W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 17, 2020 – 03:13 am BST

PDB ID	:	1SEZ
Title	:	Crystal Structure of Protoporphyrinogen IX Oxidase
Authors	:	Koch, M.; Breithaupt, C.; Kiefersauer, R.; Freigang, J.; Huber, R.; Messerschmidt, A.
Deposited on	:	2004-02-19
Resolution	:	2.90 Å(reported)

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

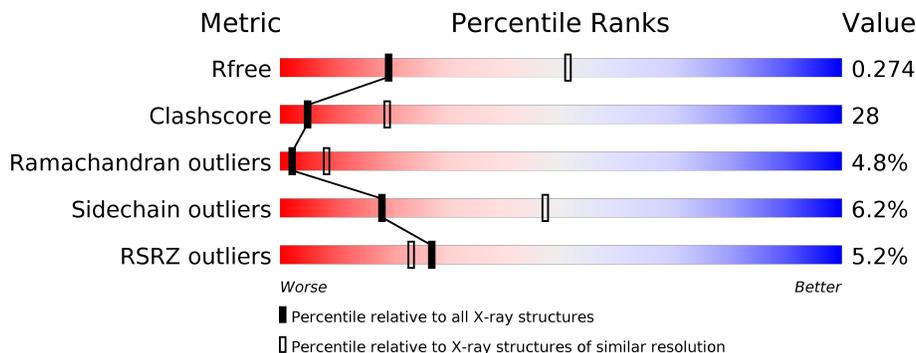|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.90 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	Percentile Ranks	Value
Rfree		0.274
Clashscore		28
Ramachandran outliers		4.8%
Sidechain outliers		6.2%
RSRZ outliers		5.2%

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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 on 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	504	5% / 53% / 35% / • / 8%
1	B	504	5% / 51% / 37% / • • / 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7295 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 Protoporphyrinogen oxidase, mitochondrial.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	465	Total	C	N	O	S	Se	41	0	0
			3480	2215	588	662	6	9			
1	B	465	Total	C	N	O	S	Se	165	0	0
			3480	2215	588	662	6	9			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	69	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	130	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	186	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	242	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	302	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	312	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	376	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	377	MSE	MET	MODIFIED RESIDUE	UNP O24164
A	456	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	69	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	130	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	186	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	242	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	302	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	312	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	376	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	377	MSE	MET	MODIFIED RESIDUE	UNP O24164
B	456	MSE	MET	MODIFIED RESIDUE	UNP O24164

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).

FAD
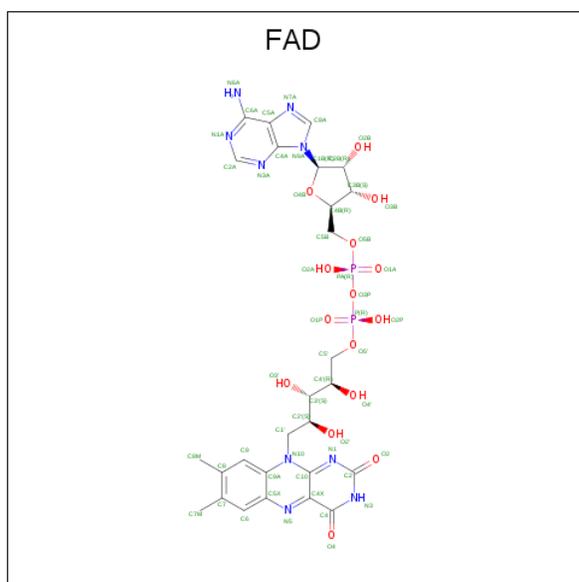
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 3 is 4-BROMO-3-(5'-CARBOXY-4'-CHLORO-2'-FLUOROPHENYL)-1-METHYL-5-TRIFLUOROMETHYL-PYRAZOL (three-letter code: OMN) (formula: $C_{12}H_6BrClF_4N_2O_2$).
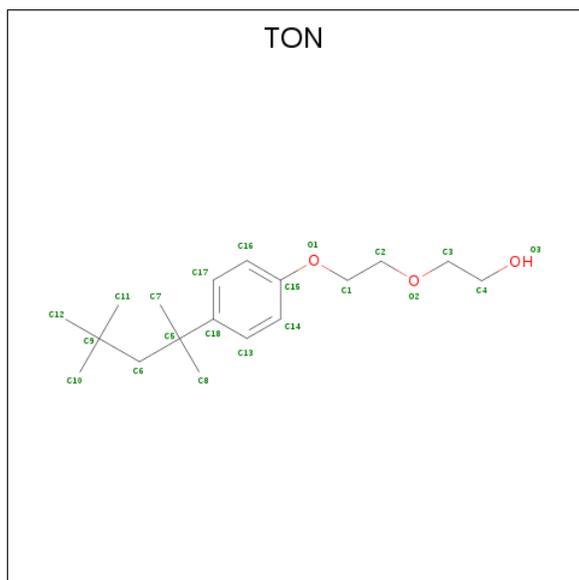
OMN
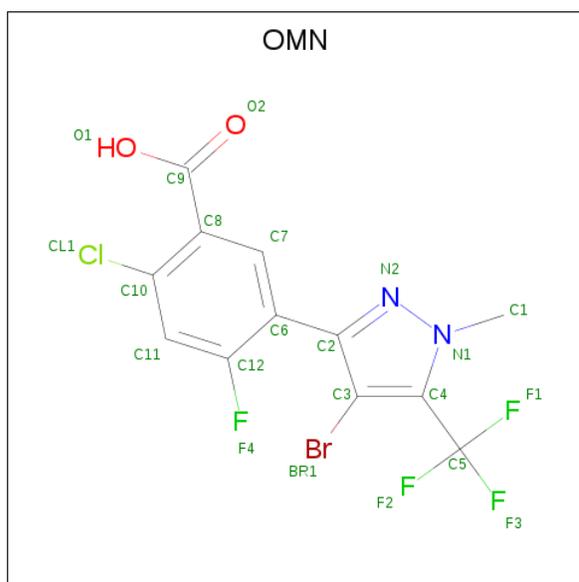
Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
3	A	1	Total	Br	C	Cl	F	N	O	0	0
			22	1	12	1	4	2	2		

Continued from previous page...

