wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:59 PM GMT

PDB ID	:	1MYP
Title	:	X-RAY CRYSTAL STRUCTURE OF CANINE MYELOPEROXIDASE AT 3 ANGSTROMS RESOLUTION
Authors	:	Fenna, R.E.; Zeng, J.
Deposited on	:	1992-04-15
Resolution	:	3.00 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:

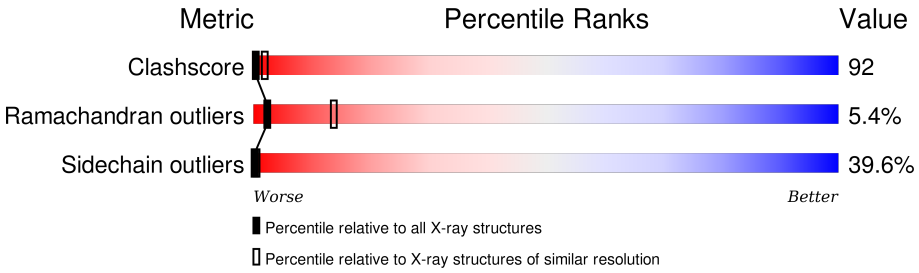
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 3.00 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	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)

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	22% 32% 28% 14%
1	B	108	19% 35% 27% 16%
2	C	466	13% 40% 35% 12%
2	D	466	16% 39% 33% 11%

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	NAG	A	586	X	-	-	-
3	NAG	B	586	X	-	-	-
5	HEM	A	580	-	-	X	-
5	HEM	B	580	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9134 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
1	A	104	Total	C	N	O	S	0	0	0
			830	525	147	153	5			
1	B	104	Total	C	N	O	S	0	0	0
			830	525	147	153	5			

- Molecule 2 is a protein called MYELOPEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	462	Total	C	N	O	S	0	0	0
			3609	2289	648	647	25			
2	D	462	Total	C	N	O	S	0	0	0
			3609	2289	648	647	25			

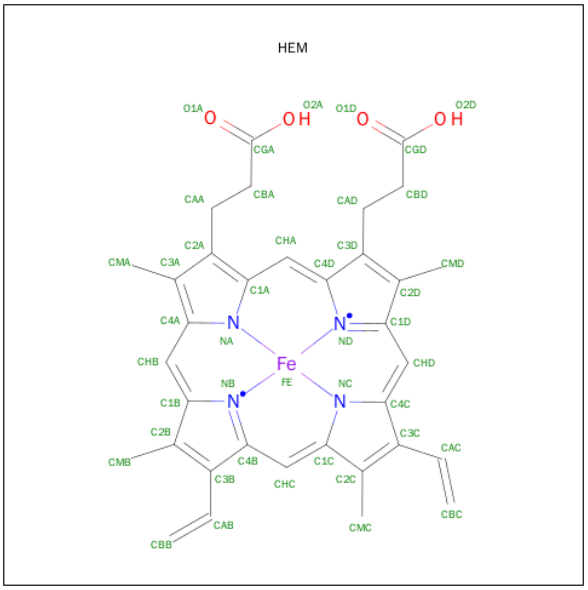
- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	B	2	Total	C	N	O	0	0
			28	16	2	10		
3	B	2	Total	C	N	O	0	0
			28	16	2	10		
3	B	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 5 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



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

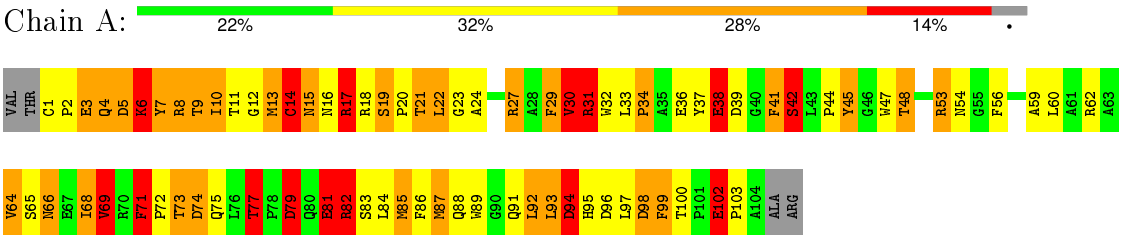
3 Residue-property plots (i)

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.

