



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 23, 2016 – 07:52 PM EDT

PDB ID : 5A3M  
Title : Structure of Cea1A in complex with Chitobiose  
Authors : Kock, M.; Brueckner, S.; Wozniak, N.; Veelders, M.; Schlereth, J.; Moesch, H.-U.; Essen, L.-O.  
Deposited on : 2015-06-02  
Resolution : 1.75 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  
<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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027790  
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) : rb-20027790

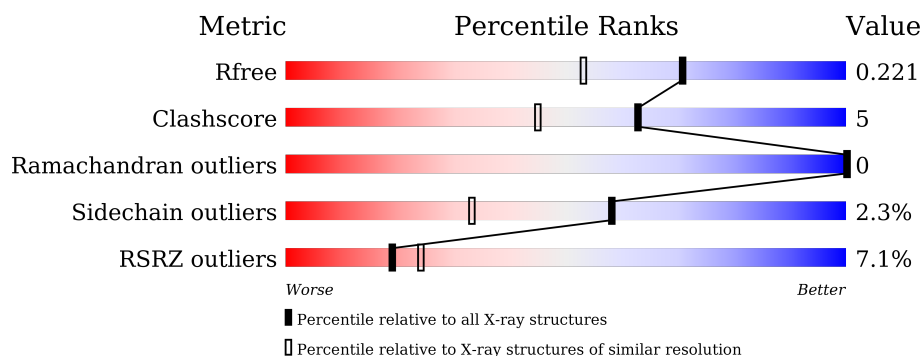
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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	240	<div> <div>3%</div> <div> <div></div> <div>76%</div> <div>10%</div> <div>13%</div> </div> </div>
1	B	240	<div> <div>2%</div> <div> <div></div> <div>80%</div> <div>6%</div> <div>13%</div> </div> </div>
1	C	240	<div> <div>10%</div> <div> <div></div> <div>77%</div> <div>5%</div> <div>16%</div> </div> </div>
1	D	240	<div> <div>10%</div> <div> <div></div> <div>75%</div> <div>8%</div> <div>16%</div> </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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	A	502	-	-	-	X
3	NAG	B	502[A]	-	-	-	X
3	NAG	B	502[B]	-	-	-	X
3	NAG	C	502	-	-	-	X
3	NAG	D	502	-	-	-	X
4	PEG	A	1242	-	-	X	X
4	PEG	A	1243	-	-	X	X
4	PEG	B	1244	-	-	-	X
5	CIT	B	1242	-	-	-	X
6	GOL	B	1243	-	-	-	X
6	GOL	D	1235	-	-	-	X
7	NA	C	1235	-	-	-	X
7	NA	C	1236	-	-	-	X

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 8031 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
1	A	209	Total	C	N	O	S	0	10	0
			1714	1100	273	334	7			
1	B	209	Total	C	N	O	S	0	11	0
			1717	1099	277	334	7			
1	C	201	Total	C	N	O	S	0	11	0
			1646	1061	265	314	6			
1	D	202	Total	C	N	O	S	0	5	0
			1622	1041	259	316	6			

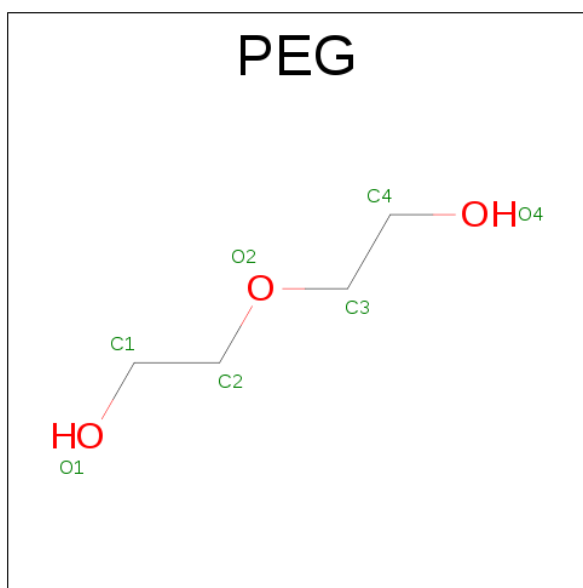
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

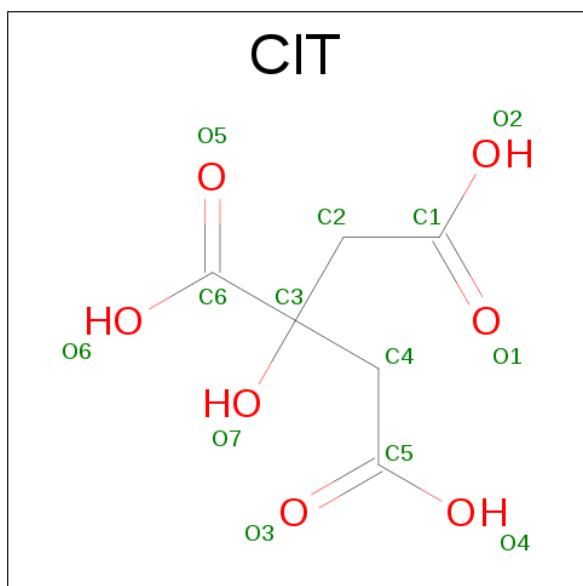
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			29	16	2	11		
3	B	2	Total	C	N	O	0	2
			58	32	4	22		
3	C	2	Total	C	N	O	0	0
			29	16	2	11		
3	D	2	Total	C	N	O	0	0
			29	16	2	11		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



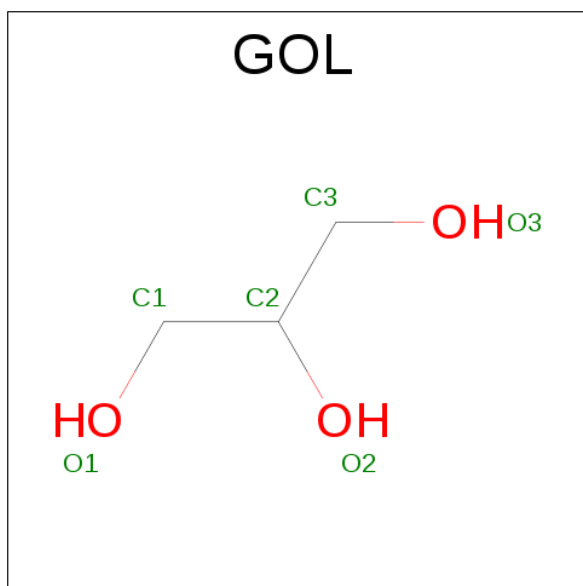
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			7	4	3		
4	A	1	Total	C	O	0	0
			7	4	3		
4	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 5 is CITRIC ACID (three-letter code: CIT) (formula:  $C_6H_8O_7$ ).



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

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	1	Total	Na	0	0
			1	1		
7	D	1	Total	Na	0	0
			1	1		
7	C	2	Total	Na	0	0
			2	2		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	297	Total	O	0	23
			320	320		

*Continued on next page...*

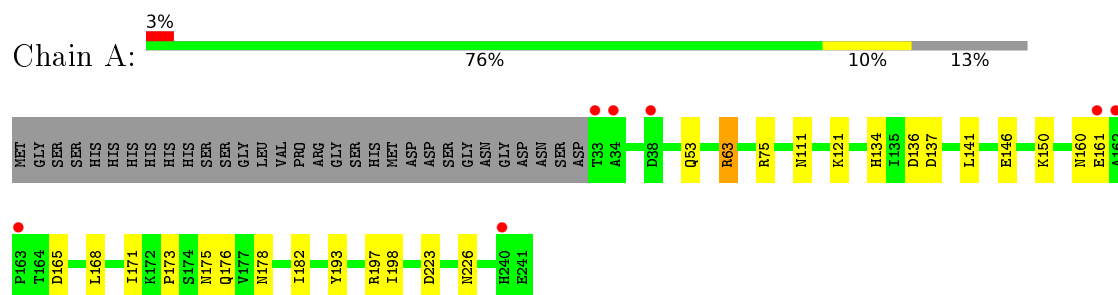
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	309	Total 325	O 325	0	16
8	C	231	Total 238	O 238	0	7
8	D	240	Total 250	O 250	0	10

