*Phytochemistry*, Vol. 27, No. 8, pp. 2397–2409, 1988.
Printed in Great Britain.

0031-9422/88 $3.00 + 0.00
Pergamon Press plc.

# REVIEW ARTICLE NUMBER 35

# POLYPHENOL COMPLEXATION—SOME THOUGHTS AND OBSERVATIONS

CATRIONA M. SPENCER, YA CAI, RUSSELL MARTIN, SIMON H. GAFFNEY, PAUL N. GOULDING, DANIELE MAGNOLATO,* TERENCE H. LILLEY and EDWIN HASLAM

Department of Chemistry, University of Sheffield, Sheffield S3 7HF, U.K.; *Nestle Research Department, Nestec Ltd, Vers-Chez-Les-Blancs CH-1000, Lausanne, Switzerland

(*Received* 26 *January* 1988)

Dedicated to the Memory of Professor T. Swain, 1922–1987.

**Key Word Index**—Vegetable tannins; polyphenols; complexation with proteins; structure–activity relationships.

**Abstract**—Current views and opinions on polyphenol (tannin)–protein complexation are outlined. 'Structure–activity' relationships are delineated as are experimental approaches which seek to define the modes and sites of binding of ligand (polyphenol) to the receptor (protein). A model for polyphenol–protein complexation is proposed.

## INTRODUCTION

Everyone has their own cherished memories of Tony Swain but he was a person about whom it could be truly said his presence was larger than life. For over 30 years he made immense contributions to the establishment, to the development and to the intellectual climate of modern phytochemistry. The outcome of this long engagement with the subject was a view of phytochemistry which was extremely well based, authoritative and comprehensive. It expressed itself in various ways most notably in his legendary erudition wherein he invariably gave expression to a highly personal and persuasive interpretation of scientific problems.

His own scientific career was certainly a form of personal pilgrimage and his paper with E.C. Bate-Smith in 1953 [1] was not only a symbolic starting point but was also a landmark in this particular area of phytochemistry. It commenced:

> The property of astringency in foods and beverages is commonly ascribed to tannins, but in very few instances have these tannins been isolated and characterised. Most of those which have been studied fall into the class of condensed tannins. . . . . several considerations have suggested to us that in many plant tissues, leuco-anthocyanins (Robinson and Robinson) are the substances responsible"

and concluded:

> "All these considerations, taken together, suggest that the leuco-anthocyanins are closely related to the catechins and are, as well as these, to be regarded as prototypes of the condensed tannins."

Retrospectively this paper represented a major intellectual breakthrough in the study of vegetable tannins and it forms the symbolic starting point for this review since polyphenols (tannins) thereafter formed a substantive element in Tony Swain's own scientific career. Bate-Smith had repeatedly drawn attention to the distinction in chemical make-up between woody and non-woody plants and in 1960 Swain wrote an important paper as a contribution to a symposium on 'Phenolics in Plants in Health and Disease' at Bristol [2]. Largely neglected this paper drew attention to the inter-relationships between lignin and leuco-anthocyanins. Nearly 30 years later it bears further detailed reading and consideration. Without doubt it highlights a relationship which has come once again into prominence [3] and will be of future significance.

Early workers had however displayed considerable interest in the biological function of tannins and it was this theme which dominated Swain's thoughts increasingly. It was known that tannins modified the effect of bacteria, fungi and viruses but it had also often been suggested that they were metabolic waste products. No doubt fired by Bate-Smith's own inspiration and conviction ("I myself started with the information that tannins were waste products deposited in the wood of trees because plants had nothing else they could do with them; a horrible thought! Convinced they had a function, I set out to try and find it by way of their systematic distribution linked to the idea of their astringency"–E. C. Bate-Smith, personal communication) Swain's work focused ever more clearly with succeeding years on this central problem and with it an attempt to understand one facet of the restless tide of evolutionary change in plants.

The key element, the *sine qua non*, of the Darwinian view of nature is that the purposeful construction of living matter can be attributed to natural selection. Several assumptions–the *ifs*, Cairns-Smith [4]–predicate that viewpoint. Most notably; *if* there are random variations in systems which can reproduce their kind; *if* such variations are inherited and confer some advantage to

these reproducing systems–then these entities will have an enhanced chance of survival in any competitive situation. The view that secondary plant constituents (alkaloids, terpenes, phenols, etc.) may play a leading role in the co-evolutionary interactions between plants and herbivores fits intuitively into this general intellectual framework [5]. The idea has a long history but only developed with explosive effect in the 1960's and because of the enviable facility with which he was able to bridge the gulf between chemistry and biology Tony Swain was a natural participant in these developments. In a key paper in 1978 he considered the central proposition and the specific influence of tannins [6]:

> "Proanthocyanidins or condensed tannins .... are undoubtedly the most useful of all plant chemical defences. Not only are they potent anti-fungal, anti-bacterial and even anti-viral agents but they bestow on plants which have the ability to synthesise them a powerful feeding deterrent to all herbivores .... These phenolic polymers, act through their ability to combine with protein and this inhibits enzymes and reduces the nutritionally available protein in foods. They are difficult to degrade and because of their molecular size this renders them almost impossible to sequester into vesicles so that their activity may be aborted".

Since a Yorkshireman is said to be a Scotsman with the generosity squeezed out, few would recognise Tony Swain as a native of the county of broad acres which takes its name from the House of York. He was the antithesis of such a figure. He was a Dalesman, invariably expansive and mellow, occasionally robust, occasionally uncertain but always eloquent and with a boyish enthusiasm for science few could equal. In Sheffield, in a corner of his native county, similar objectives and questions concerning vegetable tannins have long been entertained. The approach has been substantially different but complementary to the more biological path travelled by Swain. Initially the view was taken that it would be extraordinarily difficult, if not impossible, to assess the value of suggestions concerning the ecological/metabolic role of polyphenols until much more was known about the details of structure and biosynthesis of plant polyphenols. When this void was filled (a mere 20 years or so of intensive effort!) studies of the interaction of polyphenols with proteins, polysaccharides, nucleic acids and other substrates commenced. The objective of this phase of the work is to establish whether real specificity exists in these interactions. If it does then it may ultimately be related to some function in metabolism and/or evolution. Contemporary interpretations of the significance of plant polyphenols in evolution rest largely [6] on the presumption that they act via their capacity to bind to proteins. This ability to complex with proteins also has numerous practical consequences and the brief review which follows gives a 'state of the art' picture of polyphenol–protein complexation, as viewed from Swain's native shire, and is a tribute to his immense contributions to the study of vegetable tannins and to phytochemistry as a whole.

## DISCUSSION

Studies of the reversible association of polyphenols with proteins have a long history. One of the first scientific papers on this topic was that of Sir Humphry Davy in 1803. This early work [7] demonstrated some of the macroscopic features of polyphenol complexation and it gave rise to several, wholly empirical, definitions of the term 'vegetable tannin' [8]. However until such times as structurally defined plant polyphenols became available the molecular mechanisms which underly polyphenol–protein complexation were not amenable to systematic investigation and were, in consequence, poorly understood. Various studies of polyphenol complexation—in particular the delineation of 'structure–activity' relationships—have been conducted in the recent past and they now permit a much more detailed picture of the processes which are, most probably, involved to be outlined.

The association of polyphenols with proteins is principally a surface phenomenon (Fig. 1a, b, c). The efficacy of polyphenol binding to protein derives from the fact that polyphenols are multidentate ligands able to bind simultaneously (via different phenolic groups) at more than one point to the protein surface. When polyphenols cause precipitation of proteins from solution two situations may be envisaged. At low protein concentrations the polyphenol associates at one or more sites on the protein surface, to give a mono-layer which is less hydrophilic than the protein itself (Fig. 1a). Aggregation and precipitation then ensue. Where the protein concentration is high the relatively hydrophobic surface layer is formed by complexation of the polyphenol onto the protein and by cross-linking of different protein molecules by the multidentate polyphenols (Fig. 1b). Precipitation then follows as above. This tendency to cross-link protein molecules

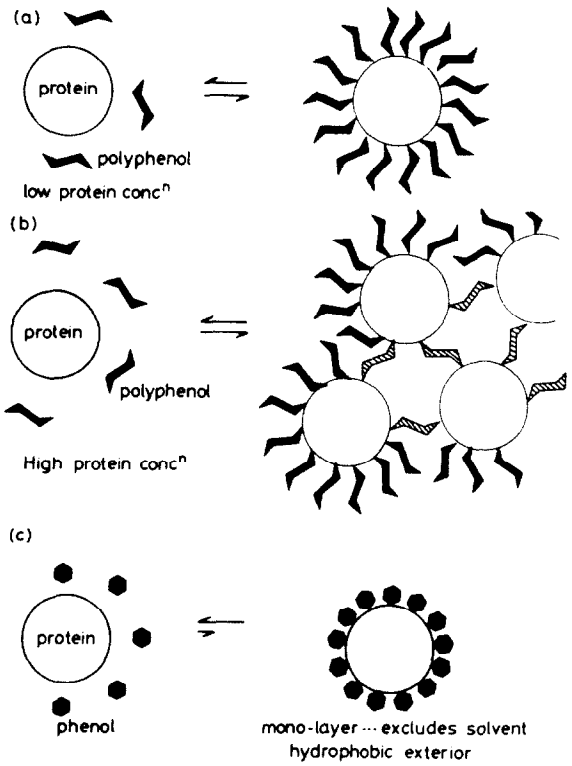



Fig. 1. Polyphenol–protein complexation and precipitation.

at higher protein concentrations explains the changing stoicheiometry of the aggregates with changing protein concentrations—an observation first hinted at by Sir Humphry Davy. More polyphenol is thus required to precipitate proteins from dilute solution than from concentrated solutions.

A significant corollary of this hypothesis is that various simple phenols such as pyrogallol should also be capable of precipitating proteins from solution if they can be maintained in solutions at concentrations sufficient to push the equilibrium in favour of the protein–phenol complex and thus form a hydrophobic layer of simple phenol molecules on the protein surface (Fig. 1c). For many simple phenols the limit is provided by their solubility in water but it can be achieved with, for example, bovine serum albumin (BSA, $3 \times 10^{-5}$ molal) and pyrogallol (1 molal). Thus when comparatively simple phenolic substrates are present in solution they will also compete with polyphenols for sites on the protein molecules in the complexation processes. Simple phenols, depending on their effective concentration in solution, may therefore be expected to modify polyphenol–protein complexation. Complexes can be dissociated by the addition of further protein (see Davy [7]), hydrophobic solvents or hydrogen bond acceptor solvents (e.g. acetone), urea, polyvinylpyrrolidone, polyethylene glycols, non-ionic detergents and nitrogen heterocycles such as caffeine. Complex formation is pH dependent; each protein has a distinctive pH optimum which usually lies at or very near to the isoelectric point of the particular protein. Where measurements have been possible the extent of complex formation rapidly declines as the pH of the medium is raised above 9.0. The inference generally drawn from all these observations is that polyphenols interact with protein by the formation of strong non-covalent bonds rather than by the elaboration of ionic or covalent linkages [8, 9].

*Structure–activity relationships*

(i) *Polyphenols.* Although as a sub-group the pro-anthocyanidins are most commonly responsible for the range of reactions generally attributed to tannins in plants many of these particular studies have generally been most conveniently pursued to date with a series of biosynthetically inter-related esters of gallic acid. They are accessible in homogeneous forms and differ systematically in phenolic content, solubility, molecular size and conformation, (e.g. Fig. 2, [10]). Polyphenol complexation has been studied in solution or by an investigation of the precipitation process which ultimately ensues after extensive complexation (Fig. 3).

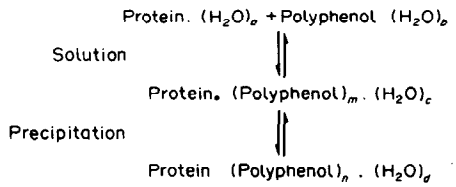



Fig. 3. Polyphenol–protein complexation.

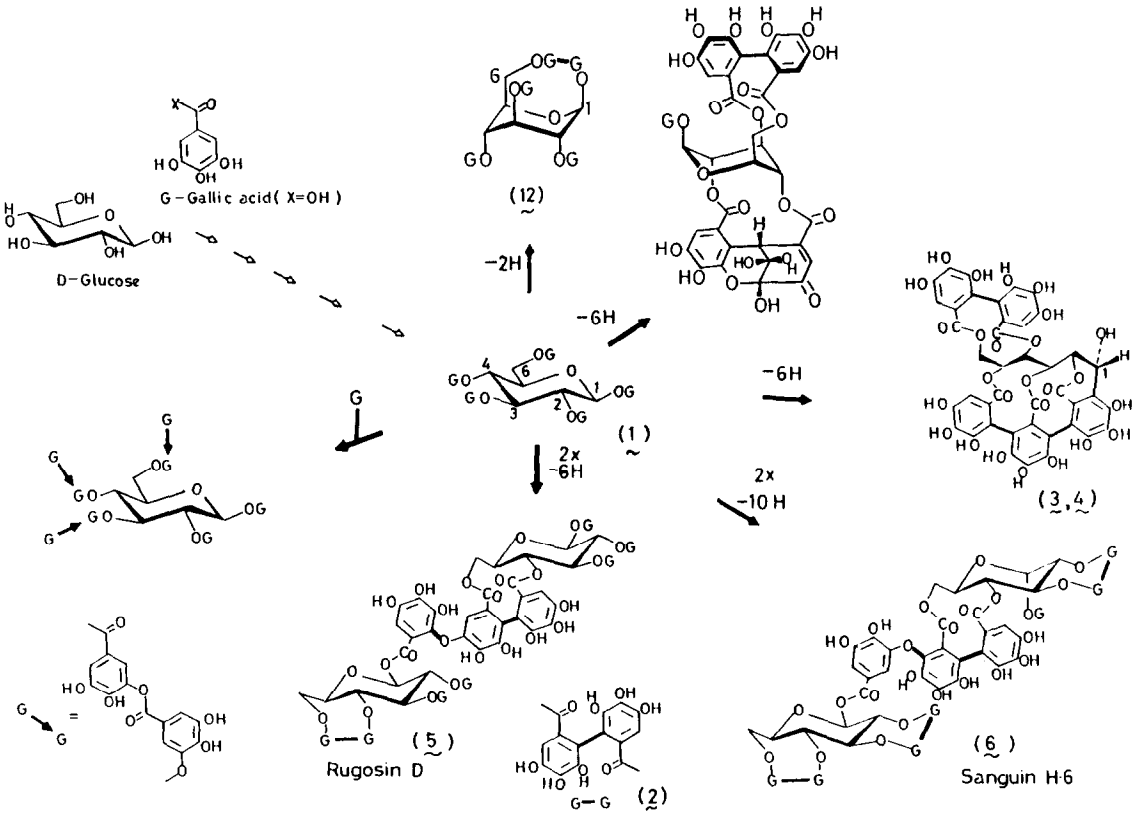



Fig. 2. Some biosynthetically inter-related esters of gallic acid.

Various physico-chemical measurements are appropriate for these various situations. Amongst those most widely employed in recent studies in solution are equilibrium dialysis and microcalorimetry [11], enzyme inhibition [12] and affinity chromatography [13]. Most familiar of the quantitative precipitation techniques are the method of haemanalysis due to Bate-Smith [14] in which precipitation of haemoglobin is followed colorimetrically, and the procedure of Hagerman and Butler [15] using radio-iodinated BSA, monitored by change of radioactivity. Mole and Waterman [16] have critically assessed some of these procedures from the point of view of their accuracy and effectiveness in ecological investigations. There is however a broad, although not exact, comparability in the information derived from the various types of measurement. The three principal features of polyphenol structure and properties which are of importance in the complexation with protein are:

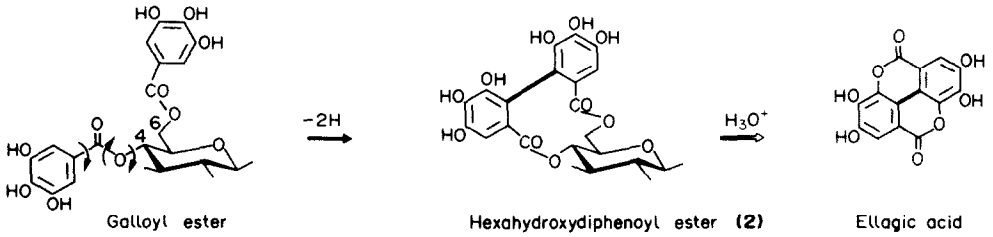
(a) Molecular size of the polyphenol.
(b) Conformational flexibility of the polyphenol. When conformational restraints are placed on the polyphenolic substrate then its capacity to complex, whatever its molecular size, is dramatically reduced.
(c) Water solubility of the polyphenol. A broadly inverse relationship exists between the strength of association with protein and the solubility of the polyphenol in water. Low solubility favours strong association.

