wwPDB NMR Structure Validation Summary Report (i)

Jun 12, 2024 – 10:55 PM EDT

PDB ID	:	1HD4
Title	:	SOLUTION STRUCTURE OF THE A-SUBUNIT OF HUMAN CHORIONIC GONADOTROPIN [MODELED WITH DIANTENNARY GLYCAN AT ASN78]
Authors	:	Erbel, P.J.A.; Karimi-Nejad, Y.; van Kuik, J.A.; Boelens, R.; Kamerling, J.P.; Vliegenthart, J.F.G.
Deposited on	:	2000-11-07

This is a wwPDB NMR Structure 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

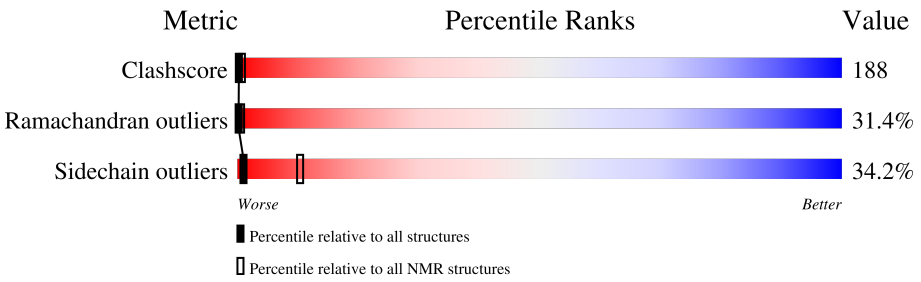
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	92	• 21% 22% • 54%
2	B	9	44% 56%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA and RNA chains that are outliers for geometric criteria:

Mol	Chain	Compound	Res	Total models with violations	
				Chirality	Geometry
2	B	NAG	2	1	-
2	B	NAG	5	1	-

2 Ensemble composition and analysis (i)

This entry contains 26 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:11-A:29, A:61-A:83 (42)	0.59	3

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 4 clusters and 1 single-model cluster was found.

Cluster number	Models
1	3, 5, 7, 9, 12, 16, 22, 23
2	1, 2, 4, 10, 13, 15, 20, 25
3	6, 8, 11, 17, 24
4	14, 19, 21, 26
Single-model clusters	18

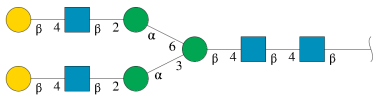
3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1595 atoms, of which 778 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CHORIONIC GONADOTROPIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	92	1381	437	675	122	134	13	0

- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	9	214	62	103	4	45	0

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: CHORIONIC GONADOTROPIN

Chain A: 21% 22% 54%

A1 P2 D3 V4 Q5 D6 C7 P8 E9 C10 T11 L12 Q13 E14 N15 P16 F17 F18 S19 Q20 P21 G22 A23 P24 I25 L26 Q27 C28 M29 G30 C31 C32 F33 S34 R35 A36 Y37 P38 T39 P40 L41 R42 S43 K44 K45 T46 M47 L48 V49 Q50 K51 N52 V53 T54 S55 E56 S57 T58 C59 C60

V61 A62 K63 S64 Y65 N66 R67 V68 T69 V70 M71 G72 G73 F74 K75 V76 E77 N78 H79 T80 A81 C82 H83 C84 S85 T86 C87 Y88 Y89 H90 K91 S92

• Molecule 2: beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B: 44% 56%

NAG1 NAG2 BMA3 MAN4 NAG5 GAL6 MAN7 NAG8 GAL9

4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 3. Colouring as in section 4.1 above.

• Molecule 1: CHORIONIC GONADOTROPIN

Chain A: 22% 17% 54%

A1 P2 D3 V4 Q5 D6 C7 P8 E9 C10 T11 L12 Q13 E14 N15 P16 F17 F18 S19 Q20 P21 G22 A23 P24 I25 L26 Q27 C28 M29 G30 C31 C32 F33 S34 R35 A36 Y37 P38 T39 P40 L41 R42 S43 K44 K45 T46 M47 L48 V49 Q50 K51 N52 V53 T54 S55 E56 S57 T58 C59 C60

V61 A62 K63 S64 Y65 N66 R67 V68 T69 V70 M71 F74 K75 V76 E77 N78 H79 T80 A81 C82 H83 C84 S85 T86 C87 Y88 Y89 H90 K91 S92

• Molecule 2: beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *DG/SA*.

