CHROM. 4365

EVALUATION OF SOLVENTS USED IN SUMMARIZED CHROMATOGRAMS

JESÚS SOUTO
Equipo de Antibióticos y Factores de Crecimiento, Instituto de Microbiología e Industrias Agropecuarias, I.N.T.A. Buenos Aires (Argentina)

AND

AMALIA GONZALEZ DE VALESI
Departamento de Biología, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (Argentina)

(Received August 11th, 1969)

SUMMARY

Based on some concepts from the Information Theory, the solvents used for antibiotic substances by ISHIDA *et al.* and BETINA were evaluated relative to each other by means of summarized chromatography.

The information that a summarized chromatogram system supplies should depend mainly on three factors: (1) distribution of the R_F of the reference substances in each solvent; (2) interrelationship between the R_F values that a given substance shows in the different solvents; and (3) experimental error.

Thus, admitting a constant error, it was found that among the 11 solvents tested, those that gave the most information were water; butanol; 3% aq. ammonium chloride; benzene–methanol (4:1) and butanol–methanol–water (4:1:2) with and without helianthin. The phenolic solvents were less informative; ethyl acetate is mediocre in all cases.

Although the information is very high in some of the 11 solvents tested (up to 85% of H_{max}), only water supplies information independent of the others, which are redundant among each other.

INTRODUCTION

The papergram or summarized chromatogram procedure is usually used in our laboratories for the rapid identification of antibiotic substances. The solvents used include those suggested by ISHIDA *et al.*[1], and, more recently, those suggested by BETINA[2] have also been employed. We have been interested in evaluating these solvents in order to determine which supplies more and which supplies less information.

According to the Information Theory, all R_F values in a solvent must be equally probable for this solvent to provide the maximum information ($H_{s(max)}$). Furthermore, to obtain the maximum information from a summarized chromatogram

TABLE I

EXPERIMENTAL R_F VALUES × 100

Solvent systems used: 1 = butanol satd. with water; 2 = 3% ammonium chloride; 3 = 80% phenol; 4 = 80% phenol in NH_3 atm.; 5 = acetone–water (1:1); 6 = butanol–methanol–water (4:1:2) + 0.15 g helianthin; 7 = butanol–methanol–water; 8 = benzene–methanol (4:1); 9 = water; 10 = ethyl acetate satd. with water; 11 = benzene satd. with water. Solvents 1 to 9 correspond to the system of ISHIDA *et al.*; solvents 9, 1, 10 and 11 correspond to BETINA's system.

Antibiotics	*Solvent system*										
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
Actinomycin	92	25	100	84	43	90	89	88	20	92	15
Aminocidin sulphate	0	87	4	77	6	23	3	0	27	0	0
Chloramphenicol	86	85	91	90	93	88	89	45	86	85	9
Chlortetracycline hydrochloride	28	51	81	82	31	64	56	0	43	0	0
Colistin sulphate	0	94	65	96	49	54	24	14	14	0	0
Cycloheximide	72	87	92	91	91	80	80	59	93	70	13
Demethyltetracycline	25	31	60	73	30	45	43	0	28	27	0
Dihydrostreptomycin sulphate	0	95	8	95	0	57	0	0	0	0	0
Etamycin	93	33	90	94	92	95	94	85	39	93	15
Filipin	95	0	97	87	94	93	92	0	0	75	0
Griseofulvin	90	0	96	91	80	90–45	92	94	0	95	89
Leukomycin tartrate	94	82–0	96	95	94	93	94	93	28	95–0	15
Lincomycin hydrochloride	38	100	90	95	89	71	66	65	100	0	0
Meticilin	34	97	95		97	49	43		95	0	0
Mikamycin A	97	0	95		92	98	97	23	3	98	15
Mikamycin B	91	13	95	81	94	96	95	55	25	97	27
Misionin	72	0	93	86	67	79	72	50	0	0	0
Neomycin	0	95	0	77	0	45	0	0	0	0	0
Nystatin	6	0	84	68	0	55	53	0	0	0	0
Oxacillin	46	88	93	96	95	66	68	0	94	58–0	0
Paromomycin	0	92	0	81	0	22	0	0	0	0	0
Penicillin G-Na	40	94	88	94	91	62	64	0	100	0	0
Penicillin V-K	37	88	88	94	94	57	62	0	100	0	0
Polymyxin B sulphate	0	97	67	96	0	67	52	0	0	0	0
Puromycin	49	65	92	95	86	73	67	43	18	0	0
Pyrrolidinemethyltetracycline	7	65	83	82	43	55	36	25	37	0	0
Rifamycin O	96	0	94	100	93	100	100	100	65–0	100–0	54–0
Rifamycin S	95	28	93	94	92	93	94	83–0	77	92	69
Rifamycin SV-Na	91	43	94	93	81	92	91	91–0	79	100	50
Spiramycin base	86	94	100	100	93	94	88	0	7	99–0	0
Staphylomycin	93	45	100	0	93	94	92	18	34	100–0	0
Streptomycin sulphate	0	100	0	85	0	36	0	0	4	0	0
Terramycin base	25	74	92	100	54	67	55	10	52	0	0
Tylosin base	92	92–0	92	100	92	87	85	35	43	100–0	0
Viomycin sulphate	0	100	0	100	0	49	0	0	0	0	0

($H_{p(max)}$), the R_F obtained with a solvent must be independent of the R_F in another solvent.

