



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:26 AM GMT

PDB ID : 2D39  
Title : Trivalent Recognition Unit of Innate Immunity System; Crystal Structure of human M-ficolin Fibrinogen-like Domain  
Authors : Tanio, M.; Kondo, S.; Sugio, S.; Kohno, T.  
Deposited on : 2005-09-26  
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  
<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

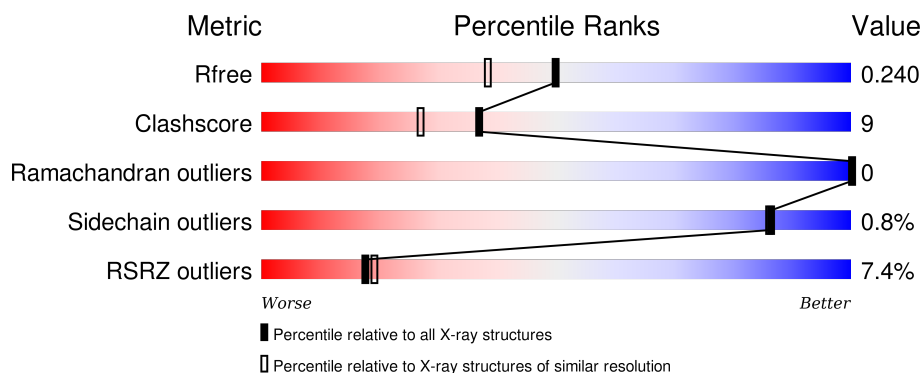
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (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 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	237	<div> <div>6%</div> <div>76%</div> <div>11%</div> <div>12%</div> </div>
1	B	237	<div> <div>6%</div> <div>61%</div> <div>24%</div> <div>16%</div> </div>
1	C	237	<div> <div>7%</div> <div>69%</div> <div>20%</div> <div>11%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5212 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 Ficolin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	209	Total	C	N	O	S	0	0	0
			1656	1049	284	314	9			
1	B	200	Total	C	N	O	S	0	0	0
			1591	1008	273	302	8			
1	C	212	Total	C	N	O	S	0	0	0
			1680	1063	289	319	9			

There are 75 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	113	GLU	-	CLONING ARTIFACT	UNP O00602
A	114	PHE	-	CLONING ARTIFACT	UNP O00602
A	327	PHE	-	CLONING ARTIFACT	UNP O00602
A	328	LEU	-	CLONING ARTIFACT	UNP O00602
A	329	GLU	-	SEE REMARK 999	UNP O00602
A	330	GLN	-	SEE REMARK 999	UNP O00602
A	331	LYS	-	SEE REMARK 999	UNP O00602
A	332	LEU	-	SEE REMARK 999	UNP O00602
A	333	ILE	-	SEE REMARK 999	UNP O00602
A	334	SER	-	SEE REMARK 999	UNP O00602
A	335	GLU	-	SEE REMARK 999	UNP O00602
A	336	GLU	-	SEE REMARK 999	UNP O00602
A	337	ASP	-	SEE REMARK 999	UNP O00602
A	338	LEU	-	SEE REMARK 999	UNP O00602
A	339	ASN	-	CLONING ARTIFACT	UNP O00602
A	340	SER	-	CLONING ARTIFACT	UNP O00602
A	341	ALA	-	CLONING ARTIFACT	UNP O00602
A	342	VAL	-	CLONING ARTIFACT	UNP O00602
A	343	ASP	-	CLONING ARTIFACT	UNP O00602
A	344	HIS	-	EXPRESSION TAG	UNP O00602
A	345	HIS	-	EXPRESSION TAG	UNP O00602
A	346	HIS	-	EXPRESSION TAG	UNP O00602
A	347	HIS	-	EXPRESSION TAG	UNP O00602

