



# Full wwPDB X-ray Structure Validation Report ⓘ

May 19, 2020 – 11:04 pm BST

PDB ID : 4R1L  
Title : Crystal structure of a Putative Acyl-CoA ligase (BT\_0428) from Bacteroides  
thetaiotaomicron VPI-5482 at 2.42 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2014-08-06  
Resolution : 2.42 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

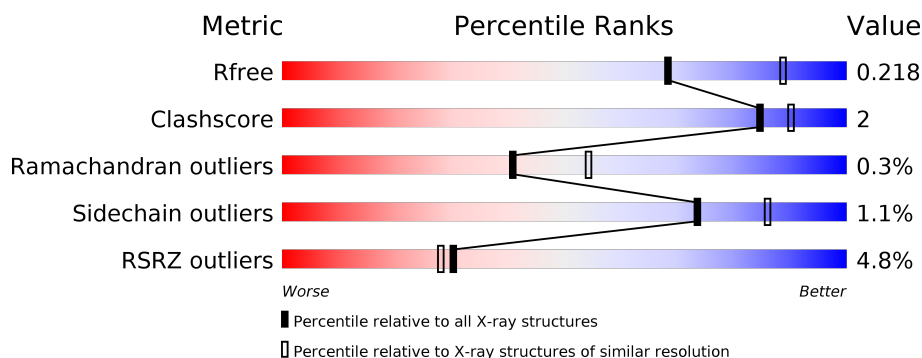
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.42 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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	436	<div> <div>2%</div> <div> <div></div> <div>94%</div> <div>5%</div> </div> </div>
1	B	436	<div> <div>11%</div> <div> <div></div> <div>89%</div> <div>8%</div> </div> </div>
1	C	436	<div> <div>3%</div> <div> <div></div> <div>92%</div> <div>7%</div> </div> </div>
1	D	436	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>6%</div> </div> </div>

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 14432 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 Phenylacetate-coenzyme A ligase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	433	Total	C	N	O	S	Se	0	10	0
			3469	2187	602	659	6	15			
1	B	422	Total	C	N	O	S	Se	0	8	0
			3336	2102	578	635	6	15			
1	C	434	Total	C	N	O	S	Se	0	4	0
			3430	2165	597	647	6	15			
1	D	425	Total	C	N	O	S	Se	0	3	0
			3340	2108	578	633	6	15			

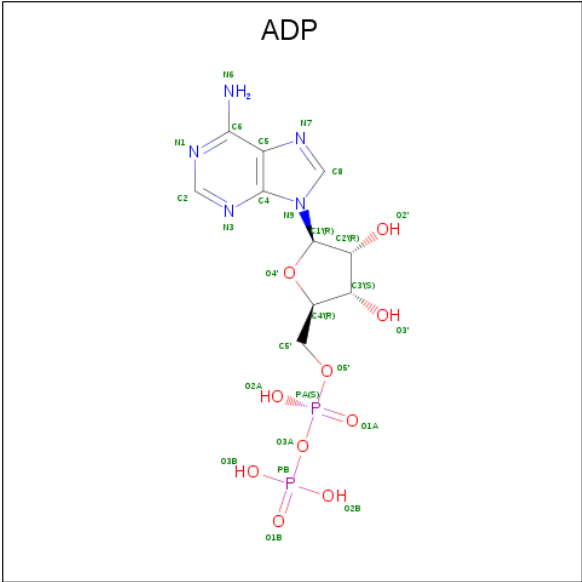
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	LEADER SEQUENCE	UNP Q8AAN6
B	0	GLY	-	LEADER SEQUENCE	UNP Q8AAN6
C	0	GLY	-	LEADER SEQUENCE	UNP Q8AAN6
D	0	GLY	-	LEADER SEQUENCE	UNP Q8AAN6

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

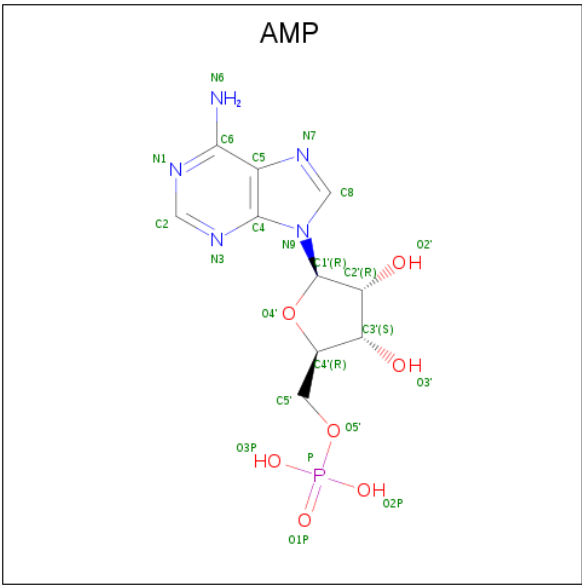
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>10</sub>P<sub>2</sub>).



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

- Molecule 4 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: C<sub>10</sub>H<sub>14</sub>N<sub>5</sub>O<sub>7</sub>P).



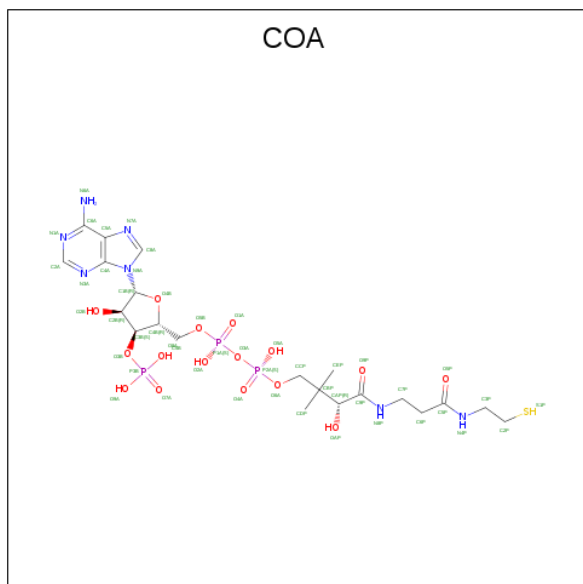
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	1
			23	10	5	7	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	C	1	Total	C	N	O	P	0	0
			23	10	5	7	1		

- Molecule 5 is COENZYME A (three-letter code: COA) (formula:  $C_{21}H_{36}N_7O_{16}P_3S$ ).

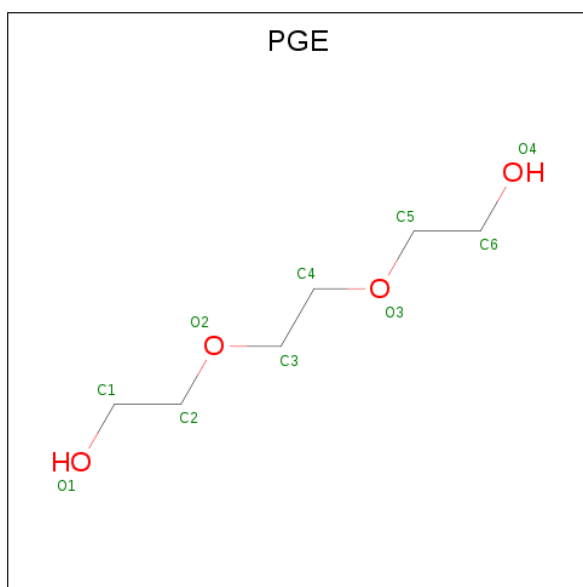


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	1
			31	10	5	13	3		
5	C	1	Total	C	N	O	P	0	1
			31	10	5	13	3		

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

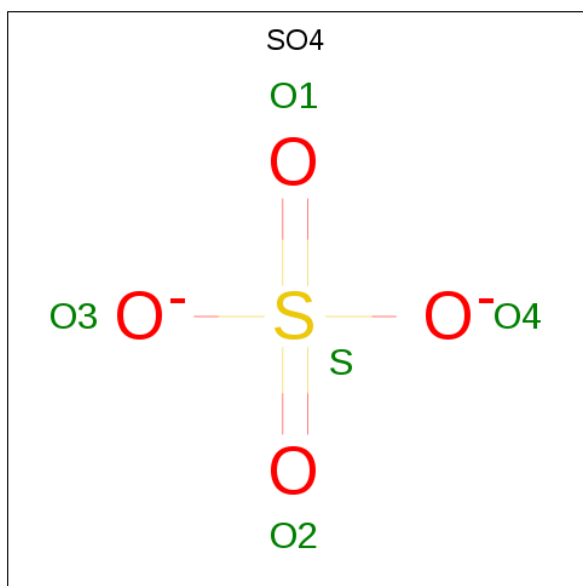
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	3	Total K 3 3	0	0
6	D	1	Total K 1 1	0	0
6	C	2	Total K 2 2	0	0

- Molecule 7 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



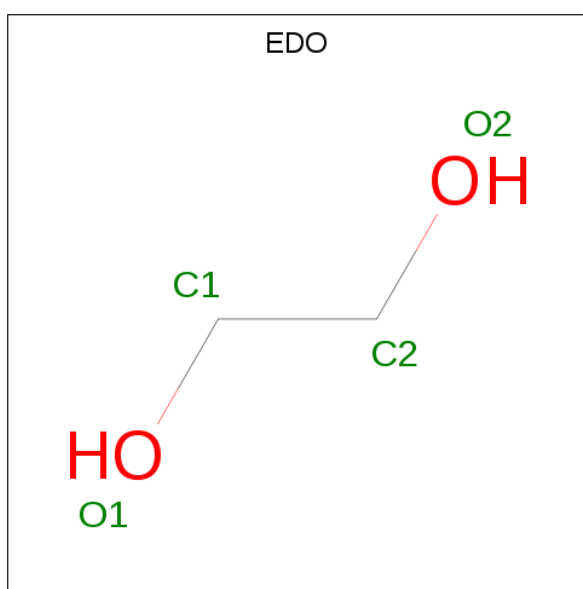
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	O	S	0	1
			5	4	1		
8	A	1	Total	O	S	0	0
			5	4	1		
8	A	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 9 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



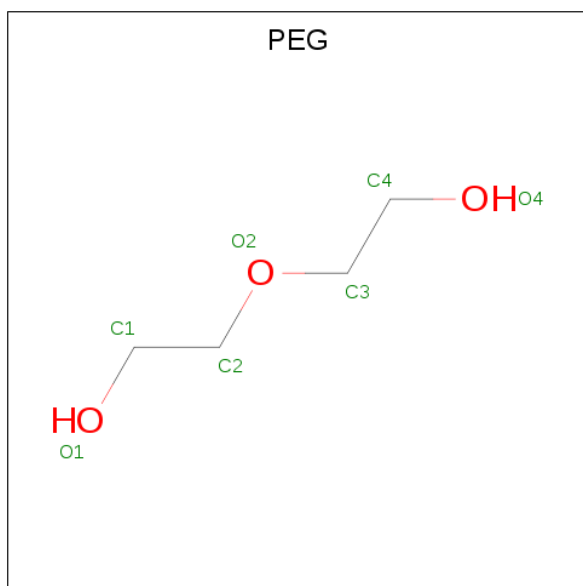
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	B	1	Total	C	O	0	0
			4	2	2		
9	B	1	Total	C	O	0	0
			4	2	2		
9	B	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	1	Total	C	O	0	0
			4	2	2		
9	C	1	Total	C	O	0	0
			4	2	2		
9	C	1	Total	C	O	0	0
			4	2	2		
9	D	1	Total	C	O	0	0
			4	2	2		
9	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 10 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	D	1	Total	C	O	0	0
			7	4	3		

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	173	Total	O	0	1
			174	174		
11	B	124	Total	O	0	0
			124	124		
11	C	114	Total	O	0	1
			115	115		

*Continued on next page...*



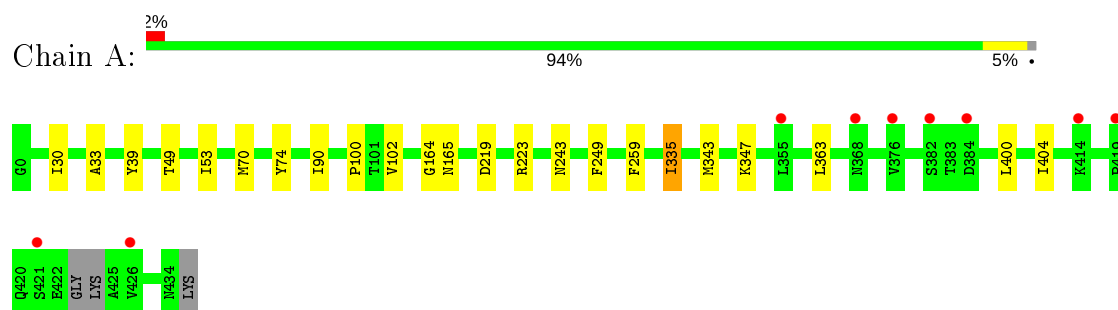
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	D	145	Total	O	0	0
			145	145		

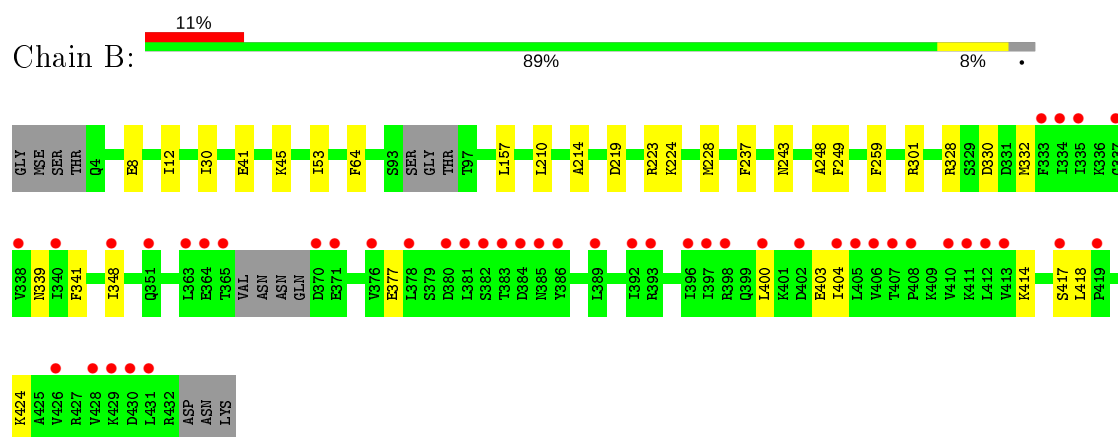
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

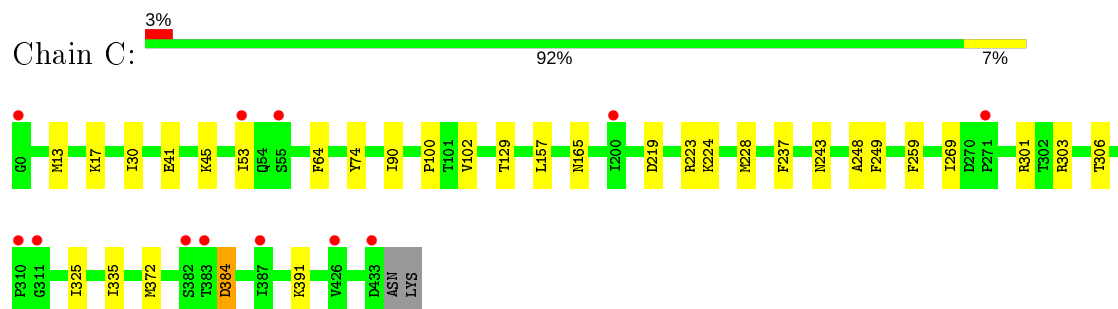
- Molecule 1: Phenylacetate-coenzyme A ligase



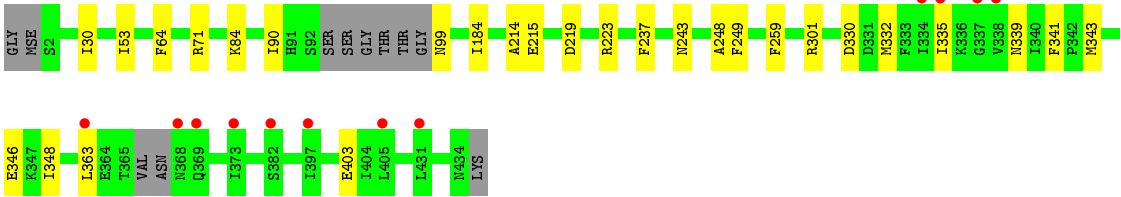
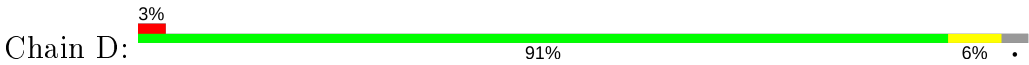
- Molecule 1: Phenylacetate-coenzyme A ligase



- Molecule 1: Phenylacetate-coenzyme A ligase



- Molecule 1: Phenylacetate-coenzyme A ligase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.70 Å   211.14 Å   71.86 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	48.78 – 2.42 48.78 – 2.42	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.78-2.42) 98.7 (48.78-2.42)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.58 (at 2.42 Å)	Xtriage
Refinement program	BUSTER-TNT 2.10.0, BUSTER 2.10.0	Depositor
R, $R_{free}$	0.172 , 0.213 0.180 , 0.218	Depositor DCC
$R_{free}$ test set	3742 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.2	Xtriage
Anisotropy	0.569	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 51.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	14432	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.81% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, PGE, ADP, K, EDO, COA, SO4, AMP, PEG

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	0.52	0/3542	0.68	0/4756
1	B	0.49	0/3402	0.66	0/4573
1	C	0.48	0/3486	0.66	0/4682
1	D	0.51	0/3392	0.66	0/4562
All	All	0.50	0/13822	0.66	0/18573

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	3469	0	3476	13	0
1	B	3336	0	3279	17	0
1	C	3430	0	3431	16	0
1	D	3340	0	3283	16	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	27	0	12	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	27	0	12	0	0
3	D	27	0	12	0	0
4	A	23	0	12	0	0
4	C	23	0	12	0	0
5	A	31	0	11	1	0
5	C	31	0	11	0	0
6	A	3	0	0	0	0
6	C	2	0	0	0	0
6	D	1	0	0	0	0
7	A	10	0	14	1	0
8	A	15	0	0	0	0
8	B	5	0	0	0	0
8	C	10	0	0	0	0
8	D	5	0	0	0	0
9	A	16	0	24	2	0
9	B	12	0	18	1	0
9	C	12	0	18	3	0
9	D	8	0	12	0	0
10	D	7	0	10	2	0
11	A	174	0	0	1	0
11	B	124	0	0	1	0
11	C	115	0	0	1	0
11	D	145	0	0	1	0
All	All	14432	0	13647	63	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 2.

All (63) 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:D:332[A]:MSE:HE2	1:D:339:ASN:HB3	1.69	0.74
1:B:332[A]:MSE:HE2	1:B:339:ASN:HB3	1.72	0.71
1:A:219:ASP:O	1:A:223:ARG:HG2	1.93	0.67
1:A:90:ILE:HG23	1:A:102:VAL:HG13	1.76	0.67
9:C:510:EDO:H11	11:C:647:HOH:O	1.95	0.66
1:A:30:ILE:HD13	1:A:53:ILE:HG13	1.77	0.66
1:D:30:ILE:HD13	1:D:53:ILE:HG13	1.77	0.65
1:C:303:ARG:HH11	9:C:510:EDO:H12	1.62	0.64
1:B:30:ILE:HD13	1:B:53:ILE:HG13	1.81	0.62
1:C:30:ILE:HD13	1:C:53:ILE:HG13	1.80	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:224:LYS:O	1:B:228:MSE:HG3	2.03	0.59
1:B:214:ALA:HB1	1:B:424:LYS:HE2	1.85	0.58
1:D:348:ILE:HD11	1:D:403:GLU:CB	2.32	0.58
10:D:504:PEG:H11	11:D:737:HOH:O	2.05	0.56
1:B:400:LEU:O	1:B:404:ILE:HG12	2.08	0.54
1:B:414:LYS:HB3	1:B:417:SER:OG	2.08	0.53
1:D:343:MSE:HE1	1:D:346:GLU:OE1	2.09	0.52
1:C:237:PHE:HB3	1:C:248:ALA:HB3	1.92	0.52
1:A:49:THR:HA	9:A:512:EDO:H21	1.90	0.52
1:C:90:ILE:HG23	1:C:102:VAL:HG13	1.92	0.52
1:D:71:ARG:HD3	1:D:99:ASN:HB2	1.91	0.51
1:A:70:MSE:O	9:A:514:EDO:H11	2.12	0.50
1:B:237:PHE:HB3	1:B:248:ALA:HB3	1.93	0.50
1:C:219:ASP:O	1:C:223:ARG:HG2	2.12	0.49
9:B:505:EDO:H11	11:B:653:HOH:O	2.12	0.49
1:D:30:ILE:CD1	1:D:53:ILE:HG13	2.41	0.49
1:D:335:ILE:HD13	1:D:363:LEU:HD22	1.94	0.49
1:A:30:ILE:CD1	1:A:53:ILE:HG13	2.42	0.49
1:A:74:TYR:CE1	1:A:100:PRO:HD2	2.49	0.48
1:D:332[A]:MSE:HE3	1:D:341:PHE:CE2	2.49	0.48
1:C:74:TYR:CE1	1:C:100:PRO:HD2	2.48	0.48
1:B:377:GLU:HG3	1:B:418:LEU:HD12	1.94	0.48
1:D:237:PHE:HB3	1:D:248:ALA:HB3	1.95	0.47
1:C:41:GLU:HG3	1:C:45:LYS:HE2	1.97	0.47
1:C:13:MSE:HG3	1:C:17:LYS:HE3	1.97	0.47
1:A:164:GLY:HA2	5:A:504[A]:COA:H1B	1.96	0.46
1:B:219:ASP:O	1:B:223:ARG:HG2	2.15	0.46
1:B:30:ILE:CD1	1:B:53:ILE:HG13	2.44	0.46
1:B:332[A]:MSE:HE3	1:B:341:PHE:CE2	2.50	0.46
1:D:64:PHE:HB3	1:D:301:ARG:HG3	1.98	0.45
1:C:335:ILE:HD13	1:C:372:MSE:HB2	1.98	0.45
1:C:30:ILE:CD1	1:C:53:ILE:HG13	2.43	0.45
1:B:8:GLU:O	1:B:12:ILE:HG12	2.16	0.45
1:D:219:ASP:O	1:D:223:ARG:HG2	2.16	0.45
1:A:90:ILE:HD11	1:B:157:LEU:HD21	1.98	0.45
1:B:41:GLU:OE2	1:B:45:LYS:HE2	2.16	0.45
1:C:64:PHE:HB3	1:C:301:ARG:HG3	1.98	0.44
1:A:400:LEU:O	1:A:404:ILE:HG12	2.17	0.44
1:A:33:ALA:O	1:A:39:TYR:HB2	2.17	0.43
1:B:64:PHE:HB3	1:B:301:ARG:HG3	2.01	0.43
1:B:348:ILE:HD11	1:B:403:GLU:HB3	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:214:ALA:HB2	10:D:504:PEG:H12	2.00	0.43
7:A:508:PGE:H6	11:A:772:HOH:O	2.19	0.42
1:C:224:LYS:O	1:C:228:MSE:HG3	2.19	0.42
1:D:184:ILE:HD12	1:D:215:GLU:OE2	2.20	0.41
1:C:129:THR:HB	1:D:84:LYS:HE2	2.02	0.41
1:C:306:THR:HB	1:C:325:ILE:HD13	2.03	0.41
1:D:348:ILE:HD11	1:D:403:GLU:HB3	2.01	0.41
1:C:303:ARG:NH1	9:C:510:EDO:H12	2.34	0.41
1:A:343[B]:MSE:HG3	1:A:347:LYS:HE3	2.03	0.41
1:A:335:ILE:HD13	1:A:363:LEU:HD22	2.02	0.41
1:C:157:LEU:HD21	1:D:90:ILE:HD11	2.02	0.41
1:B:328:ARG:HD3	1:B:330:ASP:OD2	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	439/436 (101%)	420 (96%)	18 (4%)	1 (0%)	47	61
1	B	423/436 (97%)	408 (96%)	14 (3%)	1 (0%)	47	61
1	C	436/436 (100%)	419 (96%)	15 (3%)	2 (0%)	29	40
1	D	422/436 (97%)	406 (96%)	15 (4%)	1 (0%)	47	61
All	All	1720/1744 (99%)	1653 (96%)	62 (4%)	5 (0%)	41	54

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	243	ASN
1	B	243	ASN
1	C	243	ASN

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	D	243	ASN
1	C	384	ASP

### 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	382/370 (103%)	378 (99%)	4 (1%)	76	87
1	B	360/370 (97%)	357 (99%)	3 (1%)	81	91
1	C	373/370 (101%)	367 (98%)	6 (2%)	62	78
1	D	358/370 (97%)	355 (99%)	3 (1%)	81	91
All	All	1473/1480 (100%)	1457 (99%)	16 (1%)	73	86

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	165	ASN
1	A	249	PHE
1	A	259	PHE
1	A	335	ILE
1	B	210	LEU
1	B	249	PHE
1	B	259	PHE
1	C	165	ASN
1	C	249	PHE
1	C	259	PHE
1	C	269	ILE
1	C	384	ASP
1	C	391	LYS
1	D	249	PHE
1	D	259	PHE
1	D	330	ASP

Some 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 ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 38 ligands modelled in this entry, 10 are monoatomic - leaving 28 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	$\# Z  > 2$
8	SO4	C	507	-	4,4,4	0.14	0	6,6,6	0.12	0
8	SO4	D	505	-	4,4,4	0.15	0	6,6,6	0.29	0
9	EDO	B	505	-	3,3,3	0.65	0	2,2,2	0.25	0
9	EDO	B	506	-	3,3,3	0.74	0	2,2,2	0.04	0
3	ADP	D	502	-	24,29,29	0.61	0	29,45,45	0.88	2 (6%)
5	COA	C	503[A]	-	28,33,50	0.79	1 (3%)	35,52,75	1.33	6 (17%)
8	SO4	B	503	-	4,4,4	0.20	0	6,6,6	0.15	0
10	PEG	D	504	-	6,6,6	0.27	0	5,5,5	0.33	0
8	SO4	A	509[B]	-	4,4,4	0.10	0	6,6,6	0.07	0
8	SO4	C	506[B]	-	4,4,4	0.14	0	6,6,6	0.08	0
8	SO4	A	511	-	4,4,4	0.17	0	6,6,6	0.15	0
9	EDO	C	508	-	3,3,3	0.55	0	2,2,2	0.46	0
4	AMP	C	502	6	22,25,25	0.59	0	25,38,38	0.93	1 (4%)
8	SO4	A	510	-	4,4,4	0.28	0	6,6,6	0.39	0
9	EDO	D	507	-	3,3,3	0.67	0	2,2,2	0.31	0
9	EDO	C	510	-	3,3,3	0.65	0	2,2,2	0.29	0
3	ADP	A	502[A]	6	24,29,29	0.64	0	29,45,45	0.88	2 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	EDO	D	506	-	3,3,3	0.68	0	2,2,2	0.13	0
9	EDO	A	513	-	3,3,3	0.57	0	2,2,2	0.31	0
9	EDO	A	514	-	3,3,3	0.40	0	2,2,2	0.28	0
3	ADP	B	502	-	24,29,29	0.65	0	29,45,45	0.78	2 (6%)
5	COA	A	504[A]	-	28,33,50	0.85	1 (3%)	35,52,75	1.47	7 (20%)
4	AMP	A	503[B]	6	22,25,25	0.69	0	25,38,38	0.82	1 (4%)
7	PGE	A	508	-	9,9,9	0.20	0	8,8,8	0.22	0
9	EDO	A	515	-	3,3,3	0.81	0	2,2,2	0.22	0
9	EDO	B	504	-	3,3,3	0.53	0	2,2,2	0.34	0
9	EDO	A	512	-	3,3,3	0.59	0	2,2,2	0.12	0
9	EDO	C	509	-	3,3,3	0.65	0	2,2,2	0.30	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
9	EDO	D	507	-	-	0/1/1/1	-
9	EDO	B	506	-	-	0/1/1/1	-
4	AMP	A	503[B]	6	-	0/6/26/26	0/3/3/3
7	PGE	A	508	-	-	4/7/7/7	-
9	EDO	A	515	-	-	0/1/1/1	-
9	EDO	C	510	-	-	1/1/1/1	-
3	ADP	A	502[A]	6	-	6/12/32/32	0/3/3/3
9	EDO	D	506	-	-	0/1/1/1	-
9	EDO	B	505	-	-	0/1/1/1	-
10	PEG	D	504	-	-	1/4/4/4	-
4	AMP	C	502	6	-	0/6/26/26	0/3/3/3
9	EDO	C	508	-	-	0/1/1/1	-
9	EDO	A	514	-	-	0/1/1/1	-
3	ADP	D	502	-	-	6/12/32/32	0/3/3/3
5	COA	C	503[A]	-	-	0/17/37/64	0/3/3/3
9	EDO	A	513	-	-	0/1/1/1	-
9	EDO	B	504	-	-	0/1/1/1	-
9	EDO	A	512	-	-	0/1/1/1	-
5	COA	A	504[A]	-	-	0/17/37/64	0/3/3/3
9	EDO	C	509	-	-	0/1/1/1	-
3	ADP	B	502	-	-	4/12/32/32	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	504[A]	COA	P3B-O3B	2.68	1.64	1.59
5	C	503[A]	COA	P2A-O6A	2.06	1.62	1.54

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	504[A]	COA	O6A-P2A-O3A	4.83	120.83	104.64
5	C	503[A]	COA	O6A-P2A-O3A	4.26	118.91	104.64
5	A	504[A]	COA	O5A-P2A-O4A	3.08	122.75	110.68
5	A	504[A]	COA	O6A-P2A-O5A	-2.73	97.21	107.64
4	C	502	AMP	C5-C6-N6	2.65	124.39	120.35
3	D	502	ADP	O5'-PA-O1A	2.48	118.76	109.07
5	C	503[A]	COA	O6A-P2A-O5A	-2.42	98.38	107.64
4	A	503[B]	AMP	C5-C6-N6	2.40	123.99	120.35
3	D	502	ADP	C5-C6-N6	2.39	123.98	120.35
3	B	502	ADP	O5'-PA-O1A	2.31	118.11	109.07
3	B	502	ADP	C5-C6-N6	2.29	123.84	120.35
3	A	502[A]	ADP	O5'-PA-O1A	2.28	117.98	109.07
3	A	502[A]	ADP	C5-C6-N6	2.27	123.80	120.35
5	A	504[A]	COA	P1A-O3A-P2A	-2.26	125.06	132.83
5	C	503[A]	COA	O5A-P2A-O4A	2.26	119.54	110.68
5	A	504[A]	COA	C5A-C6A-N6A	2.22	123.72	120.35
5	A	504[A]	COA	O3A-P2A-O4A	-2.12	99.45	111.19
5	C	503[A]	COA	O2A-P1A-O1A	2.10	122.62	112.24
5	C	503[A]	COA	C5A-C6A-N6A	2.10	123.54	120.35
5	A	504[A]	COA	O3B-P3B-O7A	2.09	117.47	109.39
5	C	503[A]	COA	O3A-P2A-O4A	-2.03	99.95	111.19

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	502	ADP	PA-O3A-PB-O3B
3	D	502	ADP	C5'-O5'-PA-O1A
3	D	502	ADP	C5'-O5'-PA-O2A
3	D	502	ADP	C5'-O5'-PA-O3A
3	A	502[A]	ADP	PA-O3A-PB-O2B
3	A	502[A]	ADP	C5'-O5'-PA-O1A
3	A	502[A]	ADP	C5'-O5'-PA-O2A
3	B	502	ADP	C5'-O5'-PA-O1A
3	B	502	ADP	C5'-O5'-PA-O2A
10	D	504	PEG	C1-C2-O2-C3

*Continued on next page...*

*Continued from previous page...*

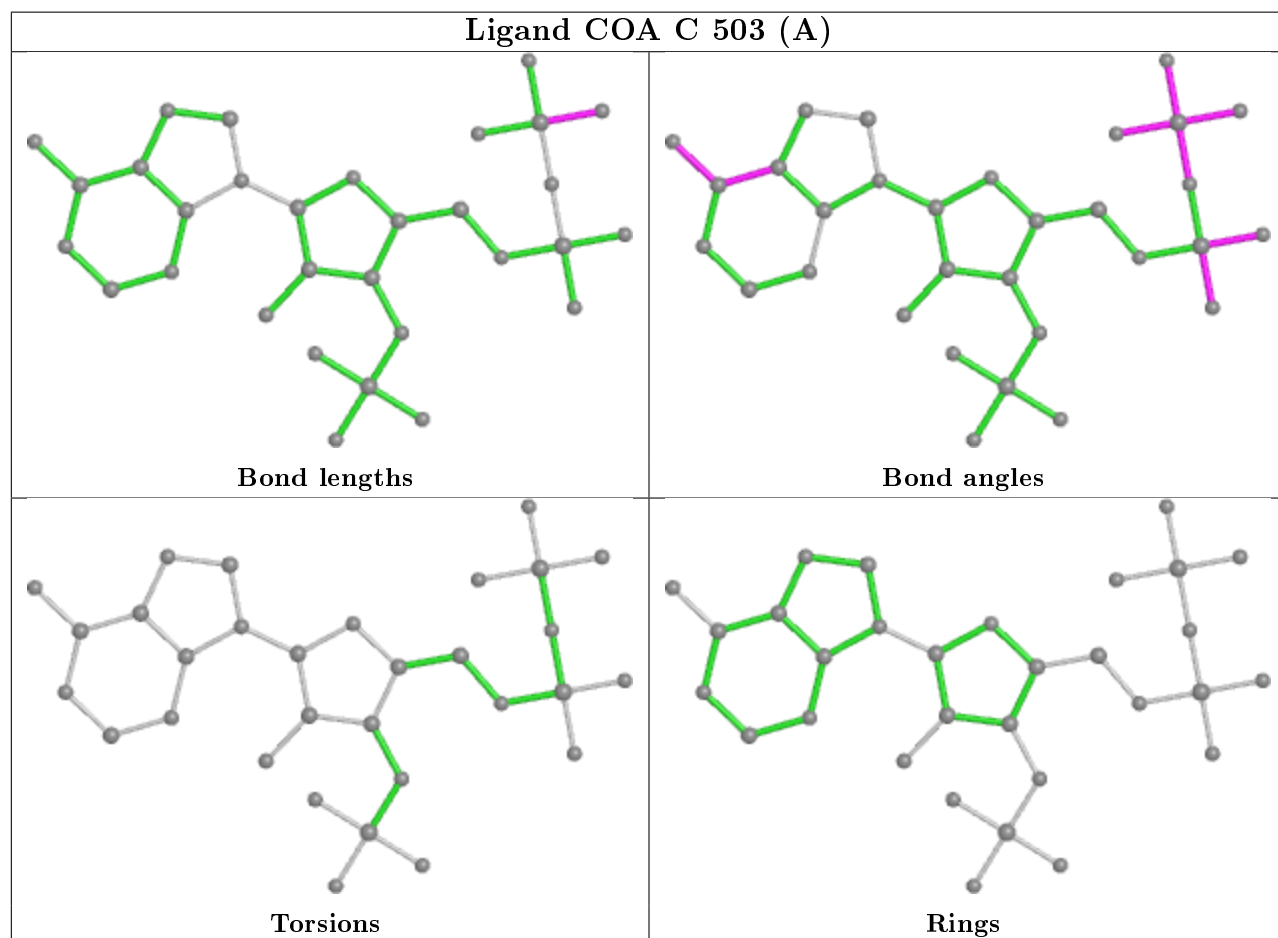
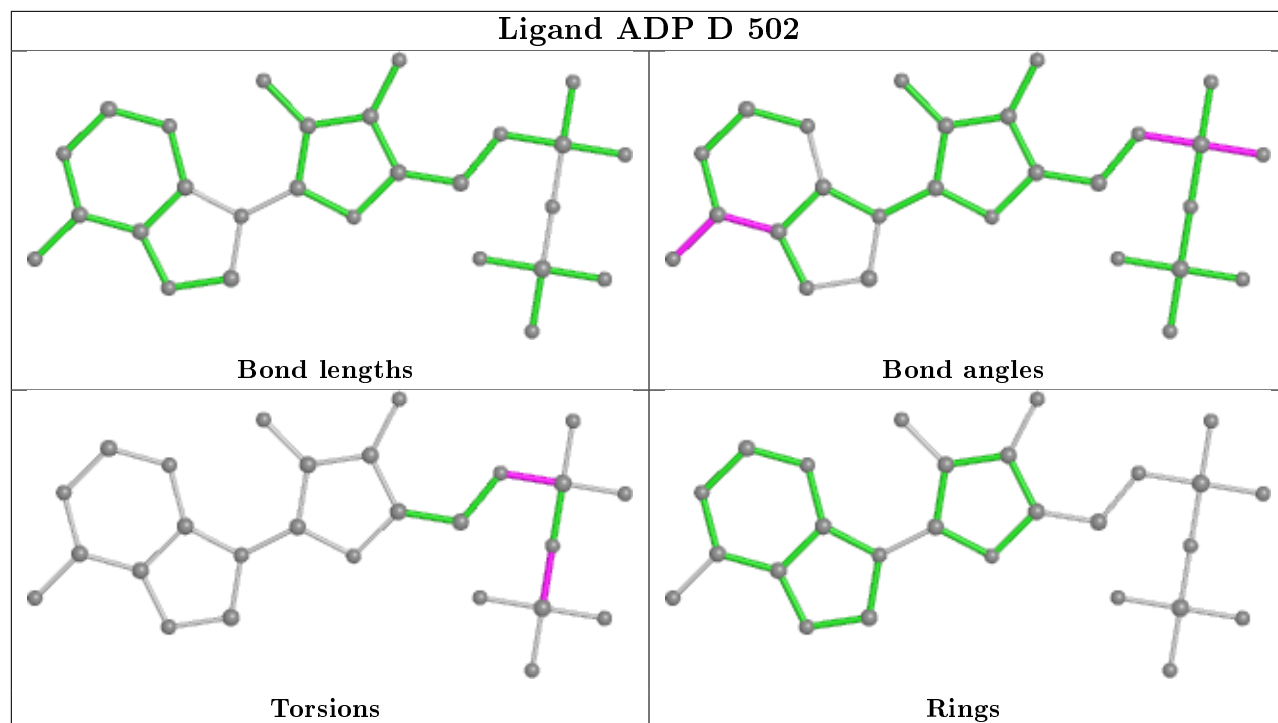
Mol	Chain	Res	Type	Atoms
7	A	508	PGE	C4-C3-O2-C2
7	A	508	PGE	C6-C5-O3-C4
3	A	502[A]	ADP	C5'-O5'-PA-O3A
7	A	508	PGE	C3-C4-O3-C5
3	D	502	ADP	PA-O3A-PB-O1B
3	A	502[A]	ADP	PA-O3A-PB-O1B
9	C	510	EDO	O1-C1-C2-O2
3	D	502	ADP	PA-O3A-PB-O2B
3	A	502[A]	ADP	PA-O3A-PB-O3B
3	B	502	ADP	PA-O3A-PB-O3B
3	B	502	ADP	C5'-O5'-PA-O3A
7	A	508	PGE	O2-C3-C4-O3

There are no ring outliers.

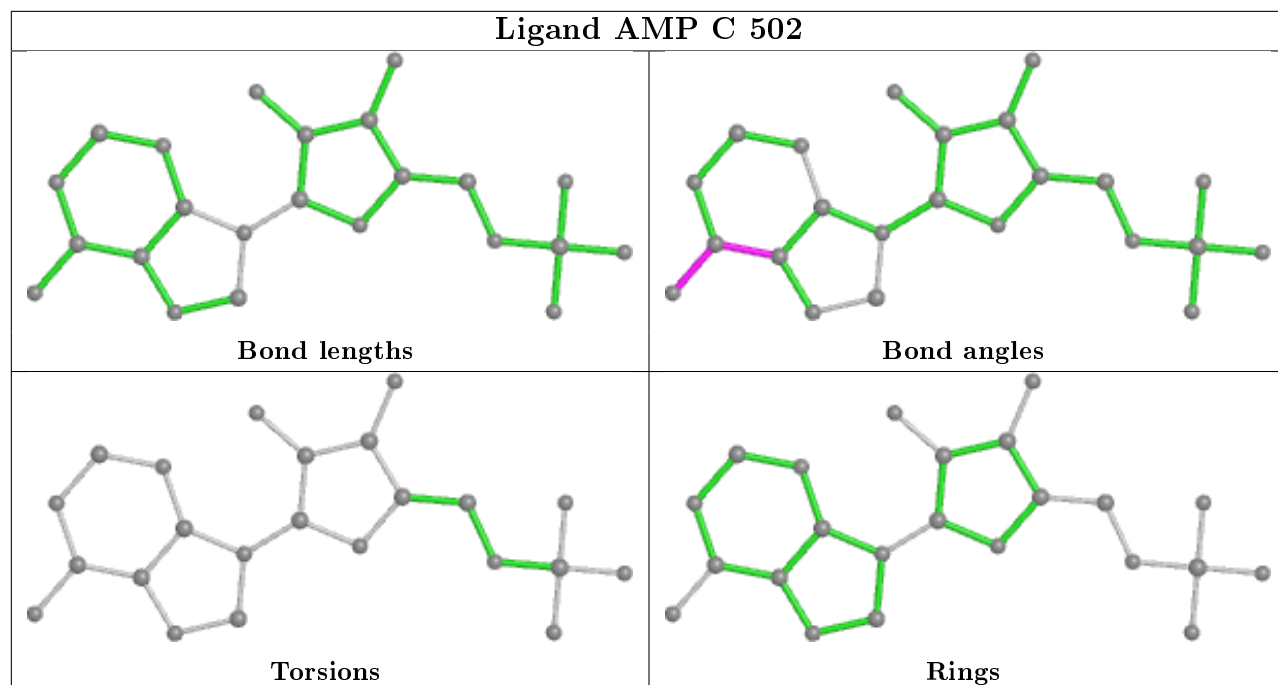
7 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	B	505	EDO	1	0
10	D	504	PEG	2	0
9	C	510	EDO	3	0
9	A	514	EDO	1	0
5	A	504[A]	COA	1	0
7	A	508	PGE	1	0
9	A	512	EDO	1	0

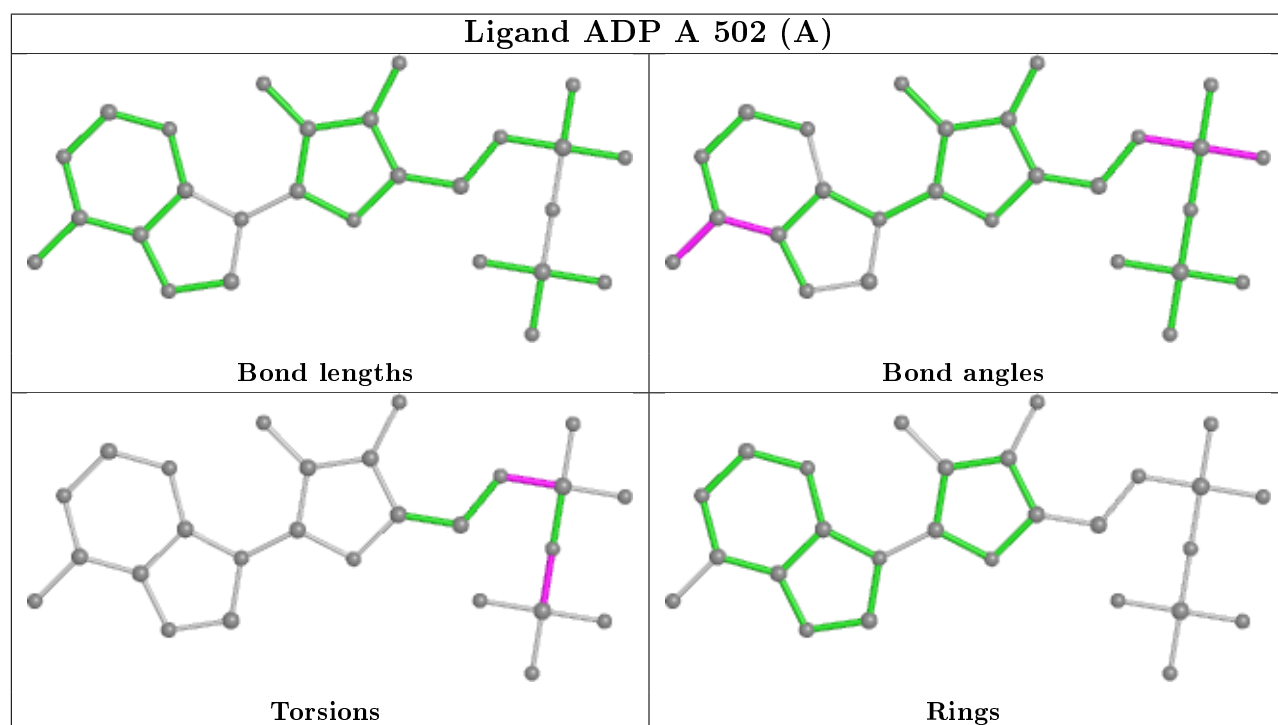
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

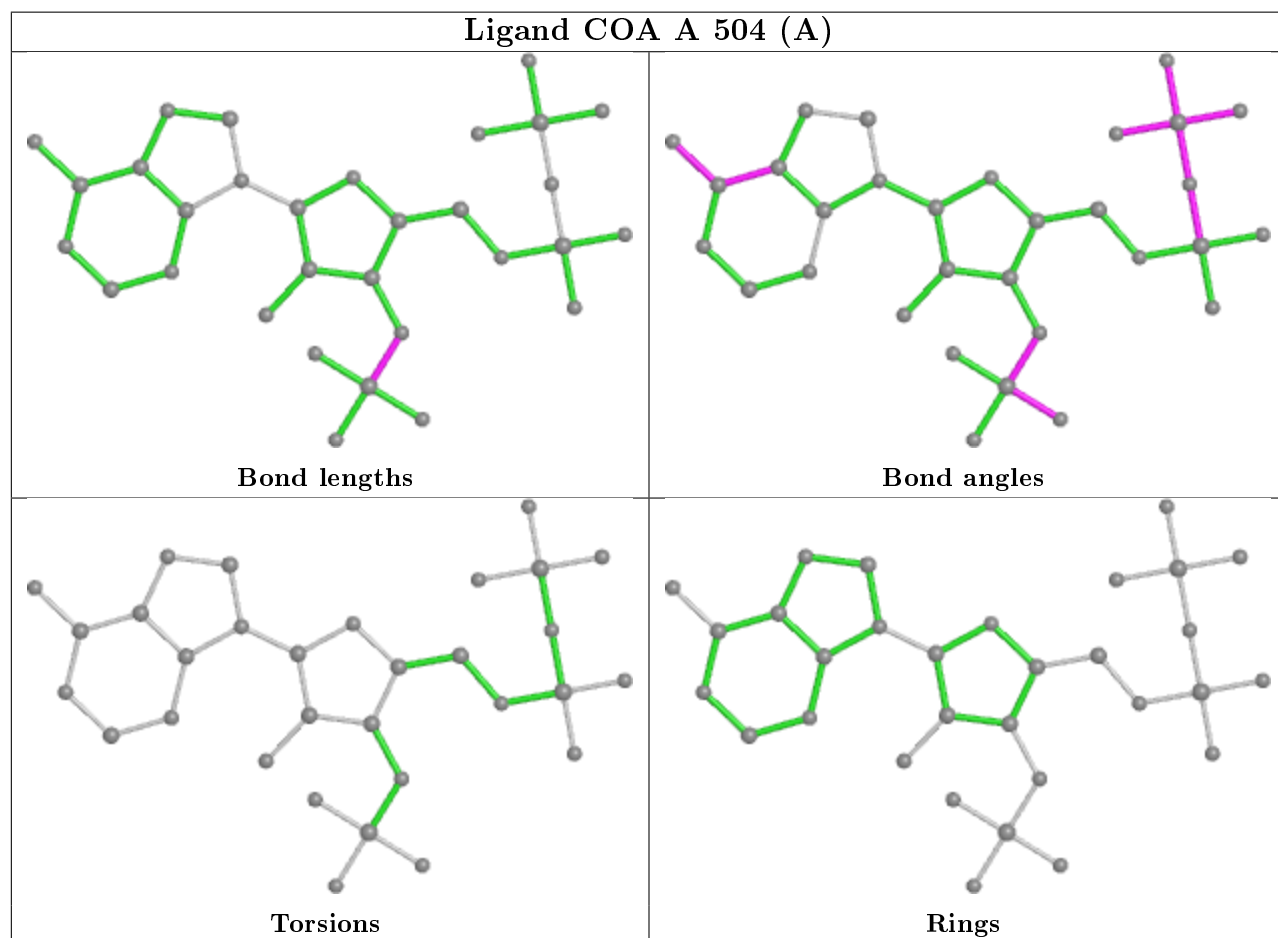
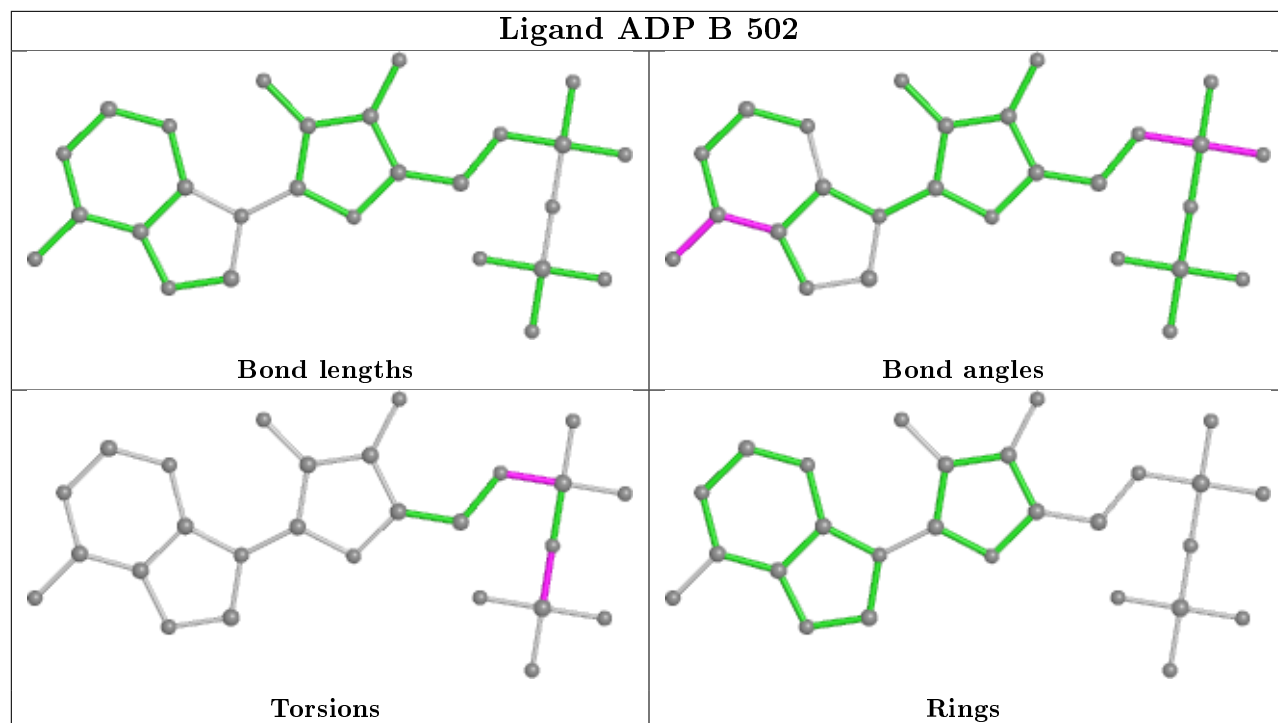


## Ligand AMP C 502

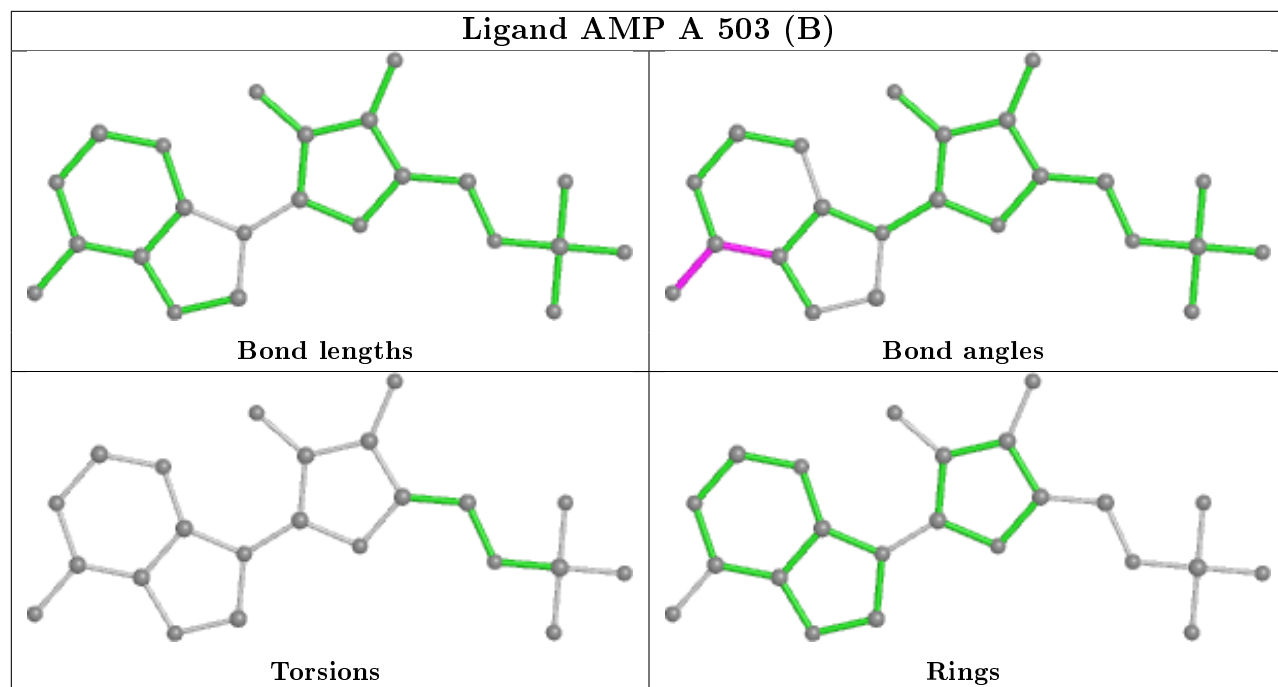


## Ligand ADP A 502 (A)









## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	418/436 (95%)	-0.28	9 (2%) 62 59	23, 45, 81, 124	0
1	B	408/436 (93%)	0.24	46 (11%) 5 4	28, 54, 107, 128	0
1	C	419/436 (96%)	-0.14	12 (2%) 51 49	29, 56, 96, 119	0
1	D	411/436 (94%)	-0.13	12 (2%) 51 49	27, 48, 94, 122	0
All	All	1656/1744 (94%)	-0.08	79 (4%) 30 28	23, 51, 97, 128	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	400	LEU	6.7
1	C	382	SER	6.5
1	B	410	VAL	4.9
1	B	405	LEU	4.9
1	B	396	ILE	4.8
1	D	338	VAL	4.6
1	B	378	LEU	4.6
1	B	333	PHE	4.5
1	A	421	SER	4.4
1	B	389	LEU	4.4
1	B	404	ILE	4.2
1	B	340	ILE	4.1
1	B	370	ASP	3.8
1	B	411	LYS	3.7
1	C	0	GLY	3.7
1	C	53	ILE	3.6
1	B	380	ASP	3.5
1	B	371	GLU	3.5
1	D	363	LEU	3.4
1	D	373	ILE	3.4
1	D	405	LEU	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	386	TYR	3.3
1	A	419	PRO	3.3
1	B	408	PRO	3.2
1	A	426	VAL	3.2
1	D	334	ILE	3.2
1	B	407	THR	3.1
1	B	363	LEU	3.1
1	D	431	LEU	3.1
1	A	368	ASN	3.1
1	B	376	VAL	3.1
1	D	368	ASN	3.0
1	B	402	ASP	3.0
1	B	430	ASP	3.0
1	B	406	VAL	3.0
1	C	271	PRO	3.0
1	B	392	ILE	2.9
1	B	426	VAL	2.9
1	D	335	ILE	2.8
1	B	337	GLY	2.8
1	B	382	SER	2.8
1	B	417	SER	2.8
1	C	433	ASP	2.8
1	B	364	GLU	2.7
1	B	381	LEU	2.7
1	B	384	ASP	2.7
1	A	355	LEU	2.7
1	D	382	SER	2.7
1	B	334	ILE	2.6
1	B	348	ILE	2.6
1	B	398	ARG	2.6
1	B	365	THR	2.6
1	C	55	SER	2.6
1	B	335	ILE	2.5
1	C	383	THR	2.5
1	C	200	ILE	2.5
1	B	383	THR	2.4
1	A	384	ASP	2.4
1	C	311	GLY	2.4
1	B	429	LYS	2.4
1	B	393	ARG	2.4
1	B	351[A]	GLN	2.4
1	B	413	VAL	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	414	LYS	2.3
1	B	419	PRO	2.3
1	D	337	GLY	2.2
1	D	397	ILE	2.2
1	B	431	LEU	2.2
1	B	428	VAL	2.2
1	C	387	ILE	2.2
1	B	412	LEU	2.2
1	D	369	GLN	2.2
1	C	310	PRO	2.2
1	A	382	SER	2.2
1	A	376	VAL	2.1
1	B	397	ILE	2.1
1	B	385	ASN	2.1
1	C	426	VAL	2.0
1	B	338	VAL	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
9	EDO	B	506	4/4	0.70	0.25	59,62,63,64	0
10	PEG	D	504	7/7	0.79	0.34	63,65,73,74	0
9	EDO	B	504	4/4	0.81	0.18	74,74,74,76	0
9	EDO	D	506	4/4	0.82	0.18	76,78,78,78	0
9	EDO	C	508	4/4	0.85	0.20	68,68,68,68	0
8	SO4	A	511	5/5	0.88	0.17	53,53,56,58	5
9	EDO	C	510	4/4	0.88	0.21	58,59,59,60	0

*Continued on next page...*

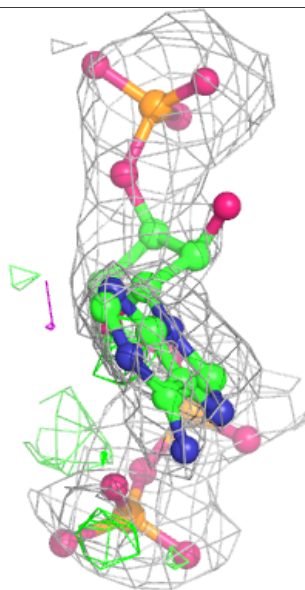
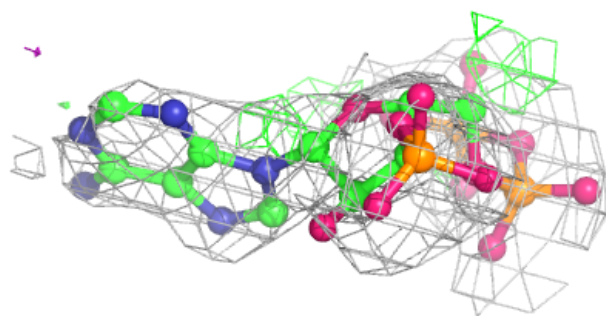
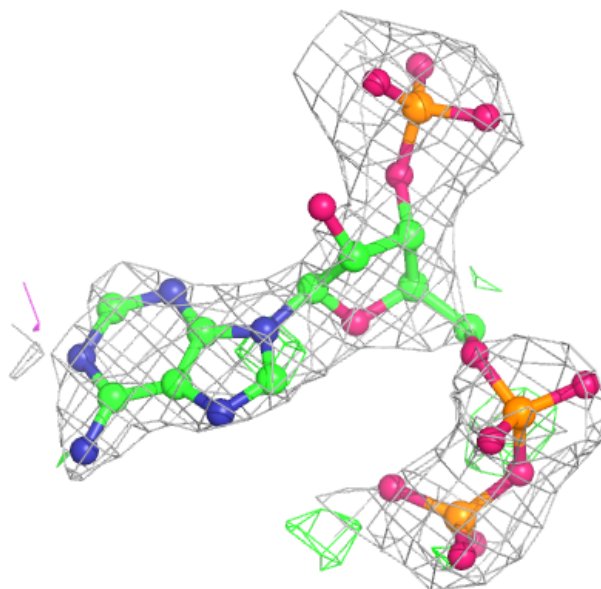
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
9	EDO	C	509	4/4	0.89	0.27	63,65,66,68	0
9	EDO	B	505	4/4	0.89	0.29	55,55,58,60	0
7	PGE	A	508	10/10	0.90	0.36	64,75,78,78	0
8	SO4	C	507	5/5	0.90	0.19	49,50,51,52	5
9	EDO	A	513	4/4	0.90	0.26	69,70,71,74	0
5	COA	A	504[A]	31/48	0.91	0.25	37,45,59,61	31
9	EDO	D	507	4/4	0.91	0.15	50,51,51,52	0
9	EDO	A	512	4/4	0.91	0.35	59,59,59,60	0
4	AMP	C	502	23/23	0.93	0.14	37,45,53,61	0
5	COA	C	503[A]	31/48	0.93	0.29	23,35,42,43	31
8	SO4	B	503	5/5	0.95	0.15	43,43,44,45	5
9	EDO	A	514	4/4	0.95	0.23	61,61,62,63	0
3	ADP	B	502	27/27	0.95	0.14	41,54,85,86	0
3	ADP	D	502	27/27	0.95	0.14	31,43,80,83	0
9	EDO	A	515	4/4	0.96	0.17	31,36,38,41	0
2	ZN	B	501	1/1	0.96	0.09	48,48,48,48	0
2	ZN	C	501	1/1	0.97	0.07	65,65,65,65	0
8	SO4	A	509[B]	5/5	0.97	0.09	102,102,102,102	5
6	K	C	505	1/1	0.97	0.07	42,42,42,42	0
3	ADP	A	502[A]	27/27	0.97	0.15	38,52,55,55	27
6	K	A	507	1/1	0.97	0.14	43,43,43,43	0
8	SO4	D	505	5/5	0.97	0.14	78,79,80,81	0
4	AMP	A	503[B]	23/23	0.98	0.13	13,21,24,25	23
8	SO4	C	506[B]	5/5	0.98	0.14	126,126,126,126	5
6	K	D	503	1/1	0.98	0.05	37,37,37,37	0
8	SO4	A	510	5/5	0.99	0.12	53,54,58,60	0
6	K	A	505	1/1	0.99	0.06	40,40,40,40	0
2	ZN	A	501	1/1	0.99	0.10	47,47,47,47	0
6	K	A	506	1/1	0.99	0.04	32,32,32,32	0
2	ZN	D	501	1/1	0.99	0.06	47,47,47,47	0
6	K	C	504	1/1	0.99	0.07	45,45,45,45	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.

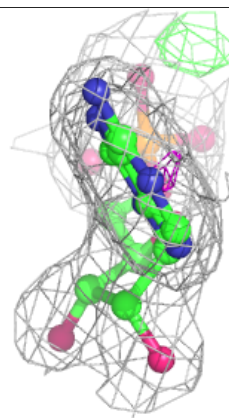
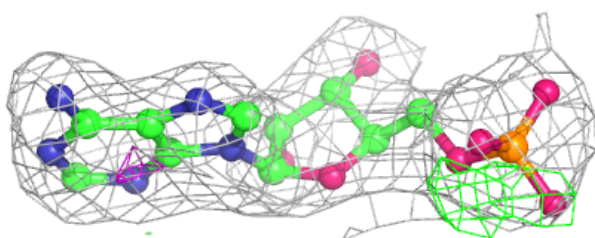
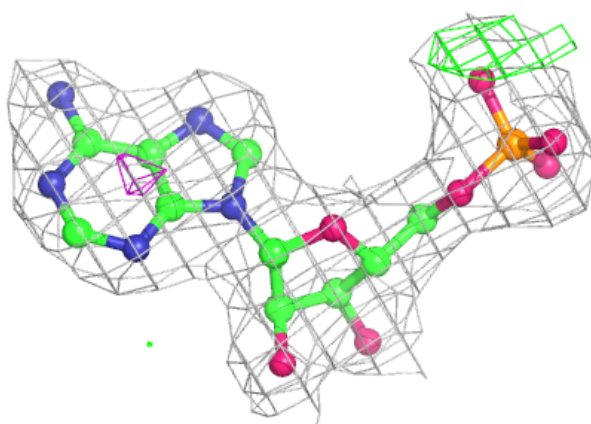
**Electron density around COA A 504 (A):**

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



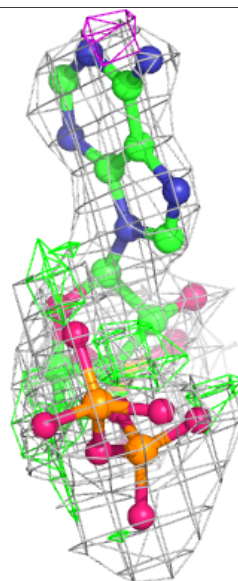
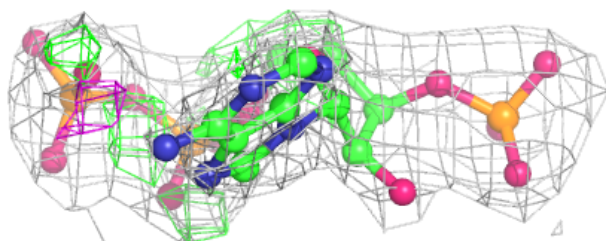
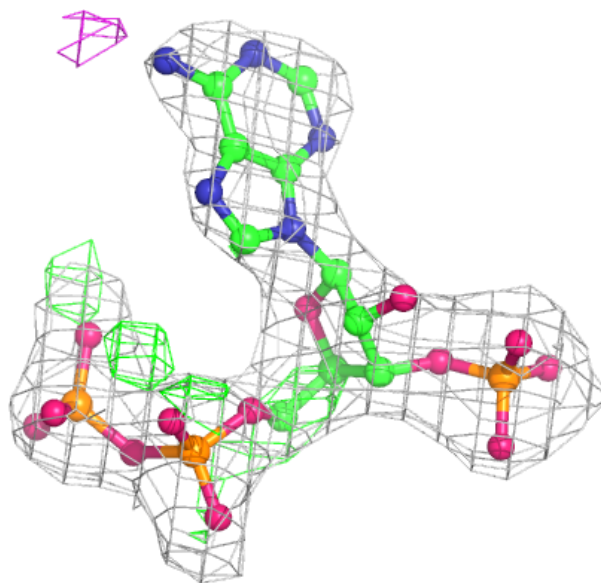
**Electron density around AMP C 502:**

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



**Electron density around COA C 503 (A):**

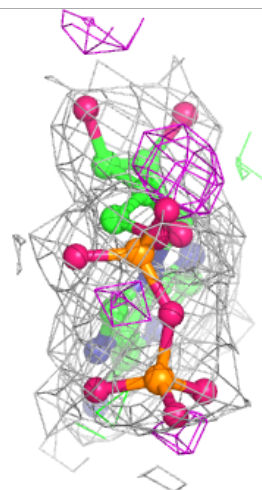
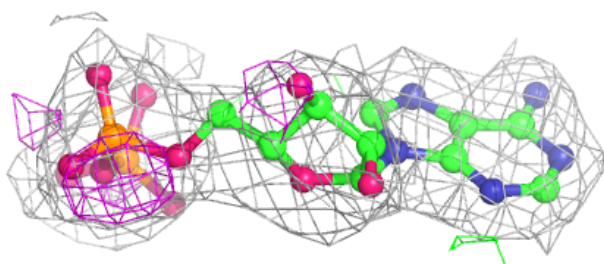
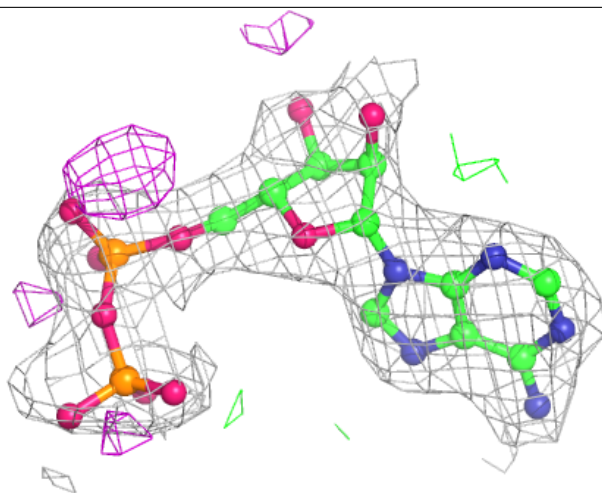
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





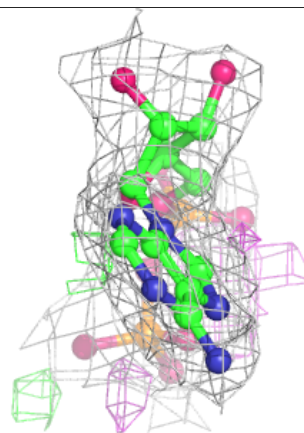
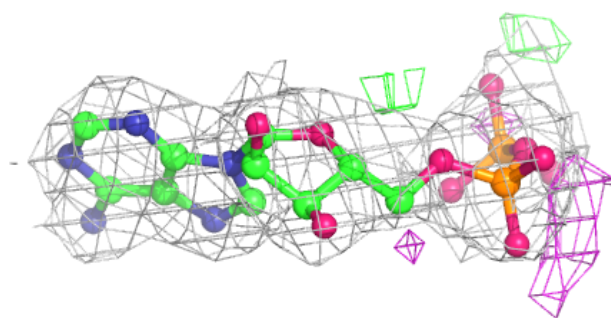
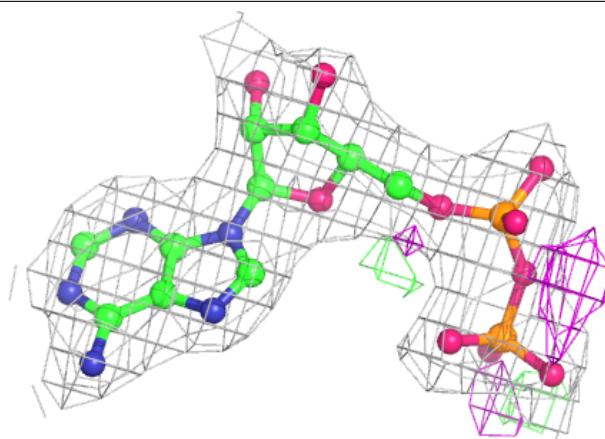
**Electron density around ADP B 502:**

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



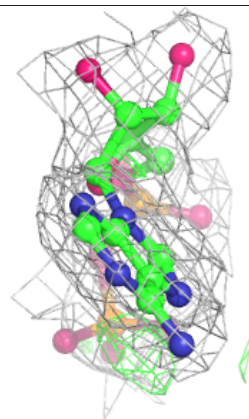
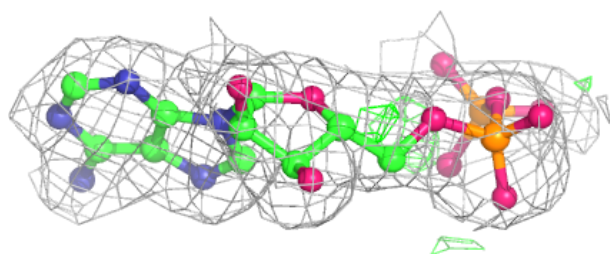
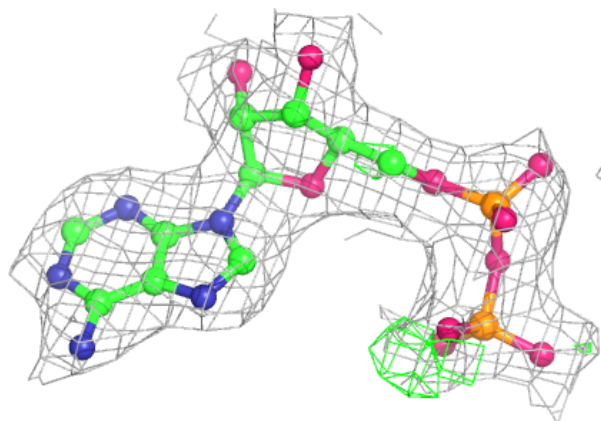
**Electron density around ADP D 502:**

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

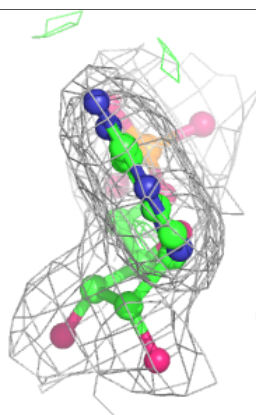
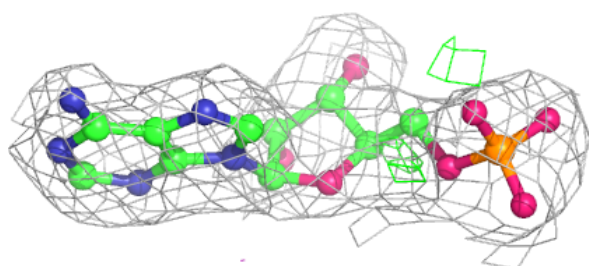
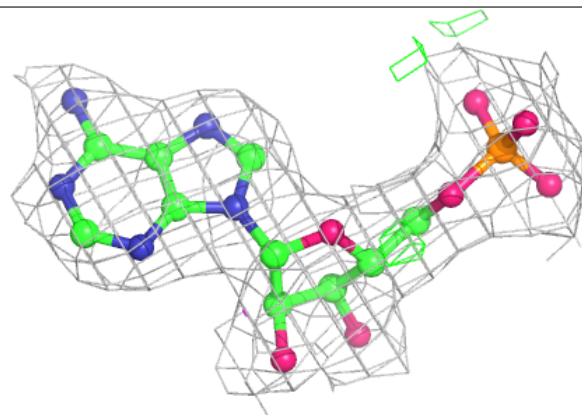


**Electron density around ADP A 502 (A):**

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

**Electron density around AMP A 503 (B):**

$2mF_o-DF_c$  (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.