



## wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

May 15, 2017 – 12:04 PM EDT

PDB ID : 5X5F  
EMDB ID: : EMD-6707  
Title : Prefusion structure of MERS-CoV spike glycoprotein, conformation 2  
Authors : Yuan, Y.; Cao, D.; Zhang, Y.; Ma, J.; Qi, J.; Wang, Q.; Lu, G.; Wu, Y.; Yan, J.; Shi, Y.; Zhang, X.; Gao, G.F.  
Deposited on : 2017-02-15  
Resolution : 4.20 Å(reported)

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report  
for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
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-20029077

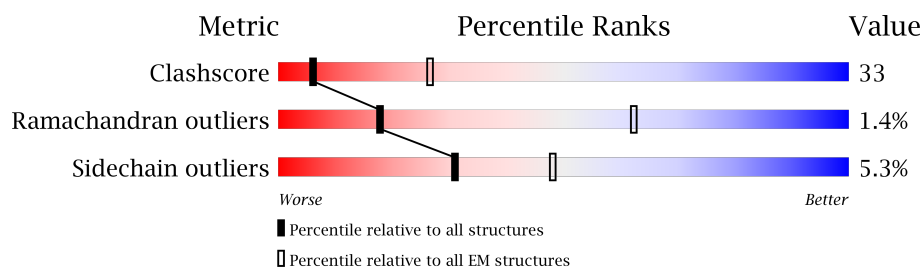
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.20 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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  $\geq 3$ , 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	1323	
1	B	1323	
1	C	1323	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 26422 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 S protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1141	Total	C	N	O	S	1	0
			8806	5599	1457	1699	51		
1	B	1141	Total	C	N	O	S	1	0
			8806	5599	1457	1699	51		
1	C	1141	Total	C	N	O	S	1	0
			8810	5601	1458	1700	51		

There are 144 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	751	SER	ARG	engineered mutation	UNP W6A028
A	1020	GLN	ARG	engineered mutation	UNP W6A028
A	1295	GLU	-	expression tag	UNP W6A028
A	1296	PHE	-	expression tag	UNP W6A028
A	1297	ARG	-	expression tag	UNP W6A028
A	1298	LEU	-	expression tag	UNP W6A028
A	1299	VAL	-	expression tag	UNP W6A028
A	1300	PRO	-	expression tag	UNP W6A028
A	1301	ARG	-	expression tag	UNP W6A028
A	1302	GLY	-	expression tag	UNP W6A028
A	1303	SER	-	expression tag	UNP W6A028
A	1304	PRO	-	expression tag	UNP W6A028
A	1305	GLY	-	expression tag	UNP W6A028
A	1306	SER	-	expression tag	UNP W6A028
A	1307	GLY	-	expression tag	UNP W6A028
A	1308	TYR	-	expression tag	UNP W6A028
A	1309	ILE	-	expression tag	UNP W6A028
A	1310	PRO	-	expression tag	UNP W6A028
A	1311	GLU	-	expression tag	UNP W6A028
A	1312	ALA	-	expression tag	UNP W6A028
A	1313	PRO	-	expression tag	UNP W6A028
A	1314	ARG	-	expression tag	UNP W6A028
A	1315	ASP	-	expression tag	UNP W6A028
A	1316	GLY	-	expression tag	UNP W6A028

