wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:40 AM GMT

PDB ID	:	2XQ1
Title	:	Crystal structure of peroxisomal catalase from the yeast Hansenula polymorpha
Authors	:	Penya-Soler, E.; Vega, M.C.; Wilmanns, M.; Williams, C.P.
Deposited on	:	2010-08-31
Resolution	:	2.90 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 (1)) were used in the production of this report:

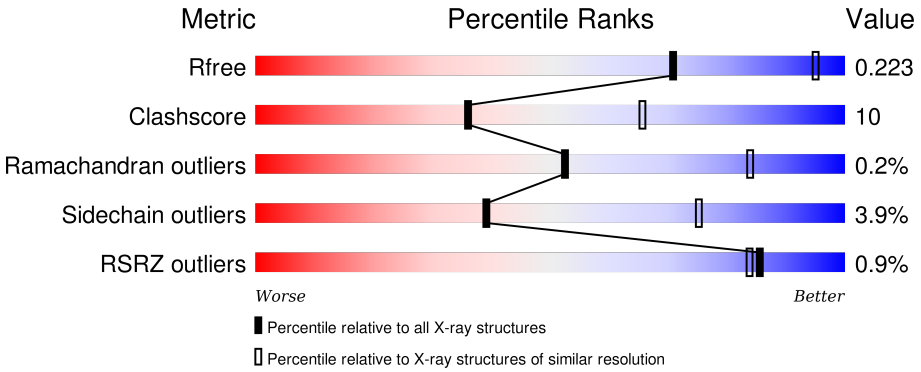
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.90 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.



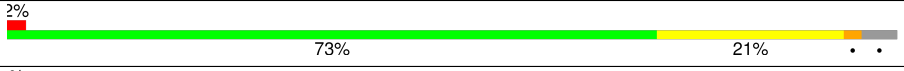
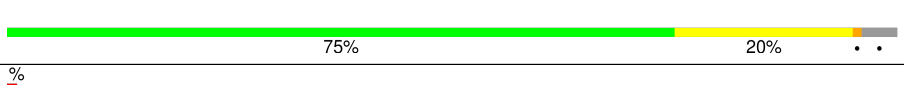
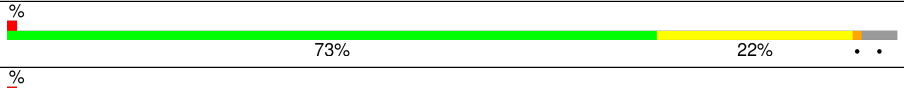
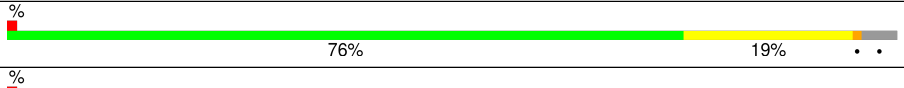
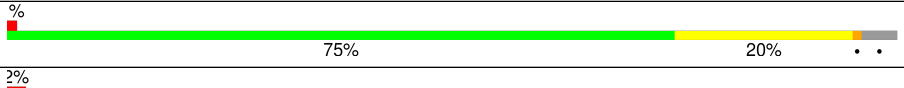
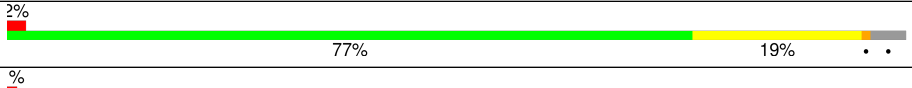
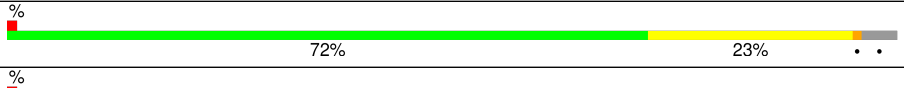
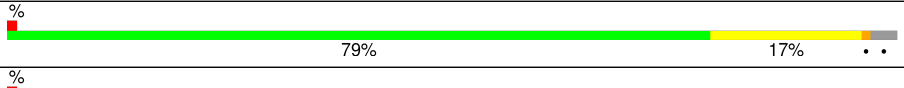
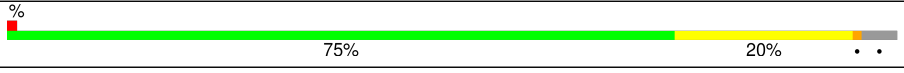
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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.

Mol	Chain	Length	Quality of chain
1	A	509	76% 20%
1	B	509	% 77% 19%
1	C	509	76% 20%
1	D	509	78% 18%
1	E	509	% 73% 22%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	509	
1	G	509	
1	H	509	
1	I	509	
1	J	509	
1	K	509	
1	L	509	
1	M	509	
1	N	509	
1	O	509	
1	P	509	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 63839 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.

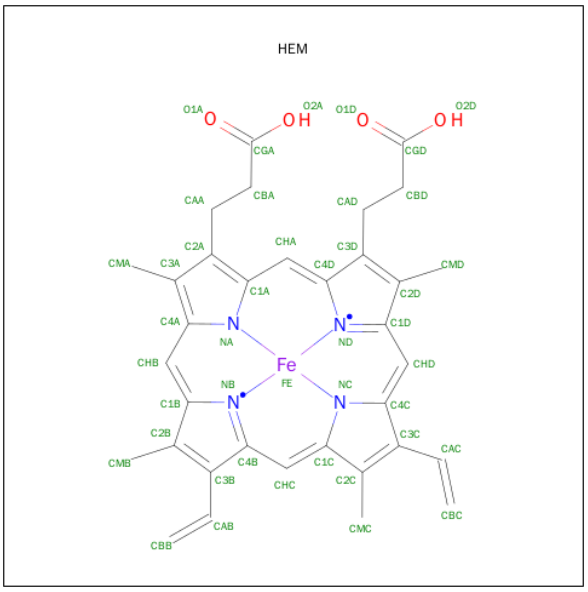
- Molecule 1 is a protein called PEROXISOMAL CATALASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	491	Total	C	N	O	S	0	0	0
			3925	2517	664	731	13			
1	B	494	Total	C	N	O	S	0	0	0
			3953	2532	671	737	13			
1	C	492	Total	C	N	O	S	0	0	0
			3960	2536	672	739	13			
1	D	494	Total	C	N	O	S	0	0	0
			3959	2534	673	739	13			
1	E	490	Total	C	N	O	S	0	0	0
			3947	2527	670	737	13			
1	F	494	Total	C	N	O	S	0	0	0
			3950	2529	671	737	13			
1	G	491	Total	C	N	O	S	0	0	0
			3951	2531	671	736	13			
1	H	487	Total	C	N	O	S	0	0	0
			3922	2511	666	732	13			
1	I	491	Total	C	N	O	S	0	0	0
			3959	2534	672	740	13			
1	J	491	Total	C	N	O	S	0	0	0
			3943	2525	670	735	13			
1	K	490	Total	C	N	O	S	0	0	0
			3949	2529	670	737	13			
1	L	489	Total	C	N	O	S	0	0	0
			3933	2519	669	732	13			
1	M	491	Total	C	N	O	S	0	0	0
			3945	2526	671	735	13			
1	N	491	Total	C	N	O	S	0	0	0
			3945	2522	669	741	13			
1	O	493	Total	C	N	O	S	0	0	0
			3957	2534	673	737	13			
1	P	491	Total	C	N	O	S	0	0	0
			3931	2518	668	732	13			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP P30263
A	0	ALA	-	EXPRESSION TAG	UNP P30263
B	-1	GLY	-	EXPRESSION TAG	UNP P30263
B	0	ALA	-	EXPRESSION TAG	UNP P30263
C	-1	GLY	-	EXPRESSION TAG	UNP P30263
C	0	ALA	-	EXPRESSION TAG	UNP P30263
D	-1	GLY	-	EXPRESSION TAG	UNP P30263
D	0	ALA	-	EXPRESSION TAG	UNP P30263
E	-1	GLY	-	EXPRESSION TAG	UNP P30263
E	0	ALA	-	EXPRESSION TAG	UNP P30263
F	-1	GLY	-	EXPRESSION TAG	UNP P30263
F	0	ALA	-	EXPRESSION TAG	UNP P30263
G	-1	GLY	-	EXPRESSION TAG	UNP P30263
G	0	ALA	-	EXPRESSION TAG	UNP P30263
H	-1	GLY	-	EXPRESSION TAG	UNP P30263
H	0	ALA	-	EXPRESSION TAG	UNP P30263
I	-1	GLY	-	EXPRESSION TAG	UNP P30263
I	0	ALA	-	EXPRESSION TAG	UNP P30263
J	-1	GLY	-	EXPRESSION TAG	UNP P30263
J	0	ALA	-	EXPRESSION TAG	UNP P30263
K	-1	GLY	-	EXPRESSION TAG	UNP P30263
K	0	ALA	-	EXPRESSION TAG	UNP P30263
L	-1	GLY	-	EXPRESSION TAG	UNP P30263
L	0	ALA	-	EXPRESSION TAG	UNP P30263
M	-1	GLY	-	EXPRESSION TAG	UNP P30263
M	0	ALA	-	EXPRESSION TAG	UNP P30263
N	-1	GLY	-	EXPRESSION TAG	UNP P30263
N	0	ALA	-	EXPRESSION TAG	UNP P30263
O	-1	GLY	-	EXPRESSION TAG	UNP P30263
O	0	ALA	-	EXPRESSION TAG	UNP P30263
P	-1	GLY	-	EXPRESSION TAG	UNP P30263
P	0	ALA	-	EXPRESSION TAG	UNP P30263

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	E	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	F	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	G	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	H	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	I	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	J	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	K	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	L	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	M	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	N	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	O	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	P	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is water.

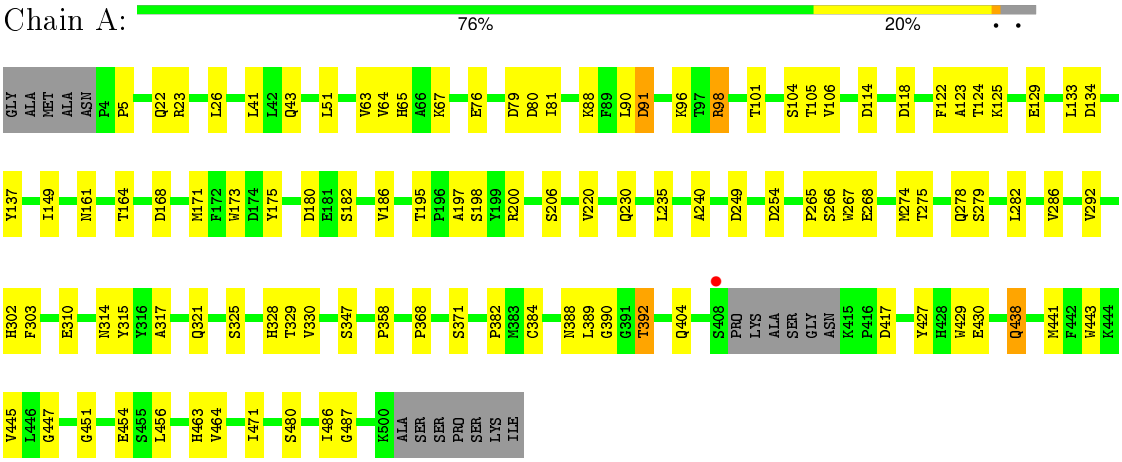
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	5	Total	O	0	0
			5	5		
3	B	7	Total	O	0	0
			7	7		
3	C	4	Total	O	0	0
			4	4		
3	D	6	Total	O	0	0
			6	6		
3	E	3	Total	O	0	0
			3	3		
3	F	6	Total	O	0	0
			6	6		
3	G	3	Total	O	0	0
			3	3		
3	H	2	Total	O	0	0
			2	2		
3	I	4	Total	O	0	0
			4	4		
3	J	6	Total	O	0	0
			6	6		
3	K	1	Total	O	0	0
			1	1		
3	L	6	Total	O	0	0
			6	6		
3	M	2	Total	O	0	0
			2	2		
3	N	1	Total	O	0	0
			1	1		
3	O	6	Total	O	0	0
			6	6		
3	P	3	Total	O	0	0
			3	3		

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.

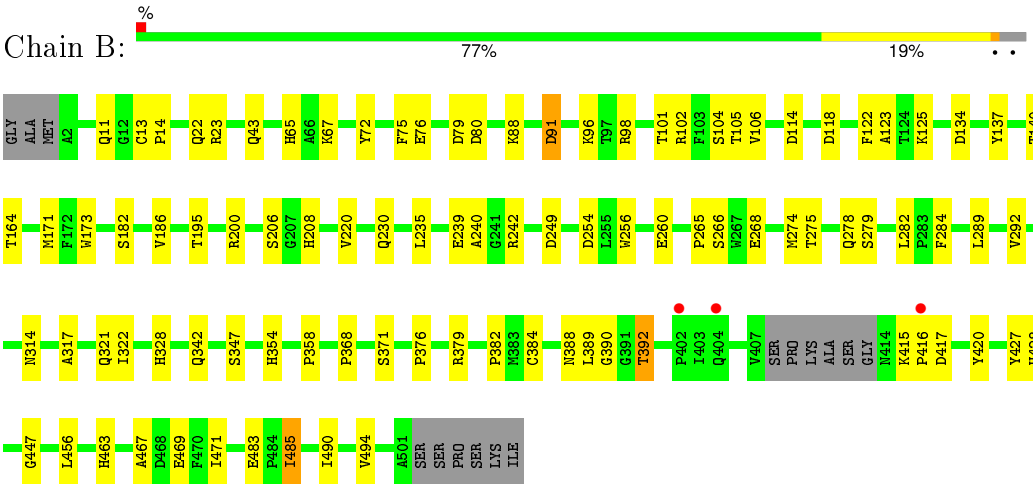
- Molecule 1: PEROXISOMAL CATALASE

Chain A:



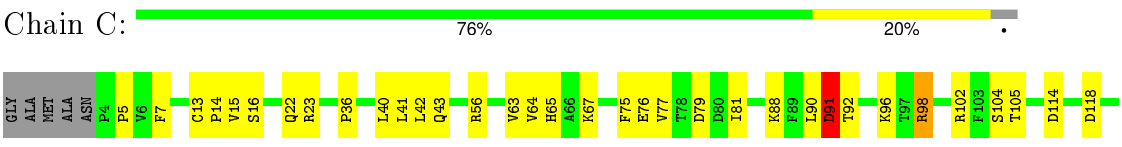
- Molecule 1: PEROXISOMAL CATALASE

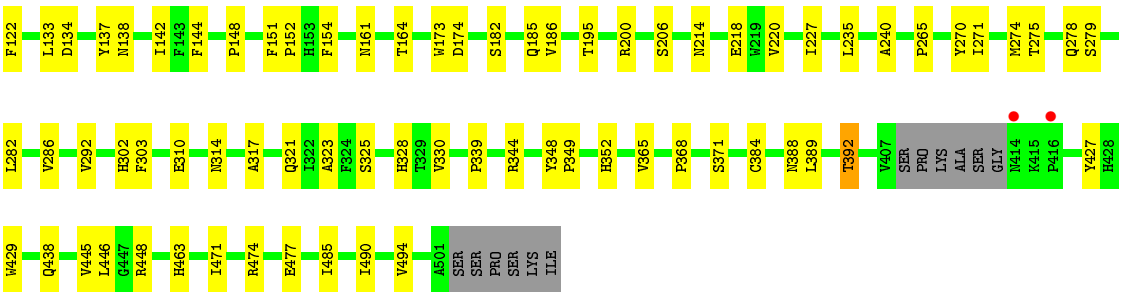
Chain B:



- Molecule 1: PEROXISOMAL CATALASE

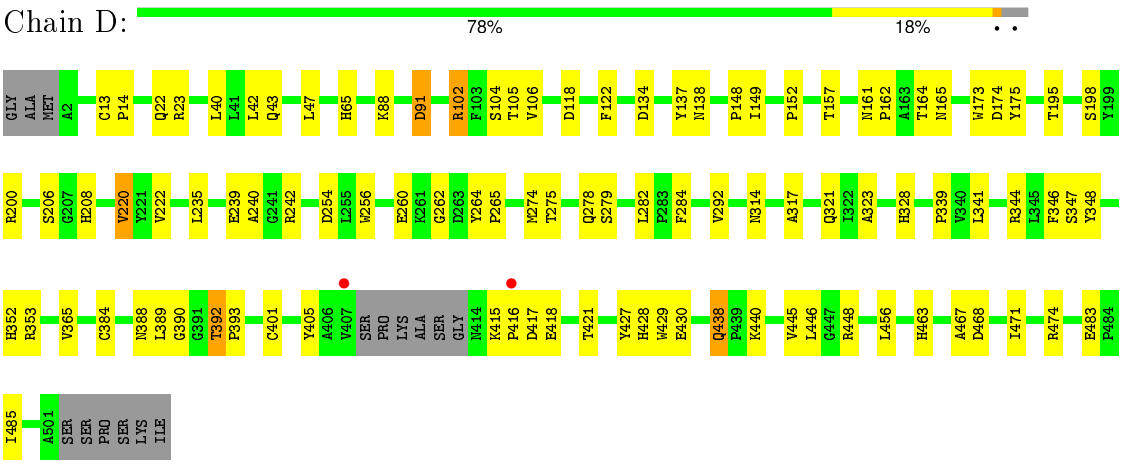
Chain C:





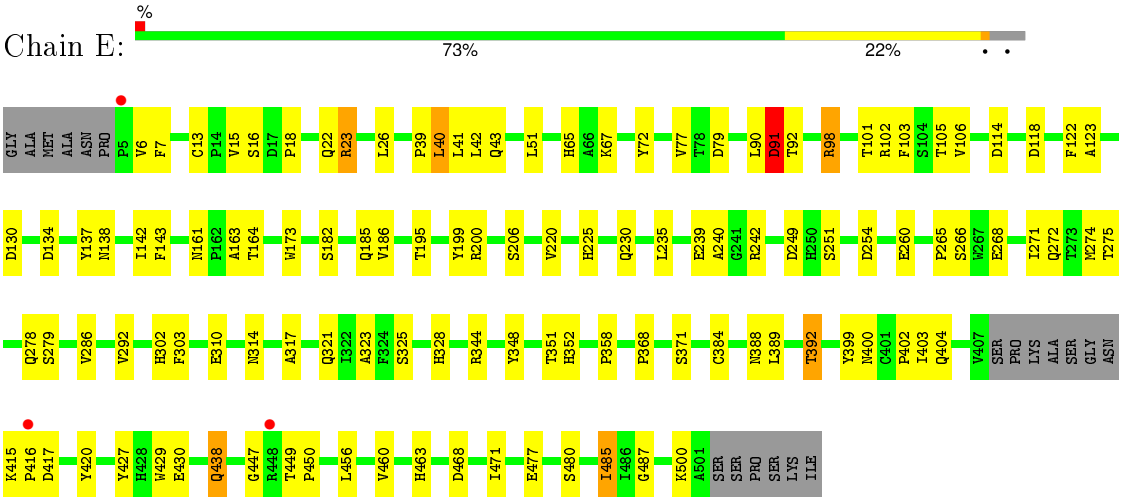
• Molecule 1: PEROXISOMAL CATALASE

Chain D:



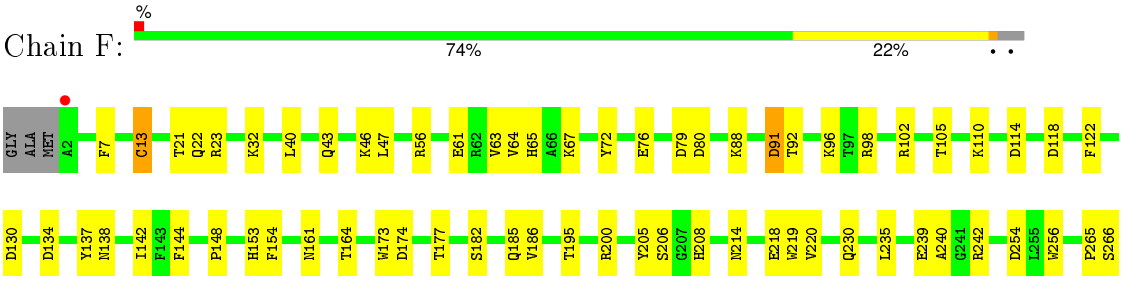
• Molecule 1: PEROXISOMAL CATALASE

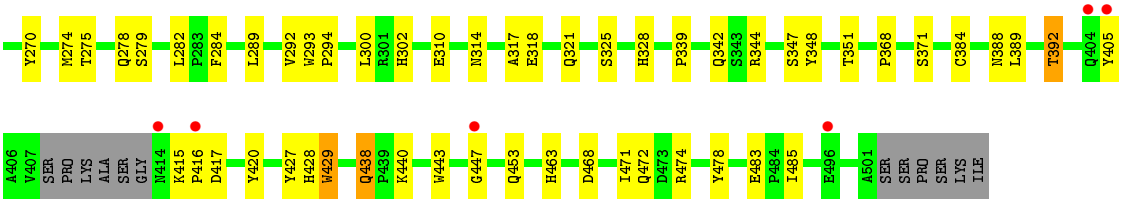
Chain E:

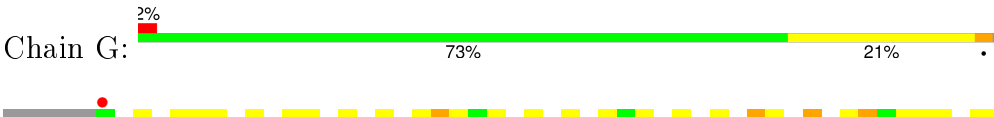


• Molecule 1: PEROXISOMAL CATALASE

Chain F:

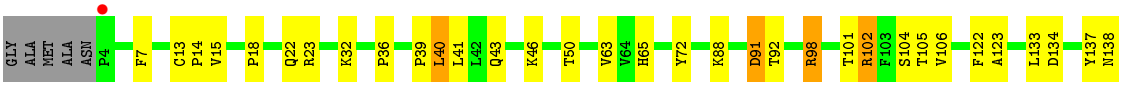


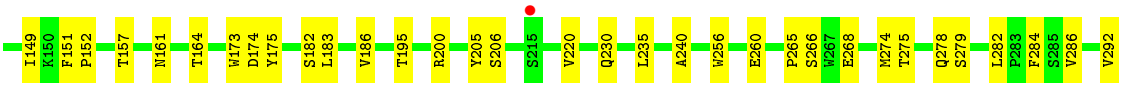


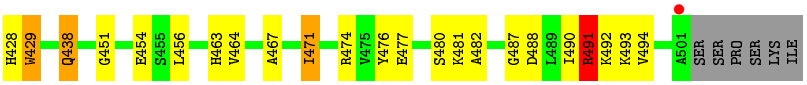


• Molecule 1: PEROXISOMAL CATALASE

Chain G:

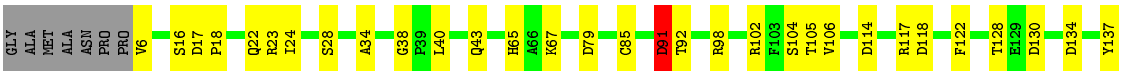


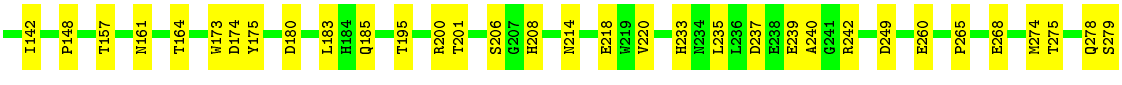


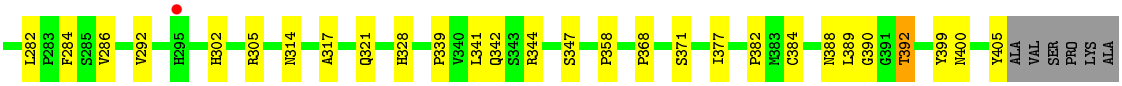


• Molecule 1: PEROXISOMAL CATALASE

Chain H:

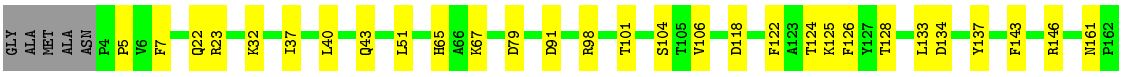


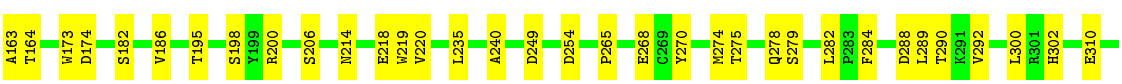


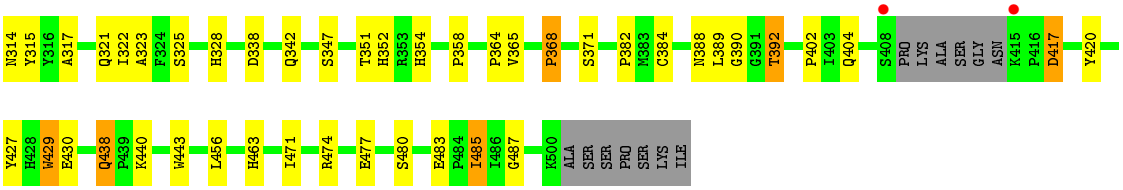


• Molecule 1: PEROXISOMAL CATALASE

Chain I:

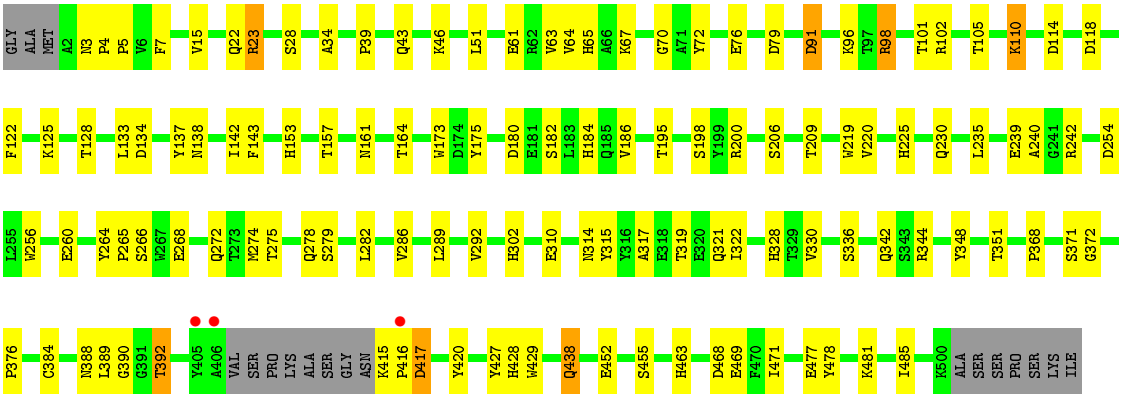






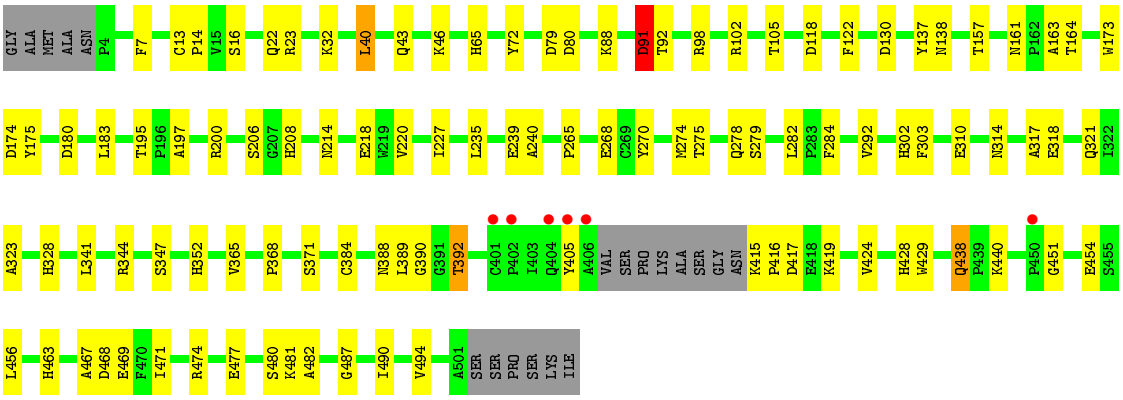
• Molecule 1: PEROXISOMAL CATALASE

Chain J:



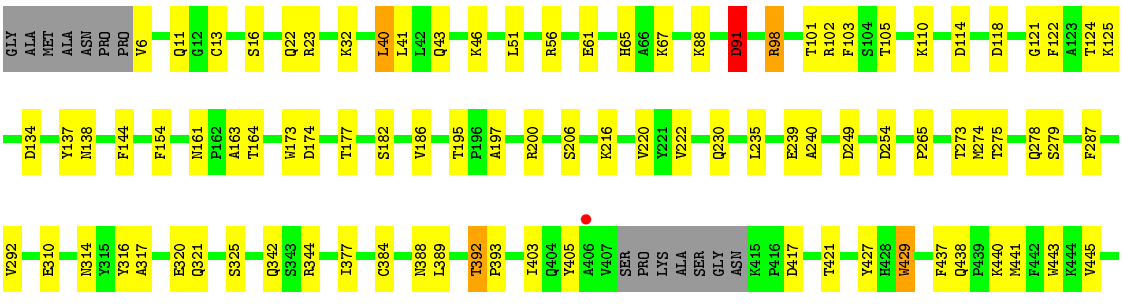
• Molecule 1: PEROXISOMAL CATALASE

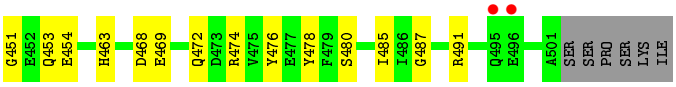
Chain K:



• Molecule 1: PEROXISOMAL CATALASE

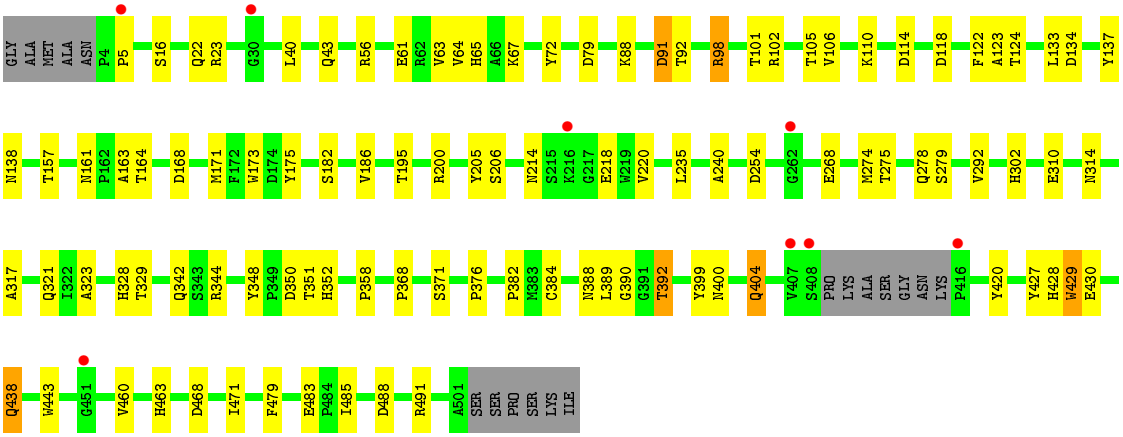
Chain L:





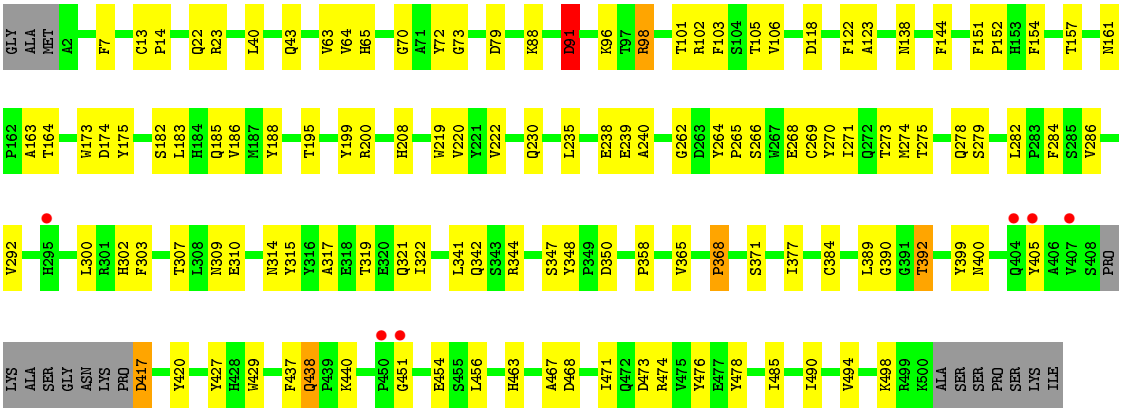
• Molecule 1: PEROXISOMAL CATALASE

Chain M:



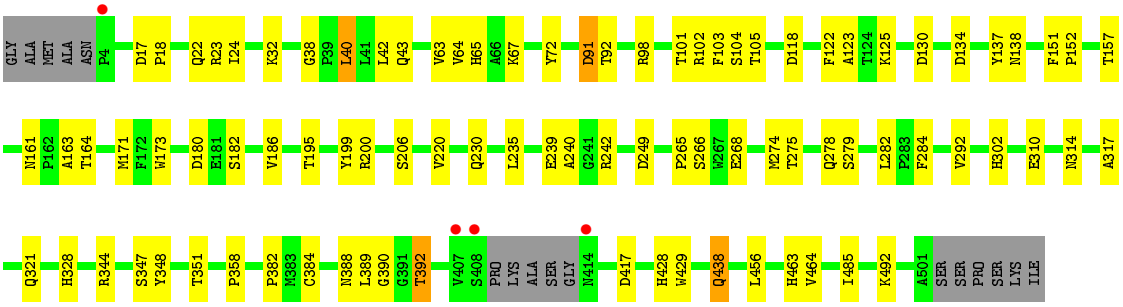
• Molecule 1: PEROXISOMAL CATALASE

Chain N:



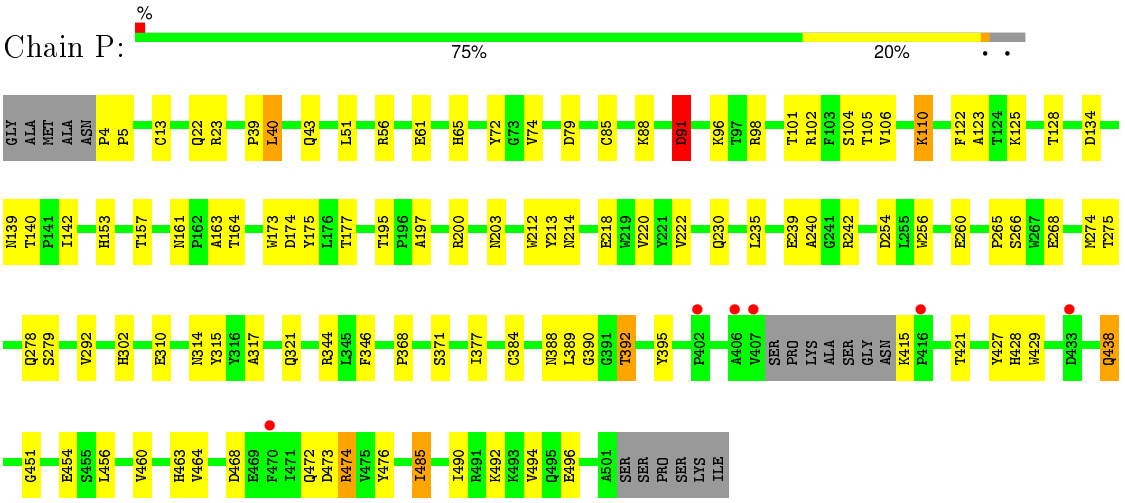
• Molecule 1: PEROXISOMAL CATALASE

Chain O:



• Molecule 1: PEROXISOMAL CATALASE

Chain P:


%
75%
20%
GLY ALA MET ALA ASN P4 P5 C13 Q22 R23 P39 L40 Q43 L51 R56 E61 H65 Y72 G73 V74 D79 C85 K88 D91 K96 T97 R98 T101 R102 F103 S104 T105 V106 K110 F122 A123 T124 K125 T128 D134
N139 T140 P141 I142 H153 T157 N161 P162 A163 T164 W173 D174 Y175 L176 T177 T195 P196 A197 R200 N203 W212 Y213 N214 E218 W219 V220 Y221 V222 Q230 L235 E239 A240 G241 R242 D254 L255 W256 E260 P265 S266 W267 E268 M274 T275
Q278 S279 V292 H302 E310 N314 Y315 Y316 A317 Q321 R344 L345 F346 P368 S371 I377 C384 N388 L389 G390 G391 T392 Y395 P402 A406 V407 SER PRO LYS ALA SER GLY ASN K415 P416 T421 Y427 H428 W429 D433 Q438
G451 E454 S455 L456 V460 H463 V464 D468 E469 F470 I471 Q472 D473 R474 V475 Y476 I485 I490 R491 K492 K493 V494 Q495 E496 A501 SER SER PRO SER LYS ILE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	132.52Å 196.68Å 170.85Å 90.00° 92.85° 90.00°	Depositor
Resolution (Å)	20.00 – 2.90 19.99 – 2.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-2.90) 100.0 (19.99-2.90)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.16 (at 2.88Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.191 , 0.223 0.193 , 0.223	Depositor DCC
R_{free} test set	976 reflections (0.51%)	DCC
Wilson B-factor (Å^2)	47.6	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 25.1	EDS
Estimated twinning fraction	0.018 for h,-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 192394 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	63839	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.20% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

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	0.68	0/4052	0.73	0/5523
1	B	0.69	0/4080	0.74	1/5562 (0.0%)
1	C	0.71	0/4087	0.74	1/5566 (0.0%)
1	D	0.70	0/4086	0.76	4/5568 (0.1%)
1	E	0.71	1/4073 (0.0%)	0.74	1/5545 (0.0%)
1	F	0.69	0/4077	0.72	0/5557
1	G	0.74	2/4078 (0.0%)	0.80	3/5553 (0.1%)
1	H	0.72	1/4047 (0.0%)	0.73	0/5511
1	I	0.73	0/4086	0.75	1/5563 (0.0%)
1	J	0.72	0/4070	0.75	1/5545 (0.0%)
1	K	0.70	0/4076	0.73	3/5550 (0.1%)
1	L	0.72	0/4058	0.74	0/5525
1	M	0.69	0/4072	0.69	0/5545
1	N	0.72	0/4071	0.72	2/5547 (0.0%)
1	O	0.67	0/4084	0.72	2/5562 (0.0%)
1	P	0.72	0/4058	0.70	0/5529
All	All	0.71	4/65155 (0.0%)	0.74	19/88751 (0.0%)

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	0	1
1	B	0	1
1	E	0	1
1	F	0	1
1	G	0	1
All	All	0	5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	491	ARG	CZ-NH1	10.08	1.46	1.33
1	E	242	ARG	CZ-NH1	5.75	1.40	1.33
1	G	491	ARG	CD-NE	5.14	1.55	1.46
1	H	85	CYS	CB-SG	-5.00	1.73	1.81

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	491	ARG	NE-CZ-NH2	19.03	129.82	120.30
1	G	491	ARG	NH1-CZ-NH2	-9.82	108.59	119.40
1	D	353	ARG	NE-CZ-NH2	-7.18	116.71	120.30
1	K	102	ARG	NE-CZ-NH2	-6.89	116.86	120.30
1	D	353	ARG	NE-CZ-NH1	6.02	123.31	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	447	GLY	Peptide
1	B	447	GLY	Peptide
1	E	447	GLY	Peptide
1	F	447	GLY	Peptide
1	G	491	ARG	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	3925	0	3662	79	0
1	B	3953	0	3696	75	0
1	C	3960	0	3734	80	0
1	D	3959	0	3716	76	0
1	E	3947	0	3720	96	0
1	F	3950	0	3693	91	0
1	G	3951	0	3725	87	0
1	H	3922	0	3690	73	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	3959	0	3735	84	0
1	J	3943	0	3703	89	0
1	K	3949	0	3725	77	0
1	L	3933	0	3706	79	0
1	M	3945	0	3719	81	0
1	N	3945	0	3691	91	0
1	O	3957	0	3726	64	0
1	P	3931	0	3687	85	0
2	B	43	0	30	6	0
2	C	43	0	30	6	0
2	D	43	0	30	4	0
2	E	43	0	30	5	0
2	F	43	0	30	3	0
2	G	43	0	30	3	0
2	H	43	0	30	5	0
2	I	43	0	30	1	0
2	J	43	0	30	5	0
2	K	43	0	30	6	0
2	L	43	0	30	6	0
2	M	43	0	30	3	0
2	N	43	0	30	5	0
2	O	43	0	30	5	0
2	P	43	0	30	3	0
3	A	5	0	0	0	0
3	B	7	0	0	2	0
3	C	4	0	0	1	0
3	D	6	0	0	0	0
3	E	3	0	0	0	0
3	F	6	0	0	0	0
3	G	3	0	0	0	0
3	H	2	0	0	0	0
3	I	4	0	0	1	0
3	J	6	0	0	3	0
3	K	1	0	0	1	0
3	L	6	0	0	0	0
3	M	2	0	0	0	0
3	N	1	0	0	0	0
3	O	6	0	0	0	0
3	P	3	0	0	1	0
All	All	63839	0	59778	1190	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 10.

The worst 5 of 1190 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:M:292:VAL:H	1:M:438:GLN:HE22	1.07	1.02
1:H:292:VAL:H	1:H:438:GLN:HE22	1.11	0.95
1:D:274:MET:CE	1:D:279:SER:HA	1.96	0.95
1:L:292:VAL:H	1:L:438:GLN:HE22	1.07	0.94
1:H:314:ASN:HD22	1:H:317:ALA:H	1.15	0.94

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	487/509 (96%)	464 (95%)	21 (4%)	2 (0%)	39	74
1	B	490/509 (96%)	471 (96%)	19 (4%)	0	100	100
1	C	488/509 (96%)	464 (95%)	22 (4%)	2 (0%)	39	74
1	D	490/509 (96%)	473 (96%)	17 (4%)	0	100	100
1	E	486/509 (96%)	463 (95%)	21 (4%)	2 (0%)	39	74
1	F	490/509 (96%)	467 (95%)	23 (5%)	0	100	100
1	G	487/509 (96%)	468 (96%)	19 (4%)	0	100	100
1	H	483/509 (95%)	456 (94%)	25 (5%)	2 (0%)	39	74
1	I	487/509 (96%)	467 (96%)	18 (4%)	2 (0%)	39	74
1	J	487/509 (96%)	466 (96%)	18 (4%)	3 (1%)	30	67
1	K	486/509 (96%)	462 (95%)	23 (5%)	1 (0%)	52	84
1	L	485/509 (95%)	464 (96%)	20 (4%)	1 (0%)	52	84
1	M	487/509 (96%)	458 (94%)	28 (6%)	1 (0%)	52	84
1	N	487/509 (96%)	466 (96%)	20 (4%)	1 (0%)	52	84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	O	489/509 (96%)	464 (95%)	24 (5%)	1 (0%)	52 84
1	P	487/509 (96%)	473 (97%)	13 (3%)	1 (0%)	52 84
All	All	7796/8144 (96%)	7446 (96%)	331 (4%)	19 (0%)	52 84

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	180	ASP
1	J	91	ASP
1	N	91	ASP
1	P	91	ASP
1	E	91	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	412/447 (92%)	397 (96%)	15 (4%)	42 78
1	B	417/447 (93%)	402 (96%)	15 (4%)	42 78
1	C	424/447 (95%)	411 (97%)	13 (3%)	47 82
1	D	421/447 (94%)	406 (96%)	15 (4%)	42 78
1	E	422/447 (94%)	407 (96%)	15 (4%)	42 78
1	F	417/447 (93%)	401 (96%)	16 (4%)	40 76
1	G	422/447 (94%)	398 (94%)	24 (6%)	25 59
1	H	419/447 (94%)	398 (95%)	21 (5%)	30 65
1	I	425/447 (95%)	408 (96%)	17 (4%)	38 74
1	J	419/447 (94%)	402 (96%)	17 (4%)	37 73
1	K	423/447 (95%)	405 (96%)	18 (4%)	35 71
1	L	419/447 (94%)	404 (96%)	15 (4%)	42 78
1	M	422/447 (94%)	409 (97%)	13 (3%)	47 82
1	N	420/447 (94%)	404 (96%)	16 (4%)	40 76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	O	422/447 (94%)	405 (96%)	17 (4%)	38 74
1	P	417/447 (93%)	401 (96%)	16 (4%)	40 76
All	All	6721/7152 (94%)	6458 (96%)	263 (4%)	39 75

5 of 263 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	104	SER
1	I	438	GLN
1	O	438	GLN
1	H	220	VAL
1	H	485	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 201 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	302	HIS
1	J	65	HIS
1	O	328	HIS
1	H	328	HIS
1	I	302	HIS

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

15 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	$\#\|Z\| > 2$
2	HEM	B	1503	1	30,50,50	2.16	7 (23%)	24,82,82	2.77	14 (58%)
2	HEM	C	1503	1	30,50,50	2.32	7 (23%)	24,82,82	2.69	12 (50%)
2	HEM	D	1503	1	30,50,50	2.62	7 (23%)	24,82,82	2.64	13 (54%)
2	HEM	E	1503	1	30,50,50	2.49	9 (30%)	24,82,82	2.77	12 (50%)
2	HEM	F	1503	1	30,50,50	2.58	8 (26%)	24,82,82	2.68	12 (50%)
2	HEM	G	1503	1	30,50,50	2.52	7 (23%)	24,82,82	2.40	10 (41%)
2	HEM	H	1503	1	30,50,50	2.41	6 (20%)	24,82,82	2.43	10 (41%)
2	HEM	I	1503	1	30,50,50	2.20	8 (26%)	24,82,82	2.64	11 (45%)
2	HEM	J	1503	1	30,50,50	2.43	10 (33%)	24,82,82	2.59	14 (58%)
2	HEM	K	1503	1	30,50,50	2.42	7 (23%)	24,82,82	2.90	13 (54%)
2	HEM	L	1503	1	30,50,50	2.38	7 (23%)	24,82,82	2.63	12 (50%)
2	HEM	M	1503	1	30,50,50	2.34	9 (30%)	24,82,82	2.69	14 (58%)
2	HEM	N	1503	1	30,50,50	2.68	6 (20%)	24,82,82	2.43	11 (45%)
2	HEM	O	1503	1	30,50,50	2.32	7 (23%)	24,82,82	2.67	12 (50%)
2	HEM	P	1503	1	30,50,50	2.28	7 (23%)	24,82,82	2.50	11 (45%)

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
2	HEM	B	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	C	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	D	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	E	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	F	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	G	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	H	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	I	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	J	1503	1	-	0/10/54/54	0/0/8/8

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEM	K	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	L	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	M	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	N	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	O	1503	1	-	0/10/54/54	0/0/8/8
2	HEM	P	1503	1	-	0/10/54/54	0/0/8/8

The worst 5 of 112 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	N	1503	HEM	C3B-C4B	-9.65	1.43	1.51
2	F	1503	HEM	C3B-C4B	-9.30	1.43	1.51
2	D	1503	HEM	C3B-C4B	-9.16	1.43	1.51
2	L	1503	HEM	C3B-C4B	-8.65	1.44	1.51
2	H	1503	HEM	C3B-C4B	-8.41	1.44	1.51

The worst 5 of 181 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	O	1503	HEM	C3C-CAC-CBC	-6.23	114.90	124.46
2	K	1503	HEM	C3C-CAC-CBC	-5.89	115.42	124.46
2	F	1503	HEM	C3C-CAC-CBC	-5.50	116.02	124.46
2	M	1503	HEM	C3C-CAC-CBC	-5.42	116.14	124.46
2	P	1503	HEM	C3C-CAC-CBC	-5.28	116.36	124.46

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

15 monomers are involved in 66 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1503	HEM	6	0
2	C	1503	HEM	6	0
2	D	1503	HEM	4	0
2	E	1503	HEM	5	0
2	F	1503	HEM	3	0
2	G	1503	HEM	3	0
2	H	1503	HEM	5	0
2	I	1503	HEM	1	0
2	J	1503	HEM	5	0
2	K	1503	HEM	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	L	1503	HEM	6	0
2	M	1503	HEM	3	0
2	N	1503	HEM	5	0
2	O	1503	HEM	5	0
2	P	1503	HEM	3	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	491/509 (96%)	-0.51	1 (0%)	95	95	21, 41, 67, 105	0
1	B	494/509 (97%)	-0.46	3 (0%)	90	89	20, 41, 68, 105	0
1	C	492/509 (96%)	-0.45	2 (0%)	93	92	20, 41, 68, 122	0
1	D	494/509 (97%)	-0.49	2 (0%)	93	92	20, 42, 69, 105	0
1	E	490/509 (96%)	-0.48	3 (0%)	90	89	20, 41, 68, 105	0
1	F	494/509 (97%)	-0.45	7 (1%)	78	76	20, 40, 69, 151	0
1	G	491/509 (96%)	-0.43	8 (1%)	74	72	20, 41, 68, 105	0
1	H	487/509 (95%)	-0.43	5 (1%)	84	82	21, 41, 68, 109	0
1	I	491/509 (96%)	-0.43	2 (0%)	93	92	20, 41, 69, 105	0
1	J	491/509 (96%)	-0.43	3 (0%)	90	89	21, 41, 68, 105	0
1	K	490/509 (96%)	-0.40	6 (1%)	81	78	20, 41, 69, 105	0
1	L	489/509 (96%)	-0.45	3 (0%)	90	89	21, 41, 69, 104	0
1	M	491/509 (96%)	-0.43	8 (1%)	74	72	22, 42, 69, 105	0
1	N	491/509 (96%)	-0.40	6 (1%)	81	78	22, 42, 69, 105	0
1	O	493/509 (96%)	-0.51	4 (0%)	87	86	21, 42, 70, 105	0
1	P	491/509 (96%)	-0.38	6 (1%)	81	78	21, 42, 70, 105	0
All	All	7860/8144 (96%)	-0.45	69 (0%)	85	84	20, 41, 69, 151	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	404	GLN	4.5
1	P	406	ALA	4.3
1	G	406	ALA	4.2
1	I	408	SER	3.7
1	E	416	PRO	3.6

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	HEM	H	1503	43/43	0.96	0.14	1.09	22,41,45,47	0
2	HEM	N	1503	43/43	0.96	0.13	0.71	28,37,39,41	0
2	HEM	O	1503	43/43	0.96	0.13	0.63	20,32,35,39	0
2	HEM	I	1503	43/43	0.97	0.13	0.52	25,34,39,43	0
2	HEM	M	1503	43/43	0.97	0.12	0.44	27,33,37,39	0
2	HEM	J	1503	43/43	0.97	0.14	0.41	24,34,39,41	0
2	HEM	K	1503	43/43	0.97	0.12	0.24	20,30,35,38	0
2	HEM	C	1503	43/43	0.97	0.12	0.06	19,25,31,34	0
2	HEM	L	1503	43/43	0.97	0.12	0.00	22,27,31,35	0
2	HEM	E	1503	43/43	0.97	0.12	-0.04	22,30,35,38	0
2	HEM	D	1503	43/43	0.97	0.12	-0.12	21,32,36,40	0
2	HEM	P	1503	43/43	0.97	0.11	-0.23	31,41,44,46	0
2	HEM	G	1503	43/43	0.97	0.12	-0.34	19,24,29,32	0
2	HEM	B	1503	43/43	0.96	0.11	-0.34	16,26,31,34	0
2	HEM	F	1503	43/43	0.98	0.11	-0.37	17,25,29,34	0

6.5 Other polymers (i)

There are no such residues in this entry.