Of the 150 calculated structures, 26 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
X-PLOR	refinement	
X-PLOR	structure solution	

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

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 (i)

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	324	313	313	153±11
2	B	111	103	94	10±4
All	All	11310	10816	10582	4125

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 188.

5 of 693 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:25:ILE:CG2	1:A:76:VAL:HG21	1.21	1.65	10	26
1:A:12:LEU:HD21	1:A:79:HIS:O	1.20	1.34	22	26
1:A:25:ILE:HG23	1:A:76:VAL:HG21	1.12	1.16	8	26
1:A:18:PHE:CD1	1:A:70:VAL:HG12	1.11	1.80	20	4
1:A:23:ALA:CB	2:B:1:NAG:H83	1.11	1.76	16	13

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	42/92 (46%)	16±2 (38±4%)	13±2 (30±5%)	13±1 (31±3%)	0 0
All	All	1092/2392 (46%)	420 (38%)	329 (30%)	343 (31%)	0 0

5 of 21 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	14	GLU	26
1	A	25	ILE	26
1	A	27	GLN	26
1	A	62	ALA	26
1	A	67	ARG	26

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	36/83 (43%)	24±2 (66±6%)	12±2 (34±6%)	1 10
All	All	936/2158 (43%)	616 (66%)	320 (34%)	1 10

5 of 27 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	26	LEU	26
1	A	77	GLU	26
1	A	70	VAL	22
1	A	66	ASN	22
1	A	67	ARG	18

6.3.3 RNA (i)

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)

9 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	NAG	B	1	2,1	14,14,15	0.38±0.02	0±0 (0±0%)
2	NAG	B	2	2	14,14,15	0.39±0.03	0±0 (0±0%)
2	BMA	B	3	2	11,11,12	0.35±0.04	0±0 (0±0%)
2	MAN	B	4	2	11,11,12	0.28±0.02	0±0 (0±0%)
2	NAG	B	5	2	14,14,15	0.44±0.02	0±0 (0±0%)
2	GAL	B	6	2	11,11,12	0.25±0.02	0±0 (0±0%)
2	MAN	B	7	2	11,11,12	0.26±0.04	0±0 (0±0%)
2	NAG	B	8	2	14,14,15	0.45±0.03	0±0 (0±0%)
2	GAL	B	9	2	11,11,12	0.25±0.02	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	NAG	B	1	2,1	17,19,21	0.68±0.04	0±0 (1±2%)
2	NAG	B	2	2	17,19,21	0.68±0.08	0±0 (0±2%)

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	BMA	B	3	2	15,15,17	0.40±0.05	0±0 (0±0%)
2	MAN	B	4	2	15,15,17	0.57±0.04	0±0 (0±0%)
2	NAG	B	5	2	17,19,21	0.67±0.04	0±0 (1±2%)
2	GAL	B	6	2	15,15,17	0.37±0.02	0±0 (0±0%)
2	MAN	B	7	2	15,15,17	0.66±0.18	0±0 (1±3%)
2	NAG	B	8	2	17,19,21	0.74±0.05	1±0 (3±2%)
2	GAL	B	9	2	15,15,17	0.36±0.02	0±0 (0±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	B	1	2,1	-	0±0,6,23,26	0±0,1,1,1
2	NAG	B	2	2	-	0±0,6,23,26	0±0,1,1,1
2	BMA	B	3	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	4	2	-	0±0,2,19,22	0±0,1,1,1
2	NAG	B	5	2	-	0±0,6,23,26	0±0,1,1,1
2	GAL	B	6	2	-	0±0,2,19,22	0±0,1,1,1
2	MAN	B	7	2	-	0±0,2,19,22	0±0,1,1,1
2	NAG	B	8	2	-	0±0,6,23,26	0±0,1,1,1
2	GAL	B	9	2	-	0±0,2,19,22	0±0,1,1,1

There are no bond-length outliers.

5 of 6 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	7	MAN	C1-O5-C5	3.94	117.54	112.19	12	4
2	B	8	NAG	C2-N2-C7	2.83	118.88	122.90	16	16
2	B	2	NAG	C2-N2-C7	2.65	119.13	122.90	25	4
2	B	5	NAG	C2-N2-C7	2.29	119.64	122.90	23	7
2	B	1	NAG	C2-N2-C7	2.16	119.82	122.90	14	7

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

Mol	Chain	Res	Type	Atoms	Models (Total)
2	B	2	NAG	C1	1

Continued on next page...

