Full wwPDB X-ray Structure Validation Report (i)

Sep 12, 2024 – 04:04 PM EDT

PDB ID	:	9CJ9
Title	:	Crystal structure of human polymerase eta with incoming dAMPnPP nucleotide opposite O4-methyl threofuranosyl thymidine in DNA template at insertion stage
Authors	:	Tomar, R.; Stone, M.P.; Egli, M.
Deposited on	:	2024-07-05
Resolution	:	2.98 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:

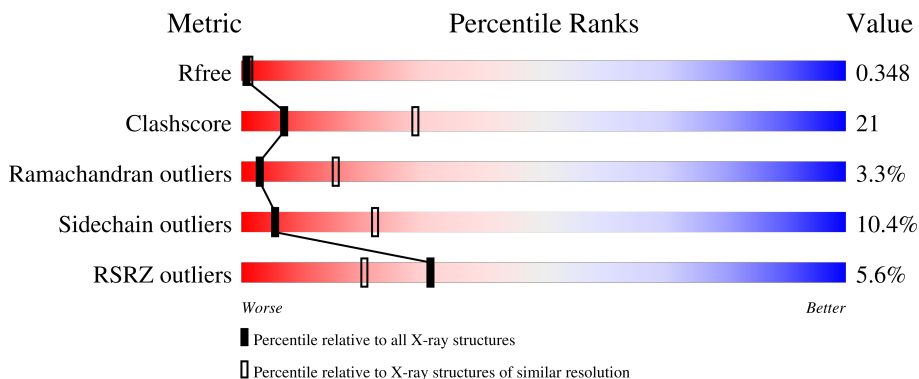|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.3

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 2.98 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	Percentile Ranks	Value
Rfree		0.348
Clashscore		21
Ramachandran outliers		3.3%
Sidechain outliers		10.4%
RSRZ outliers		5.6%

Worse → *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3360 (3.00-2.96)
Clashscore	180529	3751 (3.00-2.96)
Ramachandran outliers	177936	3628 (3.00-2.96)
Sidechain outliers	177891	3631 (3.00-2.96)
RSRZ outliers	164620	3372 (3.00-2.96)

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	A	435	6% — 51% 40% 7% ·
2	T	12	8% — 50% 42% 8%
3	P	8	38% 62%

WORLDWIDE PDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 3803 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA polymerase eta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	427	Total 3350	C 2100	N 601	O 626	S 23	11	2	0

There are 3 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q9Y253
A	-1	PRO	-	expression tag	UNP Q9Y253
A	0	HIS	-	expression tag	UNP Q9Y253

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*AP*TP*(TFT)P*AP*TP*GP*AP*CP* GP*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	12	Total 226	C 108	N 39	O 68	P 11	0	0	1

- Molecule 3 is a DNA chain called DNA (5'-D(*AP*GP*CP*GP*TP*CP*AP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	8	Total 161	C 78	N 30	O 46	P 7	0	0	0

- Molecule 4 is 2'-deoxy-5'-O-[(R)-hydroxy{[(R)-hydroxy(phosphonooxy)phosphoryl]am ino}phosphoryl]adenosine (three-letter code: DZ4) (formula: $C_{10}H_{17}N_6O_{11}P_3$) (labeled as "Ligand of Interest" by depositor).

DZ4

Mol | Chain | Residues | Atoms | ZeroOcc | AltConf
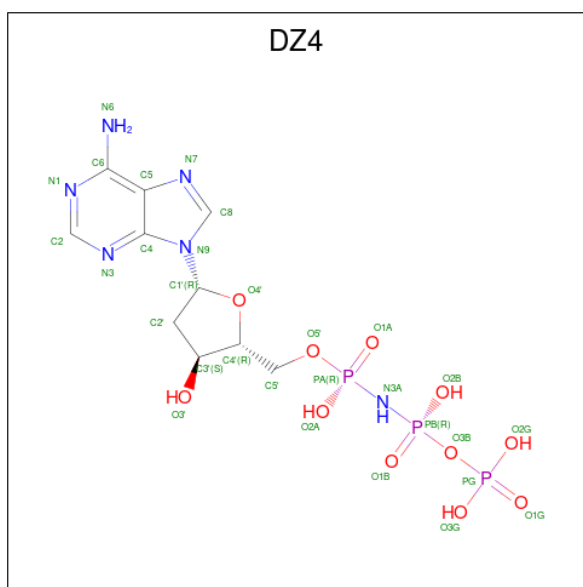
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			30	10	6	11	3		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Mg	0	0
			2	2		

- Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	30	Total	O	0	0
			30	30		
6	T	3	Total	O	0	0
			3	3		
6	P	1	Total	O	0	0
			1	1		

3 Residue-property plots ⓘ

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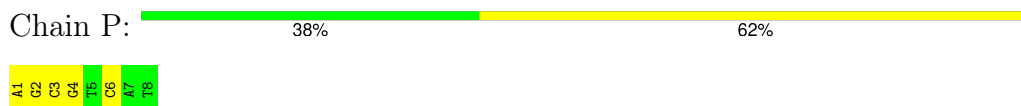
• Molecule 1: DNA polymerase eta

Chain A:
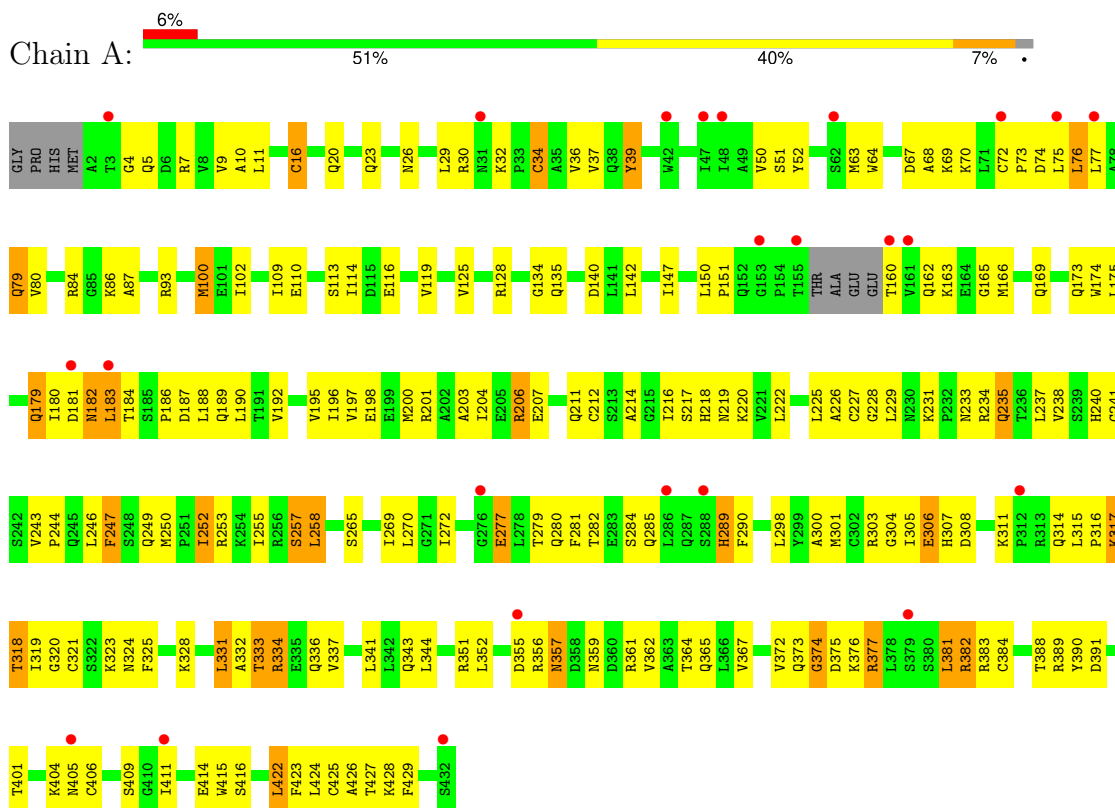
• Molecule 2: DNA (5'-D(*CP*AP*TP*(TFT)P*AP*TP*GP*AP*CP*GP*CP*T)-3')

Chain T:
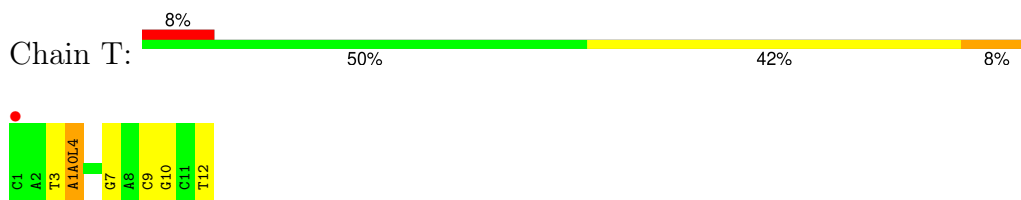
• Molecule 3: DNA (5'-D(*AP*GP*CP*GP*TP*CP*AP*T)-3')

Chain P:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	98.08Å 98.08Å 78.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.61 – 2.98 41.61 – 2.98	Depositor EDS
% Data completeness (in resolution range)	95.4 (41.61-2.98) 95.4 (41.61-2.98)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	1.22 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.220 , 0.348 0.221 , 0.348	Depositor DCC
R_{free} test set	786 reflections (4.56%)	wwPDB-VP
Wilson B-factor (Å^2)	29.8	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 49.8	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.51, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.048 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	3803	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.15% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: A1A0L, DZ4, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z| > 5$	RMSZ	#$|Z| > 5$
1	A	0.47	0/3413	0.68	1/4607 (0.0%)
2	T	1.00	0/229	0.99	0/350
3	P	0.96	0/180	1.09	1/276 (0.4%)
All	All	0.55	0/3822	0.73	2/5233 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	P	6	DC	O4'-C1'-N1	5.23	111.66	108.00
1	A	246	LEU	CA-CB-CG	5.02	126.84	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3350	0	3393	143	2
2	T	226	0	114	6	0
3	P	161	0	92	4	0
4	A	30	0	13	3	0
5	A	2	0	0	0	0
6	A	30	0	0	5	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	P	1	0	0	0	0
6	T	3	0	0	0	0
All	All	3803	0	3612	151	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 21.

All (151) 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)
2:T:4:A1A0L:O4'	2:T:4:A1A0L:C1'	1.63	1.19
1:A:128:ARG:HG3	1:A:196:ILE:HD11	1.53	0.90
1:A:4:GLY:HA3	1:A:305:ILE:HG12	1.56	0.87
1:A:319:ILE:HD12	1:A:352:LEU:HD23	1.61	0.81
1:A:351:ARG:NH2	2:T:7:DG:OP1	2.13	0.81
1:A:333:THR:HG22	1:A:336:GLN:H	1.45	0.80
1:A:179:GLN:HG2	1:A:182:ASN:HB3	1.71	0.73
1:A:217:SER:OG	1:A:218:HIS:N	2.20	0.72
1:A:52:TYR:OH	4:A:501:DZ4:O1G	2.06	0.72
1:A:5:GLN:HE21	1:A:240:HIS:CG	2.10	0.70
1:A:9:VAL:HG11	1:A:220:LYS:HG2	1.75	0.69
1:A:5:GLN:HB3	1:A:186:PRO:HB3	1.75	0.68
1:A:142:LEU:HD11	1:A:195:VAL:HG22	1.75	0.68
1:A:315:LEU:HD13	1:A:359:ASN:HB3	1.75	0.68
1:A:389:ARG:HD2	1:A:391:ASP:HB2	1.76	0.66
1:A:125:VAL:HG22	1:A:192:VAL:HG12	1.78	0.66
1:A:226:ALA:HA	1:A:229:LEU:HD12	1.79	0.65
1:A:427:THR:HG22	1:A:428:LYS:H	1.61	0.65
1:A:219:ASN:HD21	1:A:222:LEU:HD12	1.61	0.65
1:A:86:LYS:HB3	1:A:323:LYS:HD3	1.81	0.62
1:A:26:ASN:HB3	1:A:29:LEU:HG	1.82	0.61
1:A:272:ILE:HD13	1:A:277:GLU:HB2	1.82	0.61
1:A:52:TYR:OH	1:A:231:LYS:HD3	1.99	0.61
1:A:68:ALA:O	1:A:72:CYS:N	2.32	0.60
1:A:306:GLU:HG2	1:A:308:ASP:H	1.67	0.60
1:A:289:HIS:HB2	1:A:290:PHE:CE2	2.38	0.59
1:A:423:PHE:CZ	1:A:425:CYS:HB2	2.38	0.58
1:A:113:SER:HB3	1:A:116:GLU:HG2	1.86	0.58
1:A:190:LEU:HD13	1:A:218:HIS:CE1	2.40	0.57
1:A:147:ILE:HG21	1:A:150:LEU:HD12	1.87	0.57
1:A:37:VAL:HG13	1:A:79:GLN:HA	1.87	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:258:LEU:HD21	1:A:298:LEU:HD21	1.86	0.56
1:A:181:ASP:O	1:A:183:LEU:N	2.39	0.56
1:A:318:THR:HG22	1:A:427:THR:HG23	1.88	0.56
1:A:228:GLY:HA2	1:A:231:LYS:HE3	1.88	0.56
1:A:317:LYS:HG3	1:A:318:THR:HG23	1.86	0.56
1:A:332:ALA:HA	1:A:415:TRP:HB3	1.89	0.55
1:A:364:THR:HG22	1:A:428:LYS:O	2.07	0.55
1:A:26:ASN:HB2	6:A:612:HOH:O	2.06	0.55
1:A:32:LYS:HB3	1:A:76:LEU:HD13	1.90	0.54
1:A:257:SER:HB2	1:A:258:LEU:HD12	1.89	0.54
1:A:86:LYS:NZ	1:A:321[A]:CYS:SG	2.69	0.53
1:A:320:GLY:O	1:A:351:ARG:NH1	2.42	0.53
1:A:318:THR:HA	1:A:426:ALA:O	2.09	0.52
1:A:135:GLN:O	6:A:601:HOH:O	2.18	0.52
1:A:125:VAL:HG11	1:A:189:GLN:HB3	1.91	0.52
1:A:174:TRP:CH2	1:A:187:ASP:HB3	2.45	0.52
1:A:272:ILE:CD1	1:A:277:GLU:HB2	2.41	0.51
1:A:102:ILE:HD11	1:A:207:GLU:HB2	1.93	0.51
2:T:9:DC:H2''	2:T:10:DG:C8	2.45	0.51
1:A:10:ALA:HB3	1:A:119:VAL:HG13	1.93	0.51
1:A:93:ARG:HA	1:A:114:ILE:HD11	1.93	0.51
1:A:427:THR:HG22	1:A:428:LYS:N	2.26	0.50
1:A:301:MET:O	1:A:304:GLY:N	2.42	0.50
1:A:241:GLY:O	1:A:244:PRO:HD2	2.11	0.50
1:A:253:ARG:HH11	1:A:253:ARG:HB3	1.77	0.49
1:A:23:GLN:HG2	1:A:29:LEU:HB2	1.94	0.49
1:A:174:TRP:HD1	1:A:175:LEU:HD23	1.78	0.49
1:A:365:GLN:HB2	1:A:428:LYS:HD3	1.94	0.49
1:A:64:TRP:HB2	1:A:67:ASP:OD1	2.12	0.49
1:A:252:ILE:CG2	1:A:258:LEU:HD23	2.43	0.49
1:A:333:THR:O	1:A:337:VAL:HG23	2.12	0.49
3:P:3:DC:H2''	3:P:4:DG:H8	1.78	0.48
2:T:3:DT:H2'	2:T:4:A1A0L:C6	2.43	0.48
2:T:3:DT:H2'	2:T:4:A1A0L:C5	2.44	0.47
1:A:352:LEU:HD13	1:A:390:TYR:CE1	2.48	0.47
1:A:315:LEU:HD13	1:A:359:ASN:CB	2.44	0.47
1:A:409:SER:HB2	1:A:416:SER:HB3	1.95	0.47
2:T:12:DT:O2	3:P:1:DA:H2	1.96	0.47
1:A:162:GLN:HG3	1:A:165:GLY:H	1.79	0.47
1:A:39:TYR:OH	1:A:86:LYS:HA	2.15	0.47
1:A:102:ILE:HD13	1:A:204:ILE:HA	1.96	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:411:ILE:HG22	1:A:414:GLU:H	1.80	0.46
1:A:69:LYS:NZ	1:A:73:PRO:O	2.48	0.46
1:A:84:ARG:HD2	1:A:343:GLN:OE1	2.14	0.46
1:A:226:ALA:HA	1:A:229:LEU:CD1	2.44	0.46
1:A:214:ALA:O	1:A:227:CYS:HA	2.15	0.46
1:A:357:ASN:OD1	1:A:357:ASN:N	2.48	0.46
1:A:228:GLY:N	6:A:603:HOH:O	2.48	0.46
1:A:321[A]:CYS:SG	1:A:344:LEU:HD22	2.55	0.46
1:A:23:GLN:HG2	1:A:29:LEU:CB	2.46	0.45
1:A:160:THR:OG1	1:A:162:GLN:HG2	2.16	0.45
1:A:198:GLU:HG2	1:A:235:GLN:NE2	2.32	0.45
1:A:203:ALA:HA	1:A:206:ARG:HB2	1.98	0.45
1:A:270:LEU:HB3	1:A:272:ILE:HG13	1.97	0.45
4:A:501:DZ4:O2A	4:A:501:DZ4:H8	2.15	0.45
1:A:279:THR:O	1:A:303:ARG:NH2	2.48	0.45
1:A:382:ARG:O	1:A:383:ARG:HD3	2.15	0.45
1:A:16:CYS:HA	4:A:501:DZ4:O2B	2.17	0.45
1:A:110:GLU:HB2	1:A:311:LYS:O	2.17	0.45
1:A:5:GLN:HA	1:A:218:HIS:CG	2.52	0.45
1:A:67:ASP:HA	1:A:70:LYS:HD2	2.00	0.44
1:A:163:LYS:HB3	1:A:163:LYS:HE3	1.58	0.44
1:A:318:THR:HG22	1:A:427:THR:CG2	2.47	0.44
1:A:361:ARG:HG3	1:A:429:PHE:CD1	2.53	0.44
1:A:375:ASP:OD2	1:A:377:ARG:HB2	2.18	0.44
1:A:37:VAL:HG13	1:A:79:GLN:HG2	2.00	0.44
1:A:151:PRO:HG3	1:A:169:GLN:HG3	2.00	0.44
1:A:281:PHE:HA	1:A:285:GLN:OE1	2.18	0.44
1:A:86:LYS:HB2	1:A:87:ALA:H	1.60	0.44
1:A:169:GLN:O	1:A:173:GLN:HG3	2.18	0.44
1:A:69:LYS:HD2	1:A:69:LYS:HA	1.81	0.43
1:A:147:ILE:HD12	1:A:237:LEU:HD22	2.00	0.43
3:P:3:DC:H2"	3:P:4:DG:C8	2.53	0.43
1:A:325:PHE:CD1	1:A:331:LEU:HD21	2.53	0.43
1:A:243:VAL:HB	1:A:244:PRO:HD3	2.00	0.43
1:A:201:ARG:NH1	1:A:212:CYS:O	2.52	0.43
1:A:20:GLN:OE1	1:A:211:GLN:N	2.38	0.43
1:A:306:GLU:HG2	1:A:307:HIS:N	2.32	0.43
1:A:150:LEU:HD22	1:A:173:GLN:OE1	2.18	0.42
1:A:34:CYS:HB3	1:A:76:LEU:O	2.19	0.42
1:A:100:MET:HG3	1:A:314:GLN:HE21	1.84	0.42
1:A:183:LEU:HA	1:A:188:LEU:HD23	2.00	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:197:VAL:HG12	1:A:235:GLN:OE1	2.19	0.42
1:A:289:HIS:HB2	1:A:290:PHE:CD2	2.53	0.42
1:A:372:VAL:O	1:A:374:GLY:N	2.52	0.42
1:A:219:ASN:HB2	1:A:306:GLU:OE2	2.18	0.42
1:A:308:ASP:HB2	6:A:615:HOH:O	2.20	0.42
1:A:37:VAL:CG1	1:A:79:GLN:HG2	2.49	0.42
1:A:314:GLN:HB2	6:A:619:HOH:O	2.19	0.42
1:A:102:ILE:HG21	1:A:204:ILE:HG12	2.01	0.42
1:A:225:LEU:O	1:A:229:LEU:HG	2.20	0.42
1:A:162:GLN:O	1:A:166:MET:HG2	2.20	0.42
1:A:280:GLN:H	1:A:280:GLN:HG2	1.70	0.42
1:A:334:ARG:NH1	1:A:404:LYS:HE3	2.34	0.42
1:A:63:MET:SD	1:A:68:ALA:HB2	2.60	0.41
1:A:186:PRO:HA	1:A:189:GLN:OE1	2.20	0.41
1:A:200:MET:O	1:A:204:ILE:HG13	2.20	0.41
1:A:341:LEU:HG	1:A:422:LEU:HD21	2.02	0.41
1:A:427:THR:CG2	1:A:428:LYS:H	2.32	0.41
1:A:181:ASP:O	1:A:183:LEU:HD12	2.20	0.41
1:A:404:LYS:C	1:A:406:CYS:H	2.22	0.41
1:A:74:ASP:O	1:A:76:LEU:HD12	2.20	0.41
1:A:282:THR:O	1:A:285:GLN:HB3	2.21	0.41
1:A:300:ALA:HB1	1:A:305:ILE:HB	2.02	0.41
1:A:314:GLN:OE1	1:A:315:LEU:HD23	2.21	0.41
1:A:318:THR:HB	1:A:425:CYS:SG	2.61	0.41
1:A:352:LEU:HB3	1:A:390:TYR:CE2	2.56	0.41
1:A:50:VAL:HG12	1:A:51:SER:O	2.20	0.41
1:A:72:CYS:SG	1:A:75:LEU:HB2	2.61	0.41
1:A:255:ILE:HB	1:A:258:LEU:HD22	2.03	0.41
1:A:247:PHE:HA	1:A:250:MET:HB2	2.03	0.41
1:A:265:SER:O	1:A:269:ILE:HB	2.21	0.41
1:A:384:CYS:HB3	3:P:2:DG:OP2	2.21	0.41
1:A:197:VAL:HG11	1:A:216:ILE:HD11	2.03	0.41
1:A:134:GLY:O	1:A:180:ILE:HD13	2.22	0.40
1:A:174:TRP:CD1	1:A:175:LEU:HD23	2.56	0.40
1:A:404:LYS:HD2	1:A:404:LYS:O	2.20	0.40
1:A:201:ARG:HD3	1:A:214:ALA:HB2	2.04	0.40
1:A:36:VAL:HG12	1:A:80:VAL:HG23	2.03	0.40
1:A:337:VAL:O	1:A:341:LEU:HB2	2.22	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:182:ASN:OD1	1:A:376:LYS:NZ[3_454]	2.10	0.10
1:A:249:GLN:O	1:A:356:ARG:NH2[4_455]	2.13	0.07

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	
1	A	425/435 (98%)	351 (83%)	60 (14%)	14 (3%)	3	15

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	182	ASN
1	A	183	LEU
1	A	257	SER
1	A	318	THR
1	A	16	CYS
1	A	233	ASN
1	A	316	PRO
1	A	373	GLN
1	A	405	ASN
1	A	39	TYR
1	A	333	THR
1	A	76	LEU
1	A	381	LEU
1	A	374	GLY

5.3.2 Protein sidechains ⓘ

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

The Analysed column shows the number of residues for which the sidechain conformation was

analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	368/372 (99%)	330 (90%)	38 (10%)	6 22

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	ARG
1	A	11	LEU
1	A	30	ARG
1	A	34	CYS
1	A	77	LEU
1	A	79	GLN
1	A	100	MET
1	A	109	ILE
1	A	140	ASP
1	A	179	GLN
1	A	184	THR
1	A	206	ARG
1	A	234	ARG
1	A	235	GLN
1	A	238	VAL
1	A	247	PHE
1	A	252	ILE
1	A	258	LEU
1	A	277	GLU
1	A	284	SER
1	A	289	HIS
1	A	306	GLU
1	A	317	LYS
1	A	324	ASN
1	A	328	LYS
1	A	331	LEU
1	A	334	ARG
1	A	355	ASP
1	A	357	ASN
1	A	362	VAL
1	A	367	VAL
1	A	377	ARG
1	A	381	LEU
1	A	382	ARG
1	A	388	THR
1	A	401	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	422	LEU
1	A	424	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	$\#|Z| > 2$
2	A1A0L	T	4	2	18,21,22	4.80	9 (50%)	22,30,33	2.30	9 (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
2	A1A0L	T	4	2	-	4/7/23/24	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	4	A1A0L	O4'-C1'	14.67	1.63	1.42
2	T	4	A1A0L	C2'-C1'	-7.64	1.29	1.53
2	T	4	A1A0L	O4'-C4'	-6.30	1.30	1.43
2	T	4	A1A0L	O3'-C3'	-5.26	1.36	1.45

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	4	A1A0L	O4-C4	-5.05	1.21	1.34
2	T	4	A1A0L	O2'-C2'	3.71	1.52	1.43
2	T	4	A1A0L	C6-N1	-3.43	1.32	1.38
2	T	4	A1A0L	C2-N1	-2.99	1.33	1.40
2	T	4	A1A0L	C6-C5	2.27	1.38	1.34

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	4	A1A0L	C5-C6-N1	-4.46	118.47	123.31
2	T	4	A1A0L	O4'-C1'-N1	4.04	114.04	108.90
2	T	4	A1A0L	C1'-N1-C6	-4.00	114.56	121.15
2	T	4	A1A0L	C2'-C1'-N1	-3.81	102.65	113.25
2	T	4	A1A0L	C3'-C2'-C1'	3.20	106.96	99.91
2	T	4	A1A0L	O2'-C2'-C3'	2.98	119.64	111.08
2	T	4	A1A0L	C2-N3-C4	2.51	120.83	116.63
2	T	4	A1A0L	O3'-C3'-C4'	-2.37	105.37	113.24
2	T	4	A1A0L	C3-C5-C6	-2.04	120.08	122.85

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	T	4	A1A0L	C2'-C1'-N1-C6
2	T	4	A1A0L	C2'-C1'-N1-C2
2	T	4	A1A0L	O4'-C1'-N1-C6
2	T	4	A1A0L	O4'-C1'-N1-C2

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	T	4	A1A0L	3	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 2 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
4	DZ4	A	501	5	28,32,32	3.32	14 (50%)	33,50,50	2.33	6 (18%)

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	DZ4	A	501	5	-	8/15/34/34	0/3/3/3

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	501	DZ4	C2'-C3'	-8.00	1.32	1.52
4	A	501	DZ4	O4'-C1'	6.61	1.57	1.42
4	A	501	DZ4	O4'-C4'	-6.37	1.30	1.45
4	A	501	DZ4	PB-N3A	5.98	1.79	1.63
4	A	501	DZ4	C1'-N9	-4.98	1.35	1.49
4	A	501	DZ4	C6-N6	3.47	1.46	1.34
4	A	501	DZ4	PB-O1B	3.41	1.51	1.46
4	A	501	DZ4	PA-O5'	3.40	1.70	1.57
4	A	501	DZ4	C2-N3	3.30	1.37	1.32
4	A	501	DZ4	C3'-C4'	3.19	1.61	1.53
4	A	501	DZ4	PB-O2B	-2.98	1.48	1.56
4	A	501	DZ4	PA-N3A	2.74	1.70	1.63
4	A	501	DZ4	PB-O3B	2.37	1.62	1.59
4	A	501	DZ4	PG-O3G	-2.02	1.47	1.54

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	501	DZ4	C5-C6-N6	8.15	132.73	120.31
4	A	501	DZ4	N3-C2-N1	-6.89	119.32	128.67
4	A	501	DZ4	N6-C6-N1	-5.20	107.23	118.33
4	A	501	DZ4	O2A-PA-O5'	2.41	113.10	106.69
4	A	501	DZ4	O3G-PG-O3B	2.12	111.76	104.64
4	A	501	DZ4	C2'-C3'-C4'	2.03	106.91	102.80

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	501	DZ4	PB-N3A-PA-O1A
4	A	501	DZ4	PB-N3A-PA-O5'
4	A	501	DZ4	C5'-O5'-PA-O2A
4	A	501	DZ4	PA-N3A-PB-O1B
4	A	501	DZ4	PA-N3A-PB-O3B
4	A	501	DZ4	C3'-C4'-C5'-O5'
4	A	501	DZ4	O4'-C4'-C5'-O5'
4	A	501	DZ4	C5'-O5'-PA-N3A

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	501	DZ4	3	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 DZ4 A 501

Bond lengths

Bond angles

Torsions

Rings
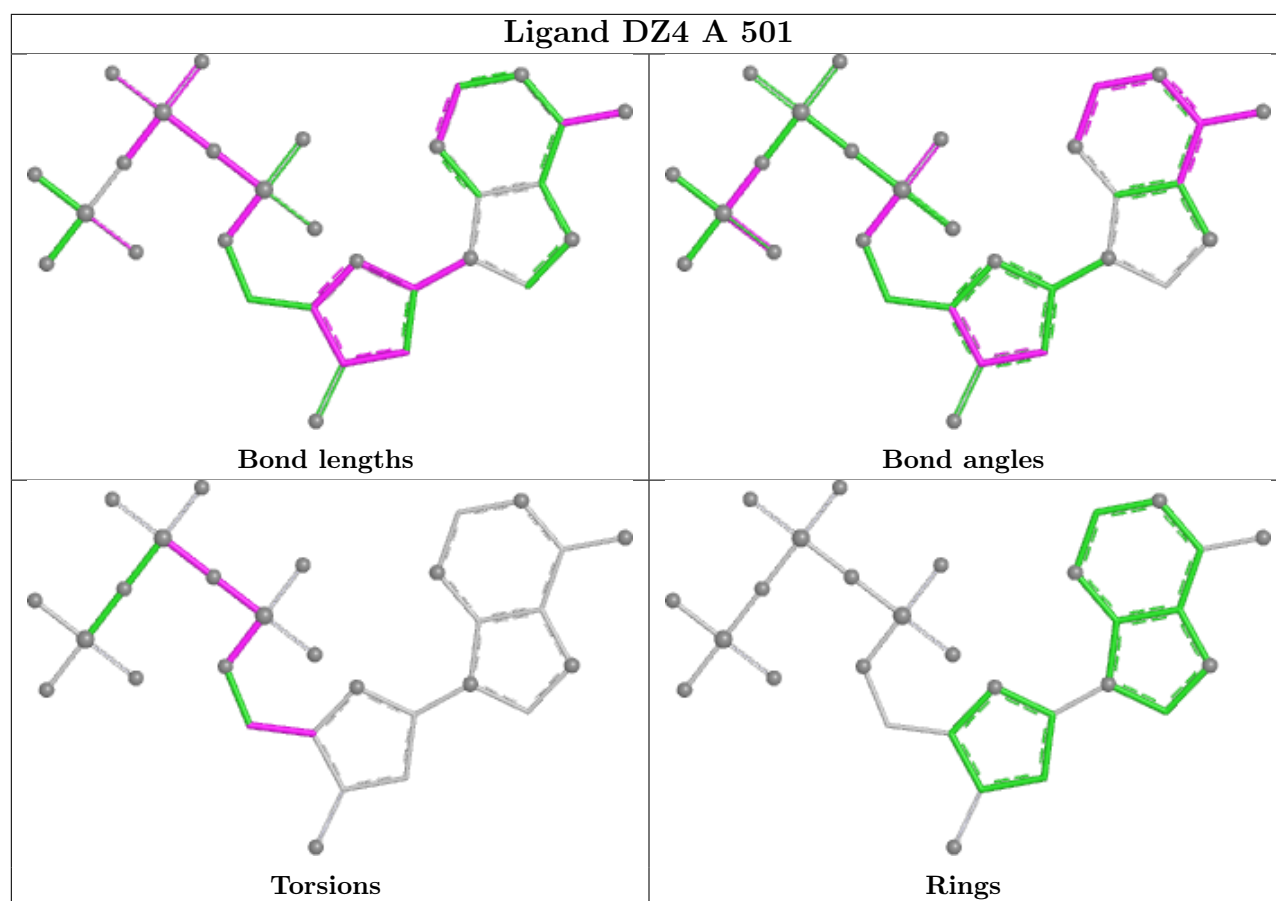
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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(Å²)	Q<0.9
1	A	427/435 (98%)	0.70	24 (5%)	31	20	11, 25, 46, 58	11 (2%)
2	T	11/12 (91%)	0.96	1 (9%)	16	11	17, 28, 76, 90	0
3	P	8/8 (100%)	-0.10	0	100	100	15, 22, 24, 24	0
All	All	446/455 (98%)	0.69	25 (5%)	31	20	11, 25, 46, 90	11 (2%)

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	1	DC	5.5
1	A	405	ASN	2.9
1	A	288	SER	2.9
1	A	161	VAL	2.8
1	A	62	SER	2.8
1	A	77	LEU	2.7
1	A	48	ILE	2.7
1	A	355	ASP	2.6
1	A	47	ILE	2.6
1	A	312	PRO	2.6
1	A	411	ILE	2.6
1	A	72	CYS	2.4
1	A	183	LEU	2.4
1	A	160	THR	2.4
1	A	181	ASP	2.3
1	A	155	THR	2.3
1	A	432	SER	2.3
1	A	379	SER	2.2
1	A	75	LEU	2.2
1	A	153	GLY	2.2
1	A	276	GLY	2.2
1	A	3	THR	2.1
1	A	286	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	31	ASN	2.1
1	A	42	TRP	2.0

## 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
2	A1A0L	T	4	20/21	0.69	0.18	61,74,100,102	0

6.3   Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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	DZ4	A	501	30/30	0.89	0.11	27,41,48,51	0
5	MG	A	503	1/1	0.97	0.10	15,15,15,15	0
5	MG	A	502	1/1	0.99	0.06	9,9,9,9	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.
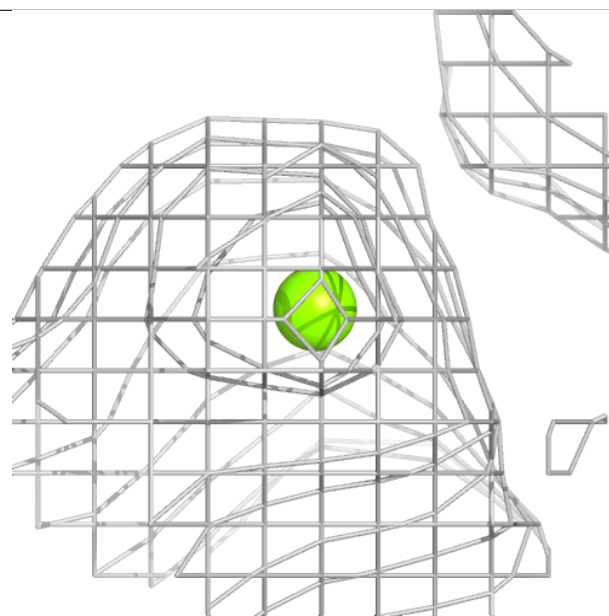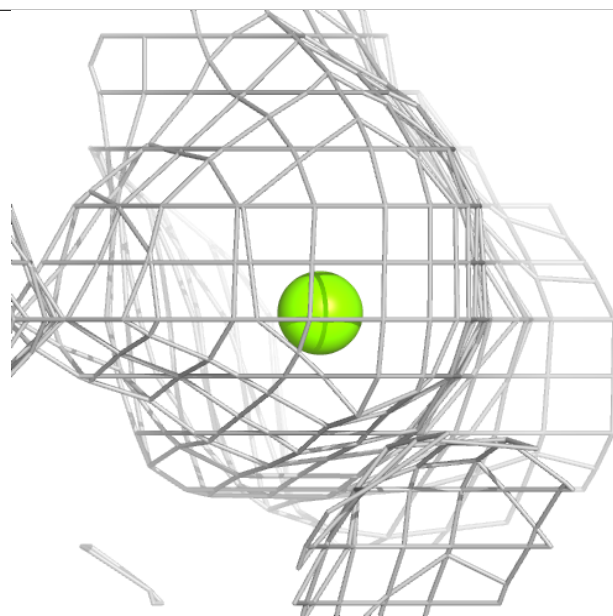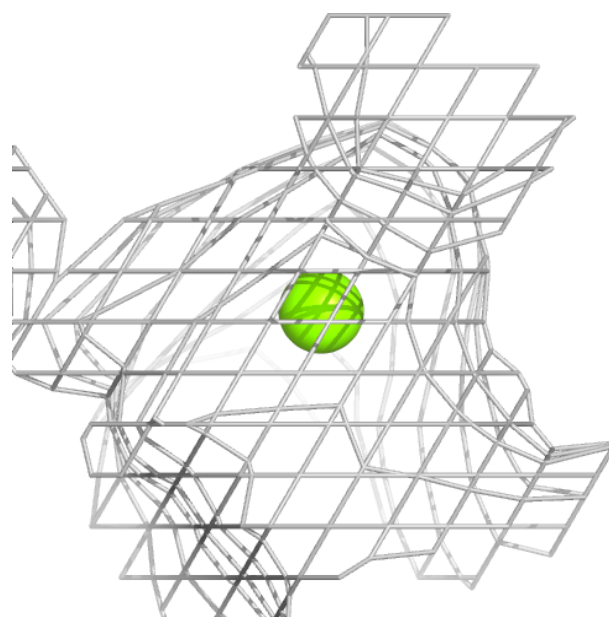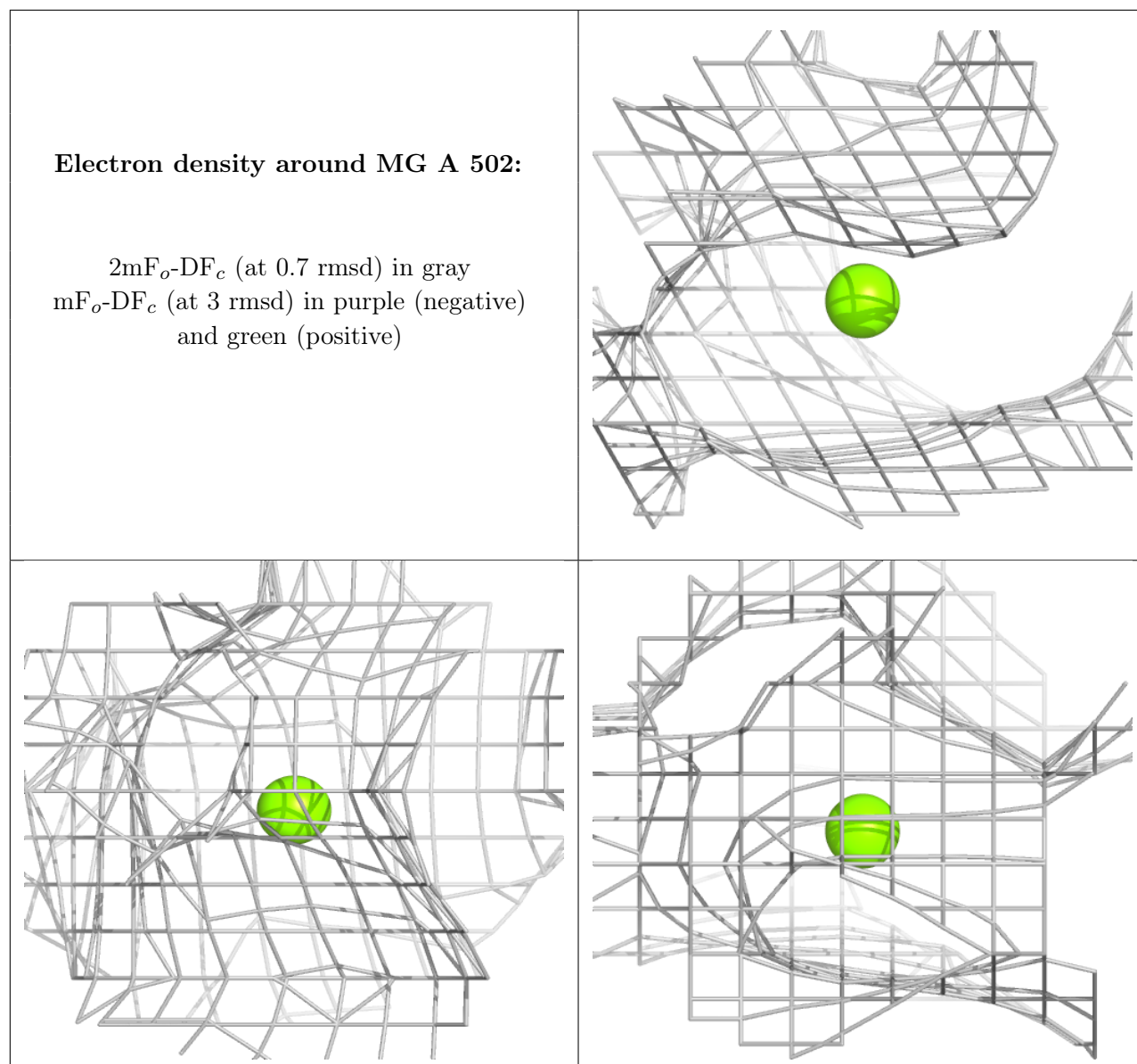
Electron density around DZ4 A 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
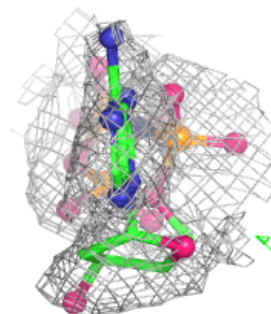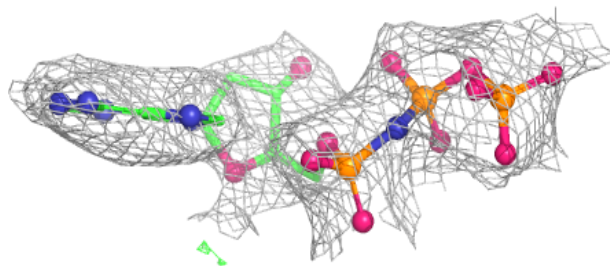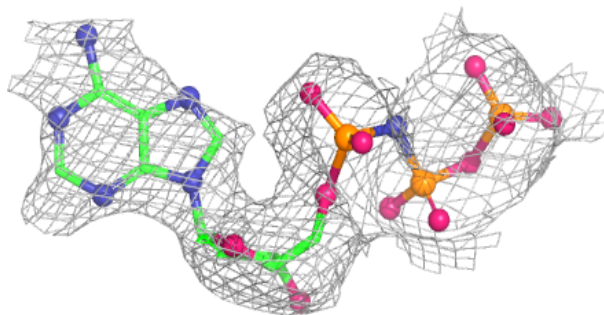
Electron density around MG A 503:

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

Electron density around MG A 502:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.