



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:17 PM GMT

PDB ID : 1SYO  
Title : N-terminal 3 domains of CI-MPR bound to mannose 6-phosphate  
Authors : Olson, L.J.; Dahms, N.M.; Kim, J.-J.P.  
Deposited on : 2004-04-01  
Resolution : 2.20 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

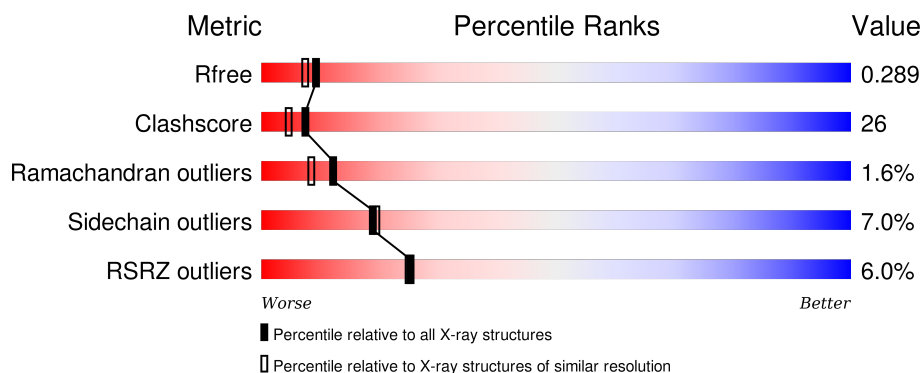
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	432	<div> <div>6%</div> <div>54%</div> <div>38%</div> <div>6%</div> </div>
1	B	432	<div> <div>6%</div> <div>57%</div> <div>35%</div> <div>6%</div> </div>

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	NDG	A	501	-	-	-	X
2	NAG	A	502	X	-	-	-
5	GOL	A	4005	-	X	-	-
5	GOL	B	4006	-	X	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6891 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.

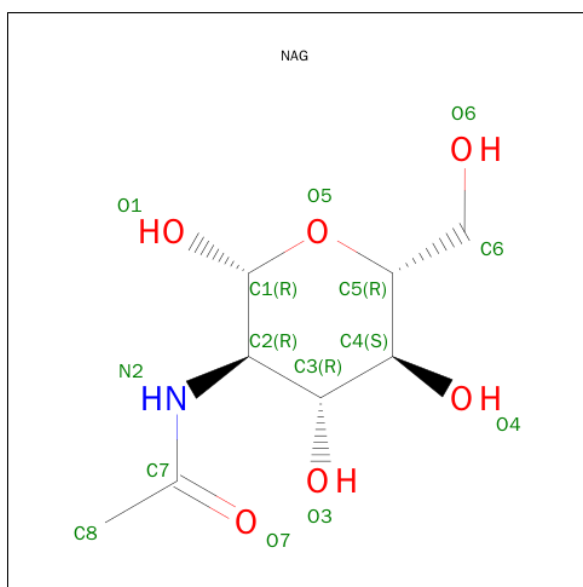
- Molecule 1 is a protein called cation-independent mannose 6-phosphate receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	0	0	0
			3268	2040	560	640	28			
1	B	408	Total	C	N	O	S	0	0	0
			3161	1980	538	616	27			

- Molecule 2 is a polymer of unknown type called SUGAR (NAG-NAG-MAN).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	3	Total	C	N	O	0	0
			40	22	2	16		

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



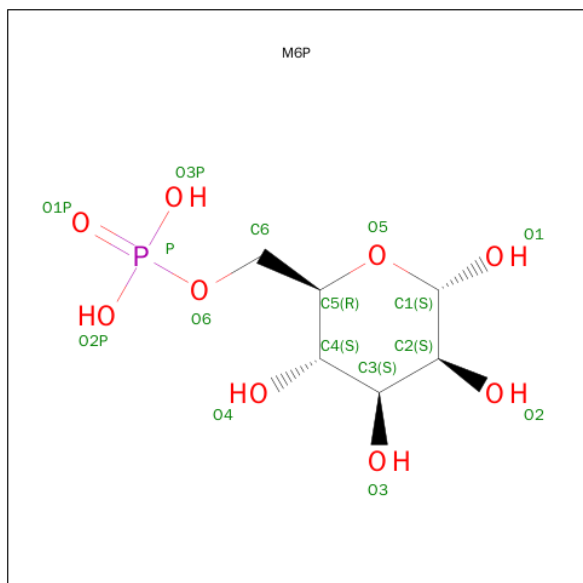
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 4 is ALPHA-D-MANNOSE-6-PHOSPHATE (three-letter code: M6P) (formula:  $C_6H_{13}O_9P$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	P	0	0
			16	6	9	1		
4	B	1	Total	C	O	P	0	0
			16	6	9	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).

