



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:03 PM BST

PDB ID : 4BZI
EMDB ID: : EMD-2428
Title : The structure of the COPII coat assembled on membranes
Authors : Zanetti, G.; Prinz, S.; Daum, S.; Meister, A.; Schekman, R.; Bacia, K.; Briggs, J.A.G.
Deposited on : 2013-07-26
Resolution : 23.00 Å(reported)
Based on PDB ID : 1M2O

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk27241

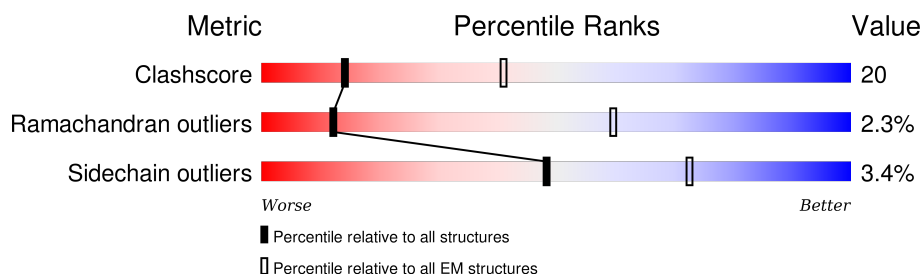
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 23.00 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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

Mol	Chain	Length	Quality of chain
1	A	768	68% 25% • 5%
1	D	768	68% 25% • 5%
1	G	768	69% 25% • 5%
2	B	190	52% 31% • 14%
2	J	190	53% 31% • 14%
2	K	190	52% 32% • 14%
3	E	926	45% 31% • 21%
3	L	926	45% 31% • 21%
3	M	926	45% 32% • 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	926	99%
4	N	926	98%
4	O	926	98%

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 39154 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 SEC23P.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	733	Total	C	N	O	S	0	0
			5780	3679	969	1109	23		
1	D	733	Total	C	N	O	S	0	0
			5780	3679	969	1109	23		
1	G	733	Total	C	N	O	S	0	0
			5780	3679	969	1109	23		

- Molecule 2 is a protein called SAR1P.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	164	Total	C	N	O	S	0	0
			1299	836	220	239	4		
2	J	164	Total	C	N	O	S	0	0
			1299	836	220	239	4		
2	K	164	Total	C	N	O	S	0	0
			1299	836	220	239	4		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	139	ALA	GLY	CONFLICT	UNP C8ZIG2
J	139	ALA	GLY	CONFLICT	UNP C8ZIG2
K	139	ALA	GLY	CONFLICT	UNP C8ZIG2

- Molecule 3 is a protein called SEC24P.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	E	735	Total	C	N	O	S	0	0
			5823	3702	999	1084	38		
3	L	735	Total	C	N	O	S	0	0
			5823	3702	999	1084	38		
3	M	735	Total	C	N	O	S	0	0
			5823	3702	999	1084	38		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	408	THR	ALA	CONFLICT	UNP C8ZAD6
E	865	VAL	ALA	CONFLICT	UNP C8ZAD6
L	408	THR	ALA	CONFLICT	UNP C8ZAD6
L	865	VAL	ALA	CONFLICT	UNP C8ZAD6
M	408	THR	ALA	CONFLICT	UNP C8ZAD6
M	865	VAL	ALA	CONFLICT	UNP C8ZAD6

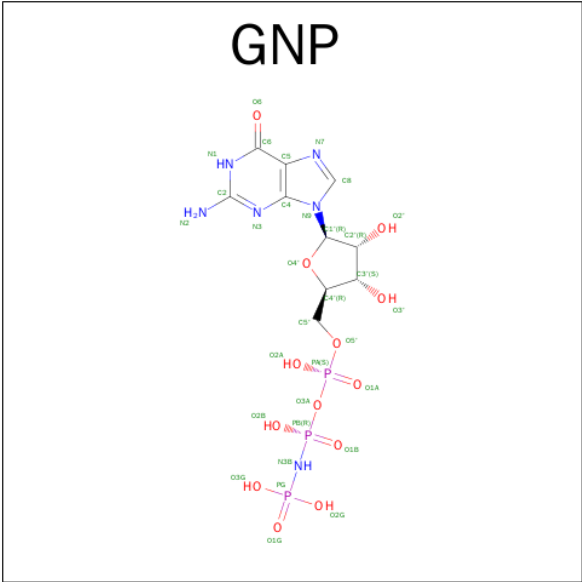
- Molecule 4 is a protein called SEC24P.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	F	13	Total	C	N	O	0	0
			109	70	19	20		
4	N	14	Total	C	N	O	0	0
			117	74	21	22		
4	O	14	Total	C	N	O	0	0
			117	74	21	22		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
5	G	1	Total	Zn	0
			1	1	
5	D	1	Total	Zn	0
			1	1	
5	E	1	Total	Zn	0
			1	1	
5	A	1	Total	Zn	0
			1	1	
5	L	1	Total	Zn	0
			1	1	
5	M	1	Total	Zn	0
			1	1	

- Molecule 6 is PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER (three-letter code: GNP) (formula: C₁₀H₁₇N₆O₁₃P₃).



Mol	Chain	Residues	Atoms					AltConf
6	B	1	Total	C	N	O	P	0
			32	10	6	13	3	
6	J	1	Total	C	N	O	P	0
			32	10	6	13	3	
6	K	1	Total	C	N	O	P	0
			32	10	6	13	3	

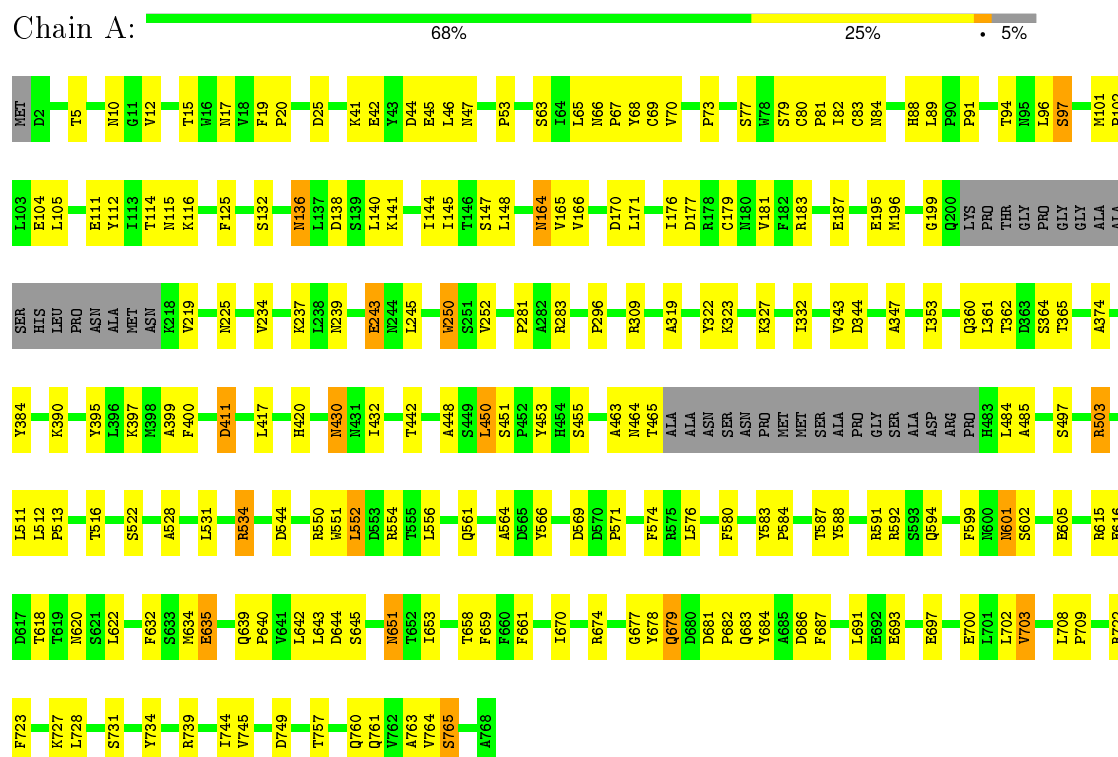
- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
7	B	1	Total	Mg	0
			1	1	
7	J	1	Total	Mg	0
			1	1	
7	K	1	Total	Mg	0
			1	1	

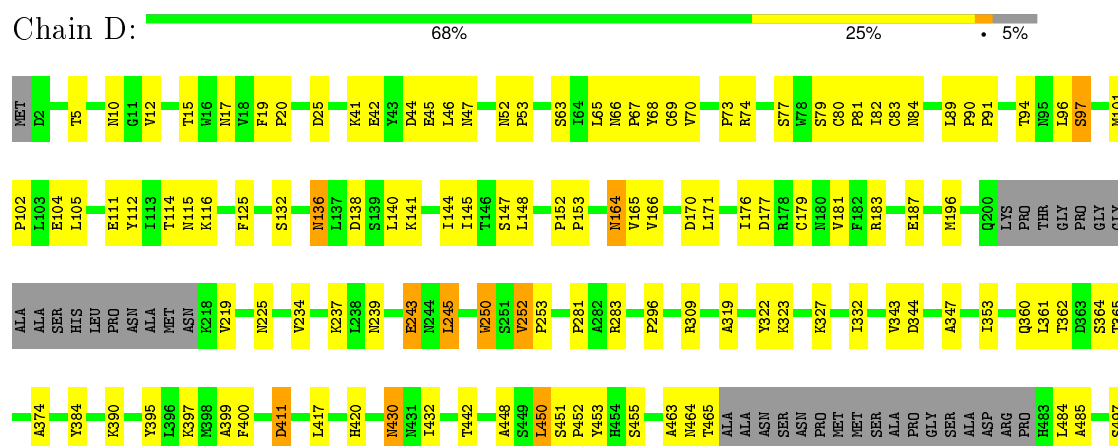
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: SEC23P



• Molecule 1: SEC23P







[illegible]

4 Experimental information

Property	Value	Source
Reconstruction method	TOMOGRAPHY	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH TILTED IMAGE WITHIN TOMOGRAM, Not provided	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	80	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3200	Depositor
Magnification	19500	Depositor
Image detector	2K X 2K MULTISCAN CHARGE-COUPLED DEVICE CAMERA	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: ZN, MG, GNP

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.38	0/5915	0.62	0/8052
1	D	0.38	0/5915	0.62	0/8052
1	G	0.38	0/5915	0.62	0/8052
2	B	0.34	0/1327	0.61	0/1800
2	J	0.34	0/1327	0.61	0/1800
2	K	0.34	0/1327	0.61	0/1800
3	E	0.39	0/5943	0.68	1/8064 (0.0%)
3	L	0.39	0/5943	0.68	1/8064 (0.0%)
3	M	0.39	0/5943	0.68	1/8064 (0.0%)
4	F	0.35	0/111	0.47	0/150
4	N	0.35	0/118	0.47	0/158
4	O	0.35	0/118	0.48	0/158
All	All	0.38	0/39902	0.64	3/54214 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	768	LEU	CA-CB-CG	5.33	127.56	115.30
3	M	768	LEU	CA-CB-CG	5.33	127.56	115.30
3	L	768	LEU	CA-CB-CG	5.32	127.55	115.30

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	5780	0	5663	201	0
1	D	5780	0	5663	205	0
1	G	5780	0	5663	193	0
2	B	1299	0	1288	65	0
2	J	1299	0	1288	59	0
2	K	1299	0	1288	61	0
3	E	5823	0	5853	291	0
3	L	5823	0	5853	293	0
3	M	5823	0	5853	301	0
4	F	109	0	105	14	0
4	N	117	0	110	14	0
4	O	117	0	110	12	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
5	L	1	0	0	0	0
5	M	1	0	0	0	0
6	B	32	0	13	5	0
6	J	32	0	13	4	0
6	K	32	0	13	5	0
7	B	1	0	0	0	0
7	J	1	0	0	0	0
7	K	1	0	0	0	0
All	All	39154	0	38776	1589	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 20.

The worst 5 of 1589 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:91:PRO:HB3	1:D:669:GLN:NE2	1.18	1.43
1:A:91:PRO:CB	1:D:669:GLN:HE21	1.35	1.38
1:G:183:ARG:NH1	3:M:383:GLU:HB2	1.53	1.22
1:D:183:ARG:NH1	3:L:383:GLU:HB2	1.53	1.22
1:A:183:ARG:NH1	3:E:383:GLU:HB2	1.53	1.21

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	727/768 (95%)	687 (94%)	35 (5%)	5 (1%)	26	71
1	D	727/768 (95%)	688 (95%)	34 (5%)	5 (1%)	26	71
1	G	727/768 (95%)	688 (95%)	34 (5%)	5 (1%)	26	71
2	B	160/190 (84%)	142 (89%)	13 (8%)	5 (3%)	5	42
2	J	160/190 (84%)	142 (89%)	13 (8%)	5 (3%)	5	42
2	K	160/190 (84%)	142 (89%)	13 (8%)	5 (3%)	5	42
3	E	723/926 (78%)	640 (88%)	55 (8%)	28 (4%)	4	36
3	L	723/926 (78%)	641 (89%)	54 (8%)	28 (4%)	4	36
3	M	723/926 (78%)	641 (89%)	54 (8%)	28 (4%)	4	36
4	F	11/926 (1%)	10 (91%)	1 (9%)	0	100	100
4	N	11/926 (1%)	10 (91%)	1 (9%)	0	100	100
4	O	11/926 (1%)	10 (91%)	1 (9%)	0	100	100
All	All	4863/8430 (58%)	4441 (91%)	308 (6%)	114 (2%)	12	48

5 of 114 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	703	VAL
2	B	139	ALA
2	B	155	THR
1	D	703	VAL
3	E	147	LEU

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	644/668 (96%)	616 (96%)	28 (4%)	35	70
1	D	644/668 (96%)	616 (96%)	28 (4%)	35	70
1	G	644/668 (96%)	616 (96%)	28 (4%)	35	70
2	B	138/159 (87%)	136 (99%)	2 (1%)	74	89
2	J	138/159 (87%)	136 (99%)	2 (1%)	74	89
2	K	138/159 (87%)	136 (99%)	2 (1%)	74	89
3	E	660/819 (81%)	640 (97%)	20 (3%)	48	77
3	L	660/819 (81%)	640 (97%)	20 (3%)	48	77
3	M	660/819 (81%)	640 (97%)	20 (3%)	48	77
4	F	12/817 (2%)	12 (100%)	0	100	100
4	N	13/817 (2%)	13 (100%)	0	100	100
4	O	13/817 (2%)	13 (100%)	0	100	100
All	All	4364/7389 (59%)	4214 (97%)	150 (3%)	48	75

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	428	LEU
1	G	166	VAL
3	M	377	ASP
3	E	557	VAL
3	E	822	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 97 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	535	ASN
1	G	233	GLN
3	M	533	ASN
3	E	732	HIS
1	G	47	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 9 are monoatomic - leaving 3 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$
6	GNP	B	1190	7	29,34,34	1.80	7 (24%)	29,54,54	2.69	4 (13%)
6	GNP	J	1190	7	29,34,34	1.80	7 (24%)	29,54,54	2.68	4 (13%)
6	GNP	K	1190	7	29,34,34	1.79	7 (24%)	29,54,54	2.68	4 (13%)

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
6	GNP	B	1190	7	-	0/13/38/38	0/3/3/3
6	GNP	J	1190	7	-	0/13/38/38	0/3/3/3
6	GNP	K	1190	7	-	0/13/38/38	0/3/3/3

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	1190	GNP	PB-O2B	-4.17	1.45	1.56
6	J	1190	GNP	PB-O2B	-4.15	1.45	1.56
6	K	1190	GNP	PB-O2B	-4.14	1.45	1.56
6	K	1190	GNP	PG-O2G	-3.28	1.48	1.56

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	1190	GNP	PG-O2G	-3.27	1.48	1.56

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	1190	GNP	C5-C6-N1	-8.85	111.95	123.52
6	K	1190	GNP	C5-C6-N1	-8.83	111.98	123.52
6	J	1190	GNP	C5-C6-N1	-8.82	111.99	123.52
6	K	1190	GNP	O3G-PG-O1G	-4.16	102.63	113.58
6	B	1190	GNP	O3G-PG-O1G	-4.15	102.65	113.58

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	1190	GNP	5	0
6	J	1190	GNP	4	0
6	K	1190	GNP	5	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.