

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

Nov 22, 2022 – 03:11 pm GMT

PDB ID : 7QAN

Title : Cytochrome P450 Enzyme AbyV Authors : Parnell, A.E.; Back, C.R.; Race, P.R.

Deposited on : 2021-11-17

Resolution : 2.01 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

Mogul : 1.8.4, CSD as541be (2020)

Xtriage (Phenix) : 1.13

EDS : 2.31.3 buster-report : 1.1.7 (2018)

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

Refmac : 5.8.0267

CCP4 : 7.1.010 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

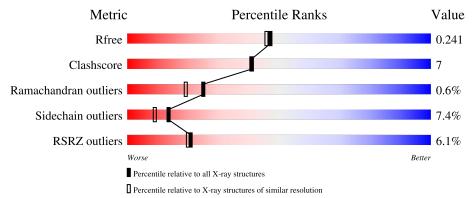
Validation Pipeline (wwPDB-VP) : 2.31.3

# 1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.01 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 $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries,\ resolution\ range(\mathring{\rm A})}) \end{array}$
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	AAA	414	79%		11%	• 9%				
1	BBB	414	64%	18%	•	16%				

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
5	PEG	AAA	406	-	-	-	X



# 2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 11477 atoms, of which 5701 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 Cytochrome P450.

$\mathbf{M}$	ol	Chain	Residues		Atoms					ZeroOcc	AltConf	Trace
]	L	AAA	378	Total 5702	C 1785	H 2863	N 521	O 527	S 6	101	0	0
]	1	BBB	346	Total 5171	C 1621	H 2600	N 466	O 480	S 4	99	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-18	MET	-	initiating methionine	UNP F4F6Q5
AAA	-17	ALA	-	expression tag	UNP F4F6Q5
AAA	-16	HIS	-	expression tag	UNP F4F6Q5
AAA	-15	HIS	-	expression tag	UNP F4F6Q5
AAA	-14	HIS	-	expression tag	UNP F4F6Q5
AAA	-13	HIS	-	expression tag	UNP F4F6Q5
AAA	-12	HIS	-	expression tag	UNP F4F6Q5
AAA	-11	HIS	-	expression tag	UNP F4F6Q5
AAA	-10	SER	-	expression tag	UNP F4F6Q5
AAA	-9	SER	-	expression tag	UNP F4F6Q5
AAA	-8	GLY	-	expression tag	UNP F4F6Q5
AAA	-7	LEU	-	expression tag	UNP F4F6Q5
AAA	-6	GLU	-	expression tag	UNP F4F6Q5
AAA	-5	VAL	-	expression tag	UNP F4F6Q5
AAA	-4	LEU	-	expression tag	UNP F4F6Q5
AAA	-3	PHE	-	expression tag	UNP F4F6Q5
AAA	-2	GLN	-	expression tag	UNP F4F6Q5
AAA	-1	GLY	-	expression tag	UNP F4F6Q5
AAA	0	PRO	-	expression tag	UNP F4F6Q5
BBB	-18	MET	-	initiating methionine	UNP F4F6Q5
BBB	-17	ALA	-	expression tag	UNP F4F6Q5
BBB	-16	HIS	-	expression tag	UNP F4F6Q5
BBB	-15	HIS	-	expression tag	UNP F4F6Q5
BBB	-14	HIS	-	expression tag	UNP F4F6Q5
BBB	-13	HIS	-	expression tag	UNP F4F6Q5

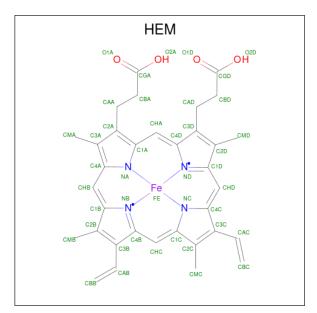
Continued on next page...



