wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 10:35 AM EDT

PDB ID	:	2ZPA
Title	:	Crystal Structure of tRNA(Met) Cytidine Acetyltransferase
Authors	:	Chimnaronk, S.; Manita, T.; Yao, M.; Tanaka, I.
Deposited on	:	2008-07-08
Resolution	:	2.35 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:

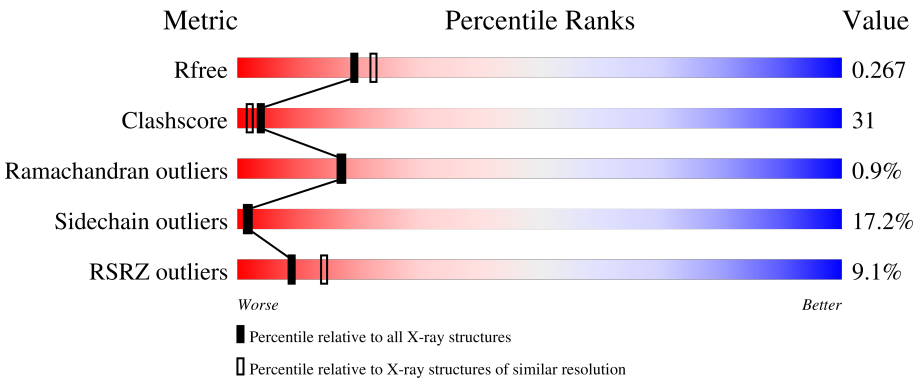
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	671	4%, 61%, 29%, 8%, •
1	B	671	14%, 48%, 38%, 11%, ••

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
2	ACO	A	700	X	-	-	-
2	ACO	B	701	X	-	-	-

2 Entry composition (i)

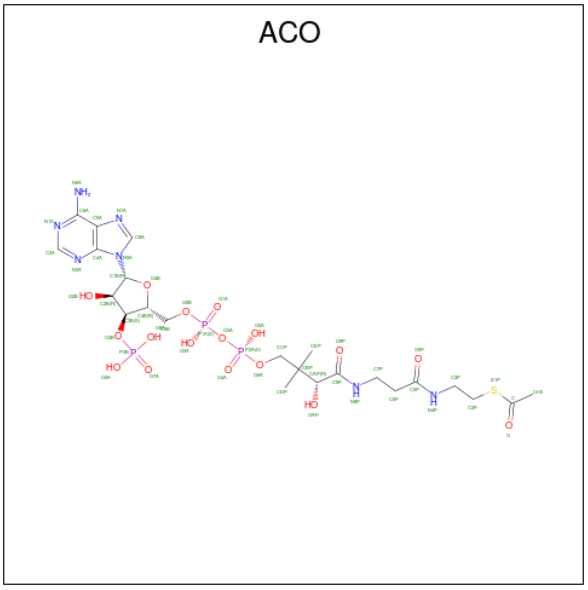
There are 5 unique types of molecules in this entry. The entry contains 10891 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Uncharacterized protein ypfI.

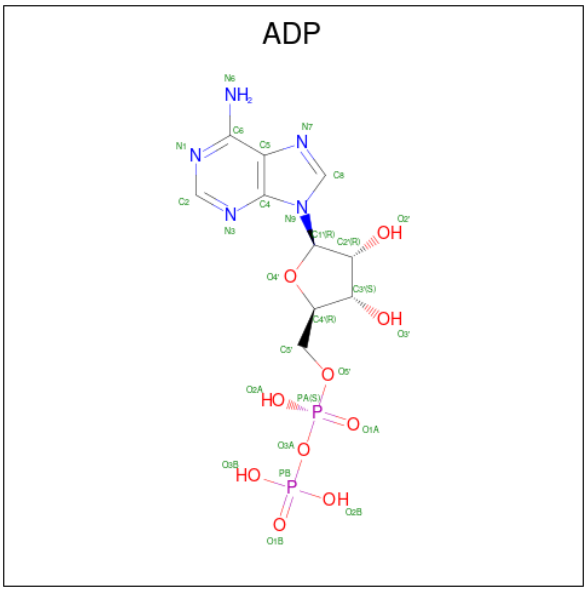
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	658	5189	3291	951	929	8	10	0	0	0
1	B	662	5218	3309	955	935	8	11	0	0	0

- Molecule 2 is ACETYL COENZYME *A (three-letter code: ACO) (formula: $C_{23}H_{38}N_7O_{17}P_3S$).



