



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 2, 2024 – 09:20 pm GMT

PDB ID : 5A3L  
Title : Structure of Cea1A in complex with N-Acetylglucosamine  
Authors : Kock, M.; Brueckner, S.; Wozniak, N.; Veelders, M.; Schlereth, J.; Moesch, H.-U.; Essen, L.-O.  
Deposited on : 2015-06-02  
Resolution : 1.66 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (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

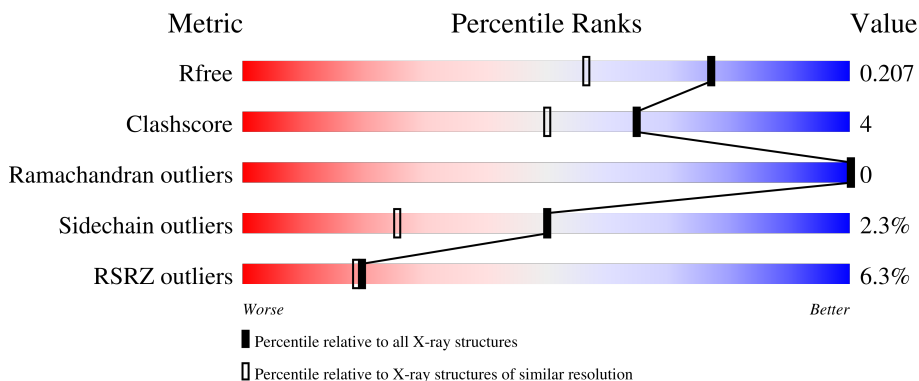
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.66 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	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	240	 2% (red), 80% (green), 7% (yellow), 13% (grey)
1	B	240	 2% (red), 82% (green), 13% (grey)
1	C	240	 10% (red), 78% (green), 5% (yellow), 16% (grey)
1	D	240	 8% (red), 78% (green), 5% (yellow), 16% (grey)

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 7992 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 CEA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	209	Total 1703	C 1091	N 270	O 334	S 8	0	10	0
1	B	209	Total 1697	C 1087	N 271	O 332	S 7	0	10	0
1	C	201	Total 1641	C 1058	N 264	O 313	S 6	0	10	0
1	D	202	Total 1624	C 1043	N 260	O 315	S 6	0	5	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

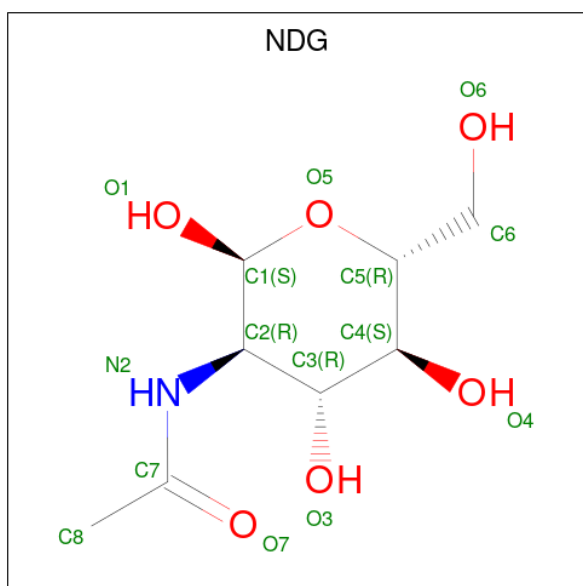
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Ca 1	0	0
2	B	1	Total 1	Ca 1	0	0
2	C	1	Total 1	Ca 1	0	0
2	D	1	Total 1	Ca 1	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



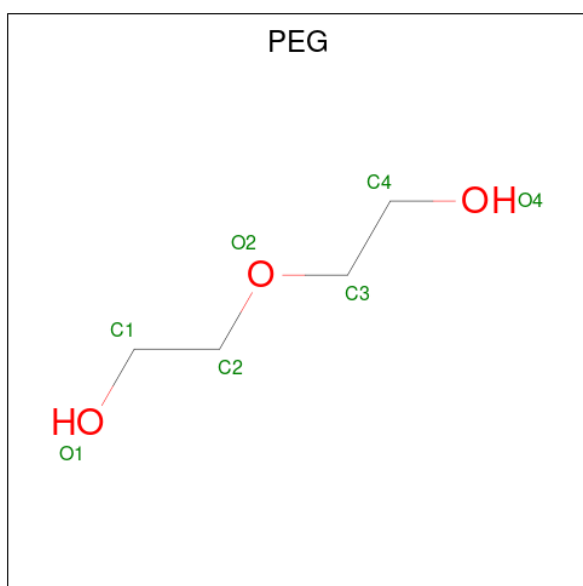
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	Total 15	C 8	N 1	O 6	0	1
3	B	1	Total 15	C 8	N 1	O 6	0	1
3	C	1	Total 15	C 8	N 1	O 6	0	1
3	D	1	Total 15	C 8	N 1	O 6	0	1

- Molecule 4 is 2-acetamido-2-deoxy-alpha-D-glucopyranose (three-letter code: NDG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	1
			15	8	1	6		
4	B	1	Total	C	N	O	0	1
			15	8	1	6		
4	C	1	Total	C	N	O	0	1
			15	8	1	6		
4	D	1	Total	C	N	O	0	1
			15	8	1	6		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).

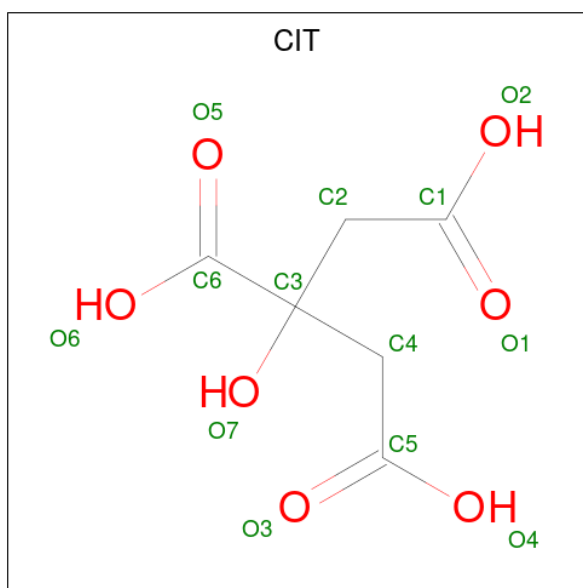


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			7	4	3		
5	D	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

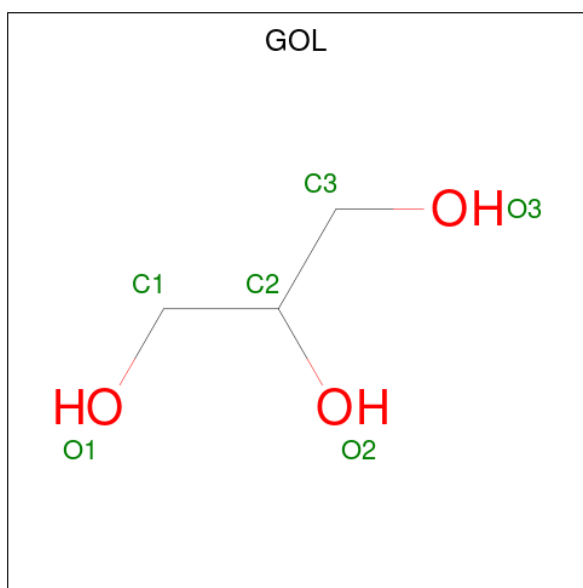
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Na	0	0
			1	1		
6	C	2	Total	Na	0	0
			2	2		
6	D	2	Total	Na	0	0
			2	2		

- Molecule 7 is CITRIC ACID (three-letter code: CIT) (formula:  $C_6H_8O_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	D	1	Total	C	O	0	1
			12	6	6		

