

# Full wwPDB X-ray Structure Validation Report (i)

#### Oct 7, 2023 – 11:36 PM EDT

PDB ID	:	6DYC
Title	:	Co(II)-bound structure of the engineered cyt cb562 variant, CH3
Authors	:	Tezcan, F.A.; Rittle, J.
Deposited on		
Resolution	:	1.33 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

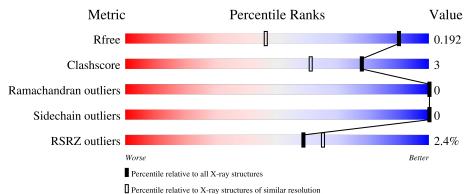
Xtriage (Phenix) EDS buster-report Percentile statistics Refmac CCP4 Ideal geometry (proteins) Ideal geometry (DNA, RNA)	: : : : :	20191225.v01 (using entries in the PDB archive December 25th 2019) 5.8.0158 7.0.044 (Gargrove) Engh & Huber (2001) Parkinson et al. (1996)
Ideal geometry (DNA, RNA) Validation Pipeline (wwPDB-VP)		Parkinson et al. (1996) 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 1.33 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	130704	1385 (1.36-1.32)
Clashscore	141614	1417 (1.36-1.32)
Ramachandran outliers	138981	1397 (1.36-1.32)
Sidechain outliers	138945	1397 (1.36-1.32)
RSRZ outliers	127900	1369 (1.36-1.32)

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 <=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	106	98%			
1	В	106	96%	•		



#### $6 \mathrm{DYC}$

# 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3729 atoms, of which 1713 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 Soluble cytochrome b562.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
1	А	106	Total 1669	C 517	H 827	N 154	0 165	S 6	0	2	0
1	В	106		С	Н	N	0 168	5 6	0	3	0

Chain	Residue	Modelled	Actual	Comment	Reference
А	59	TRP	LYS	engineered mutation	UNP P0ABE7
А	67	HIS	ILE	engineered mutation	UNP P0ABE7
А	71	HIS	GLN	engineered mutation	UNP P0ABE7
А	96	CYS	THR	engineered mutation	UNP P0ABE7
А	97	HIS	THR	engineered mutation	UNP P0ABE7
А	98	CYS	ARG	engineered mutation	UNP P0ABE7
А	101	CYS	TYR	engineered mutation	UNP P0ABE7
В	59	TRP	LYS	engineered mutation	UNP P0ABE7
В	67	HIS	ILE	engineered mutation	UNP P0ABE7
В	71	HIS	GLN	engineered mutation	UNP P0ABE7
В	96	CYS	THR	engineered mutation	UNP P0ABE7
В	97	HIS	THR	engineered mutation	UNP P0ABE7
В	98	CYS	ARG	engineered mutation	UNP P0ABE7
В	101	CYS	TYR	engineered mutation	UNP P0ABE7

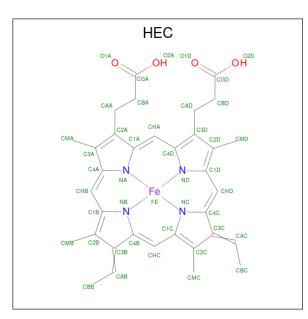
There are 14 discrepancies between the modelled and reference sequences:

• Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co) (labeled as "Ligand of Interest" by depositor).

[	Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
	2	А	1	Total Co 1 1	0	0

• Molecule 3 is HEME C (three-letter code: HEC) (formula:  $C_{34}H_{34}FeN_4O_4$ ).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
2	3 A	1	Total	С	Fe	Η	Ν	Ο	0	0
0		1	75	34	1	32	4	4	0	0
2	р	1	Total	С	Fe	Η	Ν	0	0	0
0	D	1	75	34	1	32	4	4		0

• Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	1	Total Ca 1 1	0	0

• Molecule 5 is water.

