W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB NMR Structure Validation Summary Report (i)

Aug 30, 2017 – 11:18 PM EDT

PDB ID : 5JPW
Title : Molecular basis for protein recognition specificity of the DYNLT1/Tctex1 canonical binding groove. Characterization of the interaction with activin receptor IIB
Authors : Rodriguez-Crespo, I.; Merino-Gracia, J.; Bruix, M.; Zamora-Carreras, H.
Deposited on : unknown

The following versions of software and data (see references (i)) were used in the production of this report:

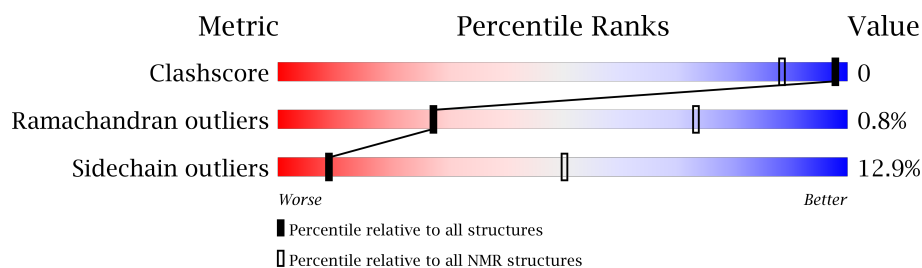|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20029824
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029824

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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		0.8%
Sidechain outliers		12.9%

Worse *Better*

▌ Percentile relative to all structures

▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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	143	72% 5% 23%
1	B	143	73% · 23%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:13-A:114, A:130-A:137, B:156-B:257, B:273-B:280 (220)	0.92	1

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 3 clusters. No single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4284 atoms, of which 2124 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Dynein light chain Tctex-type 1,Cytoplasmic dynein 1 intermediate chain 2.

Mol	Chain	Residues	Atoms						Trace
1	A	143	Total	C	H	N	O	S	0
			2142	678	1062	179	215	8	
1	B	143	Total	C	H	N	O	S	0
			2142	678	1062	179	215	8	

There are 18 discrepancies between the modelled and reference sequences:

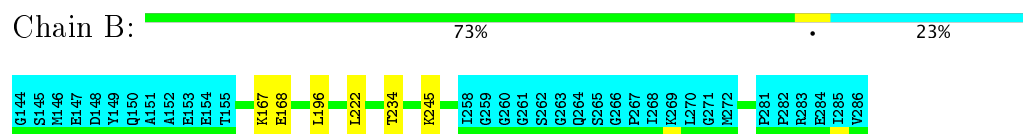| Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP P63172
A	2	SER	-	expression tag	UNP P63172
A	116	GLY	-	linker	UNP P63172
A	117	GLY	-	linker	UNP P63172
A	118	GLY	-	linker	UNP P63172
A	119	SER	-	linker	UNP P63172
A	120	GLY	-	linker	UNP P63172
A	121	GLN	-	linker	UNP P63172
A	122	SER	-	linker	UNP P63172
B	144	GLY	-	expression tag	UNP P63172
B	145	SER	-	expression tag	UNP P63172
B	259	GLY	-	linker	UNP P63172
B	260	GLY	-	linker	UNP P63172
B	261	GLY	-	linker	UNP P63172
B	262	SER	-	linker	UNP P63172
B	263	GLY	-	linker	UNP P63172
B	264	GLN	-	linker	UNP P63172
B	265	SER	-	linker	UNP P63172

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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.

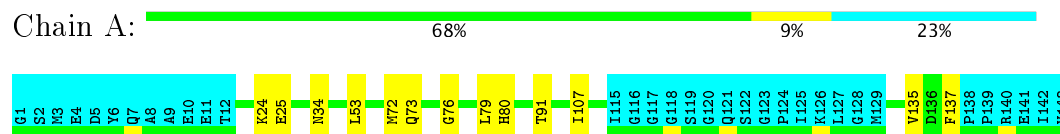- Molecule 1: Dynein light chain Tctex-type 1,Cytoplasmic dynein 1 intermediate chain 2

- Molecule 1: Dynein light chain Tctex-type 1,Cytoplasmic dynein 1 intermediate chain 2
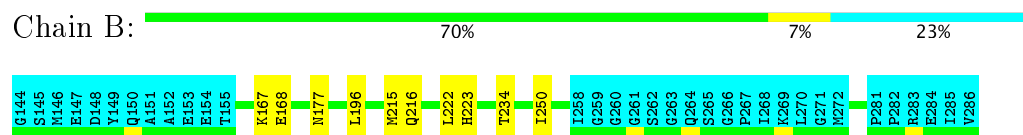

4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Dynein light chain Tctex-type 1,Cytoplasmic dynein 1 intermediate chain 2

- Molecule 1: Dynein light chain Tctex-type 1,Cytoplasmic dynein 1 intermediate chain 2

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *molecular dynamics*.

Of the 100 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
AMBER	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	5jpw_cs.cif
Number of chemical shift lists	1
Total number of shifts	2748
Number of shifts mapped to atoms	0
Number of unparsed shifts	2748
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	0%

No validations of the models with respect to experimental NMR restraints is performed at this time.

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	0.69±0.02	0±0/863 (0.0±0.0%)	0.95±0.02	0±0/1173 (0.0±0.0%)
1	B	0.69±0.02	0±0/863 (0.0±0.0%)	0.95±0.02	0±0/1173 (0.0±0.0%)
All	All	0.69	1/34520 (0.0%)	0.95	0/46920 (0.0%)

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	B	0.0±0.0	0.1±0.2
All	All	0	1

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	39	SER	CA-CB	5.34	1.60	1.52	17	1

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	B	208	TYR	Sidechain	1

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	B	846	836	834	1±2
1	A	846	836	834	1±2
All	All	33840	33440	33360	31

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 0.

5 of 21 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:85:CYS:SG	1:B:209:ILE:HD11	0.56	2.41	14	4
1:A:66:ILE:HD11	1:B:228:CYS:SG	0.55	2.42	4	2
1:A:79:LEU:HD13	1:B:222:LEU:HD13	0.55	1.77	14	1
1:A:79:LEU:HD11	1:B:220:ALA:HB1	0.52	1.82	14	1
1:B:224:THR:HB	1:B:278:VAL:HG13	0.51	1.83	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	110/143 (77%)	98±3 (89±2%)	11±3 (10±2%)	1±1 (1±1%)	30	75
1	B	110/143 (77%)	97±3 (89±3%)	12±3 (10±3%)	1±1 (1±1%)	25	72
All	All	4400/5720 (77%)	3910 (89%)	454 (10%)	36 (1%)	27	73

5 of 19 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	112	GLY	4
1	B	255	GLY	4
1	A	76	GLY	3
1	B	221	GLY	3
1	B	219	GLY	3

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	95/119 (80%)	82±3 (87±3%)	13±3 (13±3%)	8	49
1	B	95/119 (80%)	83±3 (88±3%)	12±3 (12±3%)	10	52
All	All	3800/4760 (80%)	3310 (87%)	490 (13%)	9	50

5 of 98 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	B	168	GLU	20
1	A	25	GLU	20
1	B	167	LYS	20
1	A	24	LYS	20
1	A	53	LEU	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)

There are no carbohydrates 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)

The completeness of assignment taking into account all chemical shift lists is 0% for the well-defined parts and 0% for the entire structure.

7.1 Chemical shift list 1

File name: 5jpw_cs.cif

Chemical shift list name: *Tctex-IC_dimer.bmrb*

7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	2748
Number of shifts mapped to atoms	0
Number of unparsed shifts	2748
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following errors were found when reading this chemical shift list.

- Entity instance (chain) must be specified. First 5 (of 2748) occurences are reported below.

Shift ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	?	1	GLY	CA	45.216	0.002	1
2	?	2	SER	N	115.595	0.028	1
3	?	2	SER	H	8.229	0.002	1
4	?	2	SER	CA	58.516	0.046	1
5	?	2	SER	CB	63.991	0.026	1

7.1.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 0%, i.e. 0 atoms were assigned a chemical

shift out of a possible 2578. 0 out of 34 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	0/1096 (0%)	0/438 (0%)	0/440 (0%)	0/218 (0%)
Sidechain	0/1240 (0%)	0/718 (0%)	0/476 (0%)	0/46 (0%)
Aromatic	0/242 (0%)	0/126 (0%)	0/102 (0%)	0/14 (0%)
Overall	0/2578 (0%)	0/1282 (0%)	0/1018 (0%)	0/278 (0%)

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (Tctex-IC_dimer.bmrb). RCI is only applicable to proteins.