WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 9, 2023 – 12:09 AM EST

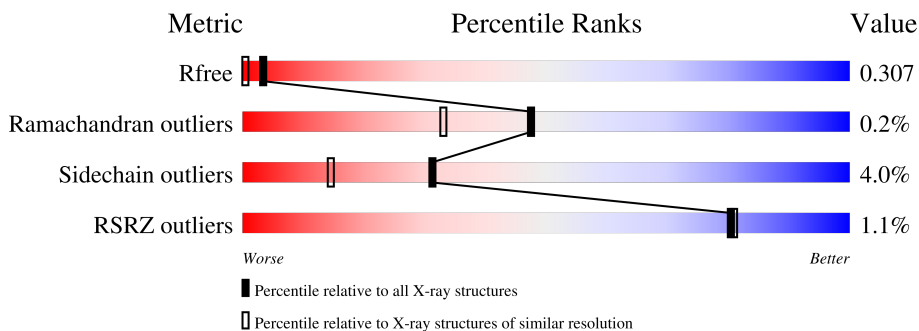|		
PDB ID	:	8F1F
Title	:	Structure of K48-linked tri-ubiquitin in complex with cyclic peptide
Authors	:	Lubkowski, J.; Fushman, D.; Lemma, B.
Deposited on	:	2022-11-05
Resolution	:	1.85 Å(reported)

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 1.85 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.
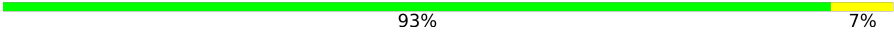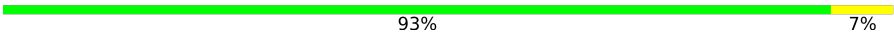
Metric	Percentile Ranks	Value
Rfree		0.307
Ramachandran outliers		0.2%
Sidechain outliers		4.0%
RSRZ outliers		1.1%

Worse — *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2469 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	77	3% / 91% / 9%
1	a	77	94% / ···
2	B	76	3% / 99% / ·
2	b	76	96% / ·
3	C	76	93% / 7%
3	c	76	.% / 92% / 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	1	14	93% 7%
4	2	14	93% 7%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 4183 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 Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	77	Total 617	C 388	N 106	O 122	S 1	0	2	0
1	a	75	Total 597	C 376	N 104	O 116	S 1	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	ASP	-	insertion	UNP P0CG47
a	77	ASP	-	insertion	UNP P0CG47

- Molecule 2 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total 601	C 378	N 105	O 117	S 1	0	0	0
2	b	76	Total 601	C 378	N 105	O 117	S 1	0	0	0

- Molecule 3 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	76	Total 612	C 386	N 107	O 118	S 1	0	2	0
3	c	76	Total 603	C 378	N 107	O 117	S 1	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	48	ARG	LYS	engineered mutation	UNP P0CG47
c	48	ARG	LYS	engineered mutation	UNP P0CG47

• Molecule 4 is a protein called Non-proteinogenic cyclic peptide (inhibitor).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	1	14	Total 107	C 74	N 15	O 17	S 1	0	0	1
4	2	14	Total 107	C 74	N 15	O 17	S 1	0	0	1

• Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
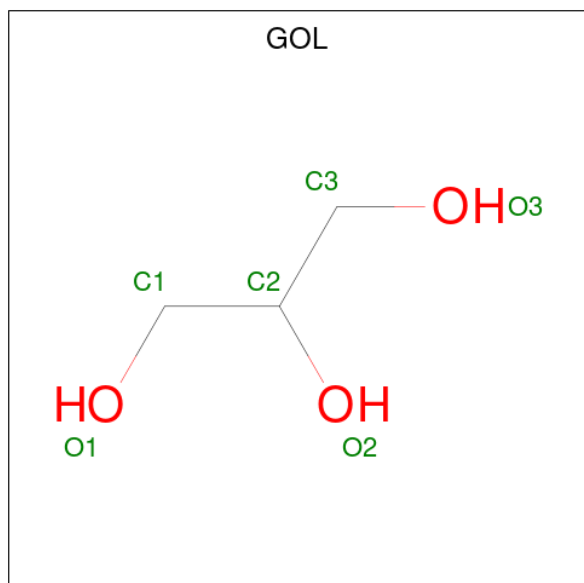
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 6	C 3	O 3	0	0

• Molecule 6 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).

EPE

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
6	b	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	55	Total	O	0	1
			56	56		
7	B	41	Total	O	0	0
			41	41		
7	C	45	Total	O	0	0
			45	45		
7	a	57	Total	O	0	1
			58	58		
7	b	43	Total	O	0	0
			43	43		
7	c	49	Total	O	0	1
			50	50		
7	1	3	Total	O	0	0
			3	3		
7	2	6	Total	O	0	0
			6	6		

3 Residue-property plots ⓘ

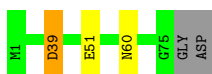These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Ubiquitin
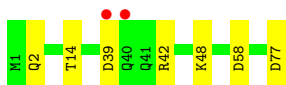
Chain A:

- Molecule 1: Ubiquitin

Chain a:

- Molecule 2: Ubiquitin

Chain B:

- Molecule 2: Ubiquitin

Chain b:

- Molecule 3: Ubiquitin

Chain C:

- Molecule 3: Ubiquitin

Chain c:

%

92%   8%

M1 K6 K11 P38 D39 R48 L67 G76

• Molecule 4: Non-proteinogenic cyclic peptide (inhibitor)

Chain 1:

93%   7%

ACE0 XC05 NH213

• Molecule 4: Non-proteinogenic cyclic peptide (inhibitor)

Chain 2:

93%   7%

ACE0 A7 NH213
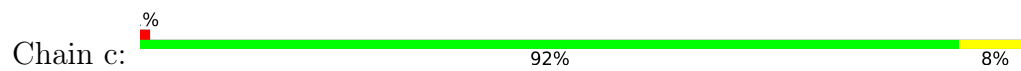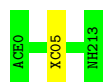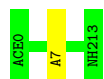
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	83.21Å 26.96Å 110.47Å 90.00° 99.85° 90.00°	Depositor
Resolution (Å)	30.00 – 1.85 29.82 – 1.85	Depositor EDS
% Data completeness (in resolution range)	93.5 (30.00-1.85) 93.6 (29.82-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	3.18 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.228 , 0.298 0.235 , 0.307	Depositor DCC
R_{free} test set	1986 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	17.0	Xtriage
Anisotropy	0.103	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 49.5	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.52, < L^2 > = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	4183	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 50.15 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.7428e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MAA, NH2, SAR, EPE, DPN, ACE, XC0

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	1.02	0/629	1.09	2/847 (0.2%)
1	a	0.93	1/603 (0.2%)	1.04	0/811
2	B	0.92	0/607	0.99	0/816
2	b	0.92	0/607	1.02	1/816 (0.1%)
3	C	0.98	1/624 (0.2%)	0.99	1/840 (0.1%)
3	c	1.02	0/609	1.07	3/819 (0.4%)
4	1	1.17	0/71	1.09	0/92
4	2	1.11	0/71	1.19	0/92
All	All	0.97	2/3821 (0.1%)	1.04	7/5133 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	a	51	GLU	CD-OE1	5.17	1.31	1.25
3	C	51	GLU	CD-OE2	5.12	1.31	1.25

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	b	54	ARG	NE-CZ-NH2	-6.89	116.86	120.30
3	c	48	ARG	NE-CZ-NH2	-6.42	117.09	120.30
3	c	48	ARG	NE-CZ-NH1	6.00	123.30	120.30
3	C	48	ARG	NE-CZ-NH2	-5.89	117.35	120.30
1	A	42	ARG	NE-CZ-NH2	-5.61	117.50	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	77/77 (100%)	76 (99%)	1 (1%)	0	100	100
1	a	73/77 (95%)	72 (99%)	0	1 (1%)	11	3
2	B	74/76 (97%)	74 (100%)	0	0	100	100
2	b	74/76 (97%)	74 (100%)	0	0	100	100
3	C	76/76 (100%)	76 (100%)	0	0	100	100
3	c	74/76 (97%)	73 (99%)	1 (1%)	0	100	100
4	1	8/14 (57%)	8 (100%)	0	0	100	100
4	2	8/14 (57%)	8 (100%)	0	0	100	100
All	All	464/486 (96%)	461 (99%)	2 (0%)	1 (0%)	47	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	a	39	ASP

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	71/69 (103%)	65 (92%)	6 (8%)	10	2
1	a	68/69 (99%)	66 (97%)	2 (3%)	42	26
2	B	68/68 (100%)	67 (98%)	1 (2%)	65	53
2	b	68/68 (100%)	66 (97%)	2 (3%)	42	26
3	C	70/68 (103%)	67 (96%)	3 (4%)	29	12
3	c	68/68 (100%)	64 (94%)	4 (6%)	19	6
4	1	7/7 (100%)	7 (100%)	0	100	100
4	2	7/7 (100%)	7 (100%)	0	100	100
All	All	427/424 (101%)	409 (96%)	18 (4%)	31	13

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	c	6	LYS
3	c	67	LEU
3	c	38	PRO
3	C	13	ILE
2	b	63	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	a	2	GLN
1	a	60	ASN
1	a	62	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues 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

