wwPDB X-ray Structure Validation Summary Report (i)

Mar 24, 2022 – 12:34 pm GMT

PDB ID	:	5LOO
Title	:	Structure of full length unliganded CodY from Bacillus subtilis
Authors	:	Wilkinson, A.J.; Levdikov, V.M.; Blagova, E.V.
Deposited on	:	2016-08-09
Resolution	:	4.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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/XrayValidationReportHelp
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:

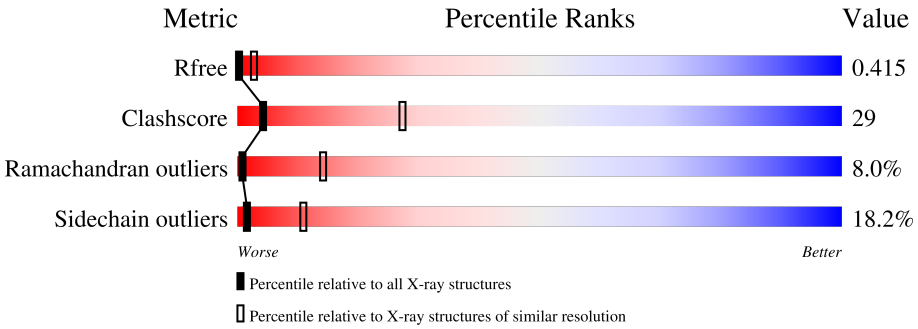
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.27
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 4.50 Å.

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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1055 (5.20-3.80)
Clashscore	141614	1123 (5.20-3.80)
Ramachandran outliers	138981	1069 (5.20-3.80)
Sidechain outliers	138945	1050 (5.20-3.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	263	30% 49% 17%
1	B	263	38% 44% 14%
1	C	263	32% 50% 16%
1	D	263	34% 49% 13%
1	E	263	39% 45% 13%
1	F	263	32% 52% 14%
1	G	263	48% 41% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	263	41% 44% 11%
1	I	263	42% 46% 9%
1	J	263	41% 44% 11%
1	K	263	40% 47% 10%
1	L	263	30% 53% 13%
1	M	263	61% 31% 6%
1	N	263	62% 29% 6%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 28154 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called GTP-sensing transcriptional pleiotropic repressor CodY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	B	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	C	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	D	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	E	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	F	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	G	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	H	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	I	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	J	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	K	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	L	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	M	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			
1	N	257	Total	C	N	O	S	0	0	0
			2011	1265	341	398	7			

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	260	HIS	-	expression tag	UNP P39779

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	261	HIS	-	expression tag	UNP P39779
A	262	HIS	-	expression tag	UNP P39779
A	263	HIS	-	expression tag	UNP P39779
A	264	HIS	-	expression tag	UNP P39779
B	260	HIS	-	expression tag	UNP P39779
B	261	HIS	-	expression tag	UNP P39779
B	262	HIS	-	expression tag	UNP P39779
B	263	HIS	-	expression tag	UNP P39779
B	264	HIS	-	expression tag	UNP P39779
C	260	HIS	-	expression tag	UNP P39779
C	261	HIS	-	expression tag	UNP P39779
C	262	HIS	-	expression tag	UNP P39779
C	263	HIS	-	expression tag	UNP P39779
C	264	HIS	-	expression tag	UNP P39779
D	260	HIS	-	expression tag	UNP P39779
D	261	HIS	-	expression tag	UNP P39779
D	262	HIS	-	expression tag	UNP P39779
D	263	HIS	-	expression tag	UNP P39779
D	264	HIS	-	expression tag	UNP P39779
E	260	HIS	-	expression tag	UNP P39779
E	261	HIS	-	expression tag	UNP P39779
E	262	HIS	-	expression tag	UNP P39779
E	263	HIS	-	expression tag	UNP P39779
E	264	HIS	-	expression tag	UNP P39779
F	260	HIS	-	expression tag	UNP P39779
F	261	HIS	-	expression tag	UNP P39779
F	262	HIS	-	expression tag	UNP P39779
F	263	HIS	-	expression tag	UNP P39779
F	264	HIS	-	expression tag	UNP P39779
G	260	HIS	-	expression tag	UNP P39779
G	261	HIS	-	expression tag	UNP P39779
G	262	HIS	-	expression tag	UNP P39779
G	263	HIS	-	expression tag	UNP P39779
G	264	HIS	-	expression tag	UNP P39779
H	260	HIS	-	expression tag	UNP P39779
H	261	HIS	-	expression tag	UNP P39779
H	262	HIS	-	expression tag	UNP P39779
H	263	HIS	-	expression tag	UNP P39779
H	264	HIS	-	expression tag	UNP P39779
I	260	HIS	-	expression tag	UNP P39779
I	261	HIS	-	expression tag	UNP P39779
I	262	HIS	-	expression tag	UNP P39779

Continued on next page...

Continued from previous page...

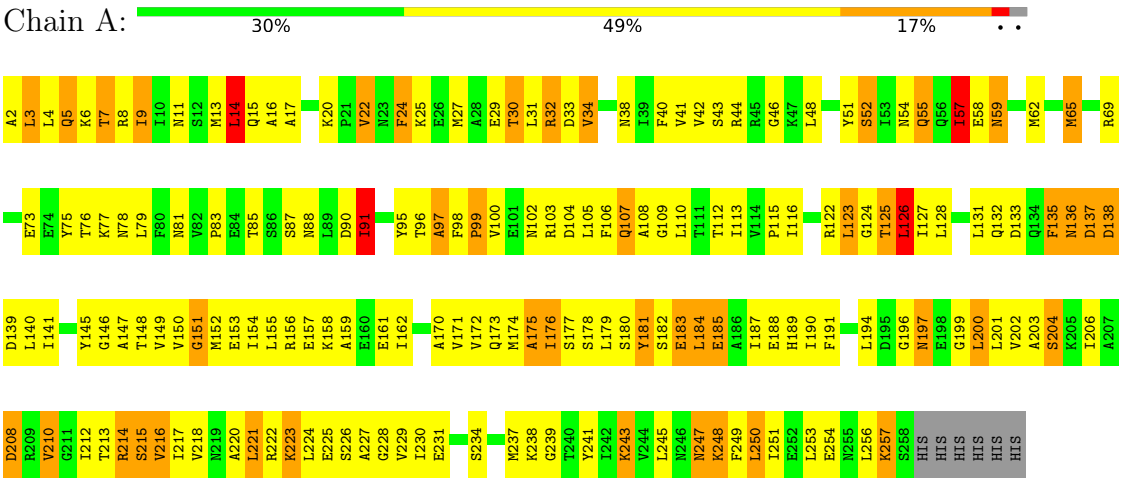
Chain	Residue	Modelled	Actual	Comment	Reference
I	263	HIS	-	expression tag	UNP P39779
I	264	HIS	-	expression tag	UNP P39779
J	260	HIS	-	expression tag	UNP P39779
J	261	HIS	-	expression tag	UNP P39779
J	262	HIS	-	expression tag	UNP P39779
J	263	HIS	-	expression tag	UNP P39779
J	264	HIS	-	expression tag	UNP P39779
K	260	HIS	-	expression tag	UNP P39779
K	261	HIS	-	expression tag	UNP P39779
K	262	HIS	-	expression tag	UNP P39779
K	263	HIS	-	expression tag	UNP P39779
K	264	HIS	-	expression tag	UNP P39779
L	260	HIS	-	expression tag	UNP P39779
L	261	HIS	-	expression tag	UNP P39779
L	262	HIS	-	expression tag	UNP P39779
L	263	HIS	-	expression tag	UNP P39779
L	264	HIS	-	expression tag	UNP P39779
M	260	HIS	-	expression tag	UNP P39779
M	261	HIS	-	expression tag	UNP P39779
M	262	HIS	-	expression tag	UNP P39779
M	263	HIS	-	expression tag	UNP P39779
M	264	HIS	-	expression tag	UNP P39779
N	260	HIS	-	expression tag	UNP P39779
N	261	HIS	-	expression tag	UNP P39779
N	262	HIS	-	expression tag	UNP P39779
N	263	HIS	-	expression tag	UNP P39779
N	264	HIS	-	expression tag	UNP P39779

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-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 outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

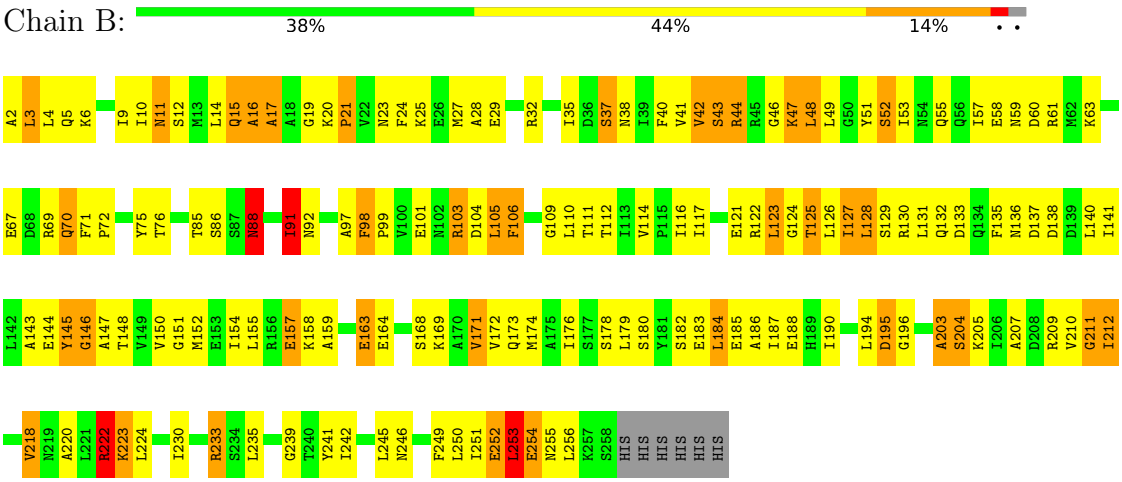
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain A:



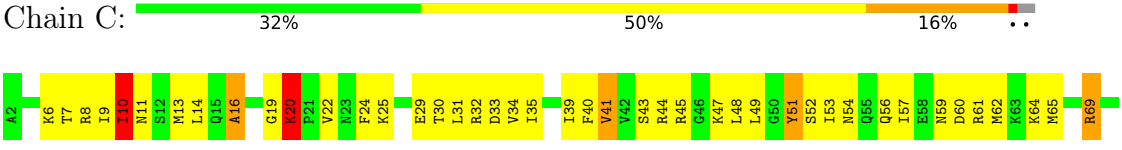
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

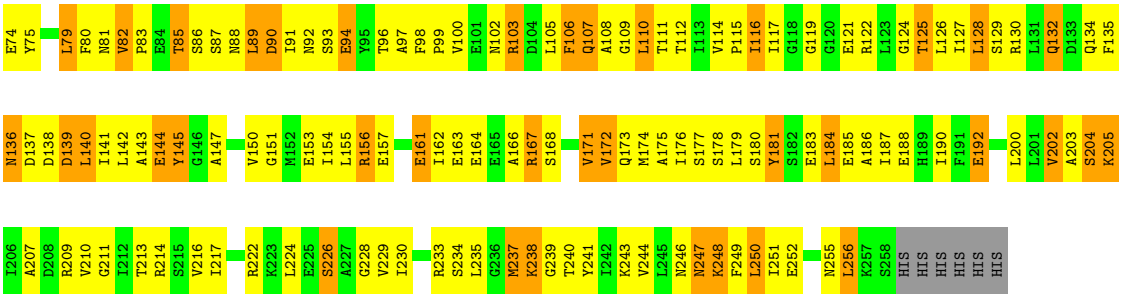
Chain B:



• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

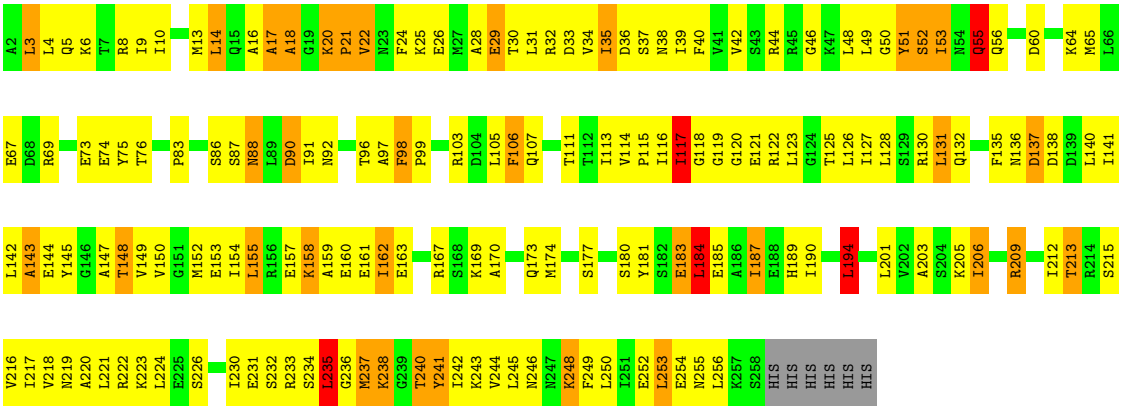
Chain C:





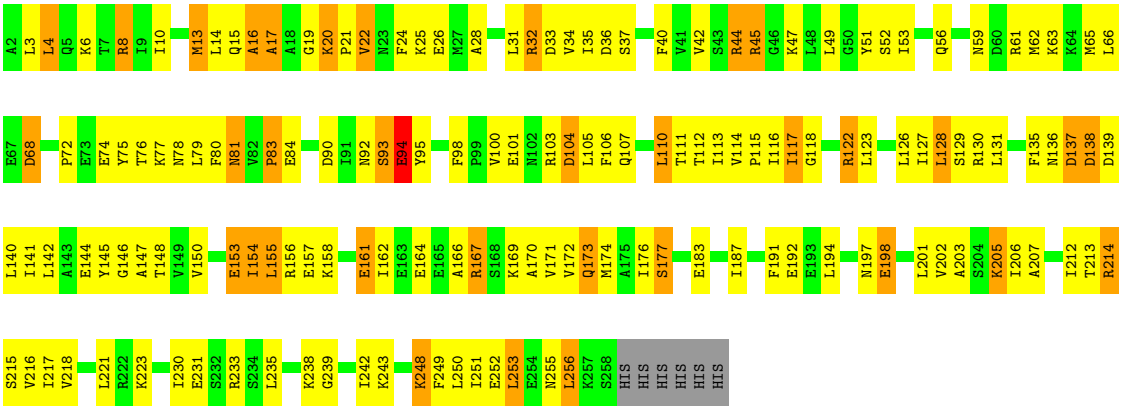
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain D:



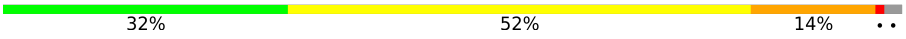
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

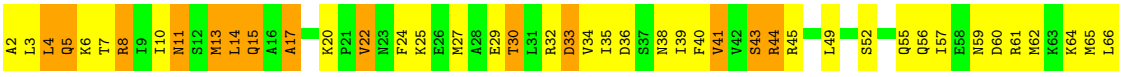
Chain E:

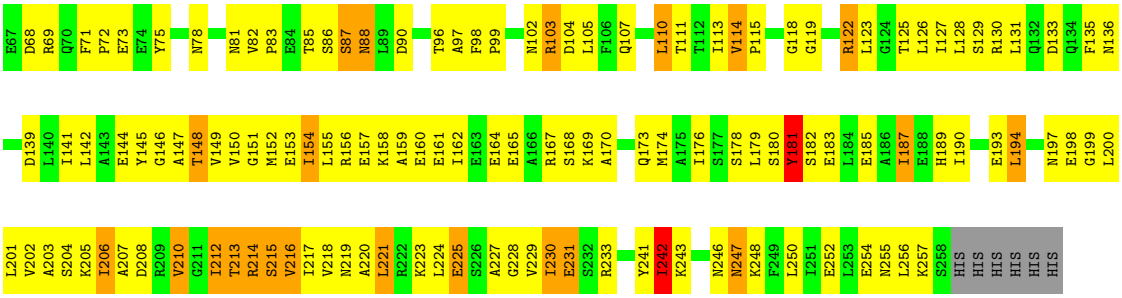


• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain F:

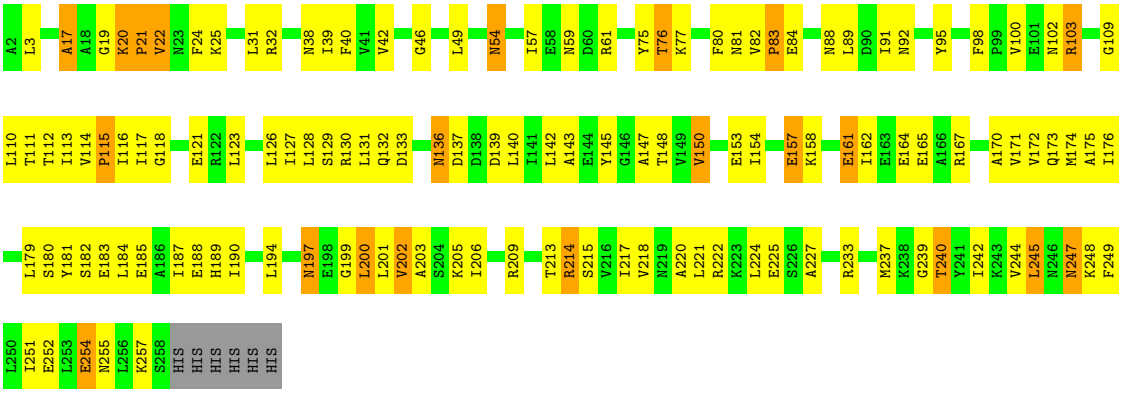






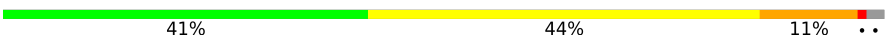
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

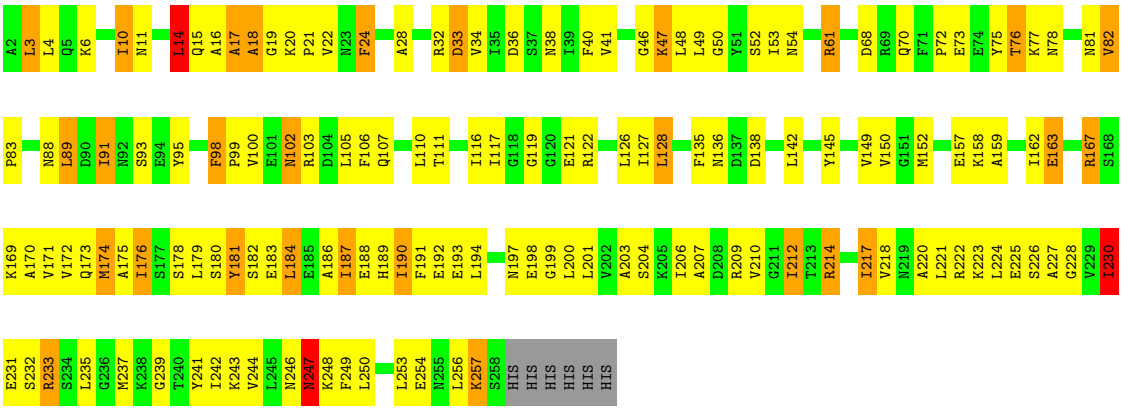
Chain G:



• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

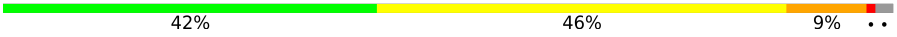
Chain H:

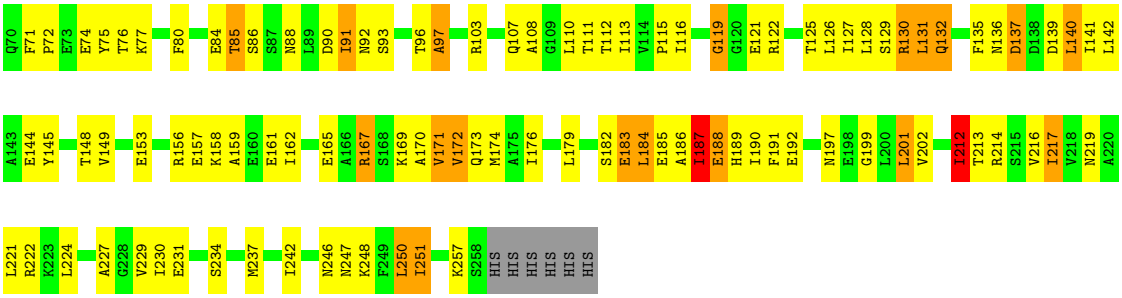




• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

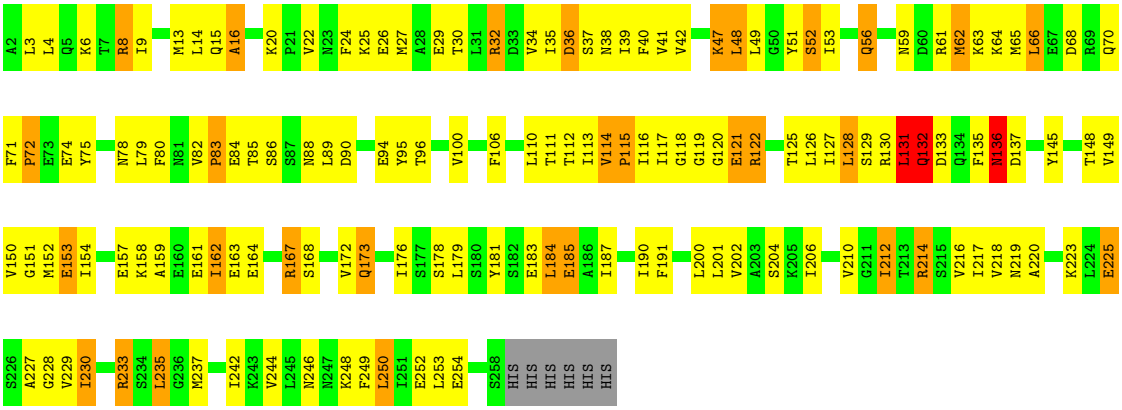
Chain I:





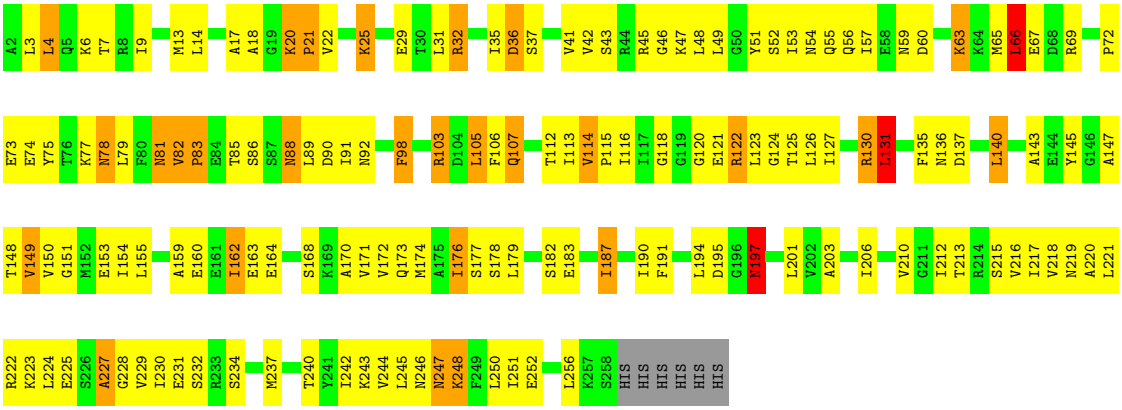
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain J:



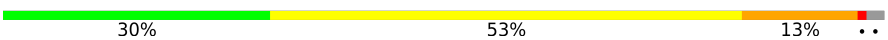
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

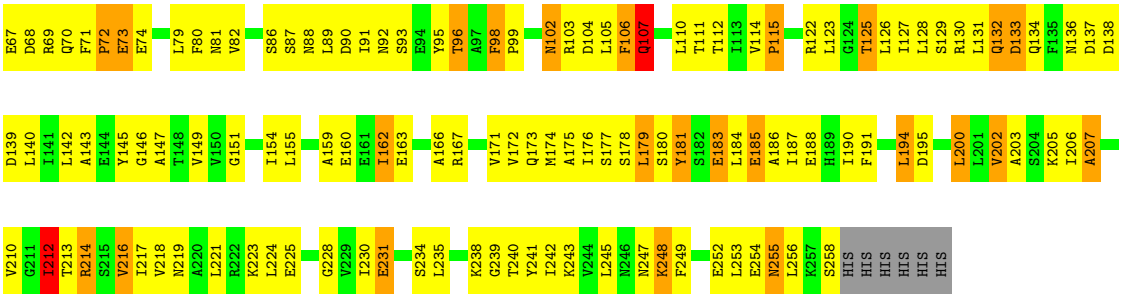
Chain K:



• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

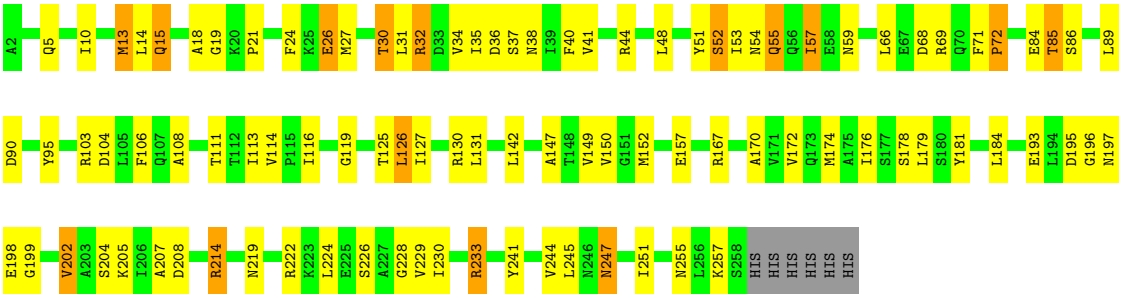
Chain L:





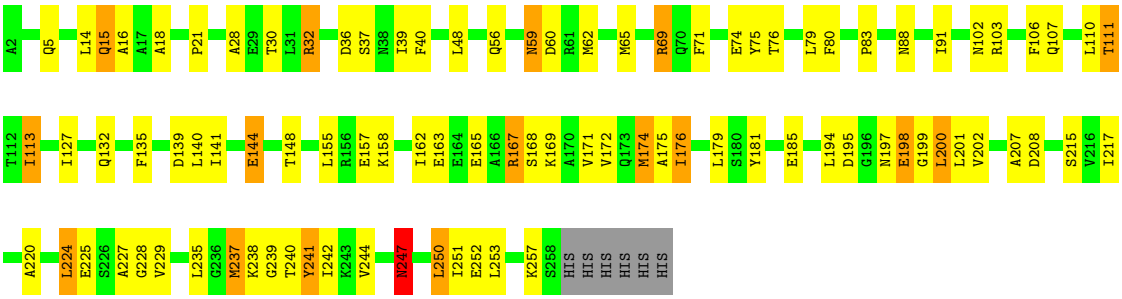
• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain M:



• Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain N:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	315.63Å 113.68Å 168.59Å 90.00° 113.23° 90.00°	Depositor
Resolution (Å)	19.88 – 4.50 19.88 – 4.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.88-4.50) 89.7 (19.88-4.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.98 (at 4.54Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.251 , 0.430 0.252 , 0.415	Depositor DCC
R_{free} test set	1479 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	199.3	Xtriage
Anisotropy	0.172	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.19$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	28154	wwPDB-VP
Average B, all atoms (Å^2)	234.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.05% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 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 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	0.57	0/2031	0.79	3/2735 (0.1%)
1	B	0.54	0/2031	0.72	1/2735 (0.0%)
1	C	0.57	0/2031	0.75	1/2735 (0.0%)
1	D	0.55	0/2031	0.75	1/2735 (0.0%)
1	E	0.61	2/2031 (0.1%)	0.73	3/2735 (0.1%)
1	F	0.54	0/2031	0.73	0/2735
1	G	0.50	0/2031	0.69	0/2735
1	H	0.53	1/2031 (0.0%)	0.67	0/2735
1	I	0.52	0/2031	0.71	1/2735 (0.0%)
1	J	0.51	0/2031	0.71	1/2735 (0.0%)
1	K	0.49	0/2031	0.68	0/2735
1	L	0.50	0/2031	0.71	1/2735 (0.0%)
1	M	0.42	0/2031	0.56	0/2735
1	N	0.49	2/2031 (0.1%)	0.58	0/2735
All	All	0.53	5/28434 (0.0%)	0.70	12/38290 (0.0%)

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 outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	D	0	2
1	F	0	2
1	I	0	1
1	J	0	1
1	L	0	1
All	All	0	9

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	20	LYS	CE-NZ	13.13	1.81	1.49
1	H	257	LYS	CE-NZ	9.42	1.72	1.49
1	N	144	GLU	CD-OE1	7.69	1.34	1.25
1	N	144	GLU	CD-OE2	7.13	1.33	1.25
1	E	20	LYS	CD-CE	5.13	1.64	1.51

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	110	LEU	CA-CB-CG	7.05	131.52	115.30
1	E	20	LYS	CD-CE-NZ	-6.61	96.51	111.70
1	A	14	LEU	CA-CB-CG	5.96	129.01	115.30
1	A	200	LEU	CA-CB-CG	5.77	128.56	115.30
1	A	126	LEU	CA-CB-CG	5.72	128.46	115.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	15	GLN	Peptide
1	B	222	ARG	Peptide
1	D	14	LEU	Peptide
1	D	16	ALA	Peptide
1	F	60	ASP	Peptide

5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2011	0	2048	216	0
1	B	2011	0	2048	159	0
1	C	2011	0	2048	174	0
1	D	2011	0	2048	152	0
1	E	2011	0	2048	129	0
1	F	2011	0	2048	155	0
1	G	2011	0	2047	117	0
1	H	2011	0	2047	129	0
1	I	2011	0	2048	107	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	J	2011	0	2048	128	0
1	K	2011	0	2047	118	0
1	L	2011	0	2048	154	0
1	M	2011	0	2048	57	0
1	N	2011	0	2048	49	0
All	All	28154	0	28669	1672	0

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 29.

The worst 5 of 1672 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:257:LYS:NZ	1:H:257:LYS:CE	1.72	1.47
1:E:20:LYS:NZ	1:E:20:LYS:CE	1.81	1.41
1:G:145:TYR:HA	1:H:145:TYR:CZ	1.63	1.31
1:K:145:TYR:CZ	1:L:145:TYR:CA	2.12	1.27
1:K:145:TYR:CZ	1:L:145:TYR:HA	1.67	1.20

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 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 X-ray entries followed by that with respect to entries of similar resolution.

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	255/263 (97%)	174 (68%)	56 (22%)	25 (10%)	0 10
1	B	255/263 (97%)	177 (69%)	54 (21%)	24 (9%)	0 12
1	C	255/263 (97%)	169 (66%)	59 (23%)	27 (11%)	0 8
1	D	255/263 (97%)	171 (67%)	55 (22%)	29 (11%)	0 7
1	E	255/263 (97%)	183 (72%)	53 (21%)	19 (8%)	1 15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	F	255/263 (97%)	169 (66%)	60 (24%)	26 (10%)	0 10
1	G	255/263 (97%)	187 (73%)	55 (22%)	13 (5%)	2 22
1	H	255/263 (97%)	178 (70%)	59 (23%)	18 (7%)	1 16
1	I	255/263 (97%)	160 (63%)	74 (29%)	21 (8%)	1 13
1	J	255/263 (97%)	179 (70%)	51 (20%)	25 (10%)	0 10
1	K	255/263 (97%)	177 (69%)	58 (23%)	20 (8%)	1 15
1	L	255/263 (97%)	169 (66%)	67 (26%)	19 (8%)	1 15
1	M	255/263 (97%)	196 (77%)	48 (19%)	11 (4%)	2 25
1	N	255/263 (97%)	198 (78%)	50 (20%)	7 (3%)	5 34
All	All	3570/3682 (97%)	2487 (70%)	799 (22%)	284 (8%)	1 14

5 of 284 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	LEU
1	A	14	LEU
1	A	52	SER
1	A	59	ASN
1	A	97	ALA

5.3.2 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 X-ray entries followed by that with respect to entries of similar resolution.

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	222/229 (97%)	182 (82%)	40 (18%)	1 11
1	B	222/229 (97%)	176 (79%)	46 (21%)	1 7
1	C	222/229 (97%)	180 (81%)	42 (19%)	1 9
1	D	222/229 (97%)	179 (81%)	43 (19%)	1 9
1	E	222/229 (97%)	169 (76%)	53 (24%)	0 5
1	F	222/229 (97%)	173 (78%)	49 (22%)	1 6
1	G	222/229 (97%)	189 (85%)	33 (15%)	3 17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	H	222/229 (97%)	182 (82%)	40 (18%)	1 11
1	I	222/229 (97%)	180 (81%)	42 (19%)	1 9
1	J	222/229 (97%)	190 (86%)	32 (14%)	3 18
1	K	222/229 (97%)	178 (80%)	44 (20%)	1 9
1	L	222/229 (97%)	181 (82%)	41 (18%)	1 10
1	M	222/229 (97%)	195 (88%)	27 (12%)	5 22
1	N	222/229 (97%)	189 (85%)	33 (15%)	3 17
All	All	3108/3206 (97%)	2543 (82%)	565 (18%)	1 11

5 of 565 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	251	ILE
1	L	69	ARG
1	K	248	LYS
1	M	157	GLU
1	E	107	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 91 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	88	ASN
1	K	197	ASN
1	I	132	GLN
1	J	136	ASN
1	L	23	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.