



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 16, 2023 – 02:13 PM EST

PDB ID : 2OM1  
Title : Structure of human insulin in presence of thiocyanate at pH 6.5  
Authors : Norrman, M.; Schluckebier, G.  
Deposited on : 2007-01-20  
Resolution : 1.97 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp>

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:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

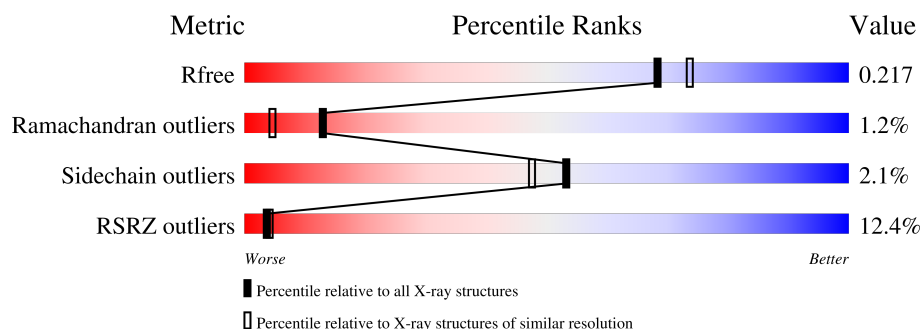
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.97 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))
$R_{free}$	130704	11647 (2.00-1.96)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 of the lower 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  $\leq 5\%$ . 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	1	21	<div> <div>10%</div> <div>100%</div> </div>
1	3	21	<div> <div>5%</div> <div>100%</div> </div>
1	A	21	<div> <div>5%</div> <div>100%</div> </div>
1	C	21	<div> <div></div> <div>100%</div> </div>
1	E	21	<div> <div>5%</div> <div>100%</div> </div>
1	G	21	<div> <div>10%</div> <div>100%</div> </div>
1	I	21	<div> <div>5%</div> <div>95%</div> <div>5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	21	<div> <div>10%</div> <div>100%</div> </div>
1	Q	21	<div> <div></div> <div>100%</div> </div>
1	S	21	<div> <div></div> <div>100%</div> </div>
1	U	21	<div> <div>5%</div> <div>100%</div> </div>
1	X	21	<div> <div></div> <div>100%</div> </div>
1	a	21	<div> <div>10%</div> <div>100%</div> </div>
1	c	21	<div> <div>5%</div> <div>95%</div> <div>5%</div> </div>
1	e	21	<div> <div>10%</div> <div>100%</div> </div>
1	g	21	<div> <div>5%</div> <div>100%</div> </div>
1	i	21	<div> <div>10%</div> <div>100%</div> </div>
1	k	21	<div> <div>24%</div> <div>100%</div> </div>
2	2	30	<div> <div>10%</div> <div>93%</div> <div>7%</div> </div>
2	4	30	<div> <div>23%</div> <div>93%</div> <div>7%</div> </div>
2	B	30	<div> <div>3%</div> <div>87%</div> <div>10%</div> </div>
2	D	30	<div> <div>7%</div> <div>97%</div> <div>3%</div> </div>
2	F	30	<div> <div>10%</div> <div>90%</div> <div>7%</div> </div>
2	H	30	<div> <div>10%</div> <div>90%</div> <div>7%</div> </div>
2	J	30	<div> <div>27%</div> <div>93%</div> <div>7%</div> </div>
2	L	30	<div> <div>20%</div> <div>93%</div> <div>7%</div> </div>
2	R	30	<div> <div>17%</div> <div>100%</div> </div>
2	T	30	<div> <div>13%</div> <div>93%</div> <div>7%</div> </div>
2	V	30	<div> <div>23%</div> <div>90%</div> <div>7%</div> </div>
2	Y	30	<div> <div>10%</div> <div>90%</div> <div>7%</div> </div>
2	b	30	<div> <div>27%</div> <div>97%</div> <div>3%</div> </div>
2	d	30	<div> <div>13%</div> <div>90%</div> <div>7%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	f	30	<div><div></div><div>17%</div><div>83%</div><div>13%</div><div></div></div>
2	h	30	<div><div></div><div>17%</div><div>93%</div><div></div><div></div></div>
2	j	30	<div><div></div><div>23%</div><div>83%</div><div>10%</div><div></div></div>
2	l	30	<div><div></div><div>23%</div><div>90%</div><div></div><div>7%</div></div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 8076 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 Insulin A chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	C	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	E	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	G	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	I	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	K	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	Q	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	S	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	U	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	X	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	1	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	3	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	a	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	c	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	e	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	g	21	Total 163	C 99	N 25	O 35	S 4	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	i	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
1	k	21	Total 162	C 99	N 25	O 34	S 4	0	0	0

- Molecule 2 is a protein called Insulin B chain.

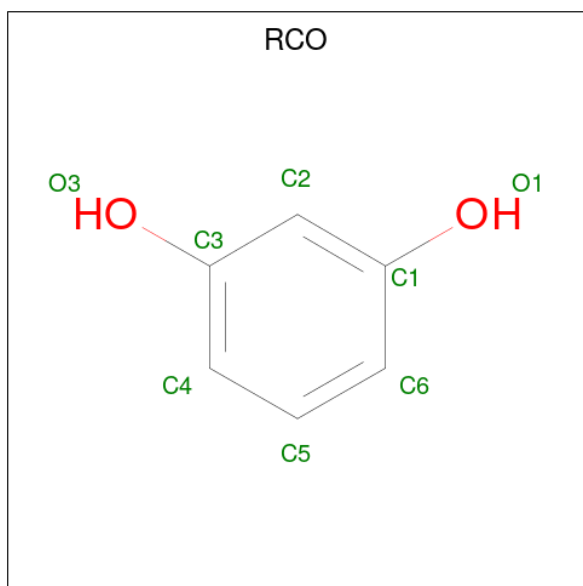
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	29	Total 229	C 150	N 38	O 39	S 2	0	0	0
2	D	29	Total 235	C 156	N 38	O 39	S 2	0	1	0
2	F	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	H	29	Total 229	C 150	N 38	O 39	S 2	0	0	0
2	J	30	Total 240	C 158	N 39	O 41	S 2	0	0	0
2	L	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	R	30	Total 241	C 158	N 40	O 41	S 2	0	0	0
2	T	30	Total 241	C 158	N 40	O 41	S 2	0	0	0
2	V	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	Y	29	Total 238	C 158	N 39	O 39	S 2	0	1	0
2	2	29	Total 234	C 155	N 38	O 39	S 2	0	1	0
2	4	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	b	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	d	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	f	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	h	29	Total 234	C 154	N 39	O 39	S 2	0	0	0
2	j	29	Total 234	C 154	N 39	O 39	S 2	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	1	28	Total	C	N	O	S	0	0	0
			221	146	37	36	2			

- Molecule 3 is RESORCINOL (three-letter code: RCO) (formula:  $C_6H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			8	6	2		
3	C	1	Total	C	O	0	0
			8	6	2		
3	E	1	Total	C	O	0	0
			8	6	2		
3	G	1	Total	C	O	0	0
			8	6	2		
3	I	1	Total	C	O	0	0
			8	6	2		
3	K	1	Total	C	O	0	0
			8	6	2		
3	Q	1	Total	C	O	0	0
			8	6	2		
3	S	1	Total	C	O	0	0
			8	6	2		
3	U	1	Total	C	O	0	0
			8	6	2		
3	X	1	Total	C	O	0	0
			8	6	2		

*Continued on next page...*

*Continued from previous page...*

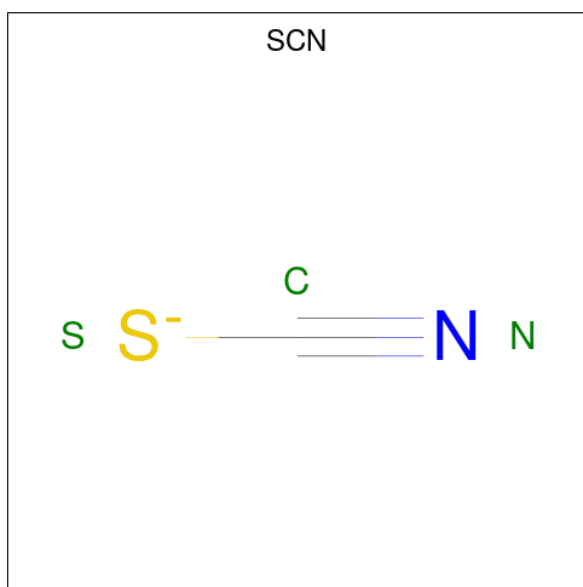
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	1	1	Total	C	O	0	0
			8	6	2		
3	3	1	Total	C	O	0	0
			8	6	2		
3	a	1	Total	C	O	0	0
			8	6	2		
3	c	1	Total	C	O	0	0
			8	6	2		
3	e	1	Total	C	O	0	0
			8	6	2		
3	g	1	Total	C	O	0	0
			8	6	2		
3	i	1	Total	C	O	0	0
			8	6	2		
3	k	1	Total	C	O	0	0
			8	6	2		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	D	1	Total	Zn	0	0
			1	1		
4	R	1	Total	Zn	0	0
			1	1		
4	Y	1	Total	Zn	0	0
			1	1		
4	b	1	Total	Zn	0	0
			1	1		
4	h	1	Total	Zn	0	0
			1	1		

