



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:27 PM GMT

PDB ID : 1FQ6  
Title : X-RAY STRUCTURE OF GLYCOL INHIBITOR PD-133,450 BOUND TO  
SACCHAROPEPSIN  
Authors : Cronin, N.B.; Badasso, M.O.; Tickle, I.J.; Dreyer, T.; Hoover, D.J.; Rosati,  
R.L.; Humblet, C.C.; Lunney, E.A.; Cooper, J.B.  
Deposited on : 2000-09-03  
Resolution : 2.70 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

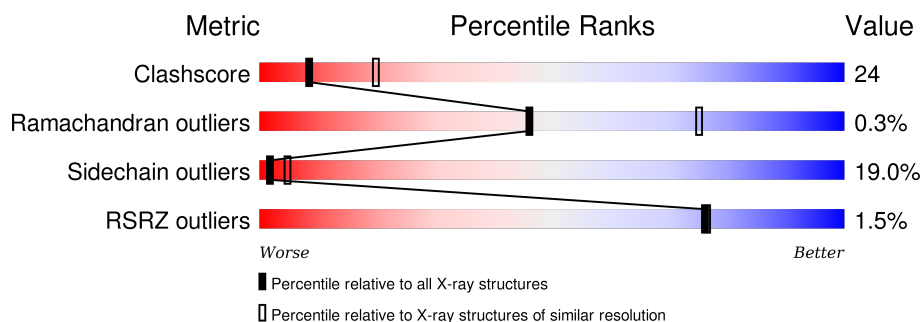
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	329	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GUP	A	336	X	-	-	-
4	NDG	A	337	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2749 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 SACCHAROPEPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	329	Total	C	N	O	S	0	0	0
			2528	1618	396	508	6			

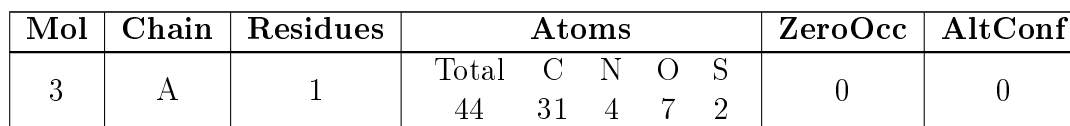
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	315	ILE	LEU	CONFLICT	UNP P07267

- Molecule 2 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 3 is N-[(1R)-2-{[(1R,2R,3S)-1-(CYCLOHEXYLMETHYL)-2,3-DIHYDROXY-5-METHYLHEXYL]AMINO}-1-(ETHYLSULFANYL)-2-OXOETHYL]-NALPHA-(MORPHOLIN-4-YLSULFONYL)-D-PHENYLALANINAMIDE (three-letter code: 0QF) (formula: C<sub>31</sub>H<sub>52</sub>N<sub>4</sub>O<sub>7</sub>S<sub>2</sub>).



- 
- The chemical structure shows a glucose molecule in its cyclic pyranose form, specifically the beta-D-glucopyranose isomer. The glucose ring is substituted with four acetyl groups (OAc) at the 2, 3, 4, and 6 positions. The acetyl groups are represented by a carbonyl group (C=O) and a methyl group (CH3). The glucose ring is labeled with C1(S), C2(R), C3(R), C4(S), C5(R), and C6. The oxygen atom of the ring is labeled O. The acetyl groups are labeled with O1L, O2, O3, O4, O5, and O6. The methyl groups are labeled with C7, C8, C9, C10, C11, and C12. The structure is shown in a chair conformation, with the acetyl groups at the 2, 3, 4, and 6 positions being in the 'down' position (axial or equatorial depending on the position). The glucose ring is shown in a perspective view, with the oxygen atom at the top. The labels C1(S), C2(R), C3(R), C4(S), C5(R), and C6 indicate the stereochemistry of the chiral centers. The labels O1L, O2, O3, O4, O5, and O6 indicate the oxygen atoms. The labels C7, C8, C9, C10, C11, and C12 indicate the carbon atoms. The structure is a chemical representation of the TGA molecule, which is a common sugar derivative used in various applications.

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

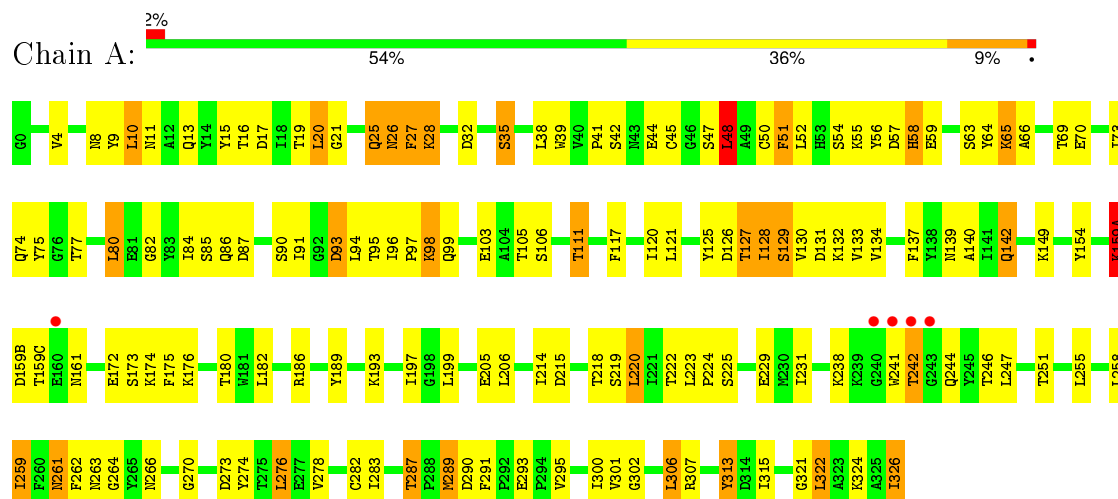
- 

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	102	Total 102	O 102	0	0

### 3 Residue-property plots

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 ( $\text{RSRZ} > 2$ ). 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: SACCHAROPEPSIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.91Å 86.91Å 110.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.60 – 2.70 15.63 – 2.75	Depositor EDS
% Data completeness (in resolution range)	100.0 (15.60-2.70) 99.2 (15.63-2.75)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.23 (at 2.73Å)	Xtriage
Refinement program	RESTRAIN	Depositor
R, $R_{free}$	0.180 , 0.270 0.171 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	45.2	Xtriage
Anisotropy	0.267	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 100.7	EDS
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 12810 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2749	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.84% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: NDG, BMA, GUP, NAG, OQF

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	0.55	0/2592	1.17	14/3526 (0.4%)

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	3	0
2	A	1	0
All	All	4	0

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	241	TRP	CA-CB-CG	12.84	138.10	113.70
1	A	4	VAL	CA-CB-CG1	7.70	122.45	110.90
1	A	93	ASP	CB-CG-OD1	-7.46	111.59	118.30
1	A	48	LEU	CA-CB-CG	7.04	131.49	115.30
1	A	220	LEU	CB-CG-CD2	6.78	122.52	111.00

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	63	SER	CA
1	A	86	GLN	CA
1	A	241	TRP	CA
2	A	336	GUP	C1



There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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	2528	0	2400	121	0
2	A	61	0	49	3	0
3	A	44	0	52	10	0
4	A	14	0	13	2	0
5	A	102	0	0	3	0
All	All	2749	0	2514	124	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 24.

The worst 5 of 124 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)
1:A:242:THR:HG22	1:A:242:THR:O	1.44	1.14
1:A:261:ASN:HD22	1:A:261:ASN:C	1.50	1.12
1:A:242:THR:CG2	1:A:242:THR:O	1.98	1.06
1:A:69:THR:HB	1:A:84:ILE:HD13	1.48	0.96
1:A:13:GLN:HE21	3:A:400:OQF:HE2	1.34	0.91

There are no symmetry-related clashes.

## 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	327/329 (99%)	304 (93%)	22 (7%)	1 (0%)	46 75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	127	THR

### 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	268/268 (100%)	217 (81%)	51 (19%)	2 5

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	111	THR
1	A	159(A)	LYS
1	A	313	TYR
1	A	129	SER
1	A	159(C)	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	86	GLN
1	A	142	GLN
1	A	232	ASN
1	A	58	HIS
1	A	161	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

5 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$
2	NAG	A	332	1,2	14,14,15	1.11	2 (14%)	15,19,21	1.55	3 (20%)
2	NAG	A	333	2	14,14,15	1.00	1 (7%)	15,19,21	1.97	4 (26%)
2	BMA	A	334	2	11,11,12	1.18	0	14,15,17	2.11	5 (35%)
2	BMA	A	335	2	11,11,12	1.15	2 (18%)	14,15,17	1.55	2 (14%)
2	GUP	A	336	2	11,11,12	1.57	2 (18%)	14,15,17	3.48	4 (28%)

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	NAG	A	332	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	333	2	-	0/6/23/26	0/1/1/1
2	BMA	A	334	2	-	0/2/19/22	0/1/1/1
2	BMA	A	335	2	-	0/2/19/22	0/1/1/1
2	GUP	A	336	2	1/1/4/5	0/2/19/22	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	336	GUP	C2-C3	-3.32	1.48	1.52
2	A	336	GUP	C1-C2	-2.01	1.47	1.52
2	A	333	NAG	O4-C4	2.02	1.47	1.43
2	A	332	NAG	C2-N2	2.12	1.50	1.46

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	335	BMA	O4-C4	2.16	1.48	1.43

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	332	NAG	C2-N2-C7	-3.57	118.45	123.04
2	A	335	BMA	O5-C1-C2	-3.54	105.11	110.86
2	A	334	BMA	C3-C4-C5	-2.32	106.14	110.20
2	A	333	NAG	C3-C4-C5	-2.18	106.40	110.20
2	A	332	NAG	O3-C3-C4	-2.18	105.44	110.34

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	336	GUP	C1

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	334	BMA	1	0
2	A	335	BMA	3	0

## 5.6 Ligand geometry

2 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	NDG	A	337	1	14,14,15	0.99	1 (7%)	15,19,21	1.32	2 (13%)
3	0QF	A	400	-	41,46,46	0.86	2 (4%)	46,62,62	3.98	9 (19%)

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	NDG	A	337	1	-	0/6/23/26	0/1/1/1
3	0QF	A	400	-	-	0/44/66/66	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	337	NDG	O4-C4	2.15	1.48	1.43
3	A	400	0QF	O1S-S	3.15	1.46	1.42
3	A	400	0QF	O2S-S	3.27	1.46	1.42

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	400	0QF	C5-N4-C3	-6.36	104.46	112.20
3	A	400	0QF	O1-C2-C3	-5.17	99.98	111.84
3	A	400	0QF	O1-C6-C5	-4.37	101.82	111.84
4	A	337	NDG	O3-C3-C4	-2.18	105.42	110.34
3	A	400	0QF	C5'-C4'-C3'	2.22	118.39	111.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	337	NDG	2	0
3	A	400	0QF	10	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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	329/329 (100%)	-0.70	5 (1%) 76 76	4, 24, 60, 130	4 (1%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	241	TRP	5.0
1	A	242	THR	4.8
1	A	240	GLY	4.4
1	A	243	GLY	2.8
1	A	160	GLU	2.7

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	NAG	A	332	14/15	0.92	0.16	0.04	29,42,108,121	0
2	BMA	A	334	11/12	0.89	0.24	-	22,33,63,67	0
2	NAG	A	333	14/15	0.97	0.10	-	6,37,55,134	0
2	GUP	A	336	11/12	0.61	0.31	-	35,45,69,157	0
2	BMA	A	335	11/12	0.93	0.22	-	0,31,71,73	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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	NDG	A	337	14/15	0.82	0.22	6.05	30,47,157,157	0
3	0QF	A	400	44/44	0.95	0.16	1.18	3,31,96,157	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.