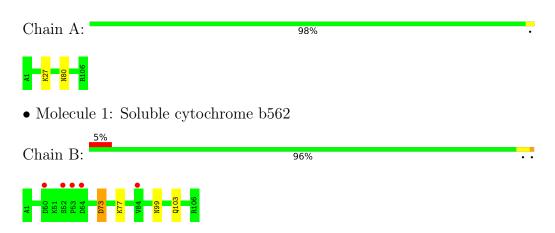
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	135	Total O 141 141	0	6
5	В	100	Total O 101 101	0	1



# 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: Soluble cytochrome b562





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	33.24Å 83.29Å 39.17Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $100.49^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	38.52 - 1.33	Depositor
Resolution (A)	38.52 - 1.33	EDS
% Data completeness	93.0 (38.52-1.33)	Depositor
(in resolution range)	93.0 (38.52-1.33)	EDS
R <sub>merge</sub>	0.07	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.49 (at 1.33 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
D D.	0.176 , $0.192$	Depositor
$R, R_{free}$	0.176 , $0.192$	DCC
$R_{free}$ test set	1999 reflections $(4.48\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	14.0	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.39, $39.4$	EDS
L-test for twinning <sup>2</sup>	$ \langle L  \rangle = 0.49, \langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3729	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 13.28% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical 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.



<sup>&</sup>lt;sup>1</sup>Intensities 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: HEC, CO, CA

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		
	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.44	0/858	0.64	0/1152	
1	В	0.39	0/863	0.57	1/1160~(0.1%)	
All	All	0.42	0/1721	0.60	1/2312~(0.0%)	

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	В	73	ASP	CB-CG-OD1	5.13	122.92	118.30

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	842	827	826	2	0
1	В	844	822	823	2	0
2	А	1	0	0	0	0
3	А	43	32	30	3	0
3	В	43	32	30	3	0
4	В	1	0	0	0	0
5	А	141	0	0	0	0
5	В	101	0	0	0	0

Continued on next page...



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	2016	1713	1709	10	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 3.

All (10) 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:B:201:HEC:HMB1	3:B:201:HEC:HBB3	1.84	0.60
3:B:201:HEC:HMC1	3:B:201:HEC:HBC3	1.85	0.58
3:A:202:HEC:HMC1	3:A:202:HEC:HBC3	1.87	0.57
3:A:202:HEC:HMB1	3:A:202:HEC:HBB3	1.86	0.57
3:A:202:HEC:HMC1	3:A:202:HEC:CBC	2.45	0.47
3:B:201:HEC:HMC1	3:B:201:HEC:CBC	2.46	0.45
1:A:27:LYS:CE	1:A:80:ASN:OD1	2.68	0.42
1:B:99[B]:ASN:OD1	1:B:103:GLN:NE2	2.50	0.42
1:A:27:LYS:HD2	1:A:80:ASN:OD1	2.20	0.41
1:B:73:ASP:O	1:B:77[B]:LYS:HE2	2.20	0.41

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	А	106/106~(100%)	106 (100%)	0	0	100	100
1	В	107/106~(101%)	106 (99%)	1 (1%)	0	100	100
All	All	213/212~(100%)	212 (100%)	1 (0%)	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	88/86~(102%)	88 (100%)	0	100 100		
1	В	89/86~(104%)	89 (100%)	0	100 100		
All	All	177/172~(103%)	177 (100%)	0	100 100		

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	А	6	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)

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



Mol	Mol Type Chain	Res	Link	Bond lengths			Bond angles			
	туре	Ullalli	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
3	HEC	А	202	1	32,50,50	1.88	3 (9%)	24,82,82	1.89	5 (20%)
3	HEC	В	201	1	32,50,50	2.01	3 (9%)	24,82,82	1.71	7 (29%)

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).

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
3	HEC	А	202	1	-	2/10/54/54	-
3	HEC	В	201	1	-	2/10/54/54	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	В	201	HEC	C2B-C3B	-6.42	1.34	1.40
3	В	201	HEC	C3C-C2C	-6.24	1.34	1.40
3	А	202	HEC	C2B-C3B	-6.18	1.34	1.40
3	А	202	HEC	C3C-C2C	-5.15	1.35	1.40
3	А	202	HEC	C3D-C2D	3.83	1.49	1.37
3	В	201	HEC	C3D-C2D	3.56	1.48	1.37

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
3	А	202	HEC	CMC-C2C-C1C	-4.48	121.58	128.46
3	В	201	HEC	CMC-C2C-C1C	-4.08	122.19	128.46
3	А	202	HEC	CMC-C2C-C3C	3.99	130.51	125.82
3	А	202	HEC	CBA-CAA-C2A	-3.98	105.89	112.60
3	А	202	HEC	CBD-CAD-C3D	-3.35	106.91	112.62
3	В	201	HEC	CMB-C2B-C3B	2.65	128.94	125.82
3	А	202	HEC	CMB-C2B-C1B	-2.45	124.70	128.46
3	В	201	HEC	CMB-C2B-C1B	-2.40	124.77	128.46
3	В	201	HEC	CBD-CAD-C3D	-2.40	108.53	112.62
3	В	201	HEC	C2B-C3B-C4B	2.36	108.90	106.35
3	В	201	HEC	C1D-C2D-C3D	-2.10	105.53	107.00
3	В	201	HEC	CMC-C2C-C3C	2.05	128.23	125.82



