

wwPDB X-ray Structure Validation Summary Report (i)

Oct 23, 2021 – 10:01 AM EDT

PDB ID : 1GGJ

Title : CRYSTAL STRUCTURE OF CATALASE HPII FROM ESCHERICHIA

COLI, ASN201ALA VARIANT.

Authors: Melik-Adamyan, W.R.; Bravo, J.; Carpena, X.; Switala, J.; Mate, M.J.; Fita,

I.; Loewen, P.C.

Deposited on : 2000-08-21

Resolution : 1.92 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 following versions of software and data (see references (1)) 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 : 2.23.2buster-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)

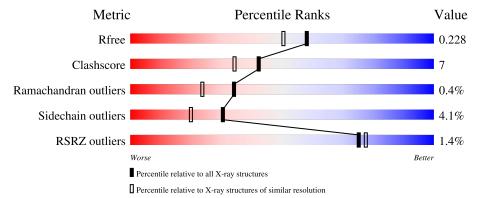
Validation Pipeline (wwPDB-VP) : 2.23.2

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 1.92 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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 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	A	753	81%	15%	• •
1	В	753	78%	17%	• •
1	С	753	78%	17%	• •
1	D	753	78%	16%	•••



2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 25962 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 HPII.

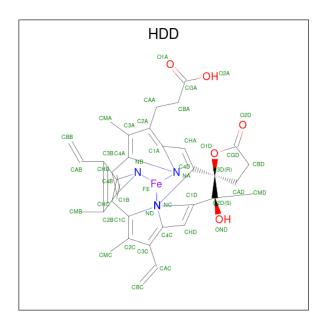
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	Λ	727	Total	С	N	О	S	0	0	0
1	A	121	5743	3646	1004	1081	12	U	U	
1	В	727	Total	С	N	О	S	0	0	0
1	Ъ	121	5743	3646	1004	1081	12		U	
1	С	727	Total	С	N	О	S	0	0	0
1		121	5743	3646	1004	1081	12	U	U	
1	D	727	Total	С	N	О	S	0	0	0
1	ע	121	5743	3646	1004	1081	12		U	U

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	ALA	ASN	engineered mutation	UNP P21179
В	201	ALA	ASN	engineered mutation	UNP P21179
С	201	ALA	ASN	engineered mutation	UNP P21179
D	201	ALA	ASN	engineered mutation	UNP P21179

• Molecule 2 is CIS-HEME D HYDROXYCHLORIN GAMMA-SPIROLACTONE (three-letter code: HDD) (formula: C₃₄H₃₂FeN₄O₅).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf		
2	Λ	1	Total	С	Fe	N	О	0	0	
	А	1	44	34	1	4	5	0		
2	В	1	Total	С	Fe	N	О	0	0	
	Б	1	44	34	1	4	5	0		
2	C	1	Total	С	Fe	N	О	0	0	
	C	1	44	34	1	4	5	0	U	
2	D	1	Total	С	Fe	N	О	0	0	
2	D	1	44	34	1	4	5		U	

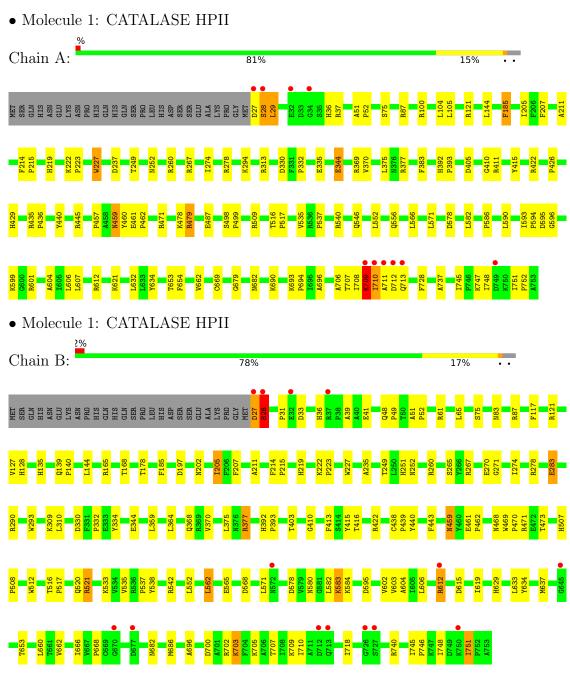
• Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	731	Total O 731 731	0	0
3	В	655	Total O 655 655	0	0
3	С	704	Total O 704 704	0	0
3	D	724	Total O 724 724	0	0



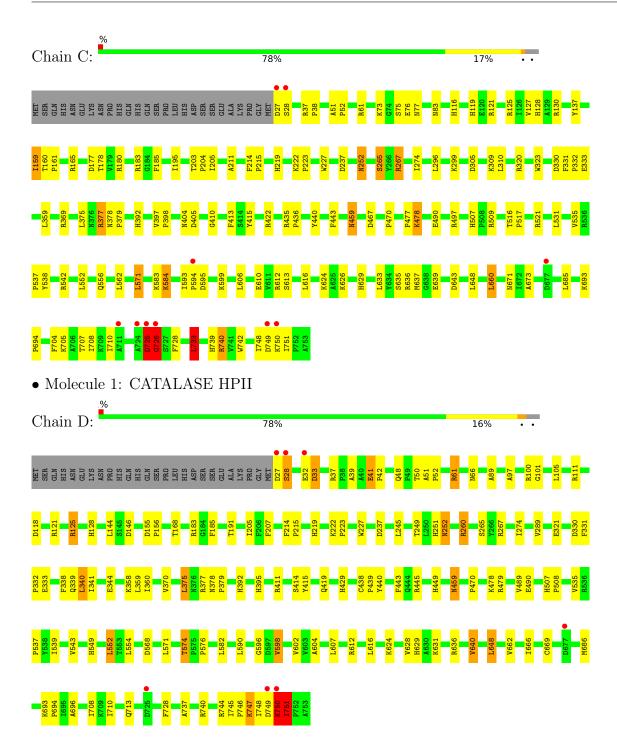
3 Residue-property plots (i)

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: CATALASE HPII







4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	93.04Å 132.34Å 121.20Å	Depositor
a, b, c, α , β , γ	90.00° 109.63° 90.00°	Depositor
Resolution (Å)	87.71 - 1.92	Depositor
Resolution (A)	19.26 - 1.92	EDS
% Data completeness	91.7 (87.71-1.92)	Depositor
(in resolution range)	91.7 (19.26-1.92)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.78 (at 1.91Å)	Xtriage
Refinement program	REFMAC	Depositor
D.D.	0.182 , 0.235	Depositor
R, R_{free}	0.181 , 0.228	DCC
R_{free} test set	9705 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	7.3	Xtriage
Anisotropy	0.467	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.35, 55.4	EDS
L-test for twinning ²	$< L >=0.50, < L^2>=0.33$	Xtriage
Estimated twinning fraction	0.027 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	25962	wwPDB-VP
Average B, all atoms (Å ²)	9.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.13% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

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

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		
MIOI		RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.43	0/5899	1.11	$24/8020 \ (0.3\%)$	
1	В	0.43	0/5899	1.11	21/8020 (0.3%)	
1	С	0.43	0/5899	1.10	19/8020 (0.2%)	
1	D	0.42	0/5899	1.14	$24/8020 \ (0.3\%)$	
All	All	0.43	0/23596	1.11	88/32080 (0.3%)	

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 maintenain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers	
1	С	0	1	

There are no bond length outliers.

The worst 5 of 88 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	D	445	ARG	NE-CZ-NH1	18.13	129.37	120.30
1	В	521	ARG	CD-NE-CZ	13.33	142.27	123.60
1	В	377	ARG	NE-CZ-NH2	-11.71	114.44	120.30
1	A	445	ARG	NE-CZ-NH1	11.37	125.98	120.30
1	D	61	ARG	NE-CZ-NH2	11.33	125.97	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	С	726	GLY	Mainchain