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	-12	HIS	-	expression tag	UNP F4F6Q5
BBB	-11	HIS	-	expression tag	UNP F4F6Q5
BBB	-10	SER	-	expression tag	UNP F4F6Q5
BBB	-9	SER	-	expression tag	UNP F4F6Q5
BBB	-8	GLY	-	expression tag	UNP F4F6Q5
BBB	-7	LEU	-	expression tag	UNP F4F6Q5
BBB	-6	GLU	-	expression tag	UNP F4F6Q5
BBB	-5	VAL	-	expression tag	UNP F4F6Q5
BBB	-4	LEU	-	expression tag	UNP F4F6Q5
BBB	-3	PHE	-	expression tag	UNP F4F6Q5
BBB	-2	GLN	-	expression tag	UNP F4F6Q5
BBB	-1	GLY	-	expression tag	UNP F4F6Q5
BBB	0	PRO	-	expression tag	UNP F4F6Q5

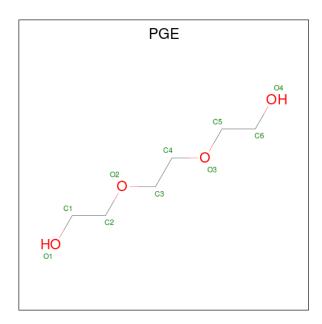
• Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf		
9	AAA	1	Total	С	Fe	Н	N	О	0	0	
	AAA	1	73	34	1	30	4	4	0	U	
2	BBB	1	Total	С	Fe	Н	N	О	0	0	
	מממ	1	73	34	1	30	4	4		U	

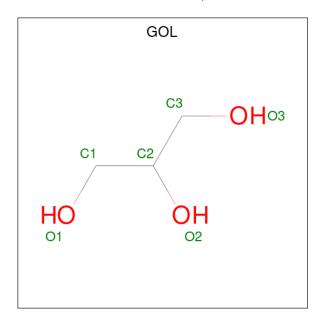
 $\bullet$  Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $\mathrm{C_6H_{14}O_4}).$ 





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	AAA	1	Total 0	C 6	H 14	O 4	1	0

 $\bullet$  Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $\mathrm{C_3H_8O_3}).$ 



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total C H O 14 3 8 3	2	0
4	AAA	1	Total C H O 14 3 8 3	2	0
4	AAA	1	Total C H O 14 3 8 3	2	0

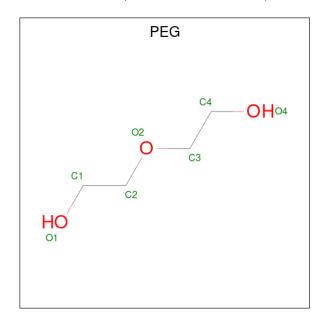
Continued on next page...



Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	BBB	1	Total		H 8		2	0

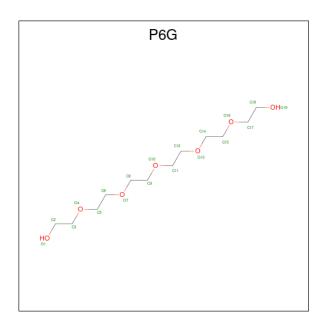
 $\bullet \ \ Molecule \ 5 \ is \ DI(HYDROXYETHYL)ETHER \ (three-letter \ code: \ PEG) \ (formula: \ C_4H_{10}O_3). \\$ 



Mol	Chain	Residues	A	Ator	ns		ZeroOcc	AltConf
5	AAA	1	Total	С	Н	О	1	0
9	AAA	1	17	4	10	3	1	0
5	AAA	1	Total	С	Η	O	1	0
9	ААА	1	17	4	10	3	1	U
5	AAA	1	Total	С	Η	O	1	0
9	АЛА	1	17	4	10	3	1	0
5	AAA	1	Total	С	Н	О	1	0
9	АЛА	1	17	4	10	3	1	0
5	AAA	1	Total	С	Н	О	1	0
	ААА	1	17	4	10	3	1	0

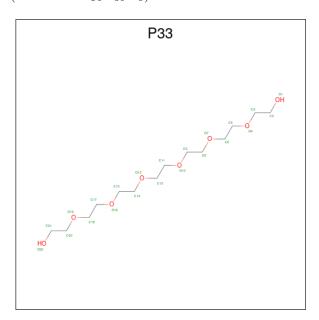
 $\bullet$  Molecule 6 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula:  $\mathrm{C_{12}H_{26}O_{7}}).$ 





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	AAA	1	Total				1	0
0	AAA	1	45	12	26	7	1	0
6	AAA	1	Total	С	Н	О	1	0
	AAA	1	45	12	26	7		U

• Molecule 7 is 3,6,9,12,15,18-HEXAOXAICOSANE-1,20-DIOL (three-letter code: P33) (formula:  $C_{14}H_{30}O_8$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	AAA	1	Total 52	C 14	H 30	O 8	1	0



• Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	AAA	1	Total Cl 1 1	0	0

• Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	AAA	3	Total Mg 3 3	0	0
9	BBB	2	Total Mg 2 2	0	0

• Molecule 10 is water.

