



Full wwPDB NMR Structure Validation Report ⓘ

May 28, 2020 – 10:37 pm BST

PDB ID : 2KMS
Title : Combined high- and low-resolution techniques reveal compact structure in central portion of factor H despite long inter-modular linkers
Authors : Schmidt, C.Q.; Herbert, A.P.; Guariento, M.; Mertens, H.D.T.; Soares, D.C.; Uhrin, D.; Rowe, A.J.; Svergun, D.I.; Barlow, P.N.
Deposited on : 2009-08-04

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 ⓘ symbol.

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

Cyrange : Kirchner and Güntert (2011)
NmrClust : Kelley et al. (1996)
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 : 2.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

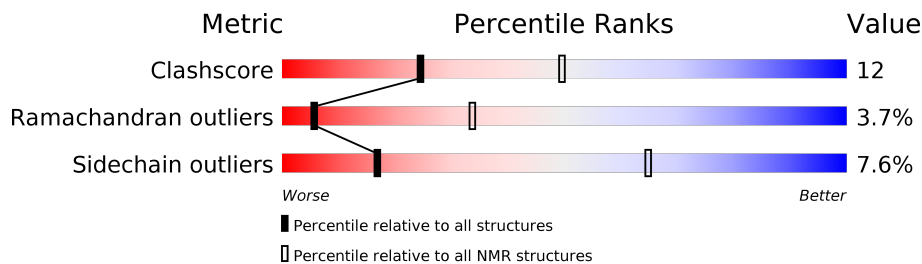
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 86%.

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	115	

2 Ensemble composition and analysis

This entry contains 20 models. Model 7 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:690-A:750 (61)	0.26	7
2	A:751-A:801 (51)	0.28	8

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 5 clusters. No single-model clusters were found.

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1815 atoms, of which 887 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Complement factor H.

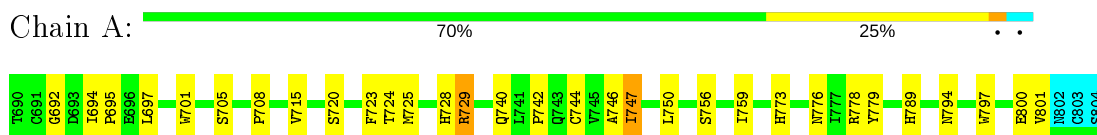
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	115	1815	582	887	164	173	9	0

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: Complement factor H

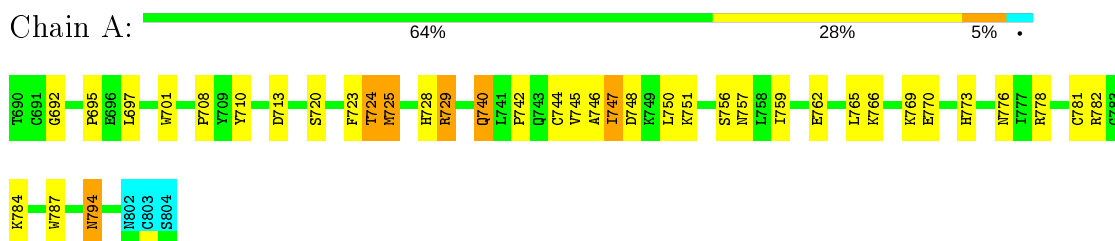


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

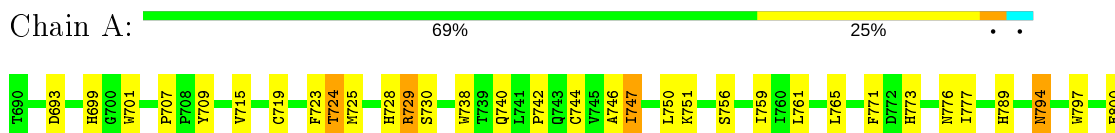
4.2.1 Score per residue for model 1

- Molecule 1: Complement factor H



4.2.2 Score per residue for model 2

- Molecule 1: Complement factor H



V801
M802
C803
S804

4.2.3 Score per residue for model 3

- Molecule 1: Complement factor H

Chain A:  62% 33%

T690 D693 D694 I694 L697 W701 S706 P707 Y708 Y709 Y710 V715 E716 F717 F723 T724 M725 H728 V737 M738 I739 Q740 I743 F742 A746 I747 L750 K753 K752 S756 I759 H764 E770 M776 I777 R778 Y779 C781 C781 R782 G783 K784 E785 G786 W787

I788 H789 T790 W797 E800 V801 M802 C803 S804

4.2.4 Score per residue for model 4

- Molecule 1: Complement factor H

Chain A:  63% 30%

T690 C691 D692 D693 D694 P695 E696 L697 W701 P708 Y709 Y710 S714 V715 E716 F717 F718 C719 F723 T724 I731 G736 V737 W738 T739 Q740 L741 A746 I747 L750 S756 I759 H773 M776 Y779 I788 H789 Y791 W794 W797 D798 F799 E800

V801
M802
C803
S804

4.2.5 Score per residue for model 5

- Molecule 1: Complement factor H

Chain A:  63% 28% 6%

T690 C691 D692 D693 I694 P695 E696 L697 W701 S706 Y709 D713 S714 V715 E716 F717 S720 F723 T724 M725 H728 R729 Q740 L741 P742 C744 V745 I746 I747 D748 K749 L750 K751 L758 I759 H773 R778 Y779 H780 C781 K784 E785 I788 H789 T790

W794 W797 V801 M802 C803 S804

4.2.6 Score per residue for model 6

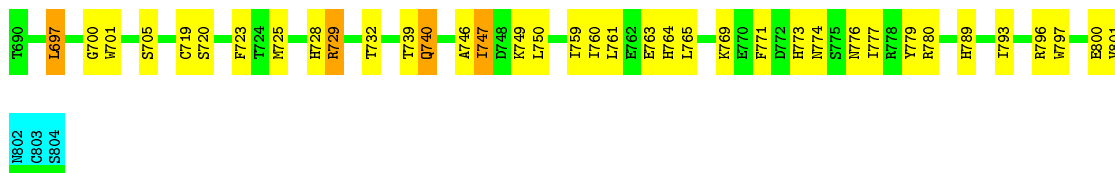
- Molecule 1: Complement factor H

Chain A:  74% 23%



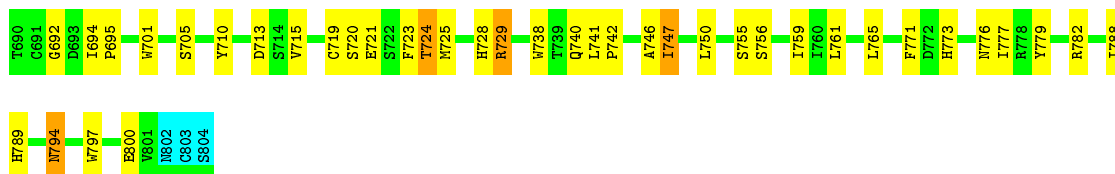
4.2.7 Score per residue for model 7 (medoid)

- Molecule 1: Complement factor H



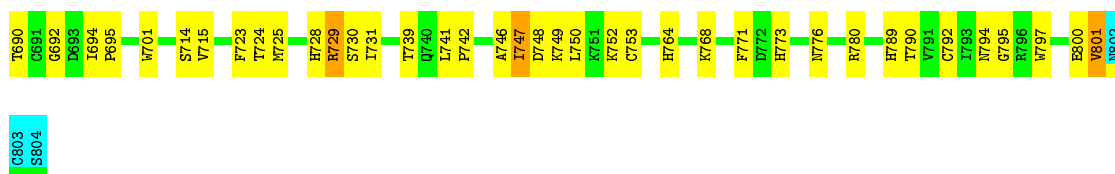
4.2.8 Score per residue for model 8

- Molecule 1: Complement factor H



4.2.9 Score per residue for model 9

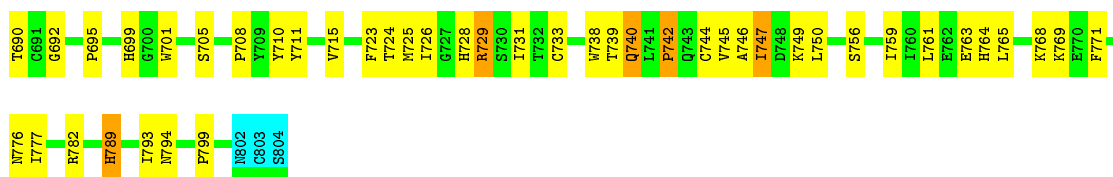
- Molecule 1: Complement factor H



4.2.10 Score per residue for model 10

- Molecule 1: Complement factor H





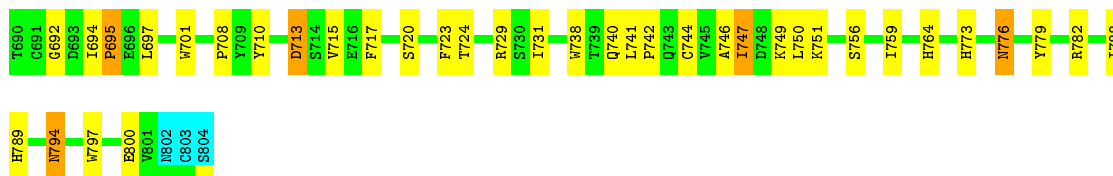
4.2.11 Score per residue for model 11

- Molecule 1: Complement factor H



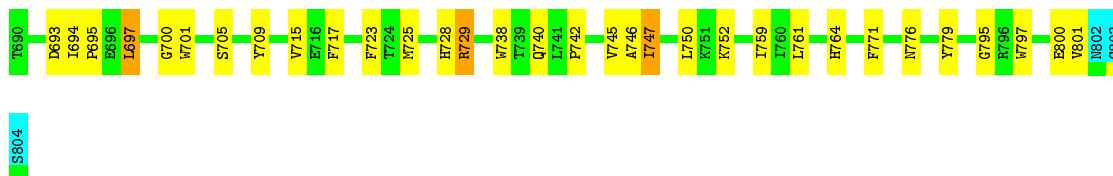
4.2.12 Score per residue for model 12

- Molecule 1: Complement factor H



4.2.13 Score per residue for model 13

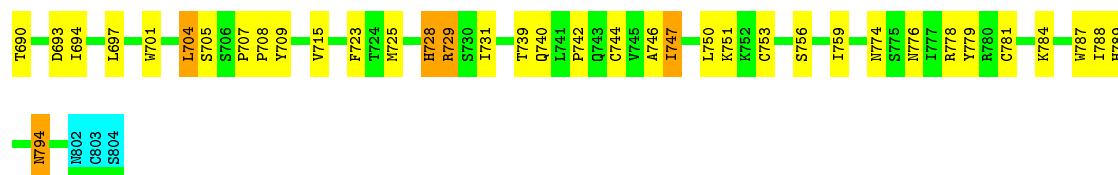
- Molecule 1: Complement factor H



4.2.14 Score per residue for model 14

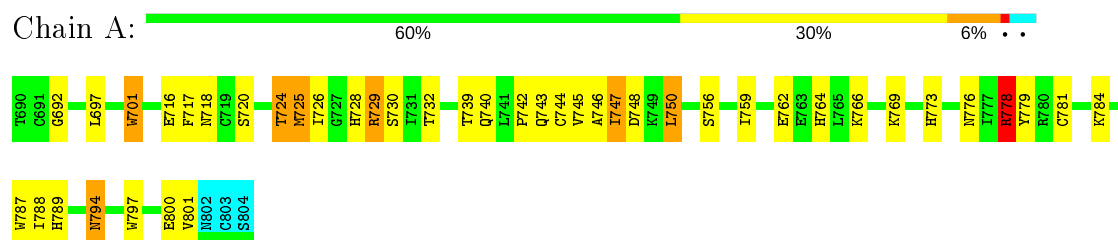
- Molecule 1: Complement factor H





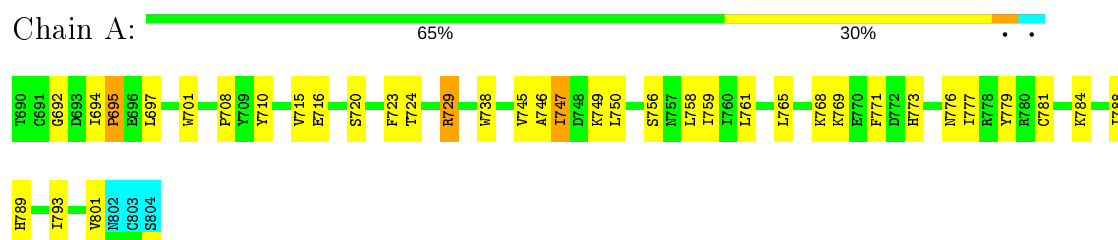
4.2.15 Score per residue for model 15

- Molecule 1: Complement factor H



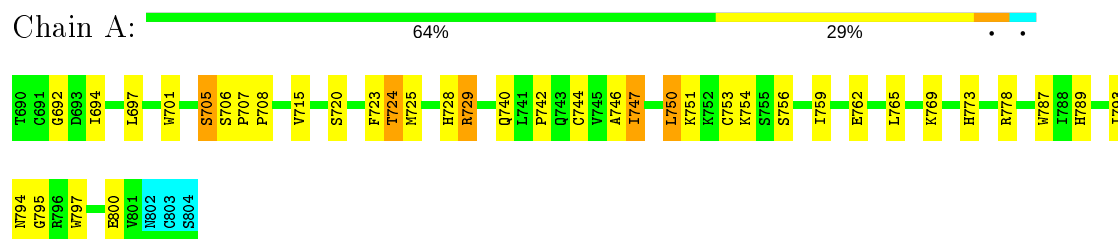
4.2.16 Score per residue for model 16

- Molecule 1: Complement factor H



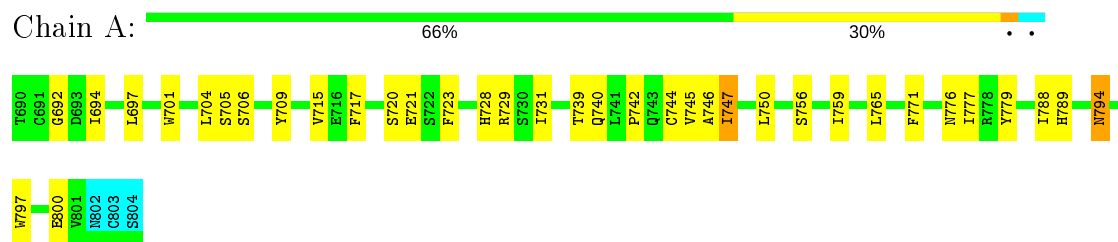
4.2.17 Score per residue for model 17

- Molecule 1: Complement factor H



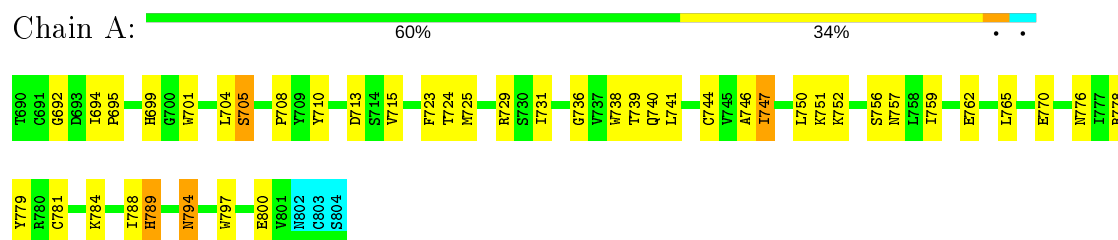
4.2.18 Score per residue for model 18

- Molecule 1: Complement factor H



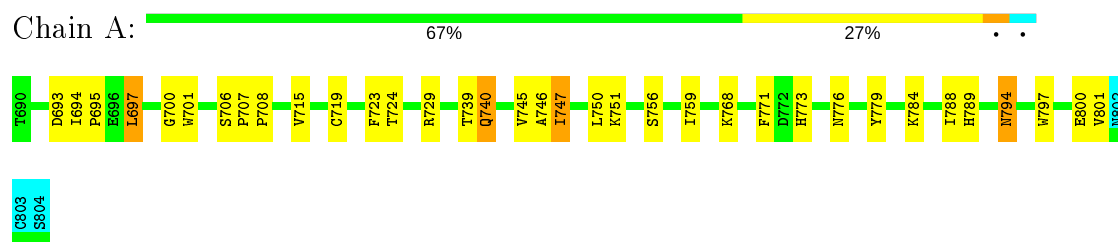
4.2.19 Score per residue for model 19

- Molecule 1: Complement factor H



4.2.20 Score per residue for model 20

- Molecule 1: Complement factor H



5 Refinement protocol and experimental data overview i

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

Of the 100 calculated structures, 20 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	refinement	1.2
CYANA	refinement	2.1
CYANA	structure solution	2.1

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	1381
Number of shifts mapped to atoms	1348
Number of unparsed shifts	0
Number of shifts with mapping errors	33
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	86%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	907	872	865	21±3
All	All	18140	17440	17300	418

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 12.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:745:VAL:HG11	1:A:750:LEU:HD21	0.93	1.37	18	1
1:A:750:LEU:HG	1:A:794:ASN:HD22	0.88	1.29	5	3
1:A:747:ILE:HA	1:A:750:LEU:HD22	0.81	1.52	17	6
1:A:750:LEU:HG	1:A:794:ASN:ND2	0.73	1.98	8	5
1:A:747:ILE:HA	1:A:750:LEU:HG	0.72	1.61	19	12
1:A:693:ASP:HA	1:A:709:TYR:CE2	0.71	2.21	5	3
1:A:750:LEU:HD12	1:A:773:HIS:CD2	0.68	2.24	20	2
1:A:750:LEU:HD23	1:A:773:HIS:HB3	0.67	1.66	17	4
1:A:725:MET:SD	1:A:729:ARG:HB3	0.67	2.30	3	14
1:A:694:ILE:HG21	1:A:715:VAL:HG21	0.65	1.68	14	15
1:A:756:SER:HB3	1:A:759:ILE:HG13	0.64	1.67	19	4
1:A:697:LEU:HD13	1:A:744:CYS:SG	0.64	2.33	1	2
1:A:762:GLU:HB2	1:A:765:LEU:HD23	0.64	1.67	19	3
1:A:750:LEU:HG	1:A:794:ASN:HB2	0.64	1.70	4	1
1:A:769:LYS:HG3	1:A:770:GLU:HG2	0.63	1.70	1	1
1:A:750:LEU:HB2	1:A:773:HIS:CD2	0.62	2.30	9	1
1:A:723:PHE:CD1	1:A:746:ALA:HA	0.61	2.30	8	19
1:A:765:LEU:CD1	1:A:777:ILE:HB	0.61	2.25	16	6
1:A:699:HIS:HB2	1:A:744:CYS:HB2	0.61	1.71	19	3
1:A:791:VAL:HG22	1:A:798:ASP:HB3	0.60	1.73	4	1
1:A:771:PHE:CE2	1:A:777:ILE:HG21	0.59	2.32	2	3
1:A:747:ILE:HA	1:A:750:LEU:CG	0.59	2.27	7	11
1:A:768:LYS:HD2	1:A:771:PHE:CD2	0.59	2.33	9	3
1:A:761:LEU:HD13	1:A:771:PHE:HZ	0.59	1.58	7	4
1:A:745:VAL:HG11	1:A:750:LEU:CD1	0.59	2.26	13	1
1:A:779:TYR:CD2	1:A:801:VAL:HG13	0.58	2.32	20	6
1:A:705:SER:HB2	1:A:715:VAL:HG22	0.58	1.76	10	2
1:A:750:LEU:HD22	1:A:794:ASN:ND2	0.58	2.13	10	1
1:A:765:LEU:HD22	1:A:771:PHE:CE2	0.57	2.34	6	7
1:A:750:LEU:HD23	1:A:773:HIS:CG	0.57	2.34	8	2
1:A:781:CYS:HB3	1:A:784:LYS:HD3	0.57	1.76	15	3
1:A:745:VAL:HG11	1:A:750:LEU:HD13	0.57	1.76	13	1
1:A:765:LEU:HD11	1:A:777:ILE:HB	0.56	1.77	16	1
1:A:756:SER:HB2	1:A:759:ILE:CG1	0.56	2.31	18	4
1:A:756:SER:HB2	1:A:759:ILE:HG13	0.55	1.76	18	4
1:A:765:LEU:HD22	1:A:771:PHE:HE2	0.55	1.62	6	4
1:A:728:HIS:O	1:A:742:PRO:HB3	0.55	2.02	3	12
1:A:704:LEU:HD13	1:A:705:SER:N	0.54	2.17	19	2
1:A:708:PRO:HG2	1:A:710:TYR:CE2	0.54	2.37	1	5
1:A:723:PHE:HA	1:A:745:VAL:O	0.54	2.03	13	1
1:A:747:ILE:HA	1:A:750:LEU:CD1	0.53	2.33	3	8
1:A:756:SER:HB3	1:A:759:ILE:CG1	0.53	2.33	19	7

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:750:LEU:HD22	1:A:794:ASN:OD1	0.53	2.04	20	6
1:A:705:SER:CB	1:A:715:VAL:HG22	0.53	2.33	17	5
1:A:724:THR:O	1:A:744:CYS:HA	0.53	2.04	19	3
1:A:694:ILE:CG2	1:A:715:VAL:HG21	0.53	2.33	14	5
1:A:761:LEU:HD13	1:A:771:PHE:CZ	0.53	2.39	7	3
1:A:779:TYR:CE1	1:A:788:ILE:HB	0.52	2.39	5	10
1:A:781:CYS:HB3	1:A:784:LYS:HB2	0.52	1.80	14	2
1:A:745:VAL:CG1	1:A:749:LYS:HG3	0.52	2.34	16	1
1:A:697:LEU:HD21	1:A:744:CYS:SG	0.52	2.45	14	5
1:A:797:TRP:HB2	1:A:800:GLU:HA	0.52	1.82	4	13
1:A:751:LYS:HB2	1:A:794:ASN:C	0.51	2.26	2	8
1:A:745:VAL:HB	1:A:750:LEU:HD21	0.51	1.82	20	2
1:A:707:PRO:HB3	1:A:709:TYR:CE2	0.51	2.40	11	1
1:A:752:LYS:HE2	1:A:770:GLU:HB3	0.51	1.82	19	2
1:A:750:LEU:HD23	1:A:773:HIS:HB2	0.51	1.82	4	1
1:A:778:ARG:HB2	1:A:787:TRP:HB3	0.50	1.83	15	4
1:A:773:HIS:HD1	1:A:793:ILE:HA	0.50	1.66	7	3
1:A:720:SER:O	1:A:721:GLU:HG3	0.50	2.05	8	1
1:A:747:ILE:CA	1:A:750:LEU:HG	0.49	2.36	19	7
1:A:751:LYS:HB3	1:A:795:GLY:N	0.49	2.21	17	1
1:A:695:PRO:HD3	1:A:738:TRP:CD1	0.49	2.42	16	6
1:A:740:GLN:NE2	1:A:740:GLN:H	0.49	2.06	20	2
1:A:708:PRO:HG2	1:A:710:TYR:CE1	0.49	2.43	10	1
1:A:746:ALA:O	1:A:747:ILE:HB	0.49	2.07	2	7
1:A:750:LEU:HD23	1:A:773:HIS:CD2	0.48	2.43	8	2
1:A:714:SER:HA	1:A:731:ILE:O	0.48	2.08	9	2
1:A:716:GLU:HG3	1:A:729:ARG:HG2	0.48	1.85	16	1
1:A:697:LEU:HD13	1:A:700:GLY:HA3	0.48	1.85	13	3
1:A:690:THR:HG21	1:A:708:PRO:HB2	0.47	1.85	10	1
1:A:726:ILE:HB	1:A:743:GLN:HB3	0.47	1.86	15	1
1:A:724:THR:HG21	1:A:750:LEU:HD21	0.47	1.86	2	5
1:A:740:GLN:O	1:A:741:LEU:HB2	0.47	2.09	4	1
1:A:778:ARG:HA	1:A:789:HIS:HA	0.47	1.86	5	2
1:A:779:TYR:CD1	1:A:801:VAL:HG22	0.47	2.45	7	2
1:A:750:LEU:CG	1:A:794:ASN:HD22	0.46	2.15	15	2
1:A:741:LEU:HB3	1:A:742:PRO:HD2	0.46	1.86	12	1
1:A:751:LYS:HB3	1:A:795:GLY:CA	0.46	2.40	17	1
1:A:750:LEU:HA	1:A:794:ASN:ND2	0.46	2.25	9	1
1:A:750:LEU:HD12	1:A:773:HIS:HD2	0.46	1.70	9	1
1:A:697:LEU:HD22	1:A:717:PHE:HB3	0.46	1.88	12	7
1:A:790:THR:HG21	1:A:801:VAL:CG2	0.46	2.41	4	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:761:LEU:HB3	1:A:765:LEU:HB2	0.46	1.85	8	2
1:A:752:LYS:O	1:A:795:GLY:HA2	0.46	2.10	13	2
1:A:747:ILE:HG23	1:A:773:HIS:NE2	0.46	2.26	20	1
1:A:750:LEU:HD23	1:A:773:HIS:CB	0.46	2.41	5	4
1:A:697:LEU:HD12	1:A:717:PHE:HB3	0.46	1.87	13	1
1:A:756:SER:HB2	1:A:759:ILE:HG12	0.45	1.86	17	2
1:A:750:LEU:HA	1:A:794:ASN:HD22	0.45	1.71	6	2
1:A:707:PRO:HA	1:A:709:TYR:N	0.45	2.27	14	2
1:A:726:ILE:HD12	1:A:793:ILE:HG23	0.45	1.86	10	1
1:A:764:HIS:CD2	1:A:765:LEU:HG	0.45	2.46	7	1
1:A:731:ILE:HG22	1:A:741:LEU:HA	0.45	1.87	12	1
1:A:745:VAL:CG1	1:A:750:LEU:HD21	0.45	2.28	18	1
1:A:710:TYR:O	1:A:713:ASP:HB2	0.45	2.12	12	2
1:A:715:VAL:O	1:A:730:SER:HA	0.45	2.12	9	2
1:A:759:ILE:HG12	1:A:801:VAL:HG11	0.45	1.89	4	3
1:A:724:THR:HG22	1:A:745:VAL:O	0.44	2.12	15	2
1:A:701:TRP:HE1	1:A:718:ASN:CG	0.44	2.15	15	1
1:A:747:ILE:O	1:A:750:LEU:HB2	0.44	2.12	13	3
1:A:773:HIS:HA	1:A:792:CYS:SG	0.44	2.53	9	1
1:A:762:GLU:O	1:A:766:LYS:HG3	0.44	2.12	1	2
1:A:731:ILE:HA	1:A:739:THR:OG1	0.44	2.13	10	5
1:A:769:LYS:HG3	1:A:770:GLU:HG3	0.44	1.90	11	1
1:A:781:CYS:HB2	1:A:784:LYS:HB2	0.44	1.90	5	2
1:A:711:TYR:HA	1:A:733:CYS:SG	0.43	2.52	10	1
1:A:759:ILE:HD12	1:A:761:LEU:HD11	0.43	1.90	10	1
1:A:690:THR:CG2	1:A:708:PRO:HB2	0.43	2.43	14	1
1:A:709:TYR:HB3	1:A:738:TRP:CZ2	0.43	2.49	2	3
1:A:789:HIS:O	1:A:799:PRO:HD2	0.43	2.14	10	1
1:A:773:HIS:ND1	1:A:793:ILE:HA	0.43	2.28	17	1
1:A:719:CYS:HB3	1:A:723:PHE:HB2	0.43	1.89	8	4
1:A:707:PRO:HA	1:A:708:PRO:C	0.43	2.35	11	1
1:A:790:THR:HG21	1:A:801:VAL:CG1	0.42	2.44	9	1
1:A:778:ARG:HB3	1:A:787:TRP:HB3	0.42	1.91	14	1
1:A:707:PRO:HB2	1:A:708:PRO:HA	0.41	1.92	17	2
1:A:771:PHE:CD2	1:A:777:ILE:HG21	0.41	2.50	8	1
1:A:779:TYR:CE1	1:A:801:VAL:HG12	0.41	2.50	13	1
1:A:768:LYS:HG2	1:A:771:PHE:CD2	0.41	2.51	20	1
1:A:762:GLU:CD	1:A:778:ARG:HE	0.41	2.19	1	1
1:A:746:ALA:C	1:A:748:ASP:H	0.41	2.17	9	2
1:A:705:SER:HB2	1:A:709:TYR:CE1	0.41	2.51	13	1
1:A:708:PRO:HG2	1:A:710:TYR:CZ	0.41	2.51	16	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:790:THR:OG1	1:A:797:TRP:HB3	0.41	2.15	5	1
1:A:760:ILE:HB	1:A:780:ARG:HG2	0.41	1.92	7	1
1:A:794:ASN:HD22	1:A:794:ASN:N	0.41	2.14	1	1
1:A:694:ILE:HG12	1:A:705:SER:CB	0.41	2.46	8	1
1:A:716:GLU:HG2	1:A:730:SER:HB3	0.41	1.93	15	1
1:A:697:LEU:HG	1:A:717:PHE:HD2	0.41	1.76	18	1
1:A:750:LEU:HD22	1:A:794:ASN:CG	0.41	2.36	18	1
1:A:776:ASN:HA	1:A:791:VAL:HG12	0.41	1.92	4	1
1:A:776:ASN:N	1:A:776:ASN:HD22	0.40	2.14	12	1
1:A:739:THR:HB	1:A:740:GLN:NE2	0.40	2.31	20	1
1:A:740:GLN:H	1:A:740:GLN:NE2	0.40	2.15	1	2
1:A:732:THR:O	1:A:739:THR:HG23	0.40	2.16	15	2
1:A:719:CYS:SG	1:A:725:MET:HA	0.40	2.56	2	1
1:A:759:ILE:HG12	1:A:801:VAL:CG1	0.40	2.46	7	1
1:A:754:LYS:HA	1:A:769:LYS:O	0.40	2.17	17	1

5.2 Torsion angles [i](#)

5.2.1 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	111/115 (97%)	91±2 (82±2%)	16±2 (14±2%)	4±1 (4±1%)	6	34
All	All	2220/2300 (97%)	1819 (82%)	318 (14%)	83 (4%)	6	34

All 17 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	747	ILE	20
1	A	789	HIS	16
1	A	720	SER	10
1	A	706	SER	7
1	A	695	PRO	7
1	A	693	ASP	4
1	A	741	LEU	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	736	GLY	3
1	A	725	MET	2
1	A	705	SER	2
1	A	724	THR	2
1	A	709	TYR	2
1	A	704	LEU	1
1	A	691	CYS	1
1	A	783	GLY	1
1	A	742	PRO	1
1	A	785	GLU	1

5.2.2 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	102/105 (97%)	94±2 (92±2%)	8±2 (8±2%)	17	65
All	All	2040/2100 (97%)	1885 (92%)	155 (8%)	17	65

All 31 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	701	TRP	20
1	A	776	ASN	18
1	A	729	ARG	17
1	A	740	GLN	17
1	A	794	ASN	11
1	A	724	THR	11
1	A	764	HIS	7
1	A	713	ASP	5
1	A	749	LYS	5
1	A	750	LEU	4
1	A	753	CYS	4
1	A	697	LEU	4
1	A	769	LYS	4
1	A	784	LYS	3
1	A	690	THR	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	728	HIS	3
1	A	763	GLU	2
1	A	757	ASN	2
1	A	782	ARG	2
1	A	774	ASN	2
1	A	748	ASP	1
1	A	741	LEU	1
1	A	785	GLU	1
1	A	704	LEU	1
1	A	778	ARG	1
1	A	737	VAL	1
1	A	781	CYS	1
1	A	716	GLU	1
1	A	721	GLU	1
1	A	705	SER	1
1	A	755	SER	1

5.2.3 RNA [i](#)

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.4 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.5 Ligand geometry [i](#)

There are no ligands in this entry.

