

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

#### Aug 8, 2020 - 02:58 AM BST

PDB ID	:	1HDS
$\operatorname{Title}$	:	MACROMOLECULAR STRUCTURE REFINEMENT BY RESTRAINED
		LEAST-SQUARES AND INTERACTIVE GRAPHICS AS APPLIED TO
		SICKLING DEER TYPE III HEMOGLOBIN
Authors	:	Amma, E.L.; Girling, R.L.
Deposited on		
Resolution	:	1.98  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:

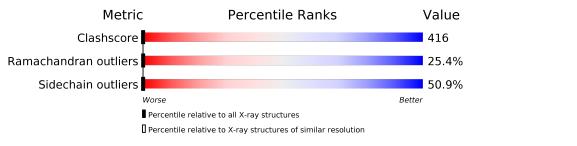
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
$\mathrm{EDS}$	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

# 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 1.98 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 $(\#Entries)$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)

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 >=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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain						
1	А	141	• 35%	63%					
1	С	141	• 38%	61%					
2	В	145	• 28%	70%					
2	D	145	• 28%	70%					

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	HEM	А	142	-	-	Х	-

Continued on next page...



Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	HEM	В	146	-	-	Х	-
3	HEM	С	142	-	-	Х	-
3	HEM	D	146	-	-	Х	-

Continued from previous page...



# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4556 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Δ	141	Total	С	Ν	Ο	S	0	0	0
	141	1076	684	199	192	1	0	0	0	
1	C	1 / 1	Total	С	Ν	0	S	0	0	0
		141	1076	684	199	192	1		0	0

• Molecule 1 is a protein called HEMOGLOBIN S (DEOXY) (ALPHA CHAIN).

Chain	Residue	Modelled	Actual	Comment	Reference
А	6	ASN	ASP	conflict	UNP P01972
А	27	GLN	GLU	conflict	UNP P01972
А	30	GLN	GLU	conflict	UNP P01972
А	55	GLN	VAL	conflict	UNP P01972
А	60	GLN	GLU	conflict	UNP P01972
А	70	GLN	VAL	conflict	UNP P01972
А	74	ASN	ASP	conflict	UNP P01972
А	82	ASN	ASP	conflict	UNP P01972
А	85	ASN	ASP	conflict	UNP P01972
А	94	ASN	ASP	conflict	UNP P01972
A	104	SER	THR	conflict	UNP P01972
А	115	THR	SER	conflict	UNP P01972
А	116	ASN	ASP	conflict	UNP P01972
А	124	ASN	SER	conflict	UNP P01972
A	126	ASN	ASP	conflict	UNP P01972
А	132	ASP	VAL	conflict	UNP P01972
С	6	ASN	ASP	conflict	UNP P01972
С	27	GLN	GLU	conflict	UNP P01972
С	30	GLN	GLU	conflict	UNP P01972
С	55	GLN	VAL	conflict	UNP P01972
С	60	GLN	GLU	conflict	UNP P01972
С	70	GLN	VAL	conflict	UNP P01972
С	74	ASN	ASP	conflict	UNP P01972
С	82	ASN	ASP	conflict	UNP P01972
С	85	ASN	ASP	conflict	UNP P01972

There are 32 discrepancies between the modelled and reference sequences:

Continued on next page...



1HDS
------

	Continuea from previous page									
Chain	Residue	Modelled	Actual	Comment	Reference					
С	94	ASN	ASP	$\operatorname{conflict}$	UNP P01972					
С	104	SER	THR	$\operatorname{conflict}$	UNP P01972					
С	115	THR	SER	$\operatorname{conflict}$	UNP P01972					
С	116	ASN	ASP	conflict	UNP P01972					
С	124	ASN	SER	conflict	UNP P01972					
С	126	ASN	ASP	conflict	UNP P01972					
С	132	ASP	VAL	$\operatorname{conflict}$	UNP P01972					

Continued from previous page...

