

with that of the K-shell fluorescent X-rays is justified.

Nuclear Science Laboratories, K. L. ALLAWADHI.
Department of Physics, S. L. VERMA.
Punjabi University, B. S. SOOD.
Patiala 147 002, March 29, 1976.

1. Allawadhi, K. L. and Sood, B. S., *Phys. Rev.*, 1975, 11 A, 1928.
2. Rakavy, G. and Ron, A., *Ibid.*, 1967, 159, 50.
3. Schmickley, R. D. and Pratt, R. H., *Lockhead Research Laboratories Report*, L.M.S.C. 5-10-67-11 NYO-3839-9.
4. Scofield, J. H., *Lawrence Livermore Laboratory Report*, 1973, UCRL 51326.
5. Allawadhi, K. L., Mohanjit Singh and Sood, B. S., *Curr. Sci.*, 1973, 42, 86.

OBSERVED CORRELATION BETWEEN COMPOSITION AND MOLECULAR WEIGHT OF AMINO ACIDS IN PROTEINS

THE composition of various proteins, in particular, the distribution of amino acids has been a topic of considerable interest. Early work^{1,2} quoted in Meister³ would appear to yield some general conclusions, although some of these need revision in the light of recent analysis (Dayhoff⁴) which is based on primary sequence analysis. The latter analysis has yielded the relative frequency of occurrence of the amino acids according to the classification of polar, apolar, heterocyclic (and aromatic) and ionisable types.

The randomness or otherwise of the occurrence of these amino acids has been treated statistically⁵⁻⁷ and tables are available to assist in assessing whether or not their occurrence in a particular protein may be considered as random. Other statistical studies⁸⁻¹⁶ are aimed at the analysis of the relative frequency of occurrence of residues in relation to the characteristic secondary structure of the protein chain, with a view to using them in algorithms for predicting protein structure from the given sequence. So also analysis from the point of view of volume occupied by side chain is also available¹⁷. However as far as we are aware of, no study seems to have been made on the relative abundance of the amino acids in relation to the molecular weight of the constituents.

We have done this and the simple linear correlation that has been observed, is felt to be sufficiently interesting to merit reporting here. Though Dayhoff's Table D 23⁴ includes 108 proteins, a few fibrous proteins and also some small peptides have been included in it. In view of this we have selected 32 proteins whose structure by x-ray crystallographic

methods have been reported and hence the sequence is uniquely known. This we believe is a sufficiently good sample to yield statistically meaningful result. Figure 1 shows the relation between relative frequency of occurrence and the molecular weight.

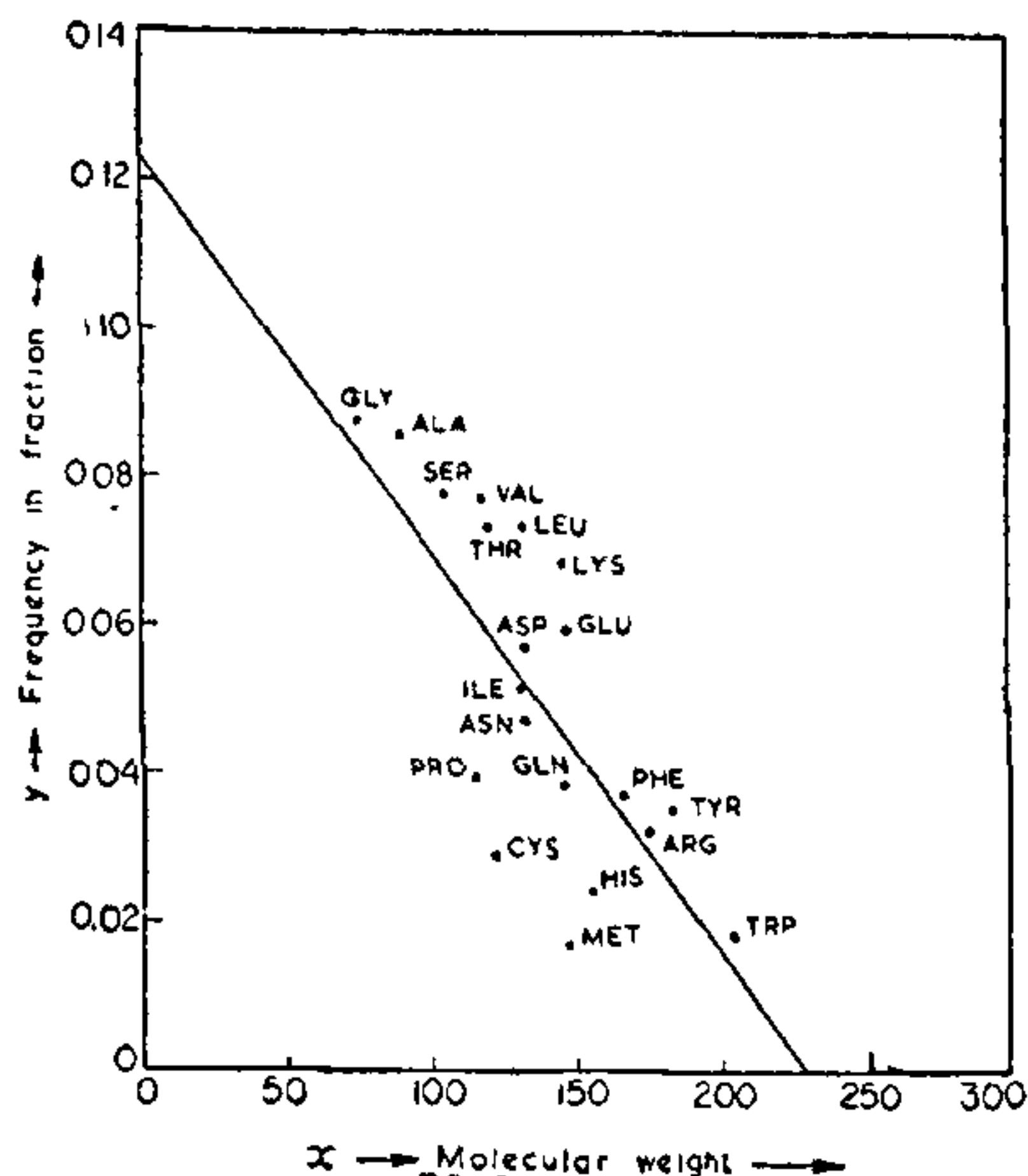


FIG. 1. Plot of molecular weight vs. the relative frequency of occurrence of amino acids in 32 proteins.

It is interesting to note that there is an approximately linear relation. The largest deviation from the above line is noticed for methionine and cystine and it is interesting that these two are amino acids containing sulphur. Among the rest, leucine (non-polar), lysine (polar) and histidine (heterocyclic) show somewhat large deviations.

The least squares equation to the line in Fig. 1 :
 $y = -0.00054 x + 0.12445$.

The equation leads to the result that nearly 67% of the amino acids are those having molecular weights less than 137 which is the mean of the 20 amino acids. At the two ends, the relative frequencies work out to be 8.4% for glycine and 1.4% for tryptophan. These are respectively about 70% more and 70% less than the mean (5%) to be expected if all the 20 amino acids are randomly distributed with equal probability.

Further analysis from this angle is in progress and results will be reported in due course.

Dept. of Crystallography and S. S. RAJAN.
Biophysics, R. SRINIVASAN.
University of Madras,
Guindy Campus, Madras 600 025,
India, July 30, 1976.

1. Tri-gram, G. R., *The Proteins* (Eds. Neurath, H. and Bailey, K., Academic Press, London, 1953), Vol. 1, Part A, p. 181.
2. —, *Advances in Protein Chemistry*, 1963, 18, 227.
3. Meister, A., *Biochemistry of Amino Acids*, Academic Press, New York, London, 1965, Vol. 1, Chap. I.
4. Davhoff, M. O. and Hunt, L. T., *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Maryland, 1972, Table D 23, D-356.
5. Holmquist, R., *J. Mol. Evol.*, 1975, 4 (4).
6. Laird, M. and Holmquist, R., *Ibid.*, 1975, 4 (4), 261.
7. Black, J. A., Stenzel, P. and Harkins, R. N., *J. Theor. Biol.*, 1975, 50 (1), 161.
8. Chou, P. Y. and Fasman, G. D., *J. Mol. Biol.*, 1973, 74, 263.
9. — and —, *Biochemistry*, 1974, 13, 211.
10. — and —, *Ibid.*, 1974, 13, 222.
11. Burges, A. W., Ponnuswamy, P. K. and Scheraga, H. A., *Israel J. Chem.*, 1974, 12, 239.
12. Lim, V. I., *J. Mol. Biol.*, 1974, 88, 857.
13. —, *Ibid.*, 1974, 88, 873.
14. Nagano, K., *Ibid.*, 1973, 75, 401.
15. —, *Ibid.*, 1974, 84, 337.
16. Ptitsyn, O. B., *Ibid.*, 1969, 42, 501.
17. Chothia, C., *Nature*, 1975, 254, 304.

THE EVALUATION OF THE ELECTRONIC TRANSITION MOMENT FOR A BAND SYSTEM OF DIATOMIC MOLECULE

It has been shown that the electronic transition moment¹ plays a very important role in the calculation of vibrational transition probabilities of diatomic molecules. The relative vibrational transition probability² $P_{v',v''}$ may be written as

$$P_{v',v''} = R_e^2(\bar{r}_{v',v''}) q_{v',v''} \quad (1)$$

where $R_e(\bar{r}_{v',v''})$ is the electronic transition moment, $\bar{r}_{v',v''}$ is the r -centroid and $q_{v',v''}$ is the Franck-Condon factor of the band in question.

The intensity in emission $I_{v',v''}$ of a band may be written as

$$I_{v',v''} = KN_{v'}E_{v',v''}^4 P_{v',v''} \quad (2)$$

in which K incorporates all the constants, $N_{v'}$ is the population of the level v' , $E_{v',v''}$ is the energy quantum of the band. Thus an array of relative $P_{v',v''}$ may be determined from $I_{v',v''}N_{v'}E_{v',v''}^4$ when the band intensities have been measured. From equations (1) and (2).

$$I_{v',v''} = KN_{v'}E_{v',v''}^4 R_e^2(\bar{r}_{v',v''}) q_{v',v''} \quad (3)$$

This equation may be used to determine the form of $R_e(\bar{r})$ for a band system.

Equation (3) may be employed to determine the relative variation of $R_e(\bar{r})$ with \bar{r} for the band system. Once a smooth curve has been obtained which shows the relative variation of $R_e(\bar{r})$ with \bar{r} , equation (1) may be employed to determine a smoothed array of $P_{v',v''}$. Values of $R_e(\bar{r}_{v',v''})$ are read from the curve at values of $\bar{r}_{v',v''}$ appropriate to each band. $R_e^2(\bar{r}_{v',v''}) q_{v',v''}$ is then a relative measure of $P_{v',v''}$ for each band. The array of $P_{v',v''}$ thereby obtained may be normalised such that $P_{00} = 1$.

In the paper of Mishra and Khanna³, the variation of $R_e(\bar{r}_{v',v''})$ with internuclear separation has been worked out by the method outlined by Nicholls⁴ and it is found that the curve shown in the figure is consistent with the relation $R_e(\bar{r}_{v',v''}) = \text{const. exp. } (0.205 r)$ over the range $1.20 \text{ \AA} < r < 1.45 \text{ \AA}$. The calculated $P_{v',v''}$ values for the bands of the A-X system of BF molecule are in fair agreement with the observed intensity distribution. Values of $P_{v',v''}$ are in good agreement with the values entered in the last column of Table I which are nothing but the product of F-C factor and r -centroid of the band. On this basis it can be inferred that vibrational transition probabilities for a band system of molecules for which $q_{00}\bar{r}_{00} \cong 1$, can be obtained without measuring the intensities of the bands. The details of this method along with the results of other band systems will be published elsewhere.

TABLE I
Data for BF (A-X) band system

Band	$\bar{r}_{v',v''}$ (Å)	$P_{v',v''}^*$	$P_{v',v''}$
(0, 0)	1957.4	1.000	0.999
(0, 1)	2001.5	0.271	0.280
(0, 2)	2068.1	0.023	0.024
(1, 0)	1911.0	0.227	0.218
(1, 1)	1962.7	0.577	0.578
(1, 2)	2016.3	0.481	0.502
(1, 3)	2071.6	0.022	0.023
(2, 1)	1917.1	0.330	0.321
(2, 2)	1968.2	0.198	0.196
(2, 3)	2021.3	0.539	0.558

* Data from Mishra and Khanna³.