References

1. ALLEN, B. M. *Biol. Bull.*, 1917, **32**, 117.
2. BLOUNT, R. F. *J. exp. Zool.*, 1932, **63**, 113.
3. *Ibid.*, 1945, **100**, 79.
4. BURCH, A. B. *Univ. Calif. Pub. Zool.*, 1946, **51**, 185.
5. SMITH, P. E. *Anat. Rec.*, 1916, **11**, 57.
6. TAYLOR, A. C. and KOLLROS, J. J. *Anat. Rec.*, 1946, **94**, 7.

A Very Water-soluble Riboflavin Derivative

George B. Stone

Technical Service Department, Chas. Pfizer & Co., Inc., Brooklyn, New York

For the past ten years a not inconsiderable amount of investigation has been carried out in an effort to increase the water solubility of riboflavin (vitamin B_2) either by the use of solubilizers or by the preparation of soluble derivatives. Despite some thirty or more references and patents, representing several hundred solubilizers or soluble derivatives, few, if any, are of practical significance for pharmaceutical application. Riboflavin is not only sparingly soluble in water, but in almost every other solvent. It is, however, relatively soluble in concentrated sulfuric acid. Investigation of this significant solubility in concentrated sulfuric acid led to the isolation of a very water-soluble riboflavin derivative.

The compound was prepared by dissolving 50 g of riboflavin, little by little, in 200 ml of concentrated sulfuric acid with vigorous stirring, while the temperature was maintained at 40°–50° C. Mixing was continued for 1–2 hr until the mixture was homogeneous and it was then quenched by pouring it over 1 kg of cracked ice. The resulting solution was neutralized with slurried calcium hydroxide to a pH of 6.5, with the temperature being maintained below 70° C. The precipitated gypsum was filtered off, washed with hot water, and then repulped with hot water, filtered, and again washed. All washings were added to the original filtrate. Assay by the fluorometric method indicated that the original 50 g of riboflavin was present in this solution. The solution was concentrated under vacuum to a volume of less than 500 ml and filtered to remove further gypsum precipitated during concentration. The filtrate was then freeze-dried to yield 114 g of a fluffy yellow-orange powder, which assayed fluorometrically 57.2% riboflavin, equivalent to a yield of 100% based on the weight of the riboflavin employed originally.

The compound is stable in air and nonhygroscopic. It is very soluble in water, and aqueous solutions containing 10% wt/vol of riboflavin have been prepared—a solubility 1,000 times greater than that of riboflavin U.S.P. It is soluble in methanol and slightly soluble in ethanol, and it decreases in solubility with the higher alcohols. It is soluble in glycerine, propylene glycol, and pyridine; slightly soluble in acetone, glacial acetic acid, and chloroform; insoluble in benzene, ether, ethyl acetate, methylethylketone, and carbon tetrachloride. Aqueous solutions are heat-stable at 15 psi for 120 min in the pH range from 1.0 to 6.5.

Preliminary chemical analysis seems to indicate that the compound may be represented by the empirical formula, $C_{17}H_{18}N_4O_{15}S_2Ca$, inasmuch as the compound contains calcium and sulfur, a portion of the sulfur being present as sulfate.

Fluorometric assay of the material yields a value of 57.2% riboflavin. The absorption spectrum is identical with riboflavin U.S.P., having the same maxima and minima, but proportionately displaced because of the lesser riboflavin content. Paper chromatographic absorption analysis indicates the material is a pure compound, much more water-soluble than riboflavin U.S.P.

Microbiological assay by the U.S.P. XIII revision method, which includes a preliminary hydrolysis at 15 psi for 30 min, yielded a value of 33.0% riboflavin. Omission of the preliminary hydrolysis gave a value of 1.5% riboflavin, whereas increase of the time of hydrolysis to 120 min gave a value of 42.2% riboflavin.

Biological assay for riboflavin by a standard rat growth method employing a basal vitamin B complex-free diet, supplemented with those members of the B complex other than riboflavin, indicated that the riboflavin potency for the rat is almost nil. There were indications of a slight antivitamin activity of this compound.

Further investigation of this material is anticipated.

A Correction for Linkage in the Computation of Number of Gene Differences

Everett R. Dempster and Leon A. Snyder[1]

Genetics Division, University of California, Berkeley

The Castle-Wright formula (*2*) for estimating the number of gene pairs differentiating two strains with respect to some quantitative character is based on the increased variance of the F_2 as compared to the parental variance. A number of postulates on which this derivation is based (*1, 5, 6, 7*) may be listed as follows: (1) The parents are homozygous. (2) All the plus alleles differentiating the two strains with respect to the character considered are in one parent and all the minus alleles in the other. (3) All gene differences affecting the character have equal effects. (4) The effects of different allelic substitutions are additive. (5) There is no linkage. Deviations from postulates 2 to 5, with the possible exception of special epistatic effects (postulate 4) will always increase the F_2 variance, and therefore, since the latter appears in the denominator of the expression for gene number, will bias the estimate toward lower values. Since actual situations are usually at variance with most of the postulates listed, the expression in general leads to minimum rather than unbiased estimates of the number of gene differences.

Some modifications have been devised for relaxing the postulates or otherwise extending the applicability of

[1] Present address: Department of Plant Sciences, University of Oklahoma, Norman, Oklahoma.

gene number estimates based on variance. Thus Serebrovsky (*6*) has presented formulas utilizing backcross data in which allowance is made for any degree of dominance, and Charles and Goodwin (*3*) have considered the extension of the method to the case of several correlated characters. The purpose of this note is to suggest a correction for linkage.

The formula given in the article by Castle (*2*) is $n = D^2/[8(\sigma_2^2 - \sigma_1^2)]$ where D is the difference between the means of the parental races, σ_1 is the standard deviation of the F_1 generation, and σ_2 is the standard deviation of the F_2 generation. The difference term in the denominator, which will be written σ_G^2 in the following discussion, represents an estimate of that portion of the F_2 variance due to segregation of genes at loci for which the parents were homozygous for different alleles.

It has been shown (*4*) that if n loci, at which nondominant alleles have equal differential effects, are distributed at random among m linkage groups within which linkage is complete, the expected factor of increase in genetic variance due to linkage is $(m+n-1)/m$. If we divide σ_G^2 by this factor, the quotient is an estimate of the genetic variance that would be obtained in the absence of linkage. We have then, for the case of complete linkage, the expression $n = [(m+n-1)D^2]/[8m\sigma_G^2]$, from which can be obtained the following:

$$n = [m-1]/[(8m\sigma_G^2/D^2) - 1].$$

For comparison the uncorrected formula may be written:

$$1/[8\sigma_G^2/D^2].$$

Since real chromosomes contain unequal numbers of genes and are subject to crossing over, their number cannot be directly substituted for m in the above formulas. Instead one should substitute the number of hypothetical chromosomes, possessing equal proportions of the total number of genes and immune to crossing over, that would increase F_2 variance to the same degree as do the real chromosomes. The following considerations lead to an approximation for m which is conservative in the sense that it somewhat underestimates the average effect that linkage would be expected to have in increasing variance. If there are k points of exchange in a bivalent chromosome, there are $(k+1)$ terminal and interstitial regions. At any exchange point two homologous chromatids are involved, the other two remaining unaltered. In the absence of chromatid interference, any of the $2^{(k+1)}$ possible combinations of the $(k+1)$ regions, with respect to origin from one or the other homolog, is equally likely to appear in any given haploid nucleus resulting from meiosis. This is equivalent, with respect to F_2 variance, to the independent assortment of $(k+1)$ pairs of chromosomes within which linkage is complete, although the lengths of the hypothetical chromosomes would be unequal. The expected effects of linkage on F_2 variance can readily be shown to be greater for chromosomes of unequal as compared to chromosomes of equal length (*4*). The over-all effect on linkage, therefore, approximates and is somewhat greater than that which would result from the division of the chromosome complement into $\Sigma_1^c(k+1)$ equal segments within which linkage is complete, where c is the actual number of chromosome pairs and k the average number of exchanges in each pair. Since fifty crossover units correspond to an average of one exchange chiasma per bivalent, the sum of one-fiftieth of the total number of map units and the number of chromosome pairs may also be employed as a rough estimate of the effective value of m. However, chromosomes that are very small compared to the others in the complement and much less than 25 map units in length would contribute very little to expected increase of variance and should be excluded in computing these approximations.

