



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 25, 2022 – 07:17 pm BST

PDB ID : 2VSU  
Title : A ternary complex of Hydroxycinnamoyl-CoA Hydratase-Lyase (HCHL) with acetyl-Coenzyme A and vanillin gives insights into substrate specificity and mechanism.  
Authors : Bennett, J.P.; Bertin, L.M.; Brzozowski, A.M.; Walton, N.J.; Grogan, G.  
Deposited on : 2008-04-29  
Resolution : 1.90 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	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28

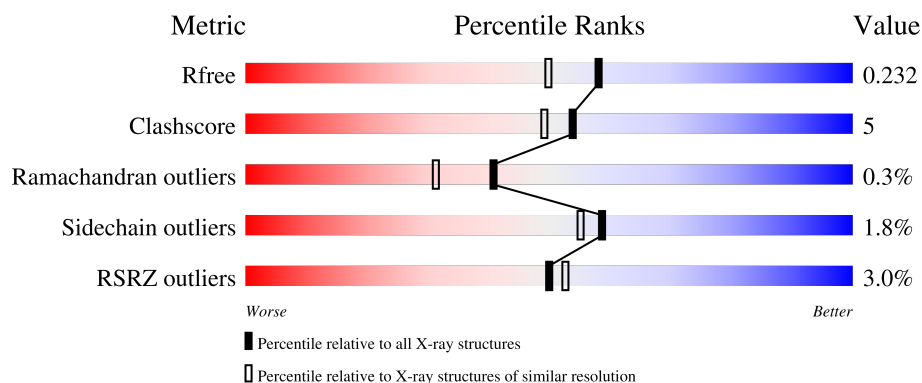
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	276	<div> <div>3%</div> <div>79%</div> <div>8%</div> <div>13%</div> </div>
1	B	276	<div> <div>4%</div> <div>78%</div> <div>11%</div> <div>10%</div> </div>
1	D	276	<div> <div>2%</div> <div>80%</div> <div>9%</div> <div>10%</div> </div>
2	C	275	<div> <div>6%</div> <div>80%</div> <div>10%</div> <div>9%</div> </div>
3	E	276	<div> <div>2%</div> <div>81%</div> <div>7%</div> <div>11%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	F	276	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	V55	D	1251	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12342 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 P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	241	Total	C	N	O	S	0	3	1
			1862	1180	322	347	13			
1	B	248	Total	C	N	O	S	0	2	1
			1958	1245	338	363	12			
1	D	248	Total	C	N	O	S	0	3	1
			1956	1244	340	360	12			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	123	ALA	SER	engineered mutation	UNP O69762
B	123	ALA	SER	engineered mutation	UNP O69762
D	123	ALA	SER	engineered mutation	UNP O69762

- Molecule 2 is a protein called P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	251	Total	C	N	O	S	0	3	0
			1942	1233	339	358	12			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	123	ALA	SER	engineered mutation	UNP O69762

- Molecule 3 is a protein called P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	245	Total	C	N	O	S	0	4	0
			1939	1232	337	357	13			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	123	ALA	SER	engineered mutation	UNP O69762
E	146	TYR	TRP	conflict	UNP O69762

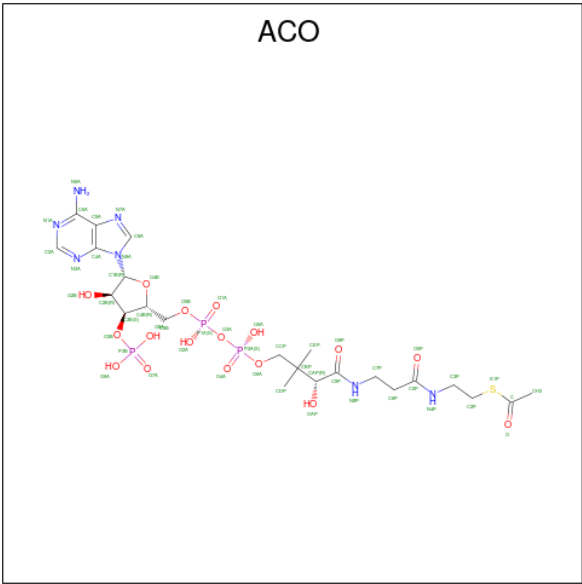
- Molecule 4 is a protein called P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	246	Total	C	N	O	S	0	2	1
			1931	1229	333	357	12			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	123	ALA	SER	engineered mutation	UNP O69762
F	29	ARG	LYS	conflict	UNP O69762

- Molecule 5 is ACETYL COENZYME \*A (three-letter code: ACO) (formula: C<sub>23</sub>H<sub>38</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>S).



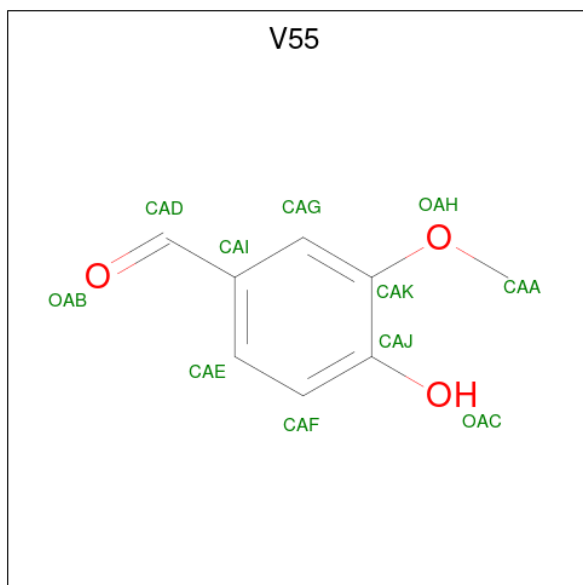
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
5	B	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
5	D	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
5	E	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	F	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

- Molecule 6 is 4-hydroxy-3-methoxybenzaldehyde (three-letter code: V55) (formula:  $C_8H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total	C	O	0	0
			11	8	3		

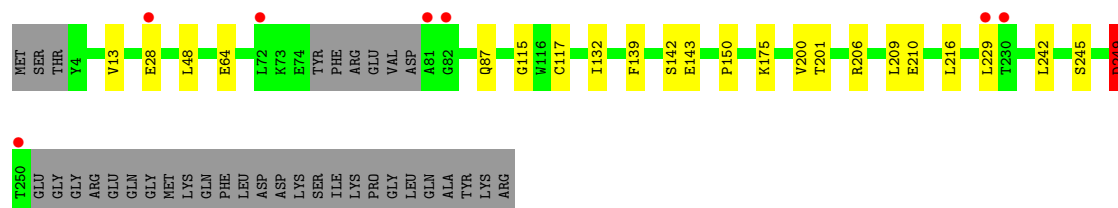
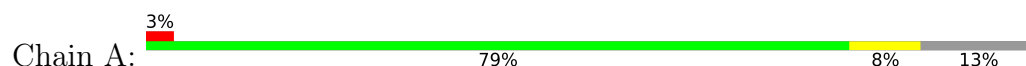
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	67	Total	O	0	0
			67	67		
7	B	84	Total	O	0	0
			84	84		
7	C	67	Total	O	0	0
			67	67		
7	D	99	Total	O	0	0
			99	99		
7	E	84	Total	O	0	0
			84	84		
7	F	87	Total	O	0	0
			87	87		