• Molecule 2 is a protein called HEMOGLOBIN S (DEOXY) (BETA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
0	р	145	Total	С	Ν	Ο	$\mathbf{S}$	0	0	0
	2 B	140	1116	719	205	189	3	0	0	0
0	2 D	D 145	Total	С	Ν	0	S	0	0	0
			1116	719	205	189	3	0		U

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
В	18	ASP	ASN	conflict	UNP P02074
В	25	GLN	GLU	conflict	UNP P02074
В	42	GLN	GLU	conflict	UNP P02074
В	46	ASN	ASP	conflict	UNP P02074
В	55	ASN	GLY	conflict	UNP P02074
В	71	THR	SER	conflict	UNP P02074
В	72	GLN	GLU	conflict	UNP P02074
В	86	GLN	GLU	conflict	UNP P02074
В	89	GLY	GLU	conflict	UNP P02074
В	98	ASN	ASP	conflict	UNP P02074
В	100	GLN	GLU	conflict	UNP P02074
В	110	ALA	VAL	conflict	UNP P02074
В	111	LEU	VAL	conflict	UNP P02074
В	113	VAL	LEU	conflict	UNP P02074
В	120	GLN	GLU	conflict	UNP P02074
В	124	ASN	LEU	conflict	UNP P02074
В	128	LEU	ASP	conflict	UNP P02074
В	143	LYS	ARG	conflict	UNP P02074
D	18	ASP	ASN	conflict	UNP P02074
D	25	GLN	GLU	conflict	UNP P02074
D	42	GLN	GLU	conflict	UNP P02074
D	46	ASN	ASP	conflict	UNP P02074
D	55	ASN	GLY	$\operatorname{conflict}$	UNP P02074

Continued on next page...

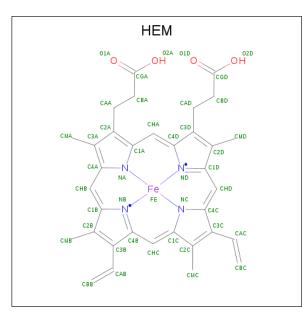


1HDS

Chain	Residue	Modelled	Actual	Comment	Reference
D	71	THR	SER	conflict	UNP P02074
D	72	GLN	GLU	conflict	UNP P02074
D	86	GLN	GLU	$\operatorname{conflict}$	UNP P02074
D	89	GLY	GLU	$\operatorname{conflict}$	UNP P02074
D	98	ASN	ASP	$\operatorname{conflict}$	UNP P02074
D	100	GLN	GLU	$\operatorname{conflict}$	UNP P02074
D	110	ALA	VAL	$\operatorname{conflict}$	UNP P02074
D	111	LEU	VAL	$\operatorname{conflict}$	UNP P02074
D	113	VAL	LEU	conflict	UNP P02074
D	120	GLN	GLU	$\operatorname{conflict}$	UNP P02074
D	124	ASN	LEU	conflict	UNP P02074
D	128	LEU	ASP	conflict	UNP P02074
D	143	LYS	ARG	conflict	UNP P02074

Continued from previous page...

• Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	Λ	1	Total	С	Fe	Ν	Ο	0	0
J	Л	I	43	34	1	4	4	0	0
3	В	1	Total	С	Fe	Ν	Ο	0	0
0	D	I	43	34	1	4	4	0	0
3	С	1	Total	С	Fe	Ν	Ο	0	0
0	U	I	43	34	1	4	4	0	0
3	Л	1	Total	С	Fe	Ν	Ο	0	0
0	D	T	43	34	1	4	4		0

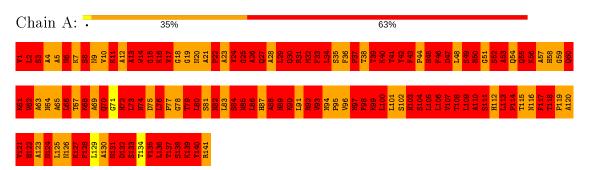


# 3 Residue-property plots (i)

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. 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.

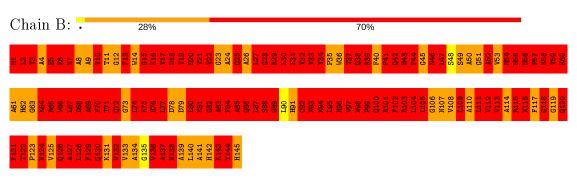
• Molecule 1: HEMOGLOBIN S (DEOXY) (ALPHA CHAIN)



• Molecule 1: HEMOGLOBIN S (DEOXY) (ALPHA CHAIN)

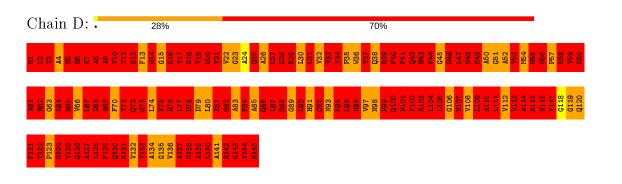
Chain C: •	38%	61%	-
V1 V1 S3 A4 A5 A5 S8 S8 V10 V10 A12 A12 A12 A12 C15	K16 V17 618 618 619 729 A21 A23 A24 A26 A26 A26 A26 A26 A26 A28 A28 A28 A28 A28 A28 A28 A28 A28 A28	F33 535 535 535 535 535 535 733 741 741 741 741 742 743 743 744 746 746 746 746 746 746 746 746 746	Q54 Q55 K56 A57 H58 G59 Q60
163 1 162 1 163 163 163 165 165 165 165 165 165 165 165 165 165	17.6 7.7 7.7 7.7 7.7 7.8 1.8 1.8 1.8 1.8 1.8 1.8 1.8 1.8 1.8 1	933 1934 1935 1935 1936 1937 1938 1938 1100 1100 1100 1100 1100 1100 1100 11	P114 T115 M116 F117 T118 P119 A120
V121 H122 A123 N126 N126 N126 F126 F126 F126 F128 N131 N131 T134 T134 V135	L136 1137 8138 1140 1141 1141		

• Molecule 2: HEMOGLOBIN S (DEOXY) (BETA CHAIN)



• Molecule 2: HEMOGLOBIN S (DEOXY) (BETA CHAIN)







# 4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	$163.50 \mathrm{\AA}$ 70.83 \mathrm{\AA} $65.95 \mathrm{\AA}$	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $94.15^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	(Not available) – 1.98	Depositor
% Data completeness	(Not available) ((Not available)-1.98)	Depositor
(in resolution range)		Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	$\operatorname{unknown}$	Depositor
$R, R_{free}$	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4556	wwPDB-VP
Average B, all atoms $(Å^2)$	13.0	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

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		
	Cham	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	10.87	704/1103~(63.8%)	9.83	796/1498~(53.1%)	
1	С	10.65	668/1103~(60.6%)	9.48	847/1498~(56.5%)	
2	В	11.08	695/1142~(60.9%)	10.07	841/1545~(54.4%)	
2	D	11.22	702/1142~(61.5%)	11.17	890/1545~(57.6%)	
All	All	10.96	2769/4490~(61.7%)	10.16	3374/6086~(55.4%)	

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	А	9	52
1	С	2	41
2	В	2	40
2	D	6	51
All	All	19	184

The worst 5 of 2769 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	104	SER	CB-OG	64.78	2.26	1.42
2	В	144	TYR	CE2-CZ	55.70	2.10	1.38
2	D	123	PRO	N-CD	53.15	2.22	1.47
2	В	32	VAL	CB-CG2	50.79	2.59	1.52
1	С	84	SER	CB-OG	49.53	2.06	1.42

The worst 5 of 3374 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	D	70	PHE	CB-CG-CD2	71.93	171.15	120.80
1	А	92	ARG	NE-CZ-NH2	-65.06	87.77	120.30
1	А	141	ARG	NE-CZ-NH1	61.72	151.16	120.30
2	D	44	PHE	CB-CG-CD2	60.92	163.44	120.80
2	D	84	PHE	CB-CG-CD1	-60.50	78.45	120.80

5 of 19 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	А	2	LEU	CA
1	А	31	ARG	CA
1	А	33	PHE	CA
1	А	93	VAL	CA
1	А	100	LEU	CA

5 of 184 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	А	1	VAL	Mainchain,Peptide
1	А	15	GLY	Peptide
1	А	22	PRO	Mainchain
1	А	3	SER	Mainchain
1	А	8	SER	Mainchain,Peptide

#### 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	А	1076	0	1032	834	80
1	С	1076	0	1034	805	3
2	В	1116	0	1070	1126	0
2	D	1116	0	1068	917	77
3	А	43	0	30	49	0
3	В	43	0	30	56	0
3	С	43	0	30	42	0
3	D	43	0	30	29	0
All	All	4556	0	4324	3698	80



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 416.

The worst 5 of 3698 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)
2:B:34:TYR:CZ	2:B:34:TYR:CE2	1.79	1.70
2:B:144:TYR:CG	2:B:144:TYR:CD2	1.76	1.69
3:A:142:HEM:CMA	3:A:142:HEM:C3A	1.75	1.67
1:C:140:TYR:CB	1:C:140:TYR:CG	1.77	1.67
2:D:14:TRP:CZ3	2:D:14:TRP:CE3	1.76	1.66

The worst 5 of 80 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)
1:A:2:LEU:O	2:D:49:SER:CB[4_555]	0.51	1.69
1:A:75:ASP:C	$2:D:55:ASN:CG[4_555]$	0.81	1.39
1:A:74:ASN:CG	2:D:54:MET:CA[4_555]	0.92	1.28
1:A:2:LEU:O	2:D:49:SER:CA[4_555]	1.05	1.15
1:A:7:LYS:NZ	2:D:51:GLY:N[4_555]	1.06	1.14

### 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	А	139/141~(99%)	74~(53%)	34~(24%)	31~(22%)	0 0
1	С	139/141~(99%)	78~(56%)	28 (20%)	33~(24%)	0 0
2	В	143/145~(99%)	75~(52%)	28~(20%)	40~(28%)	0 0
2	D	143/145~(99%)	78 (54%)	26~(18%)	39~(27%)	0 0
All	All	564/572~(99%)	305~(54%)	116~(21%)	143~(25%)	0 0

5 of 143 Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	А	2	LEU
1	А	6	ASN
1	А	13	ALA
1	А	15	GLY
1	А	17	VAL

#### 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	А	115/115~(100%)	57~(50%)	58~(50%)		0	0
1	С	115/115~(100%)	64~(56%)	51 (44%)		0	0
2	В	113/113~(100%)	50 (44%)	63~(56%)		0	0
2	D	113/113~(100%)	53~(47%)	60~(53%)		0	0
All	All	456/456~(100%)	224 (49%)	232~(51%)		0	0

5 of 232 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	В	104	LEU
1	С	20	ASN
2	D	109	LEU
2	В	112	VAL
2	В	129	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	В	145	HIS
1	С	60	GLN
2	D	76	HIS
1	С	45	HIS
1	С	70	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 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	Aol Type Chain		n Bog	Pos	Res	Bos	Dog	Dog	Dec	Dog	Dec	Link	B	ond leng	gths	B	ond ang	gles
	туре	Chain	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2								
3	HEM	В	146	2	27,50,50	10.55	22 (81%)	17,82,82	8.98	10 (58%)								
3	HEM	D	146	-	27,50,50	10.88	23 (85%)	17,82,82	9.27	15 (88%)								
3	HEM	А	142	1	27,50,50	<mark>9.67</mark>	21 (77%)	17,82,82	11.88	15 (88%)								
3	HEM	С	142	1	27,50,50	<mark>8.96</mark>	20 (74%)	17,82,82	<mark>6.87</mark>	11 (64%)								

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
3	HEM	В	146	2	-	2/6/54/54	-
3	HEM	D	146	-	-	3/6/54/54	-
3	HEM	А	142	1	-	2/6/54/54	-
3	HEM	С	142	1	-	1/6/54/54	-



Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
3	D	146	HEM	CAD-C3D	-27.74	1.03	1.52
3	В	146	HEM	C4D-C3D	23.91	1.96	1.42
3	В	146	HEM	C3B-C2B	-23.74	1.07	1.40
3	D	146	HEM	CMD-C2D	-23.16	1.03	1.51
3	В	146	HEM	CMA-C3A	20.85	1.95	1.51

The worst 5 of 86 bond length outliers are listed below:

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
3	А	142	HEM	CBD-CAD-C3D	27.20	162.59	112.48
3	D	146	HEM	C4A-C3A-C2A	25.96	125.06	107.00
3	А	142	HEM	C4C-C3C-C2C	-24.18	90.01	106.90
3	В	146	HEM	CMA-C3A-C4A	-20.36	97.18	128.46
3	В	146	HEM	CMC-C2C-C3C	17.94	158.24	124.68

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type	Atoms	
3	В	146	HEM	C2D-C3D-CAD-CBD	
3	В	146	HEM	C4D-C3D-CAD-CBD	
3	D	146	HEM	C1A-C2A-CAA-CBA	
3	D	146	HEM	C3A-C2A-CAA-CBA	
3	D	146	HEM	C3D-CAD-CBD-CGD	

There are no ring outliers.

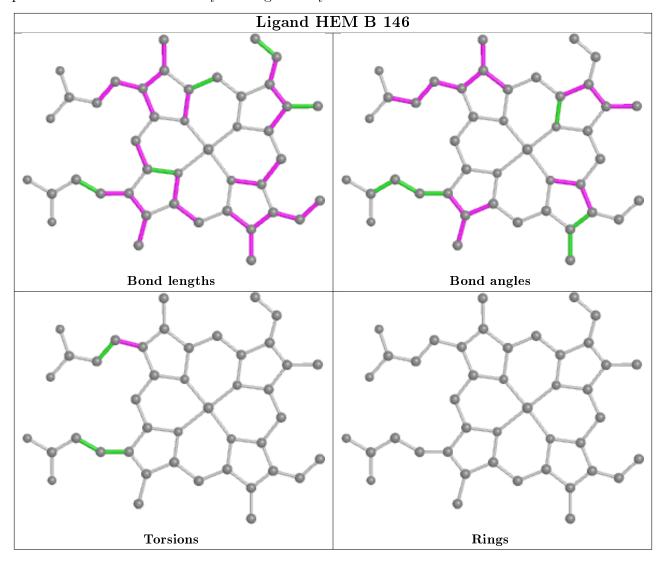
4 monomers are involved in 176 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	В	146	HEM	56	0
3	D	146	HEM	29	0
3	А	142	HEM	49	0
3	С	142	HEM	42	0

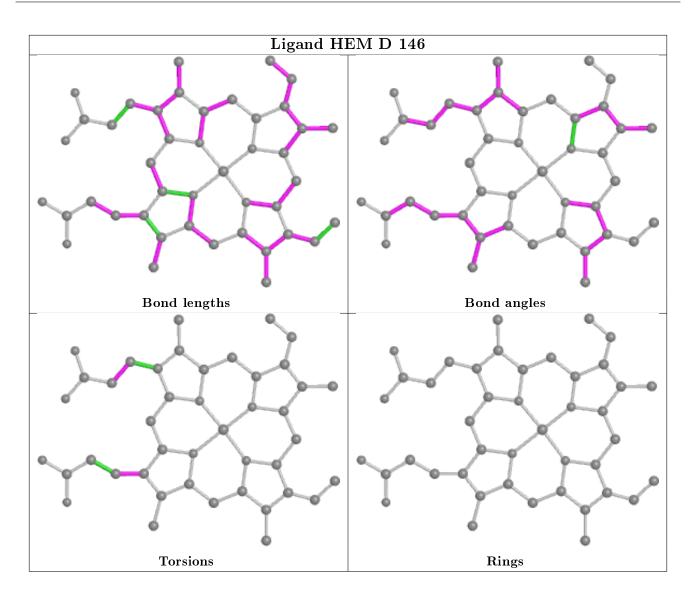
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

