



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:09 PM GMT

PDB ID : 1SDD  
Title : Crystal Structure of Bovine Factor Vai  
Authors : Adams, T.E.; Hockin, M.F.; Mann, K.G.; Everse, S.J.  
Deposited on : 2004-02-13  
Resolution : 2.80 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  
<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

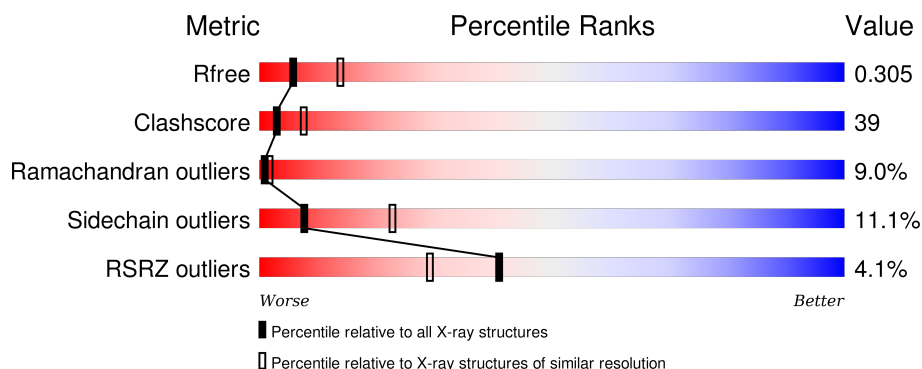
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	306	
2	B	647	

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 7271 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 Coagulation factor V.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	268	Total	C	N	O	S	0	0	0
			2132	1366	356	398	12			

- Molecule 2 is a protein called Coagulation factor V.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	601	Total	C	N	O	S	0	0	0
			4878	3127	840	888	23			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



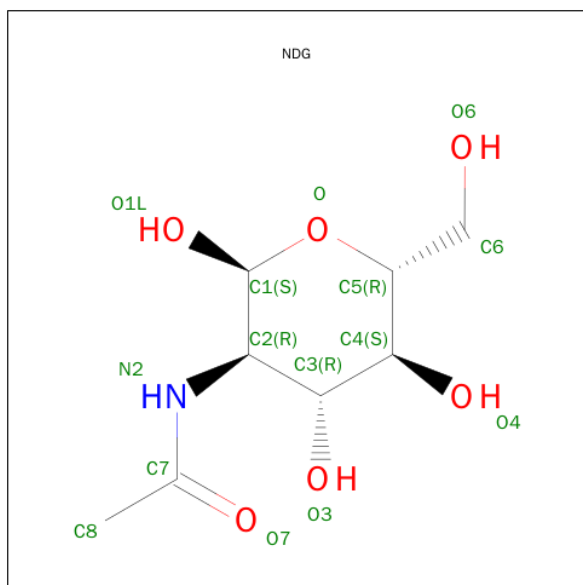
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is SUGAR (2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE) (three-letter code: NDG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Cu	0	0
			1	1		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Ca	0	0
			1	1		

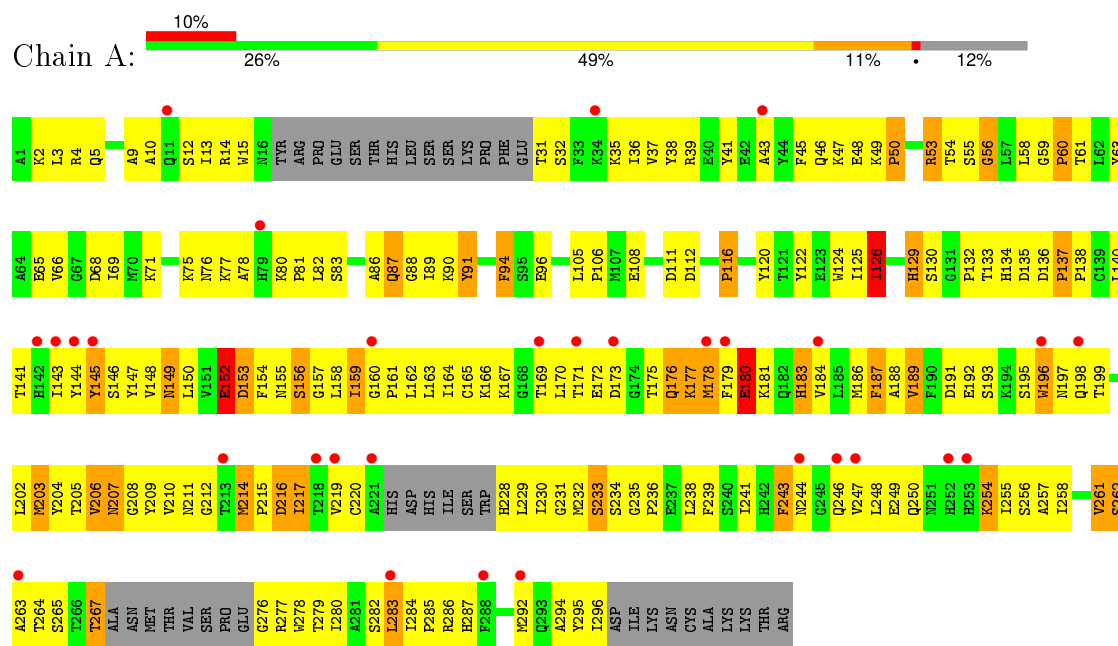
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	47	Total 47	O 47	0	0
7	B	142	Total 142	O 142	0	0

### 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 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 ( $\text{RSRZ} > 2$ ). 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: Coagulation factor V



L2171	F2178	K2088	E2016	P1920	K1837
A2172	G2179	I2089	L2017	W1921	
L2173	G2180	K2090	Q2018	M1926	L1843
R2174	N2181	K2091	G2019	Q1927	D1844
L2175	N2182	I2092		K1928	T1845
			V2022		E1846
				L1931	V1847
			C2025		G1848
			S2026	I1935	E1849
			T2027		I1850
				Q1938	Q1851
			G2030	G1939	R1852
					A1853
			G2034	H1942	G1854
			K2035	Y1943	M1855
			I2036	L1944	Q1856
			E2037	K1945	
			N2038	P1946	I1861
			K2039		
			Q2040	T1950	R1864
				E1951	
			S2044		K1867
			S2045		M1868
			F2046	D1959	P1869
			K2047	R1960	M1870
			K2048		G1871
			S2049	W1963	L1872
			W2050	R1964	
			W2051	I1965	L1876
			G2052	F1966	I1877
			N2053	K1967	
			Y2054	G1968	E1886
			W2055	M1969	F1887
				S1970	M1888
			F2058		G1889
				M1973	Y1890
			L2062	V1974	H1891
			N2063	M1975	E1892
				E1988	P1893
			K2067		K1894
			V2068	I1991	L1895
			N2069	D1992	A1896
			A2070	P1993	R1897
			W2071		L1898
			Q2072		M1899
			A2073	R1998	M1900
			K2074	Y1999	
			A2075	I2000	Y1904
			N2076	R2001	M1905
			N2077	I2002	A1906
			N2078	S2003	
			N2079	P2004	K1911
			Q2080		L1912
			W2081	Y2008	S1913
			L2082	N2009	T1914
			Q2083	K2010	E1915
			E2084	P2011	F1916
			D2085		M1917
			L2086	R2014	P1918
			L2087	L2015	E1919

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.37Å 86.56Å 229.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80 28.63 – 2.81	Depositor EDS
% Data completeness (in resolution range)	90.1 (30.00-2.80) 90.7 (28.63-2.81)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.22 (at 2.80Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.252 , 0.296 0.248 , 0.305	Depositor DCC
$R_{free}$ test set	918 reflections (3.19%)	DCC
Wilson B-factor (Å <sup>2</sup> )	70.0	Xtriage
Anisotropy	0.086	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 73.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 28745 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	7271	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CA, NDG, CU, NAG

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	0.35	0/2188	0.60	0/2960
2	B	0.46	0/5012	0.70	2/6792 (0.0%)
All	All	0.43	0/7200	0.67	2/9752 (0.0%)

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

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	1847	VAL	N-CA-C	-6.63	93.11	111.00
2	B	2014	ARG	N-CA-C	-6.47	93.53	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1890	TYR	Sidechain

### 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	A	2132	0	2074	198	0
2	B	4878	0	4767	357	0
3	A	28	0	26	3	0
3	B	14	0	13	4	0
4	B	28	0	26	6	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
7	A	47	0	0	1	0
7	B	142	0	0	3	0
All	All	7271	0	6906	548	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 39.

The worst 5 of 548 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:1992:ASP:HB3	2:B:1993:PRO:HD3	1.13	1.13
2:B:2152:ASN:HB2	2:B:2153:PRO:HD3	1.22	1.10
2:B:1721:MET:HG3	2:B:1786:ARG:HH22	1.01	1.08
2:B:1914:THR:HB	2:B:1920:PRO:HD2	1.38	1.05
1:A:264:THR:HG21	2:B:1822:PRO:HG3	1.44	0.98

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	260/306 (85%)	178 (68%)	46 (18%)	36 (14%)	<b>0</b> <b>1</b>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	593/647 (92%)	479 (81%)	73 (12%)	41 (7%)	1	3
All	All	853/953 (90%)	657 (77%)	119 (14%)	77 (9%)	1	2

5 of 77 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	50	PRO
1	A	60	PRO
1	A	180	GLU
1	A	187	PHE
1	A	206	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	A	233/269 (87%)	208 (89%)	25 (11%)	8	24
2	B	526/570 (92%)	467 (89%)	59 (11%)	7	22
All	All	759/839 (90%)	675 (89%)	84 (11%)	8	23

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1707	LEU
2	B	1808	LEU
2	B	2140	ASN
2	B	1711	ARG
2	B	1720	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1804	HIS
2	B	1817	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	2083	GLN
2	B	1814	GLN
2	B	1851	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	NAG	A	2185	1	14,14,15	0.64	0	15,19,21	0.75	0
3	NAG	A	2186	1	14,14,15	0.51	0	15,19,21	0.73	1 (6%)
4	NDG	B	2187	2	14,14,15	0.58	0	15,19,21	0.73	0
3	NAG	B	2188	2	14,14,15	0.55	0	15,19,21	0.82	1 (6%)
4	NDG	B	2189	2	14,14,15	0.62	0	15,19,21	0.73	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	NAG	A	2185	1	-	0/6/23/26	0/1/1/1
3	NAG	A	2186	1	-	0/6/23/26	0/1/1/1
4	NDG	B	2187	2	-	0/6/23/26	0/1/1/1
3	NAG	B	2188	2	-	0/6/23/26	0/1/1/1
4	NDG	B	2189	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2188	NAG	C2-N2-C7	-2.41	119.94	123.04
3	A	2186	NAG	C2-N2-C7	-2.16	120.27	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2185	NAG	3	0
4	B	2187	NDG	3	0
3	B	2188	NAG	4	0
4	B	2189	NDG	3	0

## 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	A	268/306 (87%)	0.58	30 (11%) <b>7</b> <b>3</b>	48, 89, 127, 136	0
2	B	601/647 (92%)	-0.19	6 (0%) <b>84</b> <b>77</b>	26, 54, 91, 112	0
All	All	869/953 (91%)	0.05	36 (4%) <b>41</b> <b>29</b>	26, 64, 116, 136	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	218	THR	4.6
1	A	34	LYS	3.8
1	A	247	VAL	3.7
1	A	178	MET	3.7
2	B	1760	ASN	3.7

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NDG	B	2189	14/15	0.90	0.19	-0.19	70,72,74,76	0
5	CU	B	2190	1/1	0.97	0.20	-0.46	65,65,65,65	0
3	NAG	B	2188	14/15	0.89	0.13	-1.25	73,74,76,78	0
3	NAG	A	2186	14/15	0.90	0.16	-1.61	86,88,90,91	0
6	CA	A	2184	1/1	0.95	0.08	-1.93	68,68,68,68	0
4	NDG	B	2187	14/15	0.83	0.34	-	94,98,101,101	0
3	NAG	A	2185	14/15	0.75	0.40	-	127,128,129,130	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.