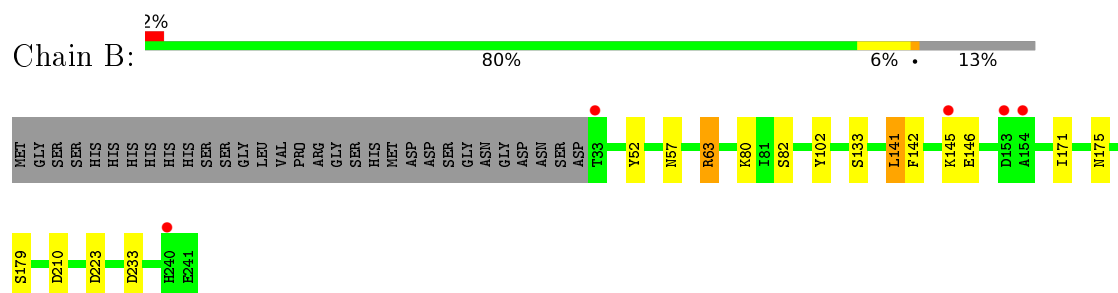
### 3 Residue-property plots

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 ( $RSRZ > 2$ ). 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.

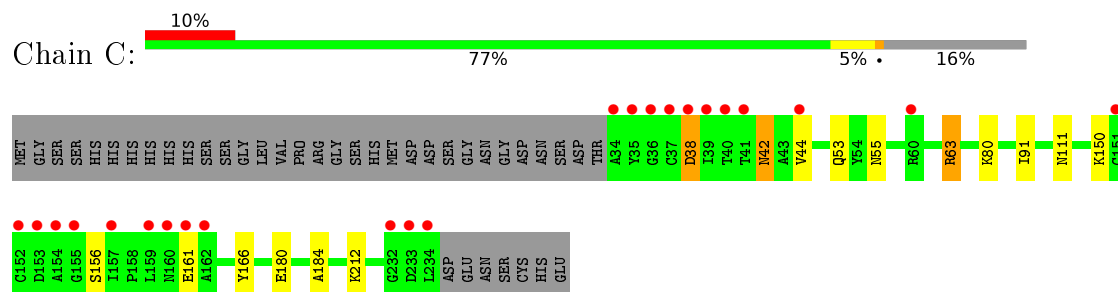
#### • Molecule 1: CEA1



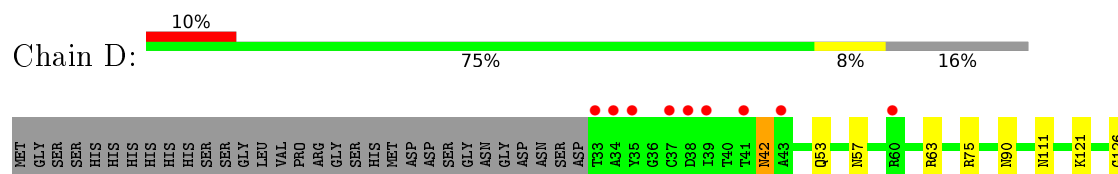
#### • Molecule 1: CEA1



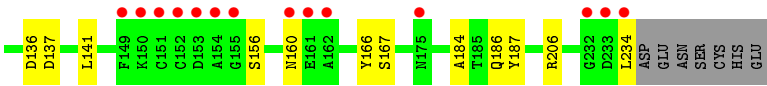
#### • Molecule 1: CEA1



#### • Molecule 1: CEA1







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.70Å 105.35Å 106.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 1.75 19.98 – 1.75	Depositor EDS
% Data completeness (in resolution range)	92.2 (19.99-1.75) 92.3 (19.98-1.75)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 1.74Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.183 , 0.211 0.191 , 0.221	Depositor DCC
$R_{free}$ test set	1619 reflections (1.54%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.3	Xtriage
Anisotropy	0.161	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 61.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.038 for -h,l,k 0.027 for -l,-k,-h 0.046 for k,h,-l 0.027 for k,l,h 0.027 for l,h,k	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8031	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NAG, NA, CA, CIT, PEG

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	0.65	0/1775	0.85	3/2411 (0.1%)
1	B	0.67	0/1783	0.85	2/2419 (0.1%)
1	C	0.63	1/1714 (0.1%)	0.79	1/2326 (0.0%)
1	D	0.61	0/1669	0.78	0/2269
All	All	0.64	1/6941 (0.0%)	0.82	6/9425 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	180	GLU	CG-CD	-5.05	1.44	1.51

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	180	GLU	OE1-CD-OE2	6.16	130.69	123.30
1	A	223	ASP	CB-CG-OD2	-5.58	113.28	118.30
1	A	146	GLU	N-CA-CB	-5.37	100.93	110.60
1	B	223	ASP	CB-CG-OD2	-5.32	113.51	118.30
1	B	80	LYS	CD-CE-NZ	5.16	123.56	111.70
1	A	197	ARG	NE-CZ-NH1	5.15	122.87	120.30

There are no chirality outliers.

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	1714	0	1656	21	0
1	B	1717	0	1673	20	0
1	C	1646	0	1626	12	0
1	D	1622	0	1568	15	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	29	0	25	0	0
3	B	58	0	50	3	0
3	C	29	0	25	0	0
3	D	29	0	25	0	0
4	A	14	0	20	8	0
4	B	7	0	10	3	0
5	B	13	0	5	1	0
6	B	6	0	8	3	0
6	D	6	0	8	2	0
7	B	1	0	0	0	0
7	C	2	0	0	0	0
7	D	1	0	0	0	0
8	A	320	0	0	9	0
8	B	325	0	0	3	0
8	C	238	0	0	4	0
8	D	250	0	0	3	0
All	All	8031	0	6699	72	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 5.

All (72) 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:52:TYR:OH	5:B:1242:CIT:O4	1.88	0.90
4:A:1243:PEG:H31	8:A:2462:HOH:O	1.72	0.90
1:B:141[A]:LEU:HD22	1:B:142:PHE:N	1.95	0.80
1:C:91:ILE:HD11	8:C:2546:HOH:O	1.82	0.78
1:D:42:ASN:HD22	1:D:42:ASN:H	1.32	0.76
1:C:42:ASN:H	1:C:42:ASN:HD22	1.34	0.75
4:A:1243:PEG:H22	8:A:2612:HOH:O	1.89	0.72
1:C:63[A]:ARG:NH1	8:C:2401:HOH:O	2.21	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:175:ASN:OD1	8:B:2401:HOH:O	2.08	0.70
1:A:63[A]:ARG:HG3	1:A:171:ILE:HG21	1.74	0.69
1:A:63[B]:ARG:HG3	1:A:171:ILE:HD13	1.74	0.69
1:B:63[B]:ARG:HH11	1:B:63[B]:ARG:HG3	1.58	0.69
1:C:53:GLN:HE22	1:C:55:ASN:HD22	1.42	0.65
1:A:226:ASN:HB3	4:A:1242:PEG:C4	2.27	0.64
1:B:233:ASP:H	6:B:1243:GOL:C3	2.10	0.63
1:C:42:ASN:HD21	1:C:156:SER:H	1.47	0.62
3:B:501[B]:NAG:C1	3:B:502[B]:NAG:H61	2.29	0.62
1:D:42:ASN:HD21	1:D:156:SER:H	1.48	0.62
1:D:160:ASN:ND2	8:D:2401:HOH:O	2.34	0.61
1:C:161:GLU:OE1	1:C:161:GLU:N	2.33	0.61
1:A:134:HIS:HE1	8:A:2618:HOH:O	1.82	0.60
1:B:63[B]:ARG:HG2	1:B:171:ILE:HG21	1.82	0.60
1:C:38:ASP:OD1	1:C:38:ASP:N	2.32	0.59
1:B:233:ASP:H	6:B:1243:GOL:H32	1.70	0.57
1:B:102:TYR:HB3	4:B:1244:PEG:H21	1.88	0.55
1:A:173:PRO:HB2	1:A:176:GLN:HG2	1.88	0.55
1:A:134:HIS:HD2	1:A:178:ASN:O	1.90	0.54
1:D:90:ASN:O	6:D:1235:GOL:H11	2.08	0.54
1:A:160:ASN:HA	8:A:2559[B]:HOH:O	2.08	0.52
1:A:226:ASN:HB3	4:A:1242:PEG:H42	1.92	0.51
1:A:226:ASN:HB3	4:A:1242:PEG:H41	1.92	0.50
1:A:75:ARG:CZ	4:A:1243:PEG:H11	2.42	0.49
1:B:145[A]:LYS:HZ2	1:B:146:GLU:H	1.61	0.48
1:B:63[C]:ARG:HG3	1:B:171:ILE:HG21	1.95	0.47
1:A:141[A]:LEU:C	1:A:141[A]:LEU:HD23	2.35	0.47
1:A:161:GLU:N	8:A:2559[B]:HOH:O	2.48	0.46
1:A:141[B]:LEU:HD12	1:A:168:LEU:HD12	1.97	0.46
1:D:57:ASN:HB3	8:D:2537:HOH:O	2.14	0.46
1:D:42:ASN:ND2	1:D:42:ASN:H	2.09	0.46
1:B:233:ASP:H	6:B:1243:GOL:H31	1.80	0.46