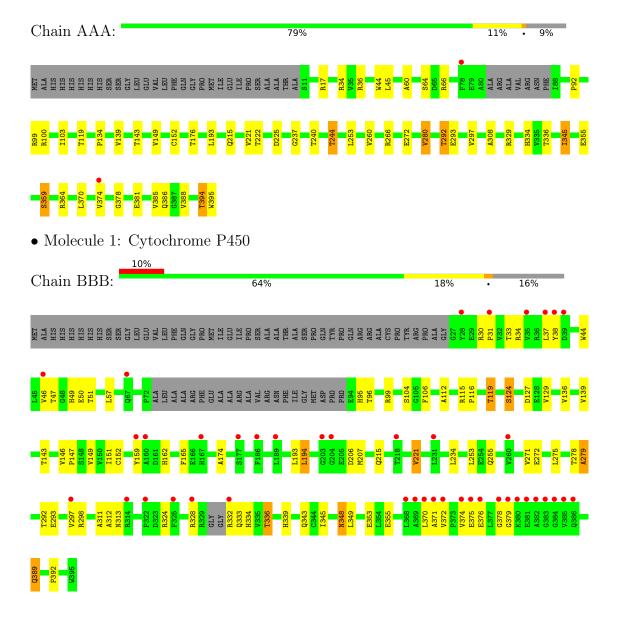
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	AAA	116	Total O 116 116	0	0
10	BBB	29	Total O 29 29	0	0



# 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: Cytochrome P450





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	50.77Å 91.83Å 82.55Å	Donogiton
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $92.36^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	61.44 - 2.01	Depositor
rtesolution (A)	61.36 - 2.01	EDS
% Data completeness	99.8 (61.44-2.01)	Depositor
(in resolution range)	99.9 (61.36-2.01)	EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.34 (at 2.02Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
D D.	0.190 , 0.238	Depositor
$R, R_{free}$	0.198 , $0.241$	DCC
$R_{free}$ test set	2409 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.5	Xtriage
Anisotropy	0.740	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	(Not available), (Not available)	EDS
L-test for twinning <sup>2</sup>	$< L > = 0.49, < L^2> = 0.33$	Xtriage
Estimated twinning fraction	0.030 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11477	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.99% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical values of <|L|>,  $<L^2>$  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: HEM, MG, PEG, GOL, CL, P6G, P33, PGE

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	ond angles
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	AAA	0.72	0/2896	0.87	1/3950 (0.0%)
1	BBB	0.71	0/2618	0.81	0/3574
All	All	0.71	0/5514	0.84	1/7524 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	${f Z}$	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	AAA	292	THR	CB-CA-C	-5.41	97.00	111.60

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	AAA	2839	2863	2846	34	0
1	BBB	2571	2600	2570	43	0
2	AAA	43	30	30	5	0
2	BBB	43	30	30	5	0
3	AAA	10	14	14	1	0
4	AAA	18	24	24	0	0
4	BBB	6	8	8	0	0
5	AAA	35	50	50	1	0

Continued on next page...



Continued	trom	mmoninonic	maaa
COHABABACA		DIEUIUU	DUIUE
0 0 1000100000			

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	AAA	38	52	52	0	0
7	AAA	22	30	30	0	0
8	AAA	1	0	0	0	0
9	AAA	3	0	0	0	0
9	BBB	2	0	0	0	0
10	AAA	116	0	0	2	0
10	BBB	29	0	0	1	0
All	All	5776	5701	5654	80	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 7.

The worst 5 of 80 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

A + a ma 1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	${\rm distance}({\rm \AA})$	overlap (Å)
1:AAA:260:VAL:HG13	1:AAA:266:ARG:HD3	1.59	0.84
1:BBB:95:HIS:HE1	2:BBB:401:HEM:O2D	1.66	0.79
1:BBB:47:THR:HG23	1:BBB:311:ALA:HB1	1.66	0.78
1:BBB:272:GLU:OE2	1:BBB:334:HIS:HE1	1.69	0.76
2:BBB:401:HEM:HBC2	2:BBB:401:HEM:HMC2	1.73	0.70

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	Percent	tiles
1	AAA	374/414 (90%)	365 (98%)	8 (2%)	1 (0%)	41 :	37
1	BBB	340/414 (82%)	319 (94%)	18 (5%)	3 (1%)	17	11
All	All	714/828 (86%)	684 (96%)	26 (4%)	4 (1%)	25	19

All (4) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	BBB	378	GLY
1	BBB	174	ALA
1	BBB	279	ALA
1	AAA	378	GLY

#### 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	AAA	290/320 (91%)	274 (94%)	16 (6%)	21 17		
1	BBB	262/320~(82%)	237 (90%)	25 (10%)	8 5		
All	All	552/640 (86%)	511 (93%)	41 (7%)	13 9		

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	255	GLN
1	BBB	348	ASN
1	BBB	293	GLU
1	BBB	328	ARG
1	BBB	374	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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 21 ligands modelled in this entry, 6 are monoatomic - leaving 15 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).

N / - 1	Т	Clasia	Das	Link	Во	ond leng	ths	В	ond ang	gles
Mol	Type	Chain	Res	Link	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	AAA	412	-	5,5,5	0.11	0	5,5,5	0.35	0
2	HEM	BBB	401	10,1	41,50,50	1.34	6 (14%)	45,82,82	1.69	10 (22%)
5	PEG	AAA	413	-	6,6,6	0.41	0	5,5,5	0.31	0
4	GOL	AAA	405	-	5,5,5	0.11	0	5,5,5	0.27	0
5	PEG	AAA	404	-	6,6,6	0.14	0	5,5,5	0.15	0
5	PEG	AAA	407	-	6,6,6	0.25	0	5,5,5	0.10	0
3	PGE	AAA	402	-	9,9,9	0.13	0	8,8,8	0.15	0
4	GOL	AAA	403	-	5,5,5	0.07	0	5,5,5	0.25	0
6	P6G	AAA	408	-	18,18,18	0.54	0	17,17,17	0.19	0
6	P6G	AAA	410	-	18,18,18	0.61	0	17,17,17	0.41	0
5	PEG	AAA	406	-	6,6,6	0.17	0	5,5,5	0.09	0
7	P33	AAA	411	-	21,21,21	0.67	0	20,20,20	0.39	0
2	HEM	AAA	401	10,1	41,50,50	1.36	7 (17%)	45,82,82	1.79	9 (20%)
5	PEG	AAA	409	-	6,6,6	0.23	0	5,5,5	0.17	0
4	GOL	BBB	402	-	5,5,5	0.09	0	5,5,5	0.25	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
4	GOL	AAA	412	-	-	2/4/4/4	-
2	HEM	BBB	401	10,1	-	0/12/54/54	-
5	PEG	AAA	413	-	-	2/4/4/4	-
4	GOL	AAA	405	-	-	1/4/4/4	-

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PEG	AAA	404	-	-	1/4/4/4	-
5	PEG	AAA	407	-	-	2/4/4/4	-
3	PGE	AAA	402	-	-	3/7/7/7	-
4	GOL	AAA	403	-	-	4/4/4/4	-
6	P6G	AAA	408	-	-	7/16/16/16	-
6	P6G	AAA	410	-	-	8/16/16/16	-
5	PEG	AAA	406	-	-	1/4/4/4	-
7	P33	AAA	411	-	-	11/19/19/19	-
2	HEM	AAA	401	10,1	-	0/12/54/54	-
5	PEG	AAA	409	-	_	2/4/4/4	_
4	GOL	BBB	402	-	-	0/4/4/4	-

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(A)
2	BBB	401	HEM	C1B-NB	-3.46	1.34	1.40
2	BBB	401	HEM	C4B-NB	-3.15	1.32	1.38
2	BBB	401	HEM	FE-NB	2.74	2.10	1.96
2	AAA	401	HEM	C3B-C4B	2.73	1.50	1.44
2	AAA	401	HEM	C1B-NB	-2.70	1.35	1.40

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
2	AAA	401	HEM	C1B-NB-C4B	5.46	110.72	105.07
2	BBB	401	HEM	C1B-NB-C4B	4.40	109.62	105.07
2	AAA	401	HEM	CHC-C4B-NB	4.31	129.12	124.43
2	BBB	401	HEM	CHC-C4B-NB	3.79	128.55	124.43
2	AAA	401	HEM	O2A-CGA-O1A	-3.51	114.56	123.30

There are no chirality outliers.

