



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 17, 2021 – 04:06 pm GMT

PDB ID : 7Q63  
Title : The tandem SH2 domains of SYK  
Authors : Bradshaw, W.J.; Katis, V.L.; Chen, Z.; Bountra, C.; von Delft, F.; Gileadi, O.; Brennan, P.E.  
Deposited on : 2021-11-05  
Resolution : 1.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

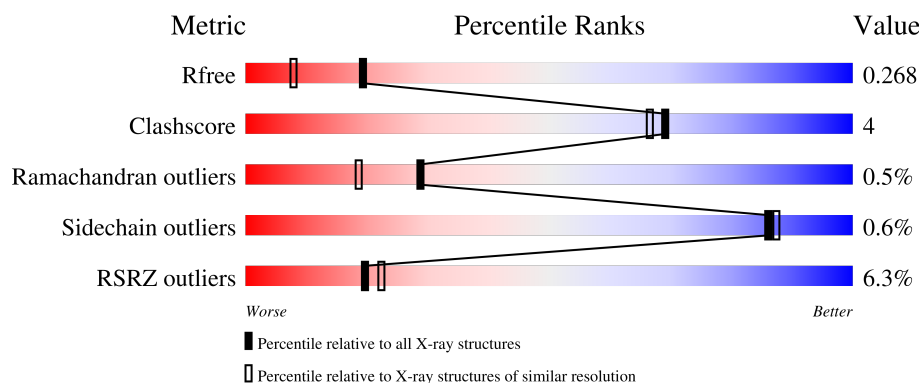
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 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	AAA	265	
1	BBB	265	
1	CCC	265	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	EDO	BBB	305	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6380 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.

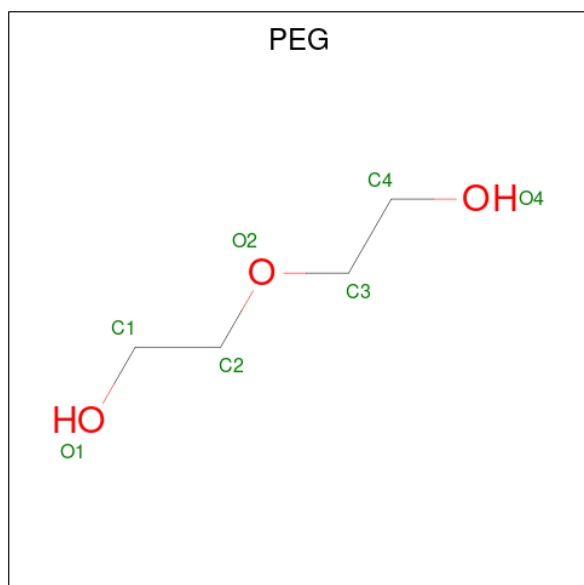
- Molecule 1 is a protein called Tyrosine-protein kinase SYK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	259	Total	C	N	O	S	0	7	0
			2102	1329	378	388	7			
1	BBB	259	Total	C	N	O	S	0	4	0
			2081	1318	373	383	7			
1	CCC	225	Total	C	N	O	S	0	6	0
			1834	1166	330	330	8			

There are 3 discrepancies between the modelled and reference sequences:

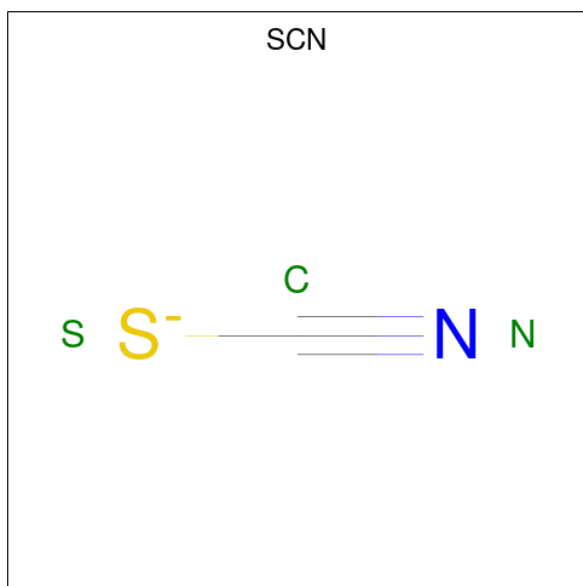
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	5	SER	-	expression tag	UNP P43405
BBB	5	SER	-	expression tag	UNP P43405
CCC	5	SER	-	expression tag	UNP P43405

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total	C	O	0	0
			7	4	3		

- Molecule 3 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).

