wwPDB X-ray Structure Validation Summary Report (i)

Nov 8, 2021 – 08:07 pm GMT

PDB ID	:	7OLH
Title	:	Bacillus subtilis Complex structure 1 of diadenylate cyclase CdaA cytoplasmic domain (CdaACD) and the phosphoglucomutase GlmM short variant (GlmMF369)
Authors	:	Pathania, M.; Grundling, A.G.; Freemont, P.
Deposited on	:	2021-05-20
Resolution	:	3.65 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:

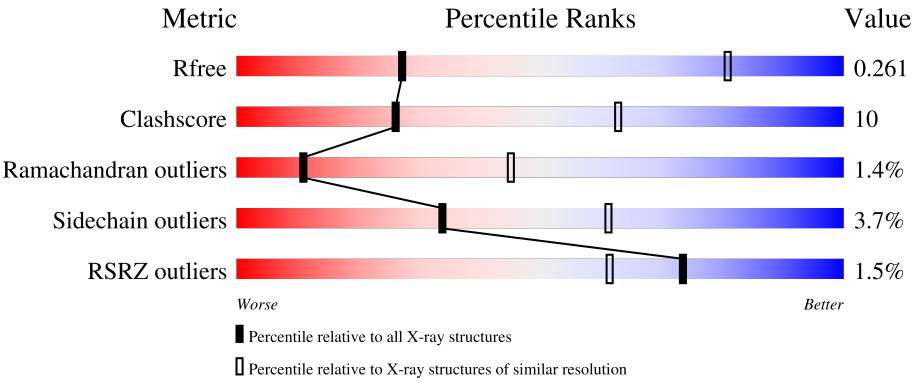
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.65 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	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	464	57% 20% • 21%
1	B	464	61% 17% 21%
1	C	464	% 60% 18% • 21%
1	D	464	2% 61% 17% • 21%
1	E	464	% 62% 17% 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	464	4% 62% 16% 21%
2	G	167	61% 20% 6% 13%
2	H	167	62% 18% 6% 14%
2	I	167	57% 28% 13%
2	J	167	55% 24% 8% 12%
2	K	167	60% 22% 5% 13%
2	L	167	66% 17% 14%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 23157 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 Phosphoglucosamine mutase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	368	Total	C	N	O	S	0	1	0
			2762	1731	464	553	14			
1	B	366	Total	C	N	O	S	0	1	0
			2747	1720	462	551	14			
1	C	368	Total	C	N	O	S	0	1	0
			2762	1731	464	553	14			
1	D	367	Total	C	N	O	S	0	1	0
			2758	1729	463	552	14			
1	E	367	Total	C	N	O	S	0	1	0
			2758	1729	463	552	14			
1	F	367	Total	C	N	O	S	0	1	0
			2751	1722	463	552	14			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	449	LEU	-	expression tag	UNP O34824
A	450	VAL	-	expression tag	UNP O34824
A	451	PRO	-	expression tag	UNP O34824
A	452	ARG	-	expression tag	UNP O34824
A	453	GLY	-	expression tag	UNP O34824
A	454	SER	-	expression tag	UNP O34824
A	455	SER	-	expression tag	UNP O34824
A	456	GLY	-	expression tag	UNP O34824
A	457	LEU	-	expression tag	UNP O34824
A	458	GLU	-	expression tag	UNP O34824
A	459	HIS	-	expression tag	UNP O34824
A	460	HIS	-	expression tag	UNP O34824
A	461	HIS	-	expression tag	UNP O34824
A	462	HIS	-	expression tag	UNP O34824
A	463	HIS	-	expression tag	UNP O34824
A	464	HIS	-	expression tag	UNP O34824
B	449	LEU	-	expression tag	UNP O34824

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	450	VAL	-	expression tag	UNP O34824
B	451	PRO	-	expression tag	UNP O34824
B	452	ARG	-	expression tag	UNP O34824
B	453	GLY	-	expression tag	UNP O34824
B	454	SER	-	expression tag	UNP O34824
B	455	SER	-	expression tag	UNP O34824
B	456	GLY	-	expression tag	UNP O34824
B	457	LEU	-	expression tag	UNP O34824
B	458	GLU	-	expression tag	UNP O34824
B	459	HIS	-	expression tag	UNP O34824
B	460	HIS	-	expression tag	UNP O34824
B	461	HIS	-	expression tag	UNP O34824
B	462	HIS	-	expression tag	UNP O34824
B	463	HIS	-	expression tag	UNP O34824
B	464	HIS	-	expression tag	UNP O34824
C	449	LEU	-	expression tag	UNP O34824
C	450	VAL	-	expression tag	UNP O34824
C	451	PRO	-	expression tag	UNP O34824
C	452	ARG	-	expression tag	UNP O34824
C	453	GLY	-	expression tag	UNP O34824
C	454	SER	-	expression tag	UNP O34824
C	455	SER	-	expression tag	UNP O34824
C	456	GLY	-	expression tag	UNP O34824
C	457	LEU	-	expression tag	UNP O34824
C	458	GLU	-	expression tag	UNP O34824
C	459	HIS	-	expression tag	UNP O34824
C	460	HIS	-	expression tag	UNP O34824
C	461	HIS	-	expression tag	UNP O34824
C	462	HIS	-	expression tag	UNP O34824
C	463	HIS	-	expression tag	UNP O34824
C	464	HIS	-	expression tag	UNP O34824
D	449	LEU	-	expression tag	UNP O34824
D	450	VAL	-	expression tag	UNP O34824
D	451	PRO	-	expression tag	UNP O34824
D	452	ARG	-	expression tag	UNP O34824
D	453	GLY	-	expression tag	UNP O34824
D	454	SER	-	expression tag	UNP O34824
D	455	SER	-	expression tag	UNP O34824
D	456	GLY	-	expression tag	UNP O34824
D	457	LEU	-	expression tag	UNP O34824
D	458	GLU	-	expression tag	UNP O34824
D	459	HIS	-	expression tag	UNP O34824

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	460	HIS	-	expression tag	UNP O34824
D	461	HIS	-	expression tag	UNP O34824
D	462	HIS	-	expression tag	UNP O34824
D	463	HIS	-	expression tag	UNP O34824
D	464	HIS	-	expression tag	UNP O34824
E	449	LEU	-	expression tag	UNP O34824
E	450	VAL	-	expression tag	UNP O34824
E	451	PRO	-	expression tag	UNP O34824
E	452	ARG	-	expression tag	UNP O34824
E	453	GLY	-	expression tag	UNP O34824
E	454	SER	-	expression tag	UNP O34824
E	455	SER	-	expression tag	UNP O34824
E	456	GLY	-	expression tag	UNP O34824
E	457	LEU	-	expression tag	UNP O34824
E	458	GLU	-	expression tag	UNP O34824
E	459	HIS	-	expression tag	UNP O34824
E	460	HIS	-	expression tag	UNP O34824
E	461	HIS	-	expression tag	UNP O34824
E	462	HIS	-	expression tag	UNP O34824
E	463	HIS	-	expression tag	UNP O34824
E	464	HIS	-	expression tag	UNP O34824
F	449	LEU	-	expression tag	UNP O34824
F	450	VAL	-	expression tag	UNP O34824
F	451	PRO	-	expression tag	UNP O34824
F	452	ARG	-	expression tag	UNP O34824
F	453	GLY	-	expression tag	UNP O34824
F	454	SER	-	expression tag	UNP O34824
F	455	SER	-	expression tag	UNP O34824
F	456	GLY	-	expression tag	UNP O34824
F	457	LEU	-	expression tag	UNP O34824
F	458	GLU	-	expression tag	UNP O34824
F	459	HIS	-	expression tag	UNP O34824
F	460	HIS	-	expression tag	UNP O34824
F	461	HIS	-	expression tag	UNP O34824
F	462	HIS	-	expression tag	UNP O34824
F	463	HIS	-	expression tag	UNP O34824
F	464	HIS	-	expression tag	UNP O34824

- Molecule 2 is a protein called Cyclic di-AMP synthase CdaA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	146	Total	C	N	O	S	121	0	0
			1106	692	187	222	5			

Continued on next page...

Continued from previous page...

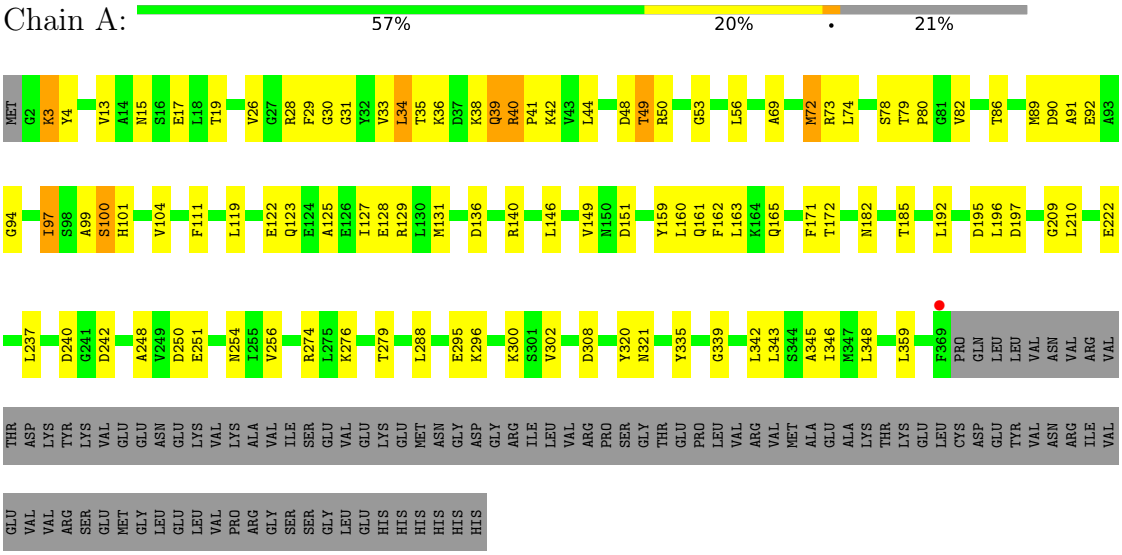
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	144	Total	C	N	O	S	103	0	0
			1092	683	185	219	5			
2	I	146	Total	C	N	O	S	121	0	0
			1106	692	187	222	5			
2	J	147	Total	C	N	O	S	103	0	0
			1117	701	188	223	5			
2	K	146	Total	C	N	O	S	121	0	0
			1106	692	187	222	5			
2	L	144	Total	C	N	O	S	103	0	0
			1092	683	185	219	5			

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

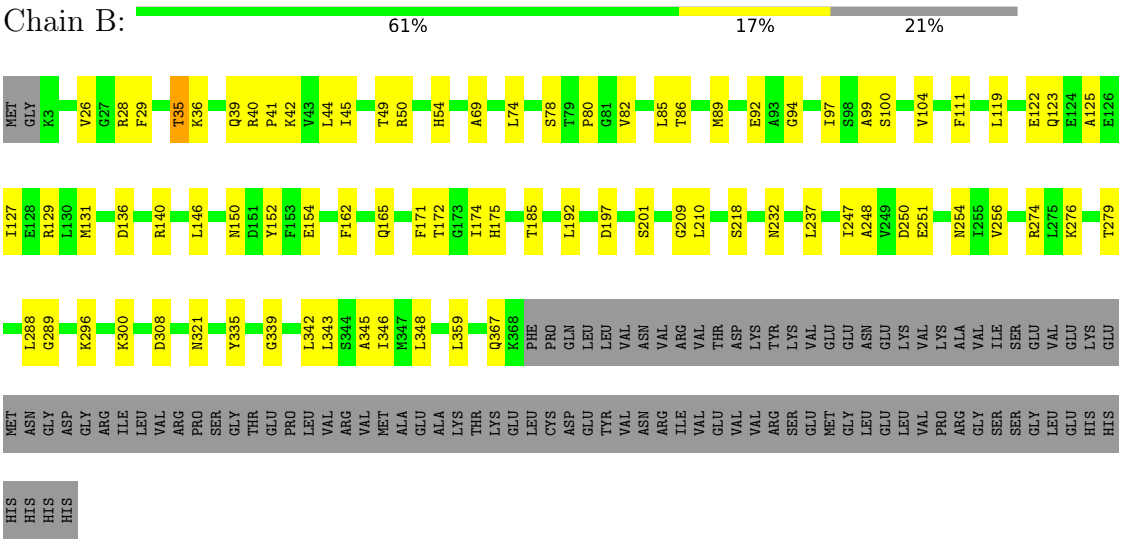
• Molecule 1: Phosphoglucosamine mutase

Chain A:



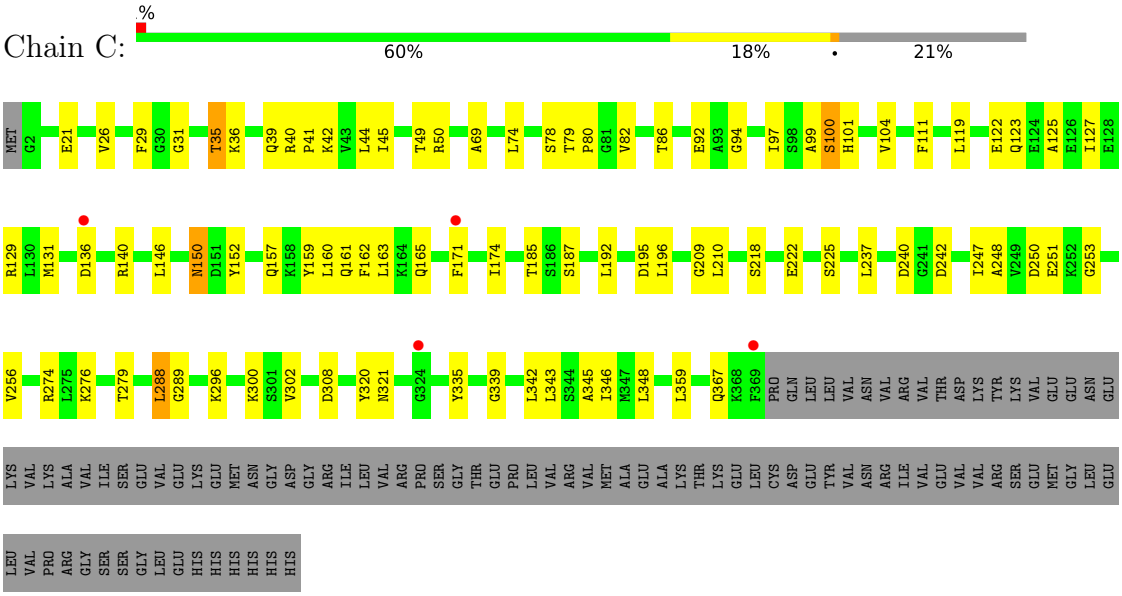
• Molecule 1: Phosphoglucosamine mutase

Chain B:



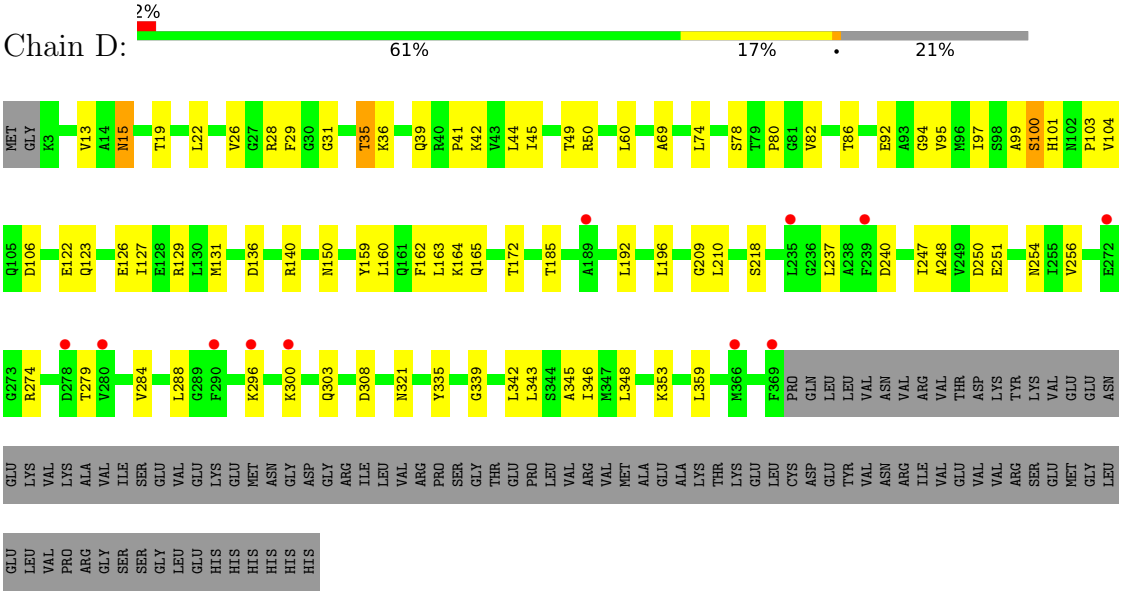
• Molecule 1: Phosphoglucosamine mutase

Chain C:



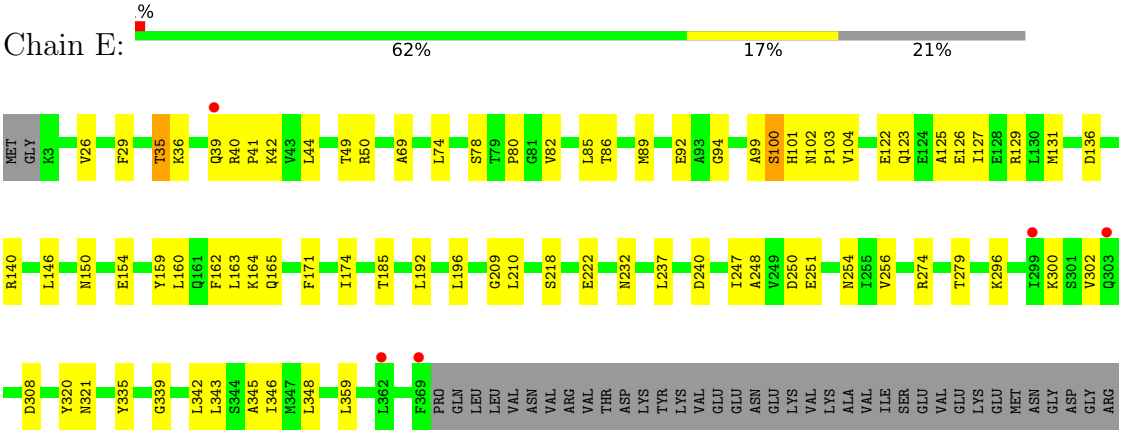
• Molecule 1: Phosphoglucosamine mutase

Chain D:



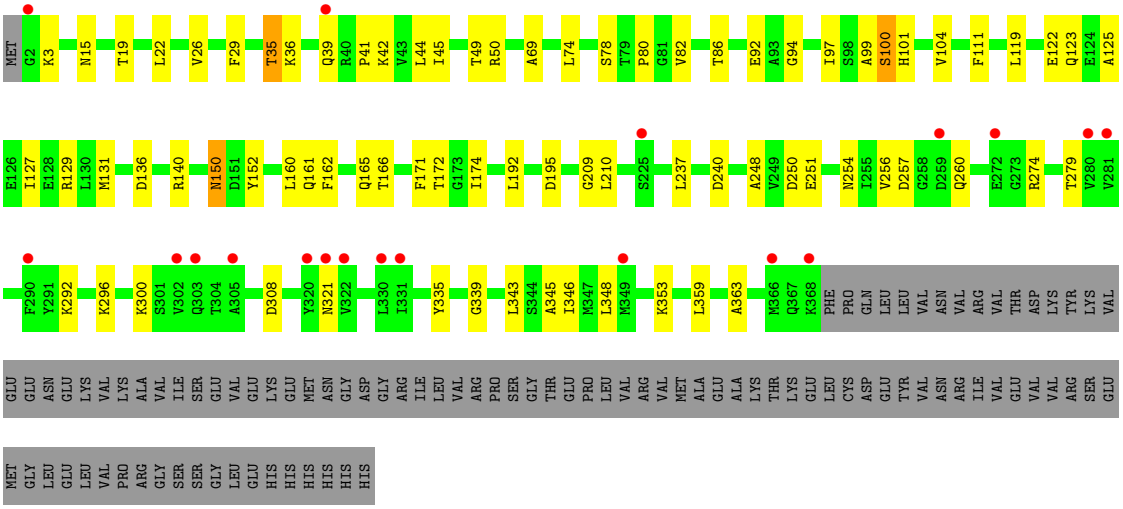
• Molecule 1: Phosphoglucosamine mutase

Chain E:



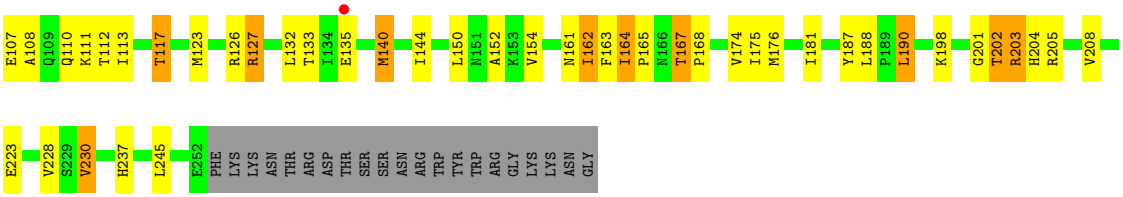
• Molecule 1: Phosphoglucosamine mutase

Chain F:



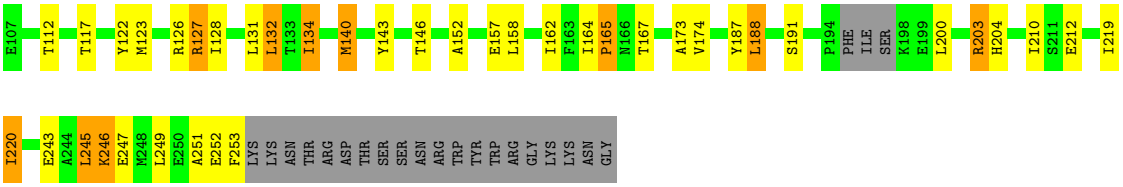
• Molecule 2: Cyclic di-AMP synthase CdaA

Chain G:



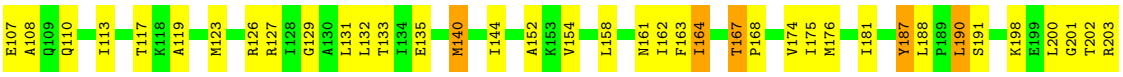
• Molecule 2: Cyclic di-AMP synthase CdaA

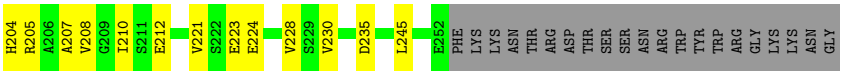
Chain H:



• Molecule 2: Cyclic di-AMP synthase CdaA

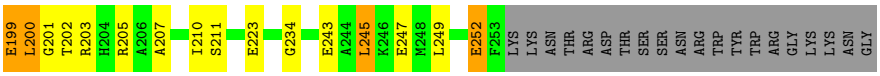
Chain I:





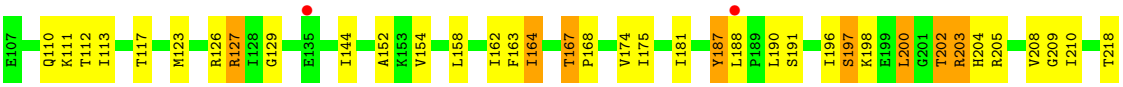
• Molecule 2: Cyclic di-AMP synthase CdaA

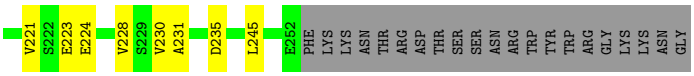
Chain J:



• Molecule 2: Cyclic di-AMP synthase CdaA

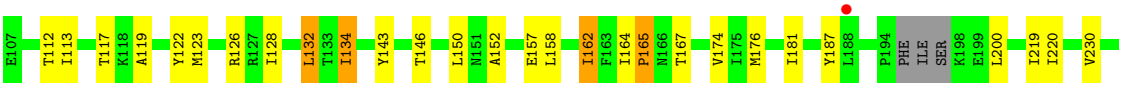
Chain K:





• Molecule 2: Cyclic di-AMP synthase CdaA

Chain L:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	62.20Å 227.56Å 151.60Å 90.00° 99.66° 90.00°	Depositor
Resolution (Å)	61.32 – 3.65 61.32 – 3.65	Depositor EDS
% Data completeness (in resolution range)	99.4 (61.32-3.65) 99.4 (61.32-3.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 3.67Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.242 , 0.262 0.242 , 0.261	Depositor DCC
R_{free} test set	2175 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å^2)	109.3	Xtriage
Anisotropy	0.294	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.043 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	23157	wwPDB-VP
Average B, all atoms (Å^2)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% 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	0.26	0/2806	0.47	0/3781
1	B	0.25	0/2790	0.45	0/3760
1	C	0.26	0/2806	0.46	0/3781
1	D	0.25	0/2802	0.45	0/3776
1	E	0.25	0/2802	0.45	0/3776
1	F	0.25	0/2794	0.45	0/3765
2	G	0.27	0/1118	0.53	0/1512
2	H	0.27	0/1103	0.52	0/1490
2	I	0.27	0/1118	0.56	1/1512 (0.1%)
2	J	0.29	0/1130	0.58	2/1528 (0.1%)
2	K	0.28	0/1118	0.54	1/1512 (0.1%)
2	L	0.26	0/1103	0.53	0/1490
All	All	0.26	0/23490	0.48	4/31683 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	187	TYR	C-N-CA	7.16	139.59	121.70
2	J	169	LEU	CA-CB-CG	5.53	128.01	115.30
2	K	187	TYR	C-N-CA	5.33	135.01	121.70
2	J	245	LEU	CA-CB-CG	5.02	126.85	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 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	2762	0	2732	68	0
1	B	2747	0	2720	53	0
1	C	2762	0	2732	57	0
1	D	2758	0	2729	51	0
1	E	2758	0	2729	48	0
1	F	2751	0	2723	51	0
2	G	1106	0	1128	32	0
2	H	1092	0	1111	31	0
2	I	1106	0	1128	32	0
2	J	1117	0	1137	27	0
2	K	1106	0	1128	29	0
2	L	1092	0	1111	24	0
All	All	23157	0	23108	468	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 10.

The worst 5 of 468 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:C:157:GLN:HE22	2:G:165:PRO:HD2	1.50	0.75
1:A:192:LEU:HD21	1:A:346:ILE:HD11	1.69	0.74
1:B:154:GLU:HG2	2:I:187:TYR:CZ	2.24	0.73
2:J:164:ILE:O	2:J:167:THR:OG1	2.07	0.73
1:E:50:ARG:NH2	1:E:99:ALA:O	2.26	0.69

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	367/464 (79%)	345 (94%)	17 (5%)	5 (1%)	11 45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	365/464 (79%)	342 (94%)	22 (6%)	1 (0%)	41	74
1	C	367/464 (79%)	345 (94%)	20 (5%)	2 (0%)	29	66
1	D	366/464 (79%)	343 (94%)	21 (6%)	2 (0%)	29	66
1	E	366/464 (79%)	343 (94%)	21 (6%)	2 (0%)	29	66
1	F	366/464 (79%)	343 (94%)	21 (6%)	2 (0%)	29	66
2	G	144/167 (86%)	132 (92%)	8 (6%)	4 (3%)	5	33
2	H	140/167 (84%)	125 (89%)	9 (6%)	6 (4%)	2	24
2	I	144/167 (86%)	132 (92%)	10 (7%)	2 (1%)	11	45
2	J	145/167 (87%)	120 (83%)	13 (9%)	12 (8%)	1	10
2	K	144/167 (86%)	132 (92%)	10 (7%)	2 (1%)	11	45
2	L	140/167 (84%)	124 (89%)	12 (9%)	4 (3%)	4	32
All	All	3054/3786 (81%)	2826 (92%)	184 (6%)	44 (1%)	11	45

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	J	199	GLU
1	A	40	ARG
1	A	100	SER
2	G	187	TYR
2	G	202	THR

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	294/380 (77%)	284 (97%)	10 (3%)	37	63
1	B	293/380 (77%)	291 (99%)	2 (1%)	84	91
1	C	294/380 (77%)	289 (98%)	5 (2%)	60	79
1	D	294/380 (77%)	290 (99%)	4 (1%)	67	82
1	E	294/380 (77%)	291 (99%)	3 (1%)	76	86

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	F	293/380 (77%)	290 (99%)	3 (1%)	76 86
2	G	120/139 (86%)	110 (92%)	10 (8%)	11 40
2	H	118/139 (85%)	109 (92%)	9 (8%)	13 43
2	I	120/139 (86%)	112 (93%)	8 (7%)	16 48
2	J	121/139 (87%)	101 (84%)	20 (16%)	2 14
2	K	120/139 (86%)	110 (92%)	10 (8%)	11 40
2	L	118/139 (85%)	111 (94%)	7 (6%)	19 51
All	All	2479/3114 (80%)	2388 (96%)	91 (4%)	34 61

5 of 91 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	126	ARG
2	J	247	GLU
2	J	132	LEU
2	J	199	GLU
2	K	164	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	157	GLN
2	L	204	HIS

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

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	368/464 (79%)	0.11	1 (0%) 94 90	59, 89, 122, 146	0
1	B	366/464 (78%)	0.10	0 100 100	71, 101, 136, 150	0
1	C	368/464 (79%)	0.20	4 (1%) 80 70	64, 104, 134, 157	0
1	D	367/464 (79%)	0.28	11 (2%) 50 36	87, 126, 158, 179	0
1	E	367/464 (79%)	0.23	5 (1%) 75 63	81, 117, 149, 168	0
1	F	367/464 (79%)	0.45	19 (5%) 27 19	93, 129, 171, 187	0
2	G	146/167 (87%)	0.19	1 (0%) 87 80	57, 85, 137, 168	27 (18%)
2	H	144/167 (86%)	0.18	0 100 100	59, 85, 113, 148	22 (15%)
2	I	146/167 (87%)	0.31	0 100 100	64, 88, 121, 135	27 (18%)
2	J	147/167 (88%)	0.13	1 (0%) 87 80	63, 88, 123, 149	22 (14%)
2	K	146/167 (87%)	0.18	2 (1%) 75 63	70, 102, 125, 144	27 (18%)
2	L	144/167 (86%)	0.22	1 (0%) 87 80	66, 95, 128, 142	22 (15%)
All	All	3076/3786 (81%)	0.22	45 (1%) 73 61	57, 105, 150, 187	147 (4%)

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	321	ASN	4.3
1	F	320	TYR	3.5
1	F	2	GLY	3.4
1	D	369	PHE	3.3
2	K	135	GLU	3.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.