



## wwPDB EM Validation Summary Report ⓘ

Oct 26, 2025 – 02:23 PM JST

PDB ID : 8XYP / pdb\_00008xyp  
EMDB ID : EMD-38780  
Title : Cryo-EM structure of SAH-bound Tetrahymena DNA methyltransferase complex MTA1c  
Authors : Xu, Q.; Shi, Z.B.  
Deposited on : 2024-01-20  
Resolution : 2.54 Å(reported)  
Based on initial model : .

This is a wwPDB EM Validation Summary 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/EMValidationReportHelp>  
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:

EMDB validation analysis : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
MolProbity : 4-5-2 with Phenix2.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
EM percentile statistics : **NOT EXECUTED**  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.46

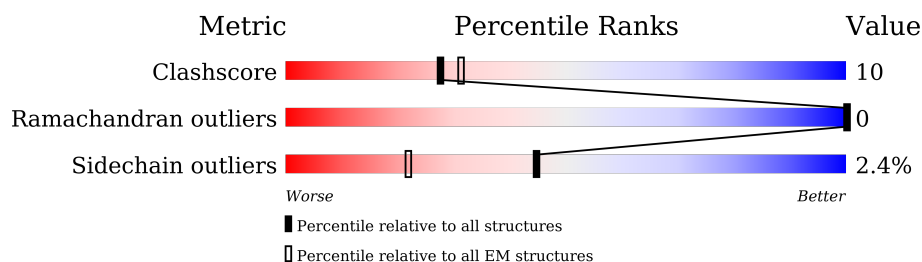
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.54 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	378	
2	B	357	
3	C	364	
4	D	146	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5317 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 MT-a70 family protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	225	Total	C	N	O	S	0	0
			1838	1173	321	334	10		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP Q22GC0
A	-4	PRO	-	expression tag	UNP Q22GC0
A	-3	GLU	-	expression tag	UNP Q22GC0
A	-2	PHE	-	expression tag	UNP Q22GC0
A	-1	LYS	-	expression tag	UNP Q22GC0
A	0	LEU	-	expression tag	UNP Q22GC0

- Molecule 2 is a protein called Methyltransferase MT, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	259	Total	C	N	O	S	0	0
			2142	1373	352	408	9		

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	GLY	-	expression tag	UNP Q22XT1
B	-3	PRO	-	expression tag	UNP Q22XT1
B	-2	GLY	-	expression tag	UNP Q22XT1
B	-1	ARG	-	expression tag	UNP Q22XT1
B	0	PRO	-	expression tag	UNP Q22XT1
B	325	TRP	-	expression tag	UNP Q22XT1
B	326	SER	-	expression tag	UNP Q22XT1
B	327	HIS	-	expression tag	UNP Q22XT1
B	328	PRO	-	expression tag	UNP Q22XT1
B	329	GLN	-	expression tag	UNP Q22XT1
B	330	PHE	-	expression tag	UNP Q22XT1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	331	GLU	-	expression tag	UNP Q22XT1
B	332	LYS	-	expression tag	UNP Q22XT1
B	333	GLY	-	expression tag	UNP Q22XT1
B	334	GLY	-	expression tag	UNP Q22XT1
B	335	GLY	-	expression tag	UNP Q22XT1
B	336	SER	-	expression tag	UNP Q22XT1
B	337	GLY	-	expression tag	UNP Q22XT1
B	338	GLY	-	expression tag	UNP Q22XT1
B	339	GLY	-	expression tag	UNP Q22XT1
B	340	SER	-	expression tag	UNP Q22XT1
B	341	GLY	-	expression tag	UNP Q22XT1
B	342	GLY	-	expression tag	UNP Q22XT1
B	343	GLY	-	expression tag	UNP Q22XT1
B	344	SER	-	expression tag	UNP Q22XT1
B	345	TRP	-	expression tag	UNP Q22XT1
B	346	SER	-	expression tag	UNP Q22XT1
B	347	HIS	-	expression tag	UNP Q22XT1
B	348	PRO	-	expression tag	UNP Q22XT1
B	349	GLN	-	expression tag	UNP Q22XT1
B	350	PHE	-	expression tag	UNP Q22XT1
B	351	GLU	-	expression tag	UNP Q22XT1
B	352	LYS	-	expression tag	UNP Q22XT1