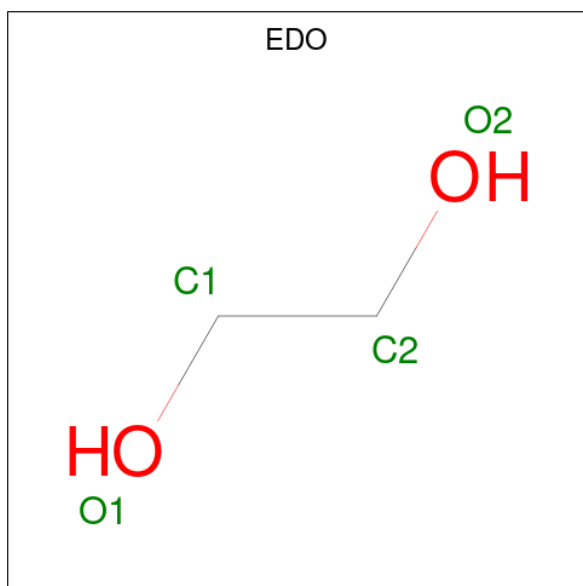


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	AAA	1	Total	C	N	S	0	0
			3	1	1	1		
3	AAA	1	Total	C	N	S	0	0
			3	1	1	1		
3	AAA	1	Total	C	N	S	0	0
			3	1	1	1		
3	BBB	1	Total	C	N	S	0	0
			3	1	1	1		
3	BBB	1	Total	C	N	S	0	0
			3	1	1	1		
3	BBB	1	Total	C	N	S	0	0
			3	1	1	1		
3	BBB	1	Total	C	N	S	0	0
			3	1	1	1		
3	CCC	1	Total	C	N	S	0	0
			3	1	1	1		
3	CCC	1	Total	C	N	S	0	0
			3	1	1	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total Cl 1 1	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	BBB	1	Total C O 4 2 2	0	0

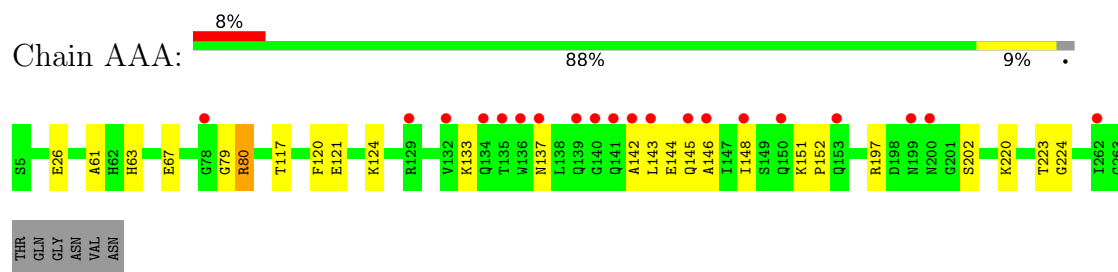
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AAA	134	Total O 134 134	0	0
6	BBB	122	Total O 122 122	0	0
6	CCC	68	Total O 68 68	0	1

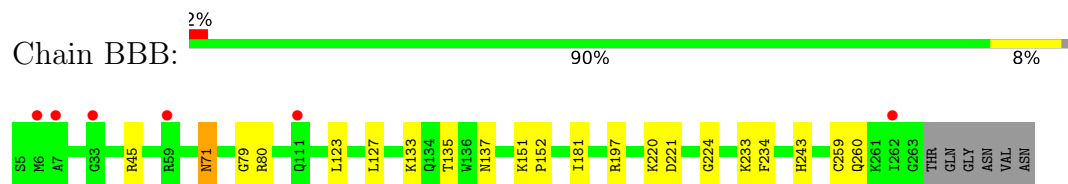
### 3 Residue-property plots [i](#)

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.

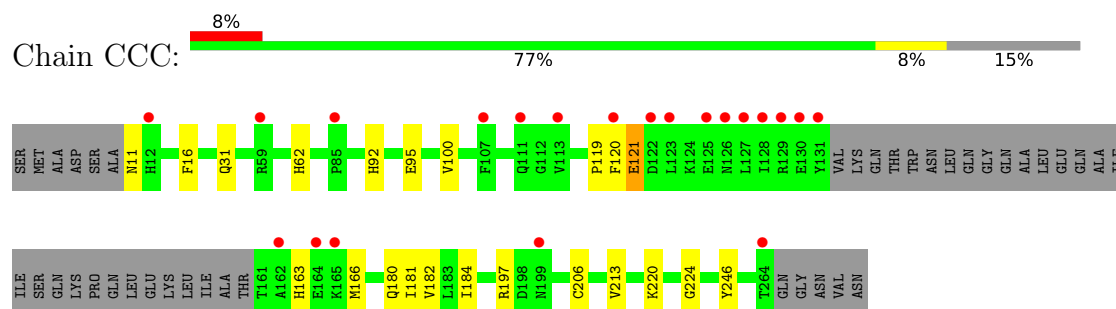
#### • Molecule 1: Tyrosine-protein kinase SYK



