



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2017 – 03:45 PM EDT

PDB ID : 5OSN
Title : Crystal Structure of Bovine Enterovirus 2 determined with Serial Femtosecond X-ray Crystallography
Authors : Roedig, P.; Ginn, H.M.; Pakendorf, T.; Sutton, G.; Harlos, K.; Walter, T.S.; Meyer, J.; Fischer, P.; Duman, R.; Vartiainen, I.; Reime, B.; Warmer, M.; Brewster, A.S.; Young, I.D.; Michels-Clark, T.; Sauter, N.K.; Kotecha, A.; Kelly, J.; Rowlands, D.J.; Sikorsky, M.; Nelson, S.; Damiani, D.S.; Alonso-Mori, R.; Ren, J.; Fry, E.E.; David, C.; Stuart, D.I.; Wagner, A.; Meents, A.
Deposited on : unknown
Resolution : 2.30 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	FAILED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029824

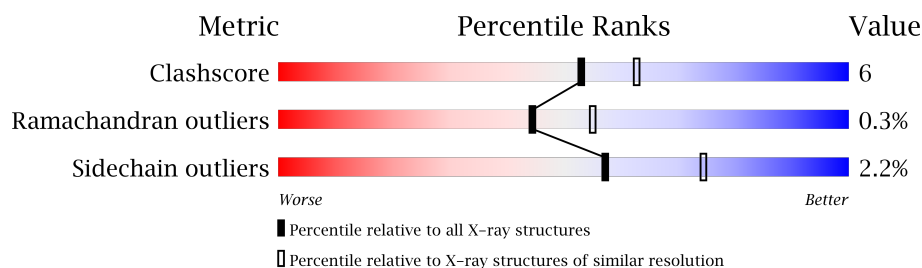
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	275	
2	B	244	
3	C	243	
4	D	71	

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 6590 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 Capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	270	Total	C	N	O	S	0	0	0
			2096	1317	360	409	10			

- Molecule 2 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	237	Total	C	N	O	S	0	0	0
			1845	1180	316	345	4			

- Molecule 3 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	243	Total	C	N	O	S	0	0	0
			1911	1232	310	357	12			

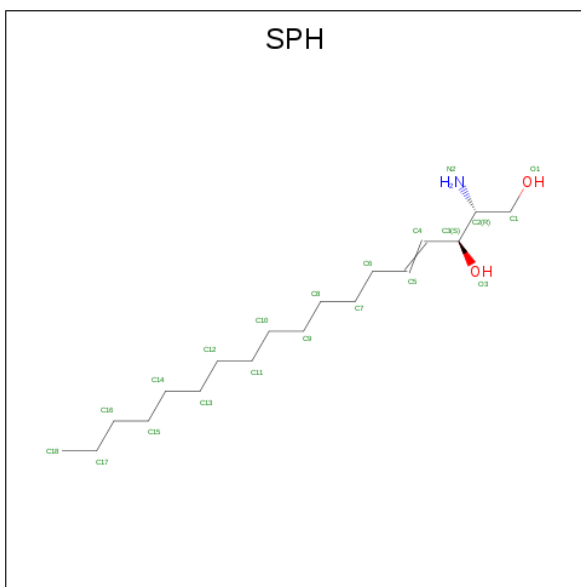
There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	102	PHE	LEU	conflict	UNP Q65480
C	103	THR	HIS	conflict	UNP Q65480
C	143	ASN	ALA	conflict	UNP Q65480
C	192	ALA	ARG	conflict	UNP Q65480
C	213	ALA	SER	conflict	UNP Q65480

- Molecule 4 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	43	Total	C	N	O	S	0	0	0
			348	220	57	70	1			

- Molecule 5 is SPHINGOSINE (three-letter code: SPH) (formula: C₁₈H₃₇NO₂).

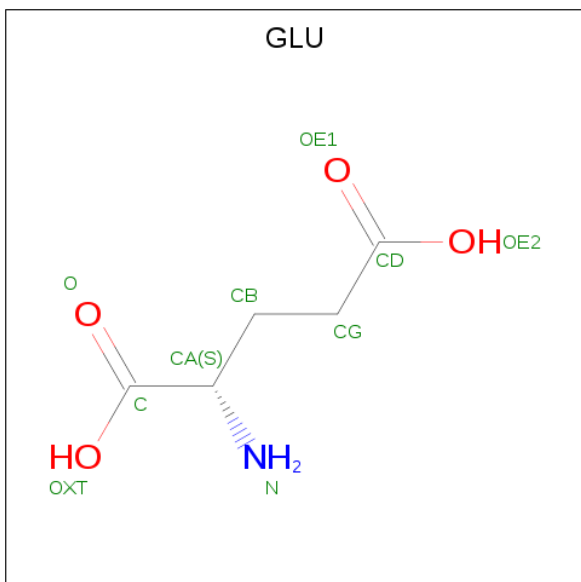


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			21	18	1	2		

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

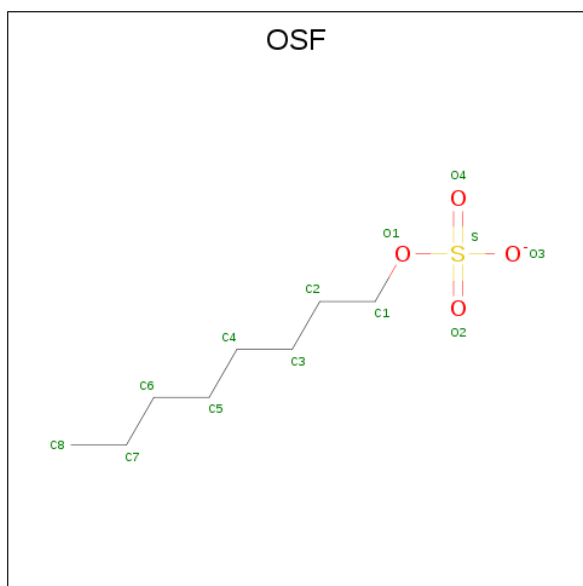
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	K	0	0
			2	2		
6	C	1	Total	K	0	0
			1	1		

- Molecule 7 is GLUTAMIC ACID (three-letter code: GLU) (formula: C₅H₉NO₄).



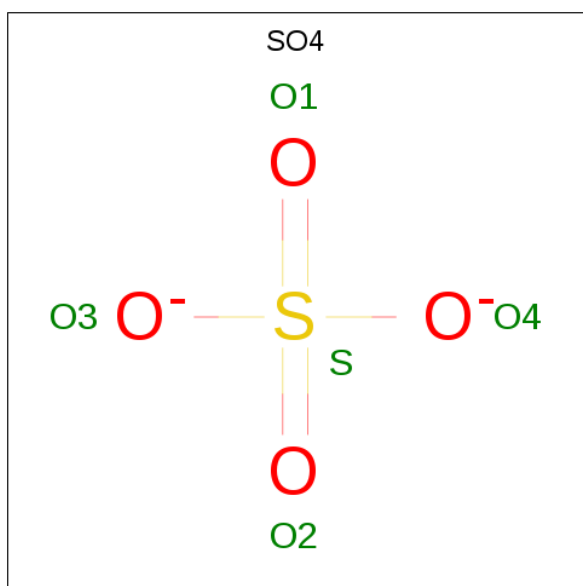
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	0	0
			10	5	1	4		

- Molecule 8 is octyl sulfate (three-letter code: OSF) (formula: $C_8H_{17}O_4S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	C	1	Total	C	O	S	0	0
			13	8	4	1		

- Molecule 9 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	1	Total	O	S	0	0
			5	4	1		

- Molecule 10 is water.

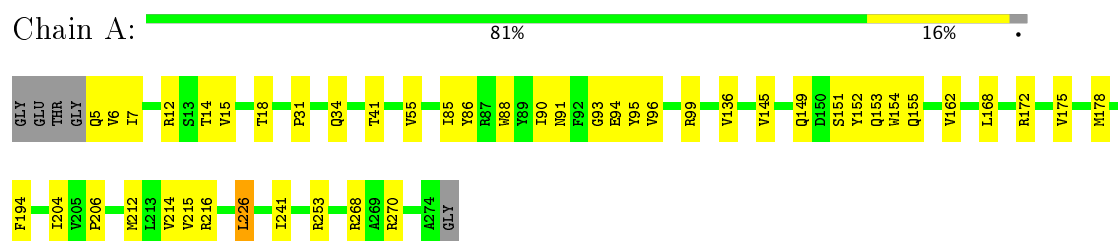
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	112	Total	O		0	0
			112	112			
10	B	100	Total	O		0	0
			100	100			
10	C	105	Total	O		0	0
			105	105			
10	D	21	Total	O		0	0
			21	21			

3 Residue-property plots [i](#)

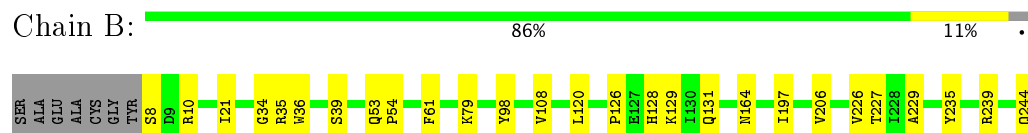
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

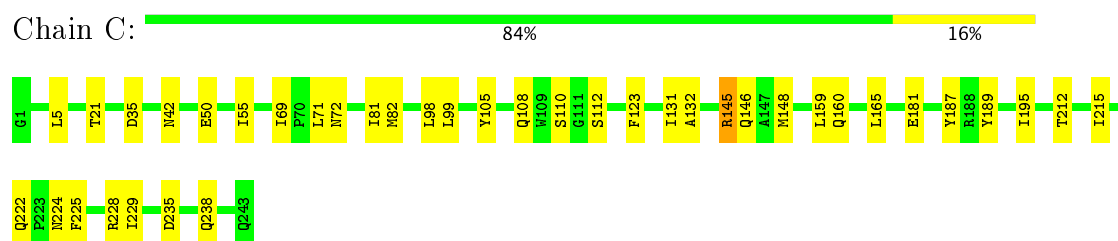
- Molecule 1: Capsid protein



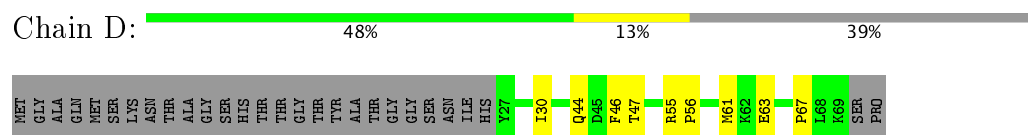
- Molecule 2: Capsid protein



- Molecule 3: Capsid protein



- Molecule 4: Capsid protein



4 Data and refinement statistics

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

Property	Value	Source
Space group	F 2 3	Depositor
Cell constants a, b, c, α , β , γ	436.60 Å 436.60 Å 436.60 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.71 – 2.30	Depositor
% Data completeness (in resolution range)	78.0 (29.71-2.30)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.98 (at 1.91 Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.233 , 0.257	Depositor
Wilson B-factor (Å ²)	21.4	Xtriage
Anisotropy	0.000	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.028 for k,h,-l	Xtriage
Total number of atoms	6590	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.38% 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: SPH, K, SO4, OSF

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	0.50	0/2144	0.73	1/2920 (0.0%)
2	B	0.51	0/1899	0.75	0/2606
3	C	0.51	0/1964	0.76	0/2681
4	D	0.53	0/355	0.70	0/482
All	All	0.50	0/6362	0.74	1/8689 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	178	MET	N-CA-C	5.15	124.91	111.00

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	2096	0	2030	40	0
2	B	1845	0	1789	24	0
3	C	1911	0	1876	27	0
4	D	348	0	335	11	0
5	A	21	0	37	2	0
6	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	1	0	0	0	0
7	A	10	0	5	0	0
8	C	13	0	17	0	0
9	C	5	0	0	0	0
10	A	112	0	0	0	0
10	B	100	0	0	1	0
10	C	105	0	0	1	0
10	D	21	0	0	0	0
All	All	6590	0	6089	77	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.

The worst 5 of 77 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:5:GLN:HB2	2:B:35:ARG:HD2	1.46	0.94
2:B:10:ARG:HD2	10:B:349:HOH:O	1.70	0.91
3:C:145:ARG:HB3	3:C:145:ARG:NH2	1.85	0.91
1:A:5:GLN:HB3	2:B:34:GLY:O	1.77	0.84
2:B:8:SER:HB3	4:D:67:PRO:O	1.78	0.84

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	268/275 (98%)	259 (97%)	8 (3%)	1 (0%)	38	47
2	B	235/244 (96%)	220 (94%)	15 (6%)	0	100	100
3	C	241/243 (99%)	233 (97%)	7 (3%)	1 (0%)	38	47

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	41/71 (58%)	38 (93%)	3 (7%)	0	100	100
All	All	785/833 (94%)	750 (96%)	33 (4%)	2 (0%)	44	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	149	GLN
3	C	229	ILE

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	229/231 (99%)	226 (99%)	3 (1%)	73	86
2	B	196/200 (98%)	194 (99%)	2 (1%)	80	90
3	C	211/211 (100%)	203 (96%)	8 (4%)	38	52
4	D	39/59 (66%)	37 (95%)	2 (5%)	28	37
All	All	675/701 (96%)	660 (98%)	15 (2%)	57	74

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	99	LEU
3	C	105	TYR
3	C	235	ASP
3	C	5	LEU
3	C	181	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	151	GLN
2	B	244	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	146	GLN
2	B	111	ASN
3	C	12	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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	SPH	A	301	-	19,20,20	0.64	0	18,21,21	0.87	2 (11%)
7	GLU	A	304	-	1,9,9	0.11	0	1,11,11	0.33	0
8	OSF	C	302	-	12,12,12	0.58	0	12,14,14	0.59	0
9	SO4	C	303	-	4,4,4	1.15	0	6,6,6	0.41	0

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	SPH	A	301	-	-	0/21/21/21	0/0/0/0
7	GLU	A	304	-	-	0/3/9/9	0/0/0/0
8	OSF	C	302	-	-	0/10/10/10	0/0/0/0
9	SO4	C	303	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	301	SPH	C3-C4-C5	2.08	128.01	125.22
5	A	301	SPH	O3-C3-C2	2.28	111.10	107.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	301	SPH	2	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 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.