Full wwPDB NMR Structure Validation Report (i)

Mar 5, 2026 – 05:46 PM UTC

PDB ID	:	1X6G / pdb_00001x6g
Title	:	Solution structures of the SH3 domain of human megakaryocyte-associated tyrosine-protein kinase.
Authors	:	Sato, M.; Koshiba, S.; Inoue, M.; Kigawa, T.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2005-05-17

This is a Full wwPDB NMR Structure Validation 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:

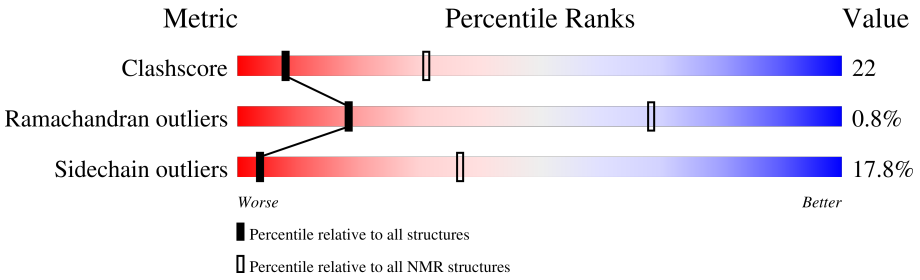
MolProbity	:	4-5-2 with Phenix2.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance

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	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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	81	35% 32% · 31%

2 Ensemble composition and analysis

This entry contains 20 models. Model 2 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:14-A:46, A:53-A:75 (56)	0.26	2

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 4 clusters and 6 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1213 atoms, of which 603 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Megakaryocyte-associated tyrosine-protein kinase.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	81	1213	371	603	121	114	4	0

There are 13 discrepancies between the modelled and reference sequences:

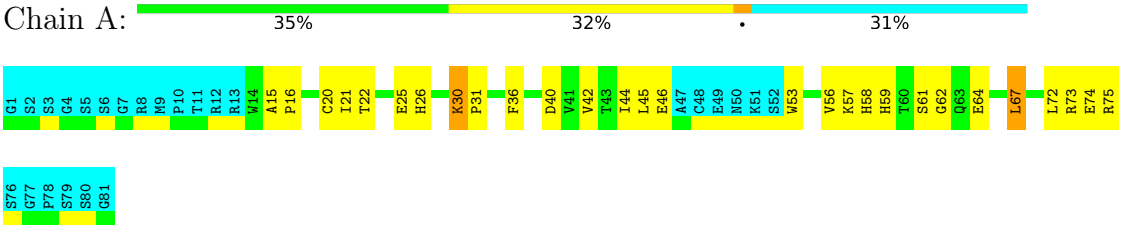
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP P42679
A	2	SER	-	cloning artifact	UNP P42679
A	3	SER	-	cloning artifact	UNP P42679
A	4	GLY	-	cloning artifact	UNP P42679
A	5	SER	-	cloning artifact	UNP P42679
A	6	SER	-	cloning artifact	UNP P42679
A	7	GLY	-	cloning artifact	UNP P42679
A	76	SER	-	cloning artifact	UNP P42679
A	77	GLY	-	cloning artifact	UNP P42679
A	78	PRO	-	cloning artifact	UNP P42679
A	79	SER	-	cloning artifact	UNP P42679
A	80	SER	-	cloning artifact	UNP P42679
A	81	GLY	-	cloning artifact	UNP P42679

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: Megakaryocyte-associated tyrosine-protein kinase