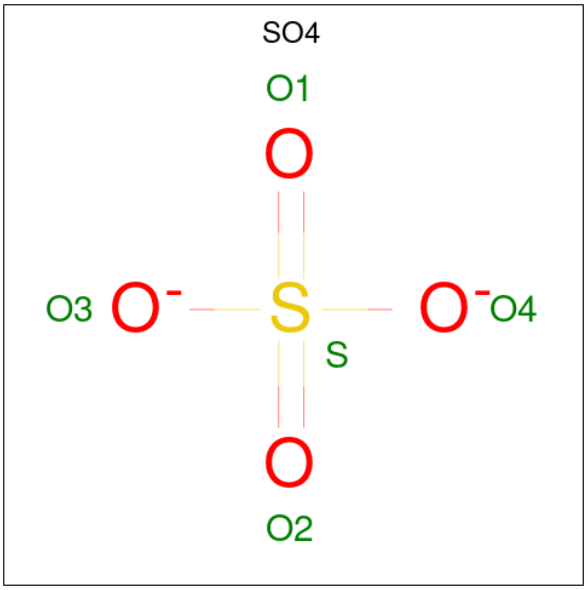
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	N	O	P	S		
2	A	1	51	23	7	17	3	1	0	0
2	B	1	51	23	7	17	3	1	0	0

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	27	10	5	10	2	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
4	B	1	5	4	1	0	0

- Molecule 5 is water.

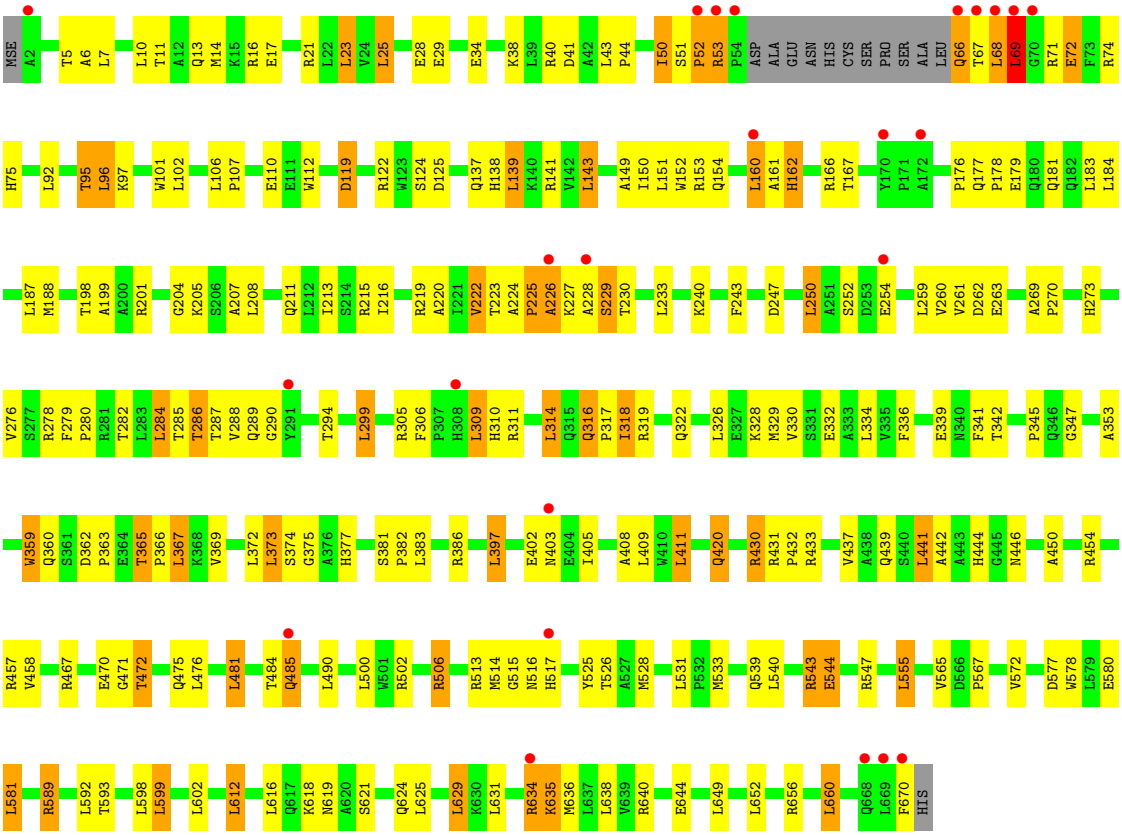
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	224	224	224	0	0
			Total	O		
5	B	126	126	126	0	0

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.

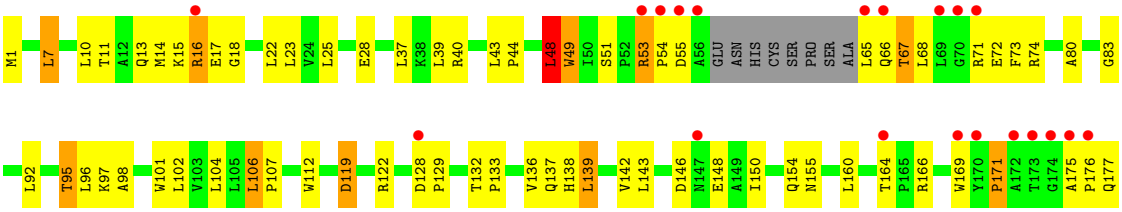
• Molecule 1: Uncharacterized protein ypfI

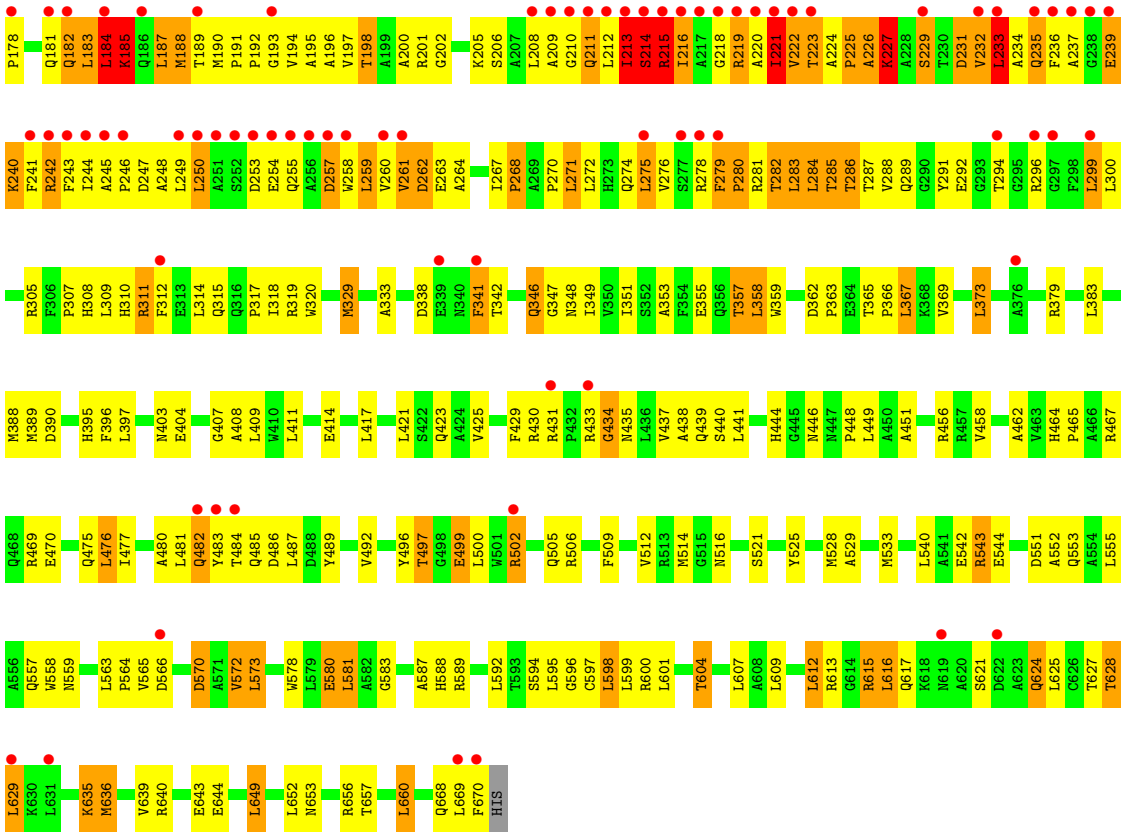
Chain A:



• Molecule 1: Uncharacterized protein ypfI

Chain B:




P178 Q181 Q182 L183 L184 K185 Q186 L187 M188 T189 M190 P191 P192 G193 V194 A195 A196 V197 T198 A199 A200 R201 G202 K205 S206 A207 L208 A209 G210 Q211 L212 I213 S214 R215 I216 A217 G218 R219 A220 I221 V222 T223 A224 P225 A226 K227 A228 S229 T230 D231 V232 L233 A234 Q235 F236 A237 G238 E239
K240 F241 R242 F243 I244 A245 P246 D247 A248 L249 L250 A251 S252 D253 E254 Q255 A256 D257 W258 L259 V260 V261 D262 E263 A264 I267 P268 A269 P270 L271 L272 H273 Q274 L275 V276 S277 R278 F279 P280 R281 T282 L283 L284 T285 T286 T287 V288 Q289 G290 Y291 E292 G293 T294 G295 R296 G297 F298 L299 L300
B305 F306 P307 H308 L309 H310 R311 F312 E313 L314 Q315 Q316 P317 I318 R319 W320 M329 A333 D338 E339 N340 F341 T342 Q346 G347 N348 I349 V350 I351 S352 A353 F354 E355 Q356 T357 L358 W359 D362 P363 E364 T365 P366 L367 K368 V369 L373 A376 B379 L383
M388 M389 D390 H395 F396 L397 N403 E404 G407 A408 L409 W410 L411 E414 L417 L421 S422 Q423 A424 V425 F429 R430 R431 P432 R433 G434 N435 L436 V437 A438 Q439 S440 L441 H444 G445 N446 N447 P448 L449 A450 A451 R456 R457 V458 A462 V463 H464 P465 A466 R467
Q468 R469 E470 Q475 L476 I477 A480 L481 Q482 Y483 T484 Q485 D486 L487 D488 Y489 V492 Y496 T497 G498 E499 L500 W501 R502 Q505 R506 F509 V512 R513 M514 G515 N516 S521 Y525 M528 A529 M533 L540 A541 E542 R543 E544 D551 A552 Q553 A554 L555
A556 Q557 W558 N559 L563 P564 V565 D566 D570 A571 V572 L573 W578 L579 E580 L581 A582 G583 A587 H588 R589 L592 T593 S594 L595 G596 C597 L598 L599 R600 L601 T604 L607 A608 L609 L612 R613 G614 R615 L616 Q617 K618 N619 A620 S621 D622 A623 Q624 L625 G626 T627 T628
L629 K630 L631 K635 M636 V639 R640 E643 E644 L649 L652 N653 R656 T657 L660 Q668 L669 F670 HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.30Å 100.99Å 263.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.35 40.99 – 2.35	Depositor EDS
% Data completeness (in resolution range)	95.8 (20.00-2.35) 96.3 (40.99-2.35)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	4.49 (at 2.37Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.235 , 0.274 0.227 , 0.267	Depositor DCC
R_{free} test set	5085 reflections (7.68%)	wwPDB-VP
Wilson B-factor (Å^2)	39.3	Xtriage
Anisotropy	0.373	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	10891	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.58% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, ADP, ACO

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	0.48	0/5306	0.77	7/7197 (0.1%)
1	B	0.49	2/5335 (0.0%)	0.82	16/7236 (0.2%)
All	All	0.48	2/10641 (0.0%)	0.79	23/14433 (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	A	0	1
1	B	0	2
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	182	GLN	C-N	5.74	1.47	1.34
1	B	617	GLN	N-CA	5.51	1.57	1.46

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	201	ARG	CD-NE-CZ	14.20	143.47	123.60
1	A	66	GLN	C-N-CA	-10.02	96.65	121.70
1	B	49	TRP	N-CA-CB	9.03	126.86	110.60
1	B	187	LEU	N-CA-C	-8.68	87.57	111.00
1	A	226	ALA	N-CA-C	-8.12	89.08	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	359	TRP	Mainchain
1	B	48	LEU	Peptide
1	B	656	ARG	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	A	5189	0	5151	230	0
1	B	5218	0	5183	413	0
2	A	51	0	30	3	0
2	B	51	0	33	7	0
3	A	27	0	12	2	0
4	B	5	0	0	1	0
5	A	224	0	0	13	0
5	B	126	0	0	8	0
All	All	10891	0	10409	643	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 31.

The worst 5 of 643 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)
1:A:67:THR:CG2	1:A:71:ARG:HD3	1.36	1.51
1:B:185:LYS:HG3	1:B:215:ARG:NH2	1.51	1.25
1:B:274:GLN:HB3	1:B:278:ARG:NH2	1.62	1.14
1:B:221:ILE:HG23	1:B:259:LEU:HA	1.34	1.08
1:B:215:ARG:NH1	1:B:215:ARG:HG3	1.60	1.08

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
1	A	654/671 (98%)	633 (97%)	18 (3%)	3 (0%)	29 32
1	B	658/671 (98%)	605 (92%)	44 (7%)	9 (1%)	11 9
All	All	1312/1342 (98%)	1238 (94%)	62 (5%)	12 (1%)	17 17

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	184	LEU
1	B	227	LYS
1	B	215	ARG
1	B	232	VAL
1	A	225	PRO

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
1	A	532/532 (100%)	460 (86%)	72 (14%)	4 3
1	B	535/532 (101%)	424 (79%)	111 (21%)	1 1
All	All	1067/1064 (100%)	884 (83%)	183 (17%)	2 1

5 of 183 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	247	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	348	ASN
1	B	255	GLN
1	B	283	LEU
1	B	433	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	371	GLN
1	B	419	GLN
1	B	446	ASN
1	A	444	HIS
1	A	423	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

4 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	ACO	A	700	-	47,53,53	3.81	17 (36%)	60,79,79	5.40	31 (51%)
2	ACO	B	701	-	47,53,53	2.95	18 (38%)	60,79,79	4.65	31 (51%)
3	ADP	A	800	-	24,29,29	1.36	2 (8%)	29,45,45	1.79	5 (17%)
4	SO4	B	801	-	4,4,4	0.35	0	6,6,6	0.15	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ACO	A	700	-	1/1/12/14	17/47/67/67	0/3/3/3
2	ACO	B	701	-	1/1/12/14	12/47/67/67	0/3/3/3
3	ADP	A	800	-	-	6/12/32/32	0/3/3/3

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	700	ACO	O5P-C5P	17.50	1.58	1.23
2	B	701	ACO	O5P-C5P	8.20	1.39	1.23
2	A	700	ACO	O9P-C9P	7.31	1.37	1.23
2	B	701	ACO	P1A-O3A	-6.97	1.52	1.59
2	A	700	ACO	P1A-O3A	-6.38	1.52	1.59

The worst 5 of 67 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	700	ACO	C2P-C3P-N4P	16.47	146.81	112.41
2	A	700	ACO	CEP-CBP-CAP	15.46	135.13	108.77
2	A	700	ACO	O5P-C5P-C6P	-15.37	94.17	122.02
2	B	701	ACO	O5P-C5P-C6P	-15.24	94.41	122.02
2	B	701	ACO	C2P-C3P-N4P	14.14	141.94	112.41

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	700	ACO	C2B
2	B	701	ACO	C2B

5 of 35 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	700	ACO	O4B-C4B-C5B-O5B
2	A	700	ACO	C5B-O5B-P1A-O3A
2	A	700	ACO	CEP-CBP-CCP-O6A
2	A	700	ACO	C9P-CAP-CBP-CCP
2	A	700	ACO	C9P-CAP-CBP-CDP

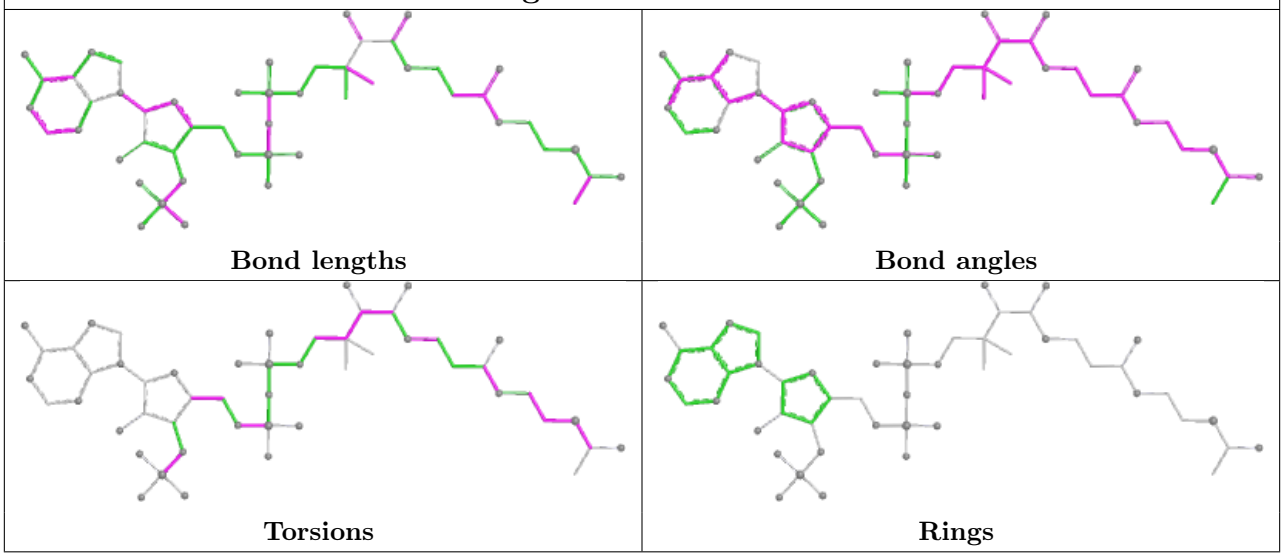
There are no ring outliers.

4 monomers are involved in 13 short contacts:

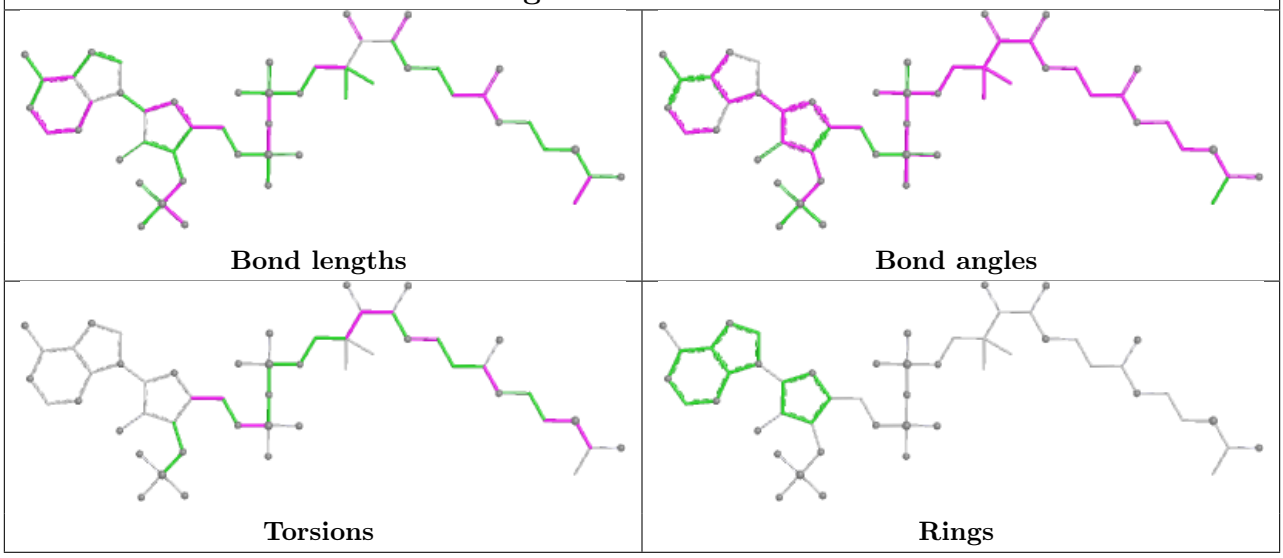
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	700	ACO	3	0
2	B	701	ACO	7	0
3	A	800	ADP	2	0
4	B	801	SO4	1	0

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 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.

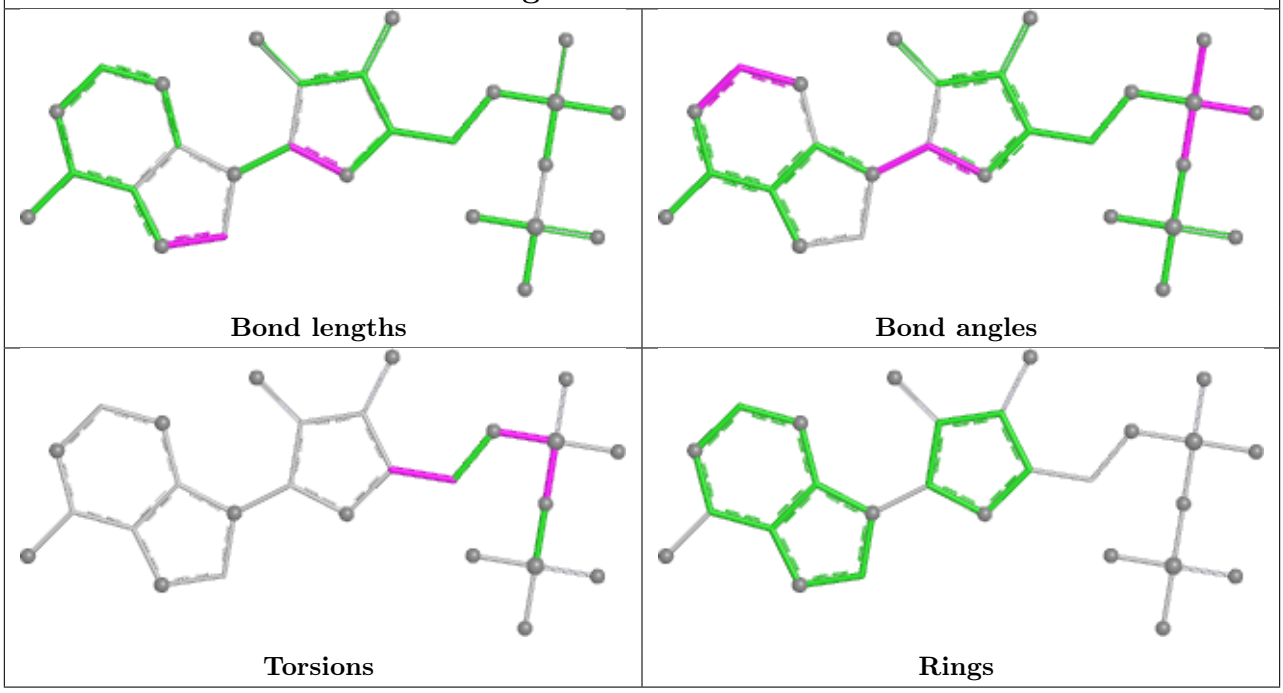
Ligand ACO A 700



Ligand ACO B 701



Ligand ADP A 800



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	648/671 (96%)	0.24	24 (3%) 41 54	18, 35, 62, 90	0
1	B	651/671 (97%)	0.95	94 (14%) 2 4	22, 49, 99, 100	0
All	All	1299/1342 (96%)	0.59	118 (9%) 9 14	18, 42, 95, 100	0

The worst 5 of 118 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	221	ILE	10.6
1	B	54	PRO	9.7
1	B	244	ILE	9.5
1	B	216	ILE	8.2
1	A	54	PRO	8.2

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

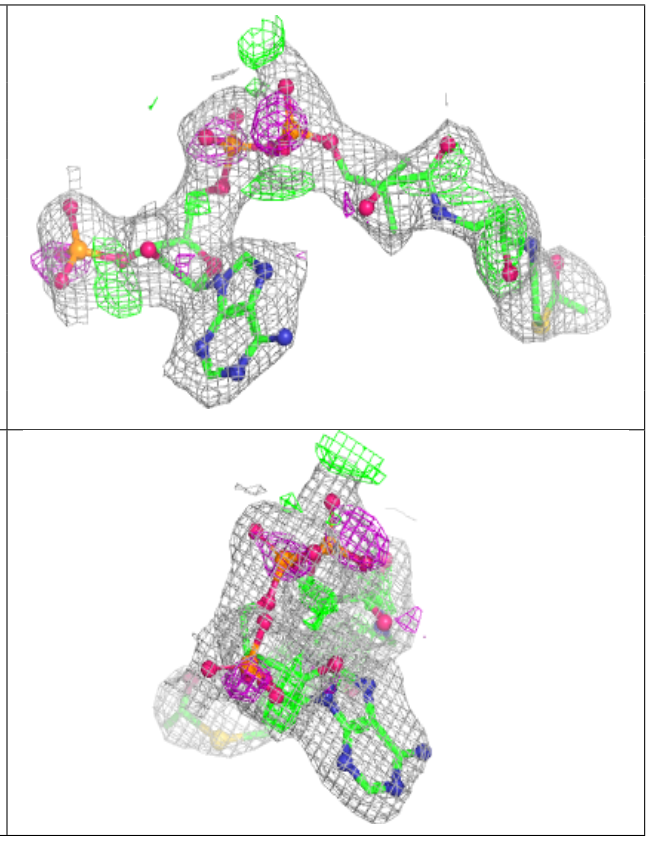
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	ACO	B	701	51/51	0.86	0.23	24,50,93,95	0
2	ACO	A	700	51/51	0.87	0.25	18,66,88,89	0
4	SO4	B	801	5/5	0.91	0.15	77,78,78,79	0
3	ADP	A	800	27/27	0.94	0.15	29,38,39,43	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.

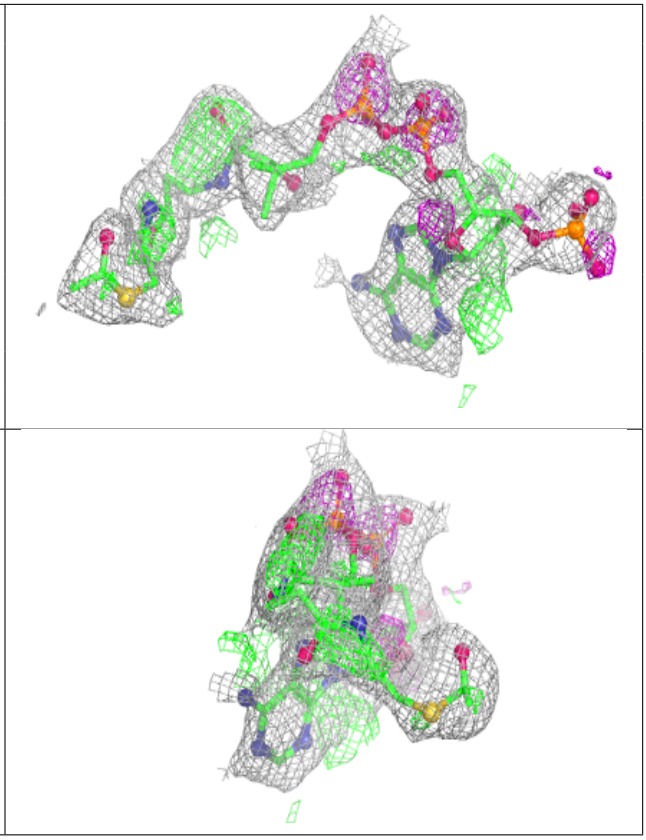
Electron density around ACO B 701:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



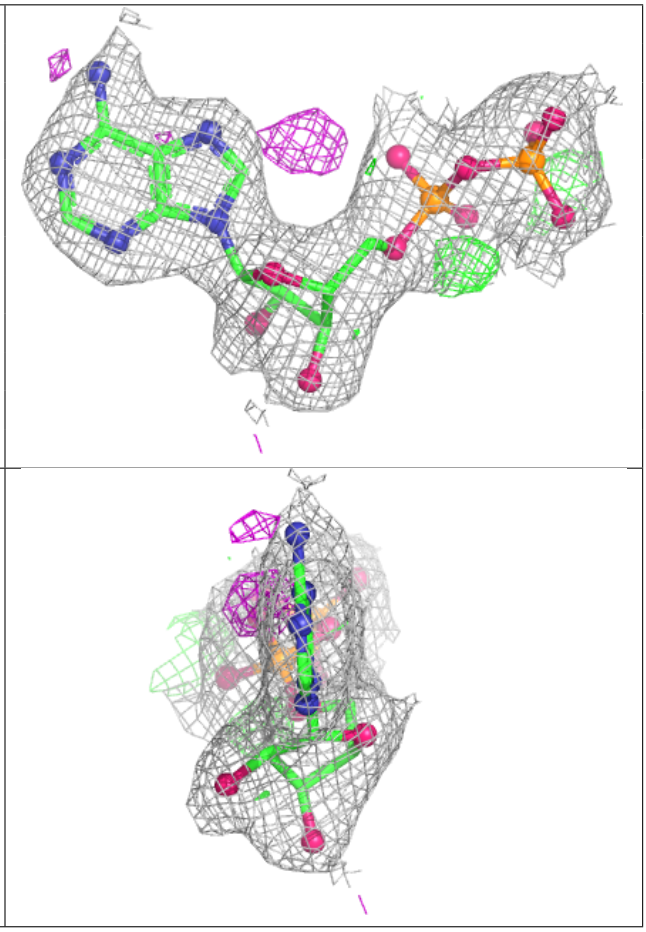
Electron density around ACO A 700:

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



Electron density around ADP A 800:

$2mF_o$-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.