

## Biotransformations in the peptide and carbohydrate fields

D.H.G. Crout,<sup>1</sup> D.A. MacManus,<sup>1</sup> J-M. Ricca,<sup>1</sup> S. Singh,<sup>1</sup> P. Critchley<sup>2</sup> and W.T. Gibson<sup>2</sup>

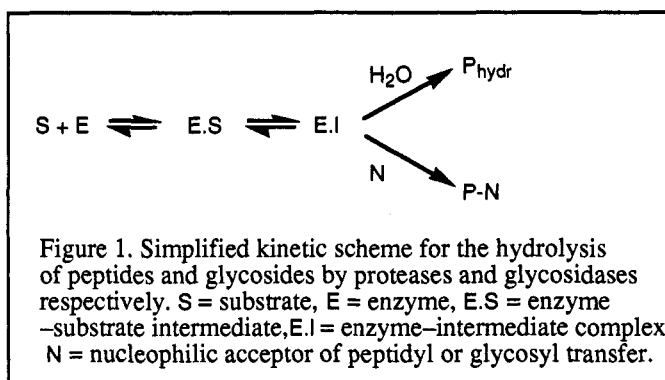
<sup>1</sup> Department of Chemistry, University of Warwick, Coventry CV4 7AL, U.K.

<sup>2</sup> The Unilever Research Laboratories, Sharnbrook, Bedford MK44 1LQ, U.K.

**Abstract** — Proteases and glycosidases are classes of hydrolytic enzymes that are used in biotransformations to catalyse the synthesis of peptides, and of glycosides and oligosaccharides, respectively, by a reversal of their normal mode of action. The practical application of both methods shares several common features, namely the shift of rate-determining step through the use of activated substrates (aminoacyl and glycosyl donors respectively), the need to reduce or eliminate competing hydrolysis in the aminoacyl and glycosyl transfer steps, and the need to minimise hydrolytic cleavage of the products. Control of these factors is illustrated by a method for peptide synthesis using  $\alpha$ -chymotrypsin suspended in dichloromethane containing low concentrations of water and a method for disaccharide synthesis using an *N*-acetylgalactosaminidase from *Aspergillus oryzae*. Molecular modelling of the oxyanion intermediate has been found to explain a number of features of reactions catalysed by  $\alpha$ -chymotrypsin and to predict others that were the subject of experimental test. The concept of “anomeric control” was used to modulate the regioselectivity of *N*-acetylgalactosaminyl transfer catalysed by the *N*-acetylgalactosaminidase.

### INTRODUCTION

Major developments in biotransformations have been the introduction of hydrolytic enzymes, proteases<sup>1</sup> and glycosidases<sup>2</sup>, for the synthesis of peptides, and glycosides and oligosaccharides, respectively. Proteases and glycosidases are related in that they are both classes of enzyme that under physiological conditions catalyse hydrolytic cleavage, of peptide and glycosidic bonds respectively. Kinetically, both follow the same general course, shown in its simplest form in Fig. 1. In peptide hydrolysis catalysed by serine proteases, E.I is the acyl-enzyme intermediate. In oligosaccharide synthesis, E.I is the glycosyl-enzyme intermediate. Clearly, for these systems to be effective in synthesis, the rate of interception of E.I by the nucleophile, N, should be as large as possible relative to the rate of interception by water. This is the first requirement. The second relates to the product P-N which, by the same mechanism whereby it is formed, is susceptible to hydrolytic cleavage. The rate of this secondary hydrolysis must be low relative to the rate of formation of P-N. Additionally, or alternatively, methods are required whereby the rate of secondary hydrolysis can be minimised. The most direct way of achieving this is by lowering the thermodynamic activity of water in the system.



In practice, formation of the enzyme intermediate complex E.I from natural substrates is relatively slow. For this reason, ‘activated’ substrates are used in biotransformations. In peptide synthesis this means in practice, using amino acid or peptide esters as substrates. In glycoside or oligosaccharide synthesis, the normal glycosidic or oligosaccharide substrates are often replaced by activated substrates such as *o*- or *p*-nitrophenylglycosides, or, in some systems, by glycosyl fluorides, as will be discussed below. A con-

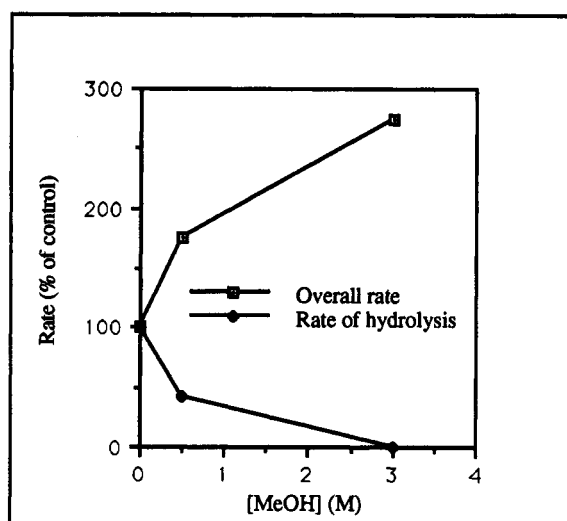


Figure 2. Rate of disappearance of *o*-nitrophenyl  $\beta$ -D-galactoside compared with the rate of hydrolysis catalysed by the  $\beta$ -galactosidase of *Escherichia coli* in aqueous methanol.<sup>4</sup>

amides or peptides in reactions catalysed by  $\alpha$ -chymotrypsin shifts the rate determining step from formation of the acyl-enzyme intermediate to capture of this by the nucleophile.<sup>5</sup> Thus again, an apparent "activation" of the enzyme by nucleophiles can be observed (Table).

#### PEPTIDE SYNTHESIS CATALYSED BY $\alpha$ -CHYMOTRYPSIN IN A LOW-WATER ENVIRONMENT

Serine proteases such as  $\alpha$ -chymotrypsin catalyse peptide cleavage *via* the well-established catalytic triad of Asp102, His57, Ser195. Peptide synthesis by reversal of the hydrolytic sequence has been much studied.<sup>1</sup> However, it is difficult to control in practice because of the concomitant hydrolysis of the product, which, by the principle of microscopic reversibility, must also be a substrate for hydrolytic cleavage by the enzyme as noted above. However, an observation by Dastoli *et al*<sup>6</sup> that a suspension of  $\alpha$ -chymotrypsin in dichloromethane was capable of catalysing peptide bond hydrolysis provided that water was present to the extent of 0.25%, prompted an investigation of this system for peptide synthesis. In this system, the thermodynamic activity of water is so low that hydrolysis of the product is not significant. Accordingly, peptides can be prepared in high yield with the associated ad-

sequence of using activated substrates is that the rate determining step in the catalytic mechanism may change. Thus in hydrolytic reactions catalysed by  $\beta$ -galactosidase, capture of the glycosyl-enzyme intermediate E.1 (Fig. 1) becomes the rate determining step.<sup>3</sup> Nucleophiles such as methanol may capture the