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	A	5743	0	5577	72	0
1	В	5743	0	5577	80	0
1	С	5743	0	5576	81	0
1	D	5743	0	5577	86	0
2	A	44	0	31	7	0
2	В	44	0	31	5	0
2	С	44	0	31	5	0
2	D	44	0	31	4	0
3	A	731	0	0	6	0
3	В	655	0	0	4	0
3	С	704	0	0	13	0
3	D	724	0	0	10	0
All	All	25962	0	22431	309	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 309 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$egin{array}{c} { m Clash} \\ { m overlap} \ ({ m \AA}) \end{array}$
1:B:392:HIS:ND1	1:B:415:TYR:HB2	1.13	1.42
1:D:392:HIS:ND1	1:D:415:TYR:HB2	1.12	1.38
1:C:392:HIS:ND1	1:C:415:TYR:HB2	1.09	1.38
1:A:392:HIS:ND1	1:A:415:TYR:HB2	1.09	1.37
2:B:760:HDD:HBC1	2:B:760:HDD:HMC1	1.33	1.07

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 nur	ıber of	f residues	for	which	the	backbone	conformation	was
analysed, and the total number of res	dues.							

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	\mathbf{ntiles}
1	A	725/753~(96%)	697 (96%)	24 (3%)	4 (1%)	25	14
1	В	725/753~(96%)	697 (96%)	26 (4%)	2 (0%)	41	31
1	C	725/753~(96%)	699 (96%)	23 (3%)	3 (0%)	34	24
1	D	725/753~(96%)	700 (97%)	21 (3%)	4 (1%)	25	14
All	All	2900/3012 (96%)	2793 (96%)	94 (3%)	13 (0%)	34	24

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	709	LYS
1	A	711	ALA
1	D	28	SER
1	D	751	ILE
1	A	28	SER

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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	$oxed{ \ \ \ \ \ \ \ \ \ \ $		Percentiles	
1	A	611/635~(96%)	596 (98%)	15 (2%)	47 39	
1	В	611/635~(96%)	587 (96%)	24 (4%)	32 22	
1	С	611/635~(96%)	580 (95%)	31 (5%)	24 13	
1	D	611/635~(96%)	582 (95%)	29 (5%)	26 15	
All	All	$2444/2540 \ (96\%)$	2345 (96%)	99 (4%)	30 20	

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	С	571	LEU
1	D	185	PHE
1	С	606	LEU

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type
1	С	660	LEU
1	D	237	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	С	459	ASN
1	С	629	HIS
1	D	546	GLN
1	С	556	GLN
1	С	671	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with |Z| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Chain Res Link Bond lengths		Bond angles					
MIOI	Type	Chain	nes	Lilik	Counts RMSZ #		# Z > 2	Counts	RMSZ	# Z > 2
2	HDD	В	760	1,3	38,52,52	3.00	14 (36%)	28,89,89	4.95	16 (57%)



Mol	Trino	Chain	Res Link		В	ond leng	Bond angles			
MIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	HDD	A	760	1,3	38,52,52	2.99	13 (34%)	28,89,89	4.75	17 (60%)
2	HDD	С	760	1,3	38,52,52	2.93	13 (34%)	28,89,89	5.27	18 (64%)
2	HDD	D	760	1	38,52,52	3.10	15 (39%)	28,89,89	4.19	17 (60%)

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
2	HDD	В	760	1,3	-	0/3/89/89	0/1/9/9
2	HDD	A	760	1,3	-	0/3/89/89	0/1/9/9
2	HDD	С	760	1,3	-	0/3/89/89	0/1/9/9
2	HDD	D	760	1	-	0/3/89/89	0/1/9/9

The worst 5 of 55 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	Ideal(A)
2	D	760	HDD	C2D-C3D	-11.31	1.27	1.55
2	С	760	HDD	C2D-C3D	-11.26	1.27	1.55
2	В	760	HDD	C2D-C3D	-11.25	1.27	1.55
2	A	760	HDD	C2D-C3D	-11.06	1.27	1.55
2	D	760	HDD	CAD-C3D	6.20	1.64	1.53

The worst 5 of 68 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$Ideal(^{o})$
2	A	760	HDD	C4B-C3B-C2B	13.67	116.45	106.90
2	С	760	HDD	C4B-C3B-C2B	13.64	116.42	106.90
2	В	760	HDD	C4B-C3B-C2B	13.20	116.11	106.90
2	В	760	HDD	CHD-C1D-ND	11.21	140.64	124.20
2	С	760	HDD	OND-C2D-CMD	-10.92	89.48	109.59

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	В	760	HDD	5	0

Continued on next page...

