Full wwPDB X-ray Structure Validation Report (i)

Jan 20, 2024 – 07:36 pm GMT

PDB ID	:	7OAP
Title	:	Nanobody H3 AND C1 bound to RBD
Authors	:	Naismith, J.H.; Mikolajek, H.
Deposited on	:	2021-04-19
Resolution	:	1.90 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 (i) 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 (1)) were used in the production of this report:

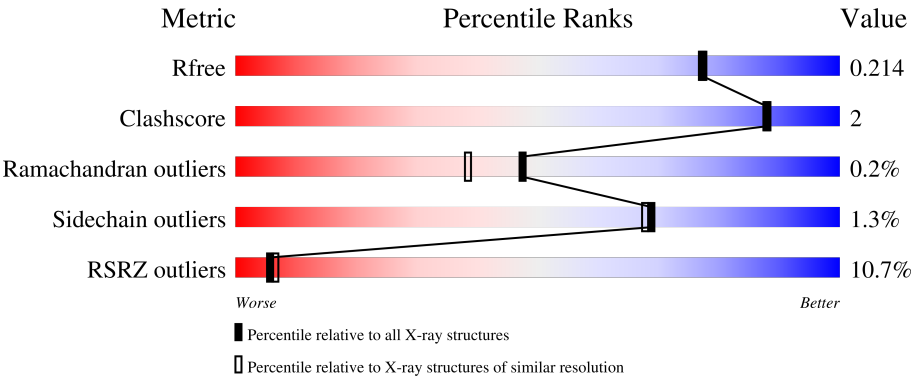
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.90 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 >=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	131	32% 89% 6% 5%
2	EEE	209	90% · 6%
3	FFF	136	4% 88% · · 7%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 3797 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 C1 nanobody.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	124	Total	C	N	O	S	0	2	0
			973	612	167	190	4			

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	EEE	197	Total	C	N	O	S	0	1	0
			1563	1002	261	292	8			

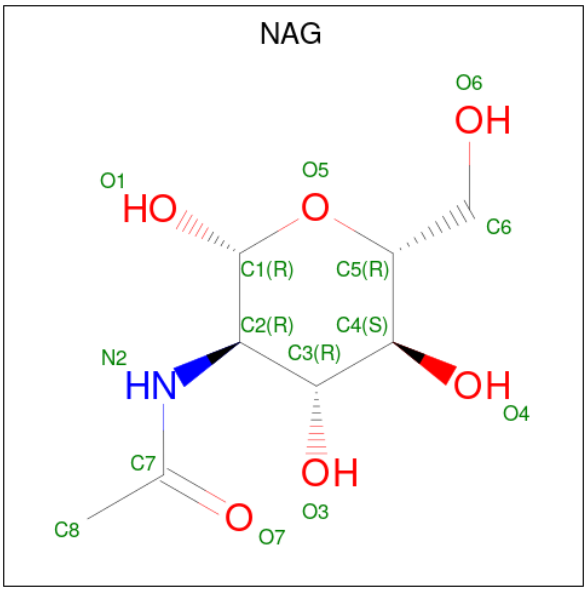
There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	533	LYS	-	expression tag	UNP P0DTC2
EEE	534	HIS	-	expression tag	UNP P0DTC2
EEE	535	HIS	-	expression tag	UNP P0DTC2
EEE	536	HIS	-	expression tag	UNP P0DTC2
EEE	537	HIS	-	expression tag	UNP P0DTC2
EEE	538	HIS	-	expression tag	UNP P0DTC2
EEE	539	HIS	-	expression tag	UNP P0DTC2

- Molecule 3 is a protein called H3 nanobody.

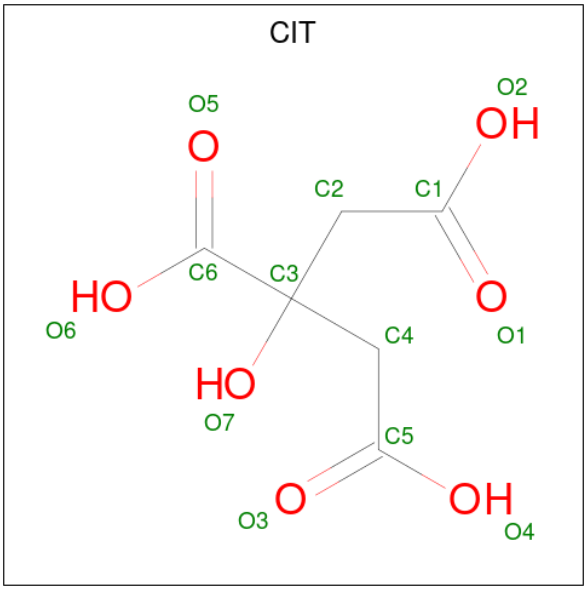
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	FFF	127	Total	C	N	O	S	0	4	0
			998	627	169	197	5			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 5 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	EEE	1	Total	C	O	0	0
			13	6	7		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	FFF	1	Total Cl 1 1	0	0

- Molecule 7 is water.

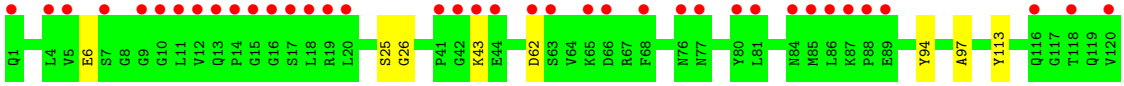
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	AAA	18	Total O 18 18	0	0
7	EEE	130	Total O 130 130	0	0
7	FFF	87	Total O 87 87	0	0

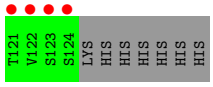
3 Residue-property plots (i)

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: C1 nanobody

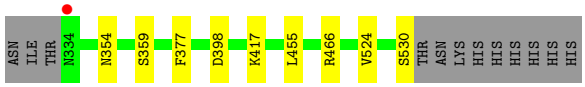
Chain AAA:





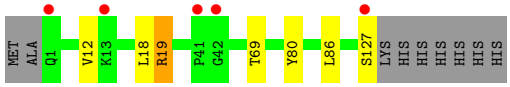
- Molecule 2: Spike protein S1

Chain EEE:



- Molecule 3: H3 nanobody

Chain FFF:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	105.68Å 105.68Å 112.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.25 – 1.90 62.25 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (62.25-1.90) 99.9 (62.25-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.180 , 0.204 0.191 , 0.214	Depositor DCC
R_{free} test set	2656 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å^2)	36.2	Xtriage
Anisotropy	0.160	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 45.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3797	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.67% 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: CL, NAG, CIT

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	0.73	0/1004	0.82	0/1366
2	EEE	0.67	0/1610	0.82	0/2189
3	FFF	0.72	0/1032	0.86	2/1396 (0.1%)
All	All	0.70	0/3646	0.83	2/4951 (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
1	AAA	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	FFF	19	ARG	CG-CD-NE	6.07	124.55	111.80
3	FFF	19	ARG	NE-CZ-NH1	5.13	122.87	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	25	SER	Peptide

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	973	0	922	2	0
2	EEE	1563	0	1488	7	0
3	FFF	998	0	963	4	0
4	EEE	14	0	13	0	0
5	EEE	13	0	5	2	0
6	FFF	1	0	0	0	0
7	AAA	18	0	0	0	0
7	EEE	130	0	0	2	1
7	FFF	87	0	0	1	0
All	All	3797	0	3391	14	1

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 2.

All (14) 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:EEE:359:SER:HA	2:EEE:524:VAL:HG22	1.77	0.66
2:EEE:417:LYS:HD3	2:EEE:455:LEU:HD12	1.80	0.62
2:EEE:359:SER:HA	2:EEE:524:VAL:CG2	2.32	0.59
2:EEE:417:LYS:HE3	7:EEE:737:HOH:O	2.03	0.58
3:FFF:19:ARG:HD2	3:FFF:80:TYR:CD1	2.40	0.57
5:EEE:602:CIT:O2	5:EEE:602:CIT:O7	2.24	0.53
2:EEE:466:ARG:HD3	5:EEE:602:CIT:H42	1.90	0.52
1:AAA:6:GLU:OE2	1:AAA:94:TYR:O	2.28	0.52
2:EEE:417:LYS:CE	7:EEE:737:HOH:O	2.62	0.46
2:EEE:354:ASN:O	2:EEE:398:ASP:HA	2.17	0.44
3:FFF:12:VAL:HG11	3:FFF:86:LEU:HD13	1.98	0.44
3:FFF:69[A]:THR:HG21	7:FFF:378:HOH:O	2.18	0.44
3:FFF:18:LEU:HD23	3:FFF:18:LEU:HA	1.93	0.42
1:AAA:97:ALA:HA	1:AAA:113:TYR:O	2.21	0.41

All (1) 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)
7:EEE:810:HOH:O	7:EEE:810:HOH:O[8_554]	1.80	0.40

