



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:49 PM GMT

PDB ID : 1HDS
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 : 1979-10-01
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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

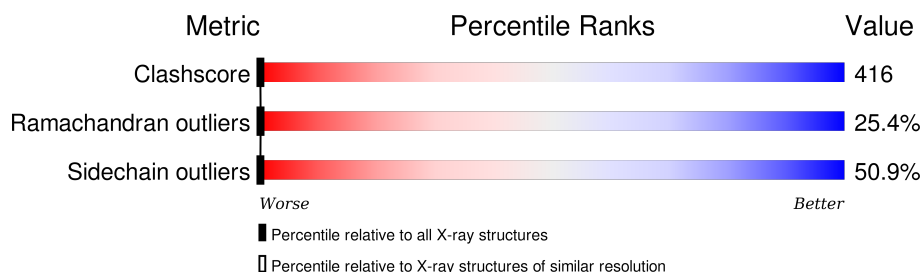
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.98 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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	9905 (2.00-1.96)
Ramachandran outliers	100387	9792 (2.00-1.96)
Sidechain outliers	100360	9791 (2.00-1.96)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	141	
1	C	141	
2	B	145	
2	D	145	

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	A	142	-	-	X	-
3	HEM	B	146	-	-	X	-
3	HEM	C	142	-	-	X	-
3	HEM	D	146	-	-	X	-

2 Entry composition

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.

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	141	Total	C	N	O	S	0	0	0
			1076	684	199	192	1			
1	C	141	Total	C	N	O	S	0	0	0
			1076	684	199	192	1			

There are 32 discrepancies between the modelled and reference sequences:

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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	94	ASN	ASP	CONFLICT	UNP P01972
C	104	SER	THR	CONFLICT	UNP P01972
C	115	THR	SER	CONFLICT	UNP P01972
C	116	ASN	ASP	CONFLICT	UNP P01972
C	124	ASN	SER	CONFLICT	UNP P01972
C	126	ASN	ASP	CONFLICT	UNP P01972
C	132	ASP	VAL	CONFLICT	UNP P01972

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	145	Total	C	N	O	S	0	0	0
			1116	719	205	189	3			
2	D	145	Total	C	N	O	S	0	0	0
			1116	719	205	189	3			

There are 36 discrepancies between the modelled and reference sequences:

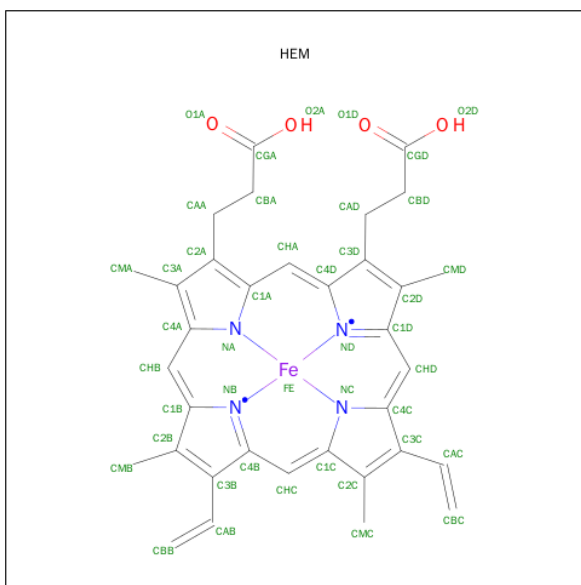
Chain	Residue	Modelled	Actual	Comment	Reference
B	18	ASP	ASN	CONFLICT	UNP P02074
B	25	GLN	GLU	CONFLICT	UNP P02074
B	42	GLN	GLU	CONFLICT	UNP P02074
B	46	ASN	ASP	CONFLICT	UNP P02074
B	55	ASN	GLY	CONFLICT	UNP P02074
B	71	THR	SER	CONFLICT	UNP P02074
B	72	GLN	GLU	CONFLICT	UNP P02074
B	86	GLN	GLU	CONFLICT	UNP P02074
B	89	GLY	GLU	CONFLICT	UNP P02074
B	98	ASN	ASP	CONFLICT	UNP P02074
B	100	GLN	GLU	CONFLICT	UNP P02074
B	110	ALA	VAL	CONFLICT	UNP P02074
B	111	LEU	VAL	CONFLICT	UNP P02074
B	113	VAL	LEU	CONFLICT	UNP P02074
B	120	GLN	GLU	CONFLICT	UNP P02074
B	124	ASN	LEU	CONFLICT	UNP P02074
B	128	LEU	ASP	CONFLICT	UNP P02074
B	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	CONFLICT	UNP P02074

Continued on next page...

Continued from previous page...

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

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



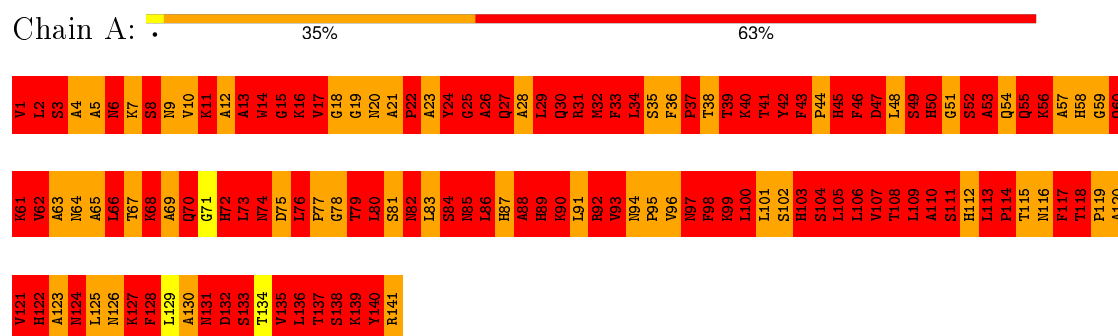
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	D	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

3 Residue-property plots

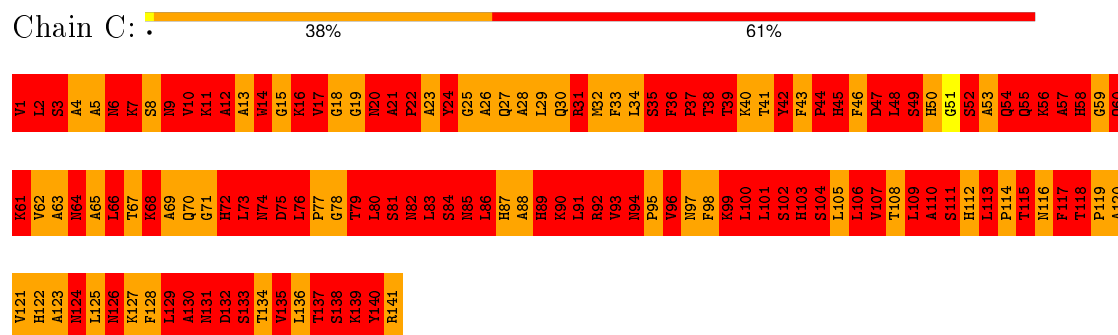
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

Note EDS was not executed.

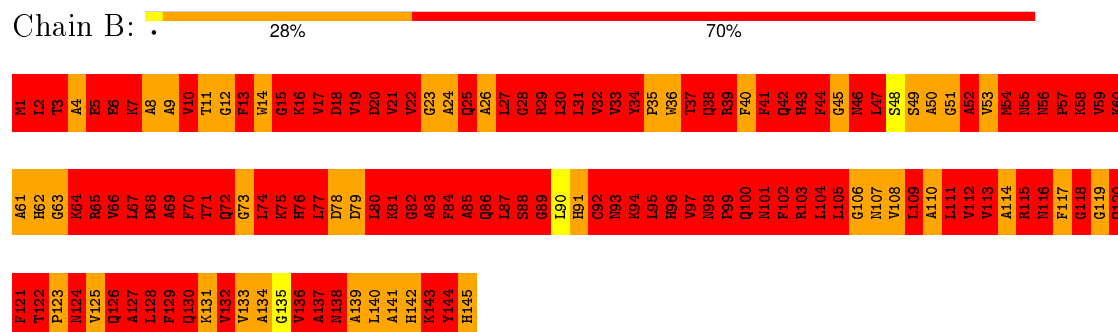
• Molecule 1: HEMOGLOBIN S (DEOXY) (ALPHA CHAIN)



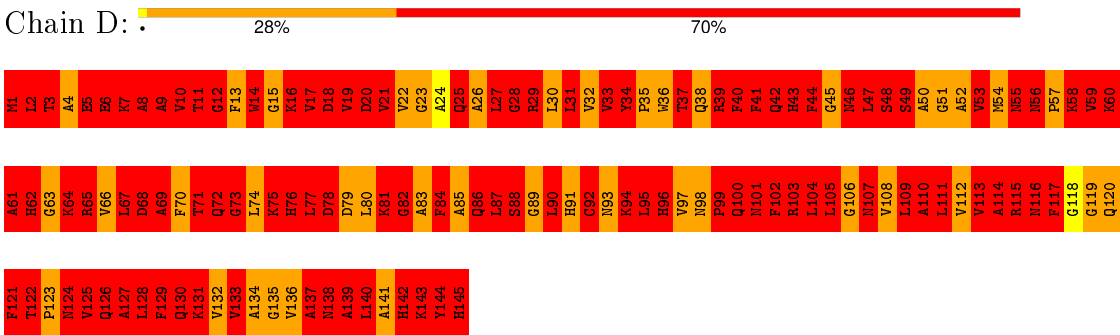
• Molecule 1: HEMOGLOBIN S (DEOXY) (ALPHA CHAIN)



