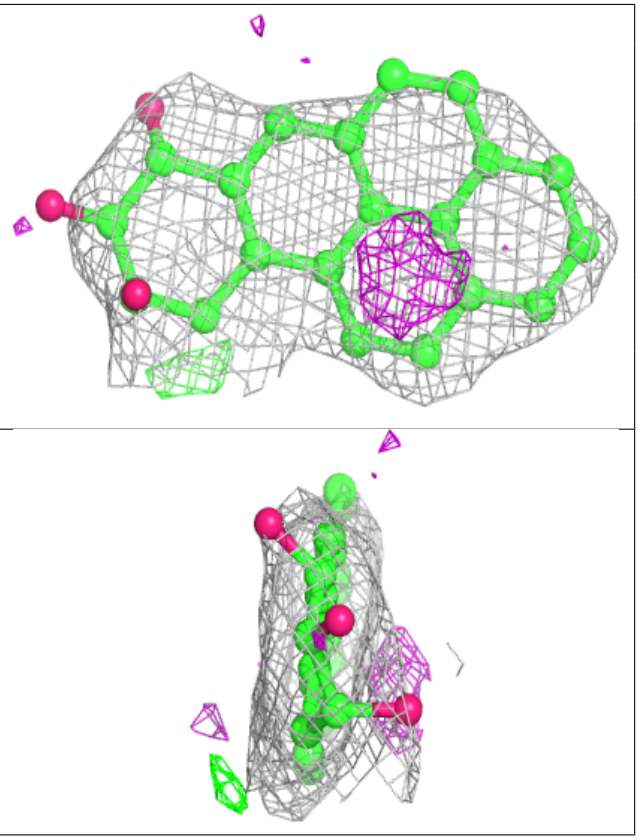
Electron density around BAP F 1920:

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



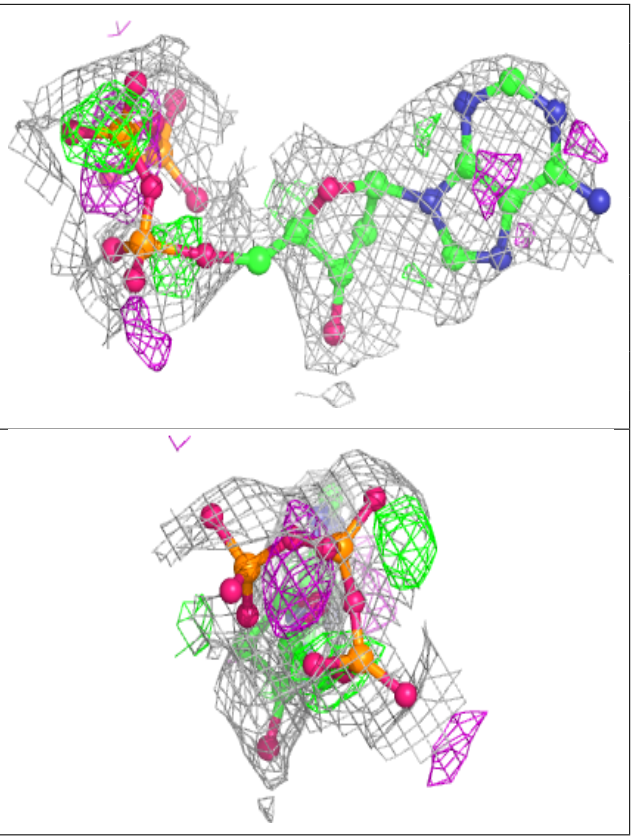
Electron density around BAP D 1920:

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



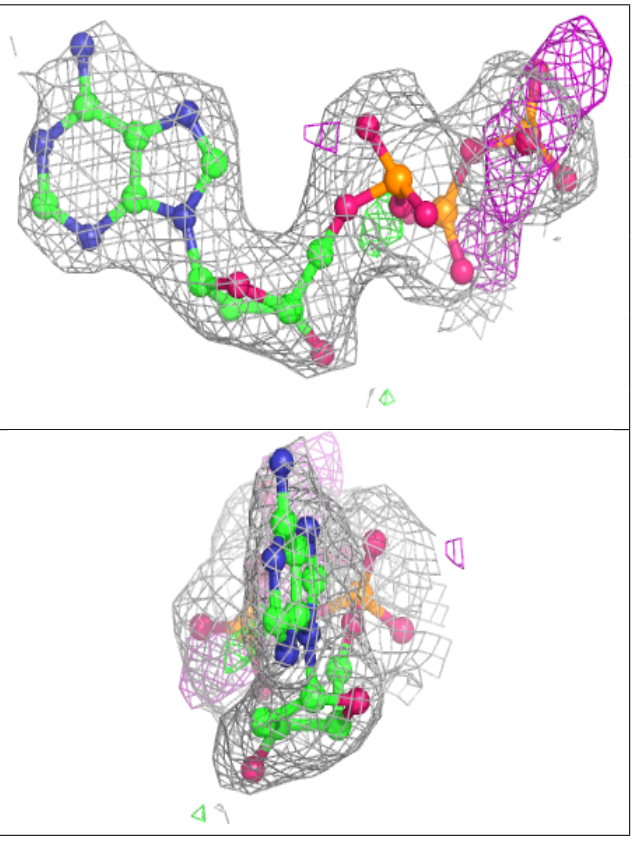
Electron density around DTP A 803:

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



Electron density around DTP B 804:

$2mF_o$-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.