The effectiveness of plant polyphenols as complexing agents derives from the fact that they are polydentate ligands with a multiplicity of potential binding sites provided by the numerous phenolic groups and aryl rings on the periphery of the molecule. Because of the propinquity of these groups in the polyphenol, cooperativity of complexation to the protein surface is observed. Also as a consequence of their molecular size and structure polyphenols form stable cross-linked structures with different protein molecules Fig. 1. Molecular size is thus critically important as a determinant of the ability of a polyphenol to bind to protein–as inferred in Bate-Smith and Swain's original definition of a vegetable tannin [8]. This is illustrated by the fact that in the simple galloyl-D-glucose series the efficacy of association with protein is enhanced with the addition of each galloyl ester group (di→tri→tetra→penta) and reaches a maximum in the flexible disc-like structure of β-penta-*O*-galloyl-D-glucose (**1**). Further metabolism of this intermediate often substantially lowers its powers of association.

Equally significant as molecular size of the polyphenol is its conformational mobility and flexibility. In the galloyl-D-glucoses when vicinal galloyl ester groups are constrained by the biosynthetic intramolecular formation of biphenyl linkages and the generation of hexahydroxydiphenoyl ester groups (**2**) the loss in conformational freedom is reflected in a reduced capacity to bind to protein. The apotheosis of this effect is seen in the case of the unique open chain D-glucose derivatives vescalagin and castalagin, metabolites of *Quercus* sp. These rigid virtually inflexible, propeller shaped molecules (**3**, **4**) are in a sense analogues of β-penta-*O*-galloyl-D-glucose (**1**), but on a molar basis they are bound much less effectively to protein than (**1**). In this context the observed 'relatively lower astringency' of the proanthocyanidins compared to other polyphenols may be explicable, in part, in terms of the conformational restraints imposed by restricted rotation about the repeating 4,8 or 4,6 interflavan bonds.

One of the key factors in this whole problem is, of course, the solvent water. To date its presence has been largely ignored. Biologists have long been intrigued by the concept of 'structured' water solvation shells around molecules and whether there are preferred sites thereon for species such as inorganic ions and water molecules. Thus whilst it is pertinent to enquire what is the effect of 'bound' water on the properties of the biological substrate (e.g. protein) and what is the effect of solvation on complexation with polyphenols, present evidence gives no simple answer to such questions. What is clear however is that polyphenolic substrates such as vescalagin and castalagin (**3**, **4**) which are quite soluble in water have a weaker affinity for proteins than those such as β-penta-*O*-galloyl-D-glucose (**1**) which are poorly soluble in water. Various observations suggest that the effects of differential solvation are clearly quite crucial. If, for example, a titration of aroyl proton chemical shift ($\delta$) versus solvent composition is made for a polyphenol such as β-pentagalloyl-D-glucose (**1**), (Fig. 4) in 100% $D_6$ acetone (or $D_4$ methanol) progressively changing the solvent until it reaches 100% $D_2O$ marked changes are observed, particularly in the region 80% $D_2O$ to 100% $D_2O$. Presumably these effects are in large part due to preferential solvation [17]. The substantially different degrees of change of chemical shift represent, it is assumed, different relative hydration propensities of the individual galloyl ester groups. However it is not yet clear how these factors are related to the individual galloyl ester group's propensity to bind to protein, although it is envisaged that relatively weak hydration would be a corollary of strong protein–galloyl group interactions. Some aspects, both empirical [18] and theoretical [19] of the competition between solvation and interaction of solutes have recently been presented.

(ii) *Proteins.* Complementary studies of proteins to determine structure–activity relationships have been made by Hagerman and Butler [15] using a competitive protein precipitation assay with bovine serum albumin (BSA) and lysozyme labelled with iodine-125, somewhat in the fashion of a competitive immuno assay. The 'competitor'



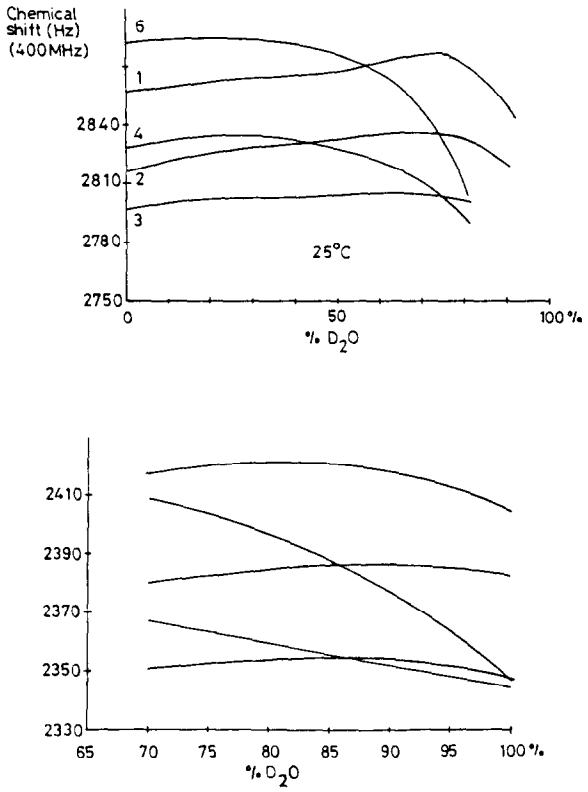



Fig. 4. Solvation induced chemical shift changes ($\Delta\delta$-$^1$H NMR) of the galloyl ester protons of $\beta$-pentagalloyl-D-glucose (**1**): 100% acetone to 100% $D_2O$.

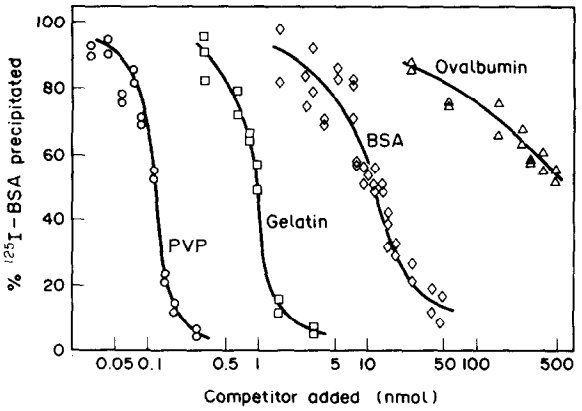



Fig. 5. Protein precipitation [15]. Inhibition of tannin–BSA precipitation at pH 4.9 by proteins and polymers.

(proteins, low $M_r$ amides and amino acids, polyamino acids) was mixed with an aliquot of the iodinated protein and the insoluble complex formed at pH 4.9 after exposure to the proanthocyanidin polymer from *Sorghum bicolor* [20, 21] assayed for radioactivity. Plots of the quantity of $^{125}$I precipitated (expressed as a percentage of the amount precipitated in the absence of the 'competitor') were plotted as a function of the logarithm of the amount of 'competitor' added (Fig. 5). The relative affinities of the various 'competitors' varied over more than four orders of magnitude at pH 4.9 clearly suggesting that the proanthocyanidin polymer interacts with proteins in a selective manner. A common feature of the proteins and polypeptides with a high affinity for the sorghum polyphenol was their high proline content and Hagerman and Butler attributed these differences to two major factors. The first was the much more open and flexible conformations of the high proline polypeptides and secondly was their increased capacity to form strong hydrogen bonds with the polyphenol brought about by the increased accessibility of the peptide linkages and by the bis-alkyl substitution of the proline amide nitrogen* (7).

The technique of Hagerman and Butler measures protein affinities via complex precipitation. Relative protein affinities have also been assessed in solution. Complexation in this situation is completely reversible (Fig. 3) and the protein-polyphenol complex is maintained in solution throughout. The assay method [12, unpublished observations] employs polyphenol induced inhibition of the enzyme $\beta$-glucosidase. 'Competitor' proteins are added and the relief of polyphenol inhibition (using the substrate, **1**) monitored by normal kinetic analysis.

Detailed studies of the enzyme $\beta$-glucosidase and its inhibition by polyphenolic substrates have demonstrated that under certain conditions the inhibition approximates to the classical non-competitive case. This occurs when enzyme, polyphenol and substrate are brought together simultaneously and without pre-equilibration [12]. The kinetics observed are then most closely correlated with the classical pattern of non-competitive inhibition in which polyphenol (inhibitor I) and substrate (S) are assumed to bind simultaneously to the enzyme. In the simplifying case, assuming Michaelis–Menten kinetics, $K_M$ remains unaffected, the [enzyme inhibitor substrate: E.I.S.] complex does not react and $K_i$ (inhibitor constant) may be determined and used as a quantitative measure of the affinity of the polyphenolic inhibitor for the enzyme–$\beta$-glucosidase. Low values of $K_i$ indicate a relatively strong affinity for the protein, high values a correspondingly weak affinity for protein. The dominant feature in the presumed mechanism of non-competitive enzyme inhibition is assumed to be that in which the polyphenolic inhibitor complexes to the protein surface cross-linking and thereby 'locking' adjacent regions of the enzyme structure such as to prevent the necessary conformational changes which must ensue for catalysis to occur. However if it is assumed that this is a process possessing a degree of randomness then complexation at or near the enzyme active-site might also be reasonably anticipated. This would then give rise to direct competitive inhibition of the enzyme. From the present kinetic data it must be presumed that, in the case of $\beta$-glucosidase *without* pre-equilibration with polyphenol, this makes a relatively small contribution in the initial stages to the overall pattern of inhibition.

If however enzyme and polyphenol (inhibitor) are pre-equilibrated in admixture prior to the addition of the substrate then mixed (competitive/non-competitive) kinetics are observed [Goulding, P. N., Lilley, T. H. and Haslam, E., unpublished]. The most probable explanation of this pattern of behaviour is that binding of the polyphenol, at or near the active site of the enzyme, is a

kinetically slow process compared to binding at other sites on the enzyme surface. Consequently under the conditions of pre-equilibration the polyphenol (inhibitor) must penetrate to the active-site of the enzyme β-glucosidase and thus compete directly with the substrate and give rise to mixed competitive/non-competitive kinetics.

The results obtained by this method, comparing different proteins and their ability to relieve polyphenol induced inhibition of the enzyme β-glucosidase, are not so clear cut and emphatic in their delineation of specificity of complexation *when comparisons are made on a weight (protein) basis*, Fig. 6. Compact, globular proteins (e.g. myoglobin, cytochrome *c*) have a poor affinity for the polyphenol inhibitor (β-pentagalloyl-D-glucose, **1**) and render little apparent relief of inhibition. Other proteins (e.g. BSA, ribonuclease, β-lactoglobulin) gave broadly similar effects–quantitative and qualitative–in the pattern of recovery of enzymic activity (Fig. 6), following a response which was approximately hyperbolic in form in its early stages. Polyvinylpyrrolidone ($M_r$ 10000 and 360000) and polyproline ($M_r$ 7000) initiated *ca* linear rates of recovery of enzyme activity (Fig. 6) but on a protein *weight basis* the extents of relief of enzyme inhibition attained were ultimately very similar to those obtained with natural proteins. It seems reasonable, at this stage, to conclude that compact globular proteins do not provide in aqueous media sufficient sites for polyphenol complexation. The full significance and the origins of the different patterns of behaviour of proline containing proteins are however not as yet clear.

*Modes and sites of binding*

Substantial efforts have been made towards the elucidation of the mode of interaction between polyphenols and proteins. Various proposals have been made and the principal suggestions are that the two species reversibly complex via (i) hydrogen bonding and (ii) hydrophobic interactions [9, 13, 22–23]. The relative quantitative importance of these two types of interactions remains, however, uncertain. Whilst the tendency has been intuitively to emphasise the part played by intermolecular hydrogen bonding several workers, notably Hoff [13], have drawn attention to the fact that hydrophobic effects may dominate the complexation process. The emphasis on hydrogen bonding as a vehicle for association was derived from the fact that polyphenols and simple phenols were bound by synthetic polymers (e.g. nylons) containing the –CO–NH– group as the only 'reactive' functionality. But evidence has also accumulated which is most reasonably interpreted in terms of hydrophobic bonding in the formation of protein–polyphenol complexes. Thus Goldstein and Swain [24] observed that the complexes could be dissociated by detergents and Hoff [13] has shown that proteins adsorbed on a column of Sepharose containing immobilized polyphenols were effectively eluted by anionic and non-ionic detergents. A study of the interaction between proanthocyanidins and polyamino-acids free in solution showed similarly that the extent of complexation was related to the number of methylene groups in the amino-acid side chain.

Some insight into this problem has recently been revealed by studies of polyphenol complexation with model systems and in particular the heterocycle caffeine (**8**) [25–27] and various model peptides.

Whilst it is not possible to extrapolate directly to behaviour in solution X-ray crystallographic analysis of various caffeine-phenol complexes confirm the importance of (i) hydrogen bonding, (ii) apolar hydrophobic interactions and, in certain situations, coordination around a metal ion as the primary intermolecular forces in caffeine-polyphenol complexation. In so far as the caffeine molecule (**8**) has structural elements which resemble a peptide-like structure these observations probably also point to the significance of these same non-covalent forces in the association of polyphenols with proteins.

X-ray crystallographic analysis of complexes of caffeine (**8**) with a variety of phenolic substrates show that in the solid state a layer-lattice structure is frequently present. In this array caffeine and aromatic molecules are arranged in stacks in alternating layers, approximately parallel, with an interplanar separation of ~ 3.3 to 3.4 Å, Fig. 7. For methyl gallate (**9**) this stacking structure is complemented (Fig. 8) by an extensive in plane system of hydrogen bonding between the three phenolic hydroxyl groups of methyl gallate and the two keto-amide groups and the basic N-9 of caffeine. The crystal structure of the caffeine complex with potassium chlorogenate first isolated from coffee beans by Gorter shows similar features to those noted with methyl gallate. However, an additional

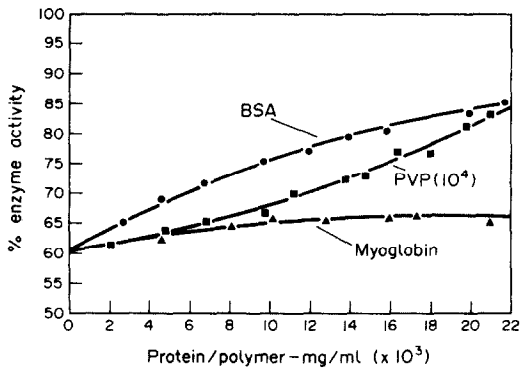



Fig. 6. Relief of polyphenol (**1**) induced inhibition of β-glucosidase.

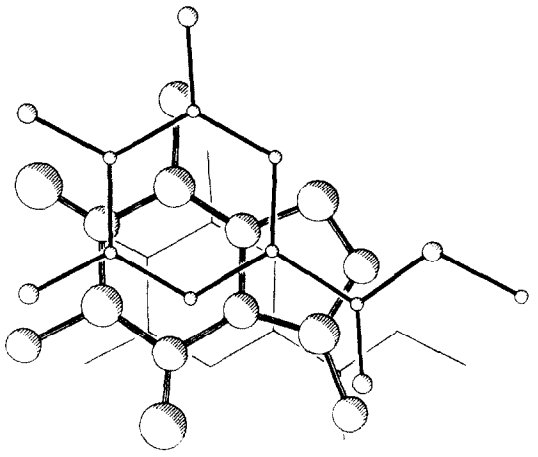


Fig. 7. Methyl gallate (**9**)–caffeine (**8**) complex–vertical stacking of 'aromatic' nuclei.

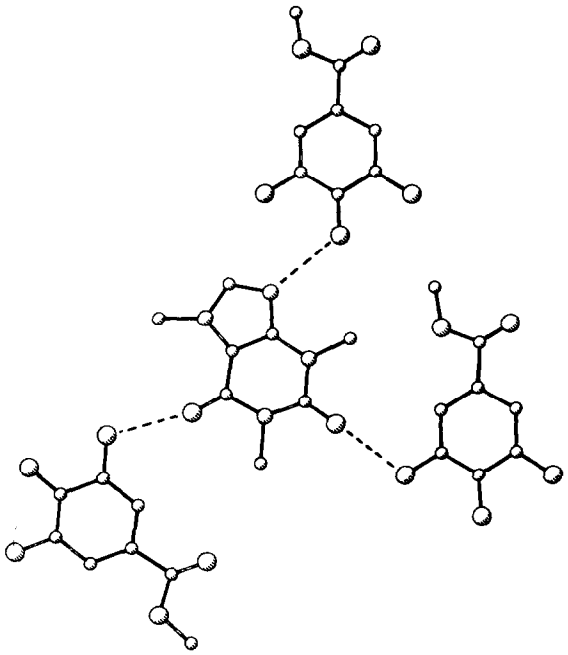
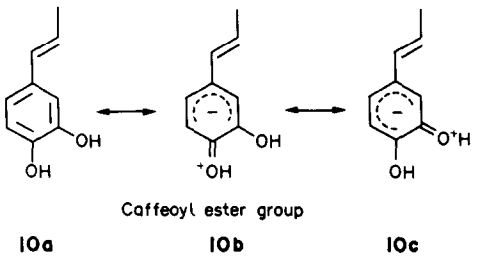
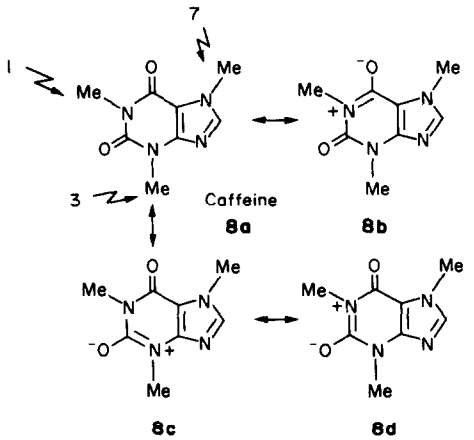


Fig. 8. Methyl gallate (**9**)–caffeine (**8**) complex–in plane hydrogen bonding.

critical stabilising factor is the coordination of seven oxygen atoms in an irregular polyhedral arrangement around the central potassium ion ($K^+$ . . . . 0 distances between 2.67 and 2.85 Å). A further feature of interest in these crystalline phenol–caffeine complexes is the relative orientation of the planar caffeine and phenolic partner in the layer lattice. The term 'polarisation bonding' has been used to describe both charge-transfer bonding and the generally weaker interactions between polar groups of one component and a polarisable second component [28]. For a weak non-covalent bonding of this type the principal feature to be expected is the juxtapositon of the polarising groups of one component and the polarisable regions of the second. It is therefore interesting to note that in the various phenol–caffeine complexes the phenolic groups and associated nuclei are generally stacked above the 6-membered ring of the caffeine molecule. This suggests that in this form of association the two components (e.g. caffeoylester and caffeine) develop complementary polar characteristics of the type shown (**8**, **10**, a–c etc.).

Previous work has shown that the heterocycle caffeine will reversibly associate with a range of substrates to form molecular complexes. The researches of Mejbaum-Katzenallenbogen [29, 30] further demonstrate that caffeine competes effectively with proteins for polyphenolic substrates and that it is possible to regenerate a wide variety of proteins, in a biologically active state, from insoluble protein–tannin complexes by treatment with caffeine. The heterocycle precipitates polyphenols from aqueous media by complexation and the ratio of caffeine to polyphenolic substrate in the complex is *higher* with *lower* initial concentrations of the polyphenolic substrate–an observation, which, in a sense, is directly analogous to the stoicheiometry of protein–polyphenol complexes formed by precipitation (*vide supra*). Studies with two polyphenolic substrates, at equivalent molar concentrations–β-penta-*O*-galloyl-D-glucose (**1**) and β-1,3,6-tri-*O*-galloyl-D-glucose (**11**) shows, rather interestingly, that the latter forms complexes with a higher caffeine–polyphenol ratio before precipitation ensues. Extrapolating from the crystal data described earlier the aggregates which form in solution, sufficient to cause precipitation, probably therefore take the general form shown in Fig. 9. In this model it is postulated that the caffeine links separate polyphenol molecules, rather as the latter cross-link proteins in the analogous protein-polyphenol precipitation process, (Fig. 1). A crystalline complex formed between β-1,2,4,6-tetra-*O*-galloyl-D-glucose (**13**) and caffeine (**8**) has been isolated but its structure has yet to be determined. The ratio of polyphenol: caffeine in the complex is 2:5. In the context of polyphenol complexation with caffeine it is interesting also to note that polyphenols may also be precipitated from solution by other substrates such as berberine, cinchonine, papaverine, quinine and strychnine and by the dyestuff Methylene Blue [31]. It is perhaps significant to note that molecules such as berberine and Methylene Blue are planar and bear a positive charge (formally located on nitrogen but extensively delocalised around



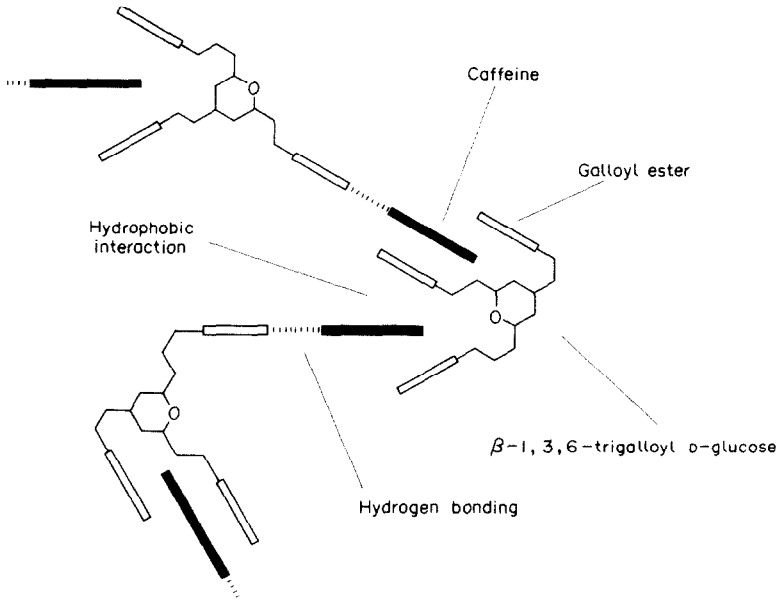



Fig. 9. Polyphenol–caffeine precipitation.

the ring system) and their interaction with polyphenols may be interpreted, in principle, in terms which are very similar to those of caffeine with phenolic groups.

The association of caffeine with various polyphenols has also been studied by microcalorimetry and by $^1$H NMR spectroscopy. The latter technique exploits the tendency of phenolic molecules and caffeine to 'stack' (*vide supra* Fig. 7) in solution. This in turn results in a deshielding of the aromatic protons of the phenol and of the three methyl groups and the single proton of the caffeine molecule and this causes upfield shifts in the respective $^1$H NMR signals, e.g. Fig. 10. The $^1$H NMR spectrum of caffeine is composed of four singlets. One from each of the three methyl groups and one from the proton attached to carbon atom (C-8). Various workers have observed that, at fixed caffeine concentration, the addition of aromatic substrates to the caffeine solution causes an 'upfield' shift of each of the four singlets due to magnetic anisotropic effects. This may be exploited to calculate the equilibrium constant for the association between caffeine and the aromatic substrate. Use has been made of eqn (1), an analogue of the Benesi–Hildebrand equation, and also making due allowances for the self-dimerisation of both caffeine and the aromatic substrate [Martin, R., Spencer, C. M., Cai, Y., Lilley, T. H. and Haslam, E. unpublished observations].

$$\frac{1}{\Delta} = \frac{1}{K.\Delta_0^{AP}} \cdot \frac{1}{[P_0]} + \frac{1}{\Delta_0^{AP}} \qquad (1)$$

where:-

$\Delta$ = Chemical shift change induced by substrate $P$ at concentration $[P_0]$

$[P_0]$ = Formal substrate concentration

$K$ = Equilibrium constant for the formation of a 1:1 complex between caffeine and the substrate $P$

$\Delta_0^{AP}$ = Chemical shift difference between the caffeine resonance in the unbound state and the state in which it is totally in the form of a 1:1 complex.

The equilibrium constants for the formation of a 1:1 complex between caffeine (**8**) and a series of natural polyphenols have been determined at various temperatures (Table 1). Significantly these results show exactly the same *trends* which were observed earlier [11] in

Table 1. Polyphenol–caffeine complexation

| Association constants ($K_{AP}$ mol$^{-1}$ l) in deuterium oxide | 60 | 45° |
|---|---|---|
| A. Galloyl esters | | |
| Methyl gallate (**9**) | 11.0 | – |
| Davidiin (**12**) | 23.2 | – |
| Corilagin (**18**) | 16.7 | – |
| Casuarictin (**23**) | 18.9 | – |
| Tellimagrandin 1 (**21**) | 21.9 | – |
| β-1,3,6-Tri-*O*-galloyl-D-glucose (**11**) | 36.4 | – |
| β-1,2,6-Tri-*O*-galloyl-D-glucose (**19**) | 42.1 | – |
| Tellimagrandin 2 (**22**) | 61.9 | – |
| β-1,2,4,6-Tetra-*O*-galloyl-D-glucose (**20**) | 52.3 | – |
| β-Penta-*O*-galloyl-D-glucose (**1**) | 97.9 | – |
| Sanguin H-6 (**6**) | 71.5 | – |
| Rugosin D (**5**) | 135.6 | – |
| B. Flavan-3-ol derivatives | | |
| (+)-3-*O*-Galloylcatechin (**16**) | 26.8 | 38.3 |
| (−)-3-*O*-Galloylepigallocatechin (**17**) | 35.1 | 52.8 |
| Procyanidin B-2 (**14**) | – | 26.1 |
| Procyanidin B-3 (**15**) | – | 22.3 |

Association constants quoted as average of four values obtained by calculation respectively from the chemical shift changes ($\Delta\delta$) of the proton at C-8 and the three *N*-methyl groups of caffeine in $D_2O$, using an external standard.

studies of polyphenol–protein complexation, namely a very strong dependence on molecular size of the polyphenol and on its conformational flexibility. They thus strongly confirm and reinforce the view, originally assumed, that polyphenol–caffeine complexation not only mirrors in many of its facets polyphenol–protein complexation, but is also a good model system with which to examine the general mechanisms of association. Particularly striking are the depressive effects on complexation which result from restriction of conformational mobility by the biosynthetic formation of the hexahydroxydiphenoyl ester group (**2**) in the polyphenolic substrate (Table 1), the strong temperature dependence of complexation (Table 1) indicative of significant entropy effects (which perhaps suggests important solvation effects), and the relatively weaker affinity of typical procyanidins (**14**, **15**) as compared to galloyl esters for caffeine (Table 1). This latter observation underlines the generally assumed weaker affinity which proanthocyanidins (condensed tannins) have for proteins when compared to galloyl esters (hydrolysable tannins). This work has also been extended to define preferred points of complexation with caffeine on the polyphenolic substrate.

