



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 12, 2024 – 07:42 AM EST

PDB ID : 2HLZ
Title : Crystal Structure of human ketohexokinase
Authors : Rabeh, W.M.; Tempel, W.; Nedyalkova, L.; Landry, R.; Arrowsmith, C.H.;
Edwards, A.M.; Sundstrom, M.; Weigelt, J.; Bochkarev, A.; Park, H.; Structural Genomics Consortium (SGC)
Deposited on : 2006-07-10
Resolution : 1.85 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtrriage (Phenix) : 1.20.1
EDS : **FAILED**
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

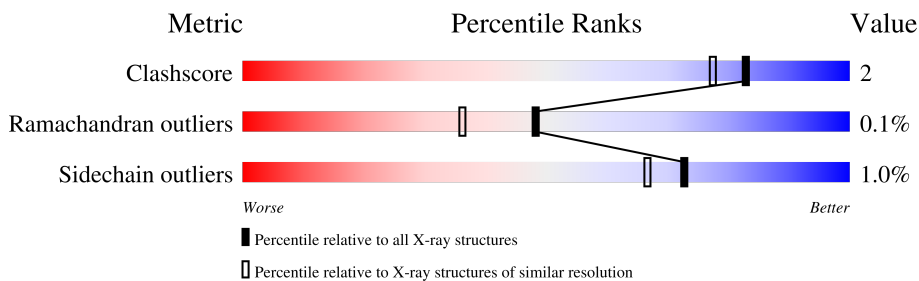
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	180529	3359 (1.86-1.86)
Ramachandran outliers	177936	3335 (1.86-1.86)
Sidechain outliers	177891	3335 (1.86-1.86)

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 $\leq 5\%$. 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 failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	312	
1	B	312	
1	C	312	
1	D	312	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9449 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 Ketohexokinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	296	2229	1418	383	418	7	3	0	1	0
1	B	296	2238	1421	383	424	7	3	0	2	0
1	C	302	2284	1450	397	427	7	3	0	1	0
1	D	301	2294	1457	397	430	7	3	0	3	0

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	GLY	-	cloning artifact	UNP P50053
A	-12	SER	-	cloning artifact	UNP P50053
A	-11	SER	-	cloning artifact	UNP P50053
A	-10	HIS	-	expression tag	UNP P50053
A	-9	HIS	-	expression tag	UNP P50053
A	-8	HIS	-	expression tag	UNP P50053
A	-7	HIS	-	expression tag	UNP P50053
A	-6	HIS	-	expression tag	UNP P50053
A	-5	HIS	-	expression tag	UNP P50053
A	-4	SER	-	cloning artifact	UNP P50053
A	-3	SER	-	cloning artifact	UNP P50053
A	-2	GLY	-	cloning artifact	UNP P50053
A	-1	LEU	-	cloning artifact	UNP P50053
A	0	VAL	-	cloning artifact	UNP P50053
A	1	PRO	-	cloning artifact	UNP P50053
A	2	ARG	-	cloning artifact	UNP P50053
A	3	GLY	-	cloning artifact	UNP P50053
A	4	SER	-	cloning artifact	UNP P50053
A	49	ILE	VAL	conflict	UNP P50053
A	60	MSE	MET	modified residue	UNP P50053
A	63	MSE	MET	modified residue	UNP P50053