5 of 44 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	AAA	403	GOL	O1-C1-C2-O2
4	AAA	412	GOL	O1-C1-C2-C3
7	AAA	411	P33	C18-C17-O16-C15
6	AAA	408	P6G	O10-C11-C12-O13
6	AAA	408	P6G	O13-C14-C15-O16



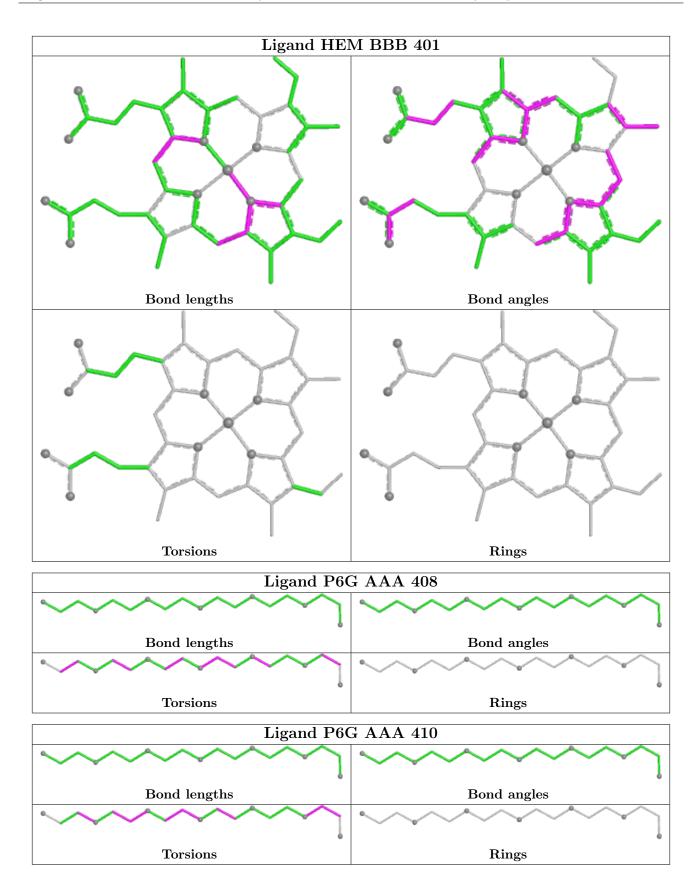
There are no ring outliers.

4 monomers are involved in 12 short contacts:

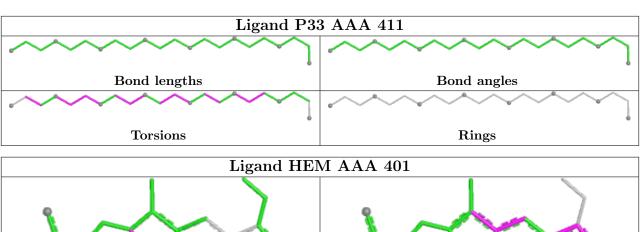
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	BBB	401	HEM	5	0
3	AAA	402	PGE	1	0
2	AAA	401	HEM	5	0
5	AAA	409	PEG	1	0

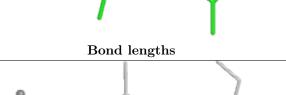
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

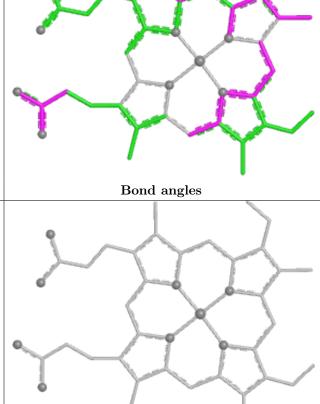




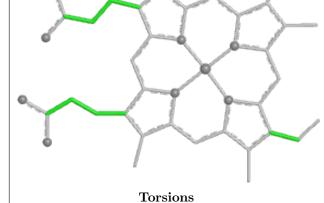


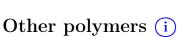






Rings





There are no such residues in this entry.

