wwPDB X-ray Structure Validation Summary Report (i)

Oct 4, 2023 – 02:38 AM EDT

PDB ID	:	6PH4
Title	:	Full length LOV-PAS-HK construct from the LOV-HK sensory protein from Brucella abortus (light-adapted, construct 15-489)
Authors	:	Rinaldi, J.; Otero, L.H.; Fernandez, I.; Goldbaum, F.A.; Shin, H.; Yang, X.; Klinke, S.
Deposited on	:	2019-06-25
Resolution	:	3.25 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 (i) 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 (i)) were used in the production of this report:

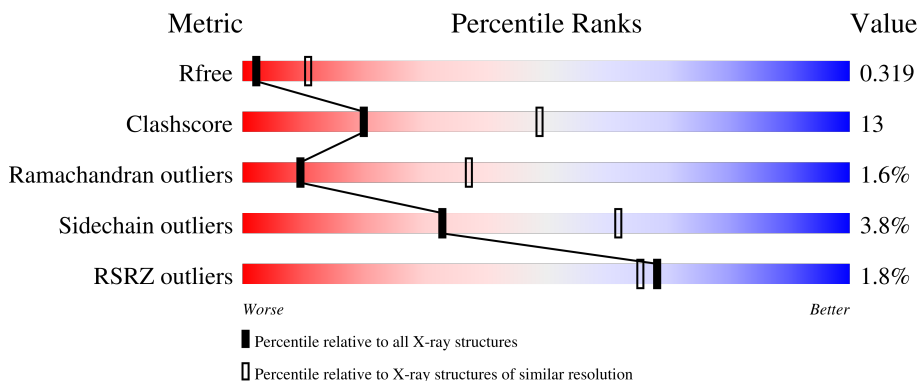|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.25 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		0.319
Clashscore		13
Ramachandran outliers		1.6%
Sidechain outliers		3.8%
RSRZ outliers		1.8%

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	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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	482	% 44% 20% 35%
1	B	482	2% 66% 28% • 5%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 6241 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 Blue-light-activated histidine kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	313	Total	C	N	O	S	0	0	0
			2503	1576	450	467	10			
1	B	460	Total	C	N	O	S	0	0	0
			3641	2306	647	671	17			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	MET	-	initiating methionine	UNP Q8YC53
A	490	HIS	-	expression tag	UNP Q8YC53
A	491	HIS	-	expression tag	UNP Q8YC53
A	492	HIS	-	expression tag	UNP Q8YC53
A	493	HIS	-	expression tag	UNP Q8YC53
A	494	HIS	-	expression tag	UNP Q8YC53
A	495	HIS	-	expression tag	UNP Q8YC53
B	14	MET	-	initiating methionine	UNP Q8YC53
B	490	HIS	-	expression tag	UNP Q8YC53
B	491	HIS	-	expression tag	UNP Q8YC53
B	492	HIS	-	expression tag	UNP Q8YC53
B	493	HIS	-	expression tag	UNP Q8YC53
B	494	HIS	-	expression tag	UNP Q8YC53
B	495	HIS	-	expression tag	UNP Q8YC53

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Cl	0	0
			1	1		
2	B	2	Total	Cl	0	0
			2	2		

- Molecule 3 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula: $C_{17}H_{21}N_4O_9P$)

(labeled as "Ligand of Interest" by depositor).

FMN

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	17	4	9	1		
3	B	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		

- Molecule 5 is PHOSPHOMETHYLPHOSPHONIC ACID ADENYLATE ESTER (three-letter code: ACP) (formula: $C_{11}H_{18}N_5O_{12}P_3$) (labeled as "Ligand of Interest" by depositor).

ACP

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

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: Blue-light-activated histidine kinase

Chain A:

• Molecule 1: Blue-light-activated histidine kinase

Chain B:

T472 P473 E474 V475 F476 P477 GLY MET GLY HIS ASN VAL PRO GLU GLN ARG ILE ALA HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	95.96Å 104.66Å 164.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.52 – 3.25 62.52 – 3.25	Depositor EDS
% Data completeness (in resolution range)	99.9 (62.52-3.25) 100.0 (62.52-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.06 (at 3.26Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.255 , 0.318 0.255 , 0.319	Depositor DCC
R_{free} test set	1345 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	133.6	Xtriage
Anisotropy	0.434	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 97.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6241	wwPDB-VP
Average B, all atoms (Å^2)	153.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.29% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: FMN, MG, CL, ACP

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	0.24	0/2553	0.47	0/3459
1	B	0.26	0/3717	0.47	0/5038
All	All	0.25	0/6270	0.47	0/8497

There are no bond length outliers.

There are no bond angle outliers.

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	2503	0	2495	77	0
1	B	3641	0	3650	94	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
3	A	31	0	19	4	0
3	B	31	0	19	3	0
4	B	1	0	0	0	0
5	B	31	0	13	0	0
All	All	6241	0	6196	157	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 13.

The worst 5 of 157 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:283:SER:HA	1:A:286:ILE:HG22	1.58	0.85
1:A:285:GLU:HA	1:A:288:HIS:HB3	1.68	0.76
1:B:63:GLU:O	1:B:67:ARG:NH1	2.19	0.76
1:B:130:SER:HB3	3:B:504:FMN:HM73	1.68	0.76
1:A:176:THR:HG1	1:A:185:THR:HG1	1.33	0.72

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	311/482 (64%)	275 (88%)	31 (10%)	5 (2%)	9	37
1	B	458/482 (95%)	401 (88%)	50 (11%)	7 (2%)	10	39
All	All	769/964 (80%)	676 (88%)	81 (10%)	12 (2%)	9	37

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ALA
1	A	270	THR
1	A	271	GLU
1	B	70	ARG
1	B	336	ASN

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	271/408 (66%)	262 (97%)	9 (3%)	38	65
1	B	390/408 (96%)	374 (96%)	16 (4%)	30	60
All	All	661/816 (81%)	636 (96%)	25 (4%)	33	62

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	217	MET
1	B	281	LEU
1	B	454	CYS
1	B	272	ARG
1	B	290	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 for Mogul analysis.

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	ACP	B	505	4	27,33,33	4.48	11 (40%)	32,52,52	1.75	5 (15%)
3	FMN	A	502	-	33,33,33	1.60	8 (24%)	48,50,50	1.48	11 (22%)
3	FMN	B	504	-	33,33,33	1.60	8 (24%)	48,50,50	1.45	11 (22%)

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	ACP	B	505	4	-	5/15/38/38	0/3/3/3
3	FMN	A	502	-	-	9/18/18/18	0/3/3/3
3	FMN	B	504	-	-	6/18/18/18	0/3/3/3

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	505	ACP	O4'-C1'	15.48	1.62	1.41
5	B	505	ACP	C2'-C1'	-12.86	1.34	1.53
5	B	505	ACP	O4'-C4'	-5.67	1.32	1.45
5	B	505	ACP	C6-N6	4.37	1.50	1.34
5	B	505	ACP	PB-O2B	-4.21	1.46	1.56

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	505	ACP	C5-C6-N6	5.65	128.94	120.35
5	B	505	ACP	N3-C2-N1	-5.48	120.11	128.68
5	B	505	ACP	N6-C6-N1	-3.60	111.10	118.57
3	A	502	FMN	C1'-N10-C9A	3.48	126.31	120.51
3	B	504	FMN	C1'-N10-C9A	3.44	126.24	120.51

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	FMN	C1'-C2'-C3'-O3'

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	502	FMN	C1'-C2'-C3'-C4'
3	A	502	FMN	O2'-C2'-C3'-O3'
3	A	502	FMN	O2'-C2'-C3'-C4'
3	B	504	FMN	C3'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	FMN	4	0
3	B	504	FMN	3	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 ACP B 505

