



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:35 PM UTC

PDB ID : 8GQN / pdb_00008gqn
Title : X-ray structure of thiolase with CoA
Authors : Hong, J.; Son, H.F.; Kim, K.J.
Deposited on : 2022-08-30
Resolution : 2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

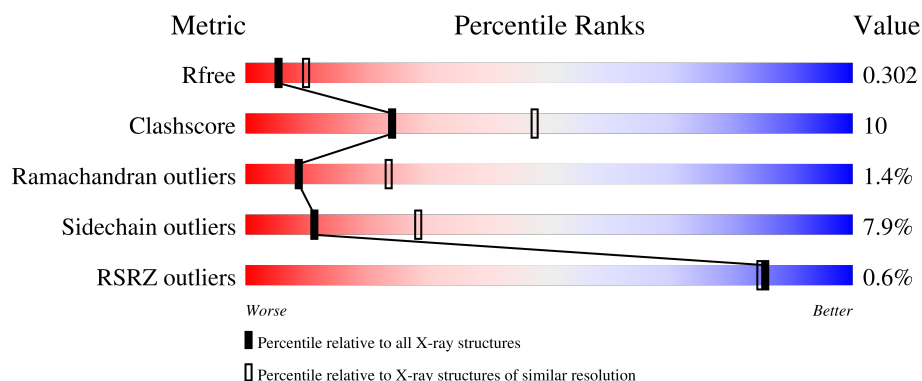
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

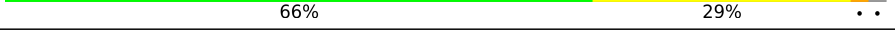
The reported resolution of this entry is 2.70 Å.

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}	180053	3538 (2.70-2.70)
Clashscore	190562	3843 (2.70-2.70)
Ramachandran outliers	187476	3778 (2.70-2.70)
Sidechain outliers	187428	3778 (2.70-2.70)
RSRZ outliers	180081	3538 (2.70-2.70)

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 $\leq 5\%$. 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	
1	B	399	
1	C	399	
1	D	399	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11502 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 Thiolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	0	0	0
			2822	1747	508	546	21			
1	B	390	Total	C	N	O	S	0	0	0
			2822	1747	508	546	21			
1	C	388	Total	C	N	O	S	0	0	0
			2802	1735	502	544	21			
1	D	390	Total	C	N	O	S	0	0	0
			2822	1747	508	546	21			

There are 32 discrepancies between the modelled and reference sequences:

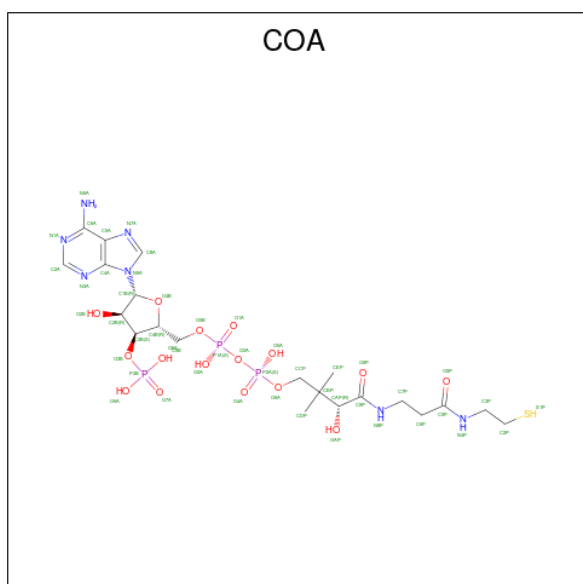
Chain	Residue	Modelled	Actual	Comment	Reference
A	392	LEU	-	expression tag	UNP G3XD40
A	393	GLU	-	expression tag	UNP G3XD40
A	394	HIS	-	expression tag	UNP G3XD40
A	395	HIS	-	expression tag	UNP G3XD40
A	396	HIS	-	expression tag	UNP G3XD40
A	397	HIS	-	expression tag	UNP G3XD40
A	398	HIS	-	expression tag	UNP G3XD40
A	399	HIS	-	expression tag	UNP G3XD40
B	392	LEU	-	expression tag	UNP G3XD40
B	393	GLU	-	expression tag	UNP G3XD40
B	394	HIS	-	expression tag	UNP G3XD40
B	395	HIS	-	expression tag	UNP G3XD40
B	396	HIS	-	expression tag	UNP G3XD40
B	397	HIS	-	expression tag	UNP G3XD40
B	398	HIS	-	expression tag	UNP G3XD40
B	399	HIS	-	expression tag	UNP G3XD40
C	392	LEU	-	expression tag	UNP G3XD40
C	393	GLU	-	expression tag	UNP G3XD40
C	394	HIS	-	expression tag	UNP G3XD40
C	395	HIS	-	expression tag	UNP G3XD40
C	396	HIS	-	expression tag	UNP G3XD40

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	397	HIS	-	expression tag	UNP G3XD40
C	398	HIS	-	expression tag	UNP G3XD40
C	399	HIS	-	expression tag	UNP G3XD40
D	392	LEU	-	expression tag	UNP G3XD40
D	393	GLU	-	expression tag	UNP G3XD40
D	394	HIS	-	expression tag	UNP G3XD40
D	395	HIS	-	expression tag	UNP G3XD40
D	396	HIS	-	expression tag	UNP G3XD40
D	397	HIS	-	expression tag	UNP G3XD40
D	398	HIS	-	expression tag	UNP G3XD40
D	399	HIS	-	expression tag	UNP G3XD40

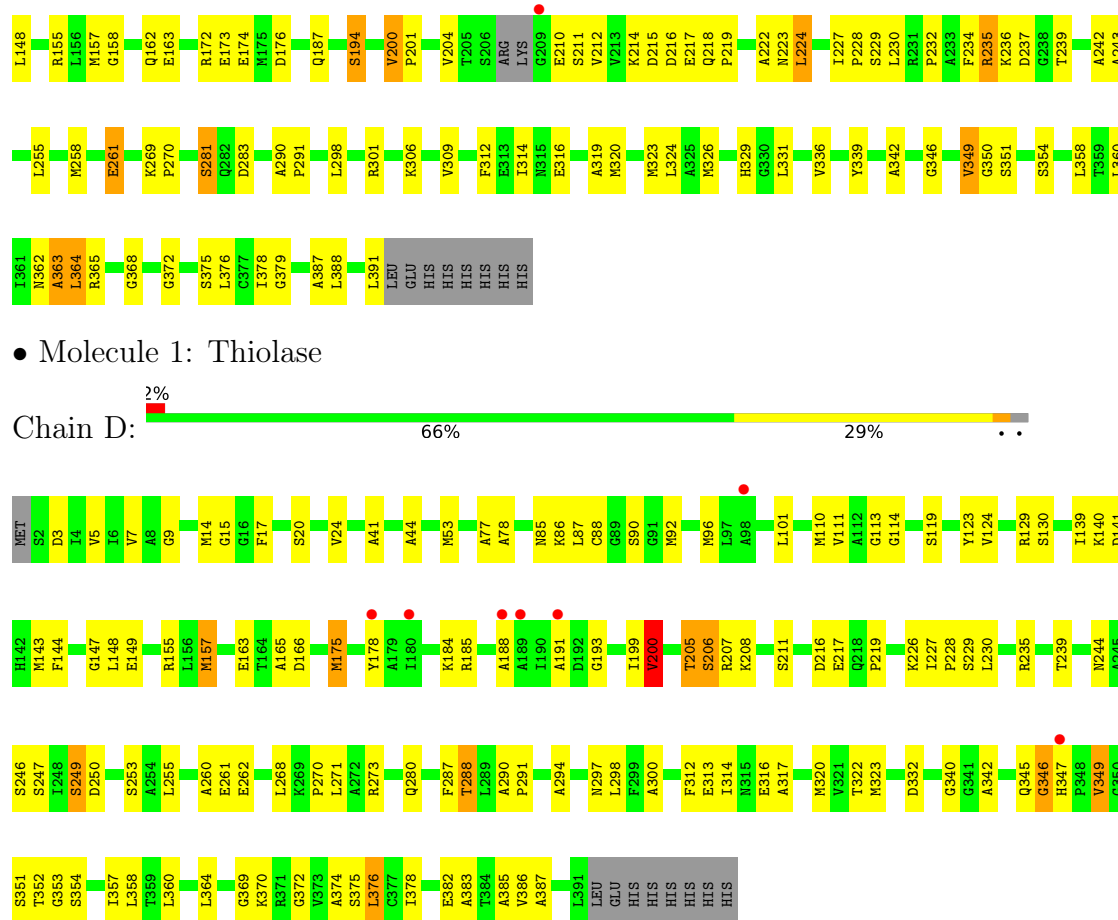
- Molecule 2 is COENZYME A (CCD ID: COA) (formula: $C_{21}H_{36}N_7O_{16}P_3S$).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	11	Total 11	O 11	0	0
3	B	10	Total 10	O 10	0	0
3	C	9	Total 9	O 9	0	0
3	D	12	Total 12	O 12	0	0



● Molecule 1: Thiolase

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	73.84Å 94.18Å 113.25Å 90.00° 107.47° 90.00°	Depositor
Resolution (Å)	34.38 – 2.70 34.38 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.0 (34.38-2.70) 97.4 (34.38-2.70)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.02 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.209 , 0.301 0.214 , 0.302	Depositor DCC
R_{free} test set	1963 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å ²)	37.6	Xtriage
Anisotropy	0.158	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 33.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.048 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	11502	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 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	1.13	0/2855	1.58	10/3854 (0.3%)
1	B	1.11	0/2855	1.57	4/3854 (0.1%)
1	C	1.13	0/2834	1.61	12/3826 (0.3%)
1	D	1.16	1/2855 (0.0%)	1.63	13/3854 (0.3%)
All	All	1.13	1/11399 (0.0%)	1.60	39/15388 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	200	VAL	N-CA	5.15	1.49	1.45

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	269	LYS	CB-CA-C	8.33	118.83	110.33
1	D	14	MET	CA-C-N	7.35	127.60	120.98
1	D	14	MET	C-N-CA	7.35	127.60	120.98
1	D	44	ALA	CA-C-N	6.68	129.12	120.44
1	D	44	ALA	C-N-CA	6.68	129.12	120.44

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	2822	0	2869	45	0
1	B	2822	0	2869	67	0
1	C	2802	0	2842	56	0
1	D	2822	0	2869	68	0
2	A	48	0	32	3	0
2	B	48	0	32	4	0
2	C	48	0	32	3	0
2	D	48	0	32	2	0
3	A	11	0	0	0	0
3	B	10	0	0	2	0
3	C	9	0	0	0	0
3	D	12	0	0	1	0
All	All	11502	0	11577	231	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 10.

The worst 5 of 231 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:B:89:GLY:HA2	1:B:376:LEU:HD11	1.64	0.78
1:C:148:LEU:O	1:C:157:MET:HG2	1.87	0.75
1:A:50:GLU:HB3	1:A:110:MET:HE3	1.70	0.72
1:B:126:GLU:OE2	1:B:140:LYS:HG3	1.94	0.67
1:B:253:SER:OG	1:B:354:SER:OG	2.12	0.67

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	388/399 (97%)	343 (88%)	42 (11%)	3 (1%)	16 37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	388/399 (97%)	350 (90%)	33 (8%)	5 (1%)	9	25
1	C	384/399 (96%)	341 (89%)	36 (9%)	7 (2%)	6	18
1	D	388/399 (97%)	330 (85%)	51 (13%)	7 (2%)	6	18
All	All	1548/1596 (97%)	1364 (88%)	162 (10%)	22 (1%)	9	23

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	133	ARG
1	C	232	PRO
1	C	319	ALA
1	C	364	LEU
1	D	206	SER

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	286/295 (97%)	269 (94%)	17 (6%)	18	42
1	B	286/295 (97%)	258 (90%)	28 (10%)	7	19
1	C	284/295 (96%)	261 (92%)	23 (8%)	11	27
1	D	286/295 (97%)	264 (92%)	22 (8%)	12	30
All	All	1142/1180 (97%)	1052 (92%)	90 (8%)	11	28

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	236	LYS
1	D	130	SER
1	C	261	GLU
1	C	351	SER
1	D	207	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	162	GLN
1	D	18	GLN
1	D	280	GLN
1	D	108	GLN
1	C	333	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

4 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	COA	D	401	-	47,50,50	0.66	1 (2%)	69,75,75	0.68	1 (1%)
2	COA	A	401	-	47,50,50	0.64	1 (2%)	69,75,75	0.98	4 (5%)
2	COA	C	401	-	47,50,50	0.51	0	69,75,75	0.94	1 (1%)
2	COA	B	401	-	47,50,50	0.57	0	69,75,75	0.87	2 (2%)

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	COA	D	401	-	-	12/48/64/64	0/3/3/3
2	COA	A	401	-	-	15/48/64/64	0/3/3/3
2	COA	C	401	-	-	7/48/64/64	0/3/3/3
2	COA	B	401	-	-	14/48/64/64	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	COA	P3B-O3B	3.08	1.64	1.59
2	D	401	COA	P3B-O3B	2.74	1.64	1.59

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	401	COA	P3B-O3B-C3B	-4.51	111.39	123.43
2	A	401	COA	P3B-O3B-C3B	-3.75	113.41	123.43
2	A	401	COA	O2B-C2B-C3B	3.49	120.95	111.19
2	B	401	COA	P3B-O3B-C3B	-3.46	114.18	123.43
2	A	401	COA	C2B-C3B-C4B	-2.59	98.71	103.24

There are no chirality outliers.

5 of 48 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	COA	C5B-O5B-P1A-O1A
2	A	401	COA	C5B-O5B-P1A-O2A
2	A	401	COA	C5B-O5B-P1A-O3A
2	A	401	COA	P2A-O3A-P1A-O5B
2	A	401	COA	CCP-O6A-P2A-O3A

There are no ring outliers.

4 monomers are involved in 12 short contacts:

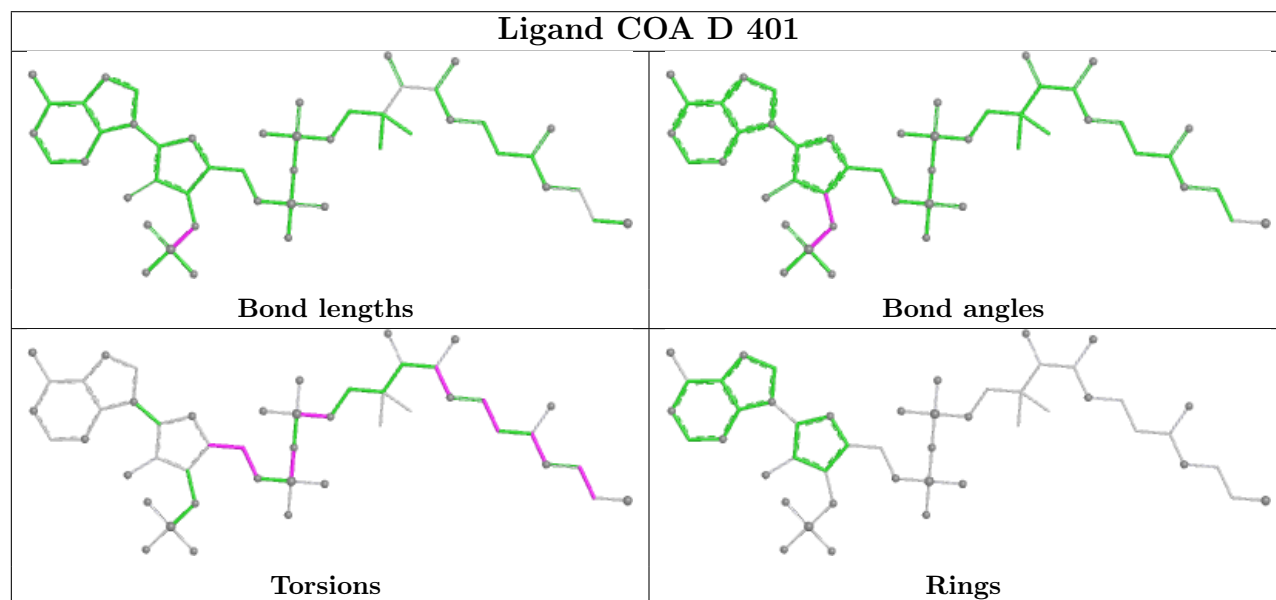
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	401	COA	2	0
2	A	401	COA	3	0
2	C	401	COA	3	0

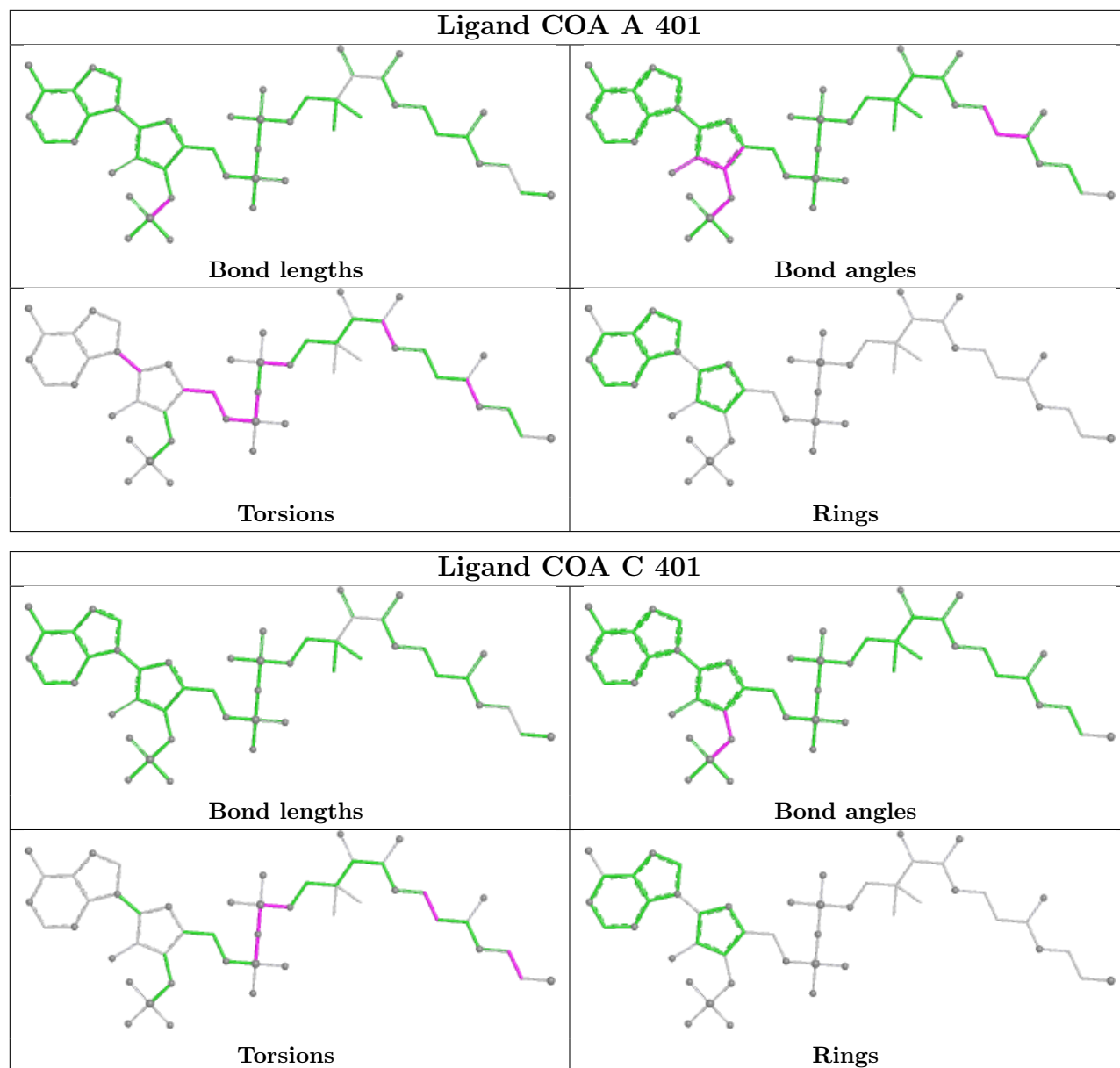
Continued on next page...

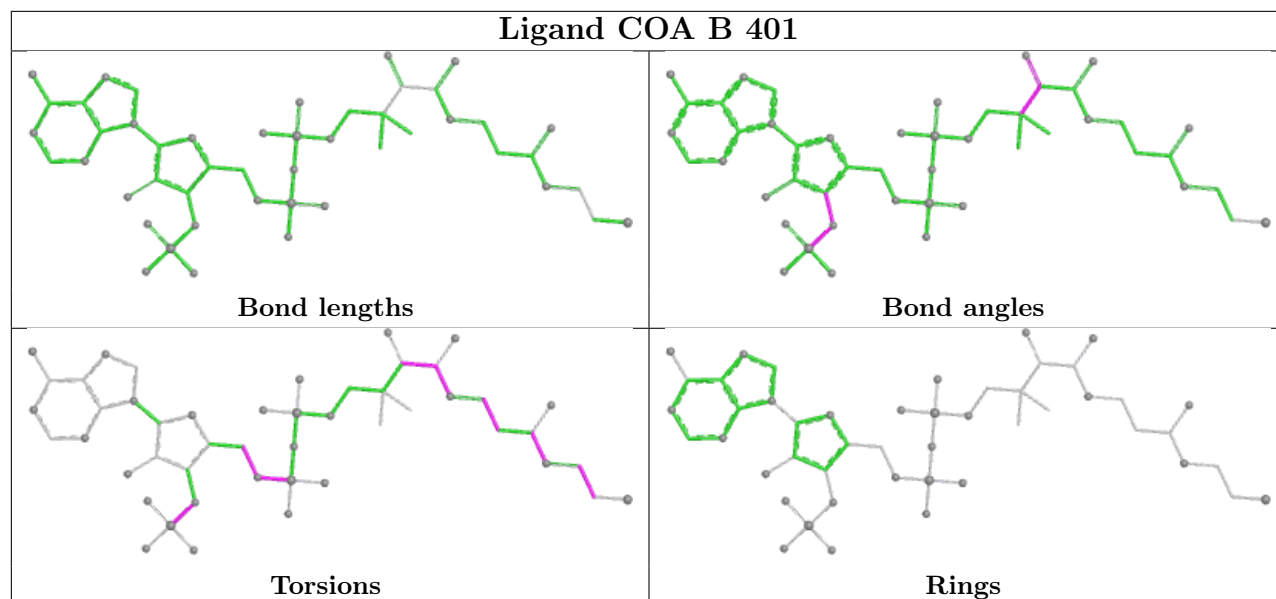
Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	COA	4	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](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	390/399 (97%)	-0.12	0 100 100	18, 33, 52, 71	0
1	B	390/399 (97%)	-0.02	1 (0%) 90 89	20, 35, 59, 92	0
1	C	388/399 (97%)	0.07	1 (0%) 90 89	19, 39, 58, 82	0
1	D	390/399 (97%)	0.35	7 (1%) 67 65	22, 45, 73, 91	0
All	All	1558/1596 (97%)	0.07	9 (0%) 85 85	18, 37, 64, 92	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	180	ILE	4.4
1	D	189	ALA	3.0
1	C	209	GLY	2.5
1	D	98	ALA	2.3
1	D	191	ALA	2.2

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

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

6.3 Carbohydrates [i](#)

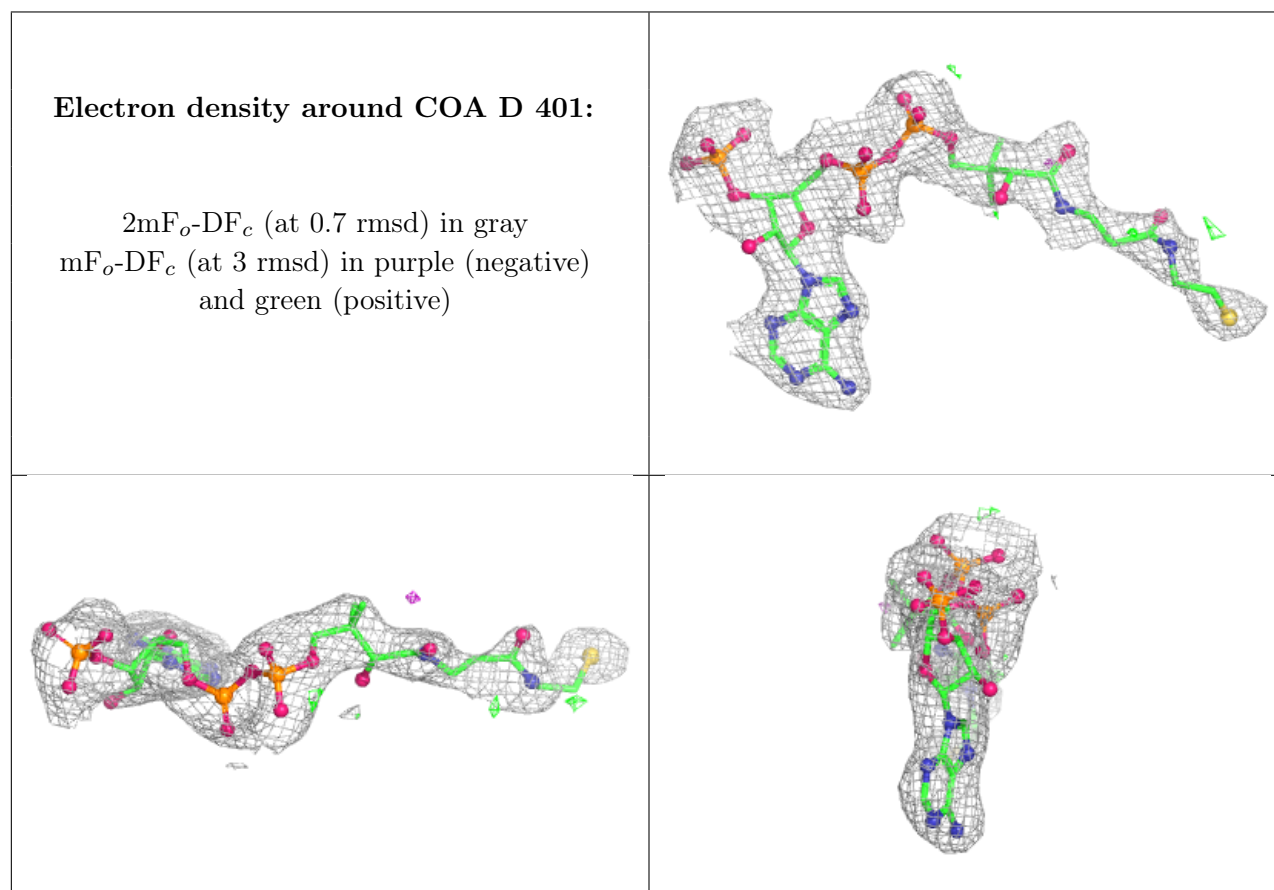
There are no oligosaccharides in this entry.

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.

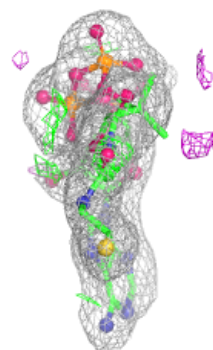
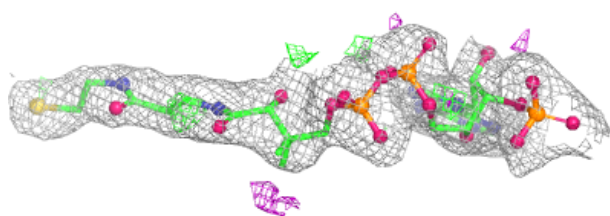
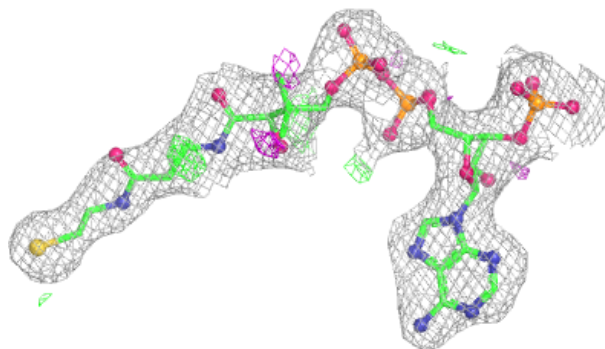
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	COA	D	401	48/48	0.90	0.11	46,63,79,87	0
2	COA	A	401	48/48	0.91	0.09	27,42,52,55	0
2	COA	C	401	48/48	0.92	0.10	32,48,58,64	0
2	COA	B	401	48/48	0.93	0.09	33,42,58,70	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

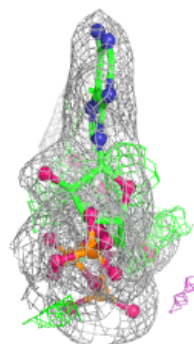
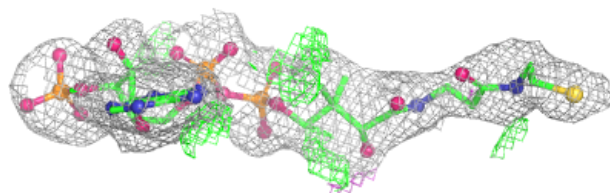
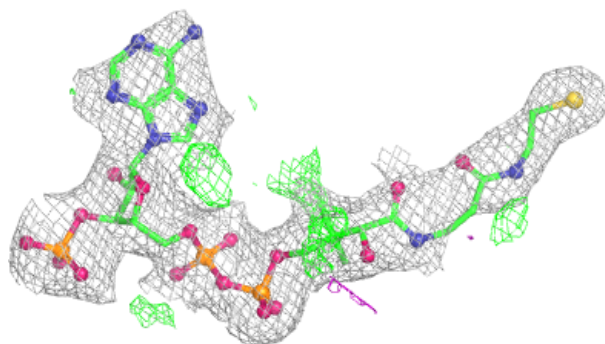


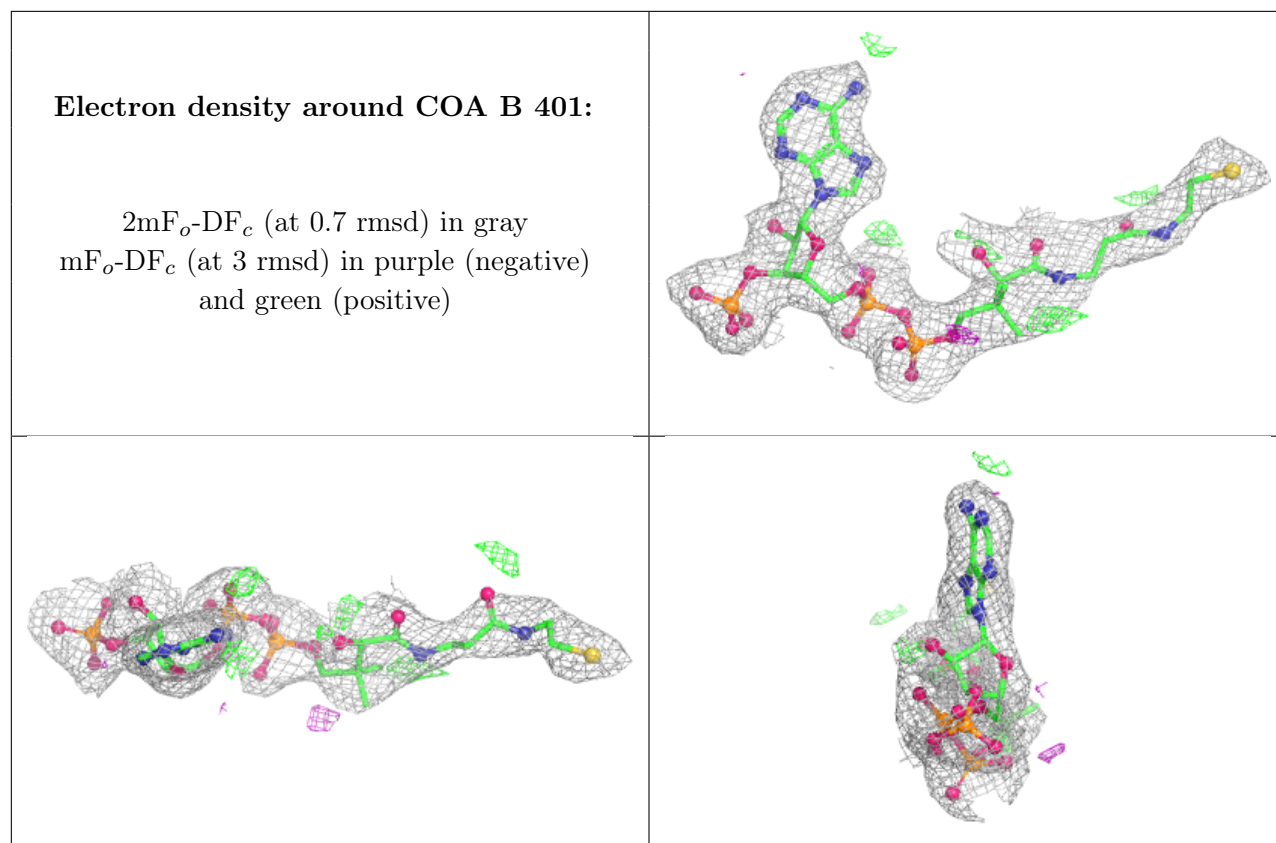
Electron density around COA A 401:

$2mF_o-DF_c$ (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)





6.5 Other polymers [i](#)

There are no such residues in this entry.