- Molecule 5 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	N	S	0	0
			3	1	1	1		
5	D	1	Total	C	N	S	0	0
			3	1	1	1		
5	R	1	Total	C	N	S	0	0
			3	1	1	1		
5	Y	1	Total	C	N	S	0	0
			3	1	1	1		
5	b	1	Total	C	N	S	0	0
			3	1	1	1		
5	h	1	Total	C	N	S	0	0
			3	1	1	1		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	T	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	30	Total	O	0	0
			30	30		
7	B	22	Total	O	0	0
			22	22		
7	C	32	Total	O	0	0
			32	32		
7	D	29	Total	O	0	0
			29	29		
7	E	11	Total	O	0	0
			11	11		
7	F	17	Total	O	0	0
			17	17		
7	G	24	Total	O	0	0
			24	24		
7	H	22	Total	O	0	0
			22	22		
7	I	13	Total	O	0	0
			13	13		
7	J	21	Total	O	0	0
			21	21		
7	K	15	Total	O	0	0
			15	15		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	L	28	Total O 28 28	0	0
7	Q	31	Total O 31 31	0	0
7	R	33	Total O 33 33	0	0
7	S	29	Total O 29 29	0	0
7	T	33	Total O 33 33	0	0
7	U	23	Total O 23 23	0	0
7	V	30	Total O 30 30	0	0
7	X	22	Total O 22 22	0	0
7	Y	21	Total O 21 21	0	0
7	1	27	Total O 27 27	0	0
7	2	15	Total O 15 15	0	0
7	3	15	Total O 15 15	0	0
7	4	21	Total O 21 21	0	0
7	a	19	Total O 19 19	0	0
7	b	10	Total O 10 10	0	0
7	c	16	Total O 16 16	0	0
7	d	20	Total O 20 20	0	0
7	e	15	Total O 15 15	0	0
7	f	23	Total O 23 23	0	0
7	g	20	Total O 20 20	0	0
7	h	22	Total O 22 22	0	0

*Continued on next page...*

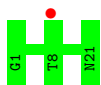
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	i	12	Total 12	O 12	0	0
7	j	19	Total 19	O 19	0	0
7	k	6	Total 6	O 6	0	0
7	l	9	Total 9	O 9	0	0

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

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 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: Insulin A chain

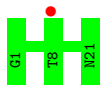


- Molecule 1: Insulin A chain

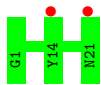


There are no outlier residues recorded for this chain.

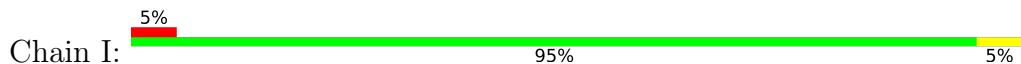
- Molecule 1: Insulin A chain



- Molecule 1: Insulin A chain

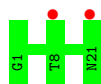


- Molecule 1: Insulin A chain



- Molecule 1: Insulin A chain





- Molecule 1: Insulin A chain

Chain Q:  100%

There are no outlier residues recorded for this chain.

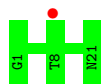
- Molecule 1: Insulin A chain

Chain S:  100%

There are no outlier residues recorded for this chain.

- Molecule 1: Insulin A chain

Chain U:  5%  
 100%



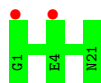
- Molecule 1: Insulin A chain

Chain X:  100%

There are no outlier residues recorded for this chain.

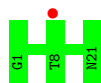
- Molecule 1: Insulin A chain

Chain 1:  10%  
 100%



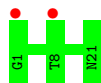
- Molecule 1: Insulin A chain

Chain 3:  5%  
 100%

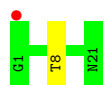
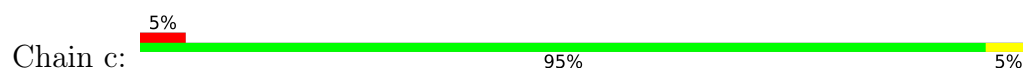