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



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



4 Data and refinement statistics

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	163.50Å 70.83Å 65.95Å 90.00° 94.15° 90.00°	Depositor
Resolution (Å)	(Not available) – 1.98	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-1.98)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	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

5.1 Standard geometry

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	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	10.87	704/1103 (63.8%)	9.83	796/1498 (53.1%)
1	C	10.65	668/1103 (60.6%)	9.48	847/1498 (56.5%)
2	B	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	A	9	52
1	C	2	41
2	B	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	A	104	SER	CB-OG	64.78	2.26	1.42
2	B	144	TYR	CE2-CZ	55.70	2.10	1.38
2	D	123	PRO	N-CD	53.15	2.22	1.47
2	B	32	VAL	CB-CG2	50.79	2.59	1.52
1	C	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($^{\circ}$)	Ideal($^{\circ}$)
2	D	70	PHE	CB-CG-CD2	71.93	171.15	120.80
1	A	92	ARG	NE-CZ-NH2	-65.06	87.77	120.30
1	A	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	A	2	LEU	CA
1	A	31	ARG	CA
1	A	33	PHE	CA
1	A	93	VAL	CA
1	A	100	LEU	CA

5 of 184 planarity outliers are listed below:

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

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	1076	0	1032	834	80
1	C	1076	0	1034	805	3
2	B	1116	0	1070	1126	0
2	D	1116	0	1068	917	77
3	A	43	0	30	49	0
3	B	43	0	30	56	0
3	C	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	A	139/141 (99%)	74 (53%)	34 (24%)	31 (22%)	0	0
1	C	139/141 (99%)	78 (56%)	28 (20%)	33 (24%)	0	0
2	B	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	A	2	LEU
1	A	6	ASN
1	A	13	ALA
1	A	15	GLY
1	A	17	VAL

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	115/115 (100%)	57 (50%)	58 (50%)	0	0
1	C	115/115 (100%)	64 (56%)	51 (44%)	0	0
2	B	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	B	104	LEU
1	C	20	ASN
2	D	109	LEU
2	B	112	VAL
2	B	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	B	145	HIS
1	C	60	GLN
2	D	76	HIS
1	C	45	HIS
1	C	70	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	HEM	A	142	1	30,50,50	9.59	25 (83%)	24,82,82	9.04	20 (83%)
3	HEM	B	146	2	30,50,50	9.63	21 (70%)	24,82,82	6.56	13 (54%)
3	HEM	C	142	1	30,50,50	12.89	25 (83%)	24,82,82	5.66	18 (75%)
3	HEM	D	146	-	30,50,50	10.73	22 (73%)	24,82,82	8.12	20 (83%)

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	A	142	1	-	0/10/54/54	0/0/8/8
3	HEM	B	146	2	-	0/10/54/54	0/0/8/8
3	HEM	C	142	1	-	0/10/54/54	0/0/8/8
3	HEM	D	146	-	-	0/10/54/54	0/0/8/8

The worst 5 of 93 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	142	HEM	C3B-C4B	-35.64	1.20	1.51
3	C	142	HEM	C3D-C4D	-30.00	1.13	1.51
3	D	146	HEM	C3B-C4B	-28.32	1.26	1.51
3	D	146	HEM	CAD-C3D	-25.24	1.03	1.54
3	D	146	HEM	CMD-C2D	-21.40	1.03	1.53

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	146	HEM	CAA-C2A-C1A	-19.79	105.52	127.01
3	A	142	HEM	C3B-CAB-CBB	-18.94	95.41	124.46
3	B	146	HEM	CMA-C3A-C4A	-18.85	97.18	128.36
3	D	146	HEM	C3C-CAC-CBC	-14.12	102.81	124.46
3	D	146	HEM	CMA-C3A-C4A	-12.29	108.03	128.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 176 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	142	HEM	49	0
3	B	146	HEM	56	0
3	C	142	HEM	42	0
3	D	146	HEM	29	0

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.