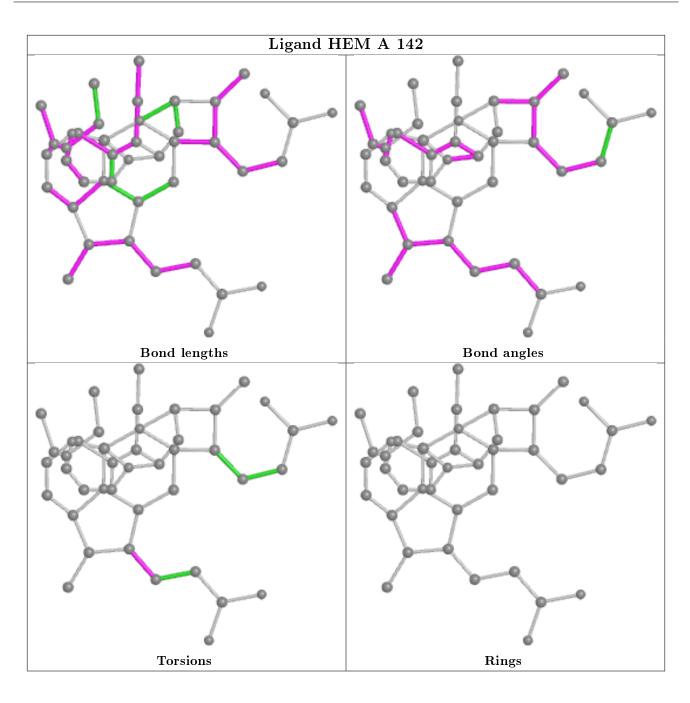






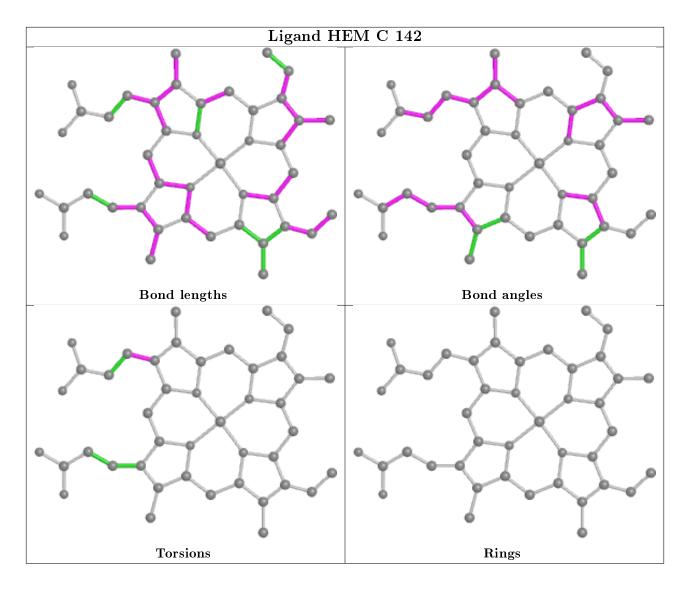












## 5.7 Other polymers (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	D	53
1	А	51
2	В	47
1	С	47

The worst 5 of 198 chain breaks are listed below:



Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	С	98:PHE	С	99:LYS	Ν	1.99
1	В	42:GLN	С	43:HIS	Ν	1.93
1	D	51:GLY	С	52:ALA	N	1.87
1	В	41:PHE	С	42:GLN	Ν	1.85
1	С	74:ASN	С	75:ASP	Ν	1.84



## 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.

