Deriving Theoretical Models in Structure Property Relations

PD Dr. U. Höweler
April 2004

The „*Real World*“ Problem

The problem, as problems tend to, completely ignores the availability of data or explanations, like:
What inhibits this enzyme?
Why is this growth factor more efficient?
Why is this ester felt more comfortable?
Why is this catalyst highly selective?

Comes with a Set of Experiments

Experiments are well established for years:
They are sufficiently reliable and the results reproducable!
But they are time-consuming and/or expensive to run!

Communicating the Problem to Theoreticians

This tedious work actually broadens/deepens the knowledge:
What does enzyme mean here? Are any inhibitors available?
What constitutes this growth factor? What is it growing?
Why are esters used and no ethers? What's about ketones?
How is selectivity determined? Does it depend on substrate?

Leads Inevitably to More Problems

Questions on the well established experiments may arise:
What do the experiments really measure? How are these phenomena related to the actual problem? Can the problem actually be split into subproblems? Is the reliability of the experiments only felt or actually measured?

Customer's Contribution Comments *CHEOPS's* Contribution

Analyzing the Problem

**All available knowledge must be compiled:
When does the problem occur? Is it a multi-step process?
What are the consequences? Which are the hypotheses on the basic physical, chemical and/or biological phenomena?**

Splits it into Subproblems

The solution to the problem may well be the validation of a hypothesis: Is there THE essential step determining the overall process? Are the steps well defined? Can specific phenomena be assigned to the steps? Can these phenomena be studied experimentally?

Understanding the Problem

All hypotheses on the essential steps of the problem are to check against up-to-date knowledge. The analysis will identify the phenomena crucial to the problem. The relevant data have to be specified to base decisions on.

Leads to Questions to be Answered

So: Which hypothesis is most promising at which step of the problem? Do experiments exist to check the inherent assumptions? Are these experiments reliable? Can they be run in-house?

Can a theoretical approach mimic this step?

Customer's Contribution Comments *CHEOPS's* Contribution

Formulating those Questions

Given the steps that define the problem are known, the right questions can be formulated: Where is the active site in the enzyme? Where can functional groups be stabilized? How is the dimer of the growth factor formed? Does the esterfunction profit from interactions with membranes? Do van der Waals interactions dominate formation of prochiral transition states?

that can be Answered

If the experimental data are closely related to the physical phenomenon and this phenomenon can be modelled, then the model result has to correspond with the experiment.

Deriving a Reliable Model

Include as many aspects of the relevant steps as possible in the model: Define the configuration space for the substrate/ enzyme complex. Establish the relative orientation of growth factor monomers to form the active dimer. Describe the structure of the ester/membrane ensemble. Construct the backbone of a substrate/catalyst transition state geometry.

Select the Theoretical Method

Molecule and ensemble properties can well be computed with the appropriate theoretical methods: molecular mechanics for interaction energies in substrate/receptor complexes or in ester/membrane ensembles together with quantum mechanics in substrate/catalyst complexes as well as „real“ thermo-dynamical data.

Customer's Contribution Comments *CHEOPS's* Contribution

Extending the Model

THE *COMPUTED* QUANTITY IS NEVER THE ONE *MEASURED*!
Find other (further computed or measured) data to include in the correlation: How well does the inhibitor pass through membranes to reach the enzyme! Are the monomers of a heterodimer equally available? Will the ester vaporize rather than rest on the membrane? Will a side reaction dominate?

to Derive a Mathematical Relation

Establish the mathematical relation that connects the model results to the experimental data. The adequate correlation method depends on the type and number of data to handle. Thus a simple linear relation may be adequate for one problem while a more complex neural network or principal component analysis are required for others.

Customer's Contribution **Comments** ***CHEOPS's* Contribution**

Validating the Model

Determine the limits of the mathematical relation: Are all aspects of the phenomenon adequately considered? Will additional data be necessary? Can exceptions be understood? Is the understanding yet incomplete? Or wrong?

to Allow for Predictions

Refine the mathematical relation: Include other data not in the initial test set. Make predictions and check the results against new measurements.

Optimize the Model Implementation

Revise the model and implementation for better performance:
Can the inhibitor/enzyme complexes be preoriented?
Can the interaction be well represented by contacts rather than energies? Can the time for computation be reduced?
Is a simpler method available? Can the procedure be better automated?

for Fast and Easy Everyday Usage

The procedure must be easy to use and as automated as possible for everyday acceptance.

Customer's Contribution Comments *CHEOPS's* Contribution

Enhance the Understanding

The results from the validated model will enhance the understanding of the phenomenon: Systematic variation of results and properties allows to establish ordering schemes.

to Devise Novel Compounds

The ordering scheme allows to develop new ideas for novel compounds with optimized properties.

And therefore even solves

the „*Real World*“ Problem

Customer's Contribution **Comments** ***CHEOPS's* Contribution**

Summary

Face „Real World“ Problem...

Formulate Adequate Questions
Check and Validate Experimental Data

Determine Phenomena to be Studied

Derive Appropriate Physical Model
Find Mathematical Correlation
Validate Correlation by Predictions

Establish Reliable Ordering Scheme to...

...Design Improved Compounds