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	348	HIS	-	EXPRESSION TAG	UNP O00602
A	349	HIS	-	EXPRESSION TAG	UNP O00602
B	113	GLU	-	CLONING ARTIFACT	UNP O00602
B	114	PHE	-	CLONING ARTIFACT	UNP O00602
B	327	PHE	-	CLONING ARTIFACT	UNP O00602
B	328	LEU	-	CLONING ARTIFACT	UNP O00602
B	329	GLU	-	SEE REMARK 999	UNP O00602
B	330	GLN	-	SEE REMARK 999	UNP O00602
B	331	LYS	-	SEE REMARK 999	UNP O00602
B	332	LEU	-	SEE REMARK 999	UNP O00602
B	333	ILE	-	SEE REMARK 999	UNP O00602
B	334	SER	-	SEE REMARK 999	UNP O00602
B	335	GLU	-	SEE REMARK 999	UNP O00602
B	336	GLU	-	SEE REMARK 999	UNP O00602
B	337	ASP	-	SEE REMARK 999	UNP O00602
B	338	LEU	-	SEE REMARK 999	UNP O00602
B	339	ASN	-	CLONING ARTIFACT	UNP O00602
B	340	SER	-	CLONING ARTIFACT	UNP O00602
B	341	ALA	-	CLONING ARTIFACT	UNP O00602
B	342	VAL	-	CLONING ARTIFACT	UNP O00602
B	343	ASP	-	CLONING ARTIFACT	UNP O00602
B	344	HIS	-	EXPRESSION TAG	UNP O00602
B	345	HIS	-	EXPRESSION TAG	UNP O00602
B	346	HIS	-	EXPRESSION TAG	UNP O00602
B	347	HIS	-	EXPRESSION TAG	UNP O00602
B	348	HIS	-	EXPRESSION TAG	UNP O00602
B	349	HIS	-	EXPRESSION TAG	UNP O00602
C	113	GLU	-	CLONING ARTIFACT	UNP O00602
C	114	PHE	-	CLONING ARTIFACT	UNP O00602
C	327	PHE	-	CLONING ARTIFACT	UNP O00602
C	328	LEU	-	CLONING ARTIFACT	UNP O00602
C	329	GLU	-	SEE REMARK 999	UNP O00602
C	330	GLN	-	SEE REMARK 999	UNP O00602
C	331	LYS	-	SEE REMARK 999	UNP O00602
C	332	LEU	-	SEE REMARK 999	UNP O00602
C	333	ILE	-	SEE REMARK 999	UNP O00602
C	334	SER	-	SEE REMARK 999	UNP O00602
C	335	GLU	-	SEE REMARK 999	UNP O00602
C	336	GLU	-	SEE REMARK 999	UNP O00602
C	337	ASP	-	SEE REMARK 999	UNP O00602
C	338	LEU	-	SEE REMARK 999	UNP O00602
C	339	ASN	-	CLONING ARTIFACT	UNP O00602

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	340	SER	-	CLONING ARTIFACT	UNP O00602
C	341	ALA	-	CLONING ARTIFACT	UNP O00602
C	342	VAL	-	CLONING ARTIFACT	UNP O00602
C	343	ASP	-	CLONING ARTIFACT	UNP O00602
C	344	HIS	-	EXPRESSION TAG	UNP O00602
C	345	HIS	-	EXPRESSION TAG	UNP O00602
C	346	HIS	-	EXPRESSION TAG	UNP O00602
C	347	HIS	-	EXPRESSION TAG	UNP O00602
C	348	HIS	-	EXPRESSION TAG	UNP O00602
C	349	HIS	-	EXPRESSION TAG	UNP O00602

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	C	1	Total Ca 1 1	0	0

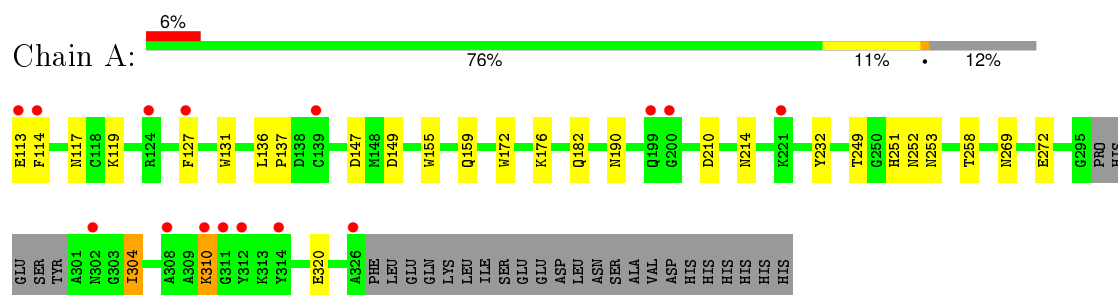
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	123	Total O 123 123	0	0
3	B	75	Total O 75 75	0	0
3	C	85	Total O 85 85	0	0

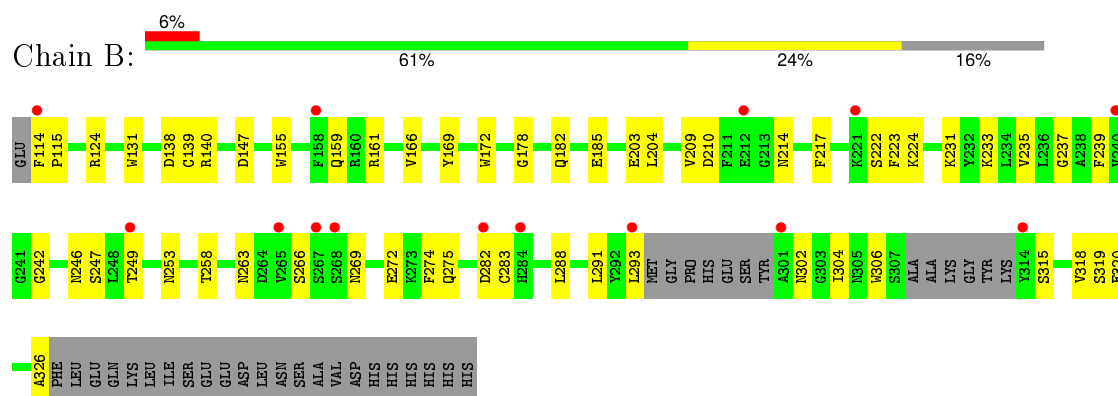
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

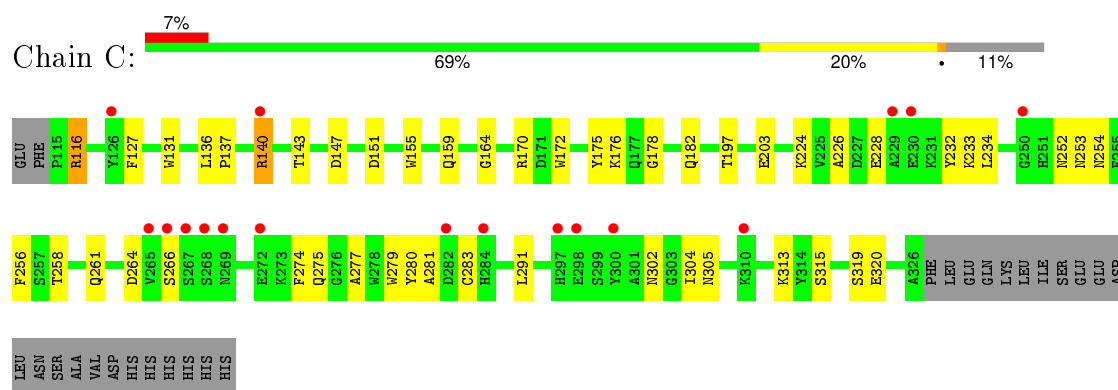
#### • Molecule 1: Ficolin-1



#### • Molecule 1: Ficolin-1



#### • Molecule 1: Ficolin-1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.16Å 117.45Å 55.19Å 90.00° 99.88° 90.00°	Depositor
Resolution (Å)	30.00 – 1.90 29.36 – 1.90	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-1.90) 98.8 (29.36-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.86 (at 1.91Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.209 , 0.240 0.209 , 0.240	Depositor DCC
$R_{free}$ test set	2717 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	21.1	Xtriage
Anisotropy	0.105	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.7	EDS
Estimated twinning fraction	0.022 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 53715 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5212	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.65% 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

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	0.36	0/1701	0.63	0/2297
1	B	0.35	0/1634	0.60	0/2208
1	C	0.34	0/1728	0.64	1/2336 (0.0%)
All	All	0.35	0/5063	0.62	1/6841 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	151	ASP	CB-CG-OD1	5.85	123.57	118.30

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	A	1656	0	1532	23	0
1	B	1591	0	1465	36	0
1	C	1680	0	1553	33	0
2	A	1	0	0	0	0
2	C	1	0	0	0	0
3	A	123	0	0	0	0
3	B	75	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	85	0	0	1	0
All	All	5212	0	4550	89	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 9.

The worst 5 of 89 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:C:136:LEU:HB3	1:C:137:PRO:HD2	1.64	0.78
1:B:239:PHE:HB2	1:B:246:ASN:ND2	1.98	0.77
1:C:254:ASN:HD21	1:C:281:ALA:HB2	1.50	0.75
1:B:239:PHE:HB2	1:B:246:ASN:HD22	1.51	0.74
1:C:176:LYS:HE2	1:C:228:GLU:OE2	1.93	0.69

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	205/237 (86%)	198 (97%)	7 (3%)	0	100	100
1	B	194/237 (82%)	185 (95%)	9 (5%)	0	100	100
1	C	210/237 (89%)	197 (94%)	13 (6%)	0	100	100
All	All	609/711 (86%)	580 (95%)	29 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 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	170/197 (86%)	168 (99%)	2 (1%)	78	76
1	B	165/197 (84%)	165 (100%)	0	100	100
1	C	173/197 (88%)	171 (99%)	2 (1%)	78	76
All	All	508/591 (86%)	504 (99%)	4 (1%)	86	86

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

Mol	Chain	Res	Type
1	A	304	ILE
1	A	310	LYS
1	C	116	ARG
1	C	140	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	275	GLN
1	C	302	ASN
1	C	254	ASN
1	B	261	GLN
1	C	251	HIS

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 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	A	209/237 (88%)	0.24	15 (7%) 18 20	12, 21, 42, 53	0
1	B	200/237 (84%)	0.41	14 (7%) 19 21	12, 26, 42, 49	0
1	C	212/237 (89%)	0.50	17 (8%) 15 17	12, 26, 48, 54	0
All	All	621/711 (87%)	0.38	46 (7%) 17 19	12, 24, 46, 54	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	114	PHE	7.0
1	B	267	SER	7.0
1	C	267	SER	5.7
1	A	114	PHE	4.3
1	A	314	TYR	4.3

### 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
2	CA	C	2401	1/1	0.97	0.08	-1.16	22,22,22,22	1
2	CA	A	1401	1/1	0.99	0.06	-2.24	20,20,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.