wwPDB X-ray Structure Validation Summary Report (i)

Oct 31, 2023 – 06:15 PM JST

PDB ID	:	5GQR
Title	:	Crystal structure of PXY-CLE41-SERK2
Authors	:	Chai, J.J.; Zhang, H.Q.
Deposited on	:	2016-08-08
Resolution	:	3.50 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*
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 (i)) were used in the production of this report:

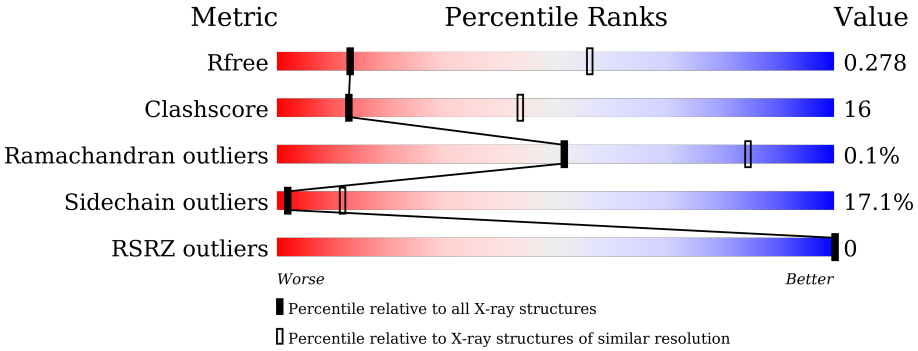
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	B	598	54% 37% 8% •
2	C	12	42% 50% 8%
3	K	185	65% 29% 5%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6160 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 Leucine-rich repeat receptor-like protein kinase TDR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	593	4613	2963	765	872	13	0	0	0

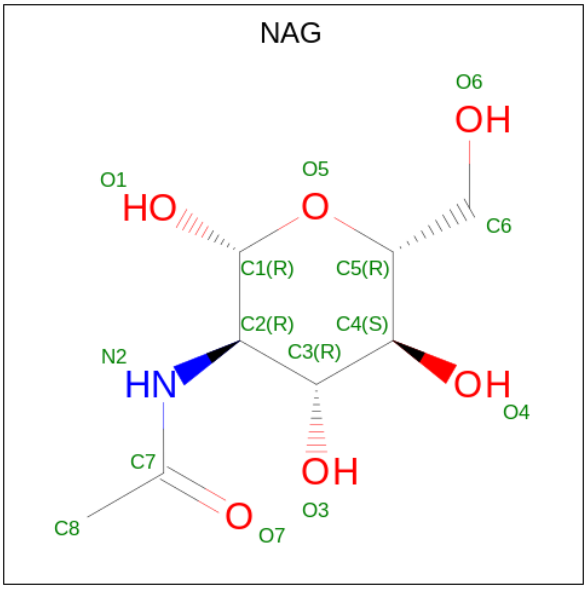
- Molecule 2 is a protein called TDIF.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	12	89	53	16	20	0	0	0

- Molecule 3 is a protein called Somatic embryogenesis receptor kinase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	K	185	1402	883	240	272	7	0	0	0

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



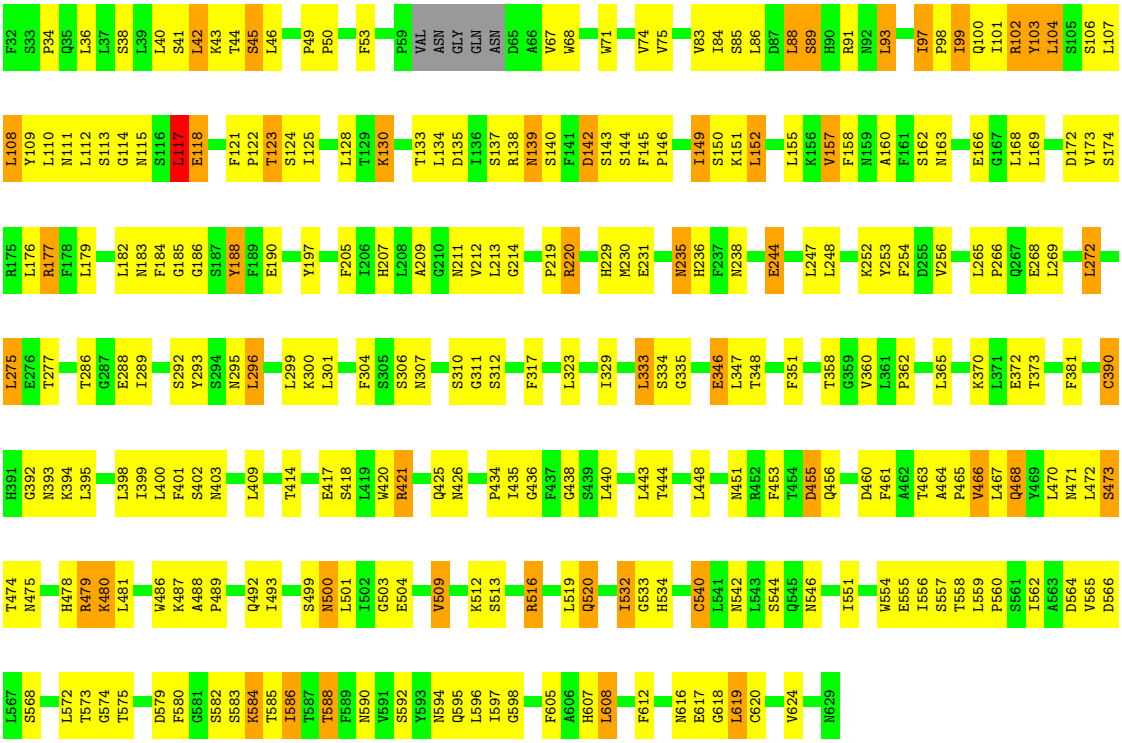
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	B	1	14	8	1	5	0	0
			Total	C	N	O		
4	B	1	14	8	1	5	0	0
			Total	C	N	O		
4	K	1	14	8	1	5	0	0
			Total	C	N	O		
4	K	1	14	8	1	5	0	0

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.

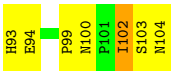
- Molecule 1: Leucine-rich repeat receptor-like protein kinase TDR

Chain B:



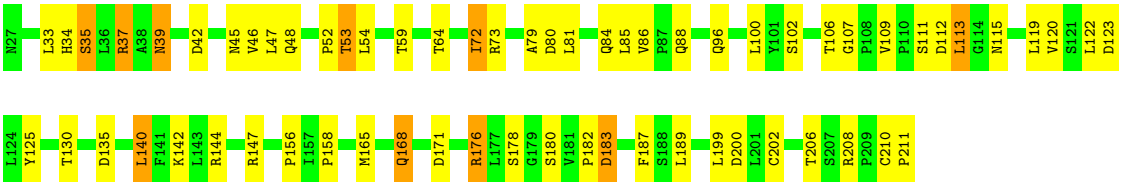
- Molecule 2: TDIF

Chain C:



- Molecule 3: Somatic embryogenesis receptor kinase 2

Chain K:


N27
L33
H34
S35
L36
R37
A38
N39
D42
M45
V46
L47
Q48
P52
T53
L54
T59
T64
I72
R73
A79
D80
L81
Q84
L85
V86
P87
Q88
Q96
L100
Y101
S102
T106
G107
P108
V109
P110
S111
D112
L113
G114
N115
L119
V120
S121
L122
D123
L124
Y125
T130
D135
L140
F141
K142
L143
R144
R147
P156
I157
P158
M165
Q168
D171
R176
L177
S178
G179
S180
V181
P182
D183
F187
S188
L189
L199
D200
L201
C202
T206
S207
R208
P209
C210
P211

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α, β, γ	162.42Å 162.42Å 187.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.47 – 3.50 40.47 – 3.50	Depositor EDS
% Data completeness (in resolution range)	95.0 (40.47-3.50) 94.4 (40.47-3.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	1.45 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.8_1069	Depositor
R, R_{free}	0.219 , 0.276 0.227 , 0.278	Depositor DCC
R_{free} test set	926 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	91.3	Xtriage
Anisotropy	0.263	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 27.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6160	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.32% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, HYP

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	B	0.47	0/4729	0.75	1/6428 (0.0%)
2	C	0.65	0/82	0.95	0/109
3	K	0.48	0/1432	0.72	0/1963
All	All	0.48	0/6243	0.75	1/8500 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	117	LEU	CA-CB-CG	5.68	128.37	115.30

There are no chirality outliers.

There are no planarity outliers.

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	B	4613	0	4546	167	0
2	C	89	0	78	4	0
3	K	1402	0	1366	35	0
4	B	28	0	26	0	0
4	K	28	0	26	0	0
All	All	6160	0	6042	201	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 16.

The worst 5 of 201 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)
3:K:64:THR:HB	3:K:73:ARG:HB2	1.60	0.81
1:B:438:GLY:HA2	1:B:464:ALA:HB2	1.63	0.81
1:B:219:PRO:HB3	1:B:244:GLU:HG3	1.62	0.80
1:B:213:LEU:HB2	1:B:235:ASN:HD22	1.46	0.78
1:B:421:ARG:HH12	3:K:54:LEU:HD11	1.47	0.78

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	B	589/598 (98%)	555 (94%)	33 (6%)	1 (0%)	47	81
2	C	9/12 (75%)	9 (100%)	0	0	100	100
3	K	183/185 (99%)	170 (93%)	13 (7%)	0	100	100
All	All	781/795 (98%)	734 (94%)	46 (6%)	1 (0%)	51	84

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	500	ASN

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	B	529/533 (99%)	435 (82%)	94 (18%)	2 10
2	C	10/10 (100%)	8 (80%)	2 (20%)	1 7
3	K	164/169 (97%)	140 (85%)	24 (15%)	3 18
All	All	703/712 (99%)	583 (83%)	120 (17%)	2 12

5 of 120 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	373	THR
3	K	130	THR
1	B	480	LYS
3	K	120	VAL
3	K	206	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	139	ASN
1	B	229	HIS
1	B	378	ASN
1	B	492	GLN
3	K	62	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	HYP	C	99	2	6,8,9	1.82	3 (50%)	5,10,12	2.18	2 (40%)

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	HYP	C	99	2	-	0/0/11/13	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	99	HYP	CB-CA	-2.53	1.48	1.54
2	C	99	HYP	CB-CG	-2.45	1.47	1.52
2	C	99	HYP	O-C	2.18	1.28	1.19

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	99	HYP	CB-CG-CD	3.52	107.58	103.27
2	C	99	HYP	O-C-CA	-2.63	117.89	124.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

4 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	$\#\|Z\| > 2$
4	NAG	B	702	1	14,14,15	0.57	0	17,19,21	1.42	2 (11%)
4	NAG	K	302	3	14,14,15	1.13	1 (7%)	17,19,21	2.27	3 (17%)
4	NAG	B	701	1	14,14,15	0.58	0	17,19,21	1.22	1 (5%)
4	NAG	K	301	3	14,14,15	0.47	0	17,19,21	1.24	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	B	702	1	-	4/6/23/26	0/1/1/1
4	NAG	K	302	3	-	2/6/23/26	0/1/1/1
4	NAG	B	701	1	-	2/6/23/26	0/1/1/1
4	NAG	K	301	3	-	4/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	K	302	NAG	O7-C7	2.32	1.28	1.23

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	K	302	NAG	C1-O5-C5	7.13	121.86	112.19
4	B	702	NAG	C1-O5-C5	4.30	118.02	112.19
4	K	302	NAG	C3-C4-C5	-3.93	103.22	110.24
4	K	301	NAG	C2-N2-C7	3.17	127.41	122.90
4	B	701	NAG	C1-O5-C5	3.09	116.38	112.19

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	K	301	NAG	C1-C2-N2-C7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	B	702	NAG	O5-C5-C6-O6
4	B	702	NAG	C8-C7-N2-C2
4	B	702	NAG	O7-C7-N2-C2
4	B	701	NAG	C8-C7-N2-C2

There are no ring outliers.

No monomer is involved in short contacts.

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 (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, 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	B	593/598 (99%)	-0.45	0	100	100	50, 78, 108, 135	0
2	C	11/12 (91%)	-0.62	0	100	100	55, 67, 75, 78	0
3	K	185/185 (100%)	-0.34	0	100	100	54, 68, 99, 129	0
All	All	789/795 (99%)	-0.43	0	100	100	50, 75, 106, 135	0

There are no RSRZ outliers to report.

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, 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
2	HYP	C	99	8/9	0.97	0.16	76,78,79,80	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
4	NAG	K	301	14/15	0.70	0.32	89,90,91,91	0
4	NAG	B	702	14/15	0.85	0.32	70,70,71,71	0
4	NAG	K	302	14/15	0.87	0.22	79,81,84,85	0
4	NAG	B	701	14/15	0.89	0.23	96,100,102,103	0

6.5 Other polymers (i)

There are no such residues in this entry.