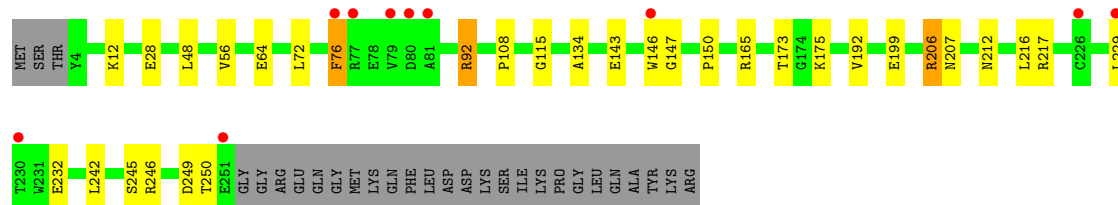
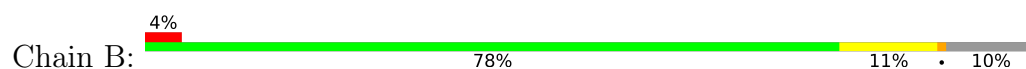
### 3 Residue-property plots [i](#)

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.

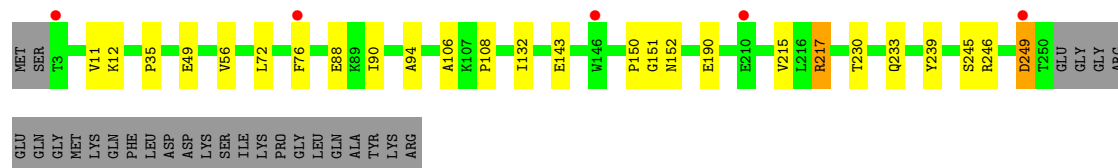
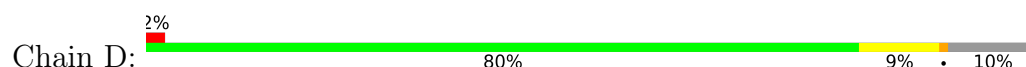
#### • Molecule 1: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



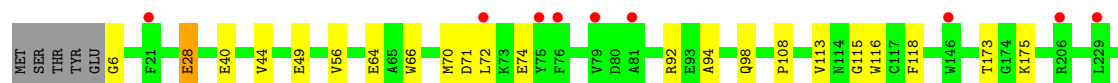
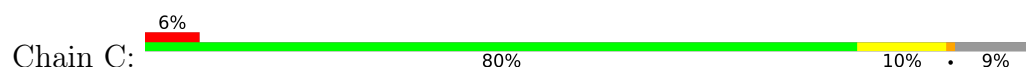
#### • Molecule 1: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

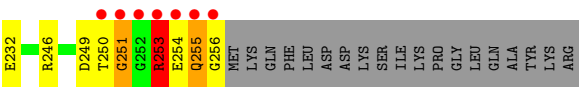


#### • Molecule 1: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

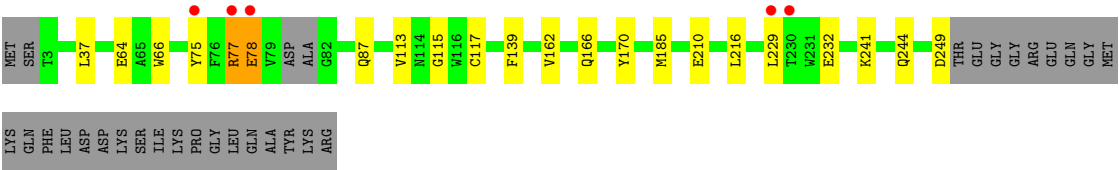
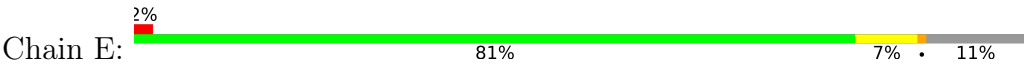


#### • Molecule 2: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE

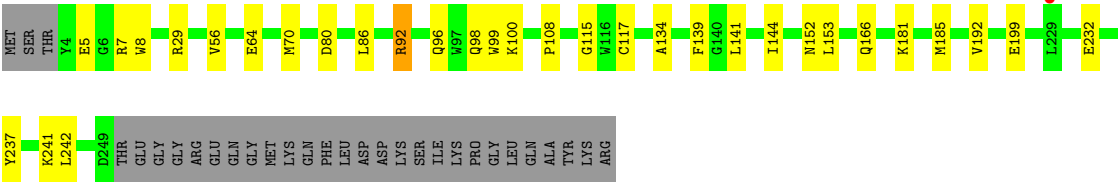
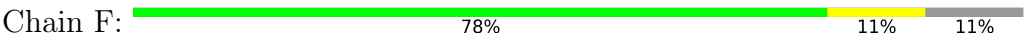




● Molecule 3: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE



● Molecule 4: P-HYDROXYCINNAMOYL COA HYDRATASE/LYASE





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.08Å 130.12Å 144.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	96.67 – 1.90 49.56 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.7 (96.67-1.90) 96.6 (49.56-1.90)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.88 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.3.0037	Depositor
R, $R_{free}$	0.188 , 0.230 0.190 , 0.232	Depositor DCC
$R_{free}$ test set	6562 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.5	Xtriage
Anisotropy	0.545	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12342	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% 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: V55, 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	0.73	1/1895 (0.1%)	0.76	0/2564
1	B	0.85	1/1997 (0.1%)	0.85	4/2706 (0.1%)
1	D	0.89	1/1998 (0.1%)	0.81	1/2706 (0.0%)
2	C	0.77	0/1980	0.74	2/2682 (0.1%)
3	E	0.82	0/1978	0.80	0/2677
4	F	0.85	0/1970	0.83	1/2668 (0.0%)
All	All	0.82	3/11818 (0.0%)	0.80	8/16003 (0.0%)

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
2	C	1	3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	206	ARG	C-N	-11.55	1.07	1.34
1	D	239	TYR	CD1-CE1	5.79	1.48	1.39
1	A	249	ASP	C-N	-5.15	1.22	1.34

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	92	ARG	NE-CZ-NH2	-13.28	113.66	120.30
1	B	92	ARG	NE-CZ-NH1	7.60	124.10	120.30
1	B	165	ARG	NE-CZ-NH2	-7.59	116.50	120.30
2	C	92	ARG	NE-CZ-NH2	-7.41	116.60	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	92	ARG	NE-CZ-NH2	-7.19	116.71	120.30
2	C	92	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	B	165	ARG	NE-CZ-NH1	5.53	123.06	120.30
1	D	217	ARG	CG-CD-NE	5.03	122.35	111.80

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	250	THR	CB

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	250	THR	Peptide
2	C	253	ARG	Peptide
2	C	254	GLU	Peptide

## 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	1862	0	1844	13	0
1	B	1958	0	1952	19	1
1	D	1956	0	1955	22	0
2	C	1942	0	1914	35	0
3	E	1939	0	1936	18	0
4	F	1931	0	1921	27	1
5	A	51	0	34	2	0
5	B	51	0	34	0	0
5	D	51	0	34	7	0
5	E	51	0	34	2	0
5	F	51	0	34	0	0
6	D	11	0	7	8	0
7	A	67	0	0	1	0
7	B	84	0	0	0	0
7	C	67	0	0	1	0
7	D	99	0	0	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	E	84	0	0	1	1
7	F	87	0	0	3	0
All	All	12342	0	11699	125	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (125) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:246:ARG:HD2	2:C:253:ARG:NH1	1.49	1.27
2:C:246:ARG:CD	2:C:253:ARG:NH1	2.11	1.13
2:C:253:ARG:O	2:C:253:ARG:HG2	1.41	1.12
2:C:246:ARG:HD2	2:C:253:ARG:HH12	1.01	1.00
5:D:1250:ACO:HH31	6:D:1251:V55:CAD	1.95	0.96
1:B:232:GLU:HG2	4:F:232:GLU:HG2	1.50	0.92
1:B:146:TRP:HE3	2:C:255:GLN:CB	1.82	0.92
1:D:245:SER:O	1:D:249:ASP:HB3	1.70	0.91
2:C:246:ARG:CD	2:C:253:ARG:HH12	1.78	0.91
4:F:99:TRP:HE1	4:F:152:ASN:HD21	1.19	0.89
3:E:75:TYR:OH	3:E:87:GLN:NE2	2.09	0.86
1:B:56:VAL:HG22	1:B:108:PRO:HG2	1.65	0.79
1:A:175:LYS:NZ	7:A:2043:HOH:O	2.19	0.75
2:C:253:ARG:O	2:C:253:ARG:CG	2.29	0.75
2:C:246:ARG:CD	2:C:253:ARG:HH11	2.01	0.74
4:F:70:MET:CE	4:F:98:GLN:HE22	2.01	0.73
4:F:70:MET:HE1	7:F:2030:HOH:O	1.90	0.72
3:E:166:GLN:NE2	3:E:185[A]:MET:HG2	2.06	0.70
4:F:152:ASN:HD22	4:F:153:LEU:H	1.38	0.70
2:C:246:ARG:CG	2:C:253:ARG:HH11	2.04	0.70
1:B:147:GLY:O	2:C:253:ARG:HB2	1.93	0.68
2:C:249:ASP:OD1	2:C:251:GLY:HA3	1.91	0.68
1:D:132:ILE:HD12	1:D:190:GLU:HG2	1.75	0.67
5:D:1250:ACO:CH3	6:D:1251:V55:OAB	2.45	0.65
1:B:232:GLU:HG2	4:F:232:GLU:CG	2.24	0.65
2:C:56:VAL:HG22	2:C:108:PRO:HG2	1.79	0.65
2:C:246:ARG:NE	2:C:253:ARG:NH1	2.46	0.64
1:D:72:LEU:HD21	5:D:1250:ACO:H22	1.79	0.64
1:D:230:THR:H	1:D:233:GLN:HE21	1.44	0.63
1:D:246:ARG:NH2	4:F:80:ASP:OD1	2.32	0.63
1:D:106:ALA:O	1:D:217:ARG:HD3	2.00	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:D:1250:ACO:HH31	6:D:1251:V55:OAB	1.99	0.62
1:B:173:THR:HB	1:B:175:LYS:HG3	1.82	0.61
2:C:232:GLU:HG2	3:E:232:GLU:HG3	1.83	0.60
2:C:246:ARG:HG2	2:C:253:ARG:HH11	1.65	0.60
1:A:87:GLN:HE22	1:B:246:ARG:HH22	1.49	0.60
1:A:142:SER:OG	5:A:1250:ACO:H31	2.01	0.60
4:F:56:VAL:HG22	4:F:108:PRO:HG2	1.84	0.58
4:F:99:TRP:HE1	4:F:152:ASN:ND2	1.98	0.58
1:B:92:ARG:NH2	1:D:49:GLU:OE2	2.25	0.57
1:D:94:ALA:HB1	6:D:1251:V55:HAA1	1.85	0.57
4:F:166:GLN:HE21	4:F:181[A]:LYS:NZ	2.03	0.57
2:C:40:GLU:O	2:C:44:VAL:HG23	2.06	0.56
1:A:142:SER:HG	5:A:1250:ACO:H31	1.69	0.56
2:C:232:GLU:HG2	3:E:232:GLU:CG	2.36	0.55
5:D:1250:ACO:CH3	6:D:1251:V55:CAD	2.78	0.54
2:C:28:GLU:CD	2:C:28:GLU:H	2.11	0.53
4:F:166:GLN:HE21	4:F:181[A]:LYS:HZ1	1.55	0.52
1:A:143:GLU:HG2	1:A:150:PRO:HA	1.90	0.52
1:A:117:CYS:O	1:A:139:PHE:HA	2.10	0.52
1:A:28:GLU:H	1:A:28:GLU:CD	2.14	0.52
4:F:29:ARG:NH1	7:F:2003:HOH:O	2.43	0.51
3:E:166:GLN:HE22	3:E:185[A]:MET:HG2	1.76	0.51
2:C:173:THR:OG1	2:C:175:LYS:HG3	2.11	0.50
1:D:56:VAL:HG22	1:D:108:PRO:HG2	1.92	0.50
1:B:146:TRP:CE3	2:C:255:GLN:CB	2.75	0.50
3:E:241:LYS:NZ	3:E:244:GLN:HE22	2.09	0.50
1:D:215:VAL:HG12	4:F:144:ILE:HD12	1.94	0.49
4:F:70:MET:HE1	4:F:98:GLN:HE22	1.77	0.49
1:A:132:ILE:HD13	1:A:200:VAL:HG12	1.95	0.49
3:E:64:GLU:O	3:E:115:GLY:HA3	2.12	0.49
2:C:71:ASP:CG	2:C:74:GLU:HB2	2.33	0.49
1:A:13:VAL:HG11	1:A:48:LEU:HD21	1.96	0.48
2:C:94:ALA:O	2:C:98:GLN:HG3	2.14	0.48
1:D:230:THR:HG23	1:D:233:GLN:HE21	1.79	0.48
1:B:64:GLU:O	1:B:115:GLY:HA3	2.14	0.47
2:C:232:GLU:CG	3:E:232:GLU:HG3	2.45	0.47
1:B:143:GLU:HG2	1:B:150:PRO:HA	1.96	0.47
1:D:72:LEU:HD13	1:D:76:PHE:HD1	1.80	0.47
3:E:249:ASP:OD1	3:E:249:ASP:C	2.53	0.47
1:D:11[A]:VAL:HG12	1:D:12:LYS:N	2.30	0.47
3:E:77:ARG:CG	3:E:78:GLU:N	2.78	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:F:134:ALA:HA	4:F:192:VAL:O	2.15	0.46
2:C:116:TRP:CE3	2:C:118:PHE:HE1	2.34	0.46
1:D:72:LEU:CD2	5:D:1250:ACO:H22	2.45	0.46
4:F:199:GLU:HG3	7:F:2067:HOH:O	2.15	0.46
5:D:1250:ACO:HH33	6:D:1251:V55:OAB	2.16	0.46
4:F:96:GLN:HA	4:F:100:LYS:HB2	1.97	0.46
1:A:132:ILE:HD12	1:A:201:THR:HA	1.97	0.45
4:F:152:ASN:ND2	4:F:153:LEU:H	2.09	0.45
4:F:181[A]:LYS:HG3	4:F:185:MET:CE	2.46	0.45
1:A:206:ARG:O	1:A:210:GLU:HG3	2.15	0.45
2:C:49:GLU:OE2	4:F:92:ARG:NH2	2.31	0.45
1:D:151:GLY:HA3	6:D:1251:V55:CAE	2.47	0.45
2:C:70:MET:SD	2:C:72:LEU:HD13	2.57	0.45
1:D:35:PRO:HB3	1:D:90:ILE:HD12	1.99	0.45
2:C:66:TRP:CD1	2:C:113:VAL:HA	2.52	0.45
5:E:1250:ACO:H61	5:E:1250:ACO:O9P	2.15	0.44
2:C:246:ARG:HD2	2:C:253:ARG:HH11	1.57	0.44
1:D:143:GLU:HG2	1:D:150:PRO:HA	2.00	0.44
1:B:76:PHE:CD2	1:B:76:PHE:C	2.92	0.43
2:C:246:ARG:CG	2:C:253:ARG:NH1	2.66	0.43
1:D:152:ASN:CB	6:D:1251:V55:HAA2	2.49	0.43
4:F:237:TYR:OH	4:F:241:LYS:HD2	2.17	0.43
2:C:64:GLU:O	2:C:115:GLY:HA3	2.17	0.43
4:F:166:GLN:NE2	4:F:181[A]:LYS:NZ	2.67	0.43
1:B:217:ARG:NH1	1:D:88:GLU:OE1	2.52	0.43
1:D:72:LEU:HD13	1:D:76:PHE:CD1	2.54	0.43
2:C:232:GLU:CB	3:E:232:GLU:HG3	2.49	0.43
2:C:255:GLN:O	2:C:256:GLY:C	2.57	0.43
1:D:230:THR:HG23	1:D:233:GLN:NE2	2.34	0.43
4:F:181[A]:LYS:HG3	4:F:185:MET:HE3	2.00	0.43
3:E:210:GLU:HB3	7:E:2068:HOH:O	2.19	0.42
1:B:134:ALA:HA	1:B:192:VAL:O	2.19	0.42
1:B:28:GLU:H	1:B:28:GLU:CD	2.23	0.42
4:F:70:MET:HE3	4:F:98:GLN:HE22	1.80	0.42
1:B:12:LYS:HB3	1:B:12:LYS:HE3	1.75	0.42
2:C:6:GLY:HA2	7:C:2001:HOH:O	2.19	0.42
2:C:116:TRP:HE3	2:C:118:PHE:HE1	1.68	0.42
4:F:64:GLU:O	4:F:115:GLY:HA3	2.20	0.42
1:A:64:GLU:O	1:A:115:GLY:HA3	2.19	0.42
3:E:170:TYR:CE2	3:E:185[A]:MET:SD	3.13	0.42
2:C:232:GLU:HG2	3:E:232:GLU:CD	2.40	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:117:CYS:O	3:E:139:PHE:HA	2.21	0.41
3:E:37:LEU:C	3:E:37:LEU:HD23	2.41	0.41
3:E:162:VAL:HG22	3:E:185[B]:MET:HG3	2.02	0.41
4:F:7:ARG:HD3	4:F:8:TRP:CH2	2.56	0.41
1:B:245:SER:O	1:B:249:ASP:HB3	2.21	0.40
4:F:117:CYS:O	4:F:139:PHE:HA	2.21	0.40
3:E:66:TRP:CD1	3:E:113:VAL:HA	2.56	0.40
1:B:48:LEU:HD23	1:B:48:LEU:HA	1.86	0.40
1:D:230:THR:H	1:D:233:GLN:NE2	2.14	0.40
1:A:242:LEU:O	1:A:245:SER:HB3	2.21	0.40
5:E:1250:ACO:O9P	5:E:1250:ACO:C6P	2.69	0.40
1:B:206:ARG:O	1:B:207:ASN:C	2.56	0.40

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:199:GLU:OE2	4:F:29:ARG:NH2[3_544]	1.67	0.53
7:D:2002:HOH:O	7:E:2066:HOH:O[4_445]	2.18	0.02

## 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	238/276 (86%)	231 (97%)	6 (2%)	1 (0%)	34	24
1	B	248/276 (90%)	239 (96%)	9 (4%)	0	100	100
1	D	248/276 (90%)	240 (97%)	8 (3%)	0	100	100
2	C	251/275 (91%)	241 (96%)	8 (3%)	2 (1%)	19	9
3	E	245/276 (89%)	235 (96%)	9 (4%)	1 (0%)	34	24
4	F	245/276 (89%)	237 (97%)	8 (3%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1475/1655 (89%)	1423 (96%)	48 (3%)	4 (0%)	41	31

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	249	ASP
2	C	251	GLY
2	C	255	GLN
3	E	78	GLU

### 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	190/226 (84%)	186 (98%)	4 (2%)	53	48
1	B	203/226 (90%)	195 (96%)	8 (4%)	32	23
1	D	202/226 (89%)	201 (100%)	1 (0%)	88	89
2	C	195/225 (87%)	193 (99%)	2 (1%)	76	76
3	E	200/226 (88%)	197 (98%)	3 (2%)	65	62
4	F	199/226 (88%)	195 (98%)	4 (2%)	55	51
All	All	1189/1355 (88%)	1167 (98%)	22 (2%)	59	53

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	209	LEU
1	A	216	LEU
1	A	229	LEU
1	A	249	ASP
1	B	72	LEU
1	B	76	PHE
1	B	212	ASN
1	B	216	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	229	LEU
1	B	242[A]	LEU
1	B	242[B]	LEU
1	B	250	THR
2	C	28	GLU
2	C	253	ARG
1	D	249	ASP
3	E	77	ARG
3	E	216	LEU
3	E	229	LEU
4	F	5	GLU
4	F	86	LEU
4	F	141	LEU
4	F	242	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	25	ASN
1	A	50	GLN
1	A	87	GLN
1	A	180	GLN
1	B	25	ASN
1	B	212	ASN
2	C	25	ASN
2	C	207	ASN
1	D	25	ASN
1	D	50	GLN
1	D	87	GLN
1	D	233	GLN
3	E	25	ASN
3	E	50	GLN
3	E	87	GLN
3	E	166	GLN
3	E	244	GLN
4	F	25	ASN
4	F	98	GLN
4	F	152	ASN
4	F	166	GLN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

6 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	$\# Z  > 2$
5	ACO	A	1250	-	45,53,53	1.15	4 (8%)	56,79,79	1.29	5 (8%)
5	ACO	B	1251	-	45,53,53	1.15	3 (6%)	56,79,79	1.43	8 (14%)
5	ACO	E	1250	-	45,53,53	1.23	4 (8%)	56,79,79	1.49	10 (17%)
6	V55	D	1251	-	11,11,11	0.95	0	14,14,14	1.17	1 (7%)
5	ACO	F	1249	-	45,53,53	1.43	6 (13%)	56,79,79	1.44	7 (12%)
5	ACO	D	1250	-	45,53,53	1.40	4 (8%)	56,79,79	1.42	9 (16%)

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
5	ACO	A	1250	-	-	14/47/67/67	0/3/3/3
5	ACO	B	1251	-	-	1/47/67/67	0/3/3/3

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	ACO	E	1250	-	-	9/47/67/67	0/3/3/3
6	V55	D	1251	-	-	0/4/4/4	0/1/1/1
5	ACO	F	1249	-	-	0/47/67/67	0/3/3/3
5	ACO	D	1250	-	-	1/47/67/67	0/3/3/3

All (21) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	1249	ACO	O4B-C1B	6.03	1.49	1.41
5	D	1250	ACO	O4B-C1B	5.83	1.49	1.41
5	D	1250	ACO	P2A-O4A	3.71	1.64	1.50
5	A	1250	ACO	P2A-O4A	3.51	1.63	1.50
5	E	1250	ACO	P1A-O1A	3.45	1.63	1.50
5	A	1250	ACO	P3B-O7A	3.38	1.61	1.50
5	E	1250	ACO	P2A-O4A	3.37	1.62	1.50
5	B	1251	ACO	P1A-O1A	3.27	1.62	1.50
5	B	1251	ACO	P2A-O4A	3.15	1.62	1.50
5	F	1249	ACO	P3B-O7A	3.15	1.60	1.50
5	E	1250	ACO	O4B-C1B	2.99	1.45	1.41
5	D	1250	ACO	P3B-O7A	2.91	1.59	1.50
5	F	1249	ACO	C-S1P	2.85	1.92	1.75
5	B	1251	ACO	P3B-O7A	2.81	1.59	1.50
5	A	1250	ACO	P1A-O1A	2.66	1.60	1.50
5	E	1250	ACO	P3B-O7A	2.63	1.59	1.50
5	F	1249	ACO	P1A-O1A	2.44	1.59	1.50
5	D	1250	ACO	P1A-O1A	2.41	1.59	1.50
5	F	1249	ACO	OAP-CAP	2.22	1.46	1.42
5	A	1250	ACO	P3B-O3B	2.05	1.63	1.59
5	F	1249	ACO	C2B-C1B	-2.03	1.50	1.53

All (40) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	1250	ACO	N3A-C2A-N1A	-5.05	120.79	128.68
5	F	1249	ACO	N3A-C2A-N1A	-4.89	121.03	128.68
5	B	1251	ACO	N3A-C2A-N1A	-4.76	121.25	128.68
5	A	1250	ACO	N3A-C2A-N1A	-4.49	121.66	128.68
5	B	1251	ACO	C7P-C6P-C5P	-3.83	105.99	112.36
5	D	1250	ACO	N3A-C2A-N1A	-3.78	122.77	128.68
5	D	1250	ACO	O3B-P3B-O7A	-3.50	95.90	109.39
5	B	1251	ACO	O4B-C1B-C2B	-3.47	101.85	106.93
5	A	1250	ACO	C6P-C7P-N8P	-3.24	105.36	111.90

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1250	ACO	O4B-C1B-C2B	-3.19	102.26	106.93
5	F	1249	ACO	CEP-CBP-CAP	3.17	114.31	108.82
5	F	1249	ACO	O9A-P3B-O8A	3.04	119.26	107.64
5	E	1250	ACO	O4B-C1B-C2B	-3.02	102.52	106.93
5	B	1251	ACO	O9A-P3B-O8A	2.89	118.68	107.64
5	E	1250	ACO	C7P-C6P-C5P	2.82	117.05	112.36
5	F	1249	ACO	C6P-C7P-N8P	-2.81	106.22	111.90
5	E	1250	ACO	C6P-C7P-N8P	-2.69	106.47	111.90
5	D	1250	ACO	O4B-C1B-C2B	-2.68	103.02	106.93
5	E	1250	ACO	O9A-P3B-O8A	2.63	117.70	107.64
5	D	1250	ACO	C4A-C5A-N7A	-2.53	106.77	109.40
5	E	1250	ACO	O5P-C5P-C6P	-2.49	117.47	122.02
5	A	1250	ACO	C6P-C5P-N4P	-2.48	112.24	116.42
5	E	1250	ACO	CDP-CBP-CAP	2.45	113.06	108.82
5	D	1250	ACO	CDP-CBP-CCP	-2.39	104.33	108.23
6	D	1251	V55	CAE-CAF-CAJ	-2.37	118.06	120.50
5	D	1250	ACO	C2P-C3P-N4P	-2.36	107.47	112.42
5	D	1250	ACO	C7P-C6P-C5P	-2.35	108.44	112.36
5	B	1251	ACO	O6A-P2A-O4A	-2.35	99.90	109.07
5	E	1250	ACO	O3B-P3B-O7A	-2.34	100.38	109.39
5	F	1249	ACO	C4A-C5A-N7A	-2.29	107.01	109.40
5	B	1251	ACO	C4A-C5A-N7A	-2.24	107.06	109.40
5	D	1250	ACO	C6P-C7P-N8P	-2.24	107.38	111.90
5	F	1249	ACO	O-C-CH3	-2.23	113.93	123.07
5	E	1250	ACO	C4A-C5A-N7A	-2.18	107.13	109.40
5	B	1251	ACO	CDP-CBP-CAP	2.13	112.52	108.82
5	E	1250	ACO	C2P-C3P-N4P	2.08	116.78	112.42
5	D	1250	ACO	C5A-C6A-N6A	2.07	123.49	120.35
5	F	1249	ACO	O4B-C1B-C2B	-2.05	103.93	106.93
5	A	1250	ACO	O5P-C5P-N4P	2.03	126.85	123.01
5	B	1251	ACO	C5A-C6A-N6A	2.01	123.41	120.35

There are no chirality outliers.

All (25) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1250	ACO	C5B-O5B-P1A-O1A
5	A	1250	ACO	P1A-O3A-P2A-O6A
5	A	1250	ACO	O-C-S1P-C2P
5	A	1250	ACO	CH3-C-S1P-C2P
5	B	1251	ACO	C3B-O3B-P3B-O7A
5	D	1250	ACO	C3B-O3B-P3B-O8A

