Full wwPDB X-ray Structure Validation Report (i)

Feb 13, 2025 – 06:07 PM EST

PDB ID	:	8FSQ
Title	:	Complex Structure of YejA with Microcin C7
Authors	:	Nair, S.K.; Dong, S.-H.
Deposited on	:	2023-01-11
Resolution	:	1.76 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:

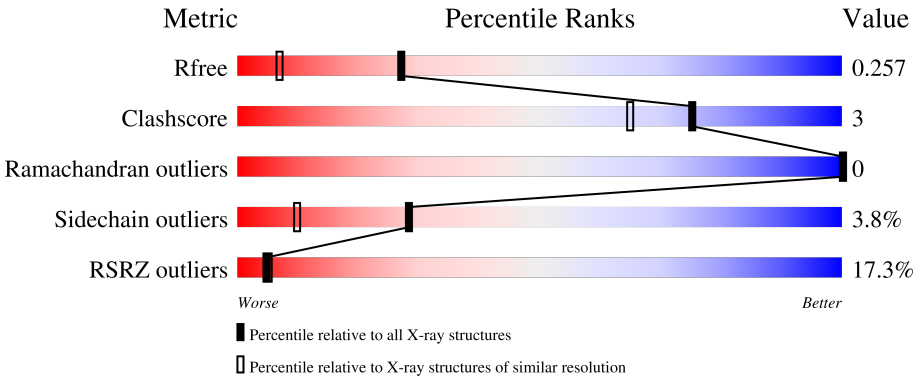
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 1.76 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}	164625	2888 (1.76-1.76)
Clashscore	180529	3097 (1.76-1.76)
Ramachandran outliers	177936	3072 (1.76-1.76)
Sidechain outliers	177891	3072 (1.76-1.76)
RSRZ outliers	164620	2887 (1.76-1.76)

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	A	585	17% 91% 6%
2	B	7	43% 57%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5428 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 YejA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	576	4725	3041	805	865	14	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	409	GLU	THR	conflict	UNP P33913
A	410	ALA	GLY	conflict	UNP P33913

- Molecule 2 is a protein called Microcin C7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
2	B	7	80	42	18	18	1	1	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	7	7MD	ASN	modified residue	UNP Q47505

- Molecule 3 is water.

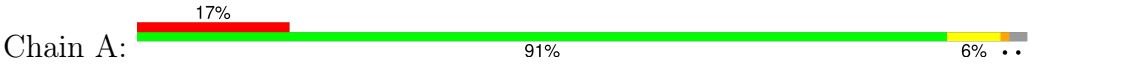
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	621	621	621	0	0
			Total	O		
3	B	2	2	2	0	0

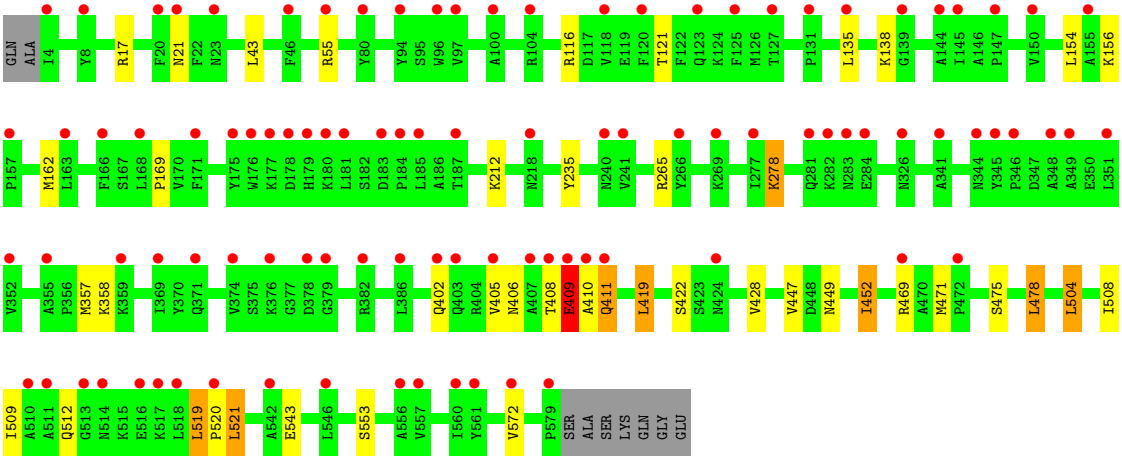
3 Residue-property plots (i)

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: YejA

Chain A:





• Molecule 2: Microcin C7

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	90.64Å 103.29Å 144.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	83.98 – 1.76 83.98 – 1.76	Depositor EDS
% Data completeness (in resolution range)	96.3 (83.98-1.76) 96.3 (83.98-1.76)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.70 (at 1.76Å)	Xtriage
Refinement program	REFMAC 5.8	Depositor
R, R_{free}	0.242 , (Not available) 0.240 , 0.257	Depositor DCC
R_{free} test set	3347 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	20.1	Xtriage
Anisotropy	0.891	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 34.0	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	0.94	EDS
Total number of atoms	5428	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.48% 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: 7MD, FME

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	0.64	0/4856	0.63	2/6587 (0.0%)
2	B	3.53	5/34 (14.7%)	1.97	2/44 (4.5%)
All	All	0.70	5/4890 (0.1%)	0.65	4/6631 (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
1	A	0	3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2	ARG	CZ-NH2	-7.66	1.23	1.33
2	B	2	ARG	NE-CZ	7.42	1.42	1.33
2	B	4	GLY	C-O	-7.28	1.12	1.23
2	B	2	ARG	CZ-NH1	7.01	1.42	1.33
2	B	2	ARG	C-O	-5.68	1.12	1.23

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	409	GLU	CB-CA-C	-8.93	92.53	110.40
1	A	278	LYS	CB-CA-C	8.18	126.76	110.40
2	B	2	ARG	CG-CD-NE	-5.62	100.00	111.80
2	B	2	ARG	CB-CA-C	-5.50	99.39	110.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	17	ARG	Sidechain
1	A	469	ARG	Sidechain
1	A	55	ARG	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	4725	0	4579	24	1
2	B	80	0	69	6	0
3	A	621	0	0	2	0
3	B	2	0	0	0	0
All	All	5428	0	4648	27	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 3.

All (27) 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:7:7MD:O4'	2:B:7:7MD:C4'	1.69	1.21
1:A:43:LEU:HD12	1:A:422:SER:HB2	1.62	0.81
1:A:235:TYR:HE2	2:B:7:7MD:H2'	1.47	0.77
1:A:419:LEU:HD21	1:A:428:VAL:HG21	1.75	0.67
1:A:409:GLU:O	1:A:410:ALA:HB3	1.96	0.64
1:A:235:TYR:CE2	2:B:7:7MD:H2'	2.32	0.64
1:A:449:ASN:HD21	2:B:7:7MD:H5'	1.63	0.63
1:A:447:VAL:HG23	1:A:452:ILE:HG12	1.81	0.62
1:A:405:VAL:HG13	1:A:410:ALA:HA	1.83	0.61
1:A:405:VAL:CG1	1:A:410:ALA:HA	2.32	0.60
1:A:154:LEU:HG	1:A:162:MSE:HE2	1.83	0.59
1:A:116:ARG:HG3	3:A:766:HOH:O	2.04	0.57
2:B:7:7MD:H33	2:B:7:7MD:O1P	2.05	0.57
1:A:509:ILE:HG23	1:A:512:GLN:HE21	1.70	0.55
1:A:508:ILE:HG13	1:A:521:LEU:HG	1.88	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:409:GLU:O	1:A:410:ALA:CB	2.58	0.51
1:A:121:THR:HG21	1:A:169:PRO:O	2.11	0.50
1:A:409:GLU:H	1:A:409:GLU:HG2	1.55	0.47
1:A:409:GLU:HB2	1:A:411:GLN:H	1.81	0.46
1:A:447:VAL:CG2	1:A:452:ILE:HG12	2.45	0.45
1:A:519:LEU:HB3	1:A:520:PRO:HD3	1.99	0.45
1:A:471:MSE:HE3	1:A:471:MSE:HB2	1.51	0.43
1:A:406:ASN:OD1	1:A:409:GLU:HG2	2.19	0.42
1:A:357:MSE:HE1	1:A:504:LEU:HD13	2.01	0.41
1:A:471:MSE:SE	1:A:478:LEU:HD13	2.71	0.41
2:B:7:7MD:HN8	2:B:7:7MD:H5'A	1.63	0.41
1:A:358:LYS:HG2	3:A:637:HOH: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)
1:A:21:ASN:ND2	1:A:408:THR:O[8_455]	1.99	0.21

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	575/585 (98%)	561 (98%)	14 (2%)	0	100	100
2	B	5/7 (71%)	5 (100%)	0	0	100	100
All	All	580/592 (98%)	566 (98%)	14 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	A	499/490 (102%)	479 (96%)	20 (4%)	27 9
2	B	3/3 (100%)	3 (100%)	0	100 100
All	All	502/493 (102%)	482 (96%)	20 (4%)	28 9

