

wwPDB X-ray Structure Validation Summary Report (i)

Sep 23, 2023 – 02:50 PM EDT

PDB ID	:	5KSI
Title	:	Crystal structure of deoxygenated hemoglobin in complex with sphingosine
		phosphate and 2,3-Bisphosphoglycerate
Authors	:	Ahmed, M.H.; Safo, M.K.; Xia, Y.
Deposited on	:	2016-07-08
Resolution	:	1.80 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 (1)) were used in the production of this report:

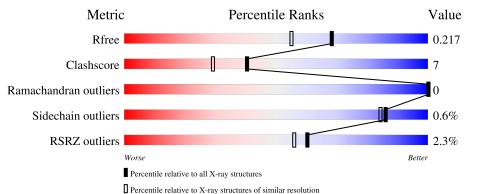
MolProbity	:	4.02b-467 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)		
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 1.80 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	А	141	.% 94%	6%
1	С	141	% • 89%	10% •
2	В	146	3% 84%	16%
2	D	146	5% 86%	14%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard



Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	S1P	А	202	-	-	-	Х
4	S1P	А	203[A]	-	-	-	Х
4	S1P	С	202	-	-	-	Х
4	S1P	С	203	-	-	-	Х
4	S1P	D	202[B]	-	-	-	Х

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:



2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5454 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.

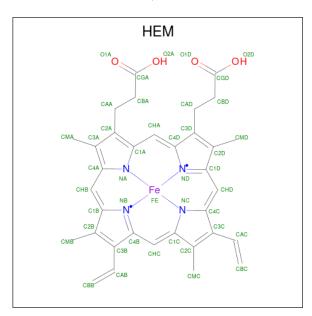
• Molecule 1 is a protein called Hemoglobin subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Λ	141	Total	С	Ν	Ο	S	0	0	0
	Л	141	1069	685	187	194	3	0		
1	С	141	Total	С	Ν	Ο	S	0	0	0
1	U	141	1069	685	187	194	3	0	0	0

• Molecule 2 is a protein called Hemoglobin subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
0	р	146	Total	С	Ν	0	S	0	0	0
	D	140	1123	724	195	201	3	0		
0	Л	146	Total	С	Ν	0	S	0	0	0
	D	140	1123	724	195	201	3	0	0	0

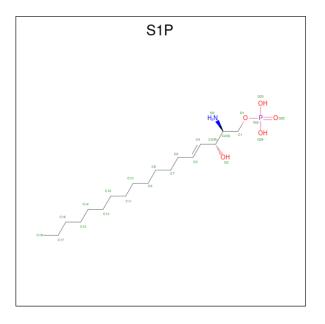
• Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: C₃₄H₃₂FeN₄O₄).





Mol	Chain	Residues		Ate	oms			ZeroOcc	AltConf
3	А	1	Total 43	C 34	Fe 1	N 4	0 4	0	0
3	В	1	Total 43	C 34	Fe 1	N 4	0 4	0	0
3	С	1	Total 43	C 34	Fe 1	N 4	0 4	0	0
3	D	1	Total 43	С 34	Fe 1	N 4	0 4	0	0

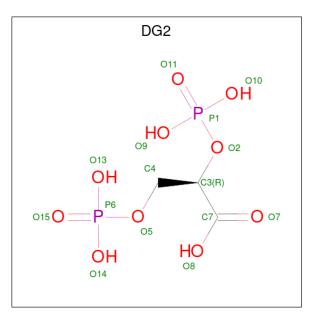
• Molecule 4 is (2S,3R,4E)-2-amino-3-hydroxyoctadec-4-en-1-yl dihydrogen phosphate (three-letter code: S1P) (formula: $C_{18}H_{38}NO_5P$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	Total C N O P 11 4 1 5 1	0	0
4	А	1	Total C N O P 25 18 1 5 1	0	1
4	С	1	$\begin{array}{cccc} \text{Total} & \text{C} & \text{O} & \text{P} \\ 7 & 2 & 4 & 1 \end{array}$	0	0
4	С	1	Total C N O P 16 9 1 5 1	0	0
4	D	1	Total C N O P 25 18 1 5 1	0	1

• Molecule 5 is (2R)-2,3-diphosphoglyceric acid (three-letter code: DG2) (formula: $C_3H_8O_{10}P_2$).





Mol	Chain	Residues	A	tor	ns		ZeroOcc	AltConf	
5	В	1	Total	С	0	Р	0	1	
		-	15	-	-			-	
5	Л	1	Total	\mathbf{C}	Ο	Р	0	1	
0	D	1	15	3	10	2	0	1	

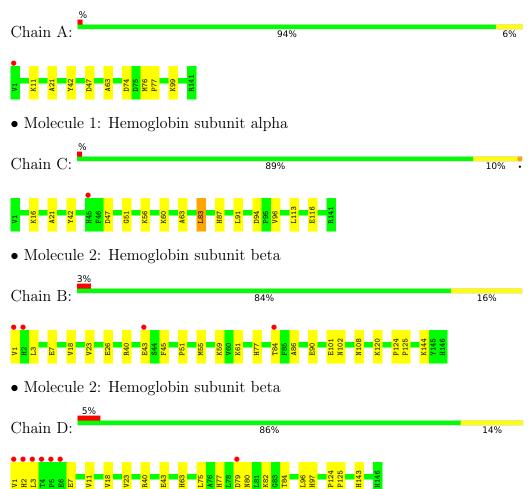
• Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	189	Total O 189 189	0	0
6	В	205	Total O 205 205	0	0
6	С	199	Total O 199 199	0	0
6	D	191	Total O 191 191	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: Hemoglobin subunit alpha



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	Depositor
Resolution (Å)	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	Depositor EDS
% Data completeness (in resolution range)	93.8 (30.00-1.80) 93.8 (29.42-1.80)	Depositor EDS
R _{merge}	0.08	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.99 (at 1.80 \text{\AA})$	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	Depositor DCC
R_{free} test set	2764 reflections $(5.11%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	19.5	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.39 , 74.9	EDS
L-test for twinning ²	$< L > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.020 for k,h,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5454	wwPDB-VP
Average B, all atoms $(Å^2)$	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 6.56% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: DG2, S1P, HEM

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		Bo	nd angles
	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	А	0.43	0/1097	0.58	1/1491~(0.1%)
1	С	0.43	0/1097	0.60	1/1491~(0.1%)
2	В	0.42	0/1153	0.58	0/1566
2	D	0.46	0/1153	0.60	0/1566
All	All	0.44	0/4500	0.59	2/6114~(0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
1	С	47	ASP	N-CA-C	-5.72	95.57	111.00
1	А	47	ASP	N-CA-C	-5.25	96.82	111.00

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	А	1069	0	1073	7	0
1	С	1069	0	1073	12	0
2	В	1123	0	1118	25	0
2	D	1123	0	1118	22	0
3	А	43	0	30	0	0

Continued on next page...



Mol	Chain	Non-H		H(added)	Clashes	Symm-Clashes
3	В	43	0	30	1	0
3	С	43	0	30	0	0
3	D	43	0	30	1	0
4	А	36	0	21	1	0
4	С	23	0	17	7	0
4	D	25	0	16	11	0
5	В	15	0	4	0	0
5	D	15	0	0	0	0
6	А	189	0	0	3	0
6	В	205	0	0	6	0
6	С	199	0	0	3	0
6	D	191	0	0	4	0
All	All	5454	0	4560	66	0

Continued from previous page...

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 7.

The worst 5 of 66 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:40:ARG:NH2	4:D:202[B]:S1P:H4	1.64	1.12
2:D:40:ARG:HH12	4:D:202[B]:S1P:H6	1.13	1.09
2:D:40:ARG:HH22	4:D:202[B]:S1P:C4	1.68	1.05
2:D:40:ARG:HH22	4:D:202[B]:S1P:H4	0.79	0.93
2:D:40:ARG:NH1	4:D:202[B]:S1P:H6	1.87	0.89

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 Al		Allowed	Outliers	Perce	ntiles
1	А	139/141~(99%)	137~(99%)	2(1%)	0	100	100
1	С	139/141 (99%)	137 (99%)	2(1%)	0	100	100
2	В	144/146~(99%)	142 (99%)	2(1%)	0	100	100
2	D	144/146~(99%)	142 (99%)	2(1%)	0	100	100
All	All	566/574~(99%)	558~(99%)	8 (1%)	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	А	113/113~(100%)	113 (100%)	0	100 100
1	С	113/113~(100%)	112~(99%)	1 (1%)	78 75
2	В	118/118 (100%)	118 (100%)	0	100 100
2	D	118/118 (100%)	116~(98%)	2(2%)	60 51
All	All	462/462~(100%)	459~(99%)	3(1%)	86 84

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	С	83	LEU
2	D	2	HIS
2	D	75	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	9	ASN
2	В	80	ASN
2	В	108	ASN
2	D	77	HIS
2	D	80	ASN



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)

11 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	Turne	Chain	Res	Link	В	ond leng	gths	Bond angles		les
MOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2
5	DG2	В	202[A]	-	13,14,14	1.51	2 (15%)	18,21,21	0.88	1 (5%)
4	S1P	А	203[A]	-	23,24,24	0.81	1 (4%)	23,28,28	1.05	0
5	DG2	D	203[B]	-	13,14,14	1.61	3 (23%)	18,21,21	0.86	1 (5%)
4	S1P	С	202	-	6,6,24	1.11	1 (16%)	8,8,28	0.74	0
3	HEM	А	201	1	41,50,50	1.34	6 (14%)	45,82,82	1.48	7 (15%)
3	HEM	С	201	1	41,50,50	1.49	6 (14%)	45,82,82	1.25	3 (6%)
3	HEM	D	201	2	41,50,50	1.61	10 (24%)	45,82,82	1.34	7 (15%)
4	S1P	С	203	-	14,15,24	1.13	3 (21%)	14,19,28	1.24	1 (7%)
4	S1P	А	202	-	9,10,24	1.02	1 (11%)	8,14,28	0.96	0
4	S1P	D	202[B]	-	23,24,24	0.95	2 (8%)	23,28,28	1.18	2 (8%)
3	HEM	В	201	2	41,50,50	1.53	10 (24%)	45,82,82	1.21	4 (8%)

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.



5KSI

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DG2	В	202[A]	-	-	6/15/15/15	-
4	S1P	А	203[A]	-	-	19/25/25/25	-
5	DG2	D	203[B]	-	-	4/15/15/15	-
4	S1P	С	202	-	-	0/4/4/25	-
3	HEM	А	201	1	-	4/12/54/54	-
3	HEM	С	201	1	-	6/12/54/54	-
3	HEM	D	201	2	-	4/12/54/54	-
4	S1P	С	203	-	-	14/16/16/25	-
4	S1P	А	202	-	-	7/10/10/25	_
4	S1P	D	202[B]	-	-	17/25/25/25	-
3	HEM	В	201	2	-	6/12/54/54	-

'-' means no outliers of that kind were identified.

The worst 5 of 45 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	С	201	HEM	C3C-CAC	-4.03	1.39	1.47
3	В	201	HEM	CHA-C4D	3.36	1.43	1.35
3	D	201	HEM	C3C-CAC	-3.35	1.40	1.47
3	С	201	HEM	CHB-C1B	3.30	1.43	1.35
3	В	201	HEM	C3C-CAC	-3.26	1.41	1.47

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\mathbf{Observed}(^{o})$	$\operatorname{Ideal}(^{o})$
3	С	201	HEM	C4B-C3B-C2B	-4.04	103.91	107.11
3	А	201	HEM	C4A-C3A-C2A	-4.03	104.19	107.00
3	А	201	HEM	C4B-C3B-C2B	-3.92	104.00	107.11
3	С	201	HEM	C3B-C2B-C1B	3.76	109.28	106.49
3	В	201	HEM	C2C-C3C-C4C	-3.43	104.51	106.90

There are no chirality outliers.

5	of	87	torsion	outliers	are	listed	below:
0	O1	01	00151011	outiful	$a_{1}c$	noucu	DCIOW.

Mol	Chain	Res	Type	Atoms
4	А	202	S1P	O1-C1-C2-N2
4	А	202	S1P	O1-C1-C2-C3
4	А	202	S1P	C1-C2-C3-O3
4	А	202	S1P	C1-C2-C3-C4
4	А	202	S1P	N2-C2-C3-O3

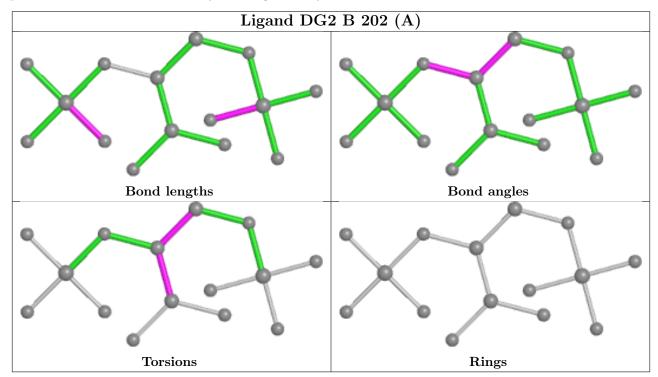


There are no ring outliers.

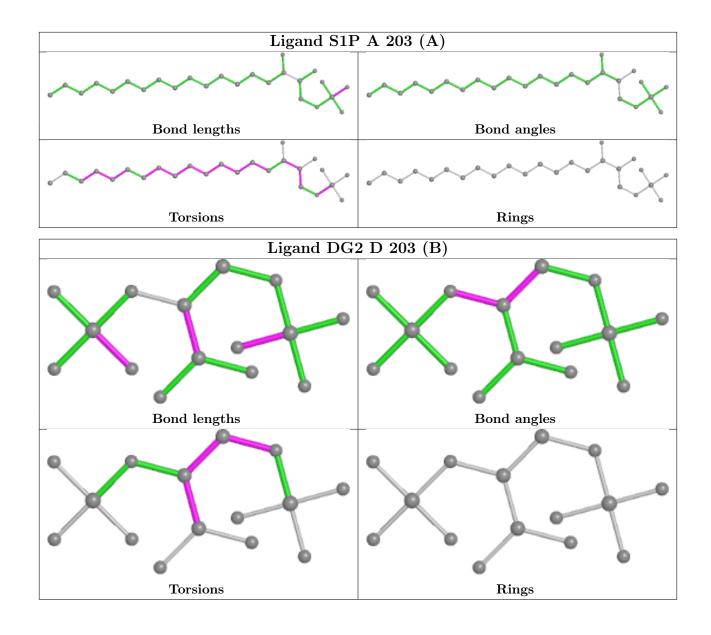
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	201	HEM	1	0
4	С	203	S1P	7	0
4	А	202	S1P	1	0
4	D	202[B]	S1P	11	0
3	В	201	HEM	1	0

5 monomers are involved in 21 short contacts:

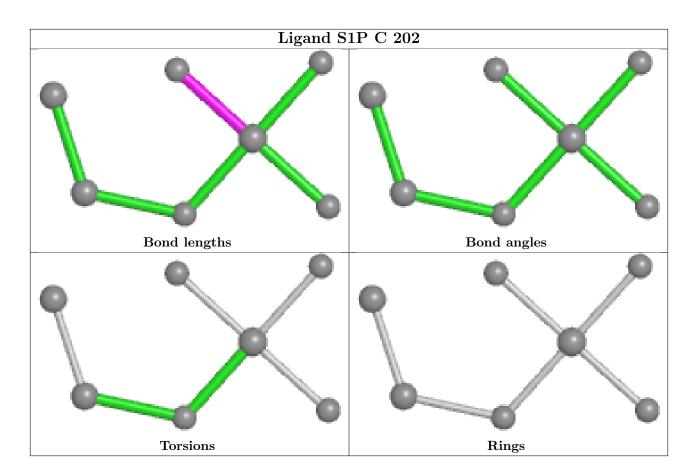
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.



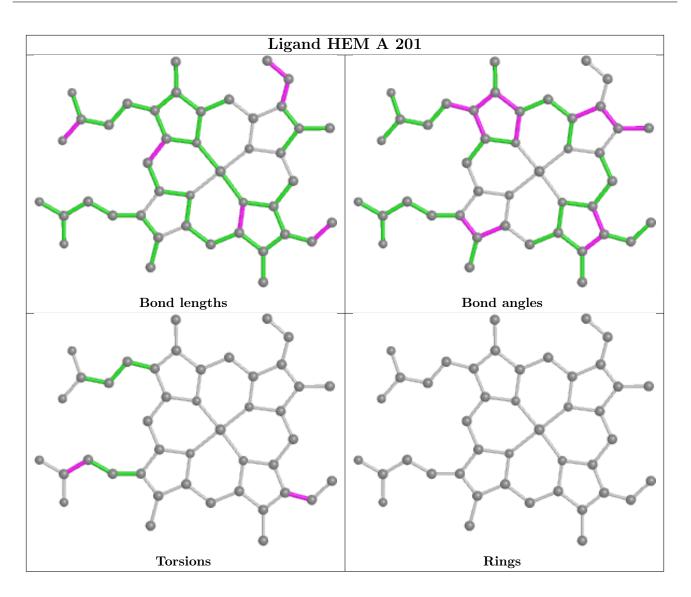




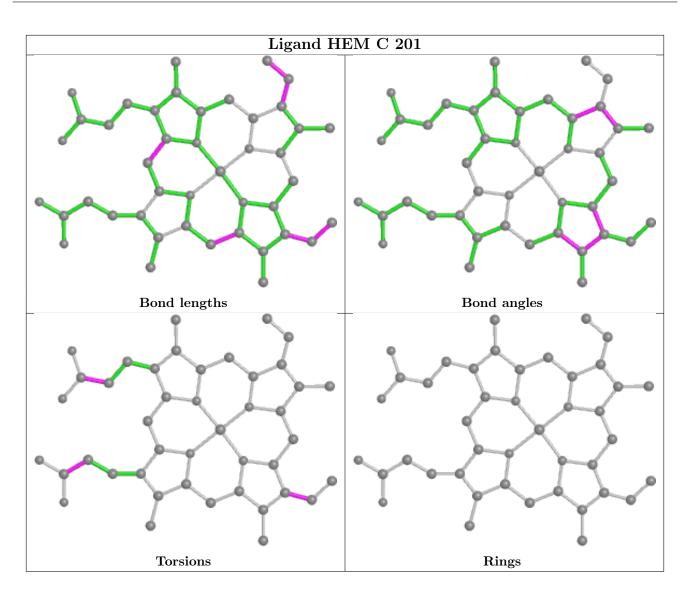




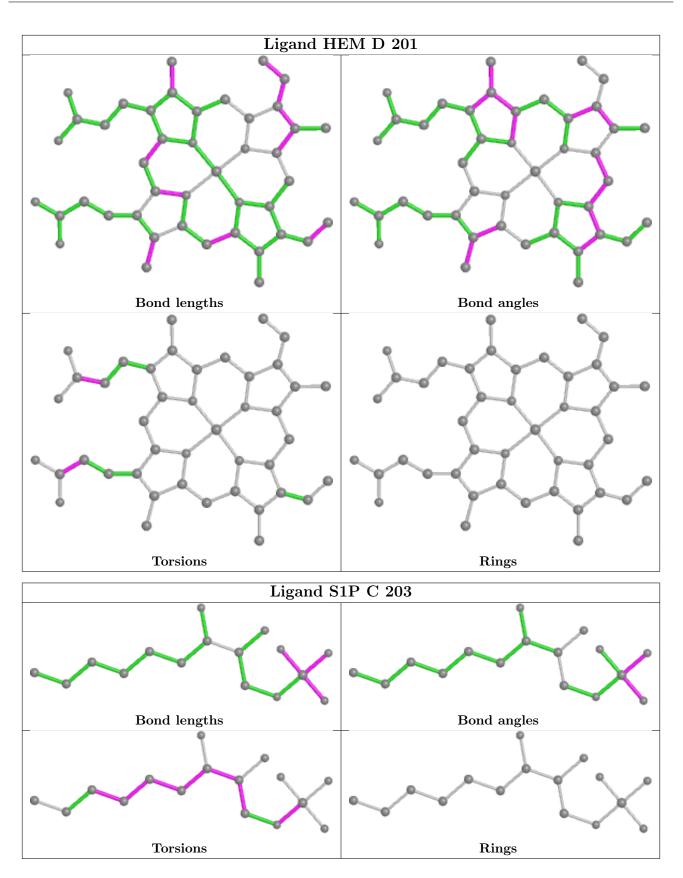




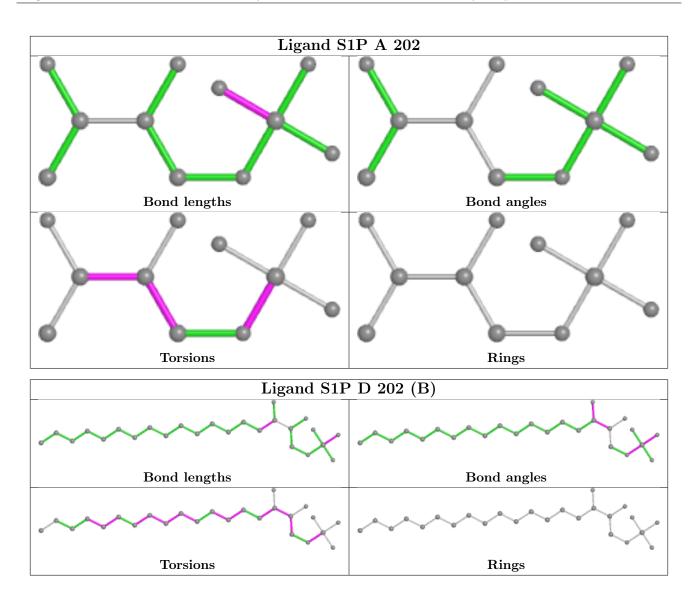




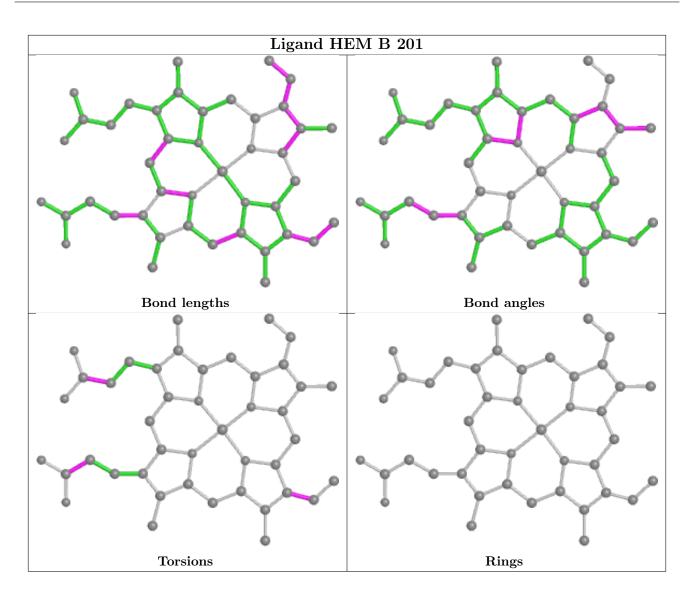












5.7 Other polymers (i)

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, 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	А	141/141~(100%)	-0.29	1 (0%) 87 86	11, 18, 28, 35	0
1	С	141/141 (100%)	-0.42	1 (0%) 87 86	13, 17, 25, 32	0
2	В	146/146~(100%)	-0.01	4 (2%) 54 49	14, 20, 32, 55	0
2	D	146/146 (100%)	-0.25	7 (4%) 30 25	11, 16, 32, 57	0
All	All	574/574~(100%)	-0.24	13 (2%) 60 56	11, 18, 29, 57	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	В	1	VAL	9.0
2	D	1	VAL	6.4
2	D	2	HIS	4.7
2	В	2	HIS	4.6
2	D	5	PRO	3.6

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.

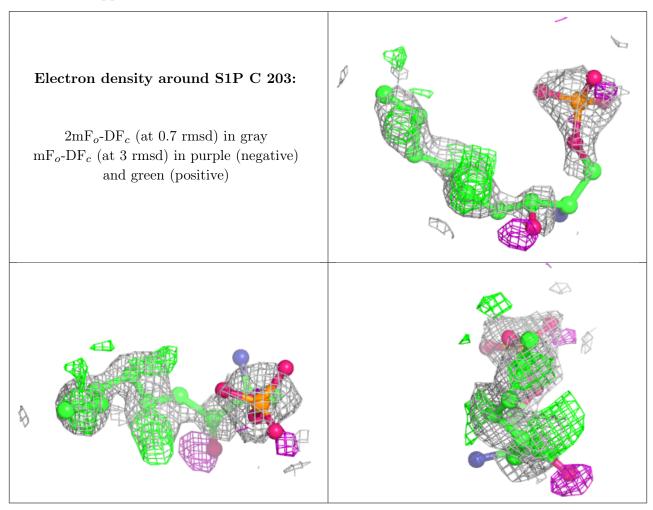
6.4 Ligands (i)

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, 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.

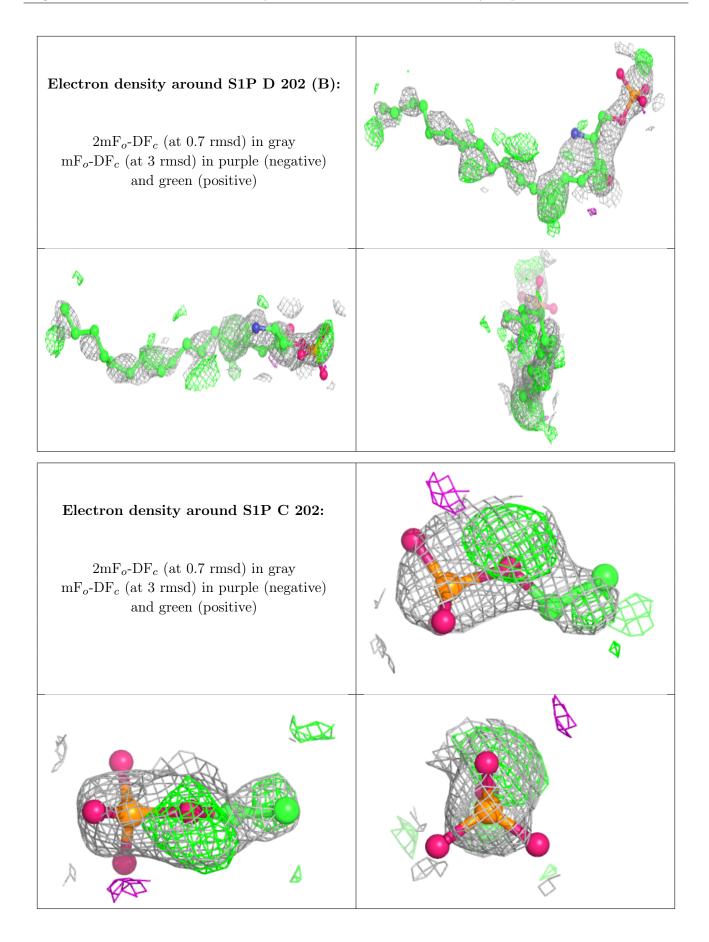


Mol	Type	Chain	Res	Atoms	RSCC	\mathbf{RSR}	B-factors(Å ²)	Q<0.9
4	S1P	С	203	16/25	0.57	0.45	41,48,54,54	16
4	S1P	D	202[B]	25/25	0.59	0.44	40,44,47,47	25
4	S1P	С	202	7/25	0.66	0.55	47,48,49,49	7
5	DG2	D	203[B]	15/15	0.70	0.20	$60,\!61,\!63,\!63$	15
4	S1P	А	203[A]	25/25	0.72	0.41	41,43,47,48	25
4	S1P	А	202	11/25	0.73	0.61	44,45,46,47	11
5	DG2	В	202[A]	15/15	0.76	0.17	54,54,55,55	15
3	HEM	В	201	43/43	0.96	0.10	12,18,33,36	0
3	HEM	С	201	43/43	0.97	0.08	$11,\!15,\!25,\!29$	0
3	HEM	А	201	43/43	0.97	0.09	12,14,23,30	0
3	HEM	D	201	43/43	0.98	0.09	10,13,18,21	0

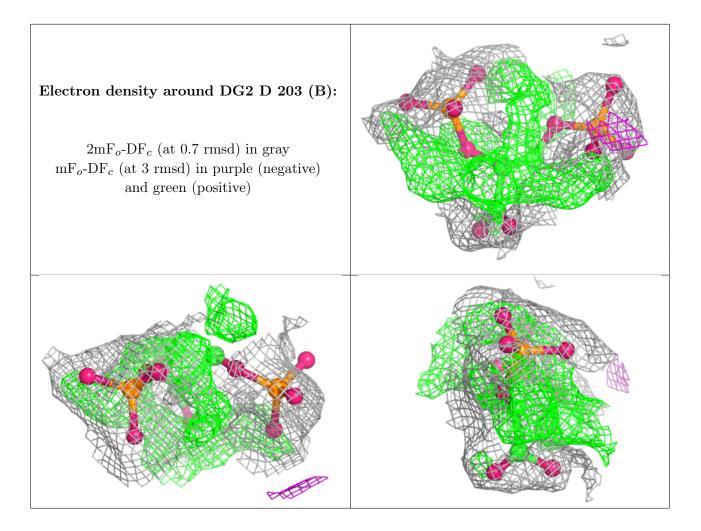
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



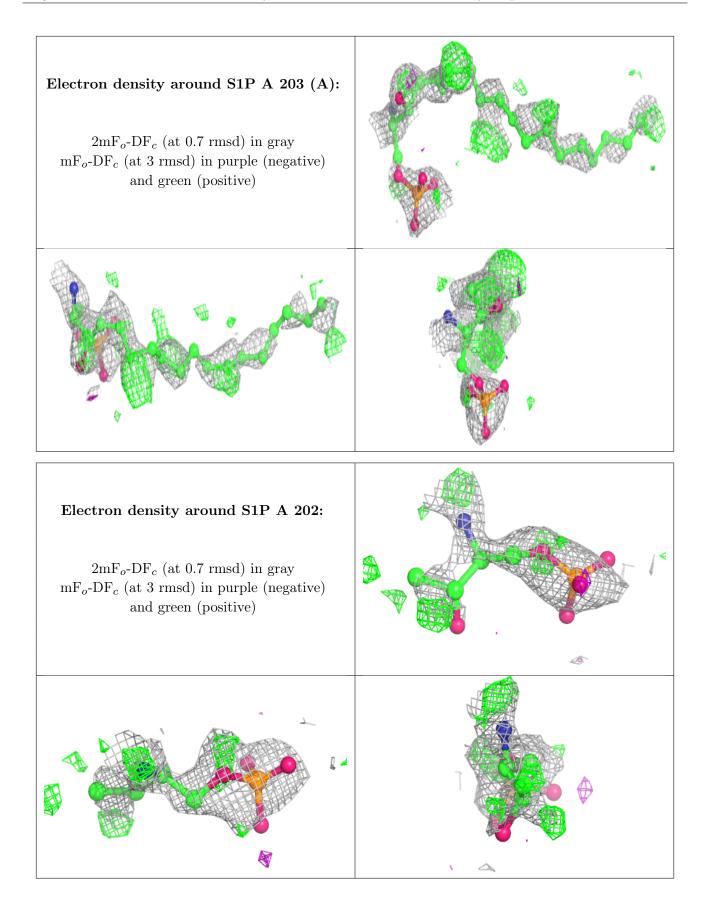




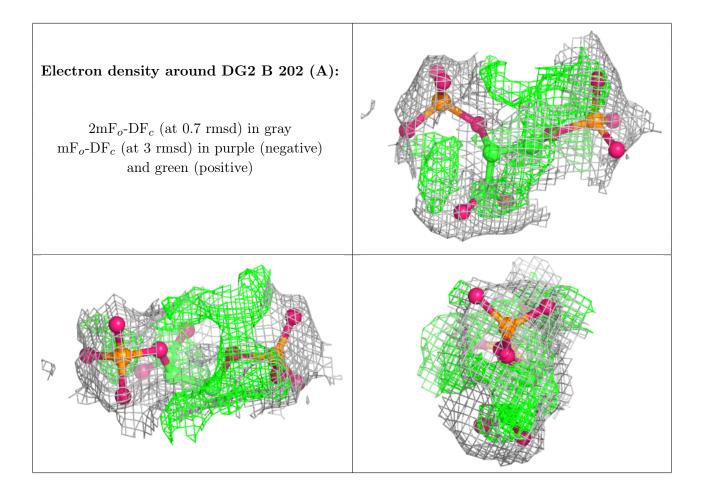




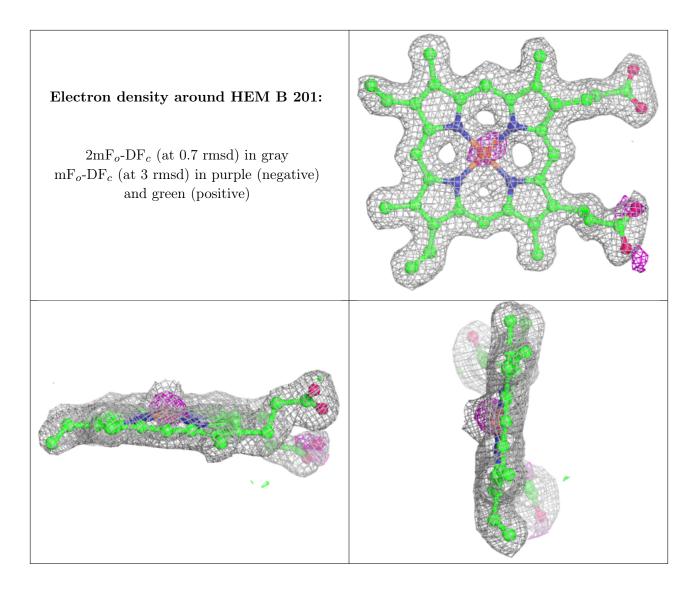