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1317	GLN	-	expression tag	UNP W6A028
A	1318	ALA	-	expression tag	UNP W6A028
A	1319	TYR	-	expression tag	UNP W6A028
A	1320	VAL	-	expression tag	UNP W6A028
A	1321	ARG	-	expression tag	UNP W6A028
A	1322	LYS	-	expression tag	UNP W6A028
A	1323	ASP	-	expression tag	UNP W6A028
A	1324	GLY	-	expression tag	UNP W6A028
A	1325	GLU	-	expression tag	UNP W6A028
A	1326	TRP	-	expression tag	UNP W6A028
A	1327	VAL	-	expression tag	UNP W6A028
A	1328	LEU	-	expression tag	UNP W6A028
A	1329	LEU	-	expression tag	UNP W6A028
A	1330	SER	-	expression tag	UNP W6A028
A	1331	THR	-	expression tag	UNP W6A028
A	1332	PHE	-	expression tag	UNP W6A028
A	1333	LEU	-	expression tag	UNP W6A028
A	1334	GLY	-	expression tag	UNP W6A028
A	1335	HIS	-	expression tag	UNP W6A028
A	1336	HIS	-	expression tag	UNP W6A028
A	1337	HIS	-	expression tag	UNP W6A028
A	1338	HIS	-	expression tag	UNP W6A028
A	1339	HIS	-	expression tag	UNP W6A028
A	1340	HIS	-	expression tag	UNP W6A028
B	751	SER	ARG	engineered mutation	UNP W6A028
B	1020	GLN	ARG	engineered mutation	UNP W6A028
B	1295	GLU	-	expression tag	UNP W6A028
B	1296	PHE	-	expression tag	UNP W6A028
B	1297	ARG	-	expression tag	UNP W6A028
B	1298	LEU	-	expression tag	UNP W6A028
B	1299	VAL	-	expression tag	UNP W6A028
B	1300	PRO	-	expression tag	UNP W6A028
B	1301	ARG	-	expression tag	UNP W6A028
B	1302	GLY	-	expression tag	UNP W6A028
B	1303	SER	-	expression tag	UNP W6A028
B	1304	PRO	-	expression tag	UNP W6A028
B	1305	GLY	-	expression tag	UNP W6A028
B	1306	SER	-	expression tag	UNP W6A028
B	1307	GLY	-	expression tag	UNP W6A028
B	1308	TYR	-	expression tag	UNP W6A028
B	1309	ILE	-	expression tag	UNP W6A028
B	1310	PRO	-	expression tag	UNP W6A028

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1311	GLU	-	expression tag	UNP W6A028
B	1312	ALA	-	expression tag	UNP W6A028
B	1313	PRO	-	expression tag	UNP W6A028
B	1314	ARG	-	expression tag	UNP W6A028
B	1315	ASP	-	expression tag	UNP W6A028
B	1316	GLY	-	expression tag	UNP W6A028
B	1317	GLN	-	expression tag	UNP W6A028
B	1318	ALA	-	expression tag	UNP W6A028
B	1319	TYR	-	expression tag	UNP W6A028
B	1320	VAL	-	expression tag	UNP W6A028
B	1321	ARG	-	expression tag	UNP W6A028
B	1322	LYS	-	expression tag	UNP W6A028
B	1323	ASP	-	expression tag	UNP W6A028
B	1324	GLY	-	expression tag	UNP W6A028
B	1325	GLU	-	expression tag	UNP W6A028
B	1326	TRP	-	expression tag	UNP W6A028
B	1327	VAL	-	expression tag	UNP W6A028
B	1328	LEU	-	expression tag	UNP W6A028
B	1329	LEU	-	expression tag	UNP W6A028
B	1330	SER	-	expression tag	UNP W6A028
B	1331	THR	-	expression tag	UNP W6A028
B	1332	PHE	-	expression tag	UNP W6A028
B	1333	LEU	-	expression tag	UNP W6A028
B	1334	GLY	-	expression tag	UNP W6A028
B	1335	HIS	-	expression tag	UNP W6A028
B	1336	HIS	-	expression tag	UNP W6A028
B	1337	HIS	-	expression tag	UNP W6A028
B	1338	HIS	-	expression tag	UNP W6A028
B	1339	HIS	-	expression tag	UNP W6A028
B	1340	HIS	-	expression tag	UNP W6A028
C	751	SER	ARG	engineered mutation	UNP W6A028
C	1020	GLN	ARG	engineered mutation	UNP W6A028
C	1295	GLU	-	expression tag	UNP W6A028
C	1296	PHE	-	expression tag	UNP W6A028
C	1297	ARG	-	expression tag	UNP W6A028
C	1298	LEU	-	expression tag	UNP W6A028
C	1299	VAL	-	expression tag	UNP W6A028
C	1300	PRO	-	expression tag	UNP W6A028
C	1301	ARG	-	expression tag	UNP W6A028
C	1302	GLY	-	expression tag	UNP W6A028
C	1303	SER	-	expression tag	UNP W6A028
C	1304	PRO	-	expression tag	UNP W6A028

