



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 09:34 am BST

PDB ID : 5XVZ  
Title : CATPO mutant - H246W  
Authors : Yuzugullu Karakus, Y.; Goc, G.; Balci, S.; Pearson, A.R.; Yorke, B.  
Deposited on : 2017-06-28  
Resolution : 1.90 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](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 : 2.11  
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) : 2.11

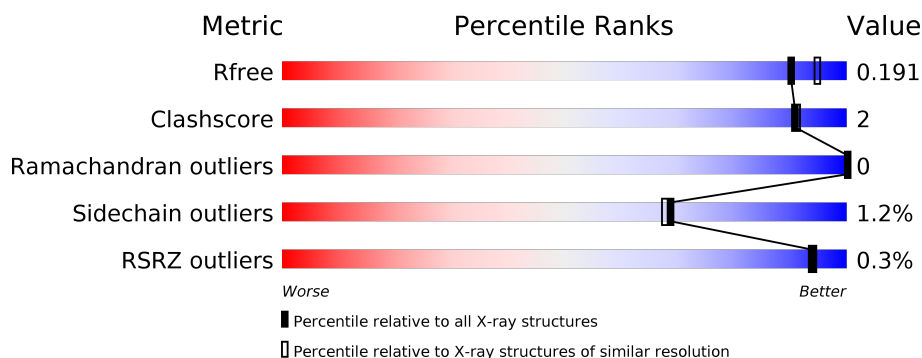
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	A	678	<div> <div></div> <div>94%5%</div> </div>
1	B	678	<div> <div>%</div> <div>94%5%</div> </div>
1	C	678	<div> <div></div> <div>92%8%</div> </div>
1	D	678	<div> <div></div> <div>92%7%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 23523 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.

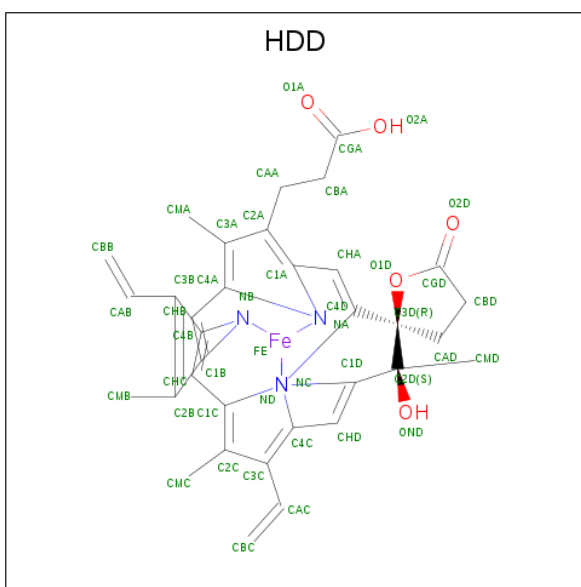
- Molecule 1 is a protein called Catalase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	675	Total	C	N	O	S	0	13	0
			5377	3397	938	1029	13			
1	B	675	Total	C	N	O	S	0	9	0
			5351	3382	936	1020	13			
1	C	675	Total	C	N	O	S	0	15	0
			5391	3408	939	1031	13			
1	D	675	Total	C	N	O	S	0	7	0
			5333	3371	934	1017	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	246	TRP	HIS	engineered mutation	UNP M4GGR7
B	246	TRP	HIS	engineered mutation	UNP M4GGR7
C	246	TRP	HIS	engineered mutation	UNP M4GGR7
D	246	TRP	HIS	engineered mutation	UNP M4GGR7

- Molecule 2 is CIS-HEME D HYDROXYCHLORIN GAMMA-SPIROLACTONE (three-letter code: HDD) (formula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 44	C 34	Fe 1	N 4	O 5	0	0
2	B	1	Total 44	C 34	Fe 1	N 4	O 5	0	0
2	C	1	Total 44	C 34	Fe 1	N 4	O 5	0	0
2	D	1	Total 44	C 34	Fe 1	N 4	O 5	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	3	Total 3 Ca 3	0	0
3	A	3	Total 3 Ca 3	0	0
3	D	2	Total 2 Ca 2	0	0
3	C	2	Total 2 Ca 2	0	0

- Molecule 4 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula:  $C_{10}H_{22}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			16	10	6		
4	B	1	Total	C	O	0	0
			16	10	6		
4	D	1	Total	C	O	0	0
			16	10	6		

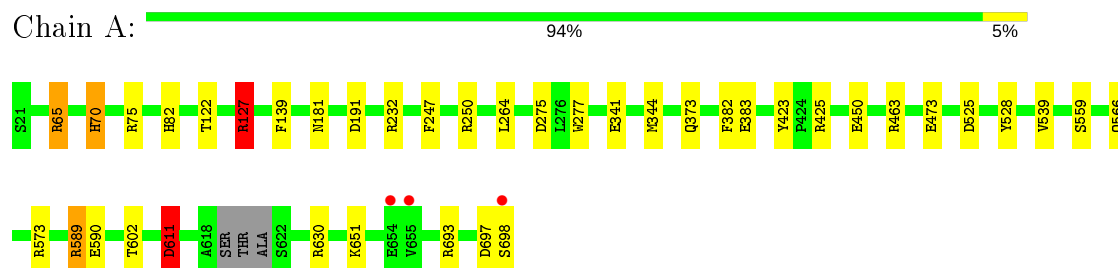
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	457	Total	O	0	0
			457	457		
5	B	458	Total	O	0	0
			458	458		
5	C	481	Total	O	0	0
			481	481		
5	D	441	Total	O	0	0
			441	441		

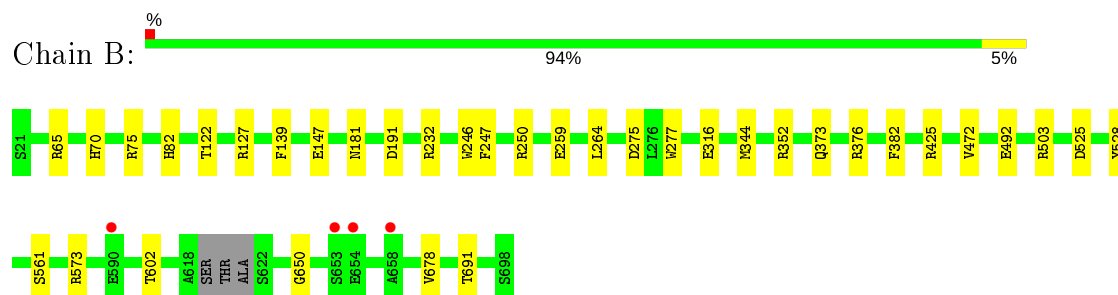
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

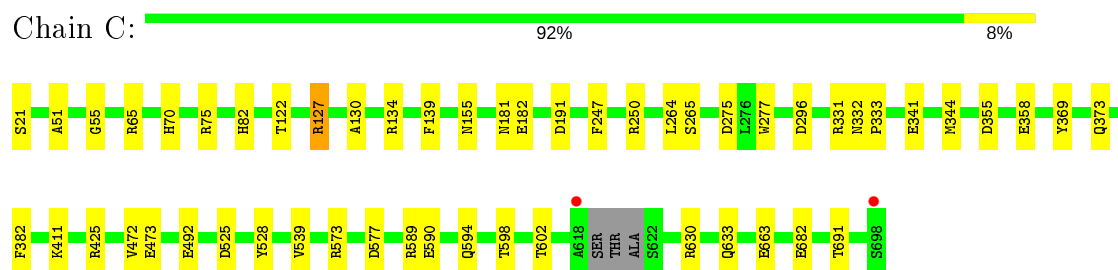
#### • Molecule 1: Catalase



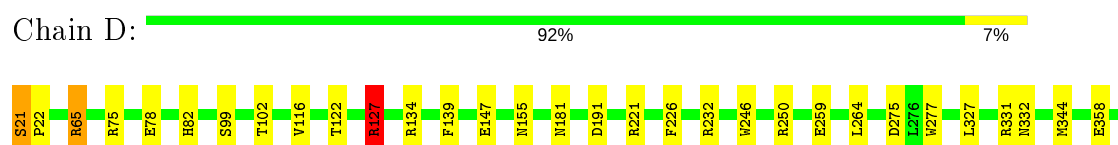
#### • Molecule 1: Catalase



#### • Molecule 1: Catalase



#### • Molecule 1: Catalase





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.34 Å   120.78 Å   185.20 Å 90.00°   102.06°   90.00°	Depositor
Resolution (Å)	100.69 – 1.90 100.49 – 1.90	Depositor EDS
% Data completeness (in resolution range)	94.5 (100.69-1.90) 94.5 (100.49-1.90)	Depositor EDS
$R_{merge}$	0.69	Depositor
$R_{sym}$	0.68	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 1.90 Å)	Xtriage
Refinement program	REFMAC 5.8.0222	Depositor
R, $R_{free}$	0.158   ,   0.187 0.165   ,   0.191	Depositor DCC
$R_{free}$ test set	10019 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.4	Xtriage
Anisotropy	0.180	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 46.2	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	0.97	EDS
Total number of atoms	23523	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 48.43 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.6115e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CA, HDD, 1PE

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	A	0.74	4/5516 (0.1%)	0.82	7/7502 (0.1%)
1	B	0.73	4/5490 (0.1%)	0.82	7/7464 (0.1%)
1	C	0.76	4/5536 (0.1%)	0.83	8/7529 (0.1%)
1	D	0.76	6/5472 (0.1%)	0.82	8/7441 (0.1%)
All	All	0.75	18/22014 (0.1%)	0.82	30/29936 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	7
1	B	0	4
1	C	0	4
1	D	0	5
All	All	0	20

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	147	GLU	CD-OE2	8.51	1.35	1.25
1	B	259	GLU	CD-OE2	8.19	1.34	1.25
1	C	358	GLU	CD-OE1	7.12	1.33	1.25
1	C	127	ARG	C-O	6.69	1.36	1.23
1	D	358	GLU	CD-OE1	6.39	1.32	1.25
1	C	663	GLU	CD-OE1	5.96	1.32	1.25
1	D	127[A]	ARG	C-O	5.95	1.34	1.23
1	D	127[B]	ARG	C-O	5.95	1.34	1.23
1	C	341	GLU	CD-OE2	5.81	1.32	1.25
1	A	383	GLU	CD-OE2	5.74	1.31	1.25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	341	GLU	CD-OE2	5.69	1.31	1.25
1	D	259	GLU	CD-OE1	5.51	1.31	1.25
1	A	127	ARG	C-O	5.48	1.33	1.23
1	B	259	GLU	CD-OE1	5.35	1.31	1.25
1	A	450	GLU	CD-OE1	5.33	1.31	1.25
1	B	147	GLU	CD-OE2	5.27	1.31	1.25
1	D	78	GLU	CD-OE1	5.19	1.31	1.25
1	B	492	GLU	CD-OE2	5.16	1.31	1.25

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	134	ARG	NE-CZ-NH2	-8.36	116.12	120.30
1	B	65	ARG	NE-CZ-NH2	-7.72	116.44	120.30
1	C	425	ARG	NE-CZ-NH1	7.54	124.07	120.30
1	C	134	ARG	NE-CZ-NH2	-7.06	116.77	120.30
1	D	449	ARG	NE-CZ-NH2	-6.91	116.84	120.30
1	D	221	ARG	NE-CZ-NH2	-6.68	116.96	120.30
1	A	232	ARG	NE-CZ-NH1	6.43	123.51	120.30
1	A	611	ASP	CB-CA-C	-6.30	97.79	110.40
1	C	127	ARG	NE-CZ-NH1	-6.30	117.15	120.30
1	B	425	ARG	NE-CZ-NH2	-6.14	117.23	120.30
1	D	232	ARG	NE-CZ-NH2	-6.12	117.24	120.30
1	A	463	ARG	NE-CZ-NH2	-6.12	117.24	120.30
1	C	630	ARG	NE-CZ-NH2	-5.87	117.36	120.30
1	A	65	ARG	NE-CZ-NH2	-5.62	117.49	120.30
1	D	65	ARG	NE-CZ-NH2	-5.61	117.49	120.30
1	B	232	ARG	NE-CZ-NH1	5.57	123.08	120.30
1	C	573	ARG	NE-CZ-NH2	-5.53	117.54	120.30
1	B	503	ARG	NE-CZ-NH2	-5.50	117.55	120.30
1	A	232	ARG	NE-CZ-NH2	-5.39	117.61	120.30
1	D	376	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	B	352	ARG	NE-CZ-NH2	-5.34	117.63	120.30
1	A	630	ARG	NE-CZ-NH2	-5.33	117.63	120.30
1	C	630	ARG	NE-CZ-NH1	5.29	122.95	120.30
1	C	425	ARG	CG-CD-NE	5.25	122.83	111.80
1	D	573	ARG	NE-CZ-NH2	-5.25	117.68	120.30
1	B	376	ARG	NE-CZ-NH2	-5.19	117.70	120.30
1	B	232	ARG	NE-CZ-NH2	-5.18	117.71	120.30
1	A	559	SER	CB-CA-C	-5.15	100.31	110.10
1	C	65	ARG	NE-CZ-NH2	-5.15	117.72	120.30
1	D	399	ARG	NE-CZ-NH2	-5.01	117.79	120.30

There are no chirality outliers.

All (20) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	127	ARG	Sidechain
1	A	250	ARG	Sidechain
1	A	573	ARG	Sidechain
1	A	589	ARG	Sidechain
1	A	65	ARG	Sidechain
1	A	693	ARG	Sidechain
1	A	75	ARG	Sidechain
1	B	127[A]	ARG	Sidechain
1	B	250	ARG	Sidechain
1	B	650	GLY	Peptide
1	B	75	ARG	Sidechain
1	C	127	ARG	Sidechain
1	C	250	ARG	Sidechain
1	C	589[A]	ARG	Sidechain
1	C	75	ARG	Sidechain
1	D	127[A]	ARG	Sidechain
1	D	250	ARG	Sidechain
1	D	391	ARG	Sidechain
1	D	65	ARG	Sidechain
1	D	75	ARG	Sidechain

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5377	0	5133	16	0
1	B	5351	0	5118	13	0
1	C	5391	0	5153	21	0
1	D	5333	0	5099	24	0
2	A	44	0	31	3	0
2	B	44	0	31	2	0
2	C	44	0	31	1	0
2	D	44	0	31	3	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	A	16	0	22	0	0
4	B	16	0	22	0	0
4	D	16	0	22	0	0
5	A	457	0	0	0	0
5	B	458	0	0	0	1
5	C	481	0	0	2	1
5	D	441	0	0	1	0
All	All	23523	0	20693	72	1

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 2.

All (72) 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:A:127:ARG:NH2	1:C:182:GLU:OE2	2.23	0.71
1:B:277[A]:TRP:CE3	1:D:181:ASN:HB3	2.32	0.64
1:A:70:HIS:HD2	5:D:961:HOH:O	1.82	0.63
1:B:277[A]:TRP:CZ3	1:D:181:ASN:HB3	2.34	0.63
2:D:701:HDD:HBC1	2:D:701:HDD:HMC1	1.80	0.62
1:A:277[A]:TRP:CE3	1:C:181:ASN:HB3	2.37	0.60
1:A:423:TYR:OH	1:A:425[B]:ARG:NH1	2.37	0.57
1:D:264:LEU:HD13	1:D:602:THR:HB	1.87	0.57
2:A:701:HDD:HBB1	2:A:701:HDD:HMB1	1.87	0.56
1:A:566:GLN:HE22	1:A:611:ASP:HB2	1.70	0.56
1:B:264:LEU:HD13	1:B:602:THR:HB	1.87	0.56
2:A:701:HDD:HMC1	2:A:701:HDD:HBC1	1.87	0.56
1:C:264:LEU:HD13	1:C:602:THR:HB	1.88	0.55
1:A:264:LEU:HD13	1:A:602:THR:HB	1.87	0.55
1:A:277[A]:TRP:CZ3	1:C:181:ASN:HB3	2.40	0.55
1:B:181:ASN:HB3	1:D:277[B]:TRP:CE3	2.42	0.54
1:D:21:SER:HB3	1:D:331:ARG:NH1	2.23	0.54
1:C:155:ASN:CG	2:C:701:HDD:HMB2	2.28	0.53
1:C:277[B]:TRP:CZ3	1:C:333:PRO:HD2	2.43	0.53
1:C:633:GLN:NE2	5:C:807:HOH:O	2.41	0.52
1:B:181:ASN:HB3	1:D:277[B]:TRP:CZ3	2.46	0.51
1:C:277[B]:TRP:CZ3	1:C:332:ASN:HB3	2.47	0.50
2:B:701:HDD:HMC1	2:B:701:HDD:HBC1	1.95	0.49
1:C:130:ALA:CB	1:C:265[B]:SER:HB2	2.43	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:226:PHE:HE1	1:D:246[B]:TRP:CE2	2.31	0.48
1:D:582:THR:HG21	1:D:594:GLN:HE21	1.76	0.48
1:A:181:ASN:HB3	1:C:277[B]:TRP:CE3	2.49	0.48
2:D:701:HDD:HBB1	2:D:701:HDD:HMB1	1.96	0.47
1:C:594[B]:GLN:HG3	1:C:598:THR:OG1	2.14	0.47
1:A:697:ASP:O	1:A:698:SER:C	2.51	0.47
1:D:21:SER:N	1:D:22:PRO:HD2	2.30	0.46
1:D:577:ASP:OD2	1:D:682:GLU:OE2	2.34	0.46
1:A:589:ARG:NH1	1:A:590:GLU:O	2.49	0.45
1:A:525:ASP:HA	1:A:528:TYR:CD2	2.52	0.45
1:C:344:MET:SD	1:C:382:PHE:HB2	2.57	0.45
1:A:425[A]:ARG:HD2	1:A:425[A]:ARG:HH21	1.69	0.44
1:B:525:ASP:HA	1:B:528:TYR:CD2	2.53	0.44
1:D:226:PHE:CE1	1:D:246[B]:TRP:CE2	3.04	0.44
1:D:82:HIS:HA	1:D:122:THR:O	2.17	0.44
1:D:344:MET:SD	1:D:382:PHE:HB2	2.58	0.44
1:C:577:ASP:OD2	1:C:682:GLU:OE2	2.36	0.44
1:B:277[A]:TRP:CE3	1:D:181:ASN:CB	3.00	0.44
1:C:473:GLU:HG3	1:C:539:VAL:CG2	2.48	0.44
1:C:82:HIS:HA	1:C:122:THR:O	2.18	0.44
1:D:116:VAL:HG21	1:D:327:LEU:HD11	2.00	0.44
1:B:573:ARG:HG3	1:B:678:VAL:HG11	2.00	0.44
2:B:701:HDD:HMB1	2:B:701:HDD:HBB1	1.99	0.43
1:D:155:ASN:CG	2:D:701:HDD:HMB2	2.39	0.43
1:A:82:HIS:HA	1:A:122:THR:O	2.18	0.43
1:C:472:VAL:HG11	1:C:691:THR:HB	1.99	0.43
1:C:525:ASP:HA	1:C:528:TYR:CD2	2.54	0.43
1:C:369:TYR:O	1:C:373:GLN:HG2	2.19	0.43
2:A:701:HDD:HBC1	2:A:701:HDD:CMC	2.49	0.43
1:D:594:GLN:HG3	1:D:598:THR:OG1	2.19	0.42
1:D:525:ASP:HA	1:D:528:TYR:CD2	2.55	0.42
1:B:82:HIS:HA	1:B:122:THR:O	2.20	0.42
1:D:99[B]:SER:HA	1:D:102:THR:O	2.20	0.42
1:A:473:GLU:HG3	1:A:539:VAL:CG2	2.50	0.42
1:A:344:MET:SD	1:A:382:PHE:HB2	2.60	0.42
1:D:277[B]:TRP:CZ3	1:D:332:ASN:HB3	2.55	0.42
1:C:355:ASP:HB2	5:C:822:HOH:O	2.19	0.42
1:D:473:GLU:HG3	1:D:539:VAL:CG2	2.50	0.41
1:C:21[B]:SER:HB2	1:C:331:ARG:NH1	2.35	0.41
1:D:277[B]:TRP:CH2	1:D:332:ASN:HB3	2.56	0.41
1:A:373:GLN:HA	1:A:373:GLN:OE1	2.20	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:246[B]:TRP:CH2	1:B:316:GLU:OE2	2.73	0.41
1:D:594:GLN:HG3	1:D:598:THR:HG1	1.86	0.41
1:B:344:MET:SD	1:B:382:PHE:HB2	2.61	0.41
1:B:373:GLN:OE1	1:B:373:GLN:HA	2.21	0.40
1:D:21:SER:N	1:D:22:PRO:CD	2.84	0.40
1:B:472:VAL:HG11	1:B:691:THR:HB	2.03	0.40
1:C:51:ALA:O	1:C:55:GLY:HA3	2.22	0.40

All (1) 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)
5:B:1173:HOH:O	5:C:1216:HOH:O[4_454]	1.20	1.00

## 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	A	684/678 (101%)	666 (97%)	18 (3%)	0	100	100
1	B	680/678 (100%)	663 (98%)	17 (2%)	0	100	100
1	C	685/678 (101%)	667 (97%)	18 (3%)	0	100	100
1	D	678/678 (100%)	663 (98%)	15 (2%)	0	100	100
All	All	2727/2712 (101%)	2659 (98%)	68 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	A	573/562 (102%)	566 (99%)	7 (1%)	71	70
1	B	568/562 (101%)	562 (99%)	6 (1%)	73	73
1	C	575/562 (102%)	565 (98%)	10 (2%)	60	57
1	D	566/562 (101%)	559 (99%)	7 (1%)	71	70
All	All	2282/2248 (102%)	2252 (99%)	30 (1%)	69	68

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	70	HIS
1	A	139	PHE
1	A	191	ASP
1	A	247	PHE
1	A	275	ASP
1	A	611	ASP
1	A	651	LYS
1	B	70	HIS
1	B	139	PHE
1	B	191	ASP
1	B	247	PHE
1	B	275	ASP
1	B	561	SER
1	C	70	HIS
1	C	139	PHE
1	C	191	ASP
1	C	247	PHE
1	C	275	ASP
1	C	296	ASP
1	C	411	LYS
1	C	492[A]	GLU
1	C	492[B]	GLU
1	C	590	GLU
1	D	21	SER
1	D	127[A]	ARG
1	D	127[B]	ARG
1	D	139	PHE
1	D	191	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	275	ASP
1	D	590	GLU

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

Mol	Chain	Res	Type
1	A	70	HIS
1	A	633	GLN
1	C	430	ASN
1	C	633	GLN
1	D	594	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 17 ligands modelled in this entry, 10 are monoatomic - leaving 7 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	1PE	B	705	-	15,15,15	0.72	0	14,14,14	0.58	0
4	1PE	D	704	-	15,15,15	0.71	0	14,14,14	0.47	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HDD	D	701	1	38,52,52	1.71	10 (26%)	28,89,89	3.45	12 (42%)
2	HDD	C	701	1	38,52,52	1.91	12 (31%)	28,89,89	3.79	16 (57%)
2	HDD	B	701	1	38,52,52	1.86	10 (26%)	28,89,89	4.09	14 (50%)
4	1PE	A	705	-	15,15,15	0.74	0	14,14,14	0.64	0
2	HDD	A	701	1	38,52,52	1.97	10 (26%)	28,89,89	3.01	14 (50%)

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	1PE	B	705	-	-	8/13/13/13	-
4	1PE	D	704	-	-	7/13/13/13	-
2	HDD	D	701	1	-	0/3/89/89	0/1/9/9
2	HDD	C	701	1	-	0/3/89/89	0/1/9/9
2	HDD	B	701	1	-	0/3/89/89	0/1/9/9
4	1PE	A	705	-	-	7/13/13/13	-
2	HDD	A	701	1	-	0/3/89/89	0/1/9/9

All (42) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	701	HDD	O1D-CGD	4.99	1.43	1.35
2	C	701	HDD	C3B-C2B	4.71	1.46	1.40
2	A	701	HDD	C3B-C2B	4.65	1.46	1.40
2	A	701	HDD	O1D-CGD	4.56	1.43	1.35
2	B	701	HDD	C3B-C2B	4.34	1.46	1.40
2	D	701	HDD	OND-C2D	4.29	1.51	1.42
2	A	701	HDD	OND-C2D	4.08	1.50	1.42
2	C	701	HDD	OND-C2D	4.03	1.50	1.42
2	A	701	HDD	CHD-C1D	-3.79	1.30	1.36
2	D	701	HDD	CAD-C3D	-3.75	1.46	1.53
2	C	701	HDD	C3C-C2C	3.73	1.45	1.40
2	B	701	HDD	CMD-C2D	-3.62	1.48	1.53
2	D	701	HDD	C4A-CHB	3.45	1.50	1.41
2	A	701	HDD	C2A-C3A	3.40	1.47	1.37
2	C	701	HDD	O1D-C3D	3.30	1.52	1.46
2	C	701	HDD	CAA-C2A	-2.94	1.47	1.52
2	D	701	HDD	CAA-C2A	-2.76	1.48	1.52
2	A	701	HDD	O1D-C3D	2.74	1.51	1.46

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	701	HDD	O1D-CGD	2.64	1.40	1.35
2	C	701	HDD	C1B-C2B	2.61	1.48	1.42
2	C	701	HDD	C4D-ND	-2.60	1.33	1.37
2	D	701	HDD	C3C-C2C	2.54	1.43	1.40
2	A	701	HDD	CMD-C2D	-2.52	1.49	1.53
2	D	701	HDD	O1D-C3D	2.48	1.50	1.46
2	A	701	HDD	C4D-ND	-2.47	1.33	1.37
2	C	701	HDD	C4C-CHD	2.45	1.47	1.41
2	D	701	HDD	C1A-CHA	2.39	1.47	1.41
2	B	701	HDD	OND-C2D	2.39	1.47	1.42
2	B	701	HDD	C4C-NC	-2.37	1.31	1.36
2	B	701	HDD	C3C-C2C	2.35	1.43	1.40
2	C	701	HDD	C4A-NA	-2.35	1.31	1.36
2	C	701	HDD	C4A-CHB	2.28	1.47	1.41
2	B	701	HDD	C1A-CHA	2.24	1.47	1.41
2	B	701	HDD	C4D-ND	-2.15	1.34	1.37
2	B	701	HDD	CAA-C2A	-2.14	1.48	1.52
2	D	701	HDD	C1A-NA	-2.11	1.31	1.36
2	A	701	HDD	C3C-CAC	-2.11	1.43	1.47
2	D	701	HDD	FE-ND	2.11	2.04	1.95
2	D	701	HDD	CMB-C2B	-2.08	1.46	1.51
2	B	701	HDD	CHA-C4D	-2.08	1.32	1.36
2	C	701	HDD	C2A-C3A	2.02	1.43	1.37
2	A	701	HDD	C4A-CHB	2.01	1.46	1.41

All (56) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	701	HDD	O1D-CGD-O2D	12.56	132.01	120.80
2	B	701	HDD	O1D-CGD-CBD	-9.84	100.27	110.19
2	D	701	HDD	O1D-CGD-CBD	-8.88	101.24	110.19
2	A	701	HDD	C4A-C3A-C2A	-8.72	100.93	107.00
2	C	701	HDD	O1D-CGD-O2D	8.58	128.46	120.80
2	C	701	HDD	O1D-CGD-CBD	-7.91	102.21	110.19
2	D	701	HDD	C4A-C3A-C2A	-7.75	101.60	107.00
2	C	701	HDD	CAA-CBA-CGA	-6.97	100.98	112.67
2	B	701	HDD	C4A-C3A-C2A	-6.86	102.23	107.00
2	C	701	HDD	OND-C2D-CMD	-6.26	98.06	109.59
2	C	701	HDD	C4A-C3A-C2A	-5.70	103.03	107.00
2	B	701	HDD	OND-C2D-CMD	-5.69	99.11	109.59
2	D	701	HDD	O1D-CGD-O2D	5.66	125.85	120.80
2	D	701	HDD	CAA-CBA-CGA	-5.46	103.51	112.67

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	701	HDD	CMC-C2C-C3C	5.44	134.86	124.68
2	A	701	HDD	C2D-C1D-CHD	-5.34	115.47	124.28
2	C	701	HDD	CMC-C2C-C3C	5.17	134.34	124.68
2	C	701	HDD	C2D-C1D-CHD	-4.99	116.05	124.28
2	D	701	HDD	CMC-C2C-C3C	4.92	133.88	124.68
2	D	701	HDD	OND-C2D-CMD	-4.91	100.55	109.59
2	A	701	HDD	C4C-CHD-C1D	-4.36	121.49	130.12
2	B	701	HDD	CAA-CBA-CGA	-4.31	105.45	112.67
2	B	701	HDD	O1D-C3D-CAD	-4.30	94.93	103.01
2	B	701	HDD	C2D-C1D-CHD	-4.19	117.36	124.28
2	A	701	HDD	OND-C2D-CMD	-3.99	102.24	109.59
2	C	701	HDD	CAD-CBD-CGD	3.95	110.68	104.56
2	C	701	HDD	CMB-C2B-C3B	3.93	132.04	124.68
2	C	701	HDD	C4B-C3B-C2B	-3.82	104.23	106.90
2	B	701	HDD	CMC-C2C-C3C	3.77	131.74	124.68
2	D	701	HDD	C2D-C1D-CHD	-3.73	118.12	124.28
2	B	701	HDD	C4C-CHD-C1D	-3.60	122.99	130.12
2	A	701	HDD	CMB-C2B-C3B	3.58	131.38	124.68
2	D	701	HDD	C1A-CHA-C4D	-3.45	123.29	130.12
2	B	701	HDD	CMD-C2D-C1D	3.17	118.26	112.63
2	B	701	HDD	CMB-C2B-C3B	3.15	130.57	124.68
2	D	701	HDD	O1D-C3D-CAD	-3.02	97.33	103.01
2	D	701	HDD	C4B-C3B-C2B	-3.00	104.80	106.90
2	C	701	HDD	CBA-CAA-C2A	-2.96	107.03	112.49
2	A	701	HDD	CMD-C2D-C1D	2.92	117.81	112.63
2	A	701	HDD	O1D-CGD-O2D	2.87	123.36	120.80
2	D	701	HDD	CAD-CBD-CGD	2.77	108.86	104.56
2	A	701	HDD	C4B-C3B-C2B	2.77	108.84	106.90
2	C	701	HDD	C1A-CHA-C4D	-2.71	124.74	130.12
2	B	701	HDD	C1A-CHA-C4D	-2.61	124.95	130.12
2	A	701	HDD	C1A-CHA-C4D	-2.58	125.00	130.12
2	A	701	HDD	CAA-CBA-CGA	-2.57	108.35	112.67
2	B	701	HDD	CAD-CBD-CGD	2.41	108.29	104.56
2	C	701	HDD	C4C-CHD-C1D	-2.33	125.50	130.12
2	C	701	HDD	CHD-C1D-ND	2.29	127.56	124.20
2	D	701	HDD	C3D-C4D-CHA	-2.27	117.60	124.34
2	C	701	HDD	C3D-C4D-CHA	-2.26	117.63	124.34
2	A	701	HDD	C3D-C4D-CHA	-2.23	117.71	124.34
2	B	701	HDD	CHD-C1D-ND	2.20	127.42	124.20
2	A	701	HDD	O1D-C3D-CAD	-2.19	98.89	103.01
2	A	701	HDD	O1D-CGD-CBD	-2.15	108.02	110.19
2	C	701	HDD	O1D-C3D-CAD	-2.14	98.99	103.01

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	704	1PE	OH2-C12-C22-OH3
4	D	704	1PE	OH6-C15-C25-OH5
4	A	705	1PE	OH4-C13-C23-OH3
4	B	705	1PE	OH4-C13-C23-OH3
4	D	704	1PE	OH7-C16-C26-OH6
4	B	705	1PE	OH7-C16-C26-OH6
4	D	704	1PE	OH5-C14-C24-OH4
4	B	705	1PE	C13-C23-OH3-C22
4	B	705	1PE	C16-C26-OH6-C15
4	B	705	1PE	C15-C25-OH5-C14
4	A	705	1PE	C23-C13-OH4-C24
4	D	704	1PE	C15-C25-OH5-C14
4	D	704	1PE	C25-C15-OH6-C26
4	A	705	1PE	C24-C14-OH5-C25
4	B	705	1PE	C23-C13-OH4-C24
4	D	704	1PE	C16-C26-OH6-C15
4	A	705	1PE	OH6-C15-C25-OH5
4	B	705	1PE	OH6-C15-C25-OH5
4	B	705	1PE	OH5-C14-C24-OH4
4	A	705	1PE	C16-C26-OH6-C15
4	A	705	1PE	C14-C24-OH4-C13
4	A	705	1PE	C13-C23-OH3-C22

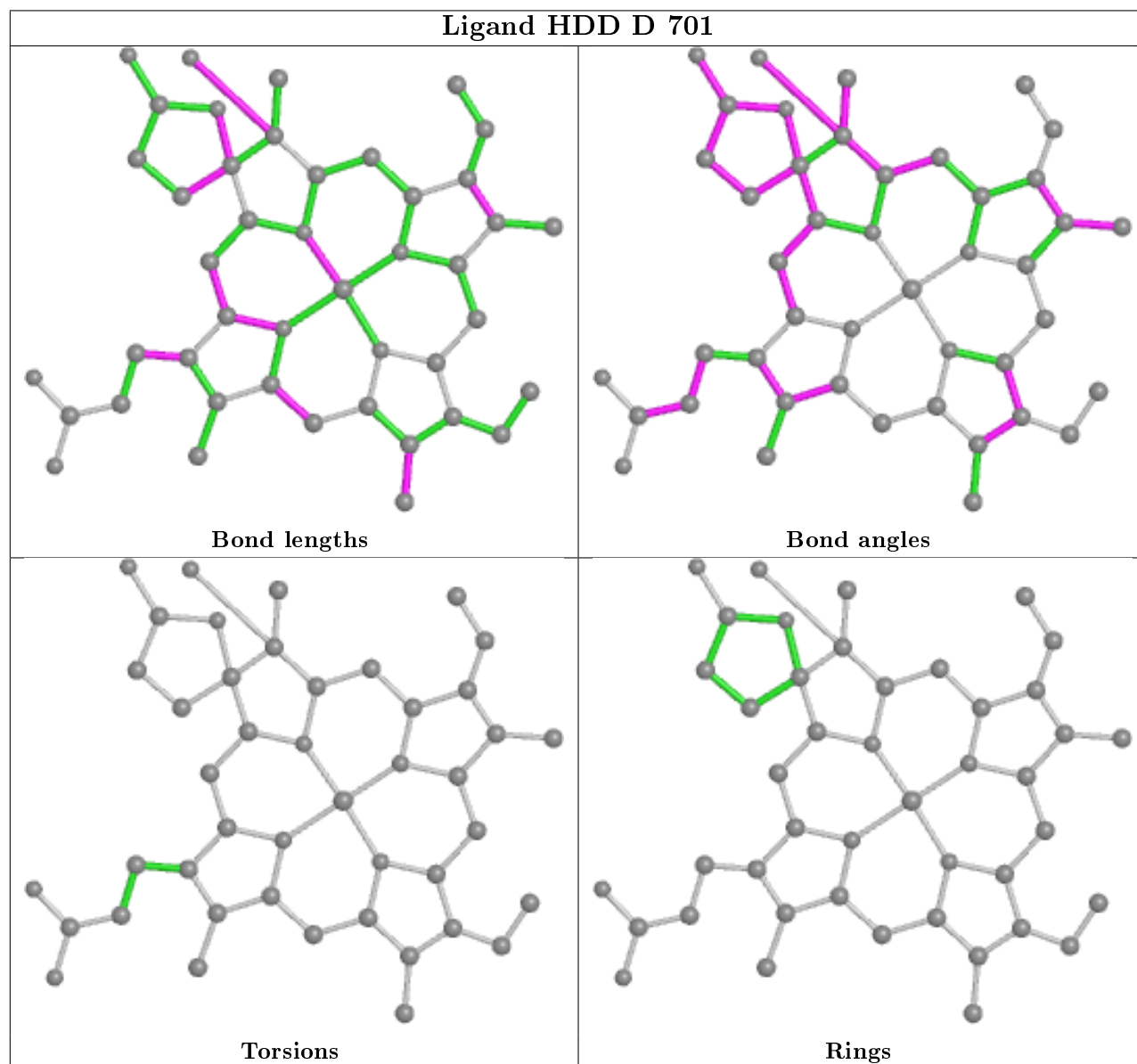
There are no ring outliers.

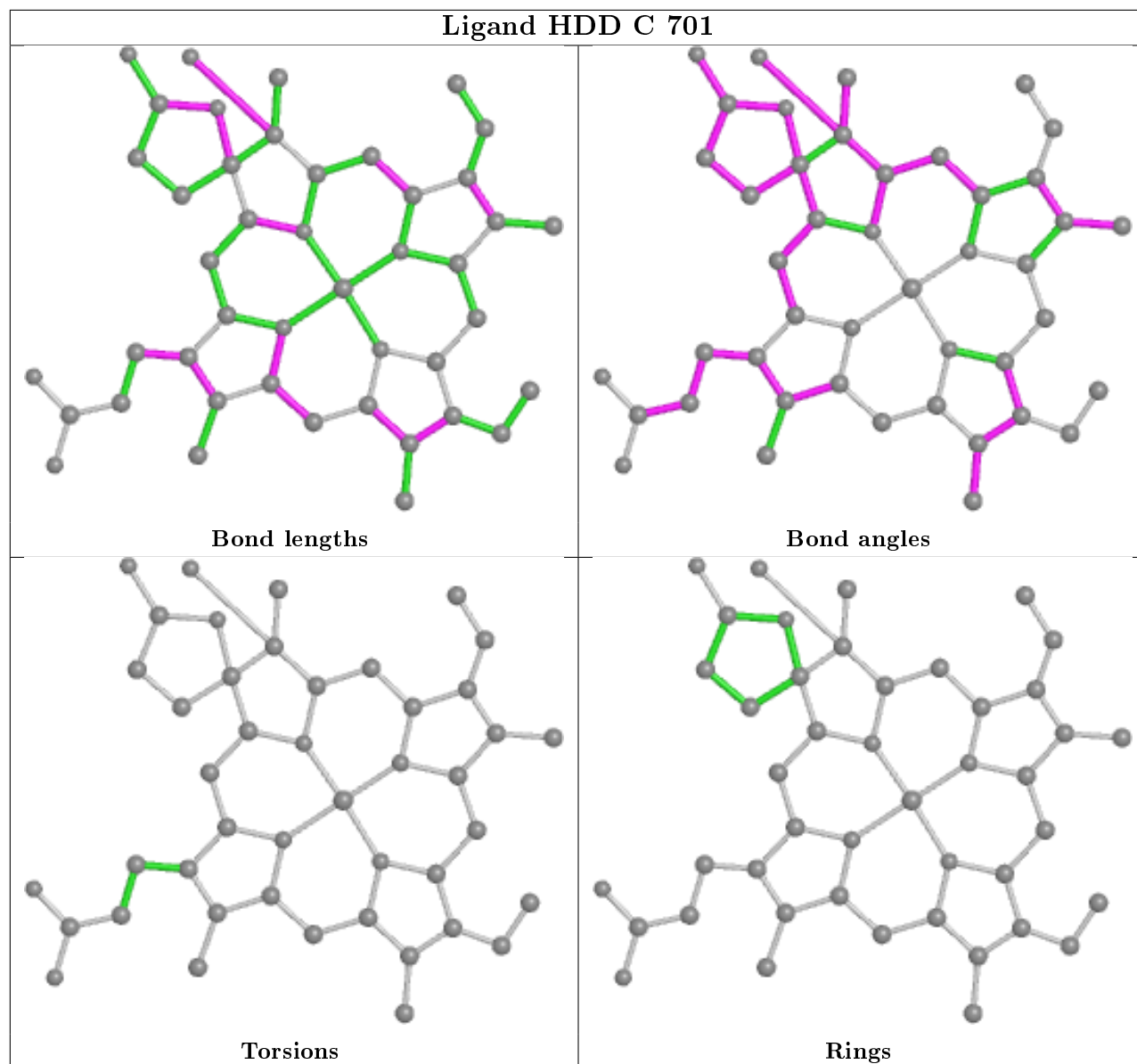
4 monomers are involved in 9 short contacts:

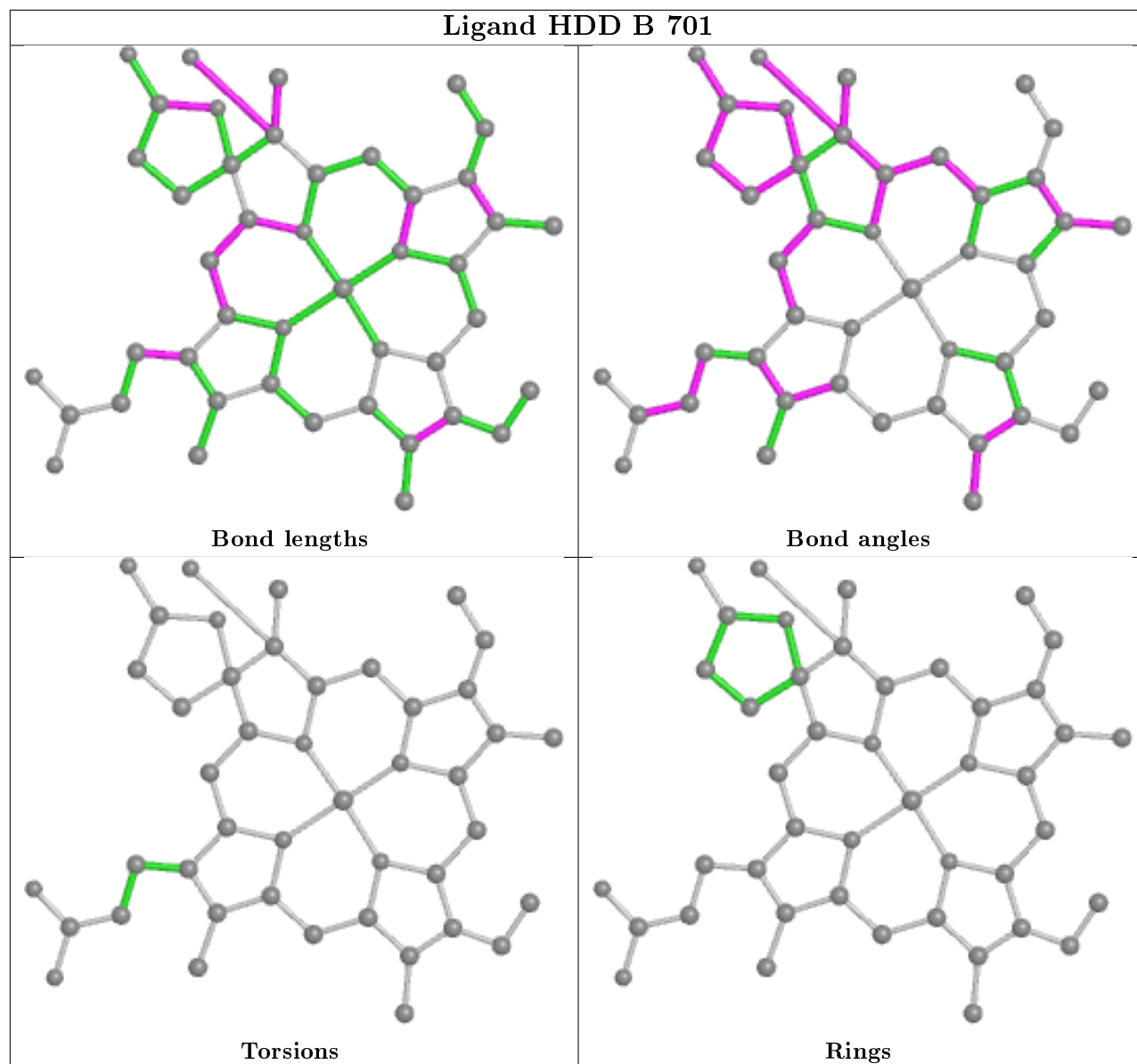
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	701	HDD	3	0
2	C	701	HDD	1	0
2	B	701	HDD	2	0
2	A	701	HDD	3	0

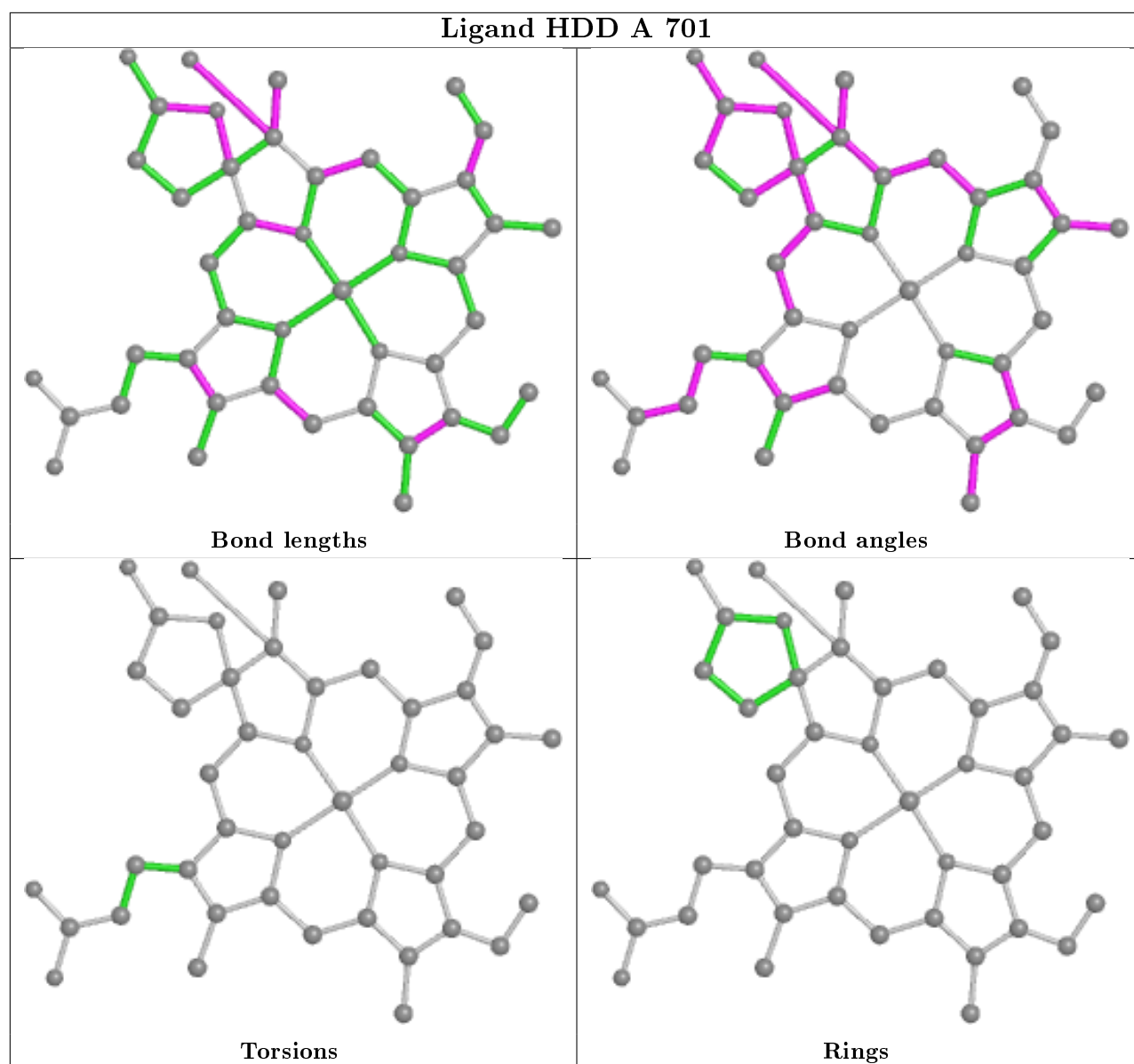
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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	A	675/678 (99%)	-0.63	3 (0%) 92 93	15, 24, 49, 89	0
1	B	675/678 (99%)	-0.65	4 (0%) 89 90	14, 24, 48, 77	0
1	C	675/678 (99%)	-0.71	2 (0%) 94 94	14, 22, 38, 55	0
1	D	675/678 (99%)	-0.70	0 100 100	14, 24, 39, 53	0
All	All	2700/2712 (99%)	-0.67	9 (0%) 94 94	14, 23, 43, 89	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	698	SER	5.2
1	A	654	GLU	3.2
1	C	698	SER	3.2
1	B	658	ALA	3.0
1	A	655	VAL	2.7
1	C	618	ALA	2.4
1	B	653	SER	2.1
1	B	590	GLU	2.1
1	B	654	GLU	2.0

### 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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

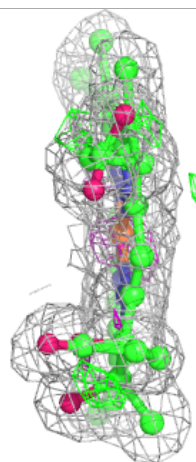
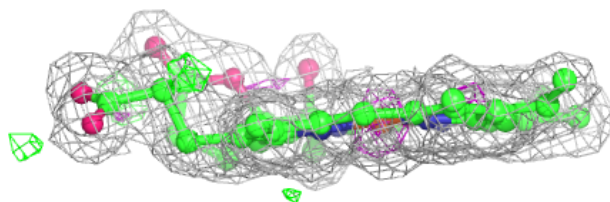
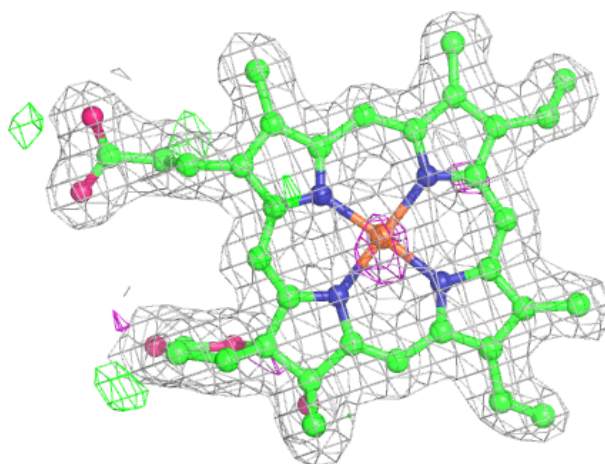
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	1PE	A	705	16/16	0.82	0.15	48,54,70,71	0
4	1PE	B	705	16/16	0.82	0.23	43,59,83,84	0
4	1PE	D	704	16/16	0.88	0.27	41,54,89,89	0
3	CA	D	703	1/1	0.96	0.06	48,48,48,48	0
2	HDD	D	701	44/44	0.98	0.07	16,18,21,23	0
3	CA	B	704	1/1	0.98	0.06	52,52,52,52	0
2	HDD	C	701	44/44	0.98	0.07	14,18,21,26	0
2	HDD	B	701	44/44	0.98	0.07	15,17,22,25	0
2	HDD	A	701	44/44	0.98	0.08	13,16,21,29	0
3	CA	C	702	1/1	0.99	0.05	31,31,31,31	0
3	CA	A	703	1/1	0.99	0.06	42,42,42,42	0
3	CA	A	702	1/1	0.99	0.04	25,25,25,25	0
3	CA	A	704	1/1	0.99	0.08	39,39,39,39	0
3	CA	C	703	1/1	0.99	0.09	41,41,41,41	0
3	CA	B	703	1/1	0.99	0.06	45,45,45,45	0
3	CA	D	702	1/1	1.00	0.04	28,28,28,28	0
3	CA	B	702	1/1	1.00	0.04	27,27,27,27	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.

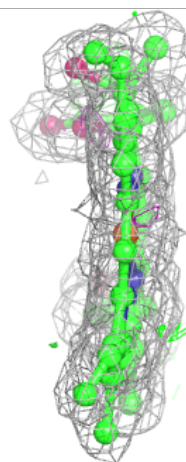
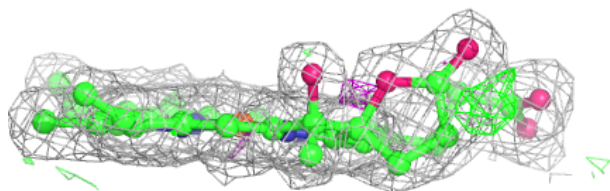
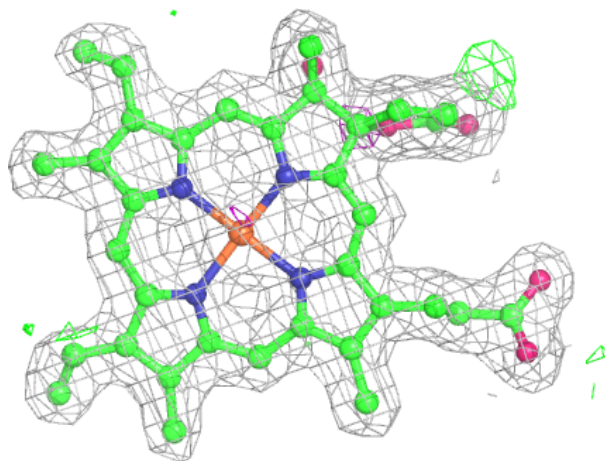
**Electron density around HDD D 701:**

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



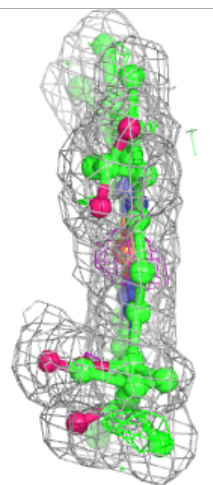
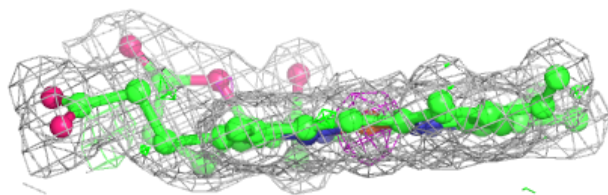
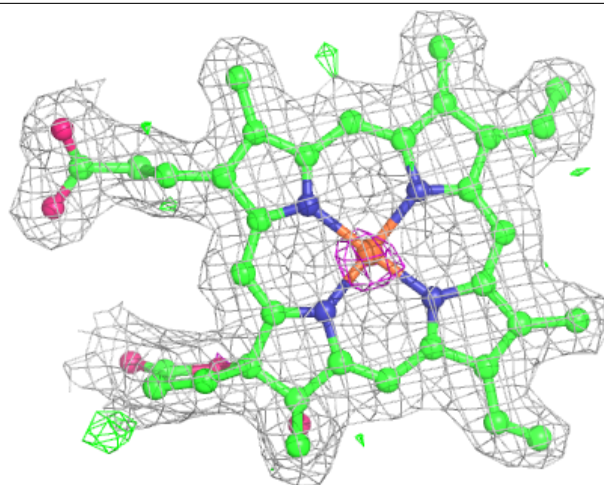
**Electron density around HDD C 701:**

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



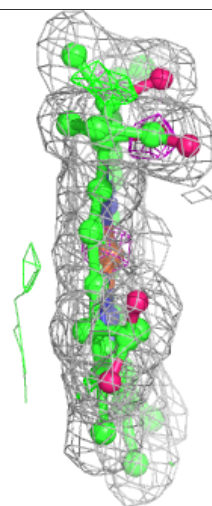
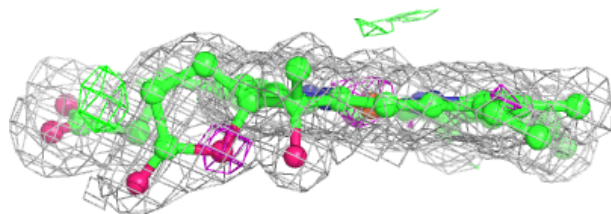
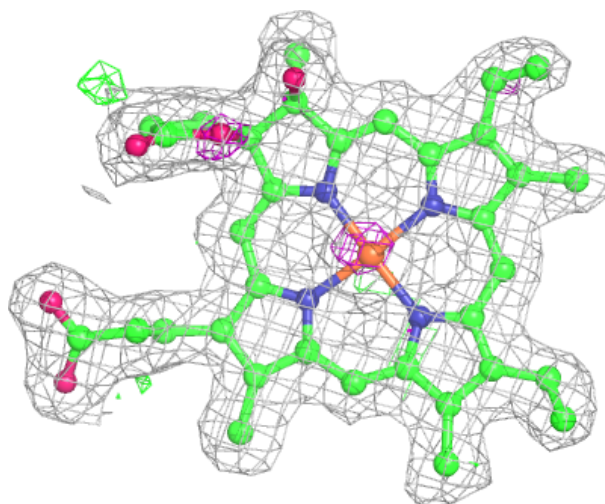
**Electron density around HDD B 701:**

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



**Electron density around HDD A 701:**

$2mF_o-DF_c$  (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.