Observation of the R_F values obtained with a group of substances and several solvents, such as the data published by BETINA[2], or those shown in Table I, suggests that the distribution of the R_F values is not equiprobable because extreme R_F values occur more frequently than median ones. In this case the graphical representation of the frequency as a function of R_F would give a U-shaped curve.

With regard to the independence of the R_F value of a solvent as regards another solvent, if the functional chromatographic mechanism is the same, we should expect

some relation between them, although it may be complicated or unknown. An example of this is the F factor found by CONNORS[3] which relates the R_F values of uracil derivatives, barbituric acids, inorganic phosphates, etc. in different solvents when the functioning mechanism is partition. Later, SOCZEWINSKI[4] used R_M values to find the same factor.

A third variable that affects the information supplied by a papergram is the experimental error, which, according to the Information Theory, establishes the number of signals, n, to be considered, that is the number of discernible R_F values.

The present report deals essentially with the study of the R_F distribution for which the experimental error can be ignored; however, it is convenient to mention it here.

In the methods described here, the error is undoubtedly greater than in conventional chromatography because data obtained with substances that are minor components in broths of variable composition are compared with those obtained from pure reference substances. In the symposium organized by the Chromatography Group of the Czechoslovak Chemical Society at Liblice, in June 1961, an error of 10% was admitted for conventional chromatograms. We have found variations of up to 0.2 R_F; in addition it was found that the error depends on the solvent. Water showed the least error (variations of 0.1 R_F) while acetone–water (1:1) showed the greatest error, which we consider a question of speed and equilibrium of the chromatographic process. We believe that expression of the R_F values for these methods would be more logical with only one decimal place ($n = 10$) rather than two decimal places ($n = 100$) which is more commonly used.

Assuming the same experimental error for all solvents, the maximum information given by each would be:

$$H_{s(\max)}\, n = \log_2 n \text{ (bits)}$$

If a chromatographic system consists of m solvents and the condition of independence is fulfilled, the maximum theoretical information that it can supply would be:

$$H_{p(\max)} = m \log_2 n$$

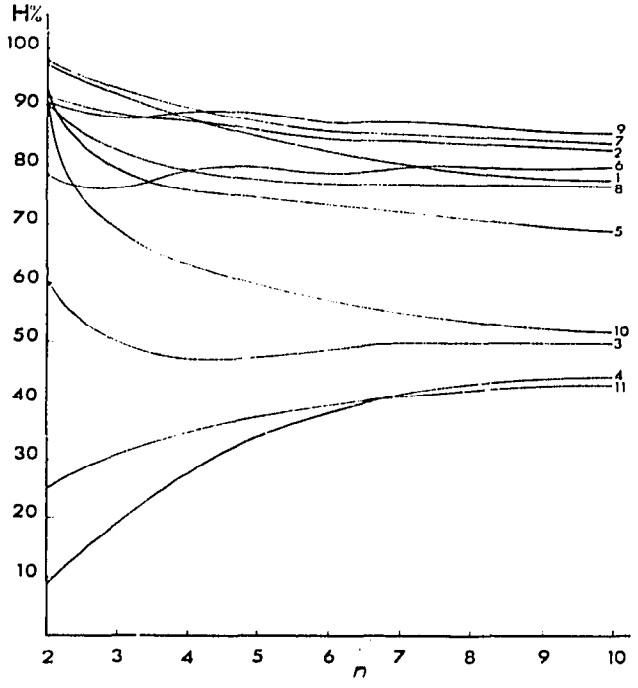
Supposing we had 2,000 reference antibiotics to identify one unknown, we would need $\log_2 2{,}000 = 10.96 \cong 11$ bits of information.

If $n = 100$, $H_{s(\max)} = \log_2 100 = 6.66$ bits so that we would only need two solvents.

If $n = 10$, $H_{s(\max)} = \log_2 10 = 3.32$ bits, and we would need four solvents to solve the problem. Experience has shown that a problem like this is unlikely to be solved with just a few solvents so that the real information obtained (H_p) would be less than the $H_{p(\max)}$ that we have just calculated.

The procedure chosen to study the distribution of the experimental data in each solvent consisted of dividing the R_F scale into a number of equal sectors and considering the R_F data within each sector as equal signals. This allowed us to calculate the information of the solvent (H_s) for n signals and to relate it to $H_{s(\max)}$ for the same n.

$$H\% = \frac{H_s}{H_{s(\max)}} \cdot 100$$



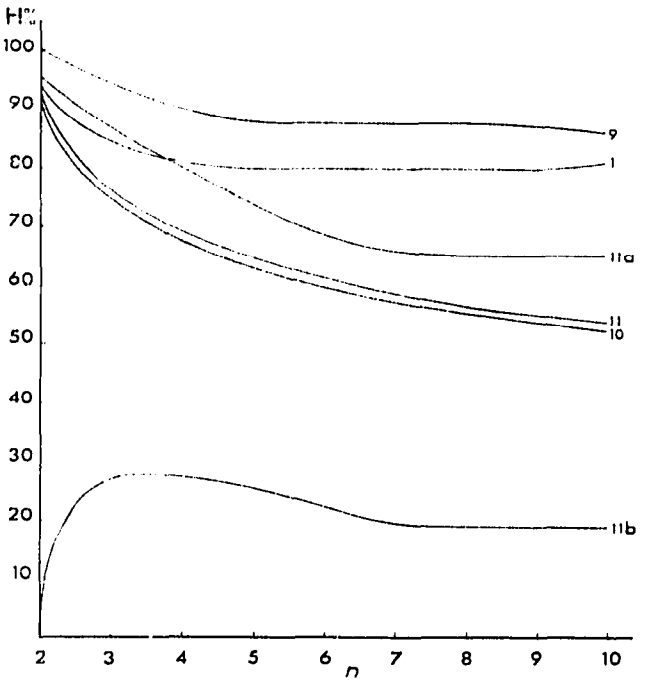


Fig. 1. $H\% = f(n)$ with our data. Numbers correspond to the solvents listed in Table I.

Fig. 2. $H\% = f(n)$ with data published by BETINA. Numbers correspond to the solvents in Table I, (11a) benzene considering only fungal antibiotics, (11b) *idem* with no fungal antibiotics.

The H_s value was calculated according to the formula of SHANNON AND WEAVER[5]:

$$H_s = \Sigma p_i \log_2 p_i$$

p_i being the probability of the appearance of each signal.

We calculated $H\%$ for each solvent and for a series of n values and plotted $H\% = f(n)$, thus obtaining curves that enabled us to evaluate the solvents relative to each other (see Figs. 1 and 2). When making the comparisons, we chose the interval of n values from 6 to 10 in order to obtain a more accurate evaluation, bearing in mind the number of experimental data and former considerations about the experimental error.

The experimental values used were partly taken from those published by BETINA (only those for the four basic solvents, 62 data each), and partly those obtained by us using the same solvents and the solvents proposed by ISHIDA *et al.* (in all 11 solvents, 35 data each), (Figs. 1 and 2).

Our data were obtained by a technique described elsewhere[6], for which the containers and quantities of solvent are smaller than the ones used by BETINA. These details are not considered important, since the problem has been studied from a comparative point of view.

The information obtained with each of the 11 solvents studied is shown in Fig. 1. It is possible to see a group of solvents that give quite a high $H\%$ (between 77 and 87%); these include water, *n*-butanol–methanol–water (4:1:2), 3% aq. ammonium chloride, *n*-butanol saturated with water and benzene–methanol (4:1). The ace-

TABLE II

χ^2 VALUES

Solvent systems compared	χ^2		
	Our data	BETINA's *data*	
Satd. butanol *vs.* water	0.00	0.53	
3% Aq. ammonium chloride *vs.* water	1.44		
Benzene–methanol (4:1) *vs.* water	1.83		
Butanol–methanol–water (4:1:2) + 0.15 g helianthin *vs. idem* without helianthin	18.48		
Butanol–methanol–water (4:1:2) *vs.* satd. butanol	10.87		
Butanol–methanol–water (4:1:2) *vs.* water	2.61		
Satd. butanol *vs.* satd. ethyl acetate	10.29	31.30	
Satd. ethyl acetate *vs.* water	0.29	0.70	
Satd. benzene *vs.* water	2.45	0.34	1.82[a]
Satd. benzene *vs.* satd. butanol	5.48	16.60	8.49[a]
3% Aq. ammonium chloride *vs.* benzene–methanol (4:1)	7.34		
3% Aq. ammonium chloride *vs.* butanol–methanol water (4:1:2)	4.81		
3% Aq. ammonium chloride *vs.* satd. butanol	12.93		
Satd. butanol *vs.* benzene–methanol (4:1)	8.68		
3% Aq. ammonium chloride *vs.* satd. ethyl acetate	7.93		
Satd. ethyl acetate *vs.* satd. benzene		18.10	

[a] Values obtained considering only fungal antibiotics.

tone–water (1:1) solvent supplies slightly less information and finally four solvents (ethyl acetate saturated with water, 80% aq. phenol, 80% aq. phenol in an ammonia atmosphere and benzene saturated with water) supply little information (between 43 and 52%).

We obtained Fig. 2 from the data published by BETINA for his four basic solvents. This is similar to Fig. 1 in the sense that water and butanol are in the group of solvents supplying more information, while ethyl acetate and benzene are in the group supplying minor information.

Almost half of the antibiotics used by BETINA are produced by fungi. Compared with this, we only used a few antibiotics of that origin. To take this difference into account we studied the data obtained by BETINA for antibiotics produced by fungi

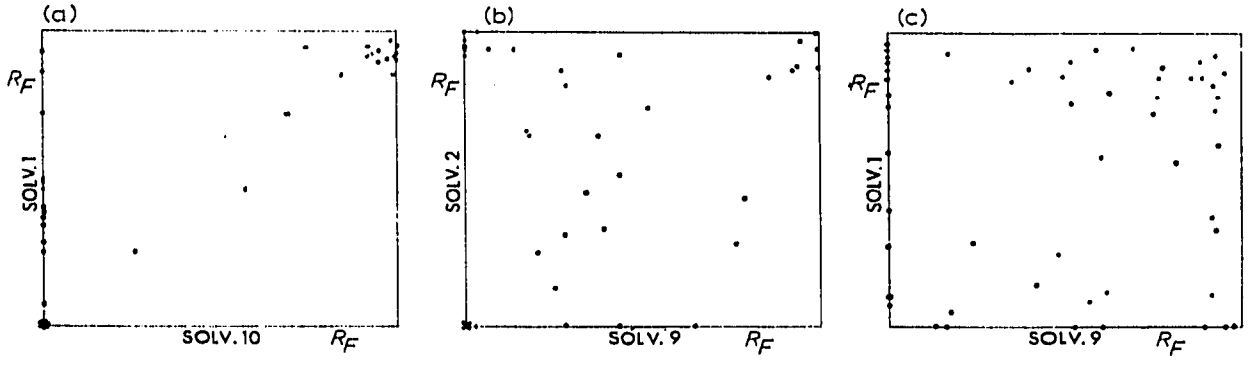


Fig. 3. Comparison of solvents. (a) Our data $\chi^2 = 10.29$, (b) our data $\chi^2 = 1.44$, (c) BETINA's data $\chi^2 = 0.53$.

and those from other sources separately. With respect to water, butanol and ethyl acetate, similar results were obtained for both groups of antibiotics, but benzene gives more information in the group of antibiotics from fungi (65%) than in the other group (20%). The explanation for this is that fungi produce simpler substances than actinomycetes and bacteria and they are often aromatic.

By applying the chains of MARKOFF[7], we can find how much the solvent information decreases due to redundance, but we do not consider that we have sufficient experimental data. However we can tell whether there is independence between the information given by two solvents if we present the R_F data of each of the substances graphically and afterwards test the independence hypothesis by the χ-square method. For this we divide the diagram into four equal sectors of 0.5 R_F per side. Fig. 3 shows the R_F distribution obtained with some pairs of solvents and Table II summarizes the χ^2 data obtained with the solvents that supplied a great deal of information.

Establishing the 0.05 significance level ($\chi^2 = 3.84$), we can only accept the independence hypothesis for the pairs when one of the solvents is water.

CONCLUSIONS

Among the solvents studied by BETINA, water and butanol supply more information. Benzene is only useful in the case of antibiotics produced by fungi; ethyl acetate is mediocre in all cases.

Some of the solvents proposed by ISHIDA *et al.* in addition to butanol and water are also very informative. These are 3% aq. ammonium chloride, benzene–methanol (4:1), butanol–methanol–water (4:1:2) with and without helianthin. The phenolic solvents were less informative.

Although much information is obtained in some of the eleven solvents tested (up to 87% of H_{max}), only water supplies information independent of the others, which are redundant among each other.

ACKNOWLEDGEMENT

We wish to express our thanks to Dr. FERMIN J. ALFONSO for his kind advice and discussion of the present paper.

REFERENCES

1 N. ISHIDA, T. SHIRATORI AND S. OKAMOTO, *J. Antibiotics (Tokyo), Ser. A*, 3 (1950) 880.
2 V. BETINA, *J. Chromatog.*, 15 (1964) 379.
3 K. A. CONNORS, *Anal. Chem.*, 37 (1965) 261.
4 E. SOCZEWINSKI, *Anal. Chem.*, 37 (1965) 439.
5 C. E. SHANNON AND W. WEAVER, *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, 1949.
6 B. L. EILBERG AND J. SOUTO, *Rev. Invest. Agrop.*, 3 (1966) 43.
7 W. FELLER, *An Introduction to Probability Theory and its Applications*, Wiley, New York, 1957.