*Continued on next page...*

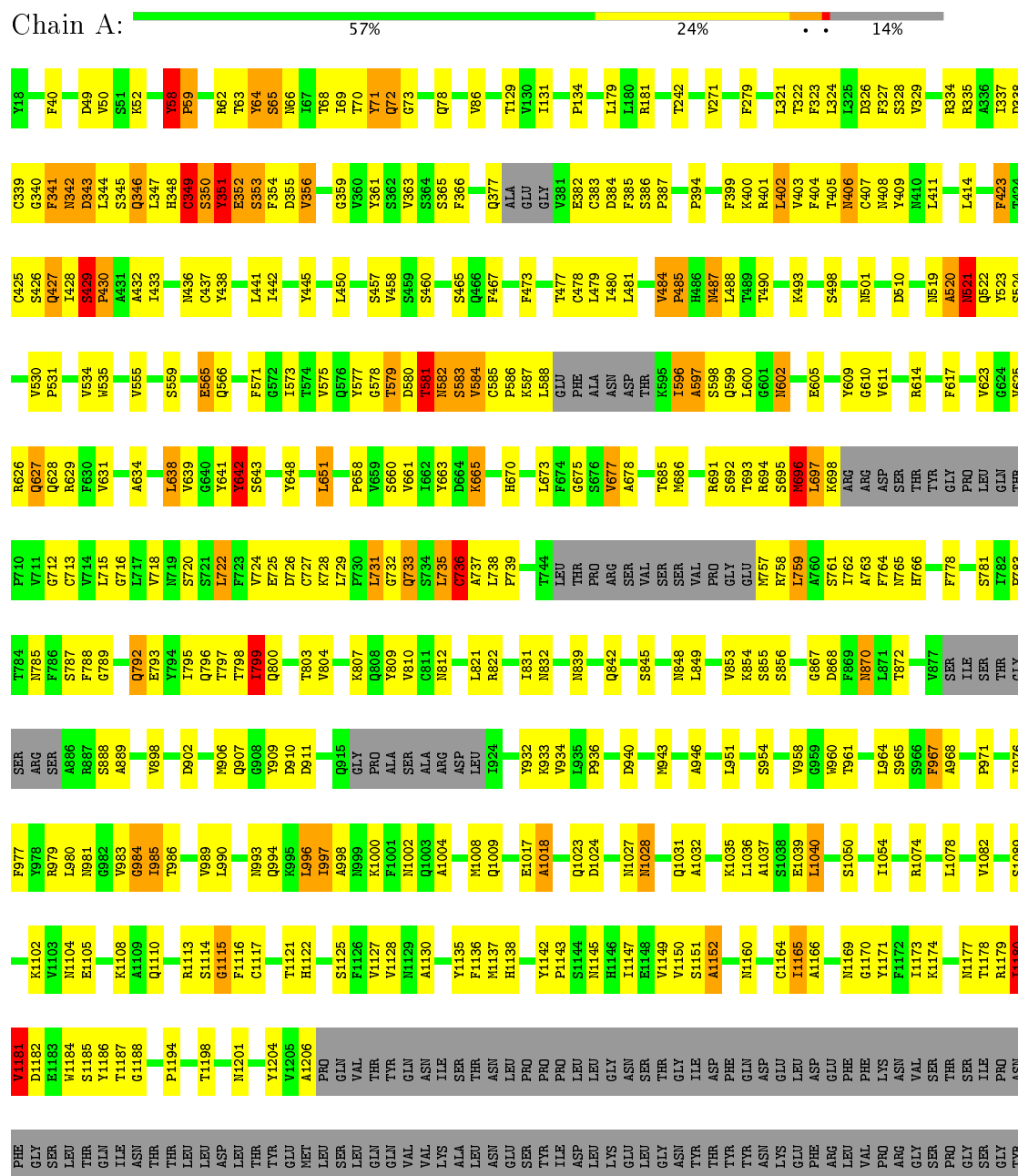
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	1305	GLY	-	expression tag	UNP W6A028
C	1306	SER	-	expression tag	UNP W6A028
C	1307	GLY	-	expression tag	UNP W6A028
C	1308	TYR	-	expression tag	UNP W6A028
C	1309	ILE	-	expression tag	UNP W6A028
C	1310	PRO	-	expression tag	UNP W6A028
C	1311	GLU	-	expression tag	UNP W6A028
C	1312	ALA	-	expression tag	UNP W6A028
C	1313	PRO	-	expression tag	UNP W6A028
C	1314	ARG	-	expression tag	UNP W6A028
C	1315	ASP	-	expression tag	UNP W6A028
C	1316	GLY	-	expression tag	UNP W6A028
C	1317	GLN	-	expression tag	UNP W6A028
C	1318	ALA	-	expression tag	UNP W6A028
C	1319	TYR	-	expression tag	UNP W6A028
C	1320	VAL	-	expression tag	UNP W6A028
C	1321	ARG	-	expression tag	UNP W6A028
C	1322	LYS	-	expression tag	UNP W6A028
C	1323	ASP	-	expression tag	UNP W6A028
C	1324	GLY	-	expression tag	UNP W6A028
C	1325	GLU	-	expression tag	UNP W6A028
C	1326	TRP	-	expression tag	UNP W6A028
C	1327	VAL	-	expression tag	UNP W6A028
C	1328	LEU	-	expression tag	UNP W6A028
C	1329	LEU	-	expression tag	UNP W6A028
C	1330	SER	-	expression tag	UNP W6A028
C	1331	THR	-	expression tag	UNP W6A028
C	1332	PHE	-	expression tag	UNP W6A028
C	1333	LEU	-	expression tag	UNP W6A028
C	1334	GLY	-	expression tag	UNP W6A028
C	1335	HIS	-	expression tag	UNP W6A028
C	1336	HIS	-	expression tag	UNP W6A028
C	1337	HIS	-	expression tag	UNP W6A028
C	1338	HIS	-	expression tag	UNP W6A028
C	1339	HIS	-	expression tag	UNP W6A028
C	1340	HIS	-	expression tag	UNP W6A028

### 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 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: S protein










## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	60000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

## 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  > 2$	RMSZ	# $ Z  > 2$
1	A	0.58	6/9006 (0.1%)	0.81	29/12245 (0.2%)
1	B	0.51	1/9006 (0.0%)	0.80	24/12245 (0.2%)
1	C	0.51	1/9010 (0.0%)	0.78	22/12250 (0.2%)
All	All	0.53	8/27022 (0.0%)	0.80	75/36740 (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
1	A	0	20
1	B	0	20
1	C	0	19
All	All	0	59

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	349	CYS	N-CA	-15.13	1.16	1.46
1	A	349	CYS	C-O	9.59	1.41	1.23
1	A	349	CYS	CB-SG	-8.35	1.68	1.82
1	A	696	MET	N-CA	6.67	1.59	1.46
1	A	59	PRO	N-CD	5.22	1.55	1.47

The worst 5 of 75 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	735	LEU	CA-CB-CG	11.61	142.00	115.30
1	B	735	LEU	CA-CB-CG	11.58	141.93	115.30
1	A	735	LEU	CA-CB-CG	11.51	141.78	115.30
1	A	349	CYS	O-C-N	-11.07	104.98	122.70
1	A	1040	LEU	CA-CB-CG	7.84	133.32	115.30

There are no chirality outliers.

5 of 59 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	520	ALA	Mainchain
1	A	639	VAL	Peptide
1	A	642	TYR	Peptide
1	A	65	SER	Peptide
1	A	733	GLN	Peptide

## 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	8806	0	8507	678	0
1	B	8806	0	8504	766	0
1	C	8810	0	8512	581	0
All	All	26422	0	25523	1686	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 33.

The worst 5 of 1686 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:339:CYS:SG	1:C:349:CYS:HB2	1.35	1.62
1:B:344:LEU:CD2	1:B:670:HIS:HB3	1.16	1.61
1:A:583:SER:HB2	1:A:609:TYR:CE1	1.37	1.60
1:C:335:ARG:HB3	1:C:354:PHE:CE2	1.34	1.60
1:B:344:LEU:HD22	1:B:670:HIS:CB	1.16	1.58

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	1128/1323 (85%)	964 (86%)	147 (13%)	17 (2%)	12	53
1	B	1128/1323 (85%)	965 (86%)	147 (13%)	16 (1%)	13	54
1	C	1128/1323 (85%)	966 (86%)	147 (13%)	15 (1%)	14	56
All	All	3384/3969 (85%)	2895 (86%)	441 (13%)	48 (1%)	18	54

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	583	SER
1	A	584	VAL
1	A	596	ILE
1	A	597	ALA
1	A	797	THR

### 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 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	973/1143 (85%)	919 (94%)	54 (6%)	25	61
1	B	973/1143 (85%)	923 (95%)	50 (5%)	28	62
1	C	974/1143 (85%)	924 (95%)	50 (5%)	28	62
All	All	2920/3429 (85%)	2766 (95%)	154 (5%)	31	61

5 of 154 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	473	PHE
1	B	627	GLN
1	C	854	LYS
1	B	481	LEU
1	B	565	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	599	GLN
1	B	832	ASN
1	C	1009	GLN
1	B	792	GLN
1	B	842	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](#)

There are no ligands in this entry.

### 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.