wwPDB NMR Structure Validation Summary Report (i)

Jun 11, 2024 – 07:14 PM EDT

PDB ID	:	1HDP
Title	:	SOLUTION STRUCTURE OF A POU-SPECIFIC HOMEODOMAIN: 3D-NMR STUDIES OF HUMAN B-CELL TRANSCRIPTION FACTOR OCT-2
Authors	:	Sivaraja, M.; Botfield, M.C.; Mueller, M.; Jancso, A.; Weiss, M.A.
Deposited on	:	1994-03-08

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 (i)) were used in the production of this report:

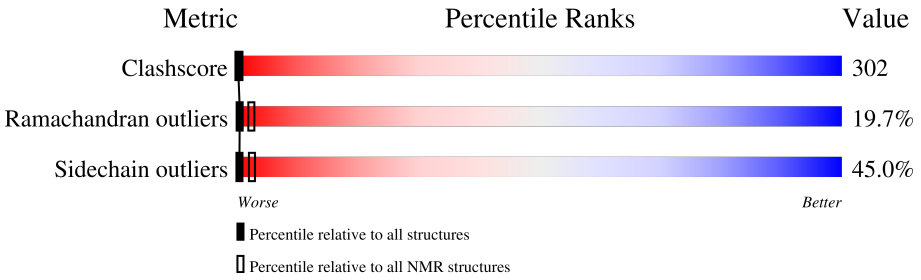
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson 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	63	5% 11% 41% 43%

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1097 atoms, of which 563 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called OCT-2 POU HOMEODOMAIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	63	1097	334	563	104	93	3	0

4 Residue-property plots (i)

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: OCT-2 POU HOMEODOMAIN

Chain A: 5% 11% 41% 43%

R1 R2 K3 K4 R5 T6 S7 I8 E9 T10 N11 V12 R13 F14 A15 L16 E17 K18 S19 F20 L21 A22 N23 Q24 K25 P26 T27 S28 E29 E30 I31 L32 L33 I34 A35 E36 Q37 L38 H39 M40 E41 K42 E43 V44 I45 R46 V47 W48 F49 C50 N51 R52 R53 Q54 K55 E56 K57 R58 I59 N60

P61 C62 S63

5 Refinement protocol and experimental data overview (i)

Of the ? calculated structures, 1 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
DGII	refinement	

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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 (average) 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	29.10	313/541 (57.9%)	15.70	259/721 (35.9%)
All	All	29.10	313/541 (57.9%)	15.70	259/721 (35.9%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	18	4
All	All	18	4

5 of 313 bond outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	61	PRO	N-CD	-98.07	0.10	1.47
1	A	56	GLU	CD-OE2	-89.77	0.26	1.25
1	A	9	GLU	CD-OE1	-89.46	0.27	1.25
1	A	30	GLU	CD-OE2	-84.24	0.33	1.25
1	A	36	GLU	CD-OE1	-83.84	0.33	1.25

5 of 259 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	2	ARG	NE-CZ-NH2	-134.60	53.00	120.30
1	A	5	ARG	NE-CZ-NH2	-101.12	69.74	120.30
1	A	56	GLU	OE1-CD-OE2	-76.28	31.76	123.30
1	A	36	GLU	OE1-CD-OE2	-76.14	31.93	123.30
1	A	5	ARG	NE-CZ-NH1	74.08	157.34	120.30

5 of 15 chiral outliers are listed below.

Mol	Chain	Res	Type	Atoms
1	A	1	ARG	CA
1	A	2	ARG	CA
1	A	3	LYS	CA
1	A	4	LYS	CA
1	A	5	ARG	CA

All planar outliers are listed below.

Mol	Chain	Res	Type	Group
1	A	5	ARG	Sidechain
1	A	13	ARG	Sidechain
1	A	46	ARG	Sidechain
1	A	52	ARG	Sidechain

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	534	563	523	319
All	All	534	563	523	319

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 302.

5 of 319 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:53:ARG:CB	1:A:53:ARG:CD	1.59	1.77
1:A:8:ILE:CD1	1:A:8:ILE:CB	1.57	1.81
1:A:38:LEU:CD1	1:A:38:LEU:CA	1.52	1.81
1:A:42:LYS:NZ	1:A:42:LYS:CD	1.50	1.71
1:A:55:LYS:NZ	1:A:55:LYS:CD	1.49	1.75

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	61/63 (97%)	42 (69%)	7 (11%)	12 (20%)	0 2
All	All	61/63 (97%)	42 (69%)	7 (11%)	12 (20%)	0 2

5 of 12 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	2	ARG
1	A	6	THR
1	A	7	SER
1	A	9	GLU
1	A	38	LEU

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	60/60 (100%)	33 (55%)	27 (45%)	0 2
All	All	60/60 (100%)	33 (55%)	27 (45%)	0 2

5 of 27 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	1	ARG
1	A	2	ARG
1	A	3	LYS
1	A	6	THR
1	A	11	ASN

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)

There are no monosaccharides 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	25

The worst 5 of 25 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	16:LEU	C	17:GLU	N	1.20
1	A	24:GLN	C	25:LYS	N	1.19
1	A	21:LEU	C	22:ALA	N	1.17
1	A	22:ALA	C	23:ASN	N	1.17
1	A	20:PHE	C	21:LEU	N	1.13

7 Chemical shift validation (i)

No chemical shift data were provided