

# Tutorial for KEGG Online Database

DI Michael Kalkusch  
[kalkusch@icg.tu-graz.ac.at](mailto:kalkusch@icg.tu-graz.ac.at)

Version 1.2

September 20<sup>th</sup>, 2006

## ***Acknowledgment:***

Thanks to Dr. Peter Abuja for his expert knowledge in the area of metabolisms and pathways.

This document was created by Michael Kalkusch after a meeting on August 17<sup>th</sup> 2006 with Dr. Abuja and Marc Streit. Thanks to Marc Streit for his contribution during the initial creation of this document.

The document was update 20<sup>th</sup> September, 2006 by Michael Kalkusch after a meeting with Dr. Abuja.

This document will be available at:

<http://www.icg.tu-graz.ac.at/research/CGIS/GENVIEW/>

## Snapshots from KEGG web interface

<http://www.genome.jp/kegg/pathway.html> visited 17-08-2006

Example: Pathway „*metionine metabolism - Reference pathway* „

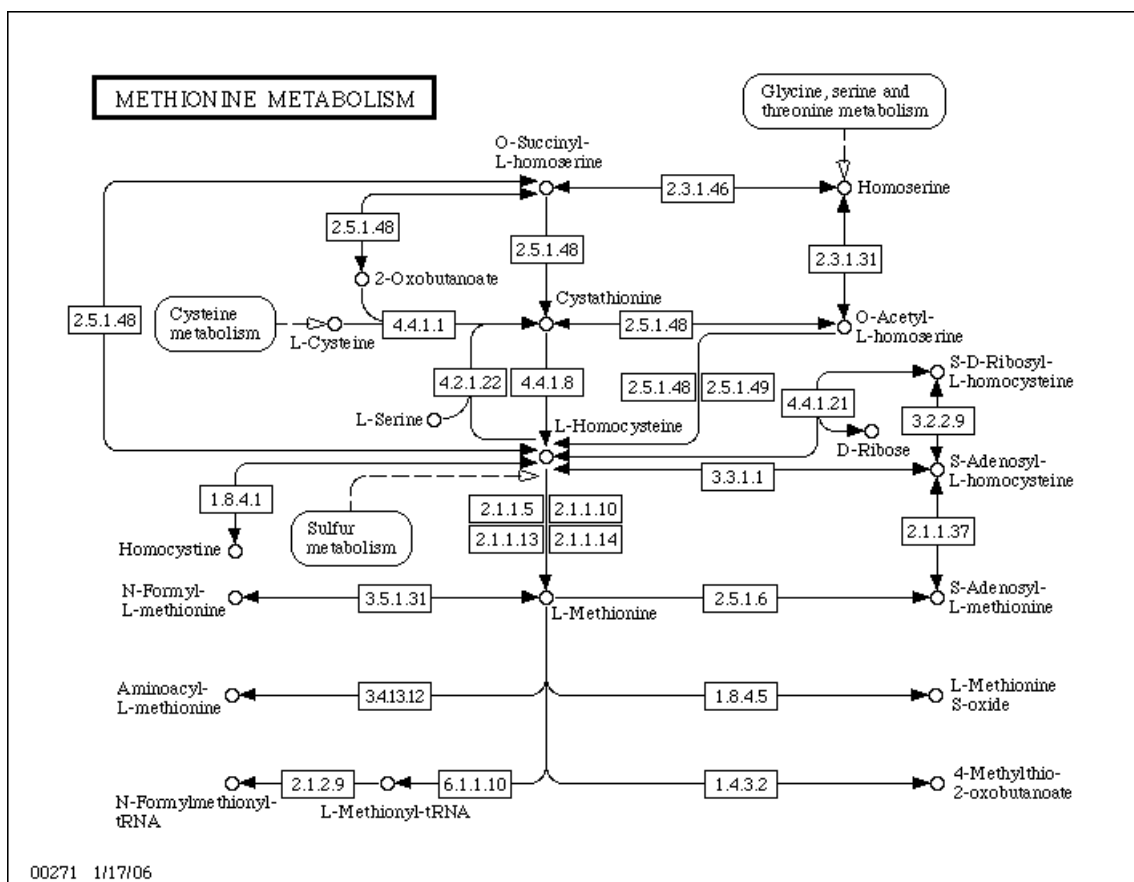


Figure 1:

Klick on [2.5.1.48] opens this:



ENZYME: 2.5.1.48

Help

<b>Entry</b>	EC 2.5.1.48	Enzyme
<b>Name</b>	cystathionine gamma-synthase; O-succinyl-L-homoserine succinate-lyase (adding cysteine); O-succinylhomoserine (thiol)-lyase; homoserine O-transsuccinylase; O-succinylhomoserine synthase; O-succinylhomoserine synthetase; cystathionine synthase; cystathionine synthetase; homoserine transsuccinylase	
<b>Class</b>	Transferases Transferring alkyl or aryl groups, other than methyl groups Transferring alkyl or aryl groups, other than methyl groups	
<b>Sysname</b>	O4-succinyl-L-homoserine:L-cysteine S-(3-amino-3-carboxypropyl)transferase	
<b>Reaction</b>	O-succinyl-L-homoserine + L-cysteine = L-cystathionine + succinate [RN: <a href="#">R00999</a> <a href="#">R01288</a> <a href="#">R02508</a> <a href="#">R03132</a> <a href="#">R03260</a> <a href="#">R04944</a> <a href="#">R04945</a> <a href="#">R04946</a> ]	
<b>Substrate</b>	O-Succinyl-L-homoserine [CPD: <a href="#">C01118</a> ]; L-Cysteine [CPD: <a href="#">C00097</a> ]	
<b>Product</b>	L-Cystathionine [CPD: <a href="#">C02291</a> ]; Succinate [CPD: <a href="#">C00042</a> ]	
<b>Cofactor</b>	Pyridoxal phosphate [CPD: <a href="#">C00018</a> ]	
<b>Comment</b>	A pyridoxal-phosphate protein. Also reacts with hydrogen sulfide and methanethiol as replacing agents, producing homocysteine and methionine, respectively. In the absence of thiol, can also catalyse beta,gamma-elimination to form 2-oxobutanoate, succinate and ammonia.	
<b>Pathway</b>	PATH: <a href="#">map00271</a> Methionine metabolism PATH: <a href="#">map00272</a> Cysteine metabolism PATH: <a href="#">map00450</a> Selenoamino acid metabolism PATH: <a href="#">map00920</a> Sulfur metabolism	
<b>Ortholog</b>	KO: <a href="#">K01739</a> cystathionine gamma-synthase	
<b>Genes</b>	XLA: <a href="#">494673</a> (LOC494673) XTR: <a href="#">394634</a> (MGC75946) ATH: <a href="#">At3g01120</a> (T4P13.19) CME: <a href="#">CMF156C</a> SCE: <a href="#">YJR130C</a> (STR2) <a href="#">YML082W</a> AGO: <a href="#">AER164C</a> (AER164Cp) CAL: <a href="#">orf19.1033</a> (STR2) <a href="#">orf19.7297</a> SPO: <a href="#">SPBC15D4.09c</a> CNE: <a href="#">CNC01220</a> LMA: <a href="#">LmjF35.3230</a> EHI: <a href="#">132.t00018</a> <a href="#">389.t00003</a> <a href="#">395.t00003</a> ECO: <a href="#">b3939</a> (metB) ECJ: <a href="#">JW3910</a> (metB) ...: reduced the list of 191 gene in total	
<b>Reference</b>	1 [PMID: <a href="#">5340123</a> ] Flavin M, Slaughter C. Enzymatic synthesis of homocysteine or methionine directly from	

O-succinyl-homoserine.  
 Biochim. Biophys. Acta. 132 (1967) 400-5.  
 2 [PMID:[5922970](#)]  
 Kaplan MM, Flavin M.  
 Cystathionine gamma-synthetase of Salmonella. Catalytic properties  
 of a new enzyme in bacterial methionine biosynthesis.  
 J. Biol. Chem. 241 (1966) 4463-71.  
 3 [PMID:[6016326](#)]  
 Wiebers JL, Garner HR.  
 Homocysteine and cysteine synthetases of Neurospora crassa.  
 Purification, properties, and feedback control of activity.  
 J. Biol. Chem. 242 (1967) 12-23.  
 4  
 Wiebers, J.L. and Garner, H.R. Acyl derivatives of homoserine as  
 substrates for homocysteine synthesis in Neurospora crassa, yeast,  
 and Escherichia coli. J. Biol. Chem. 242 (1967) 5644-5649.  
 5 [PMID:[9843488](#)]  
 Clausen T, Huber R, Prade L, Wahl MC, Messerschmidt A.  
 Crystal structure of Escherichia coli cystathionine gamma-synthase  
 at 1.5 Å resolution.  
 EMBO. J. 17 (1998) 6827-38.  
 6 [PMID:[9531508](#)]  
 Ravanel S, Gakiere B, Job D, Douce R.  
 Cystathionine gamma-synthase from Arabidopsis thaliana: purification  
 and biochemical characterization of the recombinant enzyme  
 overexpressed in Escherichia coli.  
 Biochem. J. 331 ( Pt 2) (1998) 639-48.

**Other DBs** IUBMB Enzyme Nomenclature: [2.5.1.48](#)  
 ExPASy - ENZYME nomenclature database: [2.5.1.48](#)  
 ERGO genome analysis and discovery system: [2.5.1.48](#)  
 BRENDA, the Enzyme Database: [2.5.1.48](#)  
 CAS: 9030-70-0

**LinkDB**



=> [Original format](#)

---

[DBGET](#) integrated database retrieval system, [GenomeNet](#)

**Table 1**

Klick on pathway „PATH: [map00271](#) metionine metabolism“ shows this:

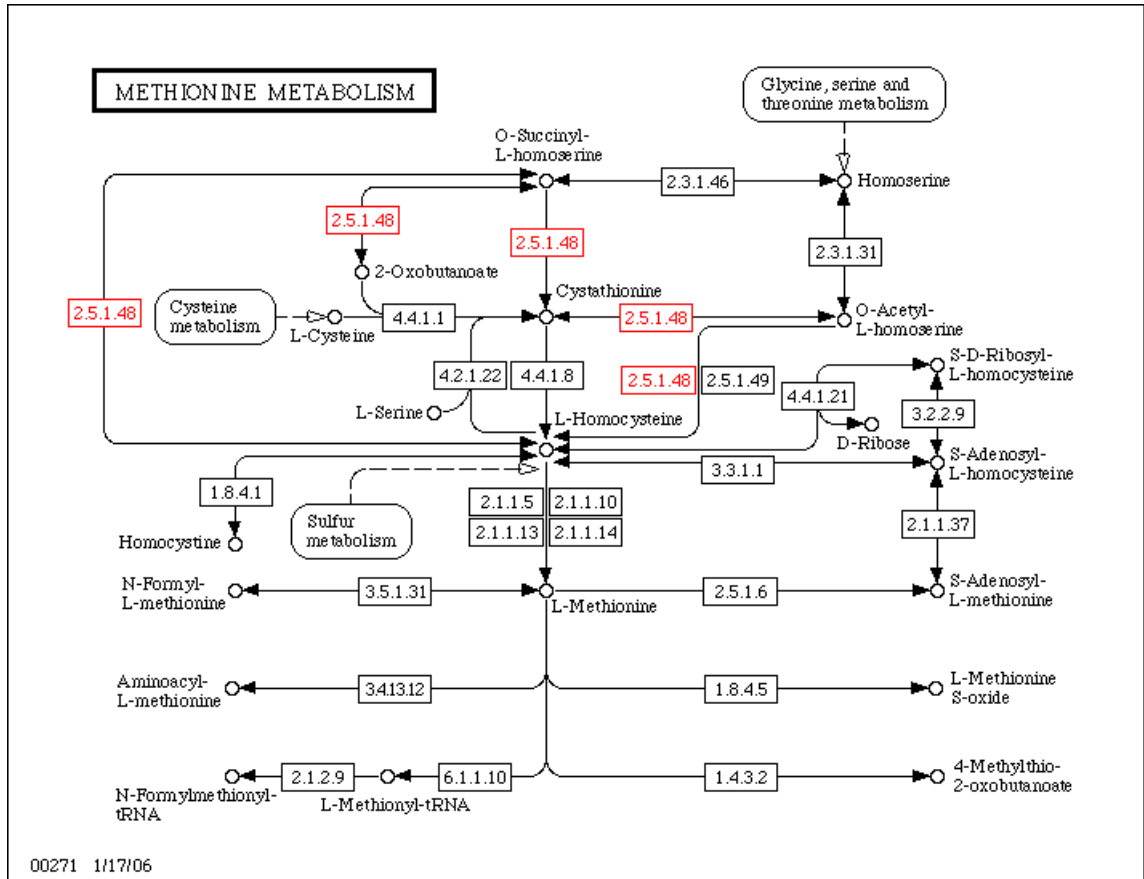


Figure 2:

Klick on Cystathionine opens this page:



COMPOUND: C02291

Help

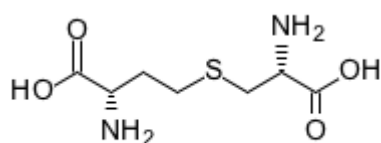
**Entry** C02291 Compound

**Name** L-Cystathionine

**Formula** C<sub>7</sub>H<sub>14</sub>N<sub>2</sub>O<sub>4</sub>S

**Mass** 222.0674

**Structure**



C02291

Mol file

KCF file

DB search

**Reaction** [R01001](#) [R01286](#) [R01290](#) [R03217](#) [R03260](#)

**Pathway** PATH: [map00260](#) Glycine, serine and threonine metabolism  
PATH: [map00271](#) Methionine metabolism

**Enzyme** [2.5.1.48](#) [2.5.1.49](#) [4.2.1.22](#) [4.4.1.1](#)  
[4.4.1.8](#)

**Other DBs** CAS: 56-88-2  
PubChem: [5347](#)  
ChEBI: [17482](#)

**LinkDB** [All DBs](#)

**KCF data** [Show](#)

=> [Original format](#)

---

[DBGET](#) integrated database retrieval system, [GenomeNet](#)

Table 2

Example for a reaction:



REACTION: R00999

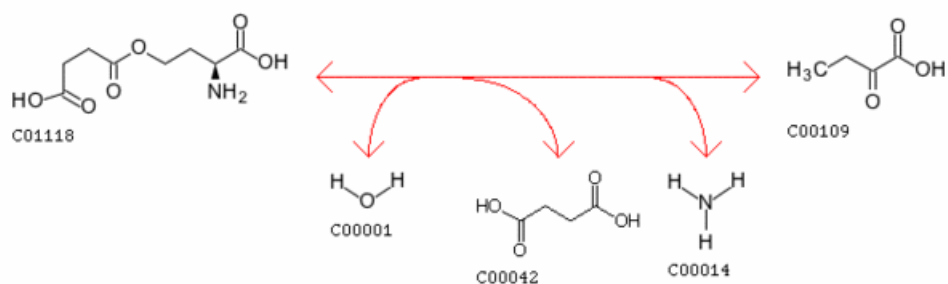
Help

**Entry** R00999 Reaction

**Name** O-Succinyl-L-homoserine succinate-lyase (adding cysteine)

**Definition** O-Succinyl-L-homoserine + H<sub>2</sub>O  $\rightleftharpoons$  2-Oxobutanoate + Succinate + NH<sub>3</sub>

**Equation** [C01118](#) + [C00001](#)  $\rightleftharpoons$  [C00109](#) + [C00042](#) + [C00014](#)



**RPair** RP: [A00132](#) C00042\_C01118 main  
RP: [A01229](#) C00109\_C01118 main

**Pathway** PATH: [rn00271](#) Methionine metabolism

**Enzyme** [2.5.1.48](#)

**Ortholog** KO: [K01739](#) cystathionine gamma-synthase

**LinkDB** [All DBs](#)

=> [Original format](#)

---

[DBGET](#) integrated database retrieval system, [GenomeNet](#)

Table 3



## Flow Charts:

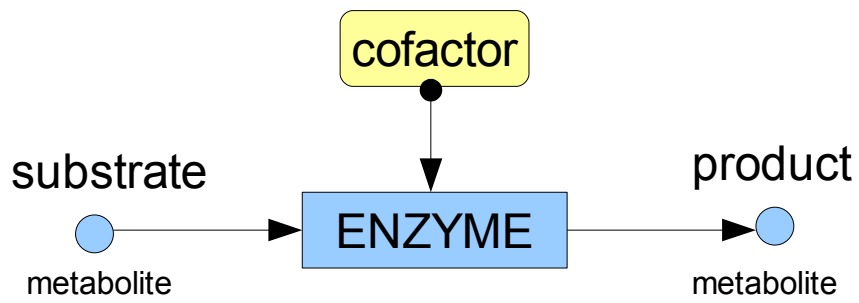


Figure 3: Basic graph with metabolite and enzyme regulated by cofactor

## Regulation of enzymes and metabolites

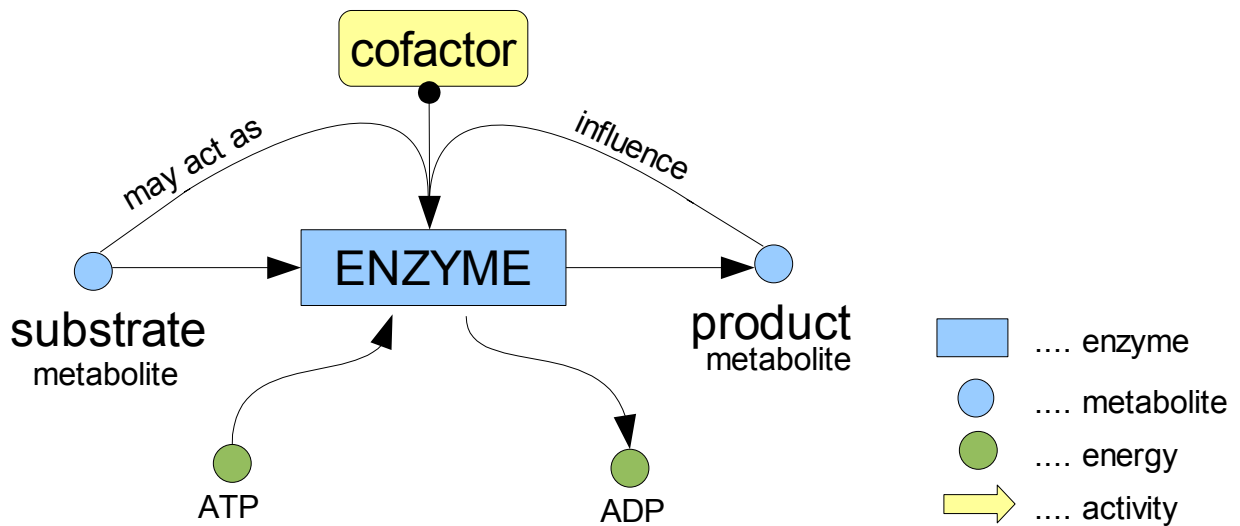


Figure 5: Extended model of pathway

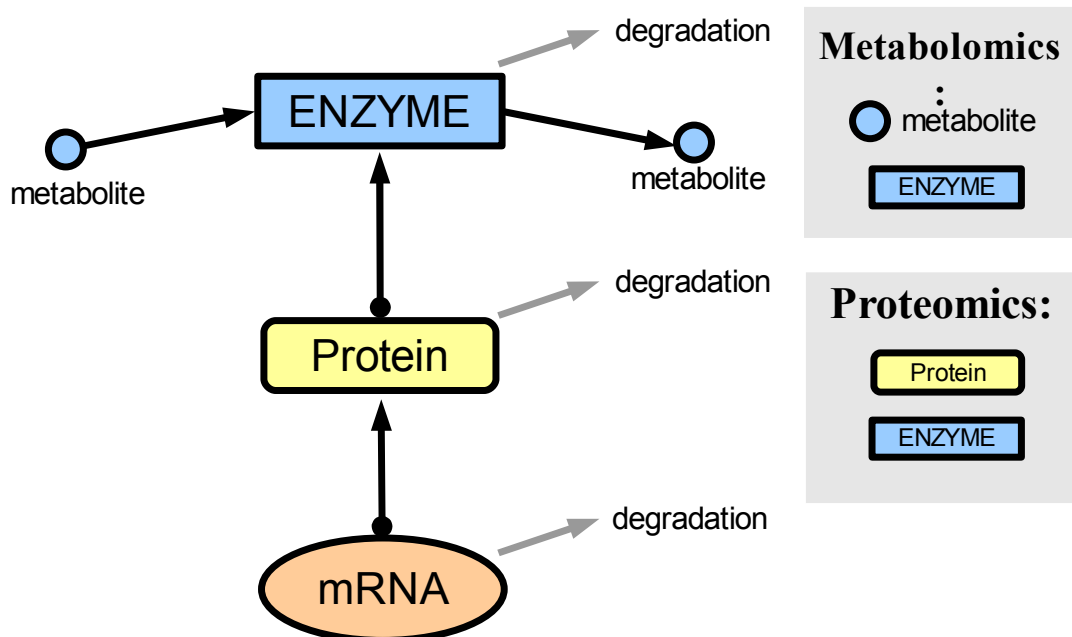


Figure 6: Interrelationship of mRNA, protein, enzyme and metabolism

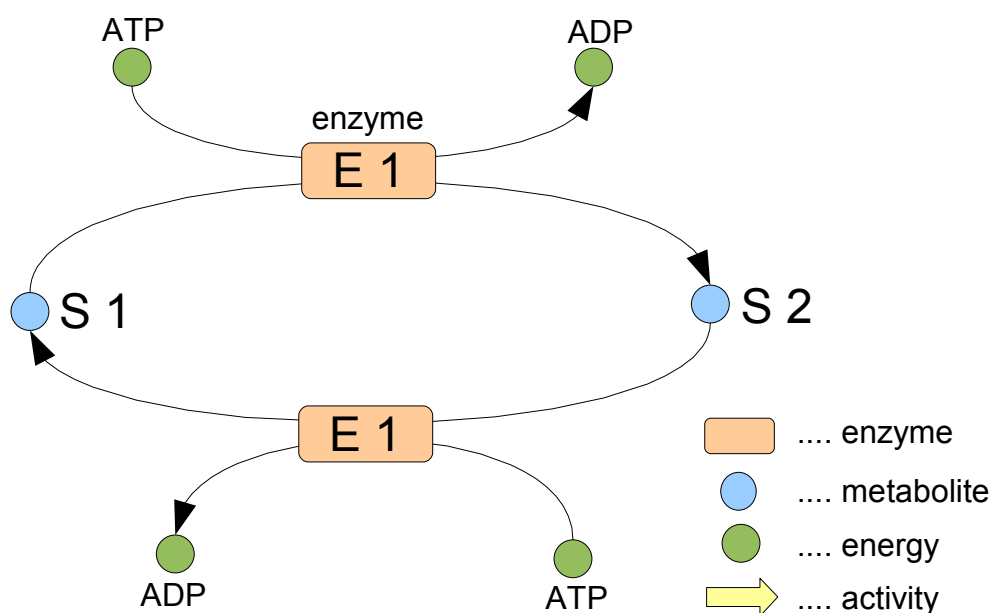


Figure 7: Two metabolites and two enzymes form a switch, that consumes energy

## Concepts:

- Inside pathway: 2 step, 3 step relations inside the graph from a given Protein or metabolit
- Take connectivity of metabolit into account (i.e. google principle)
- Use connectivity of metabolit as weight for its importance inside pathways (Spring mass model / Gravitation model)
- Interrelations/ interaction:
  - Kinase  $\implies$  Protein
  - Protein  $\iff$  Protein interrelations
- Visualize flows from:

Substrate  $\implies$   
(Enzyme/ Protein)  $\leftarrow$  using cofactor  $\implies$   
metabolit A  $\implies$   
(Enzyme / Protein)  $\implies$   
metabolit B

A possibility here could be the application of Petri-Nets ([http://en.wikipedia.org/wiki/Petri\\_net](http://en.wikipedia.org/wiki/Petri_net))

- Semantic Zooming:
  - hide / show metabolits
  - collapse protein chains
  - collapse pathway inside another pathway
  - magic lens for semantic zooming
- Measure activity of metabolids:  
required for detailed analysis, but not done within Micro Array analysis
- Position of enzyme inside cell
- Compartments the enzyme is present in
- Handle enzyme cascades  $\implies$  if top element of enzyme-cascade is active, the cascade might be active too. Thus only the result of the cascade is active, but not its intermediate enzymes.

- Important fields of „Enzymes“ in KEGG (see table 1):

Entry: unique international standardized identifier

Name: known names for this molecule and aliases for it

Reaction: chemical reaction

Substrate: Input to the Enzyme

Product: Output of the Enzyme

Cofactor: if set, it is necessary to activate the Enzyme

Pathway: shows only active pathways linked to one Enzyme but not pathways influenced by this enzyme!

Ortholog: special genes (see glossary)

Genes: ???

- Important fields of „Compounds“ in KEGG (see table 2):

Entry: unique code (used by KEGG? Only or also international unique?)

Name: known names for this molecule and aliases for it

Mass: mass of this molecule (mass unit?)

Structure: chemical structure of this molecule

Reaction: chemical reaction, where this compound is part of

Pathway: list of active pathways, where this compound can be found

Enzyme: list of enzymes capable of handling this compound, either as input or as output; also known as „Substrate“ and „Product“

- Important fields of „Reaction“ in KEGG (see table 3):

Entry: unique name (probably only in KEGG)

Name: name that defines this reaction

Definition: define this reaction

Equation: link to compounds of this reaction

Repair: ??

Pathway: list of pathways where this reaction is part of

Enzyme: list of enzymes that do this reaction

Ortholog: special genes (see glossary)

## Other databases aside from KEGG:

(suggested by Dr. Abuja)

Biomodal DB:	<a href="http://www.ebi.ac.uk/biomodals/">http://www.ebi.ac.uk/biomodals/</a>
Reactome „A curated knowledgebase of biological pathways“	<a href="http://www.reactome.org/">http://www.reactome.org/</a>
Brookhaven Protein Database for structure of enzymes	<a href="http://cds.dl.ac.uk/cds/pdb.html">http://cds.dl.ac.uk/cds/pdb.html</a>
EMBL „European Molecular Biology Laboratory“	<a href="http://www.embl.org/">http://www.embl.org/</a>
EMBO „European Molecular Biology Organization“	<a href="http://www.embo.org/">http://www.embo.org/</a>
SMBL „Systems Biology Markup Language“	<a href="http://sbml.org/">http://sbml.org/</a>

### *Other important websites:*

RCSB „Research Collaboratory for Structural Bioinformatics“	<a href="http://home.rcsb.org/">http://home.rcsb.org/</a>
RCSB PDB Protein Data Base	<a href="http://www.rcsb.org/pdb/">http://www.rcsb.org/pdb/</a>

## Software we should look at:

- „Panther“ (Applied Biosoftware)
- google: „Pathway“ as well as „Pathway Heatmap“

### Books:

David Fell, „Regulation of Metabolism“ („Kinetische Regulation“)

Homepage of David Fell:

<http://www.brookes.ac.uk/bms/research/fell.html>

## Glossary:

Term	Description	Link
Substrate	a <a href="#">molecule</a> upon which an <a href="#">enzyme</a> acts. Enzymes <a href="#">catalyze chemical reactions</a> involving the substrate(s).	<a href="http://en.wikipedia.org/wiki/Substrate_(biochemistry)">http://en.wikipedia.org/wiki/Substrate_(biochemistry)</a>
Kinase	a type of <a href="#">enzyme</a> that transfers <a href="#">phosphate</a> groups from <a href="#">high-energy</a> donor molecules, such as <a href="#">ATP</a> , to specific target molecules ( <a href="#">substrates</a> ); the process is termed <a href="#">phosphorylation</a> .	<a href="http://en.wikipedia.org/wiki/Kinase">http://en.wikipedia.org/wiki/Kinase</a>
Inhibitor	???	
Effektor	???	
Kinetische Regulation	???	
EC-Number	Enzyme Commission number (EC number)	<a href="http://en.wikipedia.org/wiki/EC_number">http://en.wikipedia.org/wiki/EC_number</a>
Homocysteine		<a href="http://en.wikipedia.org/wiki/Homocysteine">http://en.wikipedia.org/wiki/Homocysteine</a>
Ortholog	<b>Orthologs</b> are <a href="#">genes</a> in different species which <a href="#">evolved</a> from a common ancestral <a href="#">gene</a> .	<a href="http://en.wikipedia.org/wiki/Ortholog">http://en.wikipedia.org/wiki/Ortholog</a>
Proteomics		<a href="http://en.wikipedia.org/wiki/Proteomics">http://en.wikipedia.org/wiki/Proteomics</a>
Metabolomics		<a href="http://en.wikipedia.org/wiki/Metabolomics">http://en.wikipedia.org/wiki/Metabolomics</a>
mRNA	Messenger RNA, transcribed DNA that leaves the cell core.  A ribosom transcribed the mRNA to a Protein.	<a href="http://en.wikipedia.org/wiki/MRNA">http://en.wikipedia.org/wiki/MRNA</a>
Nucleotide	A,G,T,C (U=T) Adenosine Guanin Thymidine (DNA only) Cytidine Uridine (RNA only) = T	<a href="http://en.wikipedia.org/wiki/Nucleotide">http://en.wikipedia.org/wiki/Nucleotide</a>

Term	Description	Link
Codon	Group of three Nucleotides. A ribosom translates Codons of the mRNA to amino acids, that are linked to gethe to form a protein.	<a href="http://en.wikipedia.org/wiki/Codon">http://en.wikipedia.org/wiki/Codon</a>
Amino acids	Amino acids build a Protein.	<a href="http://en.wikipedia.org/wiki/List_of_standard_amino_acids">http://en.wikipedia.org/wiki/List_of_standard_amino_acids</a>
Ribosom	Translates the mRNA to a Protein by reading Codons and translating them to amion acids, that are linked together forming the protein	<a href="http://en.wikipedia.org/wiki/Ribosome">http://en.wikipedia.org/wiki/Ribosome</a>