WORLDWIDE wwPDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 4, 2024 – 01:38 AM EST

PDB ID	:	1PUE
Title	:	PU.1 ETS DOMAIN-DNA COMPLEX
Authors	:	Kodandapani, R.; Pio, F.; Ni, C.Z.; Piccialli, G.; Klemsz, M.; McKercher, S.; Maki, R.A.; Ely, K.R.
Deposited on	:	1996-07-08
Resolution	:	2.10 Å(reported)

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

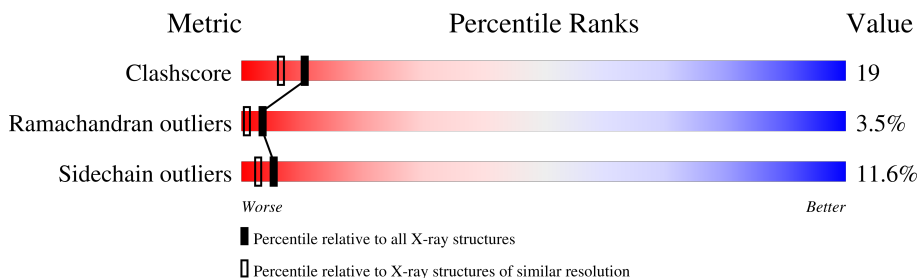|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.10 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		19
Ramachandran outliers		3.5%
Sidechain outliers		11.6%

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	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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	16	69% 31%
1	C	16	38% 62%
2	B	16	50% 44% 6%
2	D	16	56% 38% 6%
3	E	89	56% 35% 6% ··
3	F	89	54% 33% 12% ·

WORLDWIDE PDB PROTEIN DATA BANK

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2928 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 DNA chain called DNA (5'-D(*AP*AP*AP*AP*AP*GP*GP*GP*GP*AP* AP*GP*TP*GP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	16	Total 340	C 160	N 77	O 88	P 15	0	0	0
1	C	16	Total 340	C 160	N 77	O 88	P 15	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*TP*CP*CP*CP*AP*CP*TP*TP*CP*CP* CP*CP*TP*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	16	Total 310	C 152	N 43	O 100	P 15	0	0	0
2	D	16	Total 310	C 152	N 43	O 100	P 15	0	0	0

- Molecule 3 is a protein called PROTEIN (TRANSCRIPTION FACTOR PU.1 (TF PU.1)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	88	Total 738	C 475	N 135	O 125	S 3	0	0	0
3	F	89	Total 747	C 481	N 136	O 127	S 3	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	228	GLU	GLN	conflict	UNP P17433
F	228	GLU	GLN	conflict	UNP P17433

- Molecule 4 is water.

WORLDWIDE wwPDB PROTEIN DATA BANK

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	15	Total 15	O 15	0	0
4	B	21	Total 21	O 21	0	0
4	C	24	Total 24	O 24	0	0
4	D	14	Total 14	O 14	0	0
4	E	42	Total 42	O 42	0	0
4	F	27	Total 27	O 27	0	0

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: DNA (5'-D(*AP*AP*AP*AP*AP*GP*GP*GP*GP*AP*AP*GP*TP*GP*GP*G)-3')
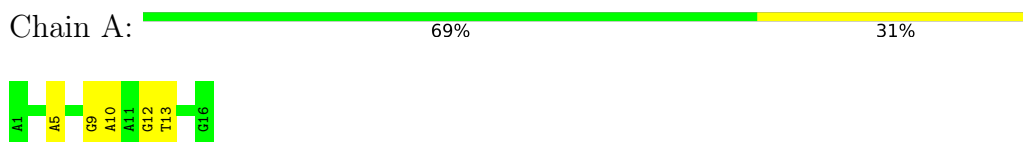
Chain A:

• Molecule 1: DNA (5'-D(*AP*AP*AP*AP*AP*GP*GP*GP*GP*AP*AP*GP*TP*GP*GP*G)-3')

Chain C:

• Molecule 2: DNA (5'-D(*TP*CP*CP*CP*AP*CP*TP*TP*CP*CP*CP*CP*TP*TP*TP*T)-3')

Chain B:

• Molecule 2: DNA (5'-D(*TP*CP*CP*CP*AP*CP*TP*TP*CP*CP*CP*CP*TP*TP*TP*T)-3')

Chain D:

• Molecule 3: PROTEIN (TRANSCRIPTION FACTOR PU.1 (TF PU.1))

Chain E:

K171 I172 R173 L174 Y175 Q176 F177 L178 L179 D180 S184 G185 D186 I191 W192 W193 V194 D195 T200 F201 Q202 L211 A212 H213 R214 W215 G216 I217 Q218 K219 Y227 E228 A231 R232 A233 L234 R235 N236 V246 K247 K248 K249 L250 T251 S255 G256 E257 V258 LEU

• Molecule 3: PROTEIN (TRANSCRIPTION FACTOR PU.1 (TF PU.1))

Chain F: 54%   33%   12%   •

K171 I172 R173 L174 Y175 Q176 F177 L181 L182 K188 D189 S190 I191 V194 D195 K196 D197 K198 G199 T200 F201 Q202 F203 K206 L211 A212 H213 R214 W215 G216 I217 Q218 K219 R222 T226 Y227 R232 A233 L234 R235 N236 Y237 K244 K247 K248 K249 L250 T251 Y252
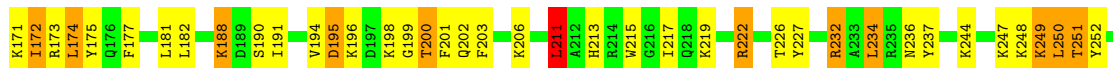
L259

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	89.10Å 101.90Å 55.60Å 90.00° 111.20° 90.00°	Depositor
Resolution (Å)	6.00 – 2.10	Depositor
% Data completeness (in resolution range)	(Not available) (6.00-2.10)	Depositor
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.225 , 0.287	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2928	wwPDB-VP
Average B, all atoms (Å²)	26.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.62	0/386	0.81	0/597
1	C	0.71	0/386	0.75	0/597
2	B	0.74	0/342	1.01	0/523
2	D	0.81	0/342	0.99	1/523 (0.2%)
3	E	0.77	1/753 (0.1%)	1.09	5/999 (0.5%)
3	F	0.61	0/762	0.86	1/1010 (0.1%)
All	All	0.71	1/2971 (0.0%)	0.93	7/4249 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	4
2	D	0	4
All	All	0	8

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	186	ASP	CB-CG	10.79	1.74	1.51

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	186	ASP	CB-CG-OD1	12.97	129.98	118.30
3	E	186	ASP	OD1-CG-OD2	-8.51	107.13	123.30
3	E	211	LEU	CA-CB-CG	7.83	133.31	115.30
3	E	250	LEU	CA-CB-CG	7.13	131.69	115.30
3	F	211	LEU	CA-CB-CG	6.26	129.71	115.30
3	E	234	LEU	CA-CB-CG	5.35	127.61	115.30
2	D	25	DC	OP2-P-O3'	5.02	116.25	105.20

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	18	DC	Sidechain
2	B	27	DC	Sidechain
2	B	28	DC	Sidechain
2	B	30	DT	Sidechain
2	D	17	DT	Sidechain
2	D	23	DT	Sidechain
2	D	25	DC	Sidechain
2	D	26	DC	Sidechain

5.2   Too-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 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	340	0	179	8	0
1	C	340	0	179	20	0
2	B	310	0	185	3	0
2	D	310	0	185	2	0
3	E	738	0	767	29	0
3	F	747	0	778	49	0
4	A	15	0	0	1	0
4	B	21	0	0	1	0
4	C	24	0	0	0	0
4	D	14	0	0	0	0
4	E	42	0	0	3	0
4	F	27	0	0	6	0
All	All	2928	0	2273	94	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 19.

All (94) 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:C:5:DA:H5"	3:F:249:LYS:NZ	1.82	0.93

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:248:LYS:HD2	3:E:251:THR:HG21	1.52	0.92
1:C:5:DA:H5"	3:F:249:LYS:HZ2	1.35	0.90
3:E:249:LYS:HG2	3:E:250:LEU:HD22	1.60	0.80
3:F:174:LEU:HD13	3:F:234:LEU:HD13	1.64	0.80
1:A:12:DG:H2"	1:A:13:DT:OP2	1.84	0.78
1:C:6:DG:H2"	1:C:7:DG:H5'	1.67	0.77
3:F:171:LYS:HG3	3:F:173:ARG:NH1	2.01	0.76
3:F:203:PHE:HB2	3:F:250:LEU:HB3	1.69	0.74
3:F:173:ARG:HD2	4:F:279:HOH:O	1.87	0.74
3:E:227:TYR:HB2	4:E:283:HOH:O	1.88	0.73
3:F:171:LYS:HZ3	3:F:173:ARG:HH12	1.36	0.73
3:F:199:GLY:HA3	3:F:259:LEU:HD11	1.71	0.73
3:F:236:ASN:ND2	4:F:266:HOH:O	2.22	0.72
1:C:5:DA:C5'	3:F:249:LYS:NZ	2.53	0.70
1:C:13:DT:H2"	1:C:14:DG:C8	2.26	0.70
3:F:171:LYS:NZ	3:F:173:ARG:HH12	1.92	0.66
3:E:192:TRP:HE1	3:E:202:GLN:HE21	1.43	0.65
3:F:213:HIS:HB2	4:F:267:HOH:O	1.97	0.64
3:E:246:VAL:HG22	3:E:251:THR:OG1	1.98	0.63
2:D:24:DT:H2"	2:D:25:DC:H5'	1.80	0.63
3:F:213:HIS:O	3:F:217:ILE:HG12	1.99	0.62
1:A:5:DA:H5"	3:E:249:LYS:HE3	1.82	0.61
3:E:255:SER:O	3:E:257:GLU:N	2.33	0.61
3:F:171:LYS:HG3	3:F:173:ARG:HH12	1.64	0.60
3:F:188:LYS:HE3	3:F:188:LYS:HA	1.82	0.60
3:E:195:ASP:HB3	3:E:200:THR:OG1	2.02	0.59
3:E:213:HIS:O	3:E:217:ILE:HG13	2.05	0.57
1:C:5:DA:H5"	3:F:249:LYS:HZ1	1.65	0.57
1:C:14:DG:C5'	3:F:222:ARG:HH12	2.18	0.56
3:E:257:GLU:HG2	3:E:257:GLU:O	2.06	0.56
1:C:9:DG:N7	3:F:232:ARG:NH1	2.52	0.56
2:B:20:DC:H2"	2:B:21:DA:C8	2.41	0.55
1:C:6:DG:OP1	3:F:249:LYS:O	2.24	0.55
1:C:15:DG:H2"	1:C:16:DG:H5"	1.89	0.54
3:F:248:LYS:O	3:F:251:THR:HG23	2.07	0.53
3:E:191:ILE:HD11	3:E:201:PHE:CD2	2.43	0.53
1:A:10:DA:N6	3:E:232:ARG:NH1	2.56	0.53
1:C:14:DG:H5'	3:F:222:ARG:NH1	2.23	0.53
3:F:196:LYS:HB2	4:F:270:HOH:O	2.09	0.52
3:E:179:LEU:HD13	3:E:258:VAL:HG21	1.90	0.52
3:E:249:LYS:HG2	3:E:250:LEU:CD2	2.34	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:174:LEU:CD1	3:F:234:LEU:HD13	2.38	0.52
1:C:15:DG:C2'	1:C:16:DG:H5"	2.40	0.52
3:E:191:ILE:HD11	3:E:201:PHE:HD2	1.75	0.52
3:E:212:ALA:HB2	4:E:283:HOH:O	2.09	0.52
3:F:171:LYS:HB3	4:F:263:HOH:O	2.09	0.52
3:F:248:LYS:HG2	3:F:251:THR:HG21	1.91	0.51
3:F:177:PHE:O	3:F:181:LEU:HD13	2.11	0.51
1:A:5:DA:C5'	3:E:249:LYS:HE3	2.40	0.51
3:F:201:PHE:CZ	3:F:252:TYR:HB2	2.45	0.51
3:F:195:ASP:HB3	3:F:200:THR:OG1	2.10	0.51
3:F:226:THR:HB	4:F:268:HOH:O	2.11	0.50
1:C:6:DG:H2"	1:C:7:DG:C5'	2.41	0.50
1:C:11:DA:H2"	1:C:12:DG:O5'	2.12	0.49
3:E:176:GLN:HG3	4:E:288:HOH:O	2.12	0.49
3:F:191:ILE:HG23	3:F:191:ILE:O	2.12	0.49
3:F:171:LYS:C	3:F:172:ILE:HG13	2.32	0.49
1:C:14:DG:H5"	3:F:222:ARG:HH12	1.77	0.48
3:F:232:ARG:HH21	3:F:236:ASN:HD21	1.61	0.48
1:A:10:DA:H62	3:E:232:ARG:NH1	2.12	0.48
3:F:215:TRP:CH2	3:F:219:LYS:HD3	2.49	0.48
2:D:27:DC:H2"	2:D:28:DC:OP2	2.13	0.48
3:E:177:PHE:HD2	3:E:178:LEU:HD13	1.79	0.47
3:E:193:TRP:HH2	3:E:258:VAL:HG11	1.78	0.47
3:E:247:LYS:H	3:E:247:LYS:CE	2.28	0.47
1:A:12:DG:C2'	1:A:13:DT:OP2	2.60	0.47
3:F:171:LYS:HZ2	3:F:173:ARG:HH22	1.63	0.46
1:C:5:DA:C5'	3:F:249:LYS:HZ2	2.14	0.46
3:F:190:SER:HB2	3:F:211:LEU:HD23	1.97	0.46
3:E:174:LEU:HD21	3:E:234:LEU:HD13	1.99	0.45
1:C:15:DG:H2"	1:C:16:DG:C5'	2.47	0.44
3:F:171:LYS:HB3	3:F:172:ILE:H	1.63	0.44
3:F:174:LEU:HD13	3:F:234:LEU:CD1	2.43	0.43
1:C:15:DG:H1'	1:C:16:DG:H5"	2.00	0.43
3:F:206:LYS:N	3:F:206:LYS:HD2	2.34	0.43
1:A:10:DA:N6	3:E:232:ARG:HH11	2.17	0.42
1:C:14:DG:C5'	3:F:222:ARG:NH1	2.82	0.42
1:A:9:DG:C6	1:A:10:DA:C6	3.07	0.42
3:F:194:VAL:HG22	3:F:200:THR:O	2.20	0.42
4:A:31:HOH:O	3:E:228:GLU:HG3	2.20	0.42
2:B:25:DC:H2'	4:B:52:HOH:O	2.18	0.42
3:E:215:TRP:CH2	3:E:219:LYS:HD3	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:202:GLN:HG3	3:F:251:THR:HG22	2.01	0.42
3:F:249:LYS:HB3	3:F:250:LEU:H	1.73	0.42
3:F:188:LYS:HE3	3:F:188:LYS:CA	2.50	0.41
1:C:5:DA:C5'	3:F:249:LYS:HZ1	2.25	0.41
3:E:236:ASN:HD22	3:E:236:ASN:HA	1.62	0.41
3:F:175:TYR:HD2	3:F:237:TYR:CZ	2.39	0.41
3:E:231:ALA:HA	3:E:234:LEU:HB2	2.01	0.41
2:B:28:DC:C2'	2:B:29:DT:H72	2.51	0.40
3:E:180:ASP:O	3:E:184:SER:OG	2.34	0.40
3:F:198:LYS:HE2	3:F:198:LYS:HB3	1.69	0.40
3:F:227:TYR:CZ	3:F:250:LEU:HD12	2.57	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	
3	E	86/89 (97%)	81 (94%)	3 (4%)	2 (2%)	6	2
3	F	87/89 (98%)	81 (93%)	2 (2%)	4 (5%)	2	0
All	All	173/178 (97%)	162 (94%)	5 (3%)	6 (4%)	3	1

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	195	ASP
3	F	249	LYS
3	E	256	GLY
3	E	257	GLU
3	F	247	LYS
3	F	250	LEU

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	
3	E	77/78 (99%)	70 (91%)	7 (9%)	9	6
3	F	78/78 (100%)	67 (86%)	11 (14%)	3	1
All	All	155/156 (99%)	137 (88%)	18 (12%)	5	3

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	173	ARG
3	E	178	LEU
3	E	211	LEU
3	E	234	LEU
3	E	247	LYS
3	E	251	THR
3	E	257	GLU
3	F	172	ILE
3	F	174	LEU
3	F	182	LEU
3	F	188	LYS
3	F	200	THR
3	F	211	LEU
3	F	222	ARG
3	F	232	ARG
3	F	234	LEU
3	F	244	LYS
3	F	251	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
3	E	202	GLN
3	E	213	HIS
3	E	221	ASN
3	E	236	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	F	236	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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 ⓘ

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.