Continued from previous page...

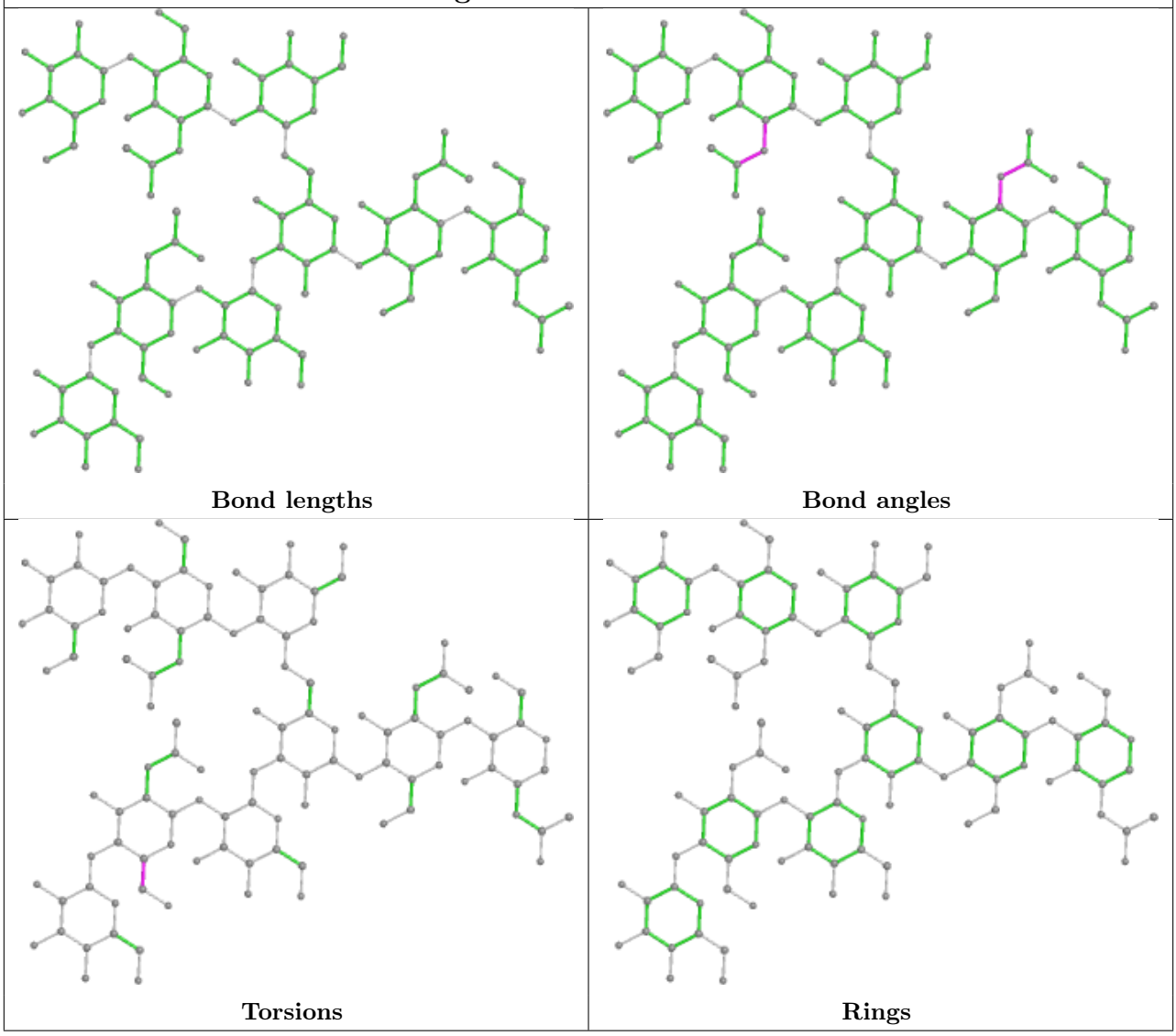
Mol	Chain	Res	Type	Atoms	Models (Total)
2	B	5	NAG	C1	1

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain B



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)

No chemical shift data were provided