



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2023 – 10:47 PM EDT

PDB ID : 1NQP  
Title : Crystal structure of Human hemoglobin E at 1.73 Å resolution  
Authors : Dasgupta, J.; Sen, U.; Choudhury, D.; Dutta, P.; Basu, S.; Chakrabarti, A.;  
Chakrabarty, A.; Dattagupta, J.K.  
Deposited on : 2003-01-22  
Resolution : 1.73 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35

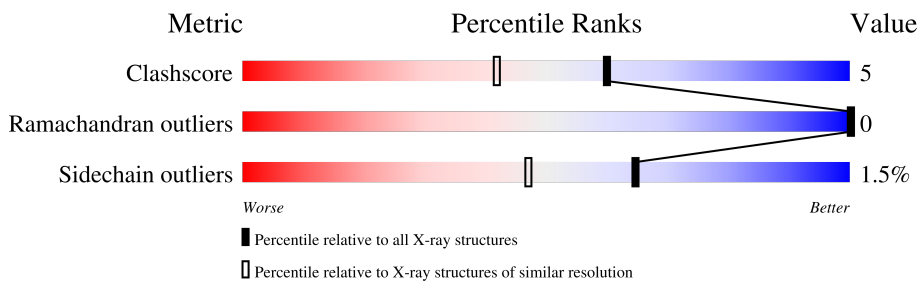
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.73 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))
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	141	
1	C	141	
2	B	146	
2	D	146	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5207 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 Hemoglobin alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	141	1060	679	184	194	3	0	0	0
1	C	141	1065	682	186	194	3	0	0	0

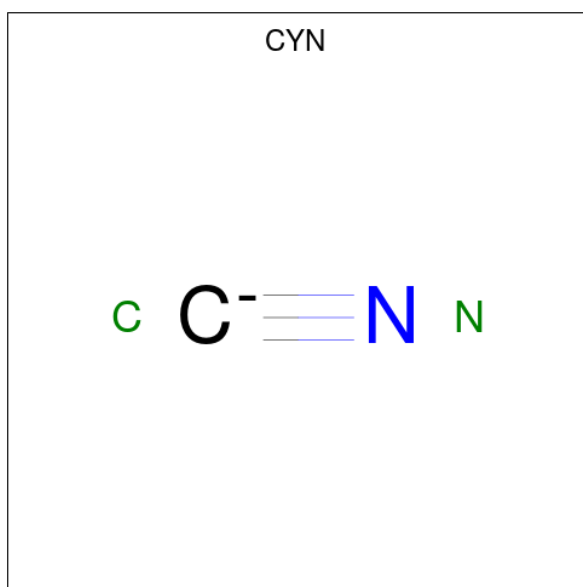
- Molecule 2 is a protein called Hemoglobin beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	146	1115	719	194	199	3	0	0	0
2	D	146	1116	720	194	199	3	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

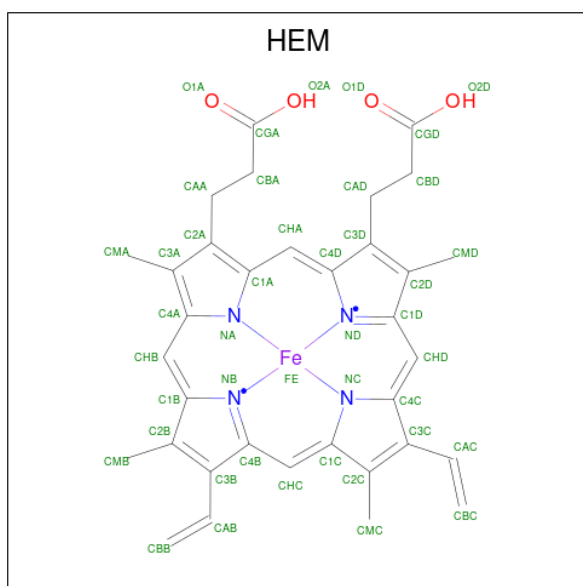
Chain	Residue	Modelled	Actual	Comment	Reference
B	26	LYS	GLU	variant	UNP P68871
D	26	LYS	GLU	variant	UNP P02023

- Molecule 3 is CYANIDE ION (three-letter code: CYN) (formula: CN).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N 2 1 1	0	0
3	B	1	Total C N 2 1 1	0	0
3	C	1	Total C N 2 1 1	0	0
3	D	1	Total C N 2 1 1	0	0

- Molecule 4 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $\text{C}_{34}\text{H}_{32}\text{FeN}_4\text{O}_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
4	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
4	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
4	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
4	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 5 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	152	Total	O	0	0
			152	152		
5	B	172	Total	O	0	0
			172	172		
5	C	178	Total	O	0	0
			178	178		
5	D	169	Total	O	0	0
			169	169		

### 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. 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. 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.


Note EDS was not executed.

- Molecule 1: Hemoglobin alpha chain

Chain A:  91% 9%



- Molecule 1: Hemoglobin alpha chain

Chain C:  87% 13%




- Molecule 2: Hemoglobin beta chain

Chain B:  94% 6%



- Molecule 2: Hemoglobin beta chain

Chain D:  86% 13%



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.89Å 95.81Å 99.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.99 – 1.73	Depositor
% Data completeness (in resolution range)	97.4 (14.99-1.73)	Depositor
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.04	Depositor
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.191 , 0.211	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	5207	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.33	0/1087	0.54	0/1479
1	C	0.32	0/1093	0.53	0/1487
2	B	0.31	0/1145	0.54	0/1557
2	D	0.31	0/1145	0.52	0/1554
All	All	0.32	0/4470	0.53	0/6077

There are no bond length outliers.

There are no bond angle outliers.

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	A	1060	0	1057	15	0
1	C	1065	0	1062	11	0
2	B	1115	0	1103	6	0
2	D	1116	0	1113	16	0
3	A	2	0	0	1	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	A	43	0	30	0	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	43	0	30	0	0
4	C	43	0	30	0	0
4	D	43	0	30	0	0
5	A	152	0	0	4	0
5	B	172	0	0	1	0
5	C	178	0	0	2	0
5	D	169	0	0	2	0
All	All	5207	0	4455	41	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 5.

The worst 5 of 41 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:141:ARG:HG2	5:A:220:HOH:O	1.60	0.98
1:A:92:ARG:NH1	2:D:40:ARG:HA	1.92	0.84
1:A:141:ARG:HA	5:A:220:HOH:O	1.76	0.84
1:C:47:ASP:H	1:C:54:GLN:NE2	1.85	0.75
2:B:66:LYS:NZ	5:B:509:HOH:O	2.31	0.64

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	139/141 (99%)	136 (98%)	3 (2%)	0	100	100
1	C	139/141 (99%)	137 (99%)	2 (1%)	0	100	100
2	B	144/146 (99%)	143 (99%)	1 (1%)	0	100	100
2	D	144/146 (99%)	137 (95%)	7 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	566/574 (99%)	553 (98%)	13 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	111/113 (98%)	110 (99%)	1 (1%)	78	67
1	C	112/113 (99%)	110 (98%)	2 (2%)	59	38
2	B	116/118 (98%)	116 (100%)	0	100	100
2	D	116/118 (98%)	112 (97%)	4 (3%)	37	13
All	All	455/462 (98%)	448 (98%)	7 (2%)	65	47

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	51	PRO
2	D	120	LYS
2	D	141	LEU
2	D	139	ASN
1	C	77	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	97	ASN
2	D	77	HIS
2	D	108	ASN
2	B	139	ASN
1	C	9	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](#)

8 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	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	HEM	C	544	1,3	41,50,50	1.37	6 (14%)	45,82,82	1.66	8 (17%)
3	CYN	A	150	4	0,1,1	-	-	-		
3	CYN	D	750	4	0,1,1	-	-	-		
4	HEM	D	748	2,3	41,50,50	1.42	7 (17%)	45,82,82	1.56	9 (20%)
3	CYN	B	350	4	0,1,1	-	-	-		
4	HEM	A	144	1,3	41,50,50	1.45	6 (14%)	45,82,82	1.53	7 (15%)
4	HEM	B	348	2,3	41,50,50	1.32	5 (12%)	45,82,82	1.53	8 (17%)
3	CYN	C	550	4	0,1,1	-	-	-		

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	HEM	A	144	1,3	-	2/12/54/54	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	HEM	C	544	1,3	-	6/12/54/54	-
4	HEM	B	348	2,3	-	3/12/54/54	-
4	HEM	D	748	2,3	-	4/12/54/54	-

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	144	HEM	C3C-CAC	-4.15	1.39	1.47
4	A	144	HEM	CBB-CAB	3.76	1.48	1.30
4	B	348	HEM	CBB-CAB	3.73	1.48	1.30
4	C	544	HEM	C3C-CAC	-3.72	1.40	1.47
4	C	544	HEM	CBB-CAB	3.64	1.48	1.30

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	544	HEM	CBA-CAA-C2A	6.16	123.13	112.62
4	A	144	HEM	CBA-CAA-C2A	5.26	121.60	112.62
4	D	748	HEM	C4B-CHC-C1C	4.11	127.98	122.56
4	B	348	HEM	C4C-CHD-C1D	3.79	127.56	122.56
4	D	748	HEM	CBA-CAA-C2A	3.58	118.73	112.62

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	544	HEM	C2B-C3B-CAB-CBB
4	D	748	HEM	C2B-C3B-CAB-CBB
4	D	748	HEM	C4B-C3B-CAB-CBB
4	C	544	HEM	C4B-C3B-CAB-CBB
4	A	144	HEM	CAD-CBD-CGD-O1D

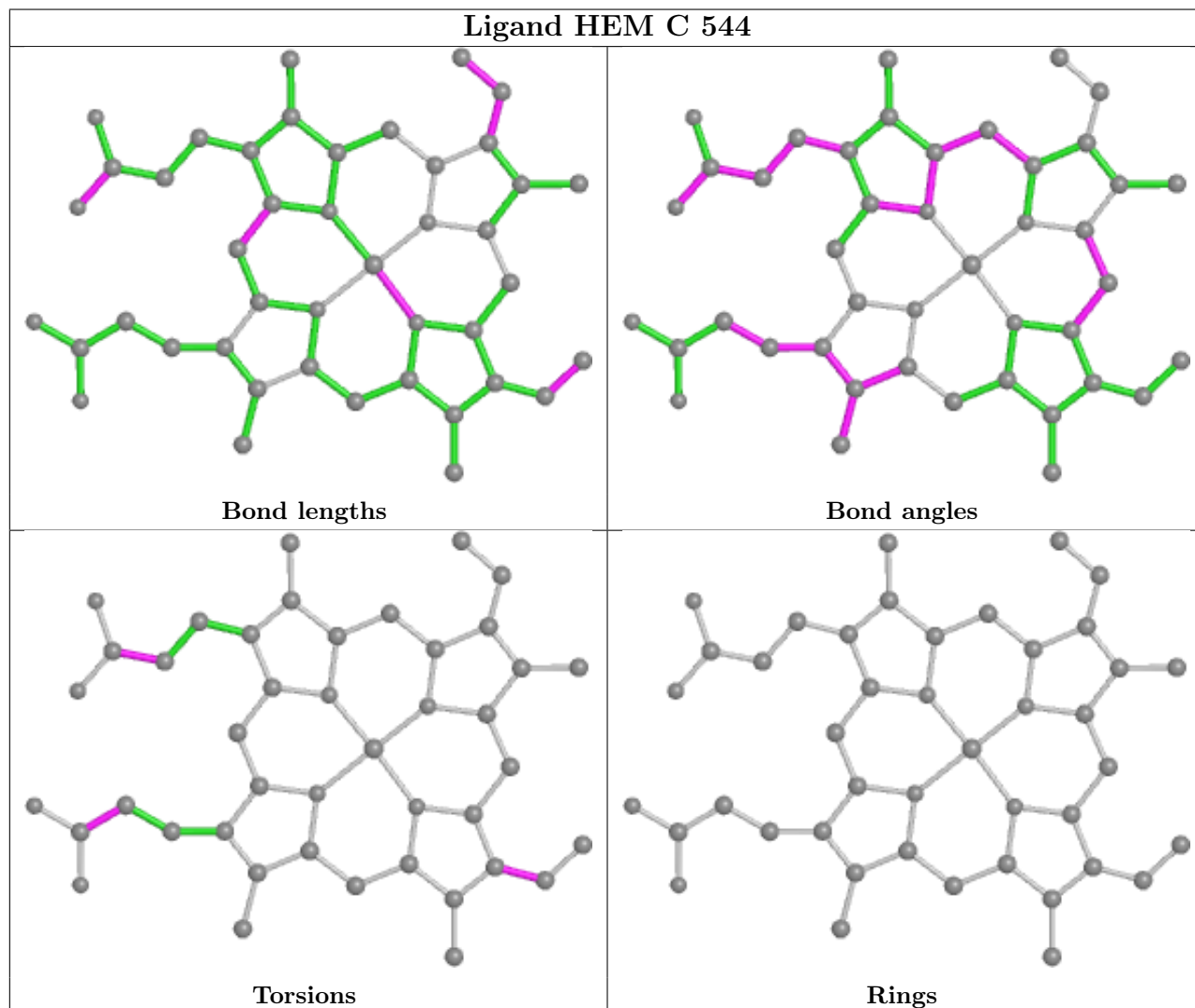
There are no ring outliers.

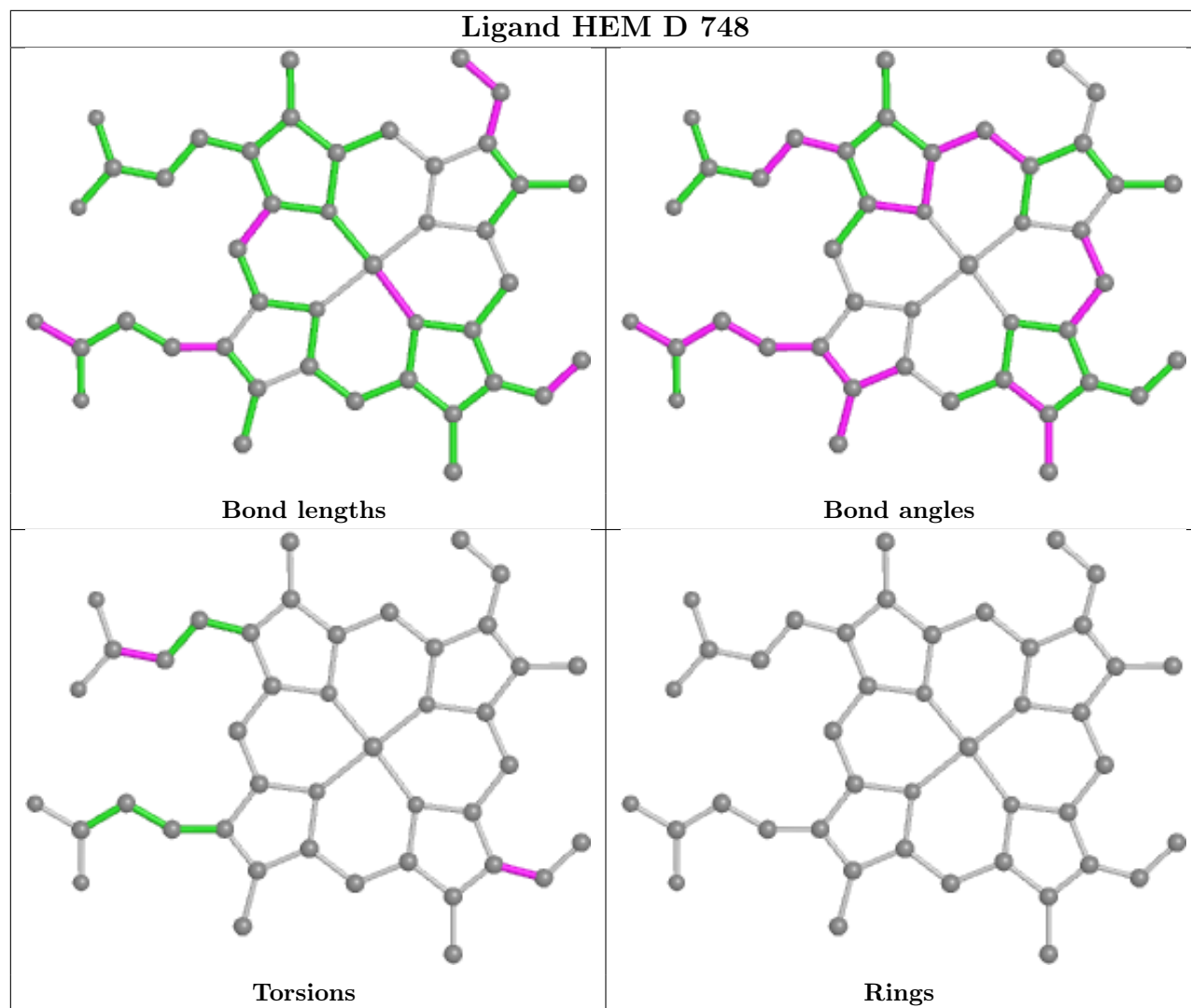
1 monomer is involved in 1 short contact:

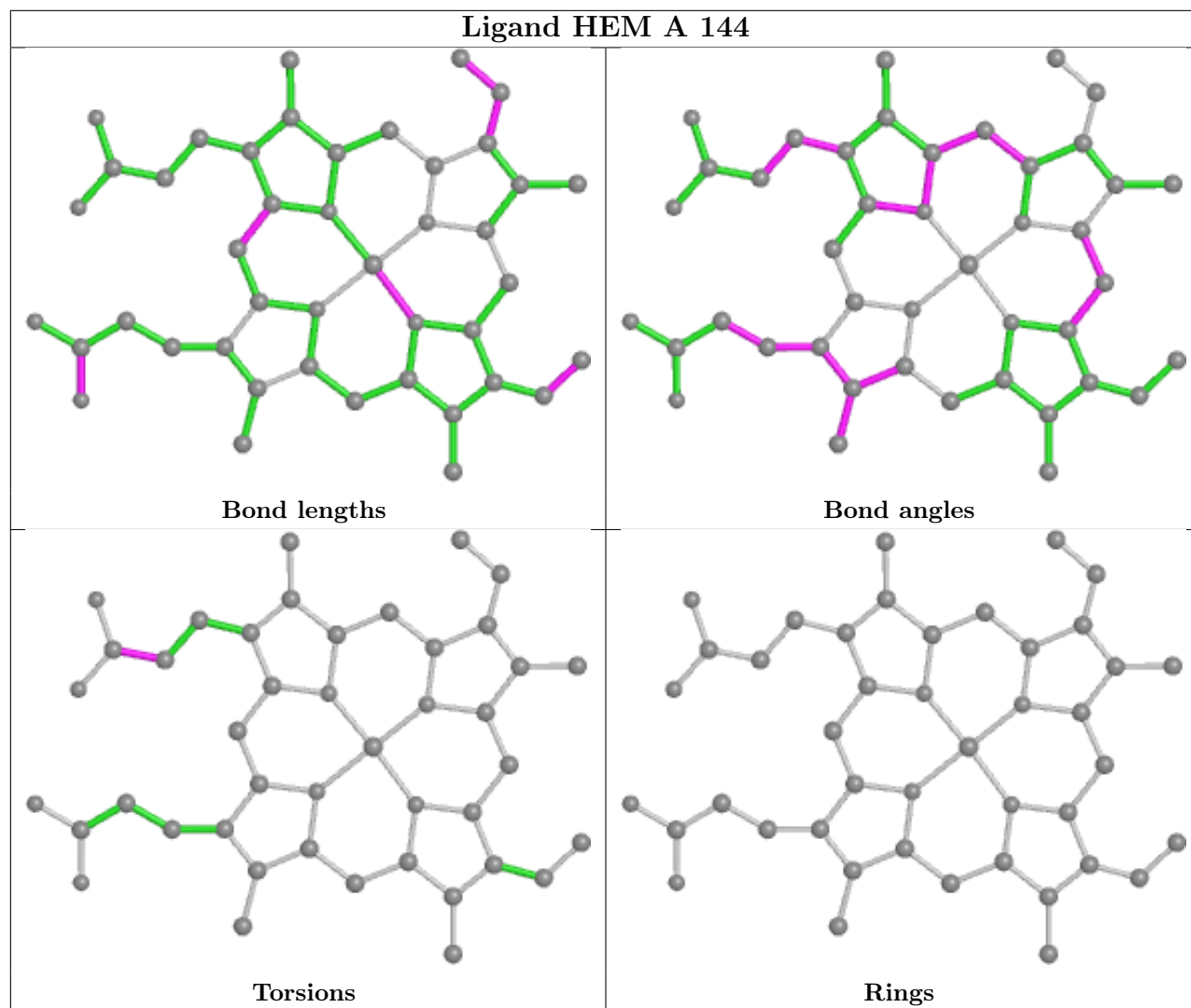
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	150	CYN	1	0

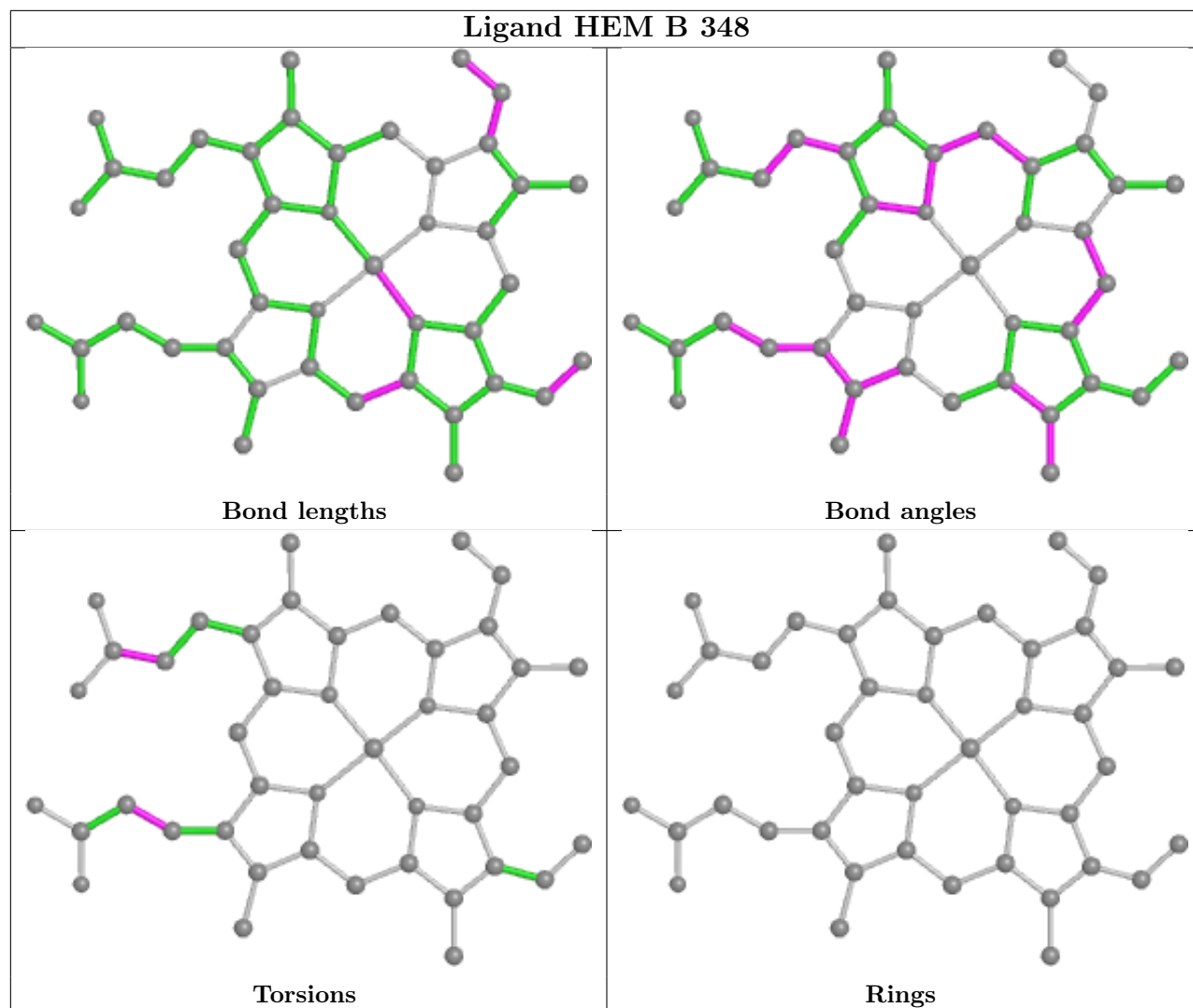
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In

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](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.