Full wwPDB EM Validation Report ⓘ

Mar 3, 2025 – 12:11 pm GMT

PDB ID	:	8QTV
EMDB ID	:	EMD-18654
Title	:	MUC5AC D'D3CysD1 domains.
Authors	:	Trillo-Muyo, S.; Hansson, G.C.
Deposited on	:	2023-10-13
Resolution	:	3.25 Å(reported)
Based on initial model	:	8QTB

This is a Full wwPDB EM Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
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:

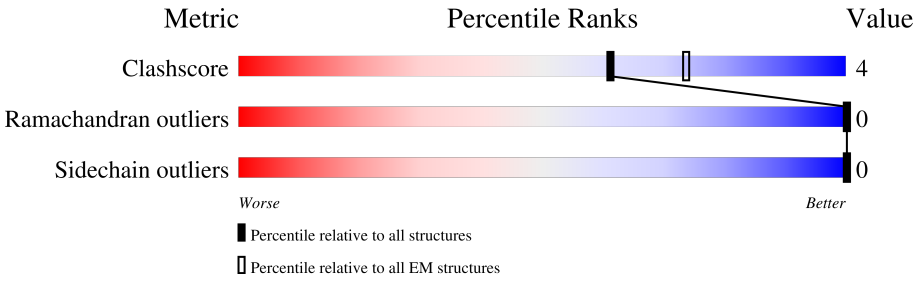
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.25 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	722	41% 5% 54%
1	B	722	41% 5% 54%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5054 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Mucin-5AC.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	331	Total	C	N	O	S	0	0
			2526	1581	427	485	33		
1	B	331	Total	C	N	O	S	0	0
			2526	1581	427	485	33		

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	778	ASP	-	expression tag	UNP P98088
A	779	ALA	-	expression tag	UNP P98088
A	780	ALA	-	expression tag	UNP P98088
A	781	GLN	-	expression tag	UNP P98088
A	782	PRO	-	expression tag	UNP P98088
A	783	ALA	-	expression tag	UNP P98088
A	784	ARG	-	expression tag	UNP P98088
A	785	ARG	-	expression tag	UNP P98088
A	786	ALA	-	expression tag	UNP P98088
A	787	VAL	-	expression tag	UNP P98088
A	788	ARG	-	expression tag	UNP P98088
A	789	SER	-	expression tag	UNP P98088
A	790	SER	-	expression tag	UNP P98088
A	791	ARG	-	expression tag	UNP P98088
A	792	HIS	-	expression tag	UNP P98088
A	793	HIS	-	expression tag	UNP P98088
A	794	HIS	-	expression tag	UNP P98088
A	795	HIS	-	expression tag	UNP P98088
A	796	HIS	-	expression tag	UNP P98088
A	797	HIS	-	expression tag	UNP P98088
A	798	GLY	-	expression tag	UNP P98088
A	799	SER	ALA	conflict	UNP P98088
A	1482	THR	SER	conflict	UNP P98088
A	1483	SER	THR	conflict	UNP P98088
A	1484	GLU	SER	conflict	UNP P98088
A	1485	GLN	SER	conflict	UNP P98088

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1486	LYS	SER	conflict	UNP P98088
A	1487	LEU	-	expression tag	UNP P98088
A	1488	ILE	-	expression tag	UNP P98088
A	1489	SER	-	expression tag	UNP P98088
A	1490	GLU	-	expression tag	UNP P98088
A	1491	GLU	-	expression tag	UNP P98088
A	1492	ASP	-	expression tag	UNP P98088
A	1493	LEU	-	expression tag	UNP P98088
A	1494	SER	-	expression tag	UNP P98088
A	1495	ARG	-	expression tag	UNP P98088
A	1496	LYS	-	expression tag	UNP P98088
A	1497	LEU	-	expression tag	UNP P98088
A	1498	THR	-	expression tag	UNP P98088
A	1499	ARG	-	expression tag	UNP P98088
B	778	ASP	-	expression tag	UNP P98088
B	779	ALA	-	expression tag	UNP P98088
B	780	ALA	-	expression tag	UNP P98088
B	781	GLN	-	expression tag	UNP P98088
B	782	PRO	-	expression tag	UNP P98088
B	783	ALA	-	expression tag	UNP P98088
B	784	ARG	-	expression tag	UNP P98088
B	785	ARG	-	expression tag	UNP P98088
B	786	ALA	-	expression tag	UNP P98088
B	787	VAL	-	expression tag	UNP P98088
B	788	ARG	-	expression tag	UNP P98088
B	789	SER	-	expression tag	UNP P98088
B	790	SER	-	expression tag	UNP P98088
B	791	ARG	-	expression tag	UNP P98088
B	792	HIS	-	expression tag	UNP P98088
B	793	HIS	-	expression tag	UNP P98088
B	794	HIS	-	expression tag	UNP P98088
B	795	HIS	-	expression tag	UNP P98088
B	796	HIS	-	expression tag	UNP P98088
B	797	HIS	-	expression tag	UNP P98088
B	798	GLY	-	expression tag	UNP P98088
B	799	SER	ALA	conflict	UNP P98088
B	1482	THR	SER	conflict	UNP P98088
B	1483	SER	THR	conflict	UNP P98088
B	1484	GLU	SER	conflict	UNP P98088
B	1485	GLN	SER	conflict	UNP P98088
B	1486	LYS	SER	conflict	UNP P98088
B	1487	LEU	-	expression tag	UNP P98088

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1488	ILE	-	expression tag	UNP P98088
B	1489	SER	-	expression tag	UNP P98088
B	1490	GLU	-	expression tag	UNP P98088
B	1491	GLU	-	expression tag	UNP P98088
B	1492	ASP	-	expression tag	UNP P98088
B	1493	LEU	-	expression tag	UNP P98088
B	1494	SER	-	expression tag	UNP P98088
B	1495	ARG	-	expression tag	UNP P98088
B	1496	LYS	-	expression tag	UNP P98088
B	1497	LEU	-	expression tag	UNP P98088
B	1498	THR	-	expression tag	UNP P98088
B	1499	ARG	-	expression tag	UNP P98088

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

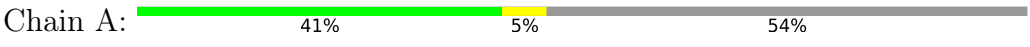
Mol	Chain	Residues	Atoms	AltConf
2	A	1	Total Ca 1 1	0
2	B	1	Total Ca 1 1	0

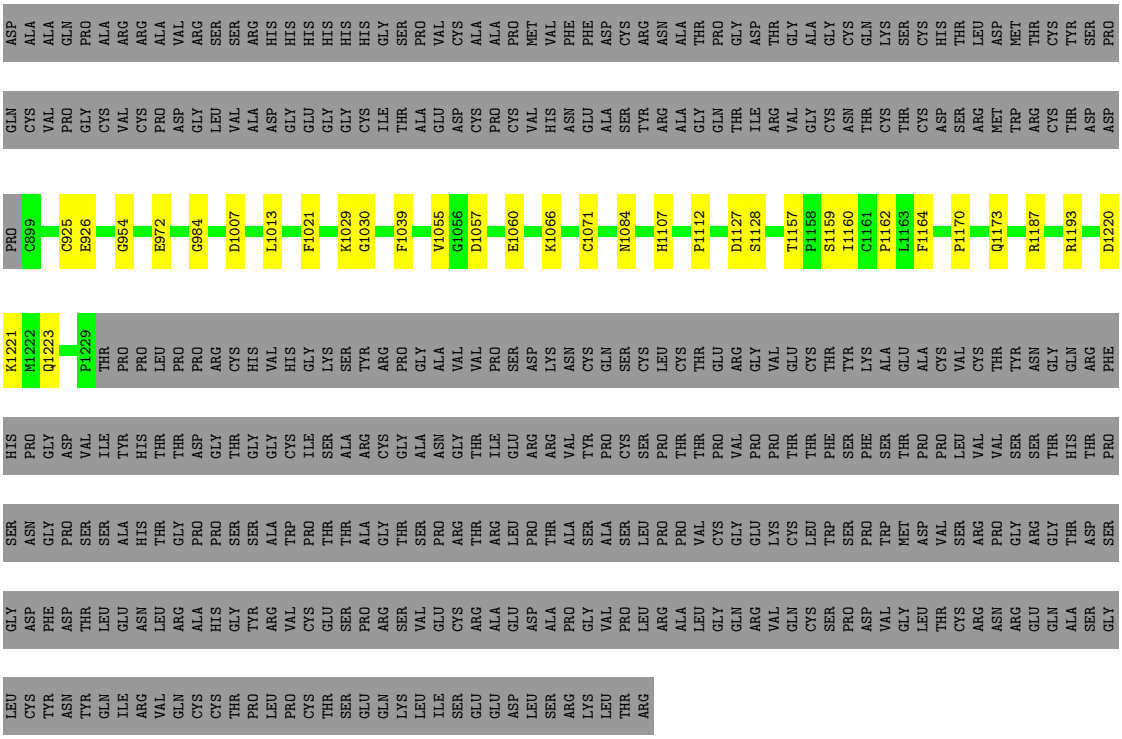
3 Residue-property plots (i)

