



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:25 PM GMT

PDB ID : 4LKX  
Title : Humanized antibody 4B12 Fab complexed with a CemX segment  
Authors : Chu, H.M.; Wright, J.; Chan, Y.H.; Lin, C.J.; Chang, T.W.; Lim, C.  
Deposited on : 2013-07-09  
Resolution : 1.92 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

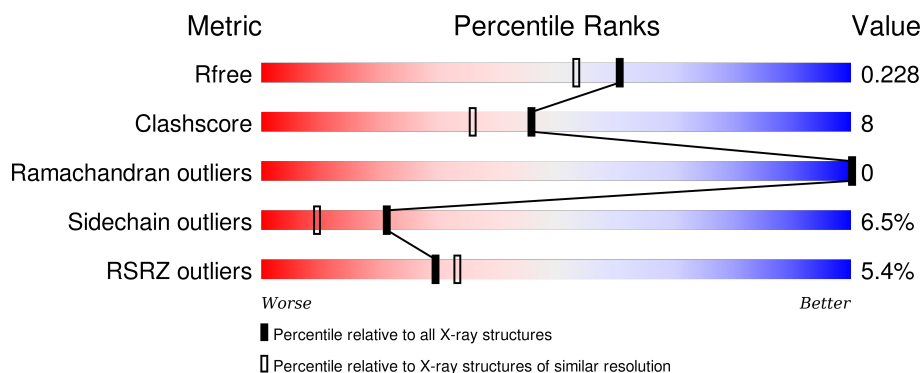
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

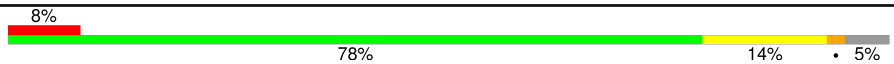
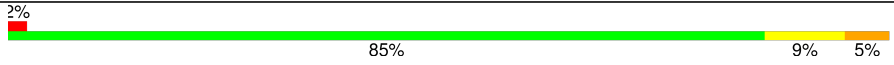
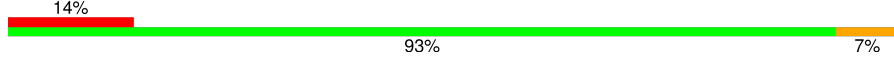
The reported resolution of this entry is 1.92 Å.

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))
$R_{free}$	91344	5832 (1.94-1.90)
Clashscore	102246	6540 (1.94-1.90)
Ramachandran outliers	100387	6464 (1.94-1.90)
Sidechain outliers	100360	6465 (1.94-1.90)
RSRZ outliers	91569	5846 (1.94-1.90)

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  $\geq 3$ , 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  $\leq 5\%$ . 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.

Mol	Chain	Length	Quality of chain
1	A	225	
2	B	219	
3	R	14	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3629 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 Fab fragment heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	214	Total	C	N	O	S	0	0	0
			1591	1008	262	316	5			

- Molecule 2 is a protein called Fab fragment light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	219	Total	C	N	O	S	0	0	0
			1679	1052	284	337	6			

- Molecule 3 is a protein called CemX segment.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	R	14	Total	C	N	O	0	0	0
			98	56	22	20			

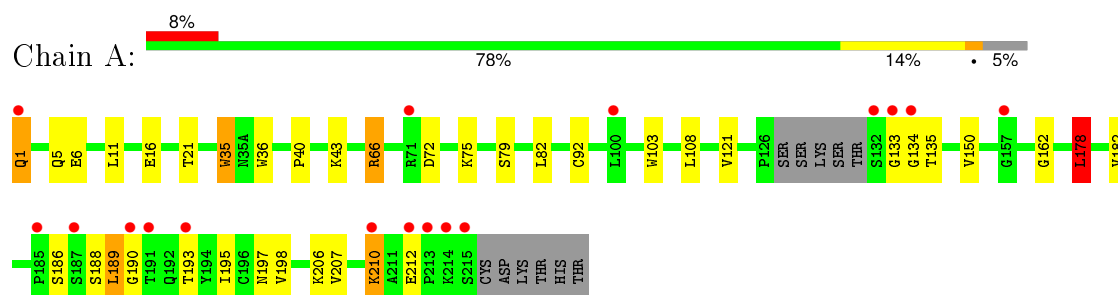
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	103	Total	O	0	0
			103	103		
4	B	140	Total	O	0	0
			140	140		
4	R	18	Total	O	0	0
			18	18		

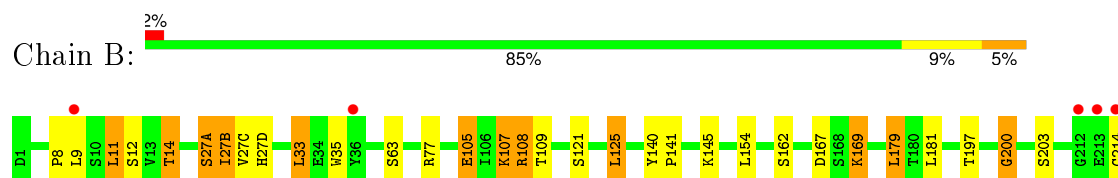
### 3 Residue-property plots [i](#)

