wwPDB X-ray Structure Validation Summary Report (i)

Feb 19, 2024 – 04:18 PM JST

PDB ID	:	8IH7
Title	:	AmnG-AmnH complex
Authors	:	Su, D.; Shi, Q.L.
Deposited on	:	2023-02-22
Resolution	:	2.48 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:

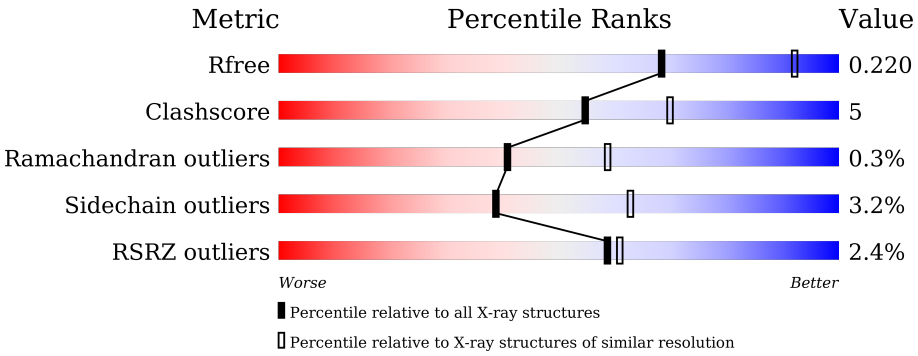
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

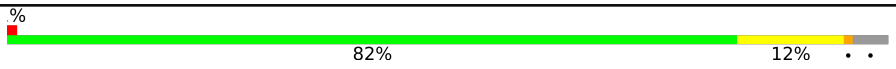
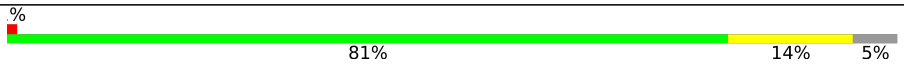
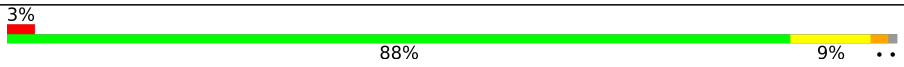
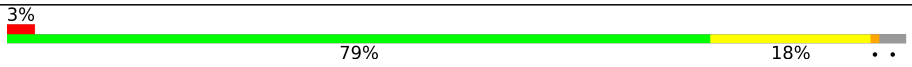
The reported resolution of this entry is 2.48 Å.

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}	130704	5857 (2.50-2.46)
Clashscore	141614	6594 (2.50-2.46)
Ramachandran outliers	138981	6469 (2.50-2.46)
Sidechain outliers	138945	6471 (2.50-2.46)
RSRZ outliers	127900	5738 (2.50-2.46)

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	355	
1	C	355	
2	B	319	
2	D	319	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	OXL	A	402	-	X	-	-
4	OXL	C	602	-	X	-	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 10057 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 4-hydroxy-2-oxovalerate aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	340	Total	C	N	O	S	0	0	0
			2565	1598	461	487	19			
1	C	339	Total	C	N	O	S	0	0	0
			2557	1594	460	484	19			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	HIS	-	expression tag	UNP Q9KWS0
A	-10	HIS	-	expression tag	UNP Q9KWS0
A	-9	HIS	-	expression tag	UNP Q9KWS0
A	-8	HIS	-	expression tag	UNP Q9KWS0
A	-7	HIS	-	expression tag	UNP Q9KWS0
A	-6	HIS	-	expression tag	UNP Q9KWS0
A	-5	LEU	-	expression tag	UNP Q9KWS0
A	-4	VAL	-	expression tag	UNP Q9KWS0
A	-3	PRO	-	expression tag	UNP Q9KWS0
A	-2	ARG	-	expression tag	UNP Q9KWS0
A	-1	GLY	-	expression tag	UNP Q9KWS0
A	0	SER	-	expression tag	UNP Q9KWS0
C	-11	HIS	-	expression tag	UNP Q9KWS0
C	-10	HIS	-	expression tag	UNP Q9KWS0
C	-9	HIS	-	expression tag	UNP Q9KWS0
C	-8	HIS	-	expression tag	UNP Q9KWS0
C	-7	HIS	-	expression tag	UNP Q9KWS0
C	-6	HIS	-	expression tag	UNP Q9KWS0
C	-5	LEU	-	expression tag	UNP Q9KWS0
C	-4	VAL	-	expression tag	UNP Q9KWS0
C	-3	PRO	-	expression tag	UNP Q9KWS0
C	-2	ARG	-	expression tag	UNP Q9KWS0
C	-1	GLY	-	expression tag	UNP Q9KWS0
C	0	SER	-	expression tag	UNP Q9KWS0

- Molecule 2 is a protein called Acetaldehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	317	Total	C	N	O	S	0	0	0
			2324	1466	399	448	11			
2	D	311	Total	C	N	O	S	0	0	0
			2297	1448	396	442	11			

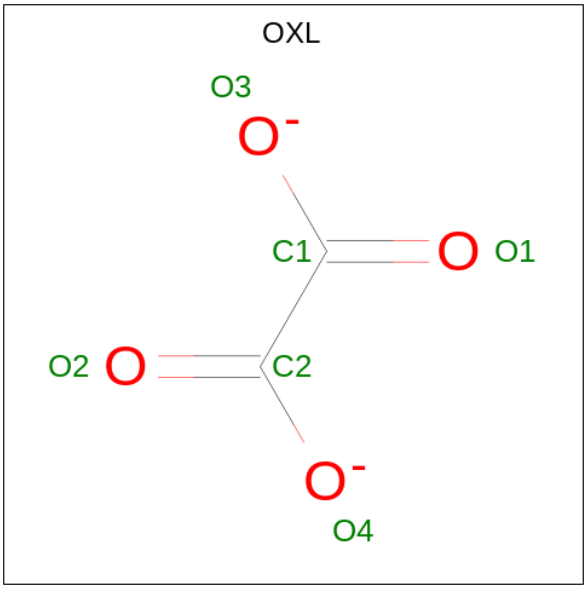
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	LEU	-	expression tag	UNP Q9KWS1
B	-1	ASN	-	expression tag	UNP Q9KWS1
B	0	TRP	-	expression tag	UNP Q9KWS1
D	-2	LEU	-	expression tag	UNP Q9KWS1
D	-1	ASN	-	expression tag	UNP Q9KWS1
D	0	TRP	-	expression tag	UNP Q9KWS1

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		

- Molecule 4 is OXALATE ION (three-letter code: OXL) (formula: C_2O_4) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 2 4	0	0
4	C	1	Total C O 6 2 4	0	0

- Molecule 5 is water.

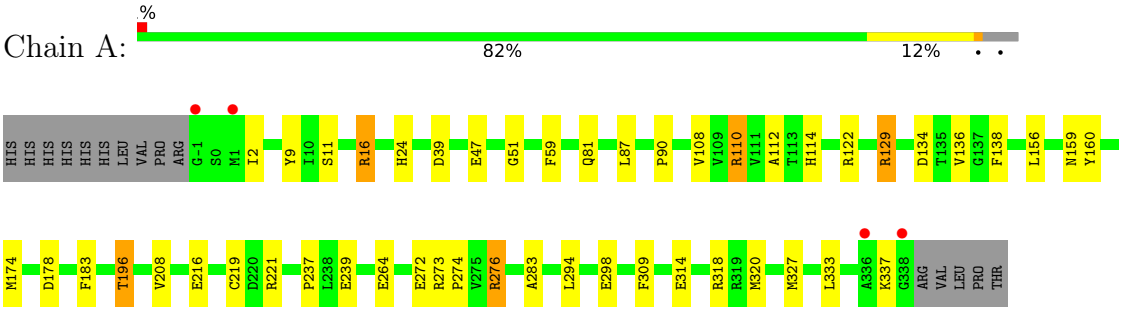
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	114	Total O 114 114	0	0
5	B	79	Total O 79 79	0	0
5	C	68	Total O 68 68	0	0
5	D	39	Total O 39 39	0	0

3 Residue-property plots (i)

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.

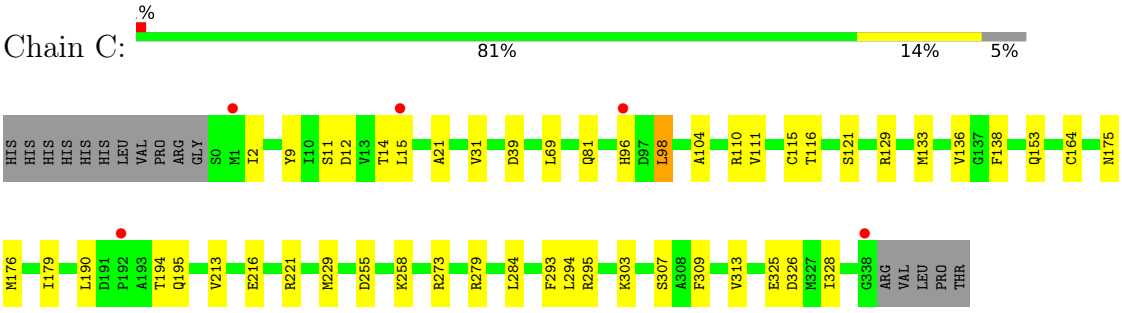
• Molecule 1: 4-hydroxy-2-oxovalerate aldolase

Chain A:



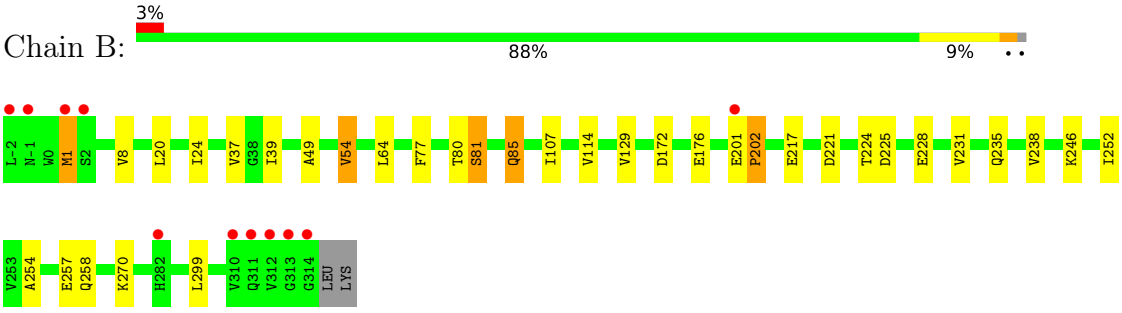
• Molecule 1: 4-hydroxy-2-oxovalerate aldolase

Chain C:



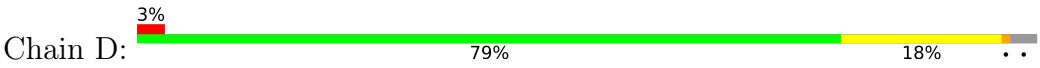
• Molecule 2: Acetaldehyde dehydrogenase

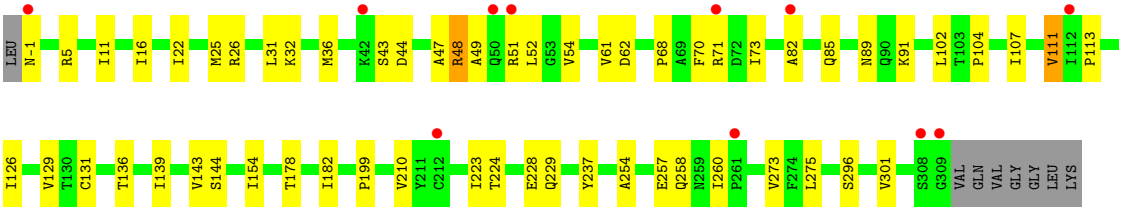
Chain B:



• Molecule 2: Acetaldehyde dehydrogenase

Chain D:




LEU N-1 R5 I11 I16 I22 M25 R26 L31 K32 M36 K42 S43 D44 A47 R48 A49 Q50 R51 L52 G53 V54 V61 D62 P68 A69 F70 R71 D72 I73 A82 Q85 M89 Q90 K91 L102 T103 P104 I107 V111 I112 P113
I126 V129 T130 C131 T136 I139 V143 S144 I154 T178 I182 P199 V210 Y211 C212 I223 T224 E228 Q229 Y237 A254 E257 Q258 N259 I260 P261 V273 F274 L275 S296 V301 S308 G309 VAL GLN VAL GLY GLY LEU LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	55.03Å 252.29Å 60.78Å 90.00° 114.94° 90.00°	Depositor
Resolution (Å)	20.11 – 2.48 20.11 – 2.48	Depositor EDS
% Data completeness (in resolution range)	99.9 (20.11-2.48) 99.9 (20.11-2.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.12 (at 2.47Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.173 , 0.223 0.177 , 0.220	Depositor DCC
R_{free} test set	1996 reflections (3.79%)	wwPDB-VP
Wilson B-factor (Å^2)	44.7	Xtriage
Anisotropy	0.559	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 34.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.013 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	10057	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% 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: OXL, ZN

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	0/2605	0.65	0/3517
1	C	0.42	0/2597	0.63	0/3507
2	B	0.45	0/2362	0.67	2/3213 (0.1%)
2	D	0.41	0/2336	0.60	0/3177
All	All	0.44	0/9900	0.64	2/13414 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
2	D	0	1
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	202	PRO	N-CA-CB	8.52	113.52	103.30
2	B	202	PRO	N-CA-C	-5.15	98.70	112.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	129	ARG	Sidechain
1	A	16	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	D	26	ARG	Sidechain

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	2565	0	2546	27	0
1	C	2557	0	2539	26	0
2	B	2324	0	2344	18	0
2	D	2297	0	2324	33	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
4	A	6	0	0	0	0
4	C	6	0	0	0	0
5	A	114	0	0	1	0
5	B	79	0	0	1	0
5	C	68	0	0	1	0
5	D	39	0	0	3	0
All	All	10057	0	9753	101	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 5.

The worst 5 of 101 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)
2:D:139:ILE:HG21	2:D:273:VAL:HG11	1.45	0.95
2:D:229:GLN:NE2	5:D:401:HOH:O	2.16	0.79
2:B:252:ILE:HD13	2:B:258:GLN:HG3	1.69	0.75
1:A:196:THR:HG22	1:A:219:CYS:HA	1.74	0.70
1:C:284:LEU:HD11	1:C:294:LEU:HB2	1.75	0.69

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	338/355 (95%)	329 (97%)	9 (3%)	0	100	100
1	C	337/355 (95%)	323 (96%)	13 (4%)	1 (0%)	41	59
2	B	315/319 (99%)	306 (97%)	7 (2%)	2 (1%)	25	40
2	D	309/319 (97%)	297 (96%)	11 (4%)	1 (0%)	41	59
All	All	1299/1348 (96%)	1255 (97%)	40 (3%)	4 (0%)	41	59

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	202	PRO
1	C	229	MET
2	B	129	VAL
2	D	129	VAL

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	261/276 (95%)	253 (97%)	8 (3%)	40	64
1	C	260/276 (94%)	254 (98%)	6 (2%)	50	74
2	B	246/252 (98%)	237 (96%)	9 (4%)	34	57
2	D	245/252 (97%)	236 (96%)	9 (4%)	34	57
All	All	1012/1056 (96%)	980 (97%)	32 (3%)	39	63

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	102	LEU
2	D	111	VAL
2	B	85	GLN
2	B	81	SER
2	D	131	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	258	GLN
2	D	93	GLN
2	D	27	ASN
2	B	229	GLN
2	D	50	GLN

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 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\#\|Z\|>2$
4	OXL	A	402	3	5,5,5	2.64	4 (80%)	6,6,6	1.80	2 (33%)
4	OXL	C	602	3	5,5,5	2.51	4 (80%)	6,6,6	1.89	3 (50%)

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
4	OXL	A	402	3	-	2/4/4/4	-
4	OXL	C	602	3	-	4/4/4/4	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	402	OXL	O2-C2	3.12	1.30	1.22
4	A	402	OXL	O4-C2	-3.09	1.21	1.30
4	C	602	OXL	O2-C2	3.08	1.30	1.22
4	A	402	OXL	O1-C1	2.87	1.30	1.22
4	C	602	OXL	O1-C1	2.85	1.30	1.22

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	402	OXL	O3-C1-C2	2.52	120.64	113.16
4	C	602	OXL	O3-C1-C2	2.49	120.56	113.16
4	A	402	OXL	O4-C2-C1	2.48	120.52	113.16
4	C	602	OXL	O4-C2-C1	2.43	120.38	113.16
4	C	602	OXL	O1-C1-C2	-2.09	114.06	120.78

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

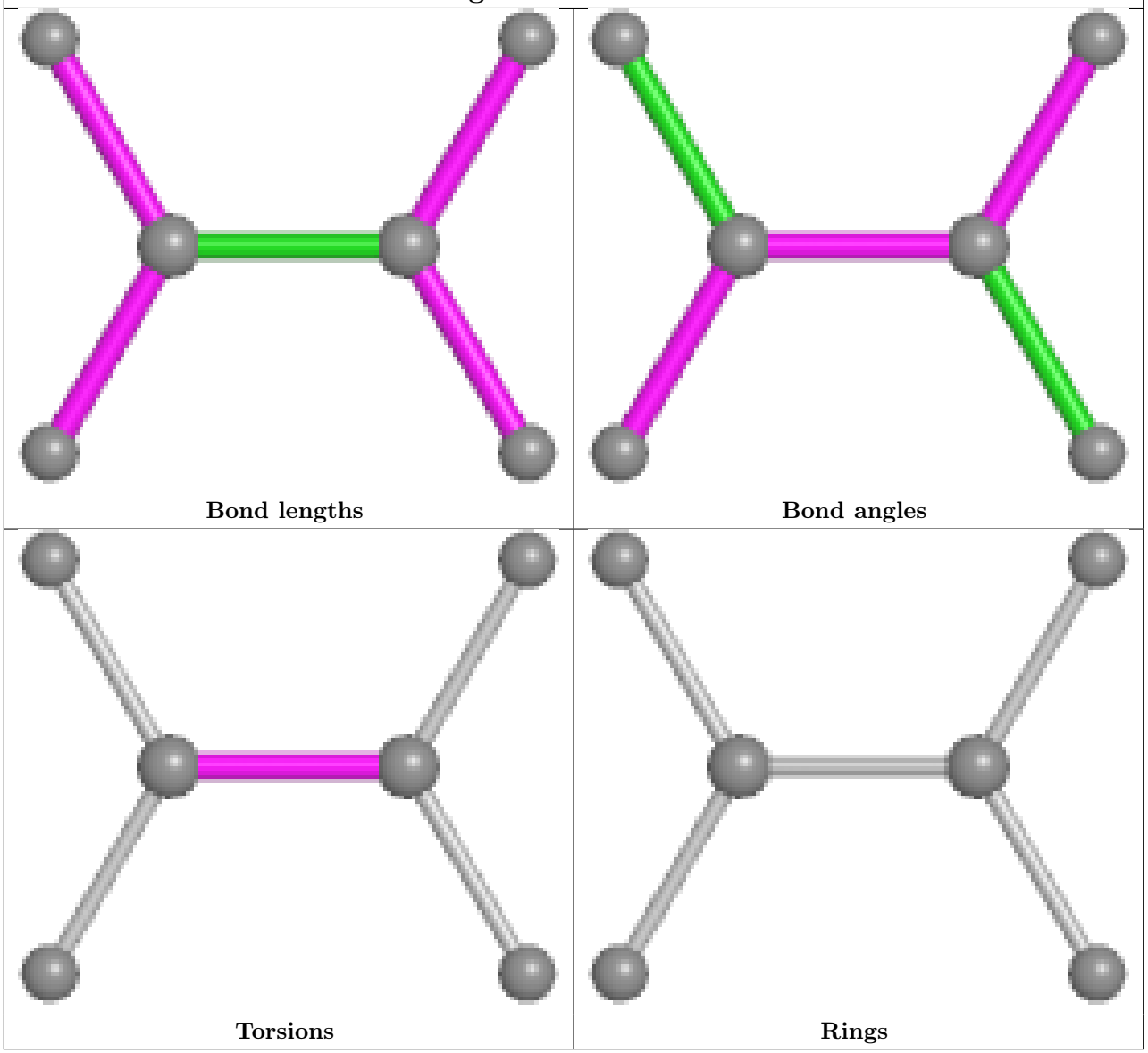
Mol	Chain	Res	Type	Atoms
4	C	602	OXL	O1-C1-C2-O2
4	C	602	OXL	O1-C1-C2-O4
4	C	602	OXL	O3-C1-C2-O4
4	C	602	OXL	O3-C1-C2-O2
4	A	402	OXL	O3-C1-C2-O4

There are no ring outliers.

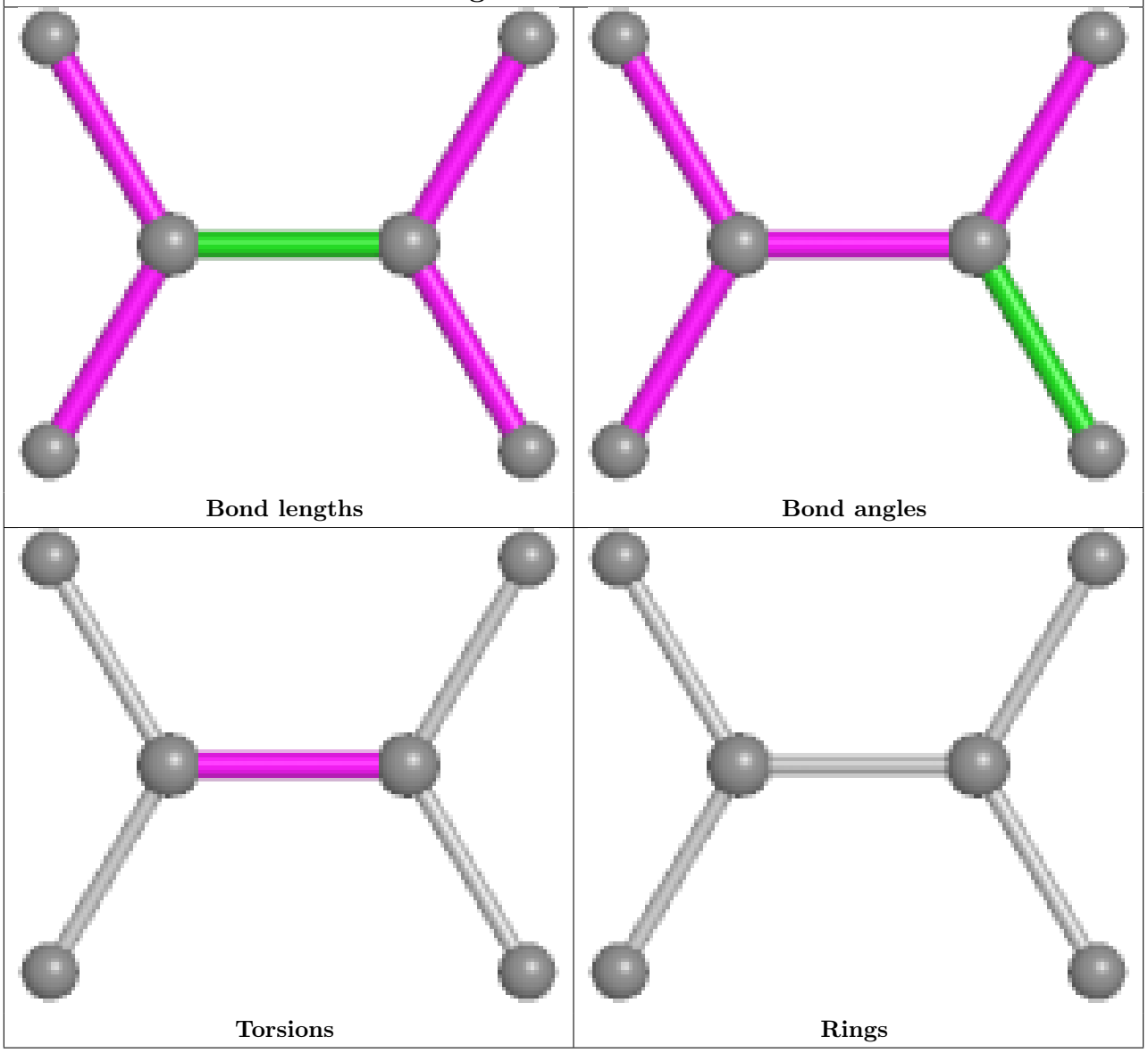
No monomer is involved in short contacts.

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 OXL A 402



Ligand OXL C 602



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, 95^{th} 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	340/355 (95%)	-0.35	4 (1%)	79	80	36, 45, 57, 80	0
1	C	339/355 (95%)	-0.28	5 (1%)	73	75	40, 52, 66, 96	0
2	B	317/319 (99%)	-0.14	11 (3%)	44	46	30, 46, 72, 116	0
2	D	311/319 (97%)	0.03	11 (3%)	44	46	47, 62, 88, 108	0
All	All	1307/1348 (96%)	-0.19	31 (2%)	59	61	30, 51, 78, 116	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	-2	LEU	7.0
2	B	314	GLY	5.5
2	B	310	VAL	5.3
2	B	312	VAL	5.1
1	C	1	MET	4.4