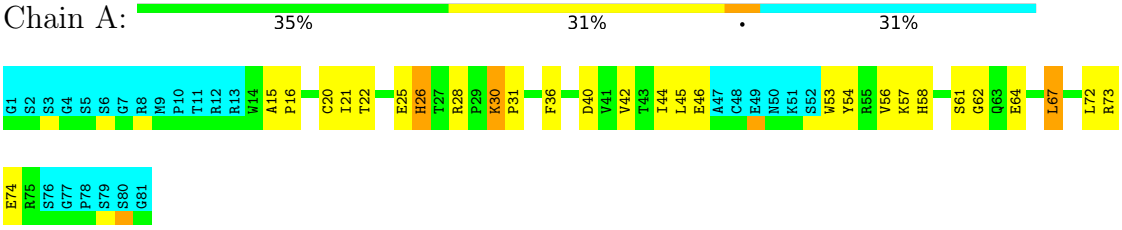


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

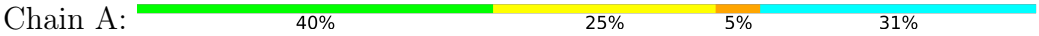
4.2.1 Score per residue for model 1

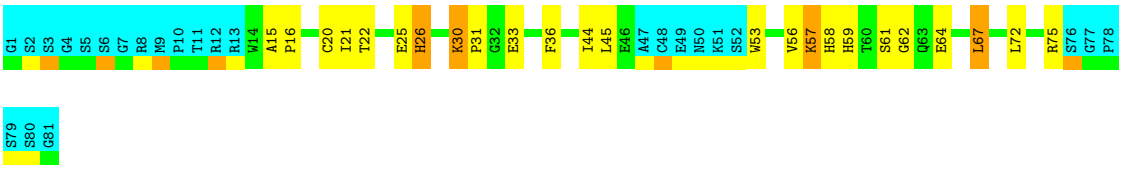
• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase



4.2.2 Score per residue for model 2 (medoid)

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

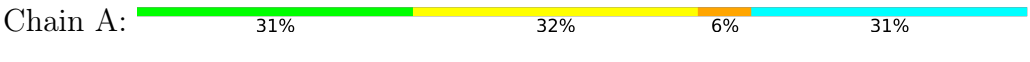


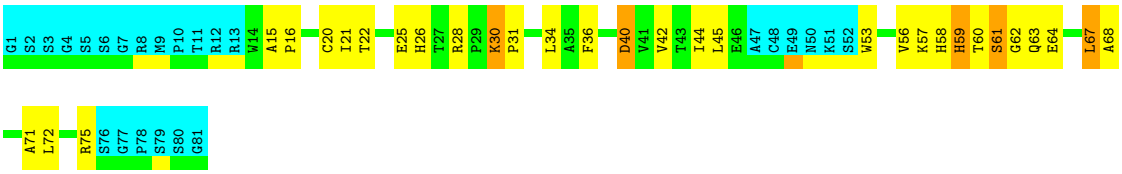


4.2.3 Score per residue for model 3

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

Chain A:

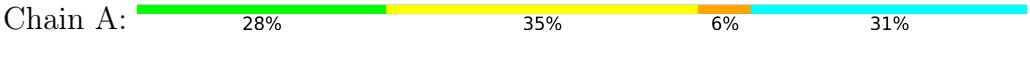


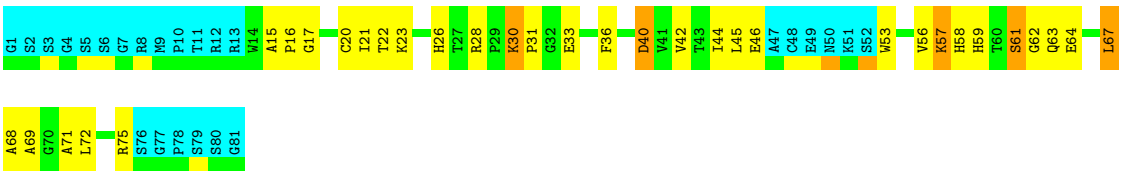


4.2.4 Score per residue for model 4

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

Chain A:

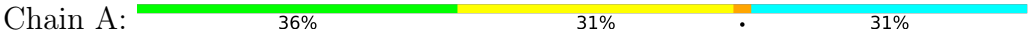


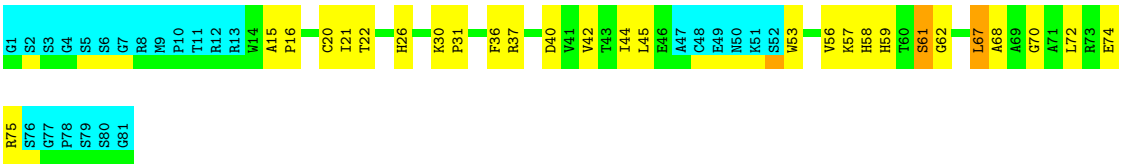


4.2.5 Score per residue for model 5

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

Chain A:

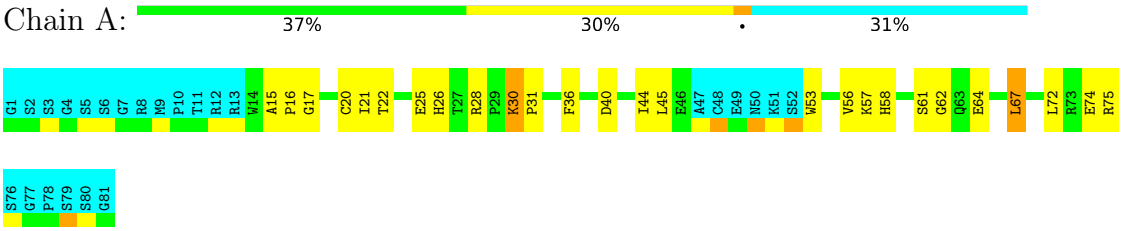




4.2.6 Score per residue for model 6

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

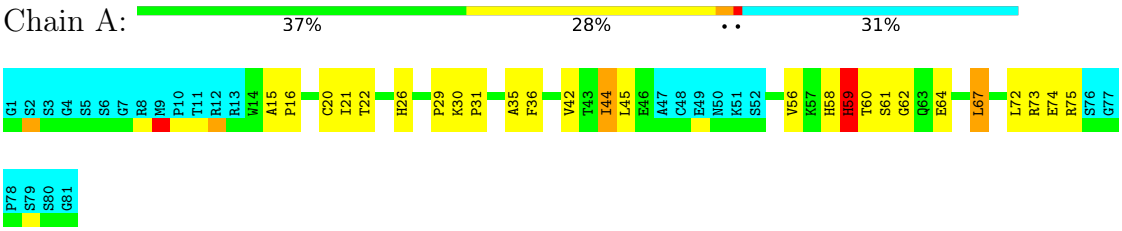
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

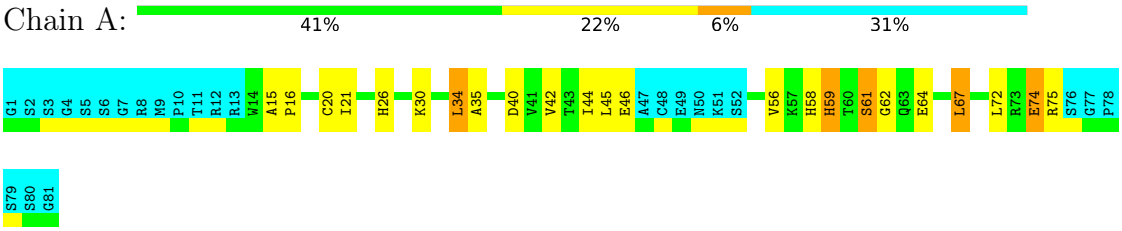
Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

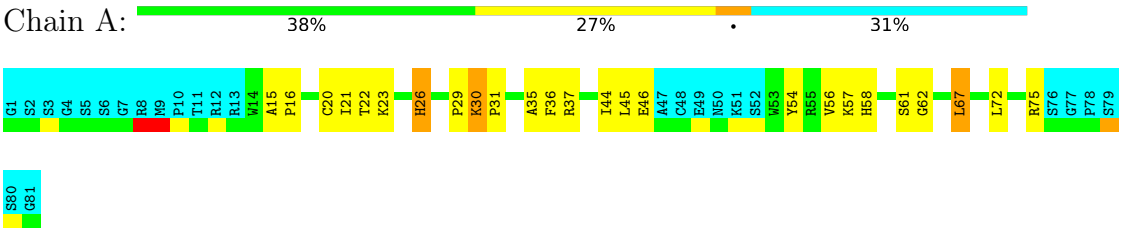
Chain A:



4.2.9 Score per residue for model 9

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

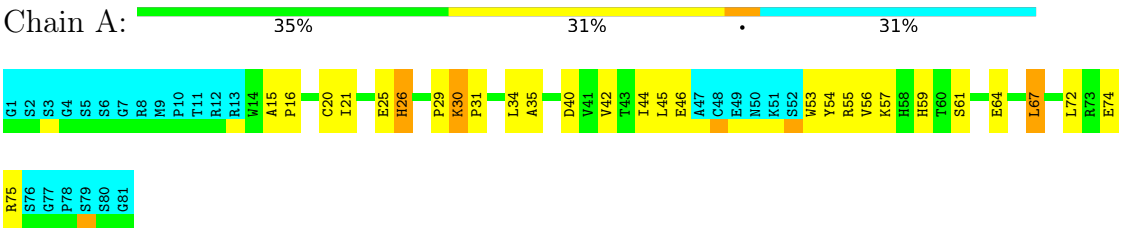
Chain A:



4.2.10 Score per residue for model 10

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

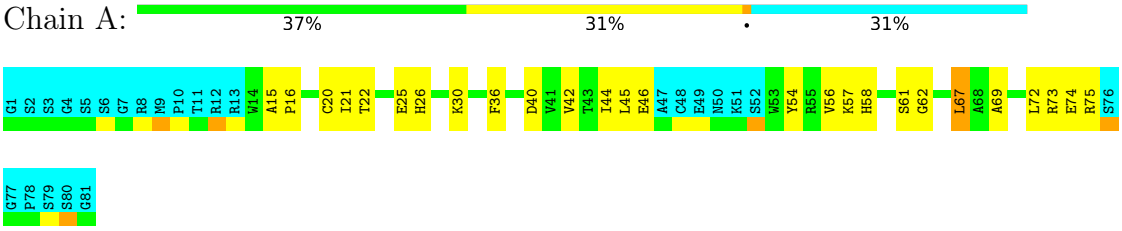
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

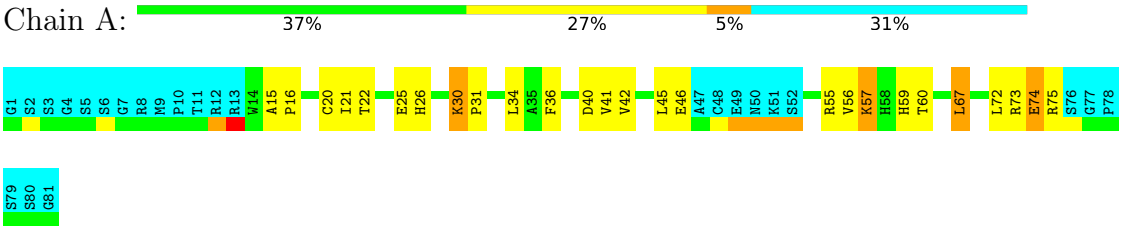
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

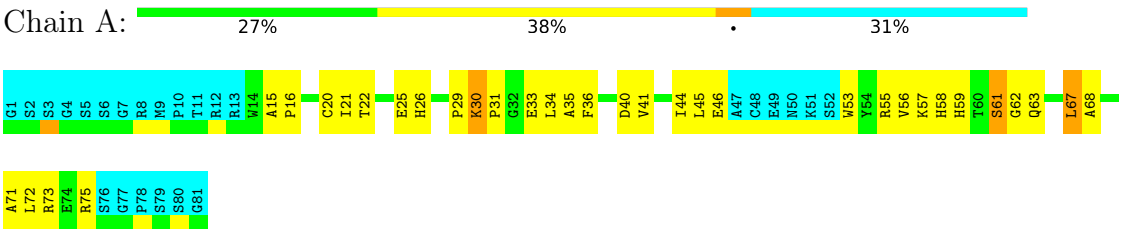
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