The value of m for maize on the basis last mentioned would be at least 27, and for *Drosophila melanogaster*, excluding the tiny fourth chromosome, approximately 9. There is, however, no crossing over in male gametogenesis in *Drosophila*, and so a more reasonable value for m in this organism would be the mean of the value just given, 9, and the number of major chromosome pairs, 3 or 6. Sex linkage would probably somewhat increase the variance among F_2 males and decrease it among F_2 females.

It may first be noted that the corrected formula leads to greatly increased estimates where the variance values are such that the uncorrected formula would indicate around m or more loci. Thus in *Drosophila melanogaster*, if m is taken as 6, some values of n given by the uncorrected and corrected formulas, respectively, are as follows: 3 and 5, 5 and 25, 6 and ∞.

The error variance of σ_G^2 is the sum of the error variances of σ_1^2 and σ_2^2 and is not likely to decrease much with increase in gene number. Therefore the accuracy with which estimates may be made decreases very rapidly as gene number increases, whether there is independent assortment or linkage. An error variance just small enough for a distinction to be made between 3 and 4 loci would barely permit discrimination between 6 and 12, or 9 and 30, or 12 and ∞ loci.

Any error in the estimated value of σ_G^2 will contribute slightly more to the inaccuracy of gene number determinations with linkage than would be the case with independent assortment. A measure of the relative effects due to this cause may be obtained by comparing $dn/[d(\sigma_G^2/D^2)]$ for the two formulas. The two values are $8n^2$ and $8n^2[m/(m-1)]$ for independent assortment and linkage respectively. Linkage thus increases the standard error by the proportion $1/(m-1)$. This increase refers, however, to a component of error that could be minimized to any desired degree by taking sufficiently large samples. Another and probably much more important effect of linkage on error is dependent on the actual distribution of loci among linkage groups in the particular parental strains considered. For any particular pair of parental strains, the loci involved would usually be either more or less uniformly distributed among the different linkage groups than mean expectation. This would be true, in general, even in the absence of special mechanisms, such as the introgression of genes from one species to another, that might favor a concentration of genetic differences in one or a few linkage groups. Thus in any particular case there may well be a serious bias in the plus or minus direction of gene number estimate, and

such a bias cannot be reduced by an increase of sample size.

In conclusion, then, it is possible to apply a simple correction for linkage to the formula for estimating the number of different genes. This correction, if the postulates on which it is based are met, applies to the average situation. In the case of any particular pair of parental strains, the actual distribution of differentiating loci on linkage groups may lead to a serious bias in estimates of gene number. Even in the absence of the complications due to linkage, the estimate of gene number is likely to be highly inaccurate except where the number of differentiating loci is relatively small.

References

1. CASTLE, W. E. *Science,* 1921, **54**, 93.
2. *Ibid.,* 1921, **54**, 223.
3. CHARLES, D. R. and GOODWIN, R. H. *Amer. Nat.,* 1943, **77**, 53.
4. DEMPSTER, E. R. *Genetics,* 1949, **34**, 272.
5. MATHER, K. and HARRISON, B. J. *Heredity,* 1949, **3**, 1.
6. SEREBROVSKY, A. S. *Z. indukt. Abstamm. Vereb.,* Lehre, 1928, **48**, 229.
7. WRIGHT, S. *Genetics,* 1934, **19**, 537.

Correlation of Certain Physical Constants of Some Alkyl Esters of *n*-Phenyl Carbamic Acid, with Their Phytotoxicity[1]

V. H. Freed

Oregon Agricultural Experiment Station, Corvallis, Oregon

The toxicity of the alkyl ester derivatives of phenyl carbamic acids to members of the Gramineae family was first demonstrated in 1929 by Freisen (*4*), dealing with ethyl *n*-phenyl carbamate. Deysson (*2*) in 1945 demonstrated that this material acted, at least in part, as a mitotic poison, inhibiting cell division and causing subsequent death of the cell. This was a direct corroboration of the work done by Lefevre (*5*) in 1939.

Templeman and Sexton (*10*) in 1945 reported on the phytotoxicity of various carbamates, with the object of controlling certain weedy plants. In 1946 the same authors (*11*) announced their discovery of the phytotoxicity of isopropyl *n*-phenyl carbamate. Allard *et al.* (*1*) in 1946 further elaborated on this phytotoxicity. Numerous investigators have since published information concerning the merits of this chemical as a phytocide.

There have been many attempts to correlate chemical structure with the biological activity of a number of compounds. Frear *et al.* (*3*) undertook the study of some 5,000 organic chemicals to elucidate certain chemical structures which could be correlated with toxicity. They concluded that certain groupings in specific types of aporadicals produces a toxic entity. Tattersfield and Roberts (*9*) reported a study of physical properties and chemical constitution of organic compounds as related to their toxicity to the wireworm. These authors concluded that in any homologous series the most toxic compound would be the one with the highest vapor pressure if it possessed a sufficiently high molecular weight. Rubbo (*7*) demonstrated the correlation of the ionization constant to toxicity for mice and bacteria of derivatives in the acridine series.

Melander (*6*) was able to correlate the physical constants of various isomers of hexachlorcyclohexane with their toxicity. He found that the gamma isomer having the highest dipole moment also has the highest insect toxicity, whereas the other isomers have considerably lower dipole moments and accordingly lower toxicity.

In the present study of the series of alkyl esters of *n*-phenyl carbamic acids, it was reasoned that since molar refractivity is a function of the geometric configuration of a molecule in any homologous series, this measurement might offer a clue to the correlation of physical chemical properties with biological activity. Accordingly, the refractive index of this series of compounds was determined with an Abbe refractometer at 91° C ± 0.1°. The compounds have previously been purified by recrystallization from petroleum ether. The density of the compound was determined by the volume displacement method at 20° C.

The biological activity of the compound was determined by planting the seeds in triplicate in gallon cans that had been previously treated with an amount of material calculated to give 1 lb of active ingredient per acre. Notes were taken on the number of seedlings that emerged, and the plants were harvested after two weeks' growth, and weighed.

From the data derived from the refractive index and density measurements, molar refractivity was calculated according to the Lorentz-Lorenz (*12*) formula.

$$N = \frac{n^2 - 1}{n^2 + 2} \cdot \frac{M}{d}$$

A further calculation was made by multiplying the molecular refractivity by the factor of the melting point of the derivative divided by the boiling point of the parent alcohol.

$$Q = N \cdot \frac{\text{mp (derivative)}}{\text{bp (alcohol)}}$$

These calculations were made prior to obtaining the data from the toxicity experiment. After obtaining the data

TABLE 1

Alkyl group	Bp (Alc) °C	MP (deriv.) °C	Density (deriv.)	Refractive index (91° C)
Methyl	66	47	1.15	1.5235
Ethyl	78	52	0.92	1.5105
Propyl	97	58	1.06	1.5056
Isopropyl	83	90	1.09	1.4989
Butyl	116	57	1.03	1.4987
Sec-butyl	99	30	1.70	1.4957
Isobutyl	108	80	1.60	1.4955
Amyl	138	46	1.01	1.4926
Isoamyl	130	55	0.98	1.4939

[1] Published as Technical Paper No. 589 with the approval of the Director of the Oregon Agricultural Experiment Station. Contribution of the Department of Agricultural Chemistry.