wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 11:54 PM BST

PDB ID	:	2KTD
Title	:	Solution structure of mouse lipocalin-type prostaglandin D synthase / substrate analog (U-46619) complex
Authors	:	Shimamoto, S.; Maruo, H.; Yoshida, T.; Kato, N.; Ohkubo, T.
Deposited on	:	2010-01-27

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
http://wwpdb.org/validation/2016/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

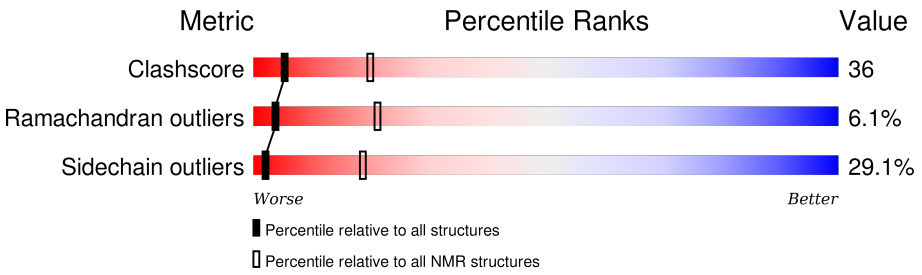
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	167	29% 51% 10% 10%

2 Ensemble composition and analysis (i)

This entry contains 15 models. Model 14 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:29-A:84, A:89-A:182 (150)	0.44	14

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, 4, 5, 6, 9, 10, 11, 12, 13, 14
2	3, 7
3	8, 15

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2651 atoms, of which 1316 are hydrogens and 0 are deuteriums.

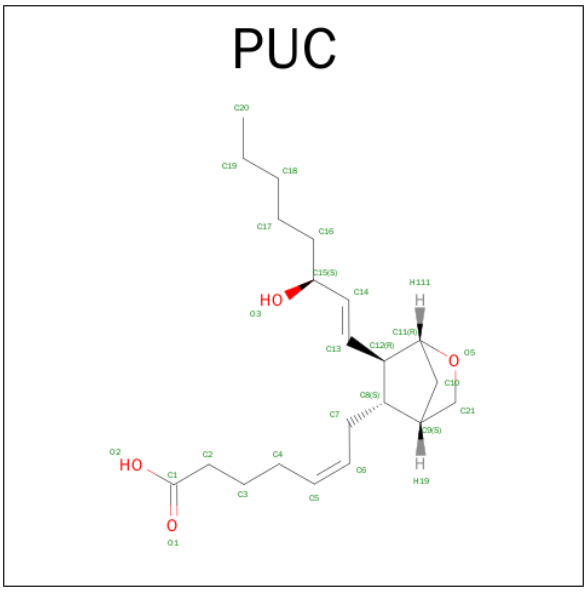
- Molecule 1 is a protein called Prostaglandin-H2 D-isomerase.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	167	2592	828	1282	223	255	4	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	GLY	-	EXPRESSION TAG	UNP O09114
A	24	SER	-	EXPRESSION TAG	UNP O09114
A	89	ALA	CYS	ENGINEERED MUTATION	UNP O09114
A	186	ALA	CYS	ENGINEERED MUTATION	UNP O09114

- Molecule 2 is (5Z)-7-{(1R,4S,5S,6R)-6-[(1E,3S)-3-HYDROXYOCT-1-EN-1-YL]-2-OXABICYCLO[2.2.1]HEPT-5-YL}HEPT-5-ENOIC ACID (three-letter code: PUC) (formula: $C_{21}H_{34}O_4$).



Mol	Chain	Residues	Atoms			
			Total	C	H	O
2	A	1	59	21	34	4

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: Prostaglandin-H2 D-isomerase

Chain A: 29% 51% 10% 10%

G23 S24 Q25 G26 H27 D28 T29 V30 Q31 P32 N33 F34 Q35 Q36 D37 K38 F39 L40 G41 R42 W43 A46 G47 L48 A49 S52 F55 K58 K59 A60 V61 L62 Y63 M64 C65 K66 T67 V68 V69 A70 P71 E74 G75 G76 L77 N78 L79 T80 F83 L84 R85 K86 N87 Q88

A89 E90 T91 K92 I93 N94 V95 L96 P102 Y105 T106 Y107 S108 S109 P110 H111 S112 G113 S114 I115 H116 S117 V118 S119 V120 V121 E122 A123 N124 Y125 D126 E127 Y128 A129 L130 L131 F132 S133 K137 Q141 M145 A146 T147 L148 Y149 S150 R151 T152 Q153 T154 L155 K156 L159 K160

E161 K162 F163 T164 T165 F166 S167 K168 A169 Q170 G171 L172 T173 E174 I177 V178 F179 L180 P181 Q182 P183 D184 K185 A186 I187 Q188 E189

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

The representative model is number 14. Colouring as in section 4.1 above.

- Molecule 1: Prostaglandin-H2 D-isomerase

Chain A: 34% 40% 15% • 10%

G23 S24 Q25 G26 H27 D28 T29 V30 Q31 P32 N33 F34 Q35 Q36 D37 K38 F39 L40 G41 R42 W43 A46 G47 L48 A49 S52 F55 K58 K59 A60 V61 L62 Y63 M64 C65 K66 T67 V68 V69 A70 P71 E74 G75 G76 L77 N78 L79 T80 F83 L84 R85 K86 N87 Q88

T91 K92 V95 L96 Q97 A101 P102 Y105 T106 Y107 S108 S109 P110 H111 S112 G113 S114 I115 H116 S117 V118 S119 V120 V121 E122 A123 N124 Y125 D126 E127 Y128 A129 L130 L131 F132 G135 Q141 D142 F143 R144 M145 A146 T147 L148 Y149 S150 R151 T152 Q153 T154 L155 K156 D157 E158

L159 K160 E161 K162 F163 T164 T165 F166 S167 K168 A169 L172 T173 I177 V178 F179 L180 P181 Q182 P183 D184 K185 A186 I187 Q188 E189

5 Refinement protocol and experimental data overview (i)

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

Of the 1500 calculated structures, 15 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
CNS	geometry optimization	
CNS	refinement	

No chemical shift data was provided. 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)

Bond lengths and bond angles in the following residue types are not validated in this section: PUC

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	1176	1155	1152	87±8
2	A	25	34	34	4±2
All	All	18015	17835	17790	1298

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 36.

5 of 541 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:148:LEU:HD13	1:A:159:LEU:HD21	0.94	1.37	3	2
1:A:121:VAL:HG23	1:A:130:LEU:HD21	0.92	1.41	2	3
1:A:48:LEU:HD22	2:A:200:PUC:H210	0.90	1.43	3	6
1:A:128:TYR:CG	1:A:159:LEU:HD12	0.89	2.02	11	2
1:A:49:ALA:HB3	1:A:146:ALA:HB3	0.87	1.45	12	11

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	150/167 (90%)	123±2 (82±1%)	18±3 (12±2%)	9±1 (6±1%)	4 21
All	All	2250/2505 (90%)	1839 (82%)	273 (12%)	138 (6%)	4 21

5 of 23 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	123	ALA	15
1	A	33	ASN	15
1	A	111	HIS	11
1	A	114	SER	10
1	A	180	LEU	10

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	128/142 (90%)	91±4 (71±3%)	37±4 (29±3%)	2 19
All	All	1920/2130 (90%)	1362 (71%)	558 (29%)	2 19

5 of 96 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	74	GLU	15
1	A	105	TYR	15
1	A	152	THR	15
1	A	112	SER	14
1	A	163	PHE	14

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)

1 ligand is 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	PUC	A	200	-	26,26,26	1.81±0.01	1±0 (3±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	PUC	A	200	-	24,33,33	2.77±0.01	1±0 (4±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	PUC	A	200	-	-	0±0,19,40,40	0±0,0,2,2

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	A	200	PUC	C21-C9	6.99	1.37	1.52	2	15

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	A	200	PUC	C9-C10-C11	10.83	86.65	103.68	11	15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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