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

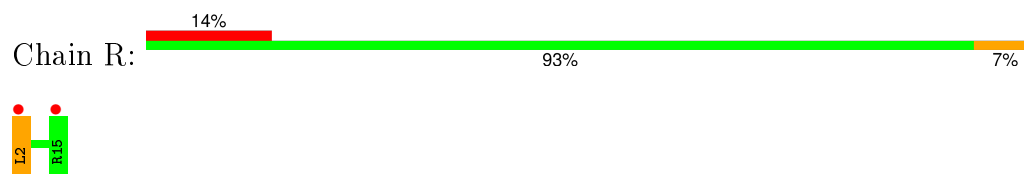
- Molecule 1: Fab fragment heavy chain



- Molecule 2: Fab fragment light chain



- Molecule 3: CemX segment



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.05Å 107.16Å 110.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.36 – 1.92 28.36 – 1.92	Depositor EDS
% Data completeness (in resolution range)	99.5 (28.36-1.92) 99.6 (28.36-1.92)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.62 (at 1.92Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.180 , 0.229 0.178 , 0.228	Depositor DCC
$R_{free}$ test set	1934 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.2	Xtriage
Anisotropy	0.914	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 53.1	EDS
Estimated twinning fraction	0.011 for -h,l,k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 38626 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3629	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.54% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.67	4/1630 (0.2%)	0.76	5/2226 (0.2%)
2	B	0.61	1/1717 (0.1%)	0.73	5/2332 (0.2%)
3	R	0.52	0/98	0.69	0/130
All	All	0.64	5/3445 (0.1%)	0.74	10/4688 (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	3
2	B	0	3
All	All	0	6

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	35	TRP	C-N	7.77	1.51	1.34
1	A	36	TRP	CD2-CE2	5.31	1.47	1.41
1	A	103	TRP	CD2-CE2	5.24	1.47	1.41
1	A	35	TRP	CD2-CE2	5.20	1.47	1.41
2	B	35	TRP	CD2-CE2	5.03	1.47	1.41

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	27(C)	VAL	O-C-N	-9.50	107.50	122.70
1	A	66	ARG	NE-CZ-NH2	-7.86	116.37	120.30
1	A	66	ARG	NE-CZ-NH1	7.64	124.12	120.30
1	A	35	TRP	CA-C-N	-6.50	102.91	117.20
2	B	27(C)	VAL	CA-C-N	6.20	130.83	117.20
2	B	27(D)	HIS	O-C-N	-6.15	112.86	122.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	178	LEU	CA-CB-CG	5.70	128.42	115.30
2	B	27(A)	SER	O-C-N	-5.60	113.73	122.70
1	A	82	LEU	C-N-CA	-5.22	108.65	121.70
2	B	33	LEU	CA-CB-CG	-5.10	103.57	115.30

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	189	LEU	Peptide
1	A	35	TRP	Mainchain
1	A	66	ARG	Sidechain
2	B	200	GLY	Peptide
2	B	27(A)	SER	Mainchain
2	B	27(B)	ILE	Mainchain

## 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	1591	0	1572	24	0
2	B	1679	0	1634	26	0
3	R	98	0	94	1	0
4	A	103	0	0	2	0
4	B	140	0	0	4	0
4	R	18	0	0	1	0
All	All	3629	0	3300	51	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 8.

All (51) 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:1:GLN:N	1:A:1:GLN:OE1	1.82	1.10

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:188:SER:C	1:A:190:GLY:HA3	1.79	1.01
1:A:16:GLU:OE2	4:A:371:HOH:O	1.84	0.95
1:A:189:LEU:N	1:A:190:GLY:HA3	1.90	0.84
1:A:1:GLN:CD	1:A:1:GLN:N	2.31	0.82
2:B:179:LEU:HD13	2:B:181:LEU:HD21	1.61	0.81
3:R:2:LEU:O	4:R:116:HOH:O	2.00	0.79
1:A:133:GLY:H	1:A:134:GLY:HA3	1.48	0.77
2:B:11:LEU:O	2:B:105:GLU:HG2	1.85	0.77
2:B:200:GLY:O	4:B:350:HOH:O	2.02	0.77
2:B:108:ARG:HG2	2:B:109:THR:N	2.01	0.74
2:B:169:LYS:HZ2	2:B:169:LYS:HB3	1.55	0.72
1:A:188:SER:O	1:A:190:GLY:HA3	1.89	0.71
1:A:186:SER:O	1:A:189:LEU:HB2	1.93	0.69
2:B:77:ARG:NH1	4:B:313:HOH:O	2.28	0.67
1:A:72:ASP:OD2	1:A:75:LYS:HD2	1.96	0.66
1:A:133:GLY:N	1:A:134:GLY:HA3	2.08	0.66
1:A:40:PRO:HB2	1:A:43:LYS:HD3	1.78	0.65
1:A:189:LEU:N	1:A:190:GLY:CA	2.61	0.63
2:B:167:ASP:OD2	2:B:169:LYS:HG3	2.01	0.60
2:B:27(B):ILE:HD11	2:B:33:LEU:HD13	1.82	0.60
2:B:8:PRO:HG3	2:B:11:LEU:HD13	1.84	0.60
2:B:108:ARG:CG	2:B:109:THR:N	2.69	0.56
1:A:198:VAL:HB	1:A:207:VAL:HG23	1.87	0.55
2:B:169:LYS:HZ2	2:B:169:LYS:CB	2.20	0.55
2:B:169:LYS:NZ	2:B:169:LYS:HB3	2.25	0.51
2:B:14:THR:HG22	2:B:107:LYS:HD2	1.91	0.51
2:B:121:SER:O	2:B:125:LEU:HD22	2.11	0.51
1:A:1:GLN:H1	1:A:1:GLN:CD	2.10	0.51
2:B:169:LYS:CB	2:B:169:LYS:NZ	2.74	0.51
2:B:63:SER:OG	4:B:414:HOH:O	2.20	0.49
1:A:21:THR:HG22	1:A:79:SER:OG	2.14	0.48
2:B:27(B):ILE:HD11	2:B:33:LEU:CD1	2.44	0.47
1:A:121:VAL:CG2	1:A:207:VAL:HG21	2.45	0.47
1:A:121:VAL:HG21	1:A:207:VAL:CG2	2.45	0.46
1:A:121:VAL:HG21	1:A:207:VAL:HG21	1.97	0.46
1:A:193:THR:HG22	1:A:195:ILE:HG13	1.97	0.46
2:B:145:LYS:HB3	2:B:197:THR:OG1	2.16	0.46
1:A:1:GLN:HB2	4:A:394:HOH:O	2.16	0.45
2:B:108:ARG:NH1	2:B:109:THR:O	2.46	0.44
1:A:210:LYS:HG2	1:A:212:GLU:CG	2.49	0.43
2:B:108:ARG:HD3	2:B:140:TYR:CG	2.53	0.43

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:162:GLY:O	1:A:182:VAL:HA	2.17	0.43
1:A:178:LEU:HD12	1:A:178:LEU:C	2.39	0.43
2:B:11:LEU:HA	2:B:11:LEU:HD12	1.79	0.43
2:B:27(B):ILE:CD1	2:B:33:LEU:HD12	2.49	0.42
2:B:140:TYR:CG	2:B:141:PRO:HA	2.55	0.42
2:B:27(B):ILE:CD1	2:B:33:LEU:CD1	2.97	0.42
1:A:6:GLU:HG3	1:A:92:CYS:SG	2.59	0.41
2:B:105:GLU:HG2	2:B:105:GLU:H	1.71	0.41
2:B:108:ARG:NH2	4:B:423:HOH:O	2.04	0.40

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	210/225 (93%)	203 (97%)	7 (3%)	0	100	100
2	B	217/219 (99%)	210 (97%)	7 (3%)	0	100	100
3	R	12/14 (86%)	12 (100%)	0	0	100	100
All	All	439/458 (96%)	425 (97%)	14 (3%)	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 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	181/192 (94%)	171 (94%)	10 (6%)	27	14
2	B	194/194 (100%)	180 (93%)	14 (7%)	18	7
3	R	9/9 (100%)	8 (89%)	1 (11%)	8	2
All	All	384/395 (97%)	359 (94%)	25 (6%)	21	9

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	GLN
1	A	5	GLN
1	A	11	LEU
1	A	108	LEU
1	A	135	THR
1	A	150	VAL
1	A	178	LEU
1	A	197	ASN
1	A	206	LYS
1	A	210	LYS
2	B	9	LEU
2	B	11	LEU
2	B	12	SER
2	B	14	THR
2	B	105	GLU
2	B	107	LYS
2	B	108	ARG
2	B	125	LEU
2	B	154	LEU
2	B	162	SER
2	B	169	LYS
2	B	179	LEU
2	B	203	SER
2	B	214	CYS
3	R	2	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

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.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	214/225 (95%)	0.32	17 (7%) 15 17	18, 29, 69, 101	0
2	B	219/219 (100%)	-0.03	5 (2%) 64 67	18, 30, 51, 112	0
3	R	14/14 (100%)	0.26	2 (14%) 4 4	19, 26, 50, 56	0
All	All	447/458 (97%)	0.14	24 (5%) 29 33	18, 29, 59, 112	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	214	CYS	10.8
1	A	132	SER	7.9
1	A	215	SER	6.1
1	A	133	GLY	5.4
1	A	187	SER	5.0
1	A	191	THR	4.8
3	R	2	LEU	4.5
2	B	213	GLU	4.3
1	A	214	LYS	4.0
1	A	213	PRO	3.7
2	B	212	GLY	3.6
1	A	193	THR	3.4
1	A	134	GLY	3.1
1	A	190	GLY	3.0
1	A	212	GLU	3.0
1	A	157	GLY	3.0
1	A	100	LEU	3.0
1	A	1	GLN	2.6
2	B	9	LEU	2.4
1	A	185	PRO	2.3
2	B	36	TYR	2.2
1	A	210	LYS	2.1
3	R	15	ARG	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	71	ARG	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.