1952

Layman Craig, 1949

The Rockefeller University

ISOLATION AND CHARACTERIZATION OF BIOLOGICALLY IMPORTANT SUBSTANCES*

LYMAN C. CRAIG

Member
The Rockefeller Institute for Medical Research,
New York, New York

THE science of organic chemistry grew from the early attempts of chemists who tried to investigate the nature and the origin of the substances so prevalent about them which they believed to be connected in some way with living things. Obviously, it could make progress as a true science only when an experimental approach to the problem of the separation, isolation, proof of purity and characterization of a given substance was developed. As a working hypothesis a substance was considered "pure" when it or some derivative of it could no longer be resolved further into fractions with definitely different properties. This viewpoint is still the correct and most reliable one and the experimental part of our science rests upon it just as much today as it has in the past.

As technical progress has been made, better and more precise methods of fractionation have come into use and the so-called "pure" preparations of an earlier time have often been shown to be mixtures. Improved resolving power not only permits us to be more critical but also permits us to extend our studies to more complicated substances. It is here we soon learn how inadequate our ability to separate mixtures still remains, particularly with the relatively complicated and sensitive compounds of biochemistry. Indeed it requires little thought for one to realize that a major problem in biochemistry is that of separating and characterizing individual substances.

We separate and characterize a substance by virtue of some distinctive physical property or combination of properties. The process is relatively easy if the substance possesses some striking property which it alone exhibits. However, the larger the molecule, the less

* Lecture delivered December 15, 1949.

likelihood there will be of a striking difference in any physical property when the substance is compared to a closely related individual. In such a case, we must search for some method of magnifying small differences. Repeating a given physical operation, a partition process, over and over again would appear to be the most fruitful approach. Fractional distillation, fractional crystallization, fractional extraction, fractional adsorption and others come into play.

Almost from the very beginning it was realized that greater separation in terms of yield and labor could be achieved from any of these operations by re-treating the various fractions in a certain sequence. It was also found in distillation that improved separations resulted when the upright stem of the distilling flask was long and there was reflux. From this observation the fractionating column probably developed. It was soon realized that the separating efficiency depended upon two phases moving counter to each other with interchange constantly taking place. Such an operation, usually carried out in a column, was called a countercurrent process.

Column processes involving extraction and adsorption also proved to be highly efficient. However, not until much later was it fully realized[1] that all such countercurrent processes have a common theoretical basis. The countercurrent effect can be obtained by a discontinuous extraction process and here the theory is easier to explain. As applied to extraction, it can be outlined briefly as follows.

Let us imagine a train of units containing equal volumes of one of the immiscible phases used in the extraction, moving counter to a train of the other phase, as shown in Figure 1. When each unit comes in contact with a fresh unit of the other phase, it remains long enough to be equilibrated before moving on to the next so that any solute which is involved will distribute itself between the phases in contact according to its partition ratio, K, the ratio of the concentration of solute in upper to lower phase. Thus if unit quantity of a solute whose partition ratio is 1 is used, the fraction in each phase on successive contacts is given by the rows of Figure 1 going from top to bottom. Each row is a transfer.

We can also think of the process in terms of a row of separatory

funnels as in Figure 2 in which one phase remains stationary and only the total in a separatory funnel is considered. The effect is the same as that of Figure 1. The funnels are numbered in the upper

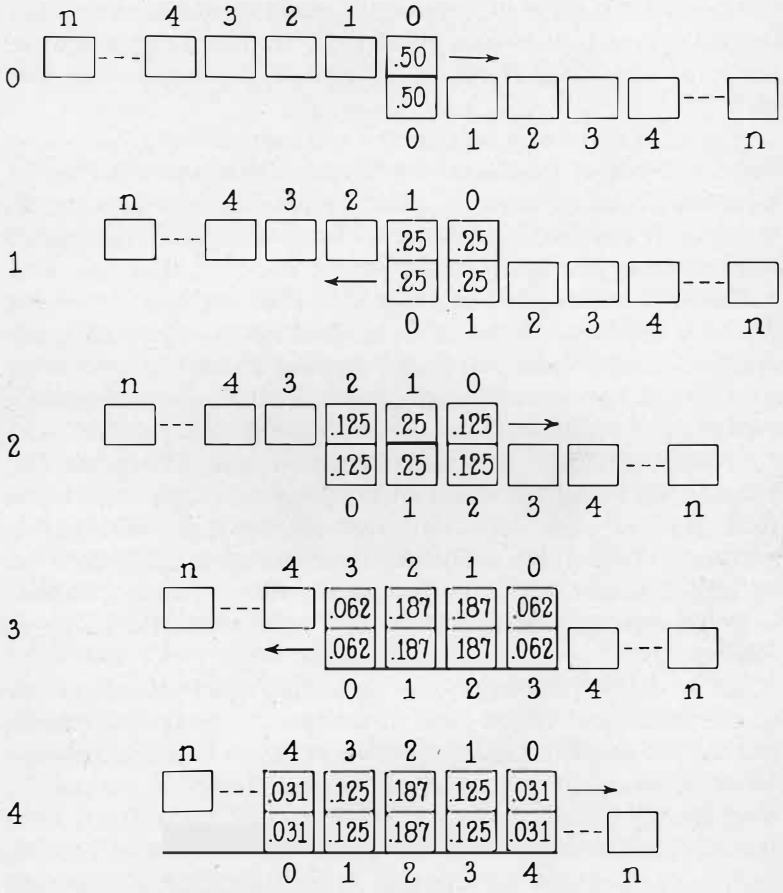

FIG. 1. Discontinuous countercurrent extraction scheme.

horizontal column, the number of transfers in the left vertical column. With a partition ratio of 1 and equal volumes of the phases, unit quantity of material is subdivided on each transfer as given by the corresponding horizontal line. Thus at 8 transfers each funnel

									
	0	1	2	3	4	5	6	7	8
0	1.000								
1	.500	.500							
2	.250	.500	.250						
3	.125	.375	.375	.125					
4	.062	.250	.375	.250	.062				
5	.031	.156	.313	.313	.156	.031			
6	.015	.093	.234	.313	.234	.093	.015		
7	.008	.054	.164	.274	.274	.164	.054	.008	
8	.004	.031	.109	.219	.274	.219	.109	.031	.004

$(x+y)^n$

Let x = Fraction in upper layer

Let y = Fraction in lower layer

$$N = n\left(\frac{K}{K+1}\right)$$

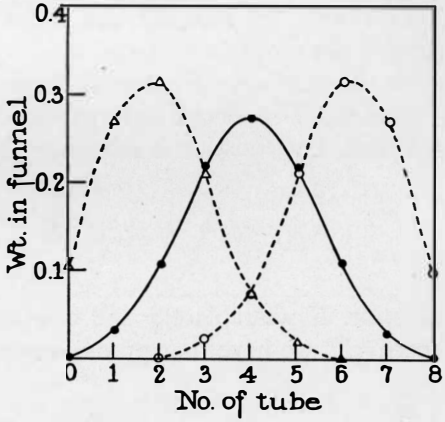

FIG. 2. Countercurrent distribution scheme.

will have the fraction in it shown. This table is a binomial series $(x+y)^n$ where x can be the fraction in the upper layer and y that in the lower layer.

If the values for 8 transfers are shown graphically, one obtains

the symmetrical curve. If K had been 3 instead of 1, the right hand dashed curve would have been obtained. If K had been 1/3 the left would occur. A mixture of the two would give a curve representing the sum of the two curves and the separation is evident at a glance. The position N of the maximum is a function of the partition ratio and the number of transfers n.

$$N = \frac{nK}{K+1} \qquad (1)$$

If an experimental curve exactly fits the curve calculated by the binomial expansion either it is the distribution of a single solute or of a mixture of solutes with identical partition ratios. On the other hand, if the experimental curve deviates from the calculated one, either the partition ratio is not constant at different concentrations or a mixture is present. The first possibility gives a skewed distribution. Determination of the actual K's on those tubes which deviate permits a decision. Progressive transformation occurring during the distribution also gives a deviation from the calculated provided a significant change in K is produced.

Obviously the accuracy and reliability of the method becomes greater when n becomes large. However, it soon becomes laborious even to calculate a theoretical curve by the binomial expansion. Fortunately statisticians have provided us with the approximations we need. The equation is familiar to biologists as the normal curve

$$y = \frac{1}{\sqrt{2\pi npq}}\, e^{\frac{-x^2}{2npq}} \qquad (2)$$

of error. If we replace p and q with K/(K+1) and 1/(K+1) respectively, we have the approximation we need for any given dis-

$$y = \frac{1}{\sqrt{2\pi nK/(K+1)^2}}\, e^{\frac{-x^2}{2nK/(K+1)^2}} \qquad (3)$$

tribution. Although this equation looks formidable, a slide rule permits calculation of a theoretical curve for any distribution in 15 minutes.

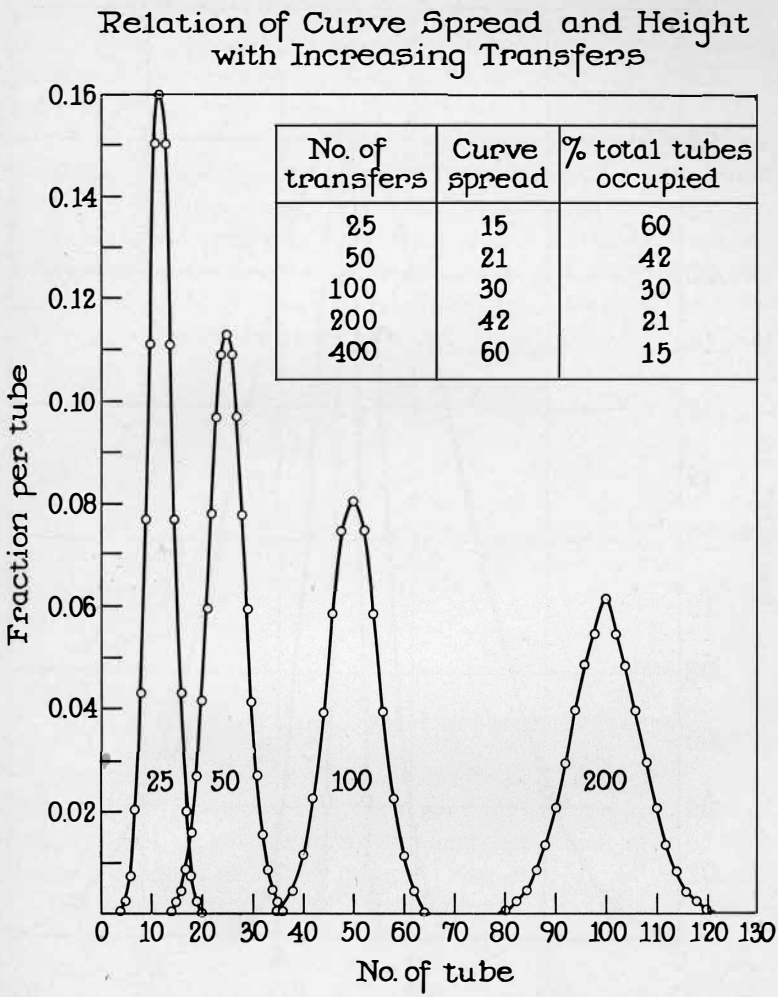

FIG. 3. Shape of distribution patterns with increasing numbers of transfers.

As n increases, the shape of the band changes as given in Figure 3. It becomes broader and lower. Of course, a tiny fraction even of a pure solute remains outside the band shown. For practical reasons, this is neglected when the amount in a single tube becomes less than 1% of the amount in the maximum tube.

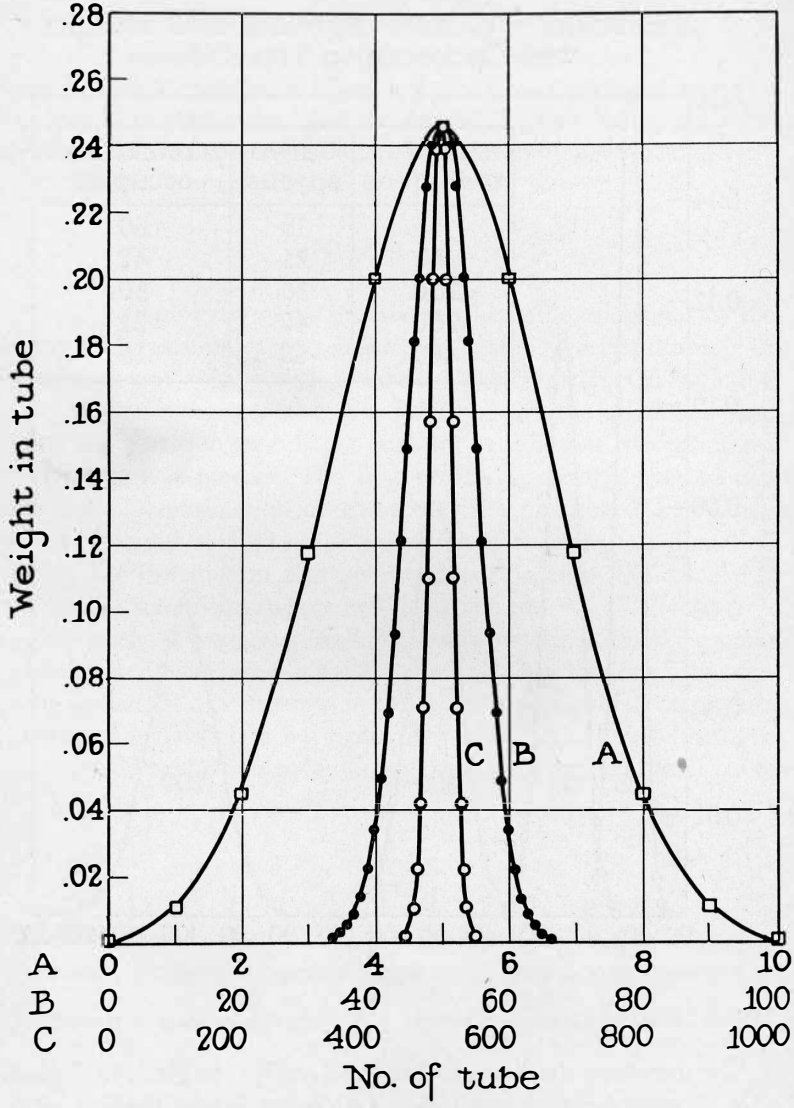

FIG. 4. Relation of band width to numbers of transfers.

The perspective of higher numbers of transfers, n, is given better in Figure 4 where the abscissa remains of constant width but the

space provided for each tube is proportionately decreased. In curve A, ($n = 10$, $K = 1$), all the tubes except 0 and 10 contain significant amounts of solute. Only one or two tubes on either side remain for throwing off impurity.

If, however, 100 transfers are applied, curve B is obtained. More sample is required so that the maximum will be of the same height as in A. Here there is room for several components on either side. 1000 transfers would give curve C which is a relatively thin band and provides space for several times the number of components that a run of 100 transfers could reveal. Thus as n increases 10 fold, the selectivity of the method increases several fold.

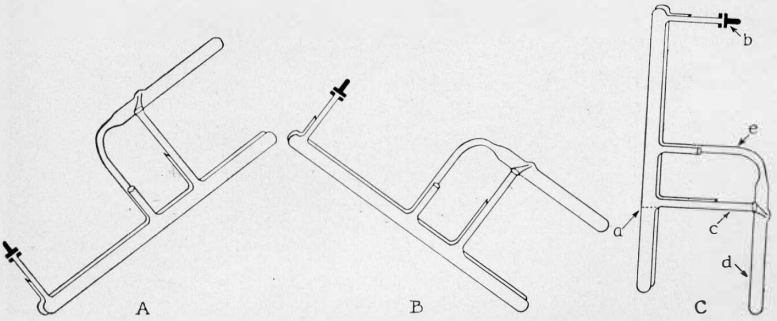

FIG. 5. Drawing of a single unit of the distribution apparatus.

The reliability of the method as an ultimate criterion of purity and as a highly selective tool for separation rests on a number of considerations which are interdependent but which may be conveniently grouped into three groups.

1. Apparatus or mechanical considerations.
2. Analysis; methods for estimating the total solute content of a given cell.
3. Systems; the discovery of two immiscible phases which show sufficient selectivity and yet give constant partition ratios with different concentrations.

In the development of apparatus for this work, most of the credit should go to Mr. Otto Post. He has produced several models made from steel which were adequate for carrying n to as high as 100 or more. I shall not take time to describe them here since they are

described in several publications.[2, 3, 4] It would appear more worthwhile to discuss the newer glass models which in their present stage of development are adequate for carrying n, the number of transfers, to 1000 or more (several hundred thousand actual extractions).

An individual unit is shown in Figure 5. A volume of the heavier phase is placed in each unit which is sufficient to fill the tube to the mark a in Figure 5, position C. An equal volume of the lighter phase is then added together with the solute through the flat ground

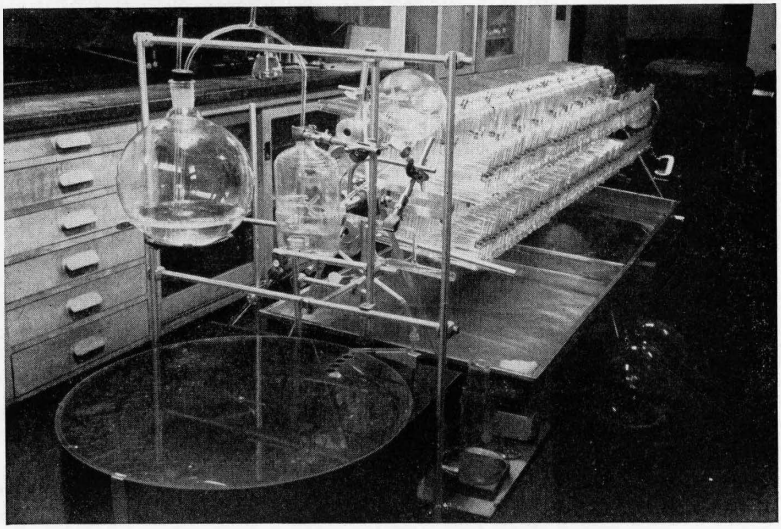

FIG. 6. Front view of the 220 tube automatic distribution apparatus.

joint at b while the unit is in position B. Equilibrium is established by rocking back and forth between positions A and B. The phases are allowed to separate in position B. The unit is then tilted to position C where the upper decants through c into d. Upon tilting the unit back again to position A, the contents of d flow out through e into the adjoining unit. Thus by simple tilting movements the upper layers are caused to move from the first tube to the last tube in the series in stepwise fashion.

The whole apparatus (Fig. 6) contains 220 such units and permits 220 quantitative extractions to be made simultaneously in approximately two minutes. It is fully automatic and is operated by

an electrical robot designed and built by Mr. Blum and Mr. Janes of the Institute instrument shop. There are two rows of the glass units. An upper phase starting at the upper far end migrates toward the robot but on reaching the end, it drops down to the lower series and migrates through these in the other direction. It leaves the train at the far end.

The apparatus can thus be operated in a manner analogous to a fractionating column but here upper layer must be added at each

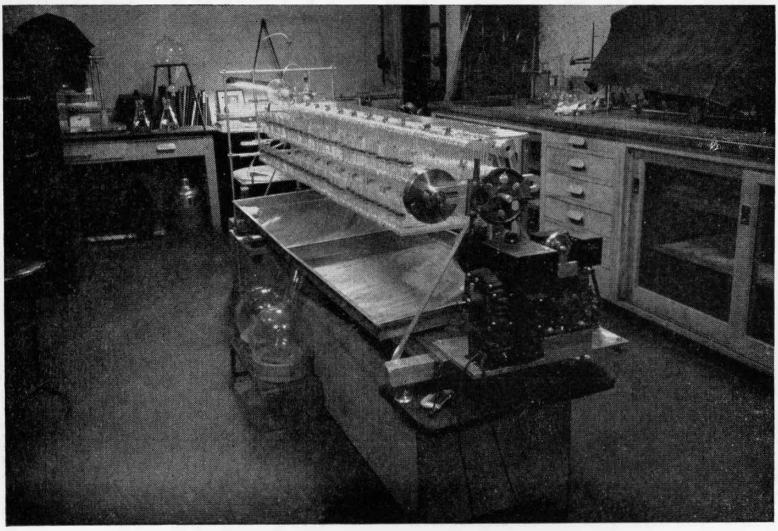

FIG. 7. Rear view of the 220 tube distribution apparatus.

transfer at the beginning of the column and a fraction removed simultaneously at the other end. Provision has been made for this. The apparatus includes an automatic filling device operating as a simple dipper which fills from a reservoir and empties into the first tube of the train on each transfer. The fraction collector is essentially the Stein and Moore[5] fraction collector. The fractions collected in this way are analogous to the effluent fractions in chromatography. Figure 7 shows the filling device and fraction collector.

Once a distribution is completed, one is confronted with many tubes to be analyzed. The problem here is much the same as in chromatography but it is easier in that higher concentrations usually

emerge and the solvents can be residue free. Therefore simple evaporation of an aliquot gives the ideal result. However, all the other analytical tools such as spectroscopy, titration, colorimetric estimations, bioassay, etc. are valuable. We attempt whenever possible to use weight as the basic method.

A good example of the separating power and the possibilities of the method is given by a trial run made by Dr. Hausmann and myself

$$y = \frac{1}{\sqrt{2\pi n \frac{1}{K}}} \cdot e^{-\left(\frac{x^2}{2n\frac{1}{K}}\right)}$$

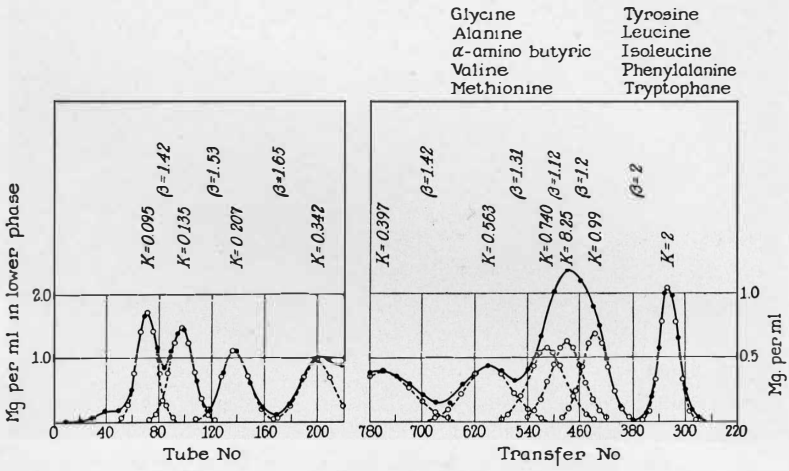

FIG. 8. Distribution of an artificial amino acid mixture.

in testing the apparatus. An artificial mixture of 10 amino acids, 300 mg. of each was prepared. The system chosen was n-butanol-5% hydrochloric acid. At the beginning the sample was equally distributed in the first 8 tubes. Figure 8 shows the result. The chart on the right is the elution or withdrawal pattern while that on the left gives the pattern of the material which remained in the apparatus after 781 transfers. The actual operating time of the run was approximately 22 hours. The curves were determined by evaporating an aliquot and weighing the residue.

Theoretical curves were calculated and fitted as shown. The

equation given at the top of the pattern for calculating a withdrawn curve is an appropriate modification of the equation for the normal curve of error.

The first band to emerge was tryptophane. It was slightly wider than the calculated because it emerged too soon to overcome completely the effect of placing the sample in a bank of 8 tubes instead of in a single tube.

The second large band is composed of three components, from right to left, phenylalanine, leucine and isoleucine. The three overlapping theoretical bands were located by paper chromatography and that of phenylalanine by ultraviolet absorption spectra. In order to resolve this band further, it could be fed in sequence back into a freshly charged machine or the fractions could be recombined, the solute recovered and distributed in a system which gives a better spread of the partition ratios. We have successfully made use of both approaches but they will not be discussed here.

The next band to emerge was tyrosine and the next was methionine. Valine was about to emerge from the train. Then came α-aminobutyric, alanine and glycine.

Paper chromatography is especially useful for interpreting overlapping bands. A broad filter paper can be spotted with material from tubes at regular intervals and all developed at once, as has been done in Figure 8 for the last named four bands. Thus the separation is confirmed.

One might logically inquire at this point what the fundamental difference between the "partition chromatography" of Martin and Synge[6] and "countercurrent distribution" is and when to use advantageously the one approach instead of the other. It can definitely be said that the first differs in being a continuous process while the latter is a strictly discontinuous process operating at equilibrium. Contrary to the views of the English school, it is our belief that liquid-liquid extraction plays a small role in such type of chromatography but that surface action is mainly responsible for the extraordinarily good separations.

Countercurrent distribution cannot at present spread the amino acid bands as far apart as the beautiful separations obtained by Moore and Stein.[7] On the other hand, its efficiency does not fall at higher concentration levels. It would be possible to get the same

pattern as shown in Figure 8 with 4 or 5 fold the amount of solute and this would yield 1-2 gms. of each pure amino acid, an amount ample for chemical study if an unknown amino acid should turn up. Furthermore, one can also apply to advantage 4 or 5 fold the number of transfers used in this run.

The paper chromatography of Consden, Gordon and Martin[8] requires much less time and labor. However, it is restricted to solutes for which sensitive color tests, spectroscopic tests and biological assays are available. Countercurrent distribution is not limited in this respect. It also shows at the present time relatively greater promise with the higher molecular weight substances as will be shown later on. These few comparative remarks might well be summed up by saying that no one technique is supreme in biochemistry but that all are needed for the work which lies ahead.

The choice of a suitable system to accomplish a given purpose is probably the greatest problem to be met in connection with this work. It requires a familiarity with the properties of the many solvents now available. Although one can learn a great deal about the suitability of a given phase pair by determination of individual partition ratios, an actual distribution gives much more information. Therefore, most of the remainder of this lecture will be in the nature of a discussion of several typical separations made with different classes of compounds difficult to separate by conventional techniques.

The first real application of the method was to the field of the synthetic quinoline antimalarials. It was found that the partition ratio was not constant at different concentration levels, presumably due to the fact that the solute was in a different state of aggregation in each of the two phases. This produces a badly skewed curve in a distribution and in most cases, partly defeats the purpose. Such a departure from the ideal was overcome in the antimalarials[9] by introducing a buffer into the aqueous phase. Time does not permit a discussion of the equilibria involved. The overall result permitted us to attain high selectivity and more than once to separate mixtures which fractional distillation, fractional crystallation of mixtures, and chromatography failed to separate.

Buffers in the systems have since proven most useful for separat-

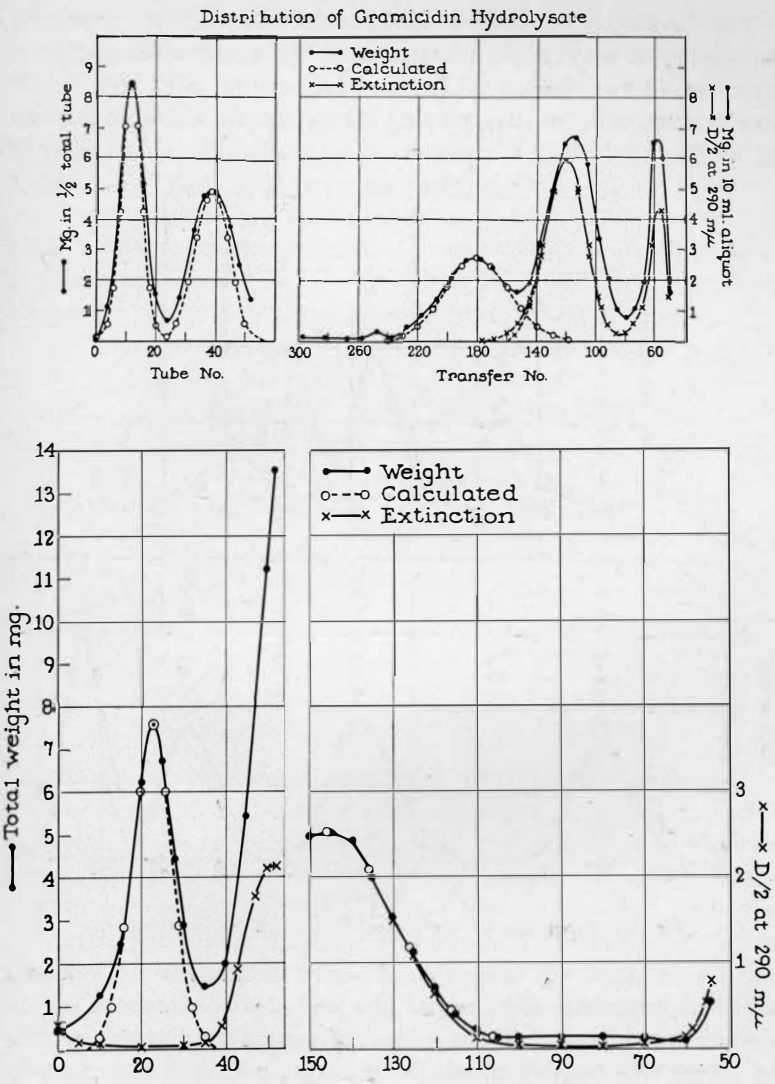

FIG. 9. Distribution of amino acids from the hydrolysis of gramicidin.

ing the lower fatty acids,[10] phenols,[11] alkaloids,[12] pyrimidines,[13] pencillins[14] and others.

On the other hand, buffers have contributed little to the specificity or ideality of amphoteric solutes such as the amino acids and polypeptides. In this class of substances addition of a salt to the system can be more effective in supplying selectivity. An example is given in Figure 9 which is the pattern of a gramicidin hydrolysate. In

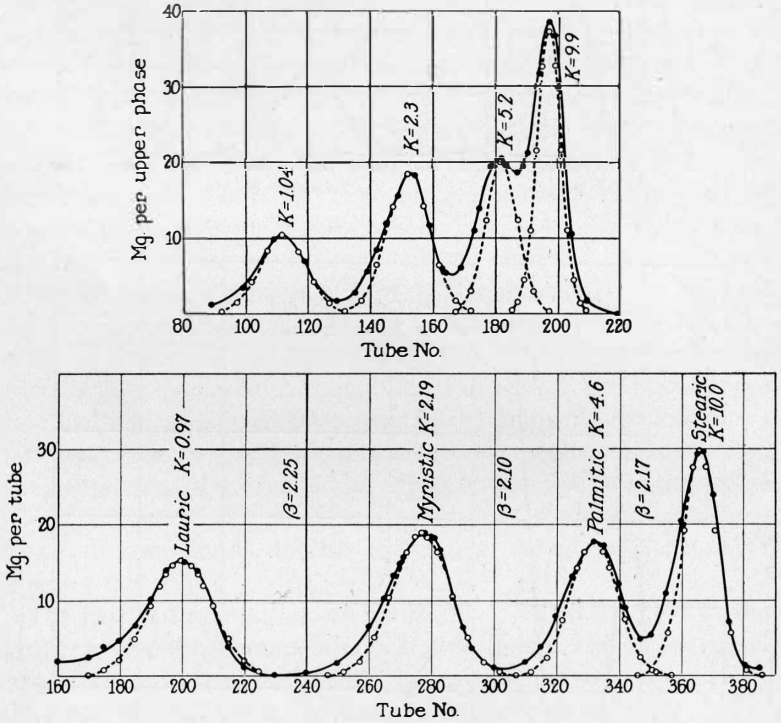

FIG. 10. Separation of the higher fatty acids by distribution.

the upper chart the system used was 2-butanol/30 % aqueous ammonium acetate. The contrast of a weight curve versus an optical density curve showed the second band from the right to be a mixture of closely overlapping bands. Redistribution of this material in a system without salt gave the lower pattern in which the peptide, valylvaline, is completely separated from tryptophane.

Buffers cannot be used in the separation of the higher fatty acids because of the formation of emulsions. Such acids are detergents

and known to form micelles of indefinite molecular size in different environments and therefore show a strong tendency to give a skewed distribution. Yet when a considerable concentration of acetic acid is incorporated into the system satisfactory separations can be made. Apparently the higher fatty acid associates preferentially with the acetic acid present in great predominance. A separation is shown in Figure 10. This was made on an artificial mixture of 500 mg. each of stearic, palmitic, myristic and lauric acids. The upper curve gives the pattern at 220 transfers while the lower was obtained by letting the machine run longer, to 400 transfers.

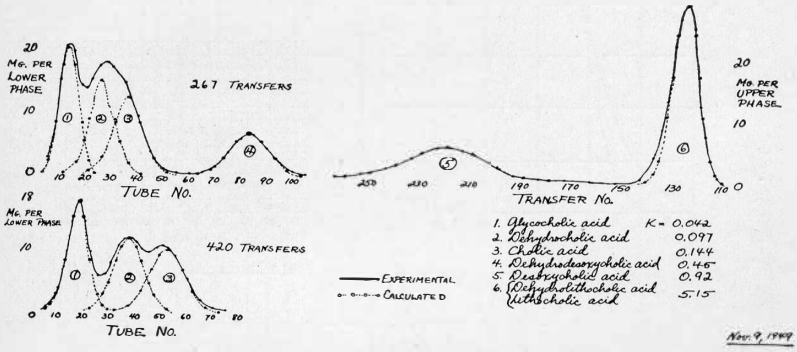

FIG. 11. Separation of bile acids by distribution.

Apparently considerable separation in the bile acid group can also be accomplished by the same approach. Dr. Ahrens and I have made a few trial runs. In Figure 11 a separation on an artificial mixture of seven members of the group is shown. Again the lower pattern shows the first three bands of the upper pattern from the left, carried to higher numbers of transfers. Only two of the seven components seriously overlap in this system.

The method is also promising in the field of the naturally-occurring phosphoric esters which are so difficult to purify. In this connection one might use as an example the excellent work of Hogeboom and Barry[15] in purifying the coenzyme, diphosphopyridine nucleotide.

The group of workers at the Squibb Institute have approached the problem of systems for strongly hydrophilic substances in an

interesting way.[16, 17, 18] The *Streptomyces* antibiotics are basic substances belonging to the polysaccharide group and are not appreciably soluble in organic solvents. Yet when lauric acid or one of the group of substances known as anionic detergents is incorporated into a system such as butanol-water, a satisfactory partition can be

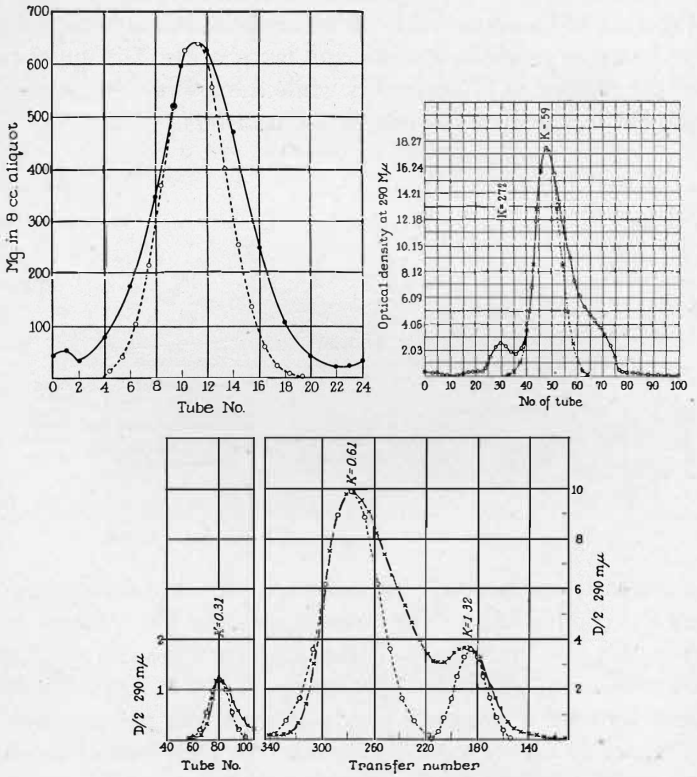

FIG. 12. Separation of gramicidins by distribution.

made. Moreover, sufficient selectivity is obtained to permit separation of the various members of the group. The same approach has even permitted separation of the high molecular weight polysulfuric acid ester, heparin, into two biologically active fractions.[19] Here lauryl amine was the active distributing solute. Such substances have been called "Carriers" by the Squibb group. Many possibilities

with the higher molecular weight substances are suggested by this approach.

For the past two years we have had the naturally-occurring polypeptides in the molecular weight range of 2000 to 5000 under study. The classes studied have been gramicidin, tyrocidine, gramicidin-S, bacitracin and more recently, polymyxin. In each case definite separations have been achieved. With increasing numbers of transfers Dr. Gregory and I found crystalline gramicidin to give

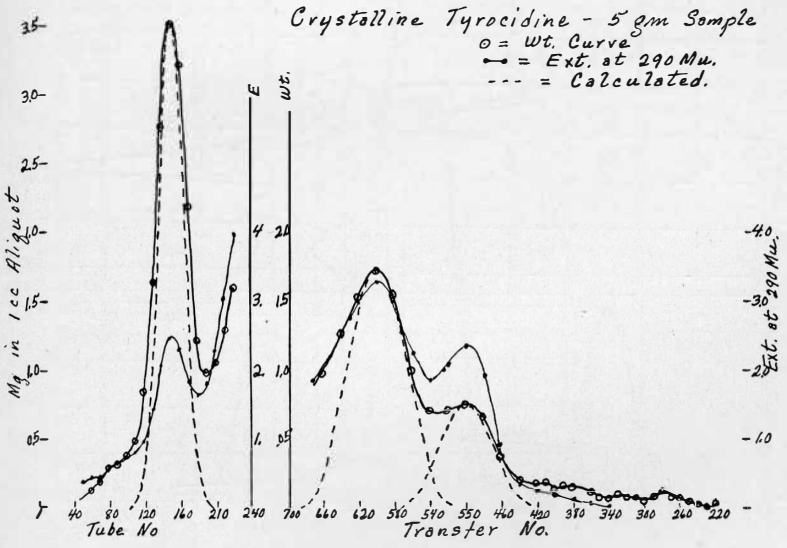

FIG. 13. Separation of tyrocidines by distribution.

the patterns shown in Figure 12. We now must deal with three crystalline gramicidins which differ slightly in their amino acid content. All these were found to be present in samples highly purified by fractional crystallization.

Tyrocidine gave a similar picture, Figure 13. Tyrocidine was known to be a mixture through chromatographic studies.[20] Here the differences in the amino acid content appear to be more quantitative in nature than qualitative. At least four major components are shown.

Gramicidin-S gave a similar result, as shown in Figure 14. A

crude preparation obtained from Sharp and Dohme gave the two upper patterns at 340 and 537 transfers. The latter was made on a 5 gram sample. A sample highly purified by fractional recrystallization and supplied by R.L.M. Synge gave the lower pattern. His sample contains only about 90% of the major component of the crude preparation. Obviously fractional crystallization alone is not a very effective method of getting a pure component from this class of substance.

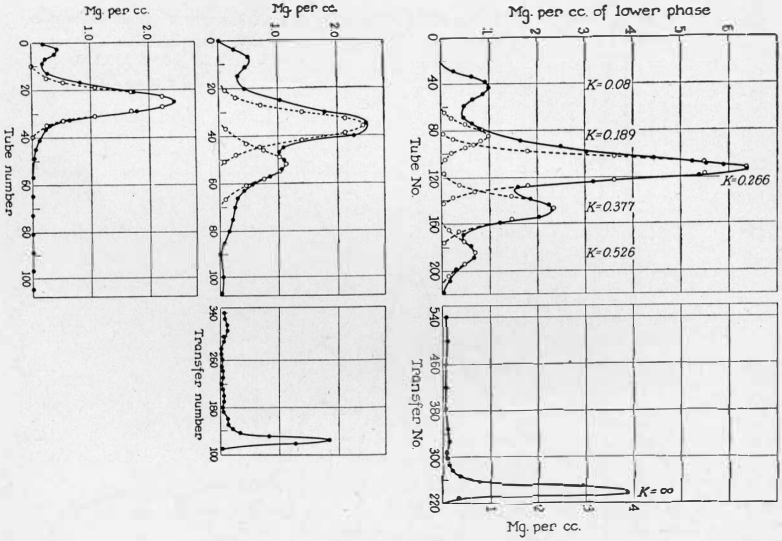

FIG. 14. Separation of gramicidin-S peptides by distribution.

Some samples of bacitracin supplied by the Commercial Solvents Co. seemed to approach purity in the systems studied. Others could be resolved into more than one component as shown in Figure 15. Four components are indicated.

The polymyxin preparation, Figure 16, we have studied so far was a crude preparation supplied by the American Cyanamid Co. It could be resolved readily into at least six or seven components which appeared to contain the same amino acids on hydrolysis and paper chromatography.

For the use of countercurrent distribution in the study of pep-

tides more closely related to higher animals, one can refer to the extremely interesting work of Livermore and du Vigneaud.[21] They have fractionated the active oxytocic principle from the hormones

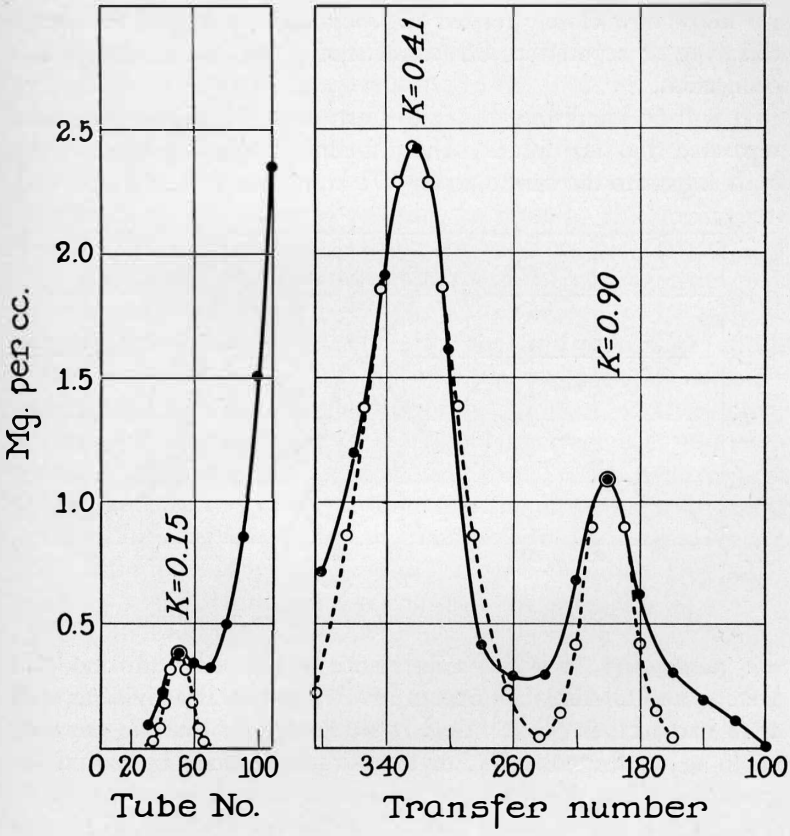

FIG. 15. Separation of bacitracin peptides by distribution.

of the posterior lobe of the pituitary. Here the most active preparation thus far isolated was obtained and at the same time, concrete evidence for purity was presented even though the substance did not crystallize.

At this stage in the study of peptides the question which is foremost would appear still to be that of "purity." Will a com-

ponent purified to the point of a single band with a thousand or more transfers in more than one system still be a mixture? We cannot answer the question with assurance as yet. We believe we have demonstrated a considerable degree of separating power but do not yet know how closely related two peptides can be and yet escape this type of separation. Structural studies will be needed in this connection.

It will be interesting to try to learn how the peptides we have separated thus far differ. Is the difference a slight qualitative one with respect to the amino acids? We know this to be the case with

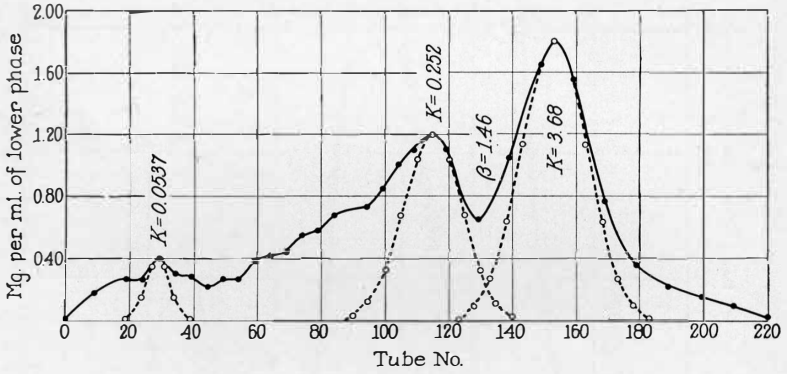

FIG. 16. Separation of polymyxin polypeptides by distribution.

the gramicidins. Is it also a difference in the order and mode of linkage and to steric differences involving optical activity as well as to many other possibilities? It would appear that we can now begin to see the possibility of true structural study by partial degradation as one approach for the answer to such questions. Chromatography, ion exchange and countercurrent distribution for this purpose are our most promising tools. Each will gain in significance by support from the others.

It seems rather striking that one after the other the polypeptides are found to occur in families. These now include the ergot alkaloids, the penicillins, the tyrocidines, the gramicidins, the gramicidin-S group, the polymyxins and others. Perhaps the enzymes which synthesize these substances are not as specific as we once thought them to be.

Finally, the subject of proteins may be mentioned. We have not made an attempt as yet to fractionate such substances. Obviously it will be a much more difficult task than the peptides and we have been trying to consolidate the gains already made with the smaller molecular weight substances.

However, since the polypeptides occur as mixtures so closely related one is tempted to extrapolate to the proteins and venture the prediction that even the most highly purified proteins are complex mixtures of closely related molecular species.

Before the organic chemist can begin serious structural work on such large molecules as proteins, he must be able to degrade them to smaller fragments. Degradation to too small a fragment such as an amino acid does not reveal much. Large fragments must be obtained. But in order to learn how to degrade proteins to large fragments we must first be able to separate and characterize them.

Our work with the peptides has been planned with this objective in mind and with the thought previously expressed by others[22] that the tyrothrycin group in fact do represent large fragments split from proteins. Obviously we are now nearly ready for some careful studies designed to try to control the splitting of a protein. It would not at all be surprising to find that our tools are yet inadequate and that a still higher order of separating power is required. This possibility, however, should not deter us from the attempt.

This has been mostly a short review of the work in our laboratory. It is the work of a group. At present Dr. W. Hausmann, Dr. E. H. Ahrens, Jr., and Dr. E. Harfenist are collaborating with me. However, the work of Dr. J. D. Gregory, Dr. G. T. Barry and Dr. Y. Sato as well as previous collaborators should also be mentioned.

BIBLIOGRAPHY

1. Saal, R. J.; and van Dyck, W. J. D.; World Petroleum Congress, 1st, *Proc., Roy. Soc. London, 2*:352, 1933.
2. Craig, L. C.: *J. Biol. Chem., 155*:519, 1944.
3. Craig, L. C.; and Post: *Anal. Chem., 21*:500, 1949.
4. Craig, L. C.; *Fortschr. d. Chem. Forschung, 1*:292, 1949.
5. Stein, W. H.; and Moore, S.: *J. Biol. Chem., 176*:337, 1949.
6. Martin, A. J. P.; and Synge, R. L. M.: *Biochem. J., 35*:1358, 1941.
7. Moore, C.; and Stein, W. H.: *J. Biol. Chem., 178*:53, 1949.
8. Consden, R.; Gordon, A. H.; and Martin, A. J. P.: *Biochem. J., 38*: 224, 1944.

9. CRAIG, L. C.; GOLUMBIC C.; MIGHTON, M.; and TITUS, E.: *J. Biol. Chem., 161*:321, 1945.
10. SATO, Y.; BARRY, G. T.; and CRAIG, L. C.: *J. Biol. Chem., 170*:501, 1947.
11. GOLUMBIC, C.; ORCHIN, M.; and WELLER, S.; *J. Am. Chem. Soc., 71*:2624, 1949.
12. FRIED, J.; WHITE, H. L.; and WINTERSTEINER, O.: *J. Am. Chem. Soc., 71*: 3260, 1949.
13. TINKER, J. F.; and BROWN, G. B.: *J. Biol. Chem., 173*:585, 1948.
14. BARRY, G. T.; SATO, Y.; and CRAIG, L. C.: *J. Biol. Chem., 174*:221, 1948.
15. HOGEBROOM, G. H.; and BARRY, G. T.: *J. Biol. Chem., 176*:935, 1948.
16. TITUS, E. D.; and FRIED J.: *J. Biol. Chem., 168*:393, 1947.
17. PLAUT, W. E.; and MCCORMACK, R. B.: *J. Am. Chem. Soc., 71*:2264, 1949.
18. O'KEEFFE, A. E.; DOLLIVER, M. A.; and STILLER, E. T.: *J. Am. Chem. Soc., 71*:2452, 1949.
19. O'KEEFFE, A. E.; RUSSO-ALESI, F. M.; DOLLIVER, M. A.; and STILLER, E. T.: *J. Am. Chem. Soc., 71*:1517, 1949.
20. SYNGE, R. L. M.; and TISELIUS, A.: *Acta chem. Scandinav., 1*:749, 1947.
21. LIVERMORE, A. H.; and DU VIGNEAUD, V.: *J. Biol. Chem., 180*:365, 1949.
22. DUBOS, R. J.; and HOTCHKISS, R. D.: *Tr. & Stud. Coll. Physicians, Philadelphia, 10*:11, 1942.