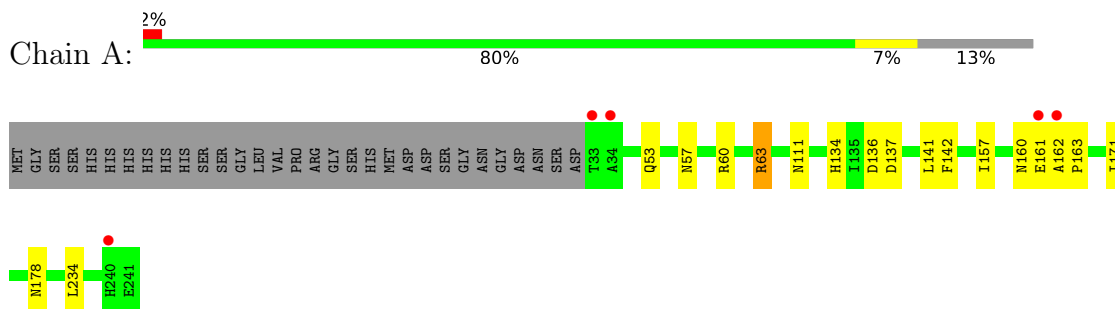
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	312	Total	O	0	16
			328	328		
9	B	313	Total	O	0	17
			330	330		
9	C	228	Total	O	0	7
			235	235		
9	D	239	Total	O	0	8
			247	247		

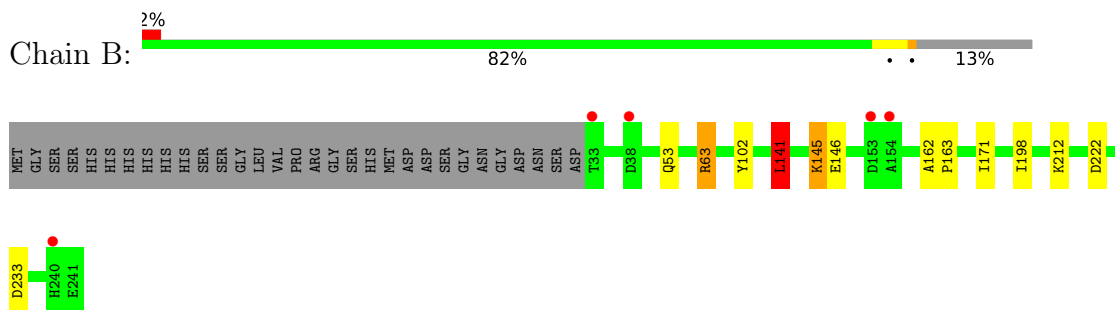
### 3 Residue-property plots [i](#)

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.

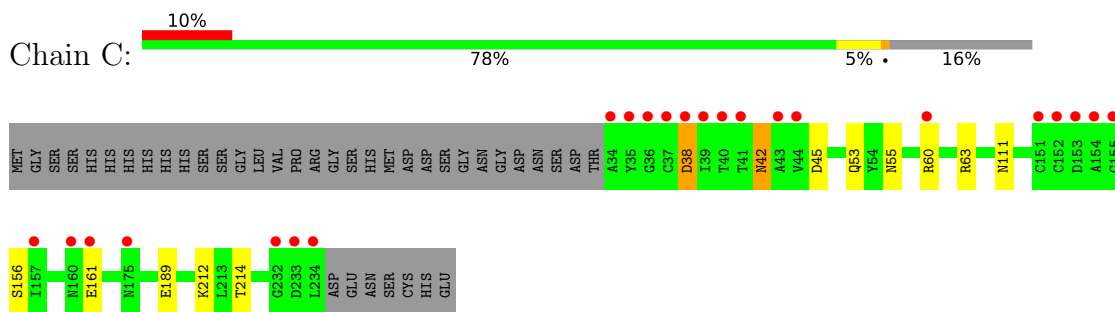
- Molecule 1: CEA1



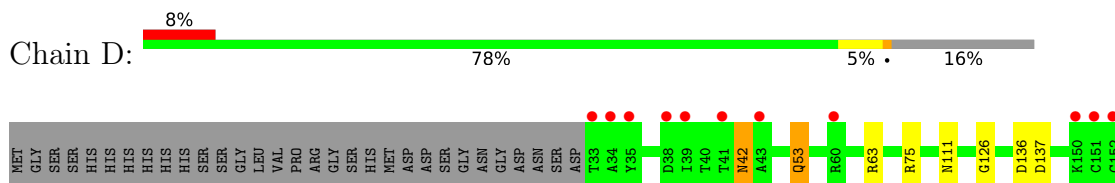
- Molecule 1: CEA1



- Molecule 1: CEA1



- Molecule 1: CEA1







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.34Å 106.21Å 107.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.84 – 1.66 19.84 – 1.66	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.84-1.66) 99.5 (19.84-1.66)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 1.66Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.171 , 0.202 0.179 , 0.207	Depositor DCC
$R_{free}$ test set	2084 reflections (1.51%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.3	Xtriage
Anisotropy	0.193	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 55.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.024 for -h,l,k 0.019 for -l,-k,-h 0.029 for k,h,-l 0.018 for k,l,h 0.018 for l,h,k	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7992	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.61% 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, 0.333 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: PEG, NAG, CA, NDG, GOL, NA, CIT

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	0.56	0/1766	0.79	0/2400
1	B	0.57	0/1763	0.86	4/2394 (0.2%)
1	C	0.54	0/1706	0.75	0/2315
1	D	0.52	0/1677	0.74	0/2279
All	All	0.55	0/6912	0.79	4/9388 (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	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	141[A]	LEU	CA-CB-CG	5.82	128.68	115.30
1	B	141[B]	LEU	CA-CB-CG	5.82	128.68	115.30
1	B	63[A]	ARG	NE-CZ-NH1	-5.27	117.66	120.30
1	B	63[B]	ARG	NE-CZ-NH1	-5.27	117.66	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	160[B]	ASN	Peptide

## 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	1703	0	1645	13	0
1	B	1697	0	1647	18	0
1	C	1641	0	1620	11	0
1	D	1624	0	1579	13	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	15	0	13	0	0
3	B	15	0	13	0	0
3	C	15	0	13	0	0
3	D	15	0	13	0	0
4	A	15	0	10	0	0
4	B	15	0	10	0	0
4	C	15	0	10	0	0
4	D	15	0	10	0	0
5	A	7	0	10	0	0
5	B	7	0	10	2	0
5	D	7	0	10	2	0
6	A	1	0	0	0	0
6	C	2	0	0	0	0
6	D	2	0	0	0	0
7	B	13	0	5	0	0
8	B	6	0	8	2	0
8	C	6	0	8	2	0
8	D	12	0	16	0	0
9	A	328	0	0	4	0
9	B	330	0	0	4	0
9	C	235	0	0	1	0
9	D	247	0	0	2	0
All	All	7992	0	6650	56	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 4.

All (56) 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:145:LYS:HE3	1:B:146:GLU:H	1.10	1.05
1:B:145:LYS:HE3	1:B:146:GLU:N	1.85	0.90
1:D:42:ASN:H	1:D:42:ASN:HD22	1.28	0.81
1:C:42:ASN:HD22	1:C:42:ASN:H	1.29	0.79
1:D:180[B]:GLU:OE1	9:D:401:HOH:O	1.99	0.79
1:A:141[B]:LEU:HD23	1:A:142:PHE:N	1.99	0.78
1:B:145:LYS:CE	1:B:146:GLU:H	1.96	0.76
1:B:212:LYS:HG2	1:B:222[B]:ASP:OD1	1.88	0.74
1:A:163:PRO:HD3	9:A:663[B]:HOH:O	1.89	0.71
1:D:53:GLN:HE22	1:D:75[B]:ARG:NH2	1.89	0.71
1:B:145:LYS:HE2	9:B:544:HOH:O	1.93	0.67
1:B:233:ASP:H	8:B:1246:GOL:C3	2.10	0.65
1:C:161:GLU:O	1:C:161:GLU:HG2	1.97	0.63
1:C:53:GLN:HE22	1:C:55:ASN:HD22	1.46	0.62
1:A:141[B]:LEU:C	1:A:141[B]:LEU:CD2	2.69	0.61
1:B:233:ASP:H	8:B:1246:GOL:H31	1.65	0.60
1:A:134:HIS:HD2	1:A:178:ASN:O	1.85	0.60
1:C:38:ASP:N	1:C:38:ASP:OD1	2.35	0.58
1:D:160:ASN:O	1:D:160:ASN:ND2	2.36	0.58
1:B:145:LYS:CE	1:B:146:GLU:HG2	2.33	0.58
1:A:134:HIS:HE1	9:A:641:HOH:O	1.88	0.57
9:B:521[B]:HOH:O	5:D:1239:PEG:H31	2.04	0.57
1:A:63:ARG:HG3	1:A:171:ILE:HG21	1.88	0.56
1:A:141[B]:LEU:CD2	1:A:142:PHE:N	2.70	0.55
1:D:53:GLN:HE22	1:D:75[B]:ARG:HH21	1.55	0.54
1:D:164:THR:HG21	5:D:1239:PEG:H32	1.89	0.54
1:C:45:ASP:HB3	9:C:580:HOH:O	2.08	0.53
1:B:63[A]:ARG:HG2	1:B:171:ILE:HD13	1.91	0.52
1:B:102:TYR:HB3	5:B:1247:PEG:H21	1.92	0.52
1:B:102:TYR:O	5:B:1247:PEG:H21	2.08	0.52
1:C:42:ASN:HD21	1:C:156:SER:H	1.58	0.52
1:D:42:ASN:HD21	1:D:156:SER:H	1.58	0.51
1:D:42:ASN:HD22	1:D:42:ASN:N	2.03	0.51
1:B:145:LYS:HE3	1:B:146:GLU:HG2	1.92	0.50
1:A:57:ASN:HB3	9:A:605:HOH:O	2.13	0.47
1:A:63:ARG:HG2	1:A:171:ILE:HD13	1.96	0.47
1:B:145:LYS:CE	9:B:544:HOH:O	2.57	0.47
1:A:60:ARG:NH1	9:A:408:HOH:O	2.46	0.47
1:A:157:ILE:HD11	1:A:234:LEU:HD11	1.97	0.46
1:D:234:LEU:C	1:D:234:LEU:HD12	2.35	0.46
1:C:42:ASN:HD22	1:C:42:ASN:N	2.03	0.45
1:A:161:GLU:HG2	1:A:162:ALA:H	1.81	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:212:LYS:NZ	9:B:408:HOH:O	2.51	0.44
1:C:212[B]:LYS:HE2	1:C:212[B]:LYS:HB2	1.32	0.44
1:B:141[A]:LEU:HD23	1:B:198:ILE:HG13	1.98	0.44
1:D:136:ASP:HA	1:D:137:ASP:HA	1.88	0.43
1:B:63[A]:ARG:HG3	1:B:171:ILE:HG21	2.01	0.43
1:A:136:ASP:HA	1:A:137:ASP:HA	1.84	0.43
1:D:53:GLN:NE2	1:D:75[B]:ARG:HH21	2.15	0.43
1:D:126:GLY:HA2	1:D:187:TYR:CE2	2.55	0.42
1:C:189[B]:GLU:HA	1:C:189[B]:GLU:OE1	2.20	0.41
1:C:214:THR:CG2	8:C:1238:GOL:H32	2.51	0.41
1:C:214:THR:HG21	8:C:1238:GOL:H32	2.02	0.41
1:B:162:ALA:HB1	1:B:163:PRO:HD2	2.04	0.40
1:D:212:LYS:HD2	9:D:594[B]:HOH:O	2.21	0.40

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	A	217/240 (90%)	214 (99%)	3 (1%)	0	100	100
1	B	217/240 (90%)	211 (97%)	6 (3%)	0	100	100
1	C	209/240 (87%)	205 (98%)	4 (2%)	0	100	100
1	D	206/240 (86%)	204 (99%)	2 (1%)	0	100	100
All	All	849/960 (88%)	834 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	191/207 (92%)	188 (98%)	3 (2%)	62	41
1	B	190/207 (92%)	186 (98%)	4 (2%)	53	29
1	C	183/207 (88%)	177 (97%)	6 (3%)	38	12
1	D	180/207 (87%)	175 (97%)	5 (3%)	43	18
All	All	744/828 (90%)	726 (98%)	18 (2%)	50	23

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	GLN
1	A	63	ARG
1	A	111	ASN
1	B	53	GLN
1	B	141[A]	LEU
1	B	141[B]	LEU
1	B	145	LYS
1	C	38	ASP
1	C	42	ASN
1	C	60	ARG
1	C	63[A]	ARG
1	C	63[B]	ARG
1	C	111	ASN
1	D	42	ASN
1	D	53	GLN
1	D	63	ARG
1	D	111	ASN
1	D	160	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	53	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	55	ASN
1	A	111	ASN
1	A	134	HIS
1	A	176	GLN
1	B	53	GLN
1	B	55	ASN
1	B	76	ASN
1	C	42	ASN
1	C	55	ASN
1	C	111	ASN
1	D	42	ASN
1	D	53	GLN
1	D	111	ASN
1	D	226	ASN

### 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](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 25 ligands modelled in this entry, 9 are monoatomic - leaving 16 for Mogul analysis.

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	1236[A]	2	15,15,15	0.45	0	21,21,21	1.49	4 (19%)
5	PEG	B	1247	-	6,6,6	0.32	0	5,5,5	0.21	0
5	PEG	A	1245	-	6,6,6	0.45	0	5,5,5	0.50	0
4	NDG	D	1237[B]	2	15,15,15	0.59	0	21,21,21	1.02	1 (4%)
3	NAG	D	1236[A]	2	15,15,15	0.51	0	21,21,21	1.33	2 (9%)
3	NAG	B	1243[A]	2	15,15,15	0.44	0	21,21,21	1.35	3 (14%)
7	CIT	B	1245	-	12,12,12	2.12	3 (25%)	17,17,17	4.09	9 (52%)
5	PEG	D	1239	-	6,6,6	0.45	0	5,5,5	0.30	0
4	NDG	A	1244[B]	2	15,15,15	0.62	0	21,21,21	1.57	5 (23%)
8	GOL	D	1238[B]	-	5,5,5	0.30	0	5,5,5	0.34	0
3	NAG	A	1243[A]	2	15,15,15	0.53	0	21,21,21	1.70	6 (28%)
8	GOL	B	1246	-	5,5,5	0.46	0	5,5,5	0.74	0
4	NDG	C	1237[B]	2	15,15,15	0.44	0	21,21,21	1.15	2 (9%)
8	GOL	C	1238	-	5,5,5	0.37	0	5,5,5	0.56	0
4	NDG	B	1244[B]	2	15,15,15	0.54	0	21,21,21	1.07	3 (14%)
8	GOL	D	1238[A]	-	5,5,5	0.25	0	5,5,5	0.17	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	1236[A]	2	-	0/6/26/26	0/1/1/1
5	PEG	B	1247	-	-	3/4/4/4	-
5	PEG	A	1245	-	-	0/4/4/4	-
4	NDG	D	1237[B]	2	-	0/6/26/26	0/1/1/1
3	NAG	D	1236[A]	2	-	0/6/26/26	0/1/1/1
3	NAG	B	1243[A]	2	-	0/6/26/26	0/1/1/1
7	CIT	B	1245	-	-	4/16/16/16	-
5	PEG	D	1239	-	-	2/4/4/4	-
4	NDG	A	1244[B]	2	-	0/6/26/26	0/1/1/1
8	GOL	D	1238[B]	-	-	4/4/4/4	-
3	NAG	A	1243[A]	2	-	0/6/26/26	0/1/1/1
8	GOL	B	1246	-	-	2/4/4/4	-
4	NDG	C	1237[B]	2	-	0/6/26/26	0/1/1/1
8	GOL	C	1238	-	-	1/4/4/4	-
4	NDG	B	1244[B]	2	-	0/6/26/26	0/1/1/1
8	GOL	D	1238[A]	-	-	2/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	1245	CIT	C4-C3	-4.48	1.48	1.53
7	B	1245	CIT	C3-C6	3.38	1.56	1.53
7	B	1245	CIT	O2-C1	-2.27	1.23	1.30

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	1245	CIT	O7-C3-C6	-10.60	93.98	108.86
7	B	1245	CIT	C2-C3-C6	-8.05	92.81	110.11
7	B	1245	CIT	O6-C6-C3	6.24	123.88	113.05
7	B	1245	CIT	O7-C3-C2	4.42	119.74	109.40
3	C	1236[A]	NAG	O3-C3-C2	4.16	118.06	109.66
3	D	1236[A]	NAG	O3-C3-C2	3.85	117.44	109.66
7	B	1245	CIT	C4-C3-C2	3.84	119.18	109.16
3	A	1243[A]	NAG	O3-C3-C2	3.74	117.21	109.66
3	A	1243[A]	NAG	O5-C1-C2	3.66	113.19	109.52
4	A	1244[B]	NDG	O5-C1-C2	3.24	112.78	109.52
3	B	1243[A]	NAG	O5-C1-C2	3.24	112.77	109.52
3	B	1243[A]	NAG	O3-C3-C2	3.17	116.07	109.66
7	B	1245	CIT	O7-C3-C4	2.98	116.36	109.40
3	A	1243[A]	NAG	C1-C2-N2	-2.85	107.42	110.73
3	D	1236[A]	NAG	O4-C4-C5	2.75	116.13	109.30
4	D	1237[B]	NDG	O3-C3-C2	2.65	115.01	109.66
4	A	1244[B]	NDG	O3-C3-C2	2.62	114.94	109.66
3	B	1243[A]	NAG	O4-C4-C3	-2.60	104.34	110.35
7	B	1245	CIT	O6-C6-O5	-2.58	115.63	123.82
3	A	1243[A]	NAG	O4-C4-C3	-2.56	104.43	110.35
3	A	1243[A]	NAG	O4-C4-C5	2.55	115.63	109.30
4	A	1244[B]	NDG	O4-C4-C5	2.54	115.60	109.30
4	B	1244[B]	NDG	O4-C4-C3	-2.45	104.69	110.35
4	C	1237[B]	NDG	O3-C3-C2	2.44	114.58	109.66
4	A	1244[B]	NDG	O4-C4-C3	-2.39	104.83	110.35
7	B	1245	CIT	C4-C3-C6	-2.35	105.06	110.11
4	B	1244[B]	NDG	C1-O5-C5	-2.30	109.33	113.66
4	A	1244[B]	NDG	O7-C7-N2	2.29	126.17	121.95
3	C	1236[A]	NAG	O4-C4-C5	2.17	114.69	109.30
4	C	1237[B]	NDG	O5-C1-C2	2.16	111.69	109.52
3	C	1236[A]	NAG	O3-C3-C4	-2.16	105.36	110.35
3	A	1243[A]	NAG	C3-C2-N2	-2.12	106.60	110.62
3	C	1236[A]	NAG	O4-C4-C3	-2.07	105.56	110.35
7	B	1245	CIT	O4-C5-C4	2.03	120.88	114.35
4	B	1244[B]	NDG	O3-C3-C2	2.00	113.70	109.66

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	1245	CIT	C1-C2-C3-O7
8	B	1246	GOL	O1-C1-C2-C3
8	D	1238[A]	GOL	O1-C1-C2-C3
8	D	1238[B]	GOL	C1-C2-C3-O3
5	D	1239	PEG	O2-C3-C4-O4
5	D	1239	PEG	O1-C1-C2-O2
5	B	1247	PEG	O2-C3-C4-O4
8	D	1238[B]	GOL	O1-C1-C2-C3
8	B	1246	GOL	O1-C1-C2-O2
8	D	1238[A]	GOL	O1-C1-C2-O2
8	D	1238[B]	GOL	O2-C2-C3-O3
8	D	1238[B]	GOL	O1-C1-C2-O2
5	B	1247	PEG	C4-C3-O2-C2
8	C	1238	GOL	C1-C2-C3-O3
7	B	1245	CIT	C3-C4-C5-O3
7	B	1245	CIT	C3-C4-C5-O4
5	B	1247	PEG	C1-C2-O2-C3
7	B	1245	CIT	O7-C3-C4-C5

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1247	PEG	2	0
5	D	1239	PEG	2	0
8	B	1246	GOL	2	0
8	C	1238	GOL	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

### 6.1 Protein, DNA and RNA chains

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	209/240 (87%)	-0.18	5 (2%) 59 59	7, 11, 26, 55	0
1	B	209/240 (87%)	-0.32	5 (2%) 59 59	7, 11, 23, 46	0
1	C	201/240 (83%)	0.24	23 (11%) 5 4	8, 13, 53, 74	0
1	D	202/240 (84%)	0.21	19 (9%) 8 7	8, 13, 50, 70	0
All	All	821/960 (85%)	-0.02	52 (6%) 20 19	7, 12, 43, 74	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	234	LEU	7.8
1	D	33	THR	7.1
1	C	39	ILE	6.7
1	A	33	THR	6.6
1	D	233	ASP	6.1
1	D	160	ASN	6.0
1	C	160	ASN	6.0
1	C	233	ASP	5.8
1	C	234	LEU	5.8
1	C	153	ASP	5.8
1	A	162	ALA	5.6
1	D	38	ASP	5.4
1	D	34	ALA	5.4
1	C	41	THR	5.2
1	A	161	GLU	4.9
1	D	153	ASP	4.8
1	C	154	ALA	4.8
1	D	35	TYR	4.5
1	C	34	ALA	4.5
1	C	37	CYS	4.2
1	D	41	THR	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	38	ASP	4.1
1	D	175	ASN	4.0
1	D	154	ALA	4.0
1	C	35	TYR	3.9
1	C	152	CYS	3.8
1	B	33	THR	3.8
1	D	151	CYS	3.7
1	D	161	GLU	3.6
1	C	175	ASN	3.6
1	C	36	GLY	3.6
1	C	151	CYS	3.5
1	B	240	HIS	3.1
1	D	150	LYS	3.0
1	D	152	CYS	3.0
1	C	40	THR	2.8
1	D	39	ILE	2.8
1	D	43	ALA	2.8
1	C	44	VAL	2.6
1	C	161	GLU	2.6
1	C	232	GLY	2.6
1	C	43	ALA	2.5
1	B	154	ALA	2.5
1	B	153	ASP	2.5
1	A	240	HIS	2.3
1	D	159	LEU	2.2
1	A	34	ALA	2.1
1	C	157	ILE	2.1
1	B	38	ASP	2.0
1	D	60	ARG	2.0
1	C	60	ARG	2.0
1	C	155	GLY	2.0

## 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](#)

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, 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	B-factors(Å <sup>2</sup> )	Q<0.9
8	GOL	C	1238	6/6	0.74	0.24	34,36,38,39	0
5	PEG	D	1239	7/7	0.77	0.23	36,37,37,38	0
8	GOL	B	1246	6/6	0.79	0.25	30,35,36,41	0
5	PEG	A	1245	7/7	0.79	0.22	30,32,33,33	0
8	GOL	D	1238[A]	6/6	0.85	0.28	20,21,22,22	6
8	GOL	D	1238[B]	6/6	0.85	0.28	25,27,28,29	6
7	CIT	B	1245	13/13	0.90	0.17	20,22,22,23	0
5	PEG	B	1247	7/7	0.92	0.18	26,28,30,31	0
3	NAG	B	1243[A]	15/15	0.97	0.07	8,8,9,9	15
3	NAG	D	1236[A]	15/15	0.97	0.06	9,10,11,11	15
4	NDG	A	1244[B]	15/15	0.97	0.07	9,10,10,11	15
4	NDG	B	1244[B]	15/15	0.97	0.07	9,9,10,10	15
4	NDG	C	1237[B]	15/15	0.97	0.07	9,11,12,12	15
4	NDG	D	1237[B]	15/15	0.97	0.07	10,11,12,12	15
3	NAG	A	1243[A]	15/15	0.97	0.07	8,8,9,9	15
6	NA	C	1240	1/1	0.98	0.11	15,15,15,15	0
3	NAG	C	1236[A]	15/15	0.98	0.06	8,8,9,10	15
6	NA	A	1246	1/1	0.99	0.08	7,7,7,7	0
2	CA	C	1235	1/1	0.99	0.04	9,9,9,9	0
6	NA	D	1240	1/1	0.99	0.05	17,17,17,17	0
2	CA	B	1242	1/1	0.99	0.04	8,8,8,8	0
2	CA	A	1242	1/1	1.00	0.05	8,8,8,8	0
2	CA	D	1235	1/1	1.00	0.04	9,9,9,9	0
6	NA	D	1241	1/1	1.00	0.12	3,3,3,3	0
6	NA	C	1239	1/1	1.00	0.14	3,3,3,3	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.