There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	А	202	HEC	CAA-CBA-CGA-O1A
3	А	202	HEC	CAA-CBA-CGA-O2A
3	В	201	HEC	CAA-CBA-CGA-O2A
3	В	201	HEC	CAA-CBA-CGA-O1A

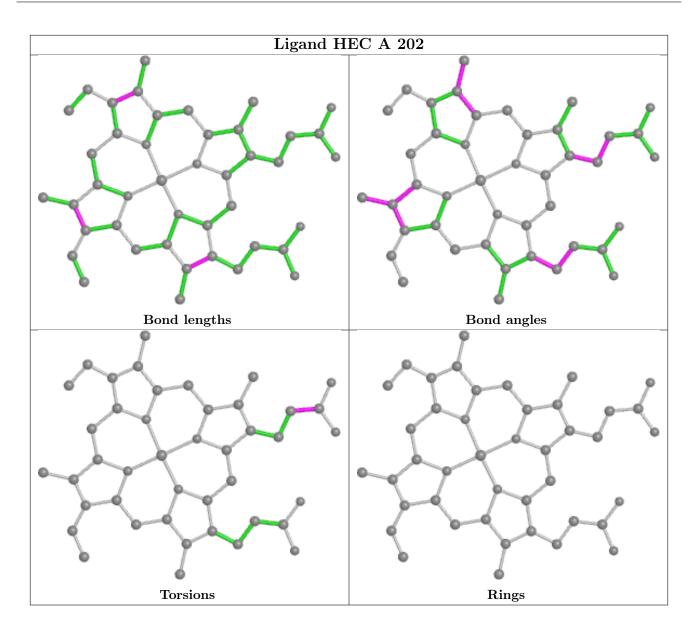
There are no ring outliers.

2 monomers are involved in 6 short contacts:

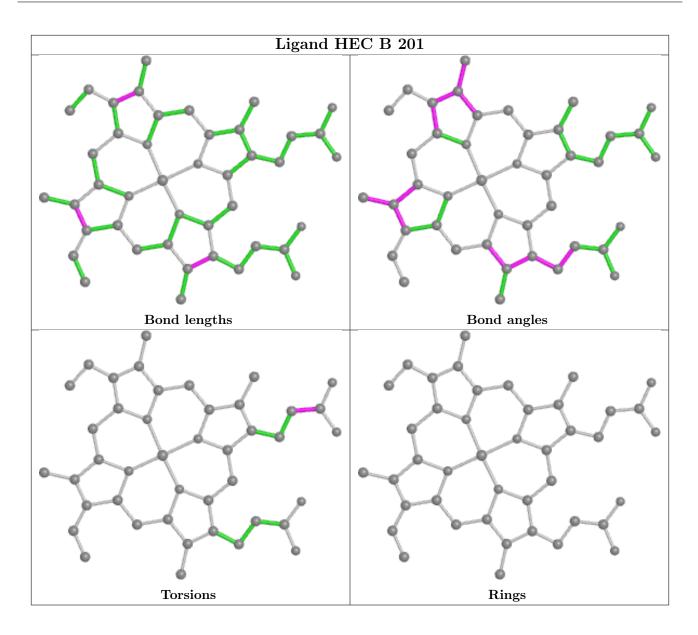
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	А	202	HEC	3	0
3	В	201	HEC	3	0

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 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	$\langle RSRZ \rangle$	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	106/106~(100%)	-0.32	0 100 100	11, 16, 28, 40	0
1	В	106/106~(100%)	0.21	5 (4%) 31 37	12, 20, 42, 58	0
All	All	212/212~(100%)	-0.06	5 (2%) 59 65	11, 18, 38, 58	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	52	SER	3.9
1	В	50	ASP	3.3
1	В	84	VAL	3.1
1	В	54	ASP	2.4
1	В	53	PRO	2.1