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
3	B	1	Total	Br	C	Cl	F	N	O	0	0
			22	1	12	1	4	2	2		

- Molecule 4 is 2-{2-[4-(1,1,3,3-TETRAMETHYLBUTYL)PHENOXY]ETHOXY}ETHANOL (three-letter code: TON) (formula: $C_{18}H_{30}O_3$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			21	18	3		
4	B	1	Total	C	O	0	0
			21	18	3		

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	75	Total	O	0	0
			75	75		
5	B	68	Total	O	0	0
			68	68		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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: Protoporphyrinogen oxidase, mitochondrial
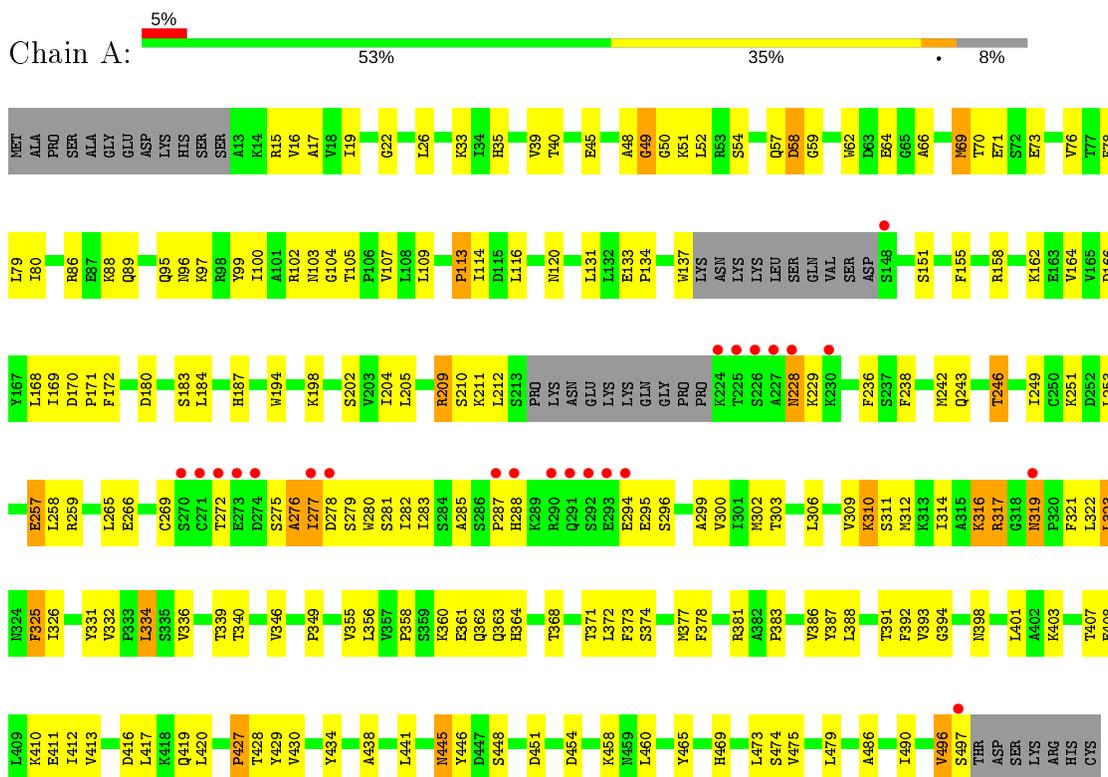
- Molecule 1: Protoporphyrinogen oxidase, mitochondrial
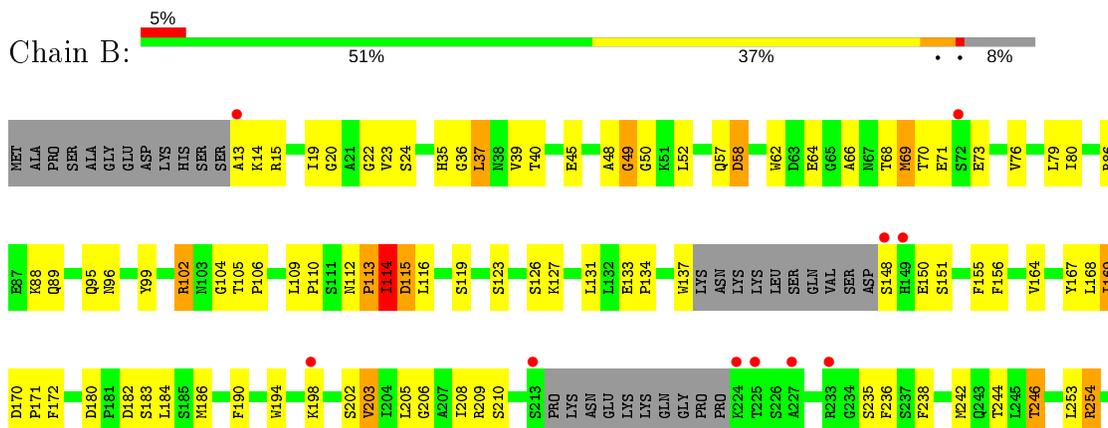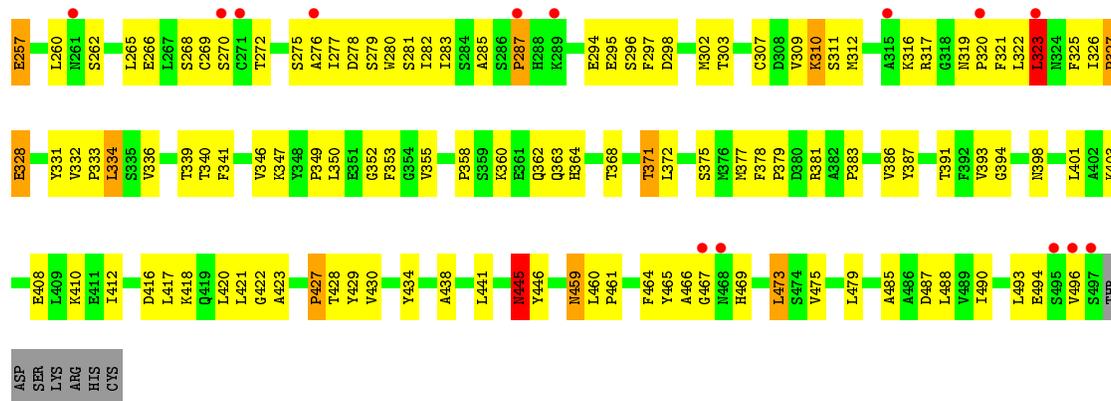
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	119.09Å 147.25Å 127.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.73 – 2.90 19.73 – 2.90	Depositor EDS
% Data completeness (in resolution range)	97.6 (19.73-2.90) 99.5 (19.73-2.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.62 (at 2.88Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.227 , 0.293 0.222 , 0.274	Depositor DCC
R_{free} test set	1209 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	77.2	Xtriage
Anisotropy	0.227	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 50.7	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7295	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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: TON, OMN, FAD

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	0.41	0/3539	0.62	1/4781 (0.0%)
1	B	0.41	0/3539	0.63	1/4781 (0.0%)
All	All	0.41	0/7078	0.62	2/9562 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	287	PRO	N-CA-CB	5.58	110.00	103.30
1	B	287	PRO	N-CA-CB	5.56	109.97	103.30

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	3480	0	3367	193	0
1	B	3480	0	3367	184	0
2	A	53	0	31	2	0
2	B	53	0	31	3	0
3	A	22	0	5	0	0
3	B	22	0	5	0	0
4	A	21	0	30	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	21	0	30	1	0
5	A	75	0	0	14	0
5	B	68	0	0	14	0
All	All	7295	0	6866	375	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 28.

The worst 5 of 375 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:A:69:MSE:HE2	1:A:242:MSE:HE1	1.44	1.00
1:B:69:MSE:HE1	1:B:238:PHE:HE1	1.25	0.99
1:B:309:VAL:HA	1:B:312:MSE:HE3	1.47	0.95
1:A:496:VAL:HG23	1:A:497:SER:H	1.32	0.95
1:B:69:MSE:HE1	1:B:238:PHE:CE1	2.06	0.90

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	459/504 (91%)	391 (85%)	50 (11%)	18 (4%)	3	12
1	B	459/504 (91%)	373 (81%)	60 (13%)	26 (6%)	1	5
All	All	918/1008 (91%)	764 (83%)	110 (12%)	44 (5%)	2	8

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	58	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	210	SER
1	A	211	LYS
1	A	272	THR
1	A	275	SER

5.3.2 Protein sidechains (i)

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	362/423 (86%)	341 (94%)	21 (6%)	20	50
1	B	362/423 (86%)	338 (93%)	24 (7%)	16	44
All	All	724/846 (86%)	679 (94%)	45 (6%)	18	47

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	473	LEU
1	B	105	THR
1	B	446	TYR
1	B	58	ASP
1	B	114	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	398	ASN
1	A	445	ASN
1	B	445	ASN
1	A	319	ASN
1	B	398	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

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	$\#|Z| > 2$
2	FAD	B	610	-	51,58,58	1.68	9 (17%)	60,89,89	2.04	7 (11%)
2	FAD	A	600	-	51,58,58	1.73	9 (17%)	60,89,89	2.06	8 (13%)
4	TON	A	603	-	21,21,21	1.90	8 (38%)	29,29,29	1.11	2 (6%)
3	OMN	B	604	-	20,23,23	2.82	5 (25%)	26,36,36	1.41	5 (19%)
4	TON	B	605	-	21,21,21	1.91	5 (23%)	29,29,29	1.05	1 (3%)
3	OMN	A	601	-	20,23,23	2.75	7 (35%)	26,36,36	1.35	5 (19%)

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	FAD	B	610	-	-	3/30/50/50	0/6/6/6
2	FAD	A	600	-	-	1/30/50/50	0/6/6/6
4	TON	A	603	-	-	10/19/19/19	0/1/1/1
3	OMN	B	604	-	-	0/8/14/14	0/2/2/2
4	TON	B	605	-	-	8/19/19/19	0/1/1/1
3	OMN	A	601	-	-	0/8/14/14	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	OMN	C2-N2	8.96	1.42	1.35
3	B	604	OMN	C2-N2	8.24	1.42	1.35
2	A	600	FAD	C4X-N5	5.61	1.41	1.33
2	B	610	FAD	C4X-C10	5.35	1.44	1.38
3	B	604	OMN	C5-C4	5.22	1.56	1.49

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	610	FAD	C4-N3-C2	11.06	124.48	115.14
2	A	600	FAD	C4-N3-C2	10.65	124.13	115.14
2	B	610	FAD	C4X-C4-N3	-4.76	116.92	123.43
2	A	600	FAD	C1'-N10-C9A	4.69	121.99	118.29
2	A	600	FAD	O4B-C1B-C2B	-4.55	100.28	106.93

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	610	FAD	N10-C1'-C2'-O2'
4	A	603	TON	C18-C5-C6-C9
4	A	603	TON	C8-C5-C6-C9
4	A	603	TON	C7-C5-C6-C9
4	B	605	TON	C18-C5-C6-C9

There are no ring outliers.
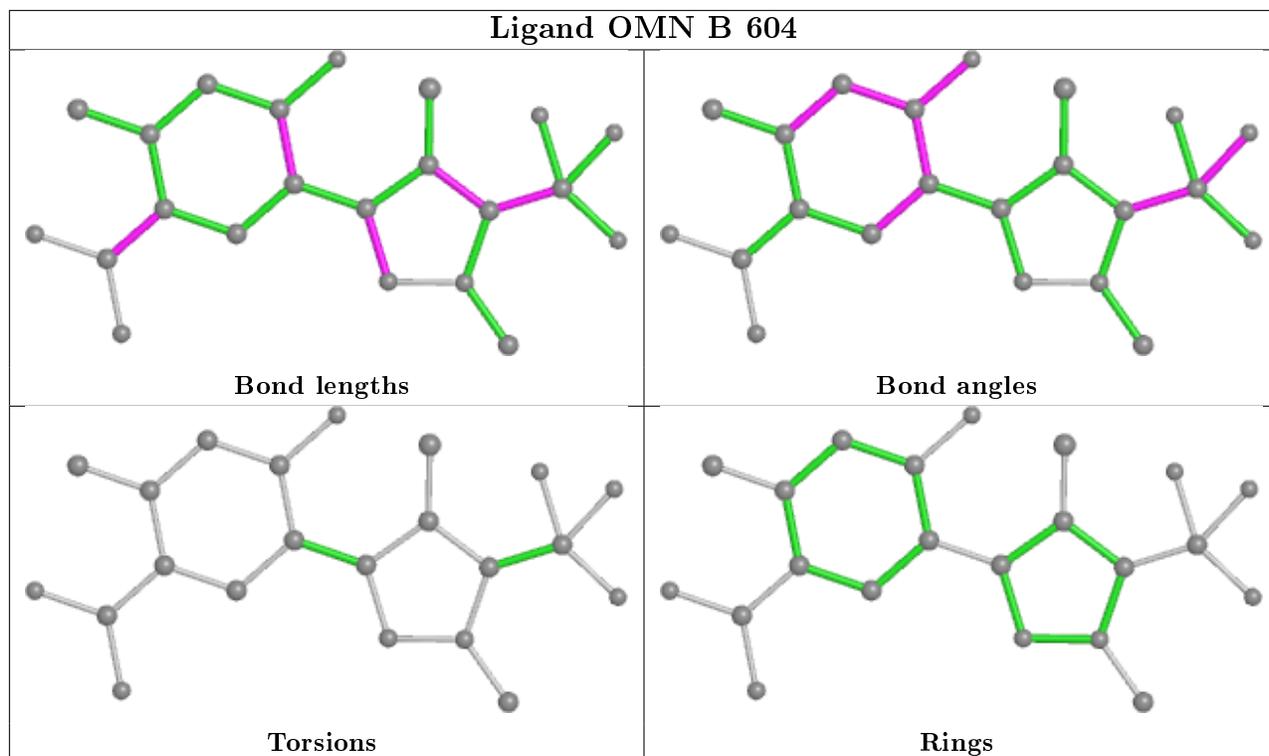
4 monomers are involved in 8 short contacts:

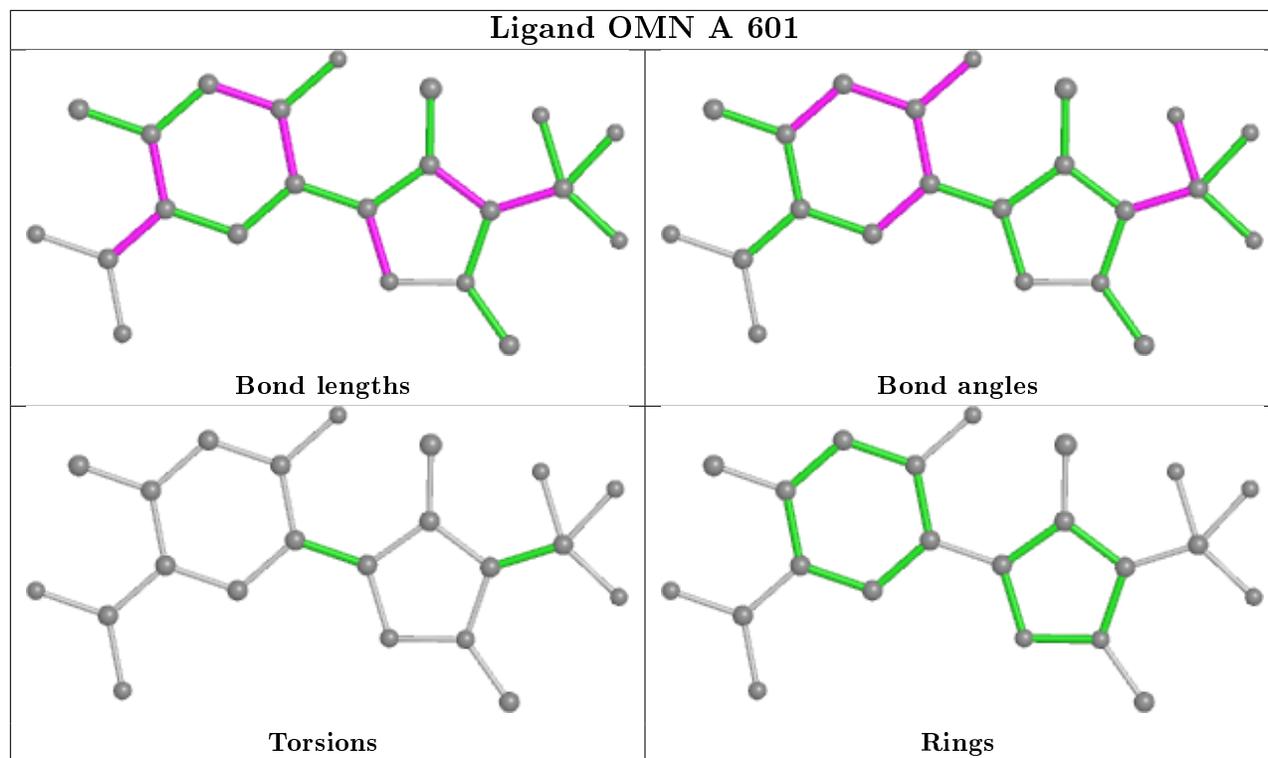| Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	610	FAD	3	0
2	A	600	FAD	2	0
4	A	603	TON	2	0
4	B	605	TON	1	0

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 FAD B 610

Bond lengths

Bond angles

Torsions

Rings

Ligand FAD A 600

Bond lengths

Bond angles

Torsions

Rings

Ligand TON A 603

Bond lengths

Bond angles

Torsions

Rings

Ligand OMN B 604

Bond lengths

Bond angles

Torsions

Rings

Ligand TON B 605

Bond lengths

Bond angles

Torsions

Rings

Ligand OMN A 601

Bond lengths

Bond angles

Torsions

Rings
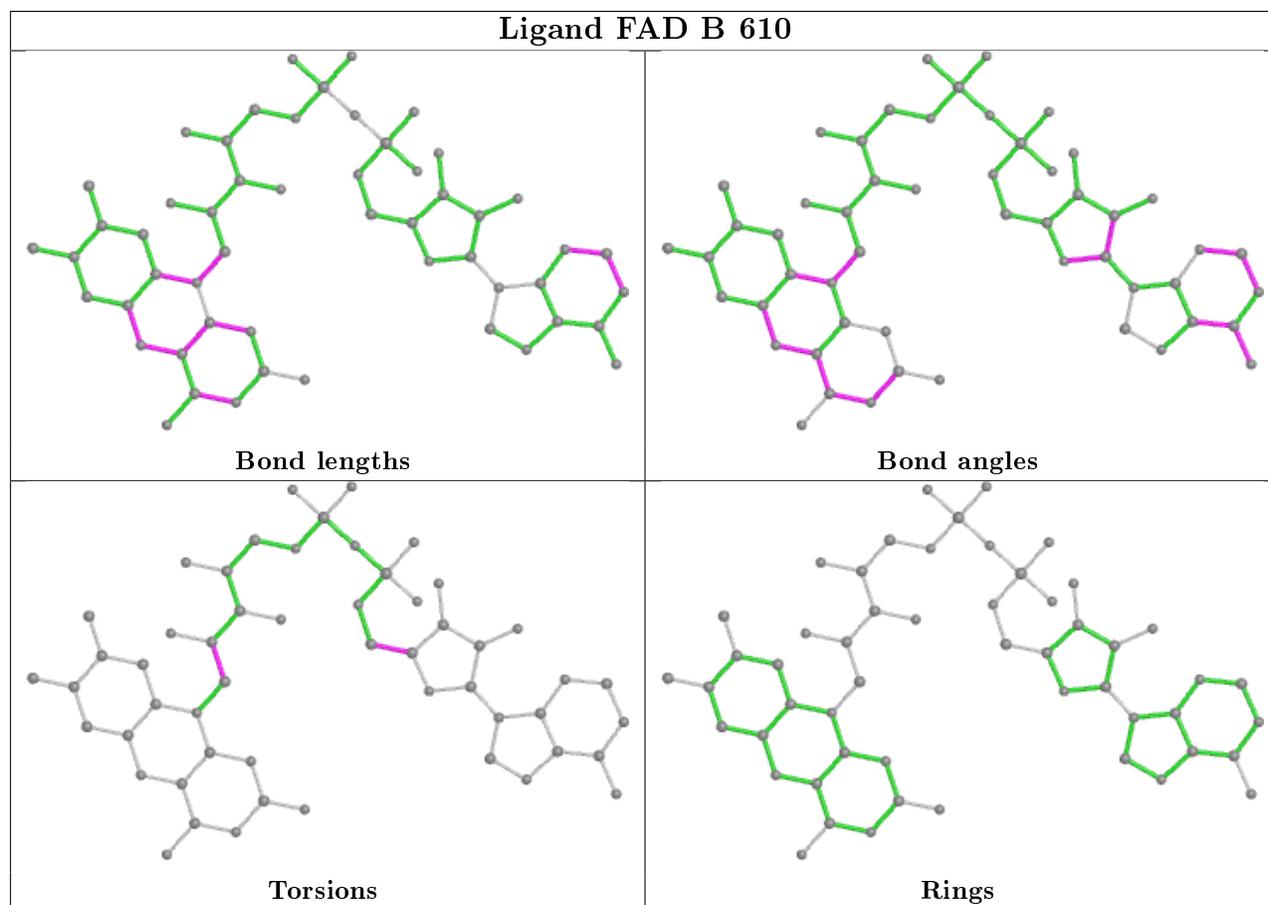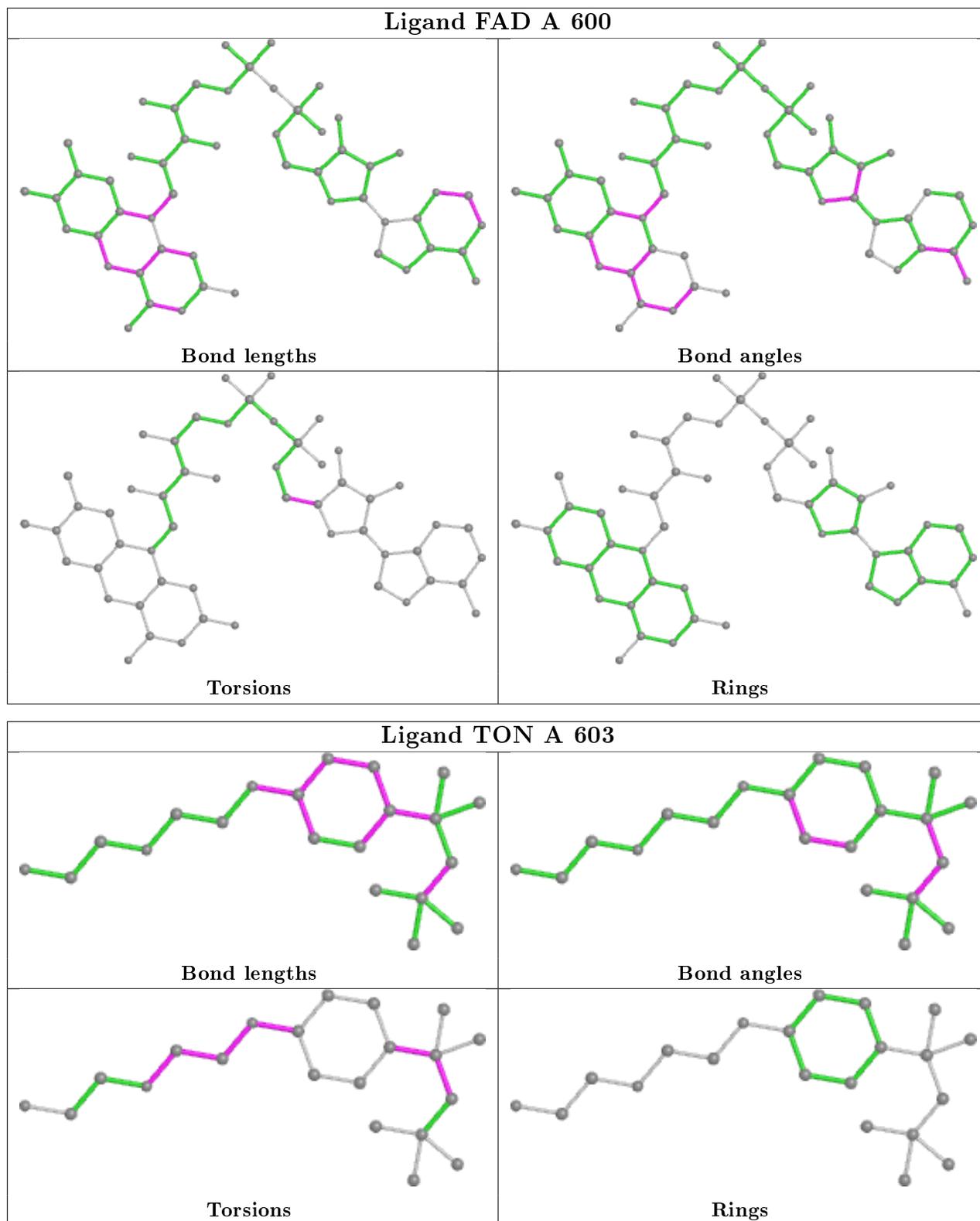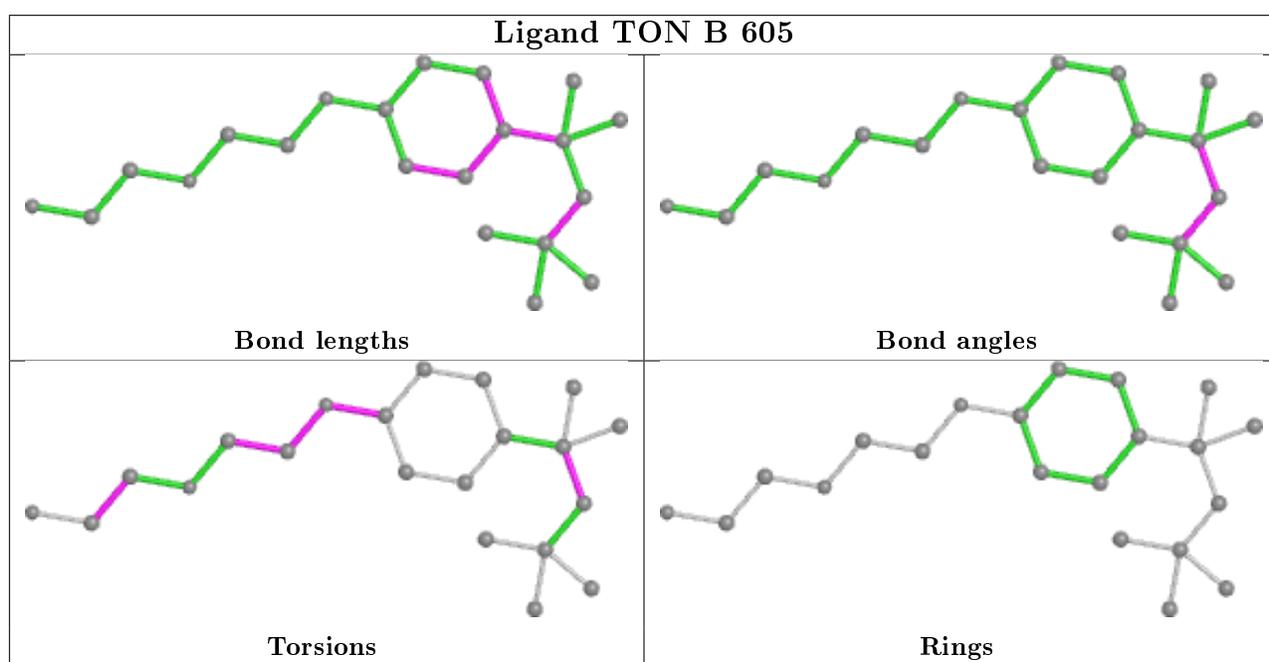
5.7 Other polymers (i)

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 ⓘ

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	455/504 (90%)	0.10	23 (5%)	28	24	41, 63, 89, 99	25 (5%)
1	B	448/504 (88%)	0.15	24 (5%)	25	22	41, 68, 96, 100	47 (10%)
All	All	903/1008 (89%)	0.13	47 (5%)	27	23	41, 65, 94, 100	72 (7%)

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	292	SER	13.5
1	A	227	ALA	12.3
1	B	271	CYS	9.7
1	A	225	THR	9.4
1	A	224	LYS	8.9

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

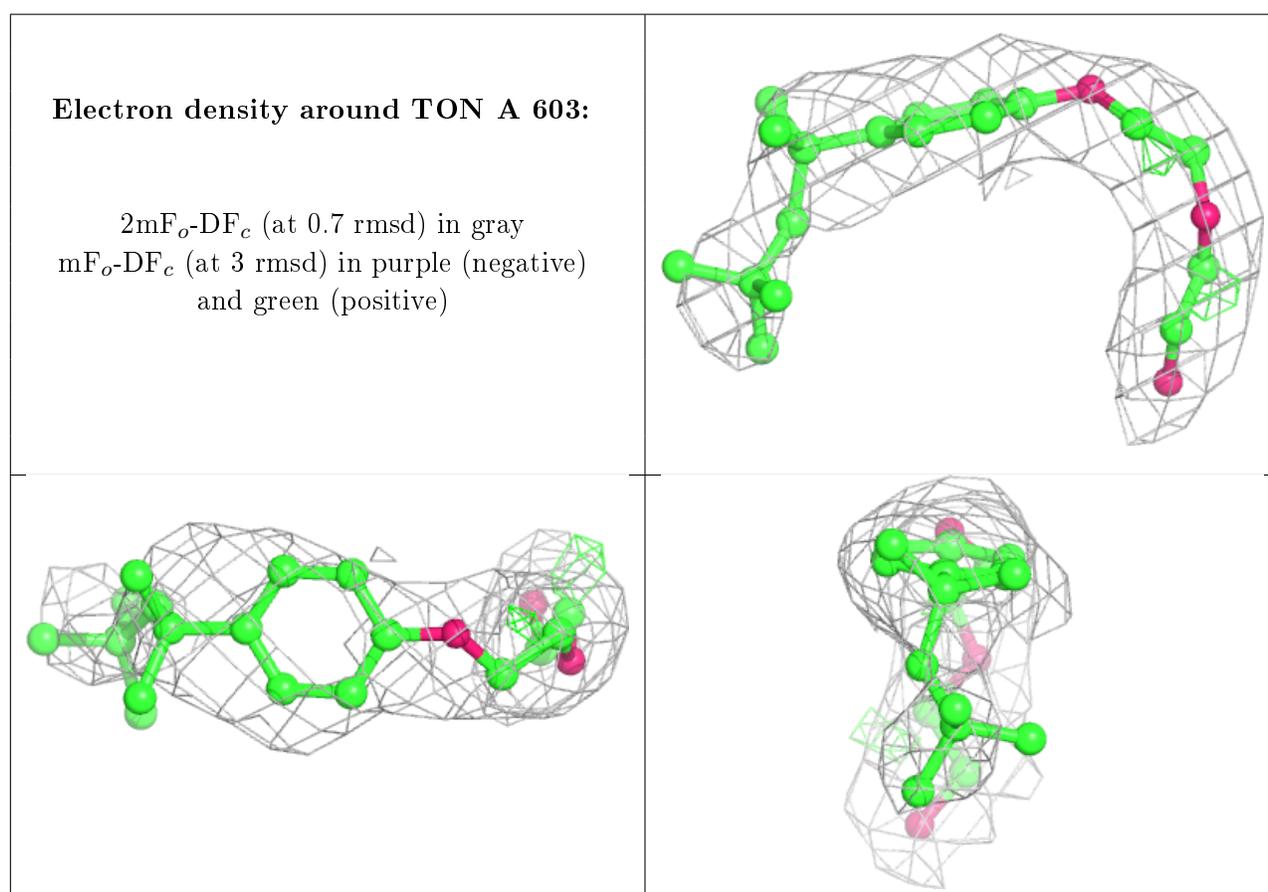There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.
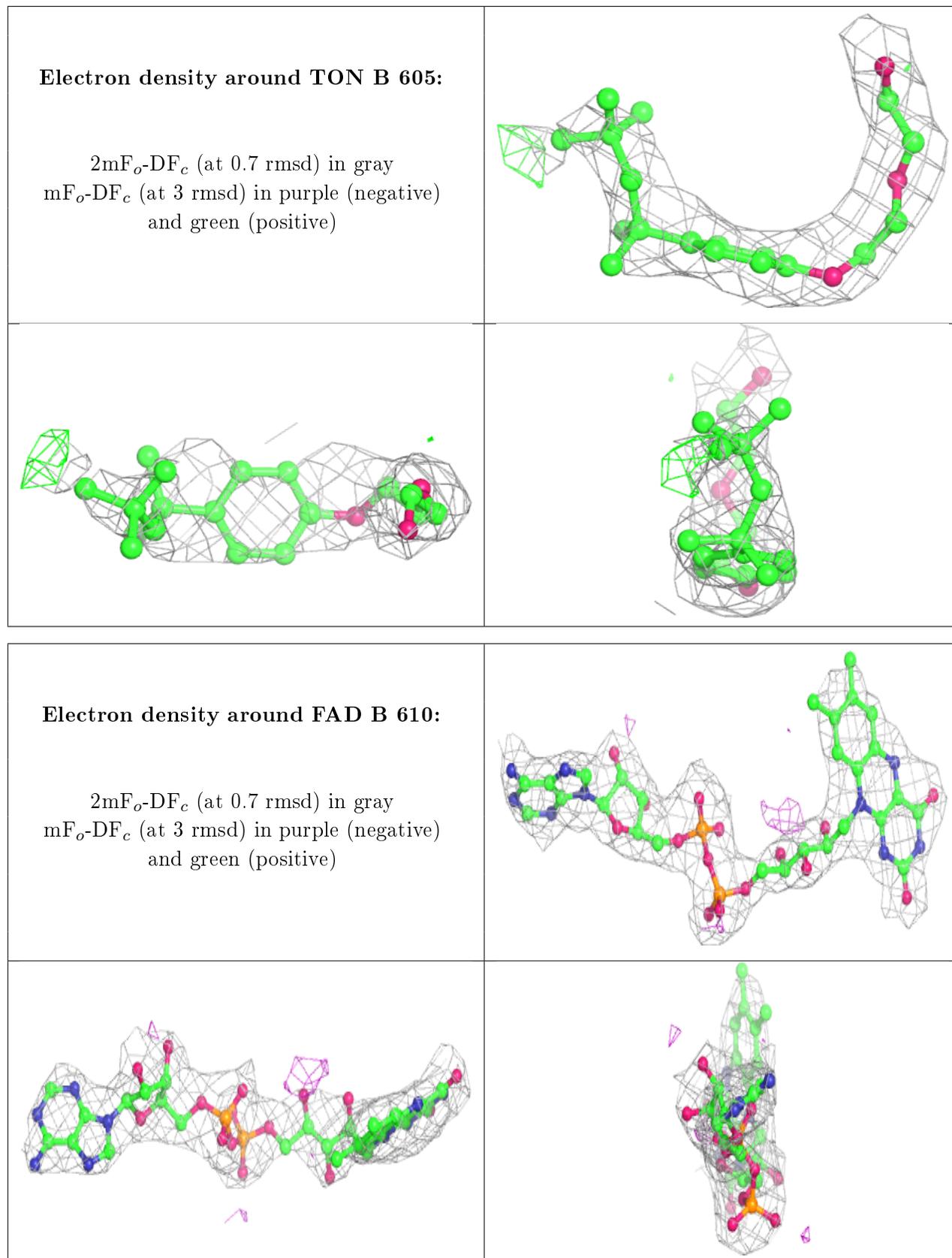
6.4 Ligands ⓘ

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, 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
4	TON	A	603	21/21	0.79	0.27	78,85,93,93	0
4	TON	B	605	21/21	0.80	0.29	83,86,90,91	0
2	FAD	B	610	53/53	0.94	0.16	49,63,72,73	0
2	FAD	A	600	53/53	0.96	0.15	27,42,50,51	0
3	OMN	B	604	22/22	0.97	0.15	61,63,66,73	0
3	OMN	A	601	22/22	0.98	0.13	52,58,60,70	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 TON A 603:

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

Electron density around TON B 605:

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

Electron density around FAD B 610:

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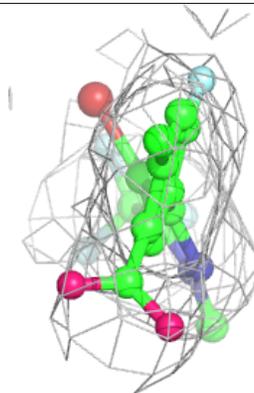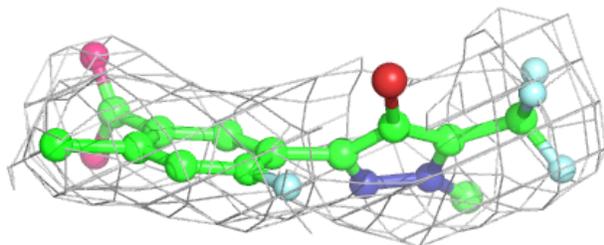mF_o-DF_c (at 3 rmsd) in purple (negative)
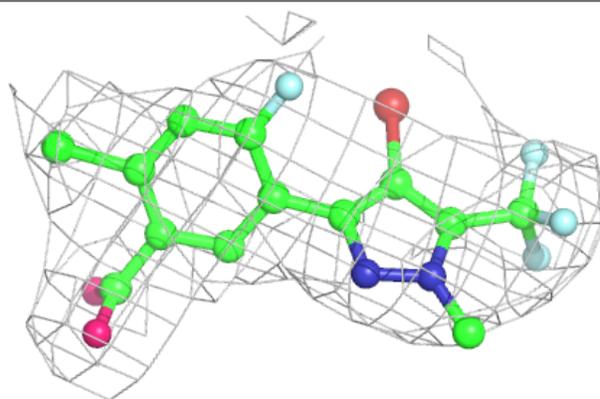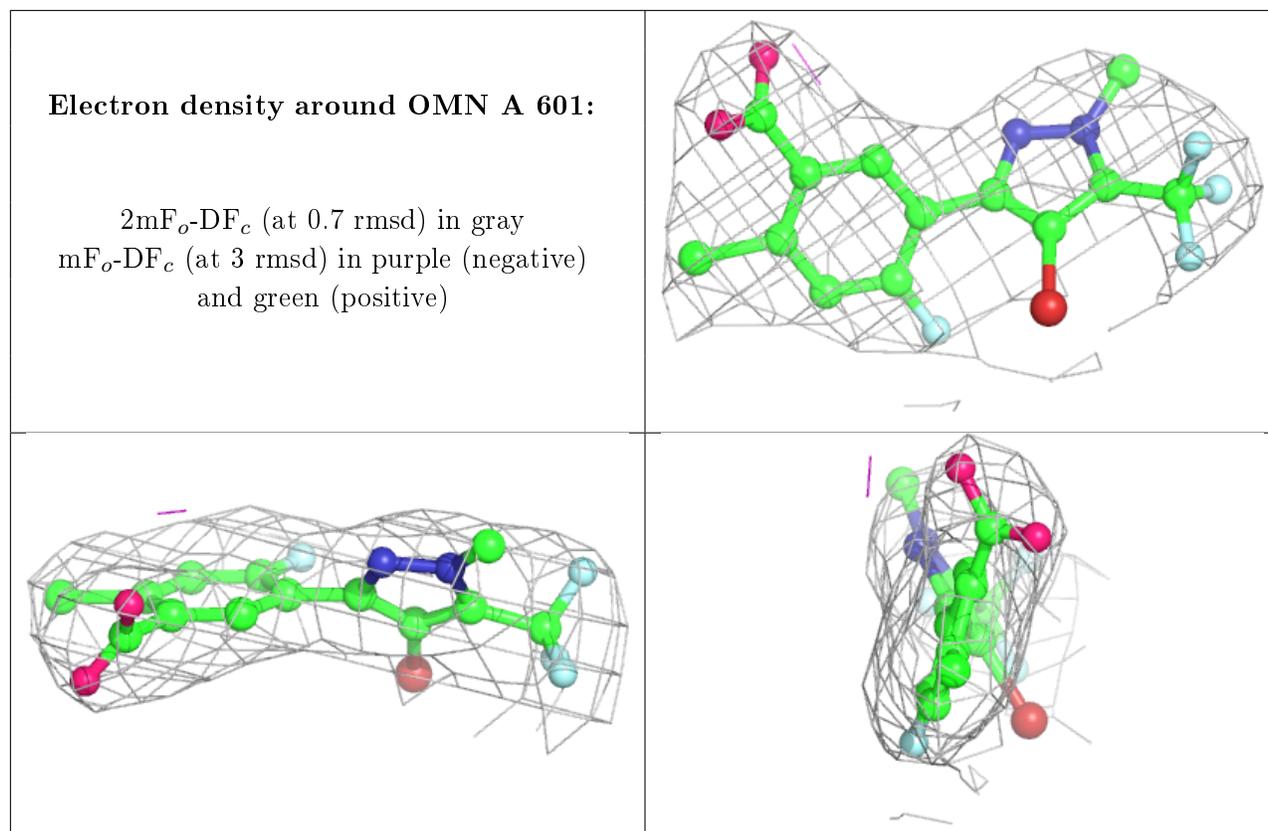and green (positive)

Electron density around FAD A 600:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
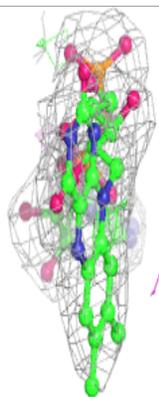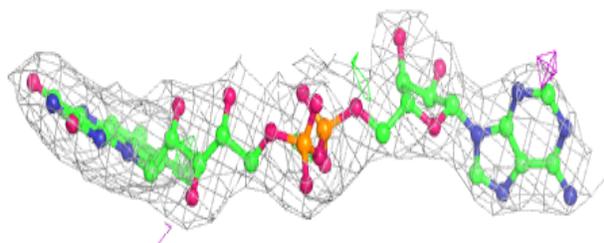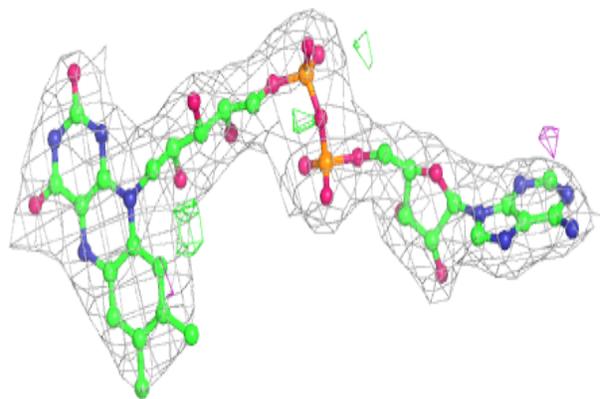
Electron density around OMN B 604:

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

Electron density around OMN A 601:

$2mF_o\text{-}DF_c$ (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.