



# Full wwPDB NMR Structure Validation Report ⓘ

May 28, 2020 – 10:56 pm BST

PDB ID : 2L8A  
Title : Structure of a novel CBM3 lacking the calcium-binding site  
Authors : Paiva, J.H.; Meza, A.N.; Sforca, M.L.; Navarro, R.Z.; Neves, J.L.; Santos, C.R.; Murakami, M.T.; Zeri, A.C.  
Deposited on : 2011-01-07

This is a Full wwPDB NMR 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/NMRValidationReportHelp>

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:

Cyrange : Kirchner and Güntert (2011)  
NmrClust : Kelley et al. (1996)  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : 2.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

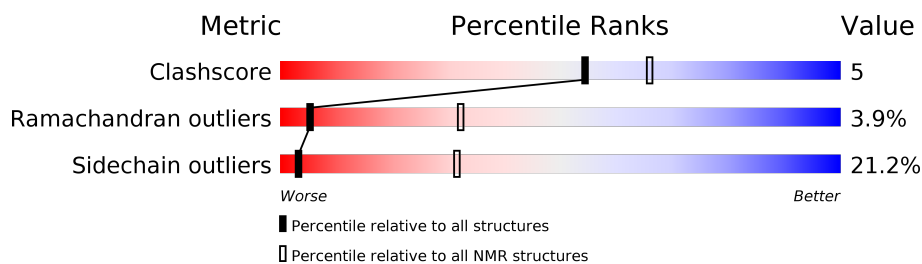
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 82%.

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)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	149	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:352-A:499 (148)	0.53	4

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20
2	9, 10

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2319 atoms, of which 1141 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Endoglucanase.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	149	2319	738	1141	208	228	4	0

There are 3 discrepancies between the modelled and reference sequences:

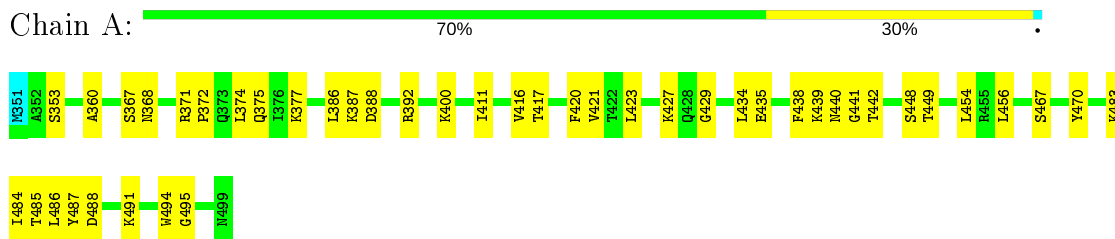
Chain	Residue	Modelled	Actual	Comment	Reference
A	351	MET	-	EXPRESSION TAG	UNP P10475
A	352	ALA	-	EXPRESSION TAG	UNP P10475
A	353	SER	-	EXPRESSION TAG	UNP P10475

## 4 Residue-property plots

### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Endoglucanase

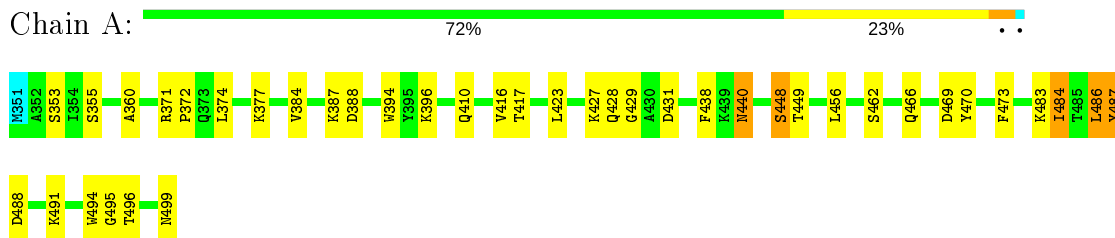


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1

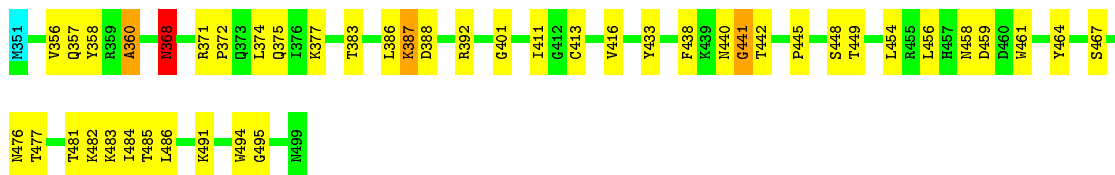
- Molecule 1: Endoglucanase



#### 4.2.2 Score per residue for model 2

- Molecule 1: Endoglucanase

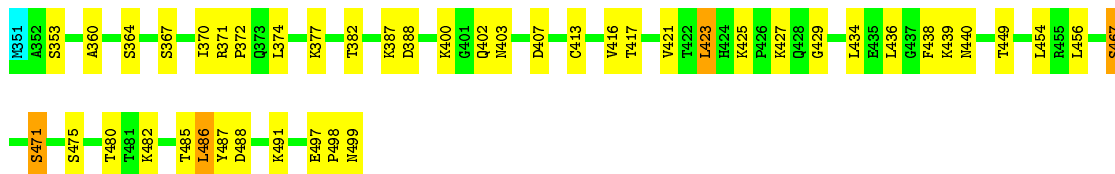




### 4.2.3 Score per residue for model 3

- Molecule 1: Endoglucanase

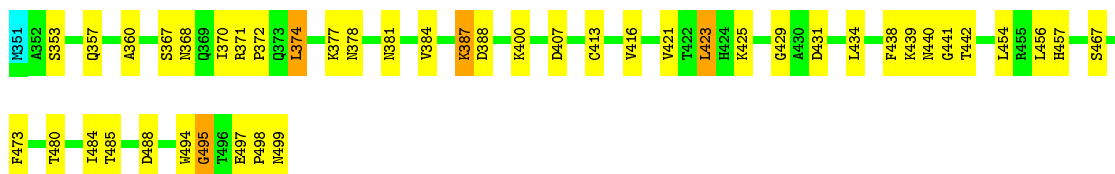
Chain A: 69% 28%



### 4.2.4 Score per residue for model 4 (medoid)

- Molecule 1: Endoglucanase

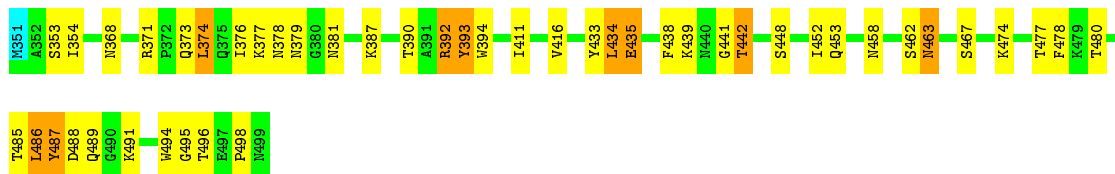
Chain A: 70% 27%



### 4.2.5 Score per residue for model 5

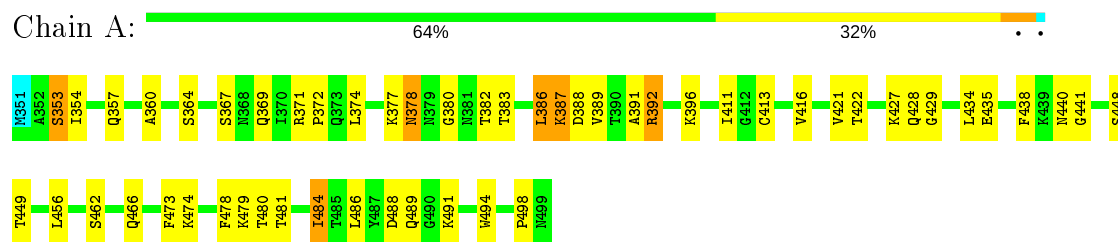
- Molecule 1: Endoglucanase

Chain A: 68% 25% 6%



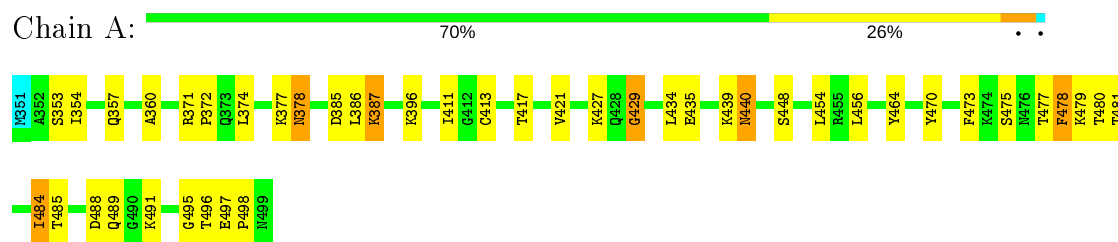
### 4.2.6 Score per residue for model 6

- Molecule 1: Endoglucanase



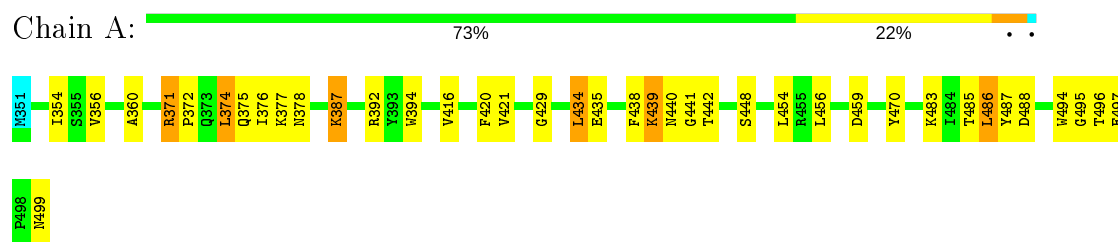
#### 4.2.7 Score per residue for model 7

- Molecule 1: Endoglucanase



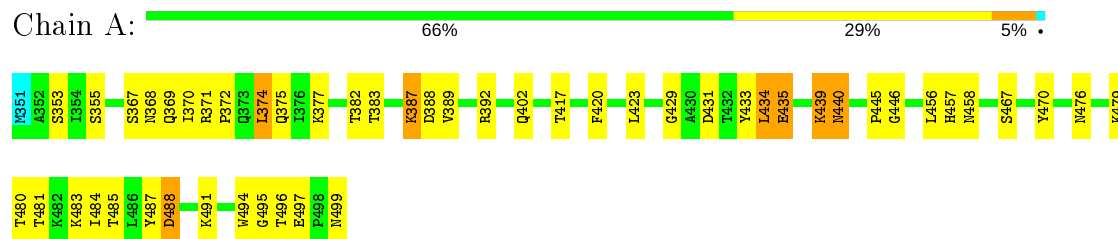
#### 4.2.8 Score per residue for model 8

- Molecule 1: Endoglucanase



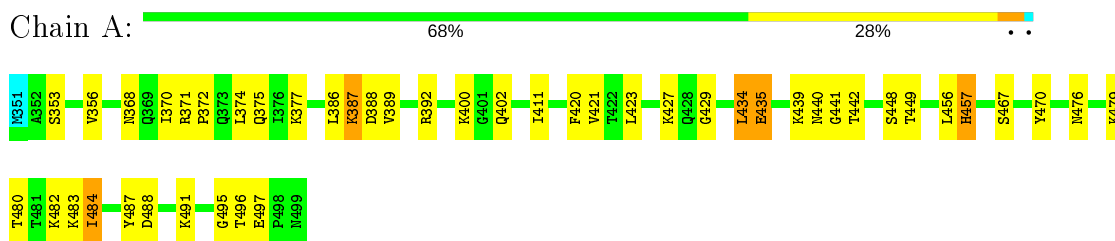
#### 4.2.9 Score per residue for model 9

- Molecule 1: Endoglucanase



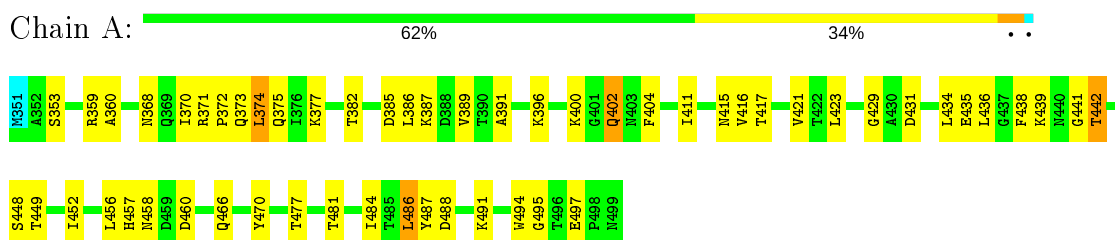
### 4.2.10 Score per residue for model 10

- Molecule 1: Endoglucanase



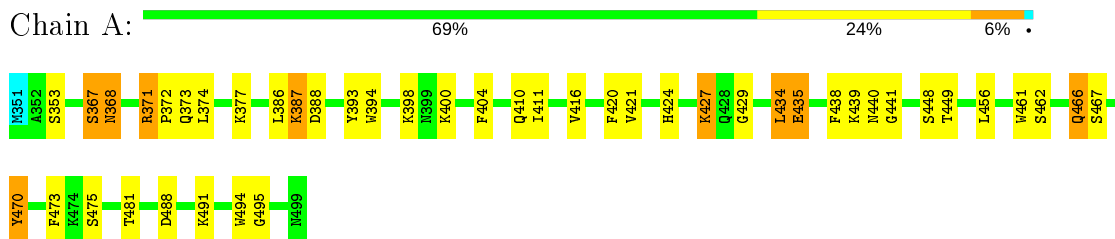
### 4.2.11 Score per residue for model 11

- Molecule 1: Endoglucanase



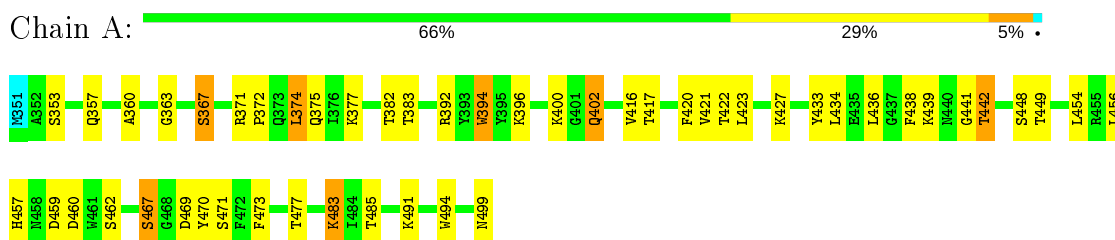
### 4.2.12 Score per residue for model 12

- Molecule 1: Endoglucanase



### 4.2.13 Score per residue for model 13

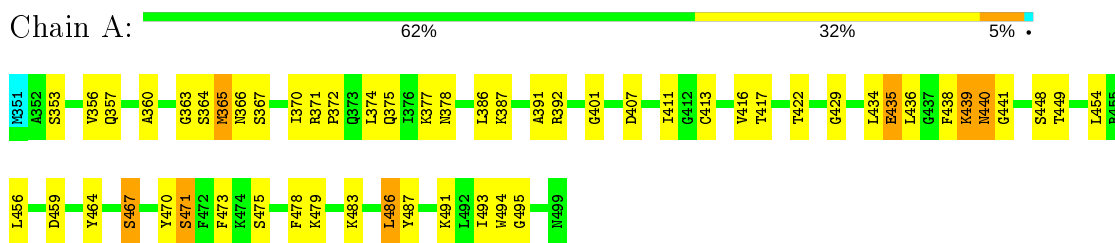
- Molecule 1: Endoglucanase





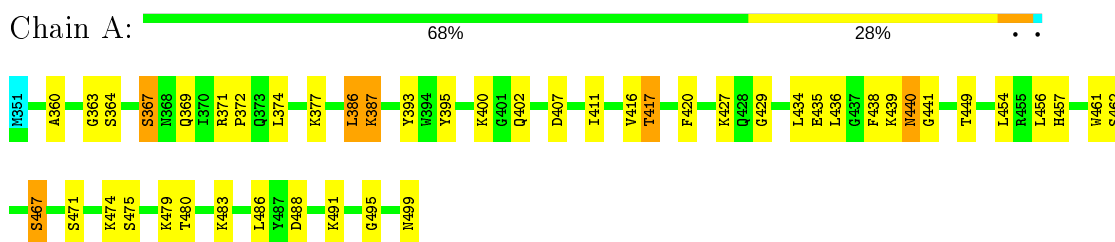
#### 4.2.14 Score per residue for model 14

- Molecule 1: Endoglucanase



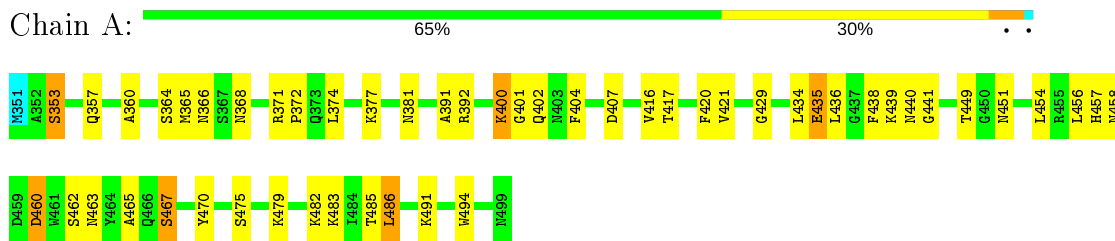
#### 4.2.15 Score per residue for model 15

- Molecule 1: Endoglucanase



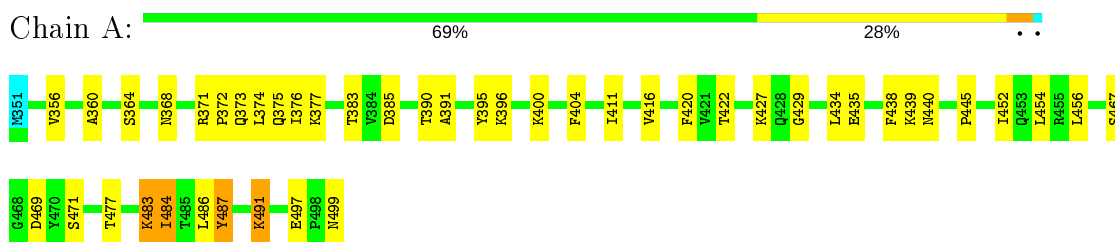
#### 4.2.16 Score per residue for model 16

