



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 12:54 am GMT

PDB ID : 2B0O  
Title : Crystal structure of UPLC1 GAP domain  
Authors : Ismail, S.; Shen, L.; Arrowsmith, C.; Edwards, A.; Sundstrom, M.; Weigelt, J.; Bochkarev, A.; Park, H.; Structural Genomics Consortium (SGC)  
Deposited on : 2005-09-14  
Resolution : 2.06 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  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

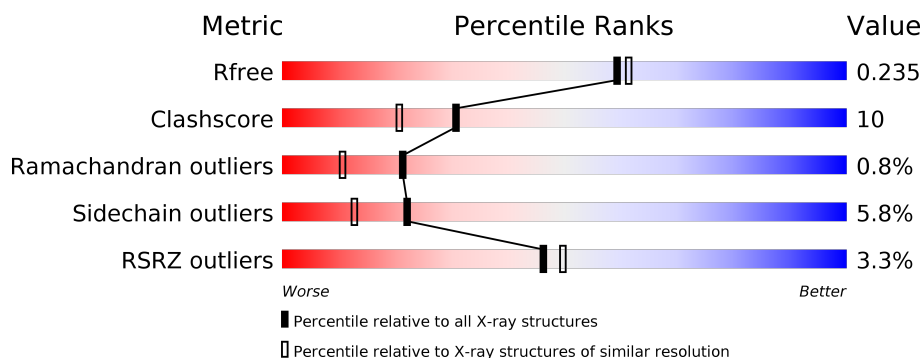
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.06 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))
$R_{free}$	111664	2250 (2.08-2.04)
Clashscore	122126	2358 (2.08-2.04)
Ramachandran outliers	120053	2339 (2.08-2.04)
Sidechain outliers	120020	2339 (2.08-2.04)
RSRZ outliers	108989	2211 (2.08-2.04)

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	E	301	<div> <div>5%</div> <div> <div></div> <div>71%</div> <div>11%</div> <div>5%</div> <div>12%</div> </div> </div>
1	F	301	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>13%</div> <div>•</div> <div>12%</div> </div> </div>
1	G	301	<div> <div>2%</div> <div> <div></div> <div>67%</div> <div>14%</div> <div>••</div> <div>16%</div> </div> </div>
1	H	301	<div> <div>4%</div> <div> <div></div> <div>67%</div> <div>17%</div> <div>••</div> <div>13%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8492 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 UPLC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	264	Total	C	N	O	S	0	0	0
			2011	1257	359	381	14			
1	F	266	Total	C	N	O	S	0	0	0
			2037	1278	361	384	14			
1	G	253	Total	C	N	O	S	0	0	0
			1929	1203	347	366	13			
1	H	262	Total	C	N	O	S	0	0	0
			2011	1264	357	377	13			

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	397	MET	-	CLONING ARTIFACT	UNP Q6P9F4
E	398	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
E	399	SER	-	CLONING ARTIFACT	UNP Q6P9F4
E	400	SER	-	CLONING ARTIFACT	UNP Q6P9F4
E	401	HIS	-	EXPRESSION TAG	UNP Q6P9F4
E	402	HIS	-	EXPRESSION TAG	UNP Q6P9F4
E	403	HIS	-	EXPRESSION TAG	UNP Q6P9F4
E	404	HIS	-	EXPRESSION TAG	UNP Q6P9F4
E	405	HIS	-	EXPRESSION TAG	UNP Q6P9F4
E	406	HIS	-	EXPRESSION TAG	UNP Q6P9F4
E	407	SER	-	CLONING ARTIFACT	UNP Q6P9F4
E	408	SER	-	CLONING ARTIFACT	UNP Q6P9F4
E	409	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
E	410	LEU	-	CLONING ARTIFACT	UNP Q6P9F4
E	411	VAL	-	CLONING ARTIFACT	UNP Q6P9F4
E	412	PRO	-	CLONING ARTIFACT	UNP Q6P9F4
E	413	ARG	-	CLONING ARTIFACT	UNP Q6P9F4
E	414	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
E	415	SER	-	CLONING ARTIFACT	UNP Q6P9F4
F	397	MET	-	CLONING ARTIFACT	UNP Q6P9F4
F	398	GLY	-	CLONING ARTIFACT	UNP Q6P9F4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	399	SER	-	CLONING ARTIFACT	UNP Q6P9F4
F	400	SER	-	CLONING ARTIFACT	UNP Q6P9F4
F	401	HIS	-	EXPRESSION TAG	UNP Q6P9F4
F	402	HIS	-	EXPRESSION TAG	UNP Q6P9F4
F	403	HIS	-	EXPRESSION TAG	UNP Q6P9F4
F	404	HIS	-	EXPRESSION TAG	UNP Q6P9F4
F	405	HIS	-	EXPRESSION TAG	UNP Q6P9F4
F	406	HIS	-	EXPRESSION TAG	UNP Q6P9F4
F	407	SER	-	CLONING ARTIFACT	UNP Q6P9F4
F	408	SER	-	CLONING ARTIFACT	UNP Q6P9F4
F	409	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
F	410	LEU	-	CLONING ARTIFACT	UNP Q6P9F4
F	411	VAL	-	CLONING ARTIFACT	UNP Q6P9F4
F	412	PRO	-	CLONING ARTIFACT	UNP Q6P9F4
F	413	ARG	-	CLONING ARTIFACT	UNP Q6P9F4
F	414	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
F	415	SER	-	CLONING ARTIFACT	UNP Q6P9F4
G	397	MET	-	CLONING ARTIFACT	UNP Q6P9F4
G	398	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
G	399	SER	-	CLONING ARTIFACT	UNP Q6P9F4
G	400	SER	-	CLONING ARTIFACT	UNP Q6P9F4
G	401	HIS	-	EXPRESSION TAG	UNP Q6P9F4
G	402	HIS	-	EXPRESSION TAG	UNP Q6P9F4
G	403	HIS	-	EXPRESSION TAG	UNP Q6P9F4
G	404	HIS	-	EXPRESSION TAG	UNP Q6P9F4
G	405	HIS	-	EXPRESSION TAG	UNP Q6P9F4
G	406	HIS	-	EXPRESSION TAG	UNP Q6P9F4
G	407	SER	-	CLONING ARTIFACT	UNP Q6P9F4
G	408	SER	-	CLONING ARTIFACT	UNP Q6P9F4
G	409	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
G	410	LEU	-	CLONING ARTIFACT	UNP Q6P9F4
G	411	VAL	-	CLONING ARTIFACT	UNP Q6P9F4
G	412	PRO	-	CLONING ARTIFACT	UNP Q6P9F4
G	413	ARG	-	CLONING ARTIFACT	UNP Q6P9F4
G	414	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
G	415	SER	-	CLONING ARTIFACT	UNP Q6P9F4
H	397	MET	-	CLONING ARTIFACT	UNP Q6P9F4
H	398	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
H	399	SER	-	CLONING ARTIFACT	UNP Q6P9F4
H	400	SER	-	CLONING ARTIFACT	UNP Q6P9F4
H	401	HIS	-	EXPRESSION TAG	UNP Q6P9F4
H	402	HIS	-	EXPRESSION TAG	UNP Q6P9F4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	403	HIS	-	EXPRESSION TAG	UNP Q6P9F4
H	404	HIS	-	EXPRESSION TAG	UNP Q6P9F4
H	405	HIS	-	EXPRESSION TAG	UNP Q6P9F4
H	406	HIS	-	EXPRESSION TAG	UNP Q6P9F4
H	407	SER	-	CLONING ARTIFACT	UNP Q6P9F4
H	408	SER	-	CLONING ARTIFACT	UNP Q6P9F4
H	409	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
H	410	LEU	-	CLONING ARTIFACT	UNP Q6P9F4
H	411	VAL	-	CLONING ARTIFACT	UNP Q6P9F4
H	412	PRO	-	CLONING ARTIFACT	UNP Q6P9F4
H	413	ARG	-	CLONING ARTIFACT	UNP Q6P9F4
H	414	GLY	-	CLONING ARTIFACT	UNP Q6P9F4
H	415	SER	-	CLONING ARTIFACT	UNP Q6P9F4

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	H	1	Total 1	Zn 1	0	0
2	G	1	Total 1	Zn 1	0	0
2	F	1	Total 1	Zn 1	0	0
2	E	1	Total 1	Zn 1	0	0

- Molecule 3 is water.

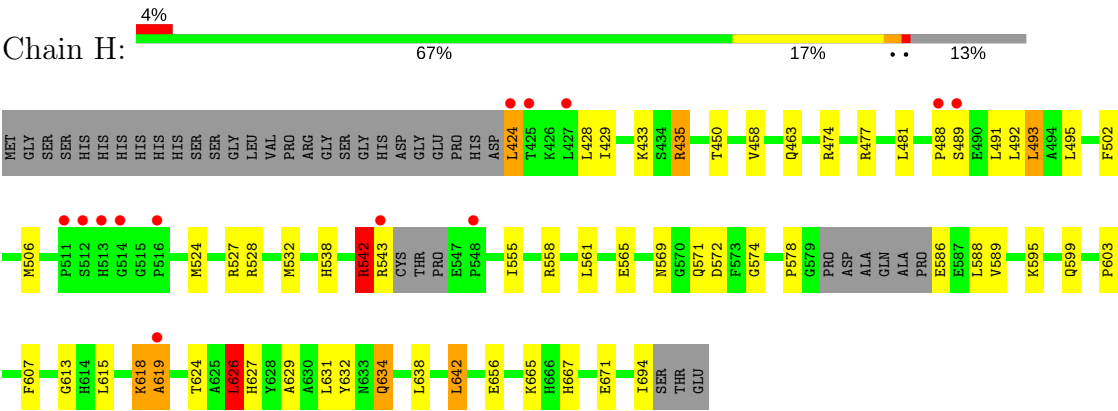
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	147	Total 147	O 147	0	0
3	F	129	Total 129	O 129	0	0
3	G	130	Total 130	O 130	2	0
3	H	94	Total 94	O 94	0	0



