

3-21-1992

## Optimal Selection of Metabolic Fluxes for in vivo Measurement. II. Application to Escherichia coli and Hybridoma Cell Metabolism

Joanne M. (Savinell) Belovich  
*Cleveland State University*

Berhard O. Palsson  
*University of Michigan*

Follow this and additional works at: [https://engagedscholarship.csuohio.edu/encbe\\_facpub](https://engagedscholarship.csuohio.edu/encbe_facpub)

 Part of the [Chemical Engineering Commons](#)

[How does access to this work benefit you? Let us know!](#)

---

### Repository Citation

(Savinell) Belovich, Joanne M. and Palsson, Berhard O., "Optimal Selection of Metabolic Fluxes for in vivo Measurement. II. Application to Escherichia coli and Hybridoma Cell Metabolism" (1992). *Chemical & Biomedical Engineering Faculty Publications*. 150.

[https://engagedscholarship.csuohio.edu/encbe\\_facpub/150](https://engagedscholarship.csuohio.edu/encbe_facpub/150)

This Article is brought to you for free and open access by the Chemical & Biomedical Engineering Department at EngagedScholarship@CSU. It has been accepted for inclusion in Chemical & Biomedical Engineering Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact [library.es@csuohio.edu](mailto:library.es@csuohio.edu).

# Optimal Selection of Metabolic Fluxes for *in vivo* Measurement. II. Application to *Escherichia coli* and Hybridoma Cell Metabolism

JOANNE M. SAVINELL AND BERNHARD O. PALSSON

A method of analysis was presented in part I of this series for determining the fluxes in a biochemical network that are the optimal choices for experimental measurement. This algorithm is applied to two important biological models: *Escherichia coli* and a hybridoma cell line (167.4G5.3). Our results show that potentially poor choices for *in vivo* measurement of metabolic fluxes exist for both model systems. For the subset of reactions in *E. coli* that was studied, the condition number of the augmented stoichiometric matrix reveals that a 60-fold amplification of experimental error during computations is possible. The biochemical network of the hybridoma cell is more complex than the *E. coli* system, and thus results in much larger possible error amplification—up to 100 000-fold. The physiological situations appear to have sensitivities that are less than 1/4 to 1/10 of those estimated by the condition number, and the maximum sensitivities are proportional to the condition number. These maximum sensitivities calculated using estimates of the fluxes and the worst possible error vector are upper bounds on the system's actual sensitivity. By examining the effect of measurement error on the sensitivity, the most probable sensitivity is calculated. These results indicate that an approximate two-fold increase in sensitivity of the *E. coli* system is likely when the worst set of fluxes are measured rather than the best set. The most likely sensitivity of the hybridoma system can range three orders of magnitude, depending on the set of fluxes that are measured. The propagation of experimental error during computations can be diminished for both systems by increasing the number of flux measurements over and above the minimum number of experimental measurements. The findings from these two model systems indicate that the calculation of the condition number can be a useful method for efficient experimental design, and that the usefulness of this method increases as the order of the system increases.

## Introduction

In part I of this series (Savinell & Palsson, 1992a) (referred to in this paper as part I) we presented an algorithm for estimating the sensitivity of calculated fluxes to experimental error when measuring the steady-state fluxes in a biochemical network. This algorithm is important to the design of sophisticated metabolic flux experiments in order to maximize the accuracy of the metabolic flux map. Further, the identification of metabolic reactions that significantly influence large segments of the metabolic flux map has clear and important implications to the study of metabolic teleology.

In this manuscript we illustrate in detail the application of the mathematical methods developed in part I to two important metabolic systems. The first is a subset of *Escherichia coli* metabolism that has been experimentally studied by Walsh & Koshland (1985). This small metabolic system has 11 metabolites and 18 reactions and represents the core of bacterial energy metabolism. The second metabolic system studied is a comprehensive representation of hybridoma metabolism. The hybridoma cell line used has been extensively studied and comprehensive experimental data are available, making the computations realistic. In addition to its clear biotechnological importance, the second example is of wider interest since the intermediary metabolism of most cultured cells is similar.

For each of these two selected metabolic systems, we calculate the condition number of the augmented stoichiometric matrix for each possible combination of flux measurements, and from these results, the combination of fluxes that is best suited for measurement is identified. These results are verified by calculating a more realistic sensitivity, using data from the literature as estimates of the fluxes. To determine the exact sensitivity, knowledge of the experimental error is also needed. Since this information is not readily available, we assume that the values of the individual error measurements are independent and random. The direction of the error vector is thus randomly selected, and the corresponding sensitivity calculated. From the frequency of each sensitivity, we can determine the sensitivity that is most likely to represent the system's true sensitivity.

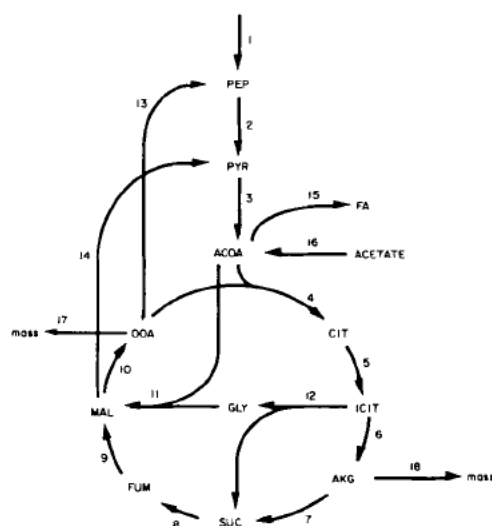


FIG. 1. Reaction network for *Escherichia coli*, from (Walsh & Koshland, 1985). Reactions are: (1), gluconeogenesis (GLCneo); (2), pyruvate kinase (PK); (3), pyruvate dehydrogenase (PDH); (4), citrate synthetase (CITsyn); (5), aconitase (ACON); (6), isocitrate dehydrogenase (ICITdh); (7), α-ketoglutarate dehydrogenase (AKGdh); (8), succinate dehydrogenase (SUCdh); (9), fumarase (FUMase); (10), malate dehydrogenase (MALdh); (11), glyoxylate lyase (GLYase); (12), glyoxylate synthetase (GLYsyn); (13), phosphoenyl pyruvate carboxykinase (PEPCK); (14), malic enzyme (ME); (15), fatty acid synthesis (FASyn); (16), acetate uptake (qACE); (17), oxaloacetate used for biosynthesis (qOOA); (18), α-ketoglutarate used for biosynthesis (qAKG). Metabolites are: PEP: phosphoenyl pyruvate; PYR: pyruvate; ACOA: acetyl CoA; CIT: citrate; ICIT: isocitrate; AKG: α-ketoglutarate; SUC: succinate; FUM: fumarate; MAL: malate; OOA: oxaloacetate; GLY: glyoxylate.

TABLE 1

*Stoichiometric matrix for the network shown in Fig. 1*

Compound																		
PEP	-1	-1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
PYR	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
ACOA	0	0	1	-1	0	0	0	0	0	0	-1	0	0	0	-1	1	0	0
CIT	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
ICIT	0	0	0	0	1	-1	0	0	0	0	0	-1	0	0	0	0	0	0
AKG	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	-1
SUC	0	0	0	0	0	0	1	-1	0	0	0	1	0	0	0	0	0	0
FUM	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0
MAL	0	0	0	0	0	0	0	0	1	-1	1	0	0	-1	0	0	0	0
OOA	0	0	0	-1	0	0	0	0	0	1	0	0	-1	0	0	0	-1	0
GLY	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	0	0
Reaction	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

**Two Model Systems**

Two metabolic networks were analyzed using the algorithm outlined in part I. A simplified biochemical network that represents a portion of *E. coli* metabolism is shown in Fig. 1 (Walsh & Koshland, 1985) and the stoichiometric matrix for this network is shown in Table 1. The flux measurements used here for estimating the

TABLE 2

*Flux measurements for reactions in the Escherichia coli network, from (Walsh & Koshland, 1985)*

Reaction		Fluxes (mmol l <sup>-1</sup> min <sup>-1</sup> )		
No.	Name	Acetate	Transition	Glucose and acetate
1	GLCneo	19.0	-20	-49
2	PK	-9.5	12	31
3	PDH	0.0	12	31
4	CITsyn	111.0	20	31
5	ACON	111.0	20	31
6	ICITdh	80.0	20	31
7	AKGdh	75.0	17	24
8	SUCdh	106.0	17	24
9	FUMase	106.0	17	24
10	MALdh	127.5	17	24
11	GLYase	31.0	0	0
12	GLYsyn	31.0	0	0
13	PEPCK	9.5	-8	-18
14	ME	9.5	0	0
15	qOOA	7.0	5	11
16	FAsyn	3.0	3	5
17	qACE	145.0	11	5
18	qAKG	5.0	3	7

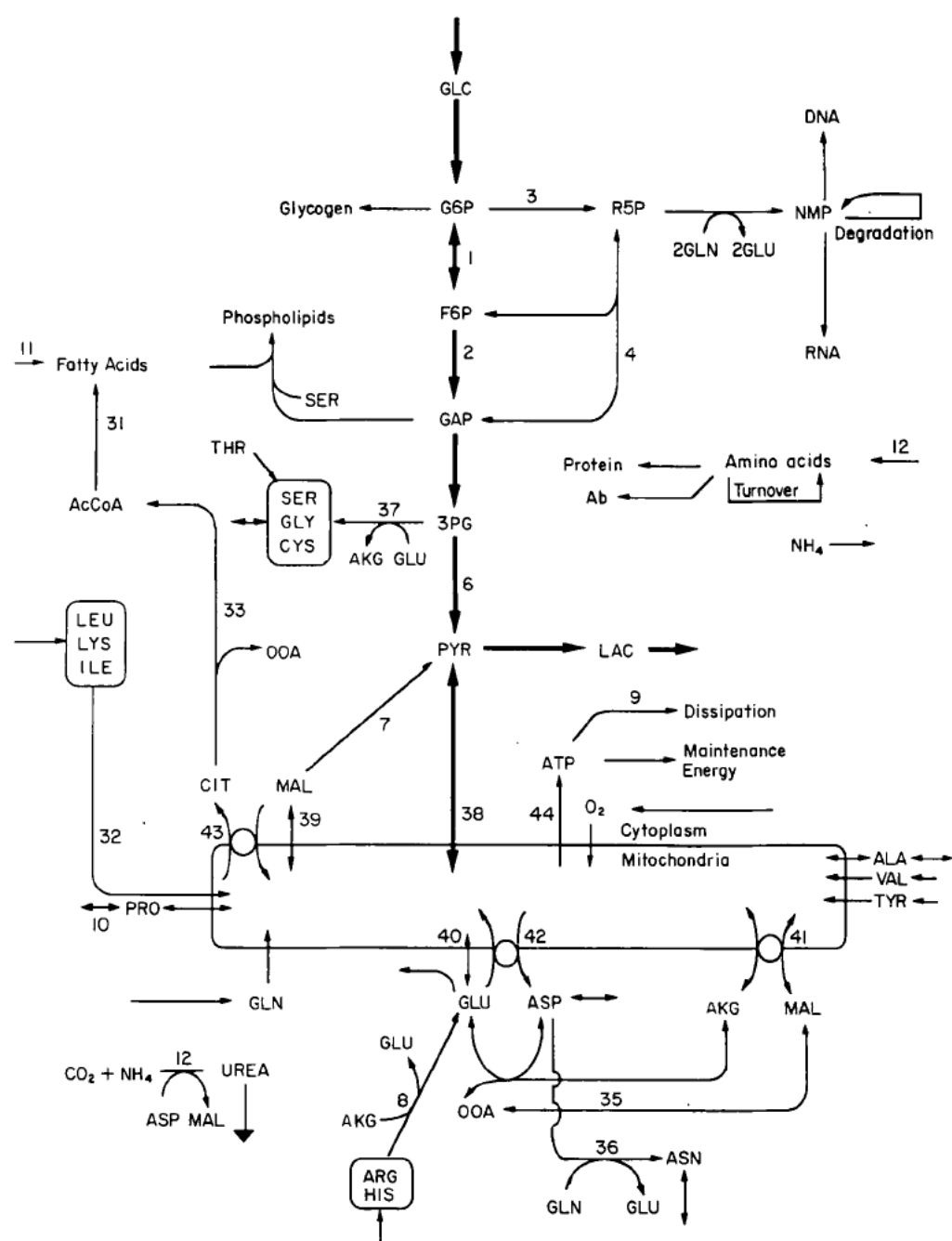


FIG. 2. Reaction network for the hybridoma cell cytoplasm. Abbreviations are: AcCoA: acetyl CoA; ADP: adenosine diphosphate; ATP: adenosine triphosphate; ARG: arginine; ASN: asparagine; ASP: aspartate; AKG: α-ketoglutarate; CIT: citrate; CYS: cysteine; DNA: deoxyribonucleic acid; F6P: fructose-6-phosphate; GLC: glucose; G6P: glucose-6-phosphate; GLN: glutamine; GLU: glutamate; GAP: glyceraldehyde-3-phosphate; GLY: glycine; HIS: histidine; ILE: isoleucine; LAC: lactate; LEU: leucine; LYS: lysine; MAL: malate; NADH: nicotinamide adenine dinucleotide; NADPH: nicotinamide adenine dinucleotide phosphate; NMP: nucleotide monophosphate; OOA: oxaloacetate; 3PG: 3-phosphoglycerate; PRO: proline; PYR: pyruvate; RNA: ribonucleic acid; R5P: ribose-5-phosphate; SER: serine; THR: threonine; TCA: tricarboxylic acid; TYR: tyrosine; VAL: valine.

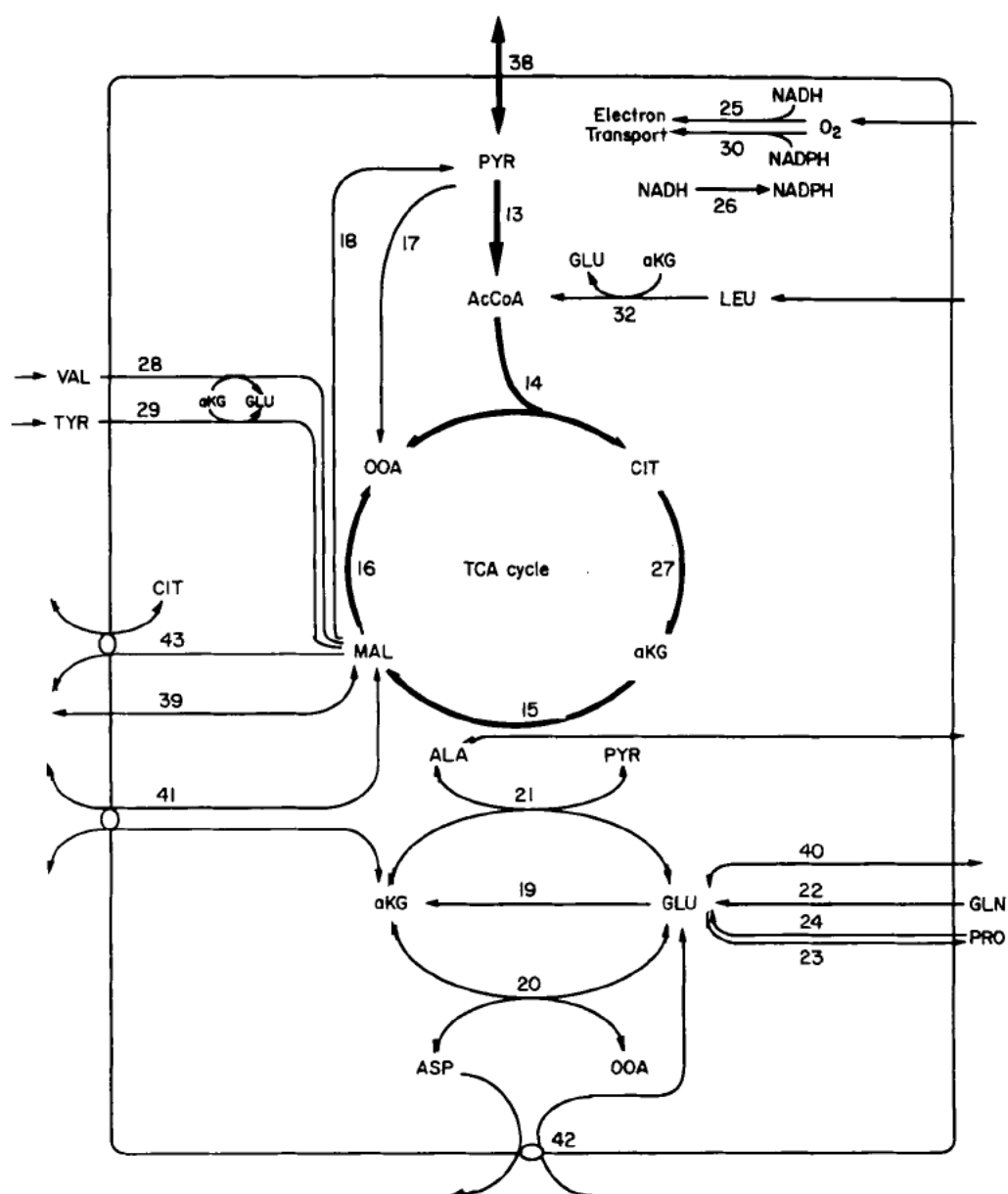


FIG. 3. Reaction network for the hybridoma cell mitochondria.

actual sensitivity,  $R$ , were obtained by Walsh & Koshland (1985). The cells were grown on media containing acetate, and the fluxes in the network measured by means of a radioactive tracer study. Glucose was then added to the media, and after 20 min the fluxes were again measured. The cells were then cultured in media containing

TABLE 3

*Stoichiometric matrix for the hybridoma cell shown in Figs. 2 and 3, modified from (Savinell & Palsson, 1992a). The subscript M denotes concentration in the mitochondria*

[illegible]

both glucose and acetate, and the steady state fluxes again measured. These measurements are shown in Table 2.

The biochemical network assumed to represent the hybridoma cell's metabolism is shown in Figs 2 and 3 (Savinell & Palsson, 1992*b*). The biochemical network for the hybridoma cell has a large number of reactions ( $m=65$ ), but the number of unknown fluxes can be reduced by incorporating measurements of metabolite uptake and production rates into the system of mass balances. Such data are available for

TABLE 3—continued

Compound																								
G6P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F6P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DHAP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3PG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
R5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FA	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PYR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
ACCOA	0	0	0	0	0	0	0	0	-8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
AKG	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	1	0	0	0	1	0	0	1	0	0
MAL	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	-1	0	-1	0	-1	0	-1	0
CIT	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	1	0	0
OOA	0	0	0	0	0	0	0	0	0	0	1	1	-1	0	0	0	0	0	0	0	0	0	0	0
GLU	0	0	0	0	0	0	0	0	0	0	0	1	0	1	-1	0	0	1	0	-1	0	0	0	0
GLN	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
SER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ASP	0	0	0	0	0	0	0	0	0	0	0	-1	0	-1	0	0	0	0	0	0	0	0	0	0
ALA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LEU	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ARG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRO	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ASN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
VAL	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TYR	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ATP	0	0	0	0	0	0	0	0	-7	0	-1	0	0	-2	0	0	0	0	0	0	0	0	0	1
NADH	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	1	0	0	0	0	0	0	0	0	0
NADPH	0	0	0	0	0	0	0	0	-14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NH <sub>4</sub>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OXY	0	0	-5	0	0	0	0	-5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PYR <sub>M</sub>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ACCOA <sub>M</sub>	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CIT <sub>M</sub>	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0
AKG <sub>M</sub>	0	0	0	0	1	-1	-1	0	0	-1	0	0	0	0	0	0	0	0	-1	0	0	0	0	0
MAL <sub>M</sub>	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0
OOA <sub>M</sub>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ASP <sub>M</sub>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0
GLU <sub>M</sub>	1	1	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	-1	0	1	0	0	0	0
NADH <sub>M</sub>	-1	1	-1	-1	0	2.4	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NADPH <sub>M</sub>	-1	0	0	1	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ATP <sub>M</sub>	-1	0	2.5	-1	0	1	0	3.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
Reaction	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44		

the hybridoma cell line 167.4G5.3 cultured at steady state in a pH and dissolved oxygen (DO) controlled bioreactor (Ozturk & Palsson, 1990). By setting these fluxes in the stoichiometric matrix equal to the measured values, the number of unknown fluxes,  $m$ , is reduced to 44. The resulting stoichiometric matrix, with 39 mass balances and thus 5 degrees of freedom, is shown in Table 3. The number of matrix rearrangements for this simplified system can be handled by a mainframe computer within a reasonable amount of time. The values in  $b$  from eqn (2) in part I are modified from (Savinell & Palsson, 1992b) since this vector must take into account the measured fluxes as well as the biosynthetic fluxes. The measurements at 30% DO and 0.1% DO



TABLE 4  
*Estimated fluxes for the hybridoma cell network*

Objective function DO	Fluxes (nmol/million cells hr <sup>-1</sup> )			
	Maximize ATP production		Minimize ATP production	
	30%	0.1%	30%	0.1%
Reaction				
1	85.28	482.79	84.40	393.32
2	84.50	483.60	83.71	402.25
3	0.00	2.37	0.12	14.48
4	-1.15	1.21	-1.04	13.33
5	167.14	966.12	165.58	807.42
6	159.91	963.59	158.35	804.89
7	0.23	24.23	0.00	0.00
8	0.93	4.73	0.00	4.73
9	0.00	184.70	0.00	2.57
10	3.19	41.55	2.14	53.66
11	2.98	0.92	2.98	0.92
12	0.00	21.00	0.00	26.86
13	18.36	0.00	17.82	0.00
14	19.08	16.42	18.54	16.42
15	39.22	19.35	38.71	19.35
16	49.25	19.98	46.23	25.84
17	0.00	0.00	0.00	0.00
18	17.27	0.44	18.92	24.67
19	0.00	7.86	0.09	2.00
20	30.17	3.57	27.69	9.42
21	11.05	13.55	11.45	13.55
22	18.00	51.14	17.91	62.85
23	0.03	38.39	0.00	50.50
24	0.00	0.00	1.02	0.00
25	0.00	0.00	134.54	22.00
26	135.68	59.95	0.00	25.83
27	19.08	0.00	18.54	0.00
28	0.00	1.99	0.89	1.99
29	0.00	3.40	0.50	3.40
30	172.00	22.00	37.46	0.00
31	0.00	2.05	0.00	2.05
32	0.72	16.42	0.72	16.42
33	0.00	16.42	0.00	16.42
34	26.67	-18.37	24.19	-18.37
35	26.67	-1.95	24.19	-1.95
36	0.66	0.00	0.66	0.00
37	7.23	2.53	7.23	2.53
38	12.14	13.11	10.35	-11.11
39	6.94	-4.56	8.09	25.52
40	7.63	13.14	9.49	18.60
41	20.37	-16.17	16.96	-16.17
42	30.17	3.57	27.69	9.42
43	0.00	16.42	0.00	16.42
44	505.51	0.00	507.06	0.00

(Ozturk & Palsson, 1990) were chosen to represent the normal aerobic condition and the nearly anaerobic condition, respectively.

Measurements of all internal metabolic fluxes in the reaction network of the hybridoma cell are not yet available. However, given that measurements of uptake and production rates are available and the low number of degrees of freedom of the resulting system, the internal metabolic fluxes through the network can be estimated using linear optimization (Savinell & Palsson, 1992b). The two objective functions of minimizing and maximizing total ATP production were chosen here for this purpose since they represent extreme points at opposite ends of the solution space. These results, used as estimates of  $b_r$ , are shown in Table 4 for both the nearly anaerobic and the aerobic conditions.

For each of these two systems, all possible configurations of  $I_p$  were derived, and the condition numbers calculated. From probability theory it is known that the

TABLE 5

*Flux combinations for Escherichia coli resulting in the ten lowest and ten highest condition numbers, where the minimum number of fluxes (7) out of 18 fluxes are measured. The number of occurrences of each flux in the flux combinations with the 50 lowest and 50 highest condition numbers are shown*

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Ten lowest																		
9-317			+		+		+			+			+	+			+	
9-317			+		+		+			+			+	+		+		
9-548	+	+			+		+			+				+			+	
9-548	+	+			+		+			+				+		+		
9-676		+			+		+			+			+	+			+	
9-676		+			+		+			+			+	+		+		
10-021	+	+		+			+			+				+			+	
10-021	+	+		+			+			+				+		+		
10-047			+	+			+			+			+	+			+	
10-047			+	+			+			+			+			+		
Frequency in																		
50 lowest	19	31	17	18	22	2	28	5	17	27	15	5	33	37	14	30	30	0
Ten highest																		
46-435	+	+					+								+	+	+	+
46-977	+				+		+							+	+	+	+	+
48-024							+						+	+	+	+	+	+
49-193	+					+				+					+	+	+	+
50-389	+					+	+			+					+	+	+	+
51-912	+						+			+					+	+	+	+
55-308	+					+								+	+	+	+	+
56-738	+									+				+	+	+	+	+
56-881	+					+	+							+	+	+	+	+
58-973	+						+							+	+	+	+	+
Frequency																		
50 highest	38	17	19	1	2	16	28	0	0	16	2	1	17	19	50	43	43	37

number of combinations possible is given by:

$$\frac{m!}{n!p!} \quad (1)$$

Since the stoichiometric matrix for the *E. coli* system has 11 rows and 18 columns, there are 31 824 possible combinations of measured fluxes, calculated by eqn (1). There are 1 370 754 possible combinations for the hybridoma system. The single value decompositions were calculated using the software package EISPACK. The calculation for the *E. coli* system required several hours on a DN 10000 Apollo workstation and the hybridoma system required about 6 hr on the IBM 3090 main-frame computer.

### Results for the *Escherichia coli* Metabolic Subsystem

#### CONDITION NUMBER CALCULATIONS

The condition numbers for the *E. coli* system were calculated and found to vary from nine to 59, indicating that the experimental error can be amplified up to a factor of 59 if the worst combination of fluxes are chosen for measurement. The lowest and highest ten values are shown in Table 5, along with the corresponding selection of fluxes. The condition numbers increase almost continuously, so that there are numerous flux combinations that are good choices for measurements, and many that are potentially poor choices for measurement.

#### UPPER BOUNDS ON SENSITIVITIES USING LIKELY VALUES OF METABOLIC FLUXES

The degree to which the condition number overestimates the system's actual sensitivity  $R$  depends on the values of  $b_i$  and  $\Delta b_i$ . If approximate values of the fluxes in  $b_i$  are available from the literature, a more refined upper bound,  $R$ , can be calculated using eqn (13) in part I. The fluxes through the *E. coli* network, obtained by Walsh & Koshland (1985) and shown in Table 2, were measured using radioactively labeled substrates, for cells grown in three different conditions: acetate as the only substrate, glucose and acetate as substrates, and during the transition between these two states. The actual sensitivities shown in Fig. 4 were calculated for each combination of fluxes, represented by the condition number of that system, using the data in Table 2 and the  $\Delta b_{i,max}$  given by eqn (17) in part I. This value of  $R$  is also an upper bound on the sensitivity, since  $\Delta b_{i,max}$  gives the maximum sensitivity for the given  $b_i$ , as discussed in part I. In all three cases, the actual sensitivities are approximately 1/4 of the condition numbers, and increase linearly with the condition number, indicating that  $C$  is a conservative estimate of the sensitivity at these physiological conditions.

#### INFLUENCE OF ERROR DIRECTION ON SENSITIVITIES

The influence of  $\Delta b_i$  on the sensitivity of the *E. coli* system was explored for the stoichiometric matrices where the best set of fluxes (given by the lowest condition

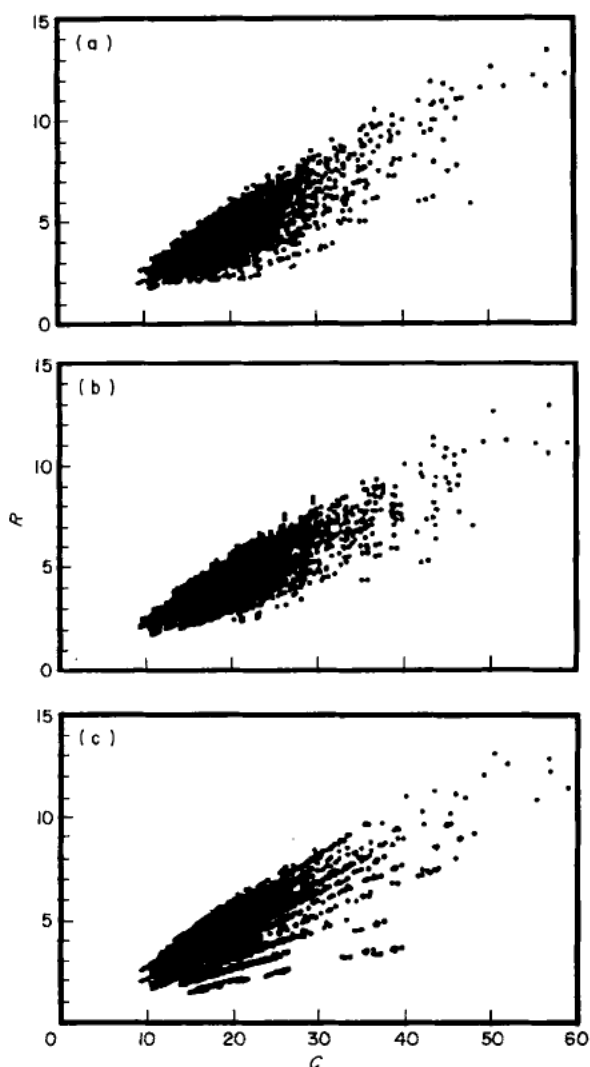


FIG. 4. Actual sensitivities for all possible combinations of flux measurements in the *Escherichia coli* system, using  $\Delta b_i = \Delta b_{i,max}$  [from eqn (17) in part I], and  $b_i$  from Table 2: (a) cells grown on acetate; (b) cells during transition to growth on glucose; (c) cells grown on glucose and acetate.

number) are measured and where the worst set of fluxes (given by the highest condition number) are measured. The values of  $b_i$  were set to the fluxes measured for growth on acetate and glucose + acetate (Table 2).

The number of points at an angle  $\phi$  with the reference vector, i.e.  $\Delta b_{i,max}$ , should increase as  $\phi$  increases up to  $90^\circ$  if the points are evenly distributed on the surface of a seven-dimensional sphere, as discussed in part I. Consequently, the frequency of  $\phi$  is greatest at  $90^\circ$  and lowest at  $0^\circ$  and  $180^\circ$ . This constraint acts as a control to

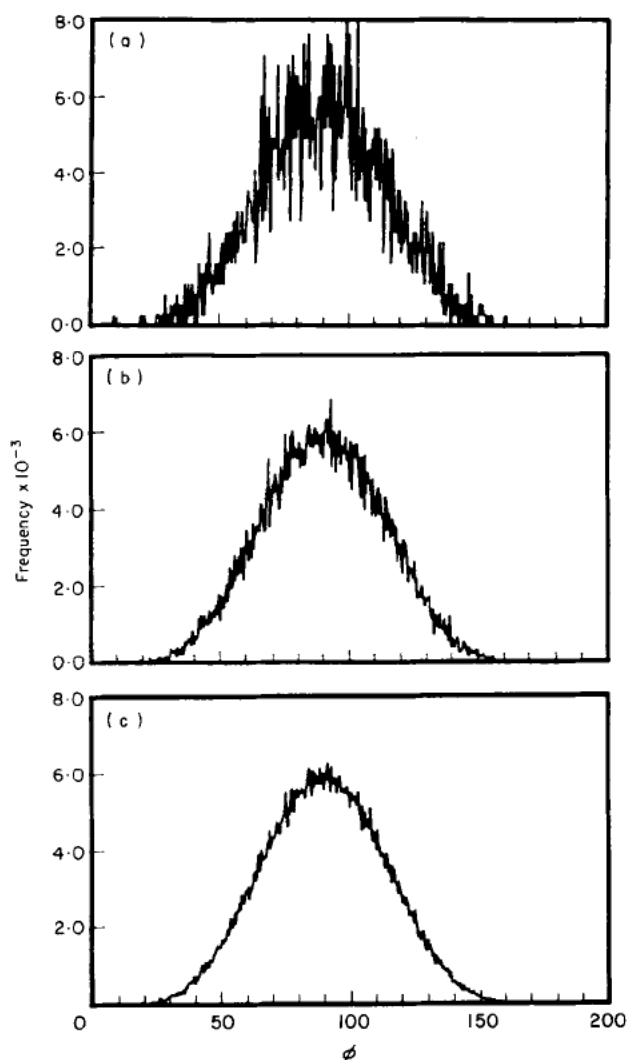


FIG. 5. Test of number of random points needed to describe all possible directions of  $\Delta b_r$ . The angle  $\phi$  between each random  $\Delta b_r$  and  $\Delta b_{r,max}$  was calculated. The fraction of the total number of trials that each  $\phi$  occurred is shown here, for different numbers of random points: (a)  $10^5$ ; (b)  $10^6$ ; (c)  $10^7$ .

insure that the computer program is generating randomly distributed  $\Delta b_r$ . Figure 5 shows the frequency distribution of  $\phi$  using three different numbers of points, in order to determine the amount of points needed to obtain a smooth distribution. The number of vectors used to calculate the sensitivities was selected to be  $10^7$ .

The sensitivities calculated using random  $\Delta b_r$ s are shown as a function of  $\phi$  in Fig. 6. The sensitivity is greatest at  $\phi = 0^\circ$  and  $180^\circ$  and the minimum value of  $R$  coincides with  $\phi = 90^\circ$ . The sensitivity tends to decrease as  $\Delta b_r$  becomes farther away from

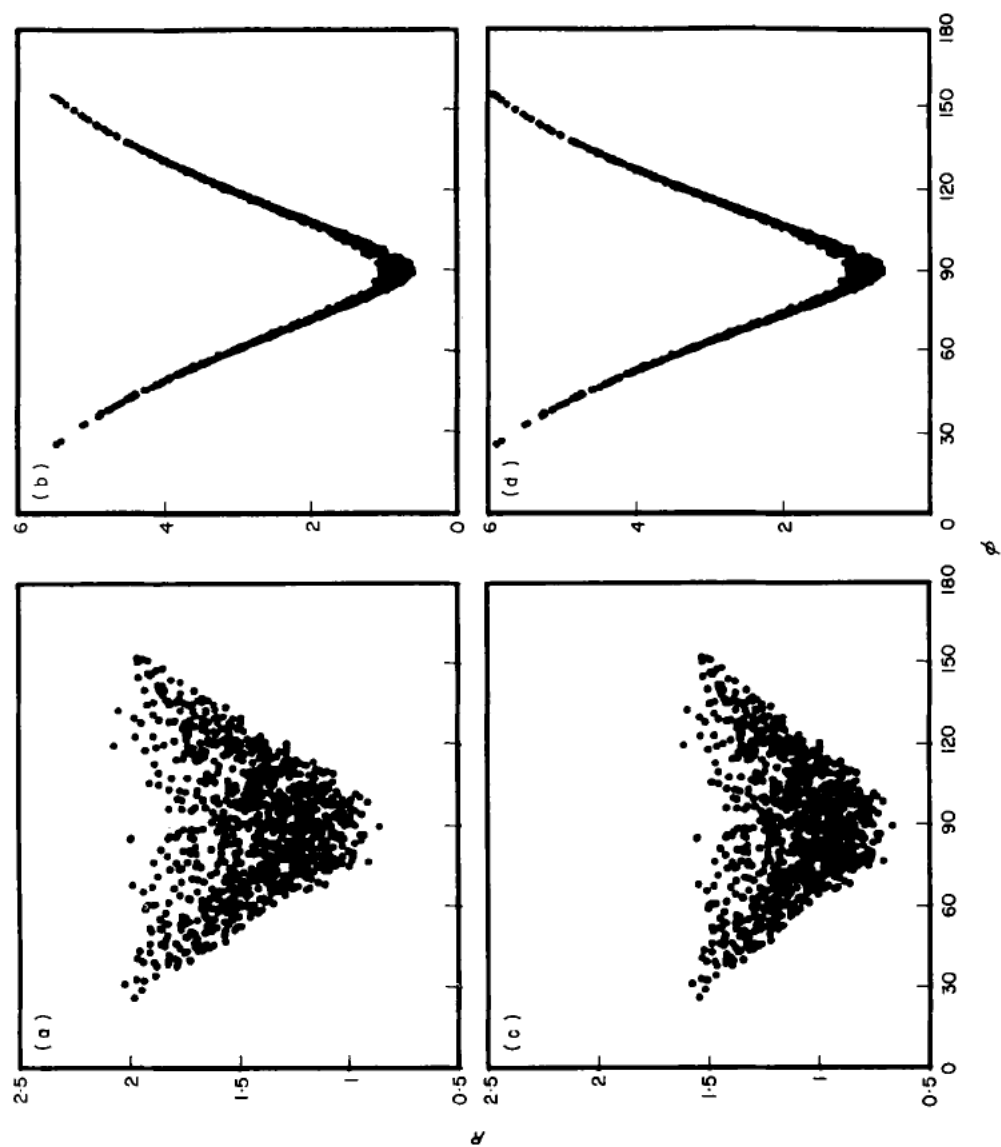


FIG. 6. The sensitivity  $R$  of the *Escherichia coli* system, calculated using  $b_i$  measured during growth on glucose + acetate (a, b) and acetate (c, d), and randomly selected  $\Delta b_i$ . (a), (c): low condition number system; (b), (d): high condition number system.

$\Delta b_{i,max}$ . The distribution of points is more narrow in the high condition number system, because the range of  $R$  is much greater than that in the low condition number system.

The frequency distributions of  $R$  for points uniformly distributed on the surface of the seven-dimensional sphere are shown in Fig. 7(a)–(d). The low condition number systems both show nearly symmetric distributions of sensitivities, and the most probable sensitivity for growth on acetate is less than the most probable sensitivity for growth on glucose + acetate. The sensitivity may be up to 6 for the high condition systems; however, the distributions are greatly skewed so that the most probable sensitivities are less than 2. The sensitivities are similarly distributed for growth on both glucose and acetate in the high condition systems.

The integral of the distributions in Fig. 7(a)–(d) are shown in Fig. 7(e)–(h). These figures show that when the low condition number fluxes are measured, there is greater than 50% probability that  $R$  is less than 1.5 for growth on glucose + acetate, and less than 1.2 for growth on acetate. When the high condition number fluxes are measured, the probability that  $R$  is less than 2.0 is greater than 50%, which is similar to the low condition number cases. However, for the high condition number system, there is also a 50% probability that  $R$  is between 2.0 and 6.0, indicating that there is a significant probability that the sensitivity is relatively high if these fluxes are measured, while the sensitivity of the low condition number system is always less than 2.0.

For the glucose + acetate data, there is a 2.5% probability that  $R$  is less than 1, indicating that it is unlikely that experimental error is compressed in the computations. On the other hand, there is a 30.7% probability that error may be compressed using the acetate data.

#### ADDITIONAL FLUX MEASUREMENTS

Figure 7 shows that even when the best combination of fluxes are measured, the value of  $R$  will usually exceed unity. In fact, the most probable result will be an amplification of experimental error by a factor of 50%. One way to overcome such error amplification is to take additional measurements and thus calculate fewer fluxes. The experimental technique that is used, either  $^{14}\text{C}$  labeling or NMR, often provides simultaneously a large number of measurements that may exceed the minimum needed for unique calculation of the metabolic flux map.

As outlined in part I, the effects of additional measurements are handled by the use of a least squares approach. The lowest condition number for *E. coli* when the minimum number of fluxes (7) are measured is 9.3 as shown in Table 5. The square of this lowest condition number, 86.8, is equal to the minimum  $C_0$ . The sensitivities,  $R_i$ , were calculated using eqn (22) in part I, using  $b_i$  set to the data for growth on glucose and acetate (Table 2) and  $S_i^T \Delta b_i$  calculated using eqn (23), part I. The minimum and maximum  $R_i$  for each number of measurements are shown in Fig. 8(a). The sensitivities decreased as additional fluxes are measured, indicating that experimental error propagates less during the computations when the number of measurements is greater than the minimum number. The sensitivity drops by more than one order of magnitude when only one additional flux is measured. The value

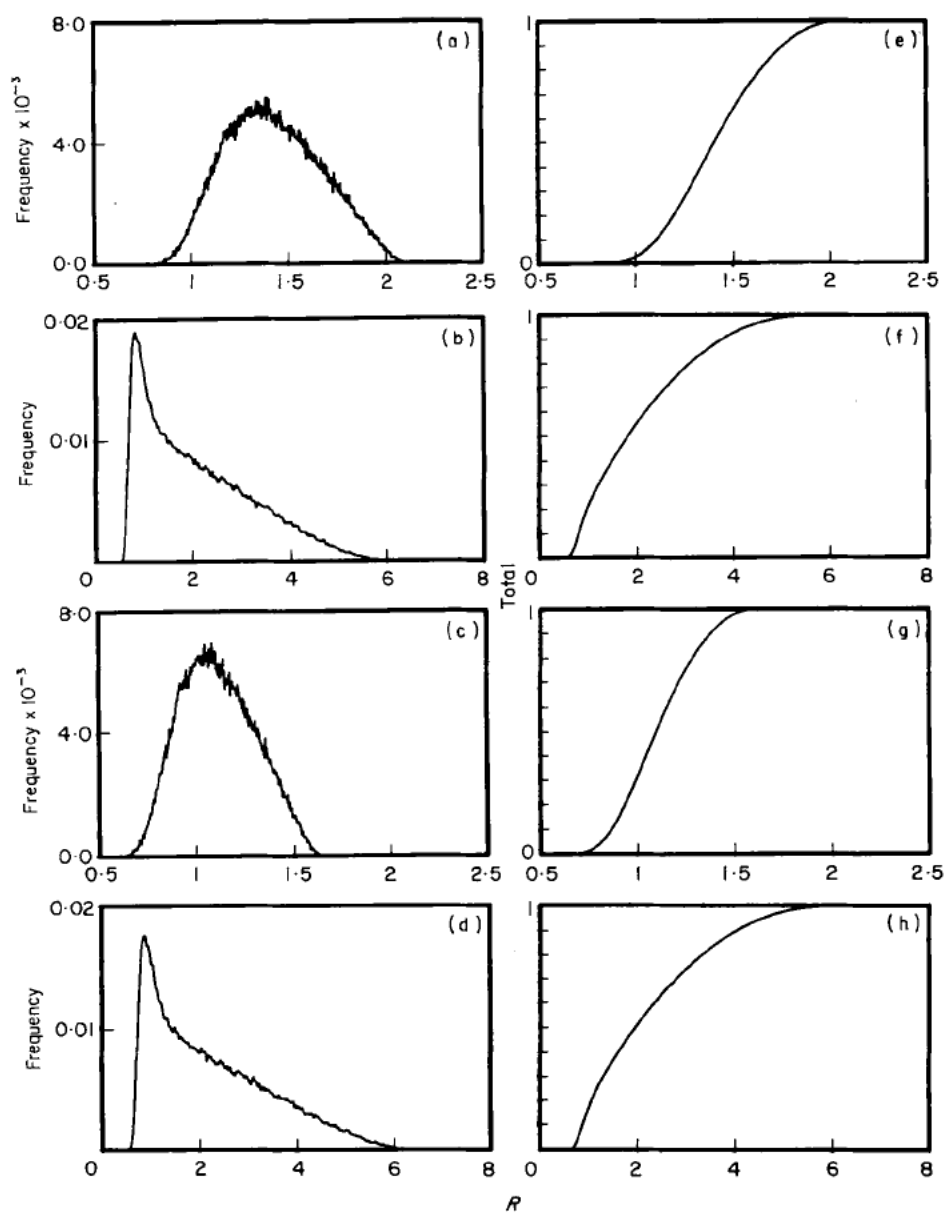


FIG. 7. (a)–(d) Are the frequency distributions of the sensitivities, and (e)–(h) are the integrals of the corresponding distributions. (a), (b):  $b$ , obtained from data from growth on glucose + acetate; (c), (d):  $b$ , obtained from data from growth on acetate; (a), (c): low condition number system; (b), (d): high condition number system.



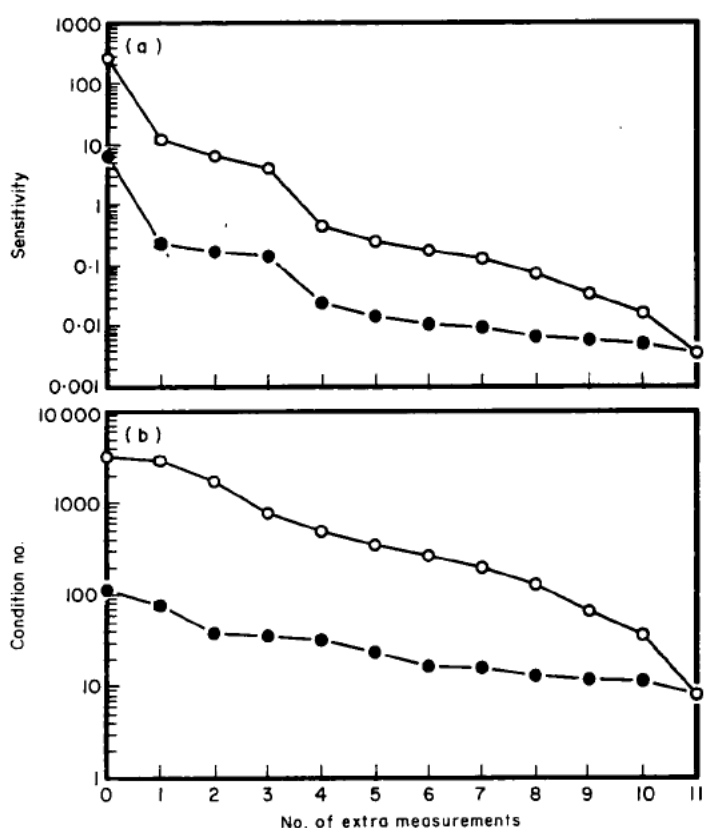


FIG. 8. (a) The lowest and highest sensitivities ( $R_\eta$ ) obtained using  $\Delta b_i = \Delta b_{i,\max}$  for *Escherichia coli* system as a function of the number of fluxes measurements greater than the minimum number; open circles: highest sensitivity; closed circles: lowest sensitivity. (b) The condition numbers that correspond to the systems in part (a).

of the minimum  $R_\eta$ , for  $\eta \geq 1$  (i.e. one or more additional measurement), is less than 1, indicating that suppression of experimental error during computation is possible. When four or more additional measurements are made, then experimental error *must* be suppressed, since the maximum  $R_\eta$  is less than 1. With more than four measurements, the sensitivity decreases by a negligible amount, indicating that additional measurements may not be worth the extra effort.

The sets of fluxes that correspond to the minimum  $R_\eta$ , for  $\eta = 0, 1, \dots, 11$ , are shown in Table 6. Of the seven fluxes in the set ( $\eta = 0$ ), from 43–100% of these also appear in the sets  $\eta = 1-11$ . This result indicates that most, but not all, of the fluxes in the minimum set are desirable to measure, regardless of the number of measurements.

The condition numbers that correspond to the sensitivities shown in Fig. 8(a) are shown in Fig. 8(b). The condition numbers decrease with additional measurements, as do the sensitivities. This result indicates that the result obtained from the calculation of  $R_\eta$ , i.e. the optimum number of measurements is greater than the minimum

TABLE 6

Flux combinations for *Escherichia coli* resulting in the lowest sensitivity, for 0 to 11 measurements in addition to the minimum (7) measurements needed. The last column is the number of fluxes from the set  $\eta=0$  that appear in the combinations  $\eta=1$  to 11

No. of additional measurements ( $\eta$ )	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	No.
0	+			+			+						+	+	+	+			7
1			+	+			+					+	+	+	+	+			6
2	+		+			+	+			+				+	+	+	+		5
3		+	+			+			+			+	+		+	+	+	+	3
4			+	+					+			+	+	+	+	+	+	+	4
5	+		+			+		+		+	+		+	+	+	+	+	+	5
6	+		+		+	+	+	+		+			+	+	+	+	+	+	6
7	+		+		+	+	+		+	+		+	+	+	+	+	+	+	6
8	+	+	+		+	+	+		+	+		+	+	+	+	+	+	+	6
9	+	+	+		+	+	+	+	+	+		+	+	+	+	+	+	+	6
10	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	7
11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	7

number of measurements, is true regardless of the values of  $b_i$ , since the value of the condition number is dependent only on the stoichiometry.

### Results for Hybridoma Cell Metabolism

#### CONDITION NUMBER CALCULATIONS

The condition numbers for the hybridoma network varied from 187 to 95 316. The lowest and highest 10 values are shown in Table 7, along with the corresponding selection of fluxes.

#### UPPER BOUNDS ON SENSITIVITIES USING LIKELY VALUES OF METABOLIC FLUXES

The sensitivities,  $R$ , of this system calculated using the estimated  $b_i$ , shown in Fig. 9, indicate that the condition number may overestimate the sensitivity by more than a factor of 10. As with the *E. coli* system,  $R$  increases linearly with  $C$ , and the trends are similar for 0.1 and 30% DO and for both objective functions.

#### INFLUENCE OF ERROR DIRECTION ON SENSITIVITIES

The influence of  $\Delta b_i$  on the sensitivity of the hybridoma system was evaluated for the stoichiometric matrices with the lowest and highest condition number. The fluxes in  $b_i$  were set equal to those measured at 30% DO. Since the hybridoma system has 5 degrees of freedom, the error vector  $\Delta b_i$  was assigned random values corresponding to the surface of a sphere in five-dimensional space, using the same algorithm

TABLE 7

*Flux combinations for the hybridoma cell network resulting in the ten lowest and ten highest condition numbers, where five of the 44 metabolic fluxes are measured. The number of occurrences of each flux in the flux combinations with the 50 lowest and 50 highest condition numbers are shown*

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Ten lowest																						
186.76																	+					
186.79																	+					
186.79																	+					
186.81							+										+					
187.05							+										+					
187.30																	+					
187.36							+										+					
187.47																	+	+				
187.59							+										+					
187.60																		+				
Frequency in 50 lowest	0	0	0	0	0	0	18	0	21	0	0	0	0	0	0	14	19	18	0	0	0	0
Ten highest																						
90.224											+						+					
90.714											+						+					
92.468											+						+					
92.970											+						+					
93.471											+							+				
93.978											+							+				
94.801											+											
94.801											+											
95.316											+											
95.316											+											
Frequency in 50 highest	0	0	0	0	0	6	0	0	10	0	21	0	0	0	0	9	25	27	0	9	0	0

described in part I. From the results in Fig. 10, the number of vectors used to calculate  $R$  was selected to be  $10^6$ , as a compromise between a smooth distribution of angles and length of computing time.

The sensitivities calculated for each of the random  $\Delta b_i$  are shown as a function of  $\phi$  in Fig. 11. These figures are similar to those for the *E. coli* system, except that the range of  $R$  is much wider.

The frequency distributions of  $R$  for high and low condition numbers and for maximum and minimum ATP production are shown in Fig. 12(a)–(d). The low condition number systems both show narrow, almost symmetric distributions of sensitivities, but the high condition systems have the highest frequency at lowest  $R$ . From the integral of these distributions [Fig. 12(e)–(h)], 50% of the possible error combinations for the low condition system result in sensitivity less than 4.0, while all the possible error combinations have sensitivity less than 6. When the

TABLE 7—continued

C	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
Ten lowest																						
186-76								+				+					+					+
186-79			+									+					+					+
186-79								+				+					+					+
186-81								+				+										+
187-05								+				+										+
187-30			+									+					+					+
187-36			+									+										+
187-47								+				+										+
187-59			+									+										+
187-60								+				+					+					
Frequency in 50 lowest																						
	0	0	17	16	0	0	0	21	0	0	0	29	0	0	0	8	18	0	0	0	0	50
Ten highest																						
90 224				+								+					+					
90 714				+													+	+				
92 468				+								+					+					
92 970				+													+	+				
93 471				+								+					+					
93 978				+													+	+				
94 801			+	+								+					+					
94 801				+				+				+					+					
95 316			+	+													+	+				
95 316				+				+									+	+				
Frequency in 50 highest																						
	0	0	17	22	0	0	0	18	0	0	0	11	6	0	0	2	24	10	11	7	0	10

fluxes from the high condition system are measured, the probability of obtaining a very sensitive system is great. For the high condition system, 50% of possible error combinations will result in  $R$  less than 2000, and 50% will result in  $R$  between 2000 and 4500.

#### ADDITIONAL FLUX MEASUREMENTS

The sensitivities  $R_\eta$  of the systems with number of measurements greater than the minimum number were calculated using eqn (22) in part I, using  $b_i$  set to the data for growth at 30% DO (Table 4) and  $S_i^T \Delta b_i$  calculated using eqn (23), part I. Due to limitations in computation time, the sensitivity of each possible flux combination for each number of additional measurements could not be calculated. The calculations were simplified by assuming that the same set of five fluxes that appear in the flux combination given by the minimum  $R_\eta$ , for  $\eta=0$ , appears in all the flux combinations with minimum  $R_\eta$ . Even with this restriction, more than 3 million combinations are possible with six additional flux measurements. Consequently, the minimum and maximum  $R_\eta$  for up to six additional fluxes were calculated. As shown in Fig. 13(a), the sensitivities decrease as additional fluxes are measured, as they did

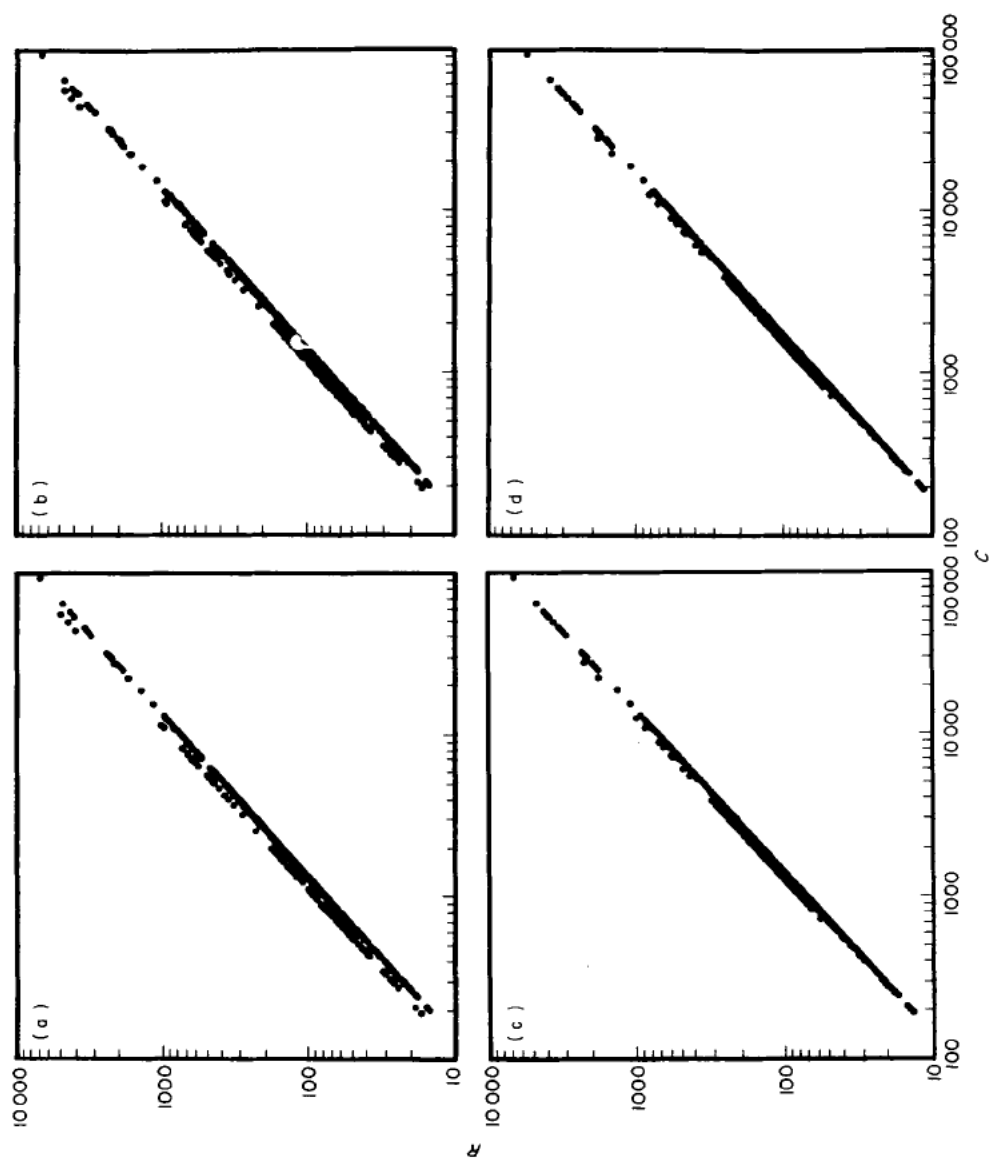


Fig. 9. Actual sensitivities for all possible combinations of flux measurements in the hybridoma system, using  $\Delta b_i = \Delta b_{i, \text{max}}$ . The values of external fluxes in  $b_i$  are obtained from cells cultured at 0.1% (a, b) and 30% DO (c, d). The values of the internal fluxes are obtained using linear optimization with the following objective functions: (a), (c): minimize total ATP production; (b), (d): maximize total ATP production.

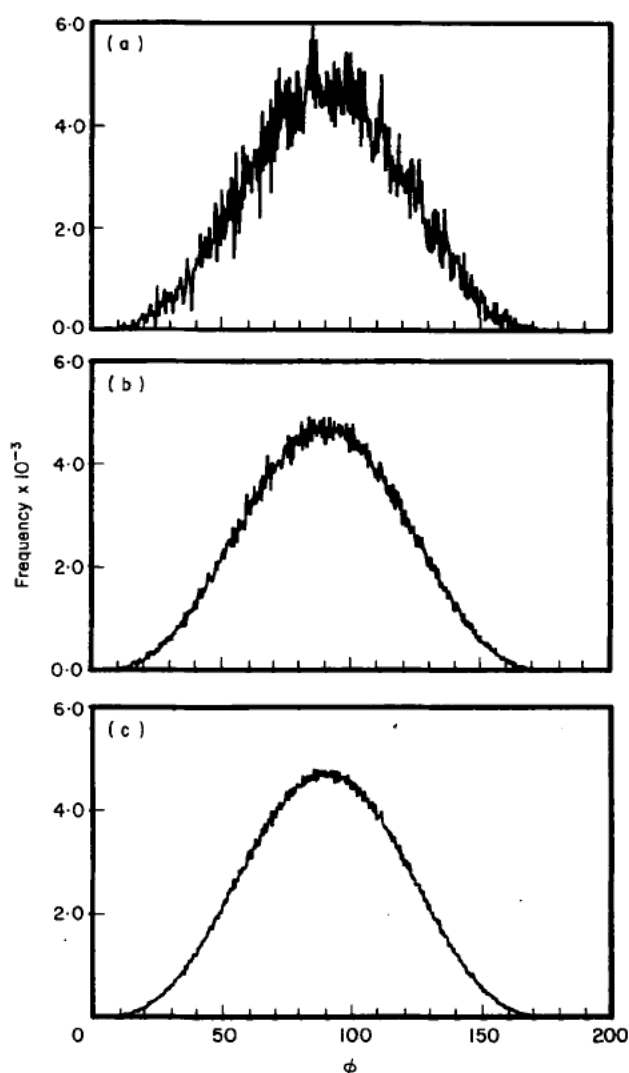


FIG. 10. Test of number of random points needed to describe all possible directions of the five-dimensional vector  $\Delta b$ , for the hybridoma system. Number of random points: (a)  $5 \times 10^5$ ; (b)  $1 \times 10^6$ ; (c)  $5 \times 10^6$ .

in the *E. coli* system. However, the decrease in sensitivity in the hybridoma system is not as dramatic as it is with the *E. coli* system. A drop in the minimum sensitivity by almost an order of magnitude results when six additional fluxes are measured. Even with six additional measurements, error amplification of more than 100 results.

The condition numbers of the flux combinations corresponding to the sensitivities shown in Fig. 13(a) are shown in Fig. 13(b). The condition numbers decrease with additional measurements, as do the sensitivities, indicating that additional flux measurements will most likely result in improved computational stability, regardless of the flux measurements.

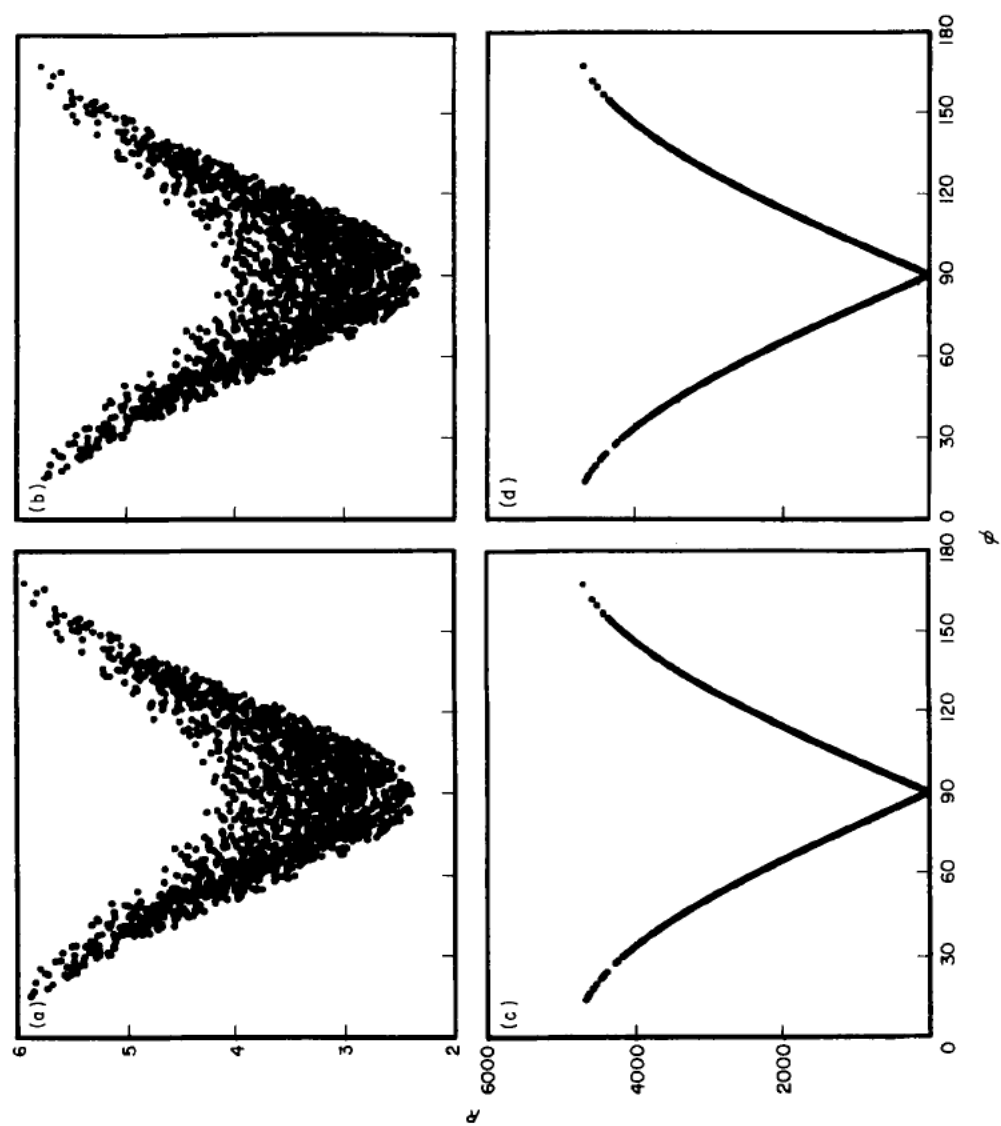


FIG. 11. The sensitivity  $R$  of the hybridoma system, calculated using  $b$ , measured at 30% DO and randomly selected  $\Delta b_i$ . Remaining fluxes were calculated using the objective functions: (a), (c): minimize total ATP production; (b), (d) maximize total ATP production. (a), (b): low condition number system; (c), (d): high condition number system.

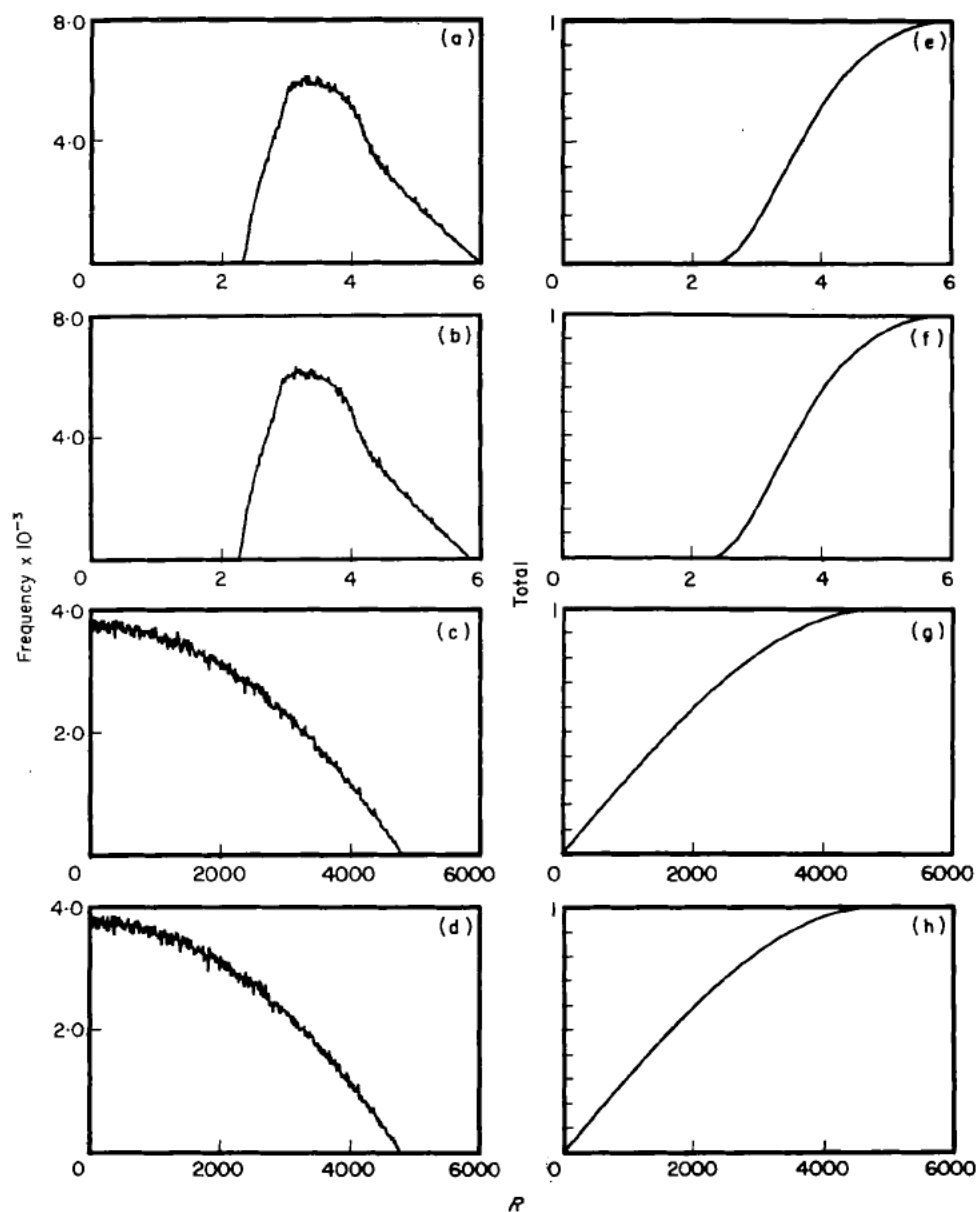


FIG. 12. Histograms of the frequency of occurrence of  $R$ , (a)–(d), and the integrals of the corresponding histograms, (e)–(h), for the hybridoma system. Remaining fluxes were calculated using the objective functions: (a), (c): minimize total ATP production; (b), (d): maximize total ATP production; (a), (b): low condition number system; (c), (d): high condition number system.



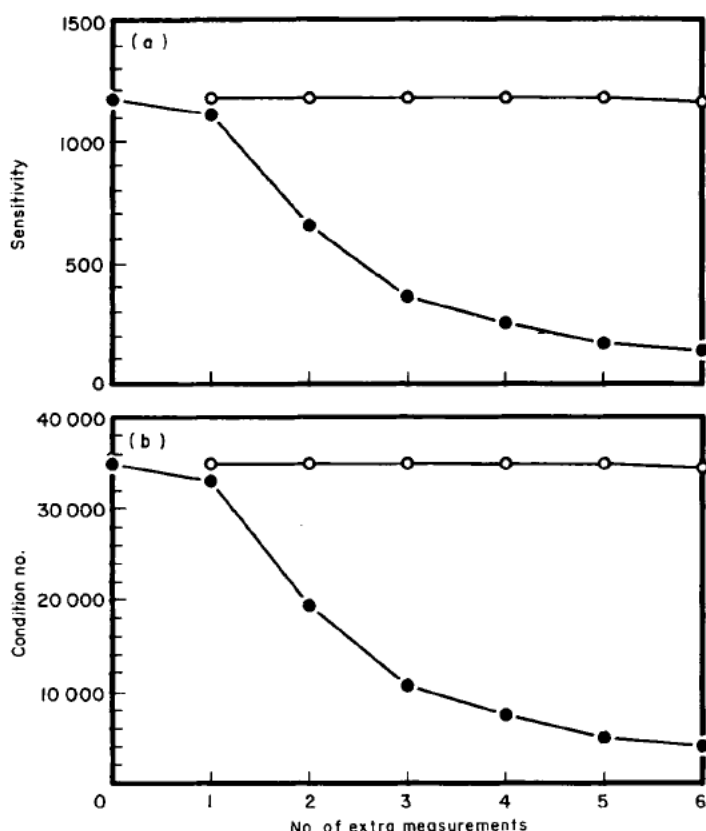


FIG. 13. (a) The lowest and highest sensitivities ( $R_n$ ) obtained using  $\Delta b_i = \Delta b_{i,max}$  for the hybridoma system as a function of the number of flux measurements; open circles: highest sensitivity; closed circles: lowest sensitivity. (b) The condition numbers that correspond to the systems in part (a). The maximum  $R_n$  and  $C_n$  were  $3.0 \times 10^8$  and  $9.1 \times 10^9$ , respectively.

### Discussion

The condition numbers that were calculated for the two biological models indicate that there are potentially very poor choices for the selection of *in vivo* metabolic fluxes to be measured. In the case of the hybridoma cell, there are choices that can lead to a 100 000-fold amplification of experimental error in the computation of the unmeasured fluxes. The possible amplification of error in the *E. coli* model (60-fold) is much smaller than in the hybridoma cell, but is still significant. The stoichiometric matrices of the hybridoma and *E. coli* systems are both sparse, but the hybridoma matrix is much larger than the *E. coli* matrix, which may be the cause of the significantly greater sensitivity of the hybridoma system. In both systems the majority of the flux combinations yield condition numbers at the low end of the spectrum (Figs 4 and 9), indicating that there are numerous choices of experimental design that result in relatively low sensitivity.

The results presented in Tables 5 and 7 show that certain reactions are prevalent in the low-condition number selections and other reactions are prevalent in the high-condition number selections. For example, in the *E. coli* model, isocitrate dehydrogenase (ICITdh) and biomass synthesis from  $\alpha$ -ketoglutarate and acetyl-CoA are prevalent in the high-condition number selections but are absent from the low-condition number selections, indicating that one should avoid measuring these reactions. The result that ICITdh should be avoided is noteworthy, since this enzyme has been suggested to be a key regulatory enzyme of the TCA cycle. Closer inspection of the role of ICITdh in the network reveals that flux through this reaction influences the fluxes throughout the TCA cycle, the glyoxylate shunt, and the biosynthetic rates. Thus, any error in ICITdh is magnified in the calculation of the other fluxes, and this characteristic causes the reaction to be a poor choice for measurement.

The reactions of pyruvate dehydrogenase, aconitase,  $\alpha$ -ketoglutarate dehydrogenase, malate dehydrogenase, PEPCK, malic enzyme, and acetate uptake result in the lowest condition number case, and thus are the best candidates for measurement. The presence of some reactions in combinations that lead to both low and high sensitivity, such as  $\alpha$ -ketoglutarate dehydrogenase and malate dehydrogenase, demonstrates that the sensitivity is a systemic property. In other words, sensitivity is a function of reaction combinations rather than individual reactions. However, if a reaction occurs only in low sensitivity systems and not in high sensitivity systems, then all combinations which contain this reaction are good choices for measurement, and consequently this reaction is always a good selection for measurement. Examples of this type of reaction are citrate synthetase and aconitase, which occur in many of the low condition number systems and in few of the high condition number systems.

In the hybridoma cell system, the reaction involving fatty acid metabolism ( $v_{11}$ ) is prevalent in the set of reactions that should not be measured. Fatty acids are very energy intensive on a molar basis, causing the network fluxes to be sensitive to errors in these measurements. The transhydrogenase ( $v_{26}$ ) and malate- $\alpha$ -ketoglutarate transporter ( $v_{41}$ ) are also poor choices for measurement. The reactions of malic enzyme ( $v_7$ ), ATP transport ( $v_{44}$ ), and ATP dissipation ( $v_9$ ) are good selections for measurements. On the other hand, ICITdh ( $v_{27}$ ), which was a poor choice in the *E. coli* system, is neither a good nor poor choice for measurement in the hybridoma system. This result may indicate that, due to the size and complexity of the hybridoma network, ICITdh does not have such a key role in this system as it did in the *E. coli* system.

The condition number is only a measure of the maximum possible error amplification. The condition number is always greater than the system's actual sensitivity by at least a factor of 2, as shown in Fig. 2(c) in part I, and this disparity between condition number and sensitivity depends on the actual flux measurements. When approximate numerical values are available, one can better estimate the system sensitivity. The results from these calculations demonstrate that under typical physiological conditions, the error amplification can be from 1/4 to less than 1/10 of that given by the condition number. Although the sensitivities are much lower than the condition numbers, there is significant variation in the sensitivities. The relationship between  $C$  and  $R$  is nearly linear, as shown in the examples examined here, and also

for the red blood cell network examined elsewhere (Savinell, 1991). The expected sensitivity indicates that the flux combinations resulting in low condition number have, in general, lower sensitivities than combinations with greater condition numbers. This proportionality between  $R$  and  $C$  demonstrates that, in the application of this algorithm, it is unnecessary to calculate  $R$  using flux estimates, since the flux combinations that result in low condition numbers are also likely to result in low sensitivities.

The sensitivities calculated using estimates of the fluxes are still upper bounds on the sensitivity, albeit more restrictive bounds than the condition numbers, since the sensitivity is also influenced by the error values. A more thorough analysis of the system sensitivity was obtained by examining the influence of the measurement error on the sensitivity. The exact  $R$  cannot be calculated since the error cannot be measured. By calculating  $R$  for randomly selected error combinations, we determined the probability that  $R$  has particular values. For the *E. coli* system, the most probable sensitivity is nearly the same for both the worst and best set of fluxes (Fig. 7). However, there is still a significant probability that sensitivities as great as 4–6 may result when the worst set of fluxes is measured. This result indicates that, in this case, a small advantage exists for measuring the combination of fluxes given by the low condition number, and this advantage must be weighed against possible advantages of measuring the high condition number fluxes. Possible advantages of measuring high condition number fluxes which counteract the harmful effects of the higher sensitivity include: (1) the fluxes may be more easily measured, and (2) the measurements may be made with greater accuracy.

The relative advantage of the low condition number flux combinations over the high condition number flux combinations is much greater with the hybridoma system. When the best set of fluxes in the hybridoma system are measured, the most likely sensitivity is relatively low ( $\approx 4$ ). When one of the worst set of fluxes is measured, there is an 80% probability that the sensitivity may be above 1000 (Fig. 12), which is an undesirably high sensitivity. Consequently, in this example, the experimental error propagation by the computations is significantly reduced by measuring fluxes given by the low condition number system, even though a sensitivity of 4.5 may still be unacceptably high. The sensitivity, and thus the possibility of erroneous results, may be reduced by making additional measurements.

The relationships between sensitivity and condition number, as shown in Fig. 4, for the *E. coli* system are similar for the three sets of data, even though fluxes through some pathways, such as the glyoxylate shunt, change significantly from growth on acetate to growth on glucose + acetate. The same correspondence between data sets is also observed between the two data sets used with the hybridoma cell system, even though the glycolytic flux changes significantly between aerobic and anaerobic conditions. It was shown in Fig. 2(c) in part I that the flux values can significantly affect the sensitivity. However, results from Figs 4 and 9 indicate that the different sets of experimental data representing large changes in cell behavior (i.e. anaerobic vs. aerobic growth) for the same system result in similar relationships between  $C$  and  $R$ . Therefore, these sets of data are actually very similar to each other, in comparison with the sets of fluxes that are mathematically, but not physiologically, possible. This

result indicates that the values of  $R$  can be estimated reasonably well before the experiment is undertaken, even if only poor estimates of the fluxes are available.

The condition numbers and sensitivities of systems with the number of flux measurements greater than the minimum number demonstrate that additional measurements reduce the experimental error propagation by the computations. In the *E. coli* system, four additional measurements guarantee that experimental error will be suppressed in the calculations. With more than four additional measurements, the decrease in sensitivity is minimal, and is probably not worth the extra effort required for the additional measurements. In the hybridoma system, a significant reduction in sensitivity also may occur with four additional measurements. However, even with this number of extra measurements, the error amplification is more than 100-fold, indicating that it is very difficult to obtain error suppression with this system.

### Conclusions

Our analysis based on two biologically important model systems reveals that, based on stoichiometry alone, certain reactions within the metabolic network have the potential to significantly influence the fluxes through other reactions in the network. Choosing these reactions for direct experimental determination and using such data for calculation may thus lead to highly erroneous results. Given the large amount of effort that is required to obtain flux data, it is worthwhile to use the algorithm developed in part I in order to design an experiment such that the propagation of experimental error by the computations is minimized. If a large spread in the condition numbers exists, these results indicate the fluxes that may give the most accurate results. To obtain a better estimate of the accuracy of the calculated fluxes, estimates of the flux data can be used to calculate the more restrictive bound on the sensitivity. It was shown that the flux combinations with the lowest condition numbers are also most likely to have the lowest sensitivities. In some instances the change in sensitivity between the high and low condition number systems is relatively small, so that there is not a significant advantage to measuring the low condition number fluxes. However, this algorithm is particularly useful when large systems are considered, since these systems have the potential to be extremely sensitive to experimental error.

This research was supported by National Science Foundation grant BCS-9009389.

### REFERENCES

- EISPACK (1979). *Eispack, Numerical Analysis and Applications Software*. Argonne: Argonne National Laboratory.
- OZTURK, S. S. & PALSSON, B. O. (1990). Effects of dissolved oxygen on hybridoma cell growth, metabolism, and antibody production kinetics in continuous culture. *Biotech. Prog.* **6**, 437–446.
- SAVINELL, J. M. (1991). *Analysis of stoichiometry in metabolic networks*. Ph.D. Thesis. The University of Michigan, Ann Arbor, MI, U.S.A.
- SAVINELL, J. M. & PALSSON, B. O. (1992a). Optimal selection of metabolic fluxes for *in vivo* measurement. I. Development of mathematical methods. *J. theor. Biol.* **155**, 201–214.

- SAVINELL, J. M. & PALSSON, B. O. (1992*b*). Network analysis of intermediary metabolism using linear optimization. II. Interpretation of Hybridoma cell metabolism. *J. theor. Biol.* **154**, 455–473.
- WALSH, K. & KOSHLAND, D. E. (1985). Branch point control by the phosphorylation state of isocitrate dehydrogenase. *J. biol. Chem.* **260**(14), 8430–8437.