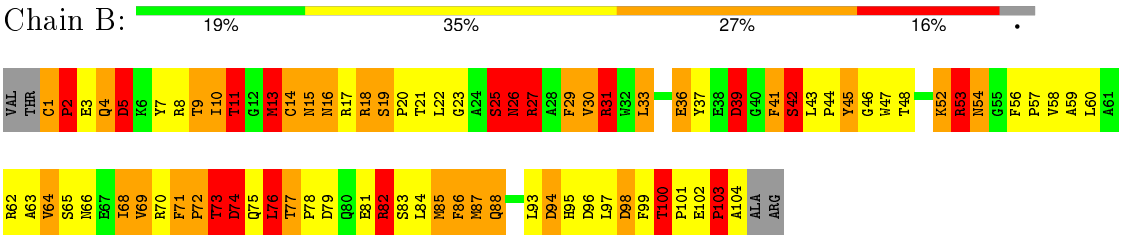
- Molecule 1: MYELOPEROXIDASE

Chain A:



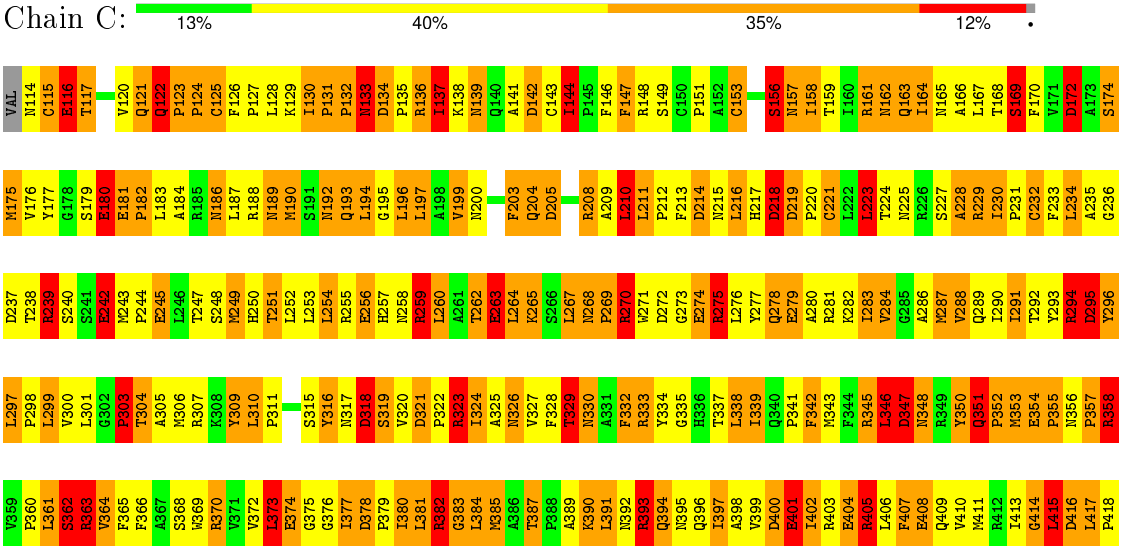
- Molecule 1: MYELOPEROXIDASE

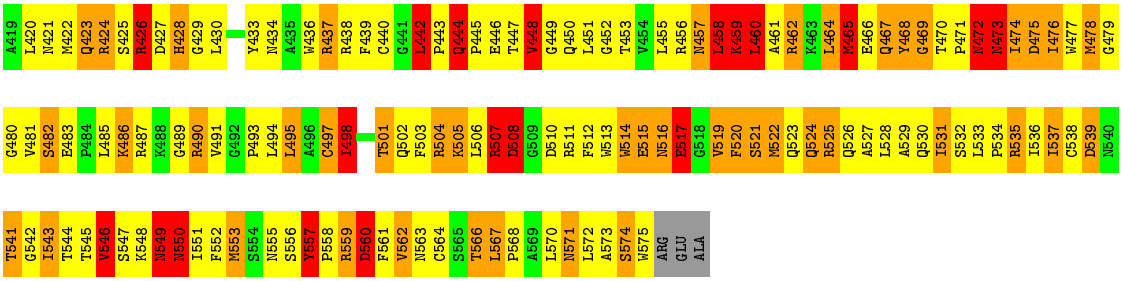
Chain B:



- Molecule 2: MYELOPEROXIDASE

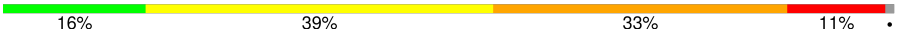
Chain C:

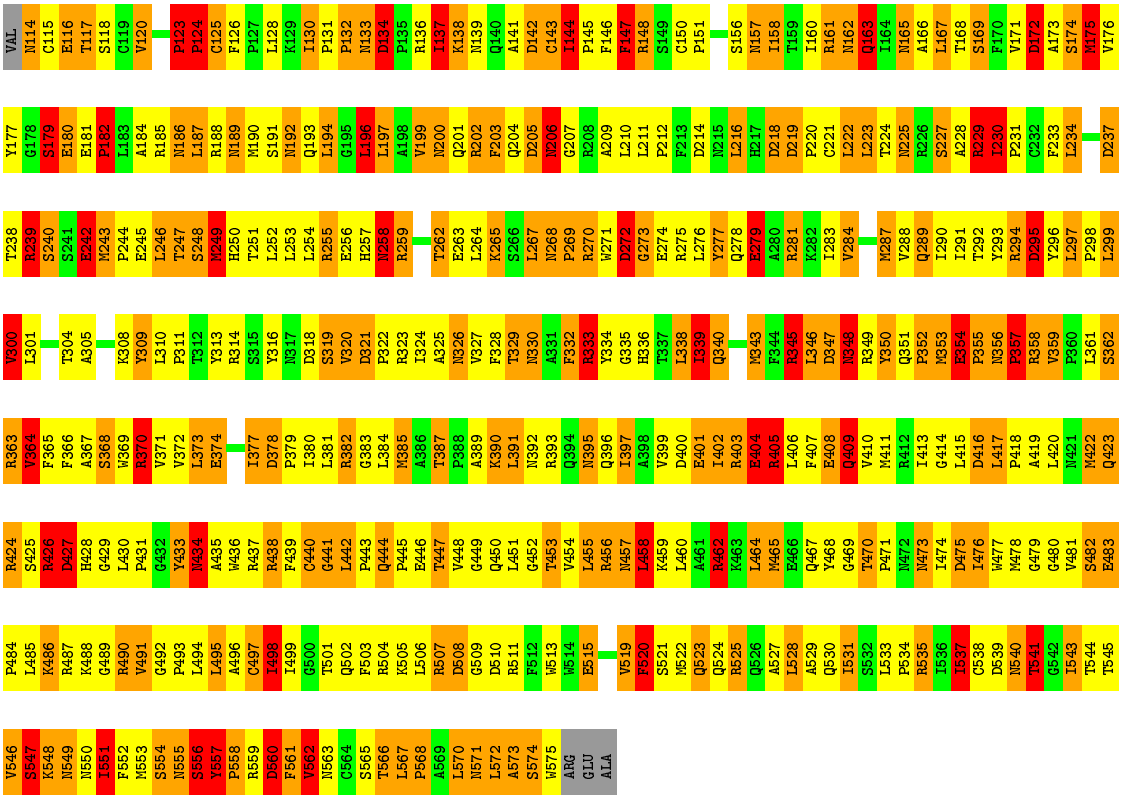




- Molecule 2: MYELOPEROXIDASE

Chain D:





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 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	133.00Å 133.00Å 203.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-3.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	TNT	Depositor
R, R_{free}	0.257 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9134	wwPDB-VP
Average B, all atoms (Å^2)	15.0	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, CA, NAG

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	0.86	5/855 (0.6%)	1.98	27/1166 (2.3%)
1	B	1.30	12/854 (1.4%)	2.10	44/1163 (3.8%)
2	C	1.03	28/3695 (0.8%)	2.16	158/5030 (3.1%)
2	D	1.13	26/3695 (0.7%)	2.15	163/5030 (3.2%)
All	All	1.08	71/9099 (0.8%)	2.13	392/12389 (3.2%)

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
1	A	2	3
1	B	0	4
2	C	6	12
2	D	5	13
3	A	1	0
3	B	1	0
All	All	15	32

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	560	ASP	C-N	26.88	1.95	1.34
2	D	179	SER	C-N	20.67	1.81	1.34
2	D	557	TYR	C-N	20.42	1.73	1.34
2	D	560	ASP	C-N	19.45	1.78	1.34
2	D	203	PHE	C-N	17.63	1.74	1.34

The worst 5 of 392 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	560	ASP	O-C-N	-21.52	88.26	122.70
2	C	560	ASP	O-C-N	-18.81	92.61	122.70
2	C	218	ASP	CB-CG-OD1	-18.33	101.80	118.30
2	D	202	ARG	O-C-N	-16.89	95.67	122.70
2	D	272	ASP	CB-CA-C	-16.82	76.77	110.40

5 of 15 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	14	CYS	CA
1	A	69	VAL	CA
3	A	586	NAG	C1
2	C	137	ILE	CA
2	C	156	SER	CA

5 of 32 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	3	GLU	Mainchain
1	A	6	LYS	Mainchain
1	A	79	ASP	Sidechain
2	C	122	GLN	Mainchain
2	C	133	ASN	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	830	0	785	198	0
1	B	830	0	784	186	0
2	C	3609	0	3529	859	0
2	D	3609	0	3529	605	0
3	A	84	0	75	8	0
3	B	84	0	75	13	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	43	0	30	35	0
5	B	43	0	30	32	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9134	0	8837	1651	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 92.

The worst 5 of 1651 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:D:557:TYR:C	2:D:558:PRO:N	1.73	1.42
2:C:560:ASP:H	2:C:561:PHE:N	1.17	1.42
2:C:516:ASN:C	2:C:517:GLU:N	1.72	1.41
2:D:203:PHE:C	2:D:204:GLN:N	1.74	1.38
2:D:560:ASP:C	2:D:561:PHE:N	1.78	1.37

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%)	75 (74%)	22 (22%)	5 (5%)	3 16
1	B	100/108 (93%)	80 (80%)	13 (13%)	7 (7%)	1 7
2	C	460/466 (99%)	349 (76%)	84 (18%)	27 (6%)	2 11
2	D	460/466 (99%)	356 (77%)	82 (18%)	22 (5%)	3 17
All	All	1122/1148 (98%)	860 (77%)	201 (18%)	61 (5%)	2 14

5 of 61 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	123	PRO
2	C	303	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	352	PRO
2	C	376	GLY
2	C	560	ASP

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	88/93 (95%)	52 (59%)	36 (41%)	0 0
1	B	88/93 (95%)	56 (64%)	32 (36%)	0 1
2	C	386/411 (94%)	228 (59%)	158 (41%)	0 0
2	D	386/411 (94%)	237 (61%)	149 (39%)	0 1
All	All	948/1008 (94%)	573 (60%)	375 (40%)	0 0

5 of 375 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	495	LEU
1	B	33	LEU
2	D	486	LYS
2	C	507	ARG
2	C	562	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 66 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	472	ASN
1	B	16	ASN
2	D	457	ASN
2	C	502	GLN
2	C	571	ASN

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

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	$\#\|Z\| > 2$
3	NAG	A	582	3,2	14,14,15	1.10	1 (7%)	15,19,21	1.77	4 (26%)
3	NAG	A	583	3	14,14,15	1.81	2 (14%)	15,19,21	1.01	1 (6%)
3	NAG	A	584	3,2	14,14,15	1.05	1 (7%)	15,19,21	1.51	2 (13%)
3	NAG	A	585	3	14,14,15	1.11	1 (7%)	15,19,21	1.19	3 (20%)
3	NAG	A	586	3,2	14,14,15	1.02	1 (7%)	15,19,21	2.03	4 (26%)
3	NAG	A	587	3	14,14,15	1.12	1 (7%)	15,19,21	1.04	2 (13%)
3	NAG	B	582	3,2	14,14,15	1.16	1 (7%)	15,19,21	1.69	3 (20%)
3	NAG	B	583	3	14,14,15	1.10	1 (7%)	15,19,21	0.99	1 (6%)
3	NAG	B	584	3,2	14,14,15	1.01	1 (7%)	15,19,21	1.70	2 (13%)
3	NAG	B	585	3	14,14,15	0.96	1 (7%)	15,19,21	1.14	2 (13%)
3	NAG	B	586	3,2	14,14,15	1.07	1 (7%)	15,19,21	1.63	4 (26%)
3	NAG	B	587	3	14,14,15	1.01	1 (7%)	15,19,21	1.00	1 (6%)

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	NAG	A	582	3,2	-	0/6/23/26	0/1/1/1
3	NAG	A	583	3	-	0/6/23/26	0/1/1/1
3	NAG	A	584	3,2	-	0/6/23/26	0/1/1/1
3	NAG	A	585	3	-	0/6/23/26	0/1/1/1
3	NAG	A	586	3,2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	587	3	-	0/6/23/26	0/1/1/1
3	NAG	B	582	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	583	3	-	0/6/23/26	0/1/1/1
3	NAG	B	584	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	585	3	-	0/6/23/26	0/1/1/1
3	NAG	B	586	3,2	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	B	587	3	-	0/6/23/26	0/1/1/1

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	586	NAG	C8-C7	3.34	1.57	1.50
3	B	583	NAG	C8-C7	3.35	1.57	1.50
3	A	587	NAG	C8-C7	3.43	1.57	1.50
3	B	584	NAG	C8-C7	3.43	1.57	1.50
3	B	585	NAG	C8-C7	3.45	1.57	1.50

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	586	NAG	O4-C4-C5	-5.46	94.78	109.24
3	A	582	NAG	O4-C4-C5	-4.42	97.53	109.24
3	B	582	NAG	O5-C5-C6	-3.80	99.11	107.35
3	B	582	NAG	C3-C2-N2	-3.74	101.61	110.56
3	A	582	NAG	O5-C5-C6	-3.24	100.34	107.35

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	586	NAG	C1
3	A	586	NAG	C1

There are no torsion outliers.

There are no ring outliers.

11 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	582	NAG	1	0
3	A	583	NAG	2	0
3	A	584	NAG	1	0
3	A	585	NAG	1	0
3	A	586	NAG	2	0
3	A	587	NAG	3	0
3	B	583	NAG	1	0
3	B	584	NAG	5	0
3	B	585	NAG	4	0
3	B	586	NAG	6	0
3	B	587	NAG	2	0

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
5	HEM	A	580	2	30,50,50	2.39	6 (20%)	24,82,82	2.35	10 (41%)
5	HEM	B	580	2	30,50,50	2.17	5 (16%)	24,82,82	2.33	10 (41%)

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
5	HEM	A	580	2	-	0/10/54/54	0/0/8/8
5	HEM	B	580	2	-	0/10/54/54	0/0/8/8

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	580	HEM	C2D-C3D	-6.79	1.34	1.54
5	A	580	HEM	C2D-C3D	-6.60	1.34	1.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	580	HEM	C3B-C4B	-6.09	1.46	1.51
5	A	580	HEM	C3B-C4B	-6.08	1.46	1.51
5	A	580	HEM	C3D-C4D	-4.03	1.46	1.51

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	580	HEM	C3B-CAB-CBB	-2.94	119.95	124.46
5	A	580	HEM	C3B-CAB-CBB	-2.91	119.99	124.46
5	A	580	HEM	C3C-CAC-CBC	-2.88	120.04	124.46
5	B	580	HEM	C3C-CAC-CBC	-2.82	120.14	124.46
5	A	580	HEM	CBA-CAA-C2A	-2.66	107.77	112.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 67 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	580	HEM	35	0
5	B	580	HEM	32	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.