- Molecule 1: Insulin A chain

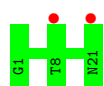
Chain a:  10%  
 100%



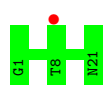
- Molecule 1: Insulin A chain



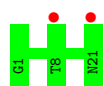
- Molecule 1: Insulin A chain



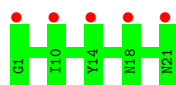
- Molecule 1: Insulin A chain



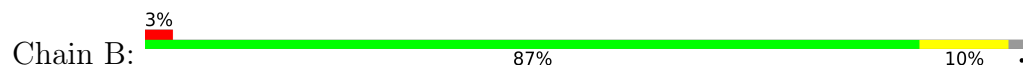
- Molecule 1: Insulin A chain



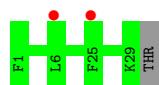
- Molecule 1: Insulin A chain



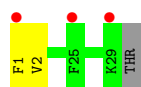
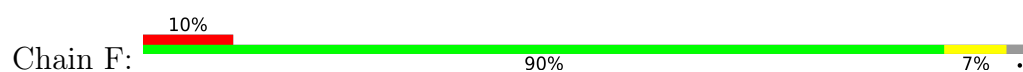
- Molecule 2: Insulin B chain



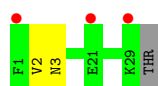
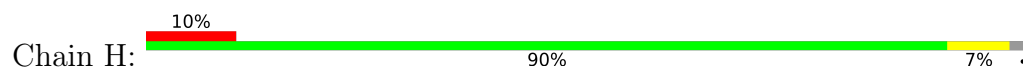
- Molecule 2: Insulin B chain



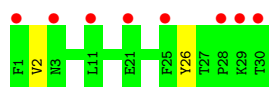
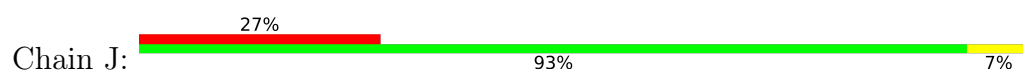
- Molecule 2: Insulin B chain



• Molecule 2: Insulin B chain



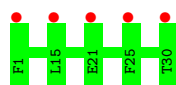
• Molecule 2: Insulin B chain



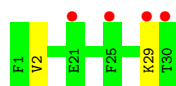
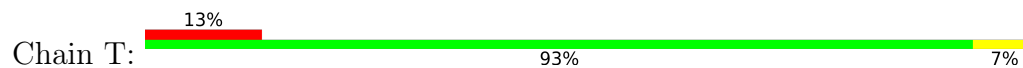
• Molecule 2: Insulin B chain



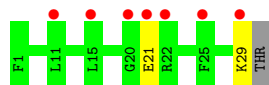
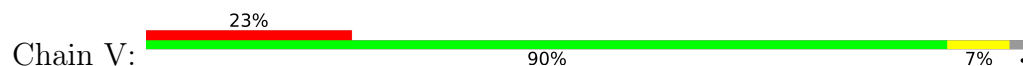
• Molecule 2: Insulin B chain



• Molecule 2: Insulin B chain

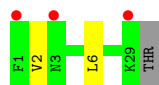
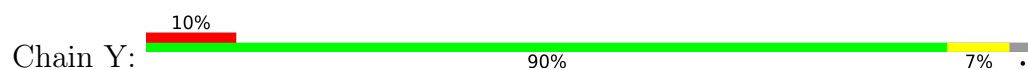


• Molecule 2: Insulin B chain

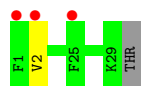


• Molecule 2: Insulin B chain

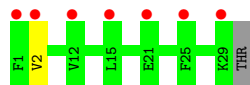




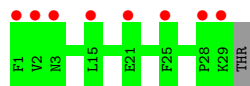
• Molecule 2: Insulin B chain



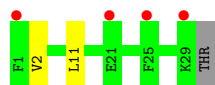
• Molecule 2: Insulin B chain



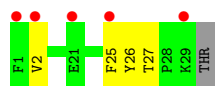
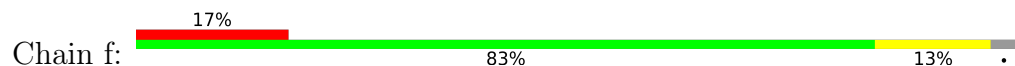
• Molecule 2: Insulin B chain



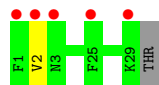
• Molecule 2: Insulin B chain



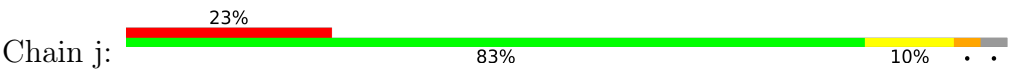
• Molecule 2: Insulin B chain



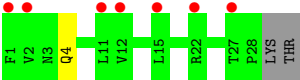
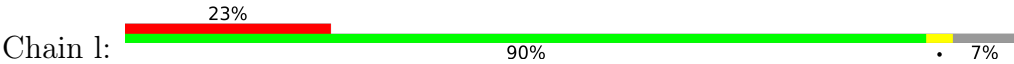
• Molecule 2: Insulin B chain



• Molecule 2: Insulin B chain



• Molecule 2: Insulin B chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.00Å 219.48Å 224.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 1.97 19.99 – 1.97	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.99-1.97) 99.5 (19.99-1.97)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.39 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.175 , 0.212 0.183 , 0.217	Depositor DCC
$R_{free}$ test set	5132 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.9	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 46.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8076	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 3.69% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SCN, ZN, GOL, RCO

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	1	0.68	0/164	0.65	0/220
1	3	0.90	0/164	0.57	0/220
1	A	0.86	0/164	0.65	0/220
1	C	0.72	0/164	0.66	0/220
1	E	0.61	0/164	0.56	0/220
1	G	0.66	0/164	0.64	0/220
1	I	0.59	0/164	0.64	0/220
1	K	0.61	0/164	0.54	0/220
1	Q	0.65	0/164	0.65	0/220
1	S	0.73	0/164	0.60	0/220
1	U	0.72	0/164	0.68	0/220
1	X	0.82	0/164	0.64	0/220
1	a	0.68	0/164	0.69	0/220
1	c	0.56	0/164	0.67	0/220
1	e	0.58	0/164	0.56	0/220
1	g	0.68	0/164	0.65	0/220
1	i	0.49	0/164	0.53	0/220
1	k	0.56	0/163	0.55	0/220
2	2	0.72	0/244	0.66	0/331
2	4	0.74	0/241	0.61	0/325
2	B	1.08	0/236	0.78	0/319
2	D	0.76	0/245	0.73	0/332
2	F	0.87	0/241	0.65	0/325
2	H	0.78	0/236	0.63	0/319
2	J	1.06	1/247 (0.4%)	0.70	0/334
2	L	0.79	0/241	0.62	0/325
2	R	0.81	0/248	0.73	0/335
2	T	0.74	0/248	0.71	0/335
2	V	0.78	0/241	0.60	0/325
2	Y	0.86	0/248	0.69	0/335
2	b	0.65	0/241	0.58	0/325
2	d	0.76	0/241	0.63	0/325

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
2	f	0.95	1/241 (0.4%)	0.68	0/325
2	h	0.77	0/241	0.71	0/325
2	j	0.78	0/241	0.71	0/325
2	l	0.69	0/228	0.57	0/309
All	All	0.77	2/7300 (0.0%)	0.65	0/9834

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
2	B	0	1
2	F	0	1
2	L	0	1
2	j	0	1
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	f	26	TYR	CD1-CE1	-5.72	1.30	1.39
2	J	26	TYR	CD1-CE1	-5.48	1.31	1.39

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1	PHE	Peptide
2	F	1	PHE	Peptide
2	L	28	PRO	Peptide
2	j	1	PHE	Peptide

## 5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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 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	1	19/21 (90%)	19 (100%)	0	0	100	100
1	3	19/21 (90%)	19 (100%)	0	0	100	100
1	A	19/21 (90%)	19 (100%)	0	0	100	100
1	C	19/21 (90%)	19 (100%)	0	0	100	100
1	E	19/21 (90%)	19 (100%)	0	0	100	100
1	G	19/21 (90%)	19 (100%)	0	0	100	100
1	I	19/21 (90%)	19 (100%)	0	0	100	100
1	K	19/21 (90%)	19 (100%)	0	0	100	100
1	Q	19/21 (90%)	19 (100%)	0	0	100	100
1	S	19/21 (90%)	19 (100%)	0	0	100	100
1	U	19/21 (90%)	19 (100%)	0	0	100	100
1	X	19/21 (90%)	19 (100%)	0	0	100	100
1	a	19/21 (90%)	19 (100%)	0	0	100	100
1	c	19/21 (90%)	19 (100%)	0	0	100	100
1	e	19/21 (90%)	19 (100%)	0	0	100	100
1	g	19/21 (90%)	19 (100%)	0	0	100	100
1	i	19/21 (90%)	19 (100%)	0	0	100	100
1	k	19/21 (90%)	19 (100%)	0	0	100	100
2	2	28/30 (93%)	26 (93%)	1 (4%)	1 (4%)	3	0
2	4	27/30 (90%)	26 (96%)	0	1 (4%)	3	0
2	B	27/30 (90%)	26 (96%)	1 (4%)	0	100	100
2	D	28/30 (93%)	28 (100%)	0	0	100	100
2	F	27/30 (90%)	25 (93%)	1 (4%)	1 (4%)	3	0
2	H	27/30 (90%)	26 (96%)	0	1 (4%)	3	0
2	J	28/30 (93%)	27 (96%)	0	1 (4%)	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	L	27/30 (90%)	27 (100%)	0	0	100	100
2	R	28/30 (93%)	27 (96%)	1 (4%)	0	100	100
2	T	28/30 (93%)	26 (93%)	1 (4%)	1 (4%)	3	0
2	V	27/30 (90%)	27 (100%)	0	0	100	100
2	Y	28/30 (93%)	27 (96%)	0	1 (4%)	3	0
2	b	27/30 (90%)	27 (100%)	0	0	100	100
2	d	27/30 (90%)	26 (96%)	0	1 (4%)	3	0
2	f	27/30 (90%)	26 (96%)	0	1 (4%)	3	0
2	h	27/30 (90%)	26 (96%)	0	1 (4%)	3	0
2	j	27/30 (90%)	26 (96%)	1 (4%)	0	100	100
2	l	26/30 (87%)	25 (96%)	1 (4%)	0	100	100
All	All	833/918 (91%)	816 (98%)	7 (1%)	10 (1%)	13	4

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	4	2	VAL
2	d	2	VAL
2	F	2	VAL
2	J	2	VAL
2	Y	2	VAL

### 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 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	1	20/20 (100%)	20 (100%)	0	100	100
1	3	20/20 (100%)	20 (100%)	0	100	100
1	A	20/20 (100%)	20 (100%)	0	100	100
1	C	20/20 (100%)	20 (100%)	0	100	100
1	E	20/20 (100%)	20 (100%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	20/20 (100%)	20 (100%)	0	100	100
1	I	20/20 (100%)	19 (95%)	1 (5%)	24	12
1	K	20/20 (100%)	20 (100%)	0	100	100
1	Q	20/20 (100%)	20 (100%)	0	100	100
1	S	20/20 (100%)	20 (100%)	0	100	100
1	U	20/20 (100%)	20 (100%)	0	100	100
1	X	20/20 (100%)	20 (100%)	0	100	100
1	a	20/20 (100%)	20 (100%)	0	100	100
1	c	20/20 (100%)	19 (95%)	1 (5%)	24	12
1	e	20/20 (100%)	20 (100%)	0	100	100
1	g	20/20 (100%)	20 (100%)	0	100	100
1	i	20/20 (100%)	20 (100%)	0	100	100
1	k	20/20 (100%)	20 (100%)	0	100	100
2	2	25/26 (96%)	25 (100%)	0	100	100
2	4	25/26 (96%)	25 (100%)	0	100	100
2	B	24/26 (92%)	22 (92%)	2 (8%)	11	3
2	D	25/26 (96%)	25 (100%)	0	100	100
2	F	25/26 (96%)	25 (100%)	0	100	100
2	H	24/26 (92%)	23 (96%)	1 (4%)	30	17
2	J	25/26 (96%)	25 (100%)	0	100	100
2	L	25/26 (96%)	25 (100%)	0	100	100
2	R	26/26 (100%)	26 (100%)	0	100	100
2	T	26/26 (100%)	25 (96%)	1 (4%)	33	21
2	V	25/26 (96%)	23 (92%)	2 (8%)	12	3
2	Y	26/26 (100%)	25 (96%)	1 (4%)	33	21
2	b	25/26 (96%)	25 (100%)	0	100	100
2	d	25/26 (96%)	24 (96%)	1 (4%)	31	19
2	f	25/26 (96%)	23 (92%)	2 (8%)	12	3
2	h	25/26 (96%)	25 (100%)	0	100	100
2	j	25/26 (96%)	21 (84%)	4 (16%)	2	0
2	l	23/26 (88%)	22 (96%)	1 (4%)	29	17

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	809/828 (98%)	792 (98%)	17 (2%)	53	47

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	j	25	PHE
2	l	4	GLN
2	Y	6	LEU
1	c	8	THR
2	d	11	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	S	21	ASN
2	h	4	GLN
2	V	3	ASN
2	l	4	GLN
2	d	3	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 31 ligands modelled in this entry, 6 are monoatomic - leaving 25 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	$\# Z  > 2$
3	RCO	i	1018	-	8,8,8	0.69	0	10,10,10	0.39	0
3	RCO	K	1004	-	8,8,8	0.64	0	10,10,10	0.69	0
3	RCO	e	1005	-	8,8,8	0.65	0	10,10,10	0.56	0
3	RCO	U	1001	-	8,8,8	0.56	0	10,10,10	0.95	0
3	RCO	c	1012	-	8,8,8	0.68	0	10,10,10	0.41	0
3	RCO	I	1013	-	8,8,8	0.63	0	10,10,10	0.54	0
3	RCO	Q	1003	-	8,8,8	0.79	0	10,10,10	1.15	1 (10%)
5	SCN	R	903	4	1,2,2	0.18	0	0,1,1	-	-
3	RCO	l	1007	-	8,8,8	0.69	0	10,10,10	0.78	0
5	SCN	h	902	-	1,2,2	0.59	0	0,1,1	-	-
3	RCO	a	1016	-	8,8,8	0.67	0	10,10,10	0.49	0
5	SCN	Y	904	4	1,2,2	0.26	0	0,1,1	-	-
6	GOL	T	1101	-	5,5,5	0.47	0	5,5,5	0.16	0
3	RCO	E	1014	-	8,8,8	0.73	0	10,10,10	0.22	0
3	RCO	C	1009	-	8,8,8	0.88	0	10,10,10	0.87	0
3	RCO	3	1006	-	8,8,8	0.90	0	10,10,10	0.69	0
5	SCN	D	906	4	1,2,2	0.20	0	0,1,1	-	-
5	SCN	B	905	4	1,2,2	0.11	0	0,1,1	-	-
3	RCO	k	1017	-	8,8,8	0.52	0	10,10,10	0.30	0
3	RCO	g	1010	-	8,8,8	0.77	0	10,10,10	0.57	0
3	RCO	A	1011	-	8,8,8	0.66	0	10,10,10	0.56	0
5	SCN	b	901	4	1,2,2	0.50	0	0,1,1	-	-
3	RCO	S	1008	-	8,8,8	0.98	0	10,10,10	0.67	0
3	RCO	X	1015	-	8,8,8	0.74	0	10,10,10	0.64	0
3	RCO	G	1002	-	8,8,8	0.83	0	10,10,10	0.51	0

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
3	RCO	i	1018	-	-	-	0/1/1/1
3	RCO	K	1004	-	-	-	0/1/1/1
3	RCO	e	1005	-	-	-	0/1/1/1
3	RCO	U	1001	-	-	-	0/1/1/1
3	RCO	c	1012	-	-	-	0/1/1/1
3	RCO	I	1013	-	-	-	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	RCO	Q	1003	-	-	-	0/1/1/1
3	RCO	1	1007	-	-	-	0/1/1/1
3	RCO	a	1016	-	-	-	0/1/1/1
6	GOL	T	1101	-	-	2/4/4/4	-
3	RCO	E	1014	-	-	-	0/1/1/1
3	RCO	C	1009	-	-	-	0/1/1/1
3	RCO	3	1006	-	-	-	0/1/1/1
3	RCO	k	1017	-	-	-	0/1/1/1
3	RCO	g	1010	-	-	-	0/1/1/1
3	RCO	A	1011	-	-	-	0/1/1/1
3	RCO	S	1008	-	-	-	0/1/1/1
3	RCO	X	1015	-	-	-	0/1/1/1
3	RCO	G	1002	-	-	-	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	Q	1003	RCO	C5-C4-C3	2.02	122.16	119.31

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	T	1101	GOL	C1-C2-C3-O3
6	T	1101	GOL	O2-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

## 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	1	21/21 (100%)	0.43	2 (9%) 8 9	23, 30, 39, 42	0
1	3	21/21 (100%)	0.44	1 (4%) 30 32	27, 34, 41, 42	0
1	A	21/21 (100%)	0.61	1 (4%) 30 32	22, 32, 40, 43	0
1	C	21/21 (100%)	0.25	0 100 100	22, 26, 32, 33	0
1	E	21/21 (100%)	0.35	1 (4%) 30 32	27, 34, 41, 42	0
1	G	21/21 (100%)	0.49	2 (9%) 8 9	27, 32, 38, 39	0
1	I	21/21 (100%)	0.72	1 (4%) 30 32	27, 36, 42, 43	0
1	K	21/21 (100%)	0.65	2 (9%) 8 9	28, 35, 40, 42	0
1	Q	21/21 (100%)	0.22	0 100 100	24, 28, 32, 34	0
1	S	21/21 (100%)	0.21	0 100 100	24, 29, 32, 35	0
1	U	21/21 (100%)	0.46	1 (4%) 30 32	25, 30, 35, 38	0
1	X	21/21 (100%)	0.22	0 100 100	24, 30, 37, 40	0
1	a	21/21 (100%)	0.50	2 (9%) 8 9	27, 35, 43, 45	0
1	c	21/21 (100%)	0.64	1 (4%) 30 32	26, 32, 46, 48	0
1	e	21/21 (100%)	0.66	2 (9%) 8 9	30, 39, 43, 45	0
1	g	21/21 (100%)	0.40	1 (4%) 30 32	24, 32, 39, 42	0
1	i	21/21 (100%)	0.79	2 (9%) 8 9	31, 40, 44, 46	0
1	k	21/21 (100%)	1.20	5 (23%) 0 0	41, 46, 50, 50	0
2	2	29/30 (96%)	1.13	3 (10%) 6 7	21, 28, 42, 48	0
2	4	29/30 (96%)	1.10	7 (24%) 0 0	26, 29, 45, 52	0
2	B	29/30 (96%)	0.90	1 (3%) 45 48	21, 27, 45, 48	0
2	D	29/30 (96%)	0.68	2 (6%) 16 18	20, 24, 35, 42	0
2	F	29/30 (96%)	1.06	3 (10%) 6 7	24, 30, 49, 55	0
2	H	29/30 (96%)	1.10	3 (10%) 6 7	25, 30, 45, 51	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
2	J	30/30 (100%)	1.60	8 (26%)	0	0	24, 33, 51, 61	0
2	L	29/30 (96%)	1.06	6 (20%)	1	0	24, 30, 44, 48	0
2	R	30/30 (100%)	1.08	5 (16%)	1	1	23, 25, 42, 48	0
2	T	30/30 (100%)	0.97	4 (13%)	3	3	22, 27, 43, 54	0
2	V	29/30 (96%)	1.14	7 (24%)	0	0	23, 27, 44, 50	0
2	Y	29/30 (96%)	1.00	3 (10%)	6	7	21, 29, 43, 51	0
2	b	29/30 (96%)	1.57	8 (27%)	0	0	29, 34, 52, 58	0
2	d	29/30 (96%)	1.23	4 (13%)	2	3	25, 29, 45, 54	0
2	f	29/30 (96%)	1.62	5 (17%)	1	1	27, 35, 49, 54	0
2	h	29/30 (96%)	1.03	5 (17%)	1	1	25, 28, 45, 50	0
2	j	29/30 (96%)	1.36	7 (24%)	0	0	29, 36, 49, 56	0
2	l	28/30 (93%)	1.55	7 (25%)	0	0	31, 37, 47, 52	0
All	All	902/918 (98%)	0.90	112 (12%)	4	4	20, 32, 47, 61	0

The worst 5 of 112 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	f	1	PHE	13.8
2	b	1	PHE	12.6
2	h	1	PHE	11.2
2	j	1	PHE	9.8
2	l	1	PHE	9.0

## 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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	GOL	T	1101	6/6	0.94	0.21	35,38,38,39	0
3	RCO	c	1012	8/8	0.95	0.10	25,26,29,30	0
3	RCO	i	1018	8/8	0.95	0.15	28,30,31,31	0
3	RCO	k	1017	8/8	0.95	0.11	29,31,34,37	0
3	RCO	I	1013	8/8	0.95	0.12	23,24,25,25	0
3	RCO	A	1011	8/8	0.96	0.10	20,23,25,26	0
5	SCN	b	901	3/3	0.96	0.10	26,26,27,31	0
3	RCO	a	1016	8/8	0.96	0.13	24,26,28,28	0
3	RCO	S	1008	8/8	0.97	0.08	19,20,23,25	0
3	RCO	X	1015	8/8	0.97	0.10	23,25,25,26	0
3	RCO	E	1014	8/8	0.97	0.14	23,25,27,28	0
3	RCO	C	1009	8/8	0.97	0.13	18,19,20,21	0
3	RCO	Q	1003	8/8	0.98	0.07	15,16,18,19	0
3	RCO	G	1002	8/8	0.98	0.06	15,17,18,19	0
3	RCO	e	1005	8/8	0.98	0.09	22,24,25,26	0
3	RCO	g	1010	8/8	0.98	0.08	20,22,23,23	0
3	RCO	U	1001	8/8	0.98	0.06	16,17,18,19	0
3	RCO	K	1004	8/8	0.98	0.07	22,23,24,25	0
5	SCN	B	905	3/3	0.98	0.11	29,29,29,31	0
5	SCN	Y	904	3/3	0.98	0.08	23,23,25,29	0
3	RCO	l	1007	8/8	0.98	0.10	20,22,23,23	0
3	RCO	3	1006	8/8	0.98	0.07	19,20,22,23	0
5	SCN	D	906	3/3	0.99	0.09	18,18,18,21	0
5	SCN	R	903	3/3	0.99	0.10	21,21,22,23	0
4	ZN	Y	804	1/1	0.99	0.07	21,21,21,21	0
4	ZN	h	805	1/1	0.99	0.09	26,26,26,26	0
5	SCN	h	902	3/3	0.99	0.12	29,29,30,32	0
4	ZN	B	801	1/1	0.99	0.08	21,21,21,21	0
4	ZN	b	806	1/1	1.00	0.06	24,24,24,24	0
4	ZN	R	803	1/1	1.00	0.09	19,19,19,19	0
4	ZN	D	802	1/1	1.00	0.09	19,19,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.