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	135	LEU
1	A	138	LYS
1	A	156	LYS
1	A	212	LYS
1	A	265	ARG
1	A	278	LYS
1	A	402	GLN
1	A	409	GLU
1	A	411	GLN
1	A	419	LEU
1	A	452	ILE
1	A	475	SER
1	A	478	LEU
1	A	504	LEU
1	A	519	LEU
1	A	521	LEU
1	A	543	GLU
1	A	553[A]	SER
1	A	553[B]	SER
1	A	572	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	57	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
2	FME	B	1	2	8,9,10	1.49	1 (12%)	8,9,11	3.44	4 (50%)

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
2	FME	B	1	2	-	5/7/9/11	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	FME	CN-N	3.77	1.45	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1	FME	CA-N-CN	6.33	132.55	122.82
2	B	1	FME	CB-CA-N	5.66	120.83	110.52
2	B	1	FME	CE-SD-CG	2.14	111.37	100.32
2	B	1	FME	O1-CN-N	-2.05	120.03	125.32

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	FME	O1-CN-N-CA
2	B	1	FME	CB-CA-N-CN
2	B	1	FME	O-C-CA-CB
2	B	1	FME	CA-CB-CG-SD
2	B	1	FME	C-CA-CB-CG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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	A	562/585 (96%)	1.19	98 (17%) 5 5	13, 23, 38, 52	1 (0%)
2	B	5/7 (71%)	0.95	0 100 100	22, 27, 29, 32	0
All	All	567/592 (95%)	1.19	98 (17%) 5 5	13, 23, 38, 52	1 (0%)

All (98) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	4	ILE	7.8
1	A	410	ALA	5.1
1	A	180	LYS	4.2
1	A	348	ALA	3.9
1	A	409	GLU	3.8
1	A	21	ASN	3.7
1	A	408	THR	3.7
1	A	345	TYR	3.5
1	A	351	LEU	3.5
1	A	176	TRP	3.5
1	A	517	LYS	3.4
1	A	326	ASN	3.3
1	A	127	THR	3.3
1	A	144	ALA	3.2
1	A	352	VAL	3.2
1	A	120	PHE	3.2
1	A	20	PHE	3.1
1	A	96	TRP	3.1
1	A	402	GLN	3.1
1	A	125	PHE	3.0
1	A	241	VAL	3.0
1	A	282	LYS	2.9
1	A	542	ALA	2.9
1	A	284	GLU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	178	ASP	2.9
1	A	157	PRO	2.8
1	A	469	ARG	2.8
1	A	376	LYS	2.8
1	A	8	TYR	2.8
1	A	349	ALA	2.8
1	A	277	ILE	2.7
1	A	369	ILE	2.7
1	A	560	ILE	2.7
1	A	579	PRO	2.7
1	A	100	ALA	2.7
1	A	510	ALA	2.7
1	A	374	VAL	2.7
1	A	379	GLY	2.7
1	A	118	VAL	2.7
1	A	572	VAL	2.7
1	A	163	LEU	2.6
1	A	407	ALA	2.6
1	A	171	PHE	2.6
1	A	386	LEU	2.6
1	A	371	GLN	2.6
1	A	518	LEU	2.6
1	A	150	VAL	2.6
1	A	175	TYR	2.6
1	A	168	LEU	2.5
1	A	403	GLN	2.5
1	A	218	ASN	2.5
1	A	283	ASN	2.5
1	A	147	PRO	2.5
1	A	346	PRO	2.5
1	A	46	PHE	2.5
1	A	145	ILE	2.5
1	A	97	VAL	2.5
1	A	281	GLN	2.5
1	A	520	PRO	2.4
1	A	131	PRO	2.4
1	A	516	GLU	2.4
1	A	359	LYS	2.4
1	A	80	TYR	2.4
1	A	179	HIS	2.3
1	A	405	VAL	2.3
1	A	181	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	341	ALA	2.3
1	A	269	LYS	2.3
1	A	166	PHE	2.3
1	A	557	VAL	2.3
1	A	546	LEU	2.3
1	A	183	ASP	2.2
1	A	513	GLY	2.2
1	A	378	ASP	2.2
1	A	556	ALA	2.2
1	A	177	LYS	2.2
1	A	424	ASN	2.2
1	A	123	GLN	2.2
1	A	139	GLY	2.2
1	A	411	GLN	2.2
1	A	185	LEU	2.2
1	A	355	ALA	2.1
1	A	511	ALA	2.1
1	A	23	ASN	2.1
1	A	472	PRO	2.1
1	A	94	TYR	2.1
1	A	135	LEU	2.1
1	A	266	TYR	2.1
1	A	184	PRO	2.1
1	A	561	TYR	2.1
1	A	344	ASN	2.1
1	A	514	ASN	2.1
1	A	155	ALA	2.1
1	A	382	ARG	2.1
1	A	187	THR	2.0
1	A	240	ASN	2.0
1	A	55	ARG	2.0
1	A	104	ARG	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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
2	FME	B	1	10/11	0.78	0.18	26,32,53,58	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.