1:B:57:ASN:HB3	8:B:2607:HOH:O	2.17	0.45
1:B:63[B]:ARG:NH1	1:B:63[B]:ARG:HG3	2.28	0.45
1:D:75:ARG:HD3	8:D:2559:HOH:O	2.15	0.45
1:B:82:SER:O	4:B:1244:PEG:C1	2.65	0.45
3:B:502[A]:NAG:H61	8:B:2432:HOH:O	2.17	0.45
1:B:175:ASN:HB2	3:B:502[B]:NAG:H82	1.98	0.44
4:A:1242:PEG:H42	8:A:2482:HOH:O	2.17	0.44
1:D:166:TYR:CE1	1:D:184:ALA:CB	3.00	0.44
1:A:150:LYS:NZ	8:A:2608[B]:HOH:O	2.52	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:53:GLN:HE22	1:D:75:ARG:NH2	2.16	0.43
1:B:141[A]:LEU:CD2	1:B:142:PHE:N	2.76	0.43
1:D:126:GLY:HA2	1:D:187:TYR:CE2	2.54	0.43
1:B:133:SER:O	1:B:210:ASP:HB3	2.18	0.43
1:B:102:TYR:HA	4:B:1244:PEG:H32	2.00	0.43
1:D:136:ASP:HA	1:D:137:ASP:HA	1.87	0.43
1:A:141[B]:LEU:HD23	1:A:198:ILE:HG13	2.01	0.42
1:A:175:ASN:ND2	8:A:2405:HOH:O	2.43	0.42
1:A:182:ILE:HG22	1:C:184:ALA:HA	2.01	0.42
1:B:179:SER:OG	1:D:186[A]:GLN:HG3	2.20	0.42
1:C:212[A]:LYS:CE	8:C:2434:HOH:O	2.67	0.42
1:A:121:LYS:HD2	1:A:193:TYR:CZ	2.55	0.42
4:A:1243:PEG:C3	8:A:2462:HOH:O	2.49	0.42
1:C:166:TYR:CE1	1:C:184:ALA:CB	3.04	0.41
1:D:53:GLN:HE22	1:D:75:ARG:HH21	1.69	0.41
1:C:212[A]:LYS:HE3	8:C:2434:HOH:O	2.20	0.41
1:D:141:LEU:O	1:D:167:SER:HA	2.21	0.41
1:C:212[B]:LYS:HE2	1:C:212[B]:LYS:HB2	1.58	0.40
1:D:206:ARG:NH1	6:D:1235:GOL:O1	2.43	0.40
1:A:136:ASP:HA	1:A:137:ASP:HA	1.85	0.40
1:A:53:GLN:HE22	1:A:75:ARG:HH11	1.70	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	219/240 (91%)	215 (98%)	4 (2%)	0	100	100
1	C	210/240 (88%)	205 (98%)	5 (2%)	0	100	100
1	D	205/240 (85%)	202 (98%)	3 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	851/960 (89%)	836 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	191/207 (92%)	188 (98%)	3 (2%)	70	52
1	B	192/207 (93%)	187 (97%)	5 (3%)	54	28
1	C	184/207 (89%)	176 (96%)	8 (4%)	35	12
1	D	179/207 (86%)	174 (97%)	5 (3%)	51	25
All	All	746/828 (90%)	725 (97%)	21 (3%)	58	25

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	63[A]	ARG
1	A	63[B]	ARG
1	A	111	ASN
1	B	63[A]	ARG
1	B	63[B]	ARG
1	B	63[C]	ARG
1	B	141[A]	LEU
1	B	141[B]	LEU
1	C	38	ASP
1	C	42	ASN
1	C	44	VAL
1	C	63[A]	ARG
1	C	63[B]	ARG
1	C	80	LYS
1	C	111	ASN
1	C	150	LYS
1	D	42	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	63	ARG
1	D	111	ASN
1	D	121	LYS
1	D	234	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	53	GLN
1	A	55	ASN
1	A	57	ASN
1	A	111	ASN
1	A	134	HIS
1	A	175	ASN
1	B	53	GLN
1	B	55	ASN
1	B	175	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	160	ASN
1	D	176	GLN
1	D	226	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 ⓘ

10 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	A	501	3,2	14,14,15	0.60	0	15,19,21	1.09	1 (6%)
3	NAG	A	502	3	15,15,15	0.47	0	17,21,21	0.61	0
3	NAG	B	501[A]	3,2	14,14,15	0.38	0	15,19,21	0.59	0
3	NAG	B	501[B]	3,2	14,14,15	0.42	0	15,19,21	0.56	0
3	NAG	B	502[A]	3	15,15,15	0.49	0	17,21,21	1.77	3 (17%)
3	NAG	B	502[B]	3	15,15,15	0.53	0	17,21,21	1.67	6 (35%)
3	NAG	C	501	3,2	14,14,15	0.53	0	15,19,21	0.81	0
3	NAG	C	502	3	15,15,15	0.68	0	17,21,21	1.01	1 (5%)
3	NAG	D	501	3,2	14,14,15	0.43	0	15,19,21	1.01	0
3	NAG	D	502	3	15,15,15	0.46	0	17,21,21	1.01	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
3	NAG	A	501	3,2	-	0/6/23/26	0/1/1/1
3	NAG	A	502	3	-	0/6/26/26	0/1/1/1
3	NAG	B	501[A]	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	501[B]	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	502[A]	3	-	0/6/26/26	0/1/1/1
3	NAG	B	502[B]	3	-	0/6/26/26	0/1/1/1
3	NAG	C	501	3,2	-	0/6/23/26	0/1/1/1
3	NAG	C	502	3	-	0/6/26/26	0/1/1/1
3	NAG	D	501	3,2	-	0/6/23/26	0/1/1/1
3	NAG	D	502	3	-	0/6/26/26	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	502[B]	NAG	C4-C3-C2	-3.45	105.13	110.37

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	502[A]	NAG	C6-C5-C4	-3.36	104.57	112.99
3	C	502	NAG	O4-C4-C3	-2.81	104.02	110.36
3	A	501	NAG	C1-O5-C5	-2.02	109.17	112.14
3	B	502[B]	NAG	C6-C5-C4	-2.02	107.93	112.99
3	B	502[B]	NAG	O3-C3-C2	2.09	113.91	109.57
3	B	502[B]	NAG	C1-O5-C5	2.12	117.59	113.54
3	B	502[A]	NAG	C3-C4-C5	2.18	114.11	110.23
3	B	502[B]	NAG	C3-C2-N2	2.53	115.94	110.67
3	B	502[B]	NAG	O5-C5-C4	2.88	115.16	109.67
3	B	502[A]	NAG	O5-C5-C4	4.80	118.83	109.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	501[B]	NAG	1	0
3	B	502[A]	NAG	1	0
3	B	502[B]	NAG	2	0

## 5.6 Ligand geometry

Of 14 ligands modelled in this entry, 8 are monoatomic - leaving 6 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
4	PEG	A	1242	-	6,6,6	0.64	0	5,5,5	0.85	0
4	PEG	A	1243	-	6,6,6	0.49	0	5,5,5	1.23	1 (20%)
5	CIT	B	1242	-	3,12,12	2.17	1 (33%)	3,17,17	3.58	2 (66%)
6	GOL	B	1243	-	5,5,5	0.39	0	5,5,5	0.65	0
4	PEG	B	1244	-	6,6,6	0.45	0	5,5,5	0.66	0
6	GOL	D	1235	-	5,5,5	0.10	0	5,5,5	0.39	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
4	PEG	A	1242	-	-	0/4/4/4	0/0/0/0
4	PEG	A	1243	-	-	0/4/4/4	0/0/0/0
5	CIT	B	1242	-	-	0/6/16/16	0/0/0/0
6	GOL	B	1243	-	-	0/4/4/4	0/0/0/0
4	PEG	B	1244	-	-	0/4/4/4	0/0/0/0
6	GOL	D	1235	-	-	0/4/4/4	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1242	CIT	C4-C3	-3.40	1.49	1.54

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1242	CIT	C3-C2-C1	-3.48	109.52	114.95
4	A	1243	PEG	C3-O2-C2	2.10	122.27	113.31
5	B	1242	CIT	C4-C3-C2	4.79	121.74	109.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1242	PEG	4	0
4	A	1243	PEG	4	0
5	B	1242	CIT	1	0
6	B	1243	GOL	3	0
4	B	1244	PEG	3	0
6	D	1235	GOL	2	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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%)	0.14	7 (3%) 50 56	7, 10, 25, 50	0
1	B	209/240 (87%)	0.03	5 (2%) 62 69	6, 10, 22, 40	0
1	C	201/240 (83%)	0.54	23 (11%) 7 9	7, 12, 50, 72	0
1	D	202/240 (84%)	0.53	23 (11%) 7 9	7, 12, 47, 62	0
All	All	821/960 (85%)	0.31	58 (7%) 19 24	6, 11, 39, 72	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	234	LEU	8.7
1	D	33	THR	7.5
1	C	233	ASP	7.5
1	A	33	THR	7.5
1	D	234	LEU	7.4
1	A	162	ALA	7.4
1	C	152	CYS	7.3
1	D	233	ASP	6.8
1	C	39	ILE	6.6
1	D	154	ALA	6.1
1	C	154	ALA	6.0
1	C	153	ASP	5.9
1	D	160	ASN	5.9
1	C	151	CYS	5.7
1	D	161	GLU	5.7
1	D	153	ASP	5.6
1	D	38	ASP	5.5
1	C	38	ASP	5.5
1	D	39	ILE	5.4
1	D	152	CYS	5.3
1	C	160	ASN	4.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	41	THR	4.6
1	C	40	THR	4.5
1	C	232	GLY	4.5
1	B	33	THR	4.5
1	D	34	ALA	4.4
1	C	36	GLY	4.4
1	C	34	ALA	4.3
1	C	60	ARG	3.8
1	D	35	TYR	3.8
1	D	150	LYS	3.5
1	C	161	GLU	3.4
1	C	35	TYR	3.4
1	D	155	GLY	3.3
1	A	161	GLU	3.3
1	C	37	CYS	3.3
1	A	163	PRO	3.2
1	D	162	ALA	3.1
1	A	240	HIS	3.1
1	B	240	HIS	3.0
1	C	44	VAL	2.8
1	D	151	CYS	2.8
1	D	41	THR	2.8
1	B	154	ALA	2.7
1	D	232	GLY	2.6
1	C	162	ALA	2.5
1	D	37	CYS	2.5
1	C	157	ILE	2.5
1	C	159	LEU	2.4
1	D	175	ASN	2.4
1	D	43	ALA	2.4
1	A	34	ALA	2.2
1	C	155	GLY	2.2
1	B	145[A]	LYS	2.2
1	B	153	ASP	2.1
1	D	149	PHE	2.1
1	D	60	ARG	2.1
1	A	38	ASP	2.0

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

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

### 6.3 Carbohydrates ⓘ

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
3	NAG	B	502[B]	15/15	0.79	0.32	7.85	13,18,24,26	15
3	NAG	B	502[A]	15/15	0.79	0.32	7.56	13,19,22,22	15
3	NAG	C	502	15/15	0.77	0.31	4.87	17,35,49,50	0
3	NAG	D	502	15/15	0.81	0.30	2.92	19,34,44,49	0
3	NAG	A	502	15/15	0.85	0.23	2.60	15,31,38,41	0
3	NAG	A	501	14/15	0.97	0.08	-0.61	9,10,11,12	0
3	NAG	D	501	14/15	0.97	0.08	-0.95	10,11,13,14	0
3	NAG	C	501	14/15	0.97	0.08	-1.00	8,10,12,13	0
3	NAG	B	501[B]	14/15	0.98	0.07	-1.61	8,9,9,10	14
3	NAG	B	501[A]	14/15	0.98	0.07	-1.61	8,9,10,11	14

### 6.4 Ligands ⓘ

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	CIT	B	1242	13/13	0.76	0.24	9.74	29,38,43,46	0
7	NA	C	1235	1/1	1.00	0.18	8.98	2,2,2,2	0
4	PEG	B	1244	7/7	0.90	0.18	8.01	24,26,28,28	0
7	NA	C	1236	1/1	0.99	0.14	7.06	12,12,12,12	0
6	GOL	D	1235	6/6	0.89	0.17	5.64	26,30,31,34	0
6	GOL	B	1243	6/6	0.83	0.17	5.29	26,29,30,37	0
4	PEG	A	1242	7/7	0.81	0.23	3.93	26,28,30,32	0
4	PEG	A	1243	7/7	0.88	0.17	2.00	21,24,29,31	0
2	CA	C	500	1/1	1.00	0.04	-3.19	8,8,8,8	0
2	CA	D	500	1/1	0.99	0.03	-3.47	8,8,8,8	0
2	CA	B	500	1/1	1.00	0.04	-4.24	7,7,7,7	0
2	CA	A	500	1/1	0.99	0.05	-4.41	7,7,7,7	0
7	NA	B	1245	1/1	0.97	0.15	-	10,10,10,10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	NA	D	1236	1/1	1.00	0.14	-	2,2,2,2	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.