



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 11:36 PM GMT

PDB ID : 1CCI  
Title : HOW FLEXIBLE ARE PROTEINS? TRAPPING OF A FLEXIBLE LOOP  
Authors : Cao, Y.; Musah, R.A.; Wilcox, S.K.; Goodin, D.B.; Mcree, D.E.  
Deposited on : 1996-12-18  
Resolution : 2.40 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

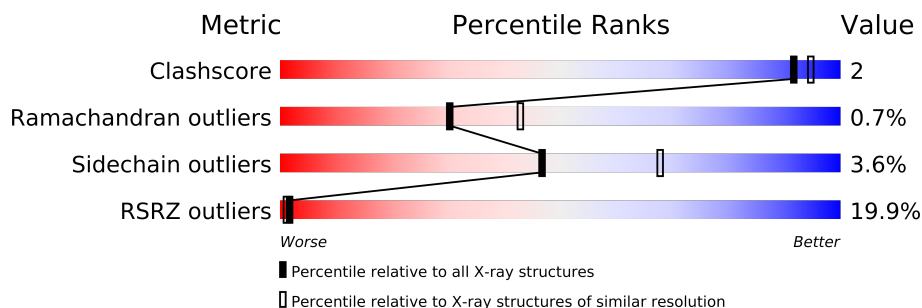
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.40 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	79885	2789 (2.40-2.40)
Ramachandran outliers	78287	2736 (2.40-2.40)
Sidechain outliers	78261	2737 (2.40-2.40)
RSRZ outliers	66119	2210 (2.40-2.40)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	294	

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 3003 atoms, of which 513 are hydrogens and 0 are deuterium.

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.

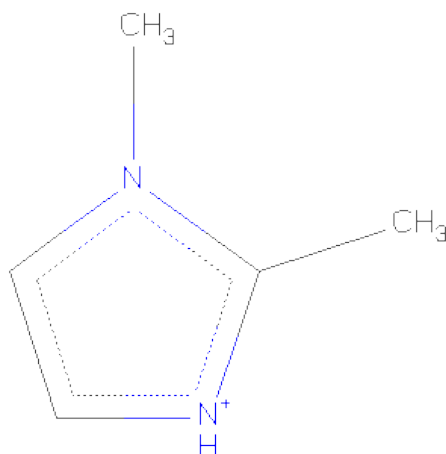
- Molecule 1 is a protein called CYTOCHROME C PEROXIDASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	291	Total	C	H	N	O	S	0	0	0
			2855	1494	512	392	451	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	53	ILE	THR	VARIANT	UNP P00431
A	152	GLY	ASP	VARIANT	UNP P00431
A	202	GLY	PHE	ENGINEERED	UNP P00431

- Molecule 2 is 2,3-DIMETHYLIMIDAZOLIUMION (three-letter code: DMI) (formula:  $C_5H_9N_2^+$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	N	0	0
			8	5	1	2		

- 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
			Total	C	Fe	N	O		
3	A	1	43	34	1	4	4	0	0

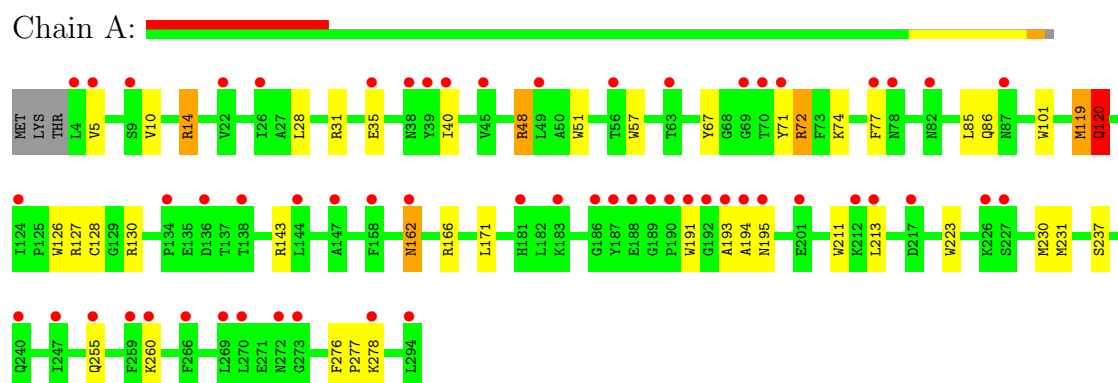
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	97	Total	O	0	0
			97	97		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

#### • Molecule 1: CYTOCHROME C PEROXIDASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.72Å 76.63Å 51.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	5.00 – 2.40 46.60 – 2.01	Depositor EDS
% Data completeness (in resolution range)	67.1 (5.00-2.40) 73.3 (46.60-2.01)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.01Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	(Not available) , (Not available) 0.388 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	19.6	Xtriage
Anisotropy	0.297	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 23.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 21339 reflections	Xtriage
$F_o, F_c$ correlation	0.76	EDS
Total number of atoms	3003	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.48% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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: DMI, 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	0.84	0/2408	1.64	46/3259 (1.4%)

There are no bond length outliers.

All (46) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	31	ARG	NE-CZ-NH2	-15.03	112.78	120.30
1	A	48	ARG	NE-CZ-NH1	11.72	126.16	120.30
1	A	166	ARG	NE-CZ-NH2	-11.50	114.55	120.30
1	A	31	ARG	NE-CZ-NH1	9.34	124.97	120.30
1	A	162	ASN	CA-C-N	-8.23	99.09	117.20
1	A	223	TRP	CD1-CG-CD2	8.19	112.85	106.30
1	A	127	ARG	CB-CG-CD	-8.16	90.38	111.60
1	A	101	TRP	CD1-CG-CD2	8.13	112.80	106.30
1	A	143	ARG	NE-CZ-NH2	-8.13	116.24	120.30
1	A	48	ARG	NE-CZ-NH2	-8.04	116.28	120.30
1	A	191	TRP	CD1-CG-CD2	8.00	112.70	106.30
1	A	72	ARG	NE-CZ-NH1	7.79	124.20	120.30
1	A	51	TRP	CD1-CG-CD2	7.75	112.50	106.30
1	A	101	TRP	CE2-CD2-CG	-7.59	101.23	107.30
1	A	211	TRP	CD1-CG-CD2	7.59	112.37	106.30
1	A	51	TRP	CE2-CD2-CG	-7.58	101.23	107.30
1	A	166	ARG	NE-CZ-NH1	7.51	124.06	120.30
1	A	191	TRP	CE2-CD2-CG	-7.42	101.37	107.30
1	A	126	TRP	CD1-CG-CD2	7.21	112.07	106.30
1	A	71	TYR	CB-CG-CD2	-7.09	116.75	121.00
1	A	223	TRP	CE2-CD2-CG	-7.07	101.65	107.30
1	A	211	TRP	CE2-CD2-CG	-7.02	101.69	107.30
1	A	127	ARG	NE-CZ-NH1	6.93	123.76	120.30
1	A	126	TRP	CE2-CD2-CG	-6.91	101.77	107.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	127	ARG	NE-CZ-NH2	-6.60	117.00	120.30
1	A	14	ARG	NE-CZ-NH2	-6.58	117.01	120.30
1	A	57	TRP	CD1-CG-CD2	6.51	111.51	106.30
1	A	57	TRP	CE2-CD2-CG	-6.37	102.20	107.30
1	A	67	TYR	CB-CG-CD1	-6.31	117.22	121.00
1	A	162	ASN	N-CA-C	6.28	127.96	111.00
1	A	195	ASN	N-CA-CB	-5.84	100.08	110.60
1	A	126	TRP	CB-CG-CD1	-5.79	119.48	127.00
1	A	126	TRP	CG-CD2-CE3	5.71	139.04	133.90
1	A	231	MET	CA-CB-CG	5.56	122.75	113.30
1	A	51	TRP	CB-CG-CD1	-5.52	119.82	127.00
1	A	166	ARG	CG-CD-NE	-5.35	100.57	111.80
1	A	223	TRP	CG-CD1-NE1	-5.34	104.76	110.10
1	A	143	ARG	NE-CZ-NH1	5.29	122.95	120.30
1	A	5	VAL	CG1-CB-CG2	-5.21	102.57	110.90
1	A	126	TRP	CG-CD1-NE1	-5.19	104.91	110.10
1	A	120	GLN	N-CA-C	5.18	125.00	111.00
1	A	162	ASN	CA-C-O	5.18	130.99	120.10
1	A	230	MET	CA-CB-CG	5.07	121.92	113.30
1	A	119	MET	CG-SD-CE	5.05	108.28	100.20
1	A	101	TRP	CG-CD1-NE1	-5.02	105.08	110.10
1	A	237	SER	CB-CA-C	-5.00	100.60	110.10

