Full wwPDB NMR Structure Validation Report (i)

Mar 5, 2022 – 09:07 AM EST

PDB ID	:	2K5B
Title	:	Human CDC37-HSP90 docking model based on NMR
Authors	:	Sreeramulu, S.; Jonker, H.R.A.; Lancaster, C.R.; Richter, C.; Langer, T.; Schwalbe, H.
Deposited on	:	2008-06-26

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:

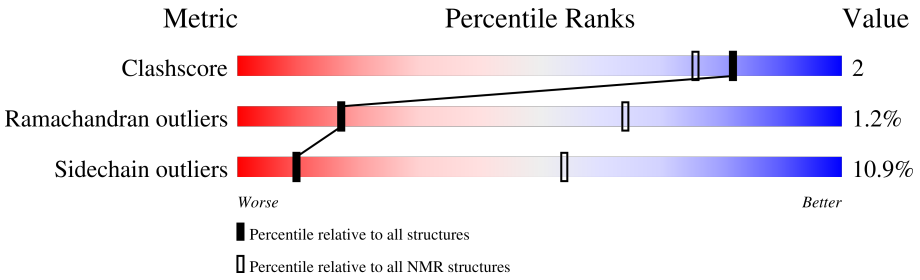
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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	210	89% 10%
2	B	129	88% 5% 8%

2 Ensemble composition and analysis

This entry contains 10 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *best haddock score*.

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:223, B:158-B:276 (329)	0.53	3

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	3, 5, 7, 8, 9
2	2, 4, 6, 10
Single-model clusters	1

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5499 atoms, of which 2757 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Heat shock protein HSP 90-alpha.

Mol	Chain	Residues	Atoms						Trace
1	A	210	Total	C	H	N	O	S	0
			3310	1049	1655	271	330	5	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	63	THR	SER	SEE REMARK 999	UNP P07900

- Molecule 2 is a protein called Hsp90 co-chaperone Cdc37.

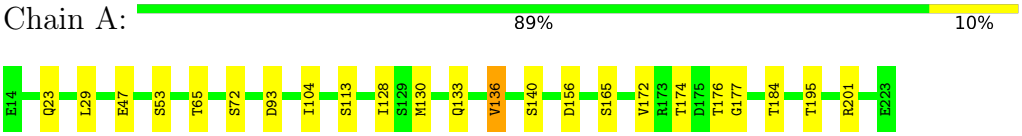
Mol	Chain	Residues	Atoms						Trace
2	B	129	Total	C	H	N	O	S	0
			2189	694	1102	190	194	9	

4 Residue-property plots

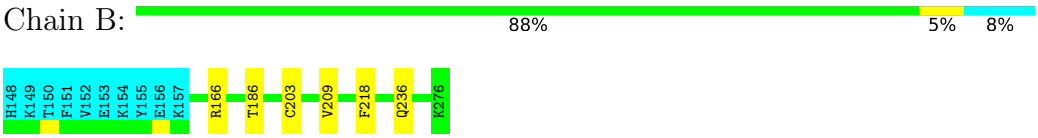
4.1 Average score per residue in the NMR ensemble

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: Heat shock protein HSP 90-alpha



• Molecule 2: Hsp90 co-chaperone Cdc37

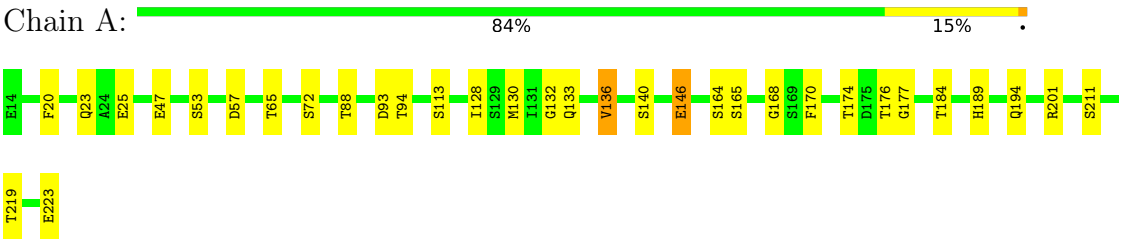


4.2 Scores per residue for each member of the ensemble

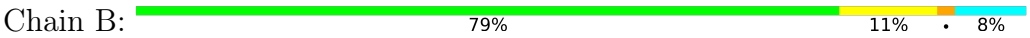
Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Heat shock protein HSP 90-alpha



• Molecule 2: Hsp90 co-chaperone Cdc37



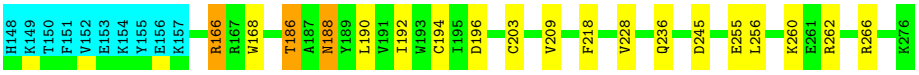
4.2.2 Score per residue for model 2

• Molecule 1: Heat shock protein HSP 90-alpha

Chain A:

• Molecule 2: Hsp90 co-chaperone Cdc37

Chain B:



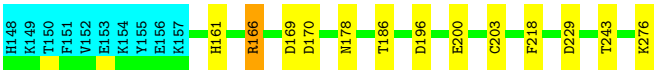
4.2.3 Score per residue for model 3 (medoid)

• Molecule 1: Heat shock protein HSP 90-alpha

Chain A:

• Molecule 2: Hsp90 co-chaperone Cdc37

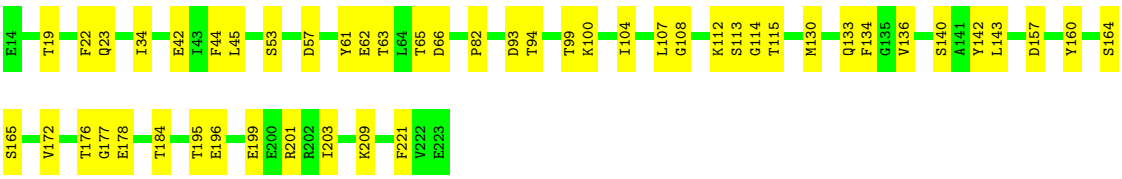
Chain B:



4.2.4 Score per residue for model 4

• Molecule 1: Heat shock protein HSP 90-alpha

Chain A:



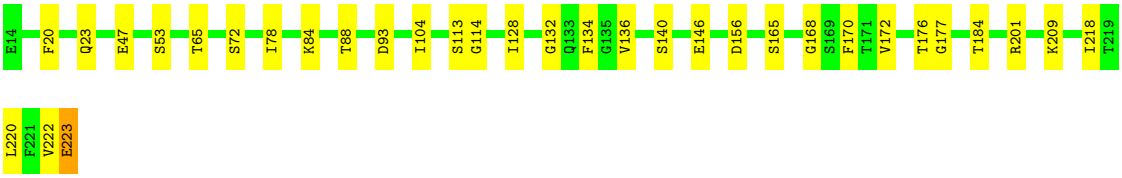
• Molecule 2: Hsp90 co-chaperone Cdc37

Chain B:

4.2.5 Score per residue for model 5

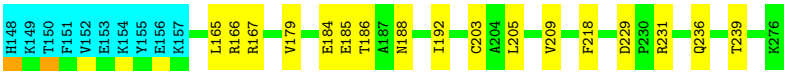
• Molecule 1: Heat shock protein HSP 90-alpha

Chain A:



• Molecule 2: Hsp90 co-chaperone Cdc37

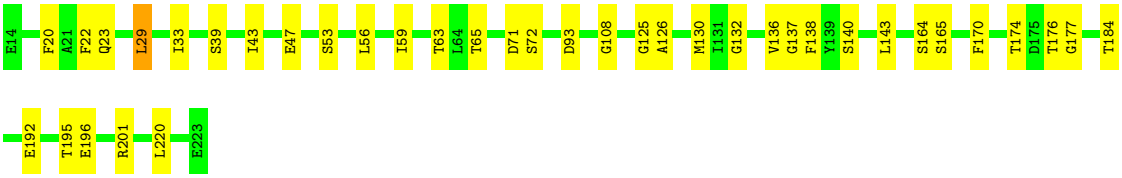
Chain B:



4.2.6 Score per residue for model 6

• Molecule 1: Heat shock protein HSP 90-alpha

Chain A:

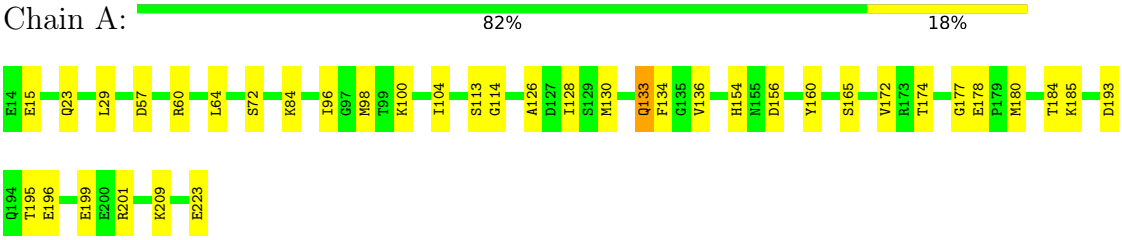


• Molecule 2: Hsp90 co-chaperone Cdc37

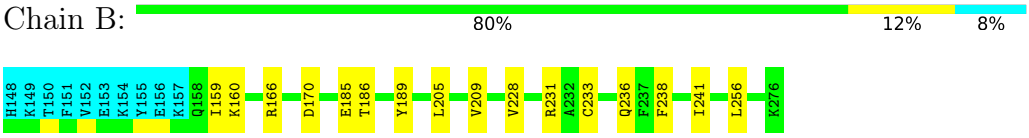
Chain B:

4.2.7 Score per residue for model 7

• Molecule 1: Heat shock protein HSP 90-alpha

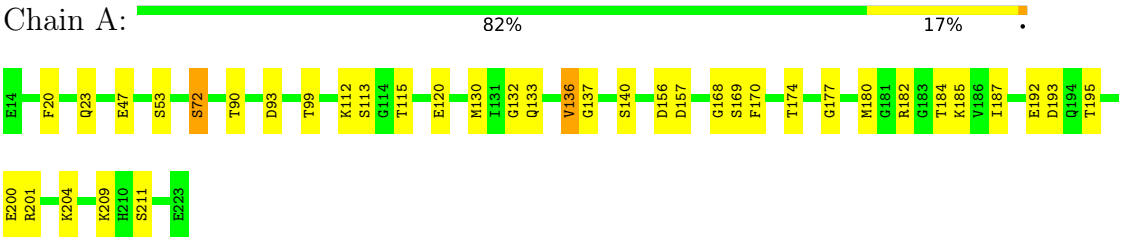


• Molecule 2: Hsp90 co-chaperone Cdc37

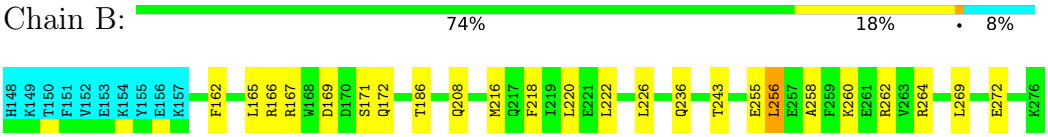


4.2.8 Score per residue for model 8

• Molecule 1: Heat shock protein HSP 90-alpha

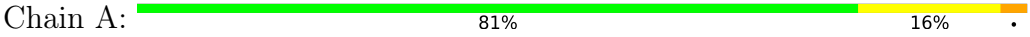


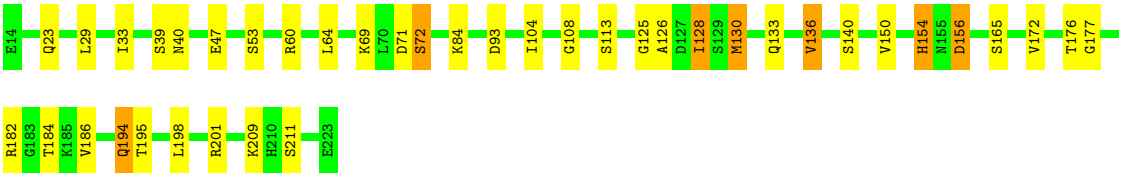
• Molecule 2: Hsp90 co-chaperone Cdc37



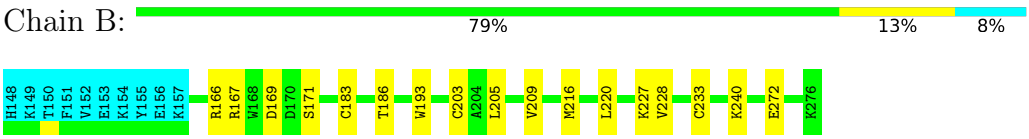
4.2.9 Score per residue for model 9

• Molecule 1: Heat shock protein HSP 90-alpha



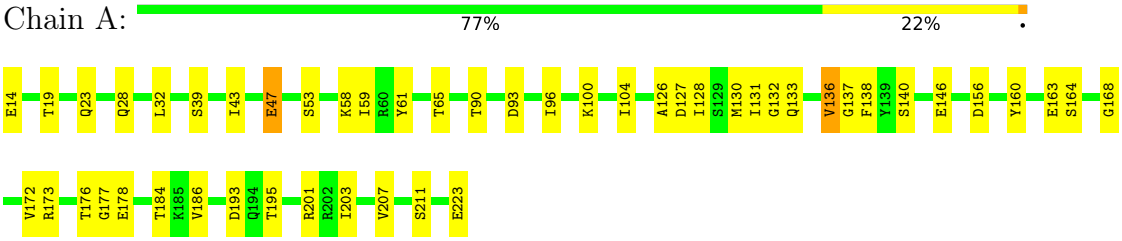


• Molecule 2: Hsp90 co-chaperone Cdc37

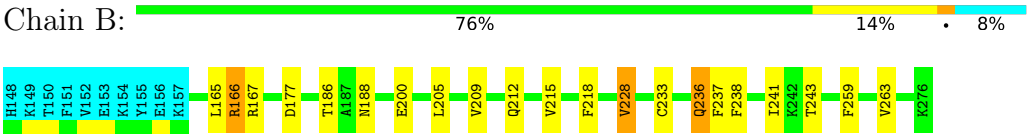


4.2.10 Score per residue for model 10

• Molecule 1: Heat shock protein HSP 90-alpha



• Molecule 2: Hsp90 co-chaperone Cdc37



5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 10 were deposited, based on the following criterion: *structures with the best HADDOCK scoring.*

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

Software name	Classification	Version
CNS	structure solution	1.1
CNS	refinement	1.1
HADDOCK	geometry optimization	2.0
HADDOCK	refinement	2.0
HADDOCK	structure solution	2.0

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	1655	1655	1648	9±3
2	B	995	1007	1004	5±2
All	All	26500	26620	26520	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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:104:ILE:HG12	1:A:172:VAL:HG21	0.70	1.62	5	6
1:A:133:GLN:NE2	2:B:170:ASP:HB3	0.58	2.14	1	1
1:A:114:GLY:HA3	1:A:134:PHE:O	0.58	1.99	5	5
2:B:228:VAL:HB	2:B:233:CYS:SG	0.57	2.39	7	3
1:A:133:GLN:O	2:B:166:ARG:HB3	0.57	2.00	4	7
1:A:180:MET:SD	1:A:185:LYS:HE2	0.57	2.40	3	3
1:A:47:GLU:OE1	1:A:132:GLY:HA3	0.55	2.02	1	3
2:B:194:CYS:SG	2:B:209:VAL:HB	0.55	2.41	1	2
1:A:126:ALA:HA	2:B:164:MET:SD	0.55	2.41	6	1
2:B:160:LYS:HE3	2:B:189:TYR:OH	0.55	2.02	7	1
2:B:222:LEU:O	2:B:226:LEU:HG	0.54	2.01	8	1
1:A:20:PHE:HB2	1:A:170:PHE:CZ	0.54	2.38	8	4
1:A:126:ALA:HB1	1:A:130:MET:SD	0.54	2.42	10	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:186:THR:O	2:B:190:LEU:HG	0.53	2.04	2	1
1:A:99:THR:HB	1:A:157:ASP:OD2	0.52	2.04	4	3
2:B:256:LEU:O	2:B:260:LYS:HG3	0.52	2.04	8	3
1:A:60:ARG:O	1:A:64:LEU:HG	0.51	2.05	7	3
2:B:238:PHE:O	2:B:241:ILE:HG22	0.51	2.05	7	3
1:A:126:ALA:HB1	1:A:130:MET:HG3	0.51	1.82	9	1
1:A:63:THR:HA	1:A:66:ASP:O	0.51	2.06	4	2
1:A:112:LYS:HB2	1:A:115:THR:OG1	0.50	2.07	8	2
1:A:56:LEU:O	1:A:59:ILE:HG22	0.50	2.07	6	1
1:A:132:GLY:HA2	1:A:137:GLY:CA	0.50	2.36	10	3
1:A:59:ILE:HD13	1:A:96:ILE:HB	0.50	1.83	10	1
2:B:216:MET:O	2:B:220:LEU:HG	0.50	2.06	8	2
1:A:128:ILE:O	1:A:131:ILE:HG22	0.50	2.06	10	1
2:B:240:LYS:HA	2:B:243:THR:OG1	0.50	2.07	1	1
2:B:268:LYS:O	2:B:272:GLU:HB2	0.49	2.06	6	1
2:B:205:LEU:O	2:B:209:VAL:HG23	0.49	2.07	6	5
1:A:72:SER:HB2	1:A:182:ARG:HB2	0.49	1.85	9	2
1:A:29:LEU:O	1:A:33:ILE:HG12	0.48	2.09	2	4
2:B:166:ARG:HD3	2:B:208:GLN:OE1	0.47	2.10	8	1
1:A:84:LYS:HD2	1:A:223:GLU:OE2	0.47	2.09	5	1
1:A:200:GLU:O	1:A:204:LYS:HG3	0.47	2.10	8	1
2:B:162:PHE:O	2:B:165:LEU:HD12	0.47	2.09	8	1
2:B:258:ALA:O	2:B:262:ARG:HG3	0.46	2.10	8	2
1:A:100:LYS:HB2	1:A:160:TYR:CD2	0.46	2.46	4	1
1:A:43:ILE:O	1:A:47:GLU:HG2	0.45	2.10	6	3
2:B:165:LEU:HB2	2:B:212:GLN:NE2	0.45	2.26	1	2
2:B:229:ASP:OD2	2:B:231:ARG:HD2	0.45	2.12	5	1
1:A:63:THR:HG22	1:A:69:LYS:HB2	0.44	1.89	3	1
2:B:166:ARG:HD2	2:B:208:GLN:OE1	0.44	2.12	4	1
1:A:194:GLN:HE21	1:A:194:GLN:HA	0.44	1.73	9	1
2:B:188:ASN:O	2:B:192:ILE:HG12	0.44	2.13	5	2
2:B:260:LYS:O	2:B:264:ARG:HD2	0.44	2.13	1	1
1:A:146:GLU:HG3	1:A:189:HIS:O	0.44	2.12	1	1
1:A:82:PRO:HD2	1:A:221:PHE:O	0.44	2.12	4	1
2:B:215:VAL:HG13	2:B:237:PHE:CE2	0.44	2.47	10	1
1:A:24:ALA:O	1:A:28:GLN:HG3	0.43	2.13	3	1
2:B:233:CYS:HA	2:B:236:GLN:OE1	0.43	2.13	10	1
1:A:34:ILE:HD11	1:A:142:TYR:CD1	0.43	2.48	4	1
1:A:130:MET:HA	1:A:133:GLN:CD	0.43	2.34	7	1
2:B:262:ARG:O	2:B:266:ARG:HG2	0.43	2.13	2	1
1:A:96:ILE:O	1:A:154:HIS:HA	0.42	2.14	7	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:84:LYS:HA	1:A:198:LEU:HD13	0.42	1.91	9	1
1:A:203:ILE:O	1:A:207:VAL:HG23	0.42	2.14	10	2
1:A:154:HIS:HD2	1:A:156:ASP:OD1	0.42	1.98	9	1
1:A:133:GLN:OE1	2:B:165:LEU:HB3	0.42	2.15	10	1
1:A:79:ASN:ND2	1:A:219:THR:HB	0.42	2.30	2	1
1:A:199:GLU:O	1:A:203:ILE:HG13	0.42	2.15	4	1
1:A:150:VAL:HG22	1:A:186:VAL:HG13	0.42	1.92	9	1
1:A:40:ASN:O	1:A:128:ILE:HD13	0.42	2.15	9	1
1:A:78:ILE:O	1:A:218:ILE:HA	0.42	2.15	5	2
1:A:84:LYS:HE3	1:A:199:GLU:OE2	0.41	2.15	7	1
1:A:200:GLU:H	1:A:200:GLU:CD	0.41	2.19	8	1
1:A:100:LYS:HE3	1:A:160:TYR:CE1	0.41	2.51	10	1
1:A:98:MET:O	1:A:154:HIS:HB2	0.41	2.15	7	1
1:A:133:GLN:HA	2:B:167:ARG:CG	0.41	2.46	10	1
1:A:44:PHE:CZ	1:A:45:LEU:HG	0.41	2.50	4	1
2:B:260:LYS:O	2:B:264:ARG:HG3	0.41	2.16	8	1
1:A:90:THR:HA	1:A:186:VAL:O	0.41	2.16	10	1
1:A:22:PHE:CD1	1:A:108:GLY:HA2	0.40	2.51	4	2
1:A:100:LYS:HB2	1:A:160:TYR:CE2	0.40	2.51	7	1
1:A:120:GLU:HA	1:A:120:GLU:OE1	0.40	2.16	2	1
2:B:163:GLY:O	2:B:212:GLN:HG3	0.40	2.17	4	1
1:A:133:GLN:NE2	2:B:165:LEU:HD22	0.40	2.31	1	1
2:B:259:PHE:O	2:B:263:VAL:HG23	0.40	2.16	10	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	208/210 (99%)	180±3 (87±1%)	25±3 (12±1%)	3±1 (2±0%)	13	57
2	B	118/129 (91%)	109±1 (92±1%)	9±2 (7±1%)	1±0 (1±0%)	29	74
All	All	3260/3390 (96%)	2887 (89%)	333 (10%)	40 (1%)	17	64

All 9 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	177	GLY	10
1	A	136	VAL	8
2	B	166	ARG	7
1	A	168	GLY	4
1	A	209	LYS	4
1	A	108	GLY	2
1	A	125	GLY	2
1	A	196	GLU	2
1	A	146	GLU	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	181/181 (100%)	159±3 (88±2%)	22±3 (12±2%)	8 50
2	B	107/117 (91%)	98±2 (91±2%)	9±2 (9±2%)	13 60
All	All	2880/2980 (97%)	2566 (89%)	314 (11%)	10 54

All 107 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	23	GLN	10
1	A	184	THR	10
1	A	201	ARG	10
1	A	93	ASP	9
1	A	140	SER	9
2	B	186	THR	9
1	A	53	SER	8
1	A	113	SER	8
1	A	136	VAL	8
1	A	176	THR	8
1	A	165	SER	7
2	B	236	GLN	7
1	A	156	ASP	7
2	B	218	PHE	7
1	A	65	THR	6
1	A	72	SER	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	128	ILE	6
1	A	130	MET	6
1	A	174	THR	6
2	B	203	CYS	6
1	A	195	THR	6
1	A	178	GLU	5
1	A	57	ASP	4
1	A	164	SER	4
1	A	211	SER	4
1	A	223	GLU	4
2	B	196	ASP	4
1	A	39	SER	4
2	B	243	THR	4
2	B	167	ARG	4
1	A	88	THR	3
1	A	194	GLN	3
2	B	170	ASP	3
2	B	276	LYS	3
1	A	47	GLU	3
1	A	71	ASP	3
2	B	169	ASP	3
2	B	172	GLN	3
1	A	193	ASP	3
1	A	94	THR	2
1	A	146	GLU	2
2	B	177	ASP	2
2	B	240	LYS	2
2	B	246	ARG	2
1	A	63	THR	2
2	B	188	ASN	2
2	B	228	VAL	2
2	B	245	ASP	2
2	B	255	GLU	2
1	A	32	LEU	2
2	B	200	GLU	2
1	A	19	THR	2
1	A	61	TYR	2
1	A	143	LEU	2
2	B	184	GLU	2
1	A	220	LEU	2
1	A	29	LEU	2
1	A	138	PHE	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	192	GLU	2
2	B	256	LEU	2
2	B	171	SER	2
2	B	272	GLU	2
1	A	25	GLU	1
1	A	219	THR	1
2	B	250	GLU	1
1	A	38	TYR	1
1	A	54	ASP	1
1	A	75	GLU	1
1	A	151	ILE	1
2	B	168	TRP	1
1	A	18	GLU	1
1	A	171	THR	1
2	B	161	HIS	1
2	B	178	ASN	1
2	B	229	ASP	1
1	A	42	GLU	1
1	A	62	GLU	1
1	A	107	LEU	1
1	A	196	GLU	1
1	A	222	VAL	1
2	B	165	LEU	1
2	B	179	VAL	1
2	B	239	THR	1
2	B	213	THR	1
2	B	225	SER	1
2	B	264	ARG	1
1	A	15	GLU	1
1	A	133	GLN	1
2	B	185	GLU	1
2	B	231	ARG	1
1	A	90	THR	1
1	A	120	GLU	1
1	A	169	SER	1
1	A	187	ILE	1
1	A	209	LYS	1
2	B	269	LEU	1
1	A	69	LYS	1
1	A	154	HIS	1
2	B	183	CYS	1
2	B	193	TRP	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	B	227	LYS	1
1	A	14	GLU	1
1	A	28	GLN	1
1	A	58	LYS	1
1	A	127	ASP	1
1	A	163	GLU	1
1	A	173	ARG	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 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