wwPDB EM Validation Summary Report

Nov 7, 2023 – 12:54 am GMT

PDB ID	:	7QTQ
EMDB ID	:	EMD-14145
Title	:	Structure of Native, iodinated bovine thyroglobulin solved on strepavidin affinity grids.
Authors	:	Marechal, N.; Weitz, J.C.; Serrano, B.P.; Zhang, X.
Deposited on	:	2022-01-15
Resolution	:	3.30 Å(reported)

This is a wwPDB EM Validation Summary 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:

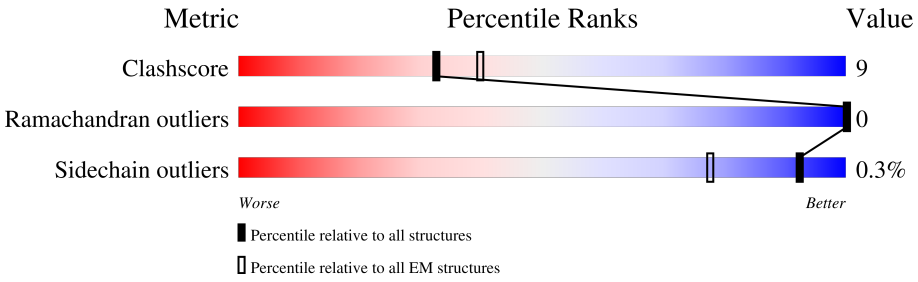
EMDB validation analysis	:	0.0.1.dev70
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.30 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 55234 atoms, of which 27101 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.

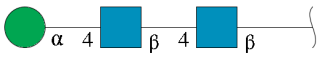
- Molecule 1 is a protein called Thyroglobulin.

Mol	Chain	Residues	Atoms							AltConf	Trace
1	A	1790	Total	C	H	I	N	O	S	0	0
			27329	8768	13410	4	2459	2587	101		
1	B	1789	Total	C	H	I	N	O	S	0	0
			27329	8768	13417	4	2455	2584	101		

There are 6 discrepancies between the modelled and reference sequences:

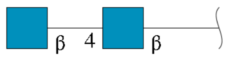
Chain	Residue	Modelled	Actual	Comment	Reference
A	196	PHE	LEU	variant	UNP P01267
A	2540	DHA	TYR	modified residue	UNP P01267
A	2573	T44	TYR	modified residue	UNP P01267
B	196	PHE	LEU	variant	UNP P01267
B	2540	DHA	TYR	modified residue	UNP P01267
B	2573	T44	TYR	modified residue	UNP P01267

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



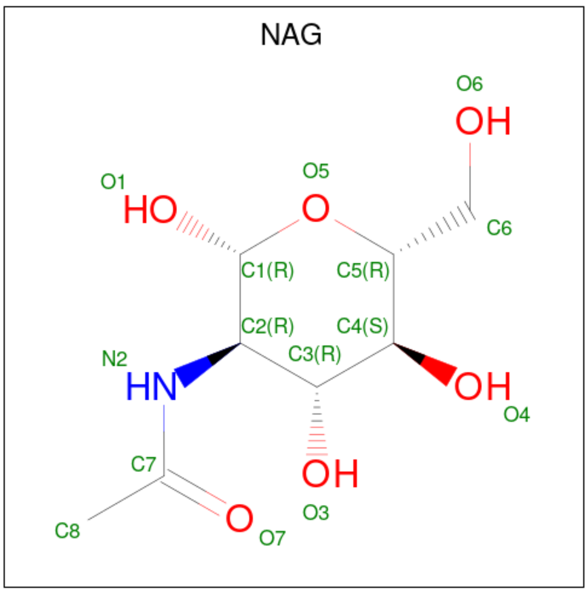
Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	3	Total	C	H	N	O	0	0
			73	22	34	2	15		
2	E	3	Total	C	H	N	O	0	0
			73	22	34	2	15		

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	H	N	O		
3	D	2	53	16	25	2	10	0	0
3	F	2	53	16	25	2	10	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	H	N	O	
4	A	1	27	8	13	1	5	0
4	A	1	27	8	13	1	5	0
4	A	1	27	8	13	1	5	0
4	A	1	27	8	13	1	5	0
4	A	1	27	8	13	1	5	0
4	A	1	27	8	13	1	5	0
4	B	1	27	8	13	1	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf
4	B	1	Total	C	H	N	O	0
			27	8	13	1	5	
4	B	1	Total	C	H	N	O	0
			27	8	13	1	5	
4	B	1	Total	C	H	N	O	0
			27	8	13	1	5	
4	B	1	Total	C	H	N	O	0
			27	8	13	1	5	
4	B	1	Total	C	H	N	O	0
			27	8	13	1	5	

