PROCEEDINGS OF THE BIOCHEMICAL SOCIETY

The 500th Meeting of the Society was held at University College London on Monday, Tuesday and Wednesday, 15, 16 and 17 December 1969, when the following papers were presented:

COMPUTER AND INSTRUMENTATION GROUP COLLOQUIUM ON 'THE FUTURE IN COMPUTING'

Information Retrieval in Chemistry and Allied Sciences

BY A. K. KENT. (*U.K. Chemical Information Service, University of Nottingham, Nottingham, U.K.*)

The 'information explosion' is almost matched by the increase in both conventional and non-conventional information retrieval services. Computer-based services, both experimental and operational, have mushroomed during the past 5 years. The majority of these, however, are of limited significance to biochemists.

Essentially three major computer-based systems are likely to be of interest and value to them. The *Chemical Abstracts Service* (CAS) system, operated in this country by the U.K. Chemical Information Service at Nottingham, provides a service in chemistry that substantially covers biochemical interests. Current-awareness searches can be made on four 'free-text' data bases: *Chemical Titles* (CT), *Chemical-Biological Activities* (CBAC), *CA Condensates* and *Polymer Science and Technology* (POST). The first three of these cover biochemical areas, and CBAC is specifically biochemical in scope. The Institute for Science Information at Philadelphia (ISI) offers a current-awareness service *Automatic Subject Citation Alert*, which is unique in its ability to retrieve information by citation linkage and is particularly valuable in fringe areas of science, and are now making available *Index Chemicus* on magnetic tape. The *Index Chemicus* service is being investigated by the Oxford Experimental Information Unit in collaboration with the U.K. Chemical Information Service and has the important characteristic that it permits substructure search by way of Wiswesser Line Notations of all novel compounds reported. The MEDLARS (Medical Literature Analysis and Retrieval Service) system, which originates from the National Library of Medicine in Washington, is operated in the U.K. through the National Lending Library at Boston Spa. It differs from the CAS and ISI systems in being a 'controlled-language' data base and consequently requires a rather different approach to search strategy.

All these systems require the preparation by the user of a 'search profile', which is a set of search terms linked in some logical way. They are all current-awareness systems with the exception of MEDLARS, which may also be, and is normally, used for retrospective search.

In this paper we shall review the main characteristics of these data bases, the problems of their use, the cost of using them and the advantages that may be gained by their use.

Computer-Aided Approaches to Biochemical Reaction Mechanisms

BY B. CHANCE. (*Johnson Research Foundation, University of Pennsylvania, Philadelphia, Pa. 19104, U.S.A.*)

The Computer-Aided Determination of Peptide Sequences

BY R. A. UPHAM. (*Department of Organic Chemistry, University of Liverpool, Liverpool L69 3BX, U.K.*)

The mass spectra of acylated peptide esters and, in particular, their permethylated derivatives (Lederer & Das, 1967) show a regular fragmentation pattern corresponding to cleavages of the amide bonds. Other peaks arising by loss of CO from the acyl ion and by fragmentations in the side chains are also observed. The regularity of these fragmentation patterns and the ease of acquisition of accurate mass data in digital form make peptide spectra ideally suited to computer-aided interpretation.

Computer programmes have been devised (Biemann, Cone & Webster, 1966; Senn, Venkataraghavan & McLafferty, 1966) that recognize the first *N*-acyl-amino acid and then reconstruct the sequence by successive additions of amino acid units to give in turn other recognizable sequence ions. An alternative programme (Barber, Powers, Wallington & Wolstenholme, 1966) deduces the molecular ion from the atomic composition. The sequence is determined by comparison of the

difference between the molecular ion and the peaks in the spectrum with the composition of the amino acids.

Work is in progress on the MS-902/Argus 500 system recently installed at the University of Liverpool to extend the scope of amino acid sequence determination of peptides by mass spectroscopy. This system is capable of recording a complete high-resolution spectrum in 25s and immediately producing a paper tape of the accurate masses.

This rapid scanning capability may, in principle, be utilized to obtain sequential spectra of mixtures

(R. A. W. Johnstone & D. W. Payling, personal communication), and hence in the case of peptide derivatives to obtain amino acid sequences of the individual components.

Barber, M., Powers, P., Wallington, M. J. & Wolstenholme, W. A. (1966). *Nature, Lond.*, **212**, 784.

Biemann, K., Cone, C. & Webster, B. J. (1966). *J. Am. chem. Soc.* **88**, 2597.

Lederer, E. & Das, B. C. (1967). *Proc. 8th Eur. Peptide Symp.* p. 131. Amsterdam: North-Holland Publishing Co.

Senn, M., Venkataraghavan, R. & McLafferty, F. W. (1966). *J. Am. chem. Soc.* **88**, 5593.

IMMUNOGLOBULIN DISCUSSION GROUP COLLOQUIUM ON 'GENES AND ANTIBODIES'

Gene Expression of Immunoglobulins

By B. PERNIS. (*Cattedra Di Immunologia, Università Degli Studi Di Milano, Milan, Italy*)

Genetic Control of the Antibody Response

By E. S. LENNOX. (*The Salk Institute, San Diego, Calif. 92112, U.S.A.*)

It should be possible by finding suitable mutants to reveal steps in the pathway of the immune response and perhaps to decide whether the many antibody specificities arise by a somatic generation of diversity. To this end, we have screened many strains of inbred mice for their ability to make antibodies against several strains of influenza virus. Several differences have been found among the mouse strains, and genetic crosses have been made to elucidate the genetic control. In addition, there are very big differences of response among individuals of the same inbred strain.

Some of the findings are as follows:

(1) Strain DBA-2 responds to injection of virus WSE by making only macroglobulin (immunoglobulin M) antibody. In response to other virus strains it does make immunoglobulin G in addition. This defect in specific immunoglobulin G production is controlled by a single gene or close cluster of genes. On second injection strain DBA-2 does make immunoglobulin G antibodies, and hence the block in primary response is at some regulatory step.

(2) Response of F₁ hybrids between two strains that respond poorly is significantly better than either parent. This would seem to indicate multi-genic control of the immune response, a hardly surprising result.

The possible uses and limitations of these kinds of experiments will be discussed.

Genetics of Human Immunoglobulins

By J. B. NATVIG. (*Research Institute of Immunology and Rheumatology, Rikshospitalet University Hospital, Oslo, Norway*)

The four γ G and two γ A heavy chain subclasses show different antigenic and biochemical structures related both to the Fc and Fd parts of the heavy chains. In addition, studies on Gm and Am (Kunkel, Smith, Joslin, Natvig & Litwin, 1969; Vyas & Fudenberg, 1969) genetic markers show that these γ G and γ A subclasses are related to genes of at least six different cistrons. Sensitive haemagglutination inhibition systems have become available for studying antigens that are subclass specific, as well as others shared by some, but not all, subclasses. Many of these antigens, particularly those studied for γ G, are isotypes present in all normal sera. For example, 'non a' and 'non g' (Natvig, Kunkel & Joslin, 1969) are shared by three and two subclasses respectively. However, although present in all γ G2 and γ G3 proteins, 'non a' is only found in Gm(a-) γ G1 proteins. Thus within γ G1 'non a' is antithetic to Gm(a). Further, its distribution completely matches that of the previously described 'non a' peptide (Frangione, Franklin, Fudenberg & Koshland, 1966). 'Non g' is present in all γ G2 proteins, but only in Gm(g-) γ G3-globulins. Thus within γ G3 'non g' is antithetic to Gm(g).

In contrast with these isotypes, the allotypes are present only in some sera and are confined to certain proteins within one particular subclass. Heavy-chain allotypes are known for four subclasses: γ G1, γ G2, γ G3 and γ A2. Most of these markers reside in the Fc portion, but two Fd markers, Gm(f) and Gm(z), are known. The 'non a' and 'non g' markers, the Fd and Fc subclass-specific antigens and the Gm and the Am allotypes have proved useful for studying γ -globulin structure and synthesis. Their