wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:52 PM GMT

PDB ID	:	1MHL
Title	:	CRYSTAL STRUCTURE OF HUMAN MYELOPEROXIDASE ISOFORM C CRYSTALLIZED IN SPACE GROUP P2(1) AT PH 5.5 AND 20 DEG C
Authors	:	Fenna, R.E.; Zeng, J.; Davey, C.
Deposited on	:	1995-06-09
Resolution	:	2.25 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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

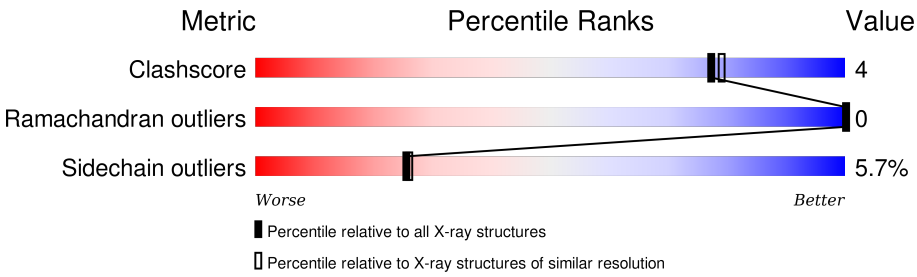
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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	1095 (2.26-2.26)
Ramachandran outliers	100387	1063 (2.26-2.26)
Sidechain outliers	100360	1063 (2.26-2.26)

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 <=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	108	75% 19%
1	B	108	81% 13%
2	C	466	79% 18%
2	D	466	83% 14%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 9847 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.

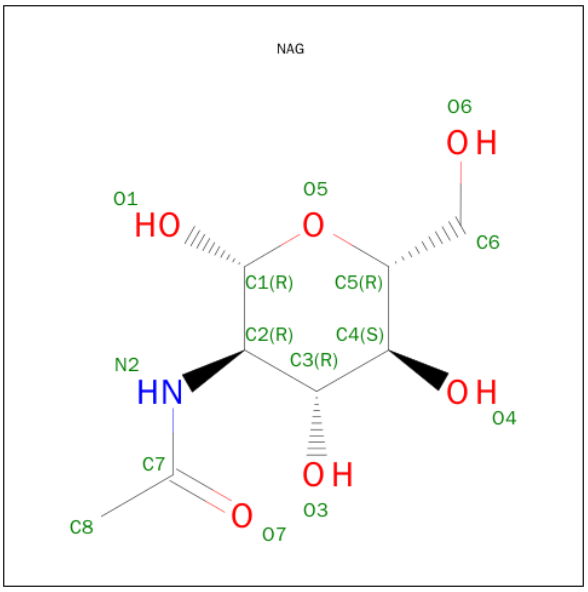
- Molecule 1 is a protein called MYELOPEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	104	837	529	148	155	5	12	0	0
1	B	104	837	529	148	155	5	7	0	0

- Molecule 2 is a protein called MYELOPEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	466	3732	2351	687	667	27	82	0	0
2	D	466	3732	2351	687	667	27	109	0	0

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	14	8	1	5	0	0
			Total	C	N	O		
3	A	1	14	8	1	5	0	0
			Total	C	N	O		
3	B	1	14	8	1	5	0	0
			Total	C	N	O		
3	B	1	14	8	1	5	0	0

- Molecule 4 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	6	71	40	2	29	0	0
			Total	C	N	O		
4	B	6	71	40	2	29	0	0

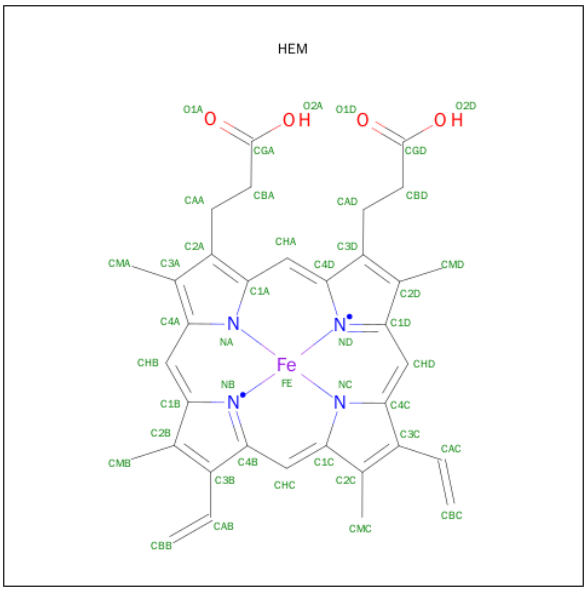
- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	B	1	1	1	0	0
			Total	Ca		
5	A	1	1	1	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
6	B	1	1	1	0	0
			Total	Cl		
6	A	1	1	1	0	0

- Molecule 7 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



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

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	55	55	55	0	0
			Total	O		
8	B	58	58	58	0	0
			Total	O		
8	C	158	158	158	0	0
			Total	O		
8	D	150	150	150	0	0

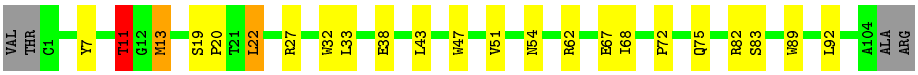
3 Residue-property plots (i)

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.

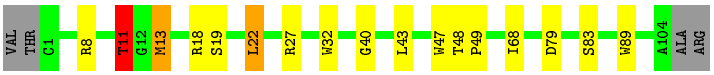
- Molecule 1: MYELOPEROXIDASE

Chain A:



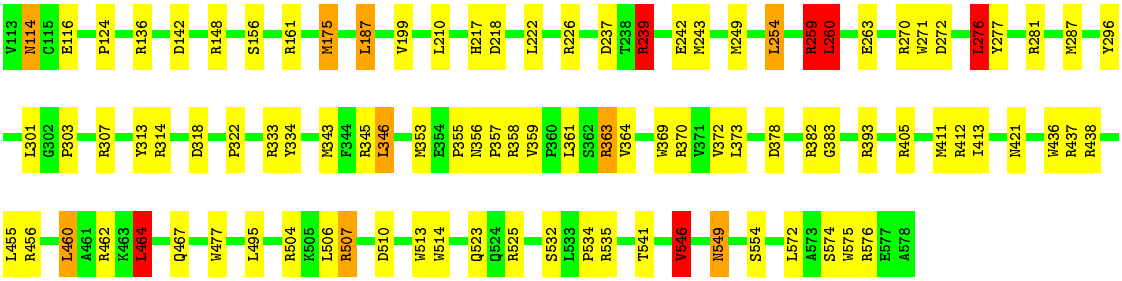
- Molecule 1: MYELOPEROXIDASE

Chain B:



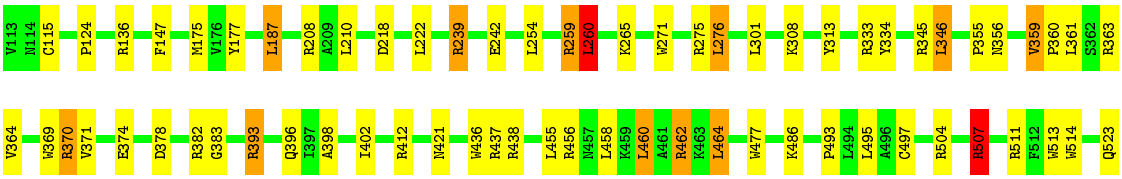
- Molecule 2: MYELOPEROXIDASE

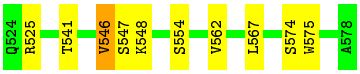
Chain C:



- Molecule 2: MYELOPEROXIDASE

Chain D:




Q524
R525
T541
V546
S547
K548
S554
V562
L567
S574
W575
A578

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	111.70Å 64.60Å 94.20Å 90.00° 97.90° 90.00°	Depositor
Resolution (Å)	8.00 – 2.25	Depositor
% Data completeness (in resolution range)	92.6 (8.00-2.25)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.0	Depositor
R, R_{free}	0.160 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9847	wwPDB-VP
Average B, all atoms (Å^2)	21.0	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: BMA, NAG, CL, CA, FUC, HEM, MAN

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	0.82	0/862	1.62	14/1174 (1.2%)
1	B	0.77	0/862	1.59	14/1174 (1.2%)
2	C	0.78	0/3818	1.58	76/5180 (1.5%)
2	D	0.75	0/3818	1.57	53/5180 (1.0%)
All	All	0.77	0/9360	1.58	157/12708 (1.2%)

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
2	C	0	1
2	D	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 157 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	507	ARG	NE-CZ-NH2	-17.73	111.44	120.30
2	D	239	ARG	NE-CZ-NH1	17.67	129.13	120.30
2	D	507	ARG	NE-CZ-NH1	15.93	128.26	120.30
2	C	507	ARG	NE-CZ-NH1	14.17	127.38	120.30
2	D	345	ARG	NE-CZ-NH2	-13.97	113.31	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	334	TYR	Sidechain
2	D	334	TYR	Sidechain

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	837	0	798	12	0
1	B	837	0	798	7	0
2	C	3732	0	3725	31	0
2	D	3732	0	3725	25	0
3	A	28	0	26	0	0
3	B	28	0	26	0	0
4	A	71	0	61	1	0
4	B	71	0	61	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	43	0	30	2	0
7	B	43	0	30	1	0
8	A	55	0	0	0	0
8	B	58	0	0	0	0
8	C	158	0	0	0	0
8	D	150	0	0	0	0
All	All	9847	0	9280	68	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 4.

The worst 5 of 68 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:19:SER:HB3	1:A:22:LEU:HD22	1.67	0.76
2:C:333:ARG:HH11	2:C:421:ASN:HD22	1.34	0.74
2:C:460:LEU:HD22	2:C:464:LEU:HD22	1.74	0.68
1:A:67:GLU:HG3	2:C:467:GLN:NE2	2.10	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:68:ILE:HD13	2:D:464:LEU:HD13	1.78	0.63

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	102/108 (94%)	99 (97%)	3 (3%)	0	100	100
1	B	102/108 (94%)	98 (96%)	4 (4%)	0	100	100
2	C	464/466 (100%)	450 (97%)	14 (3%)	0	100	100
2	D	464/466 (100%)	455 (98%)	9 (2%)	0	100	100
All	All	1132/1148 (99%)	1102 (97%)	30 (3%)	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	90/93 (97%)	87 (97%)	3 (3%)	45	56
1	B	90/93 (97%)	86 (96%)	4 (4%)	35	40
2	C	411/411 (100%)	387 (94%)	24 (6%)	25	25
2	D	411/411 (100%)	385 (94%)	26 (6%)	22	21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1002/1008 (99%)	945 (94%)	57 (6%)	25 26

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	546	VAL
2	D	124	PRO
2	D	523	GLN
2	C	574	SER
1	B	13	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	549	ASN
2	D	549	ASN
1	B	80	GLN
2	C	421	ASN
1	B	54	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)

12 carbohydrates 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	NAG	A	640	2,4	14,14,15	0.57	0	15,19,21	1.08	1 (6%)
4	NAG	A	641	4	14,14,15	0.87	0	15,19,21	0.88	1 (6%)
4	BMA	A	642	4	11,11,12	0.70	0	14,15,17	0.75	0
4	MAN	A	643	4	11,11,12	0.54	0	14,15,17	1.18	1 (7%)
4	MAN	A	644	4	11,11,12	0.61	0	14,15,17	1.02	1 (7%)
4	FUC	A	645	4	10,10,11	0.65	0	14,14,16	1.33	2 (14%)
4	NAG	B	640	2,4	14,14,15	0.69	0	15,19,21	1.09	0
4	NAG	B	641	4	14,14,15	0.40	0	15,19,21	0.77	0
4	BMA	B	642	4	11,11,12	0.36	0	14,15,17	0.98	0
4	MAN	B	643	4	11,11,12	0.97	0	14,15,17	1.42	3 (21%)
4	MAN	B	644	4	11,11,12	0.75	0	14,15,17	1.59	2 (14%)
4	FUC	B	645	4	10,10,11	0.53	0	14,14,16	1.28	1 (7%)

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	NAG	A	640	2,4	-	0/6/23/26	0/1/1/1
4	NAG	A	641	4	-	0/6/23/26	0/1/1/1
4	BMA	A	642	4	-	0/2/19/22	0/1/1/1
4	MAN	A	643	4	-	0/2/19/22	0/1/1/1
4	MAN	A	644	4	-	0/2/19/22	0/1/1/1
4	FUC	A	645	4	-	0/0/17/20	0/1/1/1
4	NAG	B	640	2,4	-	0/6/23/26	0/1/1/1
4	NAG	B	641	4	-	0/6/23/26	0/1/1/1
4	BMA	B	642	4	-	0/2/19/22	0/1/1/1
4	MAN	B	643	4	-	0/2/19/22	0/1/1/1
4	MAN	B	644	4	-	0/2/19/22	0/1/1/1
4	FUC	B	645	4	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	644	MAN	C6-C5-C4	-2.50	106.84	113.02
4	A	645	FUC	C6-C5-C4	-2.34	108.48	113.08
4	A	641	NAG	C4-C3-C2	-2.24	107.74	111.23
4	A	645	FUC	O3-C3-C4	-2.11	105.58	110.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	643	MAN	C3-C4-C5	-2.08	106.56	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	643	MAN	1	0

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	$\#\|Z\| > 2$
7	HEM	A	605	1,8,2	30,50,50	2.66	9 (30%)	24,82,82	2.10	6 (25%)
3	NAG	A	620	2	14,14,15	0.78	1 (7%)	15,19,21	0.88	1 (6%)
3	NAG	A	630	2	14,14,15	1.22	2 (14%)	15,19,21	1.93	6 (40%)
7	HEM	B	605	1,8,2	30,50,50	2.89	11 (36%)	24,82,82	2.15	6 (25%)
3	NAG	B	620	2	14,14,15	0.94	1 (7%)	15,19,21	1.62	2 (13%)
3	NAG	B	630	2	14,14,15	1.08	1 (7%)	15,19,21	1.38	3 (20%)

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
7	HEM	A	605	1,8,2	-	0/10/54/54	0/0/8/8
3	NAG	A	620	2	-	0/6/23/26	0/1/1/1
3	NAG	A	630	2	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	HEM	B	605	1,8,2	-	0/10/54/54	0/0/8/8
3	NAG	B	620	2	-	0/6/23/26	0/1/1/1
3	NAG	B	630	2	-	0/6/23/26	0/1/1/1

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	605	HEM	C3B-C4B	-8.29	1.44	1.51
7	A	605	HEM	C3B-C4B	-7.34	1.45	1.51
7	B	605	HEM	C2D-C3D	-6.16	1.36	1.54
7	A	605	HEM	C2D-C3D	-6.09	1.36	1.54
7	A	605	HEM	C3B-CAB	-5.94	1.40	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	630	NAG	C3-C4-C5	-3.56	103.98	110.20
3	A	630	NAG	O4-C4-C3	-3.03	103.52	110.34
3	B	630	NAG	C2-N2-C7	-2.98	119.20	123.04
3	B	620	NAG	C2-N2-C7	-2.48	119.85	123.04
3	B	630	NAG	C3-C4-C5	-2.28	106.22	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	605	HEM	2	0
7	B	605	HEM	1	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 (i)

6.1 Protein, DNA and RNA chains (i)

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

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

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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