



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 5, 2024 – 10:08 AM EDT

PDB ID : 8TY0
Title : Streptavidin variant S112E-K121H bound to bis-biotinylated Iron-porphyrin
Authors : Follmer, A.H.; Jakob, R.P.; Mukherjee, M.; Ward, T.R.
Deposited on : 2023-08-24
Resolution : 1.54 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 : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtrriage (Phenix) : 1.20.1
EDS : 3.0
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.002 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.38.3

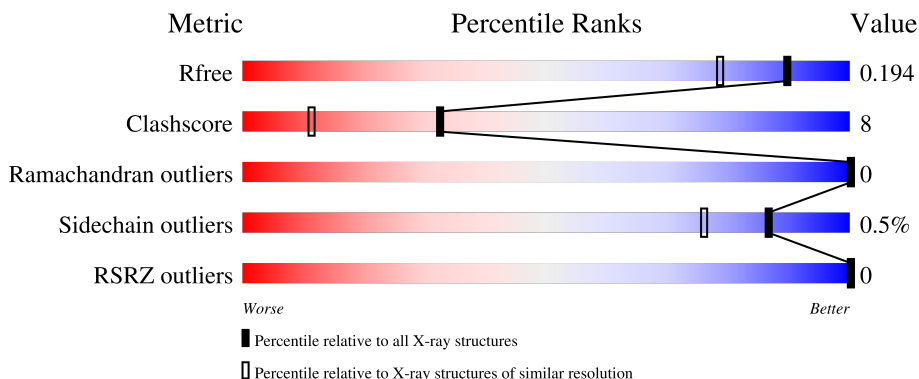
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.54 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	3511 (1.56-1.52)
Clashscore	180529	3784 (1.56-1.52)
Ramachandran outliers	177936	3720 (1.56-1.52)
Sidechain outliers	177891	3717 (1.56-1.52)
RSRZ outliers	164620	3510 (1.56-1.52)

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	158	
1	B	158	
1	C	158	
1	D	158	

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SIK	C	201[A]	X	-	-	-
2	SIK	C	201[B]	X	-	-	-
2	SIK	C	201[C]	X	-	-	-
2	SIK	D	201[A]	X	-	-	-
2	SIK	D	201[B]	X	-	-	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5142 atoms, of which 270 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 Streptavidin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	123	961	598	165	198	0	6	0
1	B	124	977	607	170	200	0	6	0
1	C	124	980	607	173	200	0	7	0
1	D	125	991	616	172	203	0	7	0

There are 56 discrepancies between the modelled and reference sequences:

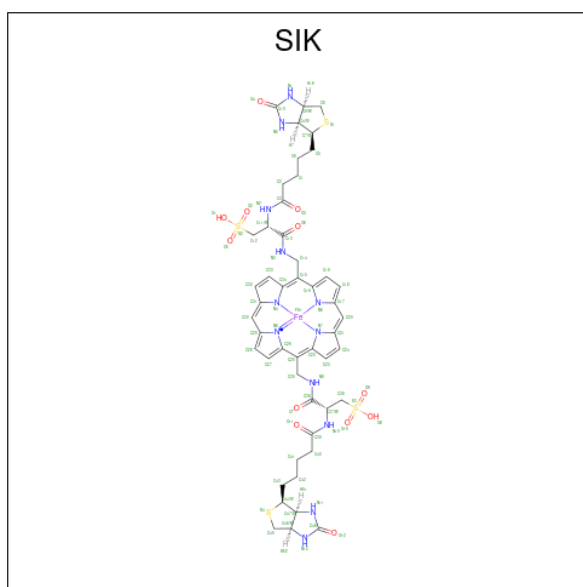
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	-	expression tag	UNP P22629
A	3	SER	-	expression tag	UNP P22629
A	4	MET	-	expression tag	UNP P22629
A	5	THR	-	expression tag	UNP P22629
A	6	GLY	-	expression tag	UNP P22629
A	7	GLY	-	expression tag	UNP P22629
A	8	GLN	-	expression tag	UNP P22629
A	9	GLN	-	expression tag	UNP P22629
A	10	MET	-	expression tag	UNP P22629
A	11	GLY	-	expression tag	UNP P22629
A	12	ARG	-	expression tag	UNP P22629
A	13	ASP	-	expression tag	UNP P22629
A	112	GLU	SER	conflict	UNP P22629
A	121	HIS	LYS	conflict	UNP P22629
B	2	ALA	-	expression tag	UNP P22629
B	3	SER	-	expression tag	UNP P22629
B	4	MET	-	expression tag	UNP P22629
B	5	THR	-	expression tag	UNP P22629
B	6	GLY	-	expression tag	UNP P22629
B	7	GLY	-	expression tag	UNP P22629
B	8	GLN	-	expression tag	UNP P22629