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2343	512	1711	9	1
2	A	7	1	8	0	0
3	A	43	0	30	0	0
4	A	97	0	0	3	0
All	All	2490	513	1749	9	1



Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (9) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:128:CYS:HA	4:A:326:HOH:O	1.81	0.81
1:A:130:ARG:NH1	4:A:326:HOH:O	2.37	0.56
1:A:276:PHE:HA	1:A:277:PRO:HD3	1.92	0.43
1:A:48:ARG:HG2	1:A:85:LEU:HD21	2.00	0.43
1:A:119:MET:O	1:A:120:GLN:HG2	2.19	0.42
1:A:48:ARG:NH1	4:A:302:HOH:O	2.52	0.41
1:A:10:VAL:HG13	1:A:128:CYS:SG	2.61	0.40
1:A:40:ILE:HD11	1:A:194:ALA:HB3	2.03	0.40
1:A:77:PHE:CE1	1:A:86:GLN:HG3	2.57	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:14:ARG:NH2	1:A:213:LEU:O[4_477]	2.00	0.20

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	289/294 (98%)	278 (96%)	9 (3%)	2 (1%)	30 43

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	162	ASN
1	A	193	ALA

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	248/251 (99%)	239 (96%)	9 (4%)	47 68

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	LEU
1	A	35	GLU
1	A	72	ARG
1	A	74	LYS
1	A	120	GLN
1	A	171	LEU
1	A	255	GLN
1	A	260	LYS
1	A	278	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	20	GLN
1	A	24	ASN
1	A	78	ASN
1	A	220	ASN
1	A	240	GLN
1	A	292	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

2 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	DMI	A	499	-	7,7,7	1.80	1 (14%)	9,9,9	2.33	3 (33%)
3	HEM	A	500	1,4	49,50,50	5.55	21 (42%)	46,82,82	1.40	7 (15%)

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	DMI	A	499	-	-	0/0/0/0	0/1/1/1
3	HEM	A	500	1,4	-	0/14/114/114	0/0/8/8

All (22) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	500	HEM	C3D-C4D	-23.62	1.38	1.44
3	A	500	HEM	C2D-C1D	-20.25	1.39	1.44
3	A	500	HEM	C2B-C1B	-18.70	1.39	1.44
3	A	500	HEM	C3B-C4B	-4.68	1.38	1.44
3	A	500	HEM	CHB-C1B	4.13	1.41	1.35
2	A	499	DMI	C4-N3	-3.88	1.30	1.37
3	A	500	HEM	CHA-C4D	3.53	1.40	1.35
3	A	500	HEM	FE-NA	3.13	2.05	1.92
3	A	500	HEM	CBC-CAC	3.10	1.46	1.28
3	A	500	HEM	FE-NC	3.07	2.09	1.97
3	A	500	HEM	CBB-CAB	3.04	1.46	1.28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	500	HEM	FE-ND	2.84	2.08	1.97
3	A	500	HEM	C4D-ND	-2.82	1.33	1.39
3	A	500	HEM	FE-NB	2.76	2.07	1.97
3	A	500	HEM	CHD-C4C	2.71	1.41	1.36
3	A	500	HEM	C1B-NB	-2.69	1.34	1.39
3	A	500	HEM	CHC-C1C	2.68	1.41	1.36
3	A	500	HEM	C1A-C2A	-2.43	1.39	1.43
3	A	500	HEM	C4C-NC	-2.36	1.34	1.38
3	A	500	HEM	CMC-C2C	2.34	1.54	1.47
3	A	500	HEM	CMD-C2D	2.16	1.54	1.47
3	A	500	HEM	CMB-C2B	2.14	1.54	1.47

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	499	DMI	CM2-C2-N3	4.34	128.72	124.88
3	A	500	HEM	C3B-C4B-NB	-4.05	111.10	114.00
2	A	499	DMI	C4-N3-C2	4.01	112.11	108.72
3	A	500	HEM	CMC-C2C-C3C	2.73	132.60	126.16
3	A	500	HEM	C4A-C3A-C2A	-2.63	105.16	107.00
2	A	499	DMI	N1-C2-N3	-2.49	105.67	109.41
3	A	500	HEM	O2A-CGA-CBA	2.19	121.95	114.22
3	A	500	HEM	C3A-C4A-NA	2.11	111.00	109.41
3	A	500	HEM	C1A-CHA-C4D	2.07	130.20	127.47
3	A	500	HEM	CMB-C2B-C3B	2.04	130.97	126.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	291/294 (98%)	1.33	58 (19%) 2 1	9, 24, 44, 58	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	38	ASN	5.4
1	A	183	LYS	4.8
1	A	188	GLU	4.0
1	A	269	LEU	3.9
1	A	63	THR	3.7
1	A	226	LYS	3.4
1	A	45	VAL	3.3
1	A	144	LEU	3.2
1	A	70	THR	3.2
1	A	213	LEU	3.2
1	A	77	PHE	3.2
1	A	270	LEU	3.1
1	A	138	THR	3.1
1	A	26	ILE	2.9
1	A	194	ALA	2.9
1	A	193	ALA	2.9
1	A	124	ILE	2.8
1	A	71	TYR	2.8
1	A	201	GLU	2.8
1	A	189	GLY	2.8
1	A	49	LEU	2.7
1	A	195	ASN	2.7
1	A	191	TRP	2.7
1	A	266	PHE	2.7
1	A	158	PHE	2.6
1	A	278	LYS	2.6
1	A	273	GLY	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	294	LEU	2.5
1	A	136	ASP	2.5
1	A	147	ALA	2.5
1	A	272	ASN	2.5
1	A	217	ASP	2.4
1	A	240	GLN	2.4
1	A	5	VAL	2.3
1	A	192	GLY	2.3
1	A	40	ILE	2.3
1	A	4	LEU	2.3
1	A	78	ASN	2.3
1	A	39	TYR	2.3
1	A	134	PRO	2.3
1	A	212	LYS	2.2
1	A	87	ASN	2.2
1	A	247	ILE	2.2
1	A	259	PHE	2.2
1	A	190	PRO	2.2
1	A	260	LYS	2.2
1	A	187	TYR	2.2
1	A	162	ASN	2.1
1	A	35	GLU	2.1
1	A	82	ASN	2.1
1	A	227	SER	2.1
1	A	255	GLN	2.1
1	A	186	GLY	2.1
1	A	9	SER	2.1
1	A	181	HIS	2.0
1	A	69	GLY	2.0
1	A	22	VAL	2.0
1	A	56	THR	2.0

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

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<sup>th</sup> 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	HEM	A	500	43/43	0.27	0.24	10,20,24,25	0
2	DMI	A	499	7/7	0.27	0.22	20,35,36,36	0

## 6.5 Other polymers

There are no such residues in this entry.