SEQUENCE-PLOTS INFOmissingINFO

3 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	242813	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	54	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	105000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

4 Model quality

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MAN, DHA, T44, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.41	0/14211	0.52	3/19249 (0.0%)
1	B	0.41	0/14204	0.51	0/19241
All	All	0.41	0/28415	0.51	3/38490 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	965	LEU	CB-CG-CD1	-5.95	100.88	111.00
1	A	965	LEU	CB-CA-C	5.36	120.38	110.20
1	A	313	HIS	CB-CA-C	5.33	121.05	110.40

There are no chirality outliers.

There are no planarity outliers.

4.2 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	13919	13410	13428	247	0
1	B	13912	13417	13430	266	0
2	C	39	34	34	3	0
2	E	39	34	34	3	0
3	D	28	25	25	0	0
3	F	28	25	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	84	78	78	3	0
4	B	84	78	78	5	0
All	All	28133	27101	27132	513	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 9.

The worst 5 of 513 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)
2:C:2:NAG:O7	2:C:2:NAG:O3	2.05	0.74
1:B:296:ARG:NH1	1:B:321:PRO:O	2.20	0.74
1:A:1977:ILE:HD12	1:A:1979:ILE:HD11	1.70	0.74
4:B:2804:NAG:O7	4:B:2804:NAG:O3	2.05	0.74
4:A:2802:NAG:O7	4:A:2802:NAG:O3	2.05	0.73

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	1726/2769 (62%)	1548 (90%)	178 (10%)	0	100	100
1	B	1725/2769 (62%)	1539 (89%)	186 (11%)	0	100	100
All	All	3451/5538 (62%)	3087 (90%)	364 (10%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	1509/2312 (65%)	1505 (100%)	4 (0%)	92 96
1	B	1509/2312 (65%)	1505 (100%)	4 (0%)	92 96
All	All	3018/4624 (65%)	3010 (100%)	8 (0%)	92 96

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2353	ASN
1	B	1244	ARG
1	B	214	PHE
1	A	979	ARG
1	B	495	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	605	GLN
1	A	2137	HIS

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	DHA	B	2540	1	4,4,5	4.43	2 (50%)	2,4,6	3.37	1 (50%)
1	DHA	A	2540	1	4,4,5	4.50	2 (50%)	2,4,6	3.31	1 (50%)

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
1	DHA	B	2540	1	-	0/0/2/4	-
1	DHA	A	2540	1	-	0/0/2/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2540	DHA	C-CA	7.61	1.57	1.45
1	B	2540	DHA	C-CA	7.45	1.57	1.45
1	B	2540	DHA	CA-N	4.55	1.46	1.35
1	A	2540	DHA	CA-N	4.48	1.46	1.35

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	2540	DHA	O-C-CA	-4.70	116.78	125.54
1	A	2540	DHA	O-C-CA	-4.46	117.22	125.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.5 Carbohydrates (i)

10 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NAG	C	1	1,2	14,14,15	0.60	0	17,19,21	2.74	7 (41%)
2	NAG	C	2	2	14,14,15	0.38	0	17,19,21	0.98	1 (5%)
2	MAN	C	3	2	11,11,12	0.26	0	15,15,17	0.96	1 (6%)
3	NAG	D	1	1,3	14,14,15	0.56	0	17,19,21	0.81	0
3	NAG	D	2	3	14,14,15	0.35	0	17,19,21	1.20	2 (11%)
2	NAG	E	1	1,2	14,14,15	0.94	1 (7%)	17,19,21	1.13	1 (5%)
2	NAG	E	2	2	14,14,15	0.27	0	17,19,21	1.35	3 (17%)
2	MAN	E	3	2	11,11,12	0.37	0	15,15,17	1.34	1 (6%)
3	NAG	F	1	1,3	14,14,15	0.29	0	17,19,21	0.72	0
3	NAG	F	2	3	14,14,15	0.28	0	17,19,21	0.56	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
2	NAG	C	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	C	2	2	-	3/6/23/26	0/1/1/1
2	MAN	C	3	2	-	1/2/19/22	0/1/1/1
3	NAG	D	1	1,3	-	4/6/23/26	0/1/1/1
3	NAG	D	2	3	-	2/6/23/26	0/1/1/1
2	NAG	E	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	E	2	2	-	2/6/23/26	0/1/1/1
2	MAN	E	3	2	-	1/2/19/22	0/1/1/1
3	NAG	F	1	1,3	-	1/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1	NAG	C1-C2	3.18	1.57	1.52

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1	NAG	O5-C1-C2	-7.58	99.31	111.29
2	C	1	NAG	C1-O5-C5	-3.94	106.85	112.19
2	E	2	NAG	O5-C1-C2	-3.55	105.69	111.29
2	C	1	NAG	C2-N2-C7	-3.49	117.94	122.90
2	E	3	MAN	O5-C5-C6	3.41	112.54	107.20

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	2	NAG	C1-C2-N2-C7
2	C	2	NAG	C1-C2-N2-C7
3	D	2	NAG	O5-C5-C6-O6
2	C	1	NAG	O5-C5-C6-O6
3	D	2	NAG	C4-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	2	NAG	2	0
2	E	2	NAG	2	0
2	C	1	NAG	1	0
2	E	3	MAN	1	0

4.6 Ligand geometry (i)

12 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	NAG	A	2805	1	14,14,15	0.43	0	17,19,21	1.28	1 (5%)
4	NAG	B	2806	1	14,14,15	0.43	0	17,19,21	1.19	1 (5%)
4	NAG	A	2804	1	14,14,15	0.35	0	17,19,21	1.29	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	A	2802	1	14,14,15	0.64	0	17,19,21	1.97	4 (23%)
4	NAG	A	2801	1	14,14,15	0.45	0	17,19,21	0.72	0
4	NAG	A	2806	1	14,14,15	0.40	0	17,19,21	1.37	1 (5%)
4	NAG	B	2801	1	14,14,15	0.46	0	17,19,21	1.03	1 (5%)
4	NAG	A	2803	1	14,14,15	0.48	0	17,19,21	1.16	2 (11%)
4	NAG	B	2803	1	14,14,15	0.42	0	17,19,21	1.27	3 (17%)
4	NAG	B	2804	1	14,14,15	0.62	0	17,19,21	1.92	4 (23%)
4	NAG	B	2805	1	14,14,15	0.63	0	17,19,21	1.14	1 (5%)
4	NAG	B	2802	1	14,14,15	0.44	0	17,19,21	0.66	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	NAG	A	2805	1	-	3/6/23/26	0/1/1/1
4	NAG	B	2806	1	-	2/6/23/26	0/1/1/1
4	NAG	A	2804	1	-	1/6/23/26	0/1/1/1
4	NAG	A	2802	1	-	2/6/23/26	0/1/1/1
4	NAG	A	2801	1	-	0/6/23/26	0/1/1/1
4	NAG	A	2806	1	-	2/6/23/26	0/1/1/1
4	NAG	B	2801	1	-	4/6/23/26	0/1/1/1
4	NAG	A	2803	1	-	2/6/23/26	0/1/1/1
4	NAG	B	2803	1	-	0/6/23/26	0/1/1/1
4	NAG	B	2804	1	-	2/6/23/26	0/1/1/1
4	NAG	B	2805	1	-	0/6/23/26	0/1/1/1
4	NAG	B	2802	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	2802	NAG	C1-O5-C5	5.32	119.41	112.19
4	B	2804	NAG	C1-O5-C5	5.23	119.27	112.19
4	A	2806	NAG	C1-O5-C5	4.82	118.72	112.19
4	A	2804	NAG	C1-O5-C5	4.45	118.22	112.19
4	A	2805	NAG	C1-O5-C5	3.90	117.48	112.19

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	2801	NAG	C1-C2-N2-C7
4	A	2805	NAG	C1-C2-N2-C7
4	A	2806	NAG	O5-C5-C6-O6
4	A	2803	NAG	O5-C5-C6-O6
4	B	2801	NAG	O5-C5-C6-O6

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	2805	NAG	2	0
4	B	2806	NAG	2	0
4	A	2802	NAG	1	0
4	B	2801	NAG	2	0
4	B	2804	NAG	1	0

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Map visualisation

This section contains visualisations of the EMDB entry EMD-14145. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

5.1 Orthogonal projections

This section was not generated.

5.2 Central slices

This section was not generated.

5.3 Largest variance slices

This section was not generated.

5.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

5.5 Orthogonal surface views

This section was not generated.

5.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

6 Map analysis

This section contains the results of statistical analysis of the map.

6.1 Map-value distribution

This section was not generated.

6.2 Volume estimate versus contour level

This section was not generated.

6.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

7 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

8 Map-model fit

This section was not generated.