## 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.

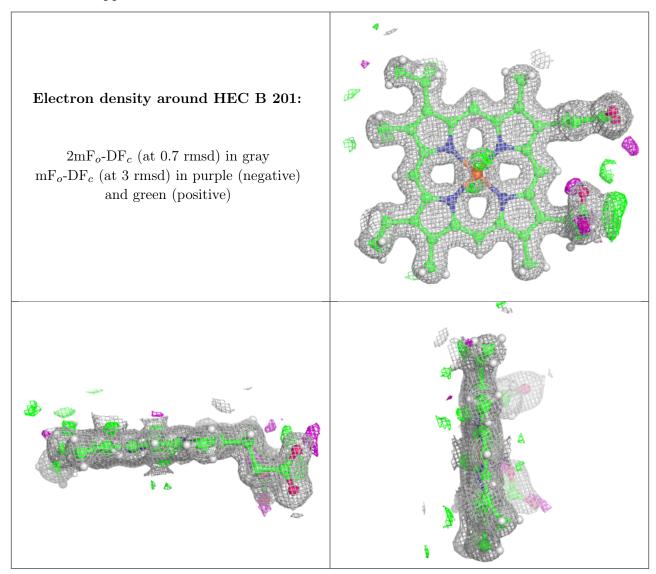
## 6.4 Ligands (i)

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.

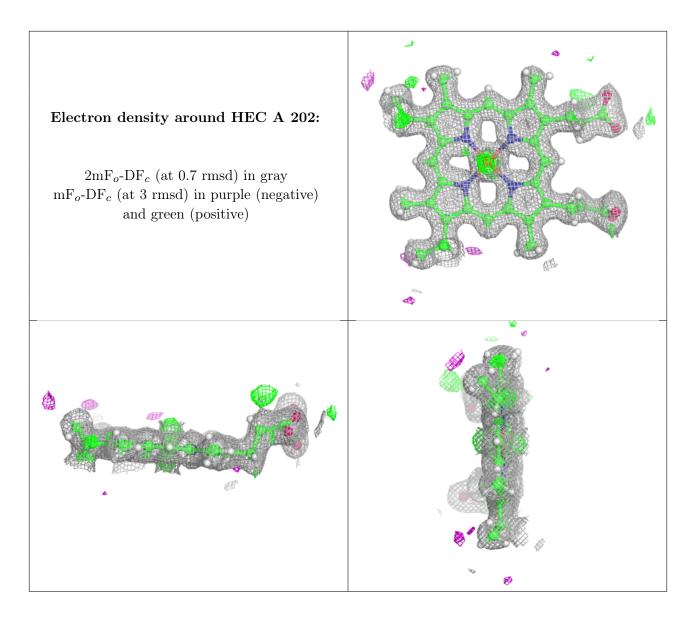


Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
4	CA	В	202	1/1	0.85	0.14	$67,\!67,\!67,\!67$	0
3	HEC	В	201	43/43	0.99	0.08	10,14,26,36	0
3	HEC	А	202	43/43	0.99	0.08	10,14,37,50	0
2	CO	А	201	1/1	1.00	0.03	12,12,12,12	0

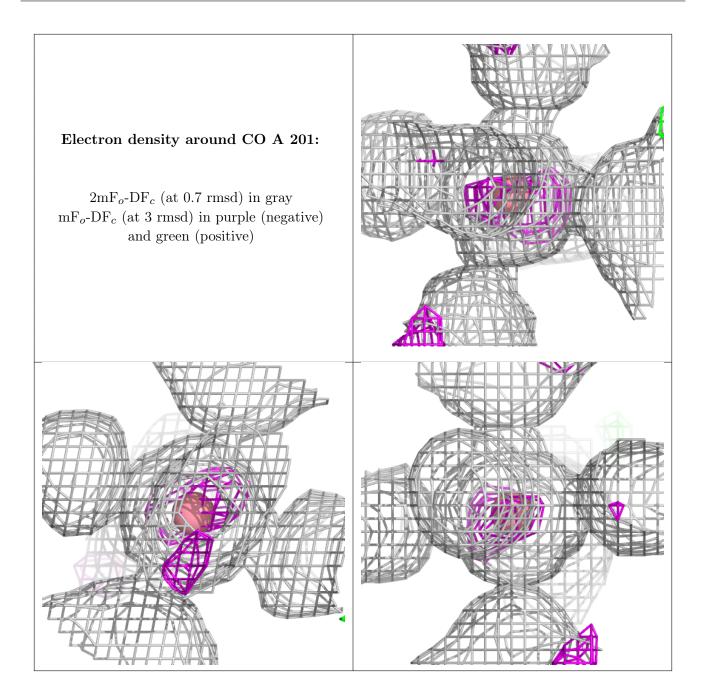
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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











# 6.5 Other polymers (i)

There are no such residues in this entry.

