



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2020 – 08:15 PM BST

PDB ID : 5BV2  
Title : Crystal structure of E. coli HP1I catalase variant  
Authors : Wang, J.; Lomkalin, I.V.  
Deposited on : 2015-06-04  
Resolution : 1.53 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13  
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) : 2.13

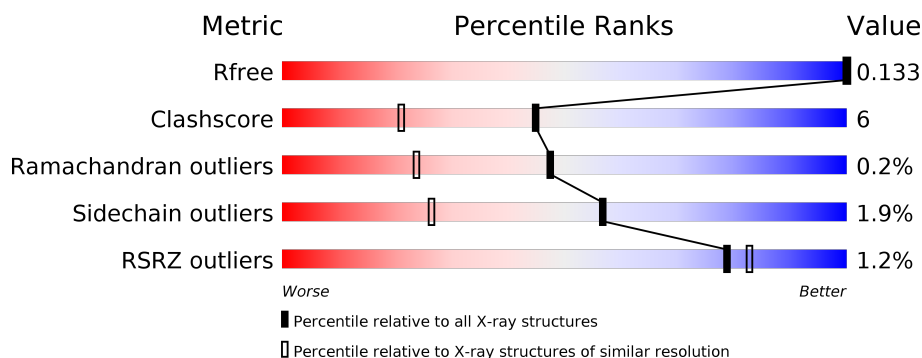
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.53 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))
$R_{free}$	130704	2556 (1.56-1.52)
Clashscore	141614	2634 (1.56-1.52)
Ramachandran outliers	138981	2580 (1.56-1.52)
Sidechain outliers	138945	2577 (1.56-1.52)
RSRZ outliers	127900	2524 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	P	753	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>9%</div> <div>.</div> </div> </div>
1	Q	753	<div> <div>%</div> <div> <div></div> <div>91%</div> <div>8%</div> <div>..</div> </div> </div>
1	R	753	<div> <div>2%</div> <div> <div></div> <div>91%</div> <div>8%</div> <div>.</div> </div> </div>
1	S	753	<div> <div>%</div> <div> <div></div> <div>91%</div> <div>7%</div> <div>..</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	P	802	-	-	X	-
3	GOL	R	802	-	-	X	-
3	GOL	S	802	-	-	X	-
5	PGE	R	807	-	-	X	-
6	PEG	P	814	-	-	X	-
6	PEG	P	817	-	-	X	-

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 53930 atoms, of which 24516 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 Catalase HPIL.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	P	747	Total	C	H	N	O	S	103	23	0
			11980	3836	5945	1059	1125	15			
1	Q	746	Total	C	H	N	O	S	106	29	0
			12077	3864	6005	1065	1128	15			
1	R	747	Total	C	H	N	O	S	106	23	0
			11933	3824	5920	1050	1124	15			
1	S	747	Total	C	H	N	O	S	107	32	0
			12118	3875	6028	1069	1131	15			

There are 24 discrepancies between the modelled and reference sequences:

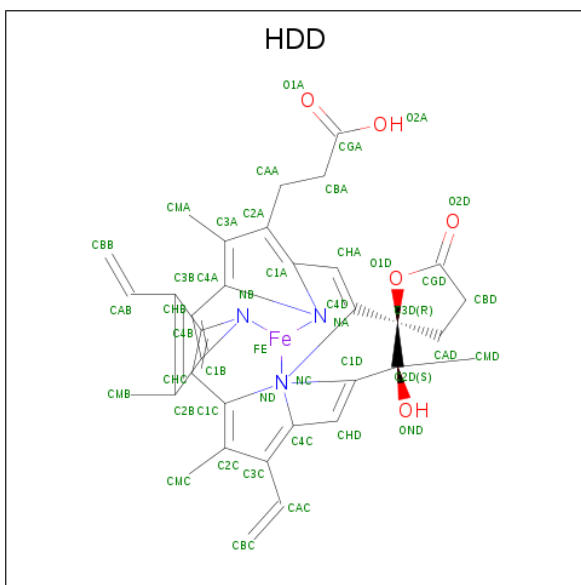
Chain	Residue	Modelled	Actual	Comment	Reference
P	99	ASP	SER	conflict	UNP P21179
P	283	SER	GLU	conflict	UNP P21179
P	372	ASN	LYS	conflict	UNP P21179
P	565	SER	GLU	conflict	UNP P21179
P	710	VAL	ILE	conflict	UNP P21179
P	750	SER	LYS	conflict	UNP P21179
Q	99	ASP	SER	conflict	UNP P21179
Q	283	SER	GLU	conflict	UNP P21179
Q	372	ASN	LYS	conflict	UNP P21179
Q	565	SER	GLU	conflict	UNP P21179
Q	710	VAL	ILE	conflict	UNP P21179
Q	750	SER	LYS	conflict	UNP P21179
R	99	ASP	SER	conflict	UNP P21179
R	283	SER	GLU	conflict	UNP P21179
R	372	ASN	LYS	conflict	UNP P21179
R	565	SER	GLU	conflict	UNP P21179
R	710	VAL	ILE	conflict	UNP P21179
R	750	SER	LYS	conflict	UNP P21179
S	99	ASP	SER	conflict	UNP P21179
S	283	SER	GLU	conflict	UNP P21179
S	372	ASN	LYS	conflict	UNP P21179

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
S	565	SER	GLU	conflict	UNP P21179
S	710	VAL	ILE	conflict	UNP P21179
S	750	SER	LYS	conflict	UNP P21179

- Molecule 2 is CIS-HEME D HYDROXYCHLORIN GAMMA-SPIROLACTONE (three-letter code: HDD) (formula:  $C_{34}H_{32}FeN_4O_5$ ).



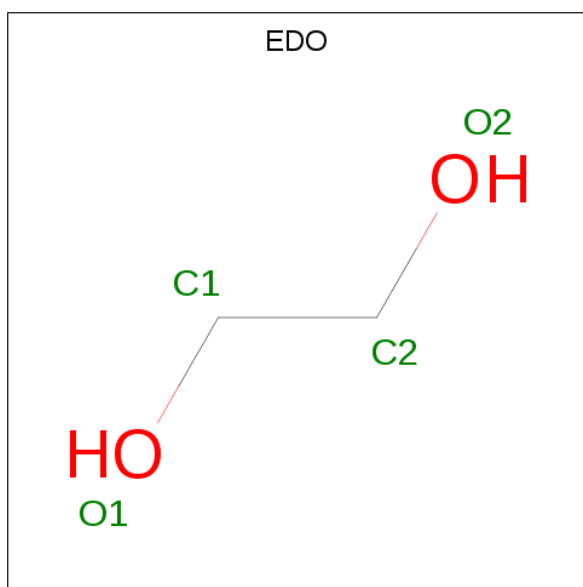
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	P	1	Total	C	Fe	H	N	O	1	0
			75	34	1	31	4	5		
2	Q	1	Total	C	Fe	H	N	O	1	0
			75	34	1	31	4	5		
2	R	1	Total	C	Fe	H	N	O	1	0
			75	34	1	31	4	5		
2	S	1	Total	C	Fe	H	N	O	1	0
			75	34	1	31	4	5		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	P	1	Total	C	H	O	0	0
			14	3	8	3		
3	P	1	Total	C	H	O	2	0
			14	3	8	3		
3	Q	1	Total	C	H	O	0	0
			14	3	8	3		
3	Q	1	Total	C	H	O	2	0
			14	3	8	3		
3	Q	1	Total	C	H	O	2	0
			14	3	8	3		
3	R	1	Total	C	H	O	0	0
			14	3	8	3		
3	R	1	Total	C	H	O	2	0
			14	3	8	3		
3	R	1	Total	C	H	O	2	0
			14	3	8	3		
3	S	1	Total	C	H	O	0	0
			14	3	8	3		
3	S	1	Total	C	H	O	2	0
			14	3	8	3		
3	S	1	Total	C	H	O	2	0
			14	3	8	3		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	1	0
4	P	1	Total 10	C 2	H 6	O 2	0	0
4	Q	1	Total 10	C 2	H 6	O 2	1	0
4	Q	1	Total 10	C 2	H 6	O 2	1	0
4	Q	1	Total 10	C 2	H 6	O 2	1	0
4	Q	1	Total 10	C 2	H 6	O 2	1	0
4	Q	1	Total 10	C 2	H 6	O 2	1	0

*Continued on next page...*

*Continued from previous page...*

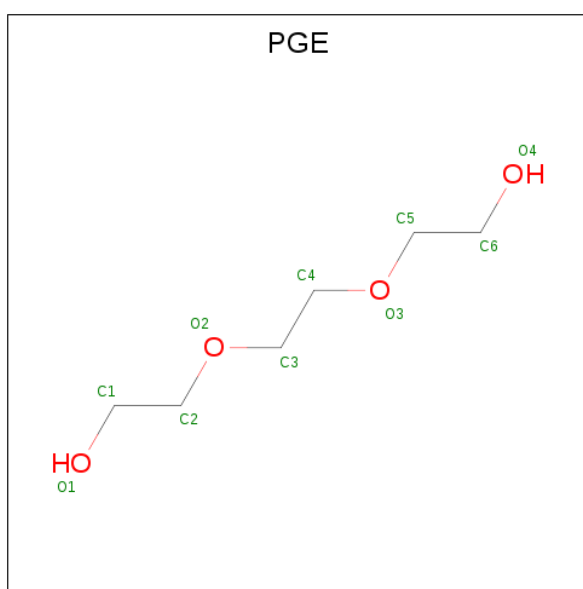
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	Q	1	Total	C	H	O	1	0
			10	2	6	2		
4	R	1	Total	C	H	O	1	0
			10	2	6	2		
4	R	1	Total	C	H	O	1	0
			10	2	6	2		
4	R	1	Total	C	H	O	1	0
			10	2	6	2		
4	R	1	Total	C	H	O	1	0
			10	2	6	2		
4	S	1	Total	C	H	O	1	0
			10	2	6	2		
4	S	1	Total	C	H	O	1	0
			10	2	6	2		
4	S	1	Total	C	H	O	1	0
			10	2	6	2		
4	S	1	Total	C	H	O	1	0
			10	2	6	2		

*Continued on next page...*

Continued from previous page...

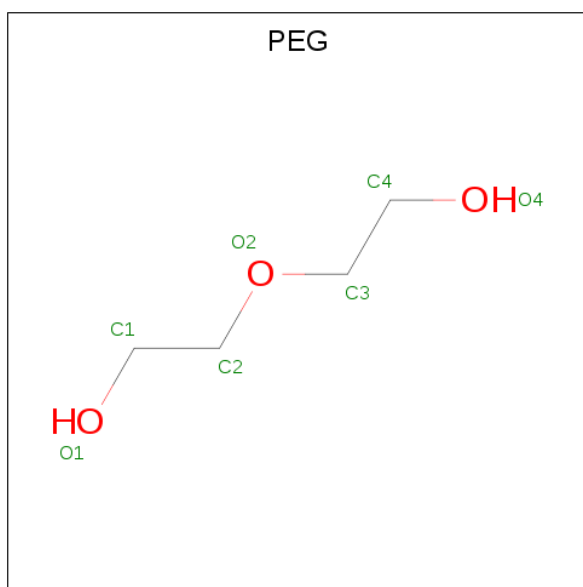
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	S	1	Total	C	H	O	1	0
			10	2	6	2		
4	S	1	Total	C	H	O	1	0
			10	2	6	2		
4	S	1	Total	C	H	O	1	0
			10	2	6	2		
4	S	1	Total	C	H	O	1	0
			10	2	6	2		

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



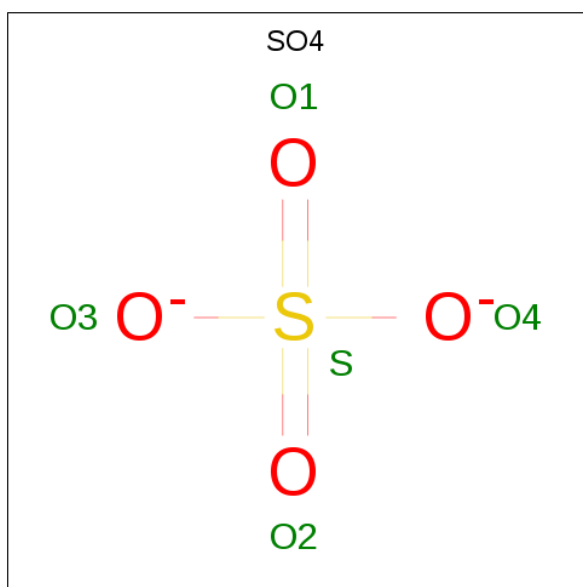
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	P	1	Total	C	H	O	1	0
			24	6	14	4		
5	R	1	Total	C	H	O	1	0
			24	6	14	4		
5	R	1	Total	C	H	O	1	0
			24	6	14	4		

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	P	1	Total	C	H	O	1	0
			17	4	10	3		
6	P	1	Total	C	H	O	1	0
			17	4	10	3		
6	P	1	Total	C	H	O	1	0
			17	4	10	3		
6	P	1	Total	C	H	O	1	0
			17	4	10	3		
6	Q	1	Total	C	H	O	1	0
			17	4	10	3		
6	Q	1	Total	C	H	O	1	0
			17	4	10	3		
6	Q	1	Total	C	H	O	1	0
			17	4	10	3		
6	Q	1	Total	C	H	O	1	0
			17	4	10	3		
6	S	1	Total	C	H	O	1	0
			17	4	10	3		
6	S	1	Total	C	H	O	1	0
			17	4	10	3		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).

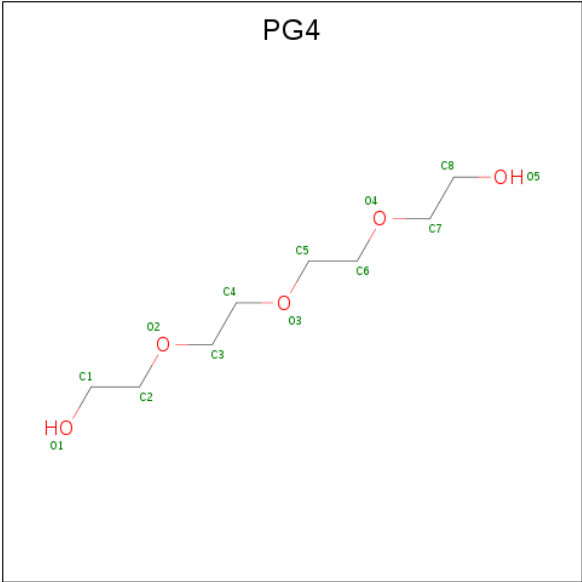


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	P	1	Total	O	S	0	0
			5	4	1		
7	S	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	R	1	Total	Mg	0	0
			1	1		
8	S	1	Total	Mg	0	0
			1	1		

- Molecule 9 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	R	1	Total	C	H	O	1	0
			31	8	18	5		

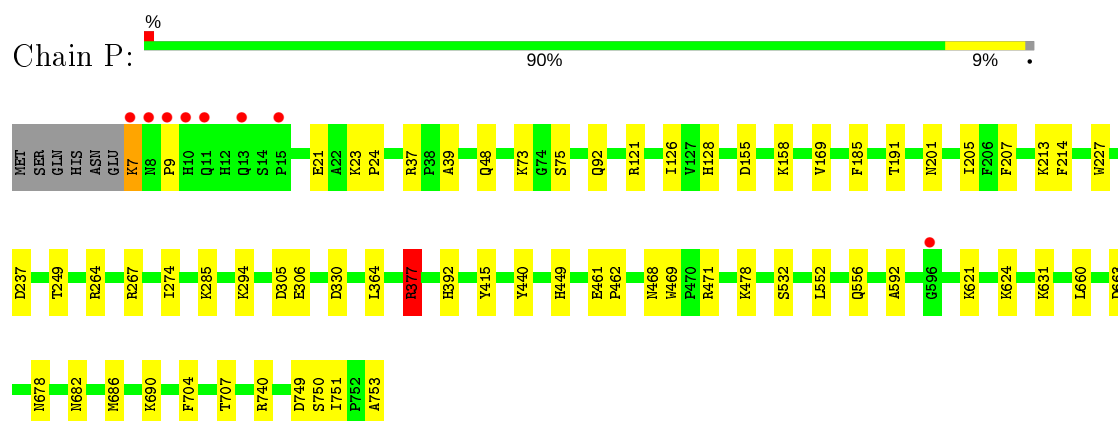
- Molecule 10 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	P	1171	Total	O		0	0
			1171	1171			
10	Q	1196	Total	O		0	0
			1196	1196			
10	R	1087	Total	H	O	5	0
			1093	6	1087		
10	S	1227	Total	H	O	5	0
			1233	6	1227		

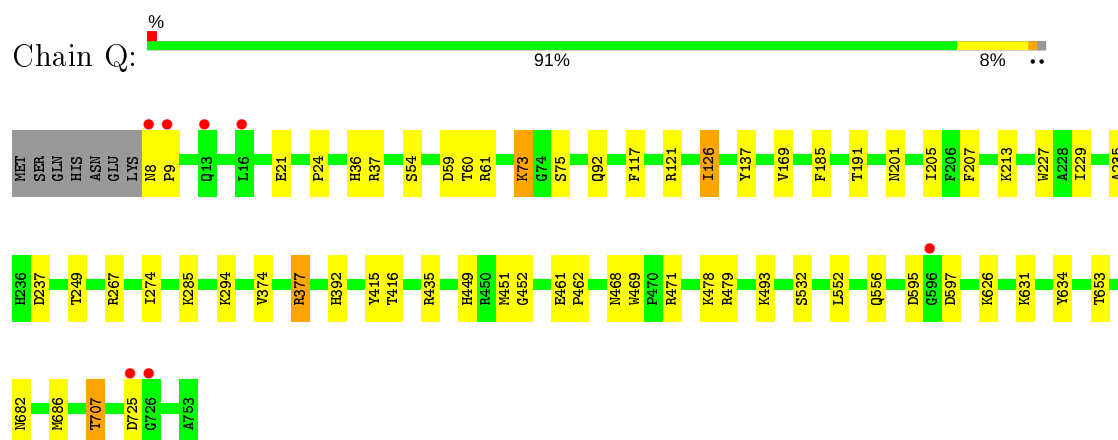
### 3 Residue-property plots [i](#)

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.

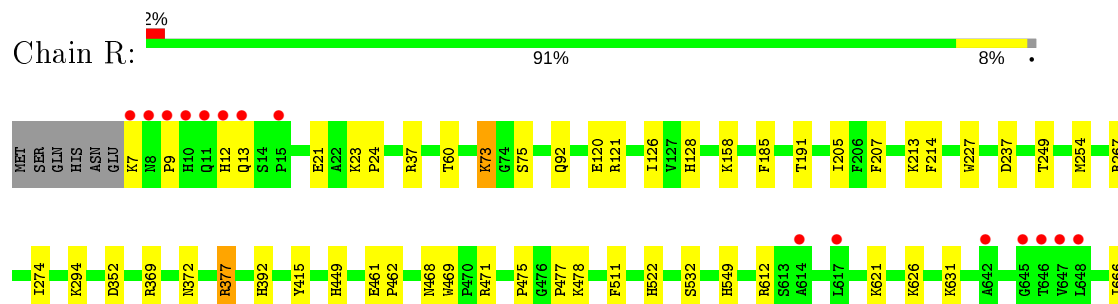
#### • Molecule 1: Catalase HP11



#### • Molecule 1: Catalase HP11

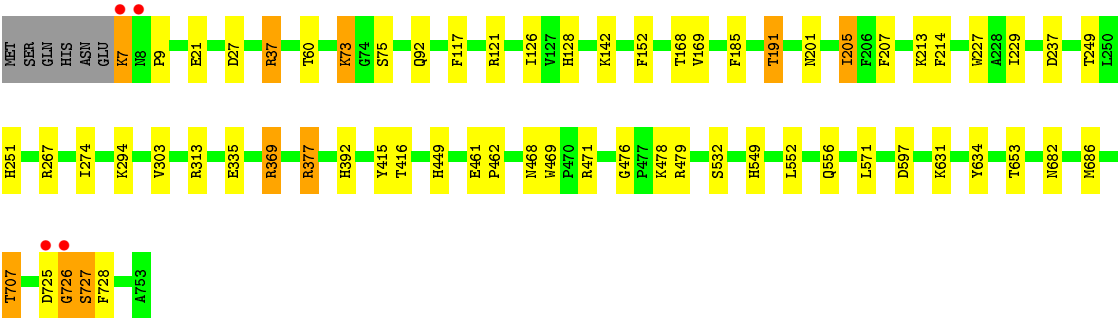
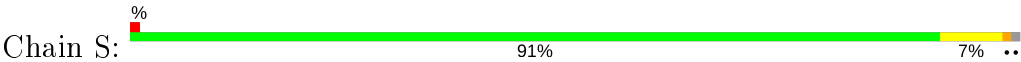


#### • Molecule 1: Catalase HP11





● Molecule 1: Catalase HP11



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	161.79Å 171.33Å 122.48Å 90.00° 121.55° 90.00°	Depositor
Resolution (Å)	50.00 – 1.53 48.62 – 1.53	Depositor EDS
% Data completeness (in resolution range)	71.8 (50.00-1.53) 71.8 (48.62-1.53)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.17 (at 1.53Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.082 , 0.132 0.084 , 0.133	Depositor DCC
$R_{free}$ test set	15186 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	13.3	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 59.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.99	EDS
Total number of atoms	53930	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.81% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<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.

## 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: GOL, MG, PGE, EDO, PG4, HDD, SO4, PEG

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	P	0.43	0/6270	0.68	5/8523 (0.1%)
1	Q	0.44	0/6331	0.65	2/8607 (0.0%)
1	R	0.43	0/6245	0.63	1/8485 (0.0%)
1	S	0.44	0/6349	0.65	3/8622 (0.0%)
All	All	0.43	0/25195	0.65	11/34237 (0.0%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	377[A]	ARG	NE-CZ-NH1	-12.04	114.28	120.30
1	P	377[B]	ARG	NE-CZ-NH1	-12.04	114.28	120.30
1	P	377[A]	ARG	NE-CZ-NH2	11.13	125.86	120.30
1	P	377[B]	ARG	NE-CZ-NH2	11.13	125.86	120.30
1	Q	725[A]	ASP	C-N-CA	5.74	134.35	122.30

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	P	6035	5945	5939	85	0
1	Q	6072	6005	6007	72	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	R	6013	5920	5914	78	1
1	S	6090	6028	6030	81	0
2	P	44	31	31	8	0
2	Q	44	31	31	7	0
2	R	44	31	31	7	0
2	S	44	31	31	9	0
3	P	12	16	16	6	0
3	Q	18	24	24	1	0
3	R	18	24	23	13	0
3	S	18	24	24	4	0
4	P	36	54	54	6	0
4	Q	68	102	102	5	2
4	R	16	24	24	1	0
4	S	36	54	54	4	0
5	P	10	14	14	0	0
5	R	20	28	28	8	0
6	P	28	40	40	21	0
6	Q	28	40	40	3	0
6	S	14	20	20	3	0
7	P	5	0	0	0	0
7	S	5	0	0	0	0
8	R	1	0	0	0	0
8	S	1	0	0	0	0
9	R	13	18	18	0	0
10	P	1171	0	0	26	0
10	Q	1196	0	0	22	0
10	R	1087	6	0	26	1
10	S	1227	6	0	37	0
All	All	29414	24516	24495	319	3

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 6.

The worst 5 of 319 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:Q:449[B]:HIS:CD2	1:Q:451[B]:MET:CE	1.77	1.61
1:Q:392:HIS:ND1	1:Q:415:TYR:CB	1.86	1.38
1:S:392:HIS:ND1	1:S:415:TYR:CB	1.86	1.36
1:R:392:HIS:ND1	1:R:415:TYR:CB	1.87	1.35
1:P:392:HIS:ND1	1:P:415:TYR:CB	1.88	1.35

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:Q:805:EDO:O1	4:Q:805:EDO:O1[2_656]	0.98	1.22
4:Q:805:EDO:O1	4:Q:805:EDO:HO1[2_656]	1.41	0.19
1:R:621:LYS:HZ3	10:R:1773:HOH:O[2_556]	1.51	0.09

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	P	770/753 (102%)	751 (98%)	18 (2%)	1 (0%)	51	26
1	Q	778/753 (103%)	761 (98%)	16 (2%)	1 (0%)	51	26
1	R	769/753 (102%)	751 (98%)	17 (2%)	1 (0%)	51	26
1	S	780/753 (104%)	761 (98%)	17 (2%)	2 (0%)	41	19
All	All	3097/3012 (103%)	3024 (98%)	68 (2%)	5 (0%)	47	24

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	S	727	SER
1	P	75	SER
1	R	75	SER
1	S	75	SER
1	Q	75	SER

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was

analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	P	655/636 (103%)	642 (98%)	13 (2%)	55	24
1	Q	662/636 (104%)	644 (97%)	18 (3%)	44	15
1	R	654/636 (103%)	640 (98%)	14 (2%)	53	22
1	S	664/636 (104%)	642 (97%)	22 (3%)	38	10
All	All	2635/2544 (104%)	2568 (98%)	67 (2%)	57	17

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	Q	707[B]	THR
1	R	205	ILE
1	S	369[B]	ARG
1	R	21	GLU
1	R	73[C]	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	Q	492	ASN
1	S	549	HIS
1	R	549	HIS
1	P	556	GLN
1	Q	549	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 72 ligands modelled in this entry, 2 are monoatomic - leaving 70 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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	Q	815	-	3,3,3	0.56	0	2,2,2	0.06	0
4	EDO	S	809	-	3,3,3	0.44	0	2,2,2	0.39	0
4	EDO	P	805	-	3,3,3	0.49	0	2,2,2	0.18	0
4	EDO	S	805	-	3,3,3	0.43	0	2,2,2	0.23	0
6	PEG	P	816	-	6,6,6	0.47	0	5,5,5	0.33	0
4	EDO	Q	805	-	3,3,3	0.69	0	2,2,2	0.11	0
6	PEG	Q	823	-	6,6,6	0.69	0	5,5,5	0.60	0
4	EDO	Q	804	-	3,3,3	0.34	0	2,2,2	0.44	0
6	PEG	S	816	-	6,6,6	0.50	0	5,5,5	0.63	0
4	EDO	S	808	-	3,3,3	0.55	0	2,2,2	0.27	0
4	EDO	P	806	-	3,3,3	0.37	0	2,2,2	0.52	0
4	EDO	P	809	-	3,3,3	0.64	0	2,2,2	0.41	0
4	EDO	S	804	-	3,3,3	0.36	0	2,2,2	0.39	0
3	GOL	Q	821	-	5,5,5	0.38	0	5,5,5	0.22	0
2	HDD	P	801	1,3	38,52,52	1.62	6 (15%)	28,89,89	2.75	11 (39%)
4	EDO	S	807	-	3,3,3	0.49	0	2,2,2	0.11	0
4	EDO	S	810	-	3,3,3	0.55	0	2,2,2	0.19	0
2	HDD	Q	801	1,3	38,52,52	1.65	8 (21%)	28,89,89	2.93	13 (46%)
4	EDO	P	804	-	3,3,3	0.57	0	2,2,2	0.11	0
5	PGE	R	808	-	9,9,9	0.60	0	8,8,8	0.39	0
4	EDO	Q	811	-	3,3,3	0.57	0	2,2,2	0.06	0
3	GOL	P	813	-	5,5,5	0.51	0	5,5,5	0.45	0
3	GOL	R	809	-	5,5,5	0.45	0	5,5,5	0.70	0
7	SO4	P	818	-	4,4,4	0.32	0	6,6,6	0.11	0
4	EDO	Q	814	-	3,3,3	0.37	0	2,2,2	0.85	0
4	EDO	R	803	-	3,3,3	0.66	0	2,2,2	0.08	0
3	GOL	R	802	2	5,5,5	0.48	0	5,5,5	1.13	0
4	EDO	Q	810	-	3,3,3	0.44	0	2,2,2	0.50	0
3	GOL	S	802	2	5,5,5	0.73	0	5,5,5	0.41	0
4	EDO	P	810	-	3,3,3	0.51	0	2,2,2	0.27	0
3	GOL	S	812	-	5,5,5	0.56	0	5,5,5	0.69	0
7	SO4	S	814	-	4,4,4	0.44	0	6,6,6	0.24	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	Q	806	-	3,3,3	0.56	0	2,2,2	0.12	0
3	GOL	R	810	-	5,5,5	0.44	0	5,5,5	0.77	0
4	EDO	Q	816	-	3,3,3	0.43	0	2,2,2	0.51	0
4	EDO	P	808	-	3,3,3	0.54	0	2,2,2	0.15	0
9	PG4	R	812	-	12,12,12	0.52	0	11,11,11	0.29	0
4	EDO	Q	817	-	3,3,3	0.49	0	2,2,2	0.24	0
4	EDO	Q	812	-	3,3,3	0.51	0	2,2,2	0.19	0
4	EDO	Q	813	-	3,3,3	0.43	0	2,2,2	0.32	0
5	PGE	P	812	-	9,9,9	0.55	0	8,8,8	0.60	0
3	GOL	P	802	2	5,5,5	0.70	0	5,5,5	1.21	1 (20%)
5	PGE	R	807	-	9,9,9	0.38	0	8,8,8	0.54	0
6	PEG	Q	825	-	6,6,6	0.53	0	5,5,5	0.45	0
3	GOL	Q	802	2	5,5,5	0.55	0	5,5,5	0.50	0
4	EDO	S	806	-	3,3,3	0.55	0	2,2,2	0.16	0
4	EDO	S	803	-	3,3,3	0.43	0	2,2,2	0.39	0
3	GOL	Q	820	-	5,5,5	0.47	0	5,5,5	0.32	0
4	EDO	R	804	-	3,3,3	0.56	0	2,2,2	0.24	0
4	EDO	S	811	-	3,3,3	0.51	0	2,2,2	0.25	0
2	HDD	S	801	1,3	38,52,52	1.68	6 (15%)	28,89,89	2.82	12 (42%)
6	PEG	P	817	-	6,6,6	0.56	0	5,5,5	1.28	0
4	EDO	R	805	-	3,3,3	0.63	0	2,2,2	0.54	0
6	PEG	P	814	-	6,6,6	0.71	0	5,5,5	0.81	0
4	EDO	Q	809	-	3,3,3	0.47	0	2,2,2	0.55	0
4	EDO	Q	808	-	3,3,3	0.44	0	2,2,2	0.51	0
3	GOL	S	813	-	5,5,5	0.32	0	5,5,5	0.75	0
2	HDD	R	801	1,3	38,52,52	1.61	7 (18%)	28,89,89	2.85	12 (42%)
4	EDO	R	806	-	3,3,3	0.45	0	2,2,2	0.22	0
4	EDO	P	811	-	3,3,3	0.37	0	2,2,2	0.37	0
6	PEG	S	817	-	6,6,6	0.70	0	5,5,5	0.72	0
6	PEG	P	815	-	6,6,6	0.39	0	5,5,5	0.37	0
4	EDO	Q	807	-	3,3,3	0.64	0	2,2,2	0.43	0
6	PEG	Q	822	-	6,6,6	0.42	0	5,5,5	0.30	0
4	EDO	Q	819	-	3,3,3	0.74	0	2,2,2	0.19	0
4	EDO	P	807	-	3,3,3	0.44	0	2,2,2	0.16	0
4	EDO	Q	803	-	3,3,3	0.50	0	2,2,2	0.21	0
4	EDO	Q	818	-	3,3,3	0.45	0	2,2,2	0.38	0
4	EDO	P	803	-	3,3,3	0.35	0	2,2,2	0.20	0
6	PEG	Q	824	-	6,6,6	0.51	0	5,5,5	0.41	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	EDO	Q	815	-	-	1/1/1/1	-
4	EDO	S	809	-	-	0/1/1/1	-
4	EDO	P	805	-	-	1/1/1/1	-
4	EDO	S	805	-	-	1/1/1/1	-
6	PEG	P	816	-	-	4/4/4/4	-
4	EDO	Q	805	-	-	1/1/1/1	-
6	PEG	Q	823	-	-	3/4/4/4	-
4	EDO	Q	804	-	-	0/1/1/1	-
6	PEG	S	816	-	-	2/4/4/4	-
4	EDO	S	808	-	-	1/1/1/1	-
4	EDO	P	806	-	-	0/1/1/1	-
4	EDO	P	809	-	-	0/1/1/1	-
4	EDO	S	804	-	-	1/1/1/1	-
3	GOL	Q	821	-	-	0/4/4/4	-
2	HDD	P	801	1,3	-	0/3/89/89	0/1/9/9
4	EDO	S	807	-	-	1/1/1/1	-
4	EDO	S	810	-	-	0/1/1/1	-
2	HDD	Q	801	1,3	-	0/3/89/89	0/1/9/9
4	EDO	P	804	-	-	1/1/1/1	-
5	PGE	R	808	-	-	5/7/7/7	-
4	EDO	Q	811	-	-	1/1/1/1	-
3	GOL	P	813	-	-	1/4/4/4	-
3	GOL	R	809	-	-	2/4/4/4	-
4	EDO	Q	818	-	-	1/1/1/1	-
4	EDO	Q	814	-	-	1/1/1/1	-
4	EDO	R	803	-	-	0/1/1/1	-
3	GOL	R	802	2	-	0/4/4/4	-
4	EDO	Q	810	-	-	0/1/1/1	-
3	GOL	S	802	2	-	2/4/4/4	-
4	EDO	P	810	-	-	1/1/1/1	-
3	GOL	S	812	-	-	4/4/4/4	-
4	EDO	Q	806	-	-	1/1/1/1	-
3	GOL	R	810	-	-	2/4/4/4	-
4	EDO	Q	816	-	-	0/1/1/1	-
4	EDO	P	808	-	-	1/1/1/1	-
9	PG4	R	812	-	-	5/10/10/10	-
4	EDO	Q	817	-	-	1/1/1/1	-
4	EDO	Q	812	-	-	1/1/1/1	-
4	EDO	Q	813	-	-	1/1/1/1	-
5	PGE	P	812	-	-	6/7/7/7	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	P	802	2	-	2/4/4/4	-
5	PGE	R	807	-	-	6/7/7/7	-
6	PEG	Q	825	-	-	3/4/4/4	-
3	GOL	Q	802	2	-	2/4/4/4	-
4	EDO	S	806	-	-	1/1/1/1	-
4	EDO	S	803	-	-	0/1/1/1	-
3	GOL	Q	820	-	-	3/4/4/4	-
4	EDO	R	804	-	-	1/1/1/1	-
4	EDO	S	811	-	-	0/1/1/1	-
2	HDD	S	801	1,3	-	0/3/89/89	0/1/9/9
6	PEG	P	817	-	-	4/4/4/4	-
4	EDO	R	805	-	-	1/1/1/1	-
4	EDO	Q	809	-	-	0/1/1/1	-
4	EDO	Q	808	-	-	1/1/1/1	-
3	GOL	S	813	-	-	2/4/4/4	-
2	HDD	R	801	1,3	-	0/3/89/89	0/1/9/9
4	EDO	R	806	-	-	1/1/1/1	-
4	EDO	P	811	-	-	0/1/1/1	-
6	PEG	S	817	-	-	2/4/4/4	-
6	PEG	P	815	-	-	2/4/4/4	-
4	EDO	Q	807	-	-	0/1/1/1	-
6	PEG	Q	822	-	-	3/4/4/4	-
4	EDO	Q	819	-	-	1/1/1/1	-
4	EDO	P	807	-	-	1/1/1/1	-
4	EDO	Q	803	-	-	1/1/1/1	-
6	PEG	P	814	-	-	3/4/4/4	-
4	EDO	P	803	-	-	1/1/1/1	-
6	PEG	Q	824	-	-	3/4/4/4	-

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	S	801	HDD	O1D-CGD	5.86	1.45	1.35
2	Q	801	HDD	O1D-CGD	5.10	1.44	1.35
2	R	801	HDD	O1D-CGD	4.72	1.43	1.35
2	P	801	HDD	O1D-CGD	4.36	1.42	1.35
2	R	801	HDD	C3B-C2B	3.97	1.45	1.40

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Q	801	HDD	O1D-CGD-O2D	9.17	128.99	120.80
2	R	801	HDD	O1D-CGD-O2D	7.72	127.69	120.80
2	S	801	HDD	O1D-CGD-O2D	7.67	127.65	120.80
2	P	801	HDD	O1D-CGD-O2D	7.40	127.41	120.80
2	P	801	HDD	C4A-C3A-C2A	-6.81	102.26	107.00

There are no chirality outliers.

5 of 97 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	R	809	GOL	C1-C2-C3-O3
3	S	802	GOL	C1-C2-C3-O3
3	S	812	GOL	O1-C1-C2-C3
3	S	812	GOL	C1-C2-C3-O3
3	S	812	GOL	O2-C2-C3-O3

There are no ring outliers.

27 monomers are involved in 104 short contacts:

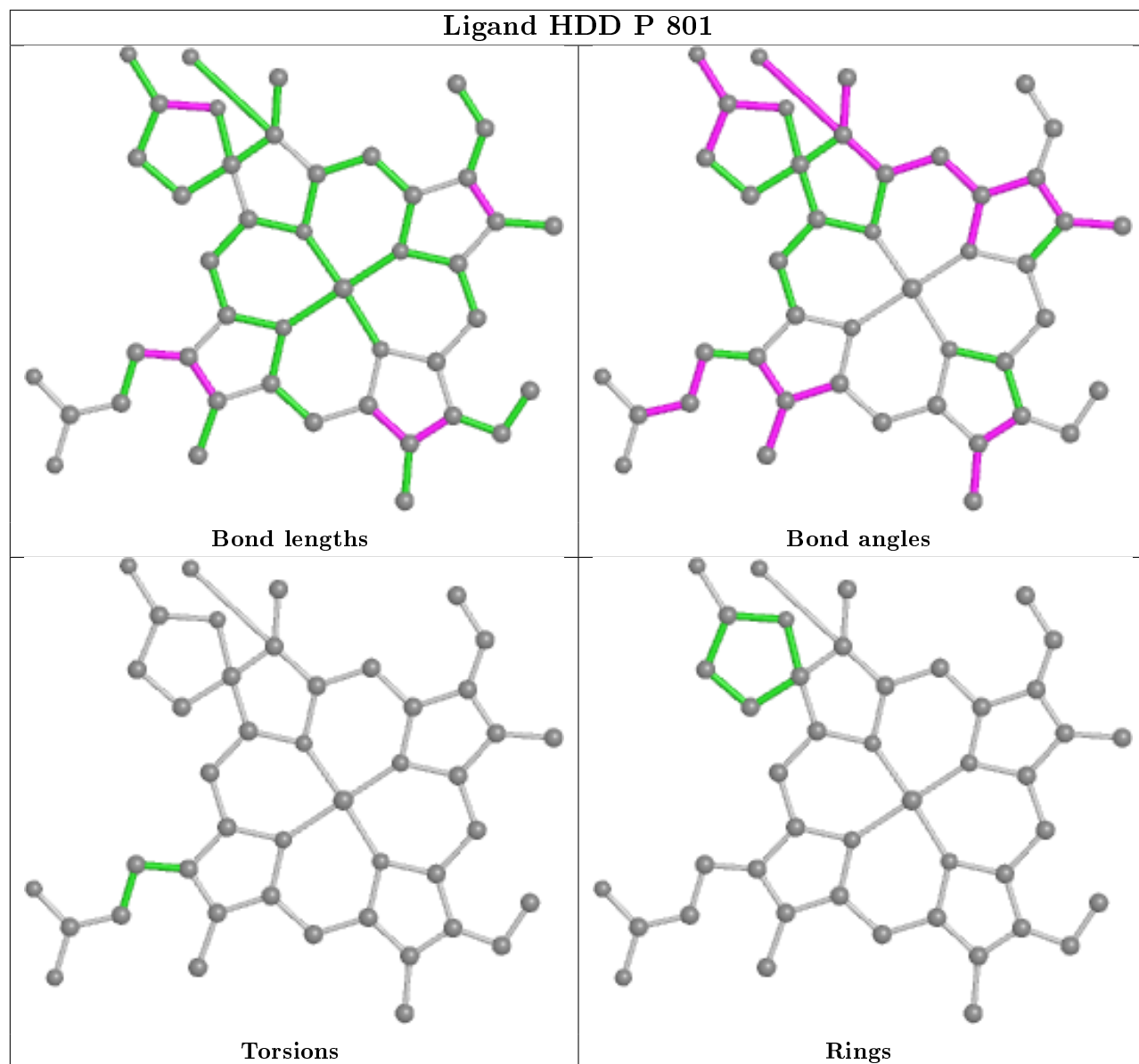
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	S	805	EDO	2	0
4	Q	805	EDO	0	2
6	S	816	PEG	1	0
4	S	804	EDO	2	0
2	P	801	HDD	8	0
2	Q	801	HDD	7	0
4	P	804	EDO	2	0
5	R	808	PGE	1	0
3	R	809	GOL	3	0
4	R	803	EDO	1	0
3	R	802	GOL	10	0
3	S	802	GOL	4	0
4	P	808	EDO	1	0
3	P	802	GOL	6	0
5	R	807	PGE	7	0
3	Q	802	GOL	1	0
2	S	801	HDD	9	0
6	P	817	PEG	13	0
6	P	814	PEG	7	0
2	R	801	HDD	7	0
6	S	817	PEG	2	0
6	P	815	PEG	1	0

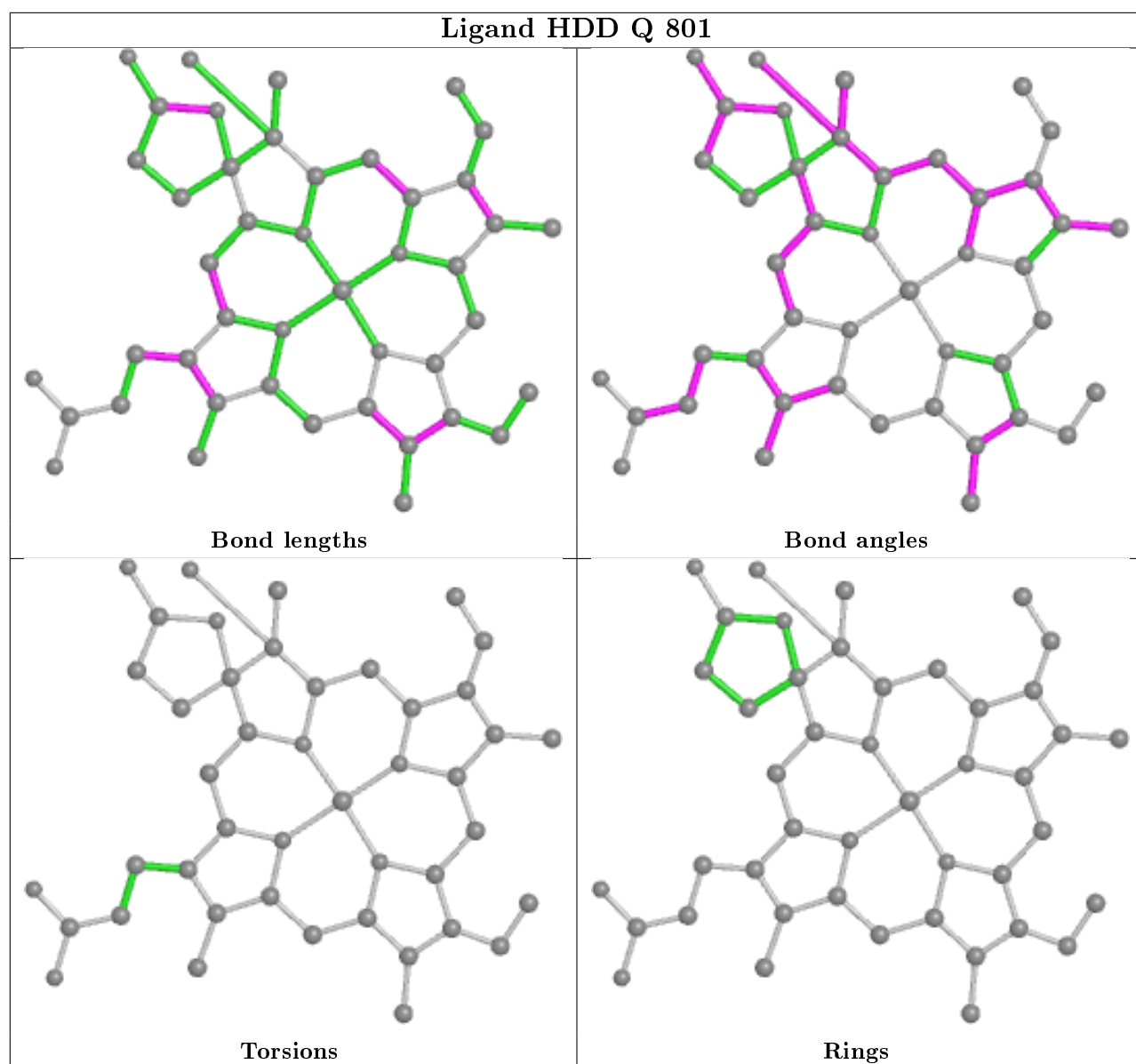
*Continued on next page...*

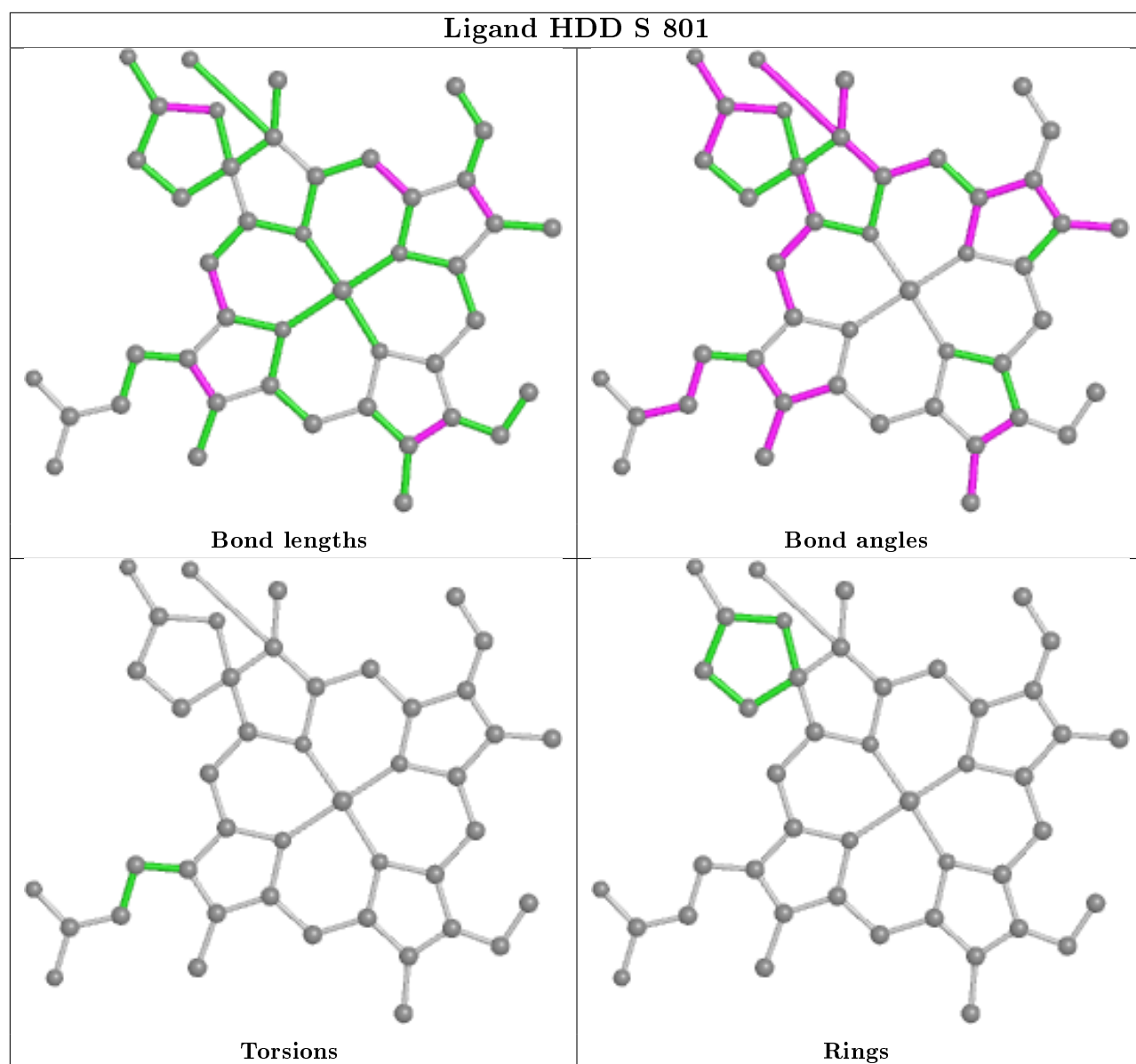
*Continued from previous page...*

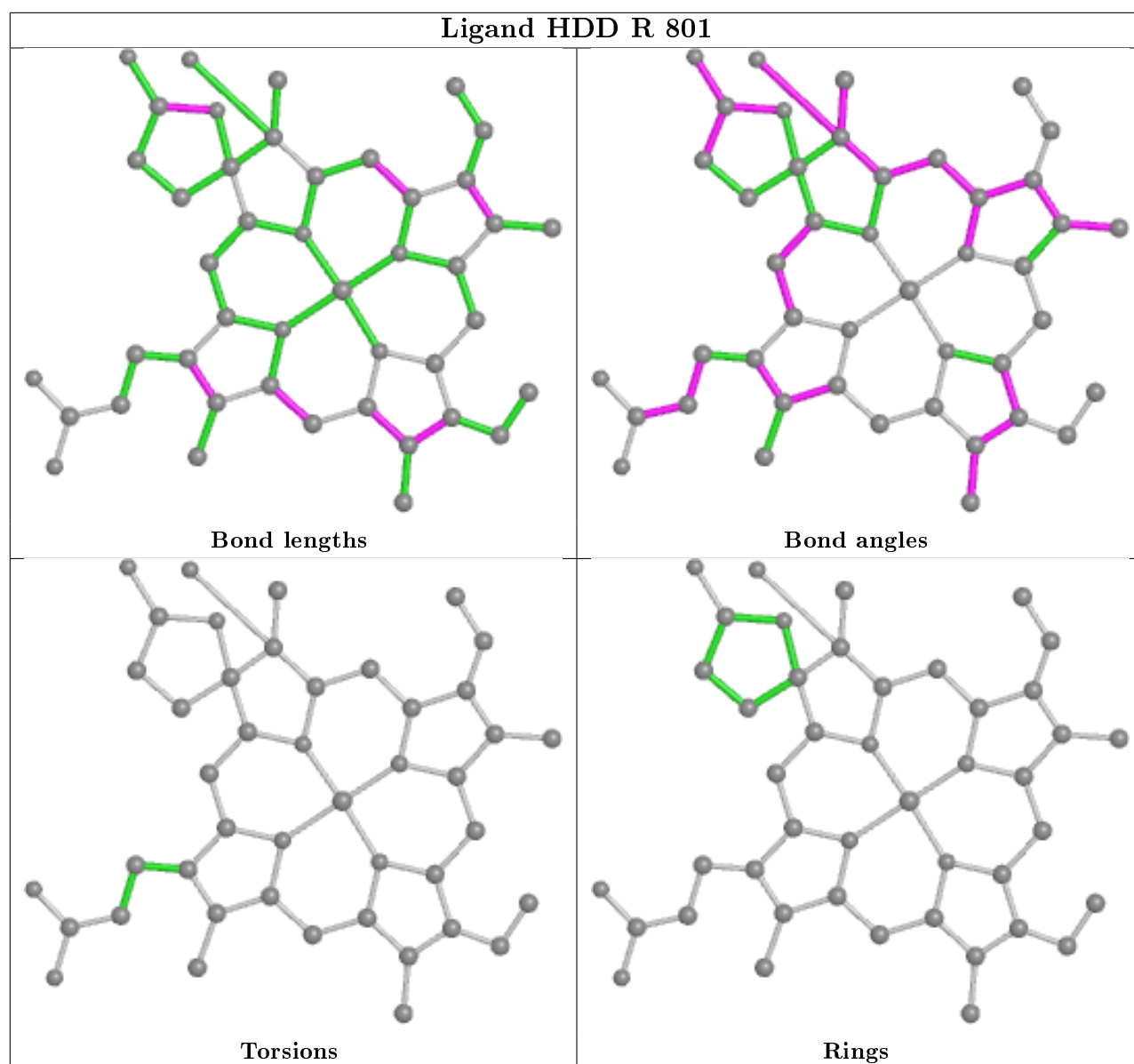
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	Q	807	EDO	2	0
6	Q	822	PEG	3	0
4	Q	819	EDO	2	0
4	P	807	EDO	3	0
4	Q	803	EDO	1	0

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 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	P	747/753 (99%)	-0.83	8 (1%) 80 83	8, 13, 30, 101	0
1	Q	746/753 (99%)	-0.80	7 (0%) 84 86	7, 13, 29, 73	0
1	R	747/753 (99%)	-0.72	16 (2%) 63 69	7, 14, 34, 95	0
1	S	747/753 (99%)	-0.82	4 (0%) 91 92	7, 13, 28, 86	0
All	All	2987/3012 (99%)	-0.79	35 (1%) 79 83	7, 13, 30, 101	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	R	7	LYS	6.3
1	R	9	PRO	6.1
1	P	7	LYS	5.9
1	P	8	ASN	5.5
1	R	8	ASN	5.0

### 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	EDO	Q	819	4/4	0.70	0.23	42,52,54,56	1
6	PEG	Q	824	7/7	0.73	0.34	38,52,73,74	1
3	GOL	S	812	6/6	0.75	0.22	32,58,64,67	2
4	EDO	Q	817	4/4	0.75	0.28	69,80,91,91	1
5	PGE	P	812	10/10	0.76	0.37	50,74,84,84	1
9	PG4	R	812	13/13	0.80	0.40	50,71,80,82	1
4	EDO	R	806	4/4	0.82	0.67	58,76,90,95	1
4	EDO	S	811	4/4	0.83	0.28	51,73,76,78	1
4	EDO	Q	806	4/4	0.83	0.27	49,69,73,79	4
6	PEG	P	817	7/7	0.84	0.20	17,21,27,32	17
4	EDO	S	806	4/4	0.84	0.21	42,56,62,68	1
3	GOL	Q	820	6/6	0.84	0.34	48,62,69,71	2
4	EDO	Q	815	4/4	0.84	0.26	45,70,82,83	1
3	GOL	Q	821	6/6	0.85	0.47	59,77,85,106	2
3	GOL	R	809	6/6	0.85	0.20	31,40,49,60	2
4	EDO	Q	818	4/4	0.86	0.17	52,55,67,68	1
4	EDO	Q	816	4/4	0.86	0.18	34,41,54,57	1
4	EDO	R	803	4/4	0.86	0.28	42,53,63,72	1
5	PGE	R	808	10/10	0.86	0.45	36,62,103,104	1
6	PEG	S	817	7/7	0.87	0.23	25,47,63,64	1
6	PEG	Q	825	7/7	0.87	0.27	41,50,62,66	1
4	EDO	S	808	4/4	0.87	0.17	30,39,52,52	1
4	EDO	P	808	4/4	0.89	0.27	43,56,62,72	1
6	PEG	P	814	7/7	0.89	0.20	29,38,48,52	1
4	EDO	Q	812	4/4	0.89	0.17	66,72,74,76	1
6	PEG	P	815	7/7	0.90	0.20	35,42,51,52	1
4	EDO	P	810	4/4	0.90	0.33	48,63,66,67	1
4	EDO	Q	811	4/4	0.91	0.13	38,43,50,51	1
4	EDO	R	805	4/4	0.92	0.12	23,26,38,38	1
4	EDO	S	810	4/4	0.92	0.17	32,42,52,52	1
4	EDO	P	806	4/4	0.92	0.14	36,50,54,58	1
4	EDO	Q	807	4/4	0.92	0.15	29,35,40,43	1
4	EDO	S	807	4/4	0.92	0.14	46,53,56,58	1
3	GOL	S	802	6/6	0.92	0.17	15,18,30,34	14
4	EDO	Q	809	4/4	0.92	0.26	34,40,43,44	1
8	MG	S	815	1/1	0.93	0.27	49,49,49,49	0
4	EDO	P	804	4/4	0.93	0.28	6,32,40,43	1
3	GOL	Q	802	6/6	0.93	0.18	16,22,35,38	0
3	GOL	R	810	6/6	0.93	0.14	22,27,39,55	2
6	PEG	S	816	7/7	0.93	0.13	21,27,38,39	1
6	PEG	Q	822	7/7	0.93	0.12	19,30,39,40	1
3	GOL	R	802	6/6	0.93	0.20	16,20,37,45	14
3	GOL	P	802	6/6	0.93	0.26	17,24,34,45	0

Continued on next page...

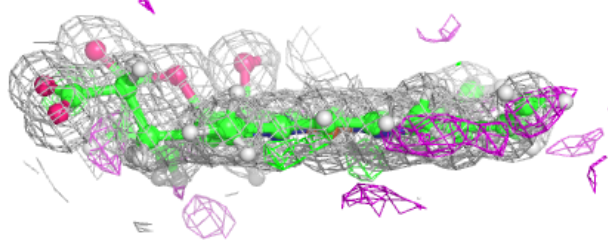
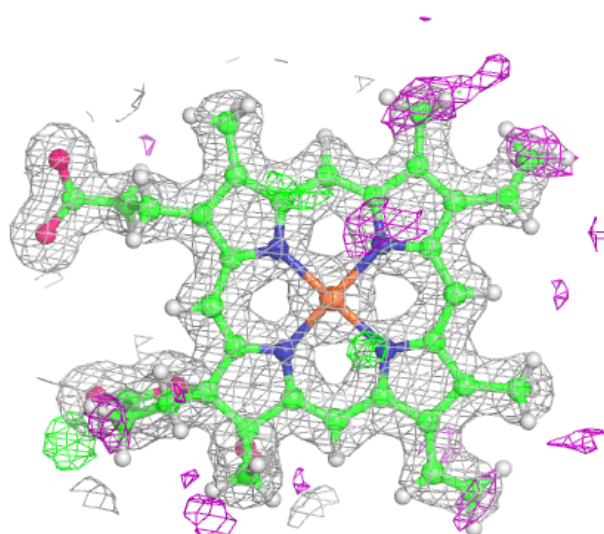
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	PGE	R	807	10/10	0.93	0.18	12,14,15,15	24
6	PEG	P	816	7/7	0.94	0.28	47,59,69,69	1
4	EDO	R	804	4/4	0.94	0.17	32,49,53,57	1
3	GOL	S	813	6/6	0.94	0.14	19,25,34,45	2
4	EDO	S	803	4/4	0.94	0.16	33,40,46,48	1
6	PEG	Q	823	7/7	0.95	0.11	18,33,55,56	1
4	EDO	Q	808	4/4	0.95	0.11	22,26,29,32	1
3	GOL	P	813	6/6	0.96	0.13	16,24,43,50	2
4	EDO	Q	810	4/4	0.96	0.14	30,35,42,44	1
4	EDO	P	809	4/4	0.96	0.12	21,24,34,35	1
4	EDO	Q	803	4/4	0.96	0.08	30,38,43,44	1
4	EDO	S	804	4/4	0.96	0.15	23,26,29,30	1
4	EDO	S	809	4/4	0.96	0.23	38,48,54,68	1
7	SO4	P	818	5/5	0.97	0.11	35,37,49,51	5
4	EDO	Q	814	4/4	0.97	0.12	20,30,42,54	1
4	EDO	Q	804	4/4	0.97	0.09	23,26,31,32	1
4	EDO	Q	813	4/4	0.97	0.25	35,38,51,54	1
7	SO4	S	814	5/5	0.98	0.25	41,50,64,68	0
4	EDO	P	807	4/4	0.98	0.17	32,35,45,46	1
8	MG	R	811	1/1	0.98	0.17	41,41,41,41	0
4	EDO	S	805	4/4	0.98	0.11	25,32,37,38	1
4	EDO	P	803	4/4	0.98	0.08	23,28,29,29	1
4	EDO	P	805	4/4	0.98	0.11	28,40,50,58	1
2	HDD	P	801	44/44	0.99	0.05	7,9,15,20	1
4	EDO	P	811	4/4	0.99	0.06	25,30,38,40	0
2	HDD	Q	801	44/44	0.99	0.07	7,9,14,18	1
2	HDD	S	801	44/44	0.99	0.08	7,9,15,17	1
4	EDO	Q	805	4/4	0.99	0.08	23,33,69,69	1
2	HDD	R	801	44/44	0.99	0.07	8,10,15,19	1

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.

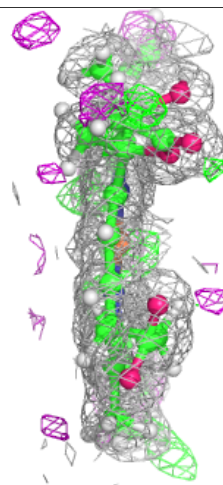
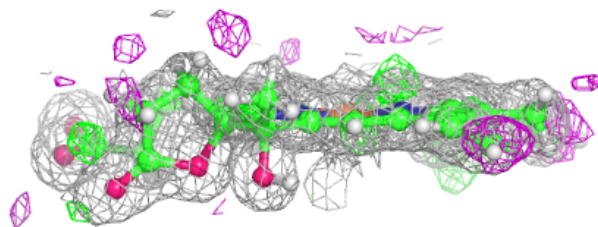
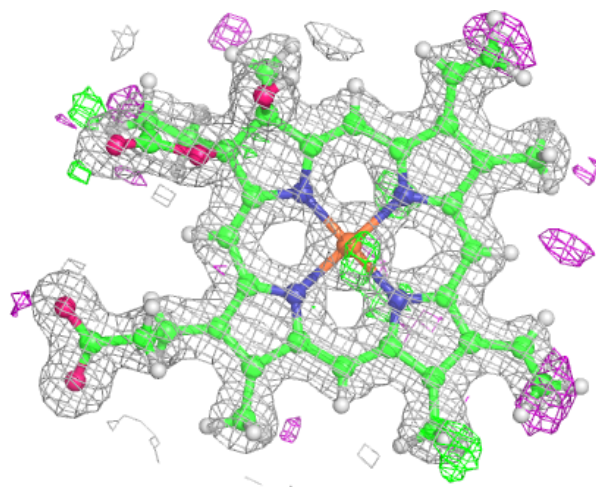
**Electron density around HDD P 801:**

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



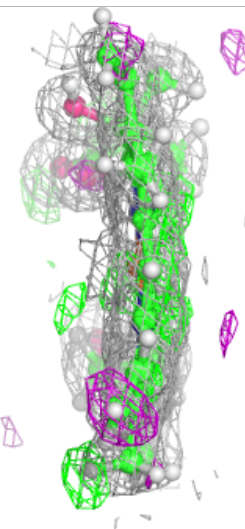
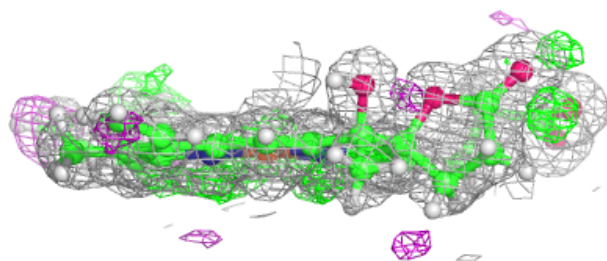
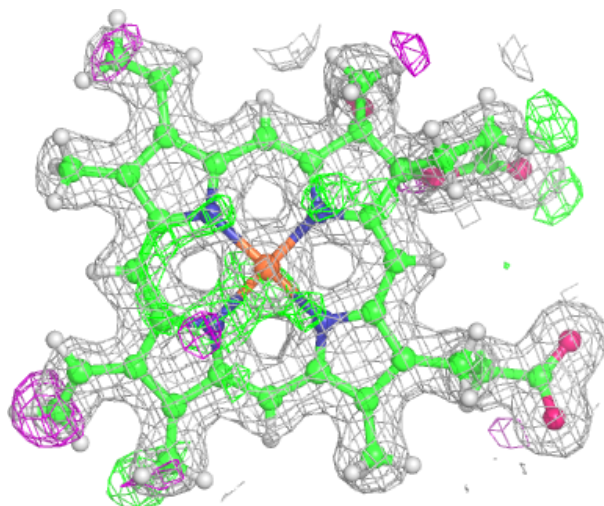
**Electron density around HDD Q 801:**

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



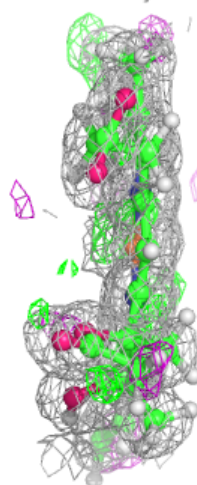
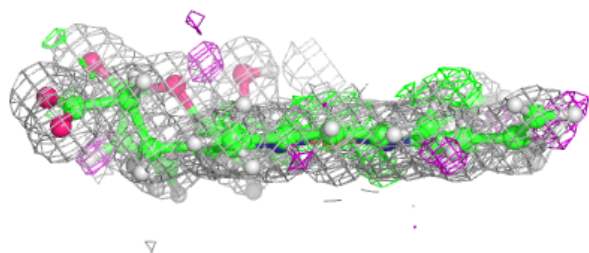
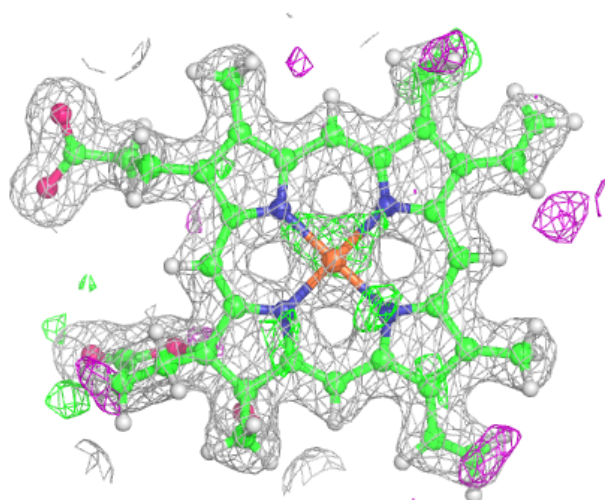
**Electron density around HDD S 801:**

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



**Electron density around HDD R 801:**

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



## 6.5 Other polymers ⓘ

There are no such residues in this entry.