WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Sep 2, 2023 – 12:25 PM EDT

PDB ID	:	3Q0J
Title	:	Crystal Structure of the Mycobacterium tuberculosis Crotonase in complex with the Inhibitor AcetoacetylCoA
Authors	:	Bruning, J.B.; Delgado, E.; Ghosh, S.; Sacchettini, J.C.; TB Structural Genomics Consortium (TBSGC)
Deposited on	:	2010-12-15
Resolution	:	2.40 Å(reported)

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

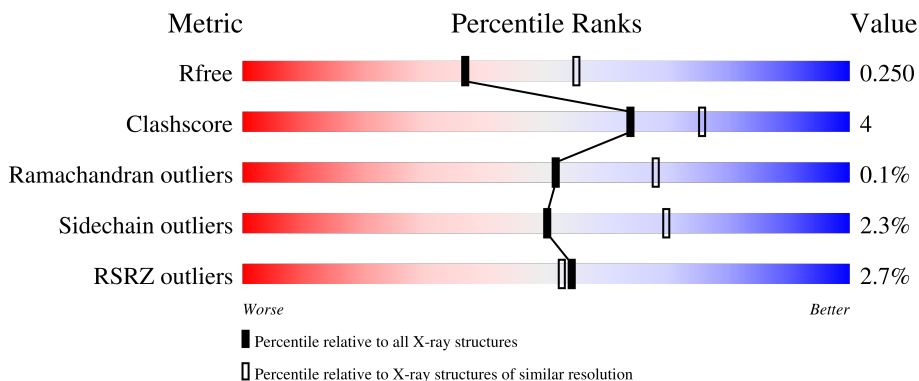|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.40 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.
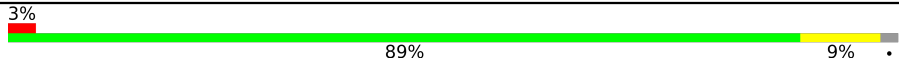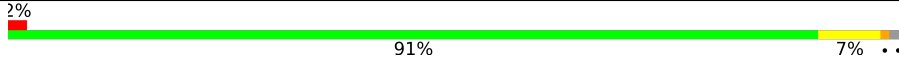
Metric	Percentile Ranks	Value
Rfree		0.250
Clashscore		4
Ramachandran outliers		0.1%
Sidechain outliers		2.3%
RSRZ outliers		2.7%

Worse ··· *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	258	2% / 92% / 7% •
1	B	258	3% / 93% / 5% •
1	C	258	2% / 88% / 11% •
1	D	258	4% / 86% / 13% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	258	3% ▮ 89% 9% ·
1	F	258	2% ▮ 91% 7% ··

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12306 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.
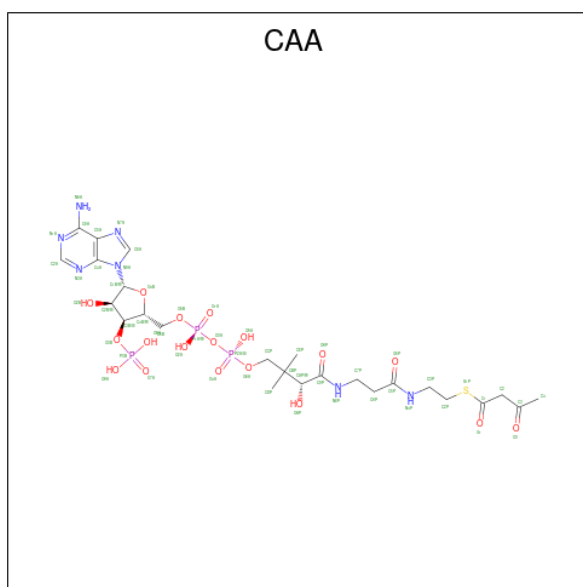
- Molecule 1 is a protein called enoyl-CoA hydratase echA8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	255	Total	C	N	O	S	0	2	0
			1872	1166	324	370	12			
1	B	255	Total	C	N	O	S	0	2	0
			1885	1175	324	374	12			
1	C	255	Total	C	N	O	S	0	0	0
			1875	1170	325	368	12			
1	D	255	Total	C	N	O	S	0	1	0
			1879	1172	326	369	12			
1	E	253	Total	C	N	O	S	0	0	0
			1852	1157	321	362	12			
1	F	254	Total	C	N	O	S	0	0	0
			1859	1160	321	366	12			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP P64016
B	0	SER	-	expression tag	UNP P64016
C	0	SER	-	expression tag	UNP P64016
D	0	SER	-	expression tag	UNP P64016
E	0	SER	-	expression tag	UNP P64016
F	0	SER	-	expression tag	UNP P64016

- Molecule 2 is ACETOACETYL-COENZYME A (three-letter code: CAA) (formula: $C_{25}H_{40}N_7O_{18}P_3S$).

WORLDWIDE PDB
PROTEIN DATA BANK

CAA

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	S	0	0
			54	25	7	18	3	1		
2	F	1	Total	C	N	O	P	S	0	0
			54	25	7	18	3	1		

• Molecule 3 is water.

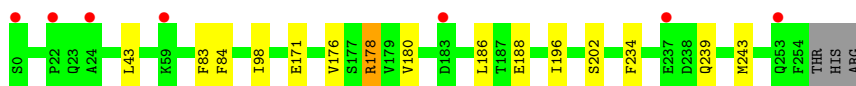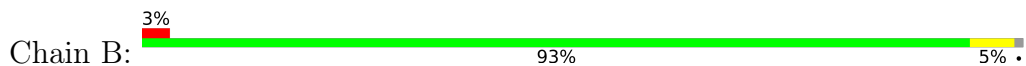| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	171	Total	O	0	0
			171	171		
3	B	173	Total	O	0	0
			173	173		
3	C	178	Total	O	0	0
			178	178		
3	D	141	Total	O	0	0
			141	141		
3	E	133	Total	O	0	0
			133	133		
3	F	180	Total	O	0	0
			180	180		

3 Residue-property plots ⓘ

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.
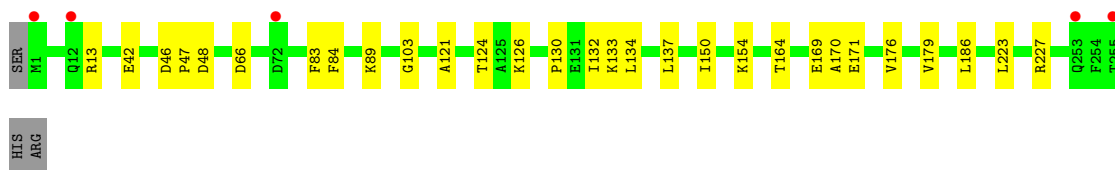
- Molecule 1: enoyl-CoA hydratase echA8

Chain A:

2%
92%
7%

SER M1 M33 D66 D72 F83 F84 A85 T86 W87 A105 C110 K133 L137 R213 A233 F234 Q239 I247 A251 P252 Q253 F254 T255 HIS ARG

- Molecule 1: enoyl-CoA hydratase echA8

Chain B:

3%
93%
5%

S0 P22 Q23 A24 L43 K59 F83 F84 I98 E171 V176 S177 R178 V179 V180 D183 L186 T187 E188 I196 S202 F234 E237 D238 Q239 M243 Q253 F254 THR HIS ARG

- Molecule 1: enoyl-CoA hydratase echA8

Chain C:

2%
88%
11%

SER M1 Q12 R13 E42 D46 P47 D48 D66 D72 F83 F84 K89 G103 A121 T124 A125 K126 P130 E131 I132 K133 L134 L137 I150 K154 T164 E169 A170 E171 V176 V179 L186 L223 R227 Q253 F254 T255

HIS ARG

- Molecule 1: enoyl-CoA hydratase echA8

Chain D:

4%
86%
13%

SER MET T2 E9 I16 R21 P22 Q23 M33 E42 D66 I67 K68 D77 F83 F84 R94 A99 A102 L119 A122 D123 V136 L137 R163 R172 R178 P181 A182 L185 L186 T187 T192 K208 R213 R227 R228

L229 A233 F234 Q239 I247 E248 K249 R250 A251 P252 Q253 F254 T255 H256 ARG

- Molecule 1: enoyl-CoA hydratase echA8

Chain E:

3%
89%
9%

• Molecule 1: enoyl-CoA hydratase echA8

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	98.96Å 134.70Å 134.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.42 – 2.40 25.99 – 2.40	Depositor EDS
% Data completeness (in resolution range)	98.4 (26.42-2.40) 98.4 (25.99-2.40)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	4.01 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.198 , 0.228 0.224 , 0.250	Depositor DCC
R_{free} test set	3522 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	29.8	Xtriage
Anisotropy	0.085	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 24.6	EDS
L-test for twinning [2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12306	wwPDB-VP
Average B, all atoms (Å^2)	6.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.16 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.0319e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CAA

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	0.39	0/1901	0.52	0/2571
1	B	0.40	0/1914	0.51	0/2589
1	C	0.40	0/1898	0.53	0/2567
1	D	0.44	1/1906 (0.1%)	0.56	2/2579 (0.1%)
1	E	0.38	0/1875	0.52	0/2538
1	F	0.39	0/1882	0.53	0/2548
All	All	0.40	1/11376 (0.0%)	0.53	2/15392 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	248	GLU	CG-CD	-6.32	1.42	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	248	GLU	CG-CD-OE2	5.55	129.40	118.30
1	D	248	GLU	CA-CB-CG	5.31	125.08	113.40

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1872	0	1853	13	0
1	B	1885	0	1878	6	0
1	C	1875	0	1873	19	0
1	D	1879	0	1867	19	0
1	E	1852	0	1842	20	0
1	F	1859	0	1845	17	0
2	C	54	0	36	6	0
2	F	54	0	36	9	0
3	A	171	0	0	3	0
3	B	173	0	0	1	0
3	C	178	0	0	4	0
3	D	141	0	0	6	0
3	E	133	0	0	8	0
3	F	180	0	0	0	0
All	All	12306	0	11230	95	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.

The worst 5 of 95 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:E:86:THR:HB	3:E:762:HOH:O	1.45	1.13
1:E:33:MET:SD	3:E:762:HOH:O	2.22	0.96
1:D:33:MET:SD	3:D:562:HOH:O	2.31	0.88
1:C:132:ILE:O	1:C:133:LYS:HG2	1.76	0.85
1:E:180:VAL:HG11	1:E:188:GLU:HG2	1.59	0.84

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	255/258 (99%)	250 (98%)	5 (2%)	0	100	100
1	B	255/258 (99%)	248 (97%)	7 (3%)	0	100	100
1	C	253/258 (98%)	247 (98%)	6 (2%)	0	100	100
1	D	254/258 (98%)	245 (96%)	8 (3%)	1 (0%)	34	48
1	E	251/258 (97%)	243 (97%)	8 (3%)	0	100	100
1	F	252/258 (98%)	243 (96%)	8 (3%)	1 (0%)	34	48
All	All	1520/1548 (98%)	1476 (97%)	42 (3%)	2 (0%)	51	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	123	ASP
1	F	136	VAL

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	186/193 (96%)	183 (98%)	3 (2%)	62	79
1	B	189/193 (98%)	184 (97%)	5 (3%)	46	66
1	C	186/193 (96%)	182 (98%)	4 (2%)	52	71
1	D	187/193 (97%)	182 (97%)	5 (3%)	44	65
1	E	182/193 (94%)	179 (98%)	3 (2%)	62	79
1	F	183/193 (95%)	177 (97%)	6 (3%)	38	57
All	All	1113/1158 (96%)	1087 (98%)	26 (2%)	50	70

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	172	ARG
1	E	137	LEU
1	F	186	LEU

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type
1	D	213	ARG
1	E	202	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	GLN
1	A	239	GLN
1	D	34	ASN

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 ⓘ

2 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	CAA	F	300	-	47,56,56	4.57	22 (46%)	60,83,83	2.41	6 (10%)
2	CAA	C	300	-	47,56,56	4.55	21 (44%)	60,83,83	2.05	5 (8%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CAA	F	300	-	-	16/50/71/71	0/3/3/3
2	CAA	C	300	-	-	17/50/71/71	0/3/3/3

The worst 5 of 43 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	300	CAA	O1-C1	13.18	1.42	1.21
2	F	300	CAA	C2A-N3A	12.80	1.52	1.32
2	C	300	CAA	O1-C1	12.68	1.41	1.21
2	C	300	CAA	C2A-N3A	12.51	1.52	1.32
2	C	300	CAA	C2A-N1A	11.43	1.55	1.33

