



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:09 AM GMT

PDB ID : 2PMV  
Title : Crystal Structure of Human Intrinsic Factor- Cobalamin Complex at 2.6 Å Resolution  
Authors : Mathews, F.S.; Gordon, M.M.; Chen, Z.; Rajashankar, K.R.; Ealick, S.E.; Alpers, D.H.; Sukumar, N.  
Deposited on : 2007-04-23  
Resolution : 2.60 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

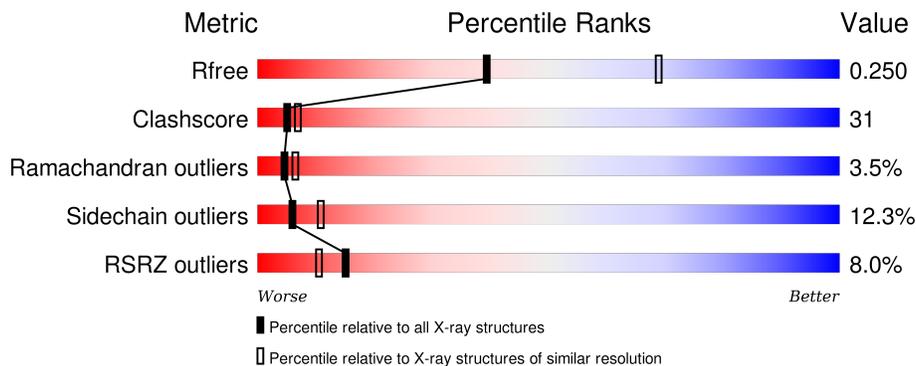
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	399	
1	B	399	
1	C	399	
1	D	399	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 10457 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 Gastric intrinsic factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	378	2871	1820	475	557	19	3	0	0
1	B	267	2009	1268	332	392	17	0	0	0
1	C	378	2871	1820	475	557	19	5	0	0
1	D	267	2009	1268	332	392	17	0	0	0

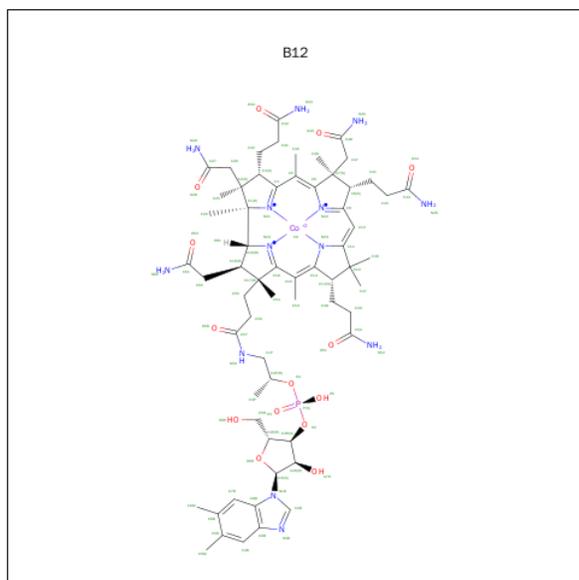
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	73	HIS	GLN	CONFLICT	UNP P27352
B	73	HIS	GLN	CONFLICT	UNP P27352
C	73	HIS	GLN	CONFLICT	UNP P27352
D	73	HIS	GLN	CONFLICT	UNP P27352

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	2	28	16	2	10	0	0
2	C	2	28	16	2	10	0	0

- Molecule 3 is COBALAMIN (three-letter code: B12) (formula: C<sub>62</sub>H<sub>89</sub>CoN<sub>13</sub>O<sub>14</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	Co	N	O			P
3	A	1	91	62	1	13	14	1	0	0
3	C	1	91	62	1	13	14	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	135	Total	O	0	0
			135	135		
4	B	100	Total	O	0	0
			100	100		
4	C	129	Total	O	0	0
			129	129		
4	D	95	Total	O	0	0
			95	95		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.10Å 67.30Å 147.70Å 90.00° 96.80° 90.00°	Depositor
Resolution (Å)	37.53 – 2.60 37.53 – 2.50	Depositor EDS
% Data completeness (in resolution range)	93.4 (37.53-2.60) 88.6 (37.53-2.50)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.44 (at 2.51Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.213 , 0.249 0.213 , 0.250	Depositor DCC
$R_{free}$ test set	1564 reflections (3.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.9	Xtriage
Anisotropy	0.558	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 49.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 55706 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10457	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.63% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: B12, 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	0.63	2/2925 (0.1%)	0.97	13/3978 (0.3%)
1	B	0.56	0/2044	0.76	1/2775 (0.0%)
1	C	0.64	5/2925 (0.2%)	0.94	12/3978 (0.3%)
1	D	0.55	0/2044	0.76	2/2775 (0.1%)
All	All	0.60	7/9938 (0.1%)	0.88	28/13506 (0.2%)

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	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	175	GLU	CB-CG	-8.85	1.35	1.52
1	C	52	LYS	CB-CG	-8.64	1.29	1.52
1	A	308	GLU	CD-OE1	-6.67	1.18	1.25
1	C	308	GLU	CD-OE1	-6.52	1.18	1.25
1	C	52	LYS	CG-CD	-6.31	1.30	1.52

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	175	GLU	N-CA-CB	-13.05	87.12	110.60
1	C	52	LYS	CD-CE-NZ	-12.87	82.11	111.70
1	A	52	LYS	CD-CE-NZ	-12.63	82.66	111.70
1	A	321	SER	N-CA-C	-11.86	78.97	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	304	LEU	N-CA-C	11.50	142.05	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	174	GLU	Mainchain

## 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	2871	0	2861	214	1
1	B	2009	0	2043	99	0
1	C	2871	0	2861	223	1
1	D	2009	0	2043	84	0
2	A	28	0	25	1	0
2	C	28	0	25	2	0
3	A	91	0	87	10	0
3	C	91	0	87	2	0
4	A	135	0	0	8	0
4	B	100	0	0	15	0
4	C	129	0	0	12	0
4	D	95	0	0	8	0
All	All	10457	0	10032	617	1

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 31.

The worst 5 of 617 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:175:GLU:OE1	1:A:177:TYR:CE1	1.89	1.24
1:C:305:ARG:O	1:C:307:VAL:N	1.80	1.14
1:A:333:GLN:HE21	1:A:340:LYS:HB2	1.09	1.13

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:304:LEU:CD1	1:C:305:ARG:HG3	1.77	1.13
1:C:302:ASN:HD22	1:C:303:GLN:N	1.46	1.12

All (1) 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:241:GLN:O	1:C:52:LYS:NZ[1_655]	1.87	0.33

## 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	374/399 (94%)	333 (89%)	24 (6%)	17 (4%)	<b>3</b> <b>4</b>
1	B	265/399 (66%)	240 (91%)	20 (8%)	5 (2%)	<b>10</b> <b>19</b>
1	C	374/399 (94%)	334 (89%)	22 (6%)	18 (5%)	<b>3</b> <b>3</b>
1	D	265/399 (66%)	243 (92%)	17 (6%)	5 (2%)	<b>10</b> <b>19</b>
All	All	1278/1596 (80%)	1150 (90%)	83 (6%)	45 (4%)	<b>4</b> <b>6</b>

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	13	ALA
1	A	291	ALA
1	A	306	GLY
1	A	311	PHE
1	A	320	LYS

### 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	322/349 (92%)	281 (87%)	41 (13%)	5	10
1	B	229/349 (66%)	203 (89%)	26 (11%)	7	12
1	C	322/349 (92%)	280 (87%)	42 (13%)	5	9
1	D	229/349 (66%)	203 (89%)	26 (11%)	7	12
All	All	1102/1396 (79%)	967 (88%)	135 (12%)	6	11

5 of 135 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	218	THR
1	C	124	LEU
1	D	174	GLU
1	B	232	THR
1	C	24	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	151	ASN
1	C	66	ASN
1	D	146	ASN
1	B	213	GLN
1	C	14	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

4 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	901	1,2	14,14,15	1.31	2 (14%)	15,19,21	1.96	4 (26%)
2	NAG	A	902	2	14,14,15	1.32	2 (14%)	15,19,21	1.39	2 (13%)
2	NAG	C	901	1,2	14,14,15	1.68	3 (21%)	15,19,21	2.09	4 (26%)
2	NAG	C	902	2	14,14,15	1.15	1 (7%)	15,19,21	1.27	2 (13%)

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	901	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	902	2	-	0/6/23/26	0/1/1/1
2	NAG	C	901	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	902	2	-	0/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	902	NAG	O5-C5	2.03	1.47	1.43
2	C	901	NAG	C2-N2	2.22	1.50	1.46
2	A	901	NAG	C4-C5	2.31	1.58	1.53
2	C	902	NAG	C4-C5	2.45	1.58	1.53
2	C	901	NAG	C4-C5	2.81	1.59	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	901	NAG	O7-C7-C8	-3.63	115.40	122.06
2	A	901	NAG	O7-C7-C8	-3.33	115.96	122.06

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	902	NAG	C1-O5-C5	-2.95	108.51	112.25
2	C	901	NAG	C4-C3-C2	-2.88	106.75	111.23
2	A	901	NAG	C4-C3-C2	-2.79	106.89	111.23

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	901	NAG	1	0
2	A	902	NAG	1	0
2	C	901	NAG	2	0
2	C	902	NAG	1	0

## 5.6 Ligand geometry [i](#)

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
3	B12	A	1001	-	74,101,101	1.70	16 (21%)	111,166,166	1.36	15 (13%)
3	B12	C	1002	-	74,101,101	1.62	10 (13%)	111,166,166	1.34	13 (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
3	B12	A	1001	-	-	0/51/223/223	0/3/11/11
3	B12	C	1002	-	-	0/51/223/223	0/3/11/11

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1001	B12	C8B-N1B	-3.36	1.34	1.38
3	C	1002	B12	C2R-C3R	-3.36	1.45	1.53
3	A	1001	B12	C2R-C3R	-3.31	1.45	1.53
3	A	1001	B12	P-O4	-2.08	1.43	1.48
3	A	1001	B12	P-O3	-2.04	1.54	1.59

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1001	B12	C37-C7-C8	-3.21	99.32	108.27
3	C	1002	B12	C41-C8-C7	-2.83	106.09	114.16
3	C	1002	B12	C37-C7-C8	-2.71	100.71	108.27
3	C	1002	B12	O58-C57-C56	-2.70	117.33	121.98
3	C	1002	B12	C48-C13-C12	-2.65	109.28	116.53

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
3	A	1001	B12	10	0
3	C	1002	B12	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	378/399 (94%)	0.27	38 (10%) 9 5	29, 49, 97, 99	1 (0%)
1	B	267/399 (66%)	0.08	8 (2%) 54 47	31, 48, 85, 99	0
1	C	378/399 (94%)	0.47	51 (13%) 4 2	30, 52, 99, 99	2 (0%)
1	D	267/399 (66%)	-0.06	6 (2%) 65 59	30, 46, 81, 99	0
All	All	1290/1596 (80%)	0.22	103 (7%) 15 10	29, 49, 97, 99	3 (0%)

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	304	LEU	7.0
1	C	272	PRO	6.6
1	A	305	ARG	6.3
1	B	174	GLU	6.2
1	C	338	MET	6.2

### 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( $\text{\AA}^2$ )	Q<0.9
2	NAG	A	901	14/15	0.88	0.20	-0.73	94,97,99,99	0
2	NAG	C	901	14/15	0.87	0.18	-0.81	86,87,91,94	0
2	NAG	C	902	14/15	0.85	0.32	-	95,98,99,99	0
2	NAG	A	902	14/15	0.78	0.43	-	97,99,99,99	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
3	B12	A	1001	91/91	0.96	0.21	0.73	30,41,53,68	0
3	B12	C	1002	91/91	0.95	0.20	0.36	37,47,60,66	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.