



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Sep 5, 2017 – 12:26 AM EDT

PDB ID : 5NMS
EMDB ID: : EMD-3459
Title : Hsp21 dodecamer, structural model based on cryo-EM and homology modelling
Authors : Rutsdottir, G.; Harmark, J.; Koeck, P.J.B.; Hebert, H.; Soderberg, C.A.G.; Emanuelsson, C.
Deposited on : unknown
Resolution : 10.00 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
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

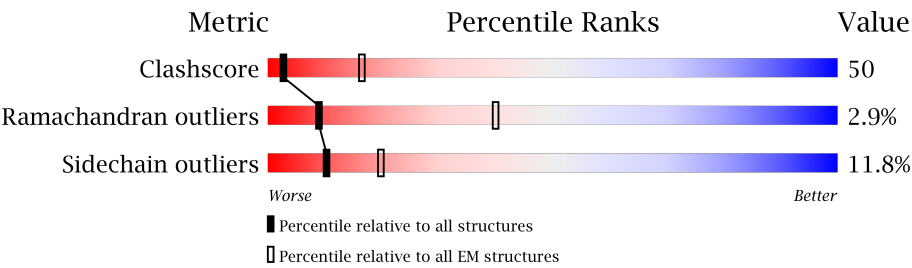
MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 10.00 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	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 ≥ 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	143	
1	C	143	
1	D	143	
1	G	143	
1	I	143	
1	J	143	
2	B	102	
2	E	102	
2	F	102	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	102	 50% 30% 9% 11%
2	K	102	 49% 32% 9% 10%
2	L	102	 48% 30% 11% 11%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 23808 atoms, of which 12006 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 25.3 kDa heat shock protein, chloroplastic.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	143	Total	C	H	N	O	S	0	0
			2301	707	1159	200	226	9		
1	C	143	Total	C	H	N	O	S	0	0
			2301	707	1159	200	226	9		
1	D	143	Total	C	H	N	O	S	0	0
			2301	707	1159	200	226	9		
1	G	143	Total	C	H	N	O	S	0	0
			2301	707	1159	200	226	9		
1	I	143	Total	C	H	N	O	S	0	0
			2301	707	1159	200	226	9		
1	J	143	Total	C	H	N	O	S	0	0
			2301	707	1159	200	226	9		

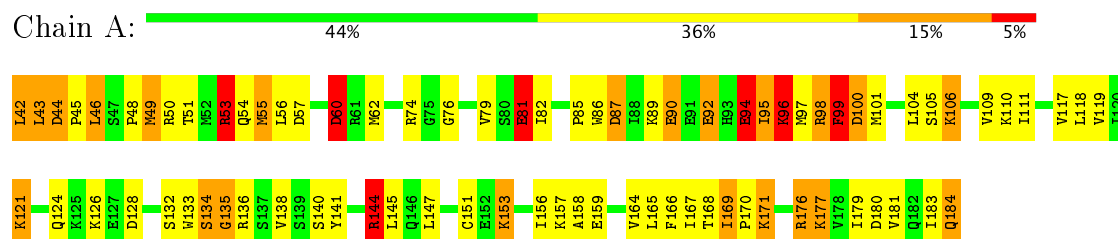
- Molecule 2 is a protein called 25.3 kDa heat shock protein, chloroplastic.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	B	102	Total	C	H	N	O	S	0	0
			1667	516	842	142	164	3		
2	E	102	Total	C	H	N	O	S	0	0
			1667	516	842	142	164	3		
2	F	102	Total	C	H	N	O	S	0	0
			1667	516	842	142	164	3		
2	H	102	Total	C	H	N	O	S	0	0
			1667	516	842	142	164	3		
2	K	102	Total	C	H	N	O	S	0	0
			1667	516	842	142	164	3		
2	L	102	Total	C	H	N	O	S	0	0
			1667	516	842	142	164	3		

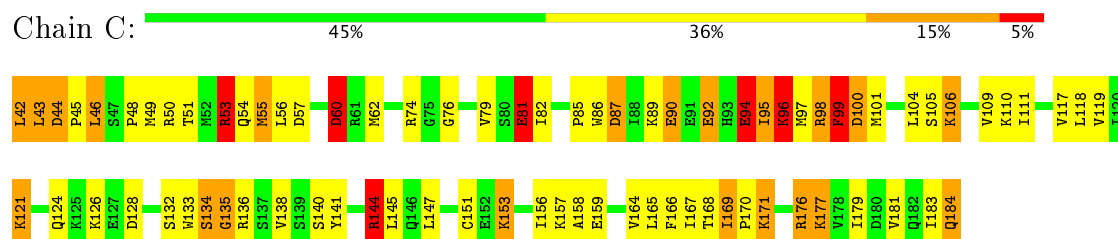
3 Residue-property plots

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.

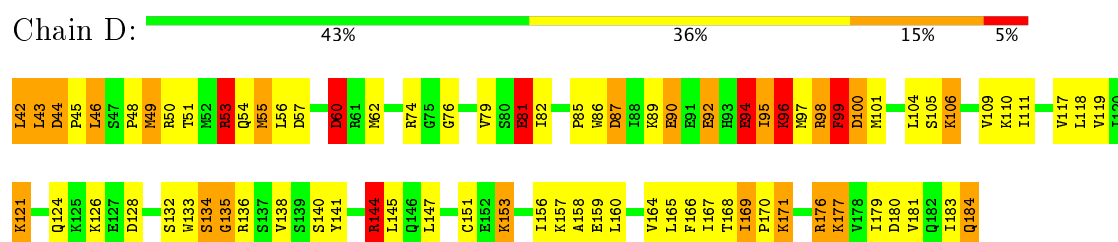
- Molecule 1: 25.3 kDa heat shock protein, chloroplastic



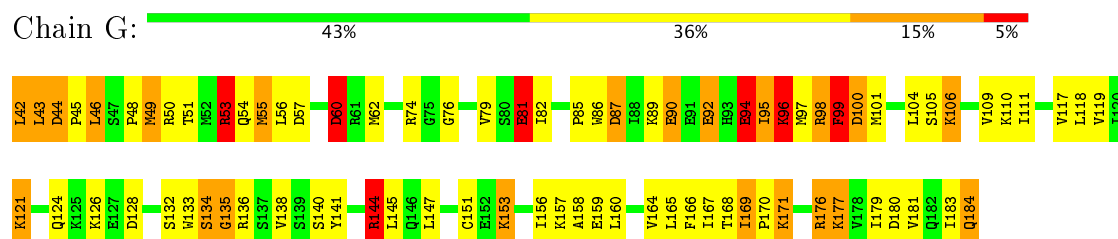
- Molecule 1: 25.3 kDa heat shock protein, chloroplastic



- Molecule 1: 25.3 kDa heat shock protein, chloroplastic

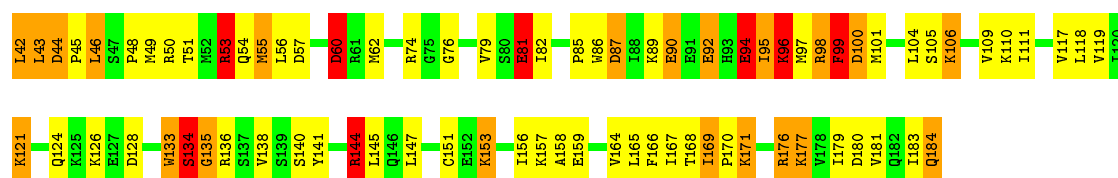


- Molecule 1: 25.3 kDa heat shock protein, chloroplastic



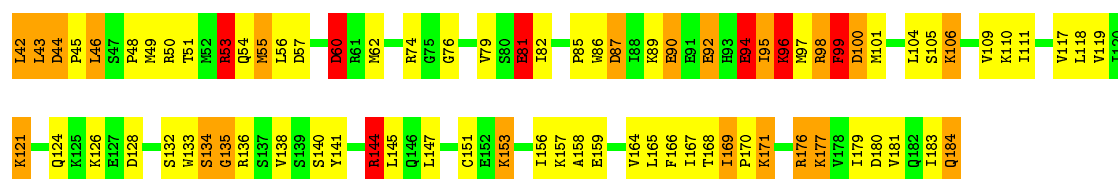
- Molecule 1: 25.3 kDa heat shock protein, chloroplastic

Chain I: 



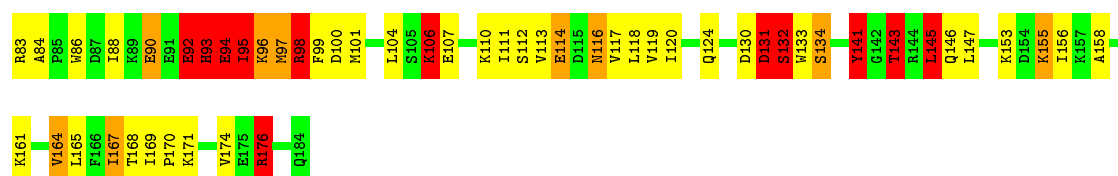
- Molecule 1: 25.3 kDa heat shock protein, chloroplastic

Chain J: 



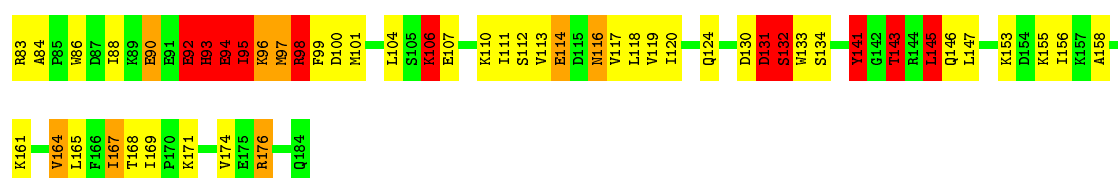
- Molecule 2: 25.3 kDa heat shock protein, chloroplastic

Chain B: 



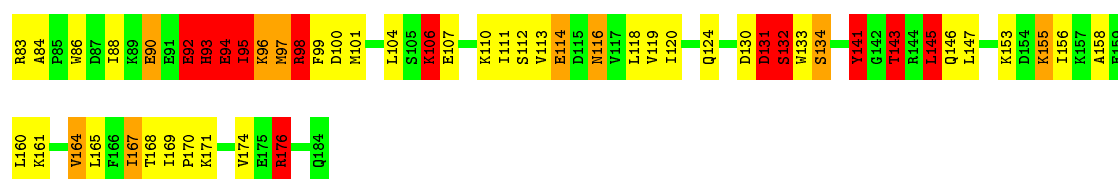
- Molecule 2: 25.3 kDa heat shock protein, chloroplastic

Chain E: 



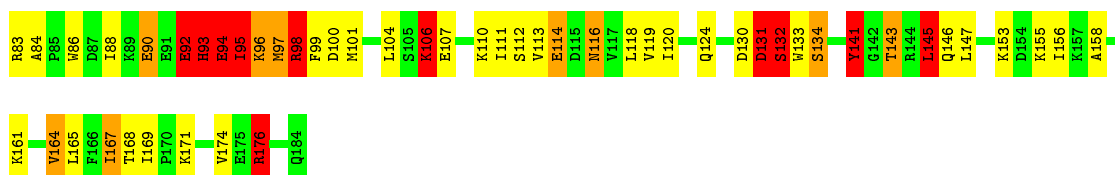
- Molecule 2: 25.3 kDa heat shock protein, chloroplastic

Chain F: 



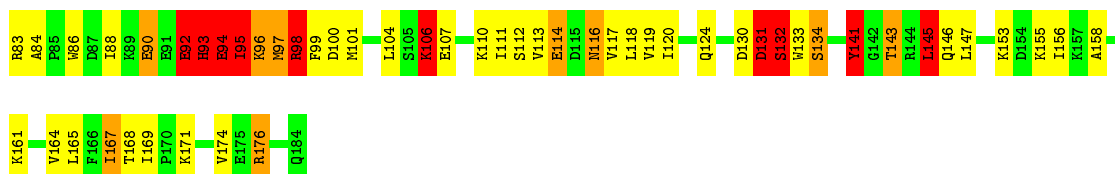
- Molecule 2: 25.3 kDa heat shock protein, chloroplastic

Chain H: 



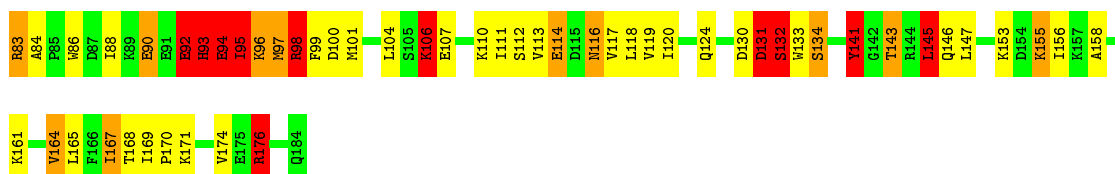
- Molecule 2: 25.3 kDa heat shock protein, chloroplastic

Chain K: 49% 32% 9% 10%



- Molecule 2: 25.3 kDa heat shock protein, chloroplastic

Chain L: 48% 30% 11% 11%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, D3	Depositor
Number of particles used	18407	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	JEOL 2100F	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	1.4	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	DIRECT ELECTRON DE-20 (5k x 3k)	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	0.76	0/1157	1.51	22/1549 (1.4%)
1	C	0.76	0/1157	1.51	22/1549 (1.4%)
1	D	0.75	0/1157	1.51	22/1549 (1.4%)
1	G	0.77	0/1157	1.51	22/1549 (1.4%)
1	I	0.77	0/1157	1.52	22/1549 (1.4%)
1	J	0.77	0/1157	1.51	22/1549 (1.4%)
2	B	0.76	0/836	2.00	15/1119 (1.3%)
2	E	0.76	0/836	2.03	15/1119 (1.3%)
2	F	0.75	0/836	2.02	14/1119 (1.3%)
2	H	0.77	0/836	2.07	13/1119 (1.2%)
2	K	0.79	0/836	1.89	15/1119 (1.3%)
2	L	0.78	0/836	1.99	15/1119 (1.3%)
All	All	0.77	0/11958	1.73	219/16008 (1.4%)

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	17
1	C	0	17
1	D	0	17
1	G	0	17
1	I	0	20
1	J	0	17
2	B	0	12
2	E	0	14
2	F	0	12
2	H	0	14
2	K	0	14
2	L	0	13
All	All	0	184

There are no bond length outliers.

The worst 5 of 219 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	H	98	ARG	NE-CZ-NH1	37.70	139.15	120.30
2	F	98	ARG	NE-CZ-NH1	35.83	138.21	120.30
2	E	98	ARG	NE-CZ-NH1	35.53	138.07	120.30
2	B	98	ARG	NE-CZ-NH1	35.22	137.91	120.30
2	L	98	ARG	NE-CZ-NH1	34.52	137.56	120.30

There are no chirality outliers.

5 of 184 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	42	LEU	Mainchain
1	A	44	ASP	Sidechain,Mainchain
1	A	49	MET	Mainchain
1	A	51	THR	Peptide
1	A	53	ARG	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	1142	1159	1156	145	0
1	C	1142	1159	1156	144	0
1	D	1142	1159	1156	145	0
1	G	1142	1159	1156	150	0
1	I	1142	1159	1156	147	0
1	J	1142	1159	1156	141	0
2	B	825	842	839	120	0
2	E	825	842	839	115	0
2	F	825	842	839	115	0
2	H	825	842	839	119	0
2	K	825	842	839	122	0
2	L	825	842	839	118	0
All	All	11802	12006	11970	1195	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 50.

The worst 5 of 1195 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)
2:B:98:ARG:HH21	2:B:164:VAL:HG21	1.21	1.05
2:H:98:ARG:HH21	2:H:164:VAL:HG21	1.20	1.05
2:F:98:ARG:HH21	2:F:164:VAL:HG21	1.21	1.04
2:E:98:ARG:HH21	2:E:164:VAL:HG21	1.23	1.04
2:L:98:ARG:HH21	2:L:164:VAL:HG21	1.24	1.00

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	141/143 (99%)	131 (93%)	5 (4%)	5 (4%)	4	34
1	C	141/143 (99%)	131 (93%)	5 (4%)	5 (4%)	4	34
1	D	141/143 (99%)	131 (93%)	5 (4%)	5 (4%)	4	34
1	G	141/143 (99%)	131 (93%)	5 (4%)	5 (4%)	4	34
1	I	141/143 (99%)	130 (92%)	6 (4%)	5 (4%)	4	34
1	J	141/143 (99%)	131 (93%)	5 (4%)	5 (4%)	4	34
2	B	100/102 (98%)	92 (92%)	6 (6%)	2 (2%)	9	46
2	E	100/102 (98%)	92 (92%)	6 (6%)	2 (2%)	9	46
2	F	100/102 (98%)	92 (92%)	6 (6%)	2 (2%)	9	46
2	H	100/102 (98%)	92 (92%)	6 (6%)	2 (2%)	9	46
2	K	100/102 (98%)	92 (92%)	6 (6%)	2 (2%)	9	46
2	L	100/102 (98%)	92 (92%)	6 (6%)	2 (2%)	9	46
All	All	1446/1470 (98%)	1337 (92%)	67 (5%)	42 (3%)	9	38

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	94	GLU
1	A	140	SER
2	B	94	GLU
1	C	94	GLU
1	C	140	SER

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	132/132 (100%)	119 (90%)	13 (10%)	9	34
1	C	132/132 (100%)	119 (90%)	13 (10%)	9	34
1	D	132/132 (100%)	119 (90%)	13 (10%)	9	34
1	G	132/132 (100%)	119 (90%)	13 (10%)	9	34
1	I	132/132 (100%)	119 (90%)	13 (10%)	9	34
1	J	132/132 (100%)	119 (90%)	13 (10%)	9	34
2	B	95/95 (100%)	81 (85%)	14 (15%)	3	20
2	E	95/95 (100%)	81 (85%)	14 (15%)	3	20
2	F	95/95 (100%)	81 (85%)	14 (15%)	3	20
2	H	95/95 (100%)	81 (85%)	14 (15%)	3	20
2	K	95/95 (100%)	82 (86%)	13 (14%)	4	23
2	L	95/95 (100%)	81 (85%)	14 (15%)	3	20
All	All	1362/1362 (100%)	1201 (88%)	161 (12%)	10	27

5 of 161 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	116	ASN
1	G	144	ARG
2	L	93	HIS
2	F	134	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	46	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	146	GLN
1	G	54	GLN
2	K	146	GLN
1	D	54	GLN
1	I	54	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.