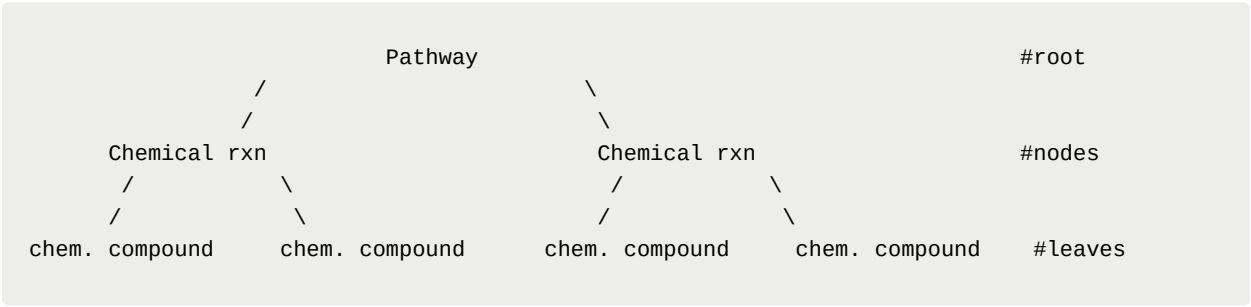
# Import guide

**Introduction**

Before begin importing anything to our database, I have provided a brief introduction on the terminology and general idea of how we will import our data.

We can imagine the database structure roughly as a tree (shown by the diagram below). Each item in our tree is independent of eachother but have relationships as indicated by the branches(dashed lines)




For our import we first need to import all the leaves within the tree. From then we will then add the chemical reactions that occur within our pathway followed by the pathway itself.

---

## Importing A Pathway:

This is a guide on how to import a pathway to 3DMetabolism's wikibase

Table of Contents:

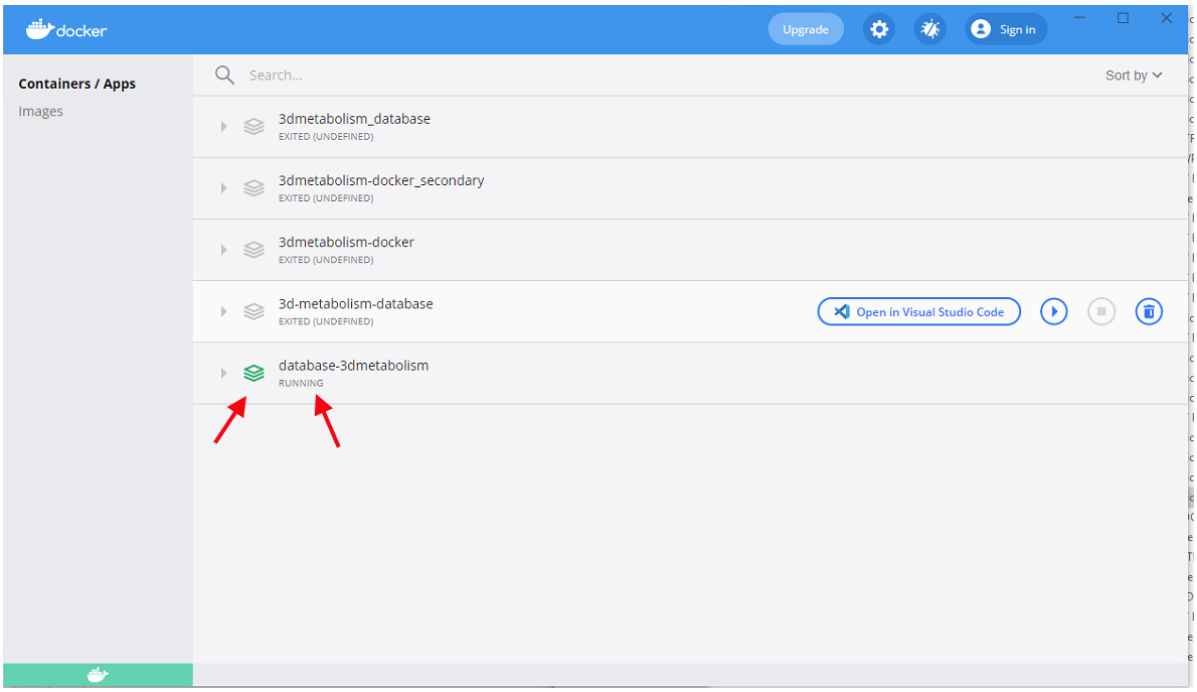


---

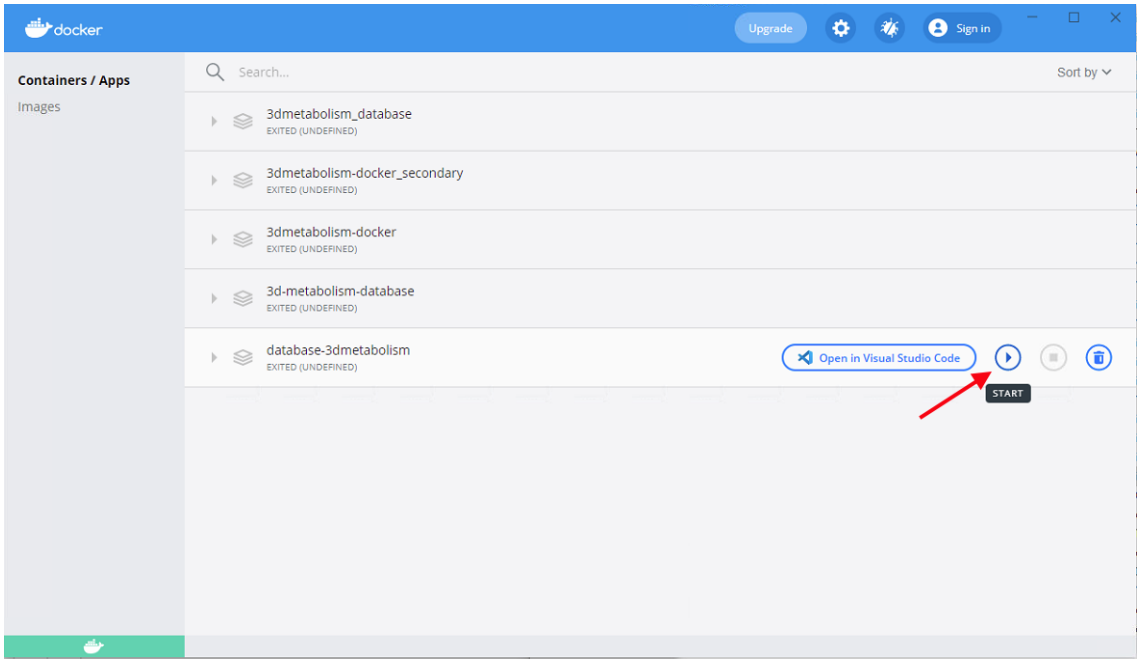
## Prior to import

Before importing anything to the database, we first need to do some preliminary setups and checks:

- Make sure you have the right Docker container running. This is done by opening the Docker application and looking at the container labelled `database-3dmetabolism`. If the container says *RUNNING* and the icon to the left is green (as indicated in the photo below), the container is running.



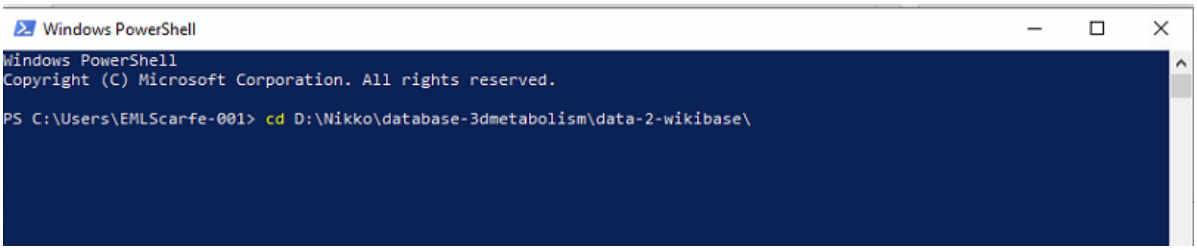

  - If the docker container is not running, hover your mouse over the container and click the *Start* button as indicated by the red arrow. The container should start up shortly(~2minutes)



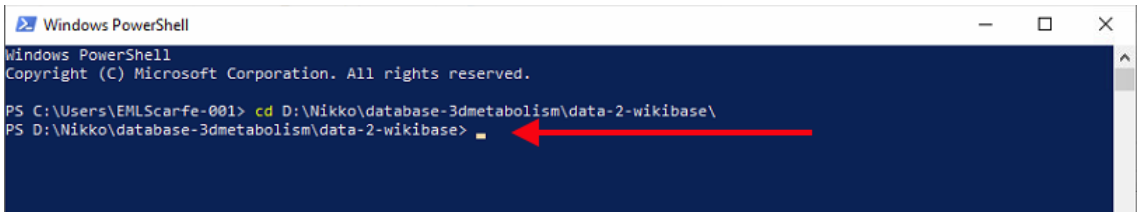

- You will need the following scripts:
  - `root_pathway_prep.py`
    - This script is responsible for creating a .csv file for a given pathway
  - `nodes_pathway_prep.py`
    - This script is responsible for creating a .csv file for a given pathways children
  - `leaves_pathway_prep.py`
    - This script is responsible for creating a .csv file for a given pathways leaves
  - `add_items.py`
    - This script is responsible for importing a .csv file to a wikibase instance
  - `add_properties.py`
    - This script is responsible for importing properties to a wikibase instance

  these scripts should already be located within the wikibase folders.
  **"D:\Nikko\database-3dmetabolism"**

- We will use Powershell as the terminal to run our python scripts. Before running, make sure the powershell is navigated to the proper directory: This is done by copying and pasting (CTRL + C to copy, CTRL + V to paste) the following command in the terminal: `cd D:\Nikko\database-3dmetabolism\data-2-wikibase\`. Then hit enter



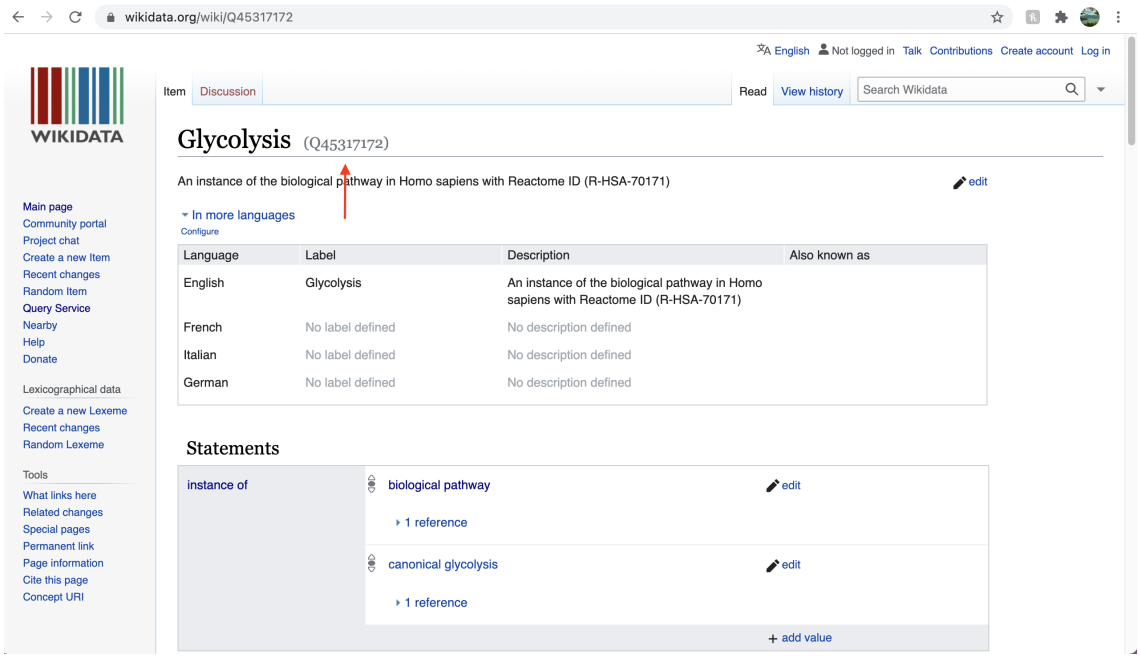

  - To confirm you are located in the proper folder, check the left side text of in the terminal. The text should look as the following.




- Generally this step is only required if we were to completely start from the beginning with a new instance of a wikibase docker container: Make sure the `add_items.py` and `add_properties.py` scripts have the proper bot name and key. This is done by opening both scripts in an IDE/text editor of choice (I recommend VS Code) and providing the proper name and key to the script. Remember to save any changes.

## Import

1. To begin adding your Pathway, we must first begin by adding the "leaves" of the pathway. This is done by using the `leaves_pathway_prep.py` script.
   - To run the script you must run the following command in your terminal: `py .\nodes_pathway_prep.py "[Wikidata QID]"` where `[Wikidata QID]` is the Pathways QID on wikidata.org




The QID of the pathway is located beside the pathways name within the brackets. Do not include the brackets when copying and pasting

```
PS D:\Nikko\3DMetabolism-Docker_Secondary\data-2-wikibase> py .\leaves_prep.py "Q45317172"
```

This is what running leaves_prep.py  with Glycolysis in your terminal looks like. Make sure the to wrap the QID in quote marks when running the command

2. After running `leaves_pathway_prep.py`, a new .csv file will appear within our directory containing the leaves of the pathway you had chosen.  You then must then add that CSV file to our database. To do so we will run the following command in our terminal: `py add_items.py .\leaves.csv` and hit enter

```
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py .\leaves_pathway_prep.py "Q45317172"
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py add_items.py .\leaves.csv
```

This is what running add_items.py on leaves.csv looks like

- running this command will display all the items within the CSV file in our terminal.

```
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py .\leaves_pathway_prep.py "Q45317172"
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py add_items.py .\leaves.csv
Q6 2,3-bisphosphoglycerate
Q7 2-phosphonato-D-glycerate(3-)
Q8 3-phosphonato-D-glycerate(3-)
Q9 3-phosphonato-D-glyceroyl phosphate(4-)
Q10 acetyl coenzyme a
Q11 adenosine cyclic phosphate
Q12 adenosine monophosphate
Q13 ADP(3-)
Q14 ADPGK:Mg2+ [endoplasmic reticulum membrane]
Q15 aldolase tetramer [cytosol]
Q16 alpha-D-glucose 6-phosphate(2-) ion
Q17 ammonium cation
Q18 ATP(4-)
Q19 beta-D-fructofuranose 1,6-bisphosphate(4-)
Q20 beta-D-fructofuranose 6-phosphate(2-)
Q21 BPGM dimer [cytosol]
Q22 citric acid
Q23 cytosolic GCK1:GKRP complex <=> glucokinase (GCK1) + glucokinase regulatory protein (GKRP)
```

3. We will then run the next script, `nodes_pathway_prep.py` . This is done by the copying this command into the terminal: `py .\nodes_pathway_prep.py "[Wikidata QID]"` where `[Wikidata QID]` is the same QID from the previous step

```
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py .\nodes_pathway_prep.py "Q45317172"
```

This is what running nodes_pathway_prep.py in your terminal should look like

4. After running `nodes_pathway_prep.py`, a new .csv file will appear within our directory that `nodes_pathway_prep.py` has created. The new csv file should be titled `[QID]_nodes_import.csv` , where `[Wikidata QID]` is the QID of the pathway you had chosen. We must now add that csv file contents to our database. To do so we will run the following command in our terminal: `py add_items.py .\[Wikidata QID]_nodes_import.csv` and hit enter

```
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py .\nodes_pathway_prep.py "Q45317172"
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py add_items.py .\Q45317172_nodes_import.csv
```

This is what running add_items.py on the output csv that nodes_pathway_prep.py created

```
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py .\nodes_pathway_prep.py "Q45317172"
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py add_items.py .\Q45317172_nodes_import.csv
Q58 Regulation of Glucokinase by Glucokinase Regulatory Protein
Q59 HK1,2,3,GCK phosphorylate Glc to form G6P
Q60 ADPGK:Mg2+ phosphorylates Glc to G6P
Q61 alpha-D-glucose 6-phosphate <=> D-fructose 6-phosphate
Q62 GNPDA1,2 hexamers deaminate GlcN6P to Fru(6)P
Q63 D-fructose 6-phosphate + ATP => D-fructose 1,6-bisphosphate + ADP
Q64 D-fructose 1,6-bisphosphate <=> dihydroxyacetone phosphate + D-glyceraldehyde 3-phosphate
Q65 dihydroxyacetone phosphate <=> D-glyceraldehyde 3-phosphate
Q66 D-glyceraldehyde 3-phosphate + orthophosphate + NAD+ <=> 1,3-bisphospho-D-glycerate + NADH + H+
Q67 BPGM dimer isomerises 1,3BPG to 2,3BPG
Q68 PGM2L1:Mg2+ phosphorylates G6P to G1,6BP
Q69 1,3-bisphospho-D-glycerate + ADP <=> 3-phospho-D-glycerate + ATP
Q70 PGP:Mg2+ dimer hydrolyses 3PG to glycerol
Q71 3-Phospho-D-glycerate <=> 2-Phospho-D-glycerate
Q72 2-Phospho-D-glycerate <=> Phosphoenolpyruvate + H2O
Q73 phosphoenolpyruvate + ADP => pyruvate + ATP
Q74 Regulation of glycolysis by fructose 2,6-bisphosphate metabolism
```

After hitting running the add_items.py script on a nodes_pathway_prep.py csv file, you will see something similar to this output. This tells you what items have been successfully added to our database.

5. The last step is to add the pathway itself. This is done by running `root_pathway_prep.py` in your terminal to create a csv. To do so, we will run the following command: `py root_pathway_prep.py [Wikidata QID]`.

   Once `root_pathway_prep.py` is done running, a new csv will be created containing the glycolysis which we will use `add_items.py` to import to our database. The csv file will have the following name, `[Wikidata QID]pathway_import.csv`

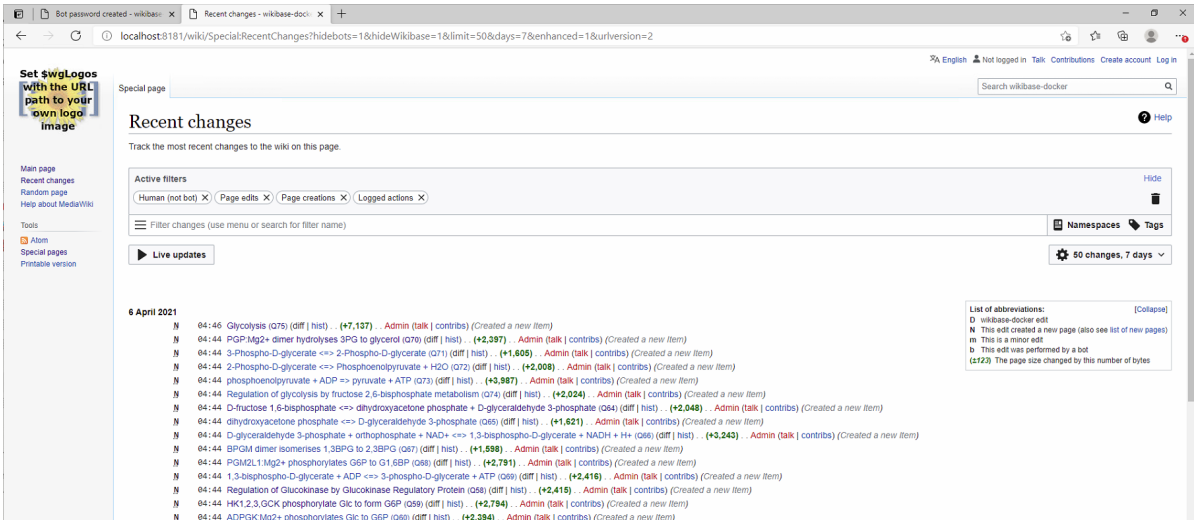
```
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py .\root_pathway_prep.py "Q45317172"
```

This is what running root_pathway_prep.py looks like in our terminal

```
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py .\root_pathway_prep.py "Q45317172"
PS D:\Nikko\database-3dmetabolism\data-2-wikibase> py add_items.py .\Q45317172pathway_import.csv
Q75 Glycolysis
PS D:\Nikko\database-3dmetabolism\data-2-wikibase>
```

This is what running add_items.py with root_pathway_prep.py output csv file looks like

6. To check the the database out, we can open a browser and type http://localhost:8181/wiki/Special:RecentChanges?hidebots=1&hideWikibase=1&limit=50&days=7&enhanced=1&urlversion=2 to see the logs of the recent changes we have done. Note: this link will only work for the computer that is running the Docker container of our database.


Bot password created - wikibase
Recent changes - wikibase-dock
localhost:8181/wiki/Special:RecentChanges?hidebots=1&hideWikibase=1&limit=50&days=7&enhanced=1&urlversion=2
Set $wgLogos with the URL path to your own logo image
English
Not logged in
Talk
Contributions
Create account
Log in
Special page
Search wikibase-docker
Recent changes
Help
Track the most recent changes to the wiki on this page.
Main page
Recent changes
Random page
Help about MediaWiki
Tools
Atom
Special pages
Printable version
Active filters
Hide
Human (not bot)
Page edits
Page creations
Logged actions
Filter changes (use menu or search for filter name)
Namespaces
Tags
Live updates
50 changes, 7 days
6 April 2021
List of abbreviations:
[Collapse]
D wikibase-docker edit
N This edit created a new page (also see list of new pages)
m This is a minor edit
b This edit was performed by a bot
(±123) The page size changed by this number of bytes
04:46 Glycolysis (Q75) (diff | hist) . . (+7,137) . . Admin (talk | contribs) (Created a new item)
04:44 PGP:Mg2+ dimer hydrolyses 3PG to glycerol (Q70) (diff | hist) . . (+2,397) . . Admin (talk | contribs) (Created a new item)
04:44 3-Phospho-D-glycerate <=> 2-Phospho-D-glycerate (Q71) (diff | hist) . . (+1,605) . . Admin (talk | contribs) (Created a new item)
04:44 2-Phospho-D-glycerate <=> Phosphoenolpyruvate + H2O (Q72) (diff | hist) . . (+2,008) . . Admin (talk | contribs) (Created a new item)
04:44 phosphoenolpyruvate + ADP => pyruvate + ATP (Q73) (diff | hist) . . (+3,987) . . Admin (talk | contribs) (Created a new item)
04:44 Regulation of glycolysis by fructose 2,6-bisphosphate metabolism (Q74) (diff | hist) . . (+2,024) . . Admin (talk | contribs) (Created a new item)
04:44 D-fructose 1,6-bisphosphate <=> dihydroxyacetone phosphate + D-glyceraldehyde 3-phosphate (Q64) (diff | hist) . . (+2,048) . . Admin (talk | contribs) (Created a new item)
04:44 dihydroxyacetone phosphate <=> D-glyceraldehyde 3-phosphate (Q65) (diff | hist) . . (+1,621) . . Admin (talk | contribs) (Created a new item)
04:44 D-glyceraldehyde 3-phosphate + orthophosphate + NAD+ <=> 1,3-bisphospho-D-glycerate + NADH + H+ (Q66) (diff | hist) . . (+3,243) . . Admin (talk | contribs) (Created a new item)
04:44 BPGM dimer isomerises 1,3BPG to 2,3BPG (Q67) (diff | hist) . . (+1,598) . . Admin (talk | contribs) (Created a new item)
04:44 PGM2L1:Mg2+ phosphorylates G6P to G1,6BP (Q68) (diff | hist) . . (+2,791) . . Admin (talk | contribs) (Created a new item)
04:44 1,3-bisphospho-D-glycerate + ADP <=> 3-phospho-D-glycerate + ATP (Q69) (diff | hist) . . (+2,416) . . Admin (talk | contribs) (Created a new item)
04:44 Regulation of Glucokinase by Glucokinase Regulatory Protein (Q58) (diff | hist) . . (+2,415) . . Admin (talk | contribs) (Created a new item)
04:44 HK1,2,3,GCK phosphorylate Glc to form G6P (Q59) (diff | hist) . . (+2,794) . . Admin (talk | contribs) (Created a new item)
04:44 ADPGK:Mg2+ phosphorylates Glc to G6P (Q60) (diff | hist) . . (+2,394) . . Admin (talk | contribs) (Created a new item)