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

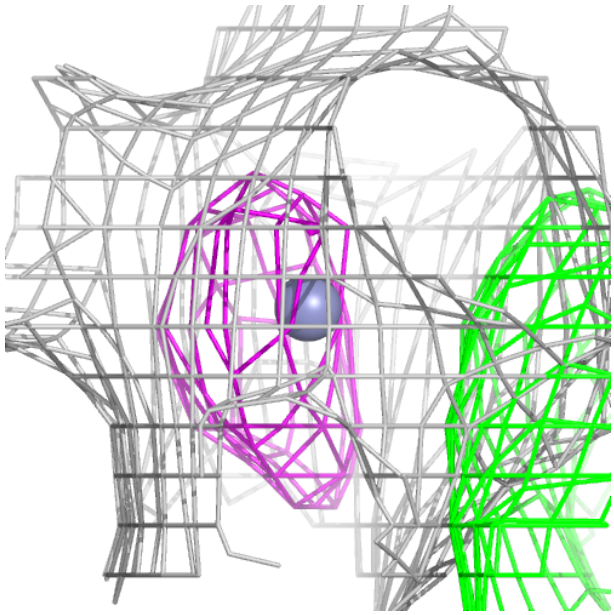
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

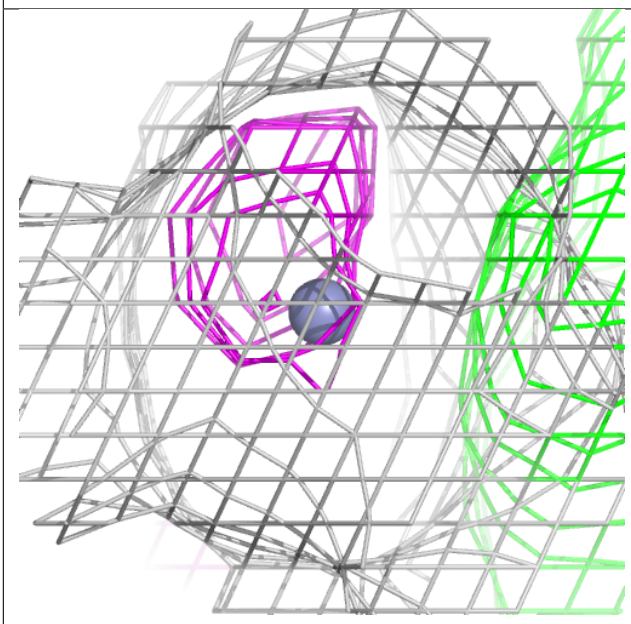
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	ZN	A	401	1/1	0.96	0.03	53,53,53,53	0
4	OXL	A	402	6/6	0.97	0.13	41,44,46,54	0
3	ZN	C	601	1/1	0.98	0.03	59,59,59,59	0
4	OXL	C	602	6/6	0.98	0.15	48,52,53,53	0

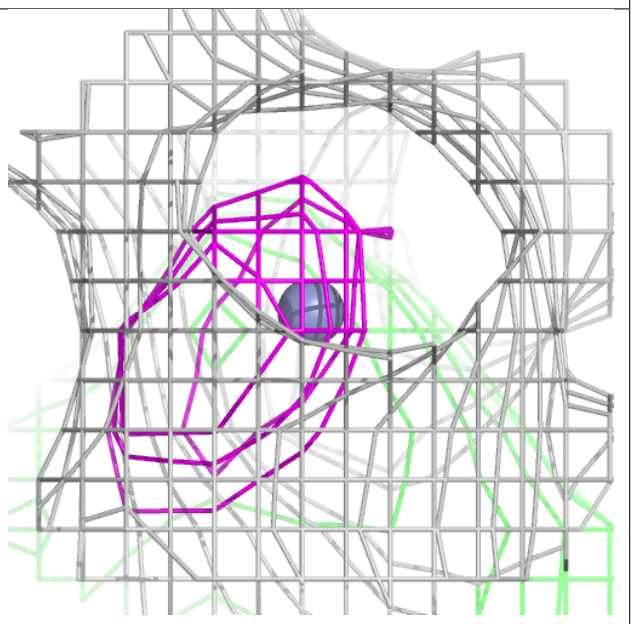
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ZN 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)

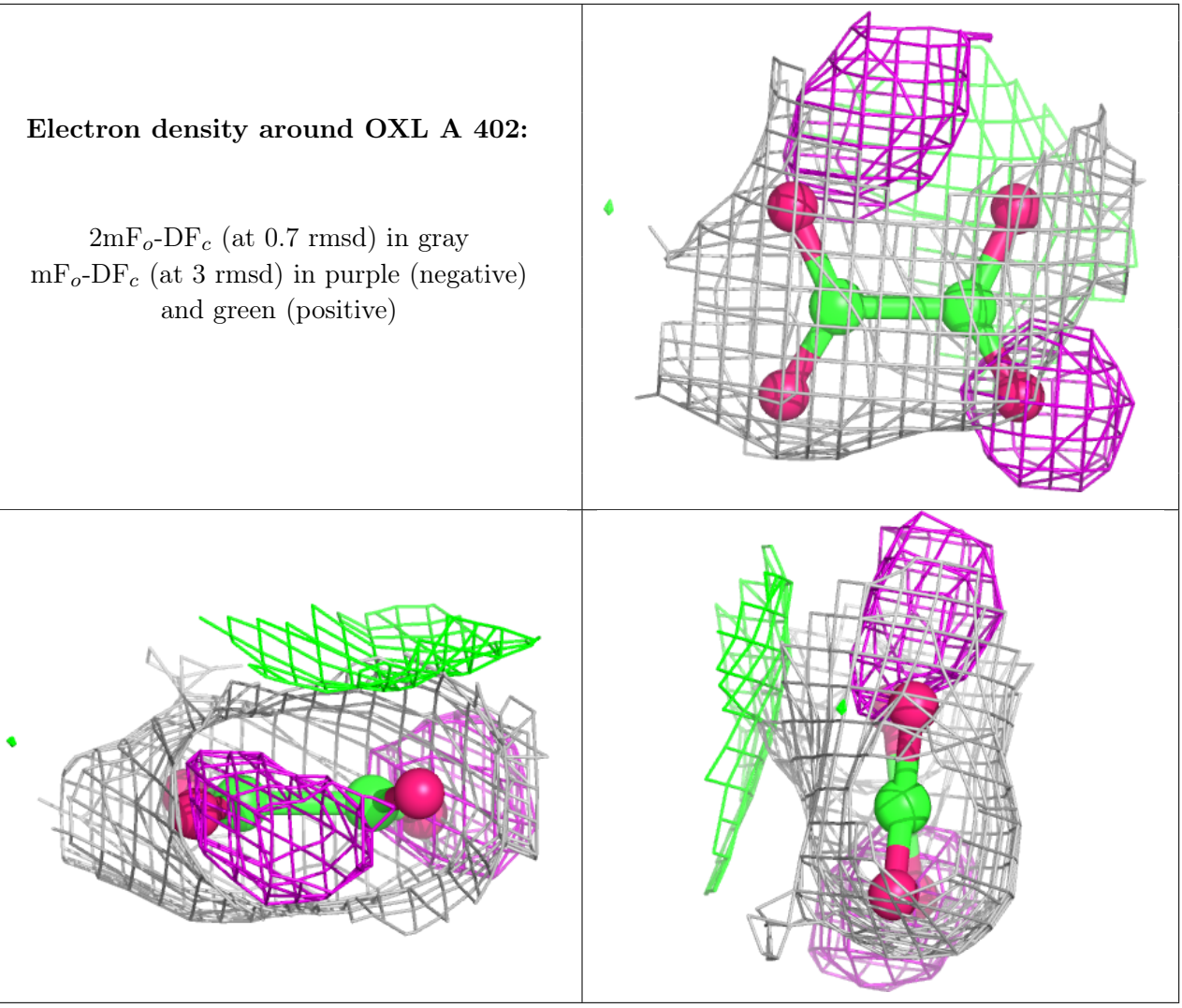






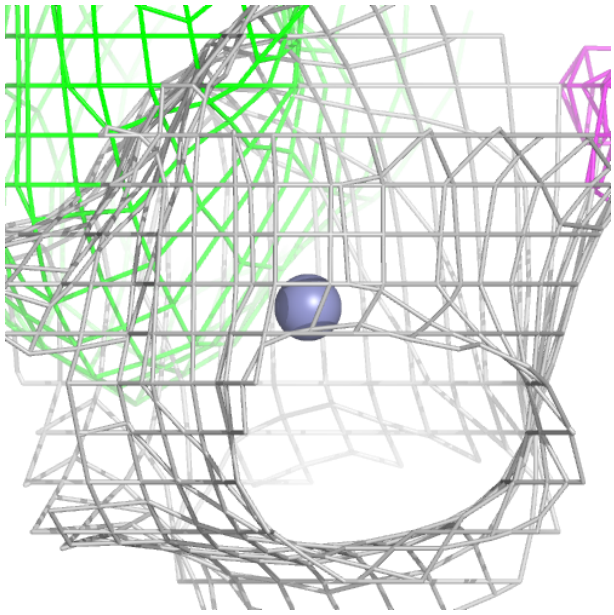
Electron density around OXL A 402:

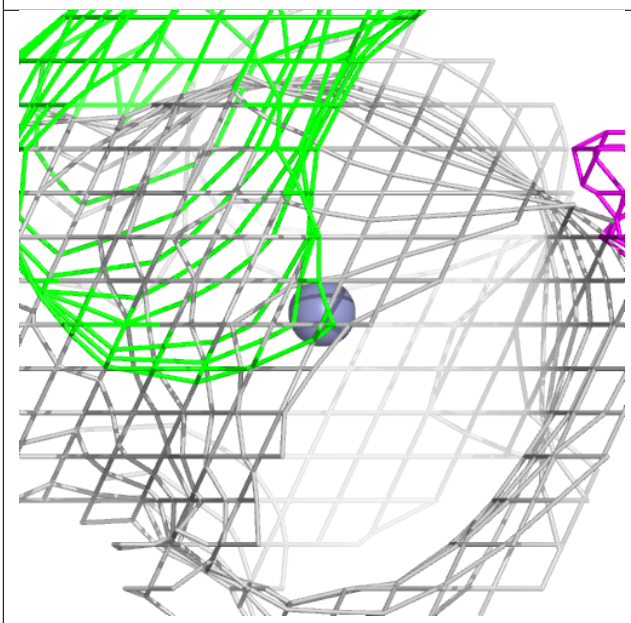
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