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. 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. 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: Mucin-5AC

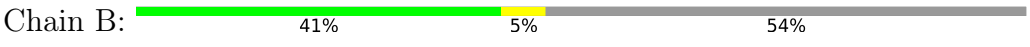
Chain A:

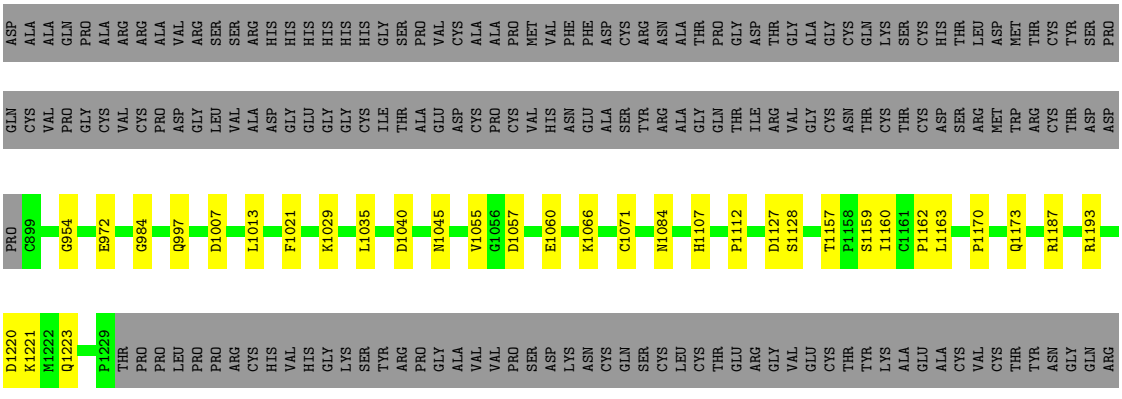


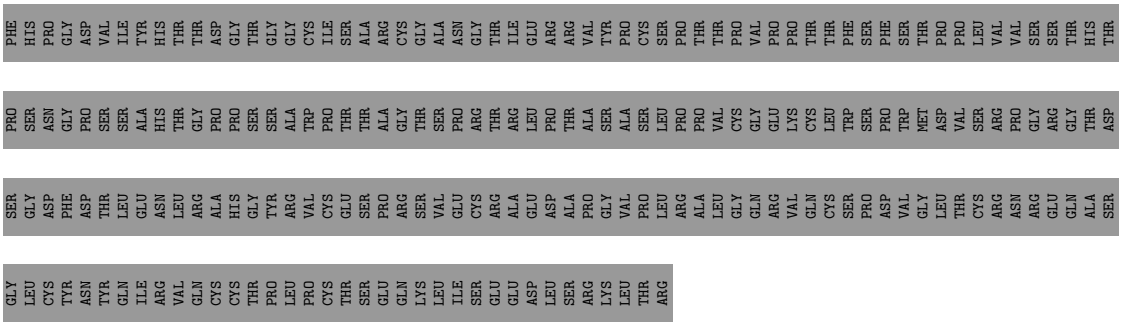


- Molecule 1: Mucin-5AC

Chain B:






PHE HIS PRO GLY ASP VAL ILE TYR HIS THR THR ASP GLY THR GLY GLY CYS ILE SER ALA ARG CYS GLY ALA ASN GLY THR ILE GLU ARG ARG VAL TYR PRO CYS SER PRO THR THR PRO VAL PRO PRO THR THR PHE SER SER PHE SER THR PRO PRO LEU VAL VAL SER SER THR HIS THR
PRO SER ASN GLY PRO SER SER ALA HIS THR GLY PRO PRO SER SER ALA TRP PRO THR THR ALA GLY THR SER PRO ARG THR ARG LEU PRO THR ALA SER ALA SER LEU PRO PRO VAL CYS GLY GLU LYS CYS LEU TRP SER PRO TRP MET ASP VAL SER ARG PRO GLY ARG GLY THR ASP
SER GLY ASP PHE ASP THR LEU GLU ASN LEU ARG ALA HIS GLY TYR ARG VAL CYS GLU SER PRO ARG SER VAL GLU CYS ARG ALA GLU ASP ALA PRO GLY VAL PRO LEU ARG ALA LEU GLY GLN ARG VAL GLN CYS SER PRO ASP VAL GLY LEU THR CYS ARG ASN ARG GLU GLN ALA SER
GLY LEU CYS TYR ASN TYR GLN ILE ARG VAL GLN CYS CYS THR PRO LEU PRO CYS THR SER GLU GLN LYS LEU ILE SER GLU GLU ASP LEU SER ARG LYS LEU THR ARG

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	216899	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	1.29	Depositor
Minimum defocus (nm)	750	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.24	0/2592	0.46	0/3528
1	B	0.24	0/2592	0.46	0/3528
All	All	0.24	0/5184	0.46	0/7056

There are no bond length outliers.

There are no bond angle outliers.

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	2526	0	2360	21	0
1	B	2526	0	2360	22	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
All	All	5054	0	4720	41	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 (41) 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:1157:THR:HG22	1:A:1159:SER:H	1.54	0.73
1:B:1157:THR:HG22	1:B:1159:SER:H	1.54	0.71
1:A:954:GLY:HA2	1:A:1128:SER:HB2	1.85	0.58
1:B:954:GLY:HA2	1:B:1128:SER:HB2	1.85	0.58
1:B:1084:ASN:ND2	1:B:1127:ASP:OD2	2.39	0.55
1:A:1084:ASN:ND2	1:A:1127:ASP:OD2	2.39	0.55
1:B:1055:VAL:HG12	1:B:1057:ASP:H	1.71	0.55
1:A:1162:PRO:HG2	1:B:1162:PRO:HG2	1.89	0.54
1:B:1170:PRO:HG2	1:B:1173:GLN:HB2	1.90	0.54
1:A:1055:VAL:HG12	1:A:1057:ASP:H	1.71	0.54
1:A:1170:PRO:HG2	1:A:1173:GLN:HB2	1.90	0.54
1:A:1007:ASP:N	1:A:1007:ASP:OD1	2.41	0.52
1:B:1013:LEU:HB2	1:B:1021:PHE:HB2	1.92	0.52
1:A:1057:ASP:HB3	1:A:1060:GLU:HB2	1.92	0.50
1:B:1057:ASP:HB3	1:B:1060:GLU:HB2	1.92	0.50
1:A:1013:LEU:HB2	1:A:1021:PHE:HB2	1.92	0.50
1:B:1007:ASP:N	1:B:1007:ASP:OD1	2.41	0.49
1:A:972:GLU:HB3	1:A:984:GLY:H	1.79	0.47
1:A:1220:ASP:OD1	1:A:1220:ASP:N	2.46	0.47
1:B:972:GLU:HB3	1:B:984:GLY:H	1.79	0.46
1:B:1157:THR:HB	1:B:1160:ILE:HB	1.98	0.46
1:A:1107:HIS:CD2	1:A:1112:PRO:HG3	2.51	0.46
1:B:1107:HIS:CD2	1:B:1112:PRO:HG3	2.51	0.45
1:A:1157:THR:HB	1:A:1160:ILE:HB	1.98	0.45
1:B:1035:LEU:HD23	1:B:1035:LEU:HA	1.88	0.44
1:A:1066:LYS:NZ	1:A:1071:CYS:O	2.51	0.44
1:A:1164:PHE:HD2	1:B:1163:LEU:H	1.66	0.43
1:A:1221:LYS:HD2	1:A:1223:GLN:HE22	1.83	0.43
1:A:1187:ARG:HG2	1:A:1193:ARG:HA	2.01	0.43
1:B:997:GLN:HE21	1:B:997:GLN:HB2	1.66	0.43
1:B:1187:ARG:HG2	1:B:1193:ARG:HA	2.01	0.43
1:B:1066:LYS:NZ	1:B:1071:CYS:O	2.51	0.42
1:B:1221:LYS:HD2	1:B:1223:GLN:HE22	1.83	0.42
1:B:1029:LYS:HE3	1:B:1029:LYS:HB2	1.87	0.42
1:A:1030:GLY:N	1:A:1039:PHE:O	2.52	0.42
1:A:1107:HIS:HD2	1:A:1112:PRO:HG3	1.86	0.41
1:B:1040:ASP:OD2	1:B:1045:ASN:ND2	2.51	0.41
1:B:1107:HIS:HD2	1:B:1112:PRO:HG3	1.85	0.41
1:B:1220:ASP:OD1	1:B:1220:ASP:N	2.46	0.41
1:A:1029:LYS:HE3	1:A:1029:LYS:HB2	1.87	0.40
1:A:925:CYS:SG	1:A:926:GLU:N	2.94	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 PDB entries followed by that with respect to all EM 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	329/722 (46%)	312 (95%)	17 (5%)	0	100	100
1	B	329/722 (46%)	312 (95%)	17 (5%)	0	100	100
All	All	658/1444 (46%)	624 (95%)	34 (5%)	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 PDB entries followed by that with respect to all EM 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	280/611 (46%)	280 (100%)	0	100	100
1	B	280/611 (46%)	280 (100%)	0	100	100
All	All	560/1222 (46%)	560 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.