Simulating Biochemistry: The eLABorate Project

John Garratt
Jane Tomlinson and Doug Clow
Department of Chemistry, The University of York, UK

We are interested in creating opportunities for students to practice applying their knowledge of science in a way which encourages them to solve the kinds of problems faced by experimental scientists. We accept Bodner's distinction[1] between an 'exercise' (which can be solved by a familiar approach) and a 'problem' (which involves "doing something when you don't know what to do"). Our goal is to create situations which are sufficiently clearly defined for solutions to be obtained, but which can be approached in different ways and which may even have different solutions. Much of our work is suitable for biochemists, and some of it has been specifically developed for biochemists; in this paper we will concentrate on the latter. Not all our work is IT-based. For example, we are enthusiastic about the value of original scientific papers in providing a context for discussions of the design and interpretation of scientific investigations[2,3,4]. One of the papers we have used successfully deals explicitly with a study of enzyme catalysis[5]. Other exercises designed to exercise the skills of thinking are collected in a recent book[6,7,8]. However, in this paper, we will confine ourselves to the description and discussion of computer based exercises. Our computer based materials suitable for biochemists are listed in Table 1. They were developed as part of the eLABorate project[9].

We describe our software as *simulations* to indicate that each program simulates some kind of experimental situation. We recognise that this is a very broad concept since virtually any activity set up to provide a learning experience for our students simulates some aspects of reality as closely as possible, but also simplifies reality by cutting out aspects which would otherwise be distracting. The importance of omitting aspects of reality is illustrated by a flight simulator used in training pilots. This simulates with great accuracy the perception of being in the cockpit of a real aeroplane, but the trainee-pilot does not get killed by 'crashing' the aircraft. The effectiveness of a simulation depends on the wisdom with which aspects of reality are included and excluded from the simulation.

enzymeLAB	An exercise in the design and planning of investigations of enzymes which obey Michaelis Menten kinetics
tracerLAB	An investigation of the incorporation of radioactive precursors into protein and nucleic acid during normal growth and after inhibition of protein or nucleic acid synthesis
proteinLAB	An exercise allowing the exploration of various procedures for purifying proteins or for designing a procedure for the purification of one from a mixture of 20 proteins
immunoLAB	Using antibodies for quantitative assays. This is really two packages, one which deals with radioimmuno assay, and one with elisa; both

	involve setting up the procedure for determining an unknown.
bindingLAB	A visual simulation of protein-ligand binding
statsLAB	An empirical investigation of the statistical principles of sampling from a normal distribution and of regression analysis
equilibLAB	Titrating a diprotic acid, offering tutor or student choice of the pK values

Table 1. eLABorate packages dealing with biochemical systems

We have previously suggested four areas in which simulations can be useful[4]. These are:

- as a pre-lab;
- to explore theory;
- to carry out a virtual investigation; and
- to gain experience with expensive equipment.

The last of these is typified by *nmrLAB* which allows students to control the 'settings' of an FT nmr spectrometer and so gain experience of the process of data collection and manipulation as well as data interpretation. In this paper we shall say no more about this type of simulation, but will use two simulations of biochemical systems to illustrate how simulations can fulfil any or all of the first three potential uses. The first of these, *enzymeLAB*, was originally designed as a virtual investigation; the second, *tracerLAB*, was commissioned as a pre-lab.

In deciding what experimental situation to simulate, we ask ourselves the following questions:

- Is there an individual teacher who intends to incorporate the final product into their teaching programme and who will be an active member of the development team?
- Is there a clearly defined educational objective?
- Does the computer provide the only (or the best) way of meeting the educational objective? (at the very least the computer must add a dimension to the student activity which cannot be provided in any other way - otherwise we might as well not use a computer); and
- Will the final product be useful to those teaching courses in other universities?

We have always adopted the principle that our simulations are packages not programs: the software does not stand alone, but is intended to be integrated into a coherent learning experience. Our reasoning was that students need guidance in the principles of experimental design and planning, and that this is best learned through interaction with a real person. As we discuss later, we are currently reviewing this principle in the light of changing needs, and also of our experience, some of which we present here.

enzymeLAB

"*Congratulations! You have isolated a new bacterial enzyme.*" Thus is the student using *enzymeLAB* greeted.

Almost all students of biochemistry will, during their laboratory course, follow a recipe to collect data from which they calculate V_{max} and K_m for a well characterised enzyme known to obey Michaelis-Menten kinetics. We wrote this simulation so that students could plan for themselves how to collect data for an enzyme for which they could feel some ownership. This meant providing each student with a different enzyme with realistic characteristics. *enzymeLAB* does this by selecting random combinations of basic values for K_m and V_{max} on which it superimposes randomly selected values for pH sensitivity to these parameters. All the enzymes obey Michaelis-Menten kinetics so that the rate of reaction (v) is given by the equation $v = (V_{max} * S) / (K_m + S)$. The software gives the students information about the specific activity of the enzyme and the sensitivity of the assay procedure since this information would necessarily be available to the isolator of a new enzyme. The computer also limits the amount of enzyme available (as would normally be the case for a newly isolated enzyme). The students are set the task of characterising the enzyme, which they are told involves investigating the optimum pH and the effect of pH on V_{max} and K_m. With these aims, the user selects values for the pH and the concentration of enzyme and substrate at which the rate of the reaction is to be measured; the resulting value of v which is displayed has a realistic experimental error with a standard deviation of 5% of the correctly calculated value.

Our view is that an effective investigation involves the use of the following rules of thumb:

- The K_m for most enzymes falls within the range 1uM to 1mM. It follows that: (i) at a substrate concentration of 20 mM, most enzymes are saturated ($v = V_{max}$); and (ii) measuring v at S = 20mM and S = 0.02mM will provide information from which a rough value for K_m can be determined in a further one or two measurements.
- Any assay procedure can measure rates only over a limited range; to maximise the range of values of S at which rates can be measured under identical conditions, the concentration of enzyme chosen should give a value of V_{max} close to the upper limit of the range.
- To determine K_m, it is desirable to make measurements of v at substrate concentrations ranging on either side of K_m (say from 0.2 to 4 K_m).

It soon became apparent that second year students lack the experience to apply their basic knowledge of enzymes in this way. We therefore prepared a set of self-test questions designed to bring out the key principles (examples are given in Table 2). Most students were insufficiently motivated to complete these. We therefore now arrange a pre-simulation class for which students prepare answers to the questions, and at which the tutor calls on students at random to provide answers which are then discussed by the class. With this pre-exercise guidance, students are able to complete the computer based simulated investigation without a tutor being present, thus allowing them the flexibility to choose when to spend time at the computer. This need for tutorial input illustrates our point that these simulations are 'packages not programs'.

| 1. | If you want to study the sensitivity of your enzyme to pH would you aim to work |

		at a substrate concentration which always saturates the enzyme or never saturates the enzyme?
	2.	You are told that the specific activity of the enzyme was measured at a substrate concentration of 20mM; why was this value chosen?
	3.	Devise a strategy for finding an approximate value for the concentration of S needed to achieve a rate of approximately half the maximum rate at pH 7 using not more than 4 measurements of rate.
	4.	Plan how you would like to space your values of S to achieve a desirable range of values of rate.

Table 2. Pre-exercise tasks in preparation for *enzymeLAB*

Although we originally wrote this software to allow students to carry out what we have called a 'virtual investigation'[4], it is in practice a much more versatile tool. We have, for example, used it for three years with a group of first year biochemistry students. For these students, the task they are set is simplified and is completed in a single day.

An important feature of the computer simulation is the speed at which it can generate data. Data which would take two or three weeks to collect at the bench can be generated by the computer in two to three hours. Furthermore, the data are of a quality which would be obtained by an experienced professional. This quantity and quality of data is sufficient to justify the requirement that each student writes a report in the style of a scientific paper. We have described our experience of this[10].

With hindsight, we can suggest improvements to the design of the software which would increase the flexibility of the package. For example, it would be useful to allow the tutor to vary the error function (or to remove it altogether). A particular regret is that we have not built in an option for the tutor to define the characteristics of the enzyme or to select a specific set of the parameters stored in the program. This facility would add a new dimension to the use of *enzymeLAB* in conjunction with laboratory work using a real enzyme.

Opinion varies whether the optimum combination of a simulation and laboratory work involves using the simulation first (as a preparation to understand the thinking behind the recipe in the laboratory manual) or after the benchwork (to allow the students to explore more fully the data available from enzyme kinetics when they have a good appreciation of the laboratory context). Booth has used his simulation of protein purification (*proteinLAB*) both to prepare students to purify a protein in the laboratory, and to broaden the experience of students who have already carried out a purification protocol. His conclusion[11] was that students generally felt that they would have gained more by doing the two exercises in the reverse order to the one they actually used!

tracerLAB

This simulation was written specifically as a pre-lab to support a laboratory class in which first year biochemistry students follow the uptake of ^{3}H-lysine and ^{14}C-adenine into protein and nucleic acid when bacteria grow normally and after the inhibition of protein synthesis or after uracil starvation. The software simulates two types of labelling experiment. The 'cumulative label' calculates the amount of label which accumulates in protein and nucleic acid when the labelled precursors are present throughout the incubation. In the 'pulse label' the culture is incubated without labelled precursors, samples are transferred at selected times into an aliquot of labelled precursors, and the program calculates the amount of label incorporated in one minute. A brief description of the program has been published[4].

In the laboratory class, students are required to select one of three strains of bacteria, TAUbar, BW113 and SH7. TAUbar cannot grow unless uracil is present in the incubation medium. The other two strains differ only in that SH7 has 'stringent' control over RNA synthesis (synthesis of tRNA and rRNA is inhibited when protein synthesis is inhibited), whereas BW113 has 'relaxed' control (it continues to synthesise tRNA and rRNA when protein synthesis is inhibited). The students also chose one of three ways of inhibiting growth (addition of valine to inhibit iso-leucine synthesis, addition of chloramphenicol, or limitation of uracil in the incubation medium). They also make their own decisions about the amount of radioactivity and carrier to add to the medium, and the times at which to add inhibitor and take samples for analysis. Identical options are provided by the simulation, which is built on realistic models of all three strains of bacteria.

The simulation allows them to test out their proposed protocol, to discuss the results with a tutor if necessary, and to revise their protocol as many times as necessary for them to obtain interpretable data. Common mistakes made in early protocols include adding insufficient carrier so that bacteria use up all the label before the end of the incubation (or so much that very little label is taken up), adding insufficient label (particularly early in the incubation period) so that too few counts are taken up, and failing to take sufficient samples to establish the pattern either of the uninhibited exponential growth or of the post-inhibition incorporation.

In order to minimise poor decisions of this sort, the students must enter the concentration of radioactive counts and of carrier they propose to use. Their next step is to enter the volumes of solutions of radioactive tracers and of carriers. The calculations involved are not trivial, and require the application of basic knowledge, which students should have, but have rarely used in this way. For example, they have to handle concepts like specific activity and counting efficiency in order to convert the specific activity of their stock solution of tracer into a count rate in their samples, and they need to use the expected growth rate of the bacteria together with the composition of protein and nucleic acid to calculate how much lysine and adenine will be incorporated into protein and nucleic acid during their chosen incubation period.

tracerLAB includes two features not incorporated into *enzymeLAB*. One is an animation which students can access at any time and which shows a diagrammatic representation of the procedure from the setting up of the incubation flasks, through the preparation of samples, to the counting of the radioactivity with a scintillation counter. This helps to create a bridge in the student's mind between the artificial nature of the simulation and the reality of the laboratory. The second feature is the hidden code which allows the tutor to decide whether to use the software as a pre-lab (when students choose the strain of bacterium they wish to study) or as a virtual investigation

(when students are allocated one of the three strains by the software, and are set the task of characterising it).

Conclusions

Writing simulations

Considerable investment of time is required to create an effective simulation. It is therefore important to maximise the eventual use of the software, and this means making it sufficiently flexible for tutors to adapt it to their own needs. Our approach to this is to write a simulation only on the specific request of a tutor (our expert advisor) who intends to incorporate it into the curriculum for a specific and defined purpose. This means that we can be confident that it will be used in at least one institution. However, we also satisfy ourselves that the subject of the simulation is of sufficiently general interest that it will be found on the syllabus of most comparable courses. The expert advisor must provide us with a realistic model on which to base the simulation, though experience shows that we can have a crucial input (as non-experts) at this stage, either because our own unfamiliarity with the system can lead us to ask 'naïve' questions which have been overlooked, or because our familiarity with the capability of the computer can help to determine whether a particular approach is too complex or could be improved. Another important role we play in the development process is to use our experience from other simulations to suggest ways in which the program can be made more flexible; we have described above how we increased the flexibility of *tracerLAB* beyond its originally specified role as a pre-lab exercise.

We aim to use what we have called a 'functional' rather than a 'flashy' interface; in our view too much software looks as though it has been written by computer enthusiasts wanting to make the maximum use of multimedia technology, without considering carefully enough the educational advantages (or more often disadvantages) of doing so[12]. We acknowledge the usefulness of high quality videos for some specific purposes, but we believe that well designed diagrammatic representations are in many cases clearer and more useful.

Student feedback and reflection

Much of the feedback we receive from students is positive. However, we have experienced two different kinds of negative feedback. One comes from analysis of the results of tests or questionnaires completed by students before and after completing a simulation. As an example, we have reported[13] the data obtained from a group of first year biochemistry students who used *enzymeLAB*. The two-part questionnaire tested student knowledge of basic aspects of enzyme kinetics, and recorded their self-assessed confidence in their ability to apply this knowledge. We found little evidence that any increase in knowledge or confidence persisted six weeks after completing the exercise. However, we regard this as understandable; we should not expect a single three-hour exercise to have a huge effect on student learning, especially since most students' primary aim is (understandably) to obtain the maximum mark for their degree for minimum work. This encourages them to look forwards (to a new piece of assessed work) rather than backwards to a completed piece of work in order to consolidate their understanding of it. Our view is that this demonstrates the importance of providing students with a formal

opportunity to reflect on the lessons learnt through a particular exercise. This is rarely if ever achieved simply by handing back a marked script, but requires a formally timetabled class in which students discuss lessons and conclusions.

The second type of negative feedback (which is rare for *enzymeLAB* and *tracerLAB*) is the view of some students that the simulations do not provide an agreeable or effective learning experience. These subjective views are, of course, highly coloured by student expectations. We believe it is significant that when a majority of students have negative views about their experience (as is usually the case with *statsLAB*), the minority who find it a valuable experience almost always comment that they were made to *think*. Our conclusion is that most students have come to regard biochemistry (or chemistry) as a subject in which correct answers to all relevant questions are known, and that the student's job is to find the correct answer. They cannot treat the simulations as an exercise (in Bodner's sense) since they have to apply their knowledge in a flexible and creative way. Many find this difficult enough to be demotivating. The solution to this must be careful preparation to ensure that they understand both how to approach the exercise and why it is an important aspect of their learning experience.

Thus our conclusion is that the effectiveness of a simulation depends not only on how well it is written in the first place, but on the individual input from the tutor both in preparing the students appropriately and in debriefing them appropriately.

Distance learning

Much computer based learning material is designed to provide a complete and self-contained learning experience for the student. In contrast to this approach, we have designed our simulations on the assumption that students will have direct contact with a tutor. We recognise that this is a limitation in so far as the assumption makes the simulations less than ideal for incorporating into programmes of distance (or fully independent) learning. We have therefore given some thought to ways in which the need for direct input from the tutor can be removed, or at least minimised. The difficulty of doing this stems from the fact that most of our simulations require the application of knowledge, rather than its acquisition.

Consider, for example, the pre-exercise tasks used as part of the preparation for using *enzymeLAB* shown in Table 2. Direct answers to these questions cannot be found in textbooks; indeed they are not the kind of question to which a single correct answer can be given, although some answers could be classed as 'better' than others. They are well suited to our current procedure of setting them for students, and then *discussing* their answers in class. It would be possible to devise material (either paper-based or computer based) which would provide support for the motivated distance-learner. But the application of knowledge is a skill which we believe involves not just 'learning by doing' but also at best it involves direct interaction with experienced practitioners. If this were not the case, there would be little point in meetings and conferences at which practicing scientists exchange ideas. We remain of the opinion that a good tutor will always be better than either a book or a computer program. As Boothroyd put it[14], in most of our teaching we invert the intellectual hierarchy of knowledge, understanding, and skill in application by spending our valuable contact time covering content (which can be learned in a variety of other ways), and have too little time left to deal adequately with the higher order skills

involved in the application of knowledge. Our aim must be, not to replace the tutor, but to create conditions in which tutorial input can be saved for the kind of learning which is the most intellectually demanding.

References

1. Bodner, G. M. and Daniel (2000) *U.Chem.Ed.* **4** (in press).
2. Garratt, J. and Aitken, M. (1994) *Biochem.Ed.* **22**, 16.
3. Garratt, J. and Overton, T. (1996) *Educ.Chem.* **33**, 137-138.
4. Garratt, J. (1997) *U.Chem.Ed.* **1**, 19-27.
5. Garratt, J. and Harrison, D. (1970) FEBS letters **11**, 17-19.
6. Garratt, J. and Overton, T. (1997) *Educ.Chem.* **34**, 79-80.
7. Overton, T. (1997) *U.Chem.Ed.* **1**, 28-30.
8. Garratt, J., Overton, T. and Threlfall, T. (1999) *A Question of Chemistry*, Longmans.
9. The eLABorate project was funded by the University Funding Councils of England, Scotland, Wales and Northern Ireland under the Teaching and Learning Technology Project.
10. Clow, D. and Garratt, J. (1998) *U.Chem.Ed.* **2**, 51-54.
11. Booth A personal communication.
12. Tomlinson, J., Clow, D., Garratt, J. and Hodgson, A. (1996) *Software Reviews* **13**, 8-10.
13. Garratt, J., Hardy, S. and Tomlinson, J. (2000) *U.Chem.Ed.* **4**, (in press).
14. Boothroyd, C. (1993) *Proceedings, Variety in Chemistry Teaching*, ed. M. Aitken, Royal Society of Chemistry.

John Garratt, Jane Tomlinson and Doug Clow
Department of Chemistry
The University of York
York YO10 5DD
United Kingdom
cjg2@york.ac.uk