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	SAR	1	9	4	4,4,5	0.67	0	1,3,5	1.59	0
4	XC0	2	5	4	8,9,10	0.71	0	4,9,11	1.10	0
4	MAA	2	7	4	4,5,6	2.07	1 (25%)	1,5,7	0.80	0
4	SAR	2	9	4	4,4,5	1.07	0	1,3,5	0.04	0
4	XC0	1	5	4	8,9,10	0.88	1 (12%)	4,9,11	0.89	0
4	MAA	1	7	4	4,5,6	1.28	0	1,5,7	0.57	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
4	SAR	1	9	4	-	1/1/2/3	-
4	XC0	2	5	4	-	0/7/8/10	-
4	MAA	2	7	4	-	0/1/4/6	-
4	SAR	2	9	4	-	1/1/2/3	-
4	XC0	1	5	4	-	1/7/8/10	-
4	MAA	1	7	4	-	0/1/4/6	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	2	7	MAA	CA-N	3.77	1.54	1.47
4	1	5	XC0	CA-N	-2.16	1.41	1.48

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	1	9	SAR	C-CA-N-CN
4	2	9	SAR	C-CA-N-CN
4	1	5	XC0	CG-CD-CE-C1

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

3 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	#$|Z| > 2$
6	EPE	b	101	-	15,15,15	1.77	1 (6%)	18,20,20	2.50	6 (33%)
5	GOL	A	101	-	5,5,5	0.52	0	5,5,5	0.84	0
6	EPE	B	101	-	15,15,15	2.01	1 (6%)	18,20,20	1.37	3 (16%)

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
6	EPE	b	101	-	-	3/9/19/19	0/1/1/1
5	GOL	A	101	-	-	2/4/4/4	-
6	EPE	B	101	-	-	2/9/19/19	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	101	EPE	C10-S	-7.35	1.67	1.77
6	b	101	EPE	C10-S	-6.08	1.68	1.77

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	b	101	EPE	O1S-S-C10	7.56	116.02	106.92
6	b	101	EPE	C5-C6-N1	3.97	118.79	110.64

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	b	101	EPE	C9-N1-C2	-2.97	103.63	111.23
6	b	101	EPE	C6-C5-N4	2.66	116.10	110.64
6	B	101	EPE	O1S-S-C10	2.64	110.09	106.92

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	101	GOL	C1-C2-C3-O3
6	B	101	EPE	C10-C9-N1-C6
6	b	101	EPE	N4-C7-C8-O8
5	A	101	GOL	O2-C2-C3-O3
6	B	101	EPE	N4-C7-C8-O8

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 ⓘ

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, 95th 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	77/77 (100%)	-0.04	2 (2%)	56	54	11, 19, 30, 44	0
1	a	75/77 (97%)	-0.15	0	100	100	10, 18, 28, 40	0
2	B	76/76 (100%)	0.04	2 (2%)	56	54	11, 19, 35, 44	0
2	b	76/76 (100%)	-0.04	0	100	100	12, 19, 35, 46	0
3	C	76/76 (100%)	0.09	0	100	100	13, 19, 33, 41	0
3	c	76/76 (100%)	0.10	1 (1%)	77	78	12, 18, 32, 37	0
4	1	8/14 (57%)	0.20	0	100	100	12, 16, 20, 23	0
4	2	8/14 (57%)	0.22	0	100	100	11, 14, 19, 22	0
All	All	472/486 (97%)	0.01	5 (1%)	80	81	10, 19, 34, 46	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	76	GLY	2.9
3	c	76	GLY	2.6
2	B	8	LEU	2.5
1	A	40	GLN	2.1
1	A	39	ASP	2.0

6.2   Non-standard residues in protein, DNA, RNA chains ⓘ

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, 95th 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	MAA	1	7	6/7	0.91	0.12	17,19,20,20	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	DPN	1	1	11/12	0.92	0.12	14,14,15,15	0
4	MAA	2	7	6/7	0.93	0.10	18,21,21,23	0
4	XC0	2	5	10/11	0.95	0.15	11,12,23,24	0
4	DPN	2	1	11/12	0.95	0.10	12,14,15,15	0
4	XC0	1	5	10/11	0.95	0.15	10,15,28,29	0
4	SAR	2	9	5/6	0.95	0.11	10,11,15,16	0
4	SAR	1	9	5/6	0.98	0.10	11,11,12,13	0

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, 95th 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	GOL	A	101	6/6	0.79	0.21	38,46,49,52	0
6	EPE	b	101	15/15	0.92	0.12	33,40,48,50	0
6	EPE	B	101	15/15	0.94	0.16	40,59,67,68	0

6.5   Other polymers (i)

There are no such residues in this entry.