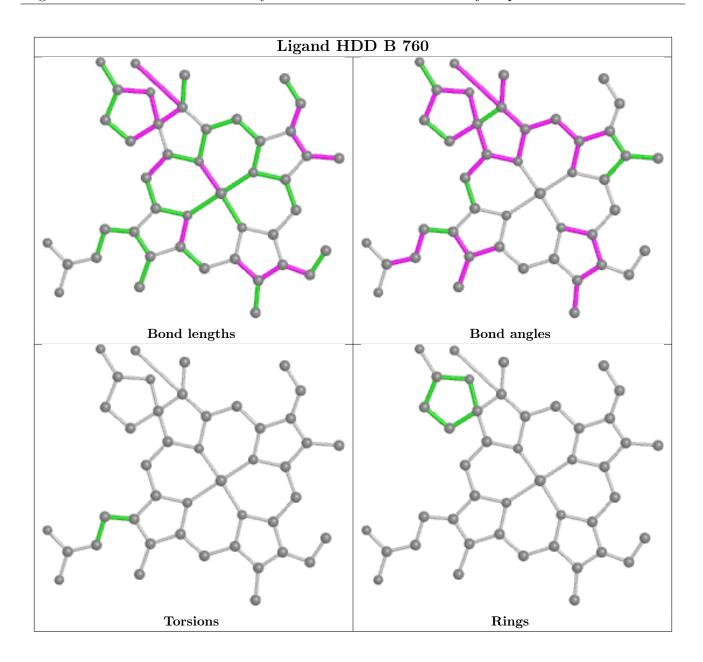


Continued from previous page...

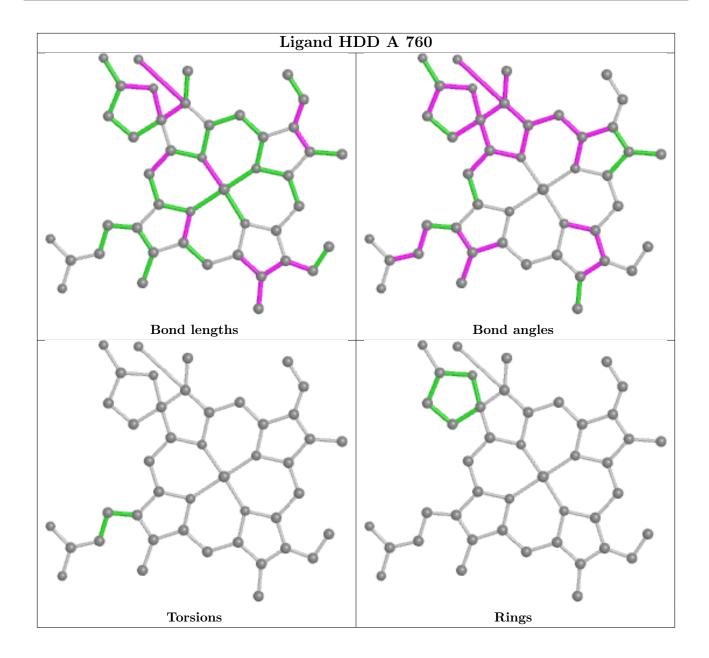
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	760	HDD	7	0
2	С	760	HDD	5	0
2	D	760	HDD	4	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

