**Parameter estimation based on protein structural information with qPIPSA**

For an introduction into the methodology and the underlying theory, the user is referred to the SYCAMORE documentation and the original scientific publications.

### 4.1. Modelling the first step of glycolysis

1st Step of Glycolysis: Glucose + ATP -> Glucose-6-phosphate + ADP

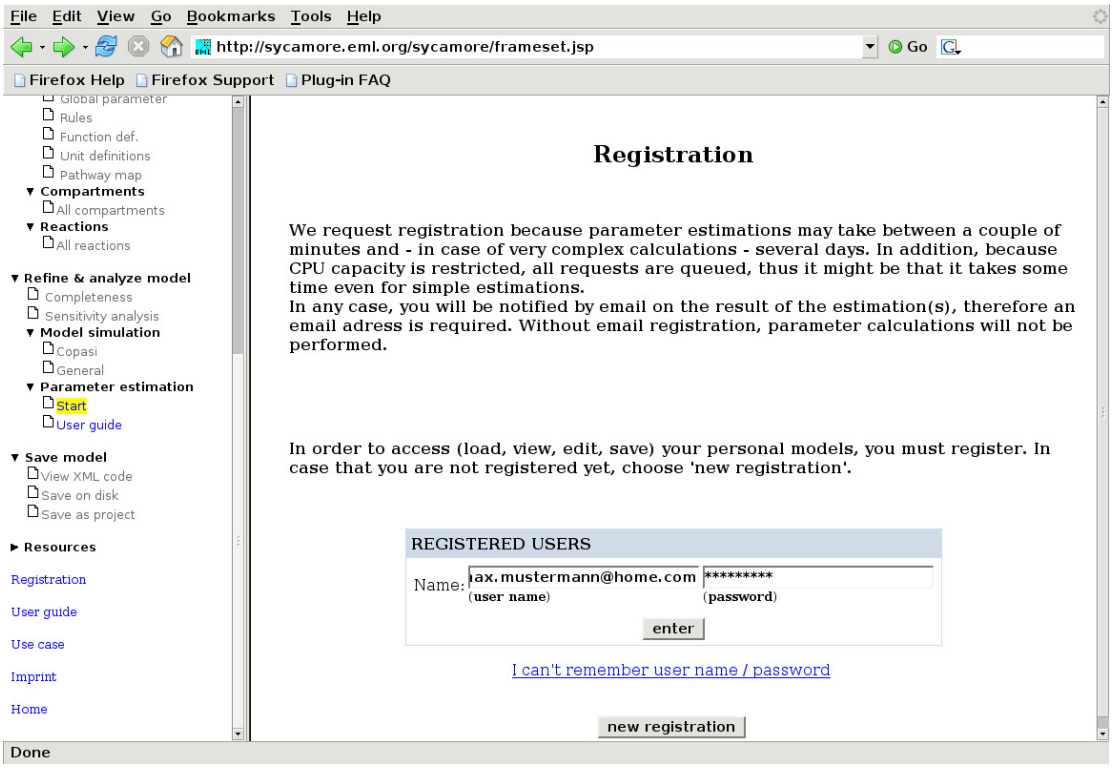
1. Modelling scenario:

    1. You are interested in modelling this reaction for the liver-specific isozyme hexokinase IV in *Homo sapiens.*

    2. You found kinetic data (for example the $K_m$ value of the substrate glucose) for the glucokinase from yeast and rat but are not sure which to use in your mathematical model.

    3. Perform an estimation of the $K_m$ parameter of the human hexokinase IV using available protein structure for that enzyme and available kinetic data for other enzymes.

    4. Register at the **Parameter Estimation** page by clicking on **New Registration** or enter your user name and password and press **Enter**




    5. In this example you are querying for the liver-specific glucokinase (hexokinase IV or D) from *Homo sapiens*.

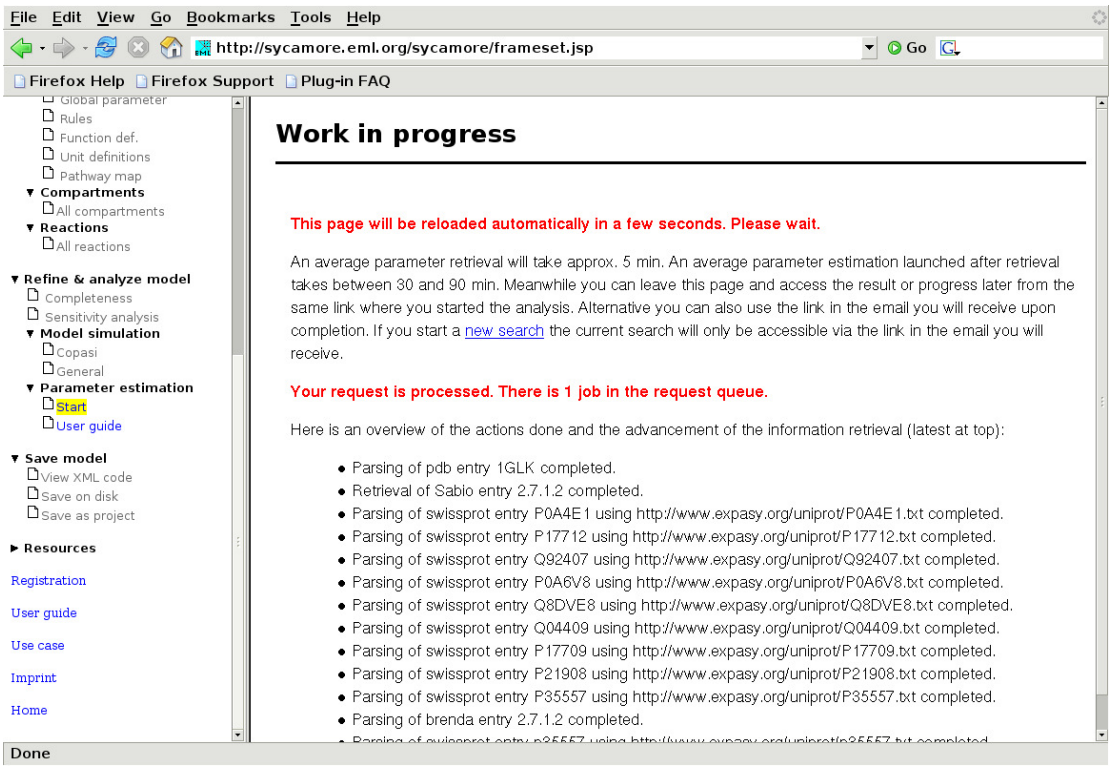
        1. Insert the SwissProt identifier of that amino acid sequence

(P35557).

2. If you do not know the SwissProt identifier of your sequence of interest, go to www.expasy.org and query for it.

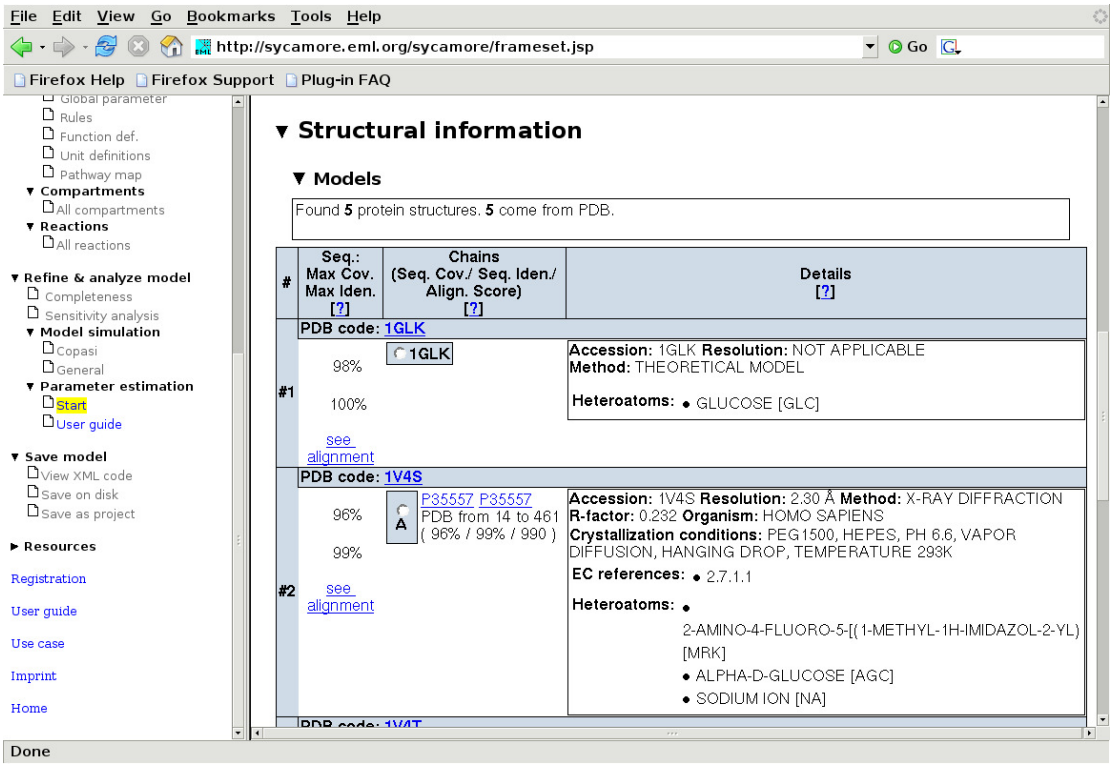
3. Click on the **Start search** button.

You will then see a screen like this:




6. The server is querying databases which contain protein structures (the PDB, Modbase and SwissModel) from X-ray analysis or theoretical models and BRENDA and SABIO-RK for existing kinetic information. Results are reported on

    1. **Structural Information**
    2. **Region selection**
    3. **Kinetic data**




5. SYCAMORE found 5 protein structures with relevance to the human hexokinase IV. 5 come from the PDB, none from ModBase or SwissModel.

   The retrieved structures are

   1. 1GLK (theoretical model of human beta-cell glucokinase)
   2. 1V4S (*H. sapiens*; closed form)
   3. 1V4T (H. sapiens; super-open form)
   4. 1Q18 (*E. coli*)
   5. 1SZ2 (*E. coli*)

6. Choose **1V4S** as a PDB template structure for subsequent structural modelling based on sequence identity with SwissProt query and completeness of X-ray structure.

7. Open the **Region Selection** submenu. Here, annotations from the SwissProt entry P35557 are reported. Select **Region Glucose-binding (Potential)** in order to also perform a comparison restricted to that region.

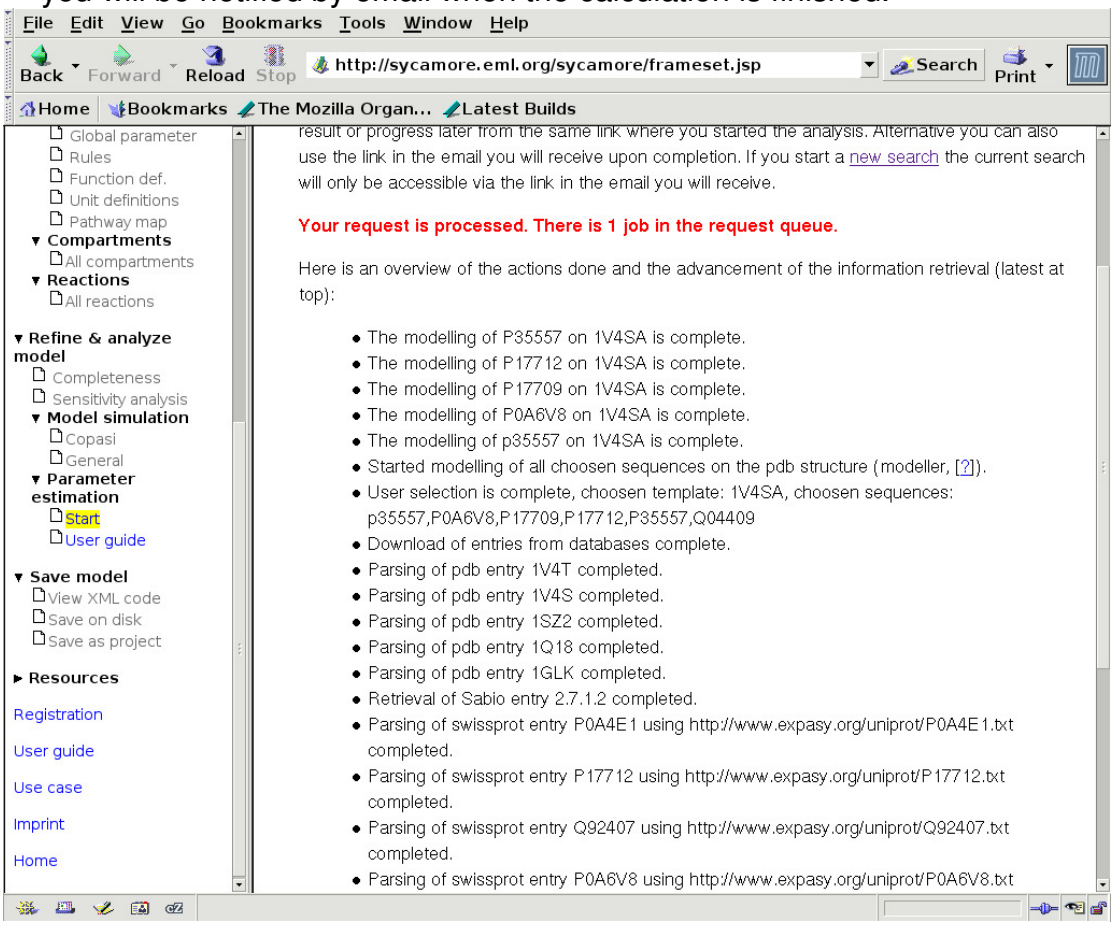
8. Scroll down further to the **Kinetic data** section.

   1. Kinetic data are retrieved from BRENDA and SABIO-RK. By default, all entries are ticked as “selected”.
   2. **Deselect** several entries from BRENDA. Leave all entries from SABIO-RK “on”:
      1. “On” should be

1. *Homo sapiens*
2. *Saccharomyces cerevisiae*
3. *E. coli*
4. *Rattus norvegicus*

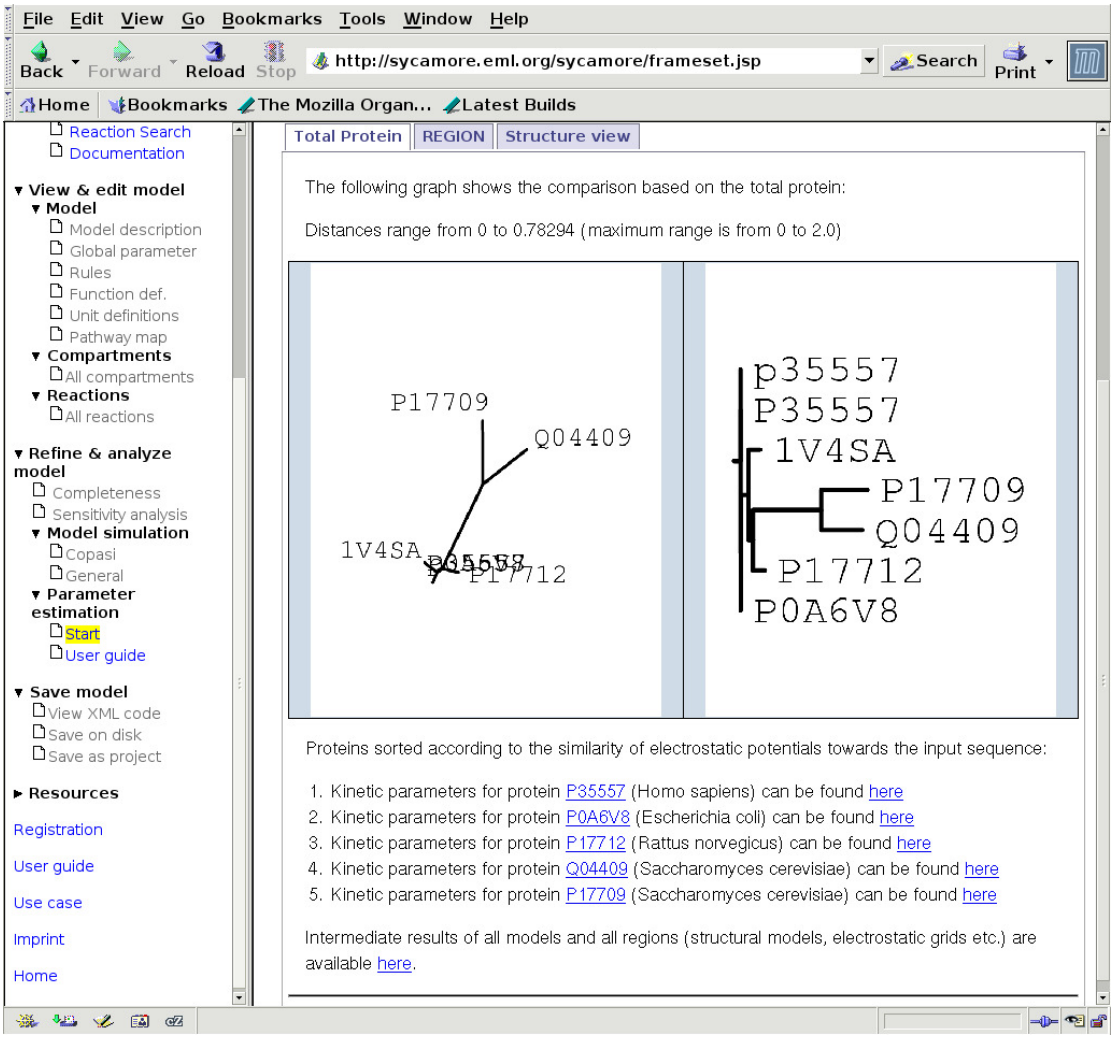
9. Go to the top of the page and click on "**Calculate electrostatically similar proteins**" to start a comparison of the selected protein structures and their properties.

10. The calculation starts. The web site will refresh automatically and you will be notified by email when the calculation is finished.

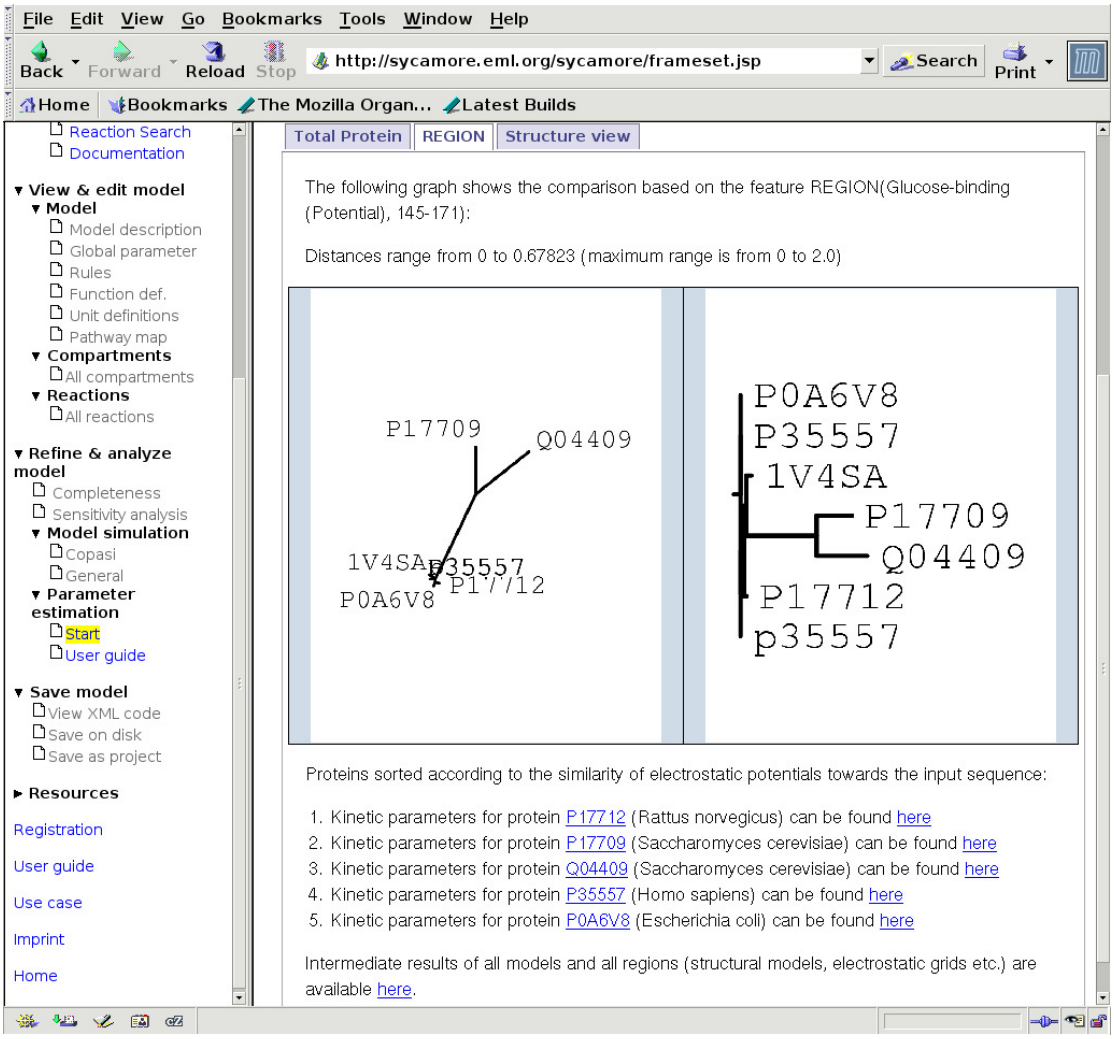



11. The output is a tree-like diagram like the following which positions your query sequence (P35557) in relation to the other chosen organisms for which kinetic data are available.

    1. The first dendogram is for the electrostatic potential of the entire protein.




2. The second dendogram is for the selected Region “glucose-binding”




The electrostatic is conserved for enzymes from human (P35557), rat (P17712), and *E. coli* (P0A6V8).
The enzymes from yeast (P17709 and Q04409) exhibit electrostatic potentials distant from the others.

This puts your query sequence P35557 very close to the hexokinase from rat (P17712), and remote from yeast (P17709, Q04409). Thus, kinetic parameters from rat could be used if you were to model the hexokinase IV from *Homo sapiens*.

This agrees with experimentally available $K_m$ values for of 7.4 and 7.7 mM for D-glucose from rat; and 6 and 8.6 mM for D-glucose from human.
The glucokinase from *E. coli* exhibits a drastically lower $K_m$ value of 0.78 mM for D-glucose and points to a different physiological role and enzymatic mechanism despite a conserved electrostatic potential near the glucose binding site.
The $K_m$ value for D-glucose from yeast is 0.028 mM is far from the ones for rat and human.

## 4.2 Modelling the second step of glycolysis

Glucose-6-phosphate isomerase (G6PI) catalyzes the reaction glucose-6-phosphate -> fructose-6-phosphate

1. You are querying for the $K_m$ value of the substrate of the glucose-6-phosphate isomerase from *Homo sapiens*.

2. Insert the SwissProt identifier for that enzyme (P06744).

3. In **Structural Information**, 30 homologuous protein structures are found in the PDB with the same EC number 5.3.1.9

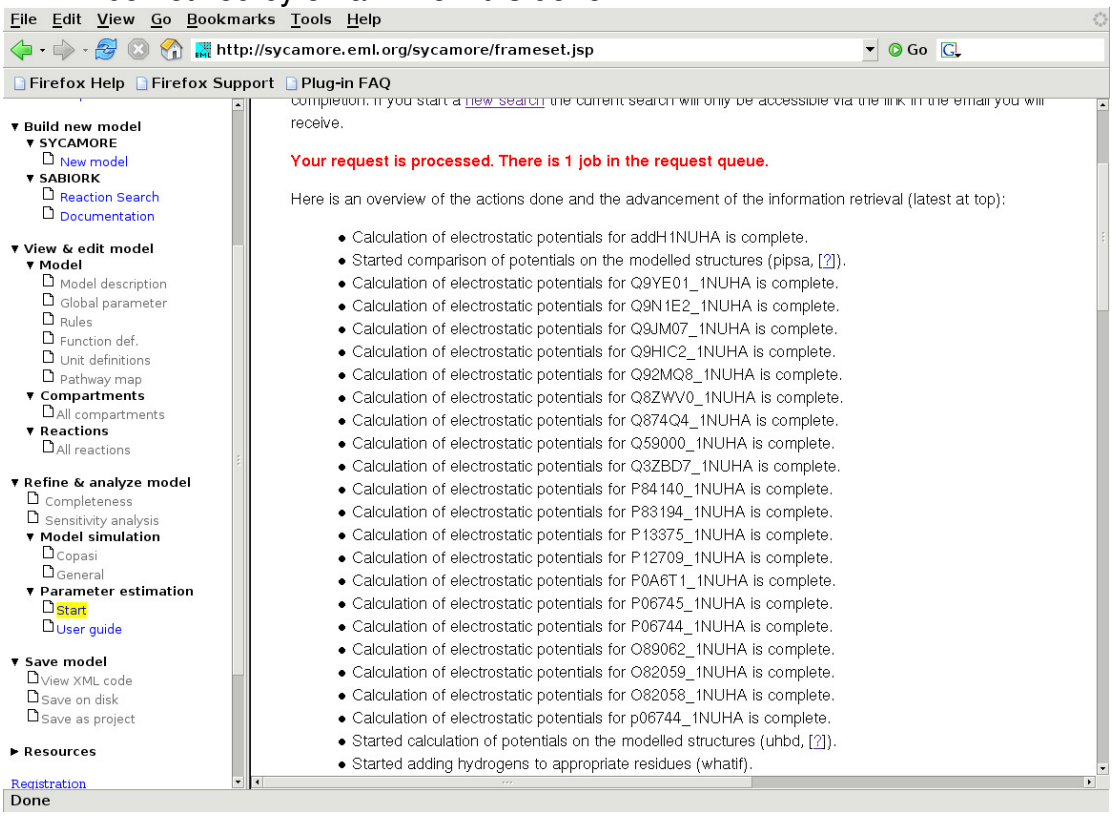
    1. Select the entry “1NUH” as a structural template based on sequence identity with query sequence (both from Homo sapiens) and co-crystallization with a transition state analogue compound.

4. In **Region Selection**, select “ACT_SITE” “By similarity” “From residue 389 to residue 389” as a centre for electrostatic potential comparison.

5. Select BRENDA and SABIO-RK entries for which $K_m$ values for the substrate glucose-6-phosphate are available.

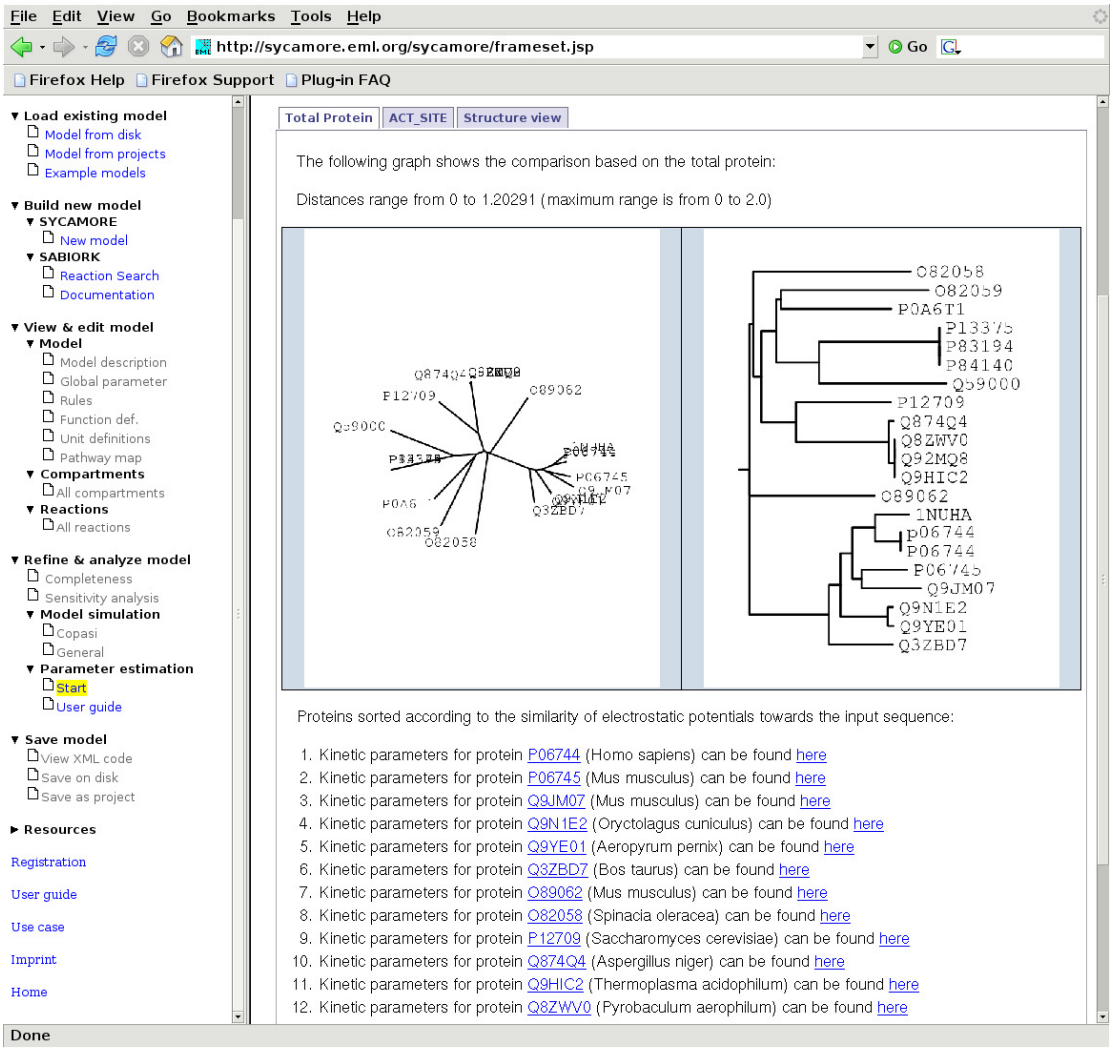
6. Click on “**Calculate electrostatically similar proteins**”

7. The calculation will start. The web site will refresh automatically and you will be notified by email when it is done.

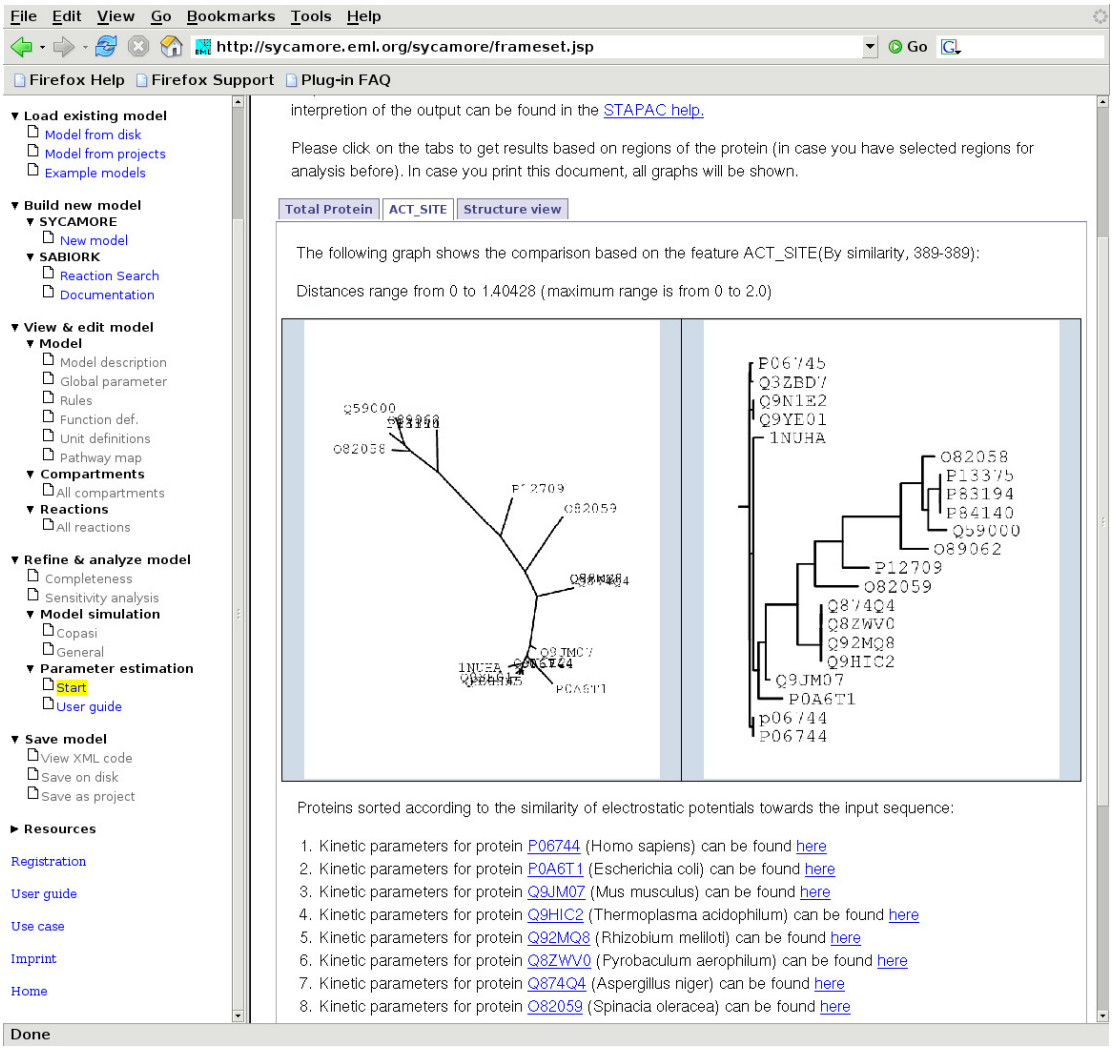



8. The results are two tree-like diagrams:

1. One for the electrostatic potential of the entire protein




2. One for the electrostatic potential around the chosen region (ACT_SITE residue 389).




9. Results
    1. The electrostatic potential at the active site of the query sequence P06744 (Homo sapiens) is closest to the ones from mouse (P06745), bovine (Q3ZBD7) and rabbit (Q9N1E2).
    2. Most distant are *M. janaschii* (Q59000), spinach (O82058), *B. stearothermophilus* (P13375), *P. furiosus* (P83194), *Th.litoralis* (P84140).
    3. This is in agreement with the experimentally measured $K_m$ values for glucose-6-phosphate:
        1. *Homo sapiens* 0.445 mM
        2. *Bos taurus* 0.45-0.58 mM
        3. *Mus musculus* 0.48 mM
        4. *Oryctolagus cuniculus* 0.6 mM
        5. *M. janaschii* 1.0 mM
        6. *P. furiosus* 1.99 mM
        7. *Th.litoralis* 11.7 mM
    4. Thus, kinetic parameters from, for example bovine, mouse and rabbit can be used with some confidence in simulations of the human enzyme.