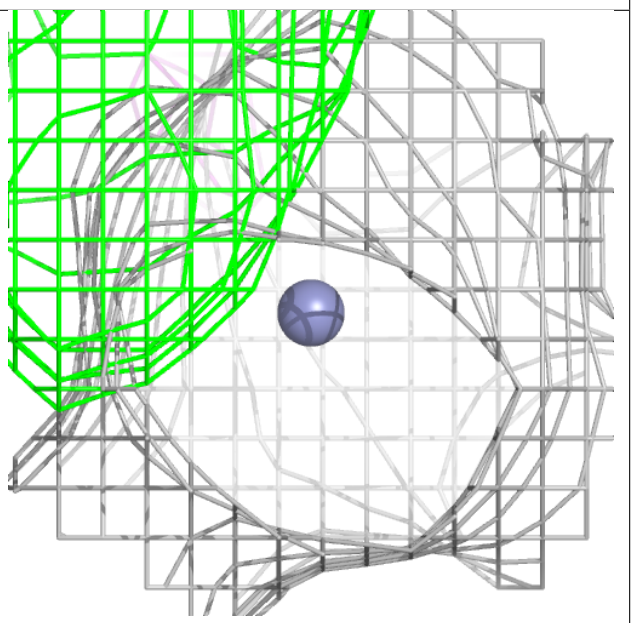


Electron density around ZN C 601:

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

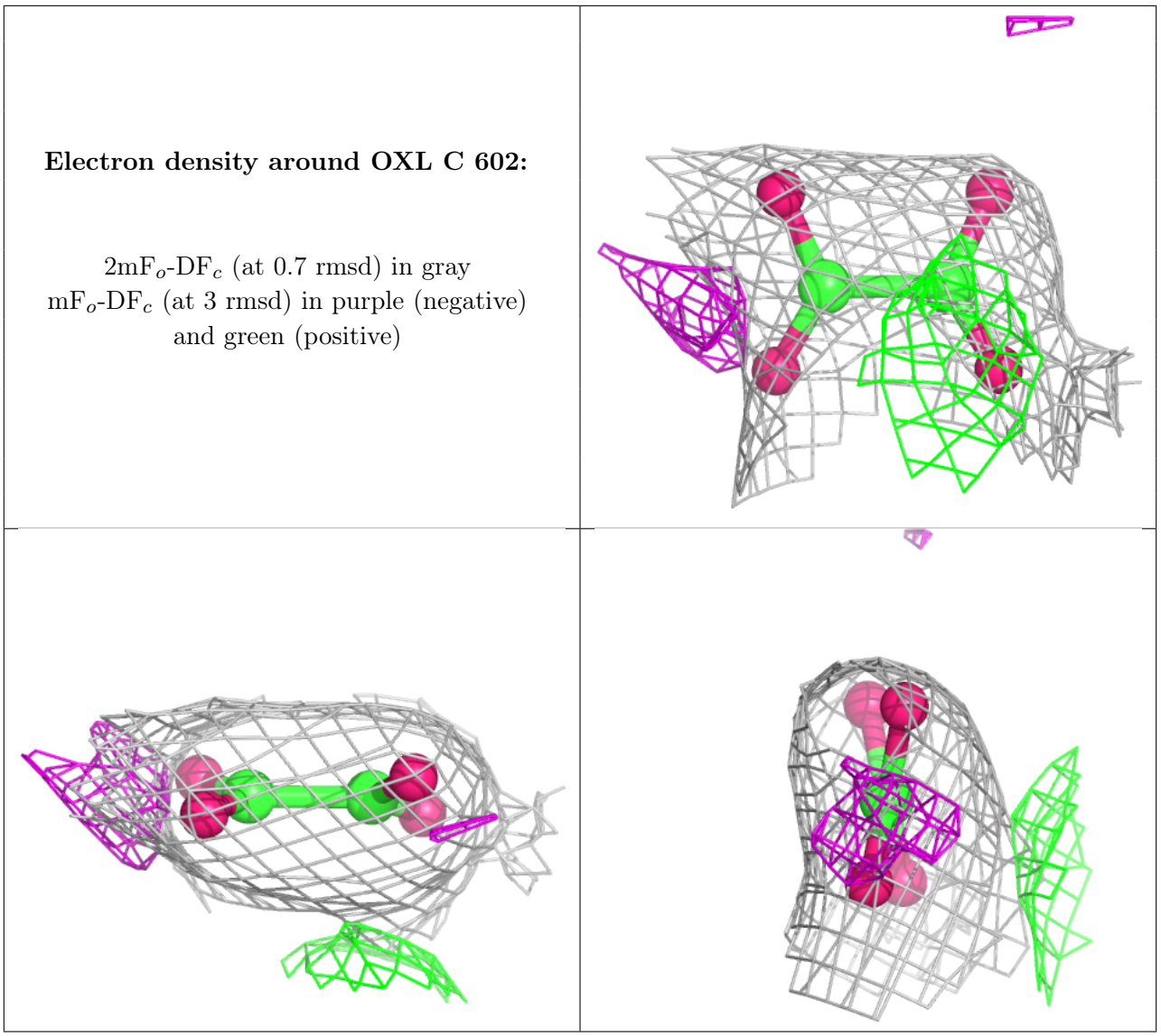






Electron density around OXL C 602:

$2mF_o$-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.