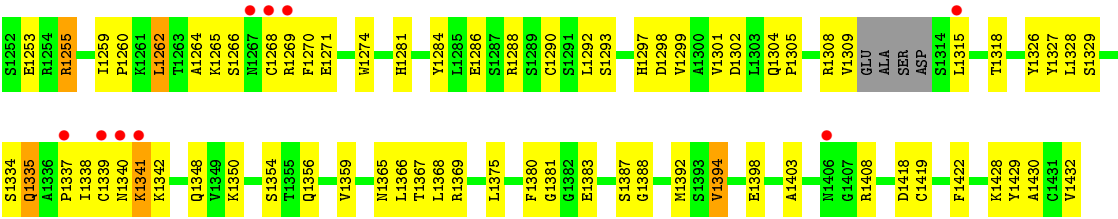


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	153	Total	O	0	0
			153	153		
6	B	183	Total	O	0	0
			183	183		







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.40 Å   86.40 Å   86.90 Å 90.00°   108.80°   90.00°	Depositor
Resolution (Å)	29.96 – 2.20 29.96 – 2.20	Depositor EDS
% Data completeness (in resolution range)	88.3 (29.96-2.20) 88.4 (29.96-2.20)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 2.20 Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.228 , 0.289 0.229 , 0.289	Depositor DCC
$R_{free}$ test set	2600 reflections (6.73%)	DCC
Wilson B-factor (Å <sup>2</sup> )	28.7	Xtriage
Anisotropy	0.473	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 50.9	EDS
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 42801 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6891	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.10% 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: M6P, GOL, BMA, NAG, NDG

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	0.39	0/3335	0.67	2/4520 (0.0%)
1	B	0.38	0/3224	0.66	0/4369
All	All	0.39	0/6559	0.67	2/8889 (0.0%)

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	A	1	0
All	All	1	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	308	ARG	NE-CZ-NH2	-6.90	116.85	120.30
1	A	203	ARG	NE-CZ-NH2	-6.45	117.07	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	502	NAG	C1

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	369	ARG	Sidechain
1	A	70	ARG	Sidechain

## 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	3268	0	3134	171	0
1	B	3161	0	3044	151	0
2	A	40	0	35	4	0
3	A	14	0	13	1	0
3	B	28	0	26	0	0
4	A	16	0	11	2	0
4	B	16	0	11	2	0
5	A	6	0	4	0	0
5	B	6	0	4	0	0
6	A	153	0	0	5	0
6	B	183	0	0	7	0
All	All	6891	0	6282	327	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 26.

The worst 5 of 327 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:B:1341:LYS:H	1:B:1341:LYS:HD3	1.15	1.06
1:B:1031:ASN:ND2	1:B:1035:ASN:H	1.53	1.06
1:A:308:ARG:HD2	1:A:408:ARG:CZ	1.94	0.97
1:B:1299:VAL:HG23	1:B:1432:VAL:HG13	1.47	0.97
1:B:1031:ASN:HD22	1:B:1035:ASN:H	1.18	0.92

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	418/432 (97%)	367 (88%)	42 (10%)	9 (2%)	8	4
1	B	400/432 (93%)	360 (90%)	36 (9%)	4 (1%)	19	16
All	All	818/864 (95%)	727 (89%)	78 (10%)	13 (2%)	12	8

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	ALA
1	A	204	ALA
1	A	40	GLN
1	A	126	ASN
1	A	165	PRO

### 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	365/370 (99%)	338 (93%)	27 (7%)	17	17
1	B	354/370 (96%)	331 (94%)	23 (6%)	21	23
All	All	719/740 (97%)	669 (93%)	50 (7%)	19	19

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	364	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	1009	PHE
1	B	1335	GLN
1	A	366	LEU
1	A	398	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1024	ASN
1	B	1113	HIS
1	B	1304	GLN
1	B	1031	ASN
1	A	281	HIS

### 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 ⓘ

3 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	NDG	A	501	1,2	14,14,15	0.66	0	15,19,21	1.06	1 (6%)
2	NAG	A	502	2	15,15,15	0.61	0	17,21,21	0.76	1 (5%)
2	BMA	A	503	2	11,11,12	0.54	0	14,15,17	0.46	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
2	NDG	A	501	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	502	2	1/1/6/7	0/6/26/26	0/1/1/1
2	BMA	A	503	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	NDG	C2-N2-C7	-2.99	119.19	123.04
2	A	502	NAG	O1-C1-O5	-2.18	104.29	110.25

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	502	NAG	C1

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	NDG	3	0
2	A	502	NAG	4	0

## 5.6 Ligand geometry

7 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
5	GOL	A	4005	-	5,5,5	4.84	5 (100%)	5,5,5	5.80	3 (60%)
4	M6P	A	500	-	16,16,16	1.56	3 (18%)	23,24,24	0.81	0
3	NAG	A	601	1	14,14,15	0.63	0	15,19,21	0.63	0
4	M6P	B	1500	-	16,16,16	1.56	4 (25%)	23,24,24	0.91	0
3	NAG	B	1501	1	14,14,15	0.47	0	15,19,21	0.75	1 (6%)
3	NAG	B	1601	1	14,14,15	0.49	0	15,19,21	0.81	1 (6%)
5	GOL	B	4006	-	5,5,5	4.83	5 (100%)	5,5,5	5.68	3 (60%)

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
5	GOL	A	4005	-	-	0/4/4/4	0/0/0/0
4	M6P	A	500	-	-	0/6/26/26	0/1/1/1
3	NAG	A	601	1	-	0/6/23/26	0/1/1/1
4	M6P	B	1500	-	-	0/6/26/26	0/1/1/1
3	NAG	B	1501	1	-	2/6/23/26	0/1/1/1
3	NAG	B	1601	1	-	0/6/23/26	0/1/1/1
5	GOL	B	4006	-	-	0/4/4/4	0/0/0/0

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	4005	GOL	C3-C2	-8.29	1.20	1.52
5	B	4006	GOL	C3-C2	-8.17	1.21	1.52
5	A	4005	GOL	C1-C2	-3.15	1.40	1.52
5	B	4006	GOL	C1-C2	-3.13	1.40	1.52
5	A	4005	GOL	O2-C2	-3.00	1.34	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1601	NAG	C2-N2-C7	-2.47	119.86	123.04
3	B	1501	NAG	C2-N2-C7	-2.25	120.15	123.04
5	A	4005	GOL	O1-C1-C2	3.00	124.75	110.18
5	B	4006	GOL	O1-C1-C2	3.25	125.96	110.18
5	A	4005	GOL	O2-C2-C3	6.78	139.74	108.65

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1501	NAG	O7-C7-N2-C2
3	B	1501	NAG	C8-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	500	M6P	2	0
3	A	601	NAG	1	0
4	B	1500	M6P	2	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	422/432 (97%)	0.36	24 (5%) 27 27	16, 40, 66, 84	0
1	B	408/432 (94%)	0.25	26 (6%) 23 22	13, 37, 64, 78	0
All	All	830/864 (96%)	0.30	50 (6%) 25 25	13, 38, 65, 84	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	308	ARG	8.3
1	A	203	ARG	6.5
1	B	1239	GLY	6.4
1	B	1340	ASN	5.1
1	B	1237	CYS	4.7

### 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	NDG	A	501	14/15	0.79	0.24	2.01	69,72,76,79	0
2	BMA	A	503	11/12	0.63	0.28	-	90,90,90,91	0
2	NAG	A	502	15/15	0.50	0.40	-	39,86,87,89	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(Å <sup>2</sup> )	Q<0.9
5	GOL	B	4006	6/6	0.86	0.21	4.10	48,49,49,51	0
4	M6P	B	1500	16/16	0.93	0.13	0.88	34,39,45,48	0
5	GOL	A	4005	6/6	0.88	0.20	0.66	63,66,66,68	0
4	M6P	A	500	16/16	0.92	0.12	-0.55	42,44,47,47	0
3	NAG	A	601	14/15	0.80	0.20	-	59,62,64,65	0
3	NAG	B	1601	14/15	0.85	0.25	-	52,54,57,62	0
3	NAG	B	1501	14/15	0.79	0.22	-	61,62,64,65	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.