

Enduring Behavioral Changes in Rats with Experimental Phenylketonuria

(phenylalanine to tyrosine ratio/*p*-chlorophenylalanine/disease model/learning ability/brain structure)

ARNOLD E. ANDERSEN AND GORDON GUROFF

Section on Intermediary Metabolism, Laboratory of Biomedical Sciences, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20014

Communicated by Sidney Udenfriend, January 24, 1972

ABSTRACT Several criticisms of past attempts to produce experimental phenylketonuria are discussed, and a model that appears to meet these criticisms is presented. This model uses inhibition in rats of phenylalanine hydroxylase (EC 1.14.3.1) by *p*-chlorophenylalanine and supplementation with phenylalanine to produce a high ratio of phenylalanine to tyrosine in their blood. By this method, an experimental subject is produced whose behavioral, neurological, and biochemical characteristics are similar to those of clinical patients with phenylketonuria.

Experimental animal models of phenylketonuria have received much attention because this disorder is associated with a type of mental retardation whose origin is genetically and biochemically understood. A valid animal model would allow one to assess the relative value of various treatment methods and to study the mechanisms by which a specific metabolic abnormality result in brain damage and behavioral derangement. Many investigators have attempted to produce phenylketonuria in experimental animals but have had only limited success, perhaps because most of them have failed to define what constitutes a valid model of this type of disease or to state how well their model approximates the clinical disorder.

The five errors that seem to occur most frequently in experimental attempts to produce phenylketonuria, and which may explain the variability of results, are as follows. First there is the chemical method of producing the model. Most workers have given phenylalanine alone in large amounts to experimental animals (1, 2). This method produces hyperphenylalaninemia, but fails to duplicate the high ratio of phenylalanine to tyrosine in blood or the absence of hepatic phenylalanine hydroxylase (EC 1.14.3.1) activity found in phenylketonuria. Some researchers have administered *p*-chlorophenylalanine (3), an inhibitor of phenylalanine hydroxylase. However, without phenylalanine supplementation this method does not produce hyperphenylalaninemia in a rat, the most commonly used animal, because of the high levels of phenylalanine hydroxylase present in this species. Neither of these methods, therefore, duplicates the biochemical characteristics of phenylketonuria: hyperphenylalaninemia with normal or lowered tyrosine concentration in blood and the absence of phenylalanine hydroxylase. Intricate behavioral work should not be based on a faulty biochemical assumption of similarity to the intended clinical disorder. The second common error concerns the time in the life of the experimental subject in which the disorder is produced. Induced at any time past the phase of rapid brain development, selective hyperphenylalaninemia certainly affects the brain differently than when such a condition is present during

the neonatal period. Various investigators may have failed to produce the behavioral changes of phenylketonuria, even when carefully approximating the biochemical characteristics of the disorder, because they used the right method at the wrong time in brain development. It is important, therefore, to induce the biochemical alteration during early development of the animal. The third error concerns the time of testing. Animals tested during the period of syndrome production are often ill simply because of the noxious effects of the diet or the injections, and experimental differences found during this period usually represent nonspecific effects of illness on performance by an animal. Animals should be tested when fully recovered from the treatment period in order to sort out these nonspecific effects. The fourth common error involves the type of behavioral testing used. Mental retardation resulting from early childhood metabolic disorders represents part of a global brain change, and is not equivalent to a selective learning deficit. All phases of behavior are changed to some degree. Therefore, a single test of learning impairment or a narrow look at the behavior of the animal is insufficient to validate the existence of experimental phenylketonuria. In addition to their profound intellectual deficit, most children with this disease are overactive, aggressive, and have poor emotional control (4). Tests used to validate an experimental phenylketonuria should reflect the global nature of the syndrome. Finally, few attempts have been made to correlate the behavioral deficits with changes in brain structure. While the neuropathology of phenylketonuria is sparser than that of some more frequently occurring or more grossly damaging brain diseases, there is general agreement that brain weight is lower than normal and that the myelination process is abnormal (5). Brain-behavior correlations are necessary to support the validity of a proposed model.

In critical discussions of model phenylketonuria, a few investigators have pointed out some of these commonly occurring errors. Karrer and Cahilly (6) criticized studies in which phenylalanine concentrations in blood were not substantially elevated, and studies in which a single behavioral test was performed to substantiate the presence of "mental deficiency." Menkes (7) objected to the use of phenylalanine feeding alone as an adequate biochemical model, and Kilbey and Harris (8) stressed the necessity of chronic administration during the entire developmental phase.

Using the following guidelines, we have attempted to produce an experimental animal model of phenylketonuria. (a) In a genetically homogeneous strain of rats (F344), we have attempted to reproduce the biochemical features of phenylketonuria by inhibiting liver phenylalanine hydroxylase

TABLE 1.

	Brain weights (grams)		
	Saline-treated	Phe-treated	CIPhe+Phe-treated
Males	1.835 ± 0.018*	1.832 ± 0.015	1.735 ± 0.019
Females	1.726 ± 0.013	1.718 ± 0.015	1.650 ± 0.010

* SE.

with CIPhe, and simultaneously administering phenylalanine, to give a high phenylalanine to tyrosine ratio in plasma and low activity of hepatic phenylalanine hydroxylase. This method was suggested by Lipton, Gordon, Guroff, and Udenfriend, who presented data showing that such a program produces the biochemical characteristics of phenylketonuria (9). (b) The treatment period began at birth and continued daily for 21 days, during which time the major development of rat brain occurs. (c) The animals were tested when mature, after a long recovery period of about 6 months. (d) To appreciate the global nature of the disorder, the subjects motor activity, autonomic activity, active and passive avoidance abilities, learning ability, and aggression were tested. (e) Finally, an attempt has been made to correlate behavioral changes with abnormalities in brain structure.

METHODS

Method for model phenylketonuria production

Pregnant F344 rats were individually housed and given food and water *ad lib*. After delivery, each litter was assigned to one of three treatment regimens. *A*. Pups from five litters received 20 μ l/g of body weight of physiological saline, injected subcutaneously. *B*. Pups from seven litters received 333 mg/kg of body weight of L-phenylalanine, given as a subcutaneous injection of 20 μ l/g of body weight of a 1.7% solution of L-phenylalanine dissolved in physiological saline. This group of animals was included to compare the effectiveness of phenylalanine administration alone with our model. *C*. Pups from 14 litters received a subcutaneous injection of 20 μ l/gram of body weight from the supernatant fluid of a 1.0% suspension of CIPhe in physiological saline containing 1.7% L-phenylalanine. The concentration of CIPhe in the supernatant fluid was 15 mM. The CIPhe was washed with water, ethanol, ethyl acetate, ethanol, and water before use. All solutions were autoclaved at 121° for 20 min at 1.3 atm (20 lb/in²) pressure.

Injections were given daily for 21 consecutive days after birth. Pups were weaned when they were about 30 days old, and were maintained thereafter on food pellets and water *ad lib*. They were kept in their litter groups until 6 weeks of age, when they were separated by sex. Before they were tested at about 7 months of age, subjects were placed in individual cages in which they were maintained throughout the testing period.

Method for behavior testing

The overall testing plan was as follows. The same subjects were used in all tests in order that intertest comparisons could be made. Predominantly observational tests were done first. Males were then tested for active and passive avoidance, and females for appetitive learning. Finally, subjects were re-

combined for aggression testing and brain examination. Wherever possible visual cues were avoided because of reports (10) that CIPhe induces cataracts.

Open-Field Test for Motor Activity. Subjects were 12 male and 8 female animals treated with saline, 7 male and 9 female rats treated with phenylalanine, and 12 male and 8 female rats treated with *p*-chlorophenylalanine plus phenylalanine. The open-field apparatus (91 × 51 cm) was divided by thin black lines into areas of 15 × 17 cm. Each rat was introduced nose first into a corner; the number of boundaries crossed over the next 5 min was recorded.

Open-Field Test for Fecal Pellets. After each open-field test for motor activity, fecal pellets were counted. The number of fecal pellets deposited on an open field has been used as one index of *automatic activity*, *emotionality*, and general reactivity to a stress situation. The important feature of this test is that fecal pellets represent a relatively constant and enduring measure of an animal's response to a novel environment. The response to constraint was tested by reintroduction of the subject back into the open field, but confined to one corner by a Plexiglas column. Fecal pellets were counted after confinement for 5 min.

Active Avoidance with Two-Way Shuttlebox (C.H. Stoelting Co., Chicago). Subjects were 9 male rats treated with saline, 6 with phenylalanine, and 6 with phenylalanine plus *p*-CIPhe. This test used a two-way shuttlebox, of dimensions 48 × 21 × 28 cm, with a floor of stainless steel bars and a central doorway between compartments. Before the first session, each animal explored the apparatus for 10 min. It then completed 10 sessions of 30 trials per session. Each trial consisted of a 23-sec pause, 3 sec of auditory stimulus (buzzer), and 4 sec

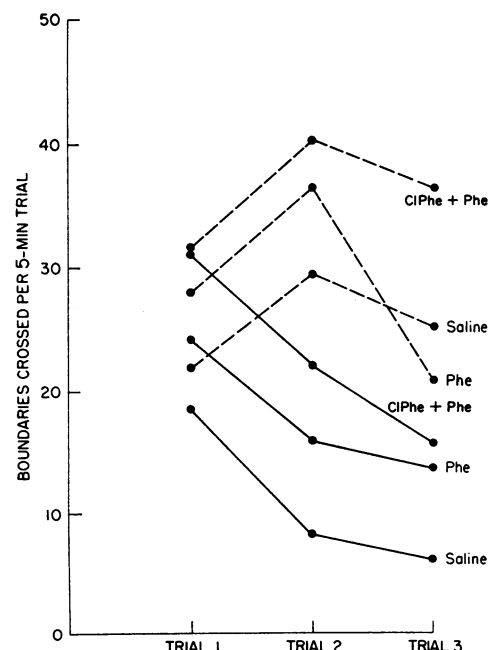


FIG. 1. Open-field motor activity. Solid lines indicate males, dotted lines indicate females. The number of boundaries crossed during three separate 5-min trials on an open-field apparatus are shown. There is a significant difference between saline and CIPhe+Phe. Phe treatment falls between saline and CIPhe+Phe except for one instance.

of combined buzzer and shock (average bar current, 0.7 mA). The task presented to the animal was to cross through the central doorway after the buzzer was sounded in order to avoid a shock.

Passive Avoidance of a Previously Learned Active Avoidance Response. The same subjects and apparatus described in active avoidance were used. 24 Hr before passive avoidance testing, each subject was given one repeat session of 30 trials of active avoidance to minimize differences in memory of the active avoidance task. The two-way shuttlebox was then modified for passive avoidance testing. One side (side A) was unchanged. The other side (side B) was converted into a small enclosed area of $17 \times 13 \times 10$ cm, opening directly off the central partition doorway, with free access between sides A and B. The task for the subject was to inhibit his previously learned active response of moving on buzzer presentation. The criterion for passive avoidance learning was defined as the number of trials required for a subject, introduced into side A, to learn to sit through four successive buzzer presentations without moving to side B. Parameters of time, buzzer loudness, and extent of shock were the same as above. If the subject moved to side B on buzzer presentation, he was shocked, removed to a holding box for 1 min, and returned to side A. 1 Week after determination of the number of trials required to reach the passive avoidance criterion, subjects were tested for the number of trials to relearn the same criterion. At this time, latency to move, defined as the number of buzzer presentations sat through before moving on the first trial, was also determined.

T-Maze Testing (Appetitive Learning). 6 Female rats treated with saline, 8 with phenylalanine, and 7 with ClPhe+Phe were accustomed to a water deprivation schedule of 1 hr of drinking time per day. As preparation for T-maze testing, rats were allowed to explore the T-maze and were trained to run for a water reward at the ends of the bar of the T. A piece of sandpaper 38 cm long by 10 cm wide was placed in the runway portion of the maze. The learning task consisted of associating the rough surface of the sandpaper with a right turn to a water reward, and the smooth side of the sandpaper with a left turn for a water reward. If all choices were incorrect at the end of a testing session, the subject was allowed to self-correct on the last trial. Trials per day were gradually increased from three to seven, with time per trial allowed at the choice point decreased from 5 min to 90 sec.

Aggression Testing. One adult male mouse was introduced into each rat home cage. Mouse killing was noted after 5 min and after 1 hr. Male rats were retested for mouse killing after they received four ascending shocks: 0.2, 0.5, 0.8, and 1.6 mA. These shocks were administered to determine the behavioral response to different shock levels.

Method of brain examination

Brain Weights. Each rat received intraperitoneally 100 units of heparin and 50 mg of pentobarbital, and was then decapitated and its brain was weighed.

Methods for Neuropathology. Subjects for neuropathology were anesthetized and perfused with heparinized saline, followed by 10% formalin. Paraffin sections were prepared and stained with hematoxylin and eosin or with Luxol Fast Blue for myelin.

RESULTS

Since we are interested in how well an experimental test matched an outcome predicted from clinical information, a one-tailed test of significance is appropriate. In other words, we are interested in rejecting the null hypothesis only if the experimental results are in the direction predicted from the clinical facts. Except where otherwise noted, probability figures refer to one-tailed values by Student's *t*-test. In one instance where a nonparametric test seemed applicable, the Wilcoxon test was used.

On casual inspection, the experimental animals groomed and ate normally. Except for appearing somewhat more active and excitable, the group treated with ClPhe+Phe did not show any gross behavioral deficits. These deficits only emerged on more sensitive testing.

Behavior tests

Open-Field Motor Activity (Fig. 1). On all observations, rats treated with ClPhe+Phe were more active than the controls. With one exception, rats treated with phenylalanine alone fell between those treated with ClPhe+Phe and saline. For males, the difference between ClPhe+Phe and saline groups is significant both on individual trials ($P < 0.05$, Wilcoxon test) and when the sums of scores on all three trials are compared ($P < 0.011$). The difference between females treated with ClPhe+Phe and saline does not reach statistical significance, although the trend is clearly in that direction. As in most other open-field tests, female rats were more active than male rats.

Open-Field Fecal Pellets As an Index of Autonomic Activity. When the sums of fecal pellets expelled on all three trials are compared, saline controls exhibited more autonomic activity than treated animals ($P < 0.04$). In the 5-min long mild-confinement situation conducted on the same apparatus, saline-treated animals showed considerably more autonomic activity than treated rats ($P < 0.004$). Animals given phenylalanine alone give intermediate results between controls and those treated with ClPhe+Phe.

Active Avoidance Response (Fig. 2). There is no statistically significant difference between any groups in active-avoidance learning ability.

Passive Avoidance Response (Fig. 2). Treated animals required nearly twice as many trials as did saline-treated controls to achieve criterion performance on initial passive-avoidance testing ($P < 0.035$). When criterion testing was repeated 1 week later, a disparity between these two groups persisted ($P < 0.012$). On the first trial at the repeat testing, the treated animals ran after 1.2 buzzer presentations, compared with saline-treated animals, who sat through 2.4 buzzer presentations before moving to side B. The difference in the latency period between treated and control rats is significant ($P < 0.015$).

T-Maze. This experiment did not show what it was intended to show: a comparison of the rate of appetitive learning in a water-deprived animal in a T-maze. No group achieved greater than a chance score of correct choices, the task apparently being too difficult even for the normal rats. What did appear significant was that the treated animals showed more no-choice events; that is, they more frequently came to the choice point of the maze and refused to move in either

TABLE 2.

Experimental test	Result predicted from clinical characteristics	Experimental results
1. Open-field test for motor activity	Phenylketonuria patients are hyperactive	Order of motor activity is CIPhe+Phe > Phe > saline
2. Autonomic activity	Information not available about phenylketonuria patients.	Order of autonomic activity is saline > Phe > CIPhe + Phe
3. Active avoidance	Information not available. Might plausibly expect hyperactive patients to have little difficulty in very simple active avoidance.	CIPhe+Phe = Phe = saline
4. Passive avoidance	Phenylketonuria patients have difficulty in inhibiting undesired hyperactivity.	Treated rats require twice as long to reach passive avoidance criterion performance as controls
5. T-maze	Retarded patients have learning deficit and may exhibit indecision when presented with a difficult choice.	No group showed learning. CIPhe + Phe have more difficulty in decision making at choice point
6. Aggression	Phenylketonuria patients aggressive with severe temper outbursts.	CIPhe+Phe more often mouse killers
7. Neurophysiology	Phenylketonuria patients have frequent seizures.	CIPhe+Phe rats have lowered seizure threshold (12)
8. Brain weights	Phenylketonuria patients have decreased brain weights.	CIPhe+Phe < Phe = saline
9. Neuropathology	Phenylketonuria associated with myelin defect, especially in young patients.	Adult CIPhe+Phe brains do not show obvious myelin defect
10. Biochemistry	Phenylketonuria patients have increased Phe to Tyr ratio and absence of phenylalanine hydroxylase activity.	Same as clinical characteristics (9)

direction. The average percentage of no-choice events per group is 21% for those treated with saline, 25% for Phe, and 35% for CIPhe+Phe. The difference between saline-treated and CIPhe+Phe-treated rats is significant ($P < 0.05$). These results may be interpreted to mean that animals

treated with CIPhe+Phe have more difficulty in decision making in difficult choice situations.

Aggression. None of 7 male and none of 8 female rats injected with saline killed mice. None of 5 male, but one of 7 female rats injected with Phe killed a mouse. 3 of 13 males and 1 of 8 females treated with CIPhe+Phe were mouse killers, all attacking within 5 min and killing within 1 hr. When rats were tested immediately after shock, no new mouse killers emerged. Although treated rats killed more mice, these numbers are not statistically significant.

Neuropathology

Brain Weights. (Table 1). There is no statistically significant difference between body weights of the three groups when all possible combinations are considered. The brain weights of the saline- and Phe-treated groups do not differ from each other, but are both very significantly heavier than those of rats treated with CIPhe+Phe. For the males: Saline against CIPhe+Phe, $P < 0.002$; Phe against CIPhe+Phe, $P < 0.001$. For the females: Saline against CIPhe+Phe, $P < 0.0005$; Phe against CIPhe+Phe, $P < 0.0005$.

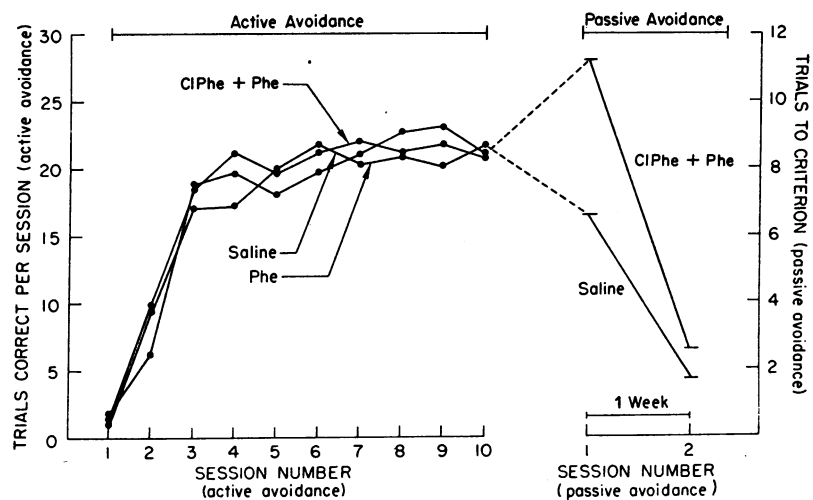
Histology. When compared with saline-treated controls, brains from rats treated with CIPhe+Phe do not show any focal myelin defects. There is a suggestion of lighter myelin staining in the CIPhe+Phe brains, but the difference is slight and may not be significant. The most severe myelin defects associated with phenylketonuria are found in young patients (11), with progressively fewer findings observed on neuropathological examination as brains from older patients are examined.

DISCUSSION

This paper reports initial observations on an animal model of phenylketonuria and offers some practical suggestions to assess its validity. Our goal is an animal model valid at the biochemical, neurophysiological, neuropathological, and behavioral levels. Table 2 lists the series of tests used and compares the outcome that might be predicted from the clinical picture with the actual experimental results. In all cases these two categories correlate well, usually with statistically significant differences between experimental and control rats, but sometimes with trends where the numbers are small.

An important question concerns the specificity of the chemical insult. Do the behavioral abnormalities found in this experiment reflect a brain change specific to phenylketonuria, or do they reflect nonspecific brain changes common to several types of neurological injuries? Two lines of thought lead us to conclude that the change is specific. First, by attempting a global comparison of experimental results with the clinical characteristics of phenylketonuria, we feel the assertion of the validity of this model is stronger than claims advanced on the basis of finding only one or a few features in common between the model and the clinical state. Better though than assertions of validity is the point by point comparison between model and disease, with the data open to critical scrutiny. Secondly, in addition to this comparison between model and clinical disorder, a comparison of the phenylketonuria model with other experimental models in which a different agent of injury was used would be useful to establish specificity. For example, the

FIG. 2. Active avoidance training, followed by passive avoidance of the active response. *Left*: the scores of male subjects over 10 sessions of active avoidance in a two-way shuttlebox. Each session consisted of 30 trials. There is no significant difference between groups on active avoidance testing. *Right*: the response of these same subjects (excluding Phe) to passive avoidance of the learned active avoidance response. The saline-treated and CIPhe+Phe-treated animals now split into two distinct groups; the statistically significant difference between groups is maintained when the subjects are retested 1 week later. The results of the first passive avoidance sessions are 6.6 ± 1.1 trials for controls and 11.2 ± 1.8 trials for CIPhe+Phe-treated rats (standard error given). The results of the repeat passive avoidance session are 1.7 ± 0.2 trials for controls and 2.6 ± 0.3 trials for treated.



results of experimental phenylketonuria should be contrasted with those of experimental malnutrition, since the actual agent of injury in experimental phenylketonuria might plausibly be an undernourishment secondary to the noxious effects of the injections. When this comparison is made, subjects with experimental malnutrition are noted to be less active, not more active, than normal rats when they are refed (13) and are permanently decreased in weight (14). Another question is whether our model produces in the experimental subject changes not found in clinical phenylketonuria. For example, cataracts sometimes occur after administration of CIPhe to rats. This represents an experimental change not found in clinical phenylketonuria. CIPhe partially inhibits tryptophan hydroxylase activity, which results in lowered serotonin concentrations. This might be considered another example of an unwanted characteristic found in the model system, but not in the clinical situation. However, phenylketonuria is associated with decreased serotonin concentrations, so in this regard the model system duplicates a biochemical characteristic of clinical phenylketonuria, rather than creating an unwanted novel effect.

Our own criteria suggest several improvements that should be made in our model. These improvements include the development of learning tasks that more effectively distinguish between experimental and control animals, a comparison of this model with others involving different types of brain damage, and a detailed study of the developmental neuropathology of our model. We feel that these efforts to achieve a valid model of phenylketonuria are worthwhile. In the field of brain-behavior research, the interaction between such a complex organ as the brain, which constantly changes in structure and function during its development, and ill-defined types of neurological injury produces a wide range of adult behavior and brain damage. By carefully exposing the developing brain to a well-defined metabolic error, this interaction may be simplified so that specific questions about the mechanism of brain damage and the resultant behavioral

abnormality may be formulated and investigated. Also, several practical suggestions for the early treatment of phenylketonuria in addition to dietary limitation of phenylalanine may be tried and the results compared.

In summary, we feel that in rats, inhibition of phenylalanine hydroxylase by CIPhe, coupled with phenylalanine supplementation, produces an experimental picture that overlaps considerably the clinical disorder.

We gratefully acknowledge the continued guidance of Dr. Jeri Sechzer throughout all stages of this project. We also appreciate the statistical analyses done by Dr. Daniel G. Seigel and Howard Hoffman. Helpful suggestions about the neuropathology work were provided by Drs. Paul R. McHugh and Jan Cammermeyer. Dr. Anatole Dekaban kindly interpreted the myelin stains. Drs. Gerard Smith and Michael Goldberg read the manuscript and gave valuable suggestions.

1. Polidora, V. J., Cunningham, R. F. & Waisman, H. A. (1966) *J. Comp. Physiol. Psychol.* **61**, 436-441.
2. Chamove, A. S. & Davenport, J. W. (1970) *Develop. Psychobiol.* **2**, 207-211.
3. Pryor, G. T. & Mitoma, C. (1970) *Neuropharmacology* **9**, 269-275.
4. Knox, W. E. (1966) in *The Metabolic Basis of Inherited Disease*, eds. Stanbury, J. B., Wyngaarden, J. B. & Fredrickson, D. S. (McGraw-Hill, New York), 2nd Ed., p. 263.
5. Malamud, N. (1966) *J. Neuropath. Exp. Neurol.* **25**, 254-268.
6. Karrer, R. & Cahilly, G. (1965) *Psychol. Bull.* **64**, 52-64.
7. Menkes, J. H. (1967) *Pediatrics* **39**, 297-308.
8. Kilbey, M. M. & Harris, R. T. (1971) *Psychopharmacologia* **19**, 334-346.
9. Lipton, M. A., Gordon, R., Guroff, G. & Udenfriend, S. (1967) *Science* **156**, 248-250.
10. Gralla, E. J. & Rubin, L. (1970) *Arch. Ophthalmol.* **83**, 734-740.
11. Alvord, E. C., Stevenson, L. D., Vogel, R. S. & Engle, R. L. (1950) *J. Neuropath. Exp. Neurol.* **9**, 298-310.
12. Prichard, J. W. & Guroff, G. (1971) *J. Neurochem.* **18**, 153-160.
13. Frankova, S. & Barnes, R. H. (1968) *J. Nutr.* **96**, 477-484.
14. Dickerson, J. W. T. & Walmsley, A. L. (1967) *Brain* **90**, 897-906.