wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 04:45 AM GMT

PDB ID	:	2O39
Title	:	Human Adenovirus type 11 knob in complex with domains SCR1 and SCR2 of CD46 (membrane cofactor protein, MCP)
Authors	:	Persson, D.B.; Reiter, D.M.; Arnberg, N.; Stehle, T.
Deposited on	:	2006-12-01
Resolution	:	2.85 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

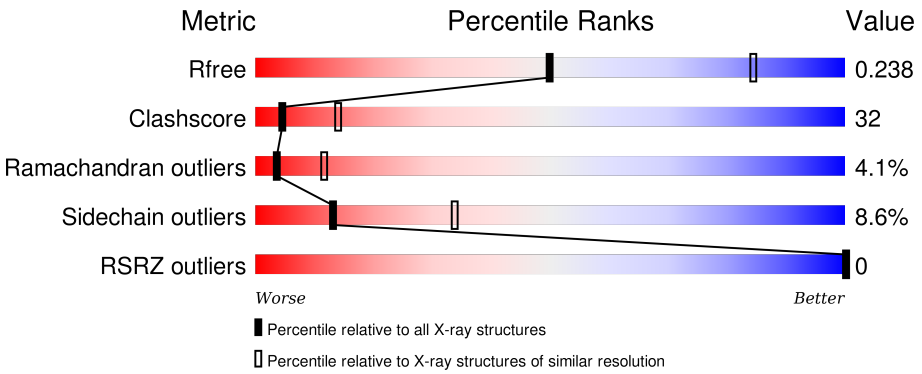
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.85 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2228 (2.90-2.82)
Clashscore	102246	2499 (2.90-2.82)
Ramachandran outliers	100387	2439 (2.90-2.82)
Sidechain outliers	100360	2442 (2.90-2.82)
RSRZ outliers	91569	2236 (2.90-2.82)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	197	54% 39% 7% •
1	B	197	56% 36% 6% •
2	C	126	44% 48% 7% •
2	D	126	44% 48% 6% •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5249 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 Fiber protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	197	1548	971	258	310	9	0	0	0
1	B	197	1548	971	258	310	9	0	0	0

- Molecule 2 is a protein called Membrane cofactor protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	126	1019	659	160	190	10	0	0	0
2	D	126	1019	659	160	190	10	0	0	0

- Molecule 3 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	C	3	39	22	2	15	0	0
3	D	3	39	22	2	15	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
4	D	2	2	2	0	0
4	C	1	1	1	0	0

- Molecule 5 is water.

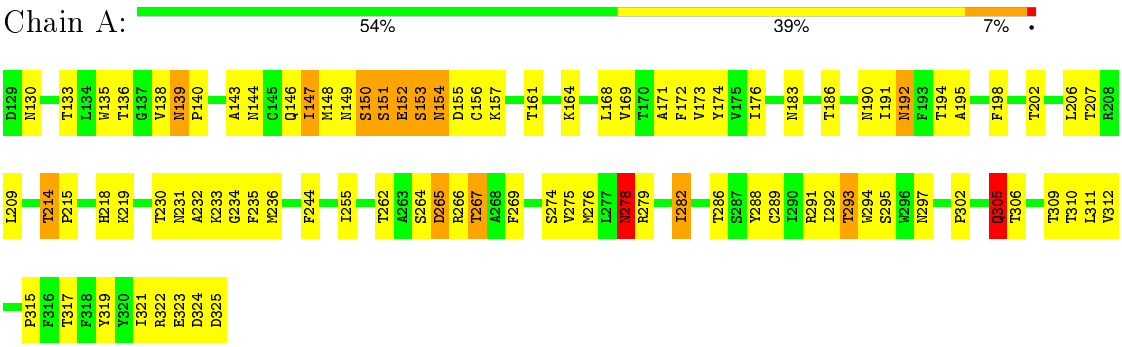
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	13	Total O 13 13	0	0
5	B	9	Total O 9 9	0	0
5	C	7	Total O 7 7	0	0
5	D	5	Total O 5 5	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

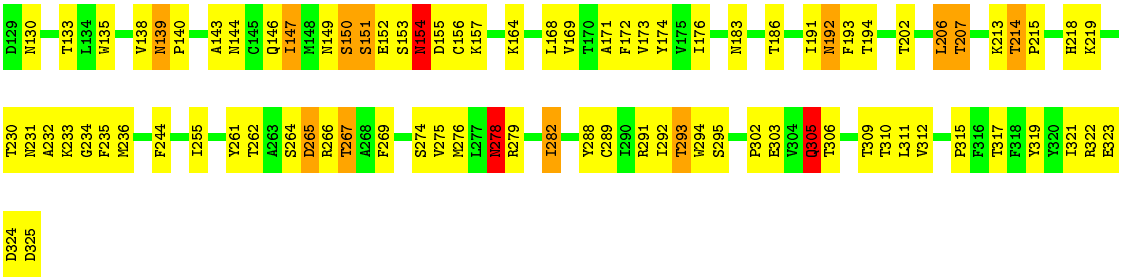
- Molecule 1: Fiber protein

Chain A:



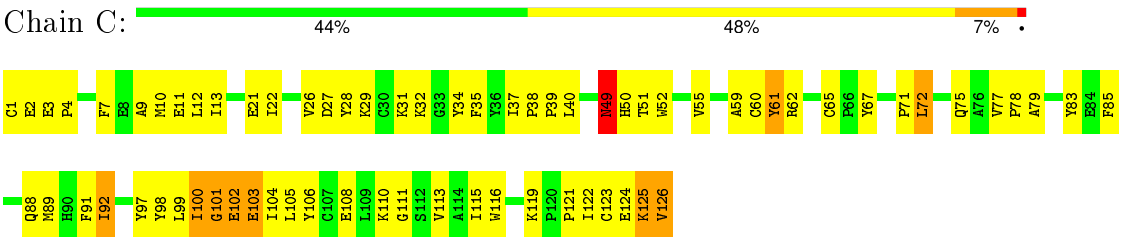
- Molecule 1: Fiber protein

Chain B:



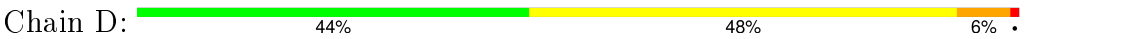
- Molecule 2: Membrane cofactor protein

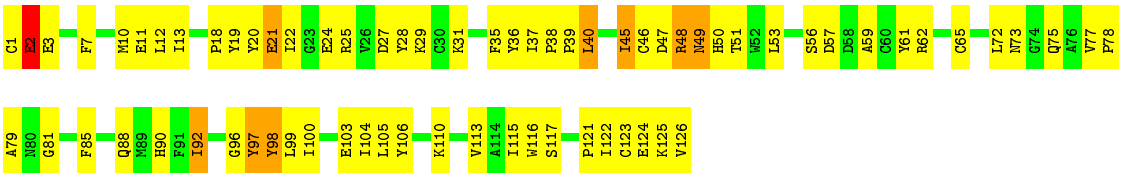
Chain C:



- Molecule 2: Membrane cofactor protein

Chain D:




C1
E2
E3
F7
M10
E11
L12
I13
P18
Y19
Y20
E21
I22
G23
E24
R25
Y26
D27
Y28
K29
C30
R31
F35
Y36
I37
P38
P39
L40
I45
C46
D47
R48
N49
H50
Y51
W52
L53
S56
D57
D58
A59
C60
Y61
R62
C65
L72
N73
G74
Q75
A76
V77
P78
A79
N80
G81
F85
Q88
M89
H90
F91
I92
G96
Y97
Y98
L99
I100
E103
I104
L105
Y106
K110
V113
A114
I115
W116
S117
P121
I122
C123
E124
K125
V126

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	106.14Å 106.14Å 68.32Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 2.85 38.13 – 2.39	Depositor EDS
% Data completeness (in resolution range)	93.1 (40.00-2.85) 77.3 (38.13-2.39)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 2.39Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.238 , 0.273 0.237 , 0.238	Depositor DCC
R_{free} test set	1778 reflections (10.04%)	DCC
Wilson B-factor (Å^2)	48.6	Xtriage
Anisotropy	0.471	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 16.3	EDS
Estimated twinning fraction	0.017 for -h,-k,l 0.457 for h,-h-k,-l 0.022 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 26502 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5249	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CA, BMA, NAG

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	0.42	0/1581	0.64	0/2156
1	B	0.43	0/1581	0.66	0/2156
2	C	0.42	1/1053 (0.1%)	0.61	0/1432
2	D	0.37	0/1053	0.61	0/1432
All	All	0.41	1/5268 (0.0%)	0.64	0/7176

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	2
2	C	0	1
2	D	0	1
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	126	VAL	C-OXT	5.05	1.32	1.23

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	152	GLU	Peptide
1	A	153	SER	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	C	49	ASN	Peptide
2	D	2	GLU	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	A	1548	0	1487	75	1
1	B	1548	0	1487	77	1
2	C	1019	0	963	92	0
2	D	1019	0	965	86	0
3	C	39	0	34	1	0
3	D	39	0	34	3	0
4	C	1	0	0	0	0
4	D	2	0	0	0	0
5	A	13	0	0	1	0
5	B	9	0	0	2	0
5	C	7	0	0	1	0
5	D	5	0	0	2	0
All	All	5249	0	4970	328	2

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 32.