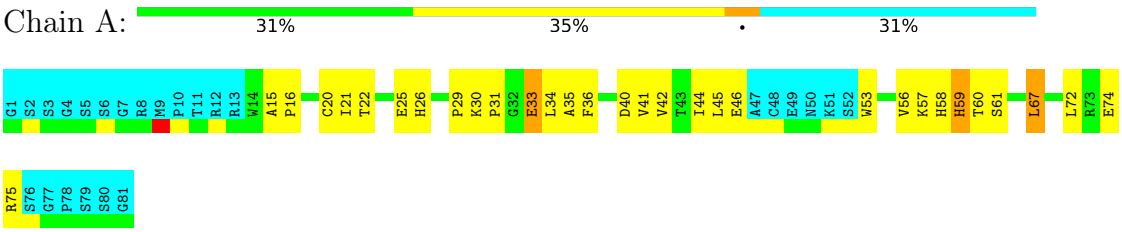
Chain A:



4.2.14 Score per residue for model 14

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

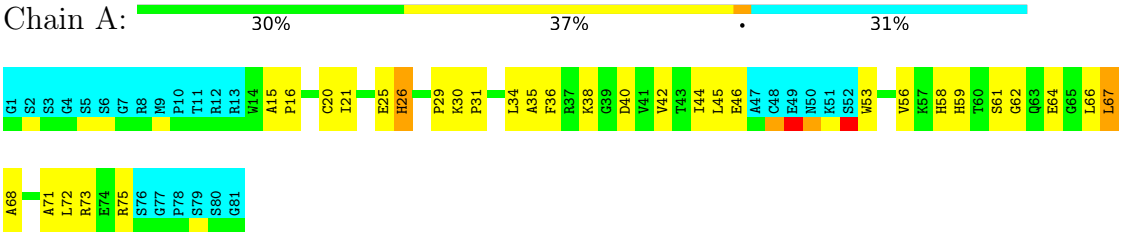
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

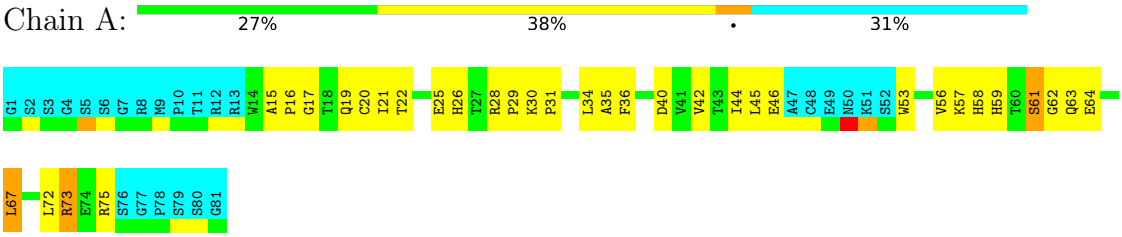
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

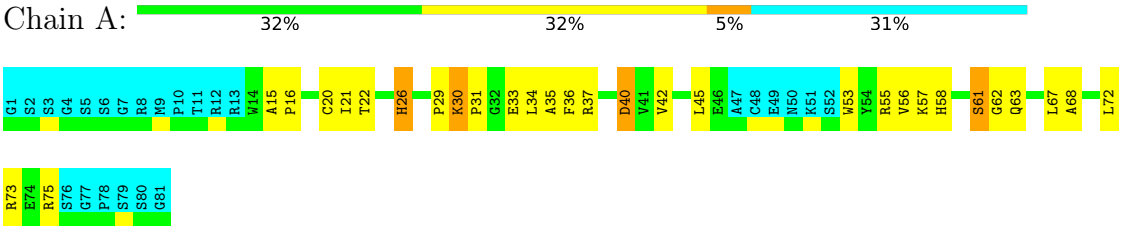
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

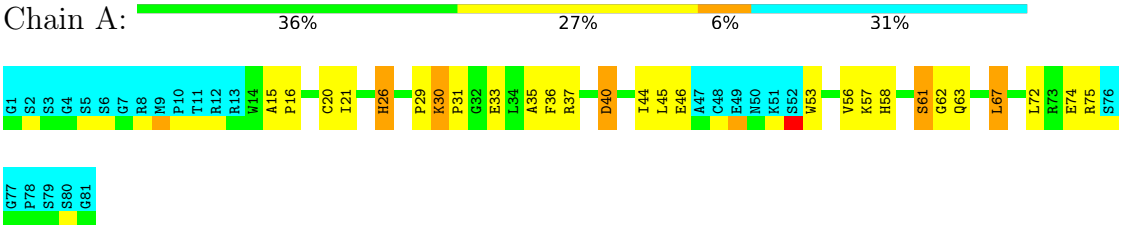
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

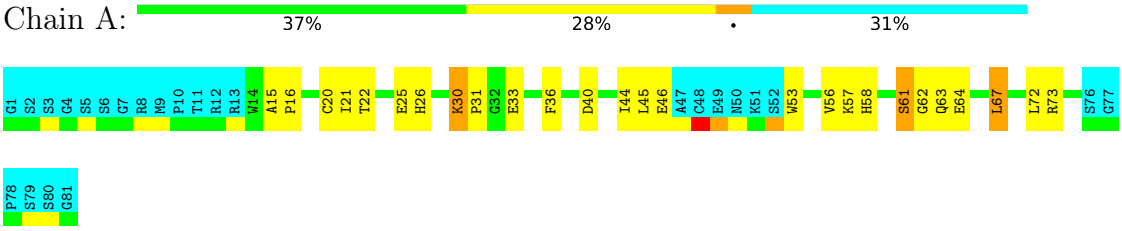
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

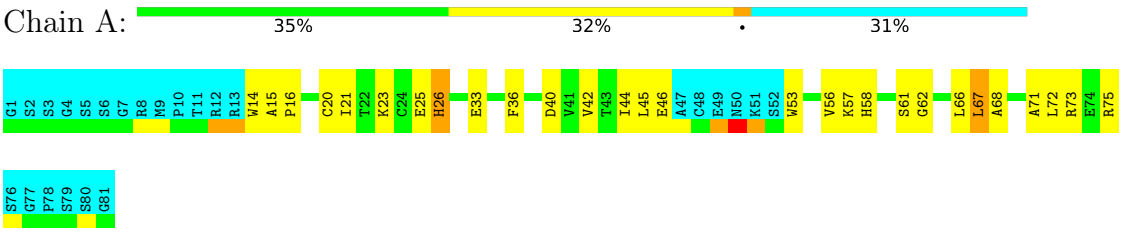
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Megakaryocyte-associated tyrosine-protein kinase

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function, structures with the least restraint violations*.

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

Software name	Classification	Version
CYANA	structure solution	1.0.7
CYANA	refinement	1.0.7

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

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

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	443	444	444	20±4
All	All	8860	8880	8880	395

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 22.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:67:LEU:HD13	1:A:72:LEU:HD21	0.97	1.36	20	20
1:A:20:CYS:O	1:A:21:ILE:HD13	0.85	1.71	17	19
1:A:44:ILE:HD13	1:A:56:VAL:HG12	0.80	1.54	19	18
1:A:20:CYS:SG	1:A:42:VAL:HG23	0.75	2.22	20	3
1:A:44:ILE:HD13	1:A:56:VAL:CG1	0.73	2.12	20	11
1:A:42:VAL:HG21	1:A:56:VAL:HG21	0.72	1.62	12	9
1:A:20:CYS:C	1:A:21:ILE:HD13	0.70	2.11	17	19
1:A:56:VAL:CG2	1:A:67:LEU:HD12	0.69	2.18	9	16
1:A:45:LEU:HD11	1:A:57:LYS:HB2	0.69	1.64	3	7
1:A:15:ALA:HB1	1:A:16:PRO:HD2	0.68	1.65	17	20
1:A:36:PHE:CE2	1:A:67:LEU:HD21	0.68	2.24	4	3
1:A:68:ALA:HB3	1:A:71:ALA:HB3	0.68	1.64	15	5
1:A:22:THR:HG21	1:A:36:PHE:CD2	0.68	2.23	9	2
1:A:56:VAL:HG21	1:A:67:LEU:HD12	0.68	1.63	13	5

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:45:LEU:HD21	1:A:57:LYS:HB2	0.67	1.67	12	11
1:A:45:LEU:HD11	1:A:64:GLU:HG2	0.64	1.69	7	10
1:A:36:PHE:CZ	1:A:67:LEU:HD21	0.63	2.28	13	7
1:A:20:CYS:SG	1:A:72:LEU:HD22	0.62	2.35	15	2
1:A:67:LEU:HD22	1:A:72:LEU:HD21	0.61	1.70	18	6
1:A:45:LEU:HD13	1:A:64:GLU:OE1	0.61	1.96	10	1
1:A:15:ALA:HB1	1:A:16:PRO:CD	0.58	2.28	17	11
1:A:22:THR:HG21	1:A:36:PHE:CD1	0.56	2.36	16	13
1:A:56:VAL:HG22	1:A:67:LEU:HD12	0.55	1.77	11	4
1:A:40:ASP:OD2	1:A:58:HIS:CE1	0.54	2.61	18	2
1:A:33:GLU:OE2	1:A:53:TRP:CH2	0.53	2.61	18	6
1:A:29:PRO:HG3	1:A:35:ALA:HB2	0.53	1.81	17	9
1:A:21:ILE:HD11	1:A:75:ARG:HB2	0.52	1.80	15	16
1:A:53:TRP:CE3	1:A:67:LEU:C	0.52	2.88	1	3
1:A:40:ASP:OD2	1:A:58:HIS:CD2	0.52	2.63	14	4
1:A:16:PRO:HA	1:A:44:ILE:HG22	0.51	1.83	16	6
1:A:33:GLU:CD	1:A:53:TRP:CH2	0.51	2.88	18	4
1:A:33:GLU:N	1:A:33:GLU:CD	0.50	2.70	14	1
1:A:68:ALA:HB3	1:A:71:ALA:CB	0.50	2.34	15	1
1:A:42:VAL:HG12	1:A:58:HIS:HA	0.49	1.83	17	4
1:A:36:PHE:CE2	1:A:67:LEU:CD2	0.49	2.95	4	2
1:A:45:LEU:CD1	1:A:56:VAL:C	0.49	2.86	9	4
1:A:33:GLU:CD	1:A:53:TRP:CZ3	0.49	2.91	18	1
1:A:40:ASP:CG	1:A:58:HIS:NE2	0.48	2.71	13	2
1:A:67:LEU:HB3	1:A:72:LEU:HD11	0.48	1.85	3	3
1:A:45:LEU:HD21	1:A:57:LYS:CB	0.47	2.38	12	1
1:A:33:GLU:OE1	1:A:53:TRP:CH2	0.47	2.67	13	2
1:A:67:LEU:CD2	1:A:72:LEU:HD21	0.47	2.40	18	1
1:A:21:ILE:HD12	1:A:75:ARG:HA	0.47	1.85	16	1
1:A:67:LEU:HD22	1:A:72:LEU:CD2	0.46	2.40	18	1
1:A:44:ILE:HA	1:A:56:VAL:HG12	0.46	1.87	20	1
1:A:66:LEU:HD22	1:A:66:LEU:N	0.46	2.26	20	1
1:A:58:HIS:O	1:A:62:GLY:N	0.45	2.50	18	17
1:A:63:GLN:N	1:A:63:GLN:CD	0.45	2.73	13	1
1:A:40:ASP:OD1	1:A:58:HIS:NE2	0.45	2.50	3	1
1:A:21:ILE:O	1:A:73:ARG:N	0.45	2.50	16	1
1:A:19:GLN:C	1:A:20:CYS:SG	0.45	3.00	16	1
1:A:61:SER:OG	1:A:63:GLN:NE2	0.45	2.50	3	6
1:A:74:GLU:O	1:A:75:ARG:C	0.45	2.60	7	7
1:A:54:TYR:O	1:A:67:LEU:N	0.45	2.50	1	3
1:A:40:ASP:OD2	1:A:58:HIS:NE2	0.45	2.50	6	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:67:LEU:HD23	1:A:68:ALA:N	0.44	2.28	17	1
1:A:17:GLY:N	1:A:44:ILE:O	0.44	2.50	16	3
1:A:37:ARG:N	1:A:40:ASP:OD2	0.44	2.50	5	1
