245

A PATH ANALYSIS OF THE CAUSAL RELATIONSHIPS BETWEEN RED CELL GLYCOLYTIC INTERMEDIATES, ADP AND ATP IN SICKLE CELL ANEMIA*

PATRICIA P. MOLL, CHARLES F. SING and GEORGE J. BREWER

Department of Human Genetics University of Michigan Ann Arbor, MI 48109, U.S.A.

The statistical relationships among the glycolytic intermediates (GIs) of the Embden-Meyerhof pathway, adenine nucleotides (ANs) and various hematological measures were estimated for 34 sickle cell anemia patients. Heterogeneity in linear and quadratic regressions of hemoglobin and hematocrit, both singly and jointly, on the GI and AN variables implied 1) that any single formula to standardize optical density measures of the GIs and ANs on a per gram hemoglobin or per liter cell water basis would not uniformly remove hemoglobin and hematocrit effects; 2) that ignoring significant hematological effects could bias the estimates of correlation among GIs and ANs; and 3) that hemoglobin and hematocrit measures do not reflect the same source of variability.

The correlations among the GIs and ANs, after adjustment for hematological variability, were analyzed by path analysis to determine which of five proposed path models for cause and effect relationships were compatible with the data. AMP had a greater influence on ADP (coefficient of determination (CD) = 23%) than all the GIs together, while G6P and ADP influenced ATP variability the most (CD = 33% and 12%). The contributions of unknown factors to ADP and ATP variability were large for all models (CD = 56—77%) possibly due to stress of sickle cell disease. The path model with AMP and the four GIs (G6P, F6P, FDP, DHAP) influencing ADP variation, and the same GIs and ADP influencing ATP was the model most compatible with the data.

1. Introduction

An earlier study of sickle cell anemia patients revealed that the correlation between pairs of red blood cell glycolytic intermediates (GI)** differed depending on whether the pair was situated across either a rate-limiting or non-rate-limiting enzymatic step

(Oelshlegel et al., 1971). The other correlations among the GIs, hematological measures and adenine nucleotides (AN) suggested possible hypotheses about glycolytic control. The goal of this study is to utilize the statistical relationships among certain red cell GIs of the Embden-Meyerhof pathway and red cell AMP, ADP, and ATP for sickle cell anemia patients to identify which GIs influence ADP and ATP variability. Glycolysis provides the energy for various red cell functions through the net synthesis of two molecules of ATP from ADP for each molecule of glucose passing through the pathway. It has also been hypothesized that glycolysis regulates oxygen transport by hemoglobin through the production of DPG (Brewer, 1974).

ATP is known to play a number of important roles in red cell structure and function. In the mature red cell ATP is involved in maintaining cell shape, transporting sodium and potassium, and modifying the oxygen affinity of hemoglobin. ATP also plays a role

---

* This paper is Part II of the series: Red Cell Glycolytic Intermediates in Sickle Cell Anemia.
** Abbreviations used in this paper: G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; FDP, fructose-1,6-diphosphate; DHAP, dihydroxyacetone phosphate; DPG, 2,3-diphosphoglycerate; 3PG, 3-phosphoglycerate; PEP, phosphenol pyruvate; PYR, pyruvate; AMP, adenosine monophosphate; ADP, adenosine diphosphate; ATP, adenosine triphosphate; Hb, hemoglobin; Hct, hematocrit; MCHC, mean cell hemoglobin concentration; RETC, reticulocytes; Chr 51, chromium 51; PGI, phosphohexose isomerase; PFK, 6-phosphofructokinase; PK, pyruvate kinase; AN, adenine nucleotides, GA3PD, glyceraldehyde 3-phosphate dehydrogenase.

in the phosphorylation of sugar, the synthesis of purine and pyridine nucleotides, and the regulation of glycolysis by allosteric inhibition of PFK (Brewer, 1974). Statistically significant ATP and DPG variation among normal individuals and among sickle cell anemia patients has been reported (Brewer, 1967; Oelshlegel et al., 1977). Analyses of correlations between relatives suggest that the individual variability of ATP and DPG may be in part due to genetic differences (Gilroy, 1977).

Since previous work has shown a relationship between the GIs, ANs and hematological variables (Oelshlegel et al., 1977), the effects of hematological variability on the GIs and ANs must be considered first. Polynomial regression analysis is utilized to adjust each GI and AN for variation in measures of hematological status. Path analysis is then employed to compare several models which we have hypothesized to explain the causal relationships between the GIs and ADP and ATP variation. Path analysis is a different approach from the theoretical modeling of human erythrocyte glycolysis by an analysis of a system of differential equations (Rapoport and Heinrich, 1975; Heinrich and Rapoport, 1975). Path analysis will help identify which apriori model of cause and effect relationships are most consistent with the observed data. The analyses described below will help answer two questions: 1) are all GIs and ANs equally influenced by hematological variability, and 2) do cause and effect relationships exist among the GIs and ANs?

2. Methods

2.1 The sample

Blood samples from 42 individuals with sickle cell anemia, taken between July 24, 1972 and September 18, 1975, were assayed. One individual was sampled 4 times, 8 were sampled 3 times, 8 were sampled twice and 25 were sampled once, for a total of 69 observa-

tion sets. Each blood sample was measured for the 8 GI variables (G6P, F6P, FDP, DHAP, DPG, 3PG, PEP and PYR), 3 ANs (AMP, ADP, and ATP) and 5 concomitant variables: Hb, Hct, MCHC RETC and Chr 51. In addition, at the time of sampling, the concomitant variables height, weight and age of each individual were recorded. Certain ANs and various concomitant variables were not measured on every blood sample. The sample of individuals, consisting of 22 males and 20 females, varied in age from 16—74.5 yr at time of sampling. Informed consent was obtained from each patient and this research was carried out according to the Declaration of Helsinki.

2.2. Methods of assay and definition of measurement

The methods of assay are the same as reported by Oelshlegel et al. (1977). The levels of GIs and ANs were measured by end-point analysis of a perchloric acid extract of homogenized red blood cells. The change in optical density per ml of a whole blood extract in the presence of an excess of enzyme was measured and then converted to mol per 100 ml whole blood, except for PYR which is in μmol/l whole blood.

This measure, rather than μmol/g Hb or the concentration of GIs and ANs per liter cell water, was chosen so that individual variability in Hb and Hct could be eliminated completely before the path analysis. The standard correction of per g Hb does not completely correct the GIs and ANs for individual variability in Hb and Hct.

2.3. Analysis of data

A path analysis is based on estimates of linear regression and linear correlation between pairs of variables in a network defined by a path diagram (model). If the distributions of 2 variables are not each normal, then their joint distributions can not be bivariate normal (Sokal and Rohlf, 1969). Probability

statements associated with tests of correlation will be appropriate only if the 2 variables are bivariate normal (Sokal and Rohlf, 1969). Skewness in the distribution of a variable can also make it difficult to distinguish a skewed distribution composed of one population from a distribution composed of comingled populations with different means and different relative contributions to the population distribution (MacLean et al., 1976). Since the analysis presented here is based on correlation among variables, rather than search for evidence of admixture, simple power transformations (Sokal and Rohlf, 1969) of the general form $y = x^a$ rather than specialized power transformations (MacLean et al., 1976; Morton et al., 1977) were employed to remove skew. Positive powers were used to avoid reversing the ranking of observations.

The correlation between 2 variables in a path model may also be due to the common influence of one of more concomitant variables which are not a part of the system of variables being studied. All transformed GIs and ANs were adjusted by multiple polynomial regression for the concomitant variables which accounted for a statistically significant portion of variability (Li, 1964).

Sing et al. (in prep., (a)) have shown that each of the GIs and ANs considered here is highly repeatable over time for each individual. Therefore, the mean of the transformed adjusted data for each of the 42 individuals has been used in an analysis of the genetic relationships among the individuals (Sing et al., in prep., a). However, 8 individuals were eliminated from the multivariate analysis presented here because AMP was not measured. The averages of the 8 GIs and 3 ANs for each of the 34 individuals were used in the path analysis. The correlations between these variables were used to estimate the causal relationships defined by specific a priori path models. The 5 models considered will be presented with the results of the path analysis. Each analysis of a path model provides a partition of the variation of ADP and ATP into the orthogonal effects of AMP and each of

the GI variables and of the joint effects of the GI variables. The mathematics and strategy for applications of path analysis to a path model are discussed in great detail elsewhere (Wright, 1968; Kempthorne, 1969; Li, 1975, 1976). Path analysis has been useful in evaluating the roles of variables in other red cell studies (Brewer et al., 1974).

2.4. Results

The sampling date was not a statistically significant source of variability among individuals for any of the variables. Also, there was no significant difference between males and females for any GI or AN. The mean, standard deviation, minimum, maximum, skew and kurtosis for each distribution of GI, AN and concomitant variable are presented in Table 1. This table also gives the measurement scale of each variable used and number of observations (N). While the mean GI and AN values are also given in per gram Hb units, these values are for references only and were not used in the analysis.

3. Regression adjustments

Using an $\alpha = 0.05$ critical value, the FDP, DHAP, PYR, ADP, AMP, age and Chr 51 distributions exhibited significant positive skewness while ATP was significantly negatively skewed (G1 in Table 1). The power of the transformations used to remove skew are also presented in Table 1. FDP, DHAP, ADP, age and MCHC had leptokurtic distributions (G2 in Table 1). After transformation, the distributions of the 34 individual means for each variable did not deviate significantly from normality.

Hb and Hct each account for the greatest portion of variability for most of the GIs and ANs. The contributions to variation of these two concomitants ranged from 0.8 to 61%. Age, RETC, and Chr 51 taken separately also accounted for statistically significant variation in DPG. Chr 51 accounted for 68% of the variability in AMP. ADP was the variable

TABLE 1

A description of the raw variables.

	N	Mean Per g Hb	Mean μmoles per 100 ml whole blood		Standard Deviation	Minimum	Maximum	G1[d]	G2[e]	Transformation power	G1 New	G2 New
G6P	69	0.245	2.072		0.568	0.909	3.497	0.288	0.211			
F6P	69	0.060	0.511		0.186	0.166	1.066	0.567	−0.032			
FDP	69	0.013	0.114		0.083	0.007	0.410	1.580[i]	2.650[i]	0.25	0.112	0.073
DHAP	69	0.063	0.536		0.387	0.062	2.323	2.595[i]	8.410[i]	0.25	0.476	1.560
DPG	69	21.391	181.700		37.129	85.305	262.010	−0.520	0.209			
3PG	69	0.193	1.626		0.581	0.083	2.971	−0.153	0.241			
PEP	69	0.054	0.463		0.151	0.108	0.847	0.043	−0.161			
PYR	69	51.900[a]	51.900[a]		23.335	3.000	116.700	0.691[g]	−0.164	0.5	0.009	0.377
AMP	47	0.125	1.039		0.393	0.246	2.317	0.755[h]	1.357		−0.008	0.693
ADP	57	0.831	6.938		3.251	2.313	21.675	2.545[i]	8.031[i]	0.5[f]	0.068	1.862
ATP	66	5.089	43.140		13.357	4.331	66.197	−0.921[i]	1.074	2.0	0.133	−0.406
AGE	68		27.41	years	14.985	16.000	74.500	1.745[i]	2.205[i]	2.0	−0.155	−1.226
Hb	69		8.527	g	1.541	5.170	11.800	−0.235	−0.244			
Hct	69		24.920	Cell vol. %	4.759	15.200	35.900	−0.049	−0.588			
MCHC[b]	69		34.326		2.934	25.400	44.400	0.467	2.447[i]			
RETC	51		15.939	%	6.202	4.500	28.600	0.163	−0.882			
CHR51[c]	20		8.050		3.179	5.000	15.500	1.339[h]	0.751	−0.6	−0.478	−0.619
Weight	56		118.190	pounds	22.191	60.250	174.000	−0.175	0.692			
Height	51		164.750	cm	9.432	144.000	185.000	−0.592	−0.224			

[a] μmoles per liter whole blood
[b] MCHC is measured in grams of H3 per 100 ml packed red cells
[c] CHR51 is measured in % labeled red cells remaining
[d] G1 is a measure of skewness
[e] G2 is a measure of kurtosis
[f] After a $\log_e$ transformation.
g = 0.05 level of probability.
h = 0.01 level of probability.
i = 0.001 level of probability.

TABLE 2

Standardized partial regression coefficients. Each independent variable is adjusted for the other concomitants in the multiple polynomial regression.

	Hb	Hb^2	Hct	Hct^2	R^2
G6P	0.85	−0.64	1.82	−1.64	0.22^b
F6P	0.01	0.28	0.17	−0.05	0.16^a
FDP	1.13	−1.45	0.26	0.05	0.06
DHAP	−0.18	−0.23	2.61	−1.94	0.18^b
DPG	1.73	−1.39	0.03	0.44	0.64^c
3PG	2.11	−1.68	−1.96	1.85	0.11
PEP	1.11	−0.58	0.15	−0.12	0.25^b
PYR	2.39	−2.33	−0.48	0.37	0.06
AMP	−1.56	1.51	7.73	−7.42	0.39^c
ADP	0.43	−0.24	2.08	−1.82	0.26^b
ATP	−0.89	1.30	0.42	−0.19	0.40^c

^{a} = 0.05 level of probability.
^{b} = 0.01 level of probability.
^{c} = 0.001 level of probability.

which had a significant quadratic regression on height. After adjustment for Hb, Hb^2, Hct and Hct^2 effects, none of the other concomitants including age, MCHC, Chr 51, RETC, weight and height accounted for a significant amount of variability in the GIs and ANs. Table 2 gives the standardized partial regres-

sion coefficients (linear and quadratic) used to adjust each GI and AN for Hb and Hct. Table 2 also has the R^2 for the total regression equation with Hb, Hb^2, Hct and Hct^2. The total R^2 values for GIs ranged from 64% for DPG to 6% for FDP and PYR. The R^2 for the ANs was approximately 35%. Only DHAP, DPG and AMP had significant contributions to R^2 by Hct after these variables were adjusted by Hb and Hb^2.

The variability in standardized regression coefficients reveal the unique relationships of each of the GIs and ANs with Hb and Hct. Hb and/or Hb^2 contributed the largest share of the total R^2 for PEP, PYR, F6P, DPG, FDP and ATP. After adjustments for Hb and Hb^2, Hct and Hct^2 still contributed a large share to the variability of ADP, AMP, G6P and DHAP. These results imply that Hb and Hct do not measure the same source of variability in these GIs and ANs.

4. Path analysis

The upper half of Table 3 presents the correlations among GIs and ANs based on the transformed unadjusted values for the 34 individuals, while the bottom half of

TABLE 3

Correlation coefficients. The bottom half of the table is based on the transformed adjusted means of 34 individuals (partial correlation coefficients). The upper half is based on transformed unadjusted means of 34 individuals (total correlation coefficients).

	G6P	F6P	FDP	DHAP	DPG	3PG	PEP	PYR	AMP	ADP	ATP
G6P	1.0	0.85^b	0.18	0.18	0.47^b	0.30	0.20	0.18	0.40^a	0.48^b	0.54^b
F6P	0.85^b	1.0	0.02	−0.06	0.30	0.26	−0.02	0.03	0.28	0.41^b	0.46^b
FDP	0.17	0.01	1.0	0.51^b	0.04	0.01	0.18	0.32	0.06	0.03	0.04
DHAP	−0.06	−0.28	0.58^b	1.0	0.51^b	0.35^a	0.50^b	0.22	0.24	0.12	0.33
DPG	0.14	0.01	0.04	0.21	1.0	0.43^a	0.57^b	0.15	0.35^a	0.41^b	0.64^b
3PG	0.09	0.12	−0.02	0.18	0.15	1.0	0.44^b	0.03	0.38^a	0.14	0.26
PEP	−0.21	−0.36^a	0.21	0.33	0.02	0.27	1.0	0.39^a	0.36^a	0.32	0.35^a
PYR	0.08	−0.03	0.33	0.27	0.12	−0.03	0.39^a	1.0	0.16	0.05	0.01
AMP	0.22	0.20	0.01	0.08	−0.17	0.08	0.02	0.07	1.0	0.59^b	0.34^a
ADP	0.25	0.29	0.01	−0.23	−0.27	−0.19	−0.20	−0.08	0.48^b	1.0	0.60^b
ATP	0.47^b	0.36^a	0.09	0.02	0.03	−0.07	−0.28	−0.06	0.26	0.40^b	1.0

^{a} Statistically significant at the 0.05 level of probability.
^{b} Statistically significant at the 0.01 level of probability.

Table 3 presents the correlations among the GIs and ANs, based on the means of transformed and regression adjusted values. The correlations in the top half of the table will be discussed later. Only the correlations in the bottom half of the table were used in the path analysis. A multivariate test of the correlation matrix (Anderson, 1958) showed

the variables are not all mutually independent and the GIs and ANs are not independent from each other. The G6P-F6P, ATP-G6P, AMP-ADP, DHAP-FDP correlations are significant at the 0.01 level of probability. Using a 0.05 test criterion, the correlations of PEP with PYR and with F6P and the correlations of ATP with F6P and with ADP were also judged to be statistically significant.

The path models considered here assume that the GIs and AMP influence ADP and that the GIs and ADP influence ATP. These assumptions are based on postulated roles of the GIs and ANs in the Embden-Meyerhof
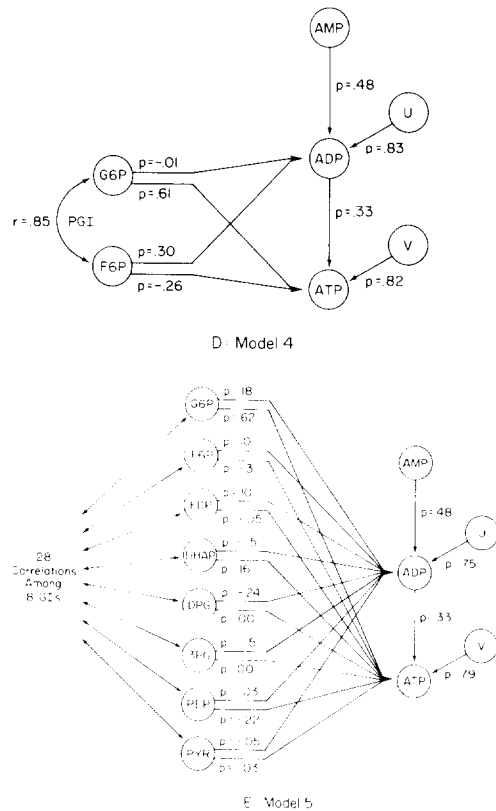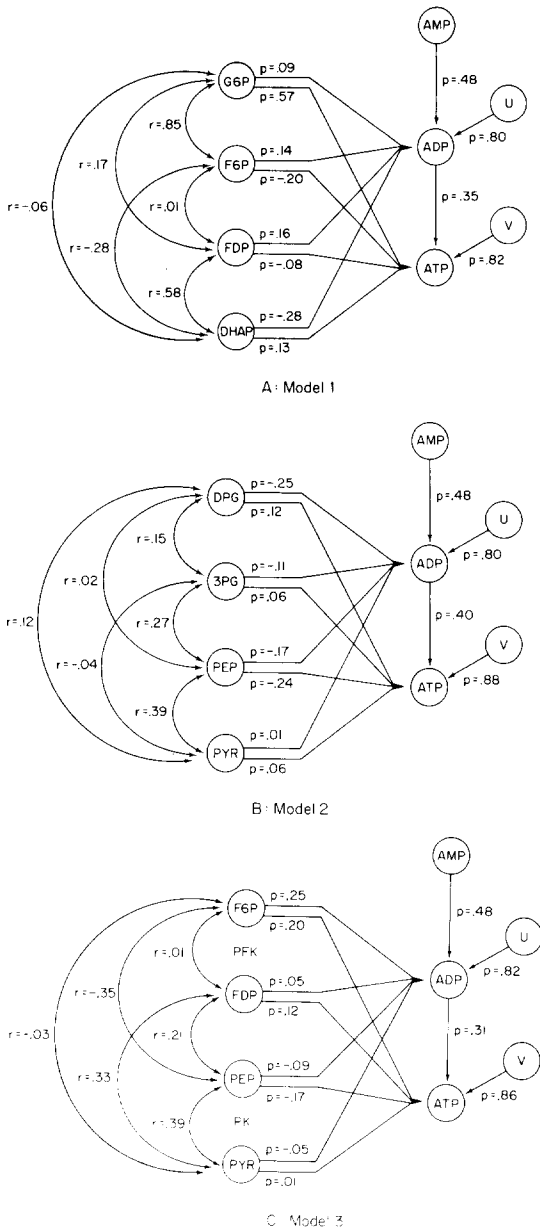

Fig. 1A—E. The models for interrelationships among potential causal variables. GIs and AMP, and ADP and ATP used in the path analysis. The Models 1 through 5 are discussed in the text. U and V represent all unknown and unmeasured sources of ADP and ATP variability, respectively. A single headed arrow indicates a path from one variable to another. A double headed arrow indicates a correlation between two variables.

pathway. AMP is not in the pathway and should thus be involved in ADP and ATP variation directly and not in association with the GIs. The first model (Fig. 1A) assumes that GIs in the front half of the pathway (G6P, F6P, FDP, DHAP) are influencing ADP and ATP, while a second model (Fig. 1B) assumes GIs in the last half of the pathway (DPG, 3PG, PEP, PYR) affect ADP and ATP. These 2 models are suggested by other analyses of these data which showed that highly significant differences exist among individuals for GIs in the front half but not the back half of the pathway (Oelshlegel et al., 1977). The third model (Fig. 1C) assumes that GIs on either side of the rate limiting enzymes PFK and PK influence ADP and ATP, while the fourth model (Fig. 1D) assumes that GIs on either side of the non-rate limiting enzyme PGI are influencing ADP and ATP. These models were chosen because the rate limiting enzyme PFK is involved in an ATP to ADP conversion, while PK is involved in an ADP to ATP conversion. The non-rate limiting enzyme PGI is not directly involved in any energy conversion. Including the GIs on either side of GA3PD (a non-rate limiting enzyme) would have made Model 4 identical to Model 1. The fifth model (Fig. 1E) assumes that all the GIs measured affect ADP and ATP. The last model is considered the complete model while the others are defined as reduced models. U and V represent all unknown and unmeasured sources of ADP and ATP variability, respectively.

The estimates of the path coefficients for each model are given in Fig. 1. The paths from G6P and F6P to ATP in Fig. 1A (Model 1) will be used to illustrate the interpretation of path coefficients. The positive path from G6P implies a like-signed deviation in the affected variable ATP on the average. The negative path from F6P indicates that a positive deviation in F6P will cause a negative deviation in ATP (and vice versa). A comparison of these 2 scale-free path coefficients reveals that G6P has a larger relative degree of hypothesized causation due to its larger

absolute influence on ATP deviations. Path coefficients from the unknown variables U and V are 0.80 and 0.82 respectively in Fig. 1A. Path coefficients from unknown variables are assumed to be non-negative (Li, 1975). The total variance of an affected variable is partitioned using the concept of degree of

TABLE 4

Coefficients of determination.

	Model 1	Model 2	Model 3	Model 4	Model 5
Single variables affecting ADP					
AMP	0.23[c]	0.23[c]	0.23[c]	0.23[c]	0.23[c]
G6P	0.01			0.00	0.03
F6P	0.02		0.06	0.09	0.01
FDP	0.03		0.00		0.01
DHAP	0.08				0.02
DPG		0.07			0.06
3PG		0.01			0.02
PEP		0.03	0.01		0.00
PYR		0.00	0.00		0.00
Fraction of determination[a]					
AMP	0.23	0.23	0.23	0.23	0.23
GIs	0.13	0.12	0.10	0.09	0.21
Unknowns	0.64	0.65	0.67	0.68	0.56
Single variables affecting ATP					
ADP	0.12[b]	0.16[b]	0.09	0.11[b]	0.11
G6P	0.33			0.37[b]	0.38
F6P	0.04		0.04	0.07	0.10
FDP	0.01		0.01		0.00
DHAP	0.02				0.02
DPG		0.01			0.00
3PG		0.00			0.00
PEP		0.06	0.03		0.05
PYR		0.00	0.00		0.00
Fraction of determination[a]					
ADP	0.13 [b]	0.16	0.12	0.13 [b]	0.13
GIs	0.20	0.06	0.13	0.19	0.24
Unknowns	0.67	0.78	0.75	0.68	0.63

[a] The fraction of determination by all GIs and ADP is the sum of all single effects and joint effects. The joint effects may be positive or negative and are not shown in this table.
[b] Statistically significant at the 0.05 level of probability.
[c] Statistically significant at the 0.01 level of probability.

determination. Table 4 gives the squared path coefficients as measures of relative contributions (coefficient of determination) of the single hypothesized casual variables on the total variance of the affected variables. G6P and F6P singly account for 33% and 4% of ATP variability in Model 1. The joint influences on the affected variables are the products of connecting path coefficients and correlation coefficients. G6P and F6P jointly account for 19% of ATP variability (twice the product of paths from G6P and F6P to ATP and the correlation between G6P and F6P). This negative joint effect implies the 2 GIs together tend to reduce the variation in ATP because they account for ATP variability in opposite ways and are positively correlated. Table 4 also has a summary of the proportion of contributions to variability by all GIs jointly and unknown variables in the different models. For Model 1, ADP accounts for 13%, the GIs in the model jointly account for 20% and the unknowns account for 67% of ATP variability. To summarize, path coefficients indicate influence on deviations and coefficients of determination measure the contribution to variance of hypothesized causal variables on hypothesized affected variables. Because they are scale-free measures, the path coefficients and coefficients of determination can be compared among the different causal variables.

Considering all 5 models, the path coefficients connecting the causal variables with ADP reveal a consistent pattern. The paths from AMP and the first 3 GIs in the pathway to ADP are either positive or zero. The last 5 GIs all have either negative or zero path coefficients. In all models, AMP determines 23% of the variance in ADP directly while the unknown factors determine between 56 and 68% (Table 4). While the size of the path coefficients for each GI vary depending on which other GIs are in the model, these differences in coefficients are relatively unimportant, since any single GI account for a small portion (9% or less) of the variance of ADP. The GIs of any specific model taken

jointly account for between 9% (Model 4) and 21% (Model 5) of the variability in ADP (Table 4).

For ATP, the signs of the path coefficient from certain GIs (F6P and FDP) were not consistent from model to model. G6P always had large, positive path coefficients ranging from 0.57 to 0.62. ADP had the next largest path coefficients, 0.30 to 0.40. PEP had a negative path coefficient in every model, -0.23 to -0.17, while DHAP had coefficients ranging from 0.12 to 0.15. The path coefficients for F6P and FDP ranged in different models from negative to positive. The path coefficients from DPG to ATP ranged from 0.0 to 0.12 while the PYR and 3PG to ATP path coefficients were essentially zero.

A comparison of the coefficients of determination for ATP among models (Table 4) shows that, except for the unknown factors, G6P and ADP account for the most of ATP

TABLE 5

A partition of the percent of sums of squares for ADP and ATP.

Sums of squares (SS)	ADP	ATP
SS due to hematological variables (Hb, Hb², Hct, Hct²)	0.26433	0.39794
SS residual	0.73567	0.60206
SS due to Path Model 1	0.26484	0.19868
SS residual Path Model 1	0.47083	0.40338
SS due to Path Model 2	0.25748	0.13245
SS residual Path Model 2	0.47819	0.46961
SS due to Path Model 3	0.24277	0.15052
SS residual Path Model 3	0.49290	0.45154
SS due to Path Model 4	0.23541	0.19266
SS residual Path Model 4	0.50026	0.40940
SS due to Path Model 5	0.32369	0.22276
SS residual Path Model 5	0.41198	0.37930
SS total	1.0000	1.0000

variance, 33—38% and 9—16%, respectively. F6P accounts for 4—10% of ATP variance, PEP accounts for 3—6% of ATP variance while all other GIs account for 2% or less. Table 4 shows that taken together the GIs jointly contribute between 0.06 to 0.24% of ATP variability in the different models. All joint effects in these models were small except for the joint effect of G6P and F6P together on ATP mentioned earlier. Table 4 also shows which effects were statistically significant.

Table 5 is a summary of all the analyses expressed in percent of sums of squares of ADP and ATP. The hematological variables account for 26 and 40% of ADP and ATP sums of squares, while the GIs, after adjustment for hematological variability, account for 24—32% and 13—22% of ADP and ATP sums of squares through the path models. The residual sums of squares from the path model, assigned to unknown variables, represent between 38 and 50% of the total sums of squares.

5. Discussion

A regression analysis has revealed a unique relationship between the hematological variables and each of the GIs and ANs. The observed linear and quadratic regression equations varied among GIs and ANs both in the total variability accounted for by Hb and Hct variables and in the relative contributions of each to variability. Any general formula used to convert the measurement scale of GIs from μmol/100 ml to per gram Hb or per liter cell water would not uniformly remove the effects of Hb and Hct from each GI nor account for any of the quadratic effects. While no simple adjustment exists, a quadratic regression adjustment as a general correction may be a good first approximation. Failure to consider the heterogeneity of the Hb and Hct effects could result in a serious bias in the estimation of GI correlations and this bias could then be carried over into the path analysis. This

heterogeneity of Hb and Hct effects suggests the relative size of the red blood cells, the water content of the cells and the spacing between Hb molecules possibly have specific effects on certain enzymes or GIs in the pathway. From Table 5 the large effects of the hematological variables Hb and Hct on ADP and ATP variability can be seen.

Twenty-three significant correlations (15 at α = 0.01 and 8 at α = 0.05) were found among the GIs and ANs before adjustments for hematological variability while only 8 significant correlations (5 at α = 0.01 and 3 at α = 0.05) remained after the hematological adjustment was made as can be seen in Table 3. The significant correlations of ATP with AMP, PEP and DPG and the correlations of ADP with G6P and FDP were no longer present after the hematological adjustments. This result suggests that the associations like the one between DPG and ATP may not be a direct influence on DPG or ATP on one another but attributable to the common influence of Hb and/or Hct. The significant correlation between PEP and F6P ($r = -0.36$) was the only correlation that did not exist in the unadjusted data. All of the significant correlations reported earlier were found in the unadjusted data of this study. The correlations between intermediates grouped between rate-limiting enzymes (G6P and F6P, DHAP and FDP) were found in the earlier study (Oelshlegel et al., 1977), and in both the adjusted and unadjusted data here as was the correlation between AMP and ADP.

The several a priori path model cannot be contrasted using standard statistical testing procedures to determine the "best" model when a totally determined system is considered. As an alternative, Li (1975) suggests a subjective evaluation of models based on 2 criteria. The "best" model is chosen 1) to have the closest agreement between correlations not used in computation of the paths and their estimates from the path model and 2) to give the minimum contribution to determination by unknown factors. With respect to the first criterion, the only correla-

tion which can be contrasted is that between AMP and ATP (0.259). From the path models, the correlation between ATP and AMP is the product of paths connecting these 2 variables and ranges from 0.148 in Model 3 to 0.192 in Model 2. Based on this criterion, Model 2 is the best fit, Model 1 the next best, and Model 3 the worse fit to the data. Combining this criterion for selection of a model with the consideration of lowest unexplained determination for both ATP and ADP indicates that Model 1 yields the best fit of the reduced models. Consequently, Model 1 (with the GIs: G6P, F6P, FDP and DHAP influencing ATP and ADP variability) is the "best" reduced model of those proposed. The complete model has 4 more GIs than Model 1 but only accounts for 5.5% more of the ADP variance and 2% more of ATP variance, which indicates that variability in DPG, 3PG, PEP or PYR has little influence on ADP and ATP variability.

The models used in this analysis all consisted of fully determined systems. The approach to causal analysis suggested by Rao et al. (1976) uses maximum likelihood methods for overdetermined systems. These techniques provide maximum likelihood estimates of path coefficients and likelihood ratio tests of hypotheses. We could make our fully determined systems overdetermined by fixing the path coefficient from ADP to ATP. The approximate chi-square for the 5 models then ranged from 0.033 (Model 4) to 0.318 (Model 5), all very close and non-significant. This non-rigorous approximation supports our choice of a parsimonious model to explain ADP and ATP variability.

The path coefficients of Model 1 summarize the major effects of the GIs and AMP on ATP and ADP in this sample of data. The independent effects of AMP contribute the greatest influence in determining ADP variability. One interpretation of this effect would be that the synthesis of AMP by the salvage pathway of the red blood cell is an important determiner of ADP levels that in turn has some influence on ATP levels. G6P was the most influential GI on ATP variability. Since G6P is the first GI pool in the pathway after the rate limiting enzyme hexokinase, this large influence of G6P is consistent with the hypothesis that hexokinase is a regulating factor in red cell glycolysis (Brewer, 1974), and is therefore an important determinant of ATP variation. The possible mechanisms for that regulation could include allelic variation in hexokinase, variation in the amount of hexokinase enzyme in each cell, or variability in affectors of the enzyme. In a parallel study of normal individuals, both G6P and ATP had high heritabilities (Gilroy, 1977). The sources of variability in ATP are important because of the functions of ATP in the red blood cell. Besides the strong influence of G6P, ADP also had a strong influence on ATP variability. In these models ATP variability is more closely influenced by variability in glycolysis while ADP variability is more closely influenced by AMP variability.

The lack of influence by pyruvate on either ADP or ATP is consistent with the physiology of the system, since muscle tissue releases a large quantity of both non-phosphorylated pyruvate and lactate which can cross the red cell membrane. The phosphorylated GIs studied here cannot cross the red cell membrane. Thus, variability in pyruvate should have little effect on ATP or ADP variability. The other GIs in the latter half of the pathway (DPG, 3PG, and PEP) make relatively small contributions to ATP and ADP variability.

There are several explanations for the size of the unknown factors in these path models. Several other GI pools were not measured; the concentrations are known to be low and therefore difficult to measure. Second, path analysis can only deal with linear relationships among variables. However, a quadratic regression of each GI on ATP and ADP supports the adequacy of a linear model for these data. Since Hb and Hct have non-linear relationships with the GIs and ANs, these sources of variability were handled first by polynomial regression and not included in these models.

Third, the path models chosen for analysis may not reflect the appropriate relationships among the variables. For instance, ATP has an effect on ADP variability not depicted in the diagrams. A path from ATP to ADP could be included to show this relationship, but this path would reduce the unknown factor for ADP but not for ATP. There are a large number of alternative path models with ATP and ADP as the affected variables which could be used for analyzing different GI and AN relationships. Those studied here were chosen to represent a class of models suggested by a priori considerations. It is important to realize that the inferences which have been made are only valid for the models studied. Finally, a possible explanation for the large contribution of the unknown factors in these models is the stress that the red cells are under due to the disease sickle cell anemia. Some of the stresses of sickle cells include increased influx of calcium and sodium and efflux of potassium. The pumping out of calcium and of the cations potassium and soidum against the concentration gradients also requires the energy of ATP. Therefore, the calcium, sodium and potassium pumping may have a great influence on both ATP and ADP variability since ATP is converted to ADP in the process.

In addition to these unmeasured contributions to variability in the data there are unknown contributions which could not be controlled by the study design. Variability in the expression of sickle cell disease exists among the small sample of patients in this study as measured by crises experience and the number of irreversible sickled cells, a variable which was not available. While the data were adjusted for Hb and Hct differences among individuals, other sources of variability could potentially influence the causal relationships between the GIs and ANs.

The path analysis used in this study allowed us to treat the GIs and ANs as a system of variables and to incorporate certain a priori information on the direction of influence within the system. The path analysis also provided an orthogonal partition of the ADP and ATP variation into the singular effects of the hypothesized causal variables and their joint effects. Despite the uncontrolled sources of variability which exist in these data, evidence for strong relationships did emerge from the application of a path analysis which were not evident from the correlations alone. From the path analyses of Model 1, G6P had the most influence on ATP and the difference between the G6P and ADP effect on ATP was stronger in the path analyses than the correlation analysis. The negative joint effect of G6P and F6P on ATP also was not evident from a simple examination of the correlations. The system of variables with AMP affecting ADP variability and G6P and ADP affecting ATP variability emerges from this analysis. However, because this study is based on a sample from the sickle cell anemia population, no inferences can be made about general causal relationships which may exist between the GIs and ANs for normal individuals. A path analysis of data collected on normal individuals is underway (Sing et al., in prep., (b)). It will soon be possible to determine if the effects reported here are unique to patients under the stress of sickle cell disease.

References

Anderson, T.W., 1958, Introduction to Multivariate Statistical Analysis (Wiley, New York) pp. 230—246.

Brewer, G.J., 1967, Genetic and population studies of quantitative levels of adenosine triphosphate in human erythrocytes. Biochem. Genet. 1, 25—34.

Brewer, G.J., 1974, Red cell metabolism and function. In: The Red Blood Cell, N. MacN. Surgenor. (Ed.) Vol. 1, 2nd Edn., (Academic Press, New York) pp. 473—507.

Brewer, G.J., C.F. Sing, J.W. Eaton, et al. 1974, Effects on hemoglobin oxygen affinity of smoking in residents of intermediate altitude. J. Lab. Clin. Med. 84, 191—205.

Gilroy, T.E., 1977, Genetic Control of Glycolysis in Human Erythrocytes. Ph.D. Dissertation, University of Michigan.

Heinrich, R. and T.A. Rapoport, 1975, Mathematical analysis of multienzyme systems. II. Steady state and transient control. BioSystems 7, 130—136.

Kempthorne, O., 1969, An Introduction to Genetic Statistics (State University Press, Ames, Iowa) pp. 270—307.

Li, C.C., 1975, Path Analysis—A Primer (Boxwood Press, Pacific Grove).

Li, C.C., 1976, First Course in Population Genetics (Boxwood Press, Pacific Grove).

Li, J.C.R., 1964, Statistical Inference, Vol. 2. (Edward Brothers, Ann Arbor).

MacLean, C.J., N.E. Morton and R.C. Elston, 1976, Skewness in comingled distributions. Biometrics 32, 695—699.

Morton, N.E., C.J. MacLean, A. Kagen et al. 1977, Admixture in distributions of lipids and related variables. Am. J. Hum. Genet. 29, 52—59.

Oelshlegel, F.J., G.J. Brewer and C.F. Sing, 1977, Patterns of red cell glycolytic intermediates in sickle cell anemia. I. Values and first analysis. J. Mol. Med. 2—51.

Rao, D.C., N.E. Morton, R.C. Elston et al., 1976, Causal analysis of academic performance. Behav. Genet. 7, 147—159.

Rapoport, T.A. and R. Heinrich, 1975, Mathematical analysis of multienzyme systems. I. Modelling of the glycolysis of human erythrocytes. BioSystems 7, 120—129.

Sing, C.F., P.P. Moll and G.J. Brewer (in prep., a) Genetic Analysis of Glycolytic Intermediate Profiles in Sickle Cell Anemia.

Sing, C.F., T.E. Gilroy, G.J. Brewer and P.P. Moll (in prep., b) A Path Analysis of the Causal Relationships between Red Cell Glycolytic Intermediates in a Normal Population.

Sokal, R.R. and F.J. Rohlf, 1969, Biometry The Principles and Practice of Statistics in Biological Research (W.H. Freeman, San Francisco).

Wright, S., 1968, Evolution and the Genetics of Populations, Vol. 1 (University of Chicago Press, Chicago) pp. 299—372.