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	9	GLN	-	expression tag	UNP P22629
B	10	MET	-	expression tag	UNP P22629
B	11	GLY	-	expression tag	UNP P22629
B	12	ARG	-	expression tag	UNP P22629
B	13	ASP	-	expression tag	UNP P22629
B	112	GLU	SER	conflict	UNP P22629
B	121	HIS	LYS	conflict	UNP P22629
C	2	ALA	-	expression tag	UNP P22629
C	3	SER	-	expression tag	UNP P22629
C	4	MET	-	expression tag	UNP P22629
C	5	THR	-	expression tag	UNP P22629
C	6	GLY	-	expression tag	UNP P22629
C	7	GLY	-	expression tag	UNP P22629
C	8	GLN	-	expression tag	UNP P22629
C	9	GLN	-	expression tag	UNP P22629
C	10	MET	-	expression tag	UNP P22629
C	11	GLY	-	expression tag	UNP P22629
C	12	ARG	-	expression tag	UNP P22629
C	13	ASP	-	expression tag	UNP P22629
C	112	GLU	SER	conflict	UNP P22629
C	121	HIS	LYS	conflict	UNP P22629
D	2	ALA	-	expression tag	UNP P22629
D	3	SER	-	expression tag	UNP P22629
D	4	MET	-	expression tag	UNP P22629
D	5	THR	-	expression tag	UNP P22629
D	6	GLY	-	expression tag	UNP P22629
D	7	GLY	-	expression tag	UNP P22629
D	8	GLN	-	expression tag	UNP P22629
D	9	GLN	-	expression tag	UNP P22629
D	10	MET	-	expression tag	UNP P22629
D	11	GLY	-	expression tag	UNP P22629
D	12	ARG	-	expression tag	UNP P22629
D	13	ASP	-	expression tag	UNP P22629
D	112	GLU	SER	conflict	UNP P22629
D	121	HIS	LYS	conflict	UNP P22629

- Molecule 2 is Bis-biotinylated Iron-porphyrin (three-letter code: SIK) (formula: C₄₈H₅₆FeN₁₂O₁₂S₄) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
			Total	C	Fe	H	N	O	S		
2	C	1	393	144	3	162	36	36	12	0	1
2	D	1	262	96	2	108	24	24	8	0	1

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	141	Total	O	0	0
			141	141		
3	B	142	Total	O	0	0
			142	142		
3	C	148	Total	O	0	0
			148	148		
3	D	147	Total	O	0	0
			147	147		

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	192.49Å 57.72Å 57.77Å 90.00° 107.44° 90.00°	Depositor
Resolution (Å)	55.11 – 1.54 55.11 – 1.54	Depositor EDS
% Data completeness (in resolution range)	94.3 (55.11-1.54) 94.1 (55.11-1.54)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.51 (at 1.54Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.173 , 0.191 0.176 , 0.194	Depositor DCC
R_{free} test set	3823 reflections (4.54%)	wwPDB-VP
Wilson B-factor (Å ²)	13.2	Xtrriage
Anisotropy	0.043	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 48.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.480 for -h+k-l,-l,-k 0.479 for -h-k-l,l,k 0.480 for -h-2*l,-k,l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5142	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.33% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SIK

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	0.76	0/993	0.92	0/1361
1	B	0.80	0/1006	0.94	0/1378
1	C	0.76	0/1016	0.96	0/1390
1	D	0.79	0/1020	0.88	0/1397
All	All	0.78	0/4035	0.92	0/5526

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	961	0	887	18	0
1	B	977	0	903	18	0
1	C	980	0	905	15	0
1	D	991	0	918	14	0
2	C	231	162	0	6	0
2	D	154	108	0	6	0
3	A	141	0	0	3	0
3	B	142	0	0	2	0
3	C	148	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	147	0	0	2	0
All	All	4872	270	3613	62	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 8.

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:40[B]:THR:HG23	3:C:313:HOH:O	1.62	0.98
1:D:40[A]:THR:HG22	3:D:303:HOH:O	1.65	0.95
1:B:134:LYS:HG2	1:B:135:PRO:HD2	1.55	0.88
1:D:103:ARG:HH11	1:D:129[A]:THR:HG21	1.41	0.85
1:B:103:ARG:HH11	1:B:129[A]:THR:HG21	1.45	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	127/158 (80%)	124 (98%)	3 (2%)	0	100	100
1	B	128/158 (81%)	125 (98%)	3 (2%)	0	100	100
1	C	129/158 (82%)	126 (98%)	3 (2%)	0	100	100
1	D	130/158 (82%)	127 (98%)	3 (2%)	0	100	100
All	All	514/632 (81%)	502 (98%)	12 (2%)	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	97/114 (85%)	97 (100%)	0	100	100
1	B	98/114 (86%)	98 (100%)	0	100	100
1	C	99/114 (87%)	98 (99%)	1 (1%)	73	52
1	D	100/114 (88%)	99 (99%)	1 (1%)	73	52
All	All	394/456 (86%)	392 (100%)	2 (0%)	86	76

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	134	LYS
1	D	101	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	121	HIS

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](#)

5 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	SIK	C	201[C]	-	79,88,88	2.02	18 (22%)	94,136,136	2.16	30 (31%)
2	SIK	C	201[A]	-	79,88,88	2.37	23 (29%)	94,136,136	2.91	35 (37%)
2	SIK	D	201[B]	-	79,88,88	2.09	19 (24%)	94,136,136	2.42	32 (34%)
2	SIK	C	201[B]	-	79,88,88	2.05	20 (25%)	94,136,136	2.36	33 (35%)
2	SIK	D	201[A]	-	79,88,88	2.00	15 (18%)	94,136,136	1.96	26 (27%)

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	SIK	C	201[C]	-	2/2/29/29	18/46/140/140	0/4/12/12
2	SIK	C	201[A]	-	2/2/29/29	18/46/140/140	0/4/12/12
2	SIK	D	201[B]	-	2/2/29/29	13/46/140/140	0/4/12/12
2	SIK	C	201[B]	-	2/2/29/29	18/46/140/140	0/4/12/12
2	SIK	D	201[A]	-	2/2/29/29	14/46/140/140	0/4/12/12

The worst 5 of 95 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	201[A]	SIK	C38-C37	8.14	1.59	1.53
2	C	201[A]	SIK	C34-C15	7.90	1.50	1.39
2	D	201[A]	SIK	C34-C15	7.85	1.50	1.39
2	C	201[C]	SIK	C34-C15	7.79	1.50	1.39
2	D	201[B]	SIK	C22-C25	7.01	1.49	1.39

The worst 5 of 156 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	201[A]	SIK	C11-C12-S2	12.84	132.76	113.55
2	C	201[A]	SIK	C37-C38-S3	11.71	131.07	113.55
2	D	201[B]	SIK	C32-C31-N4	8.99	117.00	110.17
2	D	201[A]	SIK	C24-C21-N7	7.26	115.68	110.17
2	C	201[B]	SIK	C11-C12-S2	7.23	124.37	113.55

5 of 10 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	201[A]	SIK	N7
2	C	201[A]	SIK	N4
2	C	201[B]	SIK	N7
2	C	201[B]	SIK	N4
2	C	201[C]	SIK	N7

5 of 81 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	201[A]	SIK	C13-C11-C12-S2
2	C	201[A]	SIK	N2-C11-C12-S2
2	C	201[A]	SIK	C12-C11-C13-N3
2	C	201[A]	SIK	C12-C11-C13-O6
2	C	201[A]	SIK	C13-C11-N2-C3

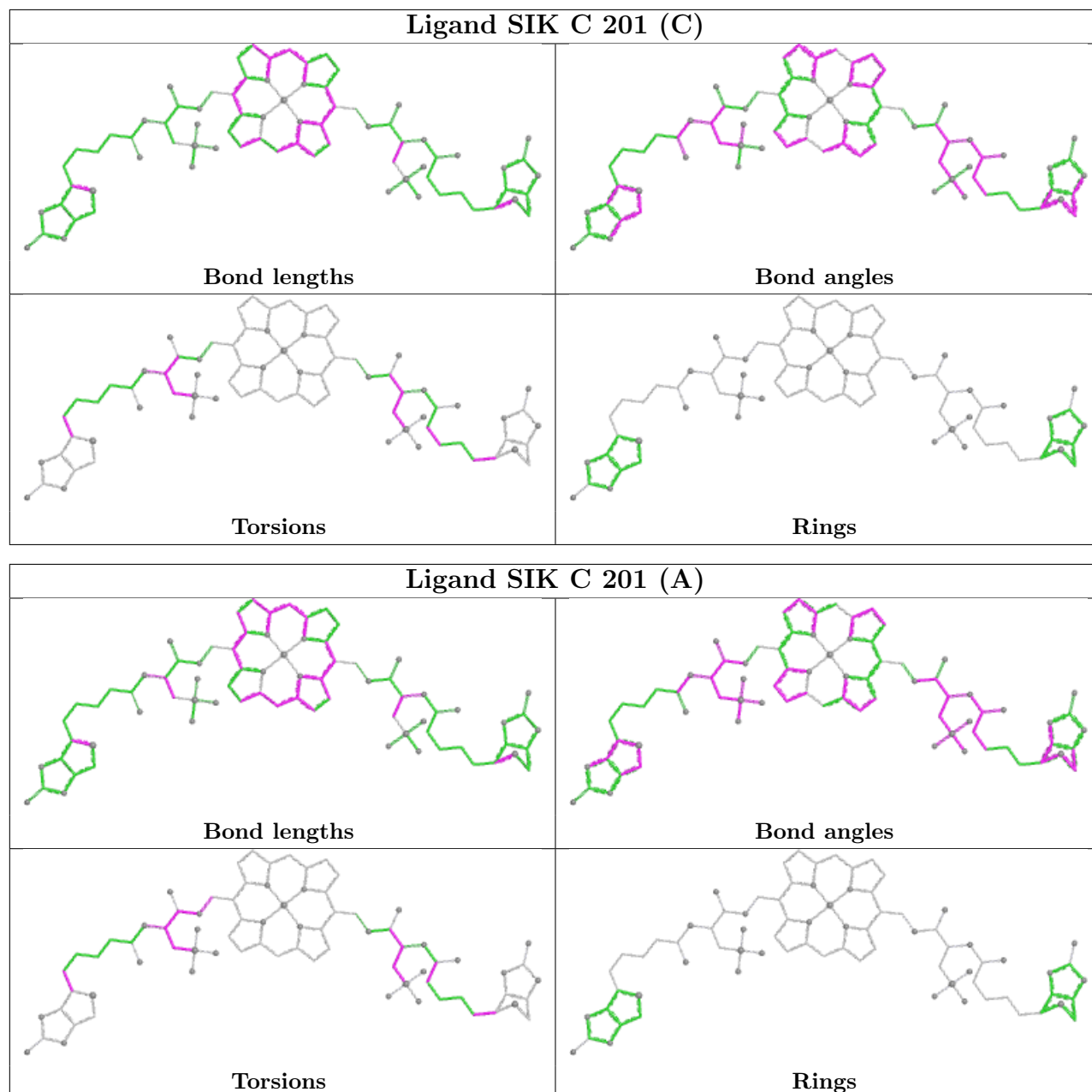
There are no ring outliers.

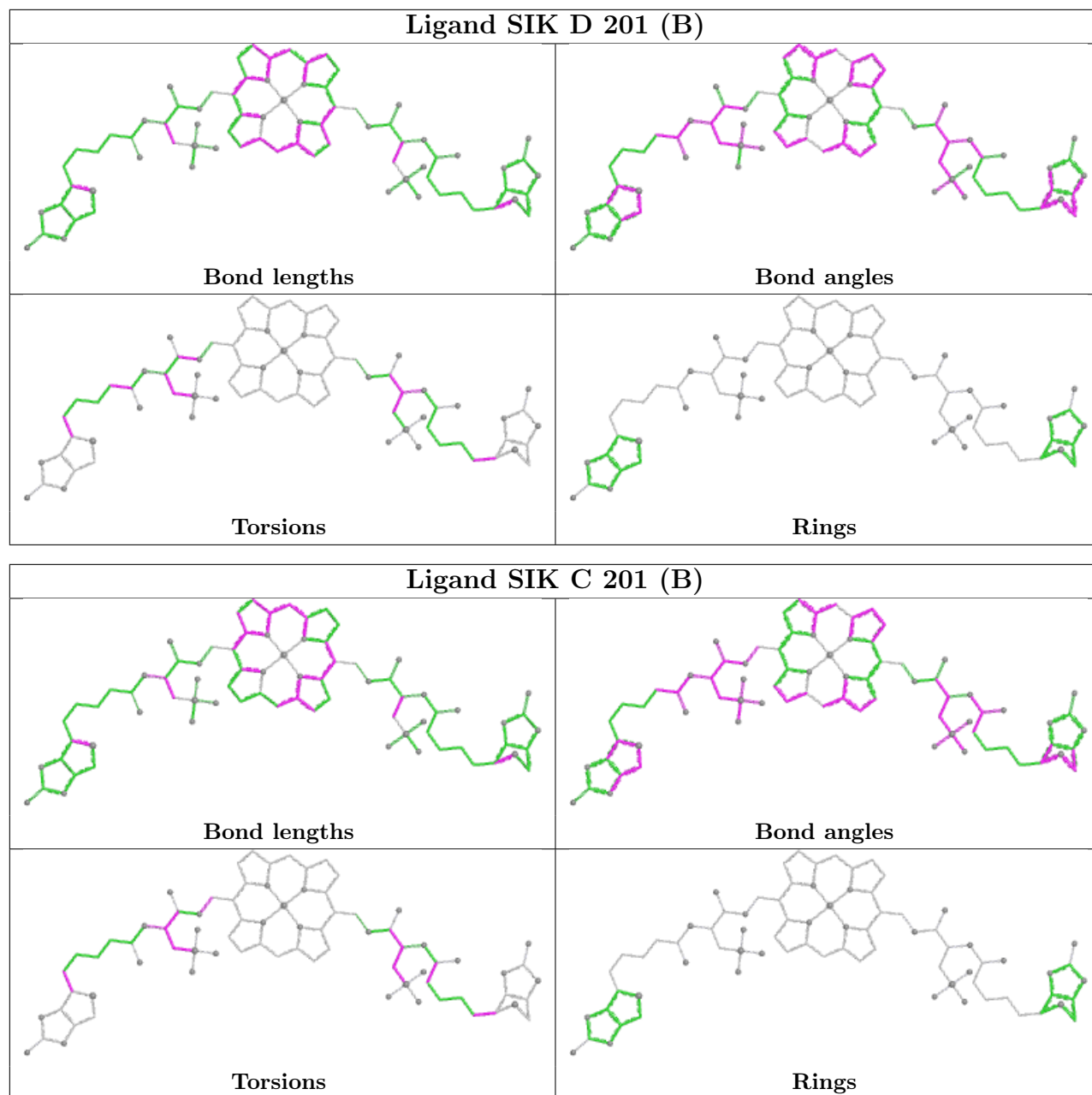
5 monomers are involved in 12 short contacts:

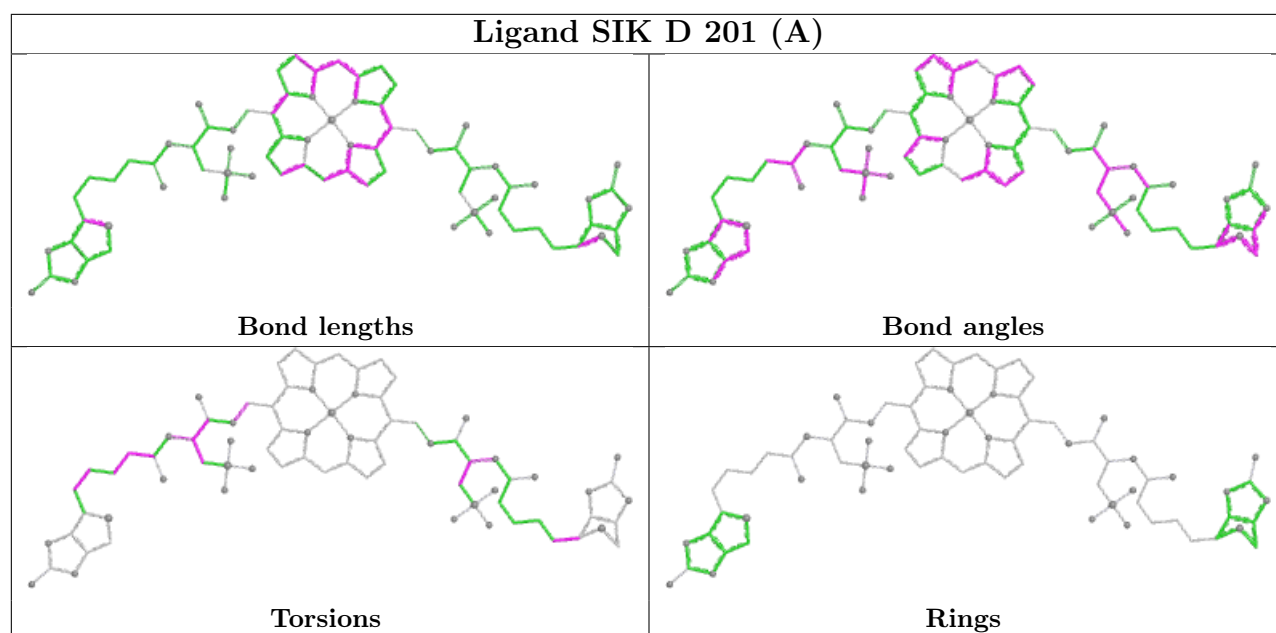
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	201[C]	SIK	2	0
2	C	201[A]	SIK	2	0
2	D	201[B]	SIK	4	0
2	C	201[B]	SIK	2	0
2	D	201[A]	SIK	2	0

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 than 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.







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 [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	123/158 (77%)	-0.93	0 100 100	6, 14, 34, 55	6 (4%)
1	B	124/158 (78%)	-0.91	0 100 100	4, 14, 34, 67	6 (4%)
1	C	124/158 (78%)	-0.92	0 100 100	5, 14, 35, 51	7 (5%)
1	D	125/158 (79%)	-0.94	0 100 100	4, 14, 36, 61	7 (5%)
All	All	496/632 (78%)	-0.93	0 100 100	4, 14, 37, 67	26 (5%)

There are no RSRZ outliers to report.

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 monosaccharides 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.

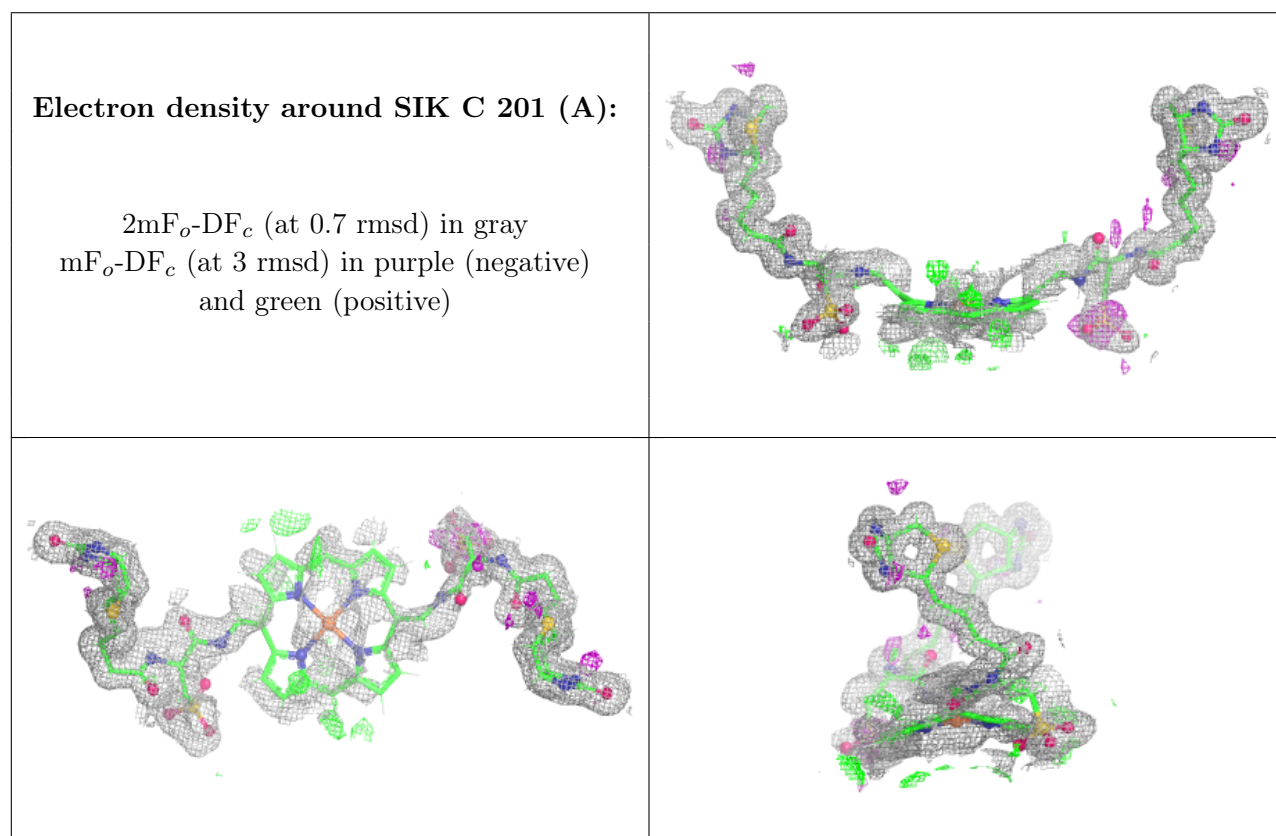
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	SIK	C	201[A]	77/77	0.98	0.08	8,31,48,54	131
2	SIK	C	201[B]	77/77	0.98	0.08	8,32,45,49	131
2	SIK	C	201[C]	77/77	0.98	0.08	8,31,45,51	131
2	SIK	D	201[A]	77/77	0.98	0.09	7,33,55,65	131

Continued on next page...

Continued from previous page...

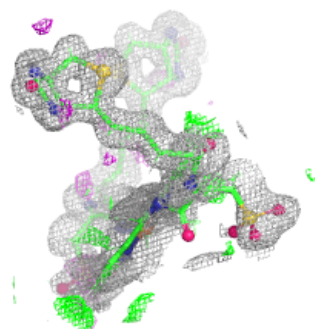
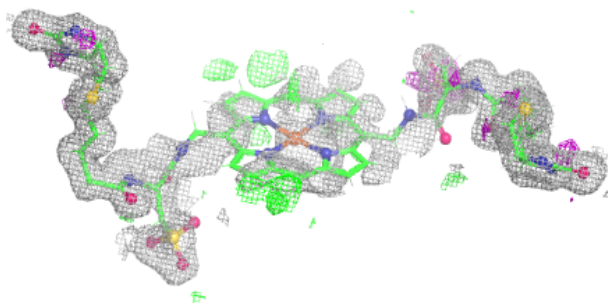
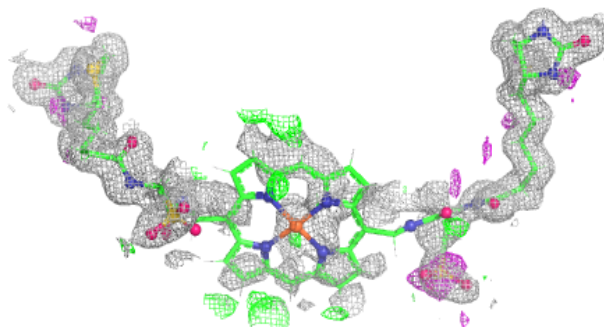
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	SIK	D	201[B]	77/77	0.98	0.09	7,30,55,68	131

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.

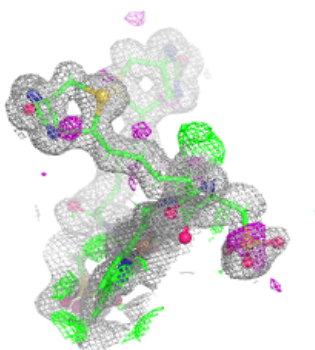
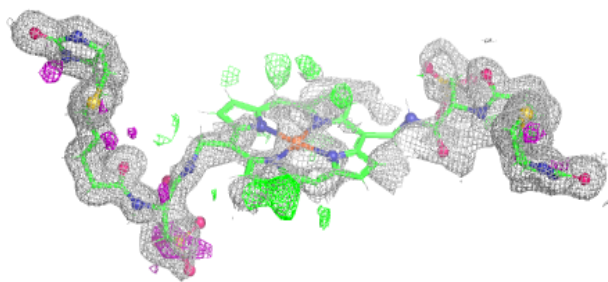
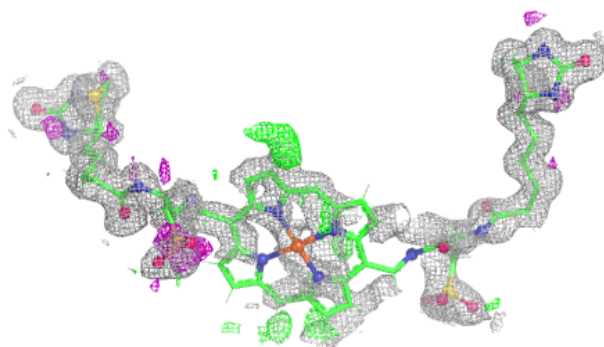


Electron density around SIK C 201 (B):

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

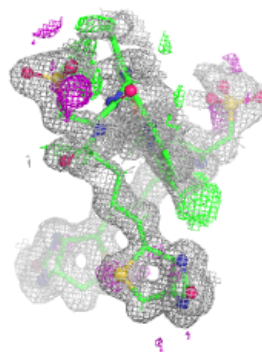
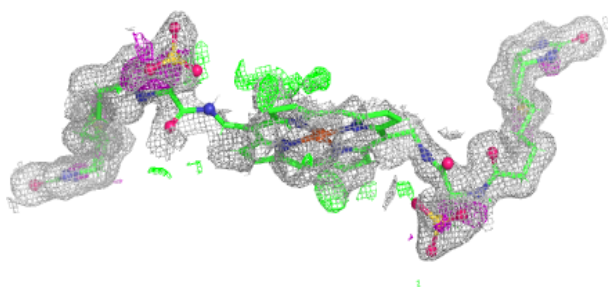
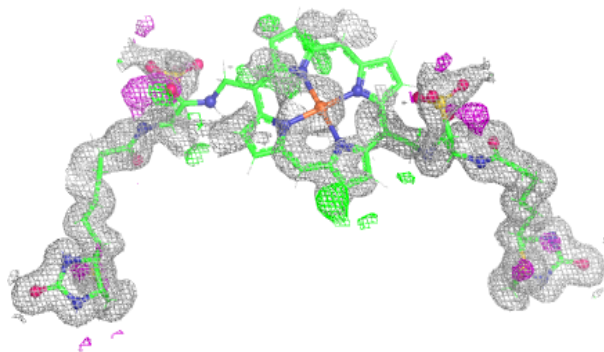
**Electron density around SIK C 201 (C):**

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

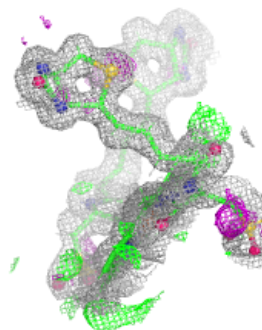
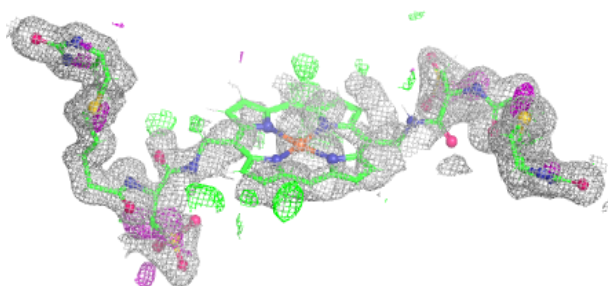
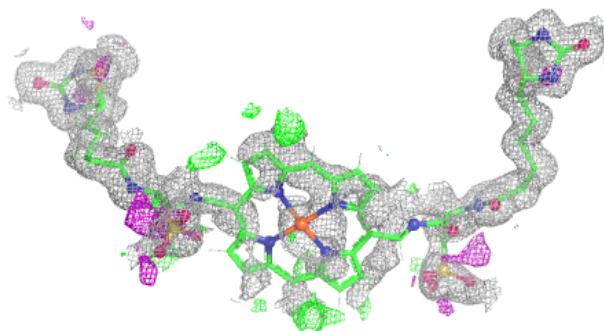


Electron density around SIK D 201 (A):

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

**Electron density around SIK D 201 (B):**

$2mF_o-DF_c$ (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.