Full wwPDB X-ray Structure Validation Report (i)

Aug 6, 2023 – 04:19 PM EDT

PDB ID	:	1KXN
Title	:	Crystal Structure of Cytochrome c Peroxidase with a Proposed Electron Transfer Pathway Excised to Form a Ligand Binding Channel.
Authors	:	Rosenfeld, R.J.; Hayes, A.M.A.; Musah, R.A.; Goodin, D.B.
Deposited on	:	2002-02-01
Resolution	:	1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

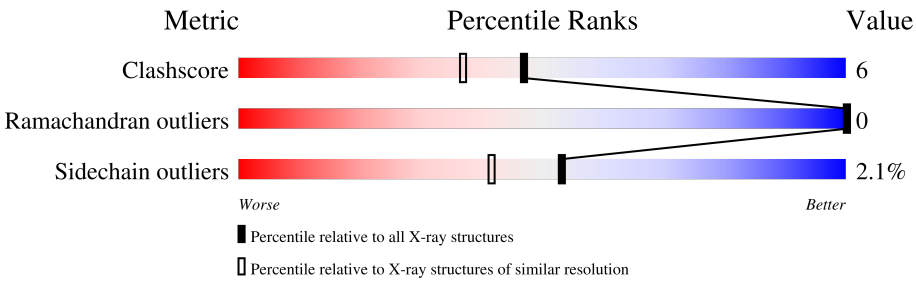
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.80 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	289	85% 13% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2706 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.

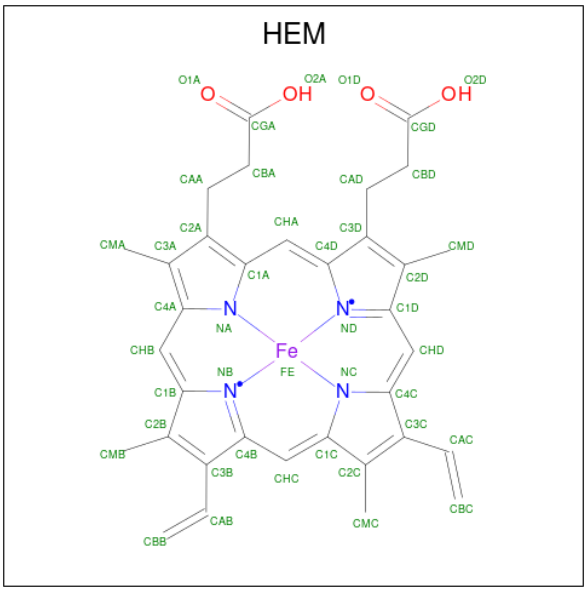
- Molecule 1 is a protein called cytochrome c peroxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	289	2315	1476	384	449	6	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	190	GLY	PRO	engineered mutation	UNP P00431
A	?	-	TRP	deletion	UNP P00431
A	?	-	ALA	deletion	UNP P00431

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	348	348	348	0	0

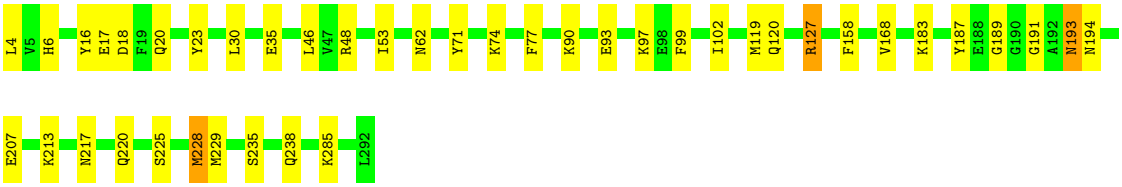
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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: cytochrome c peroxidase

Chain A:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	106.80Å 75.80Å 51.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.80	Depositor
% Data completeness (in resolution range)	46.7 (10.00-1.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
Refinement program	SHELXL	Depositor
R, R_{free}	0.187 , 0.194	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2706	wwPDB-VP
Average B, all atoms ($Å^2$)	20.0	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: 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	0.39	1/2378 (0.0%)	1.04	11/3219 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	191	GLY	C-N	-5.15	1.22	1.34

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	189	GLY	C-N-CA	-10.92	99.37	122.30
1	A	48	ARG	NE-CZ-NH1	-9.46	115.57	120.30
1	A	127	ARG	NE-CZ-NH1	-9.31	115.64	120.30
1	A	189	GLY	O-C-N	8.23	137.20	123.20
1	A	189	GLY	CA-C-N	-8.20	99.81	116.20
1	A	191	GLY	O-C-N	-8.18	109.61	122.70
1	A	23	TYR	CB-CG-CD2	6.23	124.74	121.00
1	A	23	TYR	CG-CD2-CE2	5.77	125.91	121.30
1	A	191	GLY	C-N-CA	5.49	135.42	121.70
1	A	191	GLY	CA-C-N	5.19	128.61	117.20
1	A	48	ARG	NE-CZ-NH2	5.17	122.88	120.30

There are no chirality outliers.