5.7

## 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	$\langle { m RSRZ} \rangle$	$\#\mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q<0.9
1	AAA	378/414 (91%)	0.26	2 (0%) 91 90	27, 39, 69, 94	0
1	BBB	346/414 (83%)	0.79	42 (12%) 4 3	40, 66, 103, 131	0
All	All	724/828 (87%)	0.51	44 (6%) 21 20	27, 52, 92, 131	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	382	ALA	7.7
1	BBB	378	GLY	7.3
1	BBB	372	VAL	6.2
1	BBB	384	GLY	5.0
1	BBB	383	GLY	4.9

### 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)

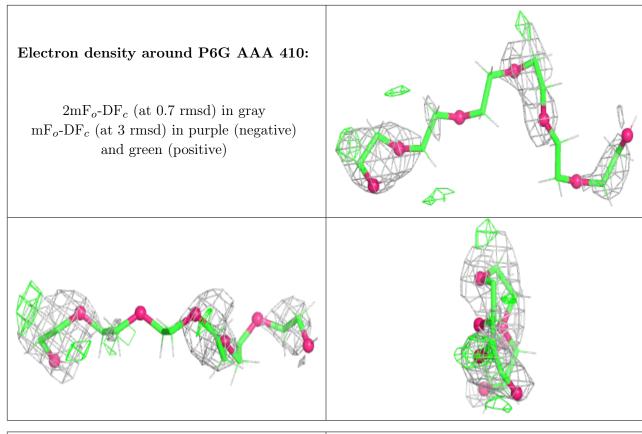
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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



Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\operatorname{B-factors}({ m \AA}^2)$	Q<0.9
6	P6G	AAA	410	19/19	0.53	0.36	94,108,116,117	1
6	P6G	AAA	408	19/19	0.54	0.26	82,103,109,109	1
4	GOL	BBB	402	6/6	0.60	0.16	86,98,104,104	2
5	PEG	AAA	406	7/7	0.77	0.59	81,93,109,109	1
5	PEG	AAA	404	7/7	0.80	0.19	63,69,75,77	1
5	PEG	AAA	409	7/7	0.81	0.27	51,63,81,81	1
4	GOL	AAA	412	6/6	0.82	0.15	77,79,96,96	2
4	GOL	AAA	405	6/6	0.82	0.17	65,70,76,76	2
5	PEG	AAA	407	7/7	0.82	0.28	58,71,76,76	1
7	P33	AAA	411	22/22	0.82	0.17	63,80,84,88	1
5	PEG	AAA	413	7/7	0.86	0.16	57,63,84,84	1
3	PGE	AAA	402	10/10	0.86	0.19	64,66,70,72	1
8	CL	AAA	414	1/1	0.91	0.20	88,88,88,88	0
4	GOL	AAA	403	6/6	0.92	0.11	56,70,73,73	2
9	MG	AAA	416	1/1	0.93	0.34	67,67,67,67	0
2	HEM	BBB	401	43/43	0.97	0.13	42,48,55,59	0
9	MG	AAA	415	1/1	0.98	0.20	44,44,44,44	0
2	HEM	AAA	401	43/43	0.98	0.15	27,32,38,40	0
9	MG	AAA	417	1/1	0.98	0.18	33,33,33,33	0
9	MG	BBB	404	1/1	0.98	0.57	64,64,64,64	0
9	MG	BBB	403	1/1	0.99	0.29	64,64,64,64	0

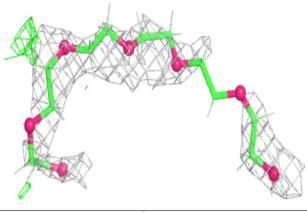
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