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	300	CAA	O1-C1-S1P	-15.33	102.70	122.61
2	C	300	CAA	O1-C1-S1P	-11.37	107.85	122.61
2	F	300	CAA	N3A-C2A-N1A	-6.11	119.13	128.68
2	C	300	CAA	N3A-C2A-N1A	-6.04	119.24	128.68
2	C	300	CAA	O1-C1-C2	-5.10	114.52	123.35

There are no chirality outliers.
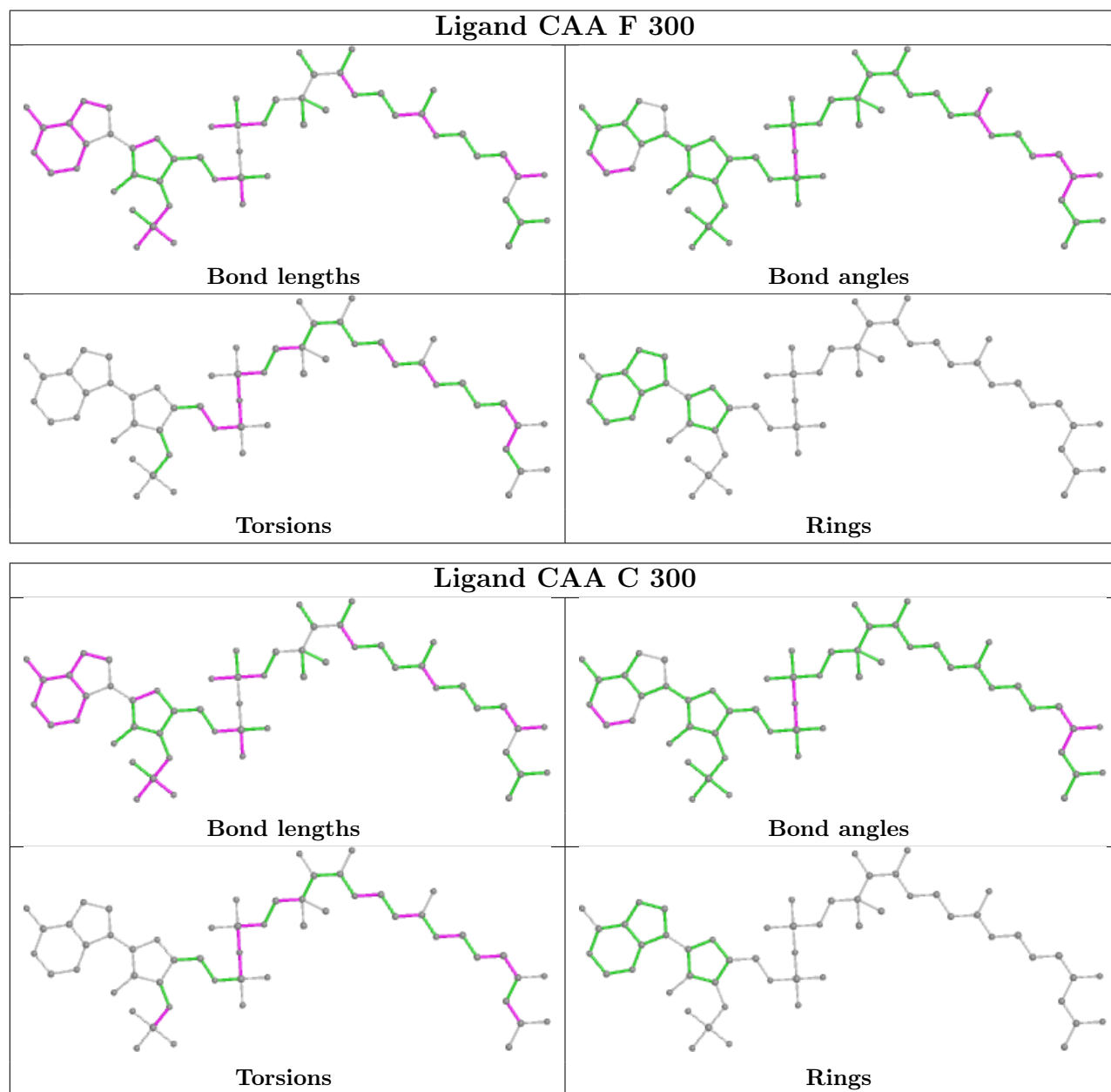
5 of 33 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	300	CAA	CCP-O6A-P2A-O4A
2	C	300	CAA	CCP-O6A-P2A-O5A
2	C	300	CAA	CDP-CBP-CCP-O6A
2	C	300	CAA	CEP-CBP-CCP-O6A
2	C	300	CAA	CAP-CBP-CCP-O6A

There are no ring outliers.

2 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	300	CAA	9	0
2	C	300	CAA	6	0

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

Ligand CAA F 300

Bond lengths   Bond angles

Torsions   Rings

Ligand CAA C 300

Bond lengths   Bond angles

Torsions   Rings

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	255/258 (98%)	-0.03	6 (2%)	59	57	2, 2, 18, 46	0
1	B	255/258 (98%)	0.02	7 (2%)	54	52	2, 2, 18, 34	0
1	C	255/258 (98%)	-0.05	5 (1%)	65	63	2, 2, 16, 33	0
1	D	255/258 (98%)	0.07	11 (4%)	35	33	2, 4, 24, 43	0
1	E	253/258 (98%)	0.04	8 (3%)	47	46	2, 5, 21, 45	0
1	F	254/258 (98%)	-0.13	4 (1%)	72	70	2, 2, 15, 27	0
All	All	1527/1548 (98%)	-0.01	41 (2%)	54	52	2, 2, 19, 46	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	255	THR	4.7
1	D	251	ALA	4.5
1	D	247	ILE	4.3
1	D	256	HIS	4.3
1	D	2	THR	3.7

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

There are no non-standard protein/DNA/RNA residues in this entry.
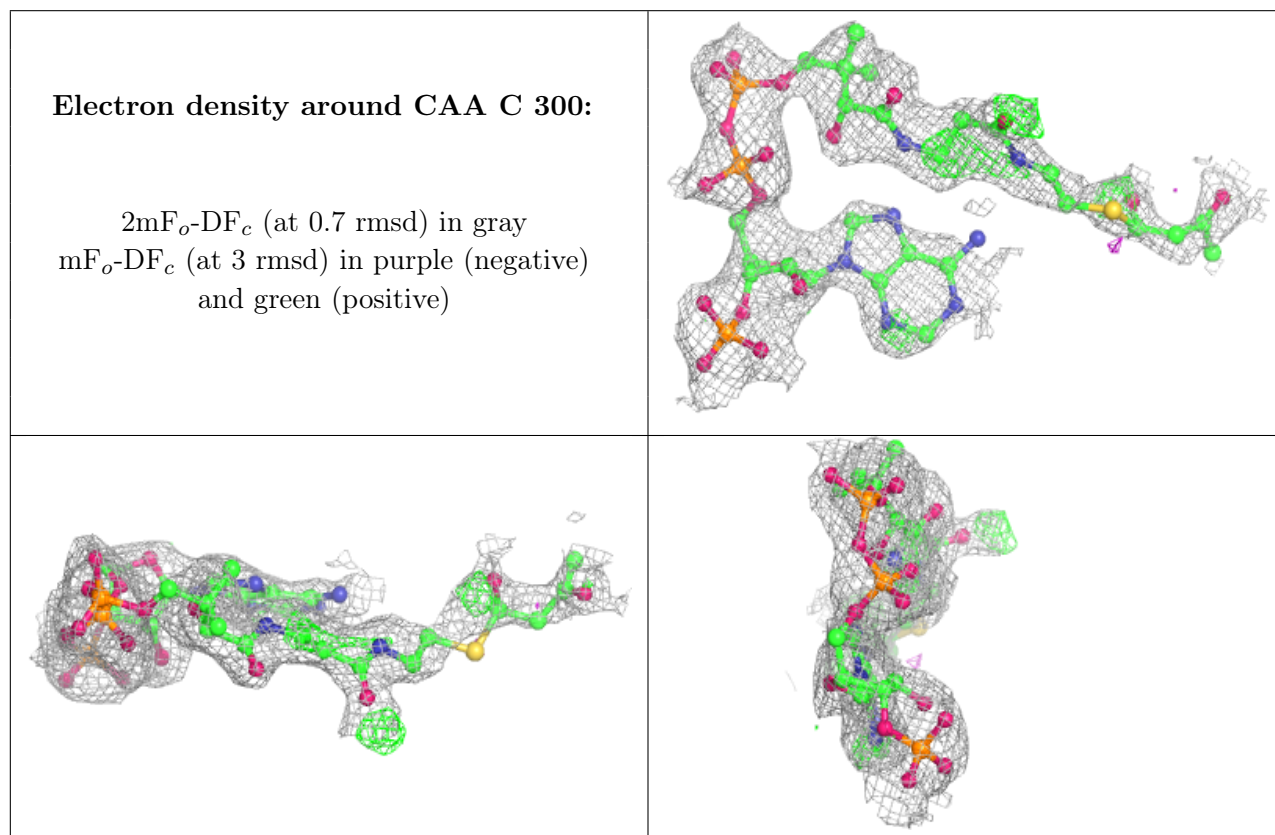
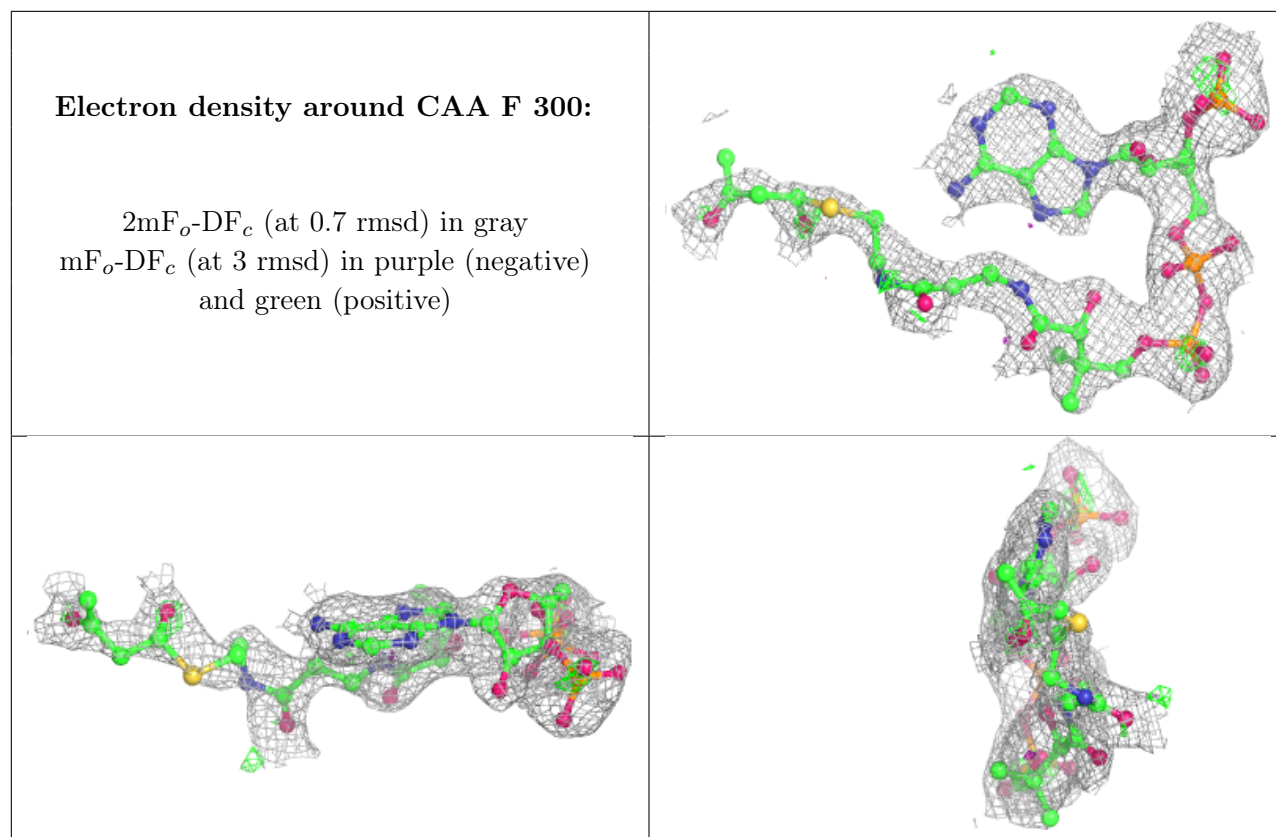## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	CAA	C	300	54/54	0.77	0.27	49,57,64,100	17
2	CAA	F	300	54/54	0.84	0.23	44,53,57,65	20

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 CAA C 300:

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

Electron density around CAA F 300:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.