W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 12, 2024 – 02:42 AM EDT

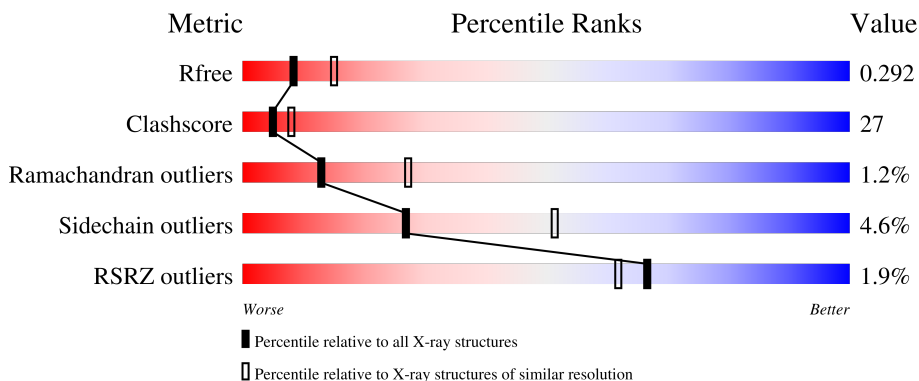|		
PDB ID	:	2NYG
Title	:	Crystal structure of YokD protein from Bacillus subtilis
Authors	:	Madegowda, M.; Eswaramoorthy, S.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2006-11-20
Resolution	:	2.60 Å(reported)

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 2.60 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		0.292
Clashscore		27
Ramachandran outliers		1.2%
Sidechain outliers		4.6%
RSRZ outliers		1.9%

Worse ◼ Percentile relative to all X-ray structures
Better ▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	273	62% 33% • •
1	B	273	% 46% 42% • 8%
1	C	273	58% 39% • •
1	D	273	64% 31% • • •
1	E	273	% 56% 39% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	273	8% ▮ 44% / 42% • 9%

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
2	COA	A	301	X	-	-	-
2	COA	B	302	X	-	-	-
2	COA	C	303	X	-	-	-
2	COA	D	304	X	-	-	-
2	COA	E	305	X	-	-	-

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 12534 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 YokD protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	270	Total	C	N	O	S	0	0	0
			2119	1359	358	395	7			
1	B	251	Total	C	N	O	S	0	0	0
			1942	1250	322	364	6			
1	C	270	Total	C	N	O	S	0	0	0
			2104	1351	348	398	7			
1	D	268	Total	C	N	O	S	0	0	0
			2089	1342	344	396	7			
1	E	260	Total	C	N	O	S	0	0	0
			2018	1295	341	375	7			
1	F	248	Total	C	N	O	S	0	0	0
			1829	1177	307	339	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	cloning artifact	UNP O32003
A	2	LEU	-	cloning artifact	UNP O32003
B	1	SER	-	cloning artifact	UNP O32003
B	2	LEU	-	cloning artifact	UNP O32003
C	1	SER	-	cloning artifact	UNP O32003
C	2	LEU	-	cloning artifact	UNP O32003
D	1	SER	-	cloning artifact	UNP O32003
D	2	LEU	-	cloning artifact	UNP O32003
E	1	SER	-	cloning artifact	UNP O32003
E	2	LEU	-	cloning artifact	UNP O32003
F	1	SER	-	cloning artifact	UNP O32003
F	2	LEU	-	cloning artifact	UNP O32003

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

COA

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 42	C 18	N 6	O 15	P 3	0	0
2	B	1	Total 42	C 18	N 6	O 15	P 3	0	0
2	C	1	Total 42	C 18	N 6	O 15	P 3	0	0
2	D	1	Total 42	C 18	N 6	O 15	P 3	0	0
2	E	1	Total 42	C 18	N 6	O 15	P 3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	39	Total 39	O 39	0	0
3	B	28	Total 28	O 28	0	0
3	C	41	Total 41	O 41	0	0
3	D	34	Total 34	O 34	0	0
3	E	43	Total 43	O 43	0	0
3	F	38	Total 38	O 38	0	0

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.
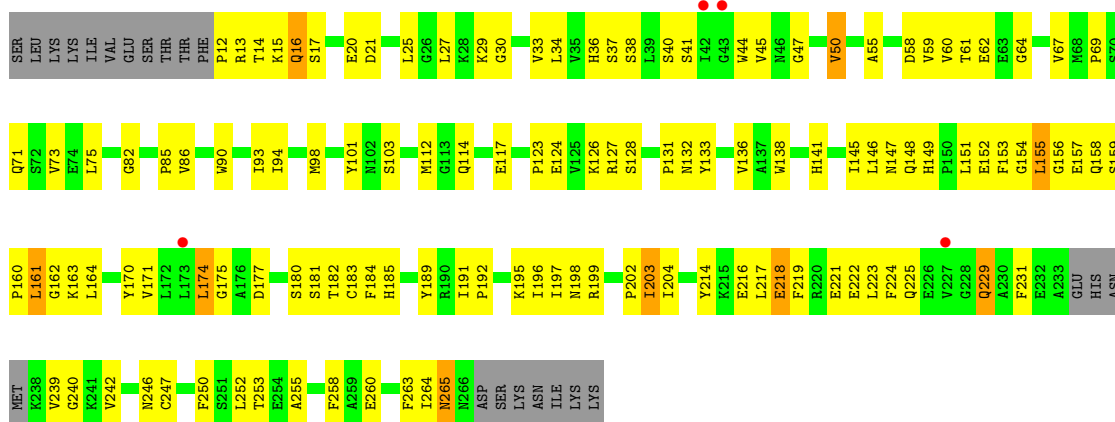
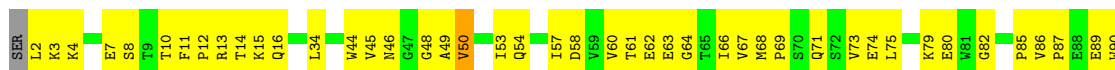• Molecule 1: YokD protein

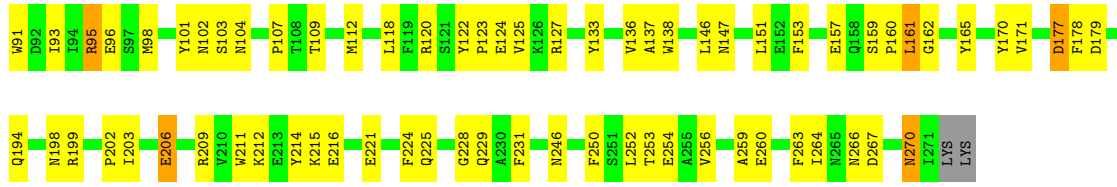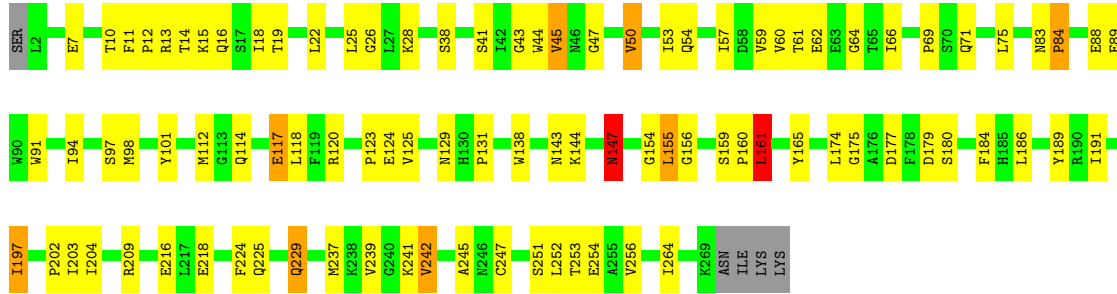Chain A:
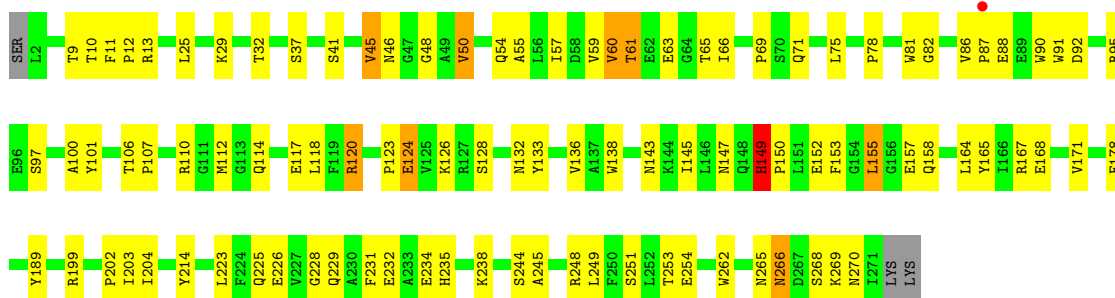
• Molecule 1: YokD protein
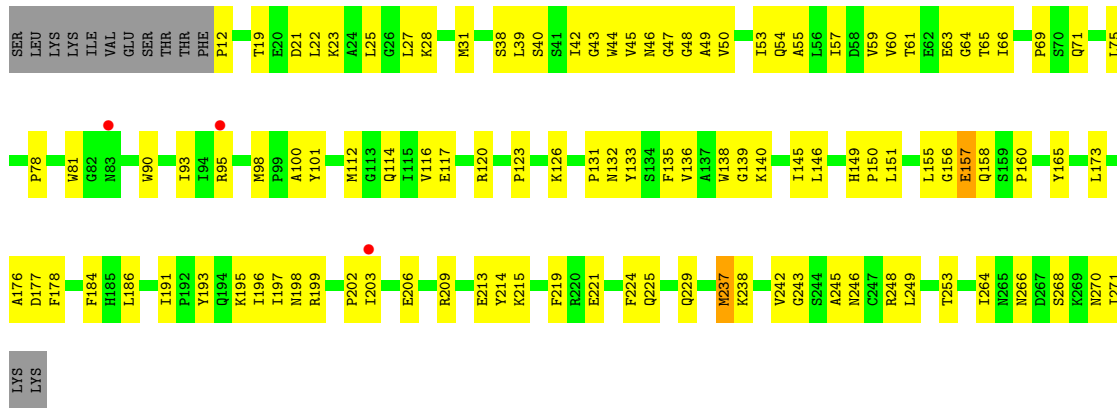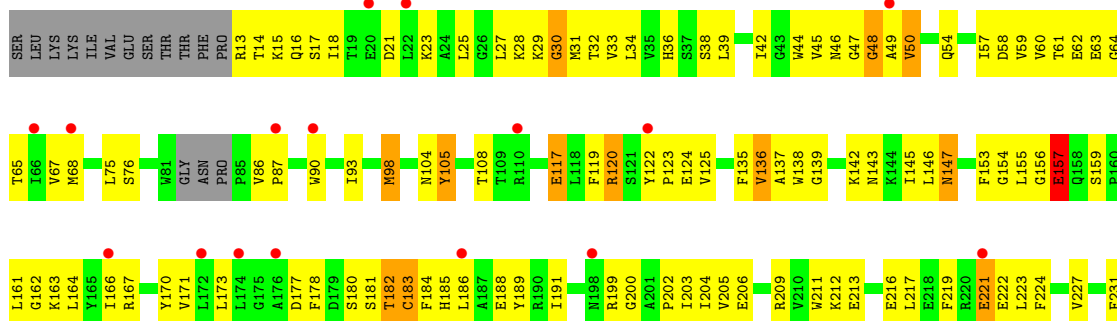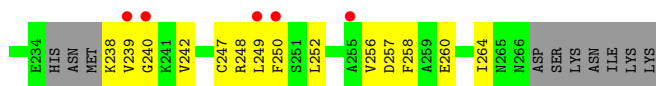
Chain B:

• Molecule 1: YokD protein

Chain C:

• Molecule 1: YokD protein

Chain D:

• Molecule 1: YokD protein

Chain E:

• Molecule 1: YokD protein

Chain F:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	73.02Å　　132.39Å　　185.80Å 90.00°　　90.00°　　90.00°	Depositor
Resolution (Å)	49.04　–　2.60 49.04　–　2.37	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.04-2.60) 96.6 (49.04-2.37)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.03 (at 2.37Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.257　,　0.296 0.251　,　0.292	Depositor DCC
R_{free} test set	1434 reflections (2.01%)	wwPDB-VP
Wilson B-factor (Å^2)	48.5	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 45.1	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	12534	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.84% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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: 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	0.46	0/2172	0.70	0/2951
1	B	0.39	0/1993	0.66	0/2716
1	C	0.45	0/2158	0.72	1/2938 (0.0%)
1	D	0.45	0/2143	0.73	3/2917 (0.1%)
1	E	0.42	0/2069	0.70	1/2812 (0.0%)
1	F	0.39	0/1872	0.63	0/2555
All	All	0.43	0/12407	0.69	5/16889 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	95	ARG	NE-CZ-NH1	-6.88	116.86	120.30
1	D	155	LEU	N-CA-C	-6.57	93.25	111.00
1	D	161	LEU	CA-CB-CG	6.40	130.03	115.30
1	D	242	VAL	N-CA-C	-5.40	96.43	111.00
1	E	243	GLY	N-CA-C	-5.19	100.13	113.10

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.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2119	0	2067	90	0
1	B	1942	0	1846	132	0
1	C	2104	0	2015	105	0
1	D	2089	0	2002	97	0
1	E	2018	0	1953	111	0
1	F	1829	0	1699	131	0
2	A	42	0	24	6	0
2	B	42	0	24	6	0
2	C	42	0	24	1	0
2	D	42	0	24	5	0
2	E	42	0	24	8	0
3	A	39	0	0	7	0
3	B	28	0	0	4	0
3	C	41	0	0	6	0
3	D	34	0	0	3	0
3	E	43	0	0	7	0
3	F	38	0	0	6	0
All	All	12534	0	11702	655	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 27.

The worst 5 of 655 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:75:LEU:HA	1:C:98:MET:HE3	1.27	1.16
1:D:14:THR:HG22	1:D:16:GLN:H	1.18	1.07
1:E:71:GLN:H	1:E:112:MET:HE2	1.19	1.03
1:A:61:THR:HG23	1:A:63:GLU:H	1.25	1.01
1:A:71:GLN:H	1:A:112:MET:HE2	1.25	1.01

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	268/273 (98%)	243 (91%)	21 (8%)	4 (2%)	10	21
1	B	247/273 (90%)	204 (83%)	41 (17%)	2 (1%)	19	39
1	C	268/273 (98%)	247 (92%)	19 (7%)	2 (1%)	22	43
1	D	266/273 (97%)	248 (93%)	17 (6%)	1 (0%)	34	57
1	E	258/273 (94%)	226 (88%)	29 (11%)	3 (1%)	13	27
1	F	242/273 (89%)	198 (82%)	37 (15%)	7 (3%)	4	7
All	All	1549/1638 (95%)	1366 (88%)	164 (11%)	19 (1%)	13	27

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	182	THR
1	C	103	SER
1	E	206	GLU
1	F	157	GLU
1	A	155	LEU

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	227/240 (95%)	218 (96%)	9 (4%)	31	57
1	B	202/240 (84%)	190 (94%)	12 (6%)	19	39
1	C	223/240 (93%)	214 (96%)	9 (4%)	31	57
1	D	222/240 (92%)	208 (94%)	14 (6%)	18	36
1	E	213/240 (89%)	207 (97%)	6 (3%)	43	69
1	F	179/240 (75%)	171 (96%)	8 (4%)	27	52
All	All	1266/1440 (88%)	1208 (95%)	58 (5%)	27	51

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	246	ASN
1	F	157	GLU
1	D	84	PRO
1	F	136	VAL
1	E	271	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	54	GLN
1	E	246	ASN
1	E	104	ASN
1	E	141	HIS
1	E	270	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 ⓘ

5 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	304	-	38,44,50	1.94	5 (13%)	49,68,75	1.95	9 (18%)
2	COA	C	303	-	38,44,50	1.80	5 (13%)	49,68,75	1.88	8 (16%)
2	COA	B	302	-	38,44,50	2.28	9 (23%)	49,68,75	1.81	6 (12%)
2	COA	A	301	-	38,44,50	1.92	7 (18%)	49,68,75	2.05	11 (22%)
2	COA	E	305	-	38,44,50	2.08	6 (15%)	49,68,75	1.88	7 (14%)

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	304	-	1/1/10/13	17/37/57/64	0/3/3/3
2	COA	C	303	-	1/1/10/13	20/37/57/64	0/3/3/3
2	COA	B	302	-	1/1/10/13	15/37/57/64	0/3/3/3
2	COA	A	301	-	1/1/10/13	17/37/57/64	0/3/3/3
2	COA	E	305	-	1/1/10/13	19/37/57/64	0/3/3/3

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	305	COA	O9P-C9P	8.69	1.40	1.23
2	B	302	COA	O9P-C9P	7.72	1.38	1.23
2	D	304	COA	O9P-C9P	7.26	1.37	1.23
2	C	303	COA	O9P-C9P	6.93	1.36	1.23
2	A	301	COA	O9P-C9P	6.92	1.36	1.23

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	303	COA	N3A-C2A-N1A	-7.42	118.60	128.67
2	A	301	COA	N3A-C2A-N1A	-7.15	118.97	128.67
2	E	305	COA	O4B-C1B-N9A	7.04	118.08	108.75
2	D	304	COA	O4B-C1B-N9A	6.98	118.01	108.75
2	E	305	COA	N3A-C2A-N1A	-6.82	119.41	128.67

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	301	COA	CAP
2	B	302	COA	CAP
2	C	303	COA	CAP
2	D	304	COA	CAP
2	E	305	COA	CAP

5 of 88 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	301	COA	CCP-O6A-P2A-O3A
2	A	301	COA	CCP-O6A-P2A-O5A
2	A	301	COA	CDP-CBP-CCP-O6A
2	A	301	COA	CEP-CBP-CCP-O6A
2	A	301	COA	CAP-CBP-CCP-O6A

There are no ring outliers.
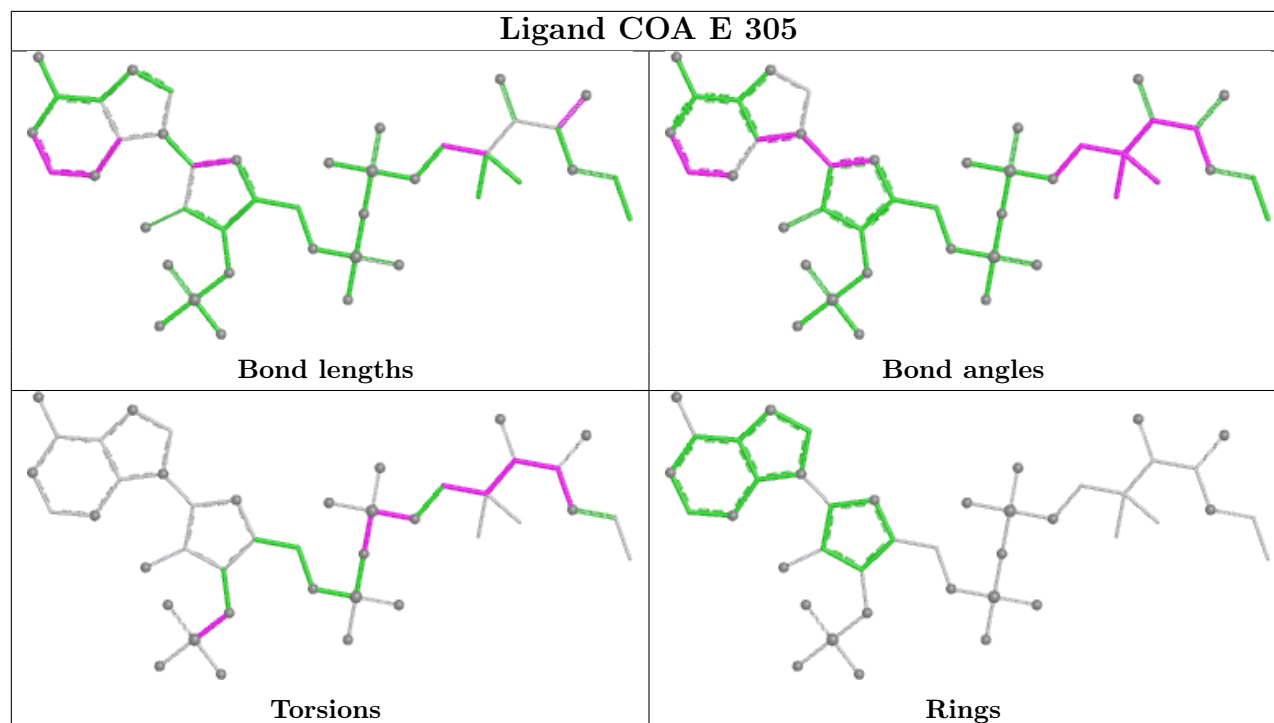
5 monomers are involved in 26 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	304	COA	5	0
2	C	303	COA	1	0
2	B	302	COA	6	0
2	A	301	COA	6	0
2	E	305	COA	8	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 COA D 304

Bond lengths

Bond angles

Torsions

Rings

Ligand COA C 303

Bond lengths

Bond angles

Torsions

Rings

Ligand COA B 302

Bond lengths

Bond angles

Torsions

Rings

Ligand COA A 301

Bond lengths

Bond angles

Torsions

Rings

Ligand COA E 305

Bond lengths

Bond angles

Torsions

Rings
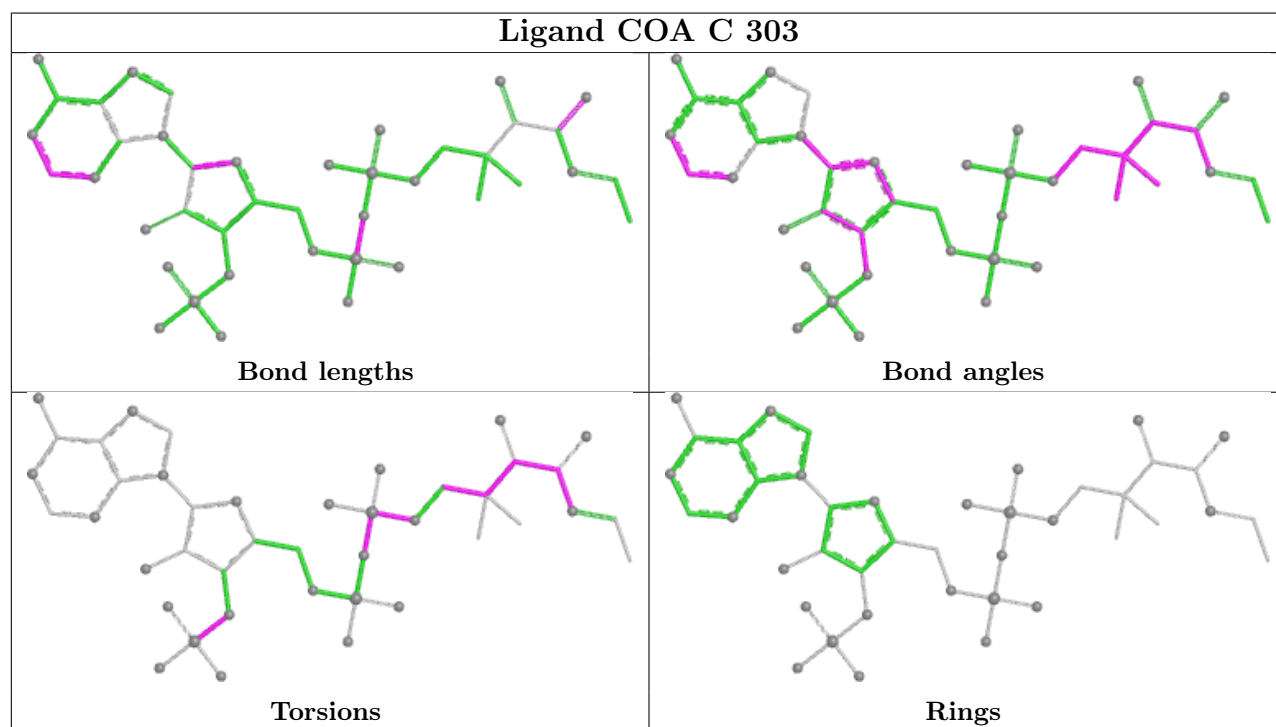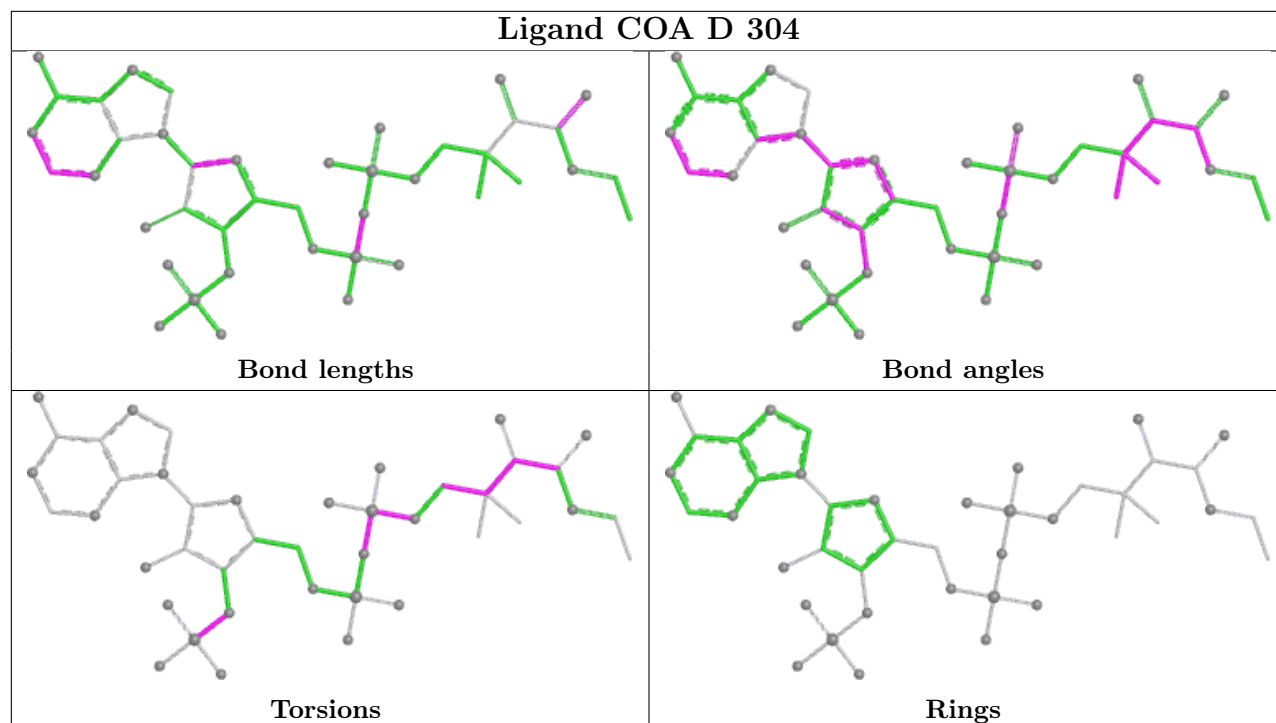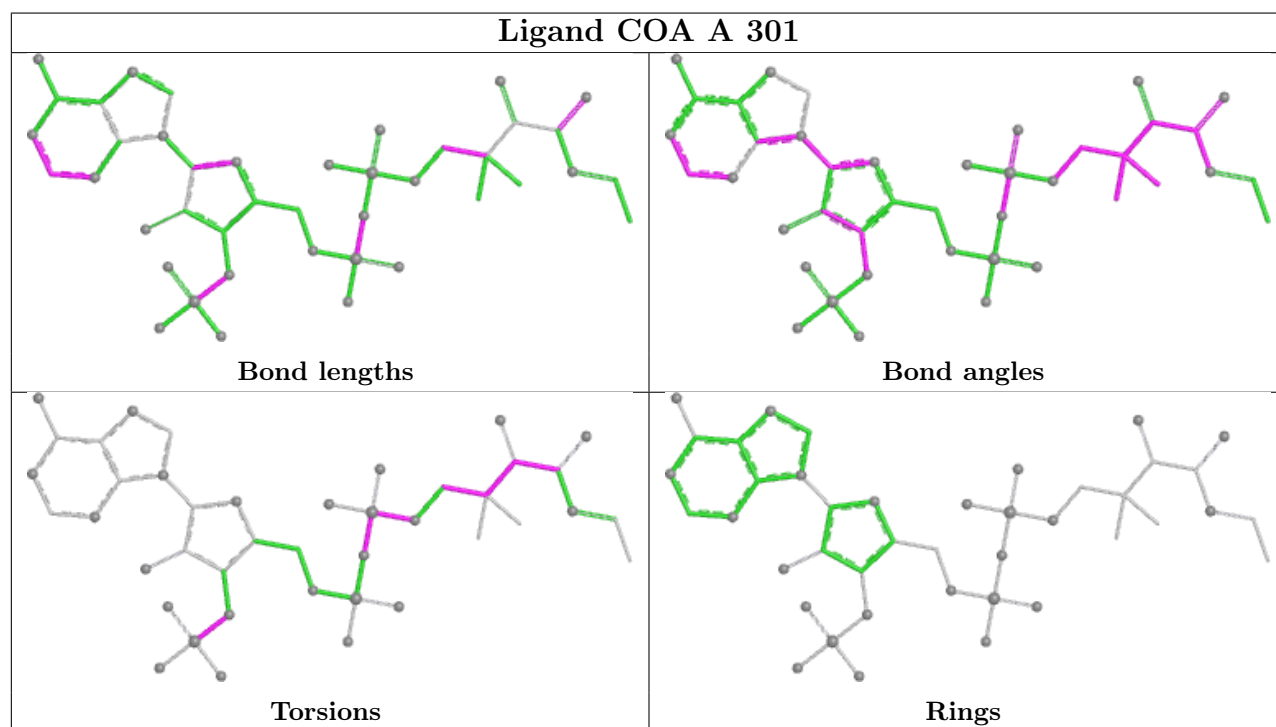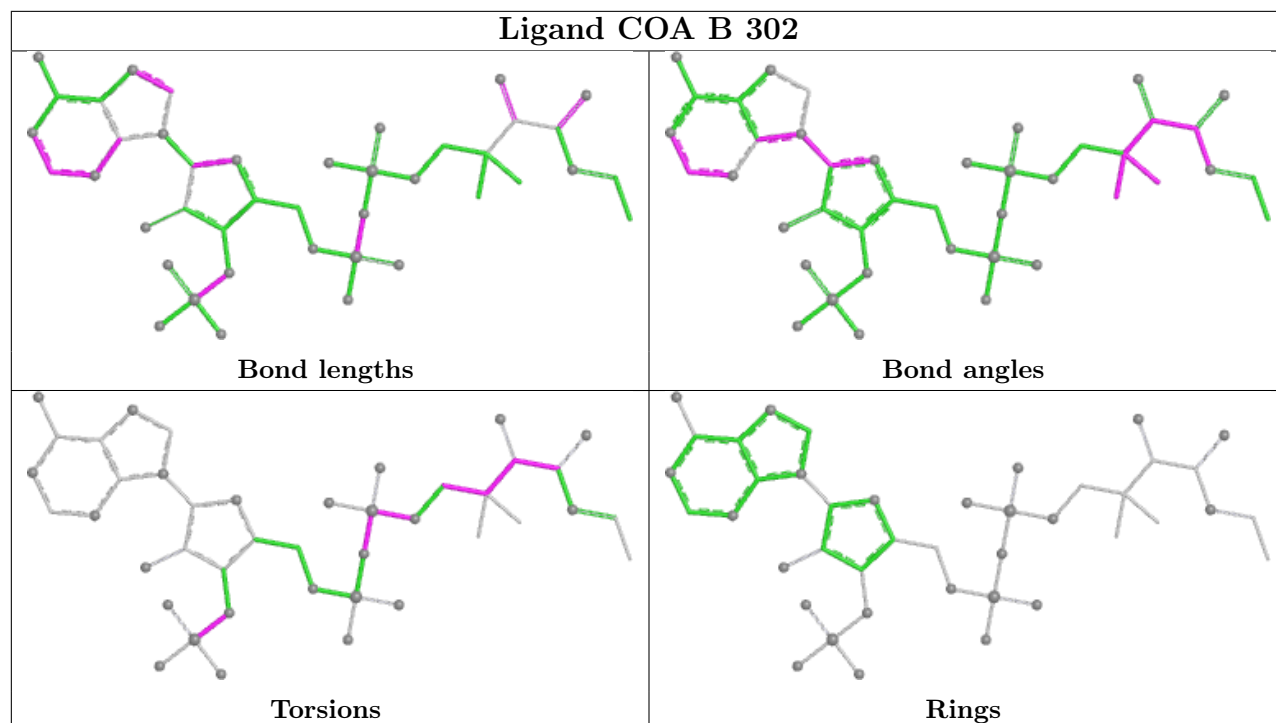
5.7 Other polymers ⓘ

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	270/273 (98%)	-0.34	1 (0%)	92	91	26, 40, 67, 92	0
1	B	251/273 (91%)	0.01	4 (1%)	72	68	33, 62, 79, 85	0
1	C	270/273 (98%)	-0.39	0	100	100	30, 43, 59, 67	0
1	D	268/273 (98%)	-0.28	0	100	100	26, 42, 64, 72	0
1	E	260/273 (95%)	-0.14	3 (1%)	79	76	33, 53, 74, 92	0
1	F	248/273 (90%)	0.35	21 (8%)	10	7	54, 71, 85, 94	0
All	All	1567/1638 (95%)	-0.14	29 (1%)	66	62	26, 51, 79, 94	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	174	LEU	4.0
1	F	240	GLY	3.9
1	F	49	ALA	3.4
1	F	172	LEU	3.3
1	E	203	ILE	3.1

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

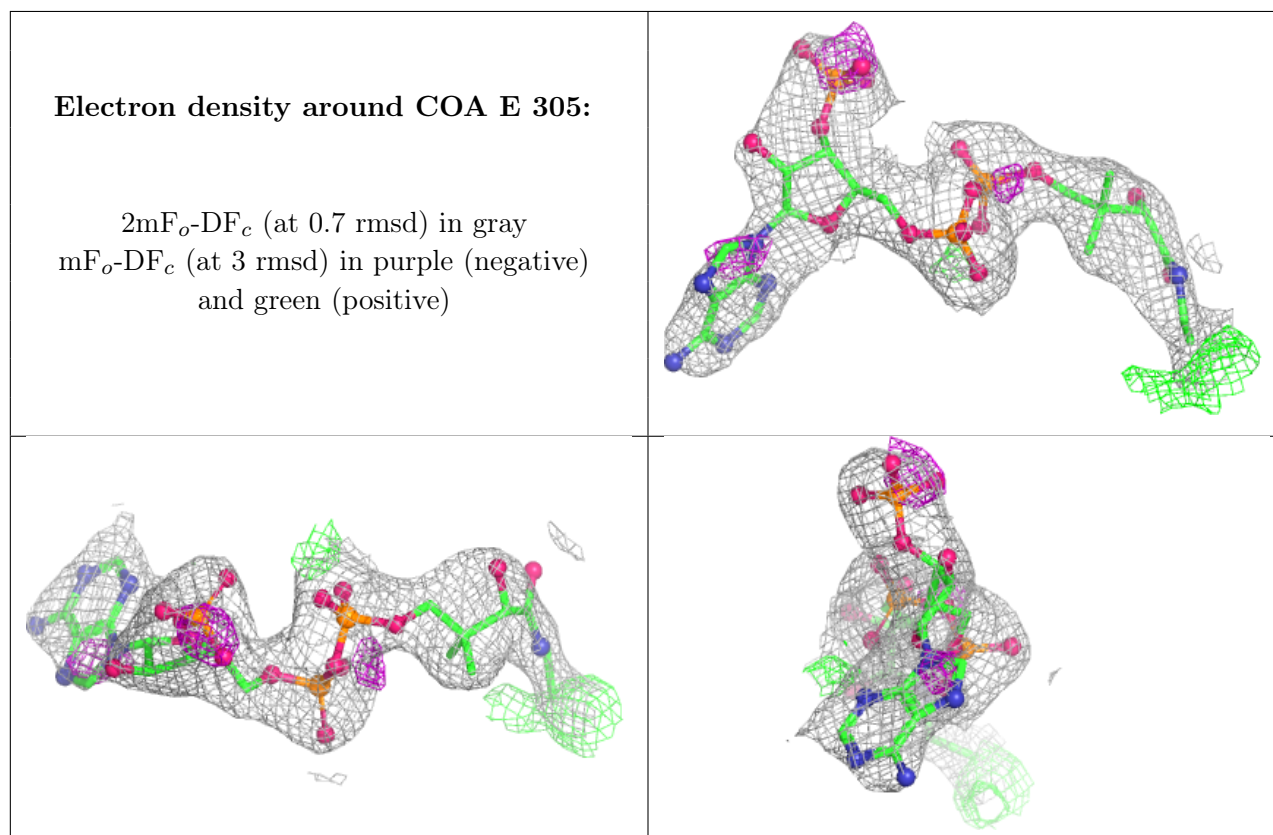There are no non-standard protein/DNA/RNA residues in this entry.

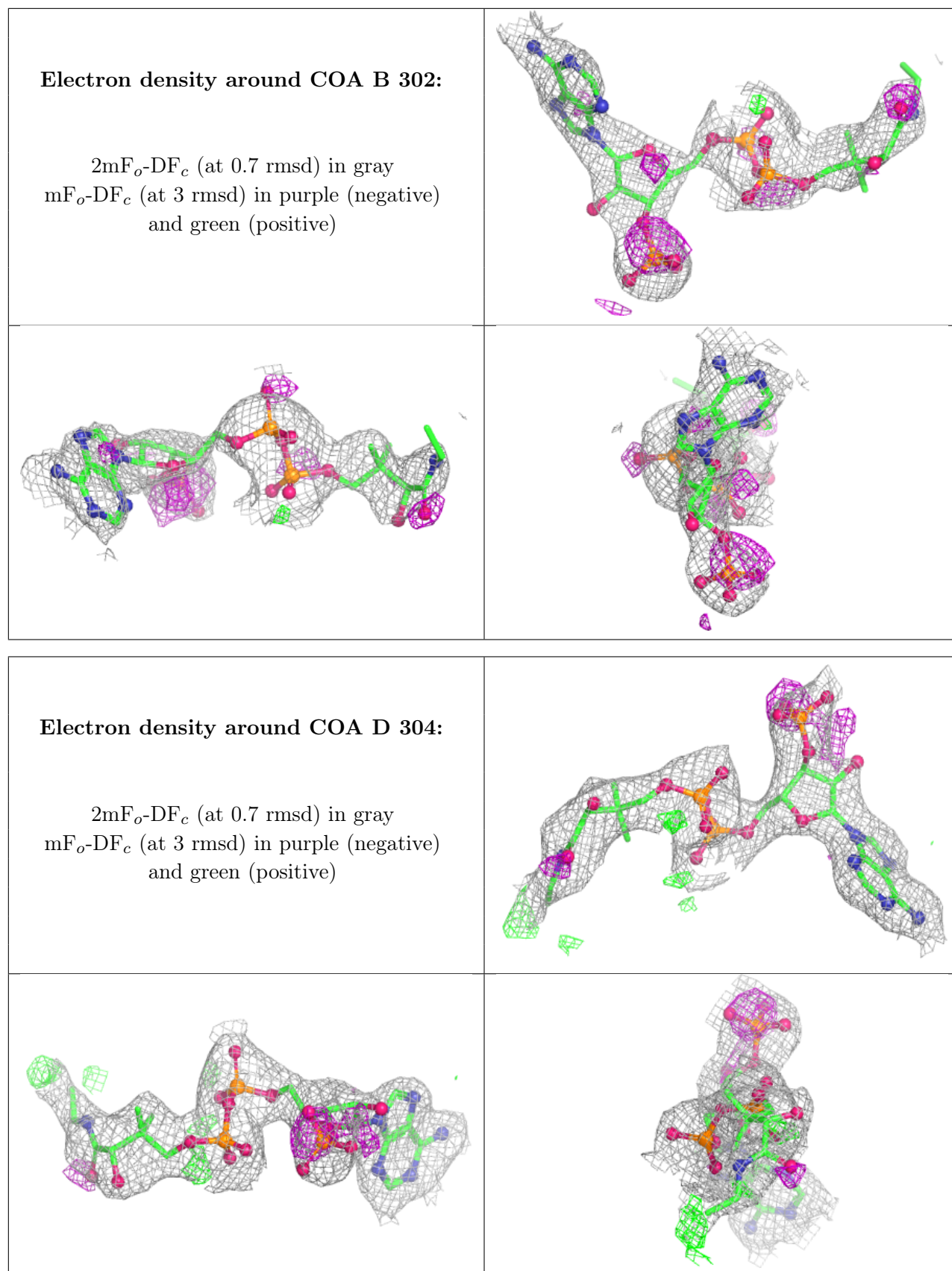## 6.3 Carbohydrates ⓘ

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	COA	E	305	42/48	0.84	0.22	58,64,76,77	0
2	COA	B	302	42/48	0.86	0.33	59,65,73,74	0
2	COA	D	304	42/48	0.90	0.18	41,50,72,74	0
2	COA	A	301	42/48	0.92	0.17	44,53,66,68	0
2	COA	C	303	42/48	0.94	0.15	43,51,67,68	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.
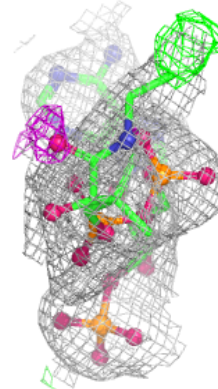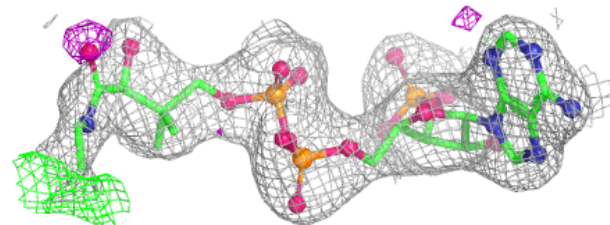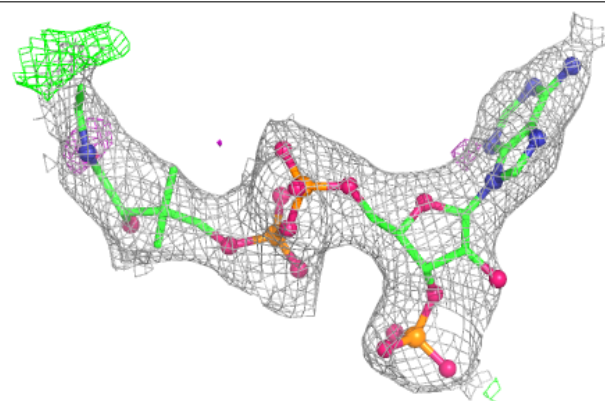
Electron density around COA E 305:

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

Electron density around COA B 302:

$2mF_o$-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 304:

$2mF_o$-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 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
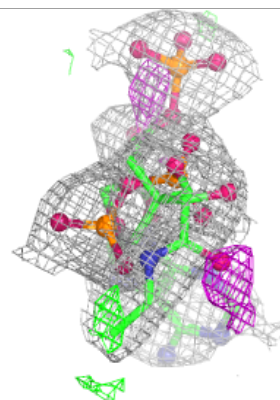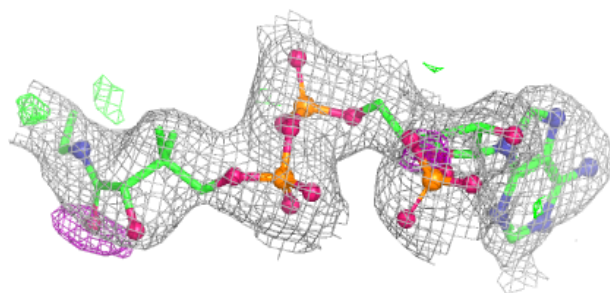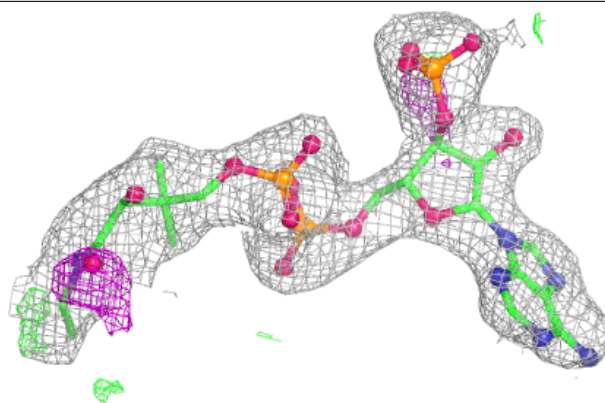
Electron density around COA C 303:

$2mF_o$-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 ⓘ

There are no such residues in this entry.