



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 25, 2021 – 10:11 AM EDT

PDB ID : 7LCL  
Title : Zoogloea ramigera biosynthetic thiolase Q183Y mutant  
Authors : Marshall, A.C.; Bruning, J.B.  
Deposited on : 2021-01-11  
Resolution : 2.29 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 : 2.23.1  
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) : 2.23.1

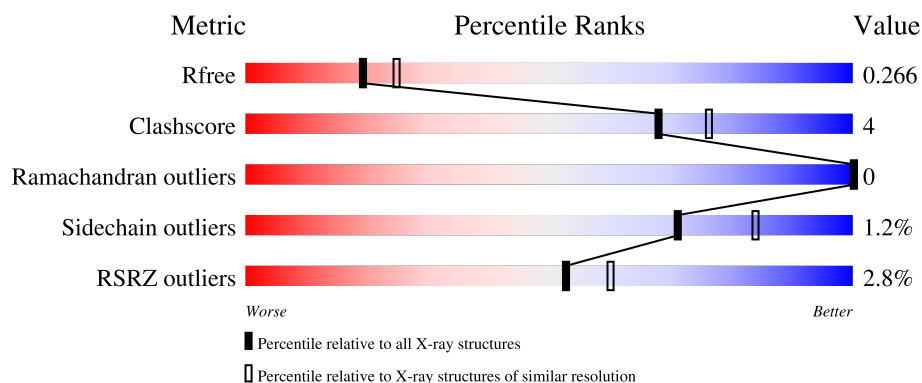
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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  $\geq 3$ , 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	399	<div> <div>8%</div> <div>87%</div> <div>10%</div> <div>.</div> </div>
1	B	399	<div> <div>89%</div> <div>9%</div> <div>.</div> </div>
1	C	399	<div> <div>3%</div> <div>87%</div> <div>11%</div> <div>.</div> </div>
1	D	399	<div> <div>8%</div> <div>88%</div> <div>9%</div> <div>.</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NH4	B	407	-	-	-	X
4	NH4	D	402	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12456 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 Acetyl-CoA acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	0	0	0
			2800	1743	501	535	21			
1	B	390	Total	C	N	O	S	0	1	0
			2816	1752	505	538	21			
1	C	390	Total	C	N	O	S	0	0	0
			2791	1736	497	537	21			
1	D	390	Total	C	N	O	S	0	0	0
			2778	1728	497	532	21			

There are 40 discrepancies between the modelled and reference sequences:

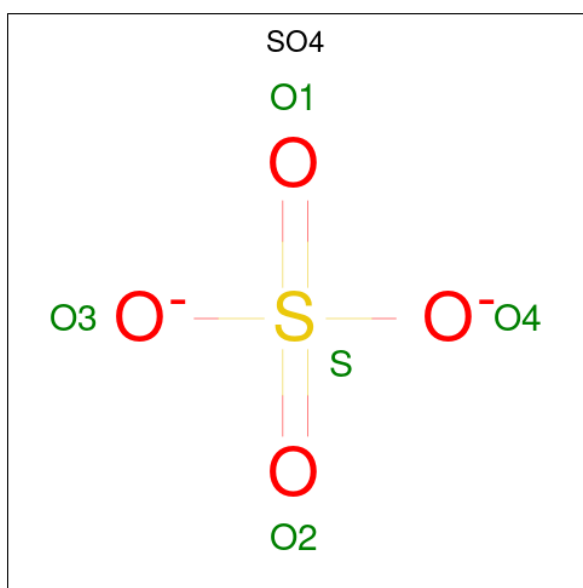
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP P07097
A	-5	HIS	-	expression tag	UNP P07097
A	-4	HIS	-	expression tag	UNP P07097
A	-3	HIS	-	expression tag	UNP P07097
A	-2	HIS	-	expression tag	UNP P07097
A	-1	HIS	-	expression tag	UNP P07097
A	0	HIS	-	expression tag	UNP P07097
A	11	ALA	-	insertion	UNP P07097
A	129	ARG	ALA	conflict	UNP P07097
A	183	TYR	GLN	engineered mutation	UNP P07097
B	-6	MET	-	initiating methionine	UNP P07097
B	-5	HIS	-	expression tag	UNP P07097
B	-4	HIS	-	expression tag	UNP P07097
B	-3	HIS	-	expression tag	UNP P07097
B	-2	HIS	-	expression tag	UNP P07097
B	-1	HIS	-	expression tag	UNP P07097
B	0	HIS	-	expression tag	UNP P07097
B	11	ALA	-	insertion	UNP P07097
B	129	ARG	ALA	conflict	UNP P07097
B	183	TYR	GLN	engineered mutation	UNP P07097
C	-6	MET	-	initiating methionine	UNP P07097

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-5	HIS	-	expression tag	UNP P07097
C	-4	HIS	-	expression tag	UNP P07097
C	-3	HIS	-	expression tag	UNP P07097
C	-2	HIS	-	expression tag	UNP P07097
C	-1	HIS	-	expression tag	UNP P07097
C	0	HIS	-	expression tag	UNP P07097
C	11	ALA	-	insertion	UNP P07097
C	129	ARG	ALA	conflict	UNP P07097
C	183	TYR	GLN	engineered mutation	UNP P07097
D	-6	MET	-	initiating methionine	UNP P07097
D	-5	HIS	-	expression tag	UNP P07097
D	-4	HIS	-	expression tag	UNP P07097
D	-3	HIS	-	expression tag	UNP P07097
D	-2	HIS	-	expression tag	UNP P07097
D	-1	HIS	-	expression tag	UNP P07097
D	0	HIS	-	expression tag	UNP P07097
D	11	ALA	-	insertion	UNP P07097
D	129	ARG	ALA	conflict	UNP P07097
D	183	TYR	GLN	engineered mutation	UNP P07097

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



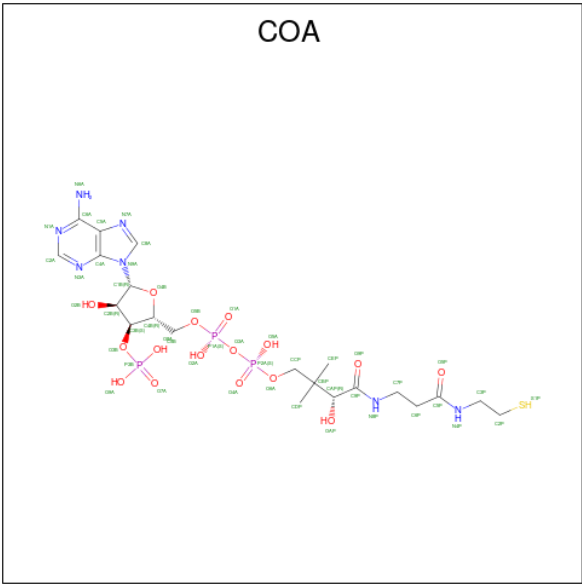
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

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

- Molecule 3 is COENZYME A (three-letter code: COA) (formula: C<sub>21</sub>H<sub>36</sub>N<sub>7</sub>O<sub>16</sub>P<sub>3</sub>S) (labeled as "Ligand of Interest" by depositor).



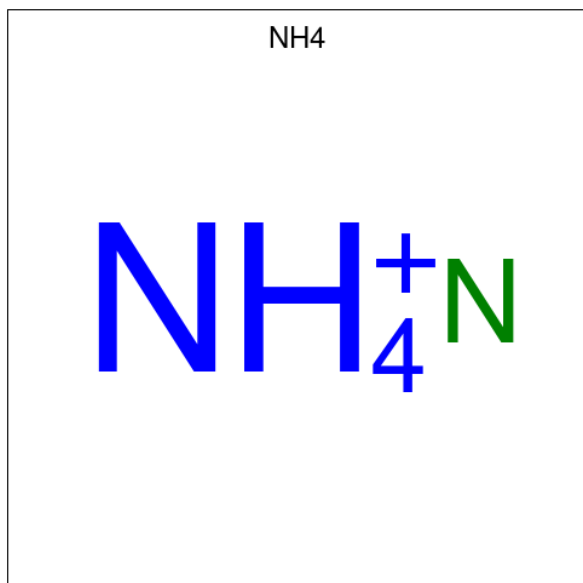
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		
3	B	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	C	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		
3	D	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		

- Molecule 4 is AMMONIUM ION (three-letter code: NH4) (formula: H<sub>4</sub>N) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	N	0	0
			1	1		
4	B	1	Total	N	0	0
			1	1		
4	C	1	Total	N	0	0
			1	1		
4	D	1	Total	N	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	350	Total	O	0	0
			350	350		
5	B	348	Total	O	0	0
			348	348		
5	C	187	Total	O	0	0
			187	187		

*Continued on next page...*

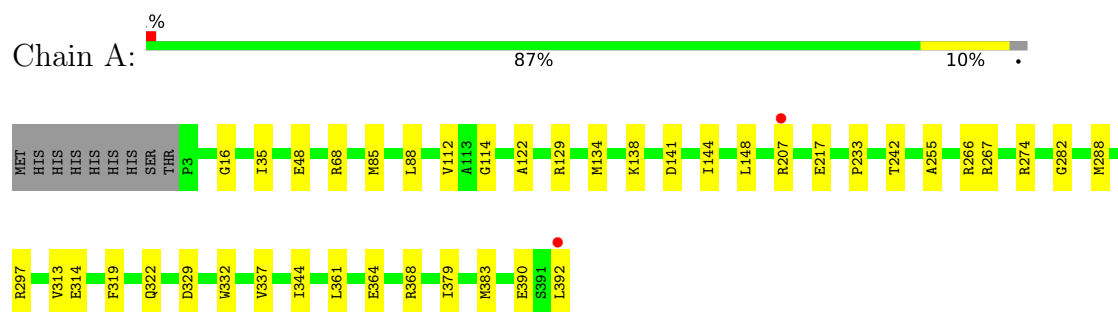
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	135	Total 135	O 135	0	0

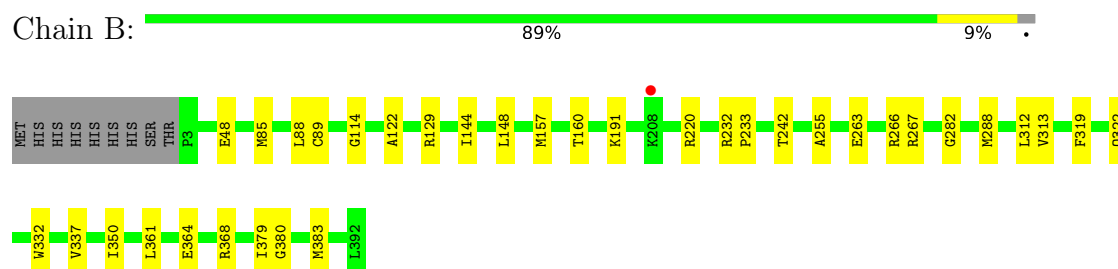
### 3 Residue-property plots [i](#)

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.

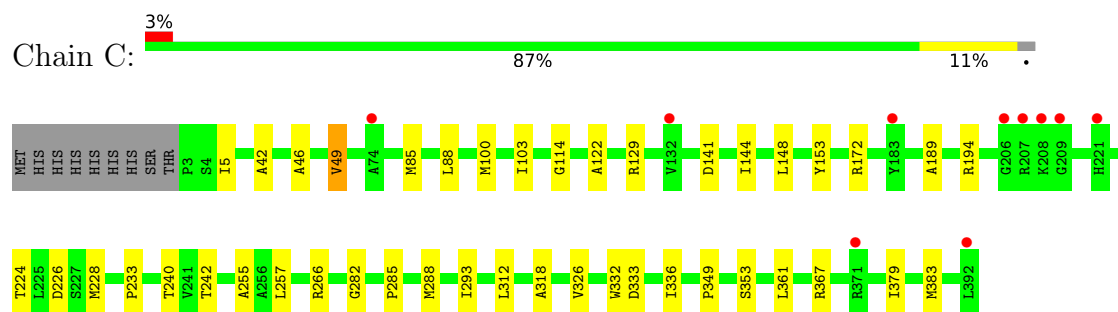
- Molecule 1: Acetyl-CoA acetyltransferase



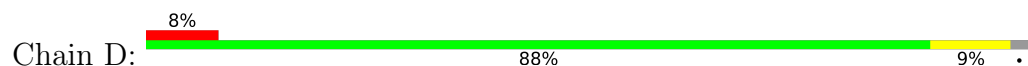
- Molecule 1: Acetyl-CoA acetyltransferase

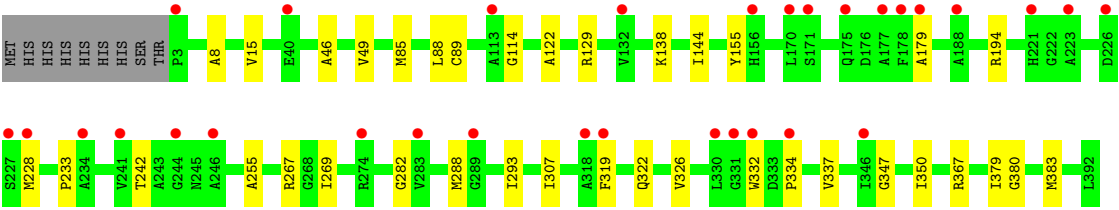


- Molecule 1: Acetyl-CoA acetyltransferase



- Molecule 1: Acetyl-CoA acetyltransferase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.17Å 79.07Å 149.63Å 90.00° 93.78° 90.00°	Depositor
Resolution (Å)	44.62 – 2.29 46.62 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (44.62-2.29) 99.8 (46.62-2.30)	Depositor EDS
$R_{merge}$	0.60	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.17 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.15_3459	Depositor
R, $R_{free}$	0.229 , 0.265 0.230 , 0.266	Depositor DCC
$R_{free}$ test set	4330 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.6	Xtriage
Anisotropy	0.557	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 38.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12456	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.88% 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: SO4, NH4, SCY, COA

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	0.24	0/2833	0.43	0/3829
1	B	0.24	0/2849	0.43	0/3849
1	C	0.24	0/2824	0.42	0/3821
1	D	0.25	0/2809	0.42	0/3800
All	All	0.24	0/11315	0.42	0/15299

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	2800	0	2782	27	0
1	B	2816	0	2800	21	0
1	C	2791	0	2748	25	0
1	D	2778	0	2735	21	0
2	A	25	0	0	0	0
2	B	25	0	0	0	0
2	C	5	0	0	0	0
3	A	48	0	32	0	0
3	B	48	0	32	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	48	0	32	2	0
3	D	48	0	32	0	0
4	A	1	0	0	1	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	350	0	0	6	0
5	B	348	0	0	4	0
5	C	187	0	0	3	0
5	D	135	0	0	2	0
All	All	12456	0	11193	86	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 4.

All (86) 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:C:224:THR:HG22	1:C:226:ASP:H	1.65	0.61
1:C:312:LEU:HD23	1:C:361:LEU:HD12	1.85	0.59
1:C:194:ARG:NH2	5:C:509:HOH:O	2.35	0.58
3:C:401:COA:H8A	3:C:401:COA:H52A	1.85	0.58
1:D:282:GLY:HA2	1:D:383:MET:HA	1.87	0.56
1:B:232:ARG:NH1	5:B:514:HOH:O	2.39	0.56
1:C:129:ARG:HH21	1:D:122:ALA:HB3	1.70	0.55
1:A:144:ILE:HD13	1:A:148:LEU:HD12	1.89	0.54
1:A:297:ARG:NH1	5:A:516:HOH:O	2.40	0.54
1:C:144:ILE:HD13	1:C:148:LEU:HD12	1.88	0.53
1:A:267:ARG:NH1	5:A:517:HOH:O	2.41	0.53
1:A:274:ARG:NH2	1:A:390:GLU:OE1	2.40	0.53
1:C:293:ILE:HG13	1:C:326:VAL:HG12	1.91	0.53
1:D:267:ARG:NH1	5:D:510:HOH:O	2.37	0.53
1:C:114:GLY:HA3	1:C:255:ALA:HA	1.90	0.53
1:B:144:ILE:HD13	1:B:148:LEU:HD12	1.92	0.52
1:D:179:ALA:HB3	1:D:228:MET:HE1	1.91	0.52
1:A:297:ARG:NH2	1:A:329:ASP:OD1	2.36	0.51
1:C:282:GLY:HA2	1:C:383:MET:HA	1.92	0.50
1:C:85:MET:HA	1:D:85:MET:HA	1.93	0.50
1:D:334:PRO:HA	1:D:337:VAL:HG12	1.93	0.50
1:C:122:ALA:HB3	1:D:129:ARG:HH21	1.77	0.49
1:A:138:LYS:NZ	5:A:522:HOH:O	2.45	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:114:GLY:HA3	1:D:255:ALA:HA	1.94	0.49
1:A:274:ARG:HB2	1:A:392:LEU:HD11	1.95	0.49
1:D:8:ALA:HB1	1:D:269:ILE:HG21	1.93	0.49
1:A:207:ARG:N	5:A:520:HOH:O	2.43	0.49
1:C:141:ASP:OD1	1:D:129:ARG:NH2	2.40	0.49
1:A:122:ALA:HB3	1:B:129:ARG:HH21	1.78	0.49
1:A:129:ARG:HH21	1:B:122:ALA:HB3	1.78	0.49
1:A:85:MET:HA	1:B:85:MET:HA	1.94	0.48
1:A:319:PHE:H	1:A:322:GLN:HE21	1.62	0.48
1:C:367:ARG:HD3	5:C:529:HOH:O	2.13	0.48
1:D:293:ILE:HG13	1:D:326:VAL:HG12	1.96	0.48
1:B:89:SCY:OCD	1:B:380:GLY:N	2.28	0.48
1:D:46:ALA:O	1:D:49:VAL:HG22	2.13	0.48
1:A:68:ARG:NH1	5:A:521:HOH:O	2.45	0.47
1:A:141:ASP:OD1	1:B:129:ARG:NH2	2.38	0.47
1:B:191:LYS:NZ	5:B:532:HOH:O	2.47	0.47
1:B:312:LEU:HD23	1:B:361:LEU:HD12	1.96	0.47
1:A:88:LEU:HB2	1:A:379:ILE:HG23	1.96	0.47
1:B:88:LEU:HB2	1:B:379:ILE:HG23	1.97	0.47
1:B:364:GLU:O	1:B:368:ARG:HG2	2.15	0.47
1:B:263:GLU:OE2	1:B:266:ARG:NH1	2.45	0.47
1:A:313:VAL:HB	1:A:337:VAL:HG22	1.96	0.47
1:D:319:PHE:H	1:D:322:GLN:HE21	1.63	0.47
1:B:114:GLY:HA3	1:B:255:ALA:HA	1.96	0.46
1:D:88:LEU:HB2	1:D:379:ILE:HG23	1.98	0.46
1:A:233:PRO:HA	1:A:242:THR:HG22	1.97	0.46
1:D:15:VAL:HG11	1:D:347:GLY:HA3	1.96	0.46
1:B:266:ARG:NH1	5:B:534:HOH:O	2.48	0.46
1:A:16:GLY:HA2	1:A:217:GLU:HG2	1.98	0.45
1:D:194:ARG:HD3	1:D:367:ARG:CZ	2.46	0.45
1:A:114:GLY:HA3	1:A:255:ALA:HA	1.98	0.45
1:C:318:ALA:HB1	3:C:401:COA:H22	1.98	0.45
1:D:89:SCY:OCD	1:D:380:GLY:N	2.33	0.45
1:A:344:ILE:O	4:A:407:NH4:N	2.50	0.45
1:B:48:GLU:OE1	1:B:267:ARG:NH1	2.43	0.45
1:A:282:GLY:HA2	1:A:383:MET:HA	1.98	0.45
1:B:220:ARG:NH1	5:B:504:HOH:O	2.28	0.45
1:D:233:PRO:HA	1:D:242:THR:HG22	1.99	0.45
1:C:46:ALA:O	1:C:49:VAL:HG23	2.16	0.44
1:C:233:PRO:HA	1:C:242:THR:HG22	1.99	0.44
1:A:266:ARG:NH2	5:A:532:HOH:O	2.51	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:134:MET:HG2	1:D:144:ILE:HD11	1.99	0.44
1:A:314:GLU:HG3	1:A:361:LEU:HB2	2.00	0.44
1:C:189:ALA:HA	1:C:194:ARG:HE	1.83	0.43
1:B:233:PRO:HA	1:B:242:THR:HG22	2.00	0.43
1:D:138:LYS:NZ	5:D:522:HOH:O	2.51	0.43
1:A:364:GLU:O	1:A:368:ARG:HG2	2.18	0.43
1:B:157:MET:O	1:B:160:THR:OG1	2.31	0.42
1:C:172:ARG:NH2	1:C:228:MET:O	2.48	0.42
1:C:5:ILE:HG12	1:C:103:ILE:HB	2.01	0.42
1:B:282:GLY:HA2	1:B:383:MET:HA	2.00	0.42
1:C:349:PRO:O	1:C:353:SER:N	2.53	0.42
1:A:48:GLU:OE1	1:A:267:ARG:NH1	2.47	0.42
1:C:88:LEU:HB2	1:C:379:ILE:HG23	2.01	0.42
1:C:153:TYR:HE2	1:C:285:PRO:HG2	1.85	0.41
1:B:319:PHE:H	1:B:322:GLN:HE21	1.68	0.41
1:C:5:ILE:HD13	1:C:100:MET:HG2	2.03	0.41
1:B:313:VAL:HB	1:B:337:VAL:HG22	2.03	0.41
1:C:333:ASP:O	1:C:336:ILE:HG12	2.20	0.41
1:A:35:ILE:HG12	1:A:112:VAL:HG11	2.03	0.41
1:C:266:ARG:NH1	5:C:537:HOH:O	2.52	0.40
1:D:307:ILE:HD12	1:D:307:ILE:H	1.87	0.40
1:C:42:ALA:HB3	1:C:257:LEU:HD13	2.03	0.40

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	387/399 (97%)	378 (98%)	9 (2%)	0	100	100
1	B	388/399 (97%)	380 (98%)	8 (2%)	0	100	100
1	C	387/399 (97%)	378 (98%)	9 (2%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	387/399 (97%)	378 (98%)	9 (2%)	0	100	100
All	All	1549/1596 (97%)	1514 (98%)	35 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	270/285 (95%)	268 (99%)	2 (1%)	84	92
1	B	272/285 (95%)	269 (99%)	3 (1%)	73	86
1	C	267/285 (94%)	263 (98%)	4 (2%)	65	79
1	D	263/285 (92%)	259 (98%)	4 (2%)	65	79
All	All	1072/1140 (94%)	1059 (99%)	13 (1%)	71	84

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	288	MET
1	A	332	TRP
1	B	288	MET
1	B	332	TRP
1	B	350	ILE
1	C	49	VAL
1	C	240	THR
1	C	288	MET
1	C	332	TRP
1	D	155	TYR
1	D	288	MET
1	D	332	TRP
1	D	350	ILE

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

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	SCY	B	89	1	7,8,9	0.90	0	3,9,11	0.78	0
1	SCY	D	89	1	7,8,9	0.89	0	3,9,11	0.81	0
1	SCY	C	89	1	7,8,9	0.92	0	3,9,11	0.82	0
1	SCY	A	89	1	7,8,9	0.93	0	3,9,11	0.85	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
1	SCY	B	89	1	-	2/5/7/9	-
1	SCY	D	89	1	-	2/5/7/9	-
1	SCY	C	89	1	-	2/5/7/9	-
1	SCY	A	89	1	-	2/5/7/9	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	D	89	SCY	OCD-CD-SG-CB
1	C	89	SCY	OCD-CD-SG-CB

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	89	SCY	OCD-CD-SG-CB
1	B	89	SCY	OCD-CD-SG-CB
1	A	89	SCY	CE-CD-SG-CB
1	B	89	SCY	CE-CD-SG-CB
1	C	89	SCY	CE-CD-SG-CB
1	D	89	SCY	CE-CD-SG-CB

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	89	SCY	1	0
1	D	89	SCY	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 4 are modelled with single atom - leaving 15 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	$\# Z  > 2$
2	SO4	A	402	-	4,4,4	0.13	0	6,6,6	0.04	0
3	COA	B	403	-	41,50,50	1.04	2 (4%)	52,75,75	1.60	7 (13%)
2	SO4	B	404	-	4,4,4	0.15	0	6,6,6	0.04	0
2	SO4	C	402	-	4,4,4	0.14	0	6,6,6	0.05	0
3	COA	A	403	-	41,50,50	1.01	2 (4%)	52,75,75	1.45	7 (13%)
2	SO4	A	405	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	A	404	-	4,4,4	0.14	0	6,6,6	0.05	0
3	COA	D	401	-	41,50,50	0.99	2 (4%)	52,75,75	1.56	7 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	B	402	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	B	405	-	4,4,4	0.14	0	6,6,6	0.05	0
3	COA	C	401	-	41,50,50	1.02	2 (4%)	52,75,75	1.47	6 (11%)
2	SO4	B	406	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	A	406	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	A	401	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	B	401	-	4,4,4	0.14	0	6,6,6	0.05	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
3	COA	B	403	-	-	1/44/64/64	0/3/3/3
3	COA	A	403	-	-	7/44/64/64	0/3/3/3
3	COA	C	401	-	-	5/44/64/64	0/3/3/3
3	COA	D	401	-	-	15/44/64/64	0/3/3/3

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	401	COA	O4B-C1B	3.29	1.45	1.41
3	A	403	COA	O4B-C1B	2.96	1.45	1.41
3	D	401	COA	O4B-C1B	2.88	1.45	1.41
3	B	403	COA	O4B-C1B	2.88	1.45	1.41
3	A	403	COA	C8A-N7A	-2.41	1.30	1.34
3	C	401	COA	C8A-N7A	-2.34	1.30	1.34
3	D	401	COA	C8A-N7A	-2.31	1.30	1.34
3	B	403	COA	C8A-N7A	-2.30	1.30	1.34

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	401	COA	P2A-O3A-P1A	-6.72	109.77	132.83
3	C	401	COA	P2A-O3A-P1A	-6.48	110.60	132.83
3	B	403	COA	P2A-O3A-P1A	-6.46	110.66	132.83
3	A	403	COA	P2A-O3A-P1A	-4.84	116.23	132.83
3	D	401	COA	N3A-C2A-N1A	-4.03	122.38	128.68
3	A	403	COA	N3A-C2A-N1A	-4.02	122.40	128.68
3	C	401	COA	N3A-C2A-N1A	-3.99	122.44	128.68

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	403	COA	N3A-C2A-N1A	-3.96	122.49	128.68
3	A	403	COA	O4B-C1B-C2B	-2.73	102.93	106.93
3	C	401	COA	P1A-O5B-C5B	-2.61	106.35	121.68
3	A	403	COA	P2A-O6A-CCP	-2.56	106.79	121.56
3	B	403	COA	P2A-O6A-CCP	-2.56	106.79	121.56
3	B	403	COA	P1A-O5B-C5B	-2.53	106.86	121.68
3	D	401	COA	P1A-O5B-C5B	-2.45	107.31	121.68
3	B	403	COA	C3B-C2B-C1B	2.44	105.30	99.89
3	A	403	COA	P1A-O5B-C5B	-2.44	107.39	121.68
3	A	403	COA	C6P-C5P-N4P	2.38	120.43	116.42
3	C	401	COA	P2A-O6A-CCP	-2.35	108.01	121.56
3	B	403	COA	C4A-C5A-N7A	-2.31	106.99	109.40
3	D	401	COA	O4B-C1B-C2B	-2.28	103.59	106.93
3	A	403	COA	C4A-C5A-N7A	-2.27	107.03	109.40
3	D	401	COA	P2A-O6A-CCP	-2.27	108.47	121.56
3	B	403	COA	C6P-C5P-N4P	2.20	120.12	116.42
3	C	401	COA	C4A-C5A-N7A	-2.17	107.13	109.40
3	D	401	COA	C6P-C7P-N8P	-2.13	107.59	111.90
3	D	401	COA	C4A-C5A-N7A	-2.10	107.21	109.40
3	C	401	COA	C6P-C5P-N4P	2.02	119.83	116.42

There are no chirality outliers.

All (28) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	403	COA	O4B-C4B-C5B-O5B
3	A	403	COA	CCP-O6A-P2A-O4A
3	A	403	COA	CCP-O6A-P2A-O5A
3	B	403	COA	C4B-C3B-O3B-P3B
3	C	401	COA	O4B-C4B-C5B-O5B
3	C	401	COA	CAP-CBP-CCP-O6A
3	D	401	COA	C3B-O3B-P3B-O7A
3	D	401	COA	C3B-C4B-C5B-O5B
3	D	401	COA	O4B-C4B-C5B-O5B
3	D	401	COA	C5B-O5B-P1A-O3A
3	D	401	COA	CCP-O6A-P2A-O4A
3	D	401	COA	CDP-CBP-CCP-O6A
3	D	401	COA	CEP-CBP-CCP-O6A
3	D	401	COA	CAP-CBP-CCP-O6A
3	D	401	COA	S1P-C2P-C3P-N4P
3	A	403	COA	C3B-C4B-C5B-O5B
3	D	401	COA	C4B-C3B-O3B-P3B

*Continued on next page...*

*Continued from previous page...*

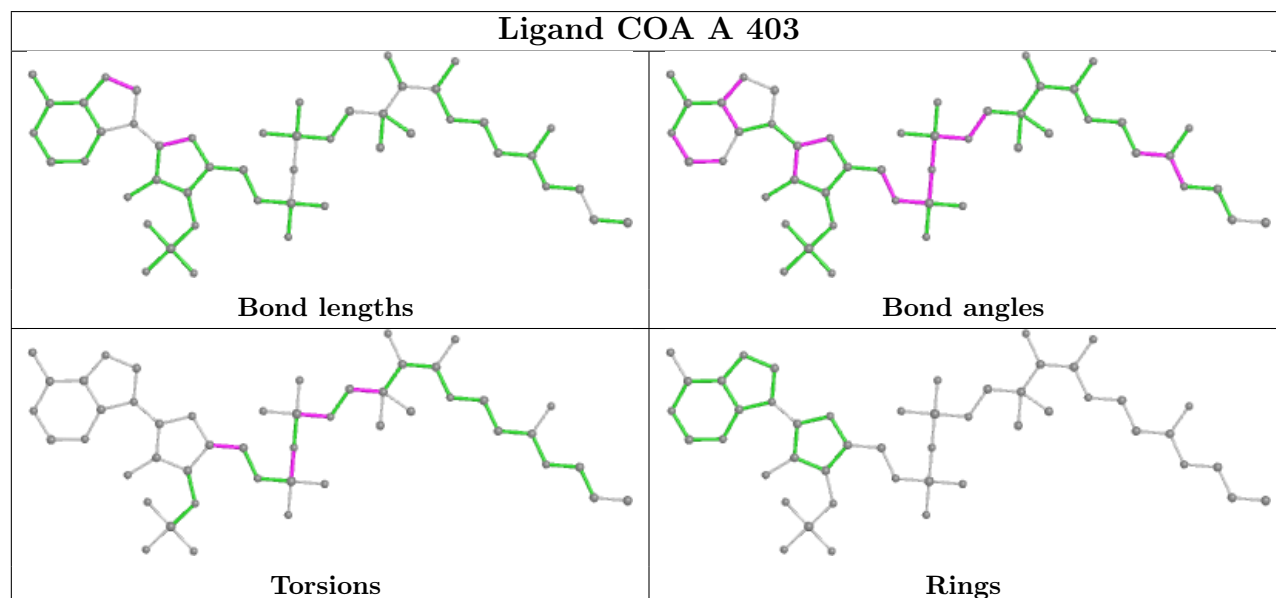
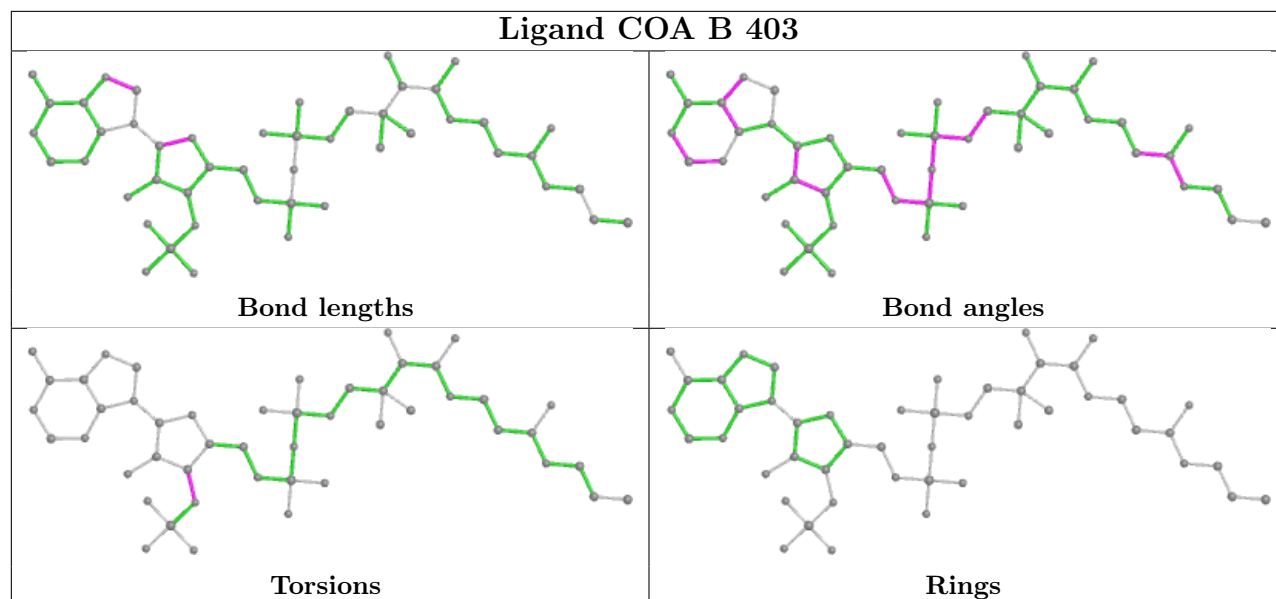
Mol	Chain	Res	Type	Atoms
3	D	401	COA	C2B-C3B-O3B-P3B
3	C	401	COA	CDP-CBP-CCP-O6A
3	C	401	COA	CEP-CBP-CCP-O6A
3	D	401	COA	CCP-O6A-P2A-O3A
3	D	401	COA	C5B-O5B-P1A-O1A
3	D	401	COA	CCP-O6A-P2A-O5A
3	A	403	COA	CDP-CBP-CCP-O6A
3	A	403	COA	CCP-O6A-P2A-O3A
3	A	403	COA	P2A-O3A-P1A-O2A
3	C	401	COA	CCP-O6A-P2A-O4A
3	D	401	COA	C5B-O5B-P1A-O2A

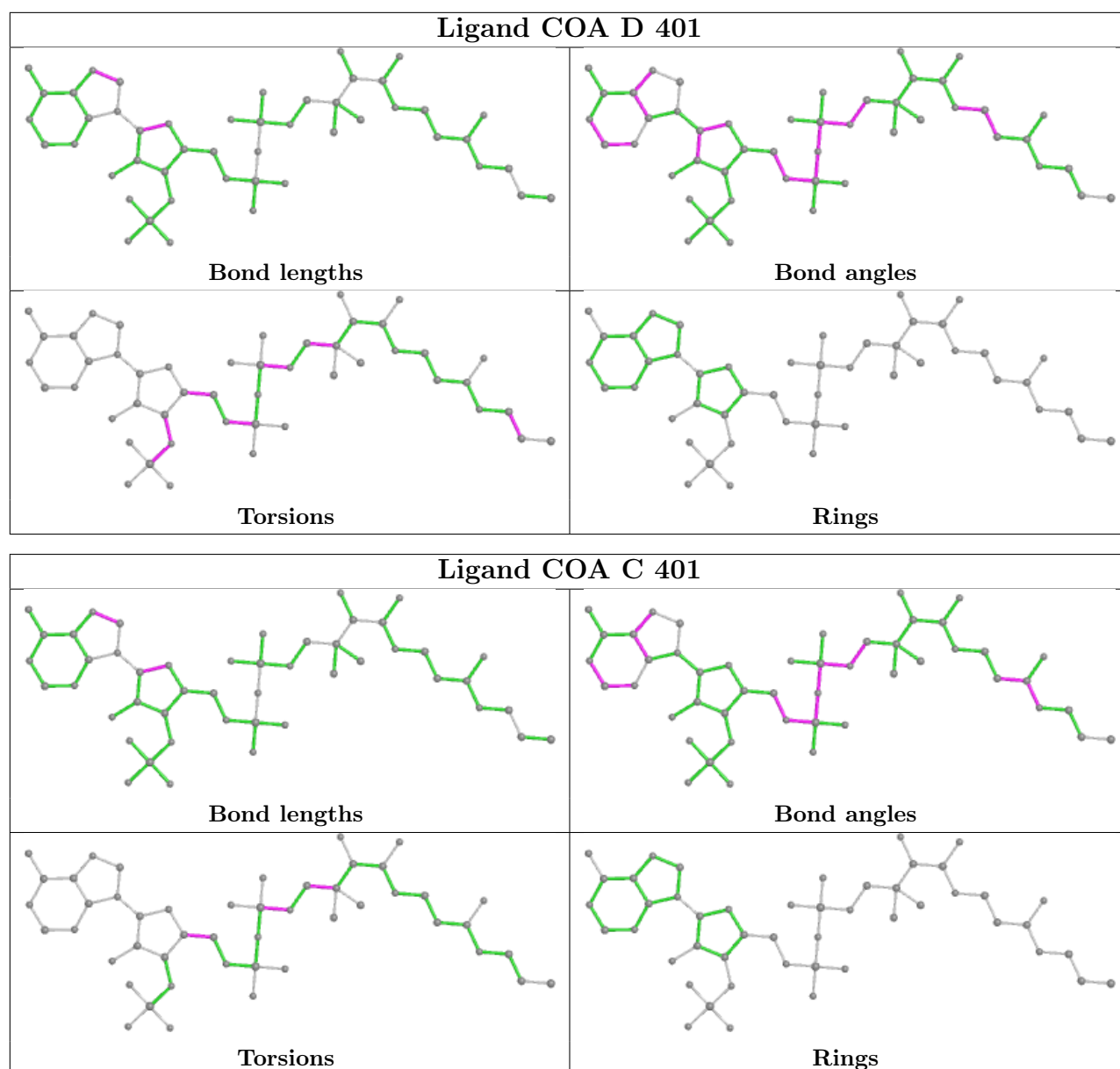
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	401	COA	2	0

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





## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	389/399 (97%)	-0.33	2 (0%) 91 94	14, 24, 44, 85	0
1	B	389/399 (97%)	-0.32	1 (0%) 94 96	13, 24, 43, 78	0
1	C	389/399 (97%)	0.13	10 (2%) 56 63	29, 48, 70, 91	0
1	D	389/399 (97%)	0.52	31 (7%) 12 16	28, 58, 91, 115	0
All	All	1556/1596 (97%)	0.00	44 (2%) 53 60	13, 39, 78, 115	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	227	SER	4.0
1	C	392	LEU	3.8
1	D	331	GLY	3.8
1	D	228	MET	3.5
1	D	156	HIS	3.3
1	D	175	GLN	3.3
1	D	223	ALA	3.3
1	D	188	ALA	3.2
1	D	234	ALA	3.1
1	D	221	HIS	3.0
1	D	179	ALA	2.9
1	D	246	ALA	2.9
1	D	330	LEU	2.9
1	D	241	VAL	2.7
1	B	208	LYS	2.7
1	C	207	ARG	2.7
1	D	132	VAL	2.7
1	D	283	VAL	2.7
1	C	371	ARG	2.6
1	C	209	GLY	2.6
1	D	332	TRP	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	226	ASP	2.5
1	D	3	PRO	2.5
1	A	392	LEU	2.4
1	C	132	VAL	2.4
1	D	289	GLY	2.4
1	D	346	ILE	2.3
1	D	318	ALA	2.2
1	A	207	ARG	2.2
1	D	113	ALA	2.2
1	D	274	ARG	2.2
1	C	206	GLY	2.2
1	D	244	GLY	2.2
1	C	221	HIS	2.1
1	C	208	LYS	2.1
1	D	319	PHE	2.1
1	D	170	LEU	2.1
1	D	40	GLU	2.1
1	C	74	ALA	2.1
1	D	177	ALA	2.1
1	D	171	SER	2.0
1	D	334	PRO	2.0
1	D	178	PHE	2.0
1	C	183	TYR	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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
1	SCY	D	89	9/10	0.91	0.15	40,47,68,70	0
1	SCY	B	89	9/10	0.96	0.16	25,31,46,48	0
1	SCY	C	89	9/10	0.97	0.13	37,43,60,62	0
1	SCY	A	89	9/10	0.97	0.16	22,35,48,52	0

## 6.3 Carbohydrates [i](#)

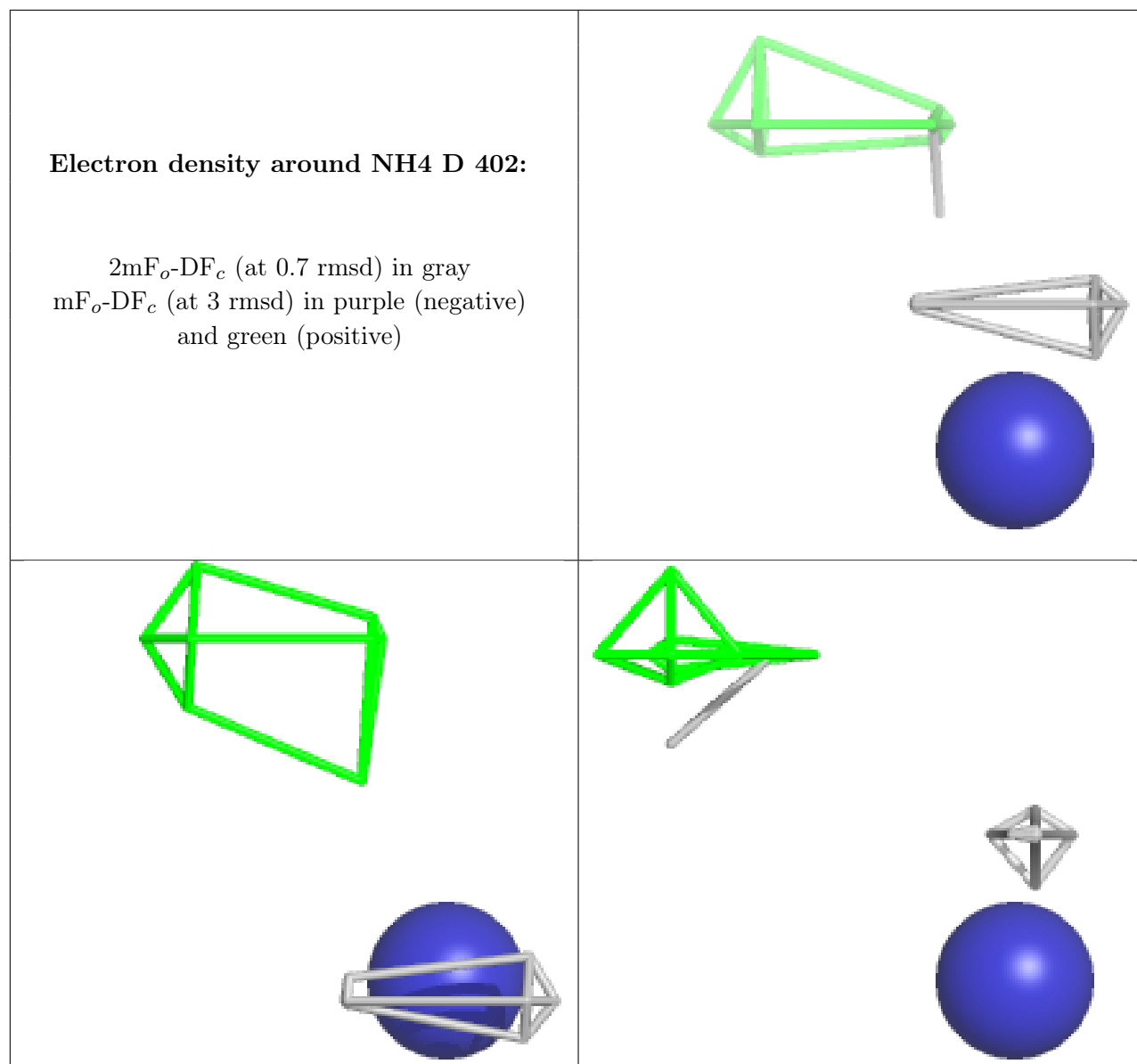
There are no monosaccharides in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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.

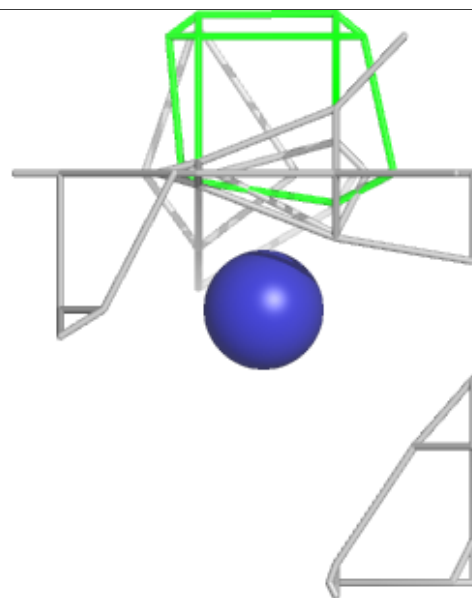
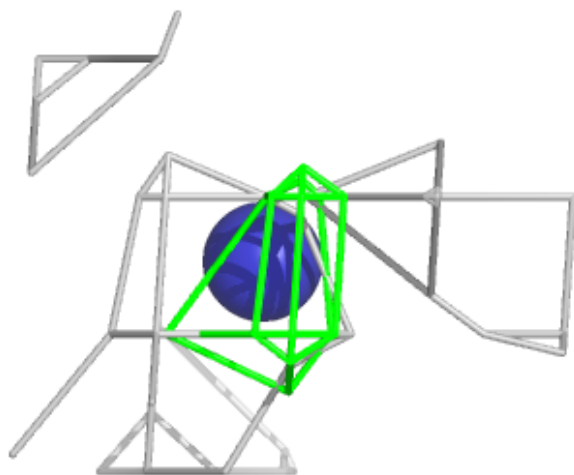
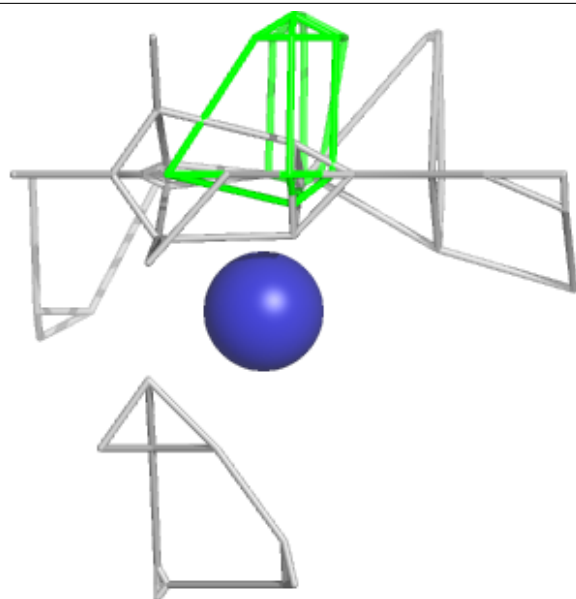
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NH4	D	402	1/1	0.54	0.65	61,61,61,61	0
4	NH4	B	407	1/1	0.72	0.54	67,67,67,67	0
4	NH4	A	407	1/1	0.75	0.26	41,41,41,41	0
4	NH4	C	403	1/1	0.84	0.16	45,45,45,45	0
2	SO4	A	404	5/5	0.86	0.17	98,100,101,102	0
2	SO4	B	404	5/5	0.87	0.25	80,82,84,88	0
2	SO4	A	406	5/5	0.88	0.16	94,96,96,97	0
3	COA	D	401	48/48	0.89	0.16	49,69,82,87	48
3	COA	C	401	48/48	0.92	0.13	40,52,85,89	48
2	SO4	B	402	5/5	0.95	0.24	75,80,83,83	0
2	SO4	B	401	5/5	0.95	0.10	65,66,67,68	0
2	SO4	C	402	5/5	0.95	0.16	99,100,101,103	0
2	SO4	B	406	5/5	0.96	0.11	94,94,95,96	0
2	SO4	A	401	5/5	0.96	0.10	50,53,55,58	0
3	COA	A	403	48/48	0.96	0.13	18,31,51,55	0
3	COA	B	403	48/48	0.96	0.13	19,29,70,72	48
2	SO4	B	405	5/5	0.96	0.15	87,87,89,90	0
2	SO4	A	402	5/5	0.97	0.15	73,74,74,74	0
2	SO4	A	405	5/5	0.98	0.12	71,72,74,74	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.



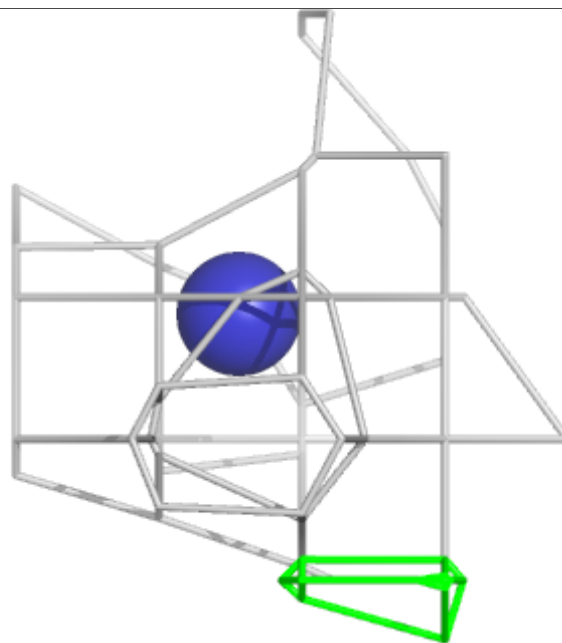
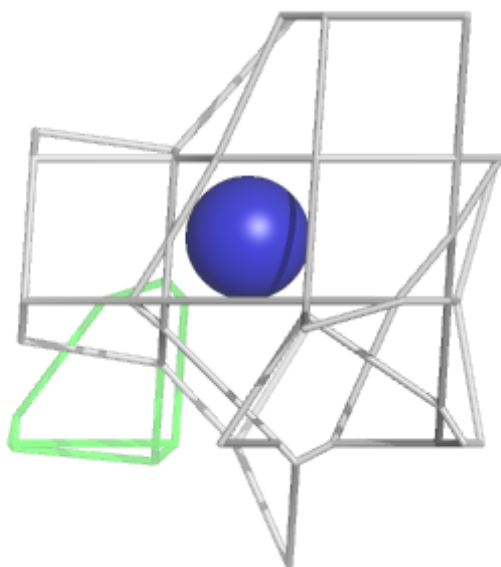
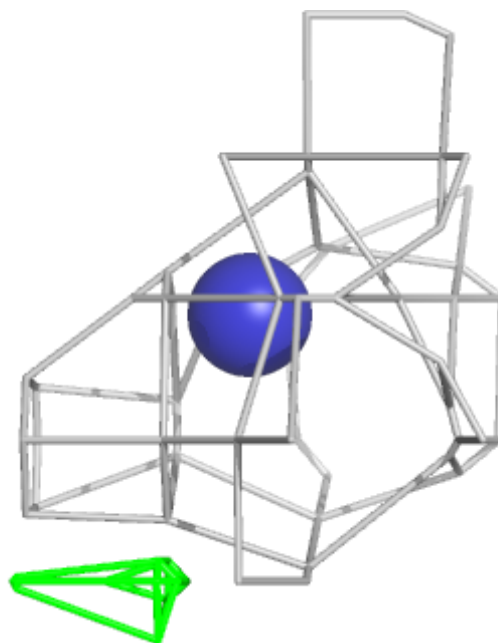
**Electron density around NH4 B 407:**

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



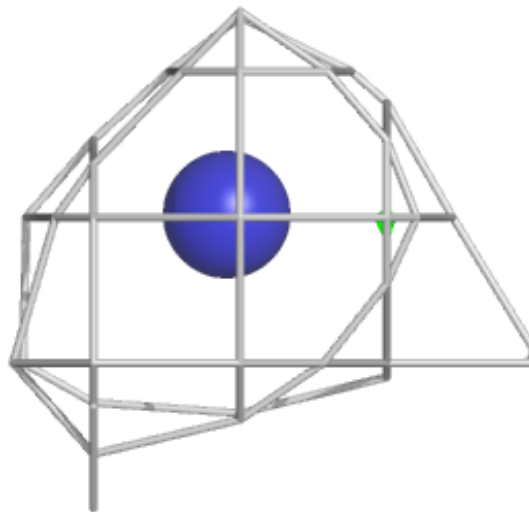
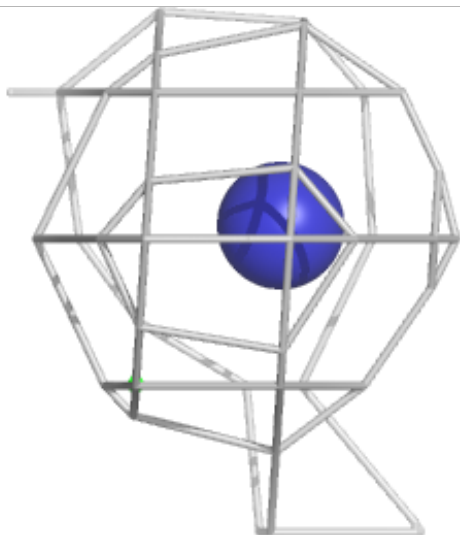
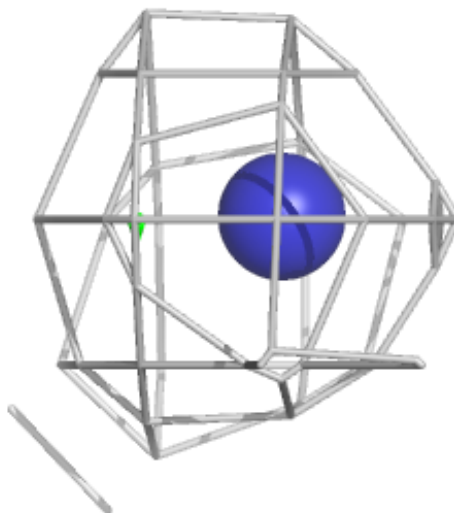
**Electron density around NH4 A 407:**

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



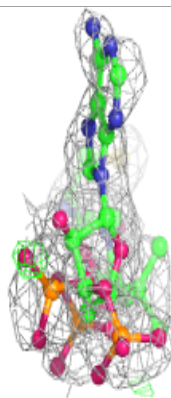
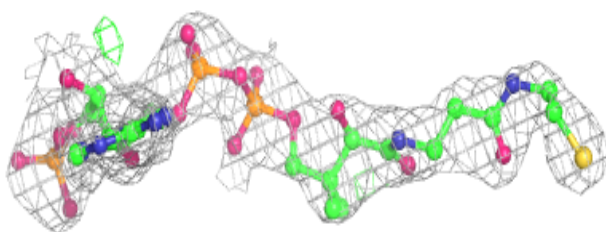
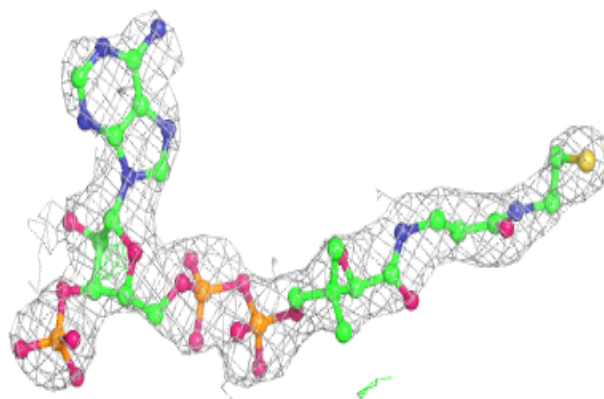
**Electron density around NH4 C 403:**

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

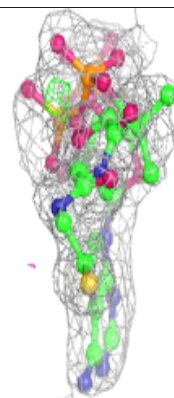
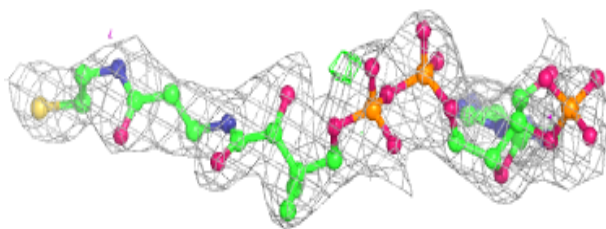
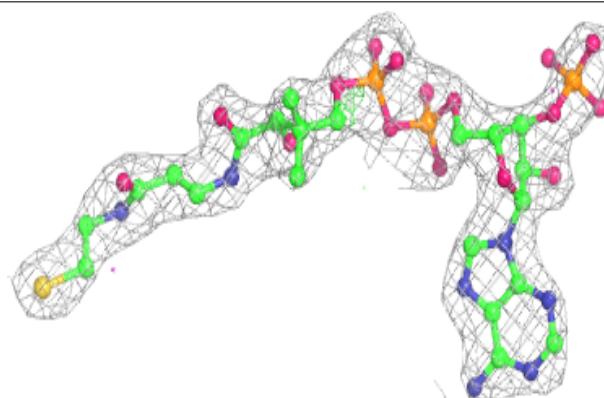


**Electron density around COA D 401:**

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

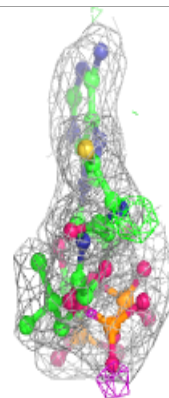
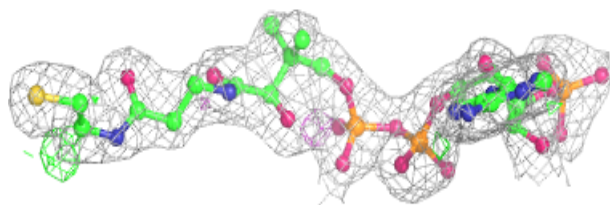
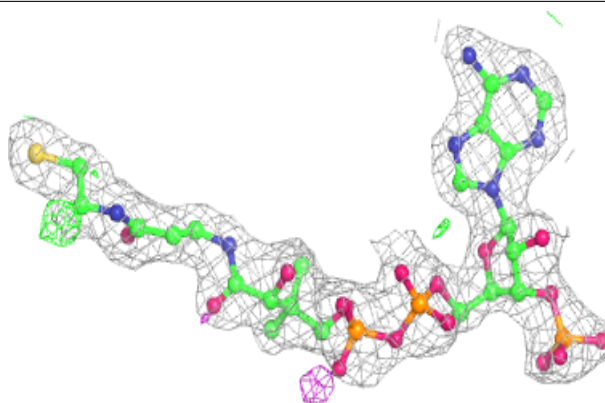
**Electron density around COA C 401:**

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

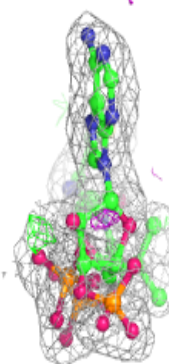
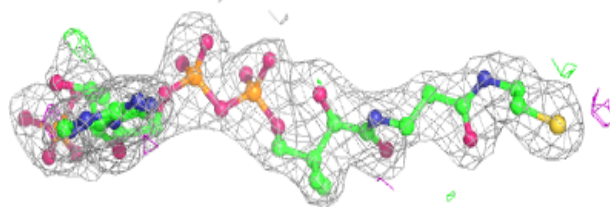
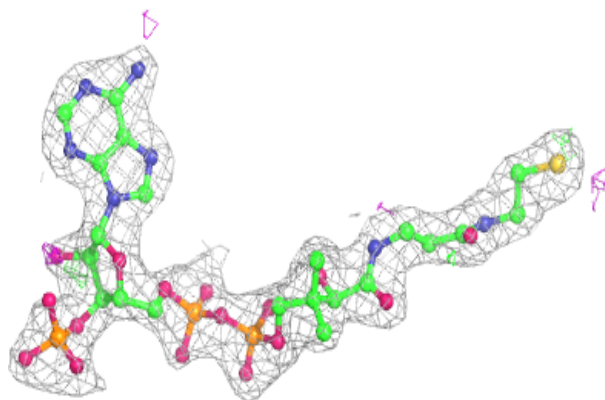


**Electron density around COA A 403:**

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

**Electron density around COA B 403:**

$2mF_o-DF_c$  (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.