wwPDB NMR Structure Validation Summary Report

Jun 12, 2024 – 01:18 PM EDT

PDB ID	:	2KX9
Title	:	Solution Structure of the Enzyme I dimer Using Residual Dipolar Couplings and Small Angle X-Ray Scattering
Authors	:	Schwieters, C.D.; Suh, J.; Grishaev, A.; Takayama, Y.; Guirlando, R.; Clore, G.
Deposited on	:	2010-04-29

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:

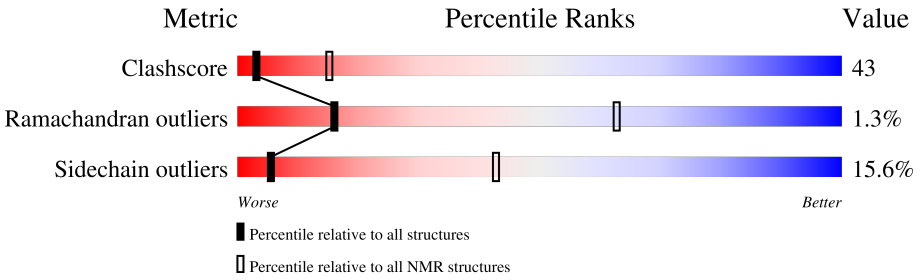
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson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

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	573	55% 33% 12%
1	B	573	56% 32% 12%

2 Ensemble composition and analysis (i)

This entry contains 2 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 17912 atoms, of which 9028 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Phosphoenolpyruvate-protein phosphotransferase.

Mol	Chain	Residues	Atoms						Trace
1	A	573	Total	C	H	N	O	S	0
			8956	2790	4514	757	874	21	
1	B	573	Total	C	H	N	O	S	0
			8956	2790	4514	757	874	21	

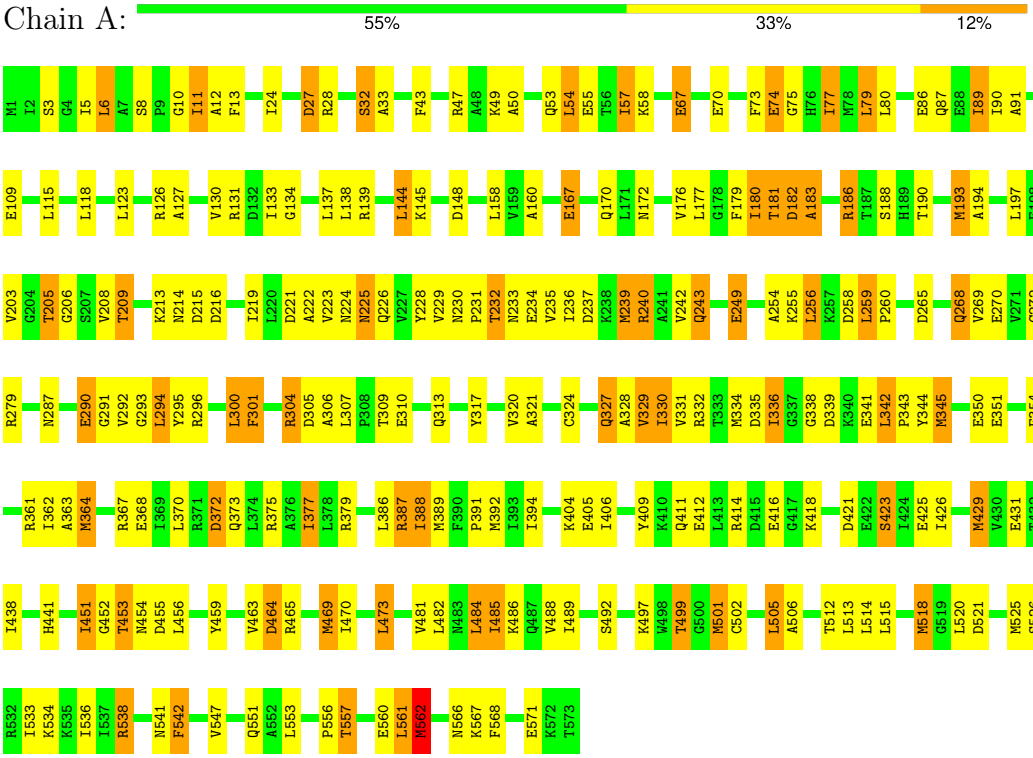
4 Residue-property plots

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.

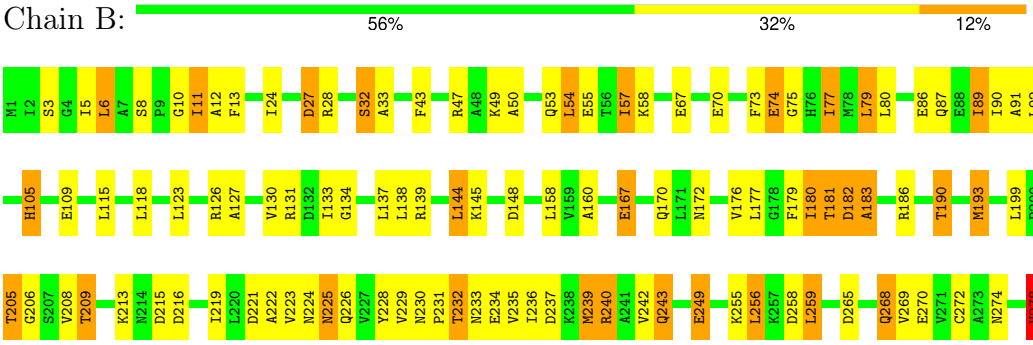
• Molecule 1: Phosphoenolpyruvate-protein phosphotransferase

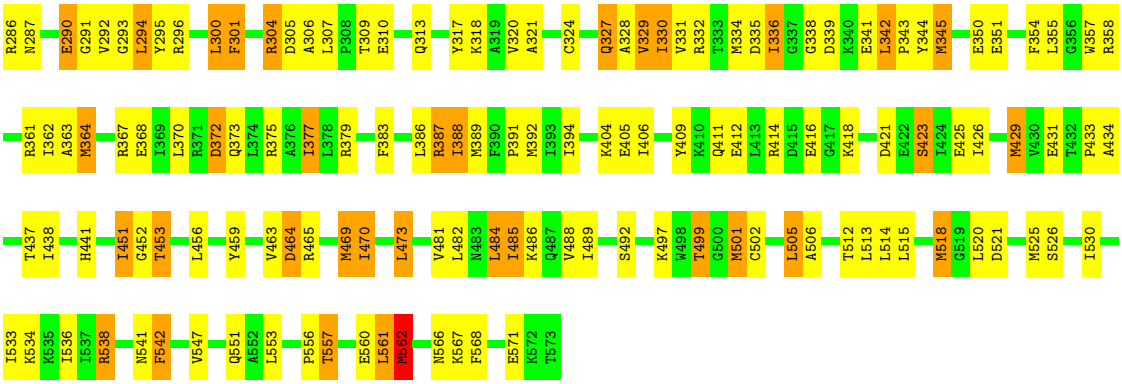
Chain A:



• Molecule 1: Phosphoenolpyruvate-protein phosphotransferase

Chain B:



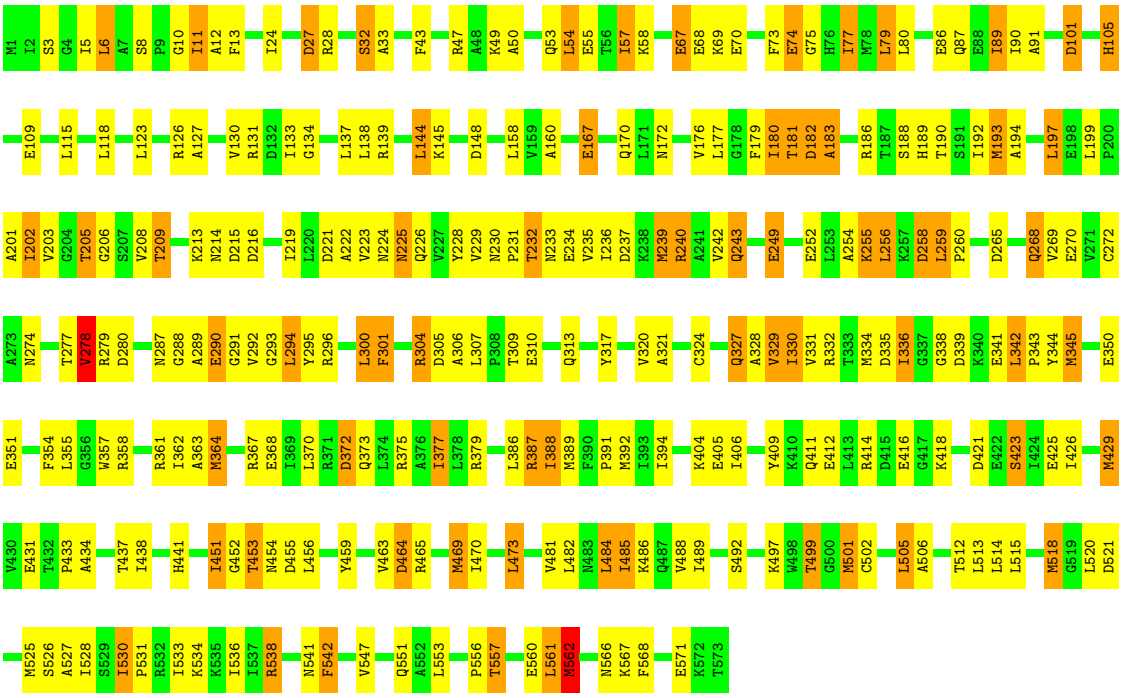


4.2 Residue scores for the representative (author defined) model from the NMR ensemble

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

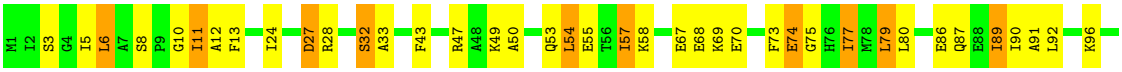
• Molecule 1: Phosphoenolpyruvate-protein phosphotransferase

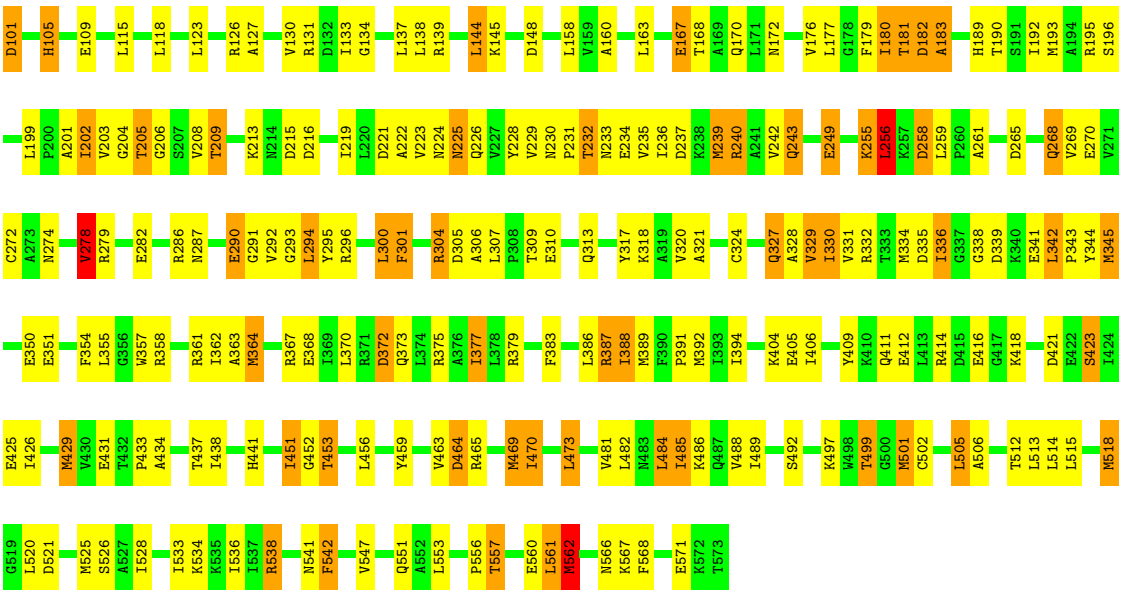
Chain A:



• Molecule 1: Phosphoenolpyruvate-protein phosphotransferase

Chain B:




D101 H105 E109 L115 L118 L123 R126 A127 V130 R131 D132 I133 G134 L137 L138 R139 L144 K145 D148 L158 Y159 A160 L163 E167 T168 A169 Q170 L171 N172 V176 L177 G178 F179 I180 T181 D182 A183 H189 T190 S191 I192 M193 A194 R195 S196
L199 P200 A201 I202 V203 G204 T205 G206 S207 V208 T209 K213 N214 D215 D216 I219 L220 D221 A222 V223 N224 N225 Q226 V227 Y228 V229 N230 P231 T232 N233 E234 V235 I236 D237 K238 M239 R240 A241 V242 Q243 E249 K255 L256 K257 D258 L259 P260 A261 D265 Q268 V269 E270 V271
C272 A273 N274 V278 R279 E282 R286 N287 E290 G291 V292 G293 L294 Y295 R296 L300 F301 R304 D305 A306 L307 P308 T309 E310 Q313 Y317 K318 A319 V320 A321 C324 Q327 A328 V329 I330 V331 R332 T333 M334 D335 I336 G337 G338 D339 K340 E341 L342 P343 Y344 M345
E350 E351 F354 L355 G356 W357 R358 R361 I362 A363 M364 R367 E368 I369 L370 R371 D372 Q373 L374 R375 A376 I377 L378 R379 F383 L386 R387 I388 M389 F390 P391 M392 I393 I394 K404 E405 I406 Y409 K410 Q411 E412 L413 R414 D415 E416 G417 K418 D421 E422 S423 I424
E425 I426 M429 V430 E431 T432 P433 A434 T437 I438 H441 I451 G452 T453 L456 Y459 V463 D464 R465 M469 I470 L473 V481 L482 M483 L484 I485 K486 Q487 V488 I489 S492 K497 W498 T499 G500 M501 C502 L505 A506 T512 L513 L514 L515 M518
G519 L520 D521 M525 S526 A527 I528 I533 K534 K535 I536 I537 R538 N541 F542 V547 Q551 A552 L553 P556 T557 E560 L561 M562 N566 K567 F568 E571 K572 T573

5 Refinement protocol and experimental data overview

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

Of the 120 calculated structures, 2 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
Xplor-NIH	structure solution	2.25
Xplor-NIH	refinement	2.25

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

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	1.04±0.00	10±0/4495 (0.2± 0.0%)	0.89±0.00	5±0/6061 (0.1± 0.0%)
1	B	1.05±0.00	10±0/4495 (0.2± 0.0%)	0.89±0.00	5±0/6061 (0.1± 0.0%)
All	All	1.04	40/17980 (0.2%)	0.89	20/24244 (0.1%)

5 of 20 unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	B	562	MET	CG-SD	7.12	1.99	1.81	1	2
1	A	562	MET	CG-SD	7.09	1.99	1.81	1	2
1	A	345	MET	SD-CE	6.55	2.14	1.77	1	2
1	B	345	MET	SD-CE	6.55	2.14	1.77	1	2
1	B	469	MET	SD-CE	6.46	2.14	1.77	1	2

5 of 10 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
1	A	561	LEU	CA-CB-CG	10.52	139.50	115.30	1	2
1	B	561	LEU	CA-CB-CG	10.52	139.49	115.30	1	2
1	A	278	VAL	CB-CA-C	-5.45	101.05	111.40	1	2
1	B	278	VAL	CB-CA-C	-5.44	101.06	111.40	1	2
1	A	300	LEU	CA-CB-CG	5.33	127.57	115.30	1	2

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts

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	4442	4514	4506	394±0
1	B	4442	4514	4506	386±4
All	All	17768	18056	18024	1524

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 43.

5 of 803 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:562:MET:SD	1:B:562:MET:CE	1.48	2.01	1	2
1:B:364:MET:SD	1:B:364:MET:CE	1.43	2.06	1	2
1:A:364:MET:SD	1:A:364:MET:CE	1.42	2.06	1	2
1:A:469:MET:SD	1:A:469:MET:CE	1.36	2.14	1	2
1:B:469:MET:CE	1:B:469:MET:SD	1.36	2.14	1	2

6.3 Torsion angles

6.3.1 Protein backbone

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	571/573 (100%)	536±0 (94±0%)	28±0 (5±0%)	7±0 (1±0%)	17	64
1	B	571/573 (100%)	536±0 (94±0%)	27±1 (5±0%)	8±0 (1±0%)	16	63
All	All	2284/2292 (100%)	2146 (94%)	109 (5%)	29 (1%)	16	63

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	6	LEU	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	148	ASP	2
1	A	183	ALA	2
1	A	232	THR	2
1	A	278	VAL	2

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	474/474 (100%)	400±2 (84±0%)	74±2 (16±0%)	5 43
1	B	474/474 (100%)	400±0 (84±0%)	74±0 (16±0%)	5 43
All	All	1896/1896 (100%)	1601 (84%)	295 (16%)	5 43

5 of 155 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	11	ILE	2
1	A	27	ASP	2
1	A	28	ARG	2
1	A	32	SER	2
1	A	54	LEU	2

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

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided