Full wwPDB NMR Structure Validation Report (i)

Apr 27, 2016 – 12:04 AM BST

PDB ID	:	2KX7
Title	:	Solution structure of the E.coli RcsD-ABL domain (residues 688-795)
Authors	:	Rogov, V.V.; Schmoee, K.; Rogova, N.Y.; Loehr, F.; Bernhard, F.; Doetsch, V.
Deposited on	:	2010-04-27

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 (1)) were used in the production of this report:

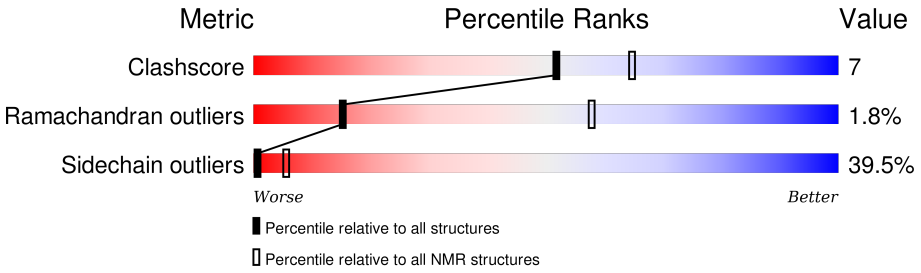
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 (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	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	117	41% 44% • 9% 5%

2 Ensemble composition and analysis

This entry contains 25 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:689-A:788 (100)	0.42	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 4 clusters and 4 single-model clusters were found.

Cluster number	Models
1	1, 4, 5, 6, 8, 11, 12, 15, 16, 17, 20, 21, 22, 24
2	2, 3, 18
3	10, 25
4	13, 14
Single-model clusters	7; 9; 19; 23

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1696 atoms, of which 838 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Sensor-like histidine kinase yojN.

Mol	Chain	Residues	Atoms						Trace
1	A	111	Total	C	H	N	O	S	0
			1696	525	838	144	182	7	

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	687	MET	-	INITIATING METHIONINE	UNP P39838
A	796	ARG	-	EXPRESSION TAG	UNP P39838
A	797	SER	-	EXPRESSION TAG	UNP P39838
A	798	HIS	-	EXPRESSION TAG	UNP P39838
A	799	HIS	-	EXPRESSION TAG	UNP P39838
A	800	HIS	-	EXPRESSION TAG	UNP P39838
A	801	HIS	-	EXPRESSION TAG	UNP P39838
A	802	HIS	-	EXPRESSION TAG	UNP P39838
A	803	HIS	-	EXPRESSION TAG	UNP P39838

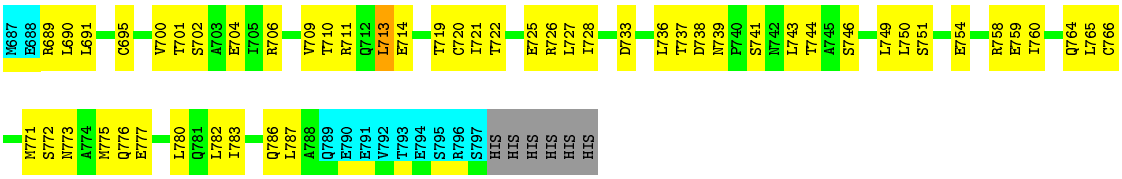
4 Residue-property plots

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: Sensor-like histidine kinase yojN

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)

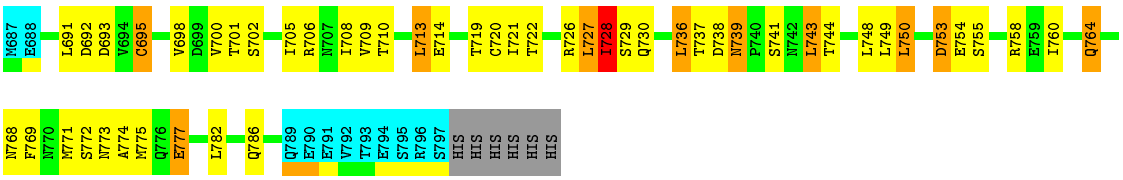
- Molecule 1: Sensor-like histidine kinase yojN

Chain A:

4.2.2 Score per residue for model 2

- Molecule 1: Sensor-like histidine kinase yojN

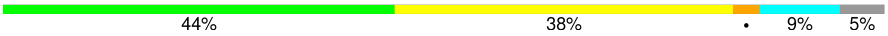
Chain A:

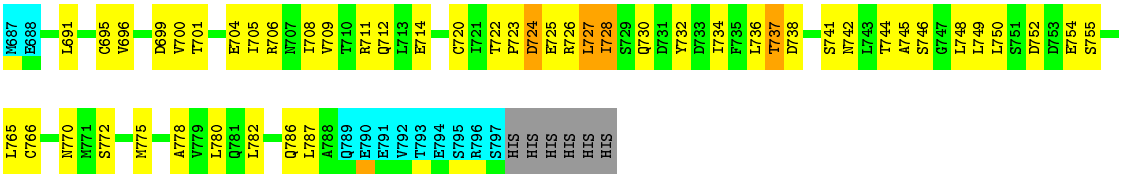


4.2.3 Score per residue for model 3

- Molecule 1: Sensor-like histidine kinase yojN

Chain A:

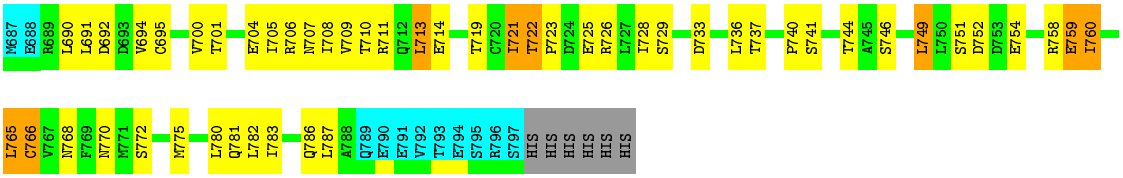




4.2.4 Score per residue for model 4

- Molecule 1: Sensor-like histidine kinase yojN

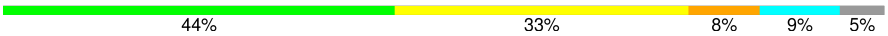
Chain A:

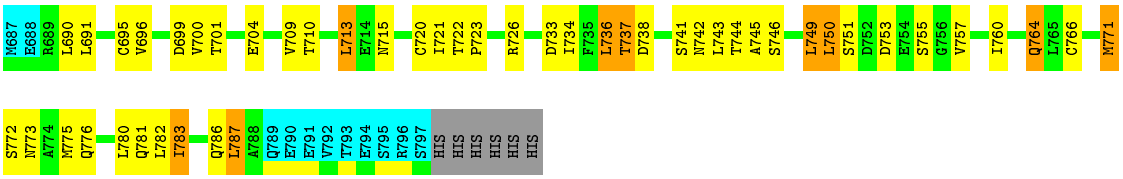


4.2.5 Score per residue for model 5

- Molecule 1: Sensor-like histidine kinase yojN

Chain A:

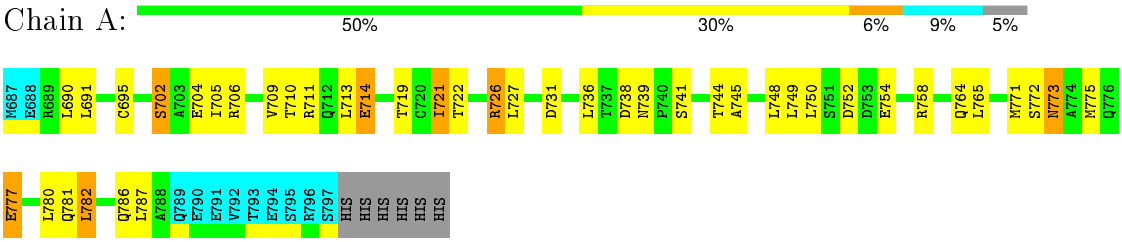




4.2.6 Score per residue for model 6

- Molecule 1: Sensor-like histidine kinase yojN

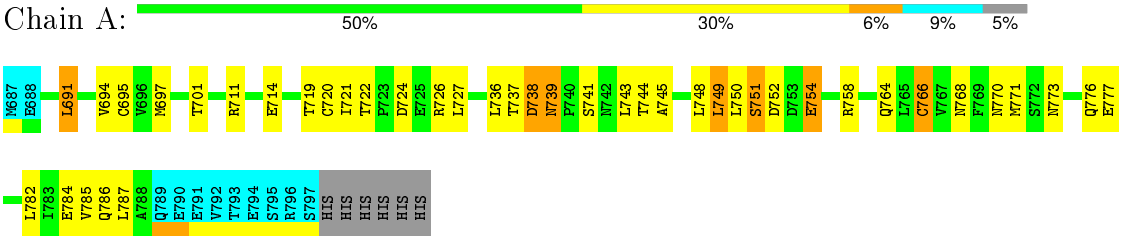
Chain A:



4.2.7 Score per residue for model 7

- Molecule 1: Sensor-like histidine kinase yojN

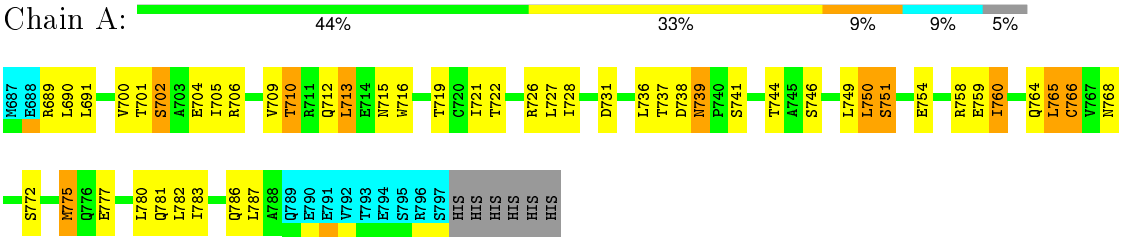
Chain A:



4.2.8 Score per residue for model 8

- Molecule 1: Sensor-like histidine kinase yojN

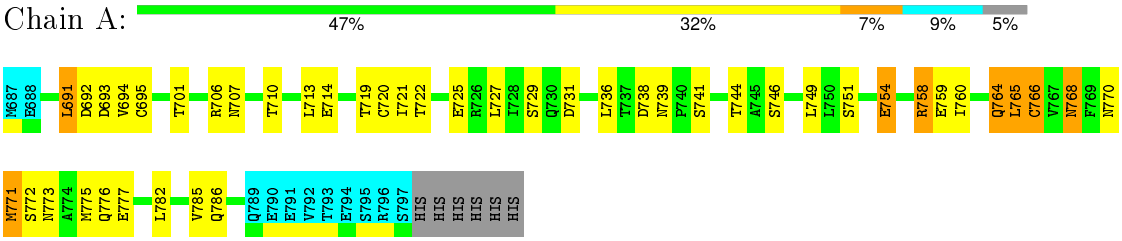
Chain A:



4.2.9 Score per residue for model 9

- Molecule 1: Sensor-like histidine kinase yojN

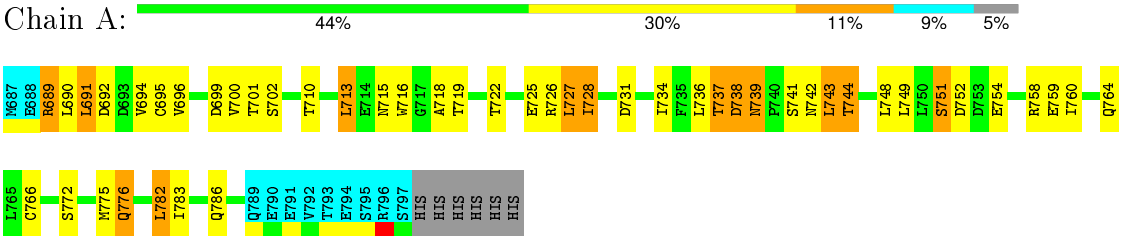
Chain A:



4.2.10 Score per residue for model 10

• Molecule 1: Sensor-like histidine kinase yojN

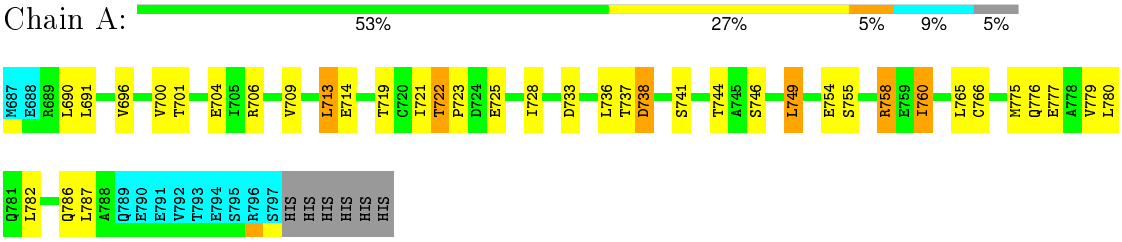
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: Sensor-like histidine kinase yojN

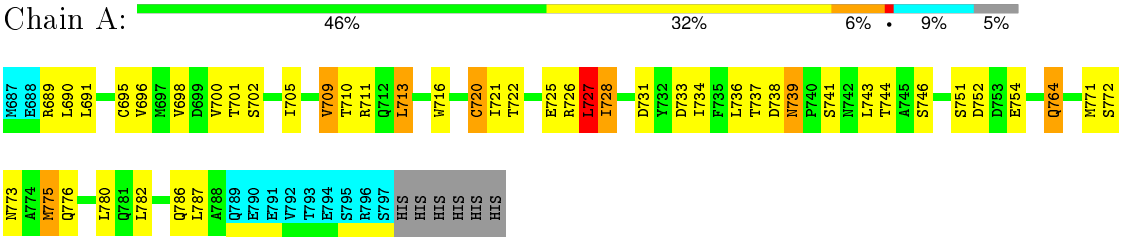
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: Sensor-like histidine kinase yojN

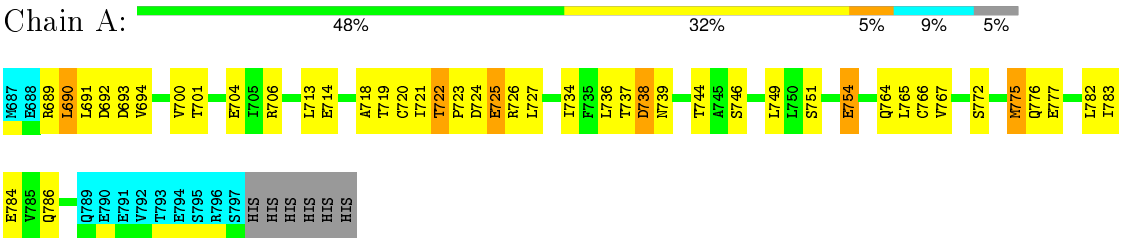
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Sensor-like histidine kinase yojN

