Supplimentary Material

Substrate bound crystal structures reveal features unique to *Mycobacterium tuberculosis* N-acetyl-glucosamine-1-phosphate uridyltransferase and a catalytic mechanism for acetyltransfer.

Pravin Kumar Ankush Jagtap[1], Vijay Soni[2], Neha Vithani[1], Gagan Deep Jhingan[2], Vaibhav Singh Bais[1], Vinay Kumar Nandicoori[2]* and Balaji Prakash[1]*

[1] *Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur 208016, India*
[2] *National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi110067, India*

* Address correspondence to: bprakash@iitk.ac.in, tel: +91-512-2594024, fax: +91-512-2594010 & vinaykn@nii.ac.in, tel: +91-11-26703789, fax: +91-11-26742125

Running Title: Catalytic mechanism of acetyl transfer reaction in *M. tuberculosis* GlmU.

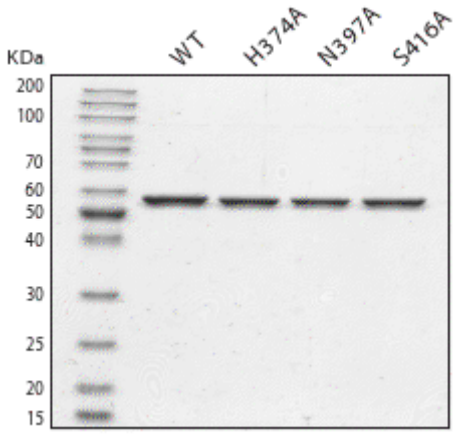


Figure S1. SDS-PAGE gel showing the purity of the mutants and wild type $GlmU^{Mtb}$ proteins employed in this study.

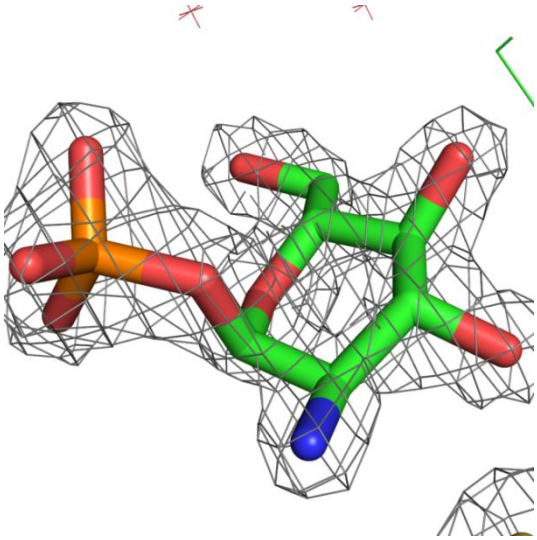
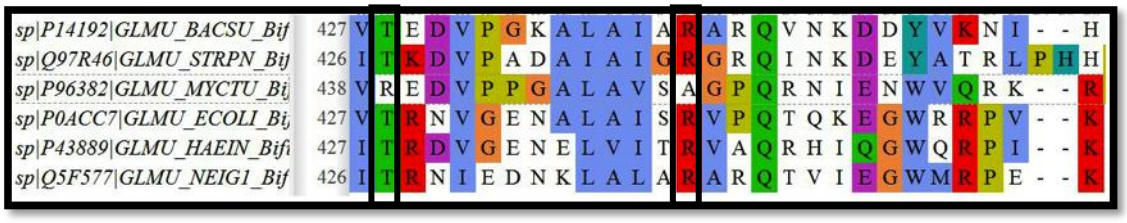
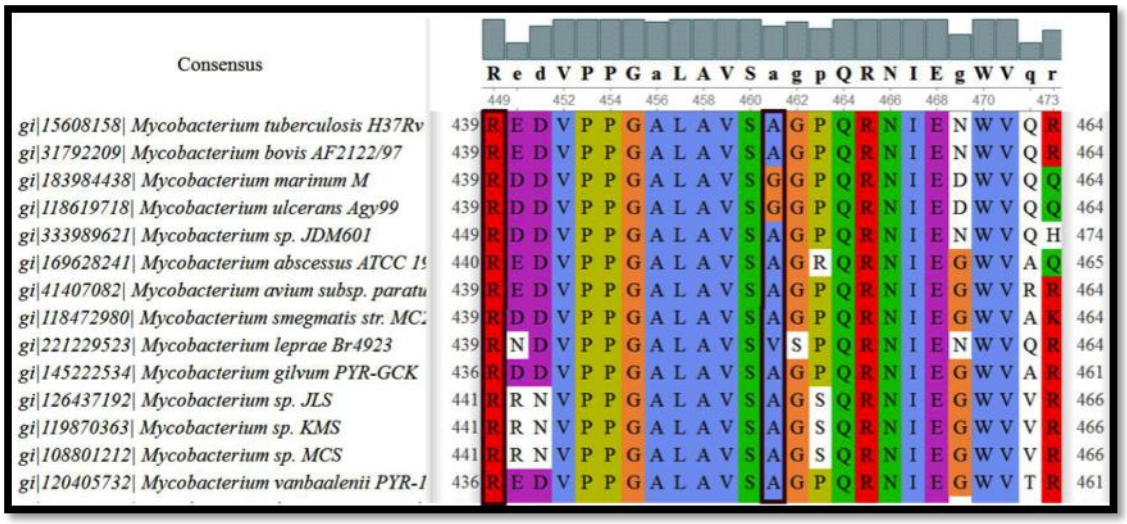


Figure S2. Glucosamine-1-phosphate (GlcN-1-P) bound at the C-terminal domain of $GlmU^{Mtb}$. Electron density for GlcN-1-P in the $2F_o$-F_c map contured at 2σ level is shown.

A



B



C

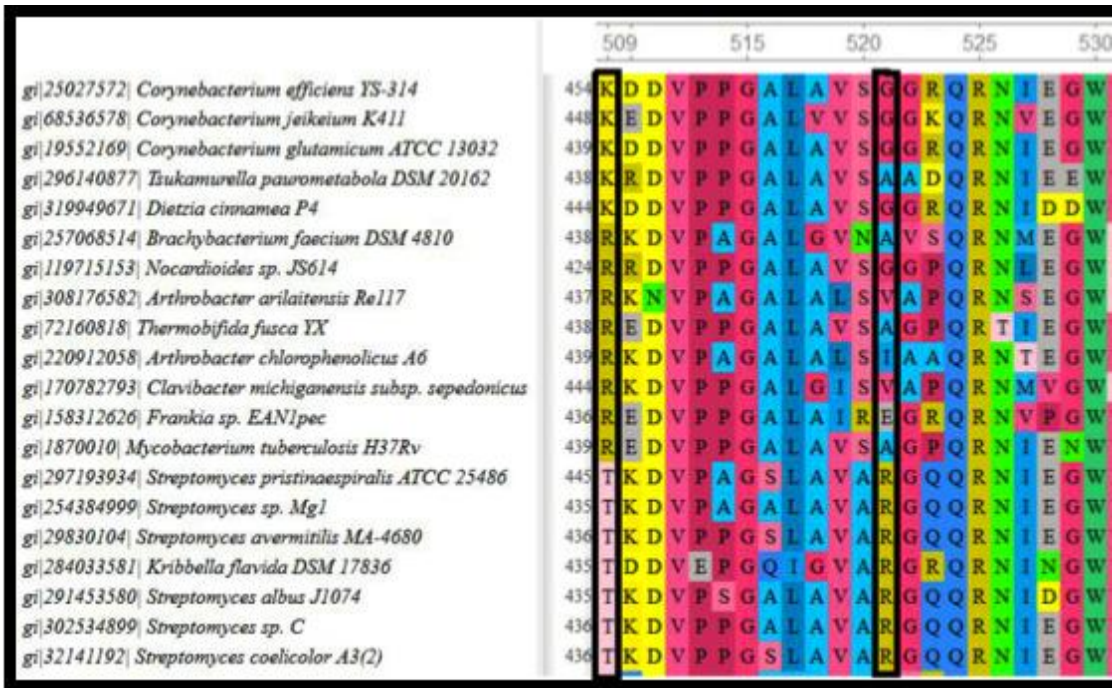


Figure S3. Sequence alignment showing the conservation of the two conformations of acetyl-CoA in the GlmU from different organisms. (A) Sequence alignment of the structures of GlmU present in Protein Data Bank reveals that the amino acids which confer U conformation to the acetyl-CoA bound to the GlmU^{Mtb} is present only in *M. tuberculosis*. (B) Sequence alignment of GlmU from

the genus *Mycobacterium* depicts that the amino acids conferring the U conformation are highly conserved in this species. (C) Analysis of GlmU in the phylum *Actinobacteria*, which *Mycobacterium* genus belongs to, reveals two groups; one having residues, responsible for the U conformation and another having those responsible for the L conformation of the bound acetyl-CoA.

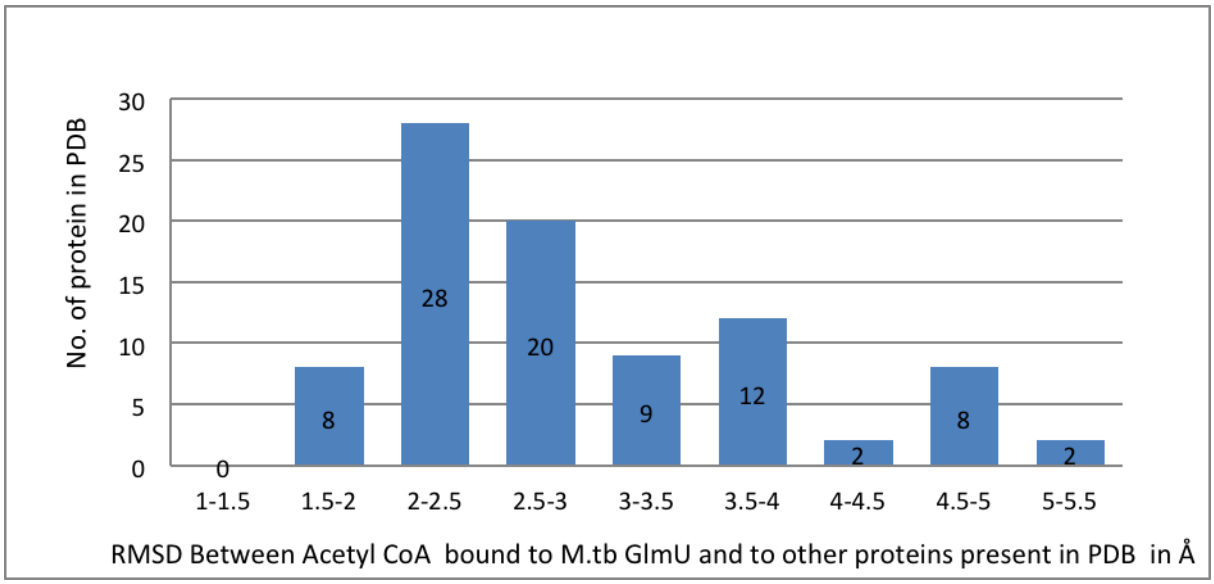


Figure S4. Comparison of the conformation of the bound acetyl-CoA in GlmU^{Mtb} and the proteins present in PDB. RMSD between the acetyl-CoA moieties bound to GlmU^{Mtb} and those bound to other proteins present in the PDB is shown. This data is presented in the histogram above. Numbers within the bars represent the number of proteins in PDB, wherein the conformation of acetyl-CoA superposes with the U conformation seen in GlmU^{Mtb} with an RMSD in the range mentioned below the bars. Only 8 proteins have RMSD less than or equal to 2Å. Visual inspection further confirmed that these adopt an almost U-like conformation.

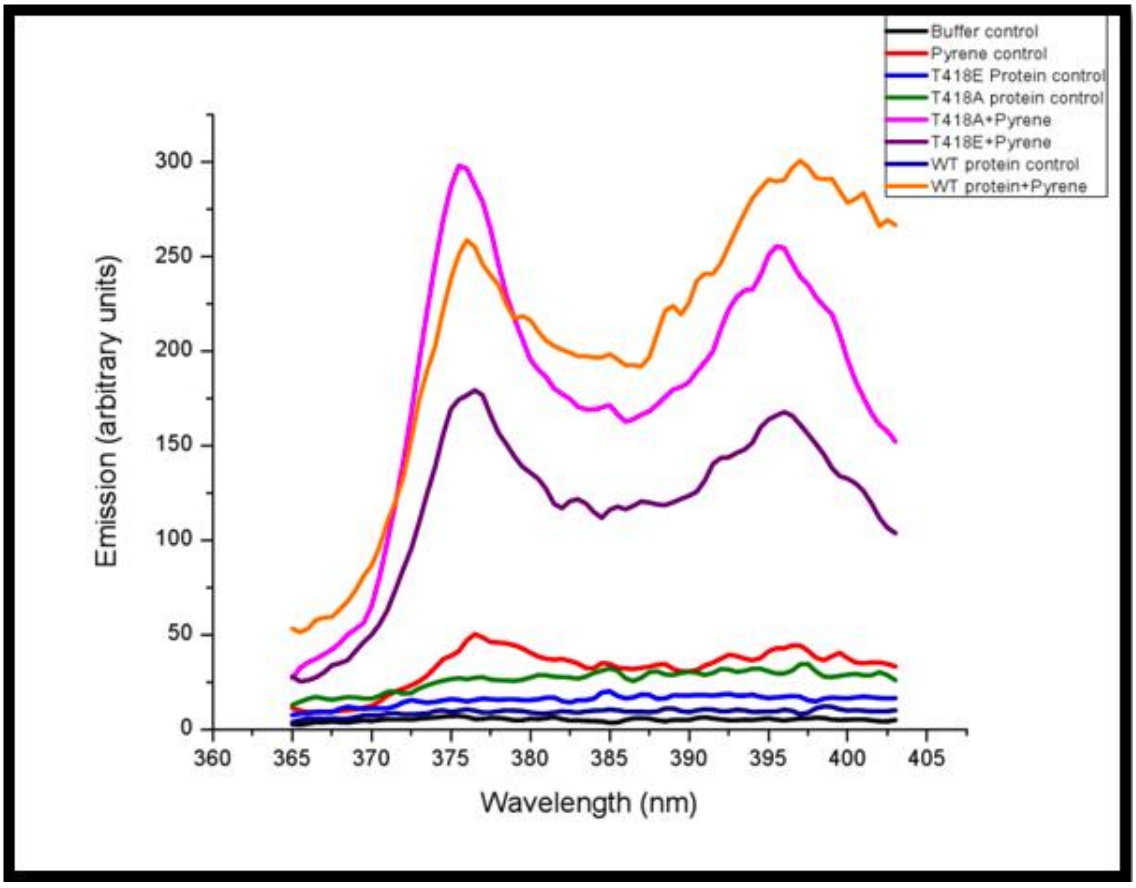


Fig S5. Phosphorylation of T418 does not affect acetyl-CoA binding by GlmU^{Mtb}. The florescence binding assay was carried out using 0.15 µM Pyrene-CoA (fluorescent analogue of CoA), 10 µM protein, 10 mM $MgCl_2$, 20 mM HEPES (pH- 7.6), 150 mM NaCl and 1 mM DTT. The pyrene group has an enhanced fluorescence in a hydrophobic environment. Both the phosphomimetic (T418E) and phosphorylation negative (T418A) mutants do not show substantial difference in pyrene-CoA binding. Buffer control - only buffer; Pyrene control - buffer and Pyrene-CoA; Protein control - buffer and protein.

Table S1: Primers used in the study.

Name	Sequence
MtGlmU F MtGlmU R	5'CACCGAATTCCATATGACGTTTCCTGGTGACACC 3' 5'GGAAGCTTATCTAGATGGTGTCTGATCAGCGTCGGG 3'
MtGlmU:1-465 F MtGlmU:1-465 R	5'CACCGAATTCCATATGACGTTTCCTGGTGACACC 3' 5'TCTAGAAGCTTACTCGATGTTGCGTTGCGGACC 3'
MtGlmU:1-458 F MtGlmU:1-458 R	5'CACCGAATTCCATATGACGTTTCCTGGTGACACC 3' 5' TCTAGAAGCTTACTCGATGTTGCGTTGCGGACC 3'
MtGlmU:R429T F MtGlmU:R429T R	5' CCGGCACAGTGGTGACCGAGGATGTCCCGCC 3' 5'GGCGGGACATCCTCGGTCACCACTGTGCCGG 3'
MtGlmU:A451R F MtGlmU:A451R R	5'CGCTGGCAGTGTCGCGGGGTCCGCAACGCAAC3' 5'GTTGCGTTGCGGACCCCGCGACACTGCCAGCG 3'
MtGlmU:W460A F MtGlmU:W460A R	5'CGCAACATCGAGAACGCGGTGCAGCGC 3' 5'GCGCTGCACCGCGTTCTCGATGTTGCG 3'
MtGlmU: K464A F MtGlmU: K464A R	5'GGTGCAGCGCGCACGCCCCGGCAGCCCAGC 3' 5'GCTGGGCTGCCGGGGCGTGCGCGCTGCACC 3'
MtGlmU:S416A F MtGlmU:S416A R	5'CACGTACGGACCGGGGCCGACACCATGTTCGTG3' 5'CACGAACATGGTGTCGGCCCCGGTCCGTACGTG3'
MtGlmU:N397A F MtGlmU:N397A R	5'TCC AGCGTGTTCGTCGCCTACGACGGTACGTCCAAA3' 5'TTTGGACGTACCGTCGTAGGCGACGAACACGCTGGA3'
MtGlmU:H374A F MtGlmU:H374A R	5'GGCACCAAGGTGCCG**GCC**CTGACCTACGTCGGCGAC3' 5'GTCGCCGACGTAGGTCAGGGCCGGCACCTTGGTGCC3'
MtGlmU:T418A F MtGlmU: T418A R	5'ACCGGGTCCGACGCTATGTTCGTGGCC3' 5'GGCCACGAACATAGCGTCGGACCCGGT 3'
MtGlmU: T418E F MtGlmU: T418E R	5'ACCGGGTCCGACGAAATGTTCGTGGCC 3' 5'GGCCACGAACATTTCGTCGGACCCGGT 3'
MtGlmU: T418S F MtGlmU: T418S R	5'ACCGGGTCCGACTCCATGTTCGTGGCC 3' 5'GGCCACGAACATGGAGTCGGACCCGGT 3'
MtGlmU:R455T+I457K F MtGlmU:R455T+I457K R	5' CCGCAAACCAACAAGGAGAACTGG 3' 5' CCAGTTCTCCTTGTTGGTTTGCGG 3'

Table S2: The conformation of acetyl-CoA bound to $GlmU^{Mtb}$ was compared with every conformation of acetyl-CoA bound to various proteins (PDB IDs of which are given in the PDB ID column). The output statistics, i.e. PDB ID in which the ligand is found, the resolution, RMSD of the superposed ligand with respect to the input ligand, RMSD bond length, RMSD bond angle, RMSD Imp (indicating Improper angle) and RMSD dihedral angle are provided below. The analysis was performed using ValiGurl server (http://eds.bmc.uu.se/eds/valligurl.php).

PDB ID	Resolution	RMSD from input structure (Å)	RMSD Bond(Å)	RMSD Angle(Å)	RMSD Imp(Å)	RMSD Dih(Å)
Ideal		**3.84**	**0.218**	**9.66**	**27.53**	**16.22**
1KHR	2.8Å	1.57	0.226	8.74	47.76	23.3
1KQA	3.2Å	1.57	0.398	11.14	47.74	31.17
1KRU	2.8Å	1.58	0.385	10.73	48.06	38.07
1KRV	2.8Å	1.64	0.417	11.53	48.05	40
2D3M	1.6Å	1.71	0.342	11.01	38.07	39.55
2D52	1.6Å	1.72	0.336	10.87	38.07	36.45
1N71	1.8Å	1.93	0.391	12.73	38.32	31.82
1IXE	2.3Å	1.98	0.406	12.26	5.06	32.2
1N8W	2.7Å	2.02	0.363	13.6	48.23	34.36
2A4N	2.2Å	2.04	0.39	12.78	38.29	30.08
1VPM	1.66Å	2.05	0.143	6.35	47.15	34.28
1BQ6	1.56Å	2.14	0.21	9.62	51.61	39.52
1I1D	1.8Å	2.19	0.254	9.15	47.45	24.57
1TIQ	1.9Å	2.21	0.171	8.92	29.05	44.8
1YLI	1.95Å	2.22	0.131	6.46	47.13	36.06
1DQ8	2.1Å	2.24	0.145	7.66	15.36	38.54
1DQA	2 Å	2.24	0.173	7.67	15.13	35.23
1D6H	2.15Å	2.25	0.394	11.84	39.28	45.74
1PU9	2.3Å	2.32	0.201	8.84	47.38	32.04
1S5K	2.4Å	2.32	0.219	8.38	30.45	37.71
2CNT	2.4Å	2.32	0.207	7.88	26.24	38.03
2CYE	1.9Å	2.32	0.311	10.66	35.35	43.77
1H16	1.53Å	2.33	0.245	9.14	37.99	10.81
1H17	1.75Å	2.33	0.238	9.16	38.08	10.85
1PUA	2.3Å	2.35	0.196	8.69	47.23	32.21
1QSN	2.2Å	2.35	0.194	8.7	47.37	30.44
1R31	2.1Å	2.36	0.255	9.93	28.15	31.35
1S3Z	2 Å	2.37	0.193	7.72	30.66	36.59
1CM0	2.3Å	2.38	0.234	9.77	38.17	37.06
1PG4	1.75Å	2.39	0.368	11.72	54.95	45.07
1PG3	2.3Å	2.4	0.349	11.46	56.17	44.54
1NL7	1.903Å	2.43	0.122	6.27	4.79	27.93
1S60	3 Å	2.44	0.284	9.59	33.05	54.03
2EIS	2.1Å	2.44	0.18	7.25	38.63	29.53

2BUE	1.7Å	2.49	0.166	6.82	25.63	28.5
2CNM	2.6Å	2.5	0.201	8.36	37.65	28.19
2C43	1.93Å	2.53	0.299	10.15	38.68	35.33
1HV9	2.1Å	2.57	0.328	9.56	47.06	36.94
1P0H	1.6Å	2.59	0.297	9.07	26.1	37.12
1WL4	1.55Å	2.61	0.424	12.78	28.27	22.85
1H9G	2.1Å	2.63	0.138	8.02	32.02	28.72
1ESM	2.5Å	2.64	0.154	8.87	54.2	28.11
1Q2C	2.25Å	2.66	1.361	29.7	51.14	32.85
1WTF	1.6Å	2.68	0.375	15.57	38.94	37.38
2AHW	2.15Å	2.69	0.135	6.86	5.1	26.55
1Q2D	2.25Å	2.7	0.191	8.62	47.5	31.36
2EFT	2 Å	2.71	0.162	7.11	54.63	42.95
1DLV	2.29Å	2.73	0.287	11.62	28.95	23.62
1QFL	1.92Å	2.73	0.287	11.65	28.97	24.55
1YRE	2.15Å	2.74	0.223	8.68	38.26	27.45
2AHV	2 Å	2.75	0.155	6.81	5.66	27.96
2CTS	2 Å	2.77	0.298	9.53	12.65	56.87
1HND	1.6Å	2.84	0.351	10.15	5.08	57.89
1F7L	1.5Å	2.88	0.346	10.94	38.88	42.58
1JKJ	2.35Å	2.94	0.158	8.05	27.81	26.61
1SCU	2.5Å	2.99	0.171	12.59	28.76	31.96
1MZJ	2.1Å	3.02	0.291	10.09	52.11	39.17
1M4I	1.5Å	3.03	0.346	10.04	25.94	36.4
1Q4S	1.95Å	3.07	0.327	14.14	38.94	31.06
1H1T	1.78Å	3.09	0.185	7.98	28.6	38.07
1M4D	1.8Å	3.09	0.337	10.14	26.12	36.31
2AF3	2.6Å	3.16	0.122	6.43	47.1	36.24
2AF4	2.147Å	3.26	0.367	12.08	26.7	30.45
1CQI	3.3Å	3.3	0.187	8.63	28.7	38.6
1YVK	3.01Å	3.46	0.194	9.07	29	38.62
1YQZ	1.54Å	3.57	0.386	12.04	38.82	36.2
2E6U	1.8Å	3.57	0.204	8.34	49.86	32.05
1XVU	2.4Å	3.62	0.339	11.17	6.83	37.75
1Y81	1.701Å	3.62	0.306	9.07	54.93	39.96
1XVT	2.3Å	3.63	0.391	12.56	4.65	30.41
1Q6Y	1.99Å	3.66	0.16	7.06	38.75	27.05
1XA4	1.9Å	3.68	0.368	12.11	54.39	38.81
1P5R	2.5Å	3.71	0.357	12.31	54.89	30.67
1VGR	2.1Å	3.81	0.367	11.87	47.33	38.61
1QR0	1.9Å	3.82	0.72	27.67	45.72	50.58
1EAB	2.6Å	3.85	0.37	13.85	26.45	27.42
1T4C	2.61Å	3.85	0.35	11.8	47.25	37.86

2B4B	2 Å	4.07	0.241	15.67	51.21	32.38
1YSL	1.9Å	4.46	0.233	7.95	28.45	36.12
2B58	1.95Å	4.7	0.226	11.87	51.04	34.99
1T7Q	1.8Å	4.73	0.352	11.04	48.43	40.51
2B4D	2 Å	4.74	0.141	7.65	50.22	25.07
1S7L	2.3Å	4.75	0.363	12.78	31.69	43.78
1S7N	2.1Å	4.76	0.382	12.66	26.53	39.29
1MJA	2.26Å	4.81	0.189	8.64	47	22.59
1NDI	2.3Å	4.85	0.373	11.51	54.74	41.28
1MJ9	2.5Å	4.86	0.191	8.71	47.14	24.75
1BO4	2.3Å	5.32	0.204	7.64	28.2	29.74
1EBL	1.8Å	5.37	0.267	9.57	39.29	57.55