W O R L D W I D E

wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:24 PM GMT

PDB ID : 1TH3
Title : Crystal structure of NADPH depleted bovine live catalase complexed with cyanide
Authors : Sugadev, R.; Balasundaresan, D.; Ponnuswamy, M.N.; Kumaran, D.; Swaminathan, S.; Sekar, K.
Deposited on : 2004-06-01
Resolution : 2.80 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) : 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)
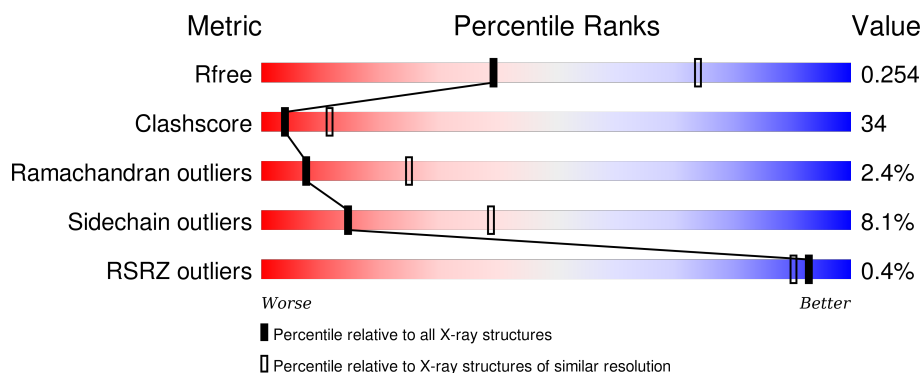Validation Pipeline (wwPDB-VP) : trunk26865

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 2.80 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	Percentile Ranks	Value
Rfree		0.254
Clashscore		34
Ramachandran outliers		2.4%
Sidechain outliers		8.1%
RSRZ outliers		0.4%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	506	52% 38% 7% ••
1	B	506	42% 47% 8% ••
1	C	506	% 44% 43% 9% ••
1	D	506	50% 44% ••

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
2	CYN	A	3000	-	-	X	-
3	HEM	B	2001	-	-	X	X
3	HEM	C	2002	-	-	X	X
3	HEM	D	2003	-	-	X	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 16932 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.
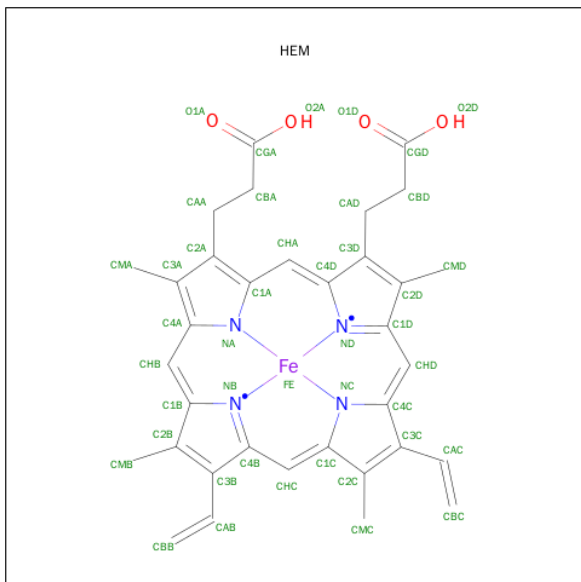
- Molecule 1 is a protein called Catalase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			
1	B	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			
1	C	499	Total	C	N	O	S	1	0	0
			4017	2548	715	740	14			
1	D	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			

- Molecule 2 is CYANIDE ION (three-letter code: CYN) (formula: CN).
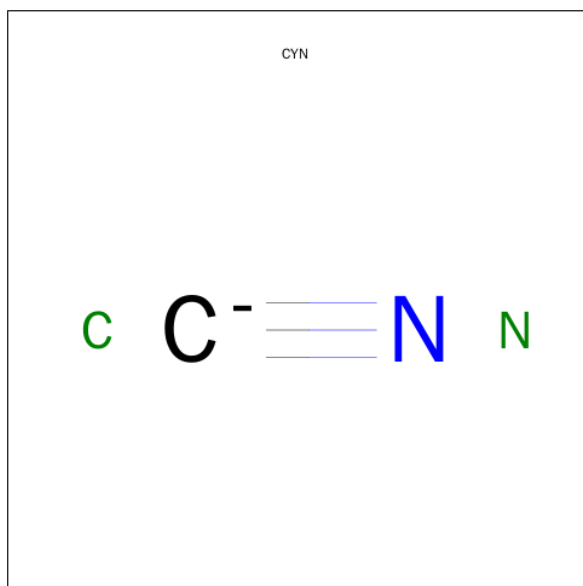
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	N	0	0
			2	1	1		
2	D	1	Total	C	N	0	0
			2	1	1		

• 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	C	Fe	N	O	0	0
			43	34	1	4	4		
3	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	201	Total	O	0	0
			201	201		
4	B	195	Total	O	0	0
			195	195		
4	C	138	Total	O	0	0
			138	138		
4	D	154	Total	O	0	0
			154	154		

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 propo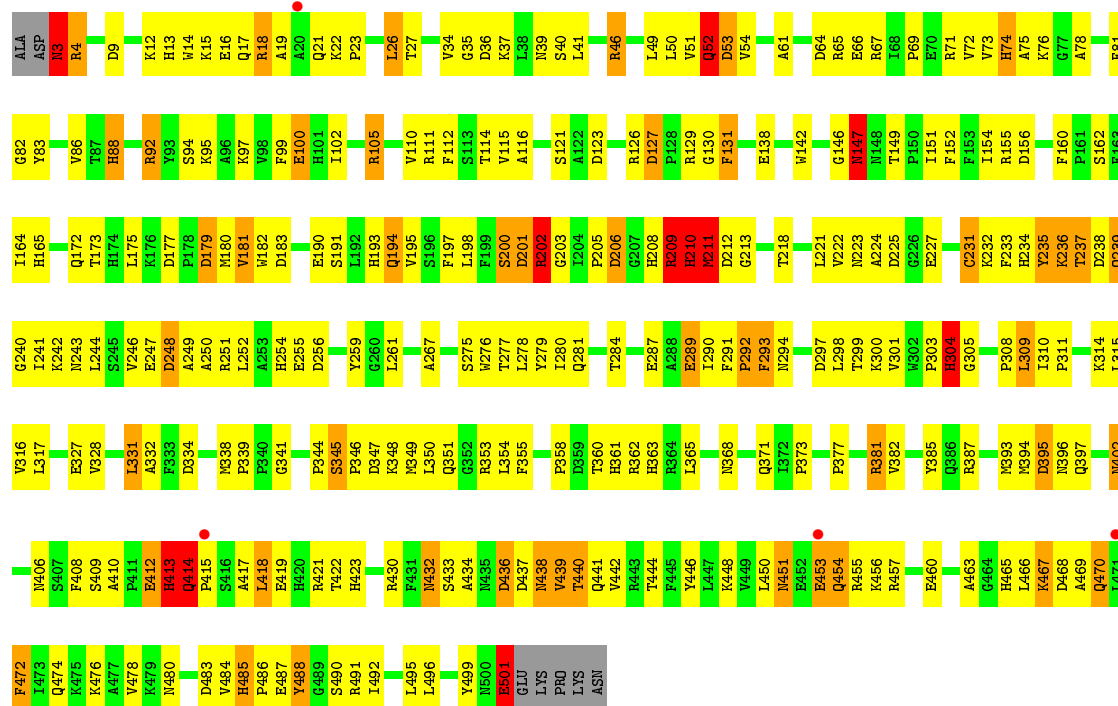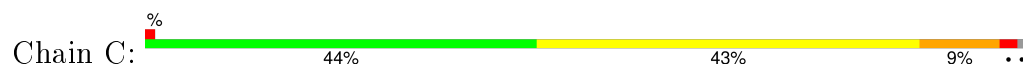rtions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 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: Catalase

- Molecule 1: Catalase

- Molecule 1: Catalase

Chain C:

- Molecule 1: Catalase

Chain D:

Q239 K242 V246 E247 L261 R262 D263 N266 T270 G271 N272 Y273 P274 S275 W276 T277 L278 Y279 I280 Q281 T284 F285 S286 E287 A288 E289 I290 F291 P292 F293 N294 P295 F296 D297 L298 T299 K300 V301 W302 P303 H304 G305 D306 Y307 P308 L309 I310 P311 V312 L317 V322 N323

E329 A332 F333 D334 P335 S336 N337 M338 P339 I342 E343 P346 D347 K348 M349 R353 Y357 P358 H361 R362 N368 Q371 I372 P373 V374 N375 C376 P377 Y378 R379 A380 R381 V382 A383 N384 D388 G389 P390 M391 C392 M393 M394 Q397 N402 Y403 Y404 P405 N406

S407 F408 S409 A410 P411 E412 H413 L418 E419 H420 R421 T422 H423 Q429 R430 F431 D437 N438 V439 T440 Q441 V442 F445 V449 L450 N451 E452 E453 Q454 R455 E460 L466 K467 L471 V478 K479 N480 V484 H485 P486 E487 Y488 I492 Q493 A494 L495 Y499 N500

E501 GLU LYS PRO LYS ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.06Å 140.11Å 226.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.21 – 2.80 42.27 – 2.80	Depositor EDS
% Data completeness (in resolution range)	76.4 (39.21-2.80) 76.5 (42.27-2.80)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.12	Depositor
$<I/\sigma(I)>$ [1]	3.07 (at 2.81Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.198 , 0.261 0.197 , 0.254	Depositor DCC
R_{free} test set	1553 reflections (3.07%)	DCC
Wilson B-factor (Å^2)	34.6	Xtriage
Anisotropy	0.578	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 53.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Outliers	1 of 52153 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	16932	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.60% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, CYN

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	0.71	8/4137 (0.2%)	1.63	67/5619 (1.2%)
1	B	0.98	5/4137 (0.1%)	1.47	29/5619 (0.5%)
1	C	0.70	12/4137 (0.3%)	1.63	50/5619 (0.9%)
1	D	0.60	1/4137 (0.0%)	0.83	8/5619 (0.1%)
All	All	0.76	26/16548 (0.2%)	1.43	154/22476 (0.7%)

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	6
1	B	0	4
1	C	3	4
All	All	5	14

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	319	ARG	CD-NE	48.75	2.29	1.46
1	D	413	HIS	CA-CB	-20.31	1.09	1.53
1	A	43	VAL	C-O	18.23	1.57	1.23
1	C	202	ARG	NE-CZ	16.64	1.54	1.33
1	B	319	ARG	NE-CZ	15.74	1.53	1.33

The worst 5 of 154 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	202	ARG	NE-CZ-NH2	-52.37	94.12	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	319	ARG	NE-CZ-NH1	-46.12	97.24	120.30
1	C	202	ARG	NE-CZ-NH1	43.99	142.29	120.30
1	B	319	ARG	CG-CD-NE	-41.07	25.55	111.80
1	A	19	ALA	O-C-N	-36.47	64.35	122.70

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	43	VAL	CA
1	A	395	ASP	CA
1	C	147	ASN	CA
1	C	453	GLU	CA
1	C	501	GLU	CA

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	19	ALA	Mainchain,Peptide
1	A	42	THR	Peptide
1	A	421	ARG	Sidechain
1	A	43	VAL	Mainchain,Peptide
1	B	3	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	4017	0	3837	240	0
1	B	4017	0	3840	343	0
1	C	4017	0	3839	327	0
1	D	4017	0	3839	270	0
2	A	2	0	0	6	0
2	D	2	0	0	0	0
3	A	43	0	30	16	0
3	B	43	0	30	25	0
3	C	43	0	30	26	0
3	D	43	0	30	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	201	0	0	33	0
4	B	195	0	0	33	0
4	C	138	0	0	16	0
4	D	154	0	0	20	0
All	All	16932	0	15475	1083	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 34.

The worst 5 of 1083 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:92:ARG:CD	1:A:92:ARG:CG	1.87	1.49
1:C:147:ASN:CG	3:C:2002:HEM:HAC	1.16	1.45
1:C:147:ASN:OD1	3:C:2002:HEM:CAC	1.63	1.42
1:C:147:ASN:ND2	3:C:2002:HEM:HAC	1.26	1.40
1:B:111:ARG:CD	3:B:2001:HEM:O1D	1.70	1.38

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	497/506 (98%)	450 (90%)	39 (8%)	8 (2%)	12	38
1	B	497/506 (98%)	425 (86%)	58 (12%)	14 (3%)	6	21
1	C	497/506 (98%)	427 (86%)	54 (11%)	16 (3%)	5	17
1	D	497/506 (98%)	438 (88%)	50 (10%)	9 (2%)	11	34
All	All	1988/2024 (98%)	1740 (88%)	201 (10%)	47 (2%)	7	25

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	43	VAL
1	A	100	GLU
1	B	4	ARG
1	B	100	GLU
1	B	124	THR

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	431/437 (99%)	400 (93%)	31 (7%)	18	45
1	B	431/437 (99%)	392 (91%)	39 (9%)	12	34
1	C	431/437 (99%)	388 (90%)	43 (10%)	9	27
1	D	431/437 (99%)	404 (94%)	27 (6%)	22	53
All	All	1724/1748 (99%)	1584 (92%)	140 (8%)	15	39

5 of 140 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	402	ASN
1	C	92	ARG
1	D	301	VAL
1	B	416	SER
1	C	4	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	435	ASN
1	C	239	GLN
1	D	337	ASN
1	B	461	ASN
1	C	52	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 ⓘ

6 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	2000	1	30,50,50	11.61	13 (43%)	24,82,82	8.90	9 (37%)
2	CYN	A	3000	-	0,1,1	0.00	-	0,0,0	0.00	-
3	HEM	B	2001	1	30,50,50	12.15	17 (56%)	24,82,82	10.90	18 (75%)
3	HEM	C	2002	1	30,50,50	2.73	10 (33%)	24,82,82	3.10	9 (37%)
3	HEM	D	2003	1,2	30,50,50	3.28	15 (50%)	24,82,82	9.07	13 (54%)
2	CYN	D	3001	3	0,1,1	0.00	-	0,0,0	0.00	-

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	2000	1	-	0/10/54/54	0/0/8/8
2	CYN	A	3000	-	-	0/0/0/0	0/0/0/0
3	HEM	B	2001	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
3	HEM	C	2002	1	-	0/10/54/54	0/0/8/8
3	HEM	D	2003	1,2	-	0/10/54/54	0/0/8/8
2	CYN	D	3001	3	-	0/0/0/0	0/0/0/0

The worst 5 of 55 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2000	HEM	C3B-C4B	-6.87	1.45	1.51
3	C	2002	HEM	C3B-C4B	-6.86	1.45	1.51
3	B	2001	HEM	C3B-CAB	-6.70	1.38	1.51
3	C	2002	HEM	C3B-CAB	-6.68	1.38	1.51
3	B	2001	HEM	C3B-C4B	-6.65	1.45	1.51

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	2000	HEM	CMC-C2C-C3C	-36.50	25.47	116.53
3	B	2001	HEM	CMC-C2C-C3C	-36.09	26.48	116.53
3	B	2001	HEM	CAA-C2A-C1A	-21.97	103.15	127.01
3	D	2003	HEM	CAA-C2A-C1A	-9.83	116.34	127.01
3	B	2001	HEM	CBA-CAA-C2A	-4.63	104.23	112.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 88 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2000	HEM	16	0
2	A	3000	CYN	6	0
3	B	2001	HEM	25	0
3	C	2002	HEM	26	0
3	D	2003	HEM	21	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)

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, 95th 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(Å²)	Q<0.9
1	A	499/506 (98%)	-0.61	2 (0%)	93	90	9, 29, 55, 85	0
1	B	499/506 (98%)	-0.49	2 (0%)	93	90	12, 37, 64, 97	0
1	C	499/506 (98%)	-0.48	4 (0%)	87	81	12, 33, 65, 90	1 (0%)
1	D	499/506 (98%)	-0.49	0	100	100	10, 35, 69, 89	0
All	All	1996/2024 (98%)	-0.52	8 (0%)	93	90	9, 34, 64, 97	1 (0%)

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	3	ASN	2.9
1	A	3	ASN	2.7
1	C	20	ALA	2.6
1	A	20	ALA	2.4
1	C	471	LEU	2.4

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, 95th 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
3	HEM	B	2001	43/43	0.92	0.25	3.09	20,37,58,100	0
3	HEM	D	2003	43/43	0.93	0.27	2.96	21,48,68,78	0
3	HEM	C	2002	43/43	0.94	0.21	2.53	22,37,54,101	0
3	HEM	A	2000	43/43	0.97	0.18	0.56	7,26,38,101	0
2	CYN	D	3001	2/2	0.99	0.12	-	57,57,57,63	0
2	CYN	A	3000	2/2	1.00	0.17	-	41,41,41,48	0

6.5 Other polymers (i)

There are no such residues in this entry.