- Molecule 3 is a protein called Myb-like DNA-binding domain protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	32	Total	C	N	O	0	0
			269	171	45	53		

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP Q22VV9
C	-2	PRO	-	expression tag	UNP Q22VV9
C	-1	GLY	-	expression tag	UNP Q22VV9
C	0	ARG	-	expression tag	UNP Q22VV9
C	1	PRO	-	expression tag	UNP Q22VV9

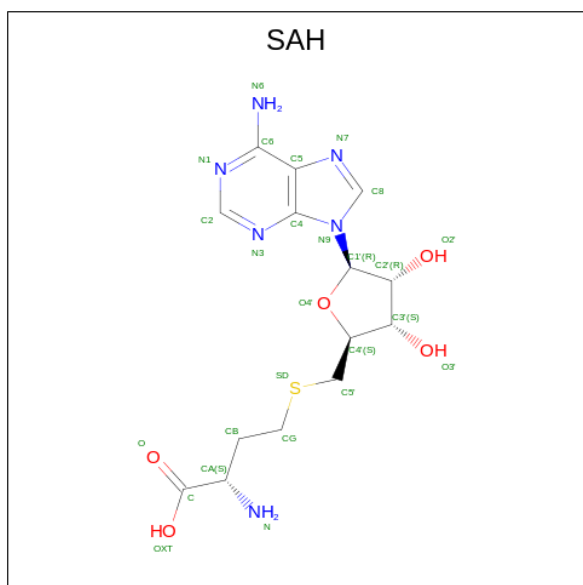
- Molecule 4 is a protein called Transmembrane protein, putative.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	127	Total	C	N	O	S	0	0
			1042	661	175	202	4		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	GLY	-	expression tag	UNP I7M8B9
D	-2	PRO	-	expression tag	UNP I7M8B9
D	-1	GLU	-	expression tag	UNP I7M8B9
D	0	PHE	-	expression tag	UNP I7M8B9

- Molecule 5 is S-ADENOSYL-L-HOMOCYSTEINE (CCD ID: SAH) (formula:  $C_{14}H_{20}N_6O_5S$ ) (labeled as "Ligand of Interest" by depositor).

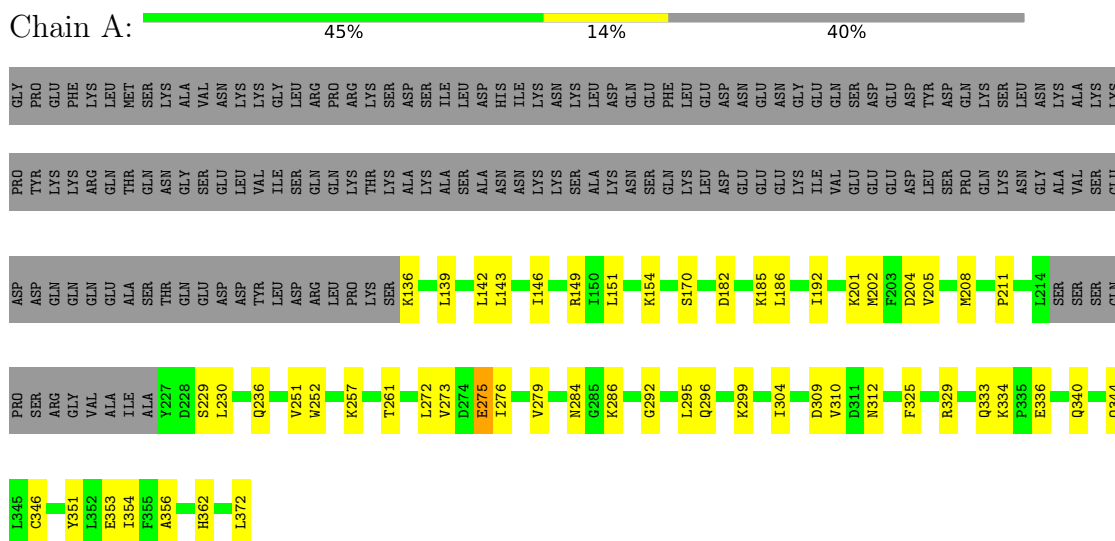


Mol	Chain	Residues	Atoms					AltConf
5	A	1	Total	C	N	O	S	0
			26	14	6	5	1	

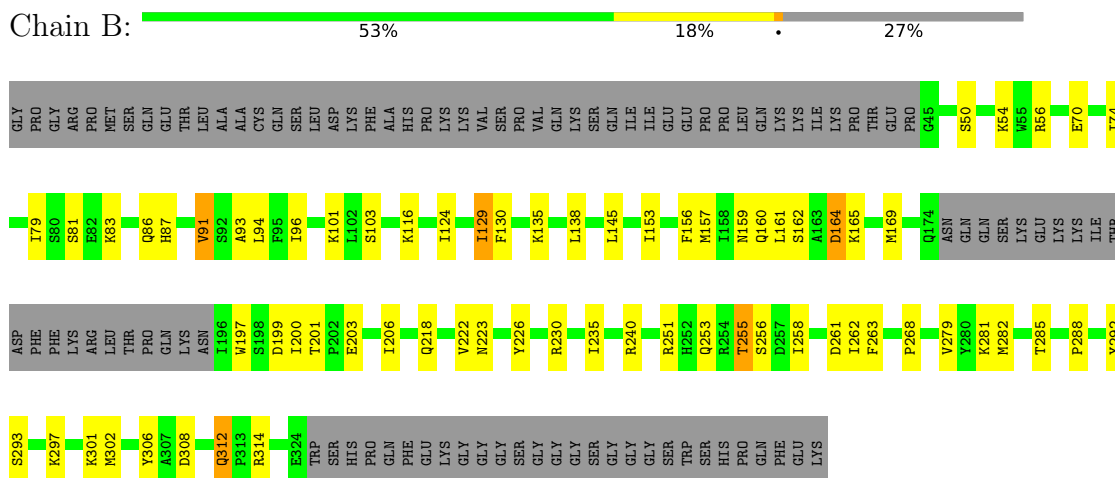
### 3 Residue-property plots

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.

- Molecule 1: MT-a70 family protein



- Molecule 2: Methyltransferase MT, putative



- Molecule 3: Myb-like DNA-binding domain protein





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	372998	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	48	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	130000	Depositor
Image detector	TFS FALCON 4i (4k x 4k)	Depositor



## 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: SAH

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	0.17	0/1876	0.33	0/2522
2	B	0.16	0/2184	0.38	0/2930
3	C	0.13	0/271	0.36	0/365
4	D	0.14	0/1059	0.35	0/1417
All	All	0.16	0/5390	0.36	0/7234

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	1838	0	1835	40	0
2	B	2142	0	2141	48	0
3	C	269	0	273	4	0
4	D	1042	0	1039	26	0
5	A	26	0	19	0	0
All	All	5317	0	5307	101	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 10.

The worst 5 of 101 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:353:GLU:HG2	1:A:356:ALA:HB2	1.68	0.75
2:B:91:VAL:HG22	2:B:124:ILE:HB	1.74	0.69
2:B:162:SER:HB3	2:B:165:LYS:HB2	1.78	0.65
4:D:56:LYS:HG3	4:D:60:ARG:HE	1.63	0.64
4:D:97:ILE:HG22	4:D:99:ASP:H	1.64	0.63

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 PDB entries followed by that with respect to all EM entries.

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	221/378 (58%)	215 (97%)	6 (3%)	0	100	100
2	B	255/357 (71%)	241 (94%)	14 (6%)	0	100	100
3	C	30/364 (8%)	30 (100%)	0	0	100	100
4	D	125/146 (86%)	120 (96%)	5 (4%)	0	100	100
All	All	631/1245 (51%)	606 (96%)	25 (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 PDB entries followed by that with respect to all EM entries.

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	200/337 (59%)	196 (98%)	4 (2%)	50	69
2	B	240/324 (74%)	233 (97%)	7 (3%)	37	54
3	C	30/335 (9%)	30 (100%)	0	100	100
4	D	113/130 (87%)	110 (97%)	3 (3%)	40	57
All	All	583/1126 (52%)	569 (98%)	14 (2%)	45	62

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	164	ASP
2	B	255	THR
4	D	119	SER
4	D	27	ASN
4	D	104	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	136	GLN
4	D	120	ASN
2	B	173	ASN
2	B	125	ASN
2	B	253	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

1 ligand is 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$
5	SAH	A	401	-	24,28,28	0.68	0	25,40,40	0.94	2 (8%)

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	SAH	A	401	-	-	0/11/31/31	0/3/3/3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	401	SAH	O4'-C1'-C2'	-2.48	103.30	106.93
5	A	401	SAH	C5-C6-N6	2.28	123.82	120.35

There are no chirality outliers.

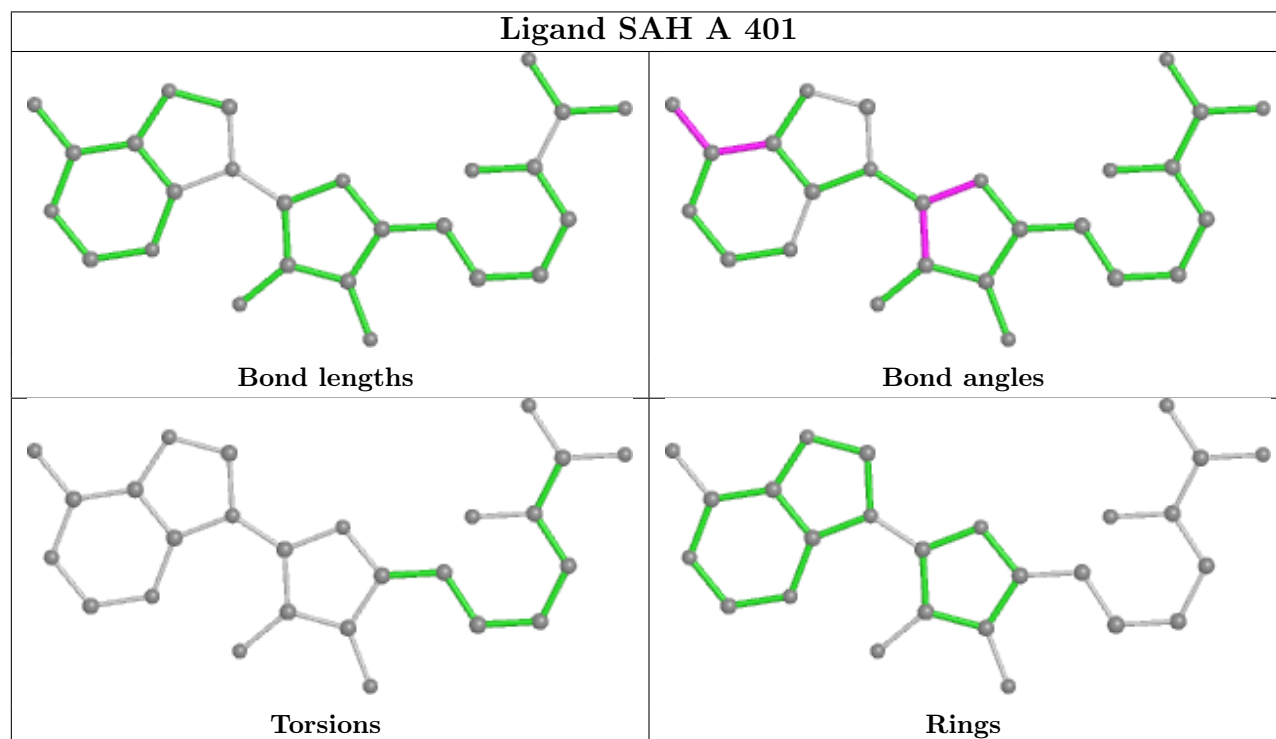
There are no torsion outliers.

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 than 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.



## 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.