



Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 09:55 am BST

PDB ID : 1SKR
Title : T7 DNA Polymerase Complexed To DNA Primer/Template and ddATP
Authors : Li, Y.; Dutta, S.; Doublié, S.; Bdour, H.M.; Taylor, J.S.; Ellenberger, T.
Deposited on : 2004-03-05
Resolution : 2.40 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 ⓘ 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

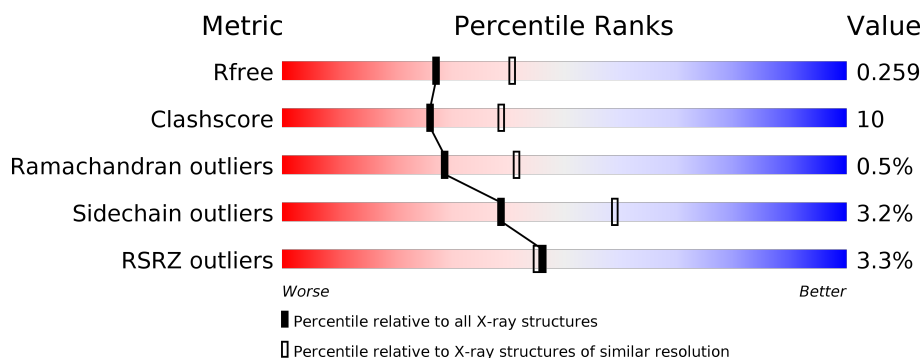
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 ≥ 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	P	22	<div> <div>9%</div> <div>27% 18% 5% 50%</div> </div>
2	T	26	<div> <div>8%</div> <div>35% 15% 50%</div> </div>
3	A	698	<div> <div>3%</div> <div>77% 17% ...</div> </div>
4	B	108	<div> <div>2%</div> <div>64% 31% ...</div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	DAD	A	4004	X	-	-	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7012 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(*CP*GP*AP*AP*AP*AP*CP*GP*AP*C*GP*GP*CP*CP*AP*GP*TP*GP*CP*CP*AP*(2DA))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	11	Total	C	N	O	P	0	0	0
			224	106	44	63	11			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	13	Total	C	N	O	P	0	0	0
			265	126	45	81	13			

- Molecule 3 is a protein called DNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	669	Total	C	N	O	S	0	0	0
			5308	3381	919	985	23			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	DELETION	UNP P00581
A	?	-	ARG	DELETION	UNP P00581
A	?	-	PHE	DELETION	UNP P00581
A	?	-	GLY	DELETION	UNP P00581
A	?	-	SER	DELETION	UNP P00581
A	?	-	HIS	DELETION	UNP P00581

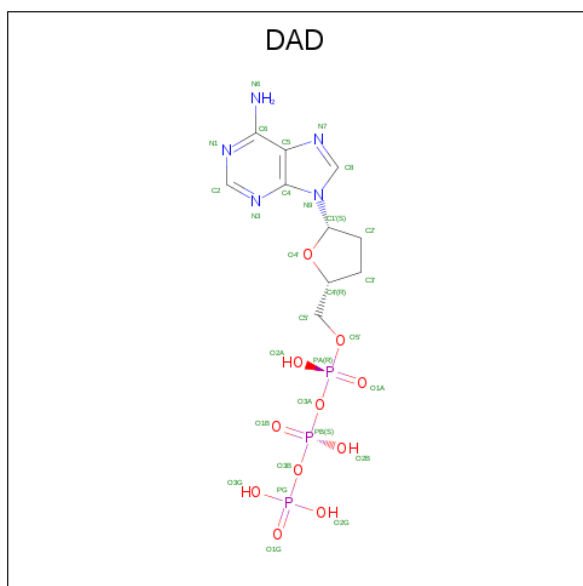
- Molecule 4 is a protein called Thioredoxin 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	105	Total	C	N	O	S	0	0	0
			802	518	129	152	3			

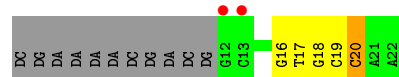
- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	3	Total	Mg	0	0
			3	3		

- Molecule 6 is 2',3'-DIDEOXYADENOSINE-5'-TRIPHOSPHATE (three-letter code: DAD) (formula: C₁₀H₁₆N₅O₁₁P₃).



● Molecule 1: 5'-D(*CP*GP*AP*AP*AP*AP*CP*GP*AP*C*GP*GP*CP*CP*AP*GP*TP*GP*CP*CP*AP*(2DA))-3'

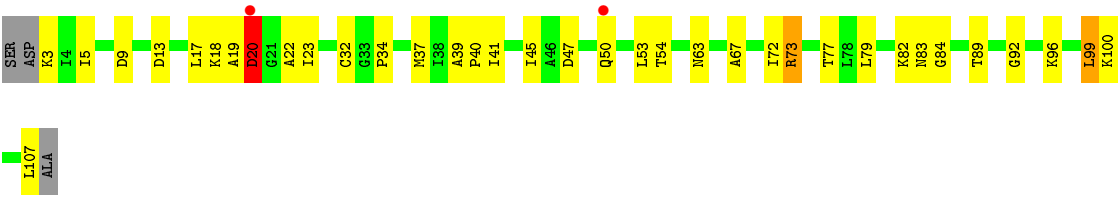


DC DC C3 T4 T5 T6 G7 G8 C15 DC DG DG DT DC DG DT DT DT DT DC DG

ALA	G442	LEU	L185	M1
GLY	V443	D316	V	R35
GLU		T317	P193	P36
GLN	G448	R318	P194	
VAL	R452	E219	E195	R52
R587			T199	G53
I593	G456	G323		F58
	A324	A324		K74
R599	V470	Y326	T204	
	Q471	T327	W222	N78
P606	I474	P328		R79
L624	E480	R339	E228	E80
M632		Q443	F232	R85
R636	Y496	P353	P233	D90
G637	A497		F234	L94
L638	R498	Y356	D235	I98
D645	I500	T357	T236	
	A513	R358	I239	R111
R653	E514	K359	E240	S112
I665	L515	G360	E241	G113
A666	P362	A361	L242	K114
Q667	D519	V364	L246	Y130
R668	T523	V364	G270	R131
V669	F524	E388	T771	
E671	I525	M391	H276	Y130
	Y526	I392	P277	R131
E675	A532	R395	R278	E134
	G533		T279	M135
R684	K536		P284	E138
F686	I537	A400		
		K404	R288	D142
P688	I540			R145
I702	K552	L407	V294	M146
C703	T559	R408	G295	L147
H704		Y409	G296	L147
		E412	I297	E148
	I562		LVS	E149
	A563		LVS	Q150
	E567	G418	LVS	G151
		S419	P80	E152
		V420	LVS	
		N421	ASN	V155
	Q570	P422	LVS	D156
			ALA	
	L573	T431	GLN	M171
	V674	H432	ARG	V172
	E575		GLU	Q173
SER		N436	GLY	
GLN	SER		ARG	V176
TRP	GLN	Q439	GLU	L181
VAL	I440	I440	CYS	L182
	P441	P441	GLU	



WORLD WIDE
PDB
PROTEIN DATA BANK



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	105.51Å 214.02Å 52.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.40 19.97 – 2.25	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.40) 91.9 (19.97-2.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.60 (at 2.26Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.226 , 0.270 0.215 , 0.259	Depositor DCC
R_{free} test set	2556 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å ²)	33.6	Xtriage
Anisotropy	0.171	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 45.5	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	7012	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% 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: DAD, MG, 2DA

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	P	0.52	0/228	0.85	0/349
2	T	0.54	0/295	0.82	0/453
3	A	0.35	0/5438	0.57	0/7364
4	B	0.30	0/817	0.57	0/1108
All	All	0.36	0/6778	0.59	0/9274

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	P	0	1
2	T	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	P	20	DC	Sidechain
2	T	7	DG	Sidechain

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	P	224	0	123	9	0
2	T	265	0	148	3	0
3	A	5308	0	5174	104	0
4	B	802	0	816	26	0
5	A	3	0	0	0	0
6	A	29	0	12	1	0
7	A	323	0	0	7	0
7	B	25	0	0	0	0
7	P	9	0	0	0	0
7	T	24	0	0	0	0
All	All	7012	0	6273	134	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 10.

All (134) 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:111:ARG:H	3:A:111:ARG:HD2	1.11	1.11
3:A:570:GLN:HE22	3:A:606:PRO:HB3	1.41	0.85
3:A:563:ALA:O	3:A:567:GLU:HG3	1.78	0.83
3:A:667:GLN:O	3:A:671:GLU:HG3	1.85	0.77
3:A:111:ARG:CD	3:A:111:ARG:H	1.89	0.76
4:B:39:ALA:HB3	4:B:40:PRO:HD3	1.71	0.71
1:P:16:DG:H2''	1:P:17:DT:C5'	2.22	0.70
3:A:570:GLN:NE2	3:A:606:PRO:HB3	2.07	0.69
3:A:111:ARG:N	3:A:111:ARG:HD2	1.97	0.69
4:B:3:LYS:HD3	4:B:47:ASP:HB3	1.72	0.69
1:P:16:DG:H2''	1:P:17:DT:H5''	1.75	0.67
4:B:13:ASP:OD1	4:B:18:LYS:HE3	1.95	0.66
3:A:35:ARG:HB3	3:A:36:PRO:HD2	1.78	0.65
3:A:195:GLU:H	3:A:195:GLU:CD	1.99	0.65
3:A:496:TYR:HA	3:A:499:GLU:HG2	1.76	0.65
3:A:645:ASP:HB3	3:A:665:ILE:HD13	1.79	0.65
4:B:96:LYS:HZ3	4:B:100:LYS:HD2	1.63	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:480:GLU:CD	6:A:4004:DAD:H2'1	2.17	0.64
4:B:17:LEU:HA	4:B:84:GLY:HA2	1.80	0.64
3:A:391:MET:HE2	3:A:391:MET:C	2.19	0.63
3:A:388:GLU:O	3:A:392:ILE:HG12	1.98	0.63
3:A:35:ARG:HB3	3:A:36:PRO:CD	2.29	0.63
3:A:391:MET:HE2	3:A:391:MET:O	1.99	0.63
3:A:79:ARG:HD3	7:A:5069:HOH:O	2.00	0.62
3:A:343:GLN:HG3	3:A:362:PRO:HG3	1.81	0.61
3:A:204:THR:HG23	7:A:5022:HOH:O	2.00	0.61
3:A:1:MET:HE1	3:A:199:THR:HG22	1.83	0.60
3:A:391:MET:HE1	3:A:392:ILE:HD13	1.85	0.59
3:A:94:LEU:HB3	3:A:185:LEU:HD13	1.84	0.59
3:A:142:ASP:O	3:A:146:MET:HG3	2.03	0.58
1:P:16:DG:C2'	1:P:17:DT:H5''	2.33	0.58
3:A:131:ARG:HA	3:A:131:ARG:HE	1.68	0.58
4:B:96:LYS:NZ	4:B:100:LYS:HD2	2.19	0.57
3:A:343:GLN:HG3	3:A:362:PRO:CG	2.34	0.57
3:A:536:LYS:O	3:A:540:ILE:HG12	2.05	0.57
3:A:145:ARG:O	3:A:149:GLU:HG3	2.04	0.57
4:B:3:LYS:CD	4:B:47:ASP:HB3	2.35	0.57
3:A:173:GLN:HA	3:A:176:VAL:CG2	2.35	0.56
4:B:77:THR:HG22	4:B:79:LEU:HD13	1.87	0.56
3:A:173:GLN:HA	3:A:176:VAL:HG22	1.86	0.56
1:P:16:DG:H4'	3:A:359:LYS:NZ	2.21	0.56
3:A:533:GLY:O	3:A:537:ILE:HG12	2.06	0.55
3:A:573:LEU:HD11	3:A:593:ILE:HD11	1.87	0.55
3:A:323:GLY:O	3:A:325:PRO:HD3	2.07	0.55
3:A:519:ASP:O	3:A:523:THR:HG22	2.07	0.55
3:A:559:THR:HG22	3:A:562:ILE:HG13	1.89	0.54
4:B:67:ALA:HB1	4:B:72:ILE:HG13	1.89	0.53
2:T:4:DT:H2'	2:T:5:DT:C6	2.43	0.53
3:A:234:PHE:CZ	3:A:239:ILE:HG13	2.42	0.53
4:B:32:CYS:SG	4:B:34:PRO:HD2	2.48	0.53
4:B:23:ILE:HD13	4:B:54:THR:HB	1.91	0.52
3:A:134:GLU:O	3:A:138:GLU:HG3	2.09	0.52
3:A:448:GLY:O	3:A:452:ARG:HB2	2.10	0.52
3:A:150:GLN:HB2	3:A:152:GLU:OE1	2.11	0.51
3:A:326:TYR:HB3	4:B:92:GLY:HA2	1.93	0.51
4:B:23:ILE:CD1	4:B:54:THR:HB	2.41	0.51
3:A:85:ARG:HD3	7:A:5248:HOH:O	2.10	0.50
4:B:100:LYS:HB3	4:B:100:LYS:NZ	2.26	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:145:ARG:HH11	3:A:145:ARG:HG3	1.77	0.50
3:A:624:LEU:HD12	3:A:684:TRP:CH2	2.45	0.50
2:T:8:DG:H4'	3:A:432:HIS:O	2.12	0.50
3:A:1:MET:CE	3:A:182:LEU:HD21	2.42	0.50
3:A:319:GLU:N	3:A:319:GLU:OE1	2.38	0.49
3:A:79:ARG:HG2	3:A:80:GLU:N	2.27	0.49
3:A:94:LEU:O	3:A:98:ILE:HG12	2.12	0.49
3:A:474:ILE:N	3:A:474:ILE:HD12	2.28	0.49
1:P:16:DG:H2''	1:P:17:DT:H5'	1.93	0.49
1:P:16:DG:H4'	3:A:359:LYS:HZ3	1.77	0.48
3:A:1:MET:CE	3:A:199:THR:HG22	2.43	0.48
3:A:698:PRO:HD2	3:A:702:ILE:HD12	1.95	0.48
4:B:45:ILE:HG13	4:B:99:LEU:HD13	1.96	0.48
3:A:499:GLU:HB2	7:A:5431:HOH:O	2.14	0.48
3:A:130:TYR:CZ	3:A:134:GLU:HG3	2.49	0.47
4:B:37:MET:O	4:B:40:PRO:HD2	2.13	0.47
3:A:443:VAL:HG13	3:A:452:ARG:HG2	1.97	0.47
3:A:391:MET:HE3	3:A:395:ARG:HG3	1.97	0.47
3:A:233:PRO:HB2	3:A:456:GLY:O	2.14	0.47
3:A:440:ILE:O	3:A:452:ARG:NH1	2.48	0.47
3:A:456:GLY:HA2	3:A:471:GLN:OE1	2.15	0.46
1:P:19:DC:C2	1:P:20:DC:C5	3.04	0.46
3:A:155:VAL:O	3:A:156:ASP:C	2.54	0.46
3:A:228:GLU:HG2	3:A:418:GLY:O	2.15	0.45
3:A:173:GLN:O	3:A:176:VAL:HG23	2.16	0.45
1:P:18:DG:H4'	3:A:339:ARG:NH2	2.31	0.45
3:A:421:ASN:HB3	3:A:431:THR:OG1	2.16	0.45
4:B:19:ALA:O	4:B:20:ASP:C	2.55	0.45
3:A:328:PRO:HB3	4:B:73:ARG:HH22	1.81	0.45
3:A:497:ALA:O	3:A:500:ILE:HG22	2.17	0.45
3:A:671:GLU:O	3:A:675:GLU:HG3	2.16	0.45
3:A:145:ARG:CZ	3:A:145:ARG:HB2	2.47	0.44
3:A:236:THR:O	3:A:240:GLU:HG3	2.17	0.44
3:A:364:VAL:O	3:A:364:VAL:HG12	2.17	0.44
3:A:400:ALA:CB	3:A:407:LEU:HD12	2.48	0.44
3:A:181:LEU:HD22	3:A:185:LEU:HD11	1.99	0.44
3:A:52:ARG:NH2	7:A:5135:HOH:O	2.50	0.44
3:A:441:PRO:O	3:A:452:ARG:HD3	2.18	0.44
3:A:559:THR:O	3:A:559:THR:HG22	2.16	0.44
3:A:353:PRO:HB2	3:A:356:TYR:CZ	2.52	0.43
3:A:665:ILE:O	3:A:669:VAL:HG23	2.19	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:41:ILE:HD12	4:B:96:LYS:HB2	2.00	0.43
3:A:276:HIS:HD2	3:A:277:PRO:HD2	1.84	0.43
4:B:82:LYS:NZ	4:B:83:ASN:ND2	2.66	0.43
3:A:532:ALA:HB3	3:A:537:ILE:HD11	2.02	0.42
3:A:242:LEU:O	3:A:246:LEU:HG	2.20	0.42
3:A:470:VAL:HG12	3:A:471:GLN:N	2.34	0.42
3:A:420:VAL:O	3:A:422:PRO:HD3	2.20	0.42
4:B:9:ASP:CG	4:B:63:ASN:HB3	2.40	0.42
3:A:400:ALA:HB2	3:A:407:LEU:HD12	2.01	0.42
3:A:276:HIS:CD2	3:A:277:PRO:HD2	2.55	0.41
3:A:74:LYS:O	3:A:78:ASN:HA	2.20	0.41
3:A:135:MET:HE1	3:A:138:GLU:OE2	2.21	0.41
3:A:193:PRO:HA	3:A:194:PRO:HD3	1.78	0.41
3:A:436:ASN:O	3:A:439:GLN:HG2	2.20	0.41
1:P:17:DT:OP1	3:A:359:LYS:HB3	2.21	0.41
4:B:53:LEU:HG	4:B:54:THR:N	2.35	0.41
3:A:145:ARG:HH11	3:A:145:ARG:CG	2.32	0.41
3:A:395:ARG:HD3	7:A:5091:HOH:O	2.20	0.41
4:B:82:LYS:NZ	4:B:83:ASN:HD21	2.19	0.41
3:A:270:GLY:HA3	3:A:288:ARG:HB2	2.02	0.41
3:A:85:ARG:HG3	3:A:222:TRP:CG	2.55	0.41
3:A:1:MET:HE3	3:A:182:LEU:HD21	2.02	0.40
3:A:638:LEU:N	3:A:638:LEU:HD12	2.35	0.40
3:A:58:PHE:O	3:A:90:ASP:HA	2.21	0.40
2:T:4:DT:H5"	3:A:532:ALA:HA	2.03	0.40
3:A:53:GLY:HA2	7:A:5103:HOH:O	2.22	0.40
3:A:638:LEU:N	3:A:638:LEU:CD1	2.84	0.40
4:B:22:ALA:O	4:B:53:LEU:HD12	2.21	0.40
3:A:288:ARG:HD3	3:A:288:ARG:HA	1.96	0.40
3:A:513:ALA:HB3	3:A:515:LEU:HG	2.03	0.40
3:A:632:MET:O	3:A:636:LYS:HG3	2.21	0.40
3:A:525:ILE:HG23	3:A:526:TYR:N	2.36	0.40
4:B:5:ILE:HD12	4:B:54:THR:HG21	2.03	0.40
3:A:404:LYS:HA	3:A:409:TYR:HE2	1.86	0.40
4:B:34:PRO:O	4:B:37:MET:HB2	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	A	663/698 (95%)	642 (97%)	19 (3%)	2 (0%)	41	55
4	B	103/108 (95%)	98 (95%)	3 (3%)	2 (2%)	8	10
All	All	766/806 (95%)	740 (97%)	22 (3%)	4 (0%)	29	41

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	156	ASP
4	B	20	ASP
3	A	653	HIS
4	B	50	GLN

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	547/579 (94%)	532 (97%)	15 (3%)	44	65
4	B	85/87 (98%)	80 (94%)	5 (6%)	19	32
All	All	632/666 (95%)	612 (97%)	20 (3%)	39	59

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	85	ARG
3	A	111	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	131	ARG
3	A	142	ASP
3	A	171	ASN
3	A	181	LEU
3	A	204	THR
3	A	232	PHE
3	A	343	GLN
3	A	391	MET
3	A	452	ARG
3	A	523	THR
3	A	552	LYS
3	A	599	ARG
3	A	686	PHE
4	B	20	ASP
4	B	73	ARG
4	B	89	THR
4	B	99	LEU
4	B	107	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	276	HIS
3	A	343	GLN
3	A	347	GLN
3	A	570	GLN
4	B	50	GLN
4	B	62	GLN
4	B	83	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	2DA	P	22	1,2	17,22,23	0.69	0	13,31,34	0.88	1 (7%)

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	2DA	P	22	1,2	-	0/3/18/19	0/3/3/3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	22	2DA	C5-C6-N6	2.24	123.76	120.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 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	$\# Z > 2$
6	DAD	A	4004	5	25,31,31	0.79	0	26,48,48	1.00	2 (7%)

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
6	DAD	A	4004	5	2/2/5/5	1/18/31/31	0/3/3/3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
6	A	4004	DAD	O2G-PG-O1G	2.32	119.78	110.68
6	A	4004	DAD	C5-C6-N6	2.19	123.68	120.35

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	4004	DAD	C4'
6	A	4004	DAD	C1'

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	4004	DAD	PG-O3B-PB-O2B

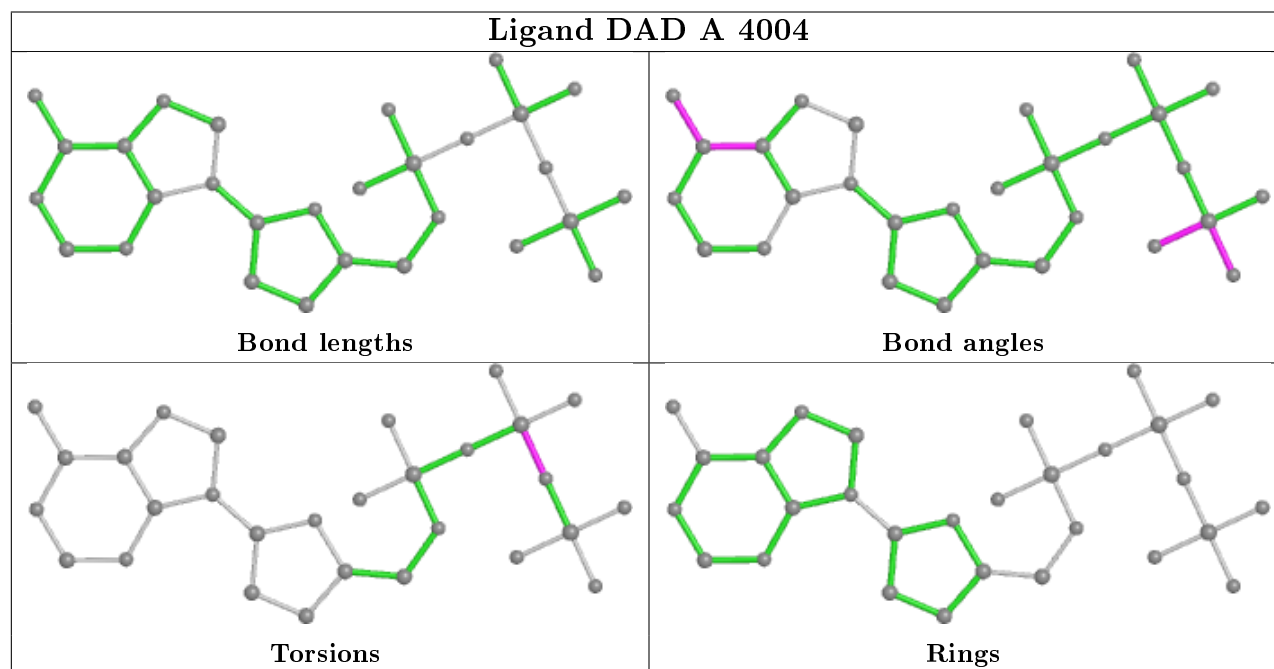
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	4004	DAD	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	P	10/22 (45%)	0.23	2 (20%) 1 0	24, 39, 69, 72	0
2	T	13/26 (50%)	-0.33	2 (15%) 2 1	23, 31, 55, 65	0
3	A	669/698 (95%)	-0.25	20 (2%) 50 49	18, 31, 55, 65	0
4	B	105/108 (97%)	0.10	2 (1%) 66 64	26, 43, 56, 60	0
All	All	797/854 (93%)	-0.20	26 (3%) 46 45	18, 33, 56, 72	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	360	GLY	5.3
3	A	156	ASP	4.9
4	B	20	ASP	4.4
3	A	279	THR	3.7
3	A	271	THR	3.7
3	A	358	ASP	3.6
3	A	294	VAL	3.4
3	A	317	THR	3.4
3	A	277	PRO	3.4
1	P	12	DG	3.2
3	A	114	LYS	3.0
3	A	359	LYS	2.7
1	P	13	DC	2.6
4	B	50	GLN	2.6
3	A	278	ARG	2.6
3	A	284	PRO	2.5
3	A	150	GLN	2.5
3	A	295	GLY	2.4
3	A	113	GLY	2.4
2	T	3	DC	2.3
3	A	111	ARG	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	A	151	GLY	2.3
3	A	361	ALA	2.3
3	A	412	GLU	2.2
2	T	15	DC	2.2
3	A	148	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	2DA	P	22	20/21	0.98	0.10	22,25,29,29	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

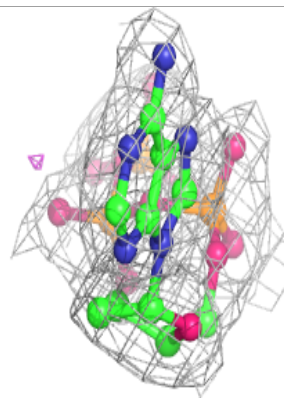
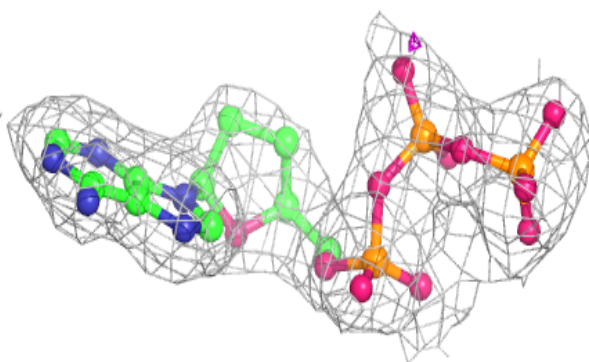
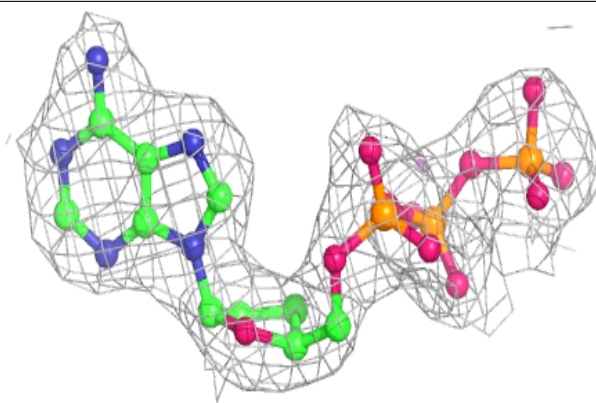
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	MG	A	4003	1/1	0.89	0.26	30,30,30,30	0
5	MG	A	4002	1/1	0.90	0.13	33,33,33,33	0
5	MG	A	4001	1/1	0.97	0.06	27,27,27,27	0
6	DAD	A	4004	29/29	0.98	0.09	27,29,32,33	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 DAD A 4004:

$2mF_o-DF_c$ (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.