



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 4, 2025 – 06:10 PM EDT

PDB ID : 9CSJ / pdb_00009csj
Title : Crystal structure of human glyoxalase domain-containing protein 4 (GLOD4) at 2.33 Å resolution.
Authors : Griswold-Prenner, I.; Dou, Y.; Jennings, A.; Kayser, F.
Deposited on : 2024-07-24
Resolution : 2.33 Å (reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

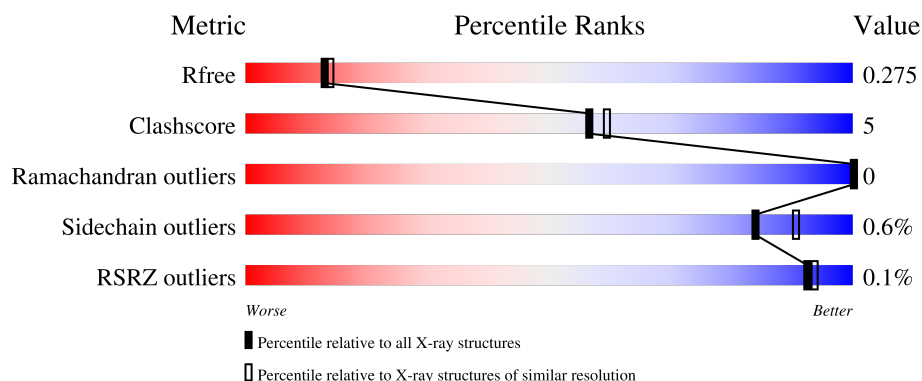
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION




The reported resolution of this entry is 2.33 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}	164625	2747 (2.36-2.32)
Clashscore	180529	2936 (2.36-2.32)
Ramachandran outliers	177936	2912 (2.36-2.32)
Sidechain outliers	177891	2912 (2.36-2.32)
RSRZ outliers	164620	2747 (2.36-2.32)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	299	
1	B	299	
1	C	299	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7425 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 Isoform 2 of Glyoxalase domain-containing protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	298	Total	C	N	O	S	0	3	0
			2364	1500	401	448	15			
1	B	297	Total	C	N	O	S	0	3	0
			2356	1495	399	448	14			
1	C	273	Total	C	N	O	S	0	0	0
			2168	1382	368	407	11			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q9HC38
B	0	SER	-	expression tag	UNP Q9HC38
C	0	SER	-	expression tag	UNP Q9HC38

- Molecule 2 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

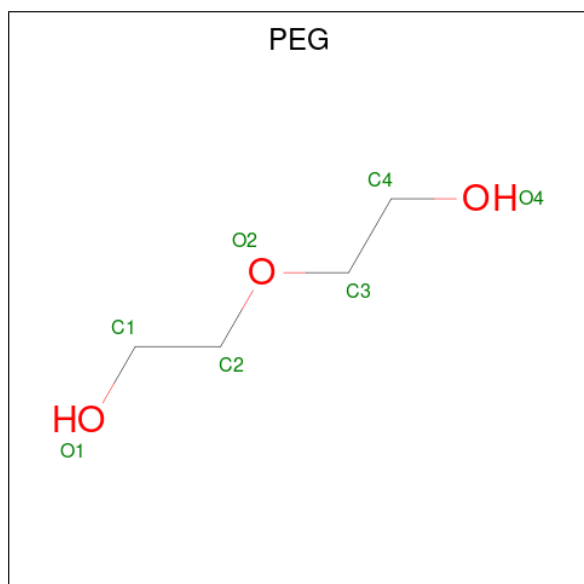
- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 4 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	7	Total	Cl	0	0
			7	7		
4	B	4	Total	Cl	0	0
			4	4		
4	C	1	Total	Cl	0	0
			1	1		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (CCD ID: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
6	C	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	191	Total	O	0	2
			191	191		
7	B	202	Total	O	0	0
			202	202		
7	C	80	Total	O	0	0
			80	80		

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: Isoform 2 of Glyoxalase domain-containing protein 4

Chain A: 




- Molecule 1: Isoform 2 of Glyoxalase domain-containing protein 4

Chain B: 



- Molecule 1: Isoform 2 of Glyoxalase domain-containing protein 4

Chain C: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	85.96Å 85.96Å 135.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.34 – 2.33 45.34 – 2.33	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.34-2.33) 99.9 (45.34-2.33)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.36 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.220 , 0.269 0.225 , 0.275	Depositor DCC
R_{free} test set	2091 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å ²)	39.4	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 29.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.487 for h,-k,-l	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	7425	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PEG, MG, EPE, ZN, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	1.03	1/2416 (0.0%)	1.31	1/3257 (0.0%)
1	B	1.03	0/2408	1.31	0/3247
1	C	1.04	0/2214	1.33	1/2984 (0.0%)
All	All	1.04	1/7038 (0.0%)	1.32	2/9488 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	294	PRO	C-O	-5.55	1.17	1.23

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	286	GLU	CB-CG-CD	6.28	123.27	112.60
1	C	79	ASP	CA-CB-CG	5.03	117.63	112.60

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	2364	0	2310	15	0
1	B	2356	0	2296	17	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2168	0	2130	32	0
2	A	1	0	0	0	0
2	B	2	0	0	0	1
3	A	3	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	7	0	0	1	0
4	B	4	0	0	1	0
4	C	1	0	0	0	0
5	A	7	0	10	2	0
5	B	7	0	10	1	0
6	B	15	0	18	0	0
6	C	15	0	18	0	0
7	A	191	0	0	1	0
7	B	202	0	0	3	1
7	C	80	0	0	2	0
All	All	7425	0	6792	62	1

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 62 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:128:GLN:OE1	1:C:130:ARG:NH2	2.08	0.86
1:C:10:VAL:HG21	1:C:201:ARG:NE	1.97	0.79
1:C:224:LYS:HB2	1:C:246:ALA:HB3	1.72	0.72
1:C:53:TRP:HZ3	1:C:55:LYS:HE3	1.55	0.71
1:B:151:SER:OG	7:B:401:HOH:O	2.08	0.70

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:302:MG:MG	7:B:507:HOH:O[3_555]	1.65	0.55

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	299/299 (100%)	289 (97%)	10 (3%)	0	100	100
1	B	298/299 (100%)	288 (97%)	10 (3%)	0	100	100
1	C	267/299 (89%)	258 (97%)	9 (3%)	0	100	100
All	All	864/897 (96%)	835 (97%)	29 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	249/247 (101%)	248 (100%)	1 (0%)	89	94
1	B	248/247 (100%)	246 (99%)	2 (1%)	79	87
1	C	230/247 (93%)	229 (100%)	1 (0%)	89	94
All	All	727/741 (98%)	723 (99%)	4 (1%)	84	90

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	247	ASP
1	B	45	CYS
1	B	247	ASP
1	C	247	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	18	GLN
1	C	74	ASN
1	C	241	GLN
1	C	149	GLN
1	B	223	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 24 ligands modelled in this entry, 20 are monoatomic - leaving 4 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	PEG	B	309	-	6,6,6	0.21	0	5,5,5	0.11	0
6	EPE	C	301	-	15,15,15	0.70	1 (6%)	19,20,20	0.72	0
6	EPE	B	303	-	15,15,15	0.66	1 (6%)	19,20,20	0.95	1 (5%)
5	PEG	A	312	-	6,6,6	0.19	0	5,5,5	0.14	0

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	PEG	B	309	-	-	2/4/4/4	-
6	EPE	C	301	-	-	6/9/19/19	0/1/1/1
6	EPE	B	303	-	-	3/9/19/19	0/1/1/1
5	PEG	A	312	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	301	EPE	O3S-S	2.32	1.55	1.47
6	B	303	EPE	O3S-S	2.28	1.55	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	303	EPE	O3S-S-C10	-2.10	101.89	106.00

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	301	EPE	S-C10-C9-N1
6	C	301	EPE	C9-C10-S-O1S
6	C	301	EPE	C9-C10-S-O2S
6	C	301	EPE	C9-C10-S-O3S
6	B	303	EPE	N4-C7-C8-O8

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	309	PEG	1	0
5	A	312	PEG	2	0

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 [i](#)

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

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å ²)	Q<0.9
1	A	298/299 (99%)	-1.53	0	100 100	19, 37, 59, 109	3 (1%)
1	B	297/299 (99%)	-1.53	0	100 100	17, 36, 58, 102	3 (1%)
1	C	273/299 (91%)	-1.03	1 (0%)	89 90	53, 75, 106, 141	0
All	All	868/897 (96%)	-1.37	1 (0%)	92 94	17, 42, 93, 141	6 (0%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	288	PHE	2.5

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

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	EPE	B	303	15/15	0.97	0.07	107,109,113,114	0
6	EPE	C	301	15/15	0.98	0.05	74,86,96,101	0

Continued on next page...

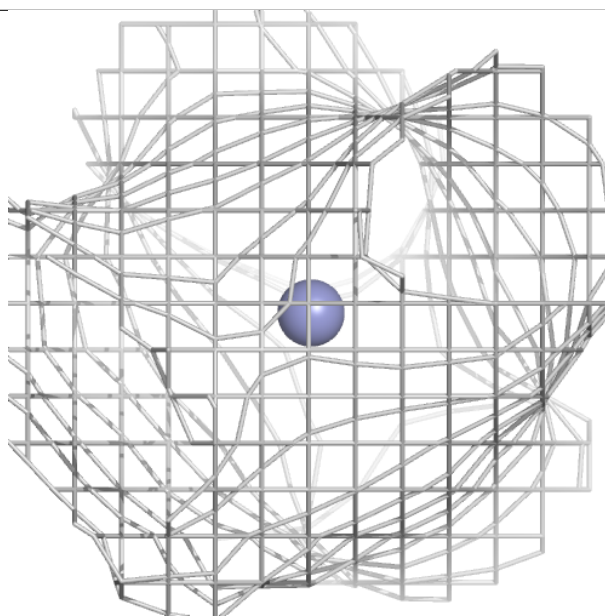
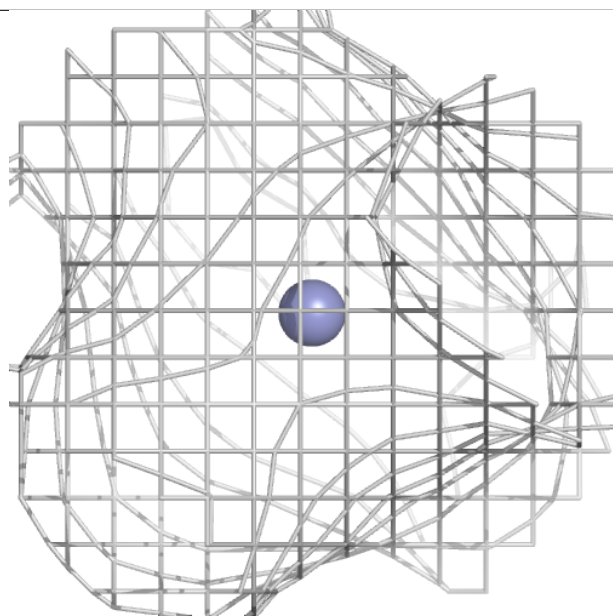
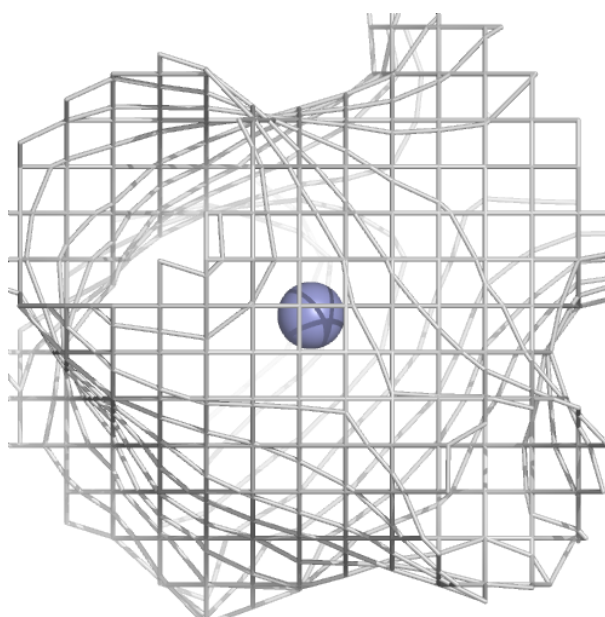
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	CL	B	308	1/1	0.99	0.07	86,86,86,86	0
5	PEG	A	312	7/7	0.99	0.04	55,56,57,57	0
5	PEG	B	309	7/7	0.99	0.06	59,60,62,66	0
2	MG	B	302	1/1	0.99	0.02	57,57,57,57	0
4	CL	B	306	1/1	0.99	0.03	48,48,48,48	0
3	ZN	C	302	1/1	1.00	0.03	57,57,57,57	1
4	CL	A	305	1/1	1.00	0.01	36,36,36,36	0
4	CL	A	306	1/1	1.00	0.01	32,32,32,32	0
4	CL	A	307	1/1	1.00	0.03	33,33,33,33	0
4	CL	A	308	1/1	1.00	0.01	36,36,36,36	0
4	CL	A	309	1/1	1.00	0.02	32,32,32,32	0
4	CL	A	310	1/1	1.00	0.03	46,46,46,46	0
4	CL	A	311	1/1	1.00	0.03	40,40,40,40	0
4	CL	B	305	1/1	1.00	0.01	35,35,35,35	0
2	MG	B	301	1/1	1.00	0.02	34,34,34,34	0
4	CL	B	307	1/1	1.00	0.06	41,41,41,41	0
2	MG	A	301	1/1	1.00	0.04	37,37,37,37	0
4	CL	C	303	1/1	1.00	0.06	68,68,68,68	0
3	ZN	A	302	1/1	1.00	0.01	35,35,35,35	0
3	ZN	A	303	1/1	1.00	0.01	30,30,30,30	0
3	ZN	A	304	1/1	1.00	0.01	29,29,29,29	0
3	ZN	B	304	1/1	1.00	0.01	35,35,35,35	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.

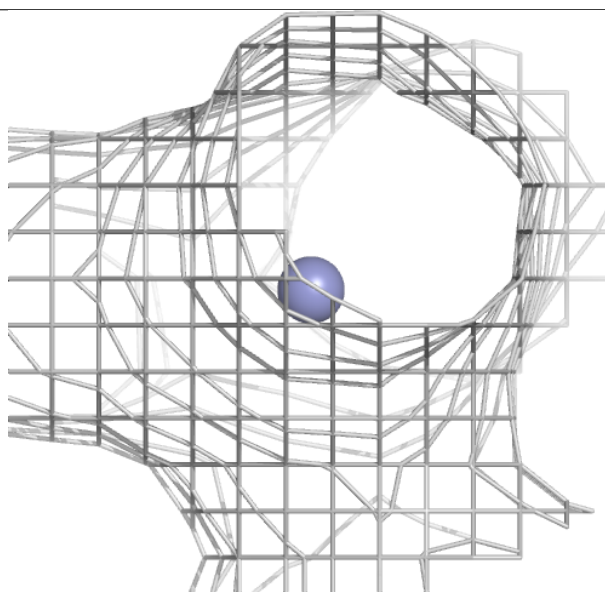
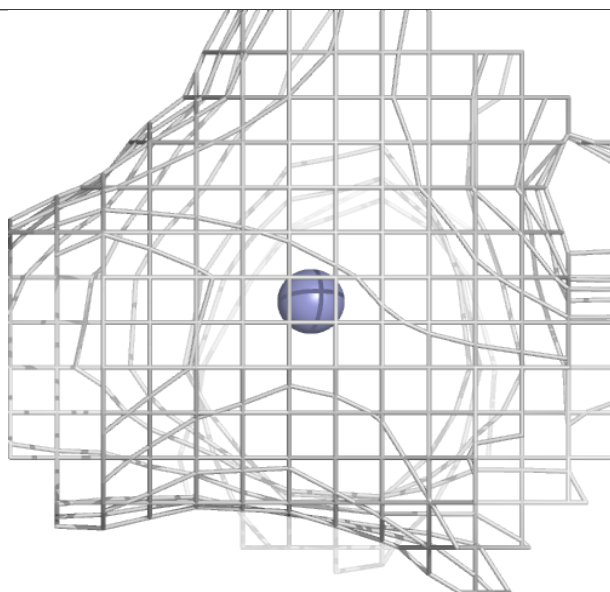
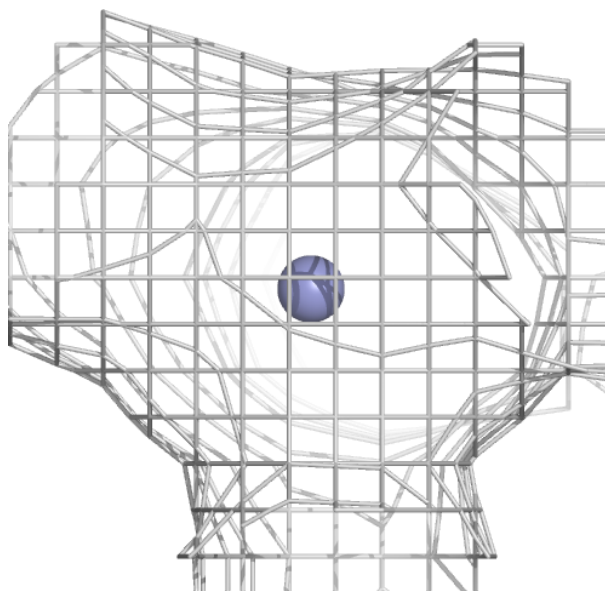
Electron density around ZN C 302:

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



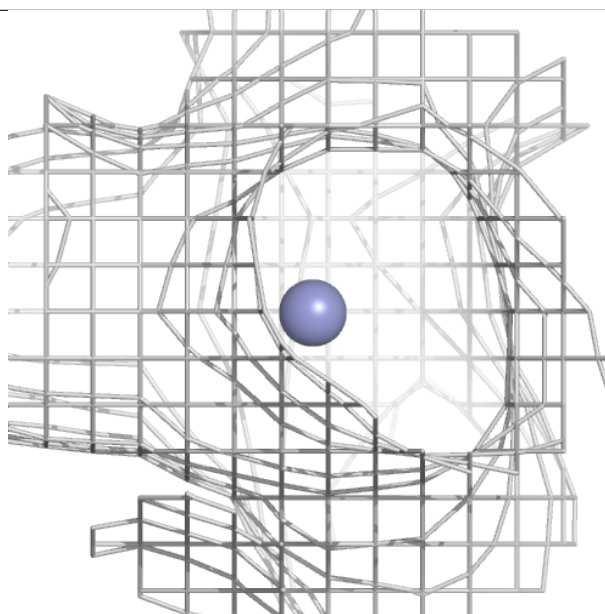
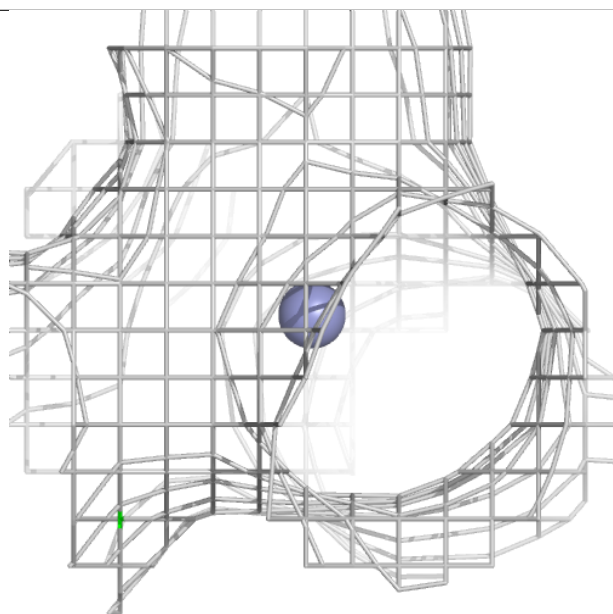
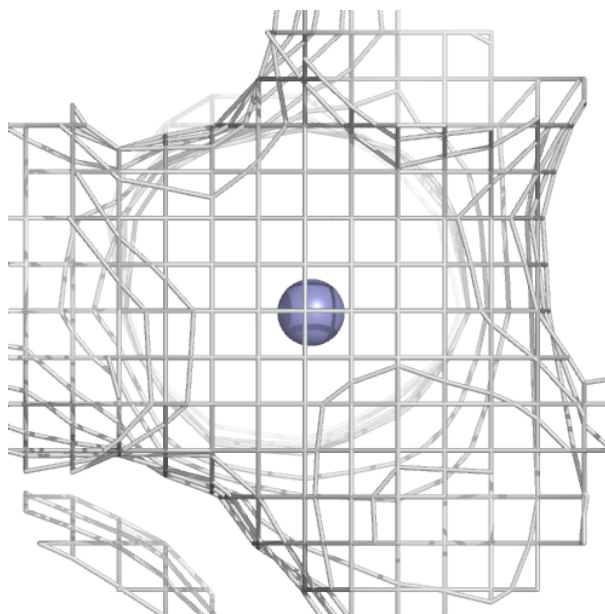
Electron density around ZN A 302:

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



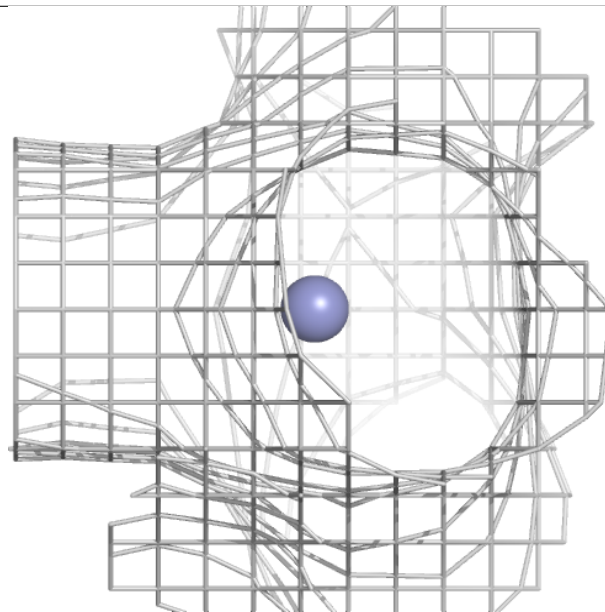
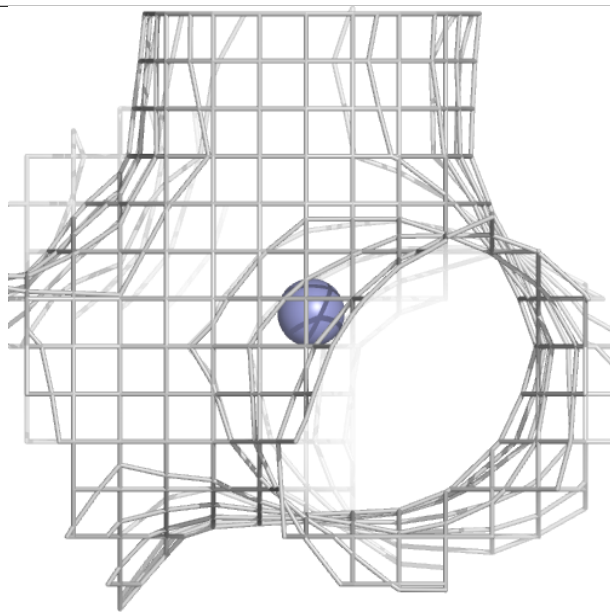
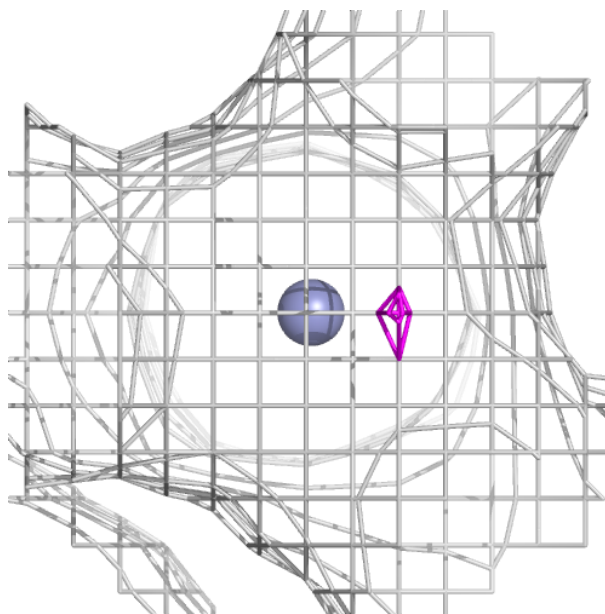
Electron density around ZN A 303:

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



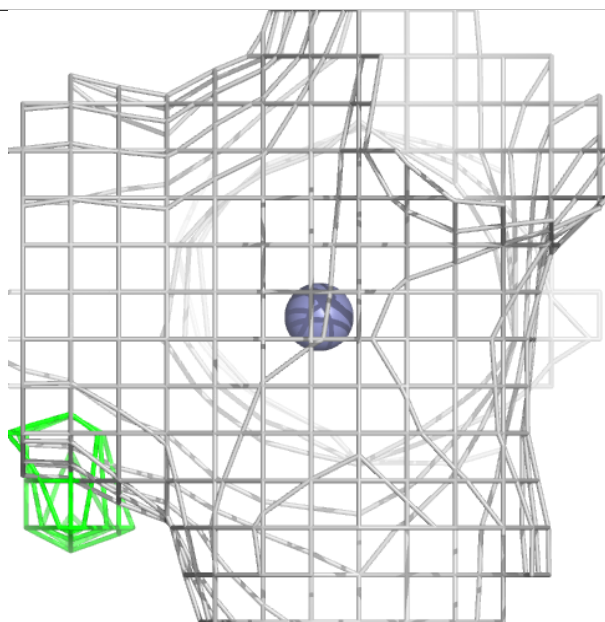
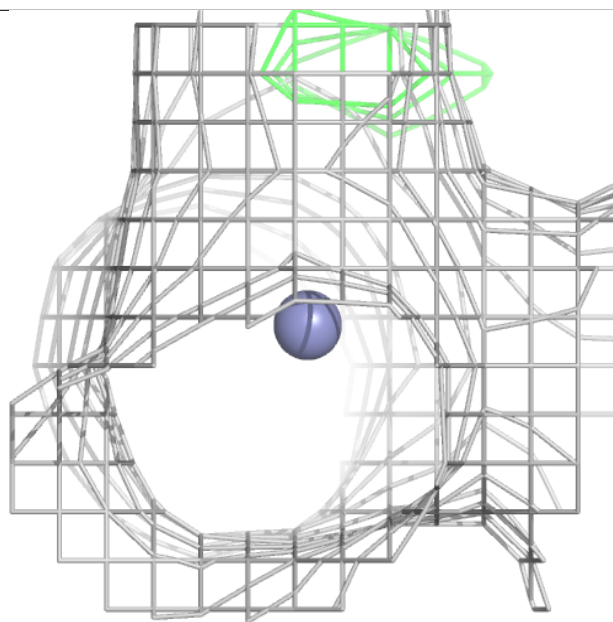
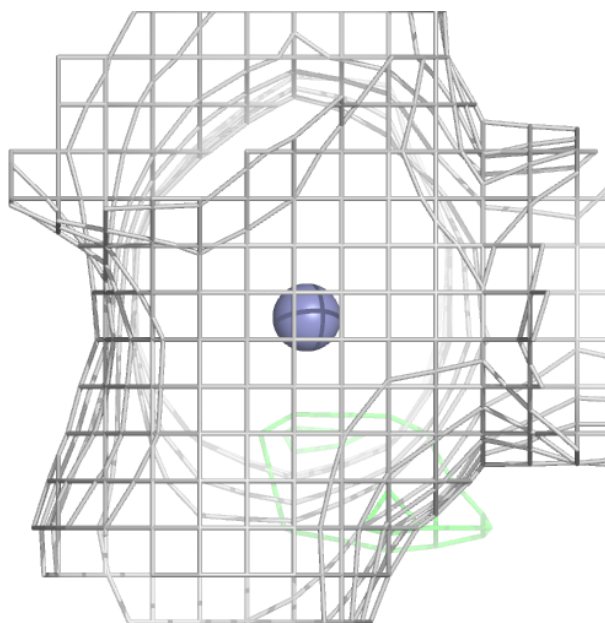
Electron density around ZN A 304:

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



Electron density around ZN B 304:

$2mF_o-DF_c$ (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.