



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 05:12 PM UTC

PDB ID : 9SKH / pdb_00009skh
Title : Biocatalytic Regioselective C-Formylation of Resorcinol Derivatives
Authors : Gal, L.; Rohan, S.; Tittmann, K.; Kroutil, W.
Deposited on : 2025-09-02
Resolution : 2.60 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 Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

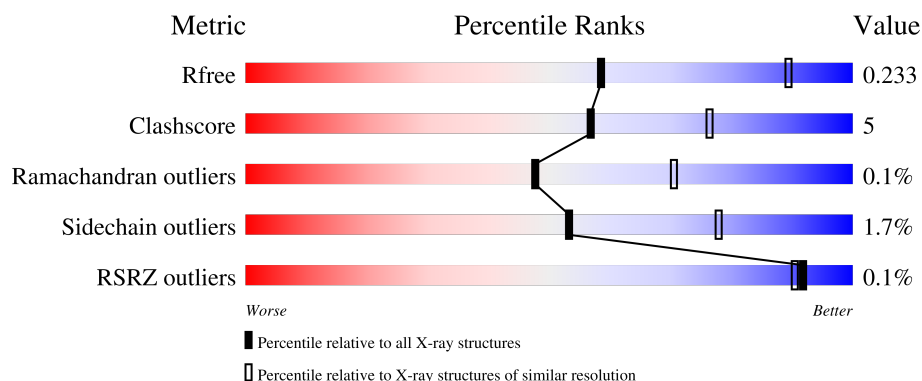
1 Overall quality at a glance

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	4008 (2.60-2.60)
Clashscore	190562	4347 (2.60-2.60)
Ramachandran outliers	187476	4277 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (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 $\leq 5\%$. 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	360	 87% 12% .
1	B	360	 88% 11% .
1	C	360	 88% 10% ..
1	D	360	 91% 8% .
2	E	145	 89% 10% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	145	 93% 6% .
2	G	145	 86% 13% .
2	H	145	 81% 15% . .
3	I	398	 90% 9%
3	J	398	 88% 11% ..
3	K	398	 88% 11% .
3	L	398	 88% 11% .

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 54119 atoms, of which 26618 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 2,4-diacetylphloroglucinol biosynthesis protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	359	Total	C	H	N	O	S	0	0	0
			5318	1669	2649	458	531	11			
1	B	359	Total	C	H	N	O	S	0	0	0
			5318	1669	2649	458	531	11			
1	C	358	Total	C	H	N	O	S	0	0	0
			5296	1663	2636	456	530	11			
1	D	359	Total	C	H	N	O	S	0	0	0
			5318	1669	2649	458	531	11			

- Molecule 2 is a protein called 2,4-diacetylphloroglucinol biosynthesis protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	E	144	Total	C	H	N	O	S	0	0	0
			2299	722	1154	213	199	11			
2	F	144	Total	C	H	N	O	S	0	0	0
			2298	722	1153	213	199	11			
2	G	144	Total	C	H	N	O	S	0	0	0
			2297	722	1152	213	199	11			
2	H	144	Total	C	H	N	O	S	0	0	0
			2298	722	1153	213	199	11			

- Molecule 3 is a protein called 2,4-diacetylphloroglucinol biosynthesis protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	I	397	Total	C	H	N	O	S	0	0	0
			5851	1875	2861	518	577	20			
3	J	396	Total	C	H	N	O	S	0	0	0
			5840	1872	2856	517	575	20			
3	K	396	Total	C	H	N	O	S	0	0	0
			5840	1872	2856	517	575	20			
3	L	395	Total	C	H	N	O	S	0	0	0
			5824	1867	2850	516	571	20			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	168	VAL	ALA	conflict	UNP A0A1S1X4K0
J	168	VAL	ALA	conflict	UNP A0A1S1X4K0
K	168	VAL	ALA	conflict	UNP A0A1S1X4K0
L	168	VAL	ALA	conflict	UNP A0A1S1X4K0

- Molecule 4 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	1	Total Zn 1 1	0	0
4	F	1	Total Zn 1 1	0	0
4	G	1	Total Zn 1 1	0	0
4	H	1	Total Zn 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	32	Total O 32 32	0	0
5	B	28	Total O 28 28	0	0
5	C	47	Total O 47 47	0	0
5	D	33	Total O 33 33	0	0
5	E	20	Total O 20 20	0	0
5	F	13	Total O 13 13	0	0
5	G	15	Total O 15 15	0	0
5	H	17	Total O 17 17	0	0
5	I	16	Total O 16 16	0	0
5	J	36	Total O 36 36	0	0

Continued on next page...

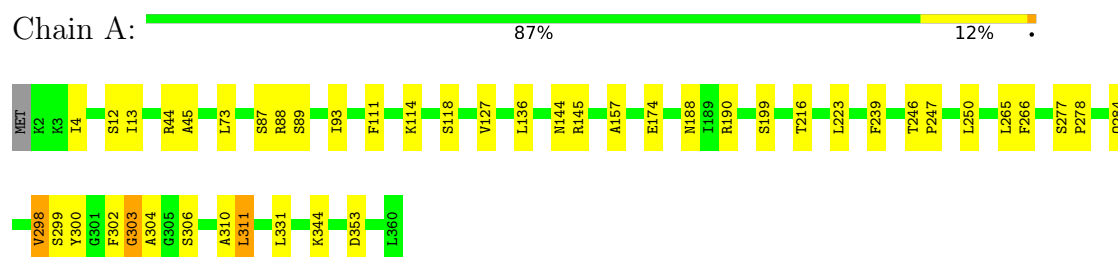
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	K	32	Total 32	O 32	0	0
5	L	29	Total 29	O 29	0	0

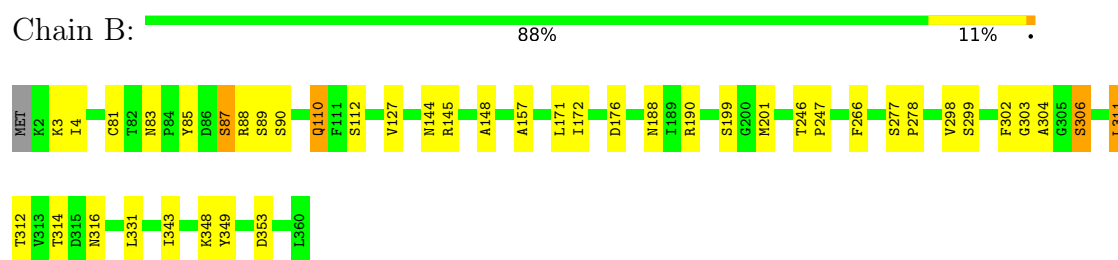
3 Residue-property plots [i](#)

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.

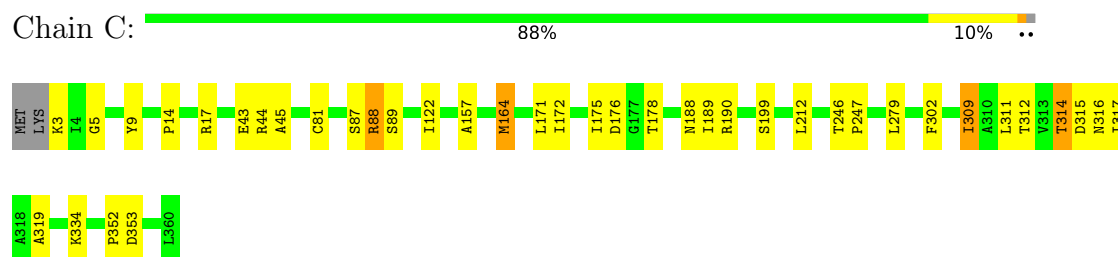
- Molecule 1: 2,4-diacetylphloroglucinol biosynthesis protein



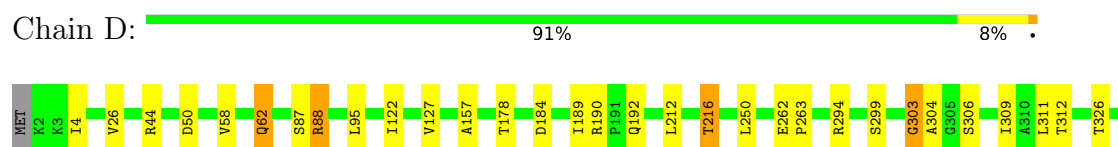
- Molecule 1: 2,4-diacetylphloroglucinol biosynthesis protein



- Molecule 1: 2,4-diacetylphloroglucinol biosynthesis protein



- Molecule 1: 2,4-diacetylphloroglucinol biosynthesis protein





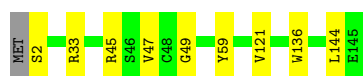
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein

Chain E: 89% 10%



- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein

Chain F: 93% 6%



- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein

Chain G: 86% 13%



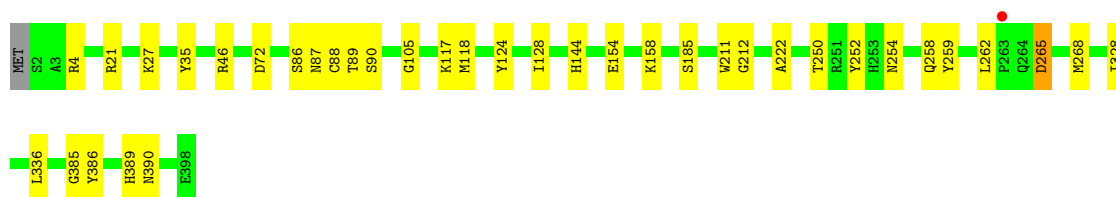
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein

Chain H: 81% 15%



- Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein

Chain I: 90% 9%



- Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein

Chain J: 88% 11%





- Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein

Chain K: 88% 11%



- Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein

Chain L: 88% 11%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	94.47Å 185.45Å 120.68Å 90.00° 103.21° 90.00°	Depositor
Resolution (Å)	91.97 – 2.60 91.97 – 2.60	Depositor EDS
% Data completeness (in resolution range)	55.9 (91.97-2.60) 55.9 (91.97-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 2.62Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.190 , 0.240 0.188 , 0.233	Depositor DCC
R_{free} test set	3438 reflections (2.78%)	wwPDB-VP
Wilson B-factor (Å ²)	40.1	Xtriage
Anisotropy	0.088	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 29.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	54119	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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: 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	0.13	0/2709	0.35	2/3673 (0.1%)
1	B	0.13	0/2709	0.35	0/3673
1	C	0.13	0/2700	0.37	0/3662
1	D	0.12	0/2709	0.36	2/3673 (0.1%)
2	E	0.17	0/1169	0.34	0/1570
2	F	0.13	0/1169	0.34	0/1570
2	G	0.15	0/1169	0.35	0/1570
2	H	0.20	0/1169	0.43	0/1570
3	I	0.13	0/3057	0.36	0/4136
3	J	0.14	0/3051	0.36	0/4128
3	K	0.15	0/3051	0.35	0/4128
3	L	0.13	0/3041	0.36	0/4116
All	All	0.14	0/27703	0.36	4/37469 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	303	GLY	CA-C-N	6.00	132.49	121.70
1	A	303	GLY	C-N-CA	6.00	132.49	121.70
1	D	303	GLY	CA-C-N	5.47	131.55	121.70
1	D	303	GLY	C-N-CA	5.47	131.55	121.70

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	2669	2649	2654	32	0
1	B	2669	2649	2654	32	0
1	C	2660	2636	2641	30	0
1	D	2669	2649	2654	25	0
2	E	1145	1154	1156	10	0
2	F	1145	1153	1156	6	0
2	G	1145	1152	1156	15	0
2	H	1145	1153	1157	22	0
3	I	2990	2861	2876	29	0
3	J	2984	2856	2871	32	0
3	K	2984	2856	2871	32	0
3	L	2974	2850	2865	28	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
5	A	32	0	0	0	0
5	B	28	0	0	0	0
5	C	47	0	0	0	0
5	D	33	0	0	0	0
5	E	20	0	0	0	0
5	F	13	0	0	0	0
5	G	15	0	0	0	0
5	H	17	0	0	1	0
5	I	16	0	0	0	0
5	J	36	0	0	1	0
5	K	32	0	0	0	0
5	L	29	0	0	0	0
All	All	27501	26618	26711	248	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 248 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)
3:I:27:LYS:NZ	3:J:133:GLU:OE2	1.87	1.07
1:B:171:LEU:O	1:B:314:THR:HG21	1.66	0.95
1:B:353:ASP:OD1	2:E:2:SER:N	2.11	0.83
1:A:87:SER:OG	1:B:81:CYS:O	1.96	0.83
3:J:259:TYR:HB3	3:J:262:LEU:HD11	1.62	0.81

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	357/360 (99%)	342 (96%)	15 (4%)	0	100	100
1	B	357/360 (99%)	343 (96%)	14 (4%)	0	100	100
1	C	356/360 (99%)	343 (96%)	12 (3%)	1 (0%)	36	58
1	D	357/360 (99%)	341 (96%)	15 (4%)	1 (0%)	36	58
2	E	142/145 (98%)	136 (96%)	6 (4%)	0	100	100
2	F	142/145 (98%)	135 (95%)	7 (5%)	0	100	100
2	G	142/145 (98%)	135 (95%)	7 (5%)	0	100	100
2	H	142/145 (98%)	135 (95%)	7 (5%)	0	100	100
3	I	395/398 (99%)	373 (94%)	22 (6%)	0	100	100
3	J	394/398 (99%)	378 (96%)	16 (4%)	0	100	100
3	K	394/398 (99%)	376 (95%)	18 (5%)	0	100	100
3	L	393/398 (99%)	372 (95%)	21 (5%)	0	100	100
All	All	3571/3612 (99%)	3409 (96%)	160 (4%)	2 (0%)	48	70

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	88	ARG
1	D	88	ARG

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	278/279 (100%)	275 (99%)	3 (1%)	65	84
1	B	278/279 (100%)	272 (98%)	6 (2%)	45	72
1	C	277/279 (99%)	271 (98%)	6 (2%)	45	72
1	D	278/279 (100%)	275 (99%)	3 (1%)	65	84
2	E	120/121 (99%)	117 (98%)	3 (2%)	42	69
2	F	120/121 (99%)	119 (99%)	1 (1%)	73	88
2	G	120/121 (99%)	120 (100%)	0	100	100
2	H	120/121 (99%)	114 (95%)	6 (5%)	22	46
3	I	305/306 (100%)	300 (98%)	5 (2%)	55	79
3	J	304/306 (99%)	299 (98%)	5 (2%)	55	79
3	K	304/306 (99%)	299 (98%)	5 (2%)	55	79
3	L	303/306 (99%)	299 (99%)	4 (1%)	61	82
All	All	2807/2824 (99%)	2760 (98%)	47 (2%)	53	78

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	88	CYS
3	J	71	SER
3	I	90	SER
3	I	390	ASN
3	J	324	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
3	J	56	HIS
3	J	381	GLN
3	K	296	GLN
3	J	390	ASN
3	J	191	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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	359/360 (99%)	-0.30	0 100 100	16, 40, 56, 72	0
1	B	359/360 (99%)	-0.36	0 100 100	24, 37, 51, 66	0
1	C	358/360 (99%)	-0.40	0 100 100	22, 34, 47, 63	0
1	D	359/360 (99%)	-0.43	0 100 100	23, 34, 48, 76	0
2	E	144/145 (99%)	-0.56	0 100 100	23, 35, 48, 60	0
2	F	144/145 (99%)	-0.47	0 100 100	25, 37, 52, 66	0
2	G	144/145 (99%)	-0.44	0 100 100	27, 38, 53, 69	0
2	H	144/145 (99%)	-0.44	0 100 100	28, 37, 52, 62	0
3	I	397/398 (99%)	-0.25	1 (0%) 90 88	29, 43, 63, 93	0
3	J	396/398 (99%)	-0.35	0 100 100	25, 40, 56, 93	0
3	K	396/398 (99%)	-0.29	1 (0%) 90 88	30, 41, 62, 101	0
3	L	395/398 (99%)	-0.27	1 (0%) 90 88	31, 43, 61, 93	0
All	All	3595/3612 (99%)	-0.35	3 (0%) 92 90	16, 39, 56, 101	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	263	PRO	2.5
3	K	295	ALA	2.2
3	I	263	PRO	2.1

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 [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

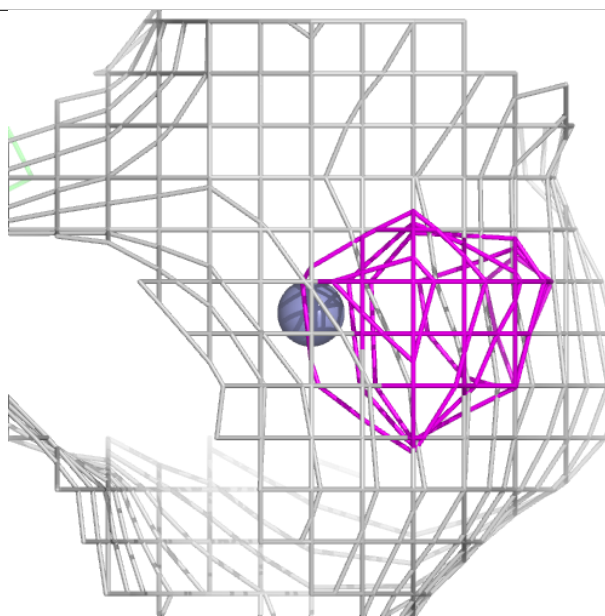
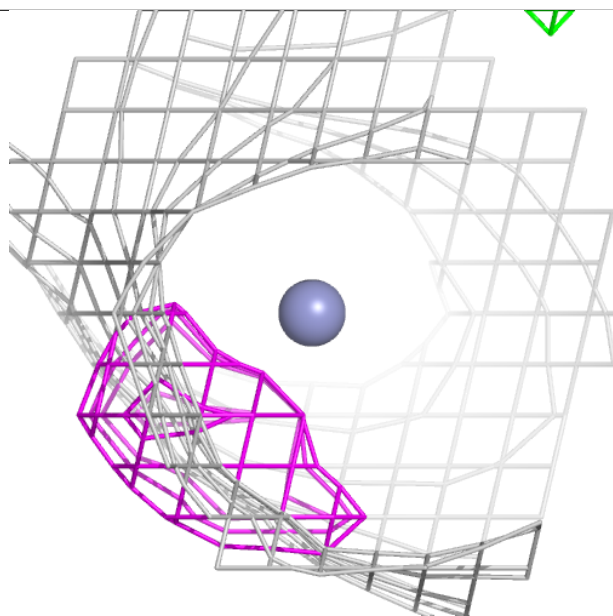
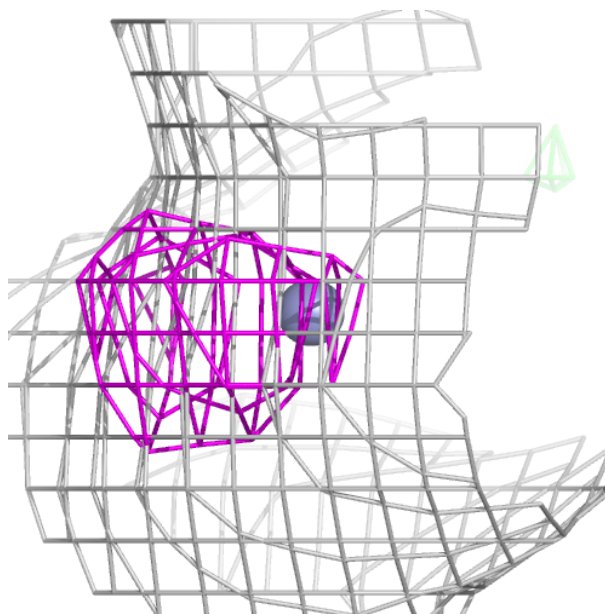
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	ZN	H	201	1/1	0.94	0.06	58,58,58,58	0
4	ZN	F	201	1/1	0.99	0.03	59,59,59,59	0
4	ZN	G	201	1/1	0.99	0.04	48,48,48,48	0
4	ZN	E	201	1/1	0.99	0.02	50,50,50,50	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.

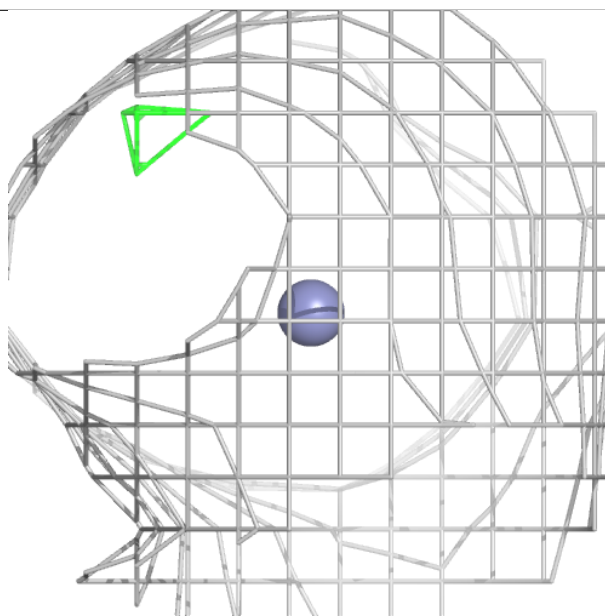
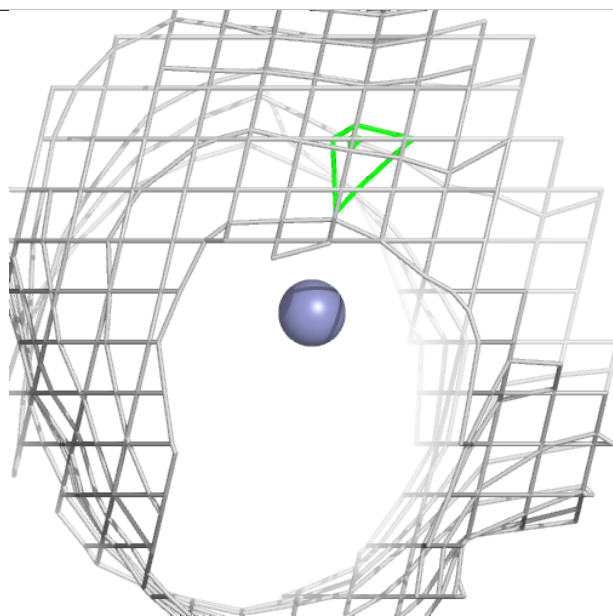
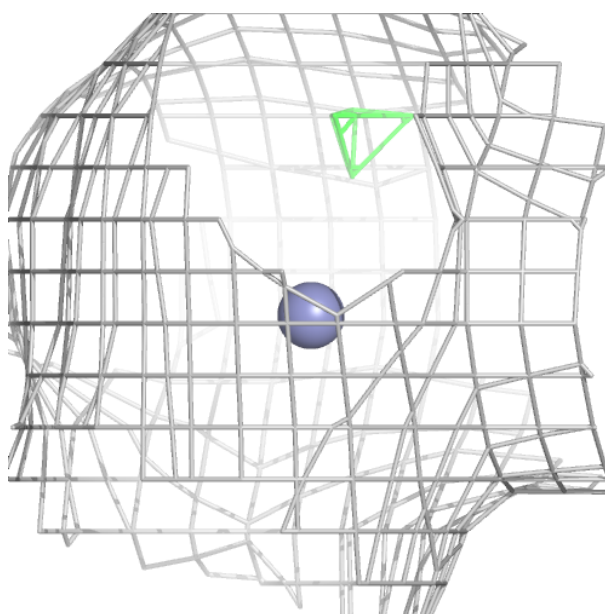
Electron density around ZN H 201:

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



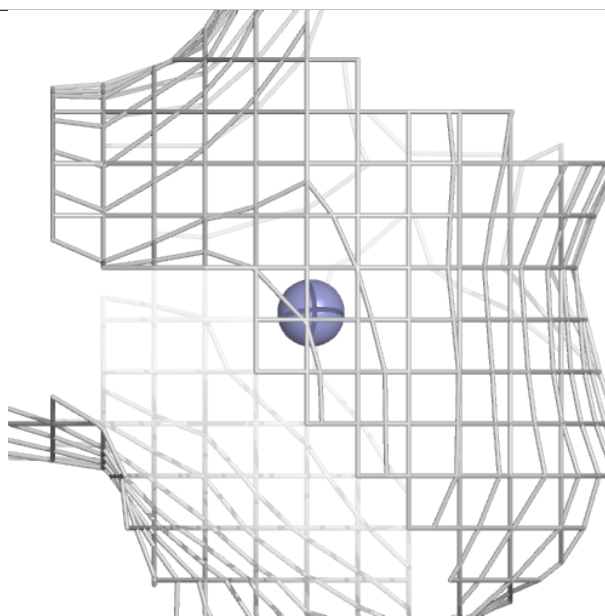
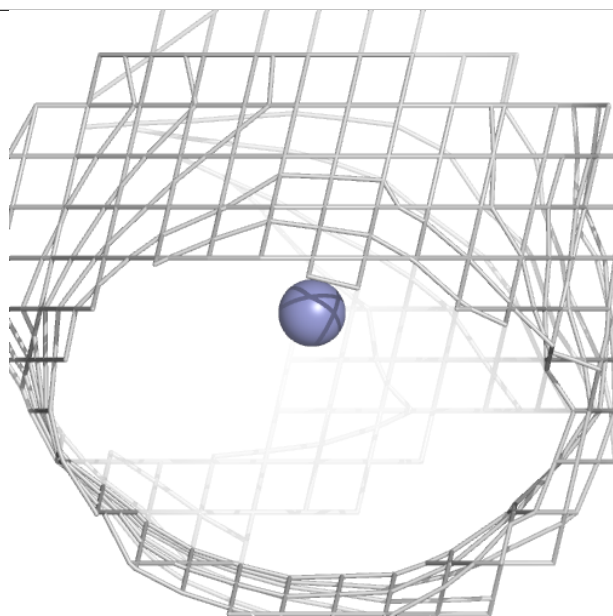
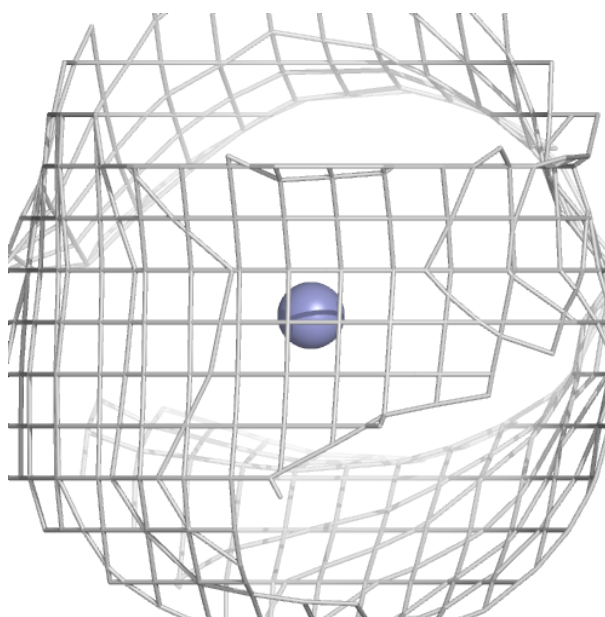
Electron density around ZN F 201:

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



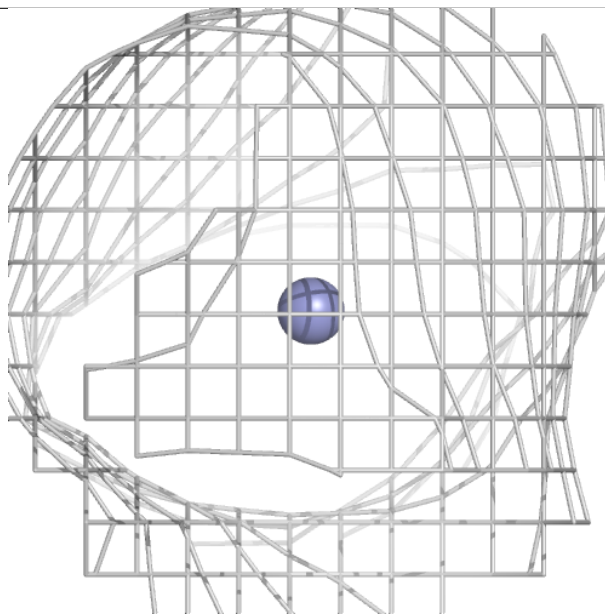
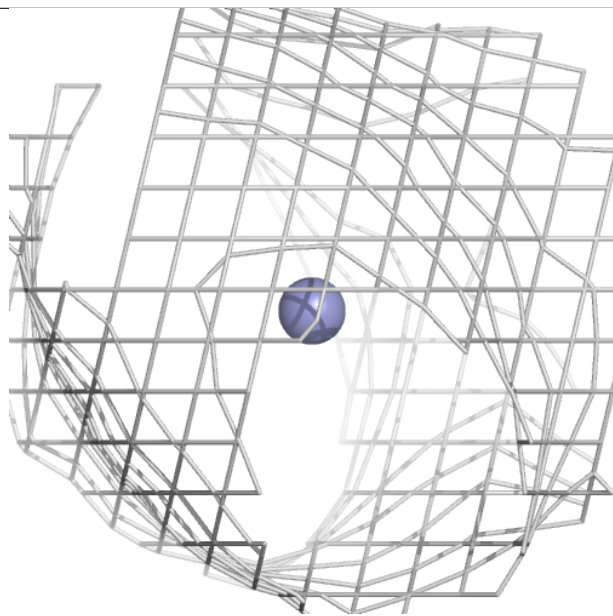
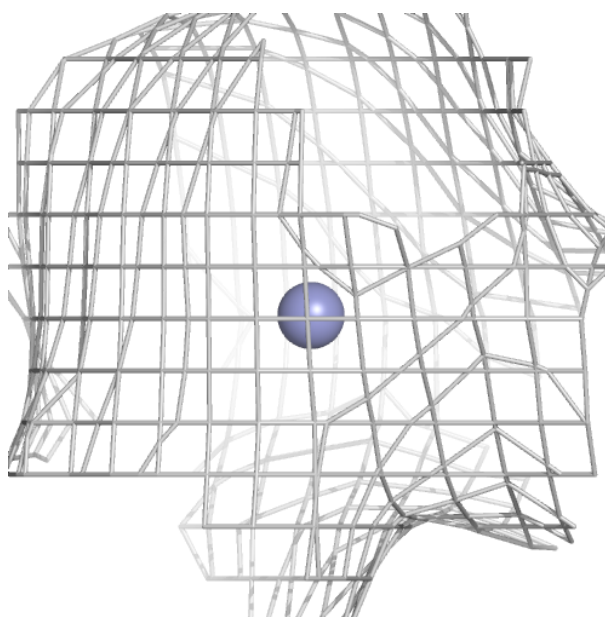
Electron density around ZN G 201:

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



Electron density around ZN E 201:

$2mF_o-DF_c$ (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.