1:A:36:PHE:CZ	1:A:67:LEU:CD2	0.44	3.00	9	1
1:A:40:ASP:OD1	1:A:41:VAL:N	0.44	2.50	14	2
1:A:45:LEU:HD12	1:A:55:ARG:HG2	0.43	1.90	12	2
1:A:30:LYS:CB	1:A:31:PRO:CD	0.43	2.97	18	17
1:A:59:HIS:O	1:A:60:THR:C	0.43	2.62	14	3
1:A:37:ARG:C	1:A:40:ASP:OD1	0.43	2.62	17	1
1:A:26:HIS:N	1:A:36:PHE:O	0.42	2.50	15	4
1:A:42:VAL:CG2	1:A:56:VAL:HG21	0.42	2.39	12	1
1:A:34:LEU:HD12	1:A:35:ALA:H	0.42	1.73	8	1
1:A:26:HIS:HB2	1:A:36:PHE:CZ	0.42	2.50	20	1
1:A:40:ASP:OD1	1:A:58:HIS:CE1	0.41	2.73	4	1
1:A:41:VAL:CG2	1:A:75:ARG:NH1	0.41	2.83	12	1
1:A:37:ARG:N	1:A:40:ASP:OD1	0.41	2.52	18	1
1:A:56:VAL:HG21	1:A:67:LEU:CD1	0.41	2.45	11	2
1:A:36:PHE:CD1	1:A:36:PHE:C	0.41	2.99	3	1
1:A:26:HIS:CD2	1:A:29:PRO:HA	0.41	2.50	10	1
1:A:21:ILE:CG1	1:A:75:ARG:HB2	0.41	2.45	8	2
1:A:29:PRO:CG	1:A:35:ALA:HB2	0.41	2.45	13	1
1:A:58:HIS:CG	1:A:61:SER:HB2	0.40	2.51	5	2
1:A:40:ASP:CG	1:A:58:HIS:CD2	0.40	2.99	13	2
1:A:67:LEU:CD1	1:A:72:LEU:HD21	0.40	2.40	17	2
1:A:68:ALA:C	1:A:70:GLY:N	0.40	2.78	5	1
1:A:20:CYS:O	1:A:42:VAL:HG22	0.40	2.16	4	1
1:A:26:HIS:O	1:A:36:PHE:O	0.40	2.40	2	2
1:A:42:VAL:HG12	1:A:57:LYS:O	0.40	2.16	20	1

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	56/81 (69%)	49±2 (87±3%)	7±2 (12±3%)	0±1 (1±1%)	18	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1120/1620 (69%)	977 (87%)	134 (12%)	9 (1%)	18 68

All 4 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	59	HIS	5
1	A	61	SER	2
1	A	44	ILE	1
1	A	30	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	45/64 (70%)	37±2 (82±4%)	8±2 (18±4%)	4 37
All	All	900/1280 (70%)	740 (82%)	160 (18%)	4 37

All 19 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	26	HIS	20
1	A	61	SER	19
1	A	67	LEU	19
1	A	46	GLU	14
1	A	25	GLU	13
1	A	30	LYS	13
1	A	40	ASP	12
1	A	73	ARG	10
1	A	59	HIS	9
1	A	34	LEU	9
1	A	28	ARG	5
1	A	74	GLU	5
1	A	57	LYS	3
1	A	23	LYS	3
1	A	55	ARG	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	37	ARG	1
1	A	60	THR	1
1	A	33	GLU	1
1	A	38	LYS	1

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no oligosaccharides 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

No chemical shift data were provided