Bond lengths

Bond angles

Torsions

Rings

Ligand FMN A 502

Bond lengths

Bond angles

Torsions

Rings

Ligand FMN B 504

Bond lengths

Bond angles

Torsions

Rings
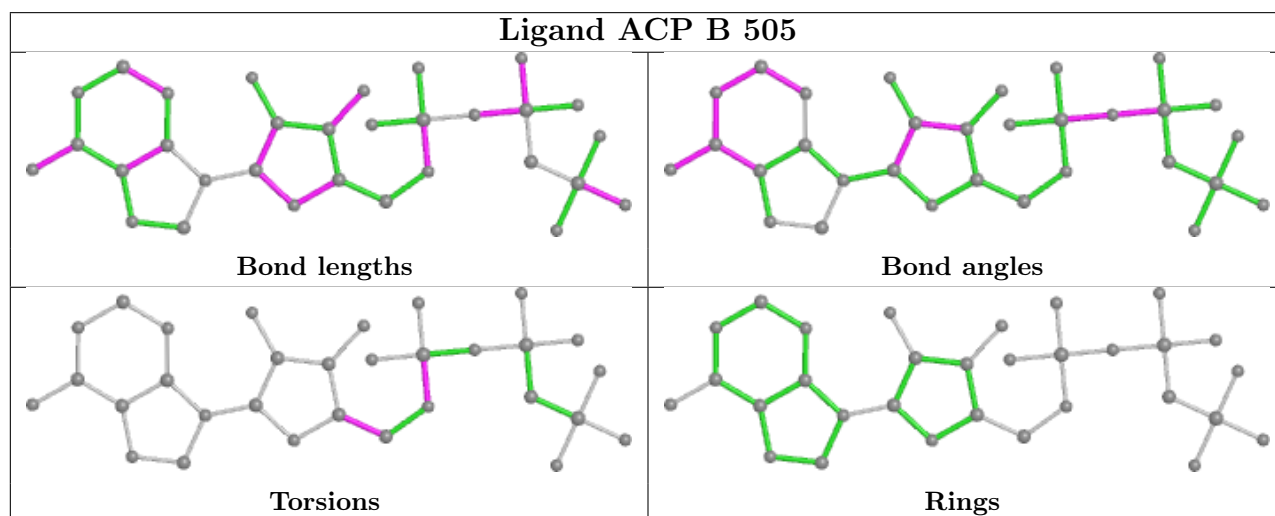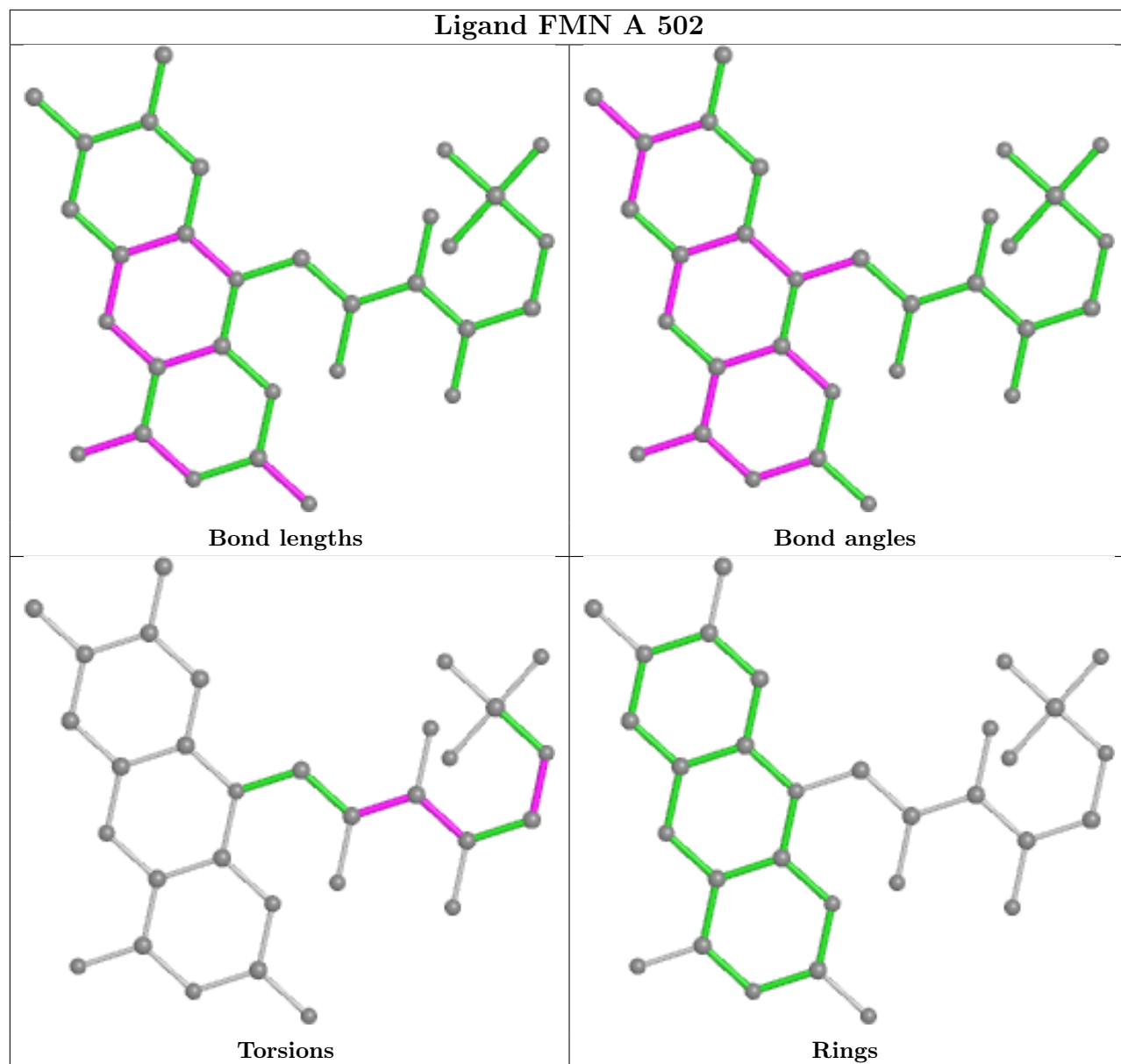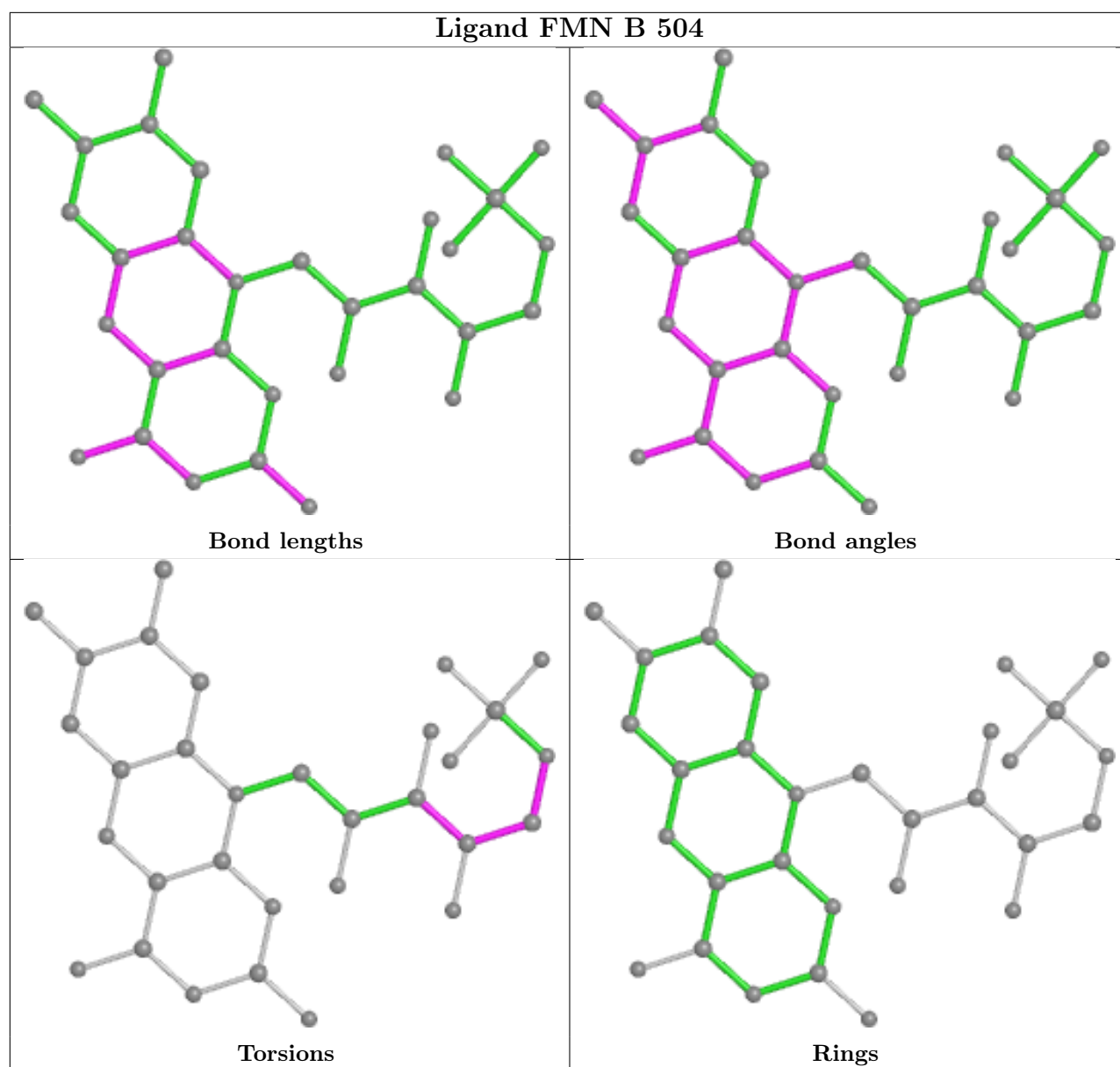
5.7 Other polymers ⓘ

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	313/482 (64%)	-0.02	5 (1%)	72	69	101, 129, 258, 276	0
1	B	460/482 (95%)	0.03	9 (1%)	65	63	101, 147, 221, 255	0
All	All	773/964 (80%)	0.01	14 (1%)	68	65	101, 139, 238, 276	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	298	GLN	4.8
1	B	454	CYS	4.6
1	B	350	LEU	4.5
1	A	322	LEU	3.3
1	B	451	LYS	3.0

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, 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	CL	B	502	1/1	0.79	0.30	117,117,117,117	0
2	CL	A	501	1/1	0.83	0.41	136,136,136,136	0
4	MG	B	503	1/1	0.88	0.24	223,223,223,223	0
3	FMN	B	504	31/31	0.92	0.17	115,125,137,141	0
3	FMN	A	502	31/31	0.92	0.19	120,134,139,147	0
5	ACP	B	505	31/31	0.93	0.21	148,163,177,182	0
2	CL	B	501	1/1	0.96	0.41	132,132,132,132	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.
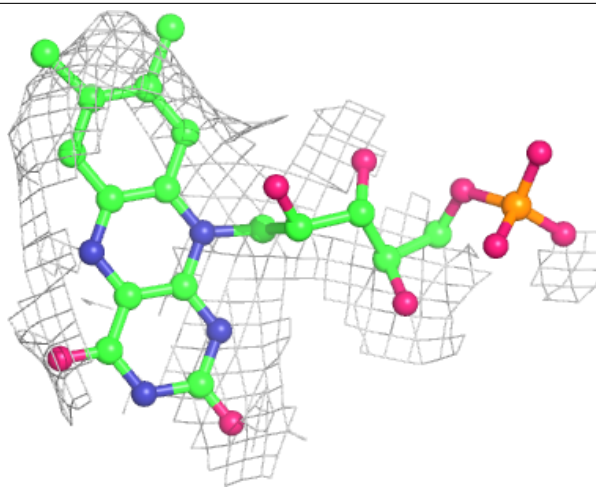
Electron density around MG B 503:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
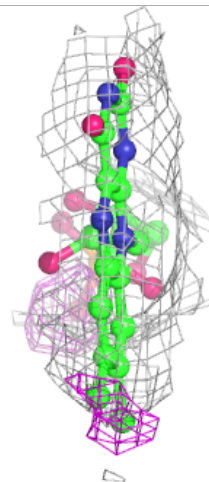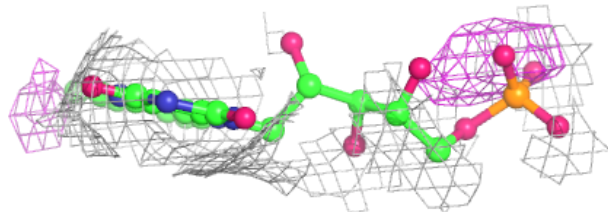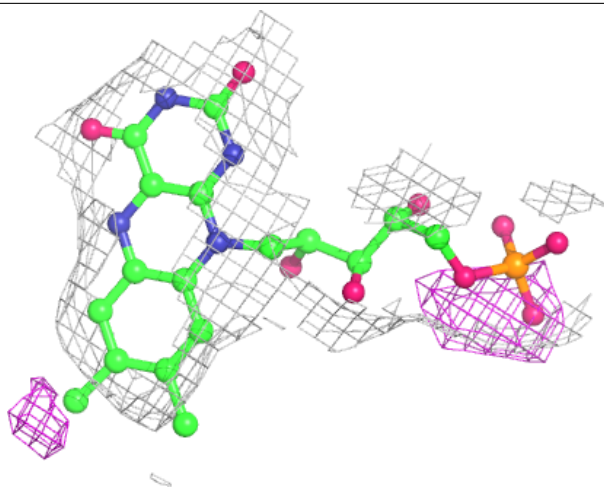$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
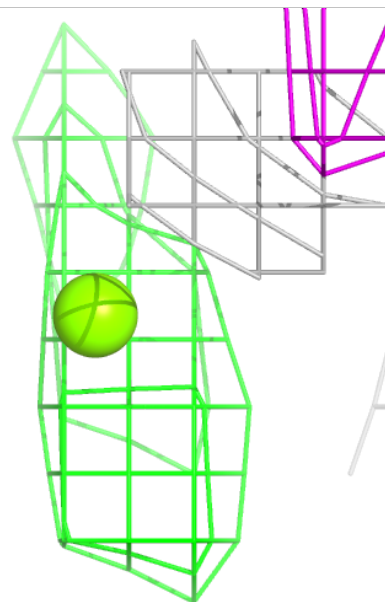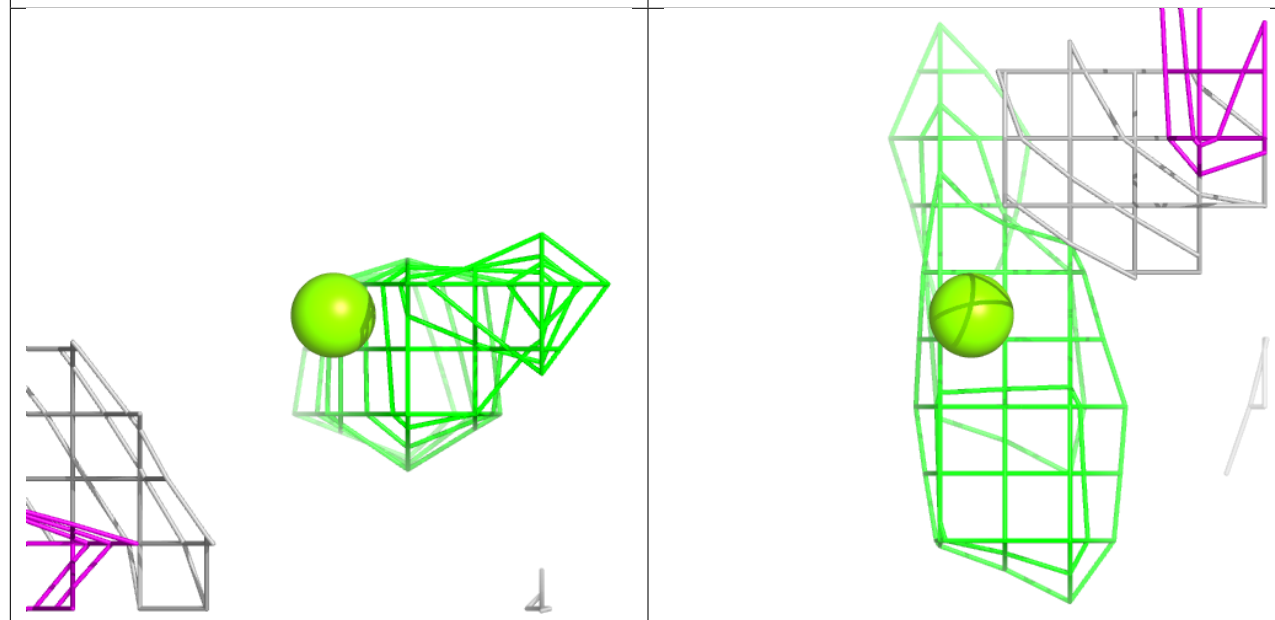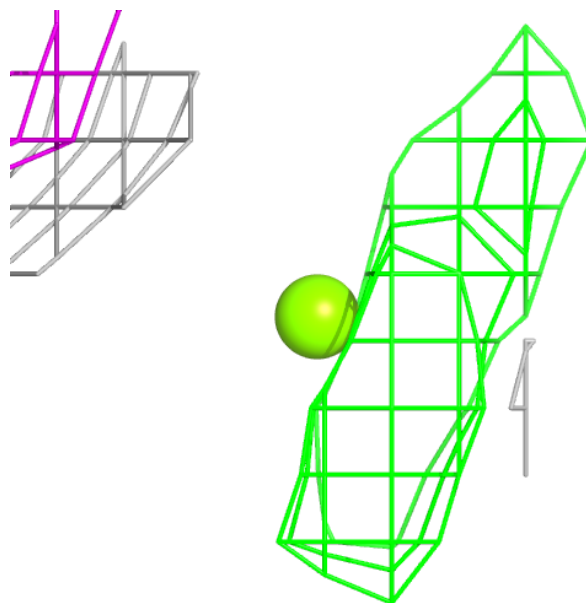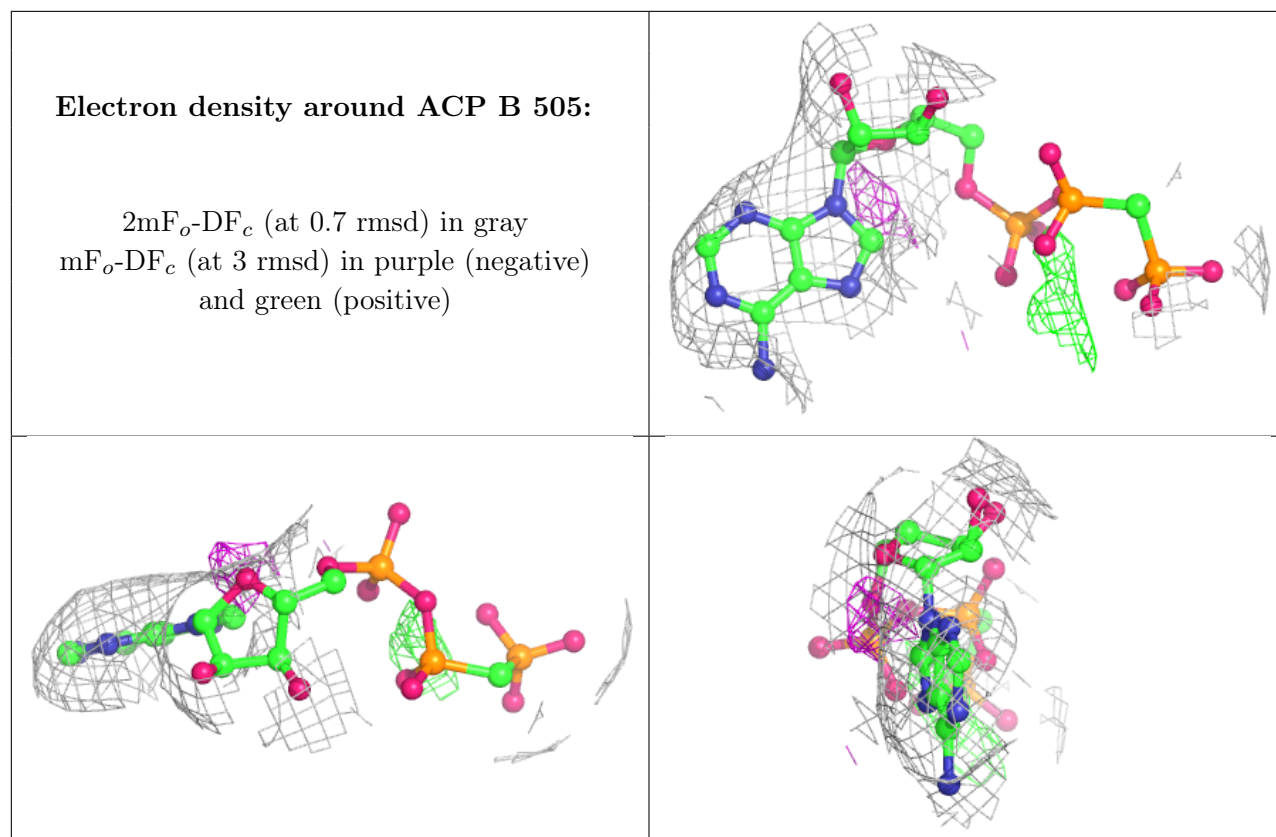
Electron density around FMN B 504:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
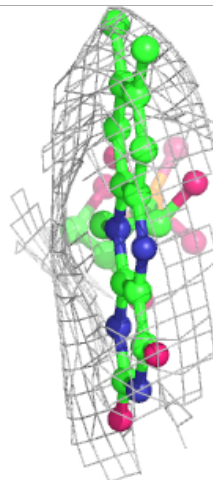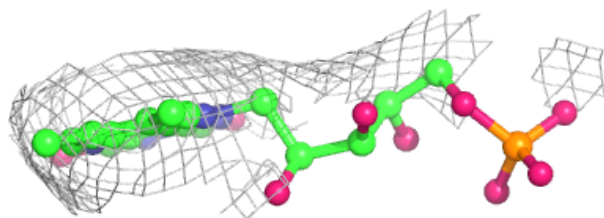
Electron density around FMN A 502:

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

Electron density around ACP B 505:

$2mF_o$-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.