Full wwPDB NMR Structure Validation Report (i)

Feb 9, 2022 – 08:05 AM EST

PDB ID	:	1DG4
Title	:	NMR STRUCTURE OF THE SUBSTRATE BINDING DOMAIN OF DNAK IN THE APO FORM
Authors	:	Pellecchia, M.; Montgomery, D.L.; Stevens, S.Y.; Van der Kooi, C.W.; Feng, H.; Gierasch, L.M.; Zuiderweg, E.R.P.
Deposited on	:	1999-11-23

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

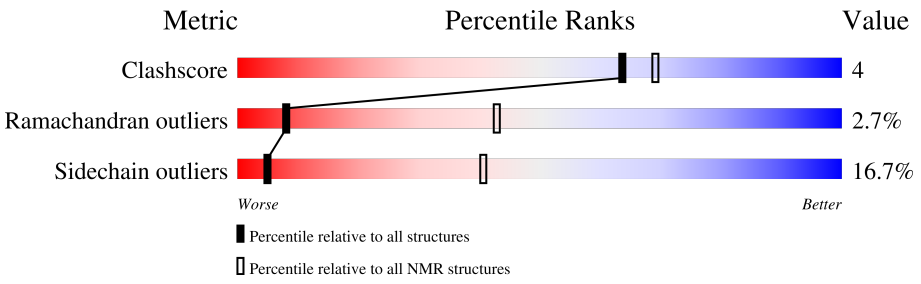
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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	115	75% 14% • 7% •

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:397-A:463, A:471-A:506 (103)	0.89	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 2 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1671 atoms, of which 843 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DNAK.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	111	1671	511	843	147	167	3	0

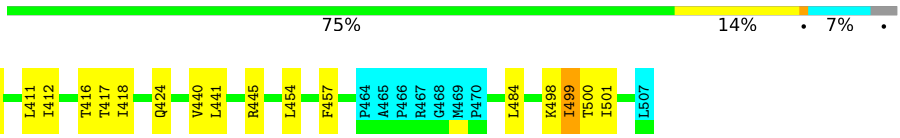
4 Residue-property plots (i)

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: DNAK

Chain A:



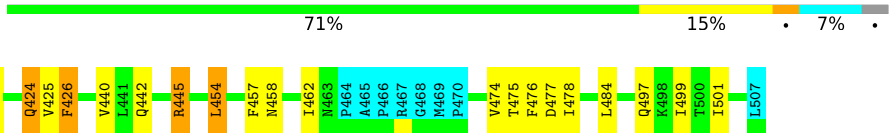
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

• Molecule 1: DNAK

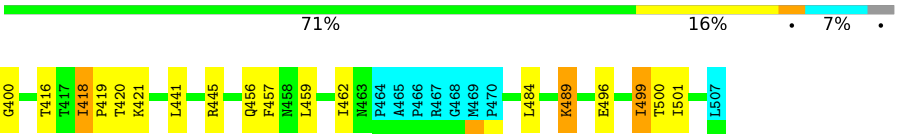
Chain A:



4.2.2 Score per residue for model 2

• Molecule 1: DNAK

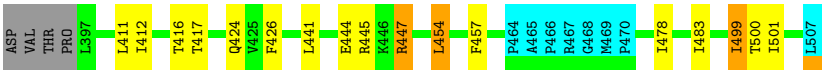
Chain A:



4.2.3 Score per residue for model 3

- Molecule 1: DNAK

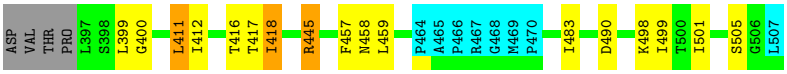
Chain A: 75% 12% • 7% •



4.2.4 Score per residue for model 4

- Molecule 1: DNAK

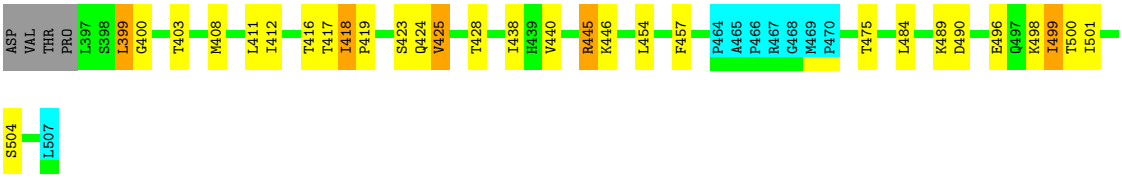
Chain A: 75% 12% • 7% •



4.2.5 Score per residue for model 5

- Molecule 1: DNAK

Chain A: 63% 22% • 7% •



4.2.6 Score per residue for model 6

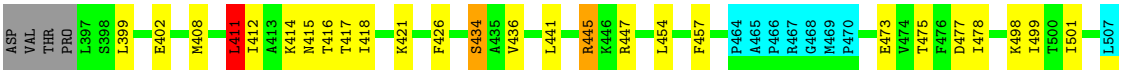
- Molecule 1: DNAK

Chain A: 72% 17% • 7% •

4.2.7 Score per residue for model 7

- Molecule 1: DNAK

Chain A: 67% 20% • • 7% •



4.2.8 Score per residue for model 8

- Molecule 1: DNAK

Chain A:

4.2.9 Score per residue for model 9

- Molecule 1: DNAK

Chain A:

4.2.10 Score per residue for model 10

- Molecule 1: DNAK

Chain A:

4.2.11 Score per residue for model 11

- Molecule 1: DNAK

Chain A:

4.2.12 Score per residue for model 12

- Molecule 1: DNAK

Chain A:

4.2.13 Score per residue for model 13

- Molecule 1: DNAK

Chain A:

4.2.14 Score per residue for model 14

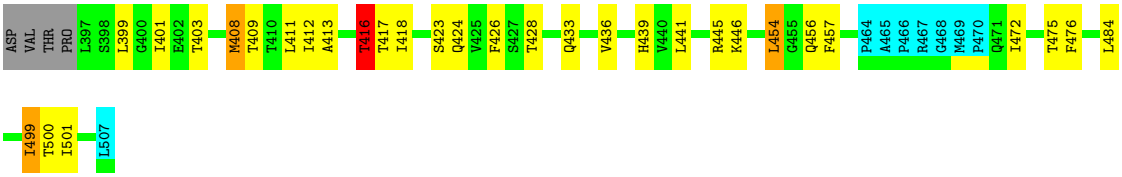
- Molecule 1: DNAK

Chain A:

4.2.15 Score per residue for model 15

- Molecule 1: DNAK

Chain A:



4.2.16 Score per residue for model 16

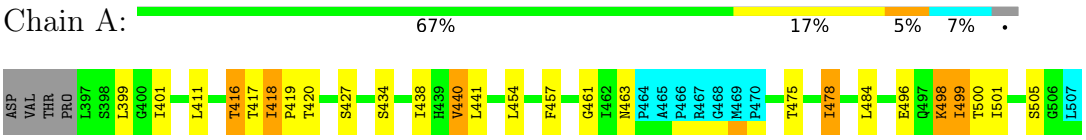
- Molecule 1: DNAK

Chain A:

4.2.17 Score per residue for model 17

- Molecule 1: DNAK

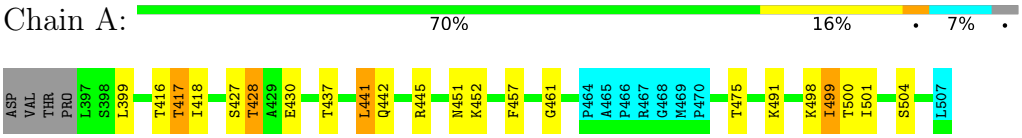
Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: DNAK

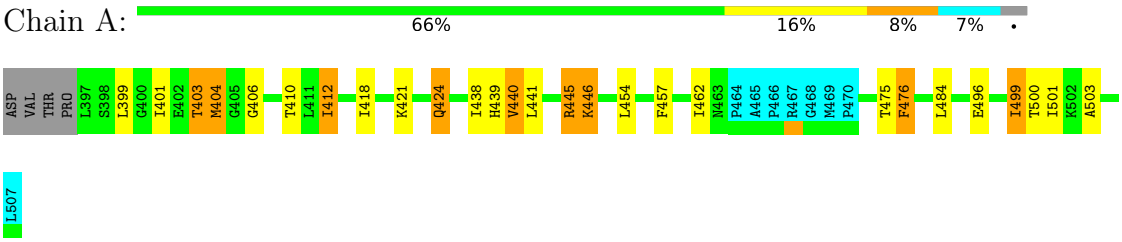
Chain A:



4.2.19 Score per residue for model 19

- Molecule 1: DNAK

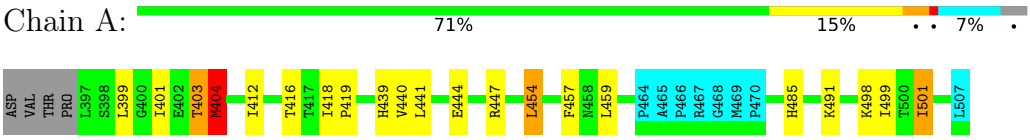
Chain A:



4.2.20 Score per residue for model 20

- Molecule 1: DNAK

Chain A:



5 Refinement protocol and experimental data overview (i)

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

Of the 50 calculated structures, 20 were deposited, based on the following criterion: *STRUCTURES WITH THE LEAST RESTRAINT VIOLATIONS,TARGET FUNCTION*.

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

Software name	Classification	Version
DYANA	structure solution	1.5
FANTOM	refinement	1.0

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

There are no covalent bond-length or bond-angle outliers.

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	A	0.0±0.0	1.8±1.4
All	All	0	36

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	445	ARG	Sidechain	9
1	A	418	ILE	Peptide	5
1	A	438	ILE	Peptide	4
1	A	447	ARG	Sidechain	3
1	A	417	THR	Peptide	2
1	A	416	THR	Peptide	2
1	A	411	LEU	Peptide	1
1	A	441	LEU	Peptide	1
1	A	473	GLU	Peptide	1
1	A	471	GLN	Peptide	1
1	A	424	GLN	Peptide	1
1	A	495	LYS	Peptide	1
1	A	472	ILE	Peptide	1
1	A	419	PRO	Peptide	1
1	A	476	PHE	Peptide	1
1	A	403	THR	Peptide	1
1	A	404	MET	Peptide	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	A	770	781	780	7±3
All	All	15400	15620	15600	132

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 4.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:412:ILE:O	1:A:412:ILE:HG23	0.74	1.80	19	1
1:A:411:LEU:O	1:A:412:ILE:HG23	0.72	1.84	7	1
1:A:424:GLN:O	1:A:425:VAL:HG12	0.67	1.89	5	1
1:A:411:LEU:O	1:A:411:LEU:HD22	0.66	1.90	4	1
1:A:440:VAL:HG21	1:A:501:ILE:HD13	0.65	1.67	11	1
1:A:436:VAL:O	1:A:436:VAL:HG23	0.63	1.94	9	3
1:A:411:LEU:HD22	1:A:411:LEU:C	0.61	2.16	4	1
1:A:454:LEU:N	1:A:454:LEU:HD13	0.61	2.11	15	5
1:A:399:LEU:HD22	1:A:440:VAL:HB	0.60	1.73	17	2
1:A:403:THR:O	1:A:404:MET:C	0.60	2.39	8	4
1:A:426:PHE:CD2	1:A:476:PHE:CE2	0.59	2.90	1	3
1:A:399:LEU:HD11	1:A:440:VAL:HG13	0.59	1.75	10	1
1:A:412:ILE:O	1:A:412:ILE:CG2	0.59	2.51	19	1
1:A:499:ILE:HD13	1:A:500:THR:N	0.56	2.15	3	11
1:A:401:ILE:HD13	1:A:476:PHE:CZ	0.56	2.36	14	1
1:A:401:ILE:HD11	1:A:426:PHE:CE1	0.55	2.37	15	1
1:A:411:LEU:HD13	1:A:411:LEU:N	0.53	2.18	7	1
1:A:483:ILE:HD12	1:A:483:ILE:N	0.53	2.19	3	3
1:A:440:VAL:HG21	1:A:501:ILE:CD1	0.53	2.32	11	1
1:A:401:ILE:HG21	1:A:476:PHE:CE1	0.52	2.40	19	1
1:A:424:GLN:O	1:A:425:VAL:C	0.52	2.48	13	2
1:A:472:ILE:O	1:A:472:ILE:HG23	0.51	2.06	14	1
1:A:411:LEU:HD13	1:A:411:LEU:H	0.51	1.66	7	1
1:A:399:LEU:HD13	1:A:400:GLY:N	0.50	2.20	6	3
1:A:416:THR:O	1:A:417:THR:C	0.50	2.50	12	5
1:A:399:LEU:HD23	1:A:400:GLY:N	0.50	2.21	4	2
1:A:454:LEU:C	1:A:454:LEU:HD22	0.50	2.27	1	3

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:437:THR:O	1:A:437:THR:HG23	0.50	2.06	9	1
1:A:428:THR:HG22	1:A:436:VAL:HG21	0.50	1.83	15	1
1:A:401:ILE:HG22	1:A:440:VAL:CG1	0.49	2.37	19	1
1:A:426:PHE:CE2	1:A:476:PHE:CZ	0.49	2.99	1	2
1:A:475:THR:OG1	1:A:489:LYS:NZ	0.49	2.45	5	1
1:A:420:THR:HG22	1:A:478:ILE:HG22	0.49	1.84	17	1
1:A:417:THR:O	1:A:418:ILE:CG1	0.48	2.62	14	2
1:A:416:THR:HG21	1:A:420:THR:HG21	0.48	1.86	17	1
1:A:417:THR:C	1:A:418:ILE:HG22	0.48	2.29	9	1
1:A:416:THR:O	1:A:418:ILE:N	0.47	2.47	4	3
1:A:411:LEU:HD23	1:A:411:LEU:C	0.47	2.31	15	1
1:A:420:THR:HG22	1:A:478:ILE:CG2	0.46	2.40	17	1
1:A:472:ILE:O	1:A:472:ILE:CG2	0.46	2.64	14	1
1:A:454:LEU:O	1:A:454:LEU:HD13	0.46	2.10	17	2
1:A:401:ILE:HD13	1:A:476:PHE:CE1	0.46	2.46	9	1
1:A:436:VAL:HG13	1:A:436:VAL:O	0.46	2.10	15	1
1:A:399:LEU:HD13	1:A:399:LEU:C	0.45	2.31	16	4
1:A:436:VAL:O	1:A:436:VAL:CG2	0.45	2.63	9	1
1:A:399:LEU:HD22	1:A:440:VAL:CB	0.45	2.41	9	1
1:A:418:ILE:HG23	1:A:419:PRO:HA	0.45	1.88	8	2
1:A:476:PHE:CD2	1:A:486:VAL:HG22	0.45	2.46	14	1
1:A:414:LYS:O	1:A:415:ASN:CB	0.44	2.65	7	1
1:A:454:LEU:N	1:A:454:LEU:CD1	0.44	2.81	1	4
1:A:489:LYS:NZ	1:A:496:GLU:OE1	0.44	2.51	2	1
1:A:399:LEU:HD21	1:A:454:LEU:HD11	0.44	1.88	9	1
1:A:445:ARG:O	1:A:446:LYS:CB	0.44	2.65	19	1
1:A:418:ILE:N	1:A:419:PRO:O	0.43	2.52	13	1
1:A:417:THR:C	1:A:418:ILE:CG2	0.43	2.85	9	1
1:A:496:GLU:OE2	1:A:498:LYS:NZ	0.43	2.50	17	1
1:A:418:ILE:HA	1:A:419:PRO:C	0.43	2.34	17	3
1:A:417:THR:O	1:A:418:ILE:C	0.43	2.57	13	1
1:A:408:MET:O	1:A:409:THR:C	0.42	2.57	15	1
1:A:399:LEU:C	1:A:399:LEU:HD23	0.42	2.34	2	1
1:A:441:LEU:CD1	1:A:441:LEU:N	0.42	2.82	18	1
1:A:401:ILE:HG22	1:A:440:VAL:HG12	0.42	1.92	19	1
1:A:424:GLN:O	1:A:426:PHE:N	0.42	2.52	10	1
1:A:426:PHE:CE2	1:A:474:VAL:HG12	0.41	2.50	1	1
1:A:399:LEU:HD11	1:A:440:VAL:HB	0.41	1.92	19	1
1:A:445:ARG:O	1:A:447:ARG:N	0.41	2.52	13	2
1:A:459:LEU:C	1:A:459:LEU:HD13	0.41	2.34	9	1
1:A:401:ILE:CD1	1:A:426:PHE:CE1	0.41	3.03	14	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:413:ALA:N	1:A:416:THR:OG1	0.41	2.54	15	1
1:A:437:THR:O	1:A:437:THR:CG2	0.41	2.69	9	1
1:A:454:LEU:C	1:A:454:LEU:CD2	0.41	2.89	1	1
1:A:411:LEU:HD13	1:A:411:LEU:C	0.41	2.35	3	1
1:A:423:SER:OG	1:A:424:GLN:N	0.41	2.54	5	1
1:A:426:PHE:CE1	1:A:476:PHE:CZ	0.41	3.09	10	1
1:A:424:GLN:N	1:A:425:VAL:HG22	0.41	2.31	13	1
1:A:427:SER:OG	1:A:428:THR:N	0.41	2.54	18	1
1:A:424:GLN:O	1:A:425:VAL:O	0.40	2.39	5	1
1:A:399:LEU:HD21	1:A:501:ILE:HG13	0.40	1.92	20	1
1:A:483:ILE:N	1:A:483:ILE:CD1	0.40	2.85	3	1
1:A:484:LEU:HD12	1:A:486:VAL:HG23	0.40	1.93	13	1
1:A:416:THR:HG22	1:A:420:THR:HG21	0.40	1.93	2	1
1:A:438:ILE:HD12	1:A:438:ILE:N	0.40	2.31	5	1
1:A:499:ILE:O	1:A:499:ILE:CG2	0.40	2.70	19	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	102/115 (89%)	86±3 (85±3%)	13±3 (13±3%)	3±2 (3±2%)	8	43
All	All	2040/2300 (89%)	1725 (85%)	260 (13%)	55 (3%)	8	43

All 21 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	418	ILE	8
1	A	417	THR	7
1	A	425	VAL	4
1	A	446	LYS	4
1	A	434	SER	4
1	A	461	GLY	4
1	A	404	MET	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	410	THR	3
1	A	398	SER	2
1	A	505	SER	2
1	A	506	GLY	2
1	A	471	GLN	2
1	A	416	THR	2
1	A	504	SER	1
1	A	427	SER	1
1	A	431	ASP	1
1	A	445	ARG	1
1	A	423	SER	1
1	A	406	GLY	1
1	A	424	GLN	1
1	A	503	ALA	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	86/96 (90%)	72±3 (83±3%)	14±3 (17±3%)	5 41
All	All	1720/1920 (90%)	1432 (83%)	288 (17%)	5 41

All 64 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	457	PHE	20
1	A	499	ILE	20
1	A	501	ILE	19
1	A	441	LEU	14
1	A	454	LEU	13
1	A	484	LEU	12
1	A	412	ILE	11
1	A	498	LYS	11
1	A	440	VAL	9
1	A	445	ARG	8
1	A	475	THR	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	411	LEU	7
1	A	424	GLN	6
1	A	478	ILE	6
1	A	408	MET	6
1	A	496	GLU	6
1	A	418	ILE	5
1	A	442	GLN	5
1	A	421	LYS	5
1	A	403	THR	5
1	A	401	ILE	5
1	A	397	LEU	4
1	A	473	GLU	4
1	A	439	HIS	4
1	A	426	PHE	3
1	A	462	ILE	3
1	A	477	ASP	3
1	A	497	GLN	3
1	A	456	GLN	3
1	A	459	LEU	3
1	A	489	LYS	3
1	A	444	GLU	3
1	A	490	ASP	3
1	A	399	LEU	3
1	A	428	THR	3
1	A	472	ILE	3
1	A	450	ASP	3
1	A	458	ASN	2
1	A	423	SER	2
1	A	416	THR	2
1	A	433	GLN	2
1	A	500	THR	2
1	A	451	ASN	2
1	A	463	ASN	2
1	A	417	THR	2
1	A	491	LYS	2
1	A	402	GLU	1
1	A	434	SER	1
1	A	495	LYS	1
1	A	398	SER	1
1	A	436	VAL	1
1	A	414	LYS	1
1	A	420	THR	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	471	GLN	1
1	A	446	LYS	1
1	A	427	SER	1
1	A	438	ILE	1
1	A	505	SER	1
1	A	430	GLU	1
1	A	437	THR	1
1	A	452	LYS	1
1	A	504	SER	1
1	A	447	ARG	1
1	A	485	HIS	1

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