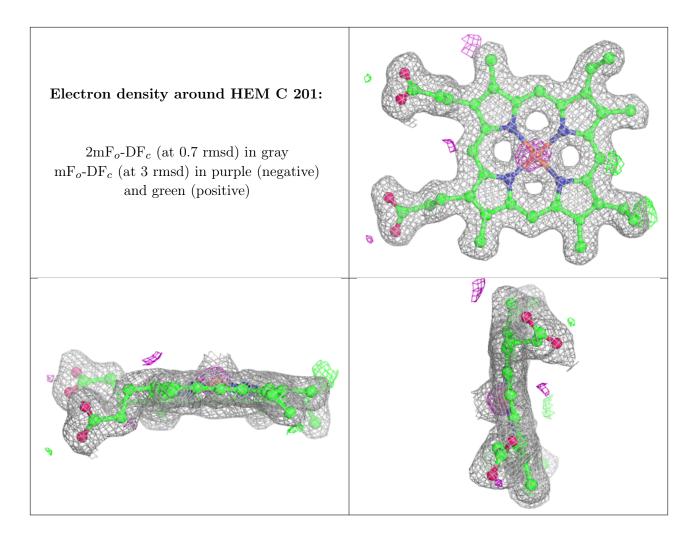




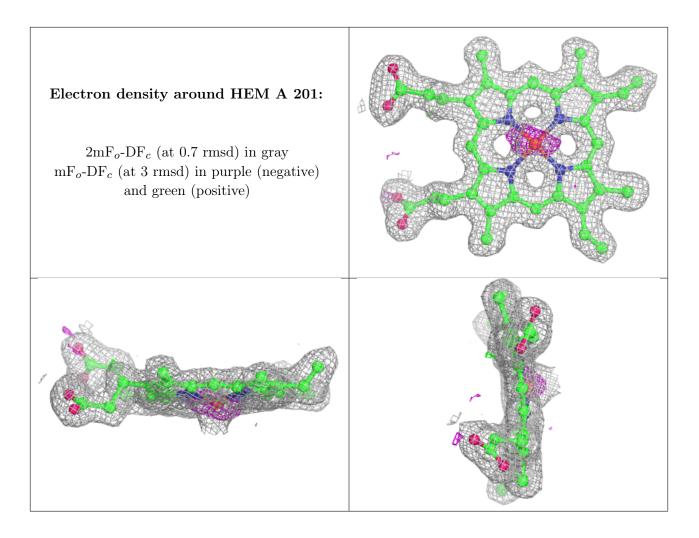




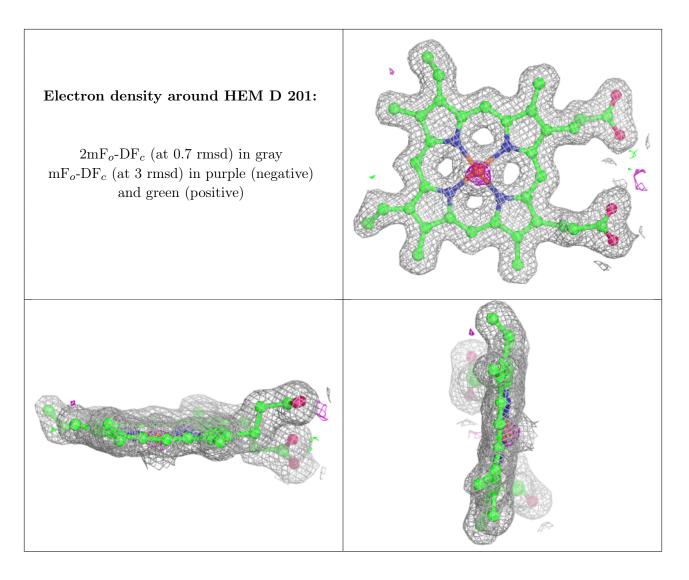












6.5 Other polymers (i)

There are no such residues in this entry.

