W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 18, 2021 – 08:01 am BST

PDB ID	:	7NQ4
Title	:	Human tRNA guanine transglycosylase (TGT), RNA-bound covalent intermediate
Authors	:	Sievers, K.; Ficner, R.
Deposited on	:	2021-03-01
Resolution	:	2.88 Å(reported)

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

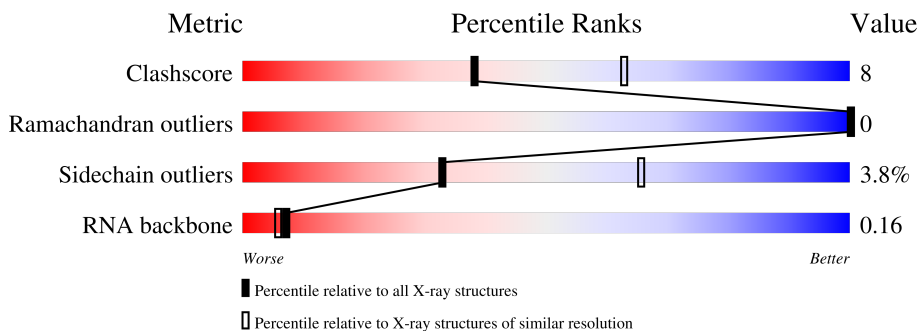|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

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 2.88 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
Clashscore		8
Ramachandran outliers		0
Sidechain outliers		3.8%
RNA backbone		0.16

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))
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RNA backbone	3102	1121 (3.16-2.60)

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	405	77% · 19% ·
2	B	415	74% · 16% · 8%
3	C	20	55% · 35% · 10%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6396 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 Queuine tRNA-ribosyltransferase catalytic subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	0	0	0
			2989	1880	531	552	26			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q9BXR0
A	0	PRO	-	expression tag	UNP Q9BXR0

- Molecule 2 is a protein called Queuine tRNA-ribosyltransferase accessory subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	381	Total	C	N	O	S	0	0	0
			2976	1885	515	554	22			

- Molecule 3 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	20	Total	C	N	O	P	0	0	0
			414	186	74	135	19			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Zn	0	0
			1	1		
4	B	1	Total	Zn	0	0
			1	1		

- Molecule 5 is 9-DEAZAGUANINE (three-letter code: 9DG) (formula: $C_6H_6N_4O$) (labeled as "Ligand of Interest" by depositor).
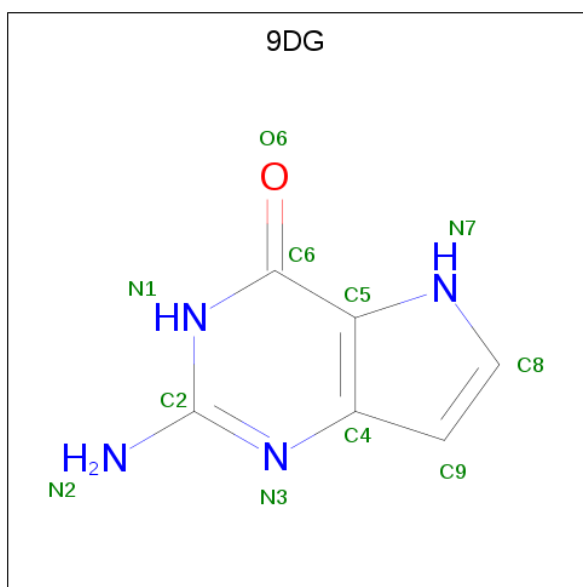
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 11	C 6	N 4	O 1	0	0

- Molecule 6 is water.

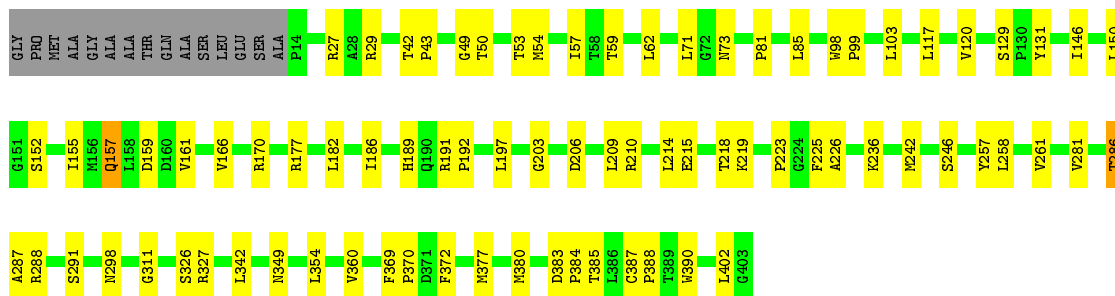| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total 1	O 1	0	0
6	B	3	Total 3	O 3	0	0

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.
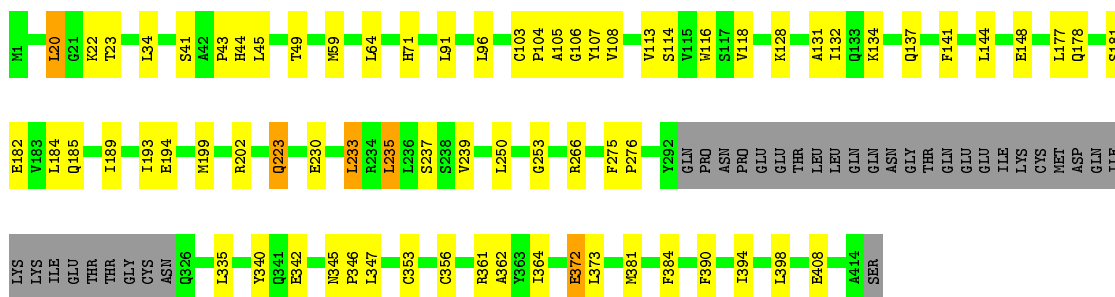
Note EDS was not executed.

• Molecule 1: Queuine tRNA-ribosyltransferase catalytic subunit 1

Chain A:

77%   19%

GLY PRO MET ALA GLY ALA ALA THR GLN ALA SER LEU GLU SER ALA P14 R27 A28 R29 T42 P43 G49 T50 T53 M54 I57 T58 T59 L62 L71 G72 N73 P81 L85 W98 P99 L103 L117 V120 S129 P130 Y131 I146 L150

G151 S152 I155 M156 Q157 L158 D159 D160 V161 V166 R170 R177 L182 I186 H189 Q190 R191 P192 L197 G203 D206 L209 R210 L214 E215 T218 K219 P223 G224 F225 A226 K236 M242 S246 Y257 L258 V261 V281 T286

A287 R288 S291 N298 G311 S326 R327 L342 N349 L354 V360 F369 P370 D371 F372 M377 M380 D383 P384 T385 L386 C387 P388 T389 W390 L402 G403

• Molecule 2: Queuine tRNA-ribosyltransferase accessory subunit 2

Chain B:

74%   16%   8%

M1 L20 G21 K22 T23 L34 S41 A42 P43 H44 L45 T49 M59 L64 H71 L91 L96 C103 P104 A105 G106 Y107 V108 V113 S114 V115 W116 S117 V118 K128 A131 I132 Q133 K134 Q137 F141 L144 E148 L177 Q178 S181

E182 V183 L184 Q185 I189 I193 E194 M199 R202 Q223 E230 L233 R234 L235 L236 S237 S238 V239 L250 G253 R266 F275 P276 Y292 GLN PRO ASN PRO GLU GLU THR LEU LEU GLN GLN ASN GLY THR GLN GLU GLU ILE LYS CYS MET ASP GLN ILE

LYS LYS ILE GLU THR THR GLY CYS ASN Q326 L335 Y340 Q341 E342 N345 P346 L347 C353 C356 R361 A362 Y363 I364 E372 L373 M381 F384 F390 I394 L398 E408 A414 SER

• Molecule 3: RNA

Chain C:

55%   35%   10%

A25 G26 C32 U33 G34 U35 A36 A37 A38 G43 C44

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	161.84Å 56.96Å 102.96Å 90.00° 124.94° 90.00°	Depositor
Resolution (Å)	42.20 – 2.88	Depositor
% Data completeness (in resolution range)	99.2 (42.20-2.88)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC v5.8.0267, PHENIX v1.19.1_4122	Depositor
R, R_{free}	0.212 , 0.248	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6396	wwPDB-VP
Average B, all atoms (Å^2)	88.0	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: ZN, 9DG

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	0.33	0/3054	0.63	0/4144
2	B	0.32	0/3036	0.60	0/4104
3	C	0.59	0/462	0.88	0/718
All	All	0.35	0/6552	0.64	0/8966

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	2989	0	2963	48	0
2	B	2976	0	2947	48	0
3	C	414	0	214	5	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	11	0	6	0	0
6	A	1	0	0	0	0
6	B	3	0	0	0	0
All	All	6396	0	6130	95	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 8.

The worst 5 of 95 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)
2:B:141:PHE:CZ	2:B:189:ILE:HG12	1.84	1.12
1:A:384:PRO:HB3	1:A:402:LEU:HB3	1.46	0.97
2:B:237:SER:HB3	2:B:266:ARG:HG2	1.49	0.93
1:A:281:VAL:HG11	3:C:35:U:H5'	1.61	0.82
1:A:206:ASP:OD2	1:A:209:LEU:HG	1.83	0.79

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	388/405 (96%)	376 (97%)	12 (3%)	0	100	100
2	B	377/415 (91%)	360 (96%)	17 (4%)	0	100	100
All	All	765/820 (93%)	736 (96%)	29 (4%)	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	324/335 (97%)	313 (97%)	11 (3%)	37	69
2	B	328/368 (89%)	314 (96%)	14 (4%)	29	60
All	All	652/703 (93%)	627 (96%)	25 (4%)	33	65

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	41	SER
2	B	182	GLU
2	B	408	GLU
2	B	118	VAL
2	B	223	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	335	HIS
2	B	71	HIS
2	B	392	HIS
2	B	223	GLN
1	A	148	ASN

5.3.3   RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	C	19/20 (95%)	6 (31%)	1 (5%)

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	C	26	G
3	C	34	G
3	C	36	A
3	C	37	A
3	C	38	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	C	32	C

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)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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
5	9DG	A	502	-	10,12,12	1.37	1 (10%)	9,17,17	3.31	3 (33%)

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
5	9DG	A	502	-	-	-	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	502	9DG	C6-N1	3.72	1.39	1.33

All (3) bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	502	9DG	C5-C6-N1	-7.49	113.18	123.43
5	A	502	9DG	C6-N1-C2	5.09	124.02	115.93
5	A	502	9DG	N3-C2-N1	-3.47	122.59	127.22

There are no chirality outliers.

There are no torsion outliers.

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 9DG A 502

Bond lengths

Bond angles

Torsions

Rings
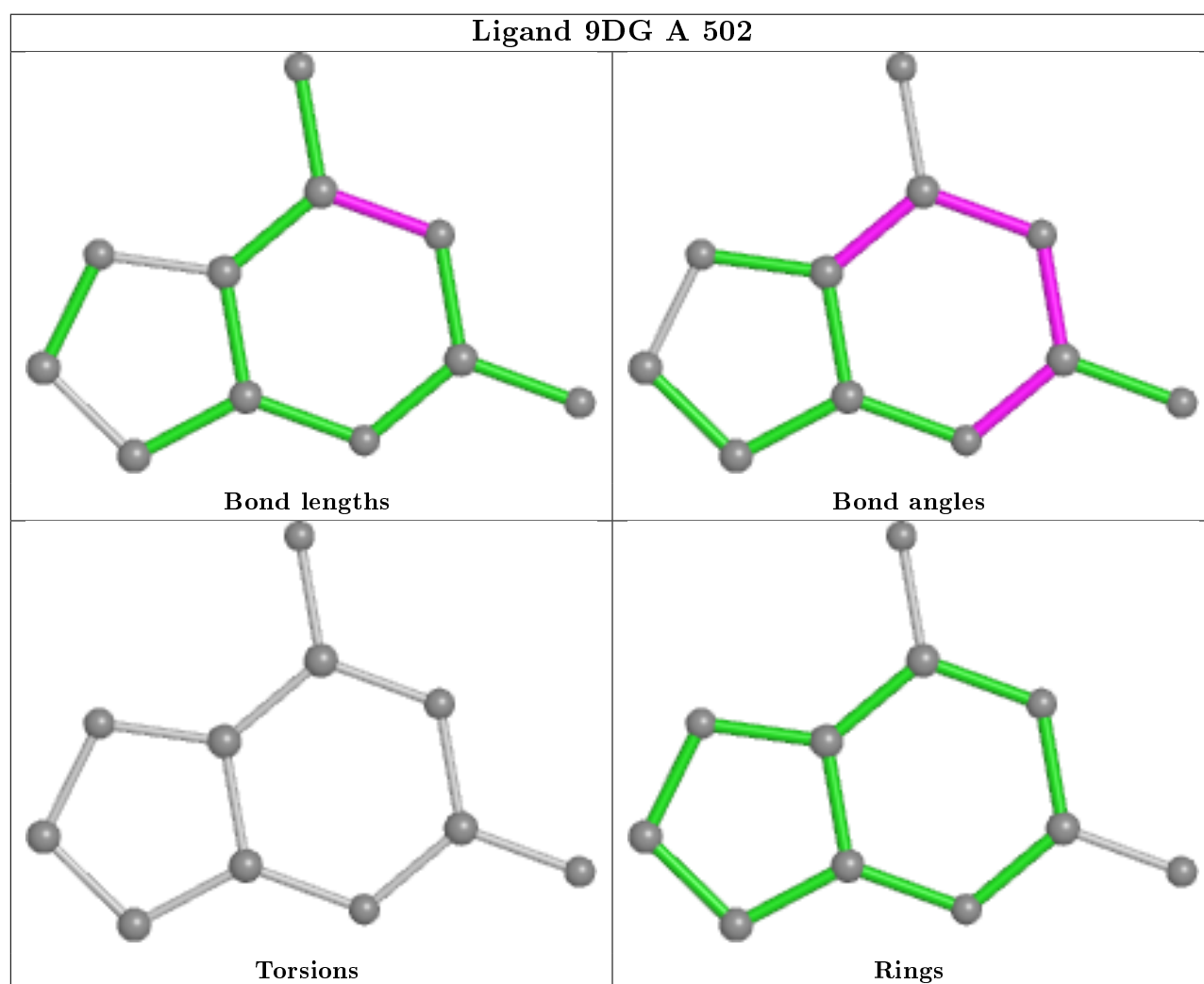
5.7 Other polymers (i)

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.