A crucial feature of the assignment of structure to polygalloyl-D-glucose esters of the type shown in Fig. 2 has been the location of the various different types of ester functionality on the D-glucopyranose molecule by means of $^1H$ and $^{13}C$ NMR spectroscopy [32]. This work has been an invaluable aid to the subsequent studies of complexation with caffeine and other substrates since it has permitted for the first time unequivocal assignment of the signals in a polygalloyl ester to the individual aroyl ester goups. The approach and methodology is illustrated for β-penta-*O*-galloyl-D-glucose (**1**). The proton NMR spectrum of this polyphenol displays a series of two proton singlets for each of the five galloyl ester groups (Figs 10, 11); the proton-coupled $^{13}C$ NMR spectrum shows an analogous sequence of five signals for the individual carbonyl carbon atoms, (Fig. 12)–four quartets ($J_{CH} \sim 5$ Hz) and a multiplet. In the $^1H$ NMR spectrum the signals from the protons of the D-glucopyranose ring may be associated unequivocally with the individual protons H-1, H-2, H-3 etc. by conventional proton decoupling techniques. Use is then made of the characteristic three-bond long range C-H coupling to establish, *via* a two dimensional NMR pulse sequence, the connectivities between individual protons on the D-glucopyranose, the carbonyl carbon atoms and the protons of each galloyl ester group. In this way each galloyl ester group two proton singlet (Fig. 11, $\sim\delta$ 6.95–7.25) and each carbonyl carbon atom signal (Fig. 12, $\sim\delta$ 166–168) may be specifically associated with a parti-

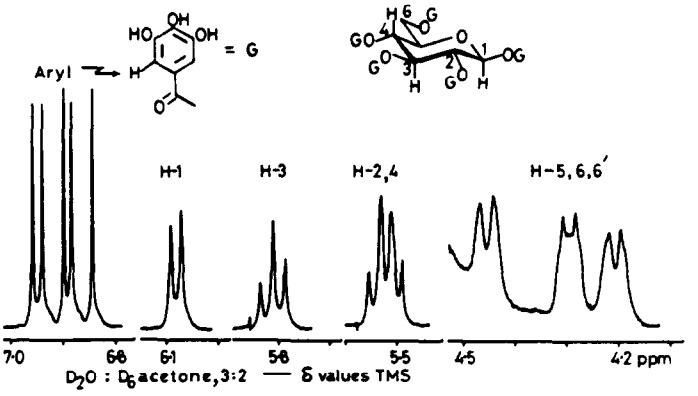



Fig. 10. $^1H$ NMR spectrum of β-pentagalloyl-D-glucose (**1**).

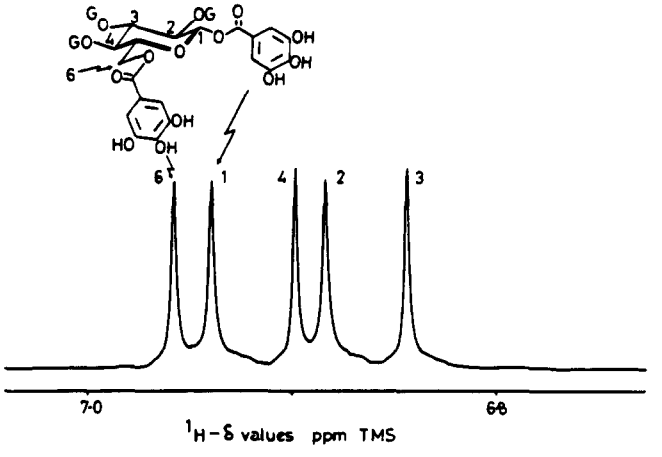



Fig. 11. $^1H$ NMR assignment of the galloyl ester proton signals.

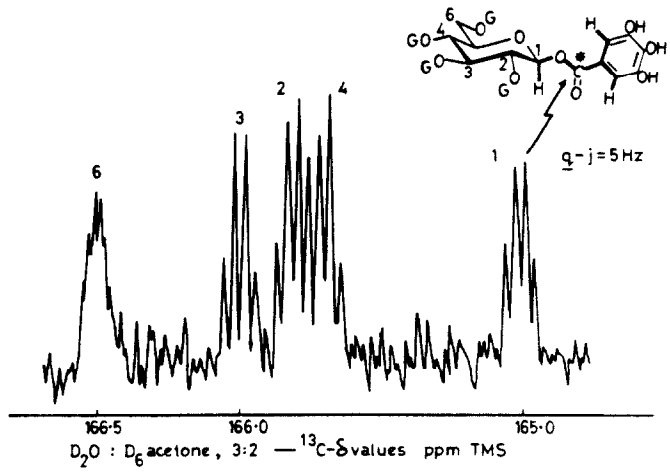



Fig. 12. $^{13}$C NMR–proton correlation, three bond couplings. Carbonyl signals of β-pentagalloyl-D-glucose (**1**).

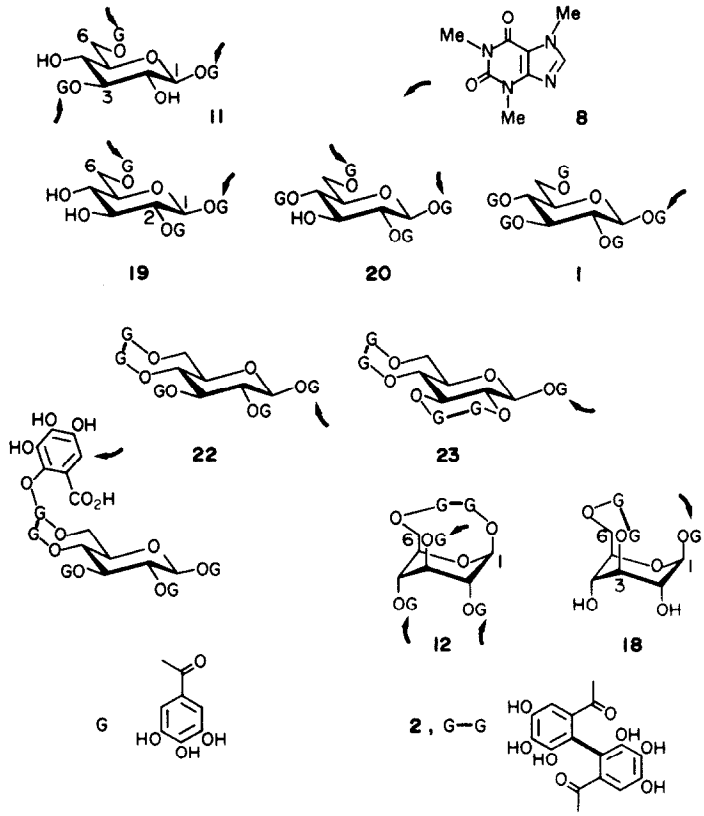



Fig. 13. Galloyl esters–preferential sites (↗) of complexation with caffeine.

cular galloyl ester group on the D-glucopyranose core. Extension of this method to a range of hexahydroxydiphenoyl and related esters (Fig. 2) has permitted complete assignments of structures.

Employing the 'reverse' experimental technique to that described previously (i.e. maintaining the concentration of the polyphenol at a fixed value, increasing the concentration of caffeine and measuring the chemical shift changes $\Delta\delta$ in the polyphenolic substrate) the preferred sites of complexation of caffeine with the polyphenol may be then provisionally identified, Figs 13, 14. Thus in β-penta-*O*-galloyl-D-glucose (**1**) the principal sites for complexation with caffeine appear to be with the galloyl ester groups at C-1 and then C-6. Interestingly the galloyl ester group is overwhelmingly the preferred site of complexation in the flavan-3-ol gallates (**16**, **17**) and, in the context of the earlier discussions, the hexahydroxydiphenoyl ester group predictably has generally a weak affinity for caffeine (**8**), Fig. 13.

Solution studies of the association of peptides with

polyphenols using $^{1}H$ and $^{13}C$ NMR spectroscopy have not, as yet, revealed information of similar detail on the corresponding complexation reactions. In studies with various simple peptides, such as glycylglycine and *N*-acetylglycylglycinamide, little evidence for association with *β*-penta-*O*-galloyl-D-glucose (**1**) can be discerned. One inference from this observation is that, in aqueous media, the peptide and amide groups do not provide, *by*

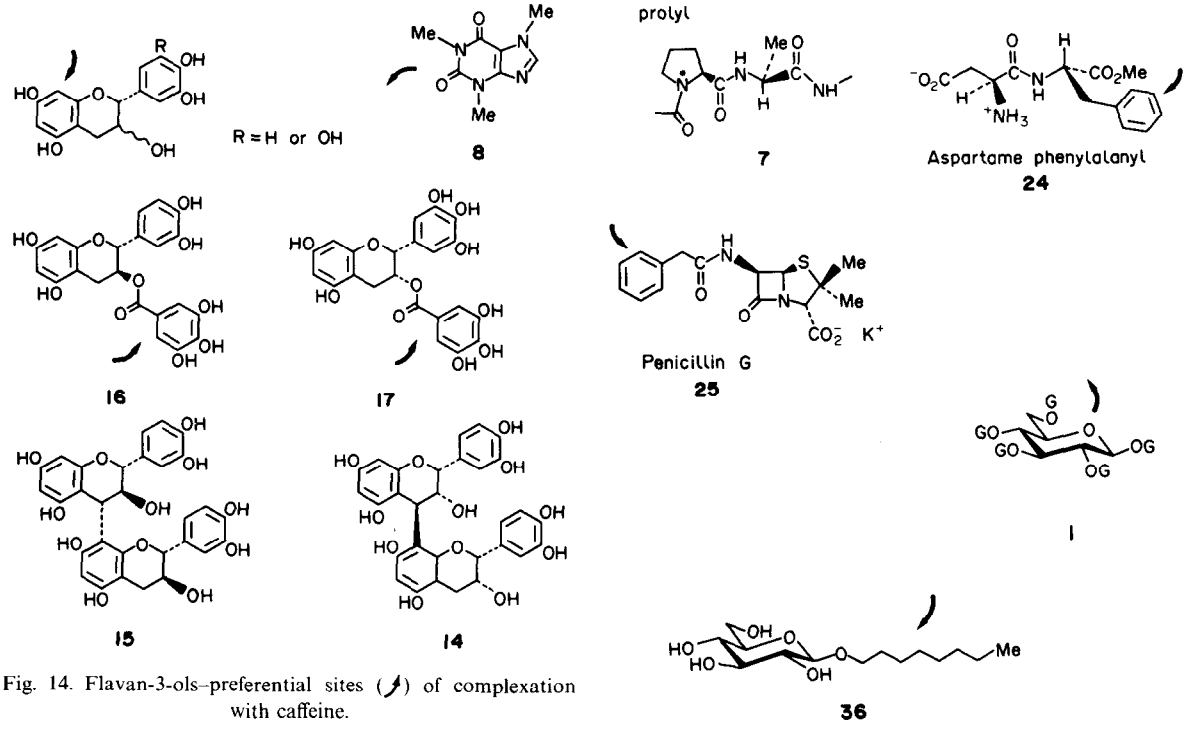
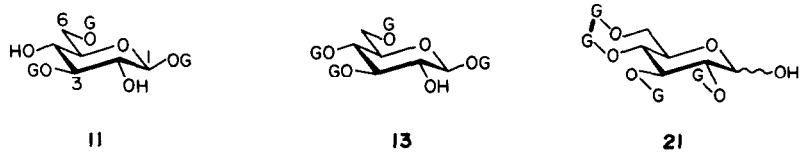
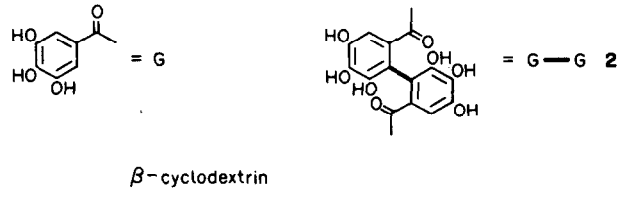
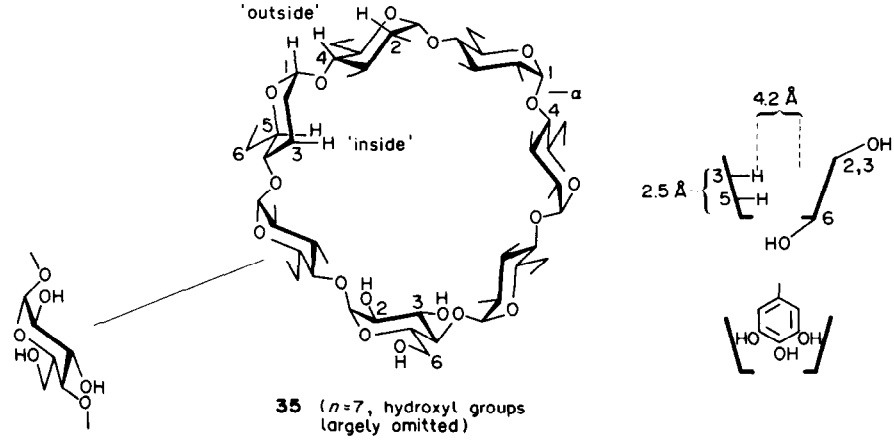



Fig. 14. Flavan-3-ols–preferential sites (↙) of complexation with caffeine.



Arg . Lys . Asp . Val . Tyr
**(26)**

Tyr . Arg
**(27)**

Thre . Lys . Pro . Arg
**(28)**

Tyr . D-Ala . Gly . D-Leu
**(29)**

Ala . Tyr . Met . Asp . Phe – $NH_2$
**(30)**

Gly . Gly . Met
**(31)**

Met . Leu . Phe
**(32)**

Arg . Pro . Pro . Gly . Phe . Ser . Pro . Phe . Arg
**(33)**

Asp . Arg . Val . Tyr . Ile . His . Pro . Phe
**(34)**

*themselves*, sites for the strong binding of polyphenols. Conversely selectivity of complexation of (**1**) with penicillin G (**25**) and aspartame (**24**) occurs in the region of the aromatic rings. However in studies with the peptides (**26**–**34**), β-1,3,6-tri-*O*-galloyl-D-glucose (**11**) and (−)-epigallocatechin gallate (**17**) induced very similar small effects in the $^1$H NMR spectra of the peptides–indicative of the general lack of influence of polyphenol structure on

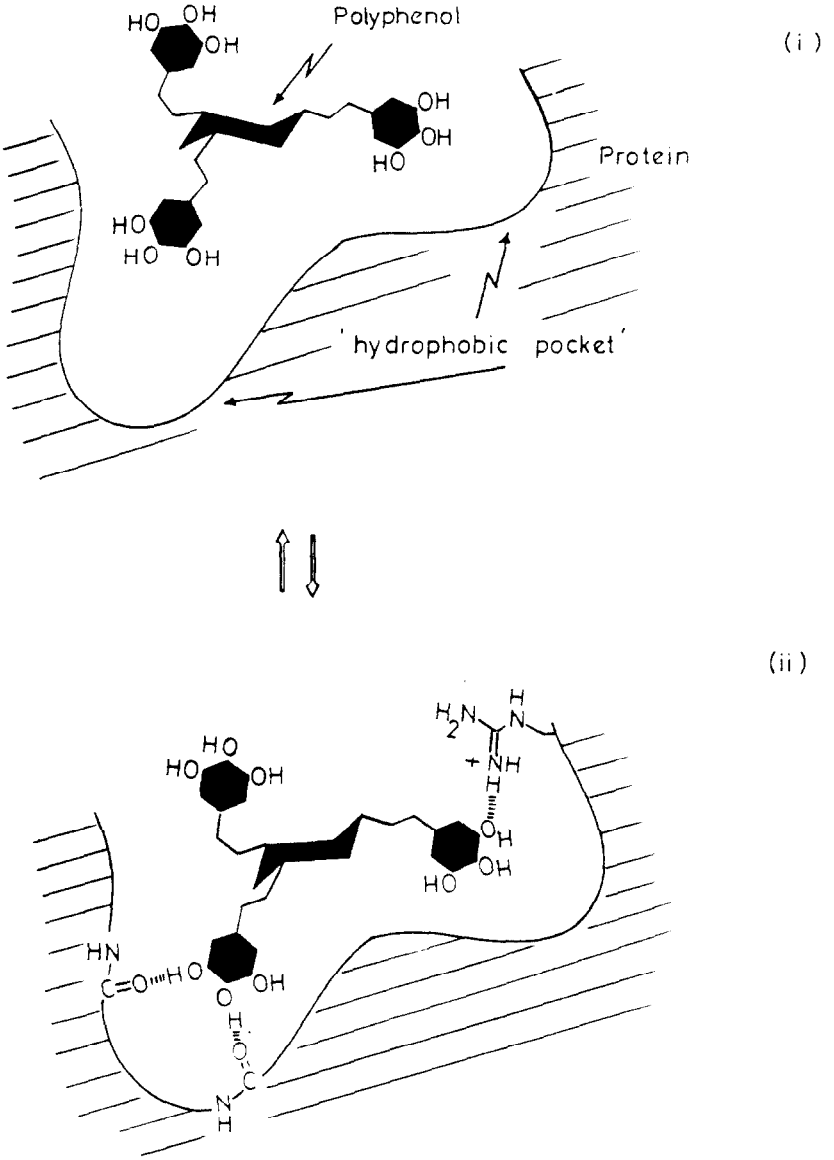



Fig. 15. Polyphenol–protein complexation, (i) docking of the polyphenol, (ii) hydrogen bonding to the protein surface.

complexation (both **11** and **17** have similar values of the association constant with caffeine, Table 1). The chemical shift changes were generally small except in the case of the peptides (**26**, **28**) where addition of the polyphenol produced strong downfield shifts of the proton resonances of the N-terminal amino-acids (L-arginine and L-threonine respectively). Several of the peptides precipitated the polyphenol (**1**) from solution ($\sim 0.05$ M in both substrates) but prior to precipitation no significant effects in the $^1$H NMR spectra were observed, with for example bradykinin (**33**), angiotensin (**34**).

The affinity of the aryl rings of polyphenols for hydrophobic cavities such as are provided by the molecule $\beta$-cyclodextrin (**35**) has been noted earlier [27, 33]. Likewise the non-ionic detergent (**36**) displays a strong affinity for various polyphenols. Association occurs highly selectively in the region of the hydrocarbon chain of (**36**). These thus observations also point to the importance of hydrophobic interactions in the complexation of polyphenols with proteins.

*Polyphenol–protein complexation*

The general picture which is now beginning to emerge is that the association of polyphenols with proteins is largely a surface phenomenon and is possibly a process which takes place in two distinct phases. The first of these is that in which the polyphenol seeks out preferred sites and regions on the protein where its numerous aromatic rings are most readily accommodated by the development of hydrophobic interactions. This selectively can be satisfied in the region of aromatic residues and particularly where prolyl residues in the polypeptide chain, by virtue of their effect on polypetide conformation, bring several such aromatic groups or hydrocarbon side chains into close juxtaposition to form a hydrophobic pocket or environment, (Fig. 15i). Presumably during the formation of this pocket some significant solvent exclusion will occur. This first stage of association is then firmly reinforced by the appropriate deployment of hydrogen bonds between phenolic residues and polar groups (e.g. guanidino, amide and peptide, amino, hydroxyl and carboxyl groups) in the vicinity of the primary site of 'docking' of the polyphenol to the polypeptide surface (Fig. 15ii). Flexibility in both protein and polyphenol is presumably of considerable benefit during this second stage. The net effect is to produce a reversible complex between protein and polyphenol and one which leaves the surface of the protein 'coated' with a much less hydrophilic layer of polyphenol molecules and which leads ultimately to aggregation and to precipitation. The efficacy of polyphenols in this complexation process derives from the fact that they function as polydentate ligands acting through several potential sites (phenolic residues) with the protein. The protein substrate is, likewise, potentially a multi-site acceptor molecule.

If this general picture and model is correct then some structural features which lead to selectivity in the complexation process might be expected in both ligand (polyphenol) and substrate (protein). Some facets of selectivity have been tentatively identified but the field remains open and ripe for further exploitation and enquiry—a feature which Tony Swain would have doubtless approved!

*Acknowledgements*—The authors thank the SERC (C.M.S), AFRC (S.H.G.), Nestec Ltd (R.M., P.N.G.) and the Government of the Peoples Republic of China (Y.C.) for financial support and the Directors of Nestec Ltd (Switzerland) for permission to publish this work.

## REFERENCES

1. Bate-Smith, E. C. and Swain, T. (1953) *Chem. Ind.* 377.
2. Swain, T. (1960) in *Phenolics in Plants in Health and Disease* (Pridham, J. B., ed.), p. 45. Pergamon Press, Oxford.
3. Stafford, H. (1988) *Phytochemistry* **27**, 1.
4. Cairns-Smith, A. G. (1985) *Seven Clues to the Origin of Life*, p. 2. Cambridge University Press, Cambridge.
5. Swain, T. (1977) *Annu. Rev. Plant Physiol.* **28**, 479.
6. Swain, T. (1978) In *Biochemical Aspects of Plant and Animal Co-evolution* (Harborne, J. B., ed.), p. 3. Academic Press, London.
7. Neirenstein, M. (1932) *Incunabula of Tannin Chemistry*. Edward Arnold, London.
8. White, T. (1957) *J. Sci. Food Agric.* **8**, 377.
9. Gustavson, K. H. and Holm, B. (1952) *J. Am. Leather Chem. Ass.* **47**, 700.
10. Haslam, E. (1982). *Fortschr. Chem. Org. Naturstoffe* **41**, 1.
11. McManus, J. P., Davis, K. G., Beart, J. E., Gaffney, S. H., Lilley, T. H. and Haslam, E. (1985) *J. Chem. Soc.* (Perkin Trans. II) 1429.
12. Ozawa, T. Lilley, T. H. and Haslam, E. (1987) *Phytochemistry* **26**, 2937.
13. Oh, H. I., Hoff, J. E., Armstrong, G. S. and Haff, L. A. (1980) *J. Agric. Food Chem.* **28**, 394.
14. Bate-Smith, E. C. (1973) *Phytochemistry* **12**, 903.
15. Hagerman, A. E. and Butler, L. G. (1981) *J. Biol. Chem.* **256**, 4494.
16. Mole, S. and Waterman, P. G. (1987) *Oecologia* **72**, 137.
17. Covington, A. K., Lilley, T. H., Newman, K. E. and Porthouse, G. A. (1973). *J. Chem. Soc. (Faraday Trans. I)*, **69**, 963.
18. Gaffney, S. H., Haslam, E., Lilley, T. H. and Ward, T. R. (1988) *J. Chem. Soc. (Faraday Trans. I)*, (in press).
19. Check, P. J. and Lilley, T. H. (1988). *J. Chem. Soc. (Faraday Trans I)*, (in press).
20. Gupta, R. K. and Haslam, E. (1978) *J. Chem. Soc. (Perkin Trans. I)*, 892.
21. Hagerman, A. E. and Butler, L. G. (1980) *Agric. Food Chem.* **28**, 944, 947.
22. Ellis, S. C. and Pankhurst, K. G. A. (1954) *Disc. Faraday Soc.* **16**, 170.
23. Loomis, W. D. (1974). *Methods Enzymol.* **31**, 528.
24. Goldstein, J. and Swain, T. (1965) *Phytochemistry* **4**, 185.
25. Martin, R., Lilley, T. H., Bailey, N. A., Falshaw, C. P., Haslam, E., Magnolato, D. and Begley, M. J. (1986) *J. Chem. Soc. Chem. Commun.* **105**.
26. Martin, R., Falshaw, C. P., Lilley, T. H., Haslam, E., Magnolato, D. and Begley, M. J. (1987) *Phytochemistry* **26**, 2733.
27. Gaffney, S. H., Martin, R. M., Lilley, T. H., Haslam, E. and Magnolato, D. (1986) *J. Chem. Soc. Chem. Commun.* 107.
28. Wallwork S. C. (1961) *J. Chem. Soc.* 494.
29. Mejbaum-Katzenellenbogen, W., Dobryszychka, W. M., Jaworska, K. and Morawiecka, B. (1959) *Nature* **184**, 1799.
30. Mejbaum-Katzenellenbogen, W. and Dobryszychka, W. M. (1962) *Nature* **193**, 1288.
31. Okuda, T., Mori, K. and Hatano, T. (1985) *Chem. Pharm. Bull.* **33**, 1424.
32. Haslam, E. (1986) *Rec. Adv. Phytochem.* **20**, 163.
33. Lilley, T. H. and Haslam, E. (1988) *CRC Crit. Rev. Food Sci. Nutrition* (in press).