Full wwPDB X-ray Structure Validation Report ⓘ

Jan 29, 2024 – 12:06 am GMT

PDB ID	:	7QNP
Title	:	Designed Armadillo repeat protein N(A4)M4C(AII) co-crystallized with hen egg white lysozyme
Authors	:	Michel, E.; Mittl, P.R.E.; Plueckthun, A.
Deposited on	:	2021-12-21
Resolution	:	1.59 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *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:

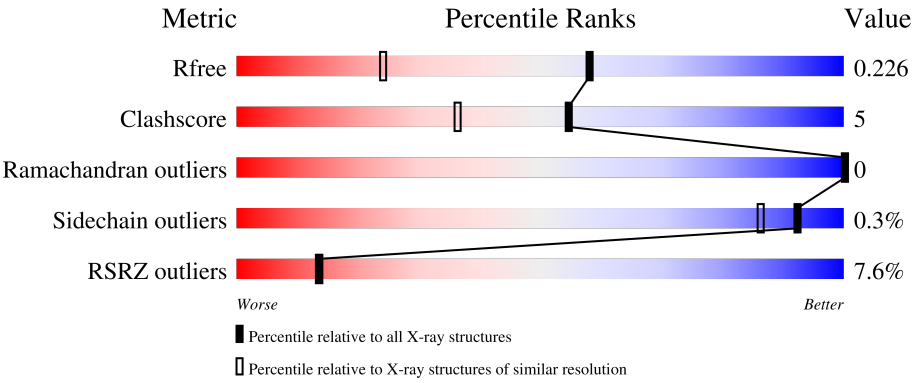
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.59 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	5534 (1.60-1.56)
Clashscore	141614	5861 (1.60-1.56)
Ramachandran outliers	138981	5708 (1.60-1.56)
Sidechain outliers	138945	5703 (1.60-1.56)
RSRZ outliers	127900	5431 (1.60-1.56)

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 >=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 <=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	240	10% 90% 9%
2	BBB	129	3% 84% 14%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3261 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.

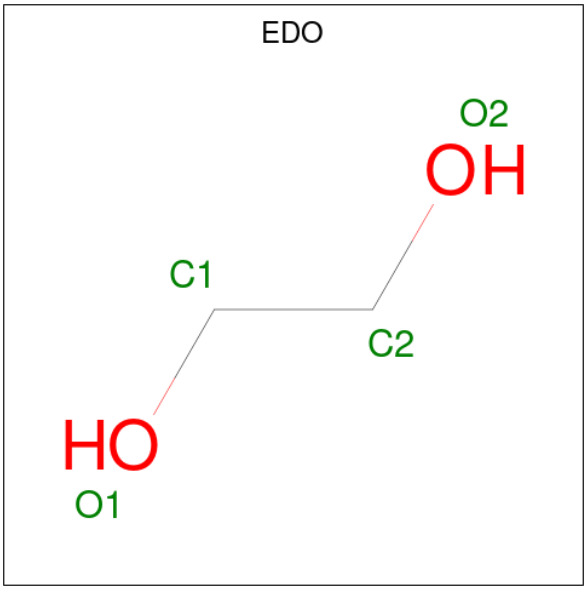
- Molecule 1 is a protein called Designed Armadillo Repeat Protein N(A4)M4C(AII).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	AAA	240	1879	1187	316	376	0	10	0

- Molecule 2 is a protein called Lysozyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	BBB	129	1043	638	203	192	10	0	5	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



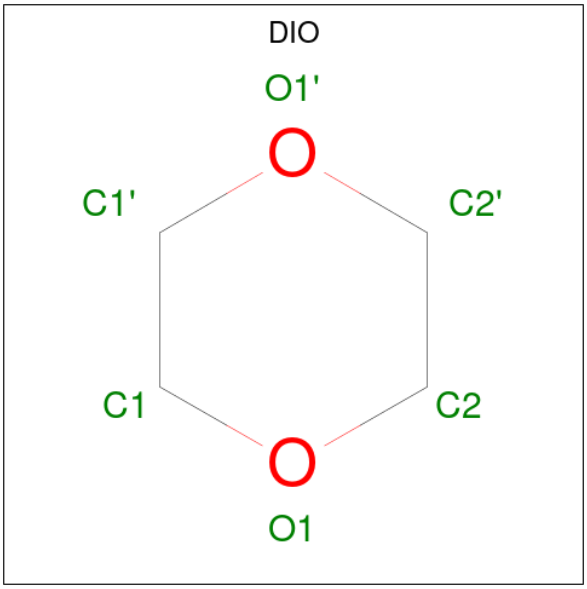
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	AAA	1	4	2	2	0	0
3	AAA	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

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

- Molecule 4 is 1,4-DIETHYLENE DIOXIDE (three-letter code: DIO) (formula: $C_4H_8O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	BBB	1	Total C O 6 4 2	0	0

- Molecule 5 is water.

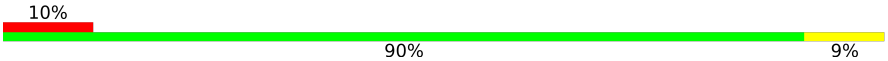
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	140	Total O 140 140	0	0
5	BBB	165	Total O 165 165	0	0

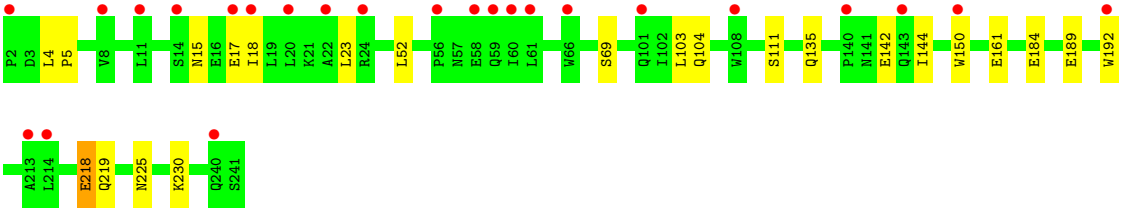
3 Residue-property plots (i)

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: Designed Armadillo Repeat Protein N(A4)M4C(AII)

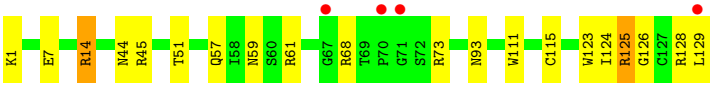
Chain AAA:





• Molecule 2: Lysozyme

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.59Å 62.66Å 108.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.00 – 1.59 42.00 – 1.59	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.00-1.59) 99.9 (42.00-1.59)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 1.58Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.186 , 0.218 0.197 , 0.226	Depositor DCC
R_{free} test set	2654 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	29.2	Xtriage
Anisotropy	0.132	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 45.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3261	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.71% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: DIO, EDO

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	0.99	2/1905 (0.1%)	0.96	0/2595
2	BBB	1.02	1/1063 (0.1%)	1.22	5/1434 (0.3%)
All	All	1.00	3/2968 (0.1%)	1.06	5/4029 (0.1%)

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	BBB	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	BBB	126	GLY	N-CA	8.29	1.58	1.46
1	AAA	69	SER	CB-OG	-7.33	1.32	1.42
1	AAA	218	GLU	CD-OE2	-5.28	1.19	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	BBB	125[A]	ARG	NE-CZ-NH2	-8.53	116.03	120.30
2	BBB	125[B]	ARG	NE-CZ-NH2	-8.53	116.03	120.30
2	BBB	14	ARG	NE-CZ-NH1	7.82	124.21	120.30
2	BBB	14	ARG	NE-CZ-NH2	-7.25	116.68	120.30
2	BBB	126	GLY	N-CA-C	-5.27	99.93	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	BBB	125[A]	ARG	Mainchain
2	BBB	125[B]	ARG	Mainchain

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	1879	0	1891	18	0
2	BBB	1043	0	1002	14	0
3	AAA	20	0	30	1	0
3	BBB	8	0	12	0	0
4	BBB	6	0	8	1	0
5	AAA	140	0	0	3	3
5	BBB	165	0	0	2	3
All	All	3261	0	2943	31	3

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 5.

All (31) 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:AAA:111:SER:OG	1:AAA:150[A]:TRP:CE3	2.16	0.98
2:BBB:93:ASN:OD1	5:BBB:301:HOH:O	1.88	0.89
1:AAA:135[A]:GLN:OE1	3:AAA:301:EDO:O1	1.94	0.84
1:AAA:218:GLU:OE2	5:AAA:401:HOH:O	2.01	0.78
1:AAA:161:GLU:OE1	2:BBB:14:ARG:HD3	1.88	0.73
2:BBB:1:LYS:NZ	2:BBB:7:GLU:OE2	2.24	0.70
1:AAA:104:GLN:HG3	1:AAA:144[A]:ILE:HD11	1.75	0.67
2:BBB:44[B]:ASN:HD22	2:BBB:57:GLN:NE2	1.95	0.64
1:AAA:189:GLU:O	1:AAA:192[B]:TRP:HD1	1.84	0.61
1:AAA:189:GLU:HA	1:AAA:192[B]:TRP:CD1	2.41	0.56
1:AAA:103[A]:LEU:HG	1:AAA:144[A]:ILE:HD12	1.89	0.55
1:AAA:23:LEU:HD11	1:AAA:52:LEU:HD13	1.89	0.54
1:AAA:104:GLN:HG3	1:AAA:144[A]:ILE:CD1	2.40	0.52
2:BBB:128:ARG:O	2:BBB:129:LEU:HD23	2.13	0.49
1:AAA:17:GLU:O	1:AAA:17:GLU:OE2	2.31	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:230:LYS:NZ	5:AAA:403:HOH:O	2.41	0.48
2:BBB:124:ILE:HD12	2:BBB:124:ILE:C	2.34	0.47
1:AAA:184:GLU:OE2	1:AAA:225:ASN:ND2	2.47	0.47
2:BBB:123:TRP:CZ3	4:BBB:202:DIO:H1'2	2.49	0.47
1:AAA:189:GLU:O	1:AAA:192[B]:TRP:CD1	2.66	0.46
2:BBB:45:ARG:CZ	2:BBB:68:ARG:HH12	2.29	0.46
1:AAA:111:SER:OG	1:AAA:150[A]:TRP:CZ3	2.59	0.45
1:AAA:4:LEU:N	1:AAA:5:PRO:CD	2.81	0.43
1:AAA:219:GLN:HA	5:AAA:418:HOH:O	2.19	0.42
2:BBB:45:ARG:HD2	2:BBB:51:THR:OG1	2.20	0.42
2:BBB:73:ARG:NH2	5:BBB:305:HOH:O	2.51	0.42
2:BBB:59:ASN:HD21	2:BBB:61:ARG:HB3	1.85	0.41
2:BBB:111:TRP:CD1	2:BBB:115:CYS:HB2	2.56	0.41
1:AAA:15:ASN:HB3	1:AAA:18:ILE:HB	2.04	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:AAA:464:HOH:O	5:BBB:461:HOH:O[3_745]	1.97	0.23
5:AAA:476:HOH:O	5:BBB:450:HOH:O[3_745]	2.02	0.18
5:AAA:523:HOH:O	5:BBB:378:HOH:O[3_645]	2.15	0.05

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	248/240 (103%)	248 (100%)	0	0	100	100
2	BBB	132/129 (102%)	128 (97%)	4 (3%)	0	100	100
All	All	380/369 (103%)	376 (99%)	4 (1%)	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	AAA	198/188 (105%)	197 (100%)	1 (0%)	88 80
2	BBB	110/105 (105%)	110 (100%)	0	100 100
All	All	308/293 (105%)	307 (100%)	1 (0%)	92 86

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	142	GLU

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

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 monosaccharides in this entry.

