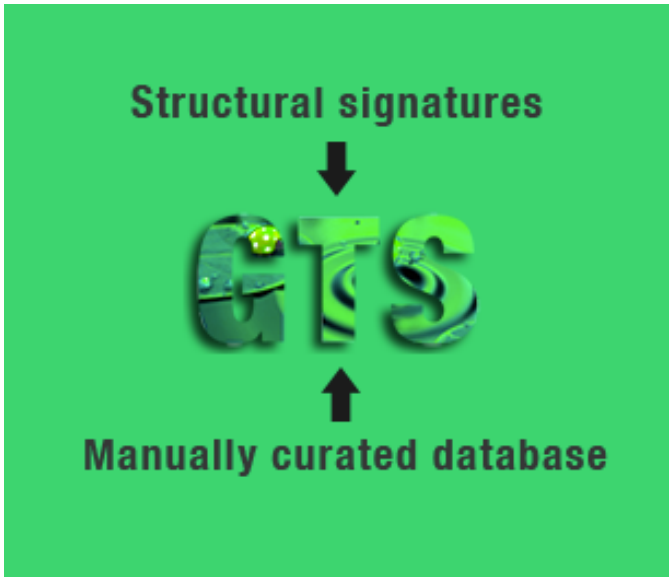
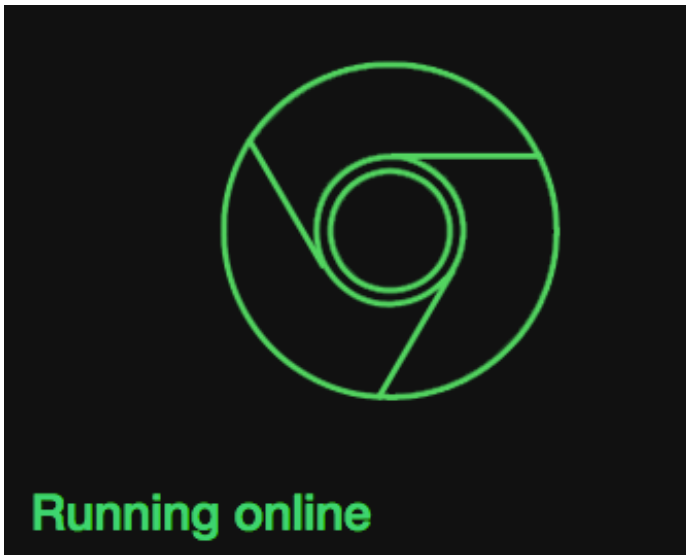
GTS docs





Glucose tolerance signature

GTS is a method to calculate if a mutation could be beneficial or not for an enzyme beta-glucosidase.

Summary

Summary 2

GTS documentation 3

What is GTS? 3

Second-Generation Biofuel problem 3

How GTS method Works? 5

Inputs 5

GTS steps 6

Running online 7

Accessing the "Run online" panel 7

Running an example (H228T) 8

Run by scripts 13

Requirements 13

Scripts 13

GTS documentation

What is GTS?

GTS (Glucose Tolerance Signature) is a method to propose mutations for β-glucosidase enzymes used in Second-Generation Biofuel production. GTS uses structural signatures to detect patterns, which can improve the activity of non-tolerant β-glucosidases, based on a manually curated database of glucose-tolerant β-glucosidases.

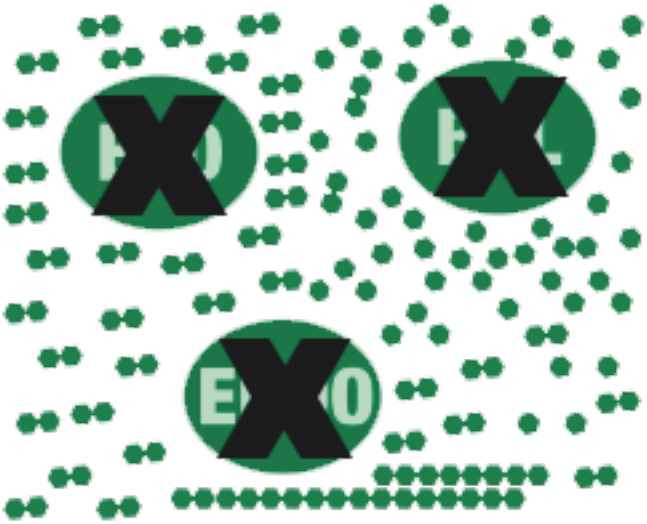
Second-Generation Biofuel problem

Second-Generation Biofuel is produced by residues obtained from the first generation biofuel production. They can be obtained from biomass, such as corn, algae, sugarcane, and so on. The production process is based on the extraction of fermentable sugars from cellulose.

Cellulose is decomposed by the action of three enzymes: endoglucanases, exoglucanases, and β-glucosidases. Endoglucanases act first, cleaving cellulose in oligosaccharides of several lengths. Then, exoglucanases cleave the oligosaccharides in disaccharides, such as cellobiose. In the end, beta-glucosidases cleaves the cellobiose in two molecules of glucose, that will be used in the fermentation process for obtain biofuel.

β-glucosidases are strongly inhibited by high glucose concentrations (same used for biofuel production).

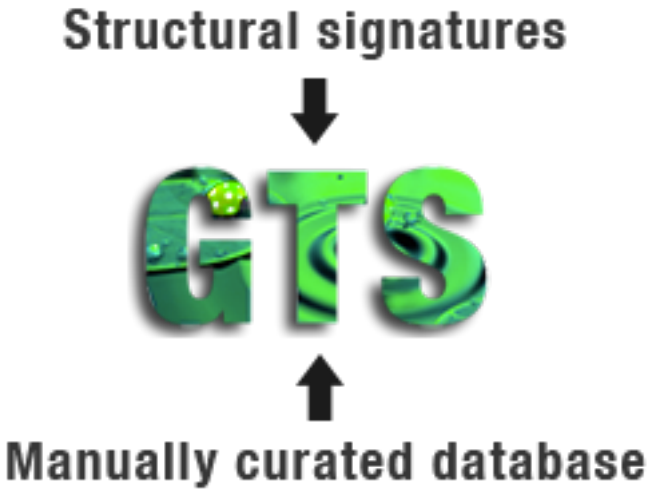
In addition, it increases cellobiose concentration that inhibits endoglucanases and exoglucanases.



Glucose-tolerant β-glucosidases: β-glucosidases of high resistance to glucose inhibition, also called glucose-tolerants, can help to improve the biofuel production. Also, some mutations can turn non-tolerant in glucose-tolerant β-glucosidases.

How GTS method Works?

GTS method constructs structural signatures for wild and mutant proteins and compares the signature's variation with a manually curated database of glucose-tolerant β-glucosidases, called **Betagdb** (available at http://bioinfo.dcc.ufmg.br/betagdb).



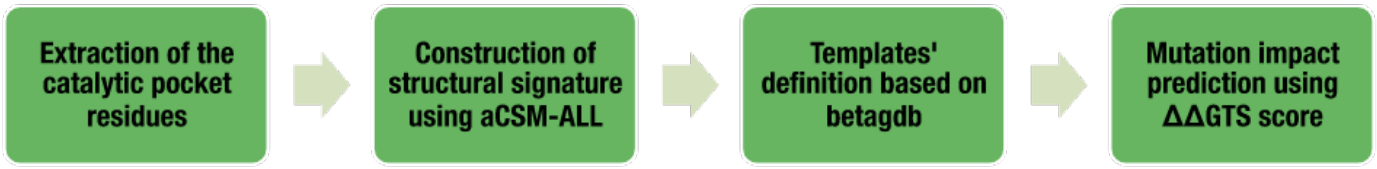
Inputs

GTS requires three beta-glucosidases:

1. Wild beta-glucosidase
2. Mutant beta-glucosidase

3. Template beta-glucosidase

GTS steps



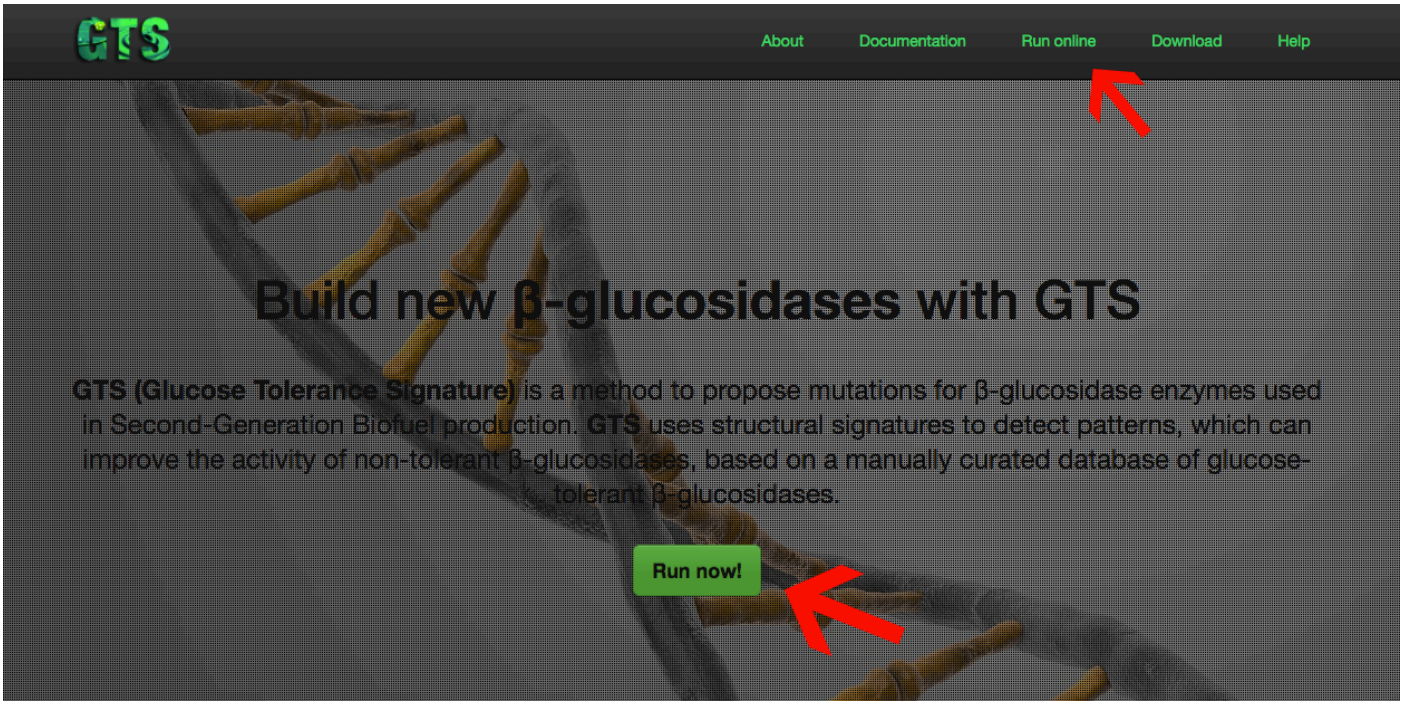
1. The first step is the extraction of the catalytic pocket residues. The mutant and wild protein are aligned with 3VIK, the only beta-glucosidase in complex with cellobiose detected in the literature. The residues of the catalytic pocket of 3VIK are known: Q45, H148, W149, N192, S193, L195, T196, D199, M207, N253, I254, N255, Y273, N335, F336, Y337, T338, L340, W374, E402, W444, E451, W452, and F460. We performed alignments using MultiProt (http://bioinfo3d.cs.tau.ac.il/MultiProt);
2. Then, we constructed structural signatures using aCSM-ALL. We used cutoff ranging of 0Å until 10Å, and step distance of 0.1Å;
3. We previously calculated the signature of 23 glucose-tolerant beta-glucosidases of betagdb. We calculated the distance among all glucose-tolerant beta-glucosidase and wild and mutant. The lowest values correspond to the templates;
4. We calculated the distance between wild and its template (ΔGTS_{Wt}), and mutant and template (ΔGTS_{Mt}). The difference between both distances is the variation of the signature variations ($\Delta\Delta GTS$).

Running online

Running online is the best way to execute GTS. It is available online at: <http://bioinfo.dcc.ufmg.br/gts>.

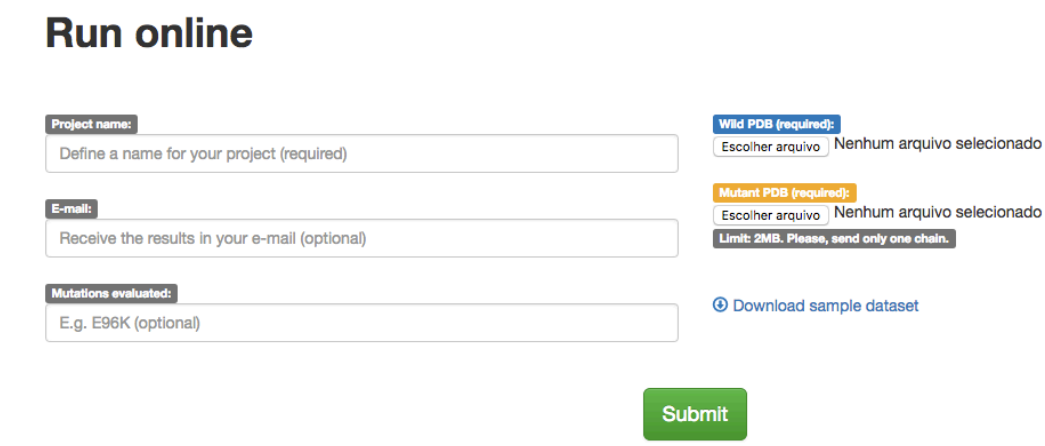
Accessing the "Run online" panel

Click in "Run online" or in the "Run now!" Button.

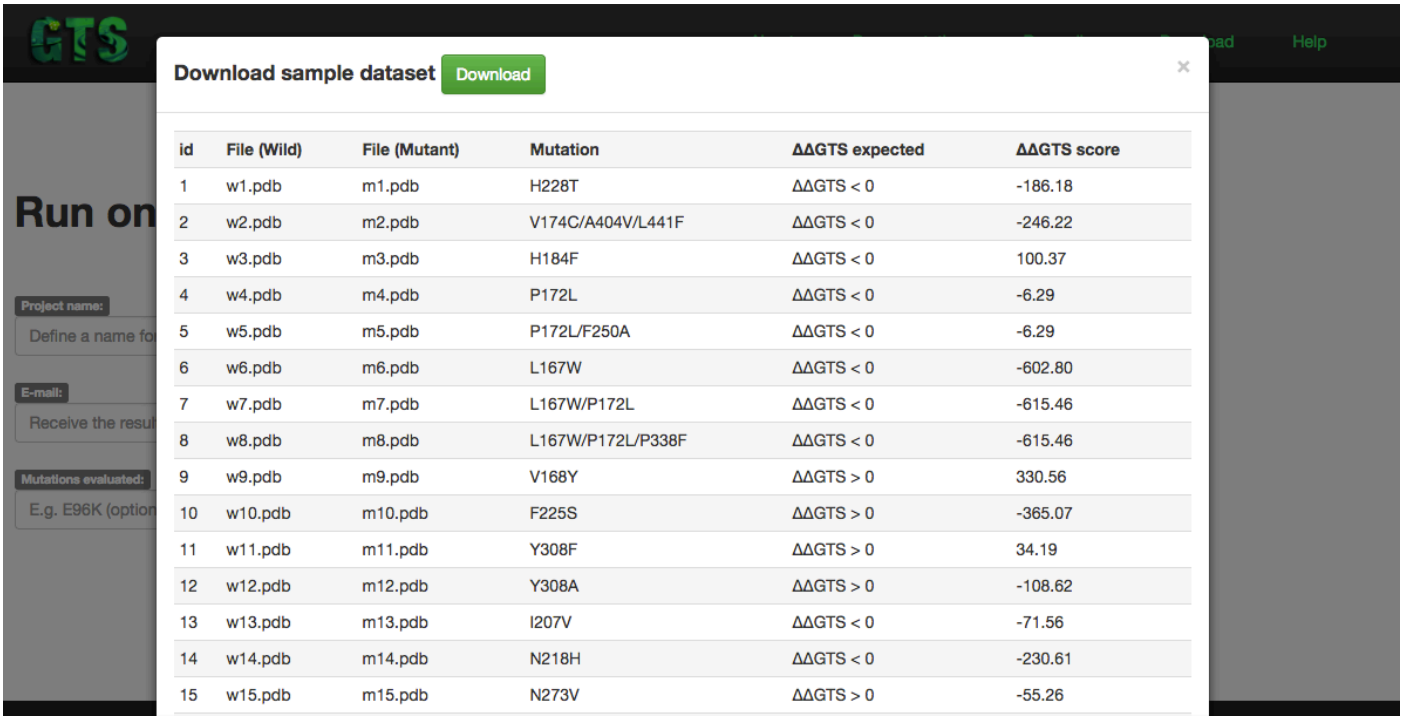


The "Run online" panel requires some information to execute:

- **Project name**: define a name for your project. You can use any name (the system will create an unique ID for each project);
- **E-mail**: declare your e-mail (optional);
- **Mutation evaluated**: insert the point mutation or multiple mutations evaluated (optional);
- **Wild PDB:** input the wild PDB file;
- **Mutant PDB:** input the mutant PDB file.



We made available a database with 27 mutations evaluated in the paper. We can download the database and perform analysis using GTS online. We also made available a table with expected values and the values found.

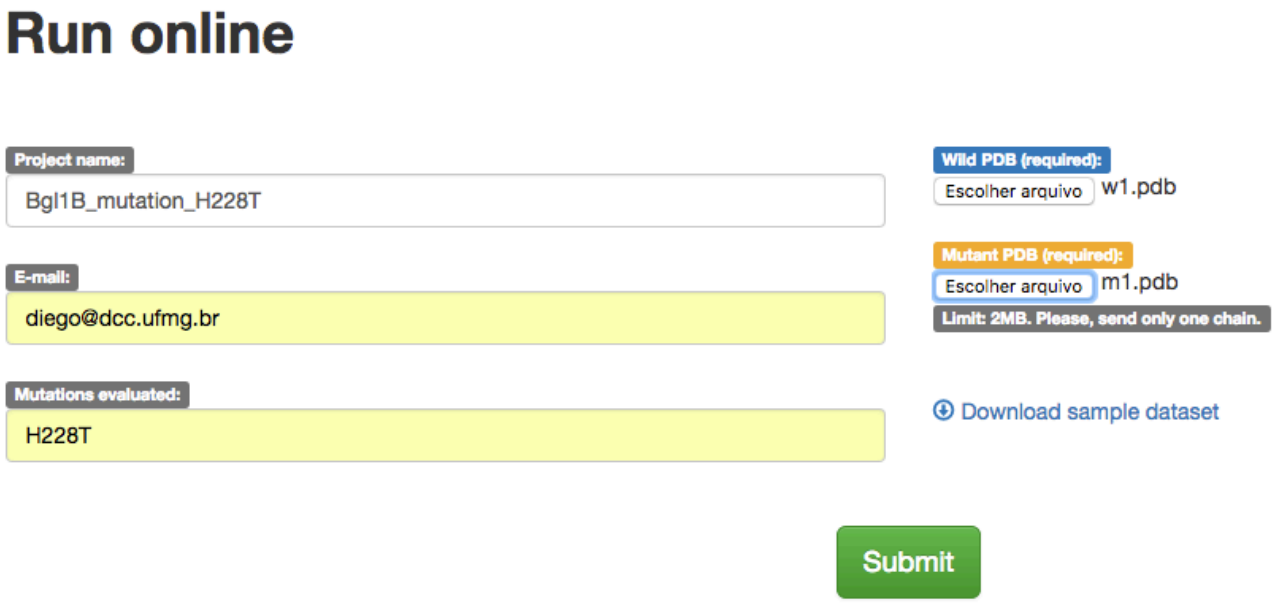


Download sample dataset Download

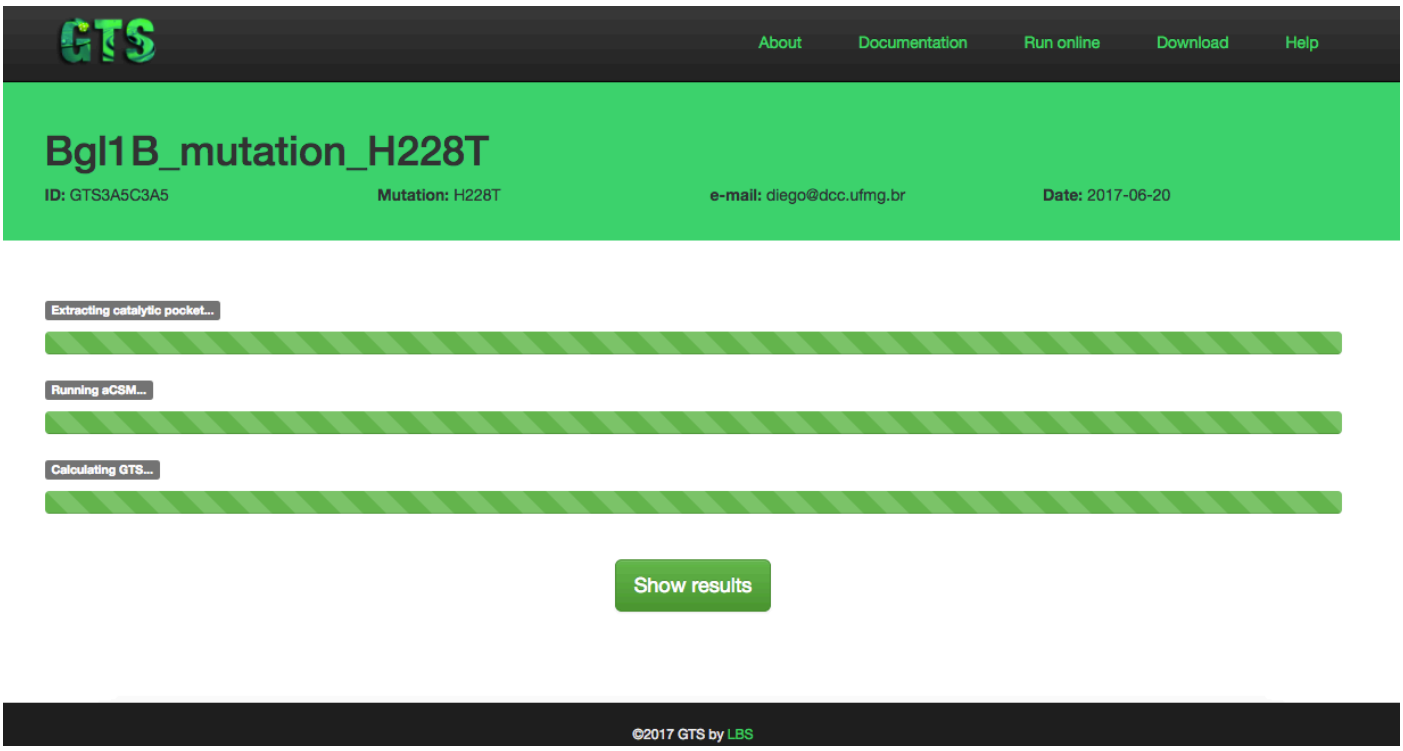
id	File (Wild)	File (Mutant)	Mutation	ΔΔGTS expected	ΔΔGTS score
1	w1.pdb	m1.pdb	H228T	ΔΔGTS < 0	-186.18
2	w2.pdb	m2.pdb	V174C/A404V/L441F	ΔΔGTS < 0	-246.22
3	w3.pdb	m3.pdb	H184F	ΔΔGTS < 0	100.37
4	w4.pdb	m4.pdb	P172L	ΔΔGTS < 0	-6.29
5	w5.pdb	m5.pdb	P172L/F250A	ΔΔGTS < 0	-6.29
6	w6.pdb	m6.pdb	L167W	ΔΔGTS < 0	-602.80
7	w7.pdb	m7.pdb	L167W/P172L	ΔΔGTS < 0	-615.46
8	w8.pdb	m8.pdb	L167W/P172L/P338F	ΔΔGTS < 0	-615.46
9	w9.pdb	m9.pdb	V168Y	ΔΔGTS > 0	330.56
10	w10.pdb	m10.pdb	F225S	ΔΔGTS > 0	-365.07
11	w11.pdb	m11.pdb	Y308F	ΔΔGTS > 0	34.19
12	w12.pdb	m12.pdb	Y308A	ΔΔGTS > 0	-108.62
13	w13.pdb	m13.pdb	I207V	ΔΔGTS < 0	-71.56
14	w14.pdb	m14.pdb	N218H	ΔΔGTS < 0	-230.61
15	w15.pdb	m15.pdb	N273V	ΔΔGTS > 0	-55.26

Running an example (H228T)

Now, we will run the first example of that database. The mutation H228T, where detected for a non-tolerant beta-glucosidase and improved its glucose tolerance. For these reason, we expected a ΔΔGTS negative.



After submit the data, GTS online will process your requisition. A unique ID will be created for your project. When the process to finish, you can click in the button "Show results".

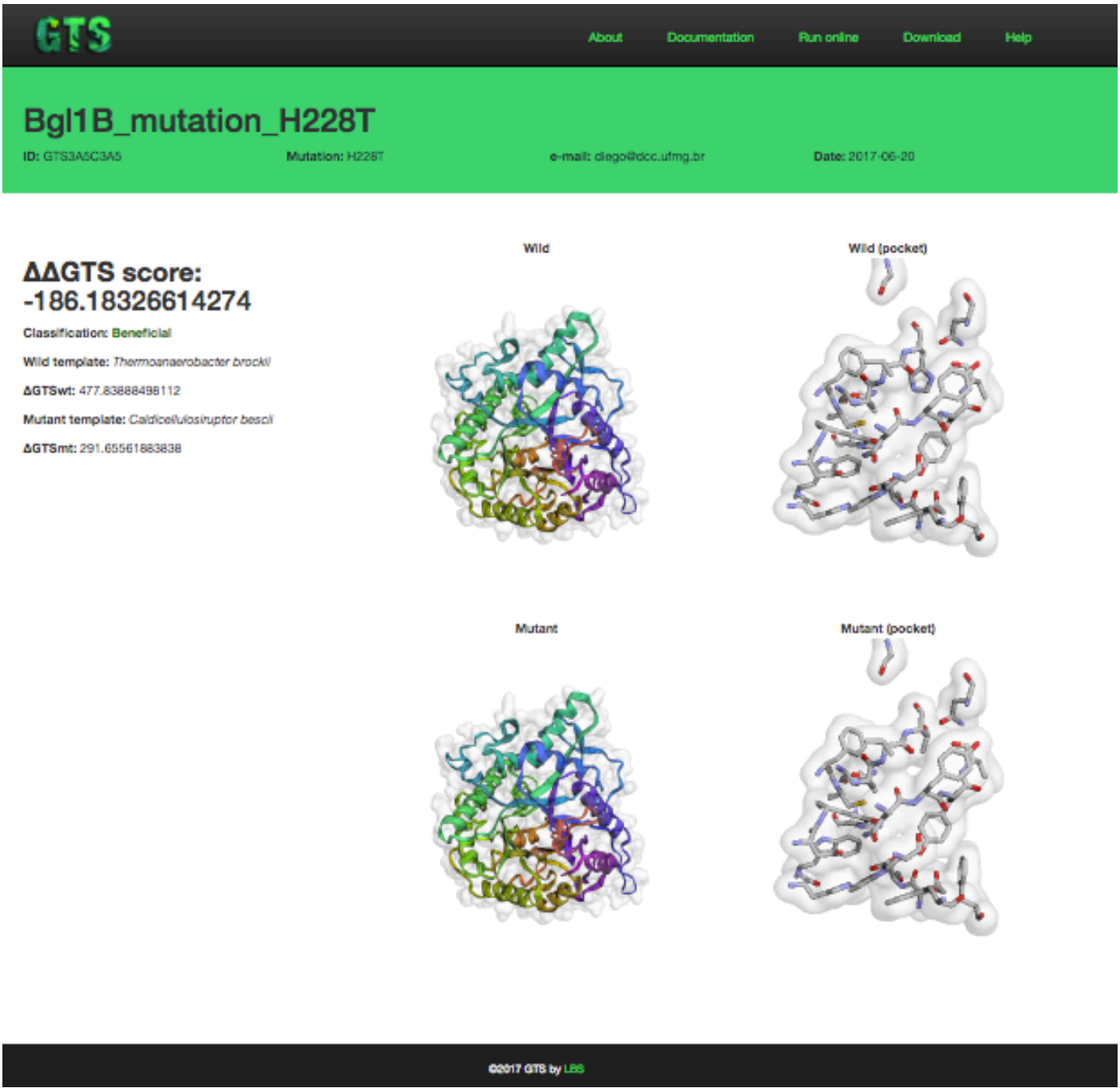


Then, you will be redirected for the individual page of the project. This page is identified by the unique ID, described below the project name (in the green sector).

Below the green sector, there are three important sectors:

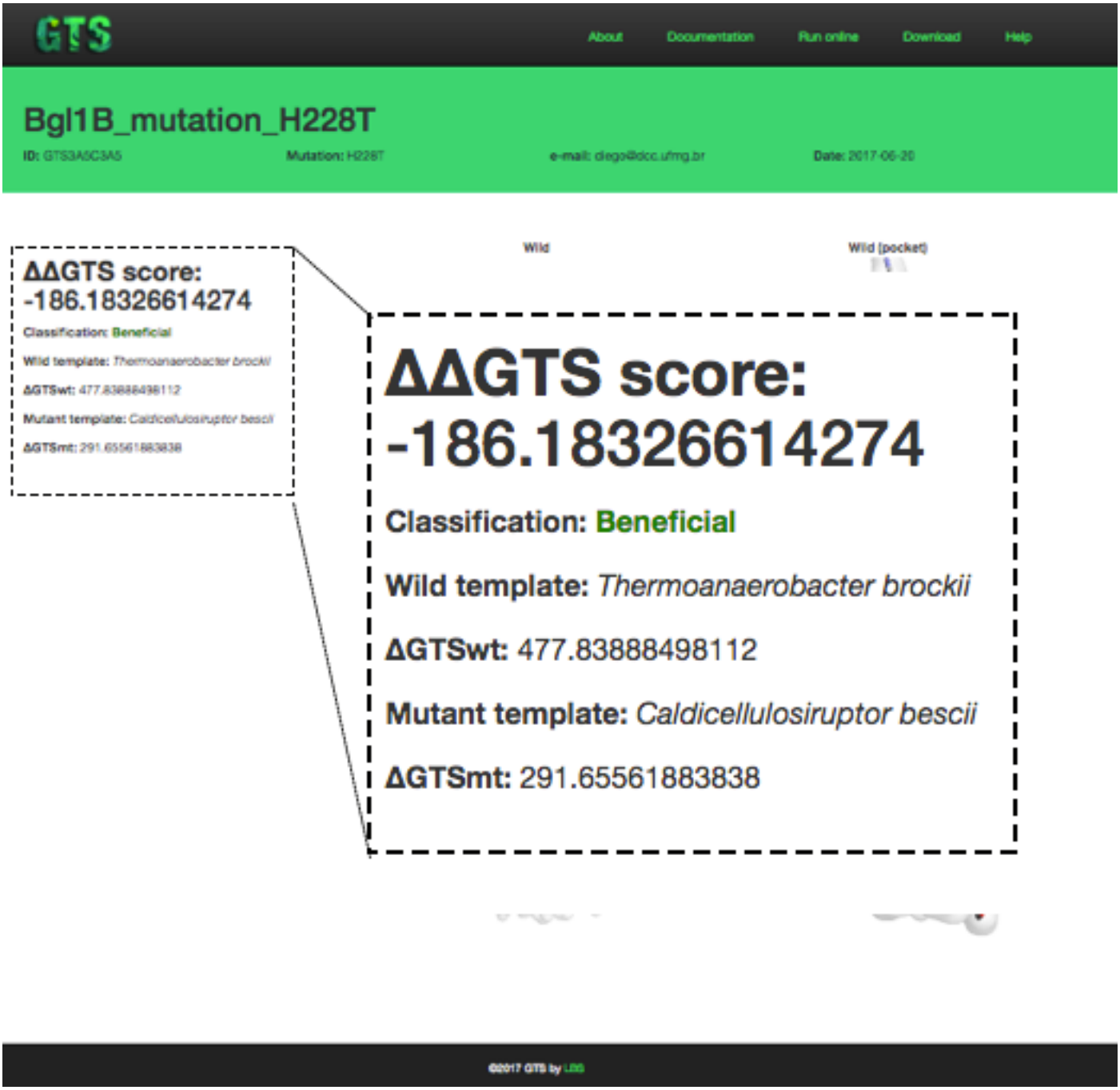
1. The results of the ΔΔGTS calculation and templates used;
2. Wild visualizations (whole protein and only pocket);

3. Mutant visualizations (whole protein and only pocket).

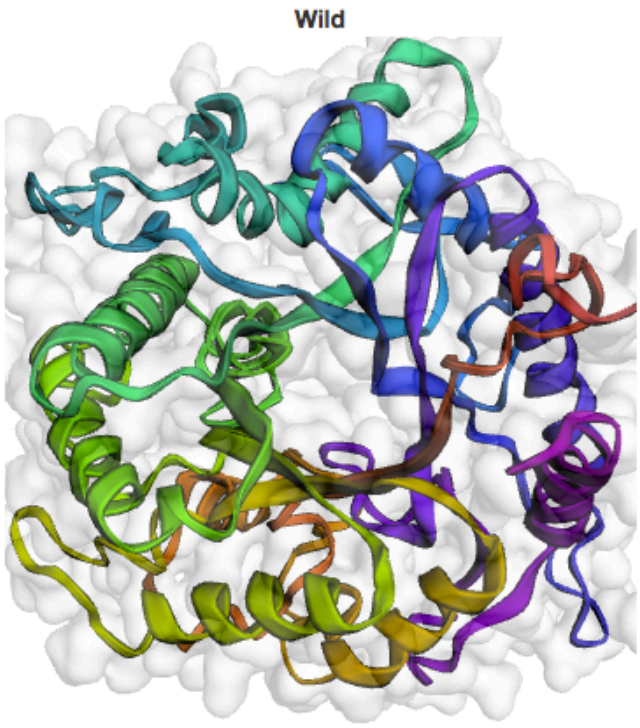


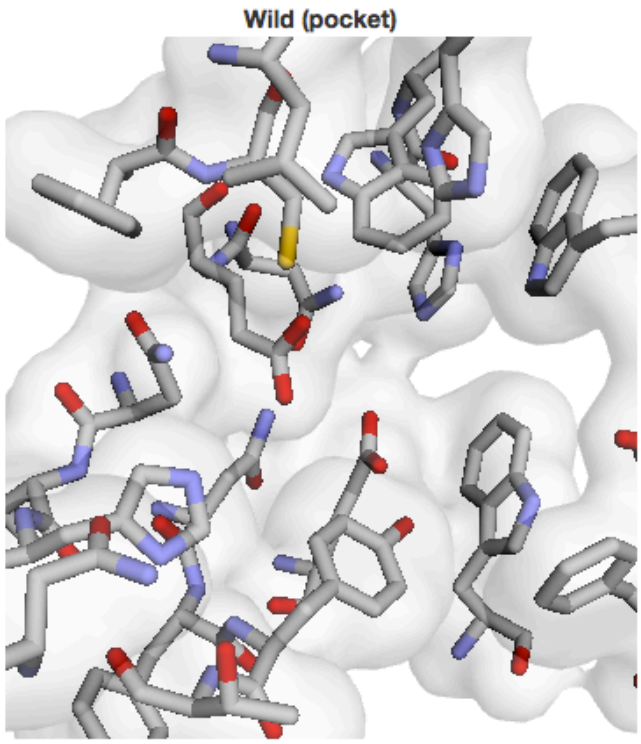
In the main panel, it is showed:

- ΔΔGTS score;
- Classification;
- Wild template;
- ΔGTS_{Wt};
- Mutant template;
- ΔGTS_{Mt}.

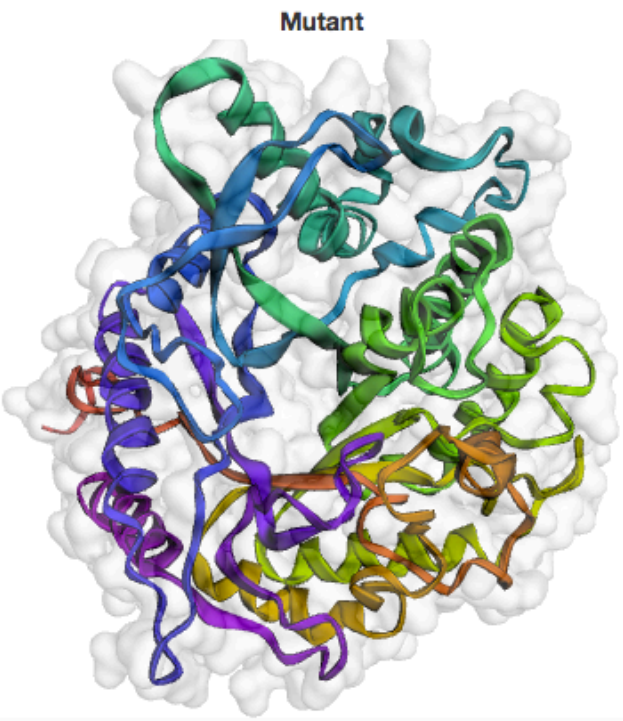


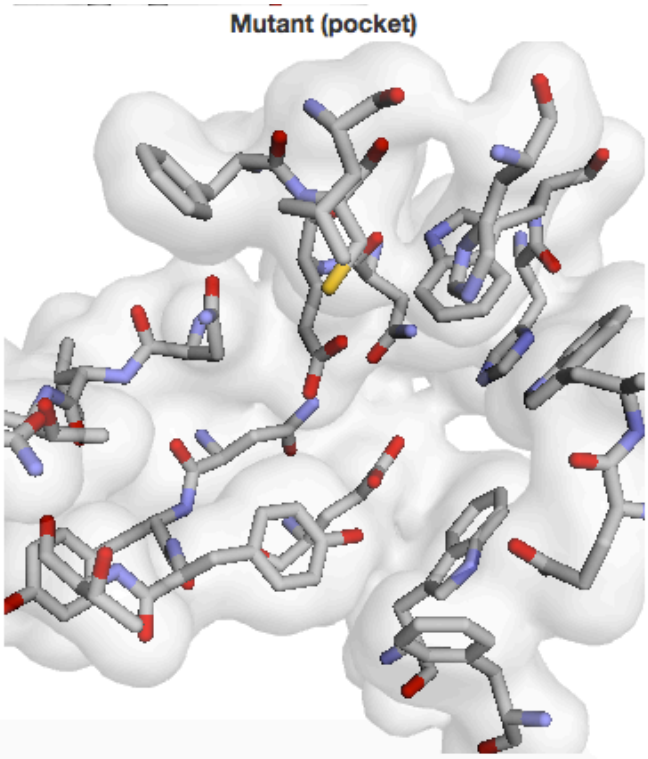
In the right, it is showed visualizations of the wild protein. We also highlighted the protein pocket and its residues.





Below, we showed a visualization of the mutant protein. We also highlighted the protein pocket and its residues.





Run by scripts

We made available the scripts to run GTS. However, we cannot have sure that you will be able to reproduce the software environment. If you have problems, run the online version.

We strongly recommend running GTS online.

Requirements

- O.S. Linux 64bit (recommended Ubuntu 16.04)
- Python (version 2)
- Library Numpy
- Perl
- aCSM
- MultiProt

Scripts

Script	Input	Output
extractSite.py	PDBs files inserted in the folder "pdb".	PDBs files with the catalytic pocket extracted.
aCSM.pl	List with address of PDBs files with the catalytic pocket extracted.	CSV file with total of atoms in the cutoff distances.
dgts3.py	CSV file obtained from aCSM. The first 23 lines must represent the beta-glucosidases from betadb in alphabetical order.	ΔΔGTS and templates for wild and mutant save in the file "result.txt".