#### • Molecule 1: Tyrosine-protein kinase SYK



#### • Molecule 1: Tyrosine-protein kinase SYK



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.90Å 152.58Å 146.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	76.29 – 1.90 76.29 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (76.29-1.90) 100.0 (76.29-1.90)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.25 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.217 , 0.264 0.224 , 0.268	Depositor DCC
$R_{free}$ test set	1972 reflections (2.59%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.9	Xtriage
Anisotropy	0.471	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.011 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.025 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6380	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, SCN, CL, PEG

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	AAA	0.68	0/2152	0.81	1/2900 (0.0%)
1	BBB	0.70	0/2131	0.83	2/2873 (0.1%)
1	CCC	0.68	0/1895	0.82	2/2550 (0.1%)
All	All	0.69	0/6178	0.82	5/8323 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	BBB	197	ARG	NE-CZ-NH2	-8.22	116.19	120.30
1	BBB	197	ARG	NE-CZ-NH1	7.71	124.16	120.30
1	AAA	197	ARG	NE-CZ-NH2	7.55	124.07	120.30
1	CCC	197	ARG	NE-CZ-NH2	-5.18	117.71	120.30
1	CCC	121	GLU	CB-CA-C	-5.09	100.22	110.40

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	AAA	2102	0	2084	16	0
1	BBB	2081	0	2066	15	0
1	CCC	1834	0	1821	15	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	AAA	7	0	10	2	0
3	AAA	9	0	0	0	0
3	BBB	12	0	0	0	0
3	CCC	6	0	0	0	0
4	AAA	1	0	0	0	0
5	BBB	4	0	6	6	0
6	AAA	134	0	0	2	0
6	BBB	122	0	0	0	0
6	CCC	68	0	0	2	0
All	All	6380	0	5987	44	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 4.

The worst 5 of 44 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:CCC:246:TYR:HE2	6:CCC:445:HOH:O	1.63	0.81
1:AAA:79[B]:GLY:O	1:AAA:80[B]:ARG:HB3	1.81	0.80
1:BBB:181:ILE:HG22	1:BBB:259:CYS:SG	2.29	0.73
1:AAA:142:ALA:O	1:AAA:146:ALA:N	2.24	0.69
1:CCC:206[B]:CYS:SG	1:CCC:213:VAL:CG1	2.83	0.66

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	AAA	264/265 (100%)	249 (94%)	12 (4%)	3 (1%)	14	5
1	BBB	261/265 (98%)	253 (97%)	6 (2%)	2 (1%)	19	9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	CCC	227/265 (86%)	222 (98%)	4 (2%)	1 (0%)	34	24
All	All	752/795 (95%)	724 (96%)	22 (3%)	6 (1%)	29	9

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	80[A]	ARG
1	AAA	80[B]	ARG
1	CCC	120	PHE
1	AAA	143	LEU
1	BBB	79[A]	GLY

### 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	AAA	222/224 (99%)	222 (100%)	0	100	100
1	BBB	220/224 (98%)	217 (99%)	3 (1%)	67	65
1	CCC	197/224 (88%)	196 (100%)	1 (0%)	88	89
All	All	639/672 (95%)	635 (99%)	4 (1%)	86	87

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	71	ASN
1	BBB	80	ARG
1	BBB	135	THR
1	CCC	181	ILE

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

### 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 12 ligands modelled in this entry, 1 is monoatomic - leaving 11 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	AAA	259/265 (97%)	0.57	20 (7%) 13 15	32, 47, 116, 157	0
1	BBB	259/265 (97%)	0.29	6 (2%) 60 63	33, 48, 73, 91	0
1	CCC	225/265 (84%)	0.71	21 (9%) 8 10	39, 57, 105, 129	0
All	All	743/795 (93%)	0.51	47 (6%) 20 22	32, 50, 98, 157	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	143	LEU	9.8
1	CCC	131	TYR	8.7
1	AAA	148	ILE	5.9
1	CCC	162	ALA	5.4
1	AAA	141	GLN	5.1

### 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( $\text{\AA}^2$ )	Q<0.9
2	PEG	AAA	301	7/7	0.82	0.17	42,53,70,70	0
5	EDO	BBB	305	4/4	0.82	0.33	48,63,64,74	0
3	SCN	BBB	301	3/3	0.89	0.13	92,92,96,97	0
3	SCN	CCC	301	3/3	0.93	0.10	38,38,49,52	0
3	SCN	BBB	302	3/3	0.96	0.20	43,43,59,66	0
3	SCN	BBB	304	3/3	0.96	0.16	56,56,60,67	0
3	SCN	AAA	304	3/3	0.97	0.28	51,51,68,71	0
3	SCN	CCC	302	3/3	0.97	0.23	56,56,69,72	0
4	CL	AAA	305	1/1	0.97	0.10	56,56,56,56	0
3	SCN	AAA	303	3/3	0.97	0.17	57,57,73,76	0
3	SCN	BBB	303	3/3	0.98	0.15	49,49,54,64	0
3	SCN	AAA	302	3/3	0.99	0.10	43,43,49,50	3

## 6.5 Other polymers ⓘ

There are no such residues in this entry.