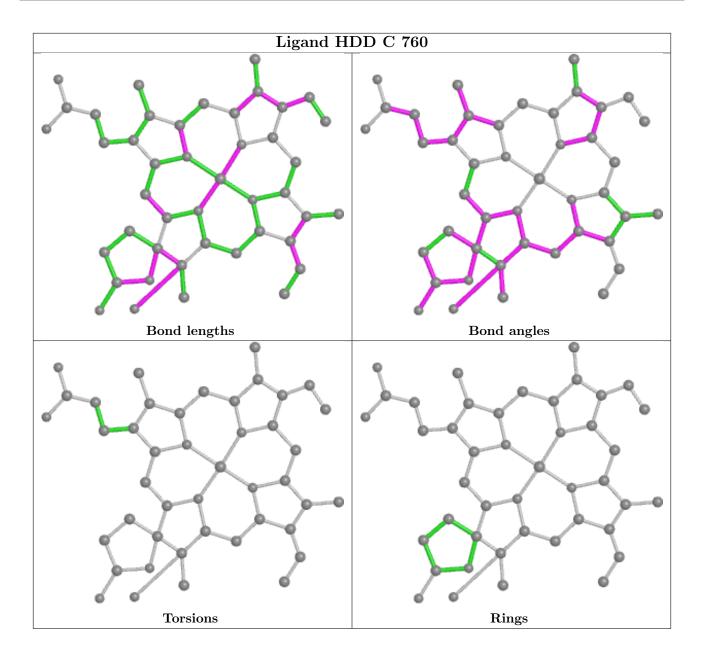




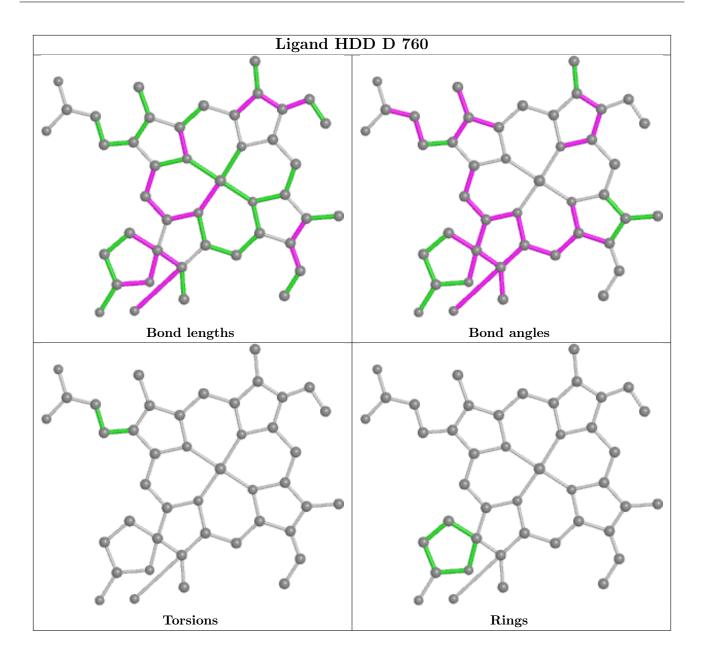












5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.



6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ>2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<rsrz></rsrz>	$\# \mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q<0.9
1	A	727/753 (96%)	-0.44	10 (1%) 75 77	2, 6, 21, 35	1 (0%)
1	В	727/753 (96%)	-0.38	14 (1%) 66 69	2, 7, 23, 33	1 (0%)
1	С	727/753 (96%)	-0.39	10 (1%) 75 77	2, 7, 23, 32	1 (0%)
1	D	727/753 (96%)	-0.47	7 (0%) 82 84	2, 7, 21, 33	1 (0%)
All	All	2908/3012 (96%)	-0.42	41 (1%) 75 77	2, 7, 22, 35	4 (0%)

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	711	ALA	8.4
1	A	27	ASP	7.9
1	A	710	ILE	4.9
1	В	27	ASP	4.7
1	В	28	SER	4.7

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

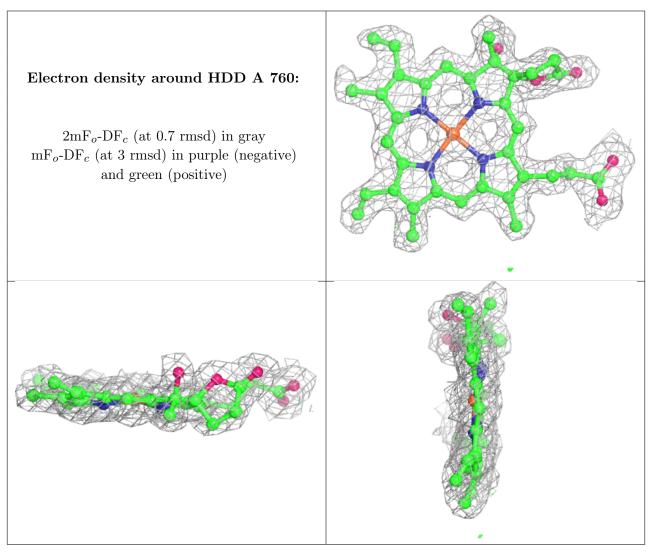
6.4 Ligands (i)

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

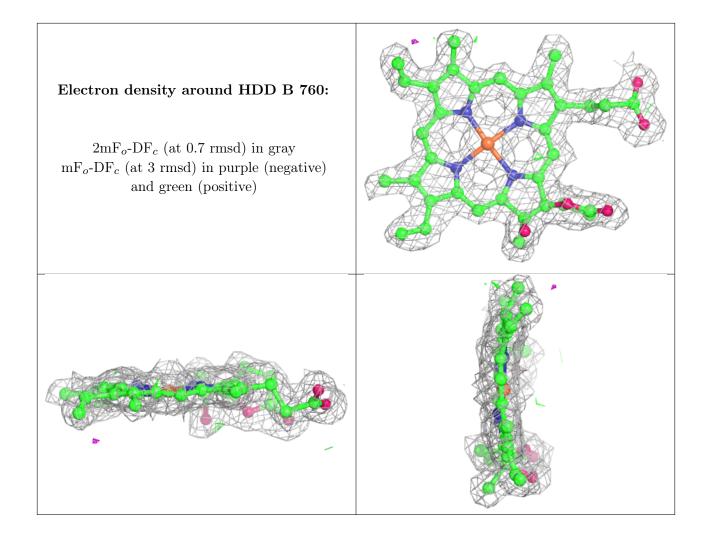


Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	HDD	A	760	44/44	0.98	0.08	8,10,12,13	0
2	HDD	В	760	44/44	0.98	0.08	2,9,13,16	0
2	HDD	С	760	44/44	0.98	0.08	7,11,13,15	0
2	HDD	D	760	44/44	0.98	0.08	2,9,12,14	0

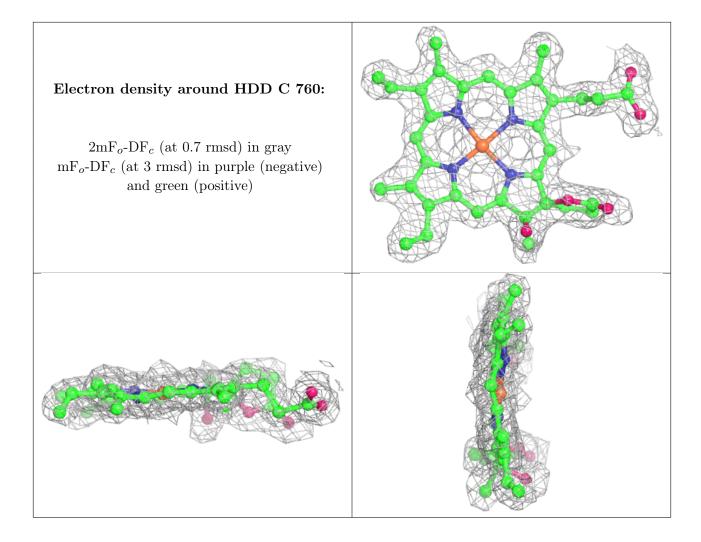
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



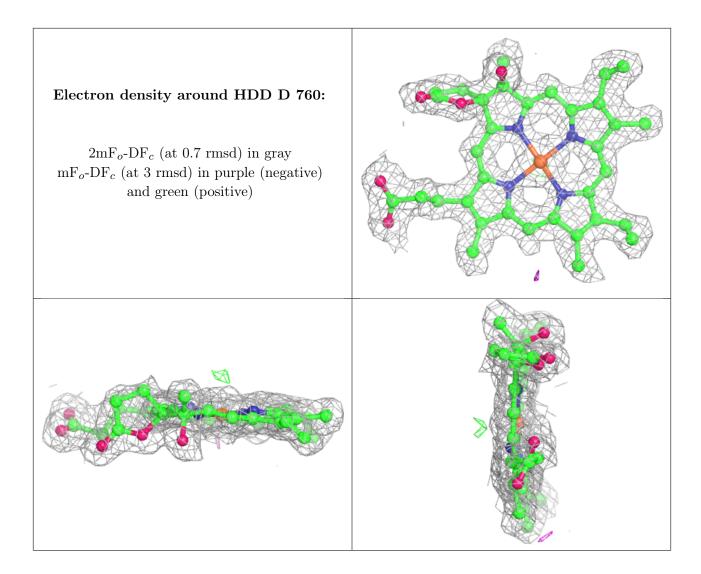












6.5 Other polymers (i)

There are no such residues in this entry.

