An artificial-intelligence technique for qualitatively deriving enzyme kinetic mechanisms from initial-velocity measurements and its application to hexokinase

Lillian GARFINKEL,* Dawn M. COHEN,† Von-Wun SOO,†‡ David GARFINKEL*§
and Casimir A. KULIKOWSKI†

*Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.,
and †Department of Computer Science, Rutgers University, New Brunswick, NJ 08903, U.S.A.

We have developed a computer method based on artificial-intelligence techniques for qualitatively analysing steady-state initial-velocity enzyme kinetic data. We have applied our system to experiments on hexokinase from a variety of sources: yeast, ascites and muscle. Our system accepts qualitative stylized descriptions of experimental data, infers constraints from the observed data behaviour and then compares the experimentally inferred constraints with corresponding theoretical model-based constraints. It is desirable to have large data sets which include the results of a variety of experiments. Human intervention is needed to interpret non-kinetic information, differences in conditions, etc. Different strategies were used by the several experimenters whose data was studied to formulate mechanisms for their enzyme preparations, including different methods (product inhibitors or alternate substrates), different experimental protocols (monitoring enzyme activity differently), or different experimental conditions (temperature, pH or ionic strength). The different ordered and rapid-equilibrium mechanisms proposed by these experimenters were generally consistent with their data. On comparing the constraints derived from the several experimental data sets, they are found to be in much less disagreement than the mechanisms published, and some of the disagreement can be ascribed to different experimental conditions (especially ionic strength).

INTRODUCTION

Over the years a traditional method of deriving enzyme kinetic mechanisms from initial-velocity data has evolved (Cleland, 1963a,b,c; Wong, 1974; Segel, 1975; Cornish-Bowden, 1976). This method is based on the qualitative patterns observed when the data are plotted with (most commonly) the Lineweaver–Burk double-reciprocal plot, specifically whether the data points define a set of straight lines which may be convergent or parallel. More recently these techniques have been supplemented by examining data on the rate of exchange of radioactive substrates and products at equilibrium, as well as non-kinetic information such as amino acid sequence and the three-dimensional structure of enzymes. Although these methods have been extensively applied, we remain in the situation described by Dixon & Webb (1979): “At the moment it is probably not too much to say that there is no enzyme whose exact mode of action can be taken as proven”.

In a recent database study of hexokinase, a moderately simple enzyme which has been extensively studied, we found that consensus on its mechanism had not yet been reached (Garfinkel *et al.*, 1987). Hexokinase has two substrates and two products, and the mammalian enzyme is strongly inhibited by one product, namely glucose 6-phosphate ($K_i = 26 \mu\text{M}$). Whether this inhibition is allosteric or not is a matter of disagreement (Solheim &

Fromm, 1981; Ureta *et al.*, 1985). However, most data sources show this inhibition to be linear, even though the inhibitory glucose 6-phosphate site may well differ from the product site. For our database we examined experiments on hexokinase from several sources, including skeletal muscle, ascites and yeast. There is considerably more information for the yeast enzyme than the mammalian enzyme, mostly because it is very difficult to study the mammalian enzyme in the reverse direction. The kinetic mechanism appears to be sequential, i.e. a central ternary complex is formed, but there is disagreement about whether glucose and MgATP^{2-} add to hexokinase in an ordered or random manner. Hexokinase appears, from recent X-ray-diffraction studies, to undergo a change in physical conformation which is described as ‘hinge bending’ or ‘hinge closing’, when the substrates bind (Bennett & Steitz, 1980a,b). The theoretical consequences of such a change have not yet found their way into the literature we examined, partly because most kinetic studies are older than these structural studies.

Some of this enzyme’s properties make it difficult to interpret the experiments included in our database. Many preparations of this enzyme are unstable, and their properties change with time. Mayer *et al.* (1966) found that the apparent inhibition constant of glucose 6-phosphate changed over several orders of magnitude with time as the enzyme from dog heart aged! Grossbard

‡ Present address: Institute of Computer Sciences, National Tsing Hua University, Hsinchu, Taiwan 30043.

§ To whom correspondence and reprint requests should be sent, at the following address: University of Pennsylvania, Department of Computer and Information Science, Moore School of Electrical Engineering, Philadelphia PA 19104-6389, U.S.A.

& Schimke (1966) noted the instability of hexokinase II and discussed its possible biological functioning. Nevertheless, at least part of the disagreement about the kinetic mechanism stems from difficulties with interpreting the kinetic data. We have studied the data of Fromm & Zewe (1962), Toews (1966), Kosow & Rose (1968), Noat *et al.* (1968), DelaFuente & Sols (1970) and Bachelard *et al.* (1971). These publications were selected because they presented large amounts of experimental data graphically, so that it was reasonable to test whether the data fitted any of the models in our model base qualitatively. Since these publications are too few to define unequivocally the hexokinase mechanism, the present paper should be considered as a description of our method as applied to hexokinase. Preliminary or previous reports of this work have been presented (Soo *et al.*, 1987; 1988).

METHODS

Here we describe a computer method for deriving a kinetic mechanism from initial-velocity data and illustrate it with data on hexokinase from yeast and mammals. This qualitative method is based on stylized descriptions of patterns of data and the traditional rules given by Cleland (1963*a,b,c*), Wong (1974) and Segel (1975). The rules as well as a library of kinetic mechanisms have been encoded in the PROLOG language. Constraints are inferred from the characteristics of the Lineweaver-Burk (or possibly other) plots and are then checked with corresponding ones derived from possible theoretical models. A statement of the constraint and whether it fits the candidate model is printed out. An unsatisfactory model is eliminated (**filtrate**), while a satisfactory model is retained for further consideration (**residue**). The system can derive all possible theoretical constraints for a model, although this is not done routinely. Unless all the theoretical constraints are known, one cannot immediately conclude from a match between data and model that a definite identification has been achieved. This method requires at least a minimal degree of data completeness. A published set of data is not necessarily extensive enough for this purpose. Data sets may be edited down to fit limited space or experimenters may investigate only one feature of a mechanism. Different experimenters may perform different experiments when studying the same problem and consequently not obtain the same results. We have prepared programs for comparing experiments reported in different publications.

The techniques developed here can be used to identify sets of **experimental data** which may not be complete enough to test all the attributes of a model and pinpoint missing data. However, it may be difficult to perform all recommended experiments because reagents are expensive or assay components are poorly soluble. This method is completely distinct from methods involving non-linear regression, where numerical data are fitted to models (see, e.g., Mannervik, 1975), and for which we have distributed a program PENNZYME (Kohn *et al.*, 1979; Schremmer *et al.*, 1984; Maclay & Garfinkel, 1986). We do not yet know how to use the combination of qualitative and quantitative methods most efficiently. Perhaps the best way to combine the two methods is to use the qualitative techniques to narrow the range of models to be tested quantitatively.

DESCRIPTION OF THE PROTOTYPE SYSTEM

Our prototype computer system is programmed to compare experimentally based constraints with corresponding model-based constraints for sets of initial-velocity data. This system is considered a prototype because it is still limited in the variety of kinetic mechanisms and types of enzyme kinetic data it can handle and needs additional capabilities. It has not been applied to specific enzymes other than hexokinase. The system requires expansion and refinement of its rule base and model base before it can be broadly distributed. A substantial programming effort is needed to make the system 'user friendly.'

This system contains the following basic features: a model base, a rule base for interpreting the results of initial-velocity experiments, a frame representation of the experimental data (Abstract Data Language or ADL), and a means for reasoning about the results (the inference engine in the PROLOG programming language). Given a data set that is extensive enough for this type of analysis, one performs the following operations:

Data-driven operations

- Specify prior knowledge; identify the experimental objects as substrates, products, or inhibitors
- Enter data into ADL; describe results of reciprocal plots according to predefined descriptors
- Infer data constraints; interpret results using rule base

Theory-driven operations

- Perform qualitative simulation; derive all the theoretical constraints from a PROLOG representation of the enzyme mechanism
- Impose experimental conditions on a model: derive the constraints resulting from experimental conditions such as saturation with a substrate or the presence or absence of a product.

Compare experimental and theoretical constraints

- Perform knowledge filtering; check if inferred constraints agree with those derived from mechanisms in the model base
- Compare different experiments; check the results of current analysis with those from other studies

An additional step in the process is to postulate discriminating experiments, experiments that have the ability to distinguish between rival models, which has not yet been routinely applied. We have a preliminary program using set-partitioning methods (Soo, 1987), which aids in designing experiments for choosing among rival models.

This system, which is written in Quintus PROLOG, runs on a Sun 4 computer located at Rutgers University.

(1) Prior knowledge specification

It is necessary to identify the experimental objects in a reaction, i.e., the substrates, products, inhibitors and possible alternative reactants. The function of each chemical must be specified; the computer term for this is **object specification**. For hexokinase the normal substrates are glucose and MgATP^{2-} , and the normal products glucose 6-phosphate and MgADP^- . It is possible to specify which alternate compounds are similar to the normal substrates or products. Thus anhydroglucitol 6-phosphate is similar to the normal product of hexokinase

and is so specified in analysing the results of Kosow & Rose (1968). Mannose, an analogue of glucose, is used as an alternative substrate in some of the experiments of Noat *et al.* (1968), but functions as a competitive inhibitor of glucose in the studies of Kosow & Rose (1968) and Fromm & Zewe (1962). The difference lies in whether both sugars are present simultaneously and whether glucose 6-phosphate or MgADP⁻ is monitored. Kosow & Rose (1968) used mannose as an inhibitor of glucose because they were monitoring glucose 6-phosphate production, whereas Noat *et al.* (1968) were monitoring MgADP⁻ production and measuring glucose (+ mannose) phosphorylation.

(2) ADL language and its application in reasoning

We have implemented in PROLOG an Abstract Data Language, ADL. ADL is a frame-like representation for describing the essential components of kinetic experiments with enzymes. In artificial intelligence a frame is a template which contains a structured description of the named objects, in this case an organized list of attribute and value pairs (or 'slots') that can capture or represent the useful concept or property of the system described.

In ADL a simple experiment has a single control variable and a single observed variable and produces only one curve. A **simple compound** experiment has two controlled variables (one primary and the other secondary) and usually produces a family of curves. PROLOG, like English, handles entities which are separated by spaces. Capitalized names represent experimental variables, whereas lower-case names represent constants; phrases are grouped with parentheses. When an entity involves several words, they are connected by underscores, e.g. **lineweaver_burk**. The familiar example of a competitive inhibition of substrate S by inhibitor I is described below in PROLOG (column 2):

Segel, 1975) and from Cleland (1963*c*, 1970). The basis of these rules is the effect that different ligands have on the slopes and intercepts of the double-reciprocal plots. There are now a total of 34 rules: 11 initial-velocity rules, six product-inhibition rules, five dead-end inhibition rules, four replot rules and eight rules that were derived from Wong's (1974) book. All of these rules are concerned with binding properties of ligands involved in enzymic reactions. These rules generally have the form: 'if a body of data having certain characteristics matches an expression for shape or intercept of the Lineweaver-Burk plot and if substrate x or substrate y is involved, then a conclusion about interaction between ligands or which enzyme form the ligands binds to follows'. This current set of 34 rules is by no means complete, because they are derived from an enzyme-kinetics literature which has neglected areas and may therefore be regarded as incomplete also. The system has recently been programmed to handle the rapid-equilibrium random reactions that were postulated in our application examples.

This set of rules has predictive ability. Although it was derived primarily from, and applied primarily to, examples involving hexokinase (i.e., situations involving Bi Bi mechanisms), we were able to predict in detail the behaviour of the (Bi Bi Uni Uni Ping Pong) Ter Ter mechanism example discussed on pp. 760-764 in Segel (1975). In this example 37 different kinetic results (states) are observed when different experimental conditions are imposed on the mechanism. These conditions are the presence of substrates under saturating or unsaturating conditions or in the absence or presence of product inhibitors. This shows that the rules are correct and being applied appropriately.

Our rules are a generalization of the conventional rules describing the effects of ligands on the slopes and intercepts of the double-reciprocal lines. They rely heavily

ADL SLOT	DESCRIPTION	EXPLANATION
Name:	data(Example_1,	Any identifier may be used.
Type:	simple_compound,	A family of curves is presented.
Experimental Conditions:	c(varying,S)& c(fixed,I),	The substrate is the primary variable; the inhibitor is the secondary variable.
Observed Variable:	v,	The velocity is the observed variable.
Plotting Method:	lineweaver_burk,	The data are presented as double-reciprocal plots.
Results:	r(shape,linear)& r(y_intercept, invariant))	The curves are linear and they converge on the vertical (y) axis.

(3) Rule base

The data given in the ADL representation are then used to infer constraints using the rule base. We have derived a set of rules for analysing initial-velocity enzyme kinetic data from a number of textbooks (Wong, 1974;

on the notion of ligands reversibly connected. This concept was introduced by Cleland (1963*c*), who defined an irreversible step as addition of a substrate at infinite concentration of that substrate, or the release of a product when it is present at zero concentration. A

reversible step in an enzyme mechanism therefore occurs in the presence of a product or of a substrate at unsaturating concentration. Two ligands are reversibly connected if no irreversible step occurs between the steps in the mechanism where the two ligands add to the enzyme. Effectively a reversible connection between two ligands means that there is a non-competitive inhibition or that two substrates form a ternary complex with the enzyme.

Reversible and irreversible connections are effectively defined according to the effects of a ligand on the double-reciprocal plots. A slope plus an intercept effect indicates a reversible connection, and an intercept effect an irreversible connection. A slope effect implies that two ligands react with the same enzyme form, or if they react with more than the one enzyme form, the two forms are reversibly connected. These rules are logically correct, although they may be difficult to explain.

Two example rules for analysing initial-velocity data are presented below. The first (**rule_i5a**) concerns the conditions where two substrates are reversibly connected, i.e. they both add to the enzyme before a product is released. Since it may be difficult to decide if the double-reciprocal lines intersect or not, this rule deals with the case where the two substrates are added in constant ratio. If the resulting curve is a parabola, the two substrates are reversibly connected and there is no product-releasing step between the addition of the two substrates. This rule in PROLOG is:

```
rule_i5a:
IF      data(Experiment_name, simple_compound,
             c(varying,S1)&c(S2/S1,constant),v,Method,Results)
AND     match_expression(Method,Results,lineweaver_burk,
                          r(shape,parabola))
AND     substrate(S1)
AND     substrate(S2)
THEN    reversibly_connected(S1,S2,c(varying,S1)&c(S2/S1,constant))
```

The second (**rule_p1**) deals with linear double-reciprocal plots which intersect at the same point on the $1/v$ axis. The PROLOG statement is:

```
rule_p1:
IF      data(Experiment_name, simple_compound,
             c(varying,S)&c(fixed,P),v,Method,Results)
AND     match_expression(Method,Results,lineweaver_burk,
                          r(shape,linear)&r(y_intercept,invariant))
AND     substrate(S)
AND     product(P)
THEN    react_with_the_same_enzyme_form(S,P,c(varying,S)&c(fixed,P))
```

The above rule describes the conditions where competitive inhibition occurs and only a slope effect is observed. The rule states that, when initial velocities are plotted as double-reciprocal plots and a substrate and a product have a common intercept on the abscissa when the substrate is varied at different constant concentrations of the product, then both the substrate and product react with the same enzyme form.

(4) Model base

The model base is primarily a set of statements in PROLOG of the chemical equations describing the enzyme kinetic mechanisms with which we are working. Basically it is a translation of the models given by Cleland (1963a), Wong (1974) and Segel (1975), which is

still evolving. There are currently 22 models in the model base, of which five are derived from Cleland (1963a), eight are derived from Wong (1974) and nine from Segel (1975). These represent mechanisms involving two or three enzyme substrates and products. Two typical enzyme mechanisms are given below. The first has been named ordered Bi Bi by Cleland (1963a) and represents a mechanism where two substrates add to the enzyme and form a central complex before the products are released in an ordered manner. The second mechanism described is the Ping Pong Bi Bi or substituted enzyme mechanism, where the addition of the first substrate to the enzyme is followed by the release of the first product before the second substrate adds to the enzyme. The substance of each model is familiar from the enzyme-kinetics literature. Its statement in PROLOG may be slightly unfamiliar, but should be intelligible, to the person knowledgeable in enzyme kinetics.

```
model_name(cle_1,ordered_bi_bi).
model_event(cle_1,bind(e,a,e*a),k1,k2).
model_event(cle_1,bind(e*a,b,e*a*b),k3,k4).
model_event(cle_1,isomerize(e*a*b,e*p*q),
             k5,k6).
model_event(cle_1,release(e*p*q,p,e*q),
             k7,k8).
model_event(cle_1,release(e*q,q,e),
             k9,k10).
```

```
model_name(cle_5,ping_pong_bi_bi).
model_event(cle_5,bind(e,a,e*a),k1,k2).
model_event(cle_5,isomerize(e*a,f*p),
             k3,k4).
model_event(cle_5,release(f*p,p,f),k5,k6).
model_event(cle_5,bind(f,b,f*b),k7,k8).
model_event(cle_5,isomerize(f*b,e*q),
             k9,k10).
model_event(cle_5,release(e*q,q,e),
             k11,k12).
```

The first PROLOG sentence of the first model states that the name of the model is **cle_1**, and that it is the

ordered Bi Bi mechanism described by Cleland (1963*a,b,c*). The second sentence describes a binding event where substrate **a** binds to an enzyme **e** to form complex **e*a** with forward rate constant **k1** and reverse rate constant **k2**. Subsequent events in this mechanism are similarly described. The rate constants are not used at this time. They are present mainly to facilitate connection to future quantitative work that may follow the qualitative analysis of the data.

The rapid-equilibrium random Bi Bi reaction is a special case. Although the rate laws for these mechanisms are convenient to derive algebraically, they were difficult to implement in our system, where the mechanism-based constraints are derived from the connections among the various enzyme intermediates. According to Cleland (1970), the rules for deriving constraints for such mechanisms differ from those for the ordinary steady-state mechanisms. We have developed rules that give the proper constraints for the ordinary rapid-equilibrium random Bi Bi mechanism as well as for mechanisms with dead-end complexes. However, the system cannot distinguish between rapid equilibrium and steady-state random Bi Bi mechanisms when dead-end complexes are formed. This is also a practical problem; Cleland (1963*a*) originally thought that steady-state random mechanisms would give non-linear double-reciprocal plots, but reported in Gulbinsky & Cleland (1968) that the rapid-equilibrium mechanism gave a good approximation to the kinetics of *Escherichia coli* galactokinase, even though the mechanism was not strictly rapid equilibrium.

(5) Comparisons

The system also contains algorithms for comparing the results of related subexperiments from different

are in disagreement. We consider the ability to compare different types of experiments and characterize how they differ an important feature of the system. Much of the debate regarding the mechanism of hexokinase appears to be about experiments that are non-comparable; this is further discussed below.

It is legitimate to compare the properties of yeast and mammalian hexokinases if one keeps in mind the similarities and differences between yeast and mammalian cells. The cells from both species are eukaryotic, so that the intracellular characteristics should be similar. Although inhibition by glucose 6-phosphate is considered to be a special property of mammalian hexokinases, it is primarily a quantitative property. We have not found many qualitative differences between yeast and mammalian hexokinases. The other kinetic properties are probably similar. Much the same mechanisms have been proposed by those studying the enzyme from both types of sources.

APPLICATION EXAMPLES

We have applied this method of analysis to six sets of real experimental data on hexokinase taken from Toews (1966), Kosow & Rose (1968), Fromm & Zewe (1962), Noat *et al.* (1968), Bachelard *et al.* (1971) and DelaFuente & Sols (1970). Each of these hexokinase reports involve 6–12 distinct experiments, which help to characterize the mechanism. The published results are represented in ADL, the abstract language described above. An ADL description of an experiment where glucose concentration is varied at several fixed concentrations of MgATP^{2-} follows. The data were taken from Fig. 1 of Fromm & Zewe (1962):

ADL Entry

```
data(fromm_62_1,
simple_compound,
c(varying,Glu) & c(fixed,Atp),
v,
lineweaver_burk,
r(shape,linear)
& r(x_intercept,invariant))
```

Description

```
Experimental Identification
Type of Experiment
Experimental variables
Observed variable
Plotting method
Description of experimental
results.
```

publications (presented as double-reciprocal plots). Types of subexperiments are defined differently according to what is being compared. For **similar experiments** the experimental objects and conditions are theoretically the same; substrate or product analogues can be defined as similar to the normal substrates and products. Results of similar experiments can be **congruent** if in agreement or **incongruent** if different. **Incongruent experiments** can also be **contradictory** if their results lead to constraints that disagree with each other. Incongruent experiments may have small differences in the exact form of the double-reciprocal lines that would not lead to qualitatively different conclusions. For **non-comparable experiments** the objects and conditions differ sufficiently that similar results are not expected. The algorithms for comparative analysis point out the ways in which results from two publications disagree with each other and indicate where additional experiments are needed to clarify points that

These data and the results of our analysis can best be described by considering how the results from different publications match. For these comparisons, intersecting double-reciprocal lines which do not correspond to competitive or uncompetitive inhibitions are all considered to be non-competitive according to the definitions of Cleland (1963*a*). Other workers use different definitions.

(1) Similar experiments

The results of running the comparative analysis algorithms are presented in Tables 1–4. Similar experiments are listed in Table 1 for mammalian and yeast hexokinase, congruent experiments in Table 2, and incongruent experiments in Table 3. A sample of the output showing that the constraints inferred from the data of Fromm & Zewe (1962) fit the rapid-equilibrium random Bi Bi reaction, but not the ordered Bi Bi or Ping Pong Bi Bi mechanisms given in Table 4. Most of the experiments reported in the different sources are similar,

Table 1. Similar experiments of different experimenters

Inhibitors or alternative substrates are indicated by superscripts as follows: ¹Fructose; ²MgITP; ³mannose, ⁴MnATP, ⁵anhydroglucitol 6-phosphate; ⁶2-deoxyglucose; ⁷N-acetylglucosamine; ⁸lyxose,

Description of experiment:		Ref. ...	Corresponding no. of Figure in reference cited				
			Mammalian hexokinase			Yeast hexokinase	
			Toews (1966)	Kosow & Rose (1968)	Bachelard <i>et al.</i> (1971)	Fromm & Zewe (1962)	Noat <i>et al.</i> (1968)
Varied substrate	Fixed substrate						DelaFuente & Sols (1970)
Glucose	MgATP		1	1	1(a)	1	1(a)
MgATP	Glucose		2	2	1(b)	2	1(b)
							4(a) ¹
							4(b) ³
							2(b) ²
							3(b) ⁴
	Fixed inhibitor						
Glucose	Glucose 6-phosphate		4	8 ⁵		8	5(a)
							5(b)
MgATP	Glucose 6-phosphate		5	5 ⁵		9	5(c)
				10 ²			
Glucose	MgADP		7	6		6	5(d)
MgATP	MgADP		8	3		7	5(e)
MgATP	Inhibitory sugars			9 ³	2(b) ⁶	5 ³	
					3(b) ⁷		
Glucose	Inhibitory sugars				2(a) ⁶	4 ³	5(b) ⁸
					3(a) ⁷		
Glucose	AMP			4	4		
MgATP	AMP			7	5		

and the vast majority of the similar experiments are congruent. Since parallel double-reciprocal lines are not observed (uncompetitive inhibition or substituted enzyme mechanisms), not every possible constraint can be inferred for any given set of similar experiments. When the results are viewed in this way, there are considerably fewer differences in reported kinetic results than there are in the mechanisms proposed to explain a single set of experiments. All of these data sets indicate that a central ternary complex is formed when both substrates add to hexokinase. Most of the inhibitions by product inhibitors, substrate analogues and other inhibitors are non-competitive. All reports find that inhibitory sugars inhibit competitively with respect to glucose and non-competitively with respect to MgATP²⁻. Inhibition by MgADP⁻ against glucose is non-competitive in all of these data sets. Only Noat *et al.* (1968) report that glucose 6-phosphate inhibition is competitive with respect to glucose. DelaFuente & Sols (1970) had expected this inhibition to be competitive, and they presented data to show that they could detect competitive inhibition by using the inhibitory substrate analogue lyxose. Only Kosow & Rose (1968) reported that anhydroglucitol 6-phosphate was competitive with respect to MgATP²⁻. Only Toews (1966) and Fromm & Zewe (1962) report that MgADP⁻ is a competitive inhibitor against MgATP²⁻. Some workers also reported the results of a number of dissimilar experiments which were designed to test enzyme properties that are beyond the scope of the present system. They are therefore not discussed further.

(2) Mechanism determination

We have tested the validity of the mechanism proposed by the several authors by comparing the experimental (data-driven) constraints with the model based (theory-driven) constraints. Unfortunately presenting all of these results would take too much space. Although the data are generally consistent with the enzyme mechanisms proposed by the authors, it is not clear to what extent they have considered alternate mechanisms. It is possible to suggest additional experiments for almost all the work considered.

The mechanisms for hexokinase determined by the experimenters whose work is studied here are:

- Toews (1966): ordered Bi Bi; A = MgATP²⁻, B = Glu, P = G6P, and Q = MgADP⁻
- Kosow & Rose (1968): rapid-equilibrium Bi Bi with A = Glu, B = MgATP²⁻, P = G6P, and Q = MgADP⁻; various dead-end complexes are formed with inhibitors
- Bachelard *et al.* (1971): rapid-equilibration Bi Bi with A = Glu, B = MgATP²⁻, P = G6P and Q = MgADP⁻; various dead-end complexes are formed with inhibitors
- Fromm & Zewe (1962): rapid-equilibrium Bi Bi with A = Glu, B = MgATP²⁻, P = G6P and Q = MgADP⁻; various dead-end complexes are formed with inhibitors
- Noat *et al.* (1968): ordered Bi Bi with EAP dead-end complex; A = Glu, B = MgATP²⁻, P = MgADP⁻, Q = G6P

Table 2. Congruent experiments yielding results compatible with those of Toews (1966)

Numbers given in column 3 correspond to Figure numbers in the cited references as closely as possible. Key to references: t, Toews (1966); kr, Kosow & Rose (1968); f, Fromm & Zewe (1962); b, Bachelard *et al.* (1971); n, Noat *et al.* (1968); d, DelaFuente & Sols (1970).

Description of experiment:		Congruent expts. (reference and figure no.)	Observation/ inferred constraint
Varied substrate	Fixed substrate		
MgATP or analogues	Glucose	t1, kr1, f1, b1a, n1a, d1a, d2a, d3a	Reciprocal lines intersect/reversibly_connected (glucose, mgatp)
Glucose or analogues	MgATP	t2, kr2, f2, b1a, n1b, n4a, n4b, d1b, d2b, d3b	Reciprocal lines intersect/reversibly_connected (glucose, mgatp)
	Fixed inhibitor		
Glucose or analogues	Glucose 6-phosphate or analogues	t4, kr8, f8, d5a (all but n5ab)	Non-competitive inhibition/react_different_forms (glucose, glucose 6-phosphate)
MgATP or analogues	Glucose 6-phosphate or analogues	t5, kr10, f9 (all but kr5)	Non-competitive inhibition/react_different_forms (mgatp, glucose 6-phosphate)
Glucose	MgATP	t7, kr5, f6, n5d (all data)	Non-competitive inhibition/react_different_forms (glucose, mgadp)
MgATP	MgADP	t8, f7	Competitive inhibition/react_same_forms (mgatp, mgadp)
MgATP	Inhibitory sugars	kr9, f5, b2b, b3b	Non-competitive inhibition/ react_different_forms (mgatp, sugars)
Glucose	Inhibitory sugars	f4, b2a, b3a, d5b	Competitive inhibition/react_same_forms (glucose, sugars)
Glucose	AMP	kr4, f4, b5	Non-competitive inhibition/react_different_forms (glucose, amp)
MgATP	AMP	kr7, b5, f5	Competitive inhibition/react_same_forms (mgatp, mgamp)

Table 3. Incongruent experiments yielding results which contradict those of Toews (1966)

Numbers correspond to Figure numbers in the cited references (see Table 2) as closely as possible.

Description of experiment:		Congruent expts. (reference and Fig. no.)	Observation/ inferred constraints
Varied substrate	Fixed inhibitor		
Glucose	Glucose 6-phosphate	t4, n5a, 5b	Competitive inhibition/react_same_forms (glucose, glucose 6-phosphate) (Noat <i>et al.</i> , 1968)
MgATP or analogue	Glucose 6-phosphate or analogue	t5, kr5	Non-competitive inhibition/react_different_forms (mgatp, glucose 6-phosphate) (Kosow & Rose, 1968)
MgATP	MgADP	t8, kr5, n5e	Non-competitive inhibition/react different forms (mgatp, mgadp) (Noat <i>et al.</i> , 1968; Kosow & Rose, 1968)

Table 4. Results of constraint satisfaction

```
/* TEST OF EXPERIMENTAL RESULTS FOR AGREEMENT WITH THE RAPID EQUILIBRIUM
RANDOM BI BI MECHANISM (Named Cle_4_eq in the system) */
```

The complete results for cle_4_eq are shown. Only results that differ from cle_4_eq are shown for other models. Results for second object assignment (MgATP = a, Glucose = b, Glucose-6-P = p, MgADP = q) are not shown.

```
/* IDENTIFICATION OF LIGANDS ACCORDING THEIR ROLE IN THEORETICAL MODELS */
/* First object assignment : Glucose = a , MgATP = b, MgADP = p,
   Glucose-6-P = q, Mannose = il(competitive inhibitor of a), Mannose-6-P =
   product similar to Glucose-6-P (= q) */
```

```
/* Rapid Equilibrium random bi-bi */
| ?- satisfy(fromm, cle_4_eq).
reversibly_connected(glu,atp,c(varying,glu)&c(fixed,atp))
                                     satisfies the model cle_4_eq
reversibly_connected(atp,glu,c(varying,atp)&c(fixed,glu))
                                     satisfies the model cle_4_eq
react_with_the_same_enzyme_form(atp,adp,c(varying,atp)&c(fixed,adp))
                                     satisfies the model cle_4_eq
react_with_the_same_enzyme_form(glu,adp,c(varying,glu)&c(fixed,adp))
   or reversibly_connected(glu,adp,c(varying,glu)&c(fixed,adp))
                                     satisfies the model cle_4_eq
react_with_the_same_enzyme_form(atp,adp,c(varying,atp)&c(fixed,adp))
   or reversibly_connected(atp,adp,c(varying,atp)&c(fixed,adp))
                                     satisfies the model cle_4_eq
react_with_the_same_enzyme_form(glu,mannose_6_p,c(varying,glu)&c(fixed,
mannose_6_p)) or reversibly_connected(glu,mannose_6_p,c(varying,glu)&
c(fixed,mannose_6_p))
                                     satisfies the model cle_4_eq
react_with_the_same_enzyme_form(atp,mannose_6_p,c(varying,atp)&c(fixed,
mannose_6_p)) or reversibly_connected(atp,mannose_6_p,c(varying,
atp)&c(fixed,mannose_6_p))
                                     satisfies the model cle_4_eq
react_with_the_same_enzyme_form(mannose,glu,c(varying,glu)&c(fixed,mannose))
                                     satisfies the model cle_4_eq
react_with_different_enzyme_forms(atp,mannose,c(varying,atp)&c(fixed,mannose))
   and reversibly_connected(atp,mannose,c(varying,atp)&c(fixed,mannose))
                                     satisfies the model cle_4_eq
binds_upstream_or_same_form_as(mannose,glu,c(varying,glu)&c(fixed,mannose))
                                     satisfies the model cle_4_eq
binds_upstream_or_same_form_as(mannose,atp,c(varying,atp)&c(fixed,mannose))
                                     satisfies the model cle_4_eq
```

```
*/ TEST OF EXPERIMENTAL RESULTS FOR AGREEMENT WITH THE ORDERED BI BI MECHANISM
WITH FIRST OBJECT ASSIGNMENT (Named Cle_1 in the system) */
```

```
/* Ordered bi-bi */
| ?- satisfy(fromm, cle_1).
```

```
react_with_the_same_enzyme_form(atp,adp,c(varying,atp)&c(fixed,adp))
                                     does not satisfy the model cle_1
```

```
binds_upstream_or_same_form_as(mannose,atp,c(varying,atp)&c(fixed,mannose))
                                     does not satisfy the model cle_1
```

```
*/ TEST OF EXPERIMENTAL RESULTS FOR AGREEMENT WITH THE PING PONG BI BI
MECHANISM WITH FIRST OBJECT ASSIGNMENT (Named Cle_5 in the system) */
```


Table 4. (cont.)

```

/* Ping pong bi-bi */
| ?- satisfy(fromm, cle_5).
reversibly_connected(glu,atp,c(varying,glu)&c(fixed,atp))
      does not satisfy the model cle_5
reversibly_connected(atp,glu,c(varying,atp)&c(fixed,glu))
      does not satisfy the model cle_5

react_with_different_enzyme_forms(atp,mannose,c(varying,atp)&c(fixed,
mannose)) and reversibly_connected(atp,mannose,c(varying,atp)&c(fixed,
mannose))
      does not satisfy the model cle_5

binds_upstream_or_same_form_as(mannose,atp,c(varying,atp)&c(fixed,mannose))
      does not satisfy the model cle_5

```

- DelaFuente & Sols (1970): ordered Bi Bi with EB dead-end complex; A = Glu, B = MgATP²⁻, P = MgADP⁻, and Q = G6P

where A and B are reactants, P and Q are products, Glu is glucose and G6P is glucose 6-phosphate. The various workers used different approaches to determining the mechanism of the enzyme. These differences are understandable in view of the difficulties in working with the enzyme. The mammalian enzyme is practically irreversible; there are just a few very recent studies of the reverse reaction. All workers except Toews (1966) have used structural analogues of glucose, MgATP²⁻ and glucose 6-phosphate as alternative substrates and/or inhibitors of the normal reactants (Bachelard *et al.*, 1971; Delafuente & Sols, 1970; Fromm & Zewe, 1962; Kosow & Rose, 1968; Noat *et al.*, 1968). Only DelaFuente & Sols (1970) reported results for the reverse reaction of yeast hexokinase. Although Noat *et al.* (1968) discussed 31 possible mechanisms, the other authors seem to have considered only one or two candidate mechanisms.

CONCLUSIONS

We have tested the mechanisms proposed by the different authors and have found that their results were generally consistent with their proposed mechanisms. The observed constraints agree with each other more than the actual mechanisms do. It might therefore be possible to infer a mechanism by attempting to reconcile the different constraints, since each data set was obtained under somewhat different experimental conditions, such as temperature, pH and ionic strength. Noat *et al.* (1968) have shown that the type of inhibition of MgADP⁻ against MgATP²⁻ depends on the ionic strength; this inhibition is competitive at an ionic strength of 0.05 M and non-competitive at 0.3 M. The results from all the data sets wherein ADP inhibitions were studied are summarized in Table 5.

Although there are no standard states for enzymic reactions in biochemistry, the most meaningful conditions for interpreting a mechanism may be those for intracellular ionic strength under physiological conditions. Although there is disagreement on the intracellular ionic strength of eukaryotic cells, values below 0.1 M are too low. Shanks & Bailey (1988) have recently shown, by comparing n.m.r. spectra in yeast with

Table 5. Types of ADP inhibition

Data source	Ionic strength (mol/l)	Type of inhibition (MgADP ⁻ versus MgATP ²⁻)
Noat <i>et al.</i> (1968)	0.30	Non-competitive
	0.05	Competitive
Kosow & Rose (1968)	0.10	Non-competitive
Toews (1966)	0.05	Competitive
Fromm & Zewe (1962)	0.08	Competitive

solutions of glycolytic intermediates *in vitro*, that the ionic strength of yeast cells is 0.15 M.

The above observations also involve different methods of calculating the concentration of MgATP²⁻, the actual reactive species of ATP. Noat *et al.* (1968) and many other researchers have reported that the value for the apparent stability constant of MgATP²⁻ depends on the ionic strength. As we have studied the subject extensively elsewhere (Garfinkel *et al.*, 1984, 1985, 1986), a detailed discussion is inappropriate here.

Although the data are generally consistent with the enzyme mechanisms proposed by the authors, this alone does not prove that a proposed mechanism is correct. Ideally one should also show that all the theoretical constraints have been satisfied by the proposed model and that other candidate models do not fit the data. It is necessary to find the experimental conditions most advantageous for determining the mechanism.

The computer-based techniques reported here can help interpret experimental results and design additional experiments in several ways. The constraint satisfaction procedures show exactly which constraints a proposed model does not satisfy. Qualitative simulation can show all the theoretical constraints inherent in a mechanism. Model discrimination based on set partitioning (Soo, 1987) methods can show how to distinguish among a number of mechanisms most efficiently (with the fewest experiments) (Soo, 1987).

It is also surprising that very few experiments were done at both unsaturating and saturating substrate concentrations. Only the experiment reported in Fig. 10 of Toews (1966), where ADP ceased to be inhibitory against glucose at high ATP concentrations, made use of conditions where observed effects change. In our system,

the results of running experiments under such conditions are predicted from the characteristics of the partial reactions resulting when different kinds of experiments are run. Even Noat *et al.* (1968), who reported a large variety of experiments and considered 31 possible models for their enzyme, did not report such experiments. It is difficult to explain these omissions when Cleland (1963*a,b*) and Segel (1975) have discussed these effects extensively. Perhaps the sheer number of measurements required makes such experimentation impractical.

In view of the finding that the discrepancies and scatter in the constraints derived from the experimental data is so much less than in the mechanisms derived from the data by the authors, and in the mechanisms reported in the literature (Garfinkel *et al.*, 1987), it appears at least conceptually feasible to design a workable small number of experiments to fill in the remaining gaps and resolve the remaining discrepancies. Although it will be necessary to analyse additional data sets for hexokinase, we have probably covered most of the mechanisms that have been proposed for the enzyme. One would then hope to have a single well-established mechanism, if only for some particular set of conditions. While considerable experimental and computer work remains to be done, the remaining effort now appears much less than what has already been expended on unravelling the kinetics of hexokinase.

This paper was taken in part from the Ph.D. Thesis of V.-W. S. (Soo, 1987) and was supported in part by National Institutes of Health grants RR2230, HL15622, and DK33016. It is also based in part upon work supported under a National Science Foundation (U.S.A.) Fellowship to D.M.C.

REFERENCES

- Bachelard, H. S., Clark, A. G. & Thompson, M. F. (1971) *Biochem. J.* **123**, 707–715
- Bennett, W. S. & Steitz, T. A. (1980*a*) *J. Mol. Biol.* **140**, 183–209
- Bennett, W. S. & Steitz, T. A. (1980*b*) *J. Mol. Biol.* **140**, 211–230
- Cha, S. (1968) *Mol. Pharmacol.* **4**, 621–629
- Cleland, W. W. (1963*a*) *Biochim. Biophys. Acta* **67**, 104–137
- Cleland, W. W. (1963*b*) *Biochim. Biophys. Acta* **67**, 173–187
- Cleland, W. W. (1963*c*) *Biochim. Biophys. Acta* **67**, 188–196
- Cleland, W. W. (1970) *Enzymes* 2nd Ed. **2**, 1–65
- Cornish-Bowden, A. (1976) *Principles of Enzyme Kinetics*, Butterworths, London
- DelaFuente, G. & Sols, A. (1970) *Eur. J. Biochem.* **16**, 234–239
- Dixon, M. & Webb, E. C. (1979) *The Enzymes*, 3rd edn., p. 271, Academic Press, New York
- Fromm, H. J. & Zewe, V. (1962) *J. Biol. Chem.* **237**, 3027–3032
- Garfinkel, L. & Garfinkel, D. (1984) *Biochemistry* **23**, 3547–3552
- Garfinkel, L. & Garfinkel, D. (1985) *Magnesium* **4**, 60–72
- Garfinkel, L., Altschuld, R. A. & Garfinkel, D. (1986) *J. Mol. Cell. Cardiol.* **18**, 1003–1013
- Garfinkel, L., Garfinkel, D., Matsiras, P. & Matschinsky, B. (1987) *Biochem. J.* **244**, 351–357
- Grossbard, L. & Schimke, R. T. (1966) *J. Biol. Chem.* **241**, 3546–3560
- Gulbinsky, J. S. & Cleland, W. W. (1967) *Biochemistry* **7**, 566–575
- Kohn, M. C., Mcnten, L. E. & Garfinkel, D. (1979) *Comput. Biomed. Res.* **12**, 461–469
- Kosow, D. P. & Rose, I. A. (1968) *J. Biol. Chem.* **243**, 3623–3630
- Maclay, J. & Garfinkel, D. (1986) *Trends Biochem. Sci.* **11**, 347
- Mannervik, B. (1975) *BioSystems* **7**, 101–109
- Mayer, S. E., Mayfield, A. S. & Hass, J. A. (1966) *Mol. Pharmacol.* **2**, 393–408
- Noat, G., Ricard, J., Borel, M. & Got, C. (1968) *Eur. J. Biochem.* **5**, 55–70
- Schremmer, S. D., Waser, M. R., Kohn, M. C. & Garfinkel, D. (1984) *Comput. Biomed. Res.* **17**, 289–301
- Segel, I. H. (1975) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, Wiley, New York
- Shanks, J. V. & Bailey, J. E. (1988) *Biotechnol. Bioeng.* **32**, 1138–1152
- Solheim, L. P. & Fromm, H. J. (1981) *Arch. Biochem. Biophys.* **211**, 92–99
- Soo, V.-W. (1987) Ph.D. Thesis, Rutgers University
- Soo, V.-W., Kulikowski, C. A., Garfinkel, L. & Garfinkel, D. (1987) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **46**, 274 (abstr.)
- Soo, V.-W., Kulikowski, C. A., Garfinkel, D. & Garfinkel, L. (1988) *Comput. Biomed. Res.* **21**, 381–403
- Toews, C. J. (1966) *Biochem. J.* **100**, 739–744
- Ureta, T., Lazo, P. A. & Sols, A. (1985) *Arch. Biochem. Biophys.* **239**, 315–319
- Wong, J. T.-F. (1974) *Kinetics of Enzyme Mechanisms*, Academic Press, New York

Received 19 April 1989/3 July 1989; accepted 14 July 1989