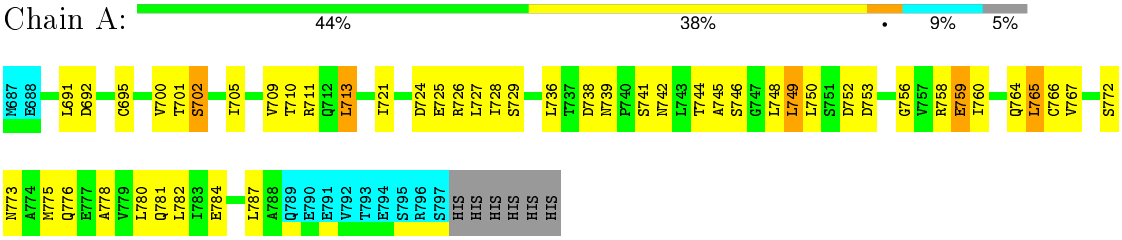
Chain A:



4.2.14 Score per residue for model 14

• Molecule 1: Sensor-like histidine kinase yojN

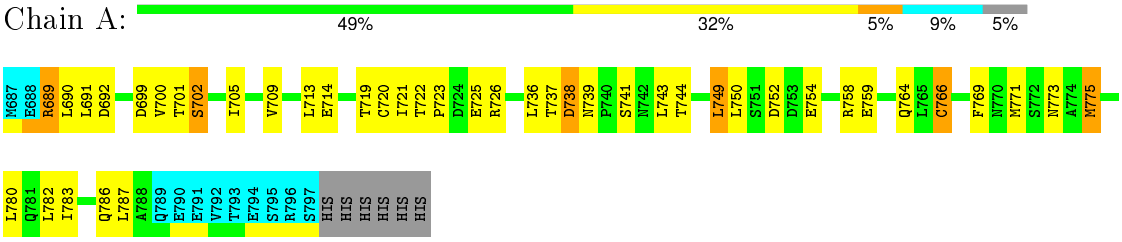
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Sensor-like histidine kinase yojN

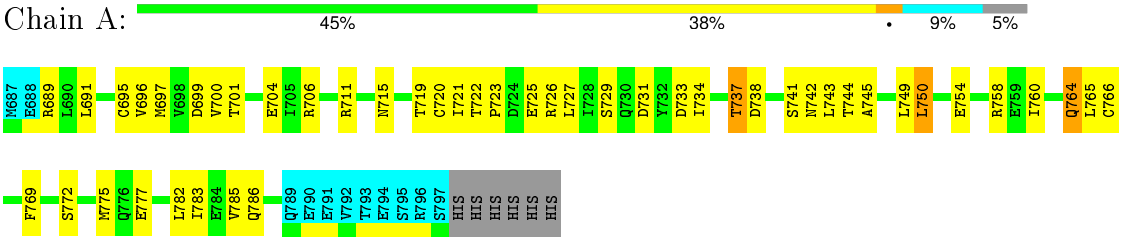
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Sensor-like histidine kinase yojN

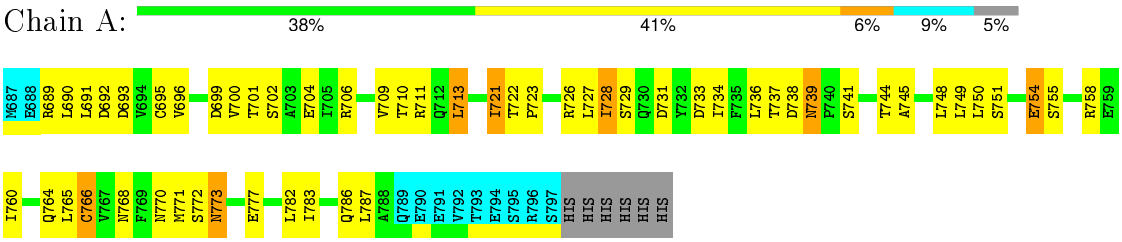
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: Sensor-like histidine kinase yojN

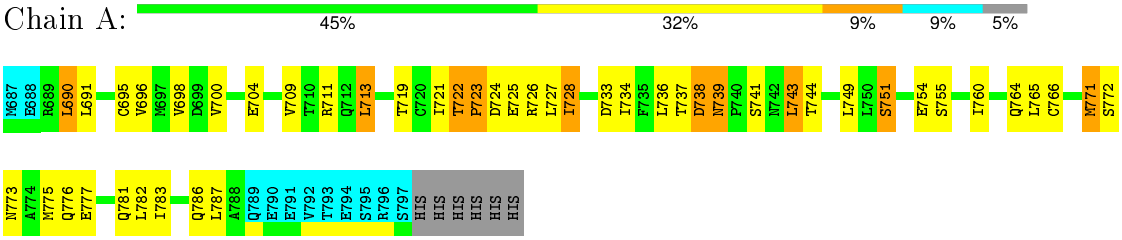
Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Sensor-like histidine kinase yojN

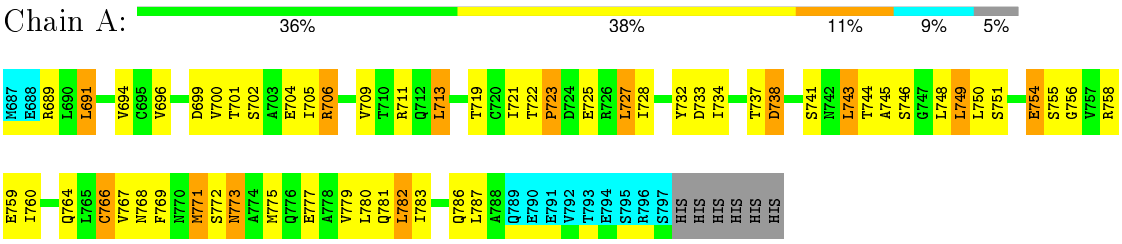
Chain A:



4.2.19 Score per residue for model 19

- Molecule 1: Sensor-like histidine kinase yojN

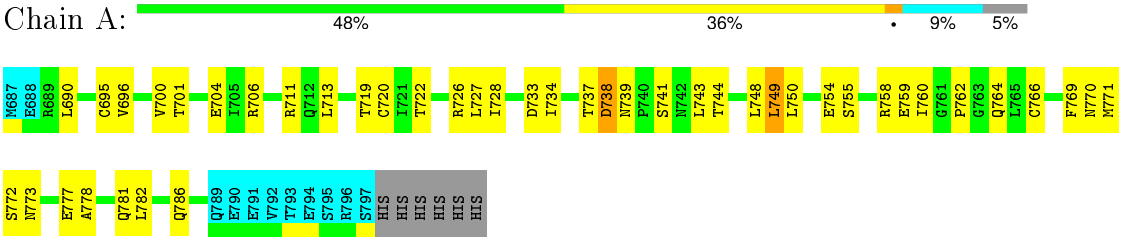
Chain A:



4.2.20 Score per residue for model 20

- Molecule 1: Sensor-like histidine kinase yojN

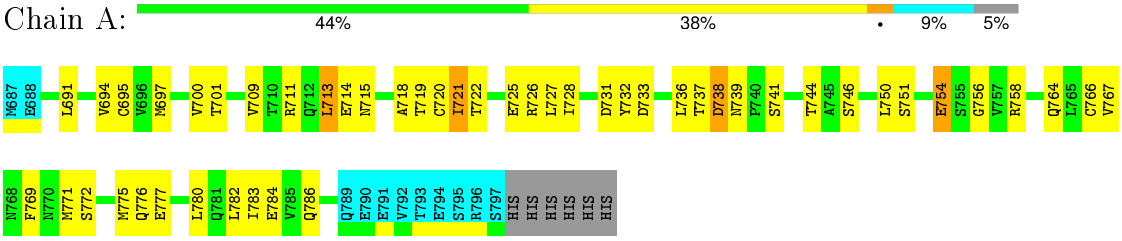
Chain A:



4.2.21 Score per residue for model 21

- Molecule 1: Sensor-like histidine kinase yojN

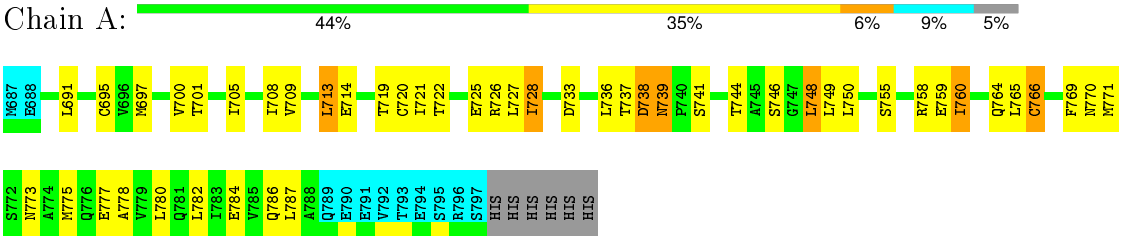
Chain A:



4.2.22 Score per residue for model 22

• Molecule 1: Sensor-like histidine kinase yojN

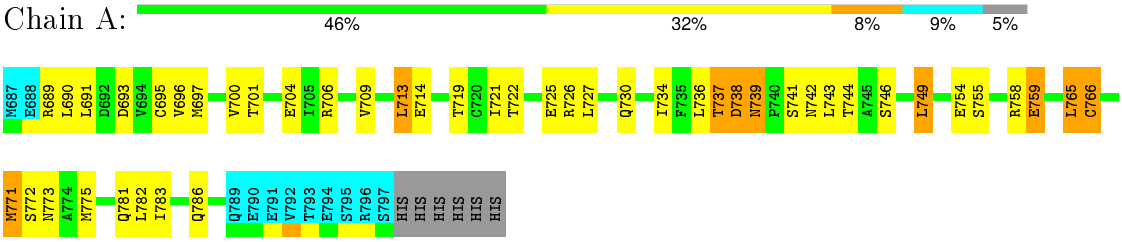
Chain A:



4.2.23 Score per residue for model 23

• Molecule 1: Sensor-like histidine kinase yojN

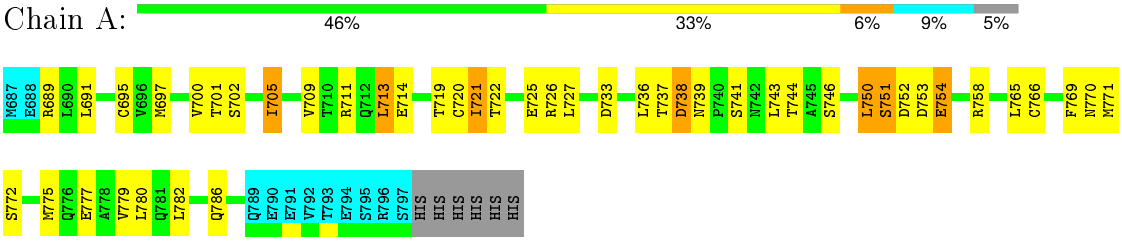
Chain A:



4.2.24 Score per residue for model 24

• Molecule 1: Sensor-like histidine kinase yojN

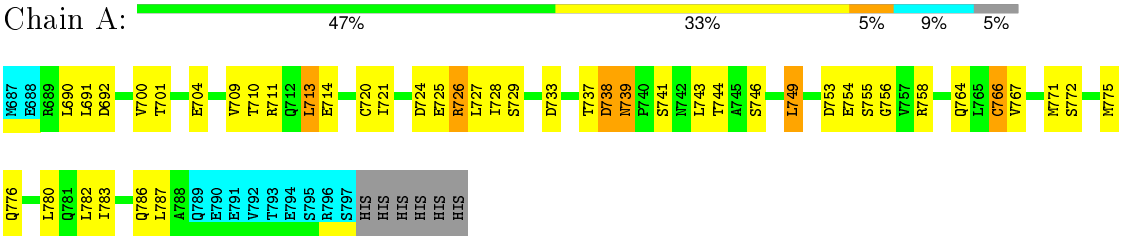
Chain A:



4.2.25 Score per residue for model 25

• Molecule 1: Sensor-like histidine kinase yojN

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 50 calculated structures, 25 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	

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

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	768	756	756	11±3
All	All	19200	18900	18900	267

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 7.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:699:ASP:HB2	1:A:723:PRO:HG2	0.90	1.43	17	3
1:A:751:SER:HB2	1:A:754:GLU:HB2	0.77	1.57	10	5
1:A:700:VAL:HG13	1:A:738:ASP:HB2	0.76	1.54	14	10
1:A:754:GLU:HG2	1:A:766:CYS:HB3	0.74	1.59	21	3
1:A:700:VAL:HA	1:A:738:ASP:HB3	0.69	1.63	10	6
1:A:695:CYS:HB3	1:A:721:ILE:HD11	0.69	1.64	5	9
1:A:700:VAL:HB	1:A:706:ARG:HG3	0.69	1.63	3	4
1:A:759:GLU:HG3	1:A:765:LEU:HG	0.66	1.66	22	5
1:A:750:LEU:HD22	1:A:769:PHE:HB3	0.64	1.70	15	3
1:A:724:ASP:HB2	1:A:730:GLN:HG2	0.63	1.69	3	1
1:A:700:VAL:HG13	1:A:738:ASP:HB3	0.62	1.71	13	6
1:A:696:VAL:HG22	1:A:734:ILE:HB	0.61	1.73	19	10
1:A:737:THR:HB	1:A:743:LEU:HD12	0.60	1.74	19	1
1:A:754:GLU:HG2	1:A:766:CYS:HB2	0.59	1.74	7	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:771:MET:SD	1:A:773:ASN:HB2	0.59	2.38	5	5
1:A:709:VAL:O	1:A:713:LEU:HG	0.58	1.98	23	15
1:A:760:ILE:HG12	1:A:766:CYS:SG	0.58	2.39	1	4
1:A:709:VAL:HG21	1:A:750:LEU:HD21	0.57	1.75	8	1
1:A:710:THR:HA	1:A:713:LEU:HD12	0.57	1.77	2	9
1:A:734:ILE:HD11	1:A:783:ILE:HG13	0.57	1.75	5	1
1:A:728:ILE:HG13	1:A:744:THR:HG21	0.57	1.76	10	1
1:A:696:VAL:HB	1:A:713:LEU:HD11	0.57	1.77	11	1
1:A:751:SER:HB3	1:A:754:GLU:HB2	0.57	1.76	13	1
1:A:700:VAL:HG13	1:A:738:ASP:CB	0.57	2.30	1	15
1:A:769:PHE:HA	1:A:775:MET:HG3	0.57	1.76	16	2
1:A:722:THR:HG22	1:A:723:PRO:HD2	0.56	1.75	11	4
1:A:702:SER:HB2	1:A:705:ILE:HB	0.56	1.76	6	8
1:A:738:ASP:HA	1:A:750:LEU:HB2	0.56	1.75	19	4
1:A:743:LEU:HD22	1:A:764:GLN:HG3	0.56	1.77	10	3
1:A:724:ASP:HB3	1:A:727:LEU:HB2	0.56	1.78	25	1
1:A:749:LEU:HG	1:A:766:CYS:HA	0.55	1.79	7	7
1:A:690:LEU:HG	1:A:776:GLN:HG3	0.55	1.78	11	4
1:A:716:TRP:HH2	1:A:775:MET:HB3	0.54	1.61	12	2
1:A:758:ARG:HD2	1:A:760:ILE:HG23	0.53	1.81	11	1
1:A:709:VAL:HG11	1:A:750:LEU:HD21	0.53	1.79	24	1
1:A:748:LEU:HD11	1:A:778:ALA:HB1	0.52	1.81	22	1
1:A:691:LEU:HB3	1:A:694:VAL:HB	0.52	1.80	9	3
1:A:773:ASN:O	1:A:777:GLU:HG3	0.52	2.04	17	2
1:A:727:LEU:O	1:A:728:ILE:HB	0.52	2.04	3	5
1:A:749:LEU:HD21	1:A:760:ILE:HD11	0.52	1.82	11	2
1:A:690:LEU:HD21	1:A:776:GLN:HG3	0.51	1.83	18	1
1:A:746:SER:HA	1:A:764:GLN:OE1	0.51	2.06	12	1
1:A:760:ILE:HG13	1:A:764:GLN:HB3	0.51	1.81	16	3
1:A:736:LEU:HG	1:A:748:LEU:HB2	0.51	1.82	2	1
1:A:754:GLU:HG3	1:A:758:ARG:HG2	0.50	1.84	9	1
1:A:740:PRO:HB3	1:A:749:LEU:HD22	0.50	1.82	4	1
1:A:699:ASP:O	1:A:737:THR:HA	0.49	2.07	16	4
1:A:709:VAL:O	1:A:713:LEU:HB2	0.49	2.08	11	1
1:A:705:ILE:HA	1:A:708:ILE:HD12	0.49	1.82	3	4
1:A:754:GLU:HB3	1:A:768:ASN:HB2	0.49	1.85	8	1
1:A:750:LEU:HB3	1:A:769:PHE:HD2	0.49	1.68	22	2
1:A:697:MET:SD	1:A:723:PRO:HA	0.49	2.47	16	1
1:A:748:LEU:HG	1:A:782:LEU:HG	0.48	1.85	6	3
1:A:714:GLU:HG2	1:A:720:CYS:SG	0.48	2.48	15	1
1:A:759:GLU:HA	1:A:765:LEU:HG	0.48	1.85	23	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:738:ASP:O	1:A:739:ASN:HB2	0.48	2.09	7	9
1:A:751:SER:O	1:A:768:ASN:HA	0.48	2.09	7	5
1:A:748:LEU:HD21	1:A:778:ALA:HB1	0.48	1.86	14	2
1:A:698:VAL:HA	1:A:736:LEU:O	0.47	2.09	2	1
1:A:724:ASP:CB	1:A:730:GLN:HG2	0.47	2.39	3	1
1:A:756:GLY:O	1:A:767:VAL:HA	0.47	2.09	25	4
1:A:689:ARG:CD	1:A:692:ASP:HA	0.47	2.39	15	1
1:A:783:ILE:O	1:A:787:LEU:HB2	0.47	2.09	5	1
1:A:767:VAL:HG21	1:A:775:MET:HA	0.47	1.86	13	1
1:A:710:THR:O	1:A:714:GLU:HB2	0.46	2.10	4	3
1:A:700:VAL:HB	1:A:706:ARG:HB2	0.46	1.86	4	4
1:A:691:LEU:HD12	1:A:694:VAL:HB	0.46	1.86	10	1
1:A:699:ASP:HB2	1:A:723:PRO:HG3	0.46	1.88	5	1
1:A:750:LEU:HD21	1:A:769:PHE:HB3	0.46	1.87	2	1
1:A:689:ARG:HE	1:A:692:ASP:HA	0.45	1.71	13	1
1:A:737:THR:O	1:A:749:LEU:HA	0.45	2.11	23	2
1:A:709:VAL:HG11	1:A:750:LEU:HD13	0.44	1.89	6	2
1:A:710:THR:HG23	1:A:720:CYS:SG	0.44	2.53	9	3
1:A:749:LEU:O	1:A:749:LEU:HG	0.44	2.13	20	1
1:A:773:ASN:O	1:A:777:GLU:HG2	0.43	2.12	6	1
1:A:697:MET:HB2	1:A:732:TYR:HB3	0.43	1.89	21	1
1:A:726:ARG:HA	1:A:726:ARG:HD3	0.43	1.56	25	1
1:A:774:ALA:HA	1:A:777:GLU:HG3	0.43	1.91	2	1
1:A:694:VAL:O	1:A:718:ALA:HA	0.42	2.13	13	2
1:A:724:ASP:OD2	1:A:727:LEU:HB3	0.42	2.14	3	1
1:A:689:ARG:HB3	1:A:716:TRP:O	0.42	2.14	10	1
1:A:748:LEU:HD22	1:A:778:ALA:HB1	0.42	1.91	20	1
1:A:723:PRO:O	1:A:725:GLU:HG3	0.42	2.15	13	1
1:A:724:ASP:HB3	1:A:727:LEU:HD12	0.42	1.92	14	1
1:A:736:LEU:HD23	1:A:750:LEU:HD21	0.42	1.91	5	1
1:A:757:VAL:HB	1:A:767:VAL:HG23	0.41	1.92	1	1
1:A:726:ARG:HD3	1:A:726:ARG:HA	0.41	1.61	6	1
1:A:750:LEU:HG	1:A:769:PHE:HB3	0.41	1.92	24	1
1:A:757:VAL:HA	1:A:766:CYS:O	0.41	2.16	5	1
1:A:749:LEU:HG	1:A:749:LEU:O	0.41	2.16	1	1
1:A:750:LEU:HA	1:A:750:LEU:HD12	0.41	1.78	16	1
1:A:691:LEU:HG	1:A:718:ALA:HB2	0.41	1.92	10	1
1:A:737:THR:HG23	1:A:742:ASN:HB3	0.41	1.93	10	1
1:A:782:LEU:HA	1:A:782:LEU:HD22	0.41	1.77	19	1
1:A:705:ILE:O	1:A:709:VAL:HG23	0.40	2.15	3	1
1:A:694:VAL:HG13	1:A:733:ASP:HB2	0.40	1.93	4	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:695:CYS:CB	1:A:721:ILE:HD11	0.40	2.45	2	1
1:A:727:LEU:HB3	1:A:728:ILE:H	0.40	1.46	2	1
1:A:759:GLU:OE2	1:A:762:PRO:HA	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	100/117 (85%)	90±2 (90±2%)	8±2 (8±2%)	2±1 (2±1%)	15	58
All	All	2500/2925 (85%)	2262 (90%)	193 (8%)	45 (2%)	15	58

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	739	ASN	16
1	A	728	ILE	10
1	A	745	ALA	8
1	A	692	ASP	4
1	A	723	PRO	3
1	A	729	SER	1
1	A	727	LEU	1
1	A	753	ASP	1
1	A	725	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	89/106 (84%)	54±3 (61±3%)	35±3 (39±3%)	1 5
All	All	2225/2650 (84%)	1347 (61%)	878 (39%)	1 5

All 77 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	782	LEU	25
1	A	744	THR	25
1	A	741	SER	24
1	A	786	GLN	24
1	A	691	LEU	24
1	A	701	THR	23
1	A	722	THR	23
1	A	713	LEU	22
1	A	749	LEU	22
1	A	772	SER	21
1	A	726	ARG	21
1	A	736	LEU	21
1	A	775	MET	21
1	A	737	THR	20
1	A	766	CYS	20
1	A	758	ARG	20
1	A	719	THR	19
1	A	721	ILE	18
1	A	754	GLU	17
1	A	727	LEU	17
1	A	764	GLN	17
1	A	771	MET	16
1	A	725	GLU	16
1	A	711	ARG	15
1	A	787	LEU	15
1	A	746	SER	15
1	A	704	GLU	14
1	A	777	GLU	14
1	A	780	LEU	14
1	A	765	LEU	14
1	A	738	ASP	13
1	A	743	LEU	13
1	A	783	ILE	13
1	A	714	GLU	12
1	A	733	ASP	12
1	A	720	CYS	12

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	755	SER	11
1	A	695	CYS	11
1	A	760	ILE	11
1	A	690	LEU	10
1	A	689	ARG	10
1	A	752	ASP	10
1	A	781	GLN	10
1	A	773	ASN	9
1	A	770	ASN	9
1	A	750	LEU	9
1	A	728	ILE	8
1	A	731	ASP	8
1	A	776	GLN	8
1	A	751	SER	8
1	A	759	GLU	7
1	A	729	SER	6
1	A	784	GLU	6
1	A	702	SER	6
1	A	715	ASN	6
1	A	706	ARG	6
1	A	693	ASP	5
1	A	753	ASP	5
1	A	742	ASN	5
1	A	739	ASN	4
1	A	697	MET	4
1	A	748	LEU	4
1	A	724	ASP	4
1	A	692	ASP	3
1	A	779	VAL	3
1	A	768	ASN	3
1	A	707	ASN	2
1	A	698	VAL	2
1	A	730	GLN	2
1	A	785	VAL	2
1	A	732	TYR	2
1	A	712	GLN	2
1	A	734	ILE	1
1	A	705	ILE	1
1	A	699	ASP	1
1	A	709	VAL	1
1	A	710	THR	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