5.6 Ligand geometry

8 ligands are modelled in this entry.

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	EDO	BBB	203	-	3,3,3	0.24	0	2,2,2	1.02	0
3	EDO	AAA	302	-	3,3,3	0.06	0	2,2,2	0.29	0
3	EDO	AAA	304	-	3,3,3	0.36	0	2,2,2	0.19	0
4	DIO	BBB	202	-	6,6,6	0.54	0	6,6,6	0.64	0
3	EDO	BBB	201	-	3,3,3	0.96	0	2,2,2	0.32	0
3	EDO	AAA	301	-	3,3,3	1.05	0	2,2,2	0.63	0
3	EDO	AAA	303	-	3,3,3	0.17	0	2,2,2	0.11	0
3	EDO	AAA	305	-	3,3,3	0.30	0	2,2,2	0.50	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	EDO	BBB	203	-	-	0/1/1/1	-
3	EDO	AAA	302	-	-	0/1/1/1	-
3	EDO	AAA	304	-	-	1/1/1/1	-
4	DIO	BBB	202	-	-	-	0/1/1/1
3	EDO	BBB	201	-	-	0/1/1/1	-
3	EDO	AAA	301	-	-	0/1/1/1	-
3	EDO	AAA	303	-	-	0/1/1/1	-
3	EDO	AAA	305	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	304	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	BBB	202	DIO	1	0
3	AAA	301	EDO	1	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

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^{th} 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^2)	Q<0.9
1	AAA	240/240 (100%)	0.40	24 (10%) 7 7	19, 35, 57, 71	0
2	BBB	129/129 (100%)	0.16	4 (3%) 49 50	18, 24, 47, 72	0
All	All	369/369 (100%)	0.32	28 (7%) 13 13	18, 31, 56, 72	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	14	SER	6.8
2	BBB	129	LEU	4.6
1	AAA	192[A]	TRP	4.3
1	AAA	56	PRO	4.3
1	AAA	150[A]	TRP	4.3
1	AAA	20	LEU	4.3
1	AAA	24	ARG	3.6
1	AAA	213	ALA	3.4
1	AAA	240	GLN	3.3
1	AAA	140	PRO	3.3
1	AAA	108[A]	TRP	3.2
1	AAA	101	GLN	3.2
1	AAA	59	GLN	2.9
1	AAA	61	LEU	2.8
2	BBB	67	GLY	2.8
1	AAA	11	LEU	2.7
1	AAA	2	PRO	2.7
1	AAA	18	ILE	2.6
1	AAA	22	ALA	2.5
1	AAA	8	VAL	2.5
1	AAA	66[A]	TRP	2.3
1	AAA	214	LEU	2.3
1	AAA	58	GLU	2.2
1	AAA	143	GLN	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	BBB	71	GLY	2.1
1	AAA	17	GLU	2.1
1	AAA	60	ILE	2.0
2	BBB	70	PRO	2.0

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^{th} 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^2)	Q<0.9
4	DIO	BBB	202	6/6	0.78	0.16	39,42,45,47	0
3	EDO	AAA	303	4/4	0.79	0.19	73,78,80,87	0
3	EDO	AAA	305	4/4	0.83	0.13	44,45,52,59	0
3	EDO	AAA	301	4/4	0.83	0.13	39,40,42,50	0
3	EDO	BBB	201	4/4	0.88	0.15	30,33,36,41	0
3	EDO	AAA	302	4/4	0.88	0.11	60,62,67,77	0
3	EDO	AAA	304	4/4	0.89	0.10	55,55,55,61	0
3	EDO	BBB	203	4/4	0.95	0.07	34,36,36,37	0

6.5 Other polymers (i)

There are no such residues in this entry.