5.6 Other polymers [i](#)

There are no such molecules in this entry.

5.7 Polymer linkage issues

There are no chain breaks in this entry.

6 Chemical shift validation i

The completeness of assignment taking into account all chemical shift lists is 86% for the well-defined parts and 86% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

6.1.1 Bookkeeping i

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1381
Number of shifts mapped to atoms	1348
Number of unparsed shifts	0
Number of shifts with mapping errors	33
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. All 33 occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	3	GLU	CB	30.19	0.037	1
A	5	ALA	N	123.847	0.015	1
A	3	GLU	H	8.525	0.003	1
A	6	GLY	C	178.709	0.003	1
A	3	GLU	N	120.983	0.035	1
A	6	GLY	N	107.415	0.013	1
A	6	GLY	HA3	3.608	0.008	2
A	3	GLU	HG2	2.234	0.003	1
A	4	ALA	HB1	1.315	0.003	1
A	5	ALA	CA	52.305	0.044	1
A	5	ALA	HA	4.221	0.007	1
A	6	GLY	H	7.778	0.01	1
A	4	ALA	HA	4.218	0.011	1
A	4	ALA	CB	19.258	0.045	1

Continued on next page...

Continued from previous page...

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	3	GLU	CG	36.24	0.052	1
A	6	GLY	CA	45.233	0.041	1
A	5	ALA	HB1	1.311	0.002	1
A	3	GLU	HB2	1.935	0.011	1
A	4	ALA	N	125.434	0.029	1
A	4	ALA	CA	52.307	0.052	1
A	5	ALA	CB	19.433	0.071	1
A	5	ALA	H	8.234	0.002	1
A	5	ALA	HB2	1.311	0.002	1
A	4	ALA	HB3	1.315	0.003	1
A	3	GLU	CA	56.68	0.04	1
A	6	GLY	HA2	3.309	0.004	2
A	5	ALA	C	174.327	0.002	1
A	3	GLU	HA	4.164	0.008	1
A	5	ALA	HB3	1.311	0.002	1
A	4	ALA	H	8.312	0.005	1
A	4	ALA	HB2	1.315	0.003	1
A	3	GLU	C	175.758	0.001	1
A	4	ALA	C	174.782	0.002	1

6.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	119	0.03 ± 0.23	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	110	0.01 ± 0.21	None needed (< 0.5 ppm)
$^{13}\text{C}'$	116	-1.67 ± 0.30	Should be applied
^{15}N	113	0.09 ± 0.51	None needed (< 0.5 ppm)

6.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 86%, i.e. 1210 atoms were assigned a chemical shift out of a possible 1413. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	545/550 (99%)	218/219 (100%)	221/224 (99%)	106/107 (99%)
Sidechain	557/708 (79%)	336/420 (80%)	212/256 (83%)	9/32 (28%)

Continued on next page...

Continued from previous page...

	Total	¹H	¹³C	¹⁵N
Aromatic	108/155 (70%)	54/79 (68%)	50/60 (83%)	4/16 (25%)
Overall	1210/1413 (86%)	608/718 (85%)	483/540 (89%)	119/155 (77%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 86%, i.e. 1236 atoms were assigned a chemical shift out of a possible 1441. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹H	¹³C	¹⁵N
Backbone	560/565 (99%)	224/225 (100%)	227/230 (99%)	109/110 (99%)
Sidechain	568/721 (79%)	343/428 (80%)	215/260 (83%)	10/33 (30%)
Aromatic	108/155 (70%)	54/79 (68%)	50/60 (83%)	4/16 (25%)
Overall	1236/1441 (86%)	621/732 (85%)	492/550 (89%)	123/159 (77%)

6.1.4 Statistically unusual chemical shifts [i](#)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	794	ASN	H	12.14	11.45 – 5.25	6.1
1	A	780	ARG	HD2	1.74	4.27 – 1.97	-6.0
1	A	718	ASN	HB2	0.99	4.36 – 1.26	-5.9

6.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:

