wwPDB NMR Structure Validation Summary Report (i)

Jun 16, 2024 – 02:03 AM EDT

PDB ID	:	2L2G
BMRB ID	:	17134
Title	:	Solution structure of Opossum Domain 11
Authors	:	Williams, C.; Hoppe, H.; Rezgui, D.; Rezgui, M.; Frago, S.; Ellis, R.Z.; Wattana-Amorn, P.; Prince, S.N.; Zaccheo, O.J.; Forbes, B.; Jones, E.Y.; Crump, M.P.; Bassim, A.H.
Deposited on	:	2010-08-18

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:

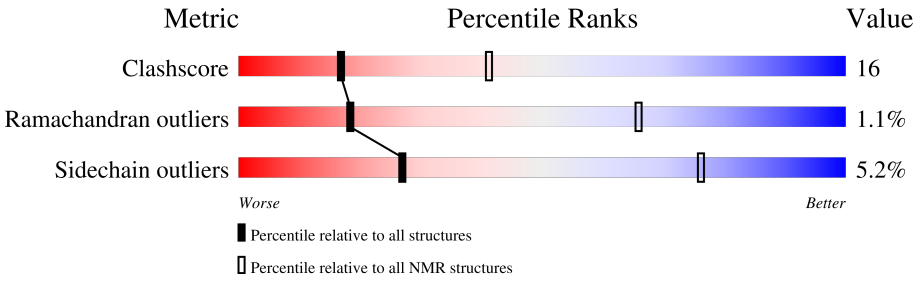
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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	151	

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 10 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:1531-A:1614, A:1620-A:1662 (127)	0.46	10

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 5 clusters and 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2215 atoms, of which 1089 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called IGF2R DOMAIN 11.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	147	2215	709	1089	190	217	10	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: IGF2R DOMAIN 11

Chain A: 56% 28% 13% •

M1524 K1525 S1526 N1527 I1528 Q1529 D1530 N1531 C1532 Q1533 F1543 D1544 L1545 K1549 V1556 K1561 G1562 L1563 I1564 Y1565 T1572 V1579 G1580 V1581 C1582 F1583 G1584 L1585 T1586 K1587 A1590 W1593 L1597 M1598 Y1599 Q1602 V1603 L1604 V1607 P1613 C1614 P1615 S1616 K1617 N1618 A1619 L1620 K1621

Y1622 V1625 I1626 S1627 F1628 N1638 K1639 P1640 V1641 F1642 L1645 D1646 T1649 C1650 T1651 L1652 Y1653 W1656 H1657 T1658 P1659 L1660 E1663 K1664 E1665 E1666 P1667 R1668 H1669 H1670 HIS HIS HIS HIS

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

The representative model is number 10. Colouring as in section 4.1 above.

• Molecule 1: IGF2R DOMAIN 11

Chain A: 47% 36% • 13% •

M1524 K1525 S1526 N1527 I1528 Q1529 D1530 N1531 C1532 Q1533 V1534 F1543 D1544 L1545 K1549 G1553 V1556 A1557 K1561 G1562 L1563 I1564 Y1565 I1568 T1572 K1573 N1574 C1575 P1576 V1579 G1580 V1581 C1582 F1583 G1584 L1585 T1586 K1587 A1590 W1593 Q1596 L1597 M1598 Y1599 Q1602 V1603 L1604

V1607 Y1608 D1609 P1613 C1614 P1615 S1616 K1617 N1618 A1619 L1620 K1621 Y1622 K1623 S1624 V1625 I1626 S1627 F1628 N1638 K1639 P1640 V1641 F1642 V1643 S1644 L1645 T1649 L1652 Y1653 F1654 S1655 W1656 H1657 T1658 P1659 L1660 E1663 K1664 E1665 E1666 P1667 R1668 H1669 H1670 HIS HIS HIS HIS

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing, simulated annealing.*

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy.*

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

Software name	Classification	Version
ARIA	structure solution	1.2
ARIA	refinement	1.2
CNS	refinement	1.2
CNS	structure solution	1.2
iCing	refinement	r765
TALOS	geometry optimization	

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	0.97±0.02	0±0/985 (0.0± 0.0%)	0.73±0.01	0±0/1343 (0.0± 0.0%)
All	All	0.97	1/19700 (0.0%)	0.73	1/26860 (0.0%)

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	1660	LEU	N-CA	-5.30	1.35	1.46	16	1

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	1660	LEU	N-CA-CB	-5.58	99.25	110.40	16	1

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	961	928	925	29±4
All	All	19220	18560	18500	586

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 16.

5 of 137 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1604:LEU:HB2	1:A:1628:PHE:HB2	0.89	1.44	9	20
1:A:1656:TRP:CD1	1:A:1658:THR:HB	0.76	2.16	19	18
1:A:1623:LYS:HB2	1:A:1651:THR:HG23	0.75	1.56	14	1
1:A:1556:VAL:HG22	1:A:1640:PRO:HB2	0.73	1.61	4	16
1:A:1560:GLU:HB3	1:A:1561:LYS:HD2	0.71	1.62	12	3

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	127/151 (84%)	113±1 (89±1%)	13±2 (10±1%)	1±1 (1±1%)	18	66
All	All	2540/3020 (84%)	2258 (89%)	254 (10%)	28 (1%)	18	66

5 of 7 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	1638	ASN	16
1	A	1560	GLU	4
1	A	1531	ASN	4
1	A	1549	LYS	1
1	A	1561	LYS	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	109/132 (83%)	103±2 (95±2%)	6±2 (5±2%)	27	76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2180/2640 (83%)	2067 (95%)	113 (5%)	27 76

5 of 26 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	1593	TRP	20
1	A	1561	LYS	17
1	A	1572	THR	15
1	A	1625	VAL	11
1	A	1585	LEU	9

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)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided