wwPDB X-ray Structure Validation Summary Report (i)

Aug 14, 2023 – 03:33 PM EDT

PDB ID	:	1S0V
Title	:	Structural basis for substrate selection by T7 RNA polymerase
Authors	:	Temiakov, D.; Patlan, V.; Anikin, M.; McAllister, W.T.; Yokoyama, S.; Vassylyev, D.G.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2004-01-05
Resolution	:	3.20 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (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:

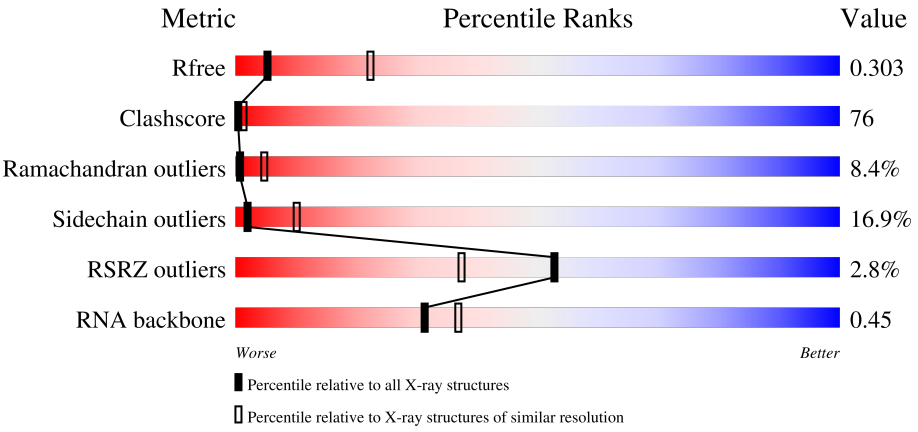
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.20 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.



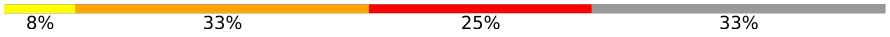
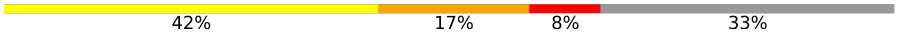
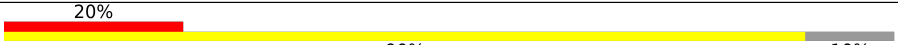
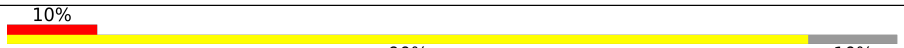
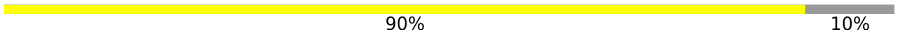
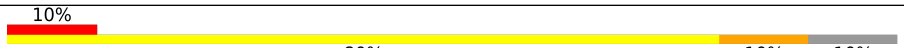
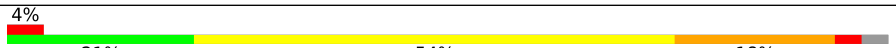
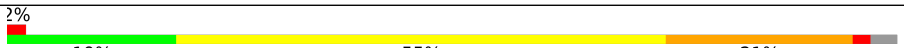
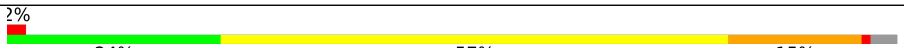
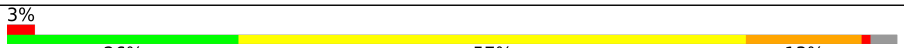
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)
RNA backbone	3102	1010 (3.50-2.90)

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	E	18	11% 39% 11% 33% 6%
1	H	18	6% 17% 28% 22% 28% 6%
1	K	18	56% 11% 28% 6%
1	N	18	6% 56% 22% 11% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	12	
2	I	12	
2	L	12	
2	O	12	
3	G	10	
3	J	10	
3	M	10	
3	P	10	
4	A	883	
4	B	883	
4	C	883	
4	D	883	

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
6	APC	A	2000	-	-	X	-
6	APC	B	2001	-	-	X	-
6	APC	C	2002	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 30899 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(*G*GP*GP*AP*AP*TP*CP*GP*AP*TP*AP*TP*CP*GP*CP*CP*GP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	17	Total	C	N	O	P	0	0	0
			346	165	66	99	16			
1	H	17	Total	C	N	O	P	0	0	0
			346	165	66	99	16			
1	K	17	Total	C	N	O	P	0	0	0
			346	165	66	99	16			
1	N	17	Total	C	N	O	P	0	0	0
			346	165	66	99	16			

- Molecule 2 is a RNA chain called 5'-R(*AP*AP*CP*U*GP*CP*GP*GP*CP*GP*AP*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			
2	I	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			
2	L	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			
2	O	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			
3	J	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			
3	M	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			

- Molecule 4 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			
4	B	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			
4	C	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			
4	D	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			

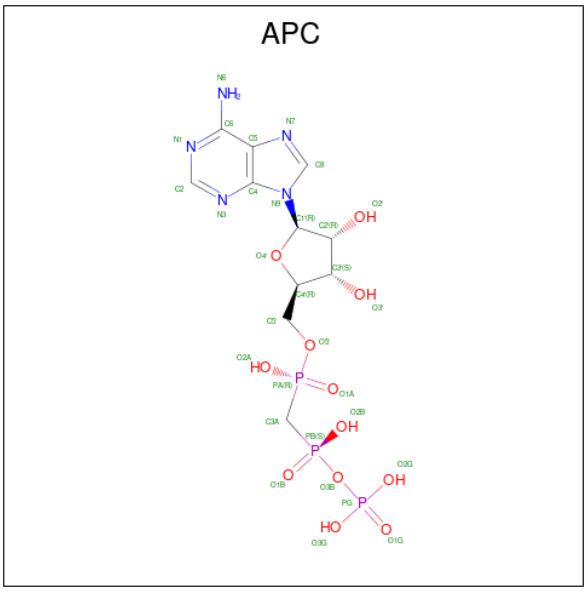
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	497	LEU	-	insertion	UNP P00573
B	497	LEU	-	insertion	UNP P00573
C	497	LEU	-	insertion	UNP P00573
D	497	LEU	-	insertion	UNP P00573

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	F	1	Total	Mg	0	0
			1	1		
5	A	1	Total	Mg	0	0
			1	1		
5	B	2	Total	Mg	0	0
			2	2		
5	C	2	Total	Mg	0	0
			2	2		
5	D	2	Total	Mg	0	0
			2	2		

- Molecule 6 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: $C_{11}H_{18}N_5O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			31	11	5	12	3		
6	B	1	Total	C	N	O	P	0	0
			31	11	5	12	3		
6	C	1	Total	C	N	O	P	0	0
			31	11	5	12	3		
6	D	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	E	39	Total	O	0	0
			39	39		
7	F	9	Total	O	0	0
			9	9		
7	G	9	Total	O	0	0
			9	9		
7	H	19	Total	O	0	0
			19	19		
7	I	14	Total	O	0	0
			14	14		
7	J	13	Total	O	0	0
			13	13		
7	K	20	Total	O	0	0
			20	20		
7	L	8	Total	O	0	0
			8	8		

Continued on next page...

Continued from previous page...

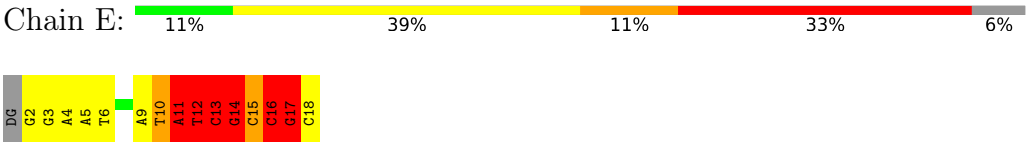
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	M	10	Total O 10 10	0	0
7	N	14	Total O 14 14	0	0
7	O	15	Total O 15 15	0	0
7	P	6	Total O 6 6	0	0
7	A	237	Total O 237 237	0	0
7	B	212	Total O 212 212	0	0
7	C	201	Total O 201 201	0	0
7	D	173	Total O 173 173	0	0

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.

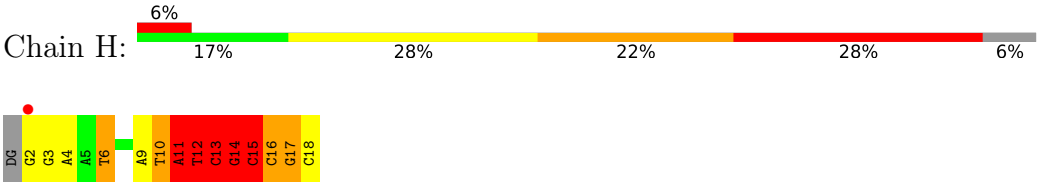
• Molecule 1: 5'-D(*G*GP*GP*AP*AP*TP*CP*GP*AP*TP*AP*TP*CP*GP*CP*CP*GP*C)-3'

Chain E:



• Molecule 1: 5'-D(*G*GP*GP*AP*AP*TP*CP*GP*AP*TP*AP*TP*CP*GP*CP*CP*GP*C)-3'

Chain H:



• Molecule 1: 5'-D(*G*GP*GP*AP*AP*TP*CP*GP*AP*TP*AP*TP*CP*GP*CP*CP*GP*C)-3'

Chain K:

• Molecule 1: 5'-D(*G*GP*GP*AP*AP*TP*CP*GP*AP*TP*AP*TP*CP*GP*CP*CP*GP*C)-3'

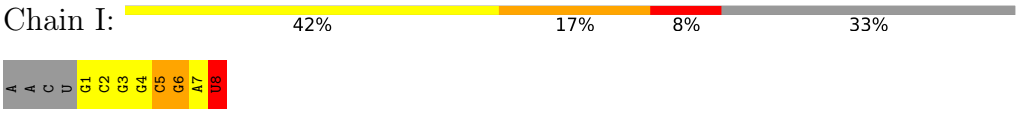
Chain N:

• Molecule 2: 5'-R(*AP*AP*CP*U*GP*CP*GP*GP*CP*GP*AP*U)-3'

Chain F:

• Molecule 2: 5'-R(*AP*AP*CP*U*GP*CP*GP*GP*CP*GP*AP*U)-3'

Chain I:



• Molecule 2: 5'-R(*AP*AP*CP*U*GP*CP*GP*GP*CP*GP*AP*U)-3'

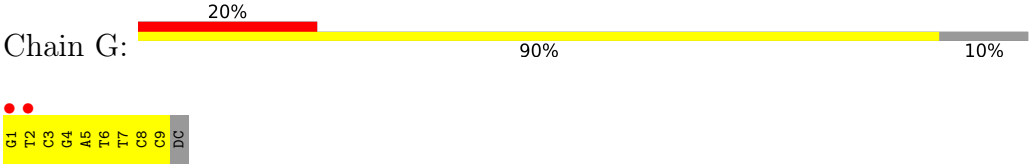
Chain L:

• Molecule 2: 5'-R(*AP*AP*CP*U*GP*CP*GP*GP*CP*GP*AP*U)-3'

Chain O:

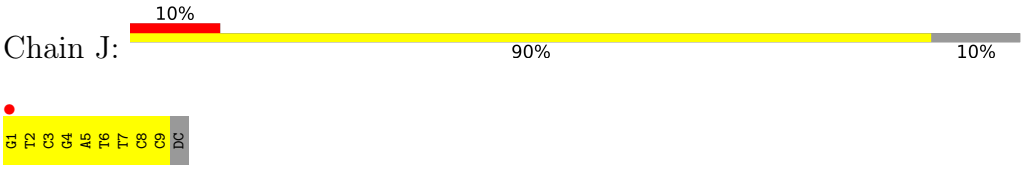
• Molecule 3: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*C)-3'

Chain G:



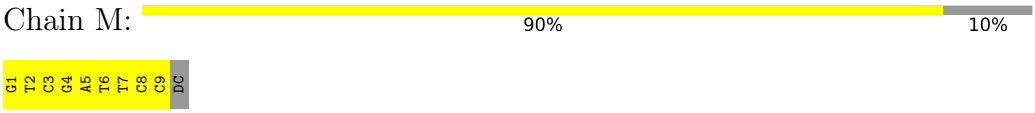
• Molecule 3: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*C)-3'

Chain J:



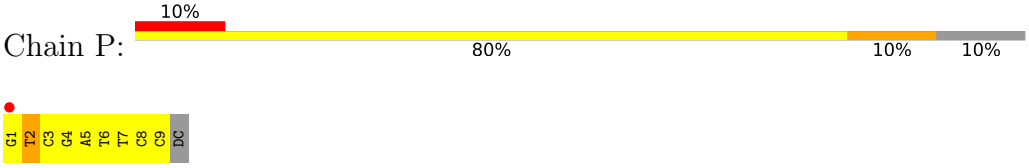
• Molecule 3: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*C)-3'

Chain M:



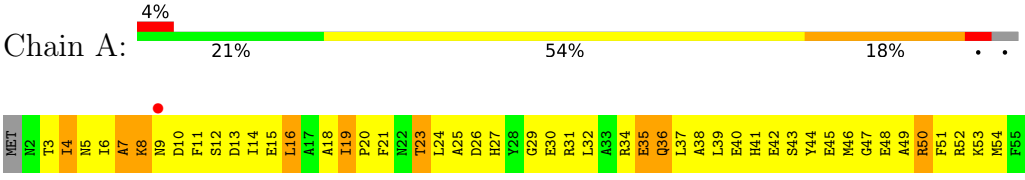
• Molecule 3: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*C)-3'

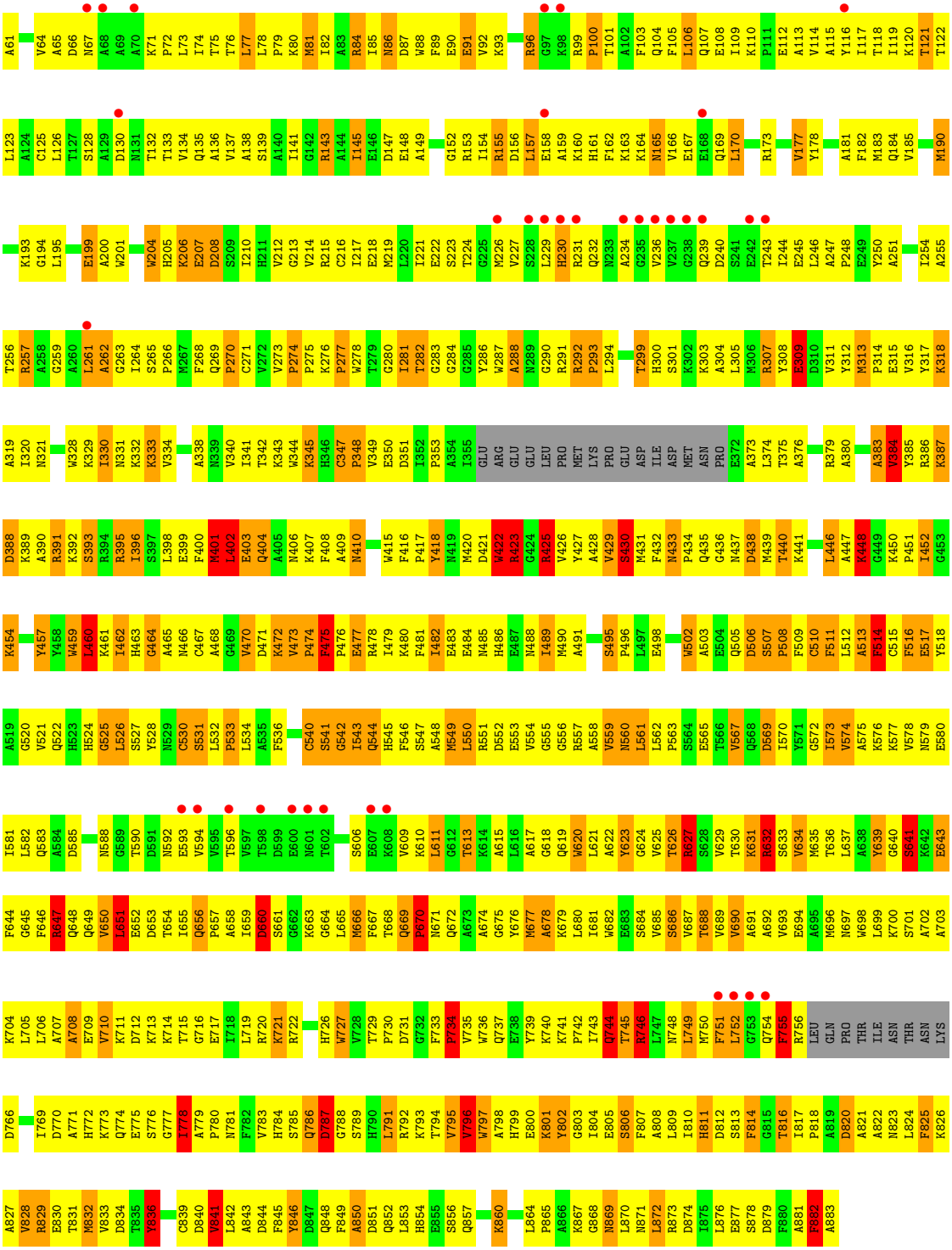
Chain P:



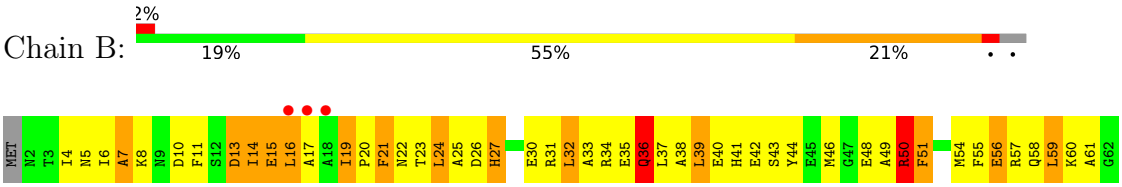
• Molecule 4: DNA-directed RNA polymerase

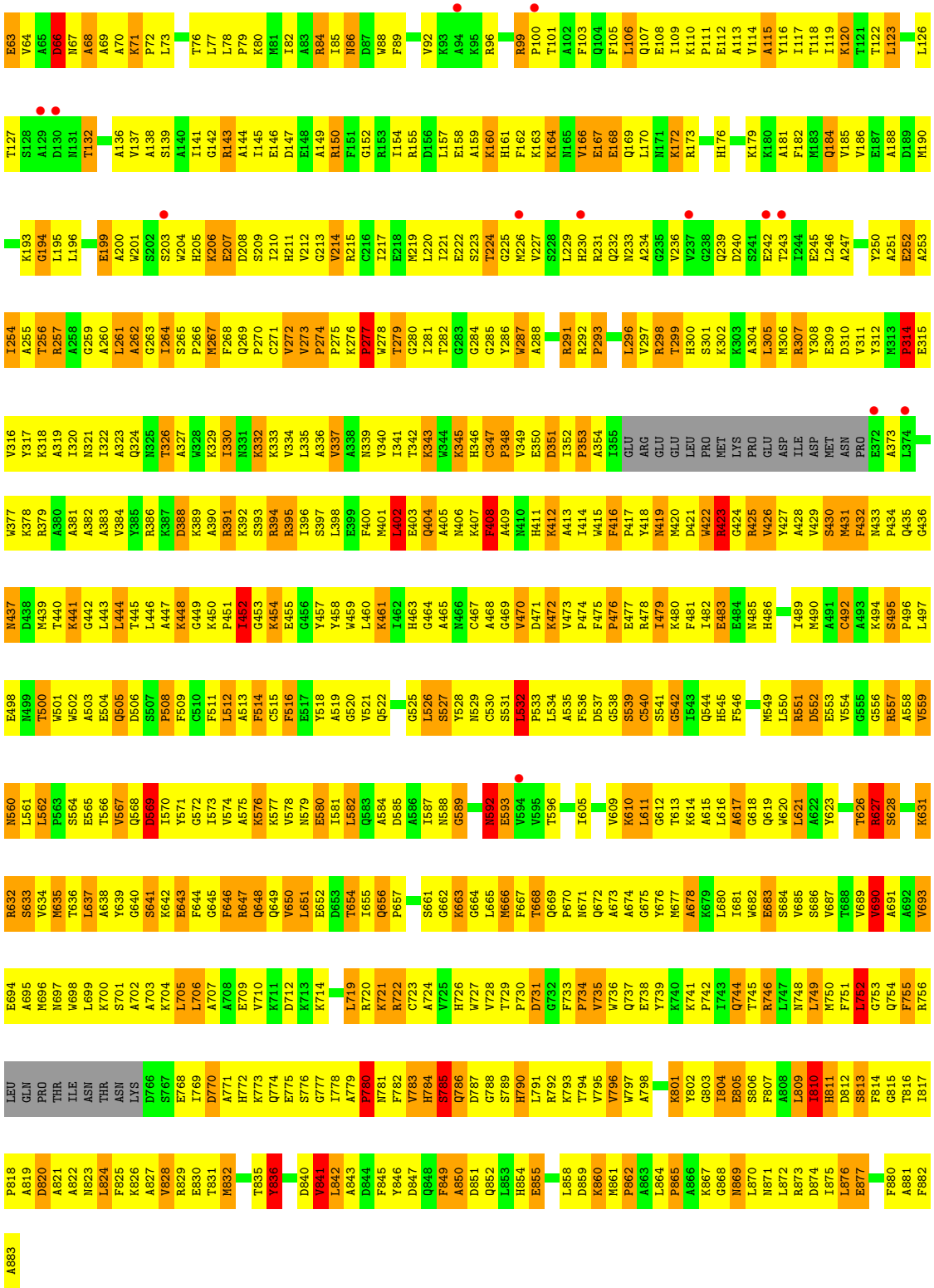
Chain A:



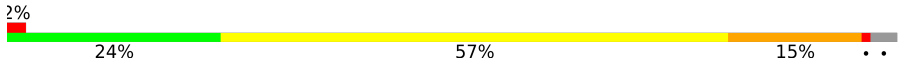


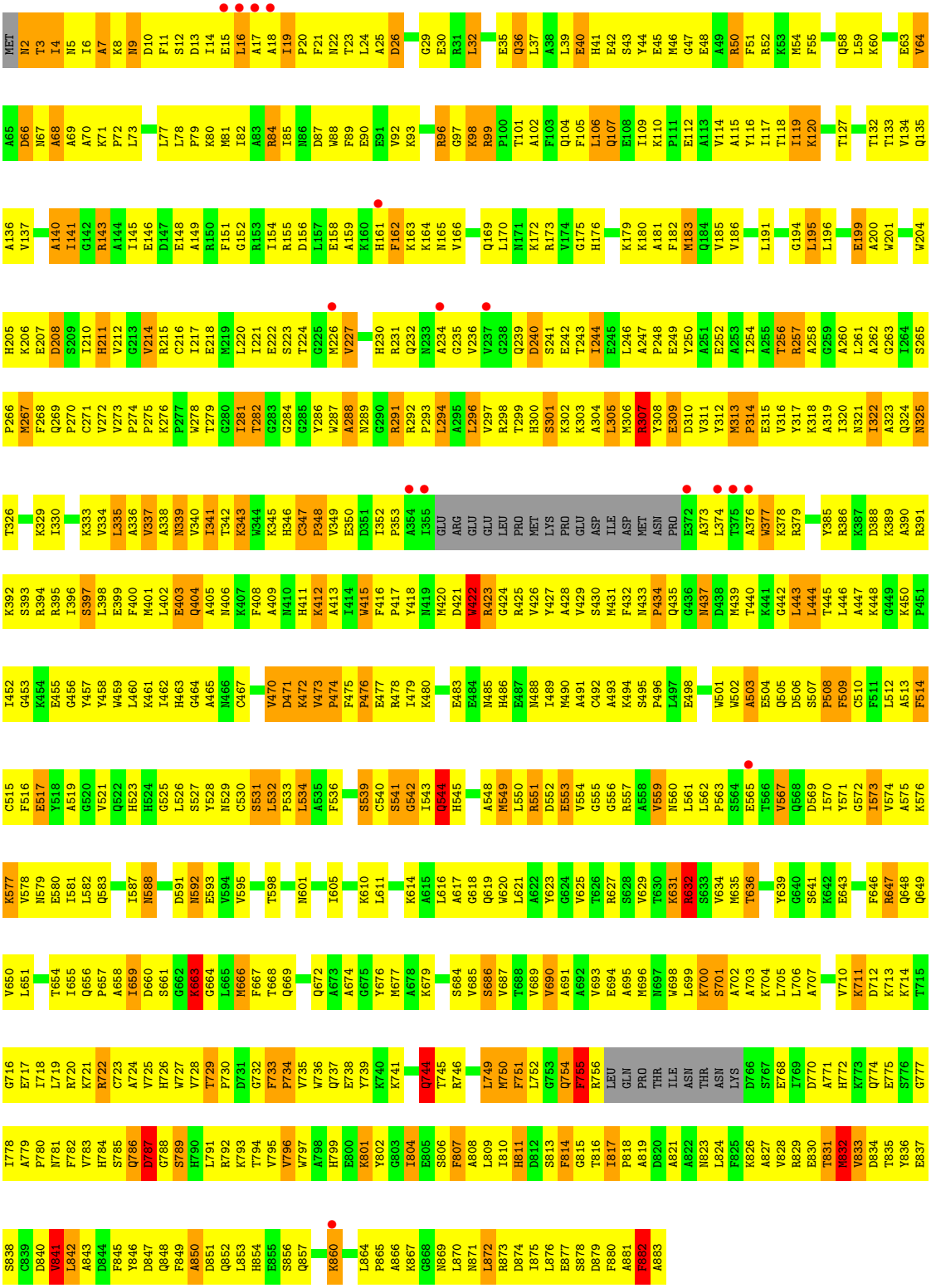
• Molecule 4: DNA-directed RNA polymerase





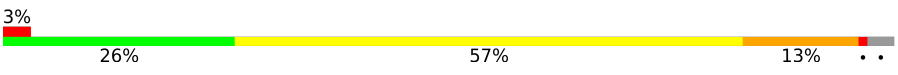
• Molecule 4: DNA-directed RNA polymerase

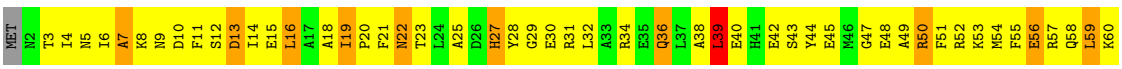


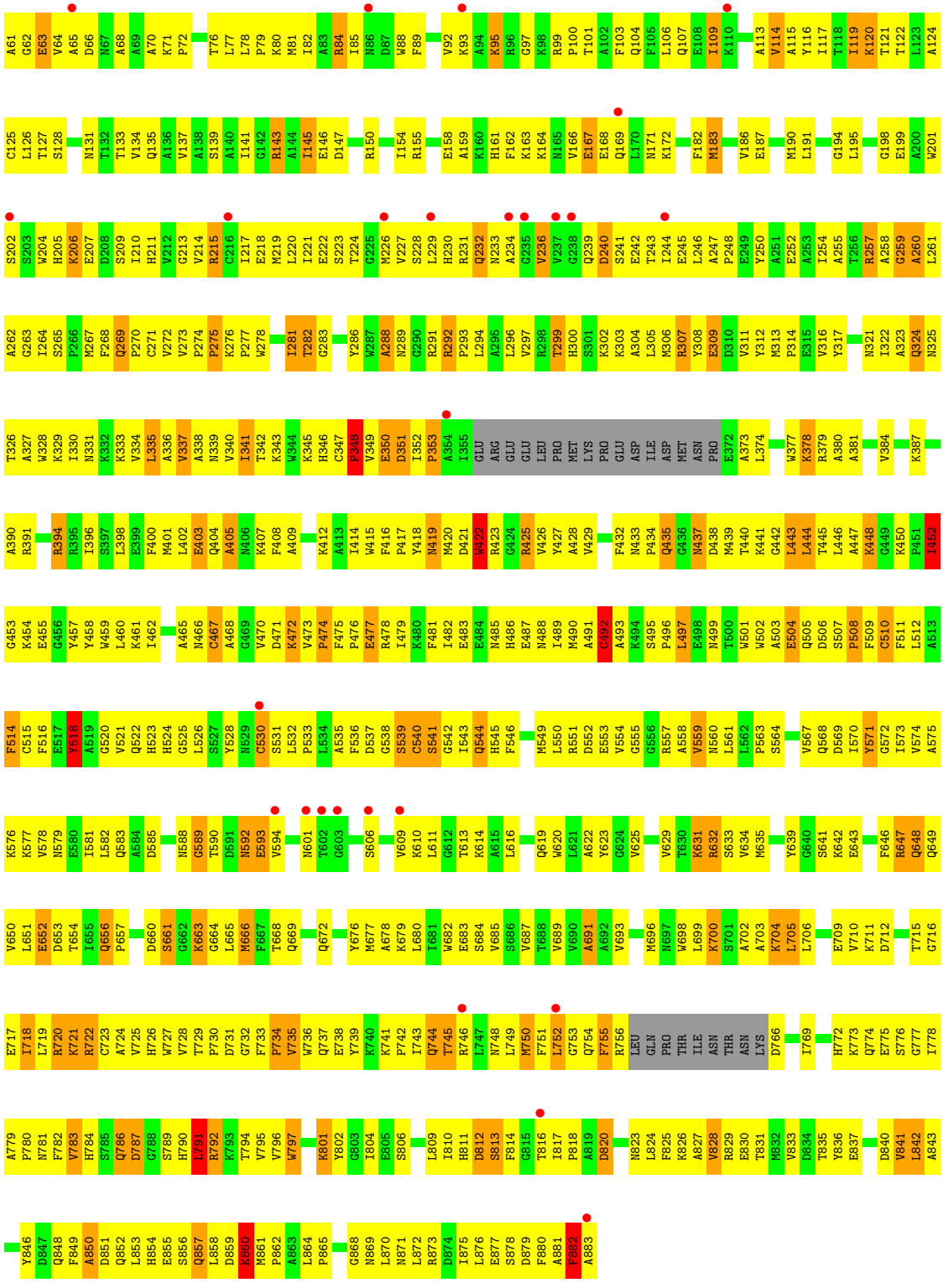


• Molecule 4: DNA-directed RNA polymerase

Chain D:






A61 G62 E63 V64 A65 D66 N67 A68 A69 A70 K71 P72 T76 L77 L78 P79 K80 M81 I82 A83 R84 I85 N86 D87 W88 F89 V92 K93 A94 K95 R96 G97 K98 R99 P100 T101 A102 F103 Q104 F105 L106 Q107 E108 I109 K110 A113 V114 A115 Y116 I117 T118 I119 K120 T121 T122 L123 A124
C125 L126 T127 S128 N131 T132 T133 V134 Q135 A136 V137 A138 S139 A140 I141 G142 R143 A144 I145 E146 D147 R150 I154 R155 E158 A159 K160 H161 F162 K163 K164 N165 V166 E167 E168 Q169 L170 N171 K172 F182 M183 V186 E187 M190 L191 G194 L195 G198 E199 A200 W201
S202 S203 W204 H205 K206 E207 D208 S209 I210 H211 V212 G213 V214 R215 C216 I217 E218 M219 L220 I221 E222 S223 T224 G225 M226 V227 S228 L229 H230 R231 Q232 N233 A234 G235 V236 V237 G238 Q239 D240 S241 E242 T243 I244 E245 L246 A247 P248 E249 Y250 A251 E252 A253 I254 A255 T256 R257 A258 G259 A260 L261
A262 G263 I264 S265 P266 M267 F268 Q269 P270 C271 V272 V273 P274 P275 K276 P277 W278 I281 T282 G283 Y286 W287 A288 N289 G290 R291 R292 P293 L294 A295 L296 V297 R298 T299 H300 S301 K302 K303 A304 L305 M306 R307 Y308 E309 D310 V311 Y312 M313 P314 E315 V316 Y317 N321 I322 A323 Q324 N325
T326 A327 W328 K329 I330 N331 K332 K333 V334 L335 A336 V337 A338 N339 V340 I341 T342 K343 W344 K345 H346 C347 P348 V349 E350 D351 I352 P353 A354 I355 GLU ARG GLU GLU LEU PRO MET LYS PRO GLU ASP ILE ASP MET ASN PRO E372 A373 L374 W377 K378 R379 A380 A381 V384 K387
A390 R391 R394 R395 I396 S397 L398 E399 F400 M401 L402 E403 Q404 A405 N406 K407 F408 A409 K412 A413 I414 W415 F416 P417 Y418 N419 M420 D421 W422 R423 G424 R425 V426 Y427 A428 V429 F432 N433 P434 Q435 G436 N437 D438 M439 T440 K441 G442 L443 L444 T445 L446 A447 K448 G449 K450 P451 T452
G453 K454 E455 G456 Y457 Y458 W459 L460 K461 I462 A465 N466 C467 A468 G469 V470 D471 K472 V473 P474 F475 P476 E477 R478 I479 K480 F481 I482 E483 E484 N485 H486 E487 N488 I489 M490 A491 C492 A493 K494 S495 P496 L497 E498 N499 T500 W501 W502 A503 E504 Q505 D506 S507 P508 F509 C510 F511 L512 A513
F514 C515 F516 E517 Y518 A519 G520 V521 Q522 H523 H524 G525 L526 S527 Y528 N529 C530 S531 L532 P533 L534 A535 F536 D537 G538 S539 C540 S541 G542 I543 Q544 H545 F546 M549 L550 R551 D552 E553 V554 G555 G556 R557 A558 V559 N560 L561 L562 P563 S564 V567 Q568 D569 I570 Y571 G572 I573 Y574 A575
K576 K577 V578 N579 E580 I581 L582 Q583 A584 D585 N588 G589 T590 D591 N592 E593 V594 N601 T602 G603 S606 V609 K610 L611 G612 T613 K614 A615 L616 Q619 W620 L621 A622 Y623 G624 V625 V629 T630 K631 R632 S633 V634 M635 Y639 G640 S641 K642 E643 F646 R647 Q648 Q649
V650 L651 E652 D653 T654 I655 Q656 P657 D660 S661 G662 K663 G664 L665 M666 F667 T668 Q669 Q672 Y676 M677 A678 K679 L680 I681 W682 E683 S684 V685 S686 V687 T688 V689 V690 A691 A692 V693 M696 N697 W698 L699 K700 S701 A702 A703 K704 L705 L706 E709 V710 K711 D712 T715 G716
E717 I718 L719 R720 K721 R722 C723 A724 V725 H726 W727 V728 T729 P730 D731 G732 F733 P734 V735 W736 Q737 E738 Y739 K740 K741 P742 I743 Q744 R746 L747 N748 L749 M750 F751 L752 G753 Q754 F755 R756 LEU GLN PRO THR ILE ASN THR ASN LYS D766 I769 H772 K773 Q774 E775 S776 G777 I778
A779 P780 N781 F782 V783 H784 S785 Q786 D787 G788 S789 H790 L791 R792 K793 T794 V795 V796 W797 K801 Y802 G803 I804 E805 S806 L809 I810 H811 D812 S813 F814 G815 T816 I817 P818 A819 D820 N823 L824 F825 K826 A827 V828 R829 E830 T831 M832 V833 D834 T835 Y836 E837 D840 V841 L842 A843
Y846 D847 Q848 F849 A850 D851 Q852 L853 H854 E855 S856 Q857 L858 D859 K860 M861 P862 A863 L864 P865 G868 N869 L870 N871 L872 R873 D874 I875 L876 E877 S878 D879 F880 A881 F882 A883

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	81.13Å 87.70Å 206.53Å 91.93° 91.02° 110.66°	Depositor
Resolution (Å)	40.00 – 3.20 39.88 – 3.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.20) 83.4 (39.88-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.255 , 0.307 0.248 , 0.303	Depositor DCC
R_{free} test set	4474 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	59.0	Xtriage
Anisotropy	0.201	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 112.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.53$, $< L^2 > = 0.37$	Xtriage
Estimated twinning fraction	0.014 for -h,-k,l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	30899	wwPDB-VP
Average B, all atoms (Å^2)	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% of the height of the origin peak. No significant pseudotranslation is detected.*

[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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: APC, 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	E	1.74	9/388 (2.3%)	1.60	9/597 (1.5%)
1	H	1.62	7/388 (1.8%)	1.47	6/597 (1.0%)
1	K	1.59	7/388 (1.8%)	1.31	1/597 (0.2%)
1	N	1.26	3/388 (0.8%)	1.27	4/597 (0.7%)
2	F	2.39	9/191 (4.7%)	2.00	11/297 (3.7%)
2	I	2.25	7/191 (3.7%)	1.68	2/297 (0.7%)
2	L	1.65	1/191 (0.5%)	1.36	0/297
2	O	1.47	0/191	1.38	0/297
3	G	0.82	0/199	0.92	0/305
3	J	0.78	0/199	0.93	0/305
3	M	0.88	0/199	0.86	0/305
3	P	0.94	0/199	1.07	0/305
4	A	1.20	22/6897 (0.3%)	1.14	24/9329 (0.3%)
4	B	1.21	24/6897 (0.3%)	1.14	24/9329 (0.3%)
4	C	0.97	5/6897 (0.1%)	0.97	7/9329 (0.1%)
4	D	0.92	3/6897 (0.0%)	0.91	4/9329 (0.0%)
All	All	1.14	97/30700 (0.3%)	1.09	92/42112 (0.2%)

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	E	0	6
1	H	0	9
1	K	0	6
1	N	0	4
2	F	0	4
2	I	0	2
2	L	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
2	O	0	3
3	P	0	1
4	A	0	1
4	B	0	1
4	D	0	1
All	All	0	39

The worst 5 of 97 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	502	TRP	CB-CG	-13.18	1.26	1.50
1	K	14	DG	C5-C6	-11.28	1.31	1.42
4	A	467	CYS	CB-SG	-10.63	1.64	1.82
2	I	8	U	N1-C6	-9.75	1.29	1.38
2	I	7	A	C5-C6	-8.93	1.33	1.41

The worst 5 of 92 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	460	LEU	CB-CG-CD1	-10.01	93.98	111.00
1	E	14	DG	O5'-P-OP1	-9.17	97.45	105.70
4	A	460	LEU	CB-CG-CD2	9.03	126.35	111.00
4	D	791	LEU	CA-CB-CG	8.79	135.52	115.30
4	A	425	ARG	NE-CZ-NH1	-8.45	116.07	120.30

There are no chirality outliers.

5 of 39 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	11	DA	Sidechain
1	E	12	DT	Sidechain
1	E	13	DC	Sidechain
1	E	14	DG	Sidechain
1	E	16	DC	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	E	346	0	190	28	0
1	H	346	0	192	66	0
1	K	346	0	190	57	0
1	N	346	0	192	38	0
2	F	171	0	89	10	0
2	I	171	0	89	11	0
2	L	171	0	89	14	0
2	O	171	0	89	5	0
3	G	179	0	104	13	0
3	J	179	0	104	17	0
3	M	179	0	104	13	0
3	P	179	0	104	8	0
4	A	6746	0	6708	1151	0
4	B	6746	0	6708	1190	0
4	C	6746	0	6708	1007	0
4	D	6746	0	6708	899	0
5	A	1	0	0	0	0
5	B	2	0	0	0	0
5	C	2	0	0	0	0
5	D	2	0	0	0	0
5	F	1	0	0	0	0
6	A	31	0	13	16	0
6	B	31	0	13	9	0
6	C	31	0	14	16	0
6	D	31	0	13	7	0
7	A	237	0	0	82	0
7	B	212	0	0	68	0
7	C	201	0	0	70	0
7	D	173	0	0	53	0
7	E	39	0	0	1	0
7	F	9	0	0	1	0
7	G	9	0	0	1	0
7	H	19	0	0	9	0
7	I	14	0	0	4	0
7	J	13	0	0	0	0
7	K	20	0	0	7	0
7	L	8	0	0	0	0
7	M	10	0	0	2	0
7	N	14	0	0	3	0
7	O	15	0	0	2	0
7	P	6	0	0	2	0
All	All	30899	0	28421	4458	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 76.

The worst 5 of 4458 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)
4:A:313:MET:CE	4:A:313:MET:SD	2.05	1.44
4:C:631:LYS:NZ	6:C:2002:APC:H3A2	1.52	1.22
2:F:1:G:H5"	7:F:3008:HOH:O	1.40	1.19
4:A:631:LYS:NZ	6:A:2000:APC:H3A2	1.58	1.18
4:A:546:PHE:CE1	4:A:783:VAL:HG22	1.78	1.16

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
4	A	851/883 (96%)	570 (67%)	198 (23%)	83 (10%)	0 3
4	B	851/883 (96%)	578 (68%)	188 (22%)	85 (10%)	0 3
4	C	851/883 (96%)	608 (71%)	178 (21%)	65 (8%)	1 7
4	D	851/883 (96%)	623 (73%)	174 (20%)	54 (6%)	1 10
All	All	3404/3532 (96%)	2379 (70%)	738 (22%)	287 (8%)	1 5

5 of 287 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	7	ALA
4	A	194	GLY
4	A	199	GLU
4	A	281	ILE
4	A	288	ALA

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
4	A	703/729 (96%)	573 (82%)	130 (18%)	1 8
4	B	703/729 (96%)	573 (82%)	130 (18%)	1 8
4	C	703/729 (96%)	594 (84%)	109 (16%)	2 12
4	D	703/729 (96%)	598 (85%)	105 (15%)	3 14
All	All	2812/2916 (96%)	2338 (83%)	474 (17%)	2 10

5 of 474 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	722	ARG
4	D	656	GLN
4	C	244	ILE
4	D	571	TYR
4	D	882	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 125 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	560	ASN
4	D	437	ASN
4	C	5	ASN
4	D	435	GLN
4	D	748	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	F	7/12 (58%)	0	0
2	I	7/12 (58%)	1 (14%)	0
2	L	7/12 (58%)	0	0
2	O	7/12 (58%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
All	All	28/48 (58%)	1 (3%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	I	8	U

There are no RNA pucker outliers to report.

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	$\#\|Z\| > 2$
6	APC	C	2002	5	27,33,33	2.02	6 (22%)	31,52,52	1.46	5 (16%)
6	APC	D	2003	5	27,33,33	1.94	5 (18%)	31,52,52	1.50	7 (22%)
6	APC	A	2000	5	27,33,33	1.89	5 (18%)	31,52,52	1.67	9 (29%)
6	APC	B	2001	5	27,33,33	1.95	7 (25%)	31,52,52	1.70	7 (22%)

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	APC	C	2002	5	-	6/15/38/38	0/3/3/3
6	APC	D	2003	5	-	9/15/38/38	0/3/3/3
6	APC	A	2000	5	-	6/15/38/38	0/3/3/3
6	APC	B	2001	5	-	7/15/38/38	0/3/3/3

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	2003	APC	PB-O3B	7.27	1.66	1.58
6	A	2000	APC	PB-O3B	6.67	1.65	1.58
6	C	2002	APC	PB-O3B	5.61	1.64	1.58
6	C	2002	APC	PA-O5'	5.59	1.65	1.57
6	B	2001	APC	PA-O5'	4.70	1.64	1.57

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	2001	APC	O1A-PA-C3A	-4.48	97.24	109.07
6	A	2000	APC	O1A-PA-C3A	3.52	118.37	109.07
6	B	2001	APC	PB-O3B-PG	-3.48	120.37	132.62
6	B	2001	APC	C2'-C3'-C4'	3.39	109.23	102.64
6	D	2003	APC	O2B-PB-O1B	3.23	120.85	110.07

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	2000	APC	PB-C3A-PA-O1A
6	A	2000	APC	PB-C3A-PA-O2A
6	A	2000	APC	PB-C3A-PA-O5'
6	A	2000	APC	O4'-C4'-C5'-O5'
6	B	2001	APC	PA-C3A-PB-O1B

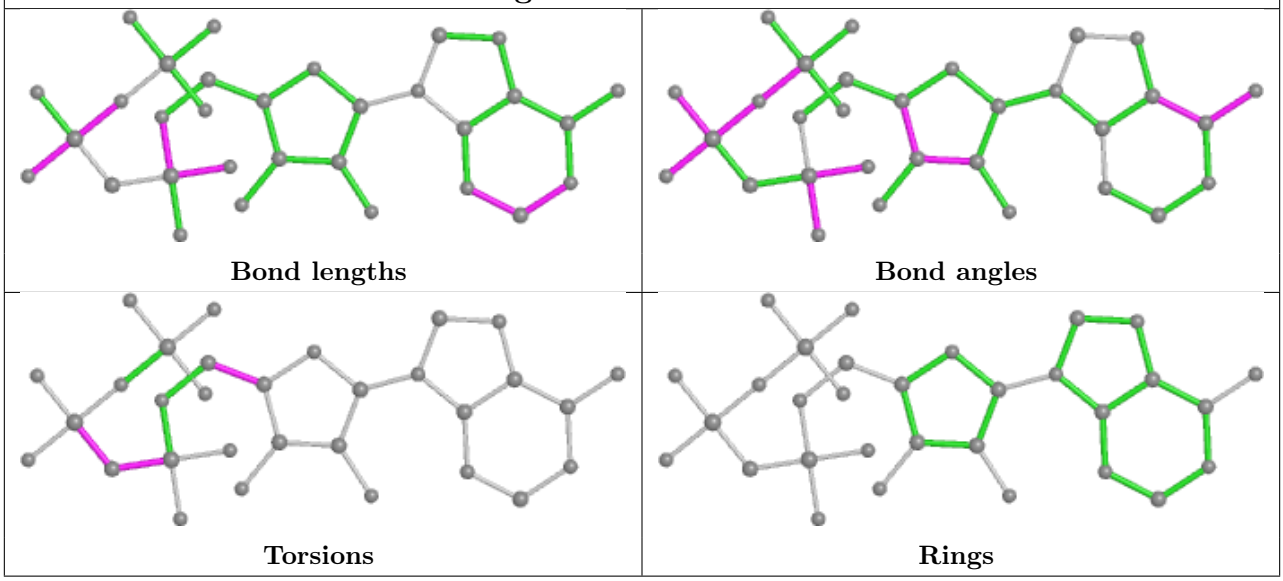
There are no ring outliers.

4 monomers are involved in 48 short contacts:

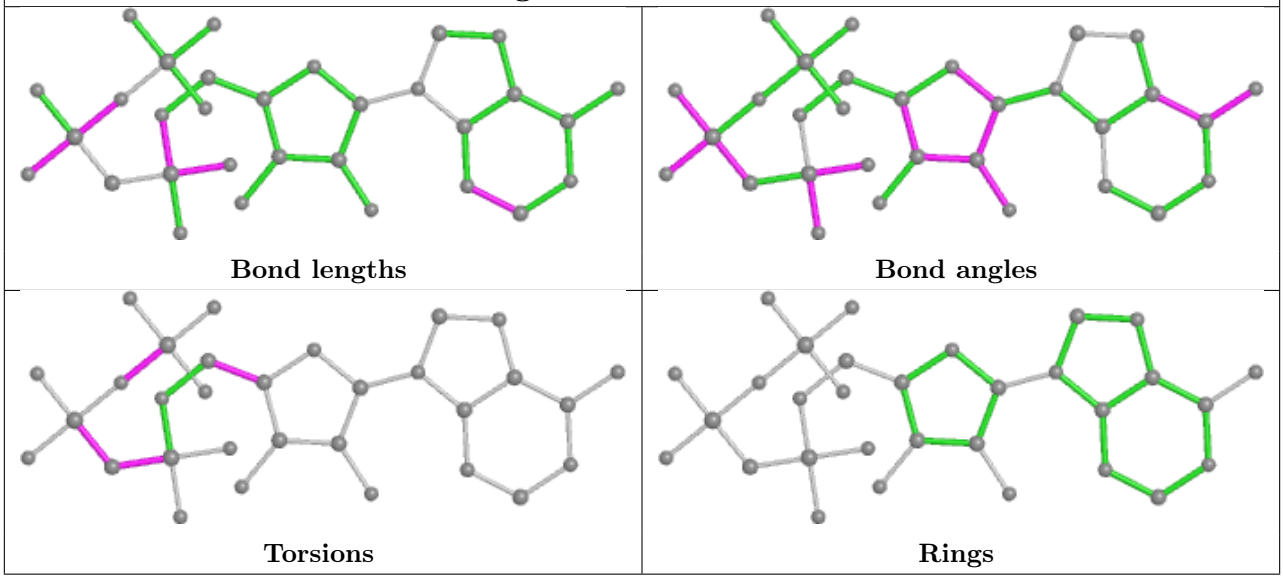
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	2002	APC	16	0
6	D	2003	APC	7	0
6	A	2000	APC	16	0
6	B	2001	APC	9	0

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.

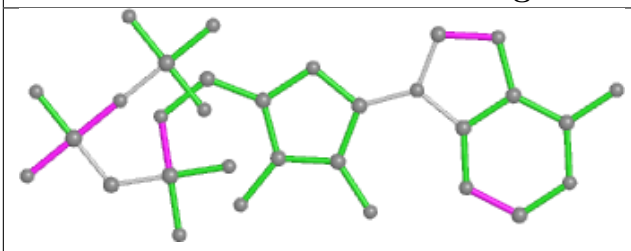
Ligand APC C 2002



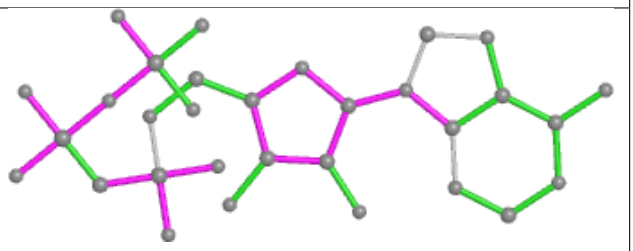
Ligand APC D 2003



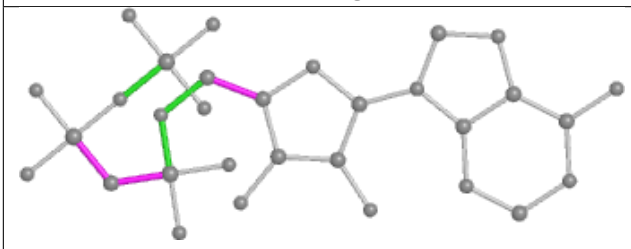
Ligand APC A 2000



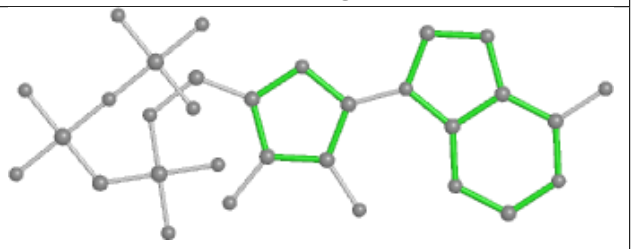
Bond lengths



Bond angles

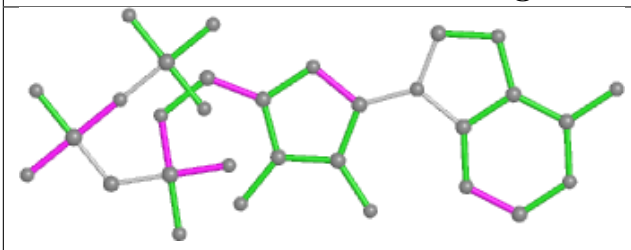


Torsions

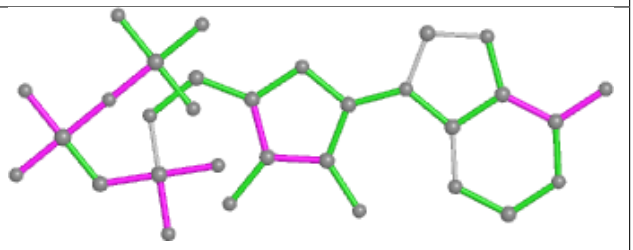


Rings

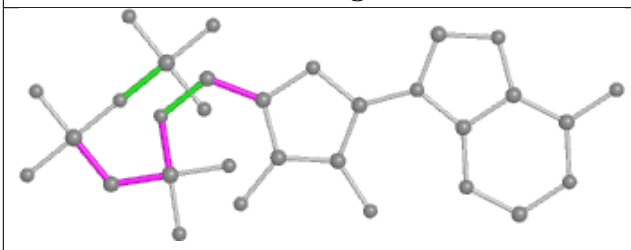
Ligand APC B 2001



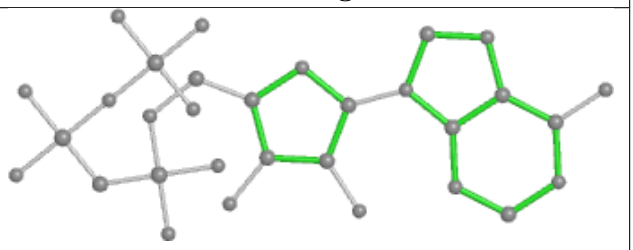
Bond lengths



Bond angles



Torsions



Rings

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, 95^{th} 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	E	17/18 (94%)	-0.25	0 100 100	19, 58, 179, 181	0
1	H	17/18 (94%)	0.17	1 (5%) 22 13	21, 75, 158, 171	0
1	K	17/18 (94%)	0.11	0 100 100	48, 94, 157, 159	0
1	N	17/18 (94%)	0.09	0 100 100	51, 102, 172, 173	0
2	F	8/12 (66%)	-0.50	0 100 100	17, 23, 82, 93	0
2	I	8/12 (66%)	-0.02	0 100 100	17, 38, 82, 92	0
2	L	8/12 (66%)	0.19	0 100 100	47, 52, 113, 118	0
2	O	8/12 (66%)	0.14	0 100 100	51, 80, 132, 144	0
3	G	9/10 (90%)	0.91	2 (22%) 0 0	146, 153, 176, 183	0
3	J	9/10 (90%)	-0.03	1 (11%) 5 3	130, 138, 162, 163	0
3	M	9/10 (90%)	-0.22	0 100 100	134, 146, 155, 158	0
3	P	9/10 (90%)	1.30	1 (11%) 5 3	157, 167, 172, 172	0
4	A	857/883 (97%)	-0.28	37 (4%) 35 22	15, 76, 141, 153	0
4	B	857/883 (97%)	-0.41	16 (1%) 66 53	10, 69, 131, 150	0
4	C	857/883 (97%)	-0.30	16 (1%) 66 53	47, 93, 134, 160	0
4	D	857/883 (97%)	-0.05	26 (3%) 50 34	41, 102, 144, 156	0
All	All	3564/3692 (96%)	-0.25	100 (2%) 53 37	10, 89, 142, 183	0

The worst 5 of 100 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	P	1	DG	8.1
4	D	235	GLY	7.0
3	G	1	DG	6.4
4	A	235	GLY	6.4
4	A	601	ASN	5.8

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)

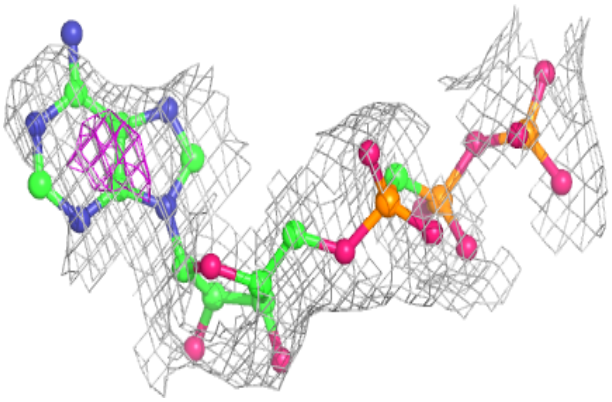
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

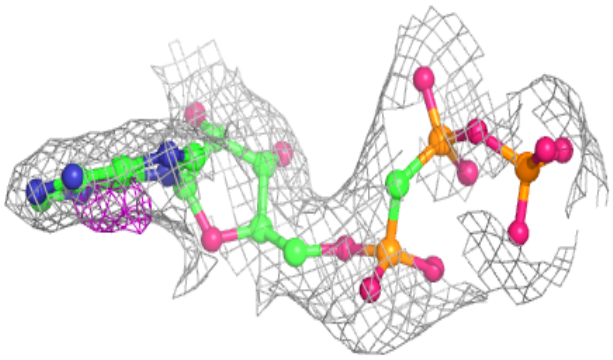
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	MG	D	3004	1/1	0.69	0.30	79,79,79,79	0
6	APC	C	2002	31/31	0.93	0.23	65,70,77,78	0
6	APC	B	2001	31/31	0.94	0.21	28,47,63,68	0
5	MG	A	3001	1/1	0.94	0.13	22,22,22,22	0
6	APC	D	2003	31/31	0.94	0.17	48,57,61,63	0
6	APC	A	2000	31/31	0.95	0.19	34,58,80,82	0
5	MG	F	3005	1/1	0.96	0.15	18,18,18,18	0
5	MG	B	3002	1/1	0.97	0.53	65,65,65,65	0
5	MG	D	3008	1/1	0.97	0.17	16,16,16,16	0
5	MG	B	3006	1/1	0.97	0.17	34,34,34,34	0
5	MG	C	3003	1/1	0.99	0.27	41,41,41,41	0
5	MG	C	3007	1/1	0.99	0.17	46,46,46,46	0

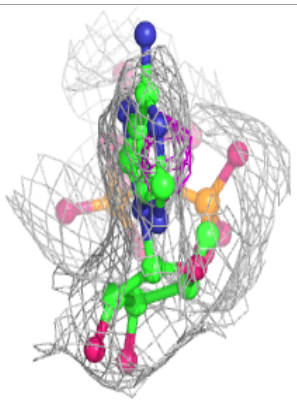
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around APC C 2002:

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

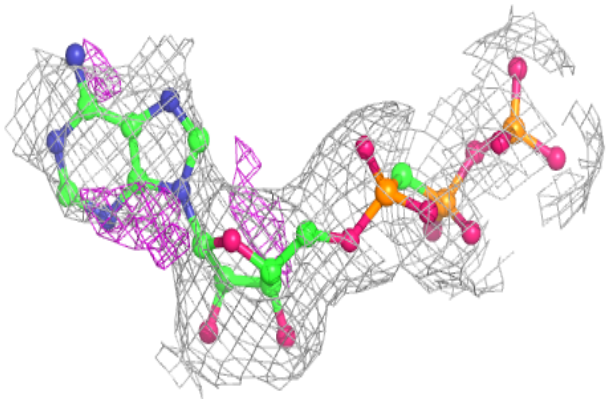


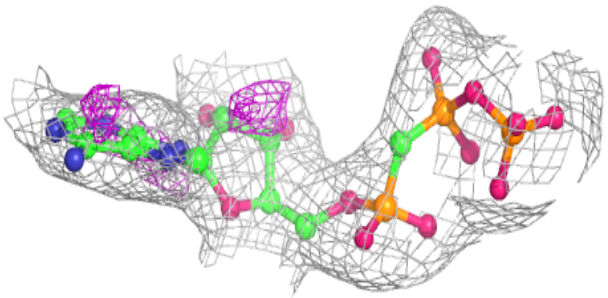


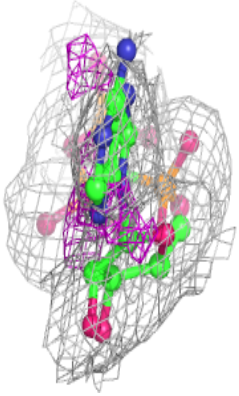


Electron density around APC B 2001:

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

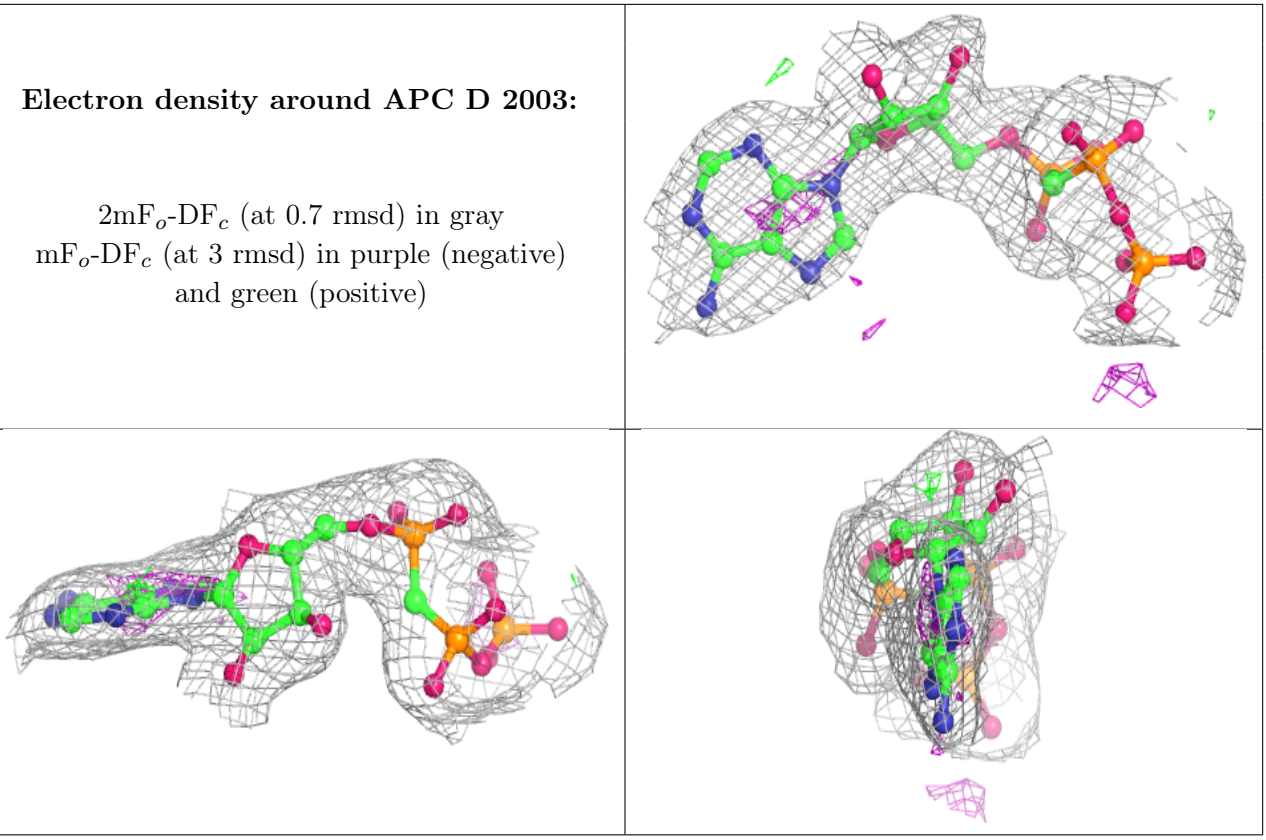






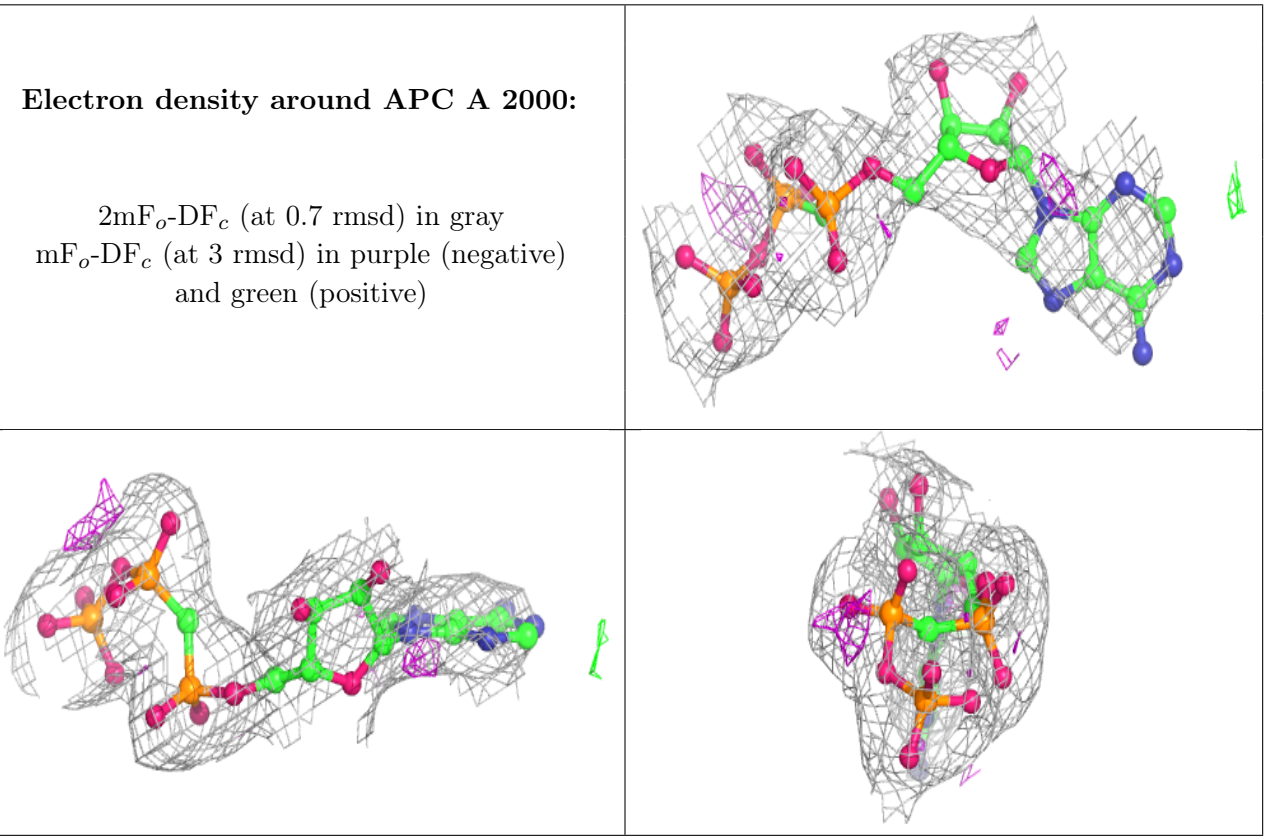
Electron density around APC D 2003:

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



Electron density around APC A 2000:

$2mF_o$-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.