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	124/131 (95%)	118 (95%)	5 (4%)	1 (1%)	19 9
2	EEE	196/209 (94%)	191 (97%)	5 (3%)	0	100 100
3	FFF	129/136 (95%)	128 (99%)	1 (1%)	0	100 100
All	All	449/476 (94%)	437 (97%)	11 (2%)	1 (0%)	47 38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	26	GLY

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	AAA	105/110 (96%)	103 (98%)	2 (2%)	57 53
2	EEE	171/182 (94%)	169 (99%)	2 (1%)	71 70
3	FFF	108/112 (96%)	107 (99%)	1 (1%)	78 79
All	All	384/404 (95%)	379 (99%)	5 (1%)	69 68

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	43	LYS
1	AAA	62	ASP
2	EEE	377	PHE
2	EEE	530	SER
3	FFF	127	SER

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 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	$\#\|Z\| > 2$
4	NAG	EEE	601	2	14,14,15	0.35	0	17,19,21	1.31	1 (5%)
5	CIT	EEE	602	-	12,12,12	1.12	1 (8%)	17,17,17	1.38	2 (11%)

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
4	NAG	EEE	601	2	-	0/6/23/26	0/1/1/1
5	CIT	EEE	602	-	-	5/16/16/16	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	EEE	602	CIT	C3-C6	2.04	1.55	1.53

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	EEE	601	NAG	C1-O5-C5	4.31	118.03	112.19
5	EEE	602	CIT	O5-C6-C3	-3.81	116.86	122.25
5	EEE	602	CIT	O6-C6-C3	2.76	117.84	113.05

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	EEE	602	CIT	C1-C2-C3-O7
5	EEE	602	CIT	C1-C2-C3-C6
5	EEE	602	CIT	C1-C2-C3-C4
5	EEE	602	CIT	O1-C1-C2-C3
5	EEE	602	CIT	O2-C1-C2-C3

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	EEE	602	CIT	2	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	124/131 (94%)	1.60	42 (33%) 0 0	36, 69, 91, 102	0
2	EEE	197/209 (94%)	0.34	1 (0%) 91 92	27, 38, 58, 85	0
3	FFF	127/136 (93%)	0.61	5 (3%) 39 42	27, 33, 62, 93	0
All	All	448/476 (94%)	0.77	48 (10%) 6 6	27, 42, 85, 102	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	65	LYS	5.9
1	AAA	122	VAL	5.7
1	AAA	1	GLN	5.5
1	AAA	10	GLY	5.5
1	AAA	85	MET	5.4
1	AAA	18	LEU	5.2
1	AAA	41	PRO	4.5
1	AAA	76	ASN	4.0
1	AAA	124	SER	3.9
1	AAA	20	LEU	3.8
1	AAA	5	VAL	3.8
1	AAA	17[A]	SER	3.7
1	AAA	16	GLY	3.6
3	FFF	127	SER	3.6
1	AAA	77	ASN	3.5
1	AAA	12	VAL	3.3
1	AAA	86	LEU	3.2
1	AAA	121	THR	3.2
1	AAA	80	TYR	3.2
1	AAA	11[A]	LEU	3.2
1	AAA	62	ASP	3.0
1	AAA	4	LEU	3.0
1	AAA	66	ASP	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	AAA	44	GLU	2.9
1	AAA	84	ASN	2.9
1	AAA	118	THR	2.8
1	AAA	68	PHE	2.8
1	AAA	9	GLY	2.8
3	FFF	41	PRO	2.8
1	AAA	120	VAL	2.7
1	AAA	123	SER	2.7
1	AAA	13	GLN	2.7
2	EEE	334	ASN	2.6
1	AAA	14	PRO	2.6
1	AAA	88	PRO	2.6
3	FFF	1	GLN	2.6
1	AAA	63	SER	2.4
1	AAA	87	LYS	2.4
1	AAA	15	GLY	2.4
1	AAA	42	GLY	2.3
1	AAA	116	GLN	2.3
1	AAA	7	SER	2.3
3	FFF	42	GLY	2.2
3	FFF	13	LYS	2.2
1	AAA	19	ARG	2.2
1	AAA	89	GLU	2.1
1	AAA	81	LEU	2.0
1	AAA	43	LYS	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
5	CIT	EEE	602	13/13	0.77	0.25	62,68,94,118	0
4	NAG	EEE	601	14/15	0.89	0.36	60,69,81,83	0
6	CL	FFF	201	1/1	0.99	0.19	32,32,32,32	0

6.5 Other polymers (i)

There are no such residues in this entry.