



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 17, 2021 – 04:47 AM EDT

PDB ID : 1K9E  
Title : Crystal structure of a mutated family-67 alpha-D-glucuronidase (E285N) from Bacillus stearothermophilus T-6, complexed with 4-O-methyl-glucuronic acid  
Authors : Golan, G.; Shallom, D.; Teplitsky, A.; Zaide, G.; Shulami, S.; Baasov, T.; Stojanoff, V.; Thompson, A.; Shoham, Y.; Shoham, G.  
Deposited on : 2001-10-29  
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](mailto: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 following versions of software and data (see [references ⓘ](#)) 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  
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) : 2.23.2

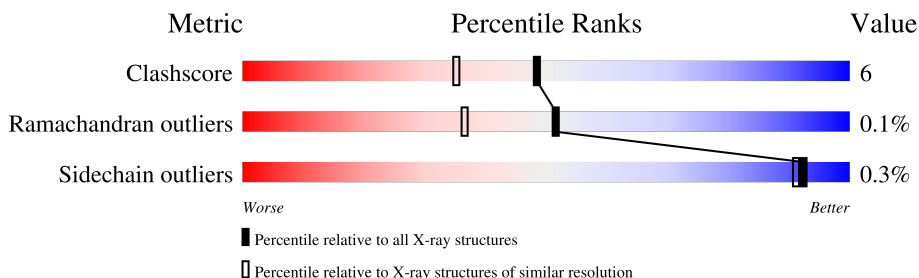
# 1 Overall quality at a glance

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	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	679	

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	GCV	A	701	X	-	-	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6175 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.

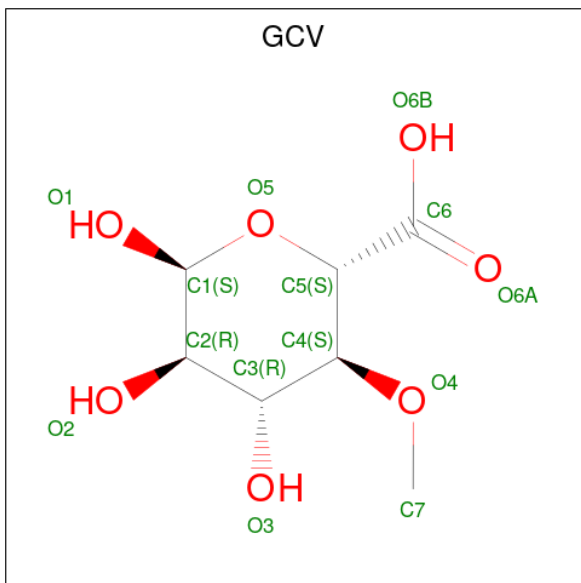
- Molecule 1 is a protein called alpha-D-glucuronidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	671	5612	3578	962	1050	22	0	20	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	285	ASN	GLU	engineered mutation	UNP Q8VVD2

- Molecule 2 is 4-O-methyl-alpha-D-glucopyranuronic acid (three-letter code: GCV) (formula:  $C_7H_{12}O_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	14	7	7	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

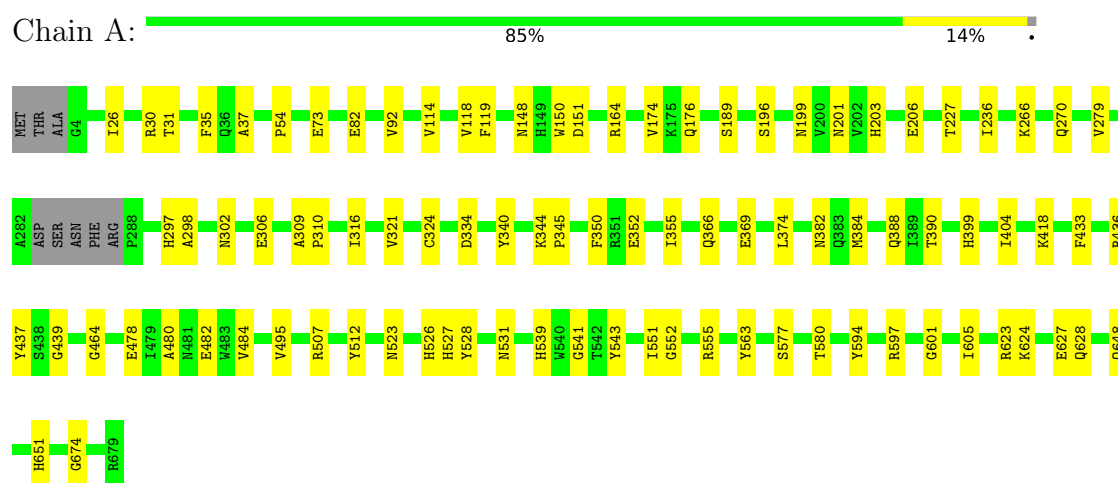
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	483	Total	O	0	0
			483	483		

### 3 Residue-property plots

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: alpha-D-glucuronidase



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.58Å 73.58Å 330.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.03 – 1.85	Depositor
% Data completeness (in resolution range)	92.1 (36.03-1.85)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.193 , 0.218	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6175	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	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: GOL, GCV

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	0.31	0/5763	0.59	0/7819

There are no bond length outliers.

There are no bond angle outliers.

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	5612	0	5345	67	0
2	A	14	0	9	1	0
3	A	66	0	88	3	0
4	A	483	0	0	5	0
All	All	6175	0	5442	67	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 (67) 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:399:HIS:HE1	1:A:651:HIS:HD2	1.28	0.81
1:A:174:VAL:HG12	1:A:176:GLN:H	1.45	0.80
1:A:399:HIS:CE1	1:A:651:HIS:HD2	2.08	0.71
1:A:399:HIS:HE1	1:A:651:HIS:CD2	2.12	0.68
1:A:526:HIS:HD2	1:A:528:TYR:H	1.41	0.67
1:A:316[A]:ILE:HD12	1:A:384[A]:MET:HE1	1.78	0.64
1:A:316[B]:ILE:HD12	1:A:384[B]:MET:HE1	1.78	0.64
1:A:526:HIS:CD2	1:A:528:TYR:H	2.16	0.63
1:A:399:HIS:CD2	1:A:648:GLN:HG3	2.36	0.60
1:A:344:LYS:HB3	1:A:345:PRO:HD3	1.86	0.58
1:A:399:HIS:HD2	1:A:512:TYR:OH	1.87	0.57
1:A:150:TRP:HA	1:A:199:ASN:HA	1.87	0.56
1:A:30:ARG:HD3	4:A:1406:HOH:O	2.05	0.56
1:A:355:ILE:HD11	1:A:437:TYR:HB3	1.87	0.56
1:A:164:ARG:HD3	4:A:1290:HOH:O	2.05	0.55
1:A:555:ARG:HB3	1:A:563:TYR:HB3	1.90	0.54
1:A:418:LYS:HE3	1:A:433:PHE:CE1	2.43	0.54
1:A:388:GLN:HG2	1:A:390:THR:O	2.09	0.53
1:A:266:LYS:O	1:A:270[B]:GLN:HG3	2.09	0.53
1:A:266:LYS:O	1:A:270[A]:GLN:HG3	2.09	0.53
1:A:366:GLN:HB2	1:A:369:GLU:HG3	1.92	0.52
1:A:541:GLY:HA2	1:A:543:TYR:CE2	2.44	0.52
1:A:555:ARG:NH2	3:A:806:GOL:H12	2.24	0.52
1:A:297:HIS:HE1	1:A:321:VAL:HG21	1.76	0.51
1:A:324:CYS:O	1:A:539:HIS:HD2	1.94	0.51
1:A:31:THR:HG22	4:A:1497:HOH:O	2.12	0.50
1:A:201:ASN:ND2	2:A:701:GCV:H71	2.28	0.49
1:A:37:ALA:HB2	1:A:227:THR:CG2	2.43	0.49
1:A:399:HIS:CE1	1:A:651:HIS:CD2	2.95	0.49
1:A:551:ILE:HG12	1:A:552:GLY:N	2.28	0.48
1:A:594:TYR:HB3	1:A:605:ILE:HB	1.95	0.47
1:A:478:GLU:O	1:A:482:GLU:HG3	2.14	0.47
1:A:623:ARG:O	1:A:627:GLU:HG3	2.14	0.47
1:A:298:ALA:HA	1:A:350:PHE:CE1	2.50	0.47
1:A:484:VAL:HG11	1:A:495:VAL:HA	1.96	0.47
1:A:651:HIS:HE1	4:A:1306:HOH:O	1.98	0.46
1:A:577:SER:HB3	1:A:580:THR:OG1	2.15	0.46
1:A:418:LYS:HE3	1:A:433:PHE:CD1	2.51	0.46
1:A:279:VAL:HA	1:A:316[B]:ILE:HB	1.98	0.46
1:A:309:ALA:HB3	1:A:310:PRO:HD3	1.98	0.45
1:A:119:PHE:CG	1:A:189:SER:HA	2.52	0.45
1:A:236:ILE:HD12	1:A:236:ILE:C	2.36	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:507:ARG:HH11	1:A:507:ARG:HG3	1.82	0.45
1:A:279:VAL:HA	1:A:316[A]:ILE:HB	1.98	0.45
1:A:624:LYS:O	1:A:628:GLN:HG3	2.17	0.45
1:A:597:ARG:HD2	1:A:601:GLY:O	2.17	0.44
1:A:384[B]:MET:HG3	1:A:439:GLY:C	2.38	0.44
1:A:334:ASP:OD1	3:A:806:GOL:H31	2.18	0.44
1:A:37:ALA:HB2	1:A:227:THR:HG22	2.00	0.44
1:A:26:ILE:HG13	1:A:54[B]:PRO:HB2	2.00	0.44
1:A:384[A]:MET:HG3	1:A:439:GLY:C	2.38	0.43
1:A:148:ASN:HA	1:A:196:SER:OG	2.18	0.43
1:A:523:ASN:O	1:A:527:HIS:HA	2.19	0.43
1:A:26:ILE:HG13	1:A:54[A]:PRO:HB2	2.00	0.43
1:A:352:GLU:HG3	1:A:436:ARG:HH12	1.83	0.42
1:A:404:ILE:HD11	1:A:464:GLY:HA3	2.01	0.42
1:A:82:GLU:HA	1:A:92:VAL:HG21	2.02	0.42
1:A:114:VAL:O	1:A:118:VAL:HG23	2.19	0.42
1:A:340:TYR:HA	1:A:374:LEU:HD21	2.01	0.42
1:A:382:ASN:HB3	1:A:437:TYR:O	2.19	0.42
1:A:73[B]:GLU:H	1:A:73[B]:GLU:CD	2.23	0.41
1:A:302:ASN:O	1:A:306:GLU:HG3	2.21	0.41
1:A:480:ALA:O	1:A:484:VAL:HG23	2.21	0.41
1:A:507:ARG:HG3	1:A:507:ARG:NH1	2.36	0.41
1:A:203:HIS:HB2	1:A:206:GLU:OE1	2.21	0.40
1:A:352:GLU:HG2	3:A:805:GOL:O3	2.21	0.40
1:A:674:GLY:HA2	4:A:1492:HOH: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	687/679 (101%)	662 (96%)	24 (4%)	1 (0%)	51	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	531	ASN

### 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	593/587 (101%)	591 (100%)	2 (0%)	92	91

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	35	PHE
1	A	151	ASP

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

Mol	Chain	Res	Type
1	A	148	ASN
1	A	357	GLN
1	A	399	HIS
1	A	526	HIS
1	A	639	HIS
1	A	651	HIS

### 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](#)

12 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	GOL	A	804	-	5,5,5	0.78	0	5,5,5	0.41	0
3	GOL	A	810	-	5,5,5	0.75	0	5,5,5	0.41	0
3	GOL	A	806	-	5,5,5	0.76	0	5,5,5	0.40	0
3	GOL	A	809	-	5,5,5	0.73	0	5,5,5	0.42	0
3	GOL	A	803	-	5,5,5	0.76	0	5,5,5	0.39	0
3	GOL	A	811	-	5,5,5	0.76	0	5,5,5	0.42	0
3	GOL	A	805	-	5,5,5	0.79	0	5,5,5	0.42	0
3	GOL	A	807	-	5,5,5	0.90	0	5,5,5	0.53	0
3	GOL	A	801	-	5,5,5	0.71	0	5,5,5	0.44	0
3	GOL	A	808	-	5,5,5	0.78	0	5,5,5	0.42	0
3	GOL	A	802	-	5,5,5	0.73	0	5,5,5	0.43	0
2	GCV	A	701	-	11,14,14	1.84	2 (18%)	14,20,20	6.47	9 (64%)

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	GOL	A	804	-	-	0/4/4/4	-
3	GOL	A	810	-	-	2/4/4/4	-
3	GOL	A	806	-	-	2/4/4/4	-
3	GOL	A	809	-	-	0/4/4/4	-
3	GOL	A	803	-	-	1/4/4/4	-
3	GOL	A	811	-	-	2/4/4/4	-
3	GOL	A	805	-	-	0/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	807	-	-	2/4/4/4	-
3	GOL	A	801	-	-	0/4/4/4	-
3	GOL	A	808	-	-	0/4/4/4	-
3	GOL	A	802	-	-	0/4/4/4	-
2	GCV	A	701	-	1/1/6/6	0/2/26/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	GCV	C3-C2	3.93	1.62	1.52
2	A	701	GCV	C4-C5	-3.53	1.46	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	701	GCV	O5-C5-C4	20.37	137.06	108.72
2	A	701	GCV	C3-C4-C5	-6.70	101.68	110.28
2	A	701	GCV	O3-C3-C2	-6.39	95.57	110.35
2	A	701	GCV	C2-C3-C4	5.58	122.43	109.68
2	A	701	GCV	O5-C1-C2	4.08	117.56	110.28
2	A	701	GCV	C1-C2-C3	-3.30	103.46	110.31
2	A	701	GCV	O2-C2-C3	2.52	116.17	110.35
2	A	701	GCV	C7-O4-C4	-2.44	108.13	114.52
2	A	701	GCV	O3-C3-C4	2.37	116.23	109.94

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	701	GCV	C5

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	806	GOL	C1-C2-C3-O3
3	A	810	GOL	O1-C1-C2-C3
3	A	807	GOL	C1-C2-C3-O3
3	A	806	GOL	O2-C2-C3-O3
3	A	807	GOL	O2-C2-C3-O3
3	A	810	GOL	O1-C1-C2-O2
3	A	811	GOL	O1-C1-C2-C3
3	A	811	GOL	O1-C1-C2-O2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	A	803	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	806	GOL	2	0
3	A	805	GOL	1	0
2	A	701	GCV	1	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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.