- Molecule 1: Endoglucanase



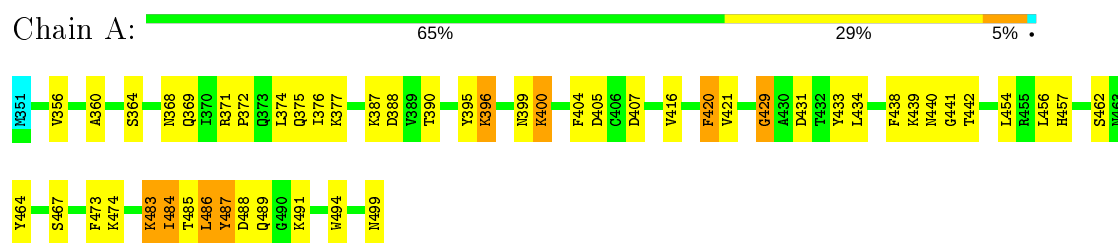
#### 4.2.17 Score per residue for model 17

- Molecule 1: Endoglucanase



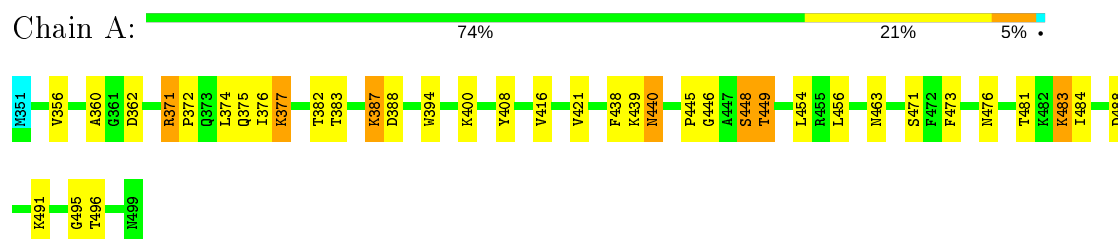
## 4.2.18 Score per residue for model 18

- Molecule 1: Endoglucanase



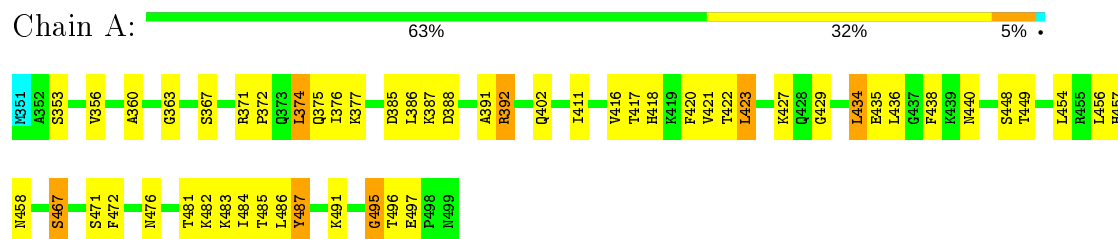
## 4.2.19 Score per residue for model 19

- Molecule 1: Endoglucanase



## 4.2.20 Score per residue for model 20

- Molecule 1: Endoglucanase



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics, molecular dynamics*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	structure solution	2.1
CNS	refinement	
CYANA	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	1647
Number of shifts mapped to atoms	1647
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	82%

No validations of the models with respect to experimental NMR restraints is performed at this time.

## 6 Model quality

### 6.1 Standard geometry

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1170	1132	1129	12±3
All	All	23400	22640	22580	248

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 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:420:PHE:HB3	1:A:434:LEU:HD23	0.76	1.58	9	5
1:A:416:VAL:HG22	1:A:438:PHE:HA	0.66	1.67	8	17
1:A:372:PRO:HD2	1:A:456:LEU:HD13	0.64	1.69	8	19
1:A:382:THR:HG23	1:A:383:THR:HG23	0.62	1.71	19	3
1:A:386:LEU:HD13	1:A:389:VAL:HB	0.62	1.71	11	2
1:A:386:LEU:HG	1:A:411:ILE:HG21	0.61	1.70	2	8
1:A:486:LEU:HD22	1:A:494:TRP:CE3	0.61	2.30	8	7
1:A:383:THR:HG22	1:A:445:PRO:HB3	0.60	1.71	9	4
1:A:391:ALA:O	1:A:435:GLU:HB3	0.58	1.99	20	4
1:A:368:ASN:HB2	1:A:458:ASN:O	0.58	1.99	11	2
1:A:467:SER:HA	1:A:470:TYR:CE2	0.57	2.35	13	1
1:A:401:GLY:HA2	1:A:458:ASN:OD1	0.56	1.99	2	2
1:A:356:VAL:HA	1:A:375:GLN:O	0.56	2.00	19	8
1:A:483:LYS:O	1:A:484:ILE:HG12	0.56	2.00	17	2

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:402:GLN:HG3	1:A:457:HIS:O	0.56	2.01	11	4
1:A:368:ASN:OD1	1:A:463:ASN:HA	0.55	2.01	5	1
1:A:491:LYS:NZ	1:A:491:LYS:HB2	0.55	2.17	17	1
1:A:396:LYS:HB2	1:A:429:GLY:O	0.55	2.01	6	3
1:A:484:ILE:HG22	1:A:485:THR:H	0.54	1.62	2	1
1:A:387:LYS:O	1:A:439:LYS:HA	0.54	2.03	18	1
1:A:480:THR:OG1	1:A:498:PRO:HB3	0.53	2.04	4	3
1:A:417:THR:O	1:A:436:LEU:HA	0.52	2.04	15	5
1:A:487:TYR:HB2	1:A:491:LYS:O	0.52	2.05	20	1
1:A:354:ILE:HD13	1:A:378:ASN:HD22	0.52	1.64	5	2
1:A:466:GLN:O	1:A:470:TYR:HB2	0.52	2.04	1	2
1:A:354:ILE:HD12	1:A:378:ASN:HD22	0.51	1.64	6	2
1:A:441:GLY:O	1:A:442:THR:HB	0.51	2.05	11	8
1:A:402:GLN:HB3	1:A:457:HIS:O	0.51	2.06	15	3
1:A:434:LEU:HD22	1:A:435:GLU:N	0.50	2.21	9	4
1:A:467:SER:O	1:A:471:SER:HB2	0.50	2.06	3	5
1:A:392:ARG:HD3	1:A:435:GLU:OE2	0.50	2.05	20	1
1:A:373:GLN:HB3	1:A:453:GLN:CD	0.50	2.26	5	1
1:A:390:THR:HB	1:A:487:TYR:CD2	0.50	2.41	5	3
1:A:420:PHE:CB	1:A:434:LEU:HG	0.49	2.37	17	3
1:A:484:ILE:HG22	1:A:485:THR:N	0.49	2.23	20	2
1:A:423:LEU:HD12	1:A:425:LYS:HB2	0.49	1.84	3	2
1:A:388:ASP:OD2	1:A:489:GLN:HA	0.49	2.07	18	1
1:A:387:LYS:HA	1:A:440:ASN:N	0.49	2.23	18	9
1:A:495:GLY:O	1:A:496:THR:HB	0.48	2.09	20	1
1:A:364:SER:O	1:A:467:SER:HA	0.47	2.09	14	1
1:A:394:TRP:HB3	1:A:433:TYR:CD2	0.47	2.44	5	1
1:A:391:ALA:O	1:A:392:ARG:HD3	0.47	2.10	6	1
1:A:411:ILE:HG13	1:A:452:ILE:HD11	0.47	1.86	17	1
1:A:367:SER:O	1:A:369:GLN:HG2	0.47	2.09	15	1
1:A:411:ILE:HB	1:A:452:ILE:HD11	0.46	1.86	5	1
1:A:374:LEU:N	1:A:374:LEU:HD13	0.46	2.25	20	4
1:A:389:VAL:HG13	1:A:488:ASP:OD1	0.46	2.11	10	2
1:A:486:LEU:HG	1:A:487:TYR:N	0.46	2.25	1	5
1:A:481:THR:HG22	1:A:494:TRP:CZ2	0.46	2.45	12	2
1:A:376:ILE:C	1:A:377:LYS:HG3	0.46	2.28	19	1
1:A:387:LYS:HA	1:A:439:LYS:C	0.46	2.30	7	7
1:A:356:VAL:HG13	1:A:483:LYS:HE2	0.46	1.86	17	1
1:A:399:ASN:O	1:A:400:LYS:HG2	0.46	2.10	18	1
1:A:434:LEU:HD23	1:A:435:GLU:N	0.46	2.25	11	2
1:A:460:ASP:OD2	1:A:463:ASN:HA	0.46	2.11	16	1

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:353:SER:HB2	1:A:381:ASN:OD1	0.46	2.11	16	1
1:A:448:SER:O	1:A:449:THR:HB	0.46	2.10	19	1
1:A:357:GLN:HG2	1:A:358:TYR:N	0.45	2.26	2	1
1:A:439:LYS:O	1:A:440:ASN:HB2	0.45	2.11	9	1
1:A:483:LYS:HD3	1:A:483:LYS:N	0.45	2.27	19	1
1:A:466:GLN:HA	1:A:469:ASP:OD1	0.45	2.11	1	1
1:A:374:LEU:HD13	1:A:374:LEU:N	0.44	2.27	4	3
1:A:483:LYS:HB3	1:A:483:LYS:NZ	0.44	2.28	18	1
1:A:480:THR:OG1	1:A:498:PRO:HA	0.44	2.13	5	1
1:A:420:PHE:HB3	1:A:434:LEU:HG	0.44	1.90	13	3
1:A:368:ASN:CG	1:A:460:ASP:HB2	0.44	2.33	16	1
1:A:485:THR:HA	1:A:495:GLY:HA2	0.44	1.88	4	1
1:A:485:THR:HA	1:A:494:TRP:O	0.44	2.13	18	4
1:A:417:THR:O	1:A:436:LEU:HD23	0.44	2.12	13	4
1:A:384:VAL:O	1:A:448:SER:HB3	0.43	2.13	1	1
1:A:368:ASN:HA	1:A:458:ASN:HB3	0.43	1.88	2	1
1:A:484:ILE:O	1:A:494:TRP:HB3	0.43	2.14	6	1
1:A:366:ASN:OD1	1:A:465:ALA:HA	0.43	2.13	16	1
1:A:411:ILE:CG1	1:A:452:ILE:HD11	0.43	2.42	17	1
1:A:378:ASN:OD1	1:A:381:ASN:HB2	0.43	2.14	4	2
1:A:374:LEU:CD2	1:A:452:ILE:HB	0.43	2.44	11	1
1:A:459:ASP:OD2	1:A:461:TRP:HB2	0.43	2.14	2	1
1:A:391:ALA:HA	1:A:486:LEU:HB2	0.43	1.91	17	1
1:A:378:ASN:N	1:A:449:THR:HG22	0.43	2.28	14	1
1:A:365:MET:O	1:A:366:ASN:HB2	0.42	2.14	14	1
1:A:367:SER:HA	1:A:464:TYR:O	0.42	2.13	14	1
1:A:365:MET:HG2	1:A:467:SER:OG	0.42	2.13	16	1
1:A:482:LYS:HB3	1:A:494:TRP:CD1	0.42	2.49	2	1
1:A:475:SER:HA	1:A:478:PHE:CE2	0.42	2.48	7	1
1:A:386:LEU:HD13	1:A:389:VAL:CB	0.42	2.43	11	1
1:A:392:ARG:HA	1:A:435:GLU:HG2	0.42	1.91	5	1
1:A:478:PHE:CD2	1:A:498:PRO:HG2	0.42	2.50	7	1
1:A:398:LYS:HB3	1:A:462:SER:OG	0.42	2.15	12	1
1:A:411:ILE:O	1:A:411:ILE:HG22	0.42	2.15	20	2
1:A:423:LEU:HA	1:A:423:LEU:HD22	0.42	1.79	20	1
1:A:394:TRP:HB3	1:A:433:TYR:CD1	0.41	2.50	13	1
1:A:386:LEU:O	1:A:438:PHE:HB2	0.41	2.14	20	1
1:A:390:THR:HB	1:A:487:TYR:CE2	0.41	2.49	5	1
1:A:393:TYR:N	1:A:435:GLU:HG3	0.41	2.30	12	1
1:A:360:ALA:HB3	1:A:476:ASN:OD1	0.41	2.15	2	1
1:A:353:SER:CB	1:A:380:GLY:HA3	0.41	2.46	6	1

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:367:SER:O	1:A:368:ASN:HB3	0.41	2.15	12	1
1:A:423:LEU:O	1:A:431:ASP:HB2	0.41	2.15	1	1
1:A:393:TYR:HB3	1:A:434:LEU:O	0.41	2.16	5	1
1:A:396:LYS:HE2	1:A:429:GLY:O	0.41	2.15	11	2
1:A:401:GLY:O	1:A:459:ASP:HB2	0.41	2.16	14	1
1:A:404:PHE:CE1	1:A:456:LEU:HG	0.41	2.51	17	2
1:A:357:GLN:C	1:A:483:LYS:HE3	0.40	2.36	13	1
1:A:354:ILE:CD1	1:A:378:ASN:HD22	0.40	2.30	8	1

## 6.3 Torsion angles [i](#)

### 6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	147/149 (99%)	113±3 (77±2%)	29±3 (20±2%)	6±2 (4±1%)	5	32
All	All	2940/2980 (99%)	2250 (77%)	574 (20%)	116 (4%)	5	32

All 24 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	360	ALA	16
1	A	495	GLY	14
1	A	429	GLY	13
1	A	484	ILE	9
1	A	427	LYS	8
1	A	440	ASN	8
1	A	441	GLY	6
1	A	413	CYS	6
1	A	368	ASN	6
1	A	449	THR	5
1	A	363	GLY	4
1	A	371	ARG	3
1	A	442	THR	3
1	A	446	GLY	2

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	367	SER	2
1	A	476	ASN	2
1	A	460	ASP	2
1	A	461	TRP	1
1	A	498	PRO	1
1	A	400	LYS	1
1	A	493	ILE	1
1	A	365	MET	1
1	A	362	ASP	1
1	A	435	GLU	1

### 6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	125/126 (99%)	99±3 (79±3%)	26±3 (21±3%)	3	31
All	All	2500/2520 (99%)	1971 (79%)	529 (21%)	3	31

All 92 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	377	LYS	20
1	A	371	ARG	20
1	A	374	LEU	20
1	A	491	LYS	17
1	A	387	LYS	14
1	A	353	SER	14
1	A	467	SER	14
1	A	448	SER	13
1	A	434	LEU	13
1	A	454	LEU	13
1	A	488	ASP	13
1	A	421	VAL	13
1	A	483	LYS	13
1	A	439	LYS	12
1	A	486	LEU	12

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	400	LYS	11
1	A	392	ARG	10
1	A	388	ASP	10
1	A	473	PHE	9
1	A	497	GLU	9
1	A	435	GLU	9
1	A	499	ASN	9
1	A	487	TYR	8
1	A	470	TYR	8
1	A	496	THR	7
1	A	423	LEU	7
1	A	479	LYS	7
1	A	367	SER	7
1	A	462	SER	7
1	A	449	THR	7
1	A	370	ILE	6
1	A	477	THR	6
1	A	481	THR	6
1	A	364	SER	6
1	A	407	ASP	6
1	A	394	TRP	5
1	A	376	ILE	5
1	A	422	THR	5
1	A	357	GLN	5
1	A	475	SER	5
1	A	440	ASN	5
1	A	485	THR	5
1	A	480	THR	4
1	A	417	THR	4
1	A	474	LYS	4
1	A	478	PHE	4
1	A	385	ASP	4
1	A	482	LYS	4
1	A	471	SER	4
1	A	484	ILE	4
1	A	431	ASP	4
1	A	427	LYS	3
1	A	433	TYR	3
1	A	466	GLN	3
1	A	395	TYR	3
1	A	457	HIS	3
1	A	396	LYS	3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	375	GLN	3
1	A	489	GLN	3
1	A	373	GLN	3
1	A	464	TYR	3
1	A	382	THR	3
1	A	402	GLN	3
1	A	369	GLN	3
1	A	404	PHE	3
1	A	420	PHE	3
1	A	386	LEU	2
1	A	469	ASP	2
1	A	410	GLN	2
1	A	378	ASN	2
1	A	355	SER	2
1	A	459	ASP	2
1	A	463	ASN	2
1	A	476	ASN	2
1	A	393	TYR	2
1	A	458	ASN	2
1	A	428	GLN	2
1	A	460	ASP	1
1	A	403	ASN	1
1	A	415	ASN	1
1	A	408	TYR	1
1	A	472	PHE	1
1	A	451	ASN	1
1	A	359	ARG	1
1	A	461	TRP	1
1	A	424	HIS	1
1	A	418	HIS	1
1	A	384	VAL	1
1	A	368	ASN	1
1	A	405	ASP	1
1	A	379	ASN	1
1	A	383	THR	1

### 6.3.3 RNA

There are no RNA molecules in this entry.

## 6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 6.7 Other polymers [i](#)

There are no such molecules in this entry.

## 6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 7 Chemical shift validation [i](#)

The completeness of assignment taking into account all chemical shift lists is 82% for the well-defined parts and 82% for the entire structure.

### 7.1 Chemical shift list 1

File name: input\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 7.1.1 Bookkeeping [i](#)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1647
Number of shifts mapped to atoms	1647
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

#### 7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	149	-0.10 $\pm$ 0.12	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	135	-0.23 $\pm$ 0.19	None needed (< 0.5 ppm)
$^{13}\text{C}'$	138	0.22 $\pm$ 0.08	None needed (< 0.5 ppm)
$^{15}\text{N}$	144	0.44 $\pm$ 0.57	None needed (< 0.5 ppm)

#### 7.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 82%, i.e. 1476 atoms were assigned a chemical shift out of a possible 1806. 10 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	720/732 (98%)	291/292 (100%)	285/296 (96%)	144/144 (100%)
Sidechain	664/896 (74%)	417/526 (79%)	247/320 (77%)	0/50 (0%)

*Continued on next page...*

Continued from previous page...

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Aromatic	92/178 (52%)	53/92 (58%)	36/77 (47%)	3/9 (33%)
Overall	1476/1806 (82%)	761/910 (84%)	568/693 (82%)	147/203 (72%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 82%, i.e. 1484 atoms were assigned a chemical shift out of a possible 1819. 10 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Backbone	723/737 (98%)	292/294 (99%)	287/298 (96%)	144/145 (99%)
Sidechain	669/904 (74%)	421/531 (79%)	248/323 (77%)	0/50 (0%)
Aromatic	92/178 (52%)	53/92 (58%)	36/77 (47%)	3/9 (33%)
Overall	1484/1819 (82%)	766/917 (84%)	571/698 (82%)	147/204 (72%)

#### 7.1.4 Statistically unusual chemical shifts [i](#)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	364	SER	CB	71.97	71.24 – 56.34	5.5
1	A	419	LYS	HE3	3.93	3.86 – 1.96	5.4
1	A	419	LYS	HE2	3.93	3.87 – 1.97	5.3

#### 7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:

