

Wed Jan 2 13:33:31 1991

1

~~(x)~~ JAN -2 1991

To: (monica riley) mriley@hoh.mbl.edu
Fcc: inbox
Reply-to: (J. Lederberg) lederberg@rockvax.rockefeller.edu
Subject: metabolism data base

Hi, Monica -- Of course I remember you! And very much appreciate your recent book and Micr. Rev. on The Bacterial Chromosome.

The data base you're contemplating needs a lot of discussion and thought to see how it might best complement existing efforts. (E.G. Mary Berlyn at Yale on E. coli genome; and all the GenBank etc stuff.) And a molecular biology bibliographic project at Stanford (Ted Shortliffe, on which I'm a co-PI).

And you should consult Dennis Smith at Molecular Design Ltd. (Calif) about representations of organic reactions; likewise look at Index Chemicus (Institute of Scientific Information).

There are some tough conceptual problems defining the boundaries of "metabolism" and development. Where would you put regulatory methylation (or any of the rest of the chemistry) of DNA?

I'm not sure you end up with anything less than the entire extant literature as the database, and a desire for assurance that one can use metabolism-oriented keys for searching that. If I'm not mistaken, Medline already uses the EC numbers in the keyword list but these are likely to be neglected for commonplace reactions. When we get to full text searchable files we'll be in shape to use much more sophistication.

So I guess the short answer/bottom line of my views is that you ought to focus your efforts on systems to aid in extracting metabolism information from the existing total fount rather than trying to download a specific "metabolism database". Believe me there'll be lots to do on that! E.G. the semantic network to search down pathways of conceptual linkage. Cf. the NLM's UMLS.

--

I'm working myself on a thorough analysis of the concepts entailed in the "central dogma" of

DNA (<)=> RNA ... => polypeptide sequence => folded active protein

and some of the feedback loops. There's lots to do with just the dozen central actors in that chain!

If you decide to go ahead anyhow (or in any event) let's stay in touch. I have some unorthodox ideas of how a database should function.

If you're interested in enzymes, you may be amused at the following:

===== DRAFT =====

C) \fI3) ENZYMES CATALYZE CHEMICAL REACTIONS., viz. are not consumed.\fR.
It is however difficult to find any experimental studies that would define the rate of "consumption" of enzymes, viz. their

thermal stability in presence and absence of their substrates. It is anecdotally suggested that substrates may stabilize enzymes in the course of purification (or perhaps enhance protein folding in the vector expected by the investigator.)

There is one situation where the enzyme consumption is understood:

In fact, experimentally contrived suicide substrates are designed to titrate the corresponding enzymes, and these are perforce "consumed". This is a natural function of the "alkylation adaptive" repair enzyme, O-6-Methylguanine- DNA Methyltransferase, which is stoichiometrically inactivated by each methyl group it extracts from target DNA.

cf: Lindahl-T Demple-B Robins-P
Suicide Inactivation of the Escherichia-Coli O-6-Methylguanine-
DNA Methyltransferase EMBO JOURNAL 1: 1359-1363 1982

Should that be called an enzyme?

Do you know any other natural suicide substrates? (leave out peroxidation).

Yours ,for an uneventful 1991, Josh

(1990 was too eventful for my taste; and I'm not too hopeful).

P.S. I tried to hire Peter Karp, but he wouldn't leave Calif. for long. Mick Noordewier is working with me on computer matters.