Full wwPDB NMR Structure Validation Report (i)

Apr 26, 2016 – 04:18 PM BST

PDB ID	:	1PBU
Title	:	Solution structure of the C-terminal domain of the human eEF1Bgamma subunit
Authors	:	Vanwetswinkel, S.; Kriek, J.; Andersen, G.R.; Guntert, P.; Dijk, J.; Canters, G.W.; Siegal, G.
Deposited on	:	2003-05-15

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
http://wwpdb.org/validation/2016/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:

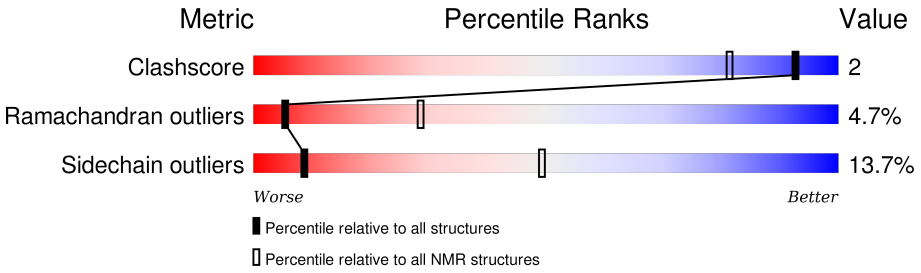
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	162	81% 14%

2 Ensemble composition and analysis

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:279-A:283, A:288-A:437 (155)	0.26	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 5 clusters and 1 single-model cluster was found.

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2628 atoms, of which 1275 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Elongation factor 1-gamma.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	162	2628	882	1275	220	248	3	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	289	ALA	VAL	ENGINEERED	UNP P26641

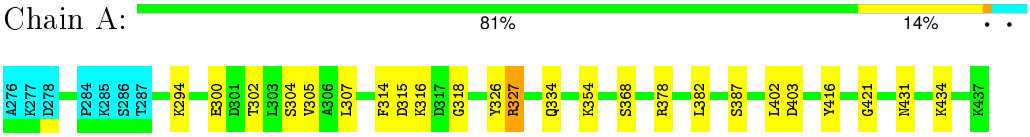
4 Residue-property plots (i)

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: Elongation factor 1-gamma

Chain A:



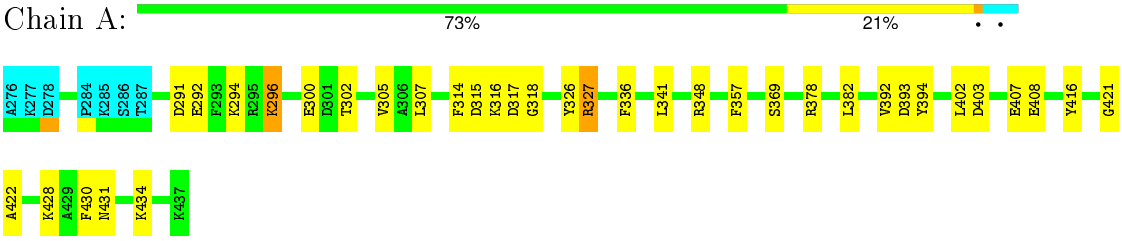
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

• Molecule 1: Elongation factor 1-gamma

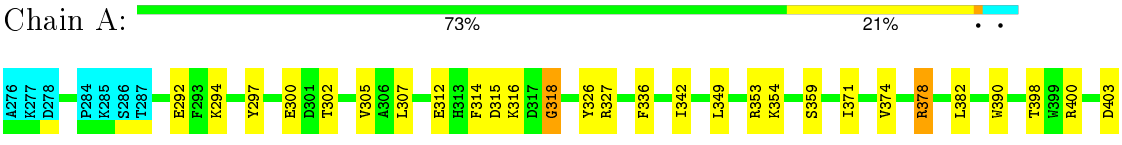
Chain A:



4.2.2 Score per residue for model 2

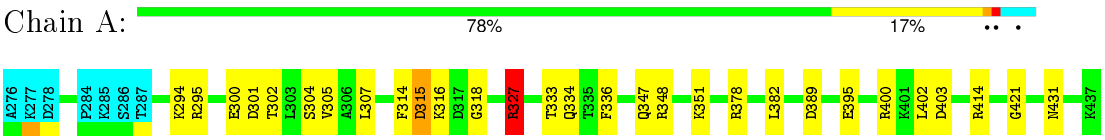
• Molecule 1: Elongation factor 1-gamma

Chain A:



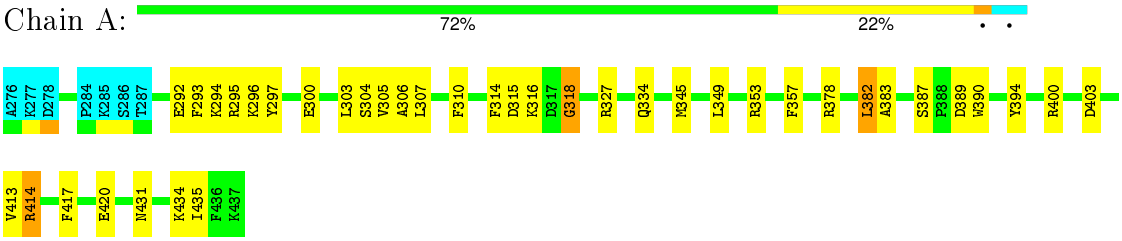
4.2.3 Score per residue for model 3

• Molecule 1: Elongation factor 1-gamma



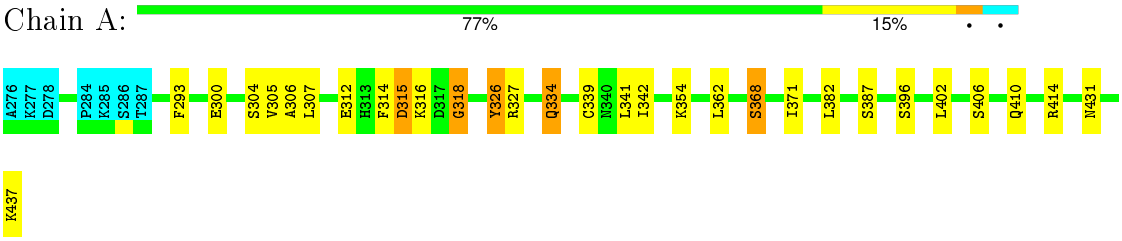
4.2.4 Score per residue for model 4

• Molecule 1: Elongation factor 1-gamma



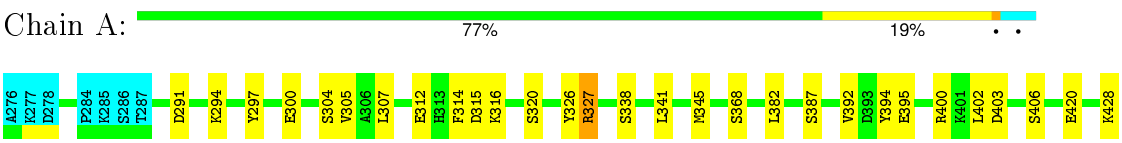
4.2.5 Score per residue for model 5

• Molecule 1: Elongation factor 1-gamma



4.2.6 Score per residue for model 6

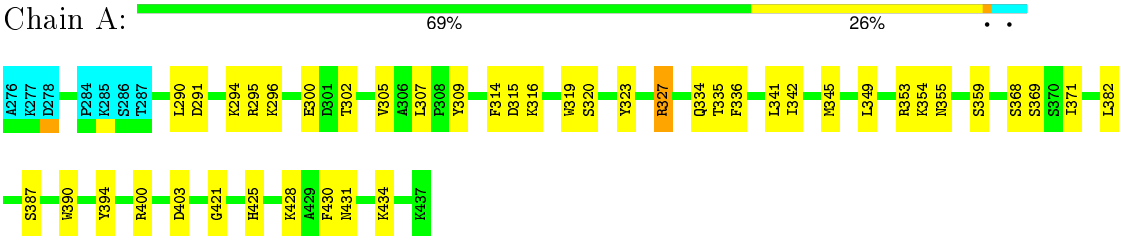
• Molecule 1: Elongation factor 1-gamma



4.2.7 Score per residue for model 7

• Molecule 1: Elongation factor 1-gamma

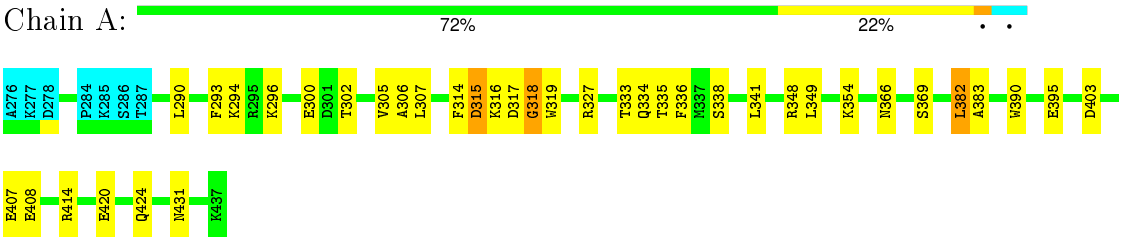
Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: Elongation factor 1-gamma

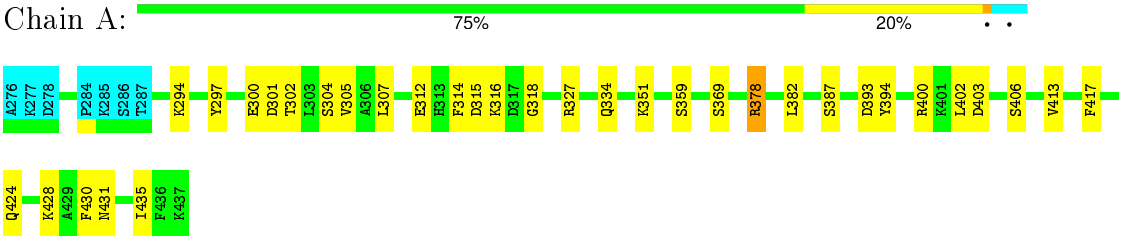
Chain A:



4.2.9 Score per residue for model 9

• Molecule 1: Elongation factor 1-gamma

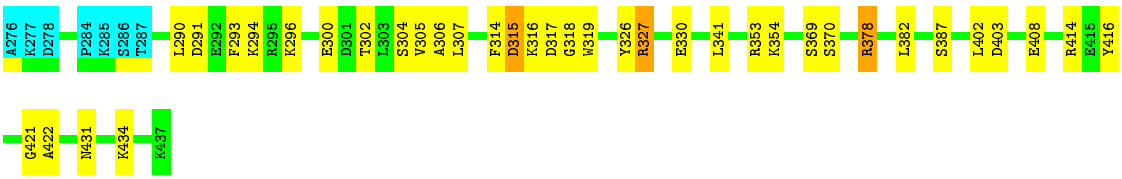
Chain A:



4.2.10 Score per residue for model 10

• Molecule 1: Elongation factor 1-gamma

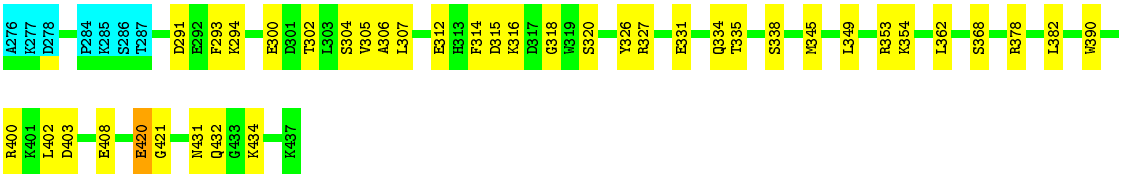
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: Elongation factor 1-gamma

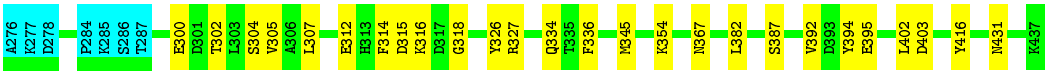
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: Elongation factor 1-gamma

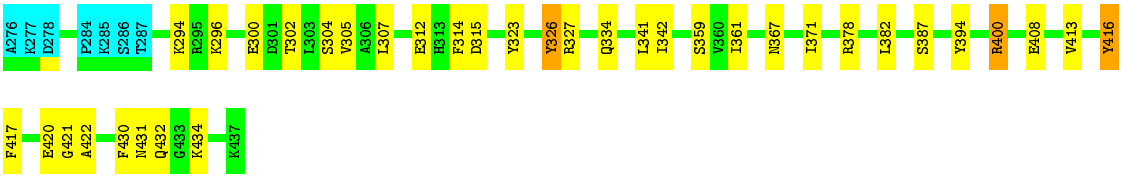
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Elongation factor 1-gamma

Chain A:



4.2.14 Score per residue for model 14

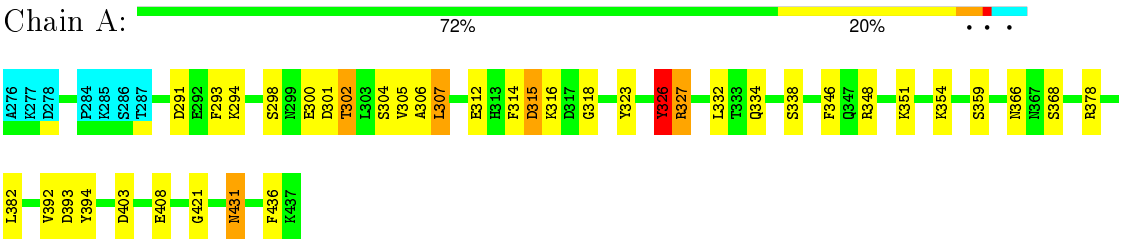
• Molecule 1: Elongation factor 1-gamma

Chain A:

4.2.15 Score per residue for model 15

- Molecule 1: Elongation factor 1-gamma

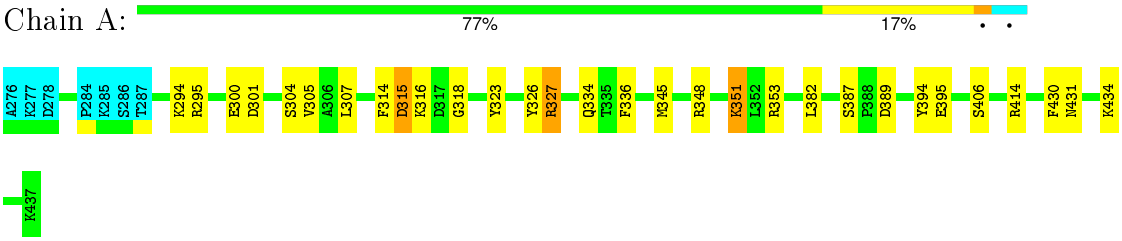
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Elongation factor 1-gamma

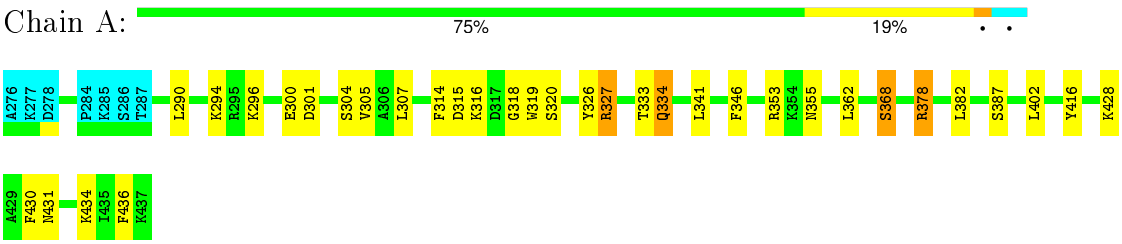
Chain A:



4.2.17 Score per residue for model 17

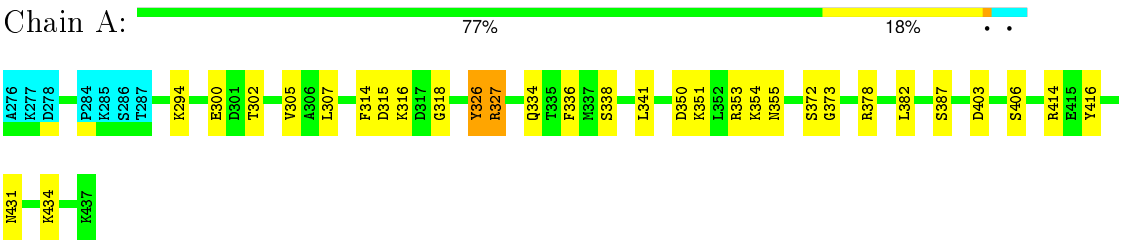
- Molecule 1: Elongation factor 1-gamma

Chain A:



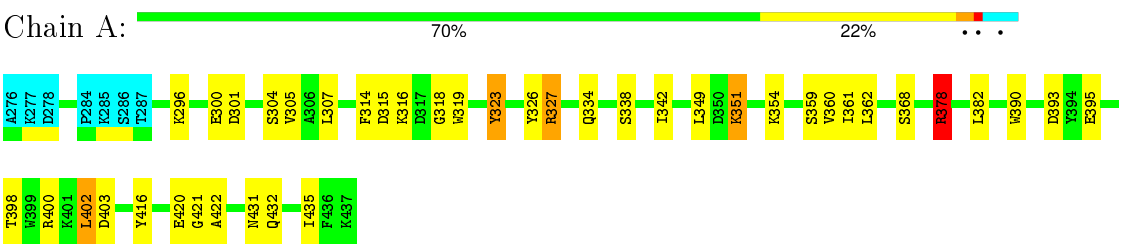
4.2.18 Score per residue for model 18

- Molecule 1: Elongation factor 1-gamma



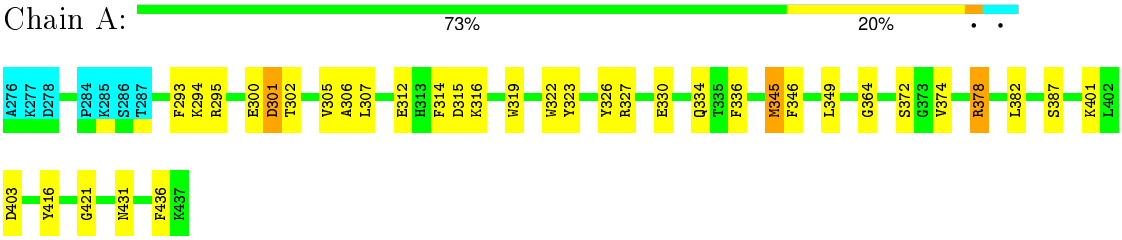
4.2.19 Score per residue for model 19

- Molecule 1: Elongation factor 1-gamma



4.2.20 Score per residue for model 20

- Molecule 1: Elongation factor 1-gamma



5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics, 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
CYANA	structure solution	1.06
OPALp	refinement	

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

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	0.70±0.00	0±0/1347 (0.0±0.0%)	1.10±0.02	2±1/1825 (0.1±0.1%)
All	All	0.70	0/26940 (0.0%)	1.10	41/36500 (0.1%)

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	2.8±1.8
All	All	0	55

There are no bond-length outliers.

All 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	348	ARG	NE-CZ-NH2	-8.28	116.16	120.30	15	3
1	A	323	TYR	CB-CG-CD2	-7.47	116.52	121.00	13	5
1	A	378	ARG	NE-CZ-NH1	6.58	123.59	120.30	15	4
1	A	353	ARG	NE-CZ-NH2	-6.43	117.09	120.30	10	2
1	A	326	TYR	CB-CG-CD2	-6.42	117.15	121.00	10	6
1	A	327	ARG	NE-CZ-NH2	-6.24	117.18	120.30	10	2
1	A	378	ARG	CD-NE-CZ	6.22	132.31	123.60	9	5
1	A	297	TYR	CB-CG-CD1	-6.18	117.29	121.00	4	2
1	A	297	TYR	CB-CG-CD2	-6.00	117.40	121.00	9	1
1	A	414	ARG	NE-CZ-NH2	-5.72	117.44	120.30	3	2
1	A	378	ARG	NE-CZ-NH2	-5.64	117.48	120.30	17	2
1	A	348	ARG	NE-CZ-NH1	5.59	123.10	120.30	15	1
1	A	350	ASP	CB-CG-OD1	5.26	123.03	118.30	18	1
1	A	295	ARG	NE-CZ-NH2	-5.25	117.67	120.30	14	1
1	A	326	TYR	CB-CG-CD1	5.17	124.10	121.00	10	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	295	ARG	NE-CZ-NH1	5.16	122.88	120.30	4	1
1	A	416	TYR	CB-CG-CD1	-5.13	117.92	121.00	12	1
1	A	317	ASP	CB-CG-OD1	5.04	122.83	118.30	14	1

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	326	TYR	Sidechain	8
1	A	430	PHE	Peptide	7
1	A	416	TYR	Sidechain	5
1	A	353	ARG	Sidechain	4
1	A	318	GLY	Peptide	4
1	A	295	ARG	Sidechain	4
1	A	378	ARG	Sidechain	4
1	A	327	ARG	Sidechain	3
1	A	323	TYR	Sidechain	3
1	A	420	GLU	Peptide	2
1	A	400	ARG	Sidechain	2
1	A	348	ARG	Sidechain	2
1	A	313	HIS	Peptide	1
1	A	364	GLY	Peptide	1
1	A	309	TYR	Sidechain	1
1	A	357	PHE	Sidechain	1
1	A	310	PHE	Sidechain	1
1	A	369	SER	Peptide	1
1	A	414	ARG	Sidechain	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	1302	1221	1221	4±2
All	All	26040	24420	24420	90

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:349:LEU:HD23	1:A:390:TRP:CD1	0.66	2.25	7	7
1:A:300:GLU:HB2	1:A:305:VAL:HG11	0.57	1.76	13	12
1:A:300:GLU:CB	1:A:305:VAL:HG11	0.57	2.30	1	19
1:A:346:PHE:CD1	1:A:436:PHE:CD1	0.55	2.94	20	3
1:A:293:PHE:CE1	1:A:306:ALA:HB1	0.54	2.37	8	7
1:A:345:MET:SD	1:A:349:LEU:HD11	0.54	2.42	20	1
1:A:326:TYR:CE2	1:A:332:LEU:HD21	0.53	2.39	15	1
1:A:338:SER:CB	1:A:362:LEU:HD13	0.51	2.36	19	1
1:A:349:LEU:HD23	1:A:390:TRP:CG	0.49	2.43	7	3
1:A:302:THR:HG23	1:A:306:ALA:HB3	0.48	1.85	8	1
1:A:359:SER:HA	1:A:435:ILE:HD13	0.47	1.85	19	2
1:A:361:ILE:HD12	1:A:416:TYR:CD1	0.47	2.45	19	2
1:A:338:SER:HB3	1:A:362:LEU:HD13	0.47	1.85	19	1
1:A:319:TRP:CD2	1:A:378:ARG:HD3	0.46	2.46	20	2
1:A:413:VAL:HG13	1:A:417:PHE:CE2	0.46	2.46	9	4
1:A:342:ILE:HD11	1:A:371:ILE:HG12	0.46	1.88	7	2
1:A:296:LYS:HD3	1:A:305:VAL:HG12	0.45	1.88	1	1
1:A:290:LEU:HD22	1:A:319:TRP:CH2	0.45	2.47	7	4
1:A:322:TRP:CZ2	1:A:401:LYS:HG3	0.44	2.47	20	1
1:A:374:VAL:HG23	1:A:416:TYR:CD2	0.44	2.47	2	2
1:A:362:LEU:HD23	1:A:432:GLN:HB2	0.44	1.88	11	1
1:A:342:ILE:HG21	1:A:360:VAL:HG11	0.44	1.88	19	1
1:A:346:PHE:CE1	1:A:436:PHE:HB2	0.43	2.48	15	2
1:A:342:ILE:CG1	1:A:371:ILE:HD13	0.43	2.44	2	2
1:A:319:TRP:CD2	1:A:378:ARG:HD2	0.42	2.49	19	1
1:A:323:TYR:HB2	1:A:402:LEU:HD11	0.41	1.93	19	1
1:A:425:HIS:H	1:A:425:HIS:CD2	0.41	2.34	7	1
1:A:382:LEU:O	1:A:383:ALA:HB3	0.40	2.15	8	2
1:A:302:THR:HG23	1:A:307:LEU:HD23	0.40	1.93	15	1
1:A:373:GLY:HA2	1:A:416:TYR:CE1	0.40	2.52	18	1
1:A:357:PHE:CZ	1:A:435:ILE:HG23	0.40	2.51	4	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	154/162 (95%)	122±3 (79±2%)	25±3 (16±2%)	7±1 (5±1%)	5 28
All	All	3080/3240 (95%)	2437 (79%)	499 (16%)	144 (5%)	5 28

All 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	431	ASN	20
1	A	327	ARG	20
1	A	315	ASP	20
1	A	314	PHE	20
1	A	318	GLY	16
1	A	421	GLY	11
1	A	428	LYS	6
1	A	422	ALA	5
1	A	334	GLN	4
1	A	368	SER	4
1	A	392	VAL	4
1	A	351	LYS	4
1	A	302	THR	2
1	A	326	TYR	2
1	A	406	SER	2
1	A	301	ASP	1
1	A	295	ARG	1
1	A	297	TYR	1
1	A	303	LEU	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	138/144 (96%)	119±3 (86±2%)	19±3 (14±2%)	9 49
All	All	2760/2880 (96%)	2382 (86%)	378 (14%)	9 49

All 62 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	382	LEU	20
1	A	307	LEU	20
1	A	316	LYS	19
1	A	294	LYS	17
1	A	403	ASP	16
1	A	334	GLN	14
1	A	304	SER	13
1	A	387	SER	12
1	A	354	LYS	11
1	A	434	LYS	11
1	A	302	THR	10
1	A	402	LEU	10
1	A	312	GLU	9
1	A	341	LEU	9
1	A	336	PHE	9
1	A	394	TYR	9
1	A	345	MET	8
1	A	296	LYS	8
1	A	400	ARG	8
1	A	408	GLU	7
1	A	378	ARG	7
1	A	368	SER	7
1	A	301	ASP	7
1	A	291	ASP	6
1	A	315	ASP	6
1	A	395	GLU	6
1	A	327	ARG	6
1	A	420	GLU	6
1	A	359	SER	5
1	A	338	SER	5
1	A	414	ARG	5
1	A	355	ASN	4
1	A	351	LYS	4
1	A	320	SER	4
1	A	393	ASP	4
1	A	369	SER	4
1	A	432	GLN	3
1	A	389	ASP	3
1	A	317	ASP	3
1	A	353	ARG	3
1	A	333	THR	3
1	A	335	THR	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	326	TYR	3
1	A	292	GLU	3
1	A	406	SER	3
1	A	398	THR	2
1	A	362	LEU	2
1	A	407	GLU	2
1	A	366	ASN	2
1	A	330	GLU	2
1	A	424	GLN	2
1	A	372	SER	2
1	A	367	ASN	2
1	A	437	LYS	1
1	A	396	SER	1
1	A	431	ASN	1
1	A	339	CYS	1
1	A	347	GLN	1
1	A	331	GLU	1
1	A	410	GLN	1
1	A	370	SER	1
1	A	298	SER	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 carbohydrates 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