There are no planarity outliers.

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	2315	0	2168	28	0
2	A	43	0	30	0	0
3	A	348	0	0	4	0
All	All	2706	0	2198	28	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 6.

All (28) 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:20:GLN:HE22	1:A:285:LYS:H	1.17	0.89
1:A:16:TYR:OH	1:A:285:LYS:HE3	1.94	0.67
1:A:220:GLN:NE2	1:A:228:MET:HG3	2.12	0.64
1:A:127:ARG:HD2	3:A:448:HOH:O	1.97	0.64
1:A:193:ASN:HD22	1:A:194:ASN:H	1.48	0.61
1:A:20:GLN:HE22	1:A:285:LYS:N	1.96	0.58
1:A:6:HIS:HE1	3:A:489:HOH:O	1.88	0.56
1:A:183:LYS:HB2	3:A:689:HOH:O	2.07	0.54
1:A:6:HIS:HD2	3:A:493:HOH:O	1.94	0.51
1:A:228:MET:HG2	1:A:229:MET:N	2.26	0.51
1:A:20:GLN:HE21	1:A:285:LYS:HZ1	1.59	0.50
1:A:20:GLN:HE21	1:A:285:LYS:NZ	2.09	0.49
1:A:93:GLU:O	1:A:97:LYS:HG2	2.12	0.49
1:A:193:ASN:ND2	1:A:194:ASN:H	2.12	0.47
1:A:53:ILE:HG22	1:A:71:TYR:HB2	1.97	0.46
1:A:4:LEU:HB3	1:A:62:ASN:HB3	1.97	0.46
1:A:74:LYS:HA	1:A:77:PHE:CE1	2.50	0.45
1:A:99:PHE:O	1:A:102:ILE:HG22	2.16	0.45
1:A:17:GLU:HG3	1:A:18:ASP:N	2.32	0.45
1:A:187:TYR:HB3	1:A:228:MET:HE1	2.00	0.43
1:A:30:LEU:HD21	1:A:46:LEU:HD12	2.00	0.43
1:A:217:ASN:HD22	1:A:217:ASN:HA	1.65	0.43
1:A:193:ASN:HD22	1:A:194:ASN:N	2.14	0.43
1:A:158:PHE:CG	1:A:168:VAL:HG22	2.54	0.42
1:A:207:GLU:OE1	1:A:225:SER:HB2	2.21	0.41
1:A:187:TYR:HB3	1:A:228:MET:CE	2.51	0.41
1:A:235:SER:HA	1:A:238:GLN:HE21	1.86	0.41
1:A:119:MET:O	1:A:120:GLN:HB2	2.21	0.40

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	287/289 (99%)	284 (99%)	3 (1%)	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	242/247 (98%)	237 (98%)	5 (2%)	53 42

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	35	GLU
1	A	90	LYS
1	A	193	ASN
1	A	213	LYS
1	A	228	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	HIS
1	A	20	GLN
1	A	24	ASN
1	A	78	ASN
1	A	120	GLN
1	A	159	GLN
1	A	162	ASN
1	A	193	ASN
1	A	217	ASN
1	A	218	ASN
1	A	238	GLN
1	A	290	GLN

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)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	A	1	1,3	41,50,50	1.91	17 (41%)	45,82,82	1.88	13 (28%)

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	A	1	1,3	-	2/12/54/54	-

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1	HEM	C1B-NB	-4.46	1.32	1.40
2	A	1	HEM	C3C-CAC	-3.17	1.41	1.47
2	A	1	HEM	C4D-ND	-2.91	1.35	1.40
2	A	1	HEM	C4B-NB	-2.89	1.32	1.38
2	A	1	HEM	CBB-CAB	2.89	1.44	1.30
2	A	1	HEM	CHB-C1B	2.57	1.41	1.35
2	A	1	HEM	CBC-CAC	2.56	1.46	1.29
2	A	1	HEM	C1D-C2D	-2.52	1.39	1.44
2	A	1	HEM	C4D-C3D	-2.47	1.40	1.45
2	A	1	HEM	O2D-CGD	-2.46	1.22	1.30
2	A	1	HEM	FE-NB	2.43	2.08	1.96
2	A	1	HEM	CAB-C3B	-2.42	1.40	1.47
2	A	1	HEM	FE-ND	2.33	2.08	1.96
2	A	1	HEM	C1D-ND	-2.30	1.34	1.38
2	A	1	HEM	CHA-C4D	2.29	1.40	1.35
2	A	1	HEM	O2A-CGA	-2.03	1.23	1.30
2	A	1	HEM	C1B-C2B	-2.00	1.40	1.44

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1	HEM	C4B-CHC-C1C	5.51	129.83	122.56
2	A	1	HEM	C4C-CHD-C1D	4.91	129.04	122.56
2	A	1	HEM	C4A-C3A-C2A	-3.38	104.64	107.00
2	A	1	HEM	CMB-C2B-C1B	-3.12	120.29	125.04
2	A	1	HEM	C1B-NB-C4B	2.94	108.11	105.07
2	A	1	HEM	CBA-CAA-C2A	2.73	117.29	112.62
2	A	1	HEM	CHD-C1D-ND	-2.70	121.49	124.43
2	A	1	HEM	O2A-CGA-O1A	-2.36	117.41	123.30
2	A	1	HEM	CMB-C2B-C3B	2.30	133.94	128.30
2	A	1	HEM	CAD-C3D-C4D	2.29	128.66	124.66
2	A	1	HEM	C4B-C3B-C2B	-2.04	105.50	107.11
2	A	1	HEM	CAD-CBD-CGD	2.03	117.97	113.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1	HEM	O1D-CGD-CBD	-2.02	116.59	123.08

There are no chirality outliers.

All (2) torsion outliers are listed below:

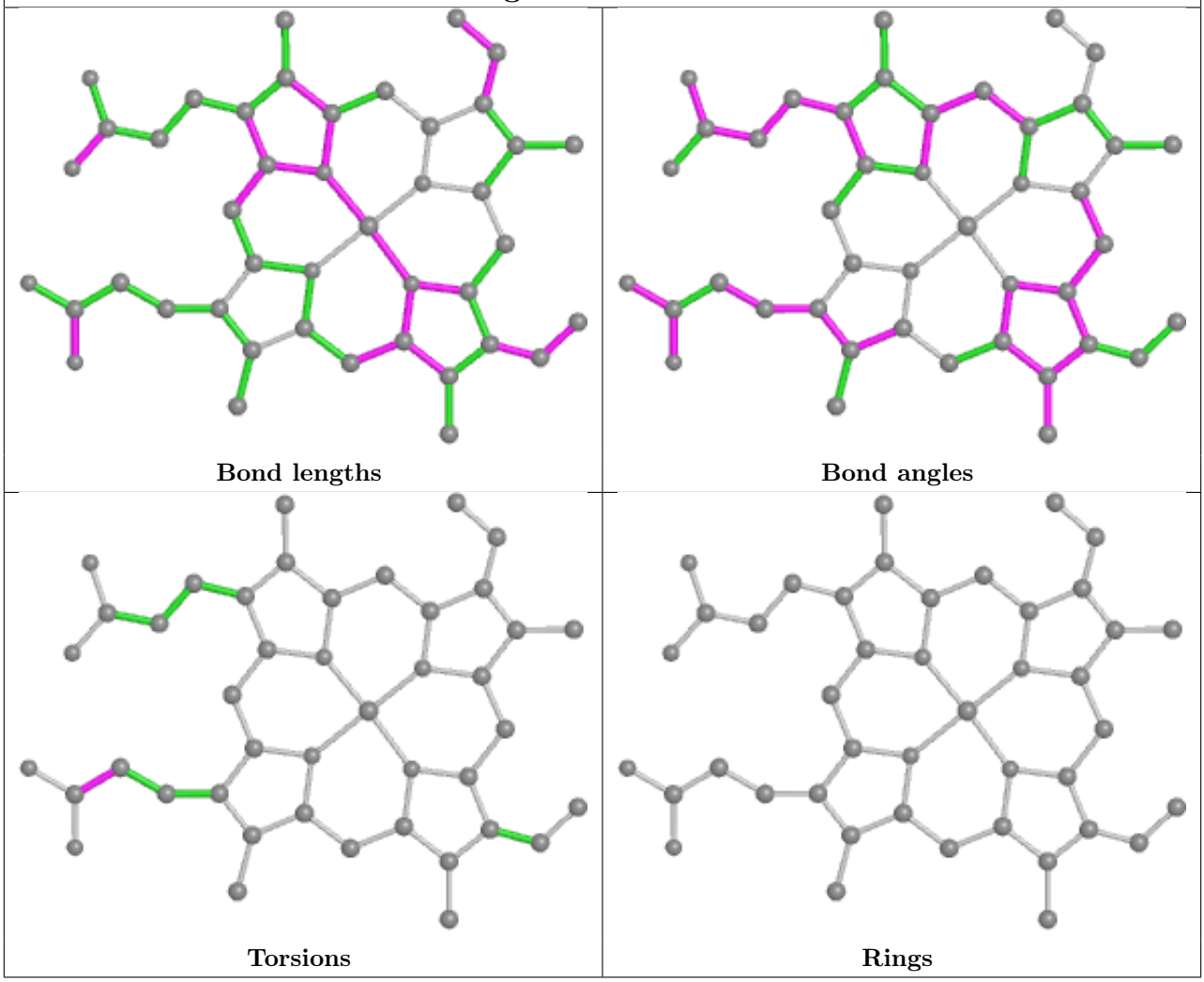
Mol	Chain	Res	Type	Atoms
2	A	1	HEM	CAA-CBA-CGA-O2A
2	A	1	HEM	CAA-CBA-CGA-O1A

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand HEM A 1



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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.