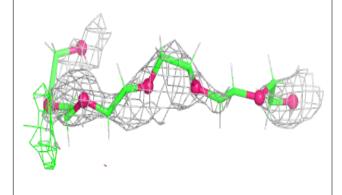


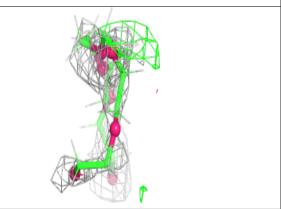


#### Electron density around P6G AAA 408:

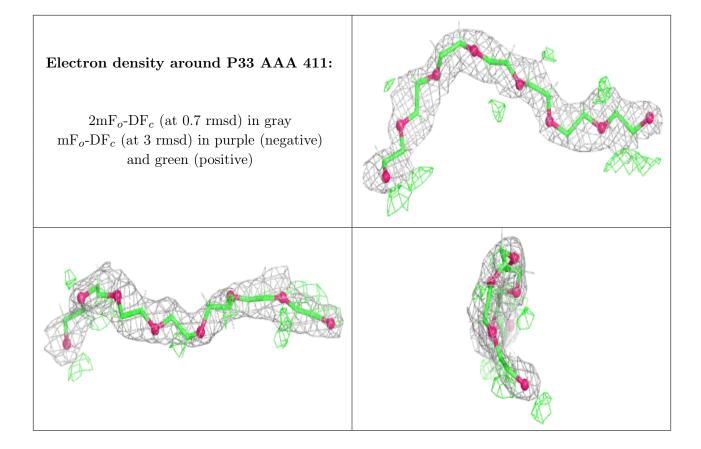
 $2 {\rm mF}_o\text{-}{\rm DF}_c$  (at 0.7 rmsd) in gray  ${\rm mF}_o\text{-}{\rm DF}_c$  (at 3 rmsd) in purple (negative) and green (positive)



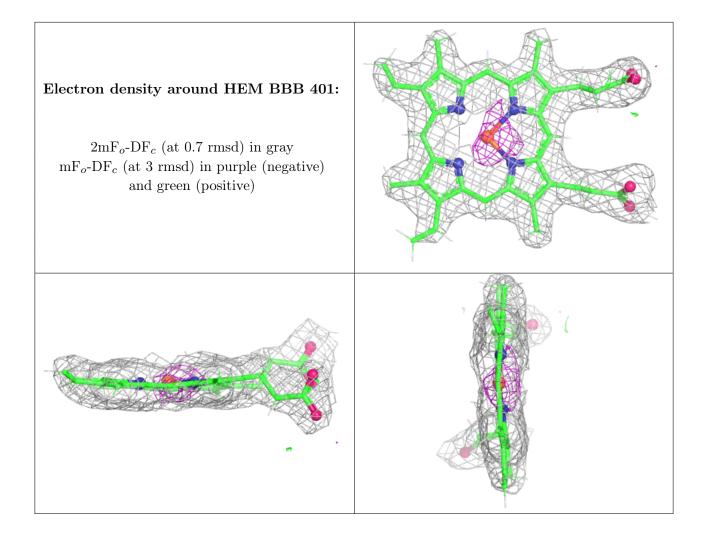




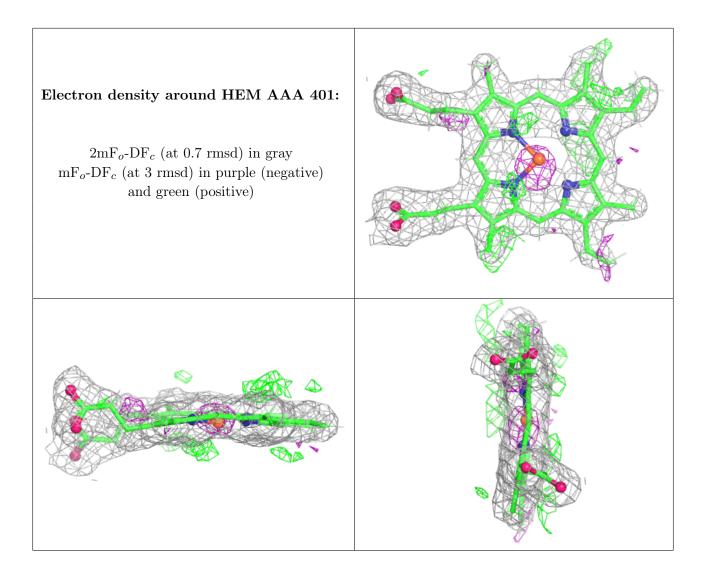












# 6.5 Other polymers (i)

There are no such residues in this entry.