- Molecule 1: UPLC1



● Molecule 1: UPLC1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.05Å 106.72Å 100.36Å 90.00° 98.97° 90.00°	Depositor
Resolution (Å)	50.00 – 2.06 44.22 – 2.06	Depositor EDS
% Data completeness (in resolution range)	97.6 (50.00-2.06) 97.5 (44.22-2.06)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 2.06Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.183 , 0.233 0.183 , 0.235	Depositor DCC
$R_{free}$ test set	3414 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.6	Xtriage
Anisotropy	0.307	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 58.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8492	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.41% 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: ZN

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	E	0.92	1/2048 (0.0%)	0.95	8/2776 (0.3%)
1	F	0.92	1/2074 (0.0%)	0.90	3/2809 (0.1%)
1	G	0.86	1/1962 (0.1%)	0.92	8/2654 (0.3%)
1	H	0.78	0/2048	0.82	3/2774 (0.1%)
All	All	0.87	3/8132 (0.0%)	0.90	22/11013 (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	E	0	3
1	F	0	2
1	G	0	1
1	H	0	2
All	All	0	8

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	613	GLY	N-CA	6.31	1.55	1.46
1	F	469	ARG	CZ-NH1	5.89	1.40	1.33
1	E	619	ALA	N-CA	5.41	1.57	1.46

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	469	ARG	NE-CZ-NH2	-13.76	113.42	120.30
1	G	528	ARG	NE-CZ-NH2	-11.34	114.63	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	469	ARG	NE-CZ-NH1	9.51	125.05	120.30
1	F	469	ARG	NE-CZ-NH2	-8.91	115.84	120.30
1	G	458	VAL	CG1-CB-CG2	7.40	122.74	110.90

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	584	ALA	Peptide
1	E	585	PRO	Peptide
1	E	683	ALA	Peptide
1	F	613	GLY	Peptide
1	F	681	THR	Peptide

## 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	E	2011	0	1992	49	0
1	F	2037	0	2021	37	0
1	G	1929	0	1918	45	0
1	H	2011	0	1996	36	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	E	147	0	0	7	0
3	F	129	0	0	5	0
3	G	130	0	0	5	0
3	H	94	0	0	5	0
All	All	8492	0	7927	158	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 10.

The worst 5 of 158 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:H:618:LYS:O	1:H:619:ALA:CB	1.90	1.14
1:F:618:LYS:O	1:F:619:ALA:HB3	1.40	1.11
1:E:618:LYS:O	1:E:619:ALA:HB3	1.37	1.10
1:E:618:LYS:O	1:E:619:ALA:CB	1.93	1.08
1:G:618:LYS:O	1:G:619:ALA:HB3	1.48	1.08

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	E	260/301 (86%)	250 (96%)	5 (2%)	5 (2%)	9	2
1	F	260/301 (86%)	253 (97%)	7 (3%)	0	100	100
1	G	247/301 (82%)	242 (98%)	5 (2%)	0	100	100
1	H	256/301 (85%)	240 (94%)	13 (5%)	3 (1%)	14	5
All	All	1023/1204 (85%)	985 (96%)	30 (3%)	8 (1%)	21	11

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	582	ALA
1	E	584	ALA
1	H	542	ARG
1	E	581	ASP
1	E	583	GLN

### 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	E	215/247 (87%)	200 (93%)	15 (7%)	16	8
1	F	219/247 (89%)	212 (97%)	7 (3%)	42	35
1	G	207/247 (84%)	194 (94%)	13 (6%)	20	11
1	H	215/247 (87%)	200 (93%)	15 (7%)	16	8
All	All	856/988 (87%)	806 (94%)	50 (6%)	22	13

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	458	VAL
1	G	558	ARG
1	H	634	GLN
1	G	481	LEU
1	G	543	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	634	GLN
1	G	538	HIS
1	H	571	GLN
1	F	633	ASN
1	H	627	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 4 ligands modelled in this entry, 4 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

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	E	264/301 (87%)	0.01	14 (5%) 26 27	10, 19, 43, 57	0
1	F	266/301 (88%)	-0.17	3 (1%) 80 82	11, 19, 39, 50	0
1	G	253/301 (84%)	-0.04	5 (1%) 65 67	11, 20, 40, 56	0
1	H	262/301 (87%)	0.07	13 (4%) 29 30	14, 25, 49, 63	0
All	All	1045/1204 (86%)	-0.03	35 (3%) 46 50	10, 21, 44, 63	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	512	SER	8.4
1	H	513	HIS	6.2
1	E	582	ALA	4.8
1	G	681	THR	4.7
1	H	514	GLY	4.6

### 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. 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	ZN	F	698	1/1	0.99	0.08	16,16,16,16	0
2	ZN	G	698	1/1	1.00	0.08	16,16,16,16	0
2	ZN	E	698	1/1	1.00	0.11	17,17,17,17	0
2	ZN	H	698	1/1	1.00	0.06	22,22,22,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.