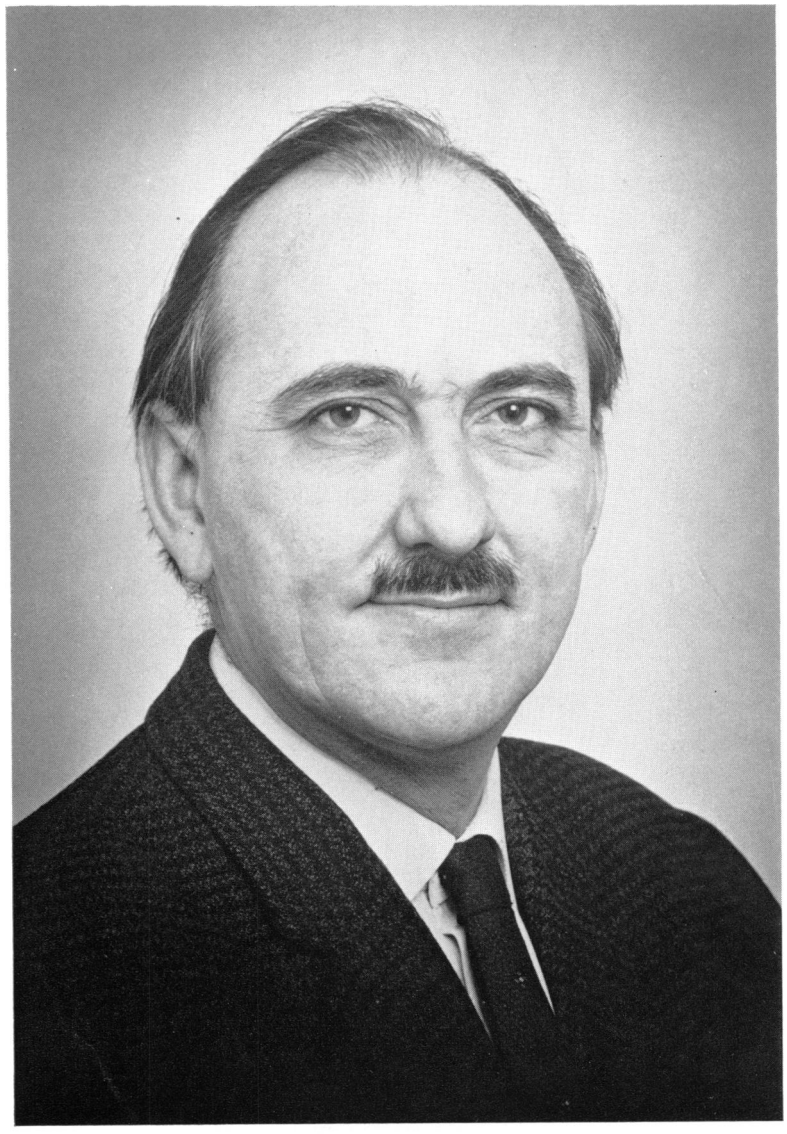


B. S. HARTLEY

Biochem. J. (1970) **119**, 805–822
Printed in Great Britain

Strategy and Tactics in Protein Chemistry

THE FIRST BDH LECTURE

By B. S. HARTLEY*

Delivered at a Meeting of the Biochemical Society on 17 *December* 1969 *at University College London, Gower Street, London W.C.*1, *U.K.*

I would like to congratulate British Drug Houses and the Biochemical Society for the idea of instituting this prize, because it may give encouragement to scientists who are thinking of developing new methods. A common theme that has run through every talk in the recent symposium on 'British Biochemistry—Past and Present' (Goodwin, 1970) is the recognition of how advances in almost every field have awaited the development of relevant methods. For example, Scheme 1 shows a way of looking at the type of studies that have contributed to our knowledge of how enzymes work. Kinetic investigations of the reaction of the pure enzyme with a range of substrates or inhibitors define the phenomena that the structure must explain. The amino acid sequence is part of the structural investigation and provides a rational framework in which residues important in the catalytic activity or specificity can be described. Chemical modification of the protein is the main way in which such active groups are discovered. The only method that can at present tell us the tertiary structure is X-ray diffraction, and for this suitable crystals must be grown. A synthesis of such information has so far provided a plausible mechanism for every enzyme studied, but further work will be needed to prove these mechanisms.

New methods are the catalysts in each of these investigations, and to be most useful should be applied at the rate-limiting steps. Kinetic information is fairly easily come by, but sequence studies have been a more serious bottleneck in the past, and I will discuss this problem shortly. Until fairly recently techniques for chemically modifying proteins were relatively few, but organic chemists have at last begun to make a contribution here with new reagents such as the affinity-labelling reagents based on substrate analogues. I will suggest some new uses of chemical modifications in this talk.

Formerly the determination of an electron-density 'map' of a crystalline protein was a daunting prospect for any crystallographer, but methods of data collection and handling have developed so fast that crystallizing the protein and preparing heavy-atom derivatives have become rate-limiting. Here I would appeal to physical chemists to help us to overcome the crystallization bottleneck by investigating the factors that allow or prevent crystallization of a protein. Organic chemists could transform our pitiful armoury of techniques for making heavy-atom derivatives. The sophisticated kinetic and physicochemical techniques that are now available for study of enzymes will surely be most usefully deployed on proteins whose tertiary structure is known.

New methods demand new ideas, and I am particularly embarrassed in receiving this prize to realize how indebted I am to others. The laboratory in which I have done most of my work has an atmosphere that is a cross between a kibbutz and a bus station, and new ideas are currency that change hands many times before they are cashed. The dansyl technique for peptide sequence analysis arose, for example, because Dr V. Massey and I were working in the same laboratory as Dr G. Weber who was already using DNS chloride to label side chains in proteins for studies of their fluorescence polarization (Weber, 1952). We decided to apply his method to chymotrypsin, but discovered that the reagent reacted specifically with a group in the active centre of the enzyme (Hartley & Massey, 1956). In abortive experiments to identify the peptide that was thus labelled we observed the stability to acid hydrolysis of the DNS-alanine that appeared as a minor side product from the *N*-terminus of the C-chain and realized that we had an

* Address: Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

$E + S \overset{m}{\leftarrow} K$: E, enzyme; S, substrates; K, kinetics; m, methods
$E + \$ + M \overset{m}{\rightarrow} S_1$: \$, money; M, manpower; S_1, primary structure
$S_1 + K + R \overset{m}{\rightarrow} G$: R, chemical modification; G, active groups
$E \rightarrow C$: C, crystals
$C + \$ + M \overset{m}{\rightarrow} F$: F, Fourier
$S_1 + F \rightarrow S_3$: S_3, tertiary structure
$S_3 + K + G \rightarrow A$: A, mechanism of action

Scheme 1. Kinetic analysis of enzyme studies.

Table 1. *Properties of DNS derivatives*

Absorption and fluorescence are measured in 0.1 M-$NaHCO_3$. $\lambda_{max.}$ and $\epsilon_{max.}$ refer to the longest-wavelength absorption maximum. $\lambda_{\frac{1}{2}}$ is the wavelength of half-maximal fluorescence on the blue side, where correction factors are minimal. Abbreviation: Mes, methionine sulphone.

	Crystals			Absorption		Fluorescence		Recovery after 18 h at 105°C with 6 M-HCl (%)
	Form	M.p. (°C)	Solvent	$\lambda_{max.}$ (nm)	$\epsilon_{max.}$	$\lambda_{\frac{1}{2}}$ (nm)	Quantum yield	
DNS-Cl	Yellow plates	—	Aq. acetone	369	3690	—	—	—
DNS-OH,$2H_2O$	White plates	316 (decomp.)	Water	314	4810	466	0.53	—
DNS-Ala	Cream prisms*	180 (decomp.)	*n*-Butyl acetate	328.5	4820	522	0.058	93
DNS-Arg	Green plates	234 (decomp.)	Aq. acetone	329	4780	—	0.082	90
DNS-Asn	Cream prisms*	187 (decomp.)	*n*-Butyl acetate	329	4820	—	0.056	—
DNS-Asp	Colourless needles*	180 (decomp.)	3-Methylbutan-1-ol	328.5	4810	524	0.055	90
$(DNS)_2$-$(Cys)_2$	White needles*	190 (decomp.)	*n*-Butyl acetate	328	—	—	—	—
DNS-Glu	Colourless plates*	186 (decomp.)	3-Methylbutan-1-ol	329	4800	521	0.060	90
DNS-Gly	Cream prisms*	184 (decomp.)	*n*-Butyl acetate	328.5	4820	522	0.049	82
DNS-Ile	White needles*	144	*n*-Butyl acetate	329	4780	518	0.071	99
DNS-Leu	White needles*	162	*n*-Butyl acetate	329	4790	518	0.076	99
DNS-Met	White needles*	175 (decomp.)	*n*-Butyl acetate	329	4790	518	0.075	—
DNS-Mes	White prisms*	—	3-Methylbutan-1-ol	329	—	521	0.069	—
DNS-Phe	White needles*	185 (decomp.)	*n*-Butyl acetate	328.5	4420	520	0.053	95
DNS-Pro	White needles*	136	*n*-Butyl acetate	329.5	4950	524	0.054	23
DNS-Ser	Yellow prisms*	205 (decomp.)	*n*-Butyl acetate	328.5	4810	523	0.053	65
DNS-Thr	Yellow prisms*	148 (decomp.)	3-Methylbutan-1-ol	329	4810	523	0.057	70
DNS-Trp	White needles*	201 (decomp.)	*n*-Butyl acetate	330	—	—	—	0
DNS-Val	White needles*	—	*n*-Butyl acetate	328.5	—	518	0.068	99
DNS-imidazole	Green plates	102	Aq. ethanol	345.5	4390	—	—	0
DNS-NH_2	Colourless needles	216 (decomp.)	Aq. ethanol	328	4690	—	—	—
DNS-$NH{\cdot}C_2H_5$	Yellow needles	134	Aq. ethanol	329	4780	—	—	—
DNS-$N(C_2H_5)_2$	Yellow needles	99	Aq. ethanol	329.5	4890	—	—	—

* Piperidinium salt.

Table 2. *Effect of solvents of a typical DNS-amino acid*

Solvent	U.v. absorption $\epsilon_{max.}$	U.v. absorption $\lambda_{max.}$ (nm)	Fluorescence Colour	Fluorescence $\lambda_{max.}$ (nm)	Fluorescence Quantum yield
0.1 M-$NaHCO_3$	4800	329	Orange–yellow	570	0.06
0.1 M-NaOH	5900	314	Orange–yellow	565	0.06
0.1 M-HCl	5000	300	Quenched	—	—
Aq. 80% acetone	4800	335	Yellow–green	520	0.33
Dry *n*-butyl acetate	?	?	Green	480	?
Paper	?	?	Yellow–green	515–525	0.5

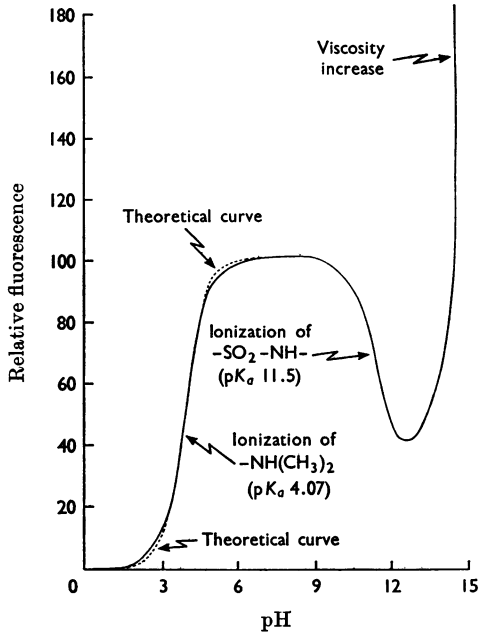


Fig. 1. Effect of pH on the fluorescence of DNS-glycine. The theoretical curve shown as a dotted line is for the ionization of a group of pK_a 4.07.

N-terminal reagent of 100 times the sensitivity of Sanger's fluorodinitrobenzene. Other problems intervened, and no-one else took up our suggestion, but in 1961 I persuaded my first research student, W. R. Gray, that this would make a good research problem. Most of our work has not been published outside Gray's (1964) Ph.D. Thesis so I will mention some of the findings here.

A good micro-method should be very simple, so we synthesized and investigated a great number of the properties of the DNS-amino acids in order to achieve this simplicity. We built a sensitive spectrophotofluorimeter that enabled us to investigate their spectral properties in solution or on paper, and Table 1 lists these. The fluorescence is very dependent on the polar nature of the solvent, as can be seen from the variation in quantum yields. Table 2 summarizes the average properties of the simple DNS-amino acids in various media, and it can be seen that paper corresponds to a medium rather like 80% acetone. Traces of moisture quench the fluorescence, so it is best to dry papers thoroughly before viewing. We had built a fluorescence scanner for paper strips with the intention of using the DNS-amino acids as a micro-method of amino acid analysis, but the above property is one that caused us to abandon this approach, though it was subsequently recommended by Boulton, Chard & Grant (1965), who used fluorescence scanning after a rather complicated automatic heating procedure. The fluorescence is also pH-dependent, as shown in Fig. 1. Ionization of the $(CH_3)_2N$– group quenches completely, but ionization of the $-SO_2-NH-$ group lowers the quantum yield somewhat. A puzzling rise at high pH was traced to the effect of the viscosity of the solvent.

To find the right reaction conditions we measured the velocity constants for the reaction of DNS chloride with water or with a large concentration excess of each amino acid, and thus obtained the results shown in Table 3. For primary amines the reaction showed a Brønsted dependence:

$$\log k = 0.52\,pK_a - 4.0$$

We were then able to predict the degree of reaction of various amino acids, as shown in Fig. 2. This shows that the concentration of DNS chloride and amino acid is vital in obtaining complete reaction rather than the mere excess of reagent over amino acid, since water and OH^- ions are competing nucleophiles. Acetone suppresses the ionization of the amino group and solubilizes the reagent, and we found that 50% acetone with bicarbonate buffer gave us optimal results (Gray & Hartley, 1963*b*). Tertiary amines catalyse the hydrolysis of DNS chloride so rapidly that the reaction may be over before proper mixing is achieved.

In our Cambridge laboratory we have largely abandoned the original electrophoretic techniques

Table 3. *Kinetic constants for reactions of DNS chloride at* 20°C *in* 0.5% *acetone*

k is the velocity constant for the free α-amino group unless otherwise indicated. Abbreviations: Mes, methionine sulphone; Cya, cysteic acid.

Reactant	pK_a	k ($M^{-1}s^{-1}$)	Reactant	pK_a	k ($M^{-1}s^{-1}$)
H_2O	—	4.1×10^{-5}	His (α)	9.18	12
OH^-	—	15	Phe	9.24	10
His (Im)	6.1	0.5	NH_3	9.26	0.5
Gly-Gly-Gly	7.91	3.0	Trp	9.39	35
Gly-Gly	8.17	3.5	Glu	9.67	5.0
Leu-Gly	8.2	2.4	Val	9.72	13
Mes	8.6	3.2	Leu	9.74	13.5
Asn	8.80	3.1	Ile	9.76	14.5
Lys (α)	8.95	3.6	Gly	9.78	13
Cya	9.0	5.7	Asp	9.82	4.4
Arg	9.04	6.4	Ala	9.87	10
Tyr (α)	9.11	~10	$N(CH_3)_3$	9.87	100
Thr	9.12	4.9	Tyr (*O*-)	10.07	280
Gln	9.13	4.8	Lys (ϵ)	10.53	42
Ser	9.15	4.6	Pro	10.60	360

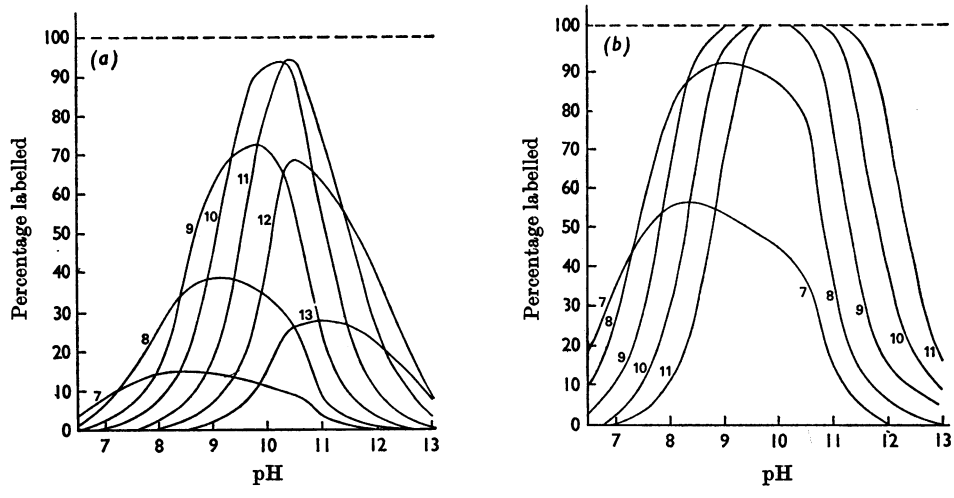


Fig. 2. Theoretical yields for the reaction of primary amino groups with excess of DNS chloride in 0.5% acetone at various pH values, calculated from the data of Table 3. (*a*) With 1 mM-DNS chloride; (*b*) with 5 mM-DNS chloride. Values on the curves show the pK_a of the amino group.

for separating DNS-amino acids in favour of thin-layer chromatography on polyamide sheets (Woods & Wang, 1967). Scheme 2 summarizes the method we now use. The sample is run alone on one side of the double-sided sheet and on the other side is mixed with a selected marker mixture. Ill-resolved pairs such as DNS-alanine and DNS-glycine can thus be unequivocally identified by seeing whether there are one or two spots in this position on the marker side. The positions of spots after solvents 1 and 2 are shown in Fig. 3(*a*). Solvent 3 (Crowshaw, Jessup & Ramwell, 1967) is then run in the same direction as solvent 2 to give the pattern shown in Fig. 3(*b*). Ambiguities between ϵ-DNS-lysine, α-DNS-histidine and DNS-arginine can be resolved by a further run in the same direction in 0.05 M-trisodium phosphate–ethanol (3:1, v/v). DNS-cysteic acid may lie below the large blue 1-dimethylaminonaphthalene-5-sulphonic acid spot, but can be resolved by using 1 M-ammonia–ethanol (1:1, v/v) in the solvent 2 direction.

The sensitivity of the method allows one to detect comfortably about 1 nmol of DNS-amino acid, so the method can be usefully used in conjunction with

(1) 1–5nmol of peptide taken in 4mm×30mm tube; evaporated
(2) 10μl of 0.2M-$NaHCO_3$ added; centrifuged; evaporated
(3) 10μl of deionized water+10μl of DNS chloride (2.5mg/ml in acetone) added to give 5mM-DNS chloride, 1mM-peptide, 50% (v/v) acetone, 'pH9.8'; mixed; sealed with Parafilm; incubated for 1h at 37°C
(4) Evaporated; 50μl of 6M-HCl added; sealed; incubated for 18h at 105°C
(5) Tubes opened; centrifuged; evaporated over NaOH
(6) 10μl of 50% (v/v) pyridine added; spotted out on both sides of thin layer in 1μl applications; marker (DNS-Gly, DNS-Glu, DNS-Ile, DNS-Phe, DNS-Pro, DNS-Ser and DNS-Arg) applied on one side
(7) Run for 50min in solvent 1 [1.5% (v/v) formic acid]; dried; run for 1h in solvent 2 [benzene–acetic acid (9:1, v/v)]; dried; examined; run for 1h in solvent 3 [ethyl acetate–methanol–acetic acid (20:1:1, by vol.)]; dried; examined.

Scheme 2. DNS *N*-terminal method.

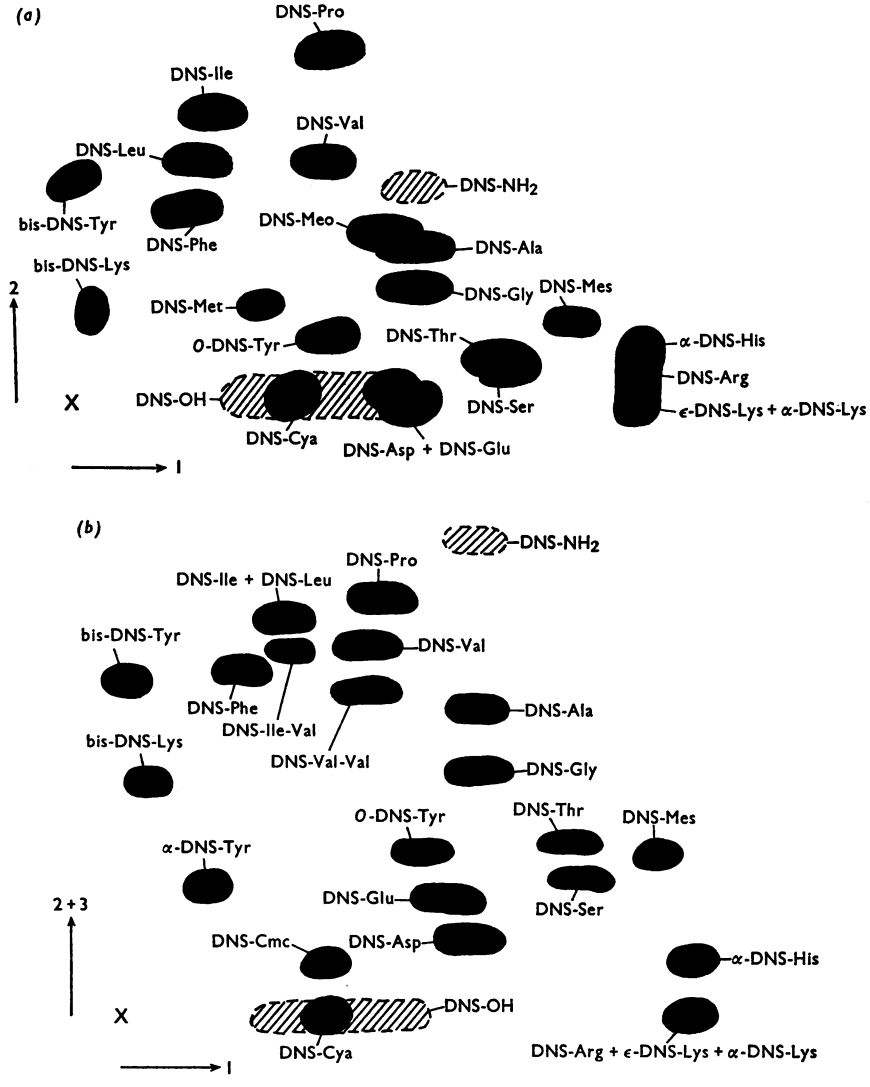


Fig. 3. Chromatography of DNS-amino acids on polyamide thin layers. Solvents: 1, 1.5% (v/v) formic acid; 2, benzene–acetic acid (9:1, v/v); 3, ethyl acetate–methanol–acetic acid (20:1:1, by vol.). Abbreviations: Cmc, *S*-carboxymethylcysteine; Cya, cysteic acid; Meo, methionine sulphoxide; Mes, methionine sulphone.

(1) 20–100nmol of peptide taken in stoppered tube; evaporated; 150μl of water added; 10μl taken for DNS *N*-terminal method
(2) 150μl of 5% (w/v) phenyl isothiocyanate in pyridine added; flushed with N_2; incubated for 1½h at 45°C
(3) evaporated over $H_2SO_4+P_2O_5+NaOH$ at 1mmHg at 60°C for about 30min
(4) 200μl of trifluoroacetic acid added; flushed with N_2; incubated for 30min at 45°C; evaporated over NaOH (10min)
(5) 200μl of water added; extracted twice with 1.5ml of *n*-butyl acetate; evaporated
(6) 150μl of water added; 10μl taken for DNS *N*-terminal method

Scheme 3. DNS–Edman method.

the stepwise degradation of peptides by reaction with phenyl isothiocyanate (Gray & Hartley, 1963*a*). One advantage is that peptides that become 'blocked' by oxidative side reactions in the phenyl isothiocyanate procedure fail to react with DNS chloride, so that a clean positive *N*-terminus is detected at each stage, albeit in decreased yield. The method also allows one to avoid another disadvantage of the classical Edman (1950*a*,*b*) procedure, namely the risk of extracting small hydrophobic phenylthiocarbamoyl-peptides in the extraction step used to remove excess of phenyl isothiocyanate. Instead one can evaporate off most of the excess of phenyl isothiocyanate and cyclize at once with trifluoroacetic acid. All organic products are readily removed from the resulting free peptide by extraction with *n*-butyl acetate (Gray, 1967). Scheme 3 summarizes the method.

Finally may I mention a development of the DNS–Edman technique described by Bruton & Hartley (1970). Following a suggestion of Dr V. Neuhoff we scaled down the DNS–Edman procedure so as to allow sequence analysis of 1nmol of peptide. The procedure is exactly as shown in Schemes 2 and 3 except that all volumes are reduced to one-tenth. Pipettes are made by drawing out capillary tubing and calibrated to deliver 0,1μl. Reaction tubes 2cm × 1mm (internal diam.) are made by cutting off the ends of sealed melting point tubes. The Edman reaction is carried out in 4mm × 30mm tubes with conical bottoms. The thin-layer sheets are 5cm × 5cm and could be smaller. The resolution is as good if not better than that on the conventional 20cm × 20cm sheets, and 0.01nmol of DNS-amino acid can easily be detected. We have used this method to determine the sequences at the *N*-terminus and around a lysine residue in the catalytic site of methionyl-tRNA synthetase from *Escherichia coli*, using only 1–2mg of enzyme in each case (Bruton & Hartley, 1970).

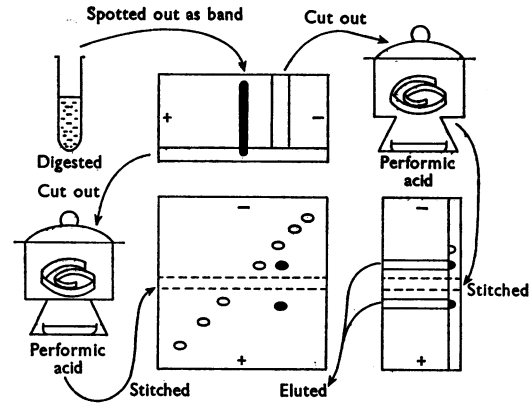


Fig. 4. Disulphide-bridge diagonal electrophoretic technique. For details see the text.

Diagonal electrophoretic techniques

Disulphide diagonal method. Improvements in amino acid analysis and sequence analysis have made peptide purification the rate-limiting step in determining the primary structure of proteins. Having determined the amino acid sequence of bovine chymotrypsinogen A in the reduced carboxymethylated form (Hartley, 1964), I was faced with the prospect of determining the disulphide bridges. This necessarily involved fractionating the complex peptic digest of the protein in order to isolate disulphide-bridged peptides.

In the precedent studies with insulin (Ryle, Sanger, Smith & Kitai, 1955) and ribonuclease (Ryle & Anfinsen, 1957; Spackman, Stein & Moore, 1960) this had proved a considerable task, so the prospect of this hard work stimulated us to develop a diagonal electrophoretic procedure for the selective purification of such disulphide-bridged peptides (Brown & Hartley, 1963, 1966).

The principle of this technique is summarized in Fig. 4. The disulphide-bridged protein is digested with a suitable proteinase. Peptic digests are to be preferred if the protein is sensitive to disulphide exchange, which may occur slowly at alkaline pH (Spackman *et al.* 1960), particularly if the protein has free thiol groups (Ryle & Sanger, 1955). The digest is applied as a band on paper and subjected to high-voltage paper electrophoresis in solvent-cooled tanks—pH 6.5 is often a suitable pH because of the information about the charge structure of the peptides that can be derived (Offord, 1966). A strip from this ionogram is subjected to the vapour of performic acid to oxidize the cystine peptides to pairs of cysteic acid peptides. The strip is then dried in a current of air to remove excess of performic acid, stitched with a sewing machine to

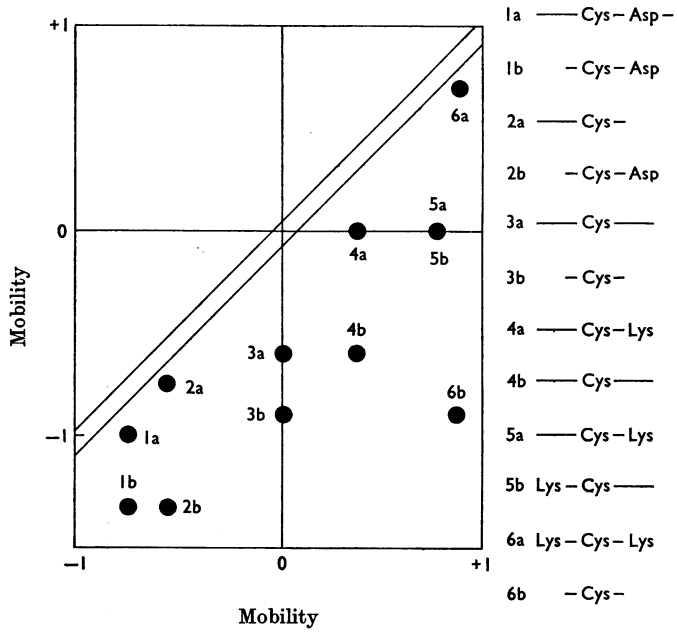


Fig. 5. Hypothetical 'map' of some cystine peptides treated by the method of Fig. 4. Mobilities of peptides are shown relative to 1-dimethylaminonaphthalene-5-sulphonic acid = −1.0 (mobility of 1-dimethylaminonaphthalene-5-sulphonic acid relative to aspartic acid = 0.65).

another sheet of paper and subjected to high-voltage electrophoresis under the original conditions but at right-angles to the original direction.

Each peptide should then lie on a 45° diagonal, but every cystine peptide has been converted into a pair of cysteic acid peptides, each with an extra negative charge, and these will lie off the diagonal but vertically below each other.

The advantage of this diagonal electrophoretic 'map' is that it allows one to recognize at once which pairs of cysteic acid peptides were originally disulphide-bridged, even if they are incompletely resolved. The 'map' can then be used as a guide to the selective purification of each cysteic acid peptide, as shown in Fig. 4. A suitable strip containing a particular cystine peptide (or peptides) is cut out, oxidized as before and re-run in the same direction under the same conditions of high-voltage electrophoresis. Most peptides will lie in the 'diagonal position', but cysteic acid peptides will have a changed electrophoretic mobility, and the relevant pair of peptides can be recognized from the diagonal 'map'. The method frequently yields pure peptides in this single operation, but further purification of the band by electrophoresis at another pH, or by chromatography, will invariably resolve any cysteic acid peptides that are cross-contaminated.

A careful examination of the diagonal electrophoretic 'map' can tell one a great deal about the charge structure of the cystine peptides. Some patterns that would be produced by various types of cystine peptides are illustrated in Fig. 5. If the electrophoretic mobilities before and after oxidation are carefully measured, it is possible to assign, within rough limits, a unique value to the molecular weight and net charge of each cysteine acid peptide, by using the mobility plots of Offord (1966). Such data are invaluable when interpreting the results of amino acid analysis or sequence studies of the peptides.

Some problems occasionally remain. For example, if two disulphide bridges occur in the same peptide, three cysteic acid peptides will result, and it is necessary to find a method of splitting between the two bridges to determine the bridge pattern. A peptide with an internal disulphide bridge will give a single peptide with two cysteic acid residues. These are sometimes difficult to spot unequivocally on the 'map', because one partner may be 'buried' in the diagonal position as in peptides 2a or 6a of Fig. 5. It is always wise to elute the band in the diagonal position and analyse for the presence of cysteic acid peptides. If present, these can generally be purified by electrophoresis at pH 2. Peptides with *N*-terminal methionine emerge a short way below

the diagonal at pH 6.5, because of the low pK_a of *N*-terminal methionine sulphone, and similar reasons cause *C*-terminal methionine peptides to lie off diagonals at pH 3.5. Peptides with *N*-terminal tryptophan give products that lie below the diagonal, perhaps because of destruction of *N*-terminal amino groups or introduction of carboxyl groups after oxidation. Such complexities are discussed by Brown & Hartley (1966).

The method proved to be a very rapid way of

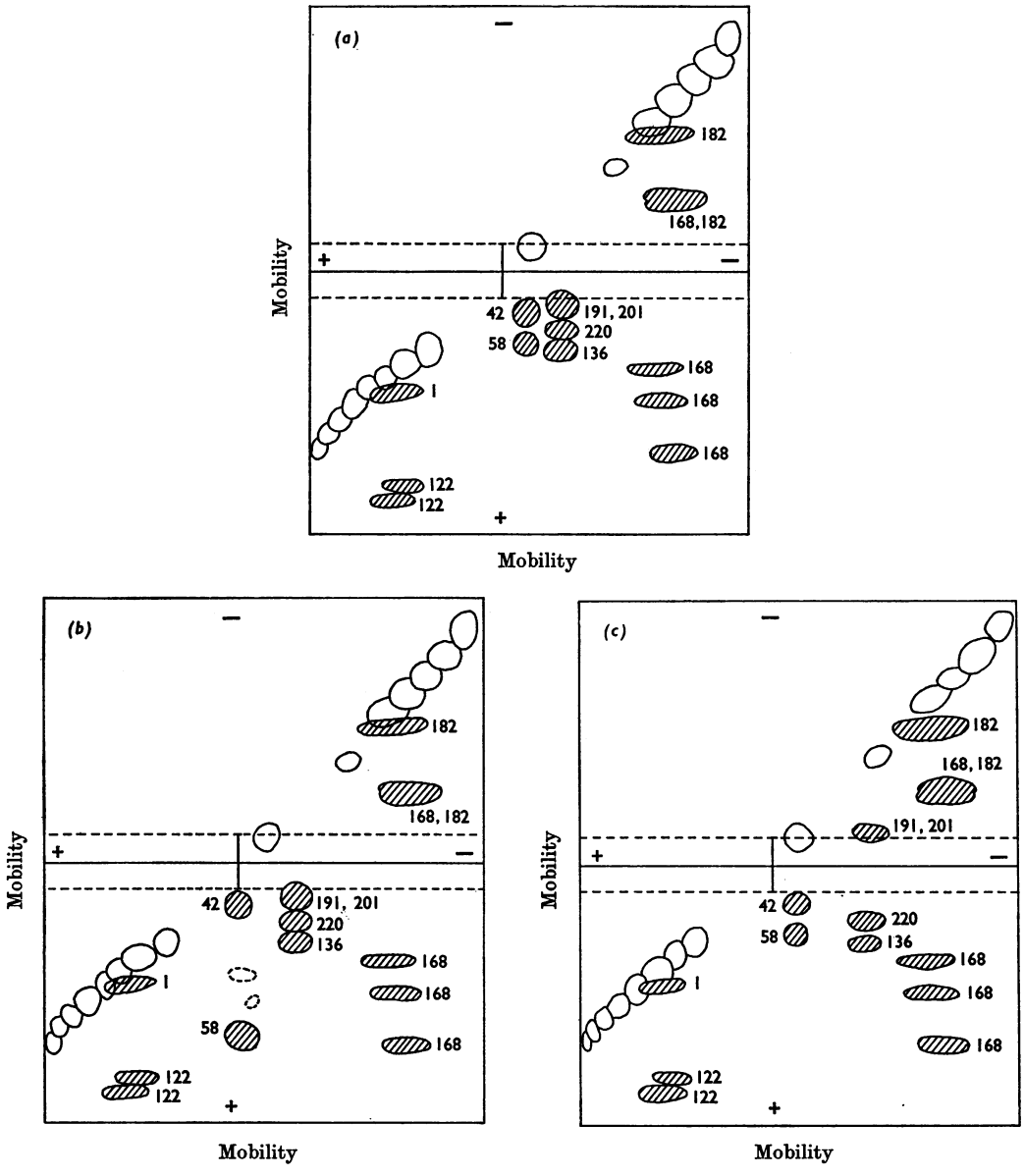


Fig. 6. Disulphide diagonal electrophoretic 'maps' of peptic digests of (*a*) α-chymotrypsin, (*b*) chymotrypsin treated with 1-chloro-4-phenyl-3-toluene-*p*-sulphonamidobutan-2-one ('*N*-tosyl-L-phenylalanyl chloromethyl ketone') and (*c*) chymotrypsin treated with 1,2-epoxy-3-phenoxypropane. Major ninhydrin spots lying off the diagonal are hatched. Numbers indicate the positions of cystine residues in the sequence of chymotrypsinogen. Electrophoresis at pH 6.5 was carried out horizontally before oxidation and vertically after oxidation.

determining disulphide-bridge patterns in proteins of known amino acid sequence, where amino acid analysis of the cysteic acid peptides is generally sufficient to describe their position in the chain. The experimental work involved in determining the bridges of bovine chymotrypsinogen A took only 6 weeks, and the bridges of egg-white lysozyme (Brown, 1964) and bovine trypsinogen (Kauffman, 1965) were also quickly elucidated. It was therefore clear that residues around disulphide bridges were no longer the most difficult parts of a protein for sequence determination and that the diagonal electrophoretic 'map' could be used as a novel 'fingerprint' of the primary structure.

Disulphide bridges of serine proteinases. The two histidine residues of bovine chymotrypsinogen A occur in the same cystine peptide in a peptic digest. It was therefore easy to determine (Smillie & Hartley, 1964) which of these reacted with 1-chloro-4-phenyl-3-toluene-*p*-sulphonamidobutan-2-one ('tosyl-L-phenylalanine chloromethyl ketone'), the substrate analogue reagent that Schoellmann & Shaw (1963) had shown to inhibit chymotrypsin by stoicheiometric reactions with a unique histidine residue. Fig. 6(*b*) shows that a diagonal electrophoretic 'map' of a peptic digest of chymotrypsin that had been inhibited with this reagent indicated that histidine-57 had selectively reacted. A similar reagent, 1,2-epoxy-3-phenoxypropane was found to our surprise to react with methionine-192 of chymotrypsin, and not histidine-57 (Brown & Hartley, 1964). The diagonal 'map' (Fig. 6*c*) shows the change in position of one peptide due to the extra positive charge of the sulphonium salt of methionine-192, and selective purification of this cysteic acid peptide allowed us to prove the structure by sequence study. Similar reagents, with chromophoretic groups, 1,2-epoxy-3-(*o*-nitrophenoxy)-propane and 1,2-epoxy-3-(α-naphthoxy)propane, were found to react with methionine-192 in a similar way (Brown & Hartley, 1964).

We naturally became interested to know whether all serine proteinases would have a similar disulphide bridged sequence homologous with the 'histidine loop' of α-chymotrypsinogen. Selective purification of disulphide-bridged peptides that also stained with the Pauly reagent for histidine offered an easy access to this problem (Smillie & Hartley, 1966). Homologous sequences were found in bovine chymotrypsinogen A, chymotrypsinogen B and trypsinogen and porcine elastase.

This success prompted us to ask whether the whole disulphide-bridge structure was homologous in the pancreatic serine proteinases (Hartley, Brown, Kauffman & Smillie, 1965). We therefore isolated by the diagonal method and determined the sequences of all the disulphide-bridged peptides from bovine chymotrypsinogen B (Smillie & Hartley, 1967) and trypsinogen (Kauffman, 1965) and porcine elastase (Brown, Kauffman & Hartley, 1967).

In each protein the bridges were homologous. With chymotrypsinogen B the five bridges gave sequences accounting for 72 of the 245 residues, and only eight differences were found. The complete sequence (Smillie, Furka, Nagabhushan, Stevenson & Parkes, 1968) shows 27 differences. The four bridges of porcine elastase were homologous with the four common to the bovine chymotrypsinogens and trypsinogen. Of the 82 residues in these peptides, 43 were identical with those of either trypsin or chymotrypsin. The complete sequence (Shotton & Hartley, 1970) shows 134 identities in the 240 residues of elastase. The latest picture of these disulphide-bridge homologies (Fig. 7) include the preliminary sequence data for bovine thrombin (S. Magnusson & B. S. Hartley, unpublished work), another mammalian serine proteinase, secreted from liver into serum as a zymogen, prothrombin. The activation of prothrombin to thrombin is the control step of blood-clotting.

All these mammalian proteinases have about 50% identity of amino acid sequence, and three disulphide bridges are common to each: A, the histidine loop (containing histidine-57), B, the methionine loop, and C, the serine loop (containing serine-195). Bridge D (136–201) is absent from thrombin, but thrombin, like chymotrypsin, has a bridge (E) at position 121 that connects the A-chain to the B-chain. Trypsin has two extra bridges (F and G) unique to itself.

We now know that these sequence homologies are matched by homologies of tertiary structure, since the folding of the chains of chymotrypsin and elastase are identical (Shotton & Watson, 1970), and convincing models of trypsin can be built to the same tertiary structure that explain its catalytic specificity (Hartley, 1970). I believe that we shall come across many more such families of enzymes with different catalytic activities based on common tertiary structures, and I think that the disulphide diagonal technique may be a rapid way of screening tell-tale homologies in such a family.

Thiol sequences in myosin. Conventional techniques of peptide purification and sequence analysis have progressed to a stage where proteins of molecular weight up to 40000 should present no insuperable problem, but the increasing complexity of digests of larger proteins stretches our fractionation techniques to their limits. We therefore sought to apply diagonal peptide purification methods to a protein that was well outside these limits, namely rabbit muscle myosin (Weeds & Hartley, 1967, 1968).

The size and dimensions of this protein are such that it stretches also the limits of physicochemical

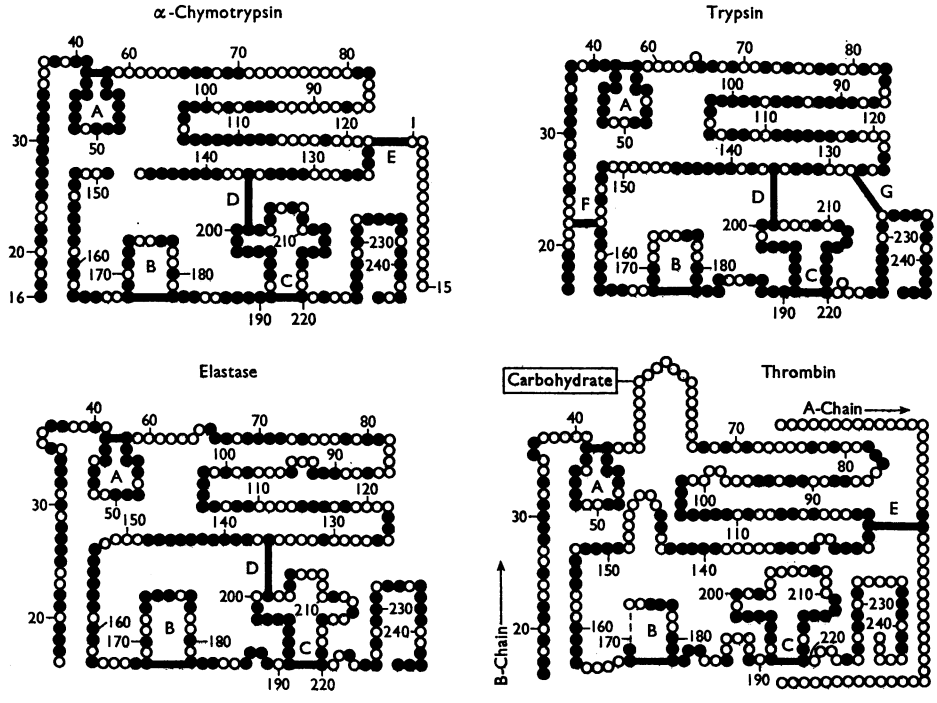


Fig. 7. Disulphide bridges in serine proteinases. Residues identical in any pair are shown as solid circles. Numbering is that of chymotrypsinogen. Insertions in one sequence relative to another are shown as kinks or small loops. The thrombin pattern is still tentative, being based on homology.

techniques for determination of molecular weight, and there had been considerable controversy about the subunit structure. Table 4 shows the amino acid composition of myosin corresponding to a molecular weight of 500 000 and a subunit structure of two or three identical chains. One sees that conventional tryptic peptide 'maps' would have to resolve mixtures of at least 230 or 340 peptides according to the subunit size. Looking at the cystine analysis, however, one sees that a structure with three identical subunits should contain 14 unique thiol sequences, whereas two subunits should give rise to 21 thiol sequences. If these peptides could be selectively purified the sequence problem would not be severe.

Unfortunately the disulphide diagonal method will not work with thiol peptides since these tend to oxidize randomly during digestion or electrophoresis. We therefore converted each of these thiol groups in myosin into a disulphide by exchanging the protein with a large excess of radioactive cystine. The radioactive peptides in a peptic digest could be fairly easily separated from each other (but not of course from the hundreds of non-radioactive peptides) by ion-exchange and electrophoresis. 'Oblique

Table 4. *Amino acid composition of rabbit myosin*

The numbers show residues/500 000 daltons assuming one, two or three identical chains.

Amino acid	Amino acid composition (residues/500 000 daltons)		
	One chain	Two chains	Three chains
Lys	445	222	148
His	79	40	26
Arg	208	104	69
Asp	437	218	146
Thr	229	115	76
Ser	228	114	76
Glu	863	431	288
Pro	113	57	38
Gly	204	102	68
Ala	389	194	130
Cys	42	21	14
Val	229	114	76
Met	123	62	41
Ile	216	108	72
Leu	416	208	139
Tyr	92	46	31
Phe	145	73	48
Trp	24	12	8
Total	4482	2241	1494

thinking', however, tells us that performic acid oxidation of each of these must automatically purify the cysteic acid peptide plus free cysteic acid from the 'diagonal contaminants'. Not all of the cystine peptides in the peptic digest of [^{35}S]cystine-exchanged myosin were purified in good amounts, because random splitting by pepsin lowers the yields of some, but Weeds & Hartley (1968) were able to determine 12 complete and four partial unique thiol sequences. Kimura & Kielley (1966) had determined 15 partial sequences around thiol groups in [^{14}C]carboxymethyl-myosin, and though nine of their partial sequences may have been the same as determined by Weeds & Hartley (1968) six were definitely different. Thus 22 unique thiol sequences had been described, and it was highly probable that myosin was a dimer.

Weeds (1967) subsequently improved this technique by exchanging light-meromyosin fraction 1 with the more soluble [^{35}S]cystamine in place of cystine. Light-meromyosin fraction 1 represents the 'tail' of myosin, which is split off by a very brief trypsin digestion. It has a molecular weight of 135000–146000 and amino acid analysis showed 6.7 thiol groups/140000 daltons. Two identical chains should give three unique thiol sequences, and Weeds (1967) isolated these in approximately equal amounts by this diagonal technique and determined their sequences.

Disulphide bridges in immunoglobulins. An alternative diagonal technique for thiol groups has been employed by Pink & Milstein (1967) in their studies of disulphide bridges of immunoglobulins. Mild thiol reduction selectively breaks interchain bridges, and these can be labelled with iodo[^{14}C]-acetic acid. Complete reduction and carboxymethylation with non-radioactive iodoacetic acid follows, and the protein is digested to give *S*-carboxymethyl-peptides. A performic acid diagonal electrophoretic 'map' at pH 3.5 causes these peptides to emerge from the diagonal because the pK_a of the *S*-carboxymethyl group is lowered when it is oxidized to the sulphone form. Intrachain sequences are recognized by the radioactive label. A combination of these methods and 'classical' disulphide diagonals have contributed greatly to our knowledge of the disulphide bridges of immunoglobulins (Milstein & Pink, 1970).

Selective purification of methionine peptides. The above examples are clearly special cases of a general principle of chemical purification, which might be nicknamed 'oblique thinking'. This tells us that a compound can be selectively purified whenever it can be selectively modified so as to change its electrophoretic or chromatographic properties.

Methionine residues are rare in proteins, so sequences containing methionine may be an easy way to detect homologies in families of proteins. Furthermore, selective cleavage at methionine residues by reaction with cyanogen bromide (Gross & Witkop, 1962) is a valuable tool in sequence studies. Peptides containing internal methionine residues will overlap such cyanogen bromide fragments and allow them to be placed in sequential order. In this way the methionine residues could be approximately located in the peptide chain merely by determining the amino acid composition of the cyanogen bromide fragments and their *N*-terminal and *C*-terminal sequences. We therefore developed a diagonal electrophoretic method for methionine peptides (Tang & Hartley, 1967).

A suitable diagonal reaction was readily available. Iodoacetamide selectively alkylates methionine residues at acid pH (free thiol groups can be previously alkylated at alkaline pH) to give the positively charged sulphonium salt (Lawson, Gross, Foltz & Witkop, 1961; Stark & Stein, 1964). We were able to show that this reaction can be carried out on paper on digests that have been separated by high-voltage electrophoresis as summarized in Fig. 8. The carbamoylmethylmethionine peptides thus selectively purified have another pleasing property. Heating these sulphonium salts at pH 7 for 2h at 100°C cleaves them quantitatively to yield an *N*-terminal portion with *C*-terminal homoserine and the *C*-terminal remainder. The former will, of course, be the *C*-terminal sequence of the appropriate cyanogen bromide fragment and the latter has the *N*-terminal sequence of another, so the overlapping problem is made even easier. We showed that this method readily purifies the two methionine peptides in a tryptic digest of *S*-aminoethyl-chymotrypsinogen.

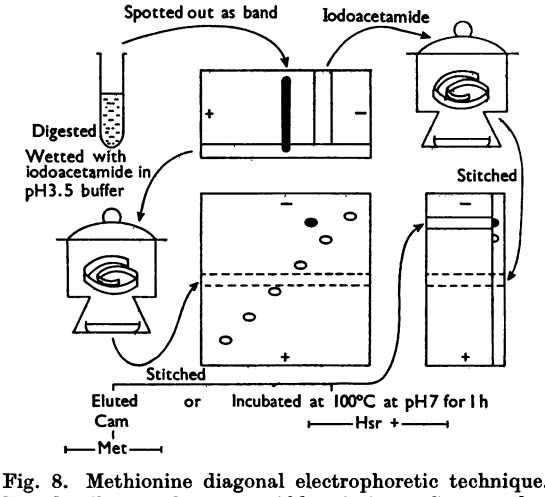


Fig. 8. Methionine diagonal electrophoretic technique. For details see the text. Abbreviations: Cam, carbamoylmethyl; Hsr, homoserine.

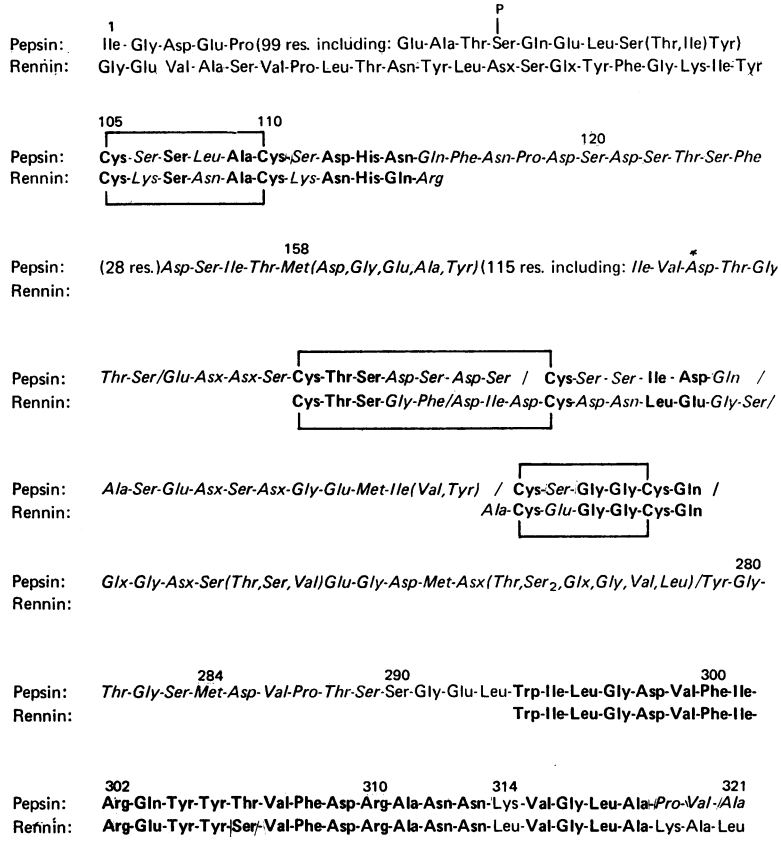


Fig. 9. Partial sequences of pig pepsin and calf rennin. Identical or chemically similar residues at homologous positions are in bold type. Peptides in *italics* were purified by diagonal electrophoretic techniques. Also shown are the sequences containing the serine phosphate residue of pepsin (Stepanov, 1968) and the active-site aspartic acid (Bayliss & Knowles, 1969) indicated Asp.

Homologies between pepsin and rennin. We applied the methionine diagonal technique to porcine pepsin, which is a single chain of 321 residues, containing only one lysine, two arginine and four methionine residues (Rajagopalan, Stein & Moore, 1966). These basic residues are clustered at the *C*-terminus, so a single tryptic peptide of 302 residues accounts for most of the molecule (Keil, 1970). Cyanogen bromide cleavage seemed an essential tool for this molecule, so we thought the methionine sequences would be particularly useful.

It proved rather difficult to get good yields of peptides from each methionine sequence, but peptic digests or papain digests eventually yielded representatives of each, which were purified by the methionine diagonal technique (Tang & Hartley, 1970). The partial sequence of these is shown in Fig. 9. Calf rennin is another gastric proteinase, and we wished to see whether it would be homologous with porcine pepsin. The classical disulphide diagonal technique showed that each protein contained two small disulphide loops, which were almost identical whereas the third bridge of pepsin or rennin showed very little homology (Fig. 9).

A diagonal electrophoretic technique for selective purification of lysine or *N*-terminal residues was developed by Perham & Jones (1967). In this method the whole polypeptide chain is trifluoroacetylated on each amino group before digestion. After electrophoresis a strip of paper is exposed to ammonia vapour to unblock the amino groups, which thereby gain a positive charge and emerge from the diagonal when the paper is run at right-angles. Perham & Jones (1967) showed that this technique gave only the expected peptides when applied to a subtilisin digest of insulin, and were readily able to isolate a chymotryptic peptide containing the only lysine residue of pepsin. Their

results, together with those of other workers (Kostka, Morávek & Šorm, 1970), established the *C*-terminus of pepsin.

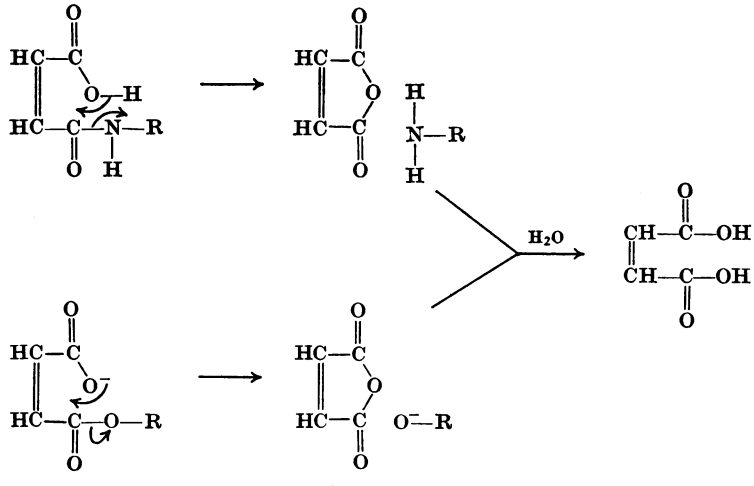
The *N*- and the *C*-terminal sequences of calf rennin (Foltmann & Hartley, 1967; Foltmann, 1970) are also shown in Fig. 9. We can see that pepsin and rennin show 20 identities in the *C*-terminal 27 residues in addition to the homologies of disulphide bridges mentioned above. Fig. 9 also shows that the cystine sequences and methionine sequences determined by the diagonal techniques have been useful in orienting fragments obtained from cyanogen bromide cleavage of pepsin or from tryptic digests of reduced aminoethylated pepsin (Keil, 1970).

Reaction of proteins with maleic anhydride. Insoluble or strongly aggregating peptides are a problem for the protein chemist. They fail to be digested completely with proteinases, they form precipitates or aggregate on gel filtration, they bind almost irreversibly to ion-exchange columns and they remain at the origin or streak badly in electrophoresis or chromatography. Solvents such as 8M-urea or 2M-guanidinium chloride are very useful, but the necessity to desalt the peptides after fractionation brings extra problems. An alternative technique is to succinylate the amino groups of the protein or peptide so that electrostatic repulsion at alkaline pH promotes disaggregation and solubilization (Habeeb, Cassidy & Singer, 1958). However, the amino groups of the lysine residues are then permanently substituted, so that digestion with trypsin is no longer possible. We decided to investigate the reaction of amino groups with maleic anhydride, since the maleyl groups should be easier to remove (Butler, Harris, Hartley & Leberman, 1969). Bender, Chow & Chloupek (1958) showed that hydrolysis of the amide group of phthalamic acid was unusually rapid at acid pH, being catalysed by the neighbouring carboxylic acid, whereas *O*-maleyltyrosine should have a half-life of only 1s at alkaline pH owing to intramolecular catalysis by the carboxylate ion (Bruice & Pandit, 1960). We found that monoamides of maleic acid were similarly labile at acid pH, presumably by the reactions shown in Scheme 4, the half-life of ϵ-maleyl-lysine being 10h at 37°C at pH3.5, but more than 10^4h above pH6.

All 15 amino groups of chymotrypsinogen appeared to react readily with maleic anhydride with optimum pH9, and the degree of substitution was unaffected by the presence of denaturing agents. Unblocking at pH3.5 appeared to occur at about the same rate as that of ϵ-maleyl-lysine. With the 18-residue peptide β-melanophore-stimulating hormone both lysine residues and the *N*-terminus reacted completely at pH9 and no substitution of tyrosine or tryptophan was detected.

A chymotryptic digest of the maleyl-(β-melanophore-stimulating hormone) was used to test a maleyl diagonal technique. After pH6.5 electrophoresis a strip was exposed to the vapour of 1% (v/v) pyridine–5% (v/v) acetic acid for 6h at 60°C and then run at right-angles in the usual way. Almost all of the expected peptides moved decisively away from the diagonal position owing to the charge change from −1 to +1 at each amino group.

It was clear that maleylation of proteins or



Scheme 4. Suggested mechanisms for the hydrolysis of maleyl derivatives catalysed by the carboxyl group.

peptides had many advantages for the protein chemist. The maleylated proteins were almost invariably soluble and disaggregated. The reaction frequently dissociates subunits, and has the advantage over denaturing solvents such as strong urea or guanidine solutions that molecular weights can be determined in the ultracentrifuge or on calibrated gel-filtration columns in the same buffers as the native protein, with consequent increase in accuracy. In this way the subunit structure and molecular weight of methionyl-tRNA synthetase from *E. coli* was determined with only 0.2mg of protein (Bruton & Hartley, 1968). The absorption of maleyl derivatives in the 240–270nm range is useful in measuring quantitatively the degree of reaction or in detecting peptides after column chromatography. Maleyl-peptides remain soluble during trypsin digestion, and only arginyl bonds will be attacked. Subsequent unblocking allows trypsin digestion at lysyl bonds. With maleyl-(yeast alcohol dehydrogenase) it was found that half the arginyl bonds split rapidly and half much more slowly presumably owing to local electrostatic effects of the maleyl groups.

Unblocking of maleyl-peptides requires rather long incubation at acid pH, and though no deamidation of asparagine was detected under these conditions some labile amides such as -Asn-Gly- sequences may be lost. Dixon & Perham (1968) showed that citraconic anhydride (2-methylmaleic anhydride) reacts similarly with amino groups, but that the half-life of the product at 20°C was about 1–2h at pH 2–3. This greater lability would generally make citraconic anhydride the reagent of choice were it not that with amino groups it yields geometrical isomers of slightly different pK_a that may complicate ion-exchange or electrophoretic separations. Disadvantages of both maleic anhydride and citraconic anhydride are that both they and the resulting acids can alkylate thiol groups (Calam & Waley, 1963; Smyth, Blumenfeld & Konigsberg, 1964), which should be protected by disulphide exchange if complete reversibility is required.

O-Maleyltyrosine, if formed, should decompose almost immediately at alkaline pH, but maleyl esters of serine and threonine would be more stable since here intramolecular catalysis by the carboxylate ion is probably much less important (Bruice & Benkovic, 1966), contrary to the results of Bender, Chloupek & Neveu (1958) quoted by Butler *et al.* (1969). Such serine or threonine esters should be formed only in minor quantities in normal conditions of maleylation, but reaction with hydroxylamine at alkaline pH should hydrolyse these without removing maleamido groups (Freedman, Grossberg & Pressman, 1968).

New sequence strategies. These diagonal electrophoretic techniques, in conjunction with methods for selective cleavage of polypeptide chains, provide the protein chemist with new strategies for determining the complete sequence, which are illustrated diagrammatically in Fig. 10.

After the blocking of amino groups by maleylation, citraconylation or trifluoroacetylation, the arginine residues can be selectively attacked by trypsin. Chymotryptic digestion of the substituted protein produces peptides with internal arginine residues that can be selectively purified, if necessary, by an arginine diagonal method. After pH 6.5 electrophoresis, a strip is sprayed with trypsin in ammonium bicarbonate buffer and then run again at right-angles. Peptides with internal arginine residues split to give an *N*-terminal and a *C*-terminal fragment that lie parallel to each other away from the diagonal position by virtue of their change in size. The sequence of these peptides overlaps the *N*-terminal and *C*-terminal sequences of the large tryptic fragments and allows them to be arranged in the correct order.

After purification each large tryptic fragment can be unblocked and redigested with trypsin at lysine residues. In this way the tryptic peptides are allocated to a particular section of the sequence. The order of tryptic peptides within each large arginine fragment is readily established by isolating peptides with internal lysine residues from a chymotryptic or peptic digest by the maleyl or the trifluoroacetyl-lysine diagonal method.

Thiol groups should be modified before digestion of the protein. If they are made to react with ethyleneimine a new site of tryptic cleavage can be generated (Raftery & Cole, 1963), and the amino group of aminoethylcysteine is a new target for selective purification of chymotryptic peptides by the maleyl, citraconyl or trifluoroacetyl diagonal techniques (Perham, 1967). Alternatively disulphide exchange with cystine or cystamine enables one to purify cysteic acid peptides by the cystine diagonal method.

As discussed above, fragments produced by cyanogen bromide cleavage of a polypeptide can be overlapped by peptides purified by the methionine diagonal technique. When these carbamoylmethylmethionine peptides are heated they yield fragments that are *C*-terminal in one cyanogen bromide piece and *N*-terminal in another.

Selective application of these new tools can be of great assistance in sequence studies. When the complete sequence is the objective, I still think that isolation of peptides from a tryptic digest of the whole protein is a sensible initial approach, but diagonal techniques are especially useful where sequences around particular residues are required. Diagonal techniques are also very useful in purifying peptides that overlap large fragments obtained by selective cleavage of the chain.

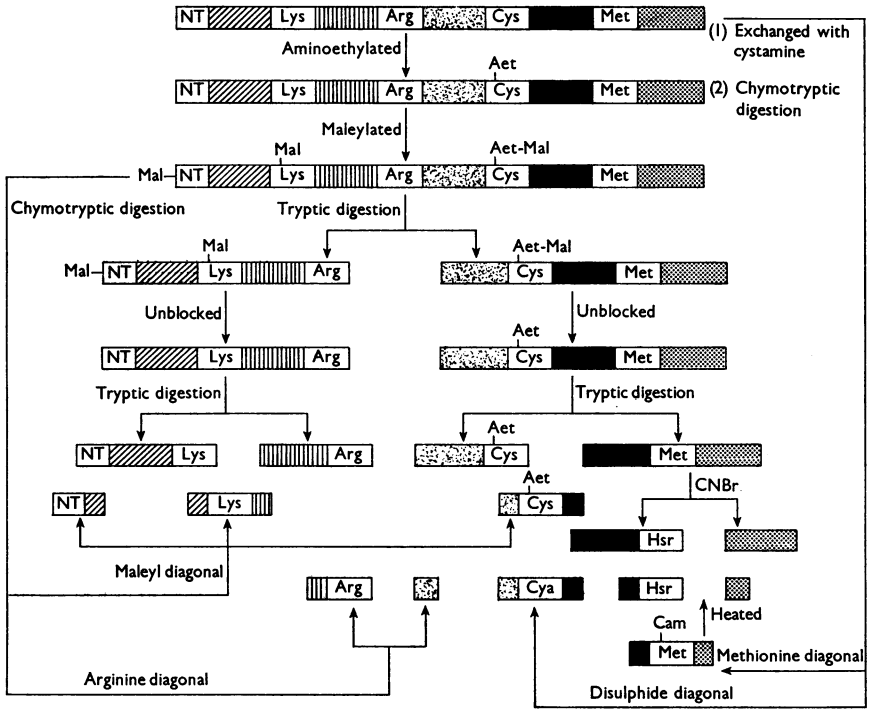


Fig. 10. Hypothetical sequence strategy illustrating selective cleavage and diagonal purification. Methods are discussed in the text. Abbreviations: NT, *N*-terminus; Aet, aminoethyl; Cam, carbamoylmethyl; Hsr, homoserine; Cya, cysteic acid; Mal, maleyl.

Determination of pK_a *and reactivity of amino groups*

The protein biochemist's job is by no means finished by the time the tertiary structure of a protein has been determined. The electron-density 'map' gives, at best, only a picture of the most probable conformation in the crystal. One needs to know how far this structure represents that of the major species in solution, and what equilibria are possible between other conformational states. One also needs to know the precise charge distribution, which ideally means the pK_a of each ionizing group, and the extent to which any single group may be specially reactive. Investigations of the chemical reactivity of individual groups with different reagents can help to answer these questions, but biochemists are often rather naive when discussing reactivity of functional groups. A particular amino group, for example, may be specifically acylated because it has a low pK_a or a high nucleophilicity, or because it is uniquely oriented to the particular reagent. A minimum description of its reactivity therefore demands a 'Brønsted plot' of log$k$ (the velocity constant for the uncharged amino group) versus pK_a compared with that for a series of simple primary amines with the same reagent.

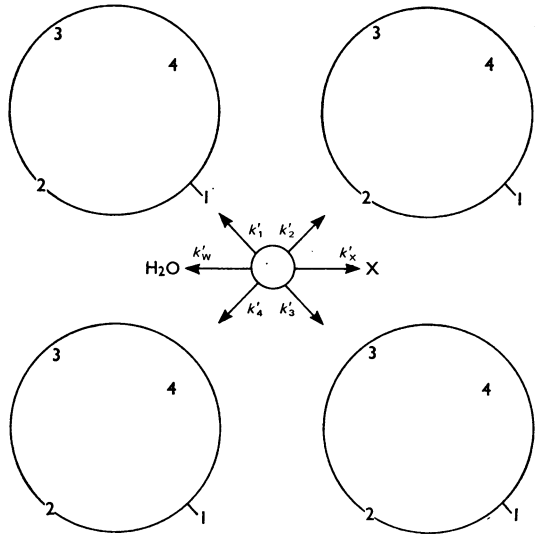
This requirement has invariably been dodged because no method was available, but a 'competitive labelling' technique has been devised by Kaplan, Hartley & Stevenson (1969) that makes it possible.

The principle of this method is shown diagrammatically in Fig. 11. An electrophilic reagent, R, such as acetic anhydride is made to react to completion with a large excess of protein in the presence of an excess of added 'calibrating nucleophile' X, such as the amino group of phenylalanine. Fig. 11 shows four amino groups that will react at any pH according to their pK_a and to the reactivity of the uncharged species, represented in this figure by varying degrees of 'buriedness'. Since all competing nucleophiles remain in excess throughout the reaction, the degree of acetylation of each when reaction is complete will be in proportion to their apparent velocity constants, i.e.:

$$\frac{[\text{Ac-1}]}{k_1'[\text{E}]} = \frac{[\text{Ac-2}]}{k_2'[\text{E}]} = \text{etc.} = \frac{[\text{Ac-X}]}{k_X'[\text{X}]}$$

Therefore:

$$\frac{[\text{Ac-1}]}{[\text{E}]} \Big/ \frac{[\text{Ac-X}]}{[\text{X}]} = \frac{k_1'}{k_X'} = \frac{k_1'(1+[\text{H}^+]/K_X)}{k_X'(1+[\text{H}^+]/K_1)}$$



$$\frac{k'_1}{k'_X} = \frac{k_1(1+[H^+]/K_X)}{k_X(1+[H^+]/K_1)}$$

Fig. 11. Schematic representation of the reaction of amino groups of a protein with a small amount of acetic anhydride. 1, 2, 3 and 4 are four amino groups, progressively more 'buried'. X is an added calibrating nucleophile such as phenylalanine. $k'_1 \ldots k'_w$ are the pH-dependent apparent velocity constants for reaction of each group, and K_1 and K_X are acid dissociation constants.

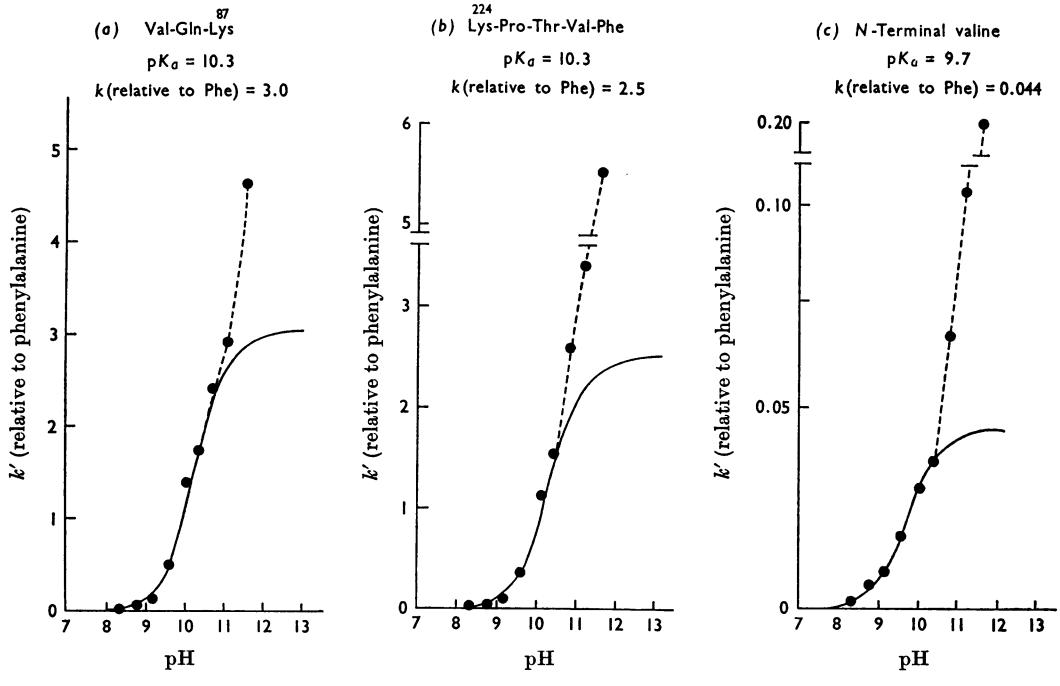


Fig. 12. pH-dependence of the reaction of acetic anhydride with three of the four amino groups of porcine elastase. The solid circles show experimental points determined by competition kinetics and the solid lines are theoretical dissociation curves for groups of pK_a and k as indicated.

where k_1 and k_X are velocity constants and K_1 and K_X the ionization constants for group 1 and phenylalanine respectively. The reacting species in excess is always the unmodified protein, so we need not worry in this case that acetylation of one group will affect the reactivity of another. Knowing pK_X and k_X for phenylalanine, we can calculate pK_1 and k_1 for each group by measuring the relative degrees of acetylation at a series of different pH values.

To do this we make an excess of protein plus phenylalanine react with [^{3}H]acetic anhydride at a series of different pH values. After complete reaction we add an excess of unlabelled acetic anhydride to produce chemically homogeneous but isotopically heterogeneous protein. The acetylphenylalanine is extracted and purified and its specific radioactivity determined. The protein is digested with a suitable mixture of proteolytic enzymes and the radioactive peptides are purified. Their composition and sequence identifies their position in the peptide chain and the specific radioactivity at each pH gives pK_1 and k_1.

Porcine elastase has four amino groups: the *N*-terminus, valine-16, and lysine-87, lysine-177 and lysine-224 (Shotton & Hartley, 1970). It was used as the model for this method since we expected valine-16 to be unreactive because it might form an internal ion pair with aspartic acid-194, like isoleucine-16 and aspartic acid-194 of α-chymotrypsin (Matthews, Sigler, Henderson & Blow, 1967). Reaction with radioactive acetic anhydride, as above, followed by peptic plus tryptic plus chymotryptic digestion gave peptides derived from each amino group except lysine-177. Non-specific proteolytic cleavage around this residue produced a number of peptides in poor yield, which were difficult to purify. The other labelled peptides were purified from mixtures reacted at a series of pH values from 8 to 11, and their specific radioactivities were determined.

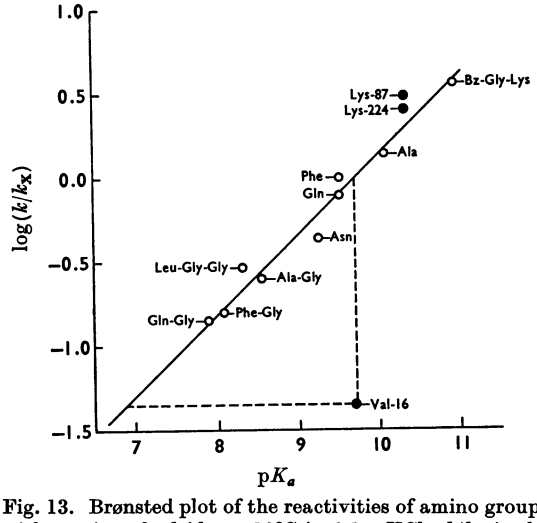


Fig. 13. Brønsted plot of the reactivities of amino groups with acetic anhydride at 10°C in 0.1 M-KCl. k/k_X is the ratio of the velocity constant of the unprotonated group to that of unprotonated phenylalanine. Abbreviation: Bz, α-*N*-benzoyl. The reactivities of amino groups of native elastase are indicated ●. The slope is 0.46 ± 0.03.

The results are shown in Fig. 12. The apparent velocity constants each fit a theoretical titration curve up to pH 10.5, but above this there is a discontinuity, possibly due to denaturation of the protein. Fig. 13 shows the Brønsted plot determined for some simple amino acids and peptides by the competitive labelling method. One sees that the pK_a and reactivity of each lysine residue is that expected for an ε-amino group freely accessible in solution to the reagent. The terminal amino group, in contrast, has pK_a 9.7 against an expected value of 7.5–8.0 (Tanford, 1962) and a reactivity only 4% that for a free amino group of this pK_a. One can conclude, as expected, that valine-16 forms an internal ion pair with aspartic acid-194 in elastase in solution, as was subsequently shown to be the case for the structure in the crystal (Shotton & Watson, 1970).

I feel that this competitive labelling approach could be widely exploited to explore the reactivity and pK_a of functional groups. The Bohr effect in haemoglobin, for example, can be defined as the change in pK_a of ionizing groups on oxygenation, so that a complete description of this phenomenon demands determination of all these pK_a values. Local conformational changes in a protein may well be revealed by changes in reactivity of one ionizing group relative to others. The competitive labelling method does not demand that the protein be pure so long as the relevant peptides can be purified, so it may be possible to ask precise questions about the relative accessibility of amino groups in complex cellular structures, such as cell membranes. The protein chemist has plenty of work ahead apart from the traditional determination of structures, and will be grateful for even more tools to do his job.

REFERENCES

Bayliss, R. S. & Knowles, J. R. (1969). *Biochem. J.* **113**, 377.

Bender, M. L., Chloupek, F. & Neveu, M. C. (1958). *J. Am. chem. Soc.* **80**, 5384.

Bender, M. L., Chow, Y.-L. & Chloupek, F. (1958). *J. Am. chem. Soc.* **80**, 5380.

Boulton, A. A., Chard, N. E. & Grant, L. (1965). *Biochem. J.* **96**, 82P.

Brown, J. R. (1964). *Biochem. J.* **92**, 13P.
Brown, J. R. & Hartley, B. S. (1963). *Biochem. J.* **89**, 59P.
Brown, J. R. & Hartley, B. S. (1964). *Abstr. FEBS 1st Meet., London*, p. 25.
Brown, J. R. & Hartley, B. S. (1966). *Biochem. J.* **101**, 214.
Brown, J. R., Kauffman, D. L. & Hartley, B. S. (1967). *Biochem. J.* **103**, 497.
Bruice, T. C. & Benkovic, S. J. (1966). *Bio-organic Mechanisms*, vol. 1, p. 174. New York: W. A. Benjamin Inc.
Bruice, T. C. & Pandit, U. K. (1960). *Proc. natn. Acad. Sci. U.S.A.* **46**, 402.
Bruton, C. J. & Hartley, B. S. (1968). *Biochem. J.* **108**, 281.
Bruton, C. J. & Hartley, B. S. (1970). *J. molec. Biol.* (in the Press).
Butler, P. J. G., Harris, J. I., Hartley, B. S. & Leberman, R. (1969). *Biochem. J.* **112**, 679.
Calam, D. H. & Waley, S. G. (1963). *Biochem. J.* **86**, 226.
Crowshaw, K., Jessup, S. & Ramwell, P. W. (1967). *Biochem. J.* **103**, 79.
Dixon, H. B. F. & Perham, R. N. (1968). *Biochem. J.* **109**, 312.
Edman, P. (1950*a*). *Acta chem. scand.* **4**, 277.
Edman, P. (1950*b*). *Acta chem. scand.* **4**, 283.
Foltmann, B. (1970). *Phil. Trans. R. Soc.* B, **257**, 147.
Foltmann, B. & Hartley, B. S. (1967). *Biochem. J.* **104**, 1064.
Freedman, M. H., Grossberg, A. L. & Pressman, D. (1968). *Biochemistry, Easton*, **7**, 1941.
Goodwin, T. W. (Ed.) (1970). *British Biochemistry—Past and Present.* London: Academic Press (Inc.) Ltd.
Gray, W. R. (1964). Ph.D. Thesis: University of Cambridge.
Gray, W. R. (1967). In *Methods in Enzymology*, vol. 11, p. 469. Ed. by Hirs, C. H. W. New York: Academic Press Inc.
Gray, W. R. & Hartley, B. S. (1963*a*). *Biochem. J.* **89**, 379.
Gray, W. R. & Hartley, B. S. (1963*b*). *Biochem. J.* **89**, 59P.
Gross, E. & Witkop, B. (1962). *J. biol. Chem.* **237**, 1856.
Habeeb, A. F. S. A., Cassidy, H. G. & Singer, S. J. (1958). *Biochim. biophys. Acta*, **29**, 587.
Hartley, B. S. (1964). *Nature, Lond.*, **201**, 1284.
Hartley, B. S. (1970). *Phil. Trans. R. Soc.* B, **257**, 77.
Hartley, B. S., Brown, J. R., Kauffman, D. L. & Smillie, L. B. (1965). *Nature, Lond.*, **207**, 1157.
Hartley, B. S. & Massey, V. (1956). *Biochim. biophys. Acta*, **21**, 58.
Kaplan, H., Hartley, B. S. & Stevenson, K. J. (1969). *Fedn Proc., Fedn Am. Socs exp. Biol.* **28**, 533.
Kauffman, D. L. (1965). *J. molec. Biol.* **12**, 969.
Keil, B. (1970). *Phil. Trans. R. Soc.* B, **257**, 125.
Kimura, M. & Kielley, W. W. (1966). *Biochem. Z.* **345**, 188.
Kostka, V., Morávek, L. & Šorm, F. (1970). *Eur. J. Biochem.* **13**, 447.
Lawson, W. B., Gross, E., Foltz, C. M. & Witkop, B. (1961). *J. Am. chem. Soc.* **83**, 1509.
Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. (1967). *Nature, Lond.*, **214**, 652.
Milstein, C. & Pink, J. R. L. (1970). In *Progress in Biophysics and Molecular Biology*, vol. 21, p. 209. Ed. by Butler, J. A. V. & Noble, D. Oxford: Pergamon Press Ltd.
Offord, R. E. (1966). *Nature, Lond.*, **211**, 591.
Perham, R. N. (1967). *Biochem. J.* **105**, 1203.
Perham, R. N. & Jones, G. M. T. (1967). *Eur. J. Biochem.* **2**, 84.
Pink, J. R. L. & Milstein, C. (1967). *Nature, Lond.*, **214**, 92.
Raftery, M. A. & Cole, R. D. (1963). *Biochem. biophys. Res. Commun.* **10**, 467.
Rajagopalan, T. G., Stein, W. H. & Moore, S. (1966). *J. biol. Chem.* **241**, 4940.
Ryle, A. P. & Anfinsen, C. B. (1957). *Biochim. biophys. Acta*, **24**, 633.
Ryle, A. P. & Sanger, F. (1955). *Biochem. J.* **60**, 535.
Ryle, A. P., Sanger, F., Smith, L. F. & Kitai, R. (1955). *Biochem. J.* **60**, 541.
Schoellmann, G. & Shaw, E. (1963). *Biochemistry, Easton*, **2**, 252.
Shotton, D. M. & Hartley, B. S. (1970). *Nature, Lond.*, **225**, 802.
Shotton, D. M. & Watson, H. C. (1970). *Nature, Lond.*, **225**, 811.
Smillie, L. B., Furka, A., Nagabhushan, N., Stevenson, K. J. & Parkes, C. O. (1968). *Nature, Lond.*, **218**, 343.
Smillie, L. B. & Hartley, B. S. (1964). *Abstr. FEBS 1st Meet., London*, p. 26.
Smillie, L. B. & Hartley, B. S. (1966). *Biochem. J.* **101**, 232.
Smillie, L. B. & Hartley, B. S. (1967). *Biochem. J.* **105**, 1125.
Smyth, D. G., Blumenfeld, O. O. & Konigsberg, W. (1964). *Biochem. J.* **91**, 589.
Spackman, D. H., Stein, W. H. & Moore, S. (1960). *J. biol. Chem.* **235**, 648.
Stark, G. R. & Stein, W. H. (1964). *J. biol. Chem.* **239**, 3755.
Stepanov, V. M. (1968). *Abstr. FEBS 5th Meet., Prague*, no. 1089.
Tanford, C. (1962). *Adv. Protein Chem.* **17**, 69.
Tang, J. & Hartley, B. S. (1967). *Biochem. J.* **102**, 593.
Tang, J. & Hartley, B. S. (1970). *Biochem. J.* **118**, 611.
Weber, G. (1952). *Biochem. J.* **51**, 155.
Weeds, A. G. (1967). *Biochem. J.* **104**, 44P.
Weeds, A. G. & Hartley, B. S. (1967). *J. molec. Biol.* **24**, 307.
Weeds, A. G. & Hartley, B. S. (1968). *Biochem. J.* **107**, 531.
Woods, K. R. & Wang, K. T. (1967). *Biochim. biophys. Acta*, **133**, 369.