The worst 5 of 328 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:D:1:CYS:SG	2:D:46:CYS:CB	2.25	1.24
2:C:99:LEU:HD11	2:C:103:GLU:HB3	1.35	1.07
2:C:1:CYS:N	2:C:50:HIS:HB3	1.72	1.03
2:D:62:ARG:HG3	2:D:85:PHE:CE2	2.01	0.96
2:C:101:GLY:O	2:C:102:GLU:O	1.85	0.94

All (2) 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:319:TYR:OH	1:A:319:TYR:OH[2_665]	1.90	0.30
1:B:319:TYR:OH	1:B:319:TYR:OH[3_555]	1.90	0.30

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	195/197 (99%)	173 (89%)	13 (7%)	9 (5%)	3	10
1	B	195/197 (99%)	173 (89%)	13 (7%)	9 (5%)	3	10
2	C	124/126 (98%)	107 (86%)	13 (10%)	4 (3%)	5	18
2	D	124/126 (98%)	106 (86%)	14 (11%)	4 (3%)	5	18
All	All	638/646 (99%)	559 (88%)	53 (8%)	26 (4%)	3	12

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	154	ASN
2	C	102	GLU
1	A	278	ASN
1	B	154	ASN
1	B	278	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	A	175/175 (100%)	160 (91%)	15 (9%)	13	34
1	B	175/175 (100%)	159 (91%)	16 (9%)	12	31
2	C	110/110 (100%)	102 (93%)	8 (7%)	17	42
2	D	110/110 (100%)	100 (91%)	10 (9%)	12	31
All	All	570/570 (100%)	521 (91%)	49 (9%)	13	34

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	207	THR
1	B	282	ILE
2	D	92	ILE
1	B	262	THR
1	B	292	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	146	GLN
1	B	149	ASN
1	B	305	GLN
1	A	305	GLN
2	C	75	GLN

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)

6 carbohydrates 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	NAG	C	1080	3,2	14,14,15	0.66	0	15,19,21	1.16	2 (13%)
3	NAG	C	2080	3	14,14,15	0.51	0	15,19,21	0.73	1 (6%)
3	BMA	C	3080	3	11,11,12	0.50	0	14,15,17	0.97	1 (7%)
3	NAG	D	1080	3,2	14,14,15	0.61	0	15,19,21	0.88	1 (6%)
3	NAG	D	2080	3	14,14,15	0.71	0	15,19,21	1.45	2 (13%)
3	BMA	D	3080	3	11,11,12	0.45	0	14,15,17	1.20	2 (14%)

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	NAG	C	1080	3,2	-	0/6/23/26	0/1/1/1
3	NAG	C	2080	3	-	0/6/23/26	0/1/1/1
3	BMA	C	3080	3	-	0/2/19/22	0/1/1/1
3	NAG	D	1080	3,2	-	0/6/23/26	0/1/1/1
3	NAG	D	2080	3	-	0/6/23/26	0/1/1/1
3	BMA	D	3080	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	1080	NAG	C2-N2-C7	-2.68	119.60	123.04
3	D	2080	NAG	C2-N2-C7	-2.25	120.15	123.04
3	C	1080	NAG	C2-N2-C7	-2.23	120.17	123.04
3	C	2080	NAG	C2-N2-C7	-2.02	120.45	123.04
3	D	3080	BMA	C1-O5-C5	2.51	115.43	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1080	NAG	1	0
3	D	1080	NAG	1	0
3	D	2080	NAG	2	0

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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	A	197/197 (100%)	-0.89	0	100	100	27, 39, 60, 75	0
1	B	197/197 (100%)	-0.91	0	100	100	23, 38, 60, 75	0
2	C	126/126 (100%)	-0.84	0	100	100	34, 52, 71, 78	0
2	D	126/126 (100%)	-0.84	0	100	100	26, 52, 74, 83	0
All	All	646/646 (100%)	-0.88	0	100	100	23, 44, 68, 83	0

There are no RSRZ outliers to report.

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)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
3	NAG	D	1080	14/15	0.90	0.12	-0.20	67,73,77,83	0
3	NAG	C	1080	14/15	0.90	0.10	-1.63	68,72,75,80	0
3	BMA	C	3080	11/12	0.88	0.10	-	95,97,97,97	0
3	NAG	C	2080	14/15	0.93	0.11	-	84,88,90,93	0
3	BMA	D	3080	11/12	0.83	0.12	-	98,99,100,100	0
3	NAG	D	2080	14/15	0.91	0.09	-	87,92,95,96	0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
4	CA	D	3082	1/1	0.84	0.14	1.27	67,67,67,67	0
4	CA	D	3081	1/1	0.98	0.17	-	53,53,53,53	0
4	CA	C	3081	1/1	0.98	0.18	-	54,54,54,54	0

6.5 Other polymers (i)

There are no such residues in this entry.