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	149	MSE	MET	modified residue	UNP P50053
B	-13	GLY	-	cloning artifact	UNP P50053
B	-12	SER	-	cloning artifact	UNP P50053
B	-11	SER	-	cloning artifact	UNP P50053
B	-10	HIS	-	expression tag	UNP P50053
B	-9	HIS	-	expression tag	UNP P50053
B	-8	HIS	-	expression tag	UNP P50053
B	-7	HIS	-	expression tag	UNP P50053
B	-6	HIS	-	expression tag	UNP P50053
B	-5	HIS	-	expression tag	UNP P50053
B	-4	SER	-	cloning artifact	UNP P50053
B	-3	SER	-	cloning artifact	UNP P50053
B	-2	GLY	-	cloning artifact	UNP P50053
B	-1	LEU	-	cloning artifact	UNP P50053
B	0	VAL	-	cloning artifact	UNP P50053
B	1	PRO	-	cloning artifact	UNP P50053
B	2	ARG	-	cloning artifact	UNP P50053
B	3	GLY	-	cloning artifact	UNP P50053
B	4	SER	-	cloning artifact	UNP P50053
B	49	ILE	VAL	conflict	UNP P50053
B	60	MSE	MET	modified residue	UNP P50053
B	63	MSE	MET	modified residue	UNP P50053
B	149	MSE	MET	modified residue	UNP P50053
C	-13	GLY	-	cloning artifact	UNP P50053
C	-12	SER	-	cloning artifact	UNP P50053
C	-11	SER	-	cloning artifact	UNP P50053
C	-10	HIS	-	expression tag	UNP P50053
C	-9	HIS	-	expression tag	UNP P50053
C	-8	HIS	-	expression tag	UNP P50053
C	-7	HIS	-	expression tag	UNP P50053
C	-6	HIS	-	expression tag	UNP P50053
C	-5	HIS	-	expression tag	UNP P50053
C	-4	SER	-	cloning artifact	UNP P50053
C	-3	SER	-	cloning artifact	UNP P50053
C	-2	GLY	-	cloning artifact	UNP P50053
C	-1	LEU	-	cloning artifact	UNP P50053
C	0	VAL	-	cloning artifact	UNP P50053
C	1	PRO	-	cloning artifact	UNP P50053
C	2	ARG	-	cloning artifact	UNP P50053
C	3	GLY	-	cloning artifact	UNP P50053
C	4	SER	-	cloning artifact	UNP P50053
C	49	ILE	VAL	conflict	UNP P50053

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	60	MSE	MET	modified residue	UNP P50053
C	63	MSE	MET	modified residue	UNP P50053
C	149	MSE	MET	modified residue	UNP P50053
D	-13	GLY	-	cloning artifact	UNP P50053
D	-12	SER	-	cloning artifact	UNP P50053
D	-11	SER	-	cloning artifact	UNP P50053
D	-10	HIS	-	expression tag	UNP P50053
D	-9	HIS	-	expression tag	UNP P50053
D	-8	HIS	-	expression tag	UNP P50053
D	-7	HIS	-	expression tag	UNP P50053
D	-6	HIS	-	expression tag	UNP P50053
D	-5	HIS	-	expression tag	UNP P50053
D	-4	SER	-	cloning artifact	UNP P50053
D	-3	SER	-	cloning artifact	UNP P50053
D	-2	GLY	-	cloning artifact	UNP P50053
D	-1	LEU	-	cloning artifact	UNP P50053
D	0	VAL	-	cloning artifact	UNP P50053
D	1	PRO	-	cloning artifact	UNP P50053
D	2	ARG	-	cloning artifact	UNP P50053
D	3	GLY	-	cloning artifact	UNP P50053
D	4	SER	-	cloning artifact	UNP P50053
D	49	ILE	VAL	conflict	UNP P50053
D	60	MSE	MET	modified residue	UNP P50053
D	63	MSE	MET	modified residue	UNP P50053
D	149	MSE	MET	modified residue	UNP P50053

- Molecule 2 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	8	Total X 8 8	0	0
2	B	8	Total X 8 8	0	0
2	C	8	Total X 8 8	0	0
2	D	8	Total X 8 8	0	0

- Molecule 3 is water.

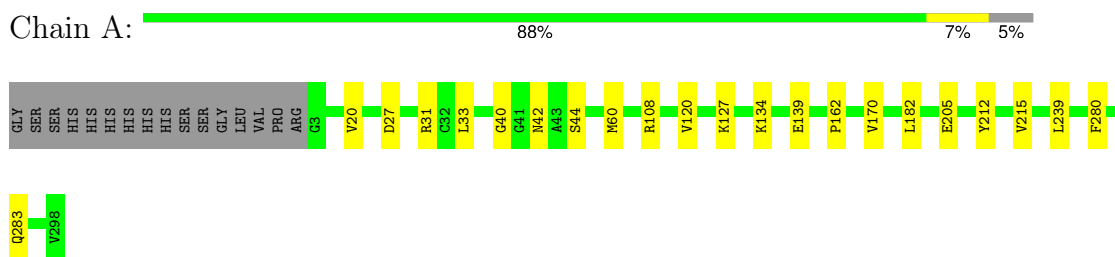
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	67	Total O 67 67	0	0
3	B	86	Total O 86 86	0	0
3	C	105	Total O 105 105	0	0
3	D	114	Total O 114 114	0	0

3 Residue-property plots [i](#)

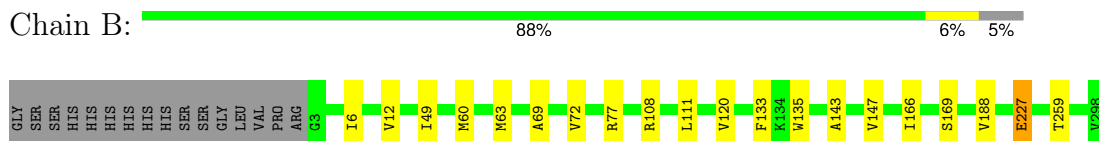
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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 failed to run properly.

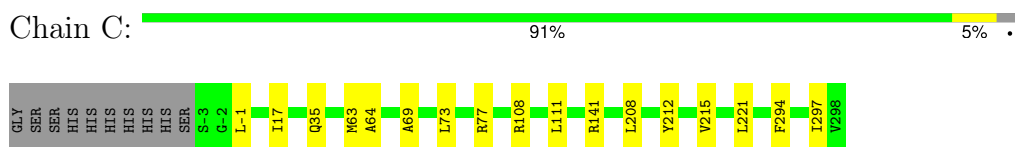
- Molecule 1: Ketohexokinase



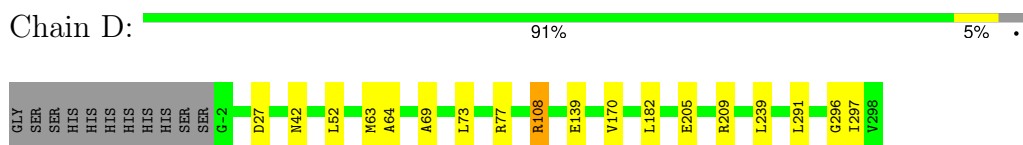
- Molecule 1: Ketohexokinase



- Molecule 1: Ketohexokinase



- Molecule 1: Ketohexokinase



4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	66.93Å 71.67Å 82.69Å 105.06° 107.33° 93.09°	Depositor
Resolution (Å)	29.66 – 1.85	Depositor
% Data completeness (in resolution range)	94.6 (29.66-1.85)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.29 (at 1.85Å)	Xtrriage
Refinement program	REFMAC refmac_5.2.0019	Depositor
R, R_{free}	0.214 , 0.249	Depositor
Wilson B-factor (Å ²)	18.8	Xtrriage
Anisotropy	0.003	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	9449	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.55% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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: UNX

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	0.73	0/2270	0.68	0/3071
1	B	0.77	0/2283	0.71	1/3091 (0.0%)
1	C	0.78	0/2329	0.72	1/3153 (0.0%)
1	D	0.78	0/2345	0.73	0/3175
All	All	0.77	0/9227	0.71	2/12490 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	77	ARG	NE-CZ-NH1	5.49	123.05	120.30
1	B	77	ARG	NE-CZ-NH1	5.06	122.83	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	2229	0	2204	11	0
1	B	2238	0	2197	13	0
1	C	2284	0	2250	9	0
1	D	2294	0	2265	11	0
2	A	8	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	8	0	0	0	0
2	C	8	0	0	0	0
2	D	8	0	0	0	0
3	A	67	0	0	0	0
3	B	86	0	0	0	0
3	C	105	0	0	0	0
3	D	114	0	0	0	0
All	All	9449	0	8916	42	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 2.

All (42) 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:B:60:MSE:HE2	1:B:120:VAL:HG11	1.60	0.83
1:A:205:GLU:HG3	1:A:239:LEU:HD21	1.69	0.73
1:D:205:GLU:HG3	1:D:239:LEU:HD13	1.74	0.69
1:B:60:MSE:HE2	1:B:120:VAL:CG1	2.22	0.68
1:B:12:VAL:HG21	1:B:72[B]:VAL:HG11	1.83	0.60
1:C:64:ALA:O	1:C:69:ALA:HB2	2.02	0.60
1:B:12:VAL:HG21	1:B:72[B]:VAL:CG1	2.35	0.56
1:D:170[A]:VAL:HG11	1:D:182:LEU:HB3	1.87	0.56
1:C:294:PHE:O	1:C:297:ILE:HG23	2.06	0.56
1:A:127:LYS:HB2	1:C:-1:LEU:HD23	1.88	0.56
1:B:227[A]:GLU:H	1:B:227[A]:GLU:CD	2.09	0.55
1:A:20:VAL:HA	1:A:33[B]:LEU:HD13	1.91	0.53
1:A:212:TYR:O	1:A:215:VAL:HG22	2.09	0.53
1:C:212:TYR:O	1:C:215:VAL:HG22	2.09	0.51
1:A:42:ASN:HB2	1:A:139:GLU:CD	2.32	0.49
1:A:60:MSE:HE2	1:A:120:VAL:CG1	2.43	0.49
1:B:69:ALA:HA	1:B:72[A]:VAL:HG22	1.95	0.48
1:B:111:LEU:HD12	1:B:111:LEU:N	2.29	0.47
1:A:280:PHE:HA	1:A:283:GLN:HE21	1.81	0.46
1:A:60:MSE:HE2	1:A:120:VAL:HG11	1.98	0.45
1:C:111:LEU:HD12	1:C:111:LEU:N	2.31	0.45
1:A:27:ASP:HA	1:D:108:ARG:O	2.16	0.45
1:B:63:MSE:HE1	1:B:72[B]:VAL:HG12	1.99	0.45
1:D:73:LEU:O	1:D:77:ARG:HG3	2.17	0.44
1:D:64:ALA:O	1:D:69:ALA:HB2	2.17	0.44
1:A:170:VAL:HG11	1:A:182:LEU:HB3	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:63:MSE:HE3	1:C:73:LEU:CD2	2.48	0.44
1:D:63:MSE:HE3	1:D:73:LEU:HD21	1.99	0.43
1:D:205:GLU:HG3	1:D:239:LEU:CD1	2.44	0.43
1:C:63:MSE:HE3	1:C:73:LEU:HD21	2.00	0.43
1:D:209:ARG:NH2	1:D:239:LEU:HD12	2.34	0.43
1:B:133:PHE:O	1:B:166:ILE:HG12	2.19	0.42
1:D:291:LEU:HD13	1:D:296:GLY:HA3	2.02	0.42
1:C:17:ILE:O	1:C:35:GLN:HA	2.20	0.42
1:C:208:LEU:HD23	1:C:221:LEU:HG	2.02	0.41
1:B:143:ALA:O	1:B:147:VAL:HG23	2.21	0.41
1:D:52:LEU:HD13	1:D:297:ILE:HG12	2.02	0.41
1:D:42:ASN:HB2	1:D:139:GLU:CD	2.41	0.41
1:B:6:ILE:HG12	1:B:135:TRP:HB3	2.03	0.41
1:A:40:GLY:O	1:A:44:SER:HB3	2.21	0.40
1:B:49:ILE:HG13	1:B:259:THR:HA	2.04	0.40
1:B:169:SER:HA	1:B:188:VAL:O	2.20	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	295/312 (95%)	291 (99%)	3 (1%)	1 (0%)	37	25
1	B	296/312 (95%)	289 (98%)	7 (2%)	0	100	100
1	C	301/312 (96%)	299 (99%)	2 (1%)	0	100	100
1	D	302/312 (97%)	298 (99%)	4 (1%)	0	100	100
All	All	1194/1248 (96%)	1177 (99%)	16 (1%)	1 (0%)	48	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	162	PRO

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	232/257 (90%)	229 (99%)	3 (1%)	65	55
1	B	234/257 (91%)	231 (99%)	3 (1%)	65	55
1	C	239/257 (93%)	237 (99%)	2 (1%)	79	74
1	D	242/257 (94%)	240 (99%)	2 (1%)	79	74
All	All	947/1028 (92%)	937 (99%)	10 (1%)	73	62

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	ARG
1	A	108	ARG
1	A	134	LYS
1	B	108	ARG
1	B	227[A]	GLU
1	B	227[B]	GLU
1	C	108	ARG
1	C	141	ARG
1	D	27	ASP
1	D	108	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	283	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 32 ligands modelled in this entry, 32 are unknown - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.