*Continued on next page...*

*Continued from previous page...*

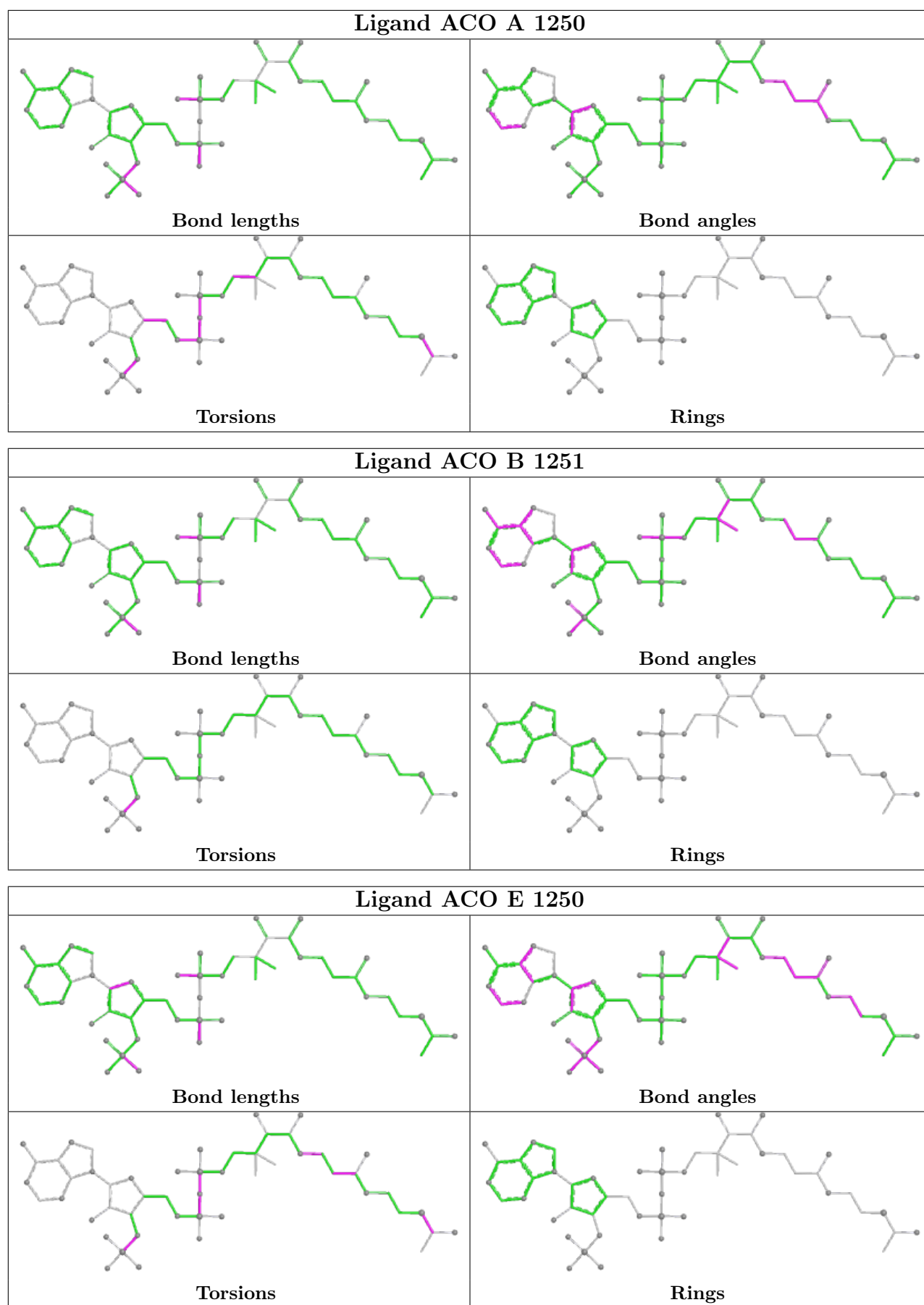
Mol	Chain	Res	Type	Atoms
5	E	1250	ACO	C3B-O3B-P3B-O7A
5	E	1250	ACO	O-C-S1P-C2P
5	E	1250	ACO	CH3-C-S1P-C2P
5	E	1250	ACO	C6P-C7P-N8P-C9P
5	A	1250	ACO	O4B-C4B-C5B-O5B
5	A	1250	ACO	C3B-C4B-C5B-O5B
5	E	1250	ACO	P1A-O3A-P2A-O6A
5	E	1250	ACO	O5P-C5P-C6P-C7P
5	A	1250	ACO	C5B-O5B-P1A-O3A
5	E	1250	ACO	N4P-C5P-C6P-C7P
5	A	1250	ACO	C5B-O5B-P1A-O2A
5	A	1250	ACO	CAP-CBP-CCP-O6A
5	A	1250	ACO	CDP-CBP-CCP-O6A
5	A	1250	ACO	P2A-O3A-P1A-O2A
5	A	1250	ACO	CEP-CBP-CCP-O6A
5	A	1250	ACO	C3B-O3B-P3B-O8A
5	A	1250	ACO	P2A-O3A-P1A-O1A
5	E	1250	ACO	P2A-O3A-P1A-O1A
5	E	1250	ACO	P2A-O3A-P1A-O2A

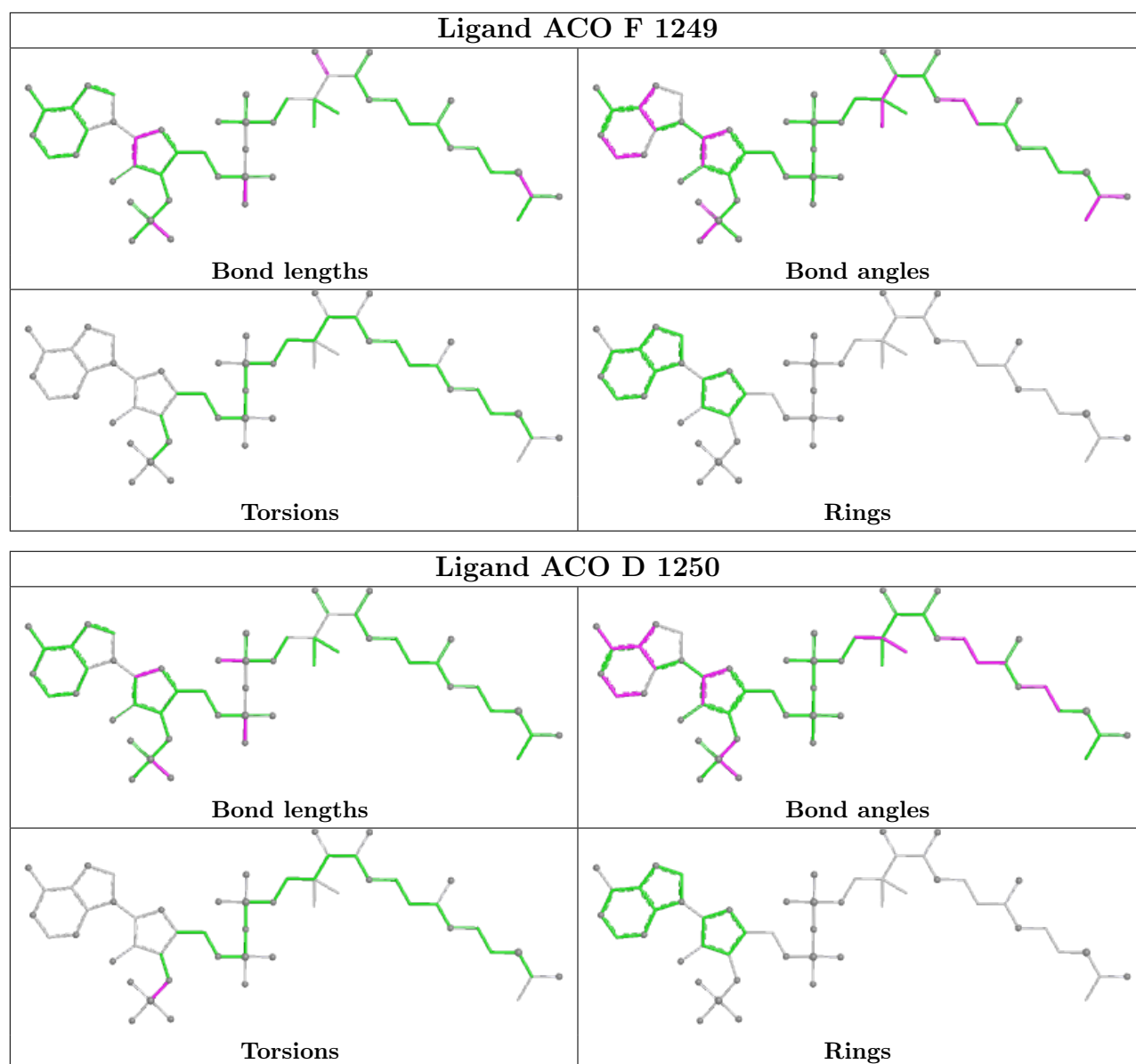
There are no ring outliers.

4 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1250	ACO	2	0
5	E	1250	ACO	2	0
6	D	1251	V55	8	0
5	D	1250	ACO	7	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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	206:ARG	C	207:ASN	N	1.07



## 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	241/276 (87%)	0.24	7 (2%) 51 54	19, 28, 37, 47	4 (1%)
1	B	248/276 (89%)	0.06	10 (4%) 38 41	17, 23, 36, 51	6 (2%)
1	D	248/276 (89%)	-0.08	5 (2%) 65 68	15, 21, 33, 46	6 (2%)
2	C	251/275 (91%)	0.53	16 (6%) 19 22	19, 28, 40, 58	4 (1%)
3	E	245/276 (88%)	-0.05	5 (2%) 65 68	16, 22, 35, 54	4 (1%)
4	F	246/276 (89%)	-0.15	1 (0%) 92 93	17, 23, 31, 37	3 (1%)
All	All	1479/1655 (89%)	0.09	44 (2%) 50 53	15, 24, 36, 58	27 (1%)

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	250	THR	7.9
2	C	255	GLN	7.2
2	C	252	GLY	7.0
2	C	250	THR	6.8
2	C	254	GLU	6.5
1	B	81	ALA	6.4
2	C	253	ARG	6.4
2	C	251	GLY	6.3
2	C	256	GLY	6.3
1	B	76	PHE	5.6
2	C	81	ALA	5.0
2	C	75	TYR	4.9
3	E	75	TYR	4.9
1	B	251	GLU	4.6
2	C	146	TRP	4.5
1	D	146	TRP	3.8
2	C	76	PHE	3.6
1	B	229	LEU	3.3
2	C	229[A]	LEU	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	F	229	LEU	3.1
2	C	72	LEU	3.0
2	C	79	VAL	2.9
3	E	229	LEU	2.9
1	A	82	GLY	2.7
1	D	3	THR	2.7
1	B	77	ARG	2.6
3	E	77	ARG	2.6
1	B	79	VAL	2.6
3	E	230	THR	2.6
1	B	230	THR	2.5
1	A	81	ALA	2.5
1	B	80	ASP	2.4
3	E	78	GLU	2.4
1	D	210	GLU	2.4
1	B	146	TRP	2.3
1	A	28	GLU	2.2
1	D	76	PHE	2.2
1	D	249	ASP	2.2
1	A	229	LEU	2.1
2	C	206	ARG	2.1
1	A	230	THR	2.1
1	A	72[A]	LEU	2.1
1	B	226	CYS	2.1
2	C	21	PHE	2.0

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

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

## 6.3 Carbohydrates ⓘ

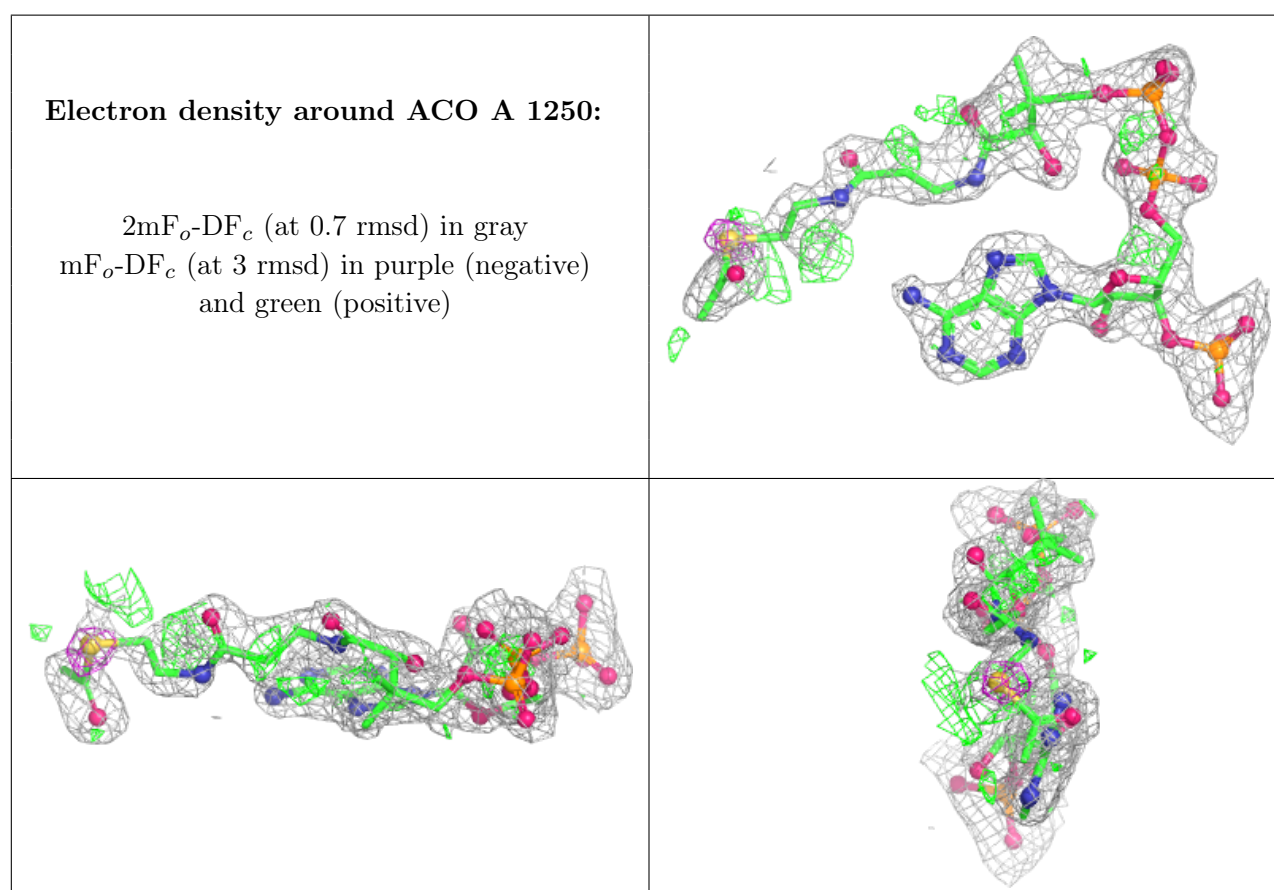
There are no monosaccharides in this entry.

## 6.4 Ligands ⓘ

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

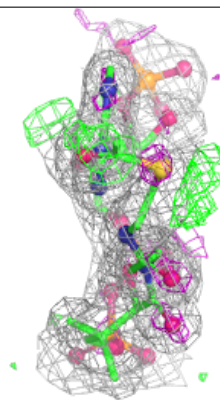
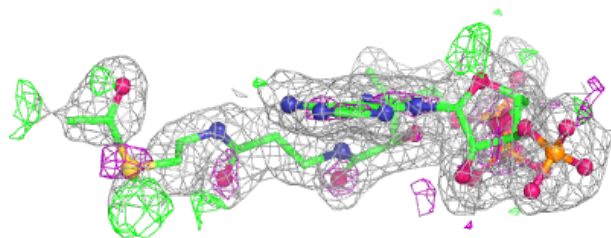
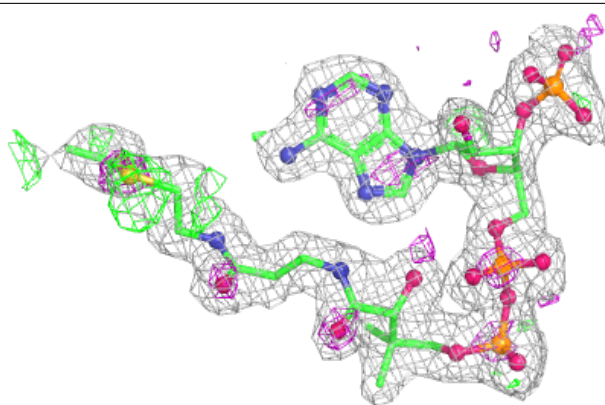
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	ACO	A	1250	51/51	0.85	0.17	15,32,37,40	49
6	V55	D	1251	11/11	0.93	0.26	22,29,31,32	3
5	ACO	F	1249	51/51	0.94	0.14	12,29,36,40	2
5	ACO	E	1250	51/51	0.95	0.13	15,32,39,40	2
5	ACO	D	1250	51/51	0.96	0.15	19,28,34,38	1
5	ACO	B	1251	51/51	0.96	0.14	26,32,36,39	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.

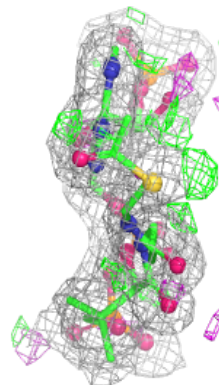
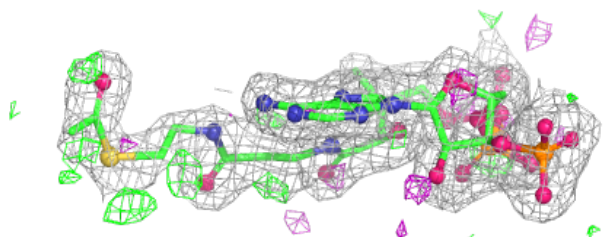
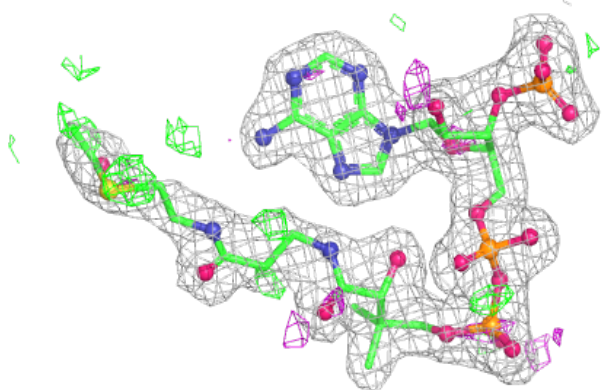


**Electron density around ACO F 1249:**

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

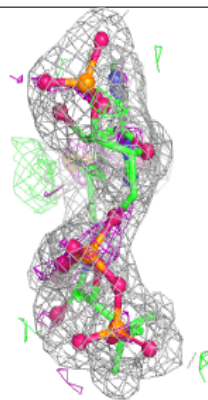
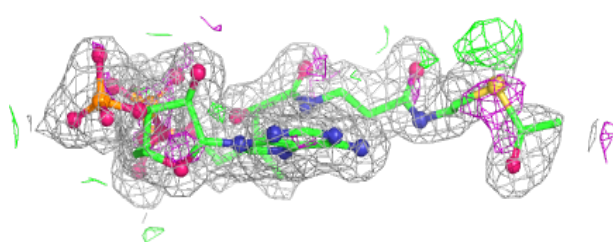
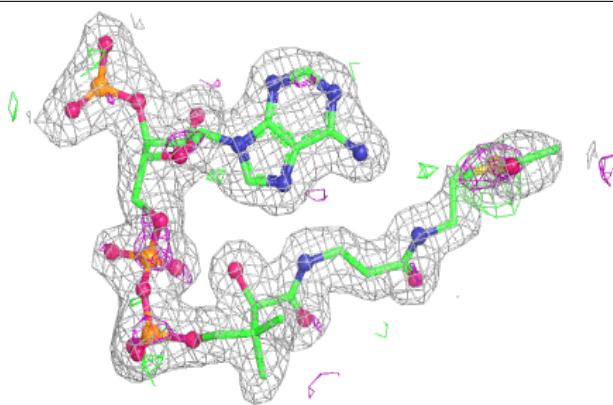
**Electron density around ACO E 1250:**

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

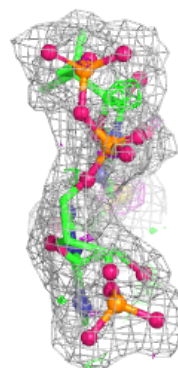
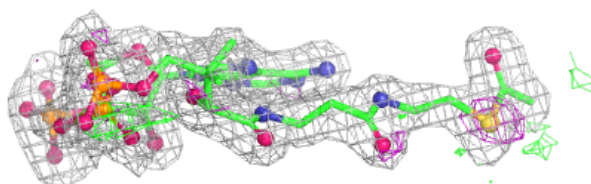
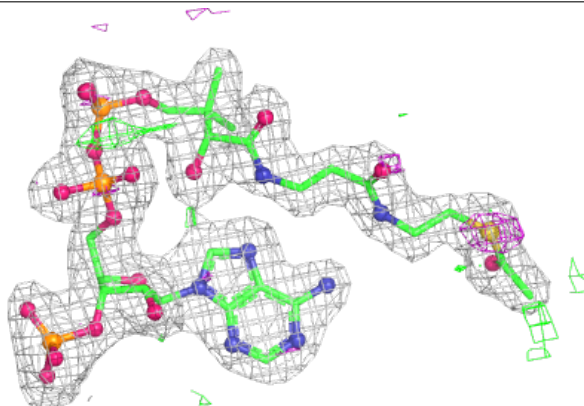


**Electron density around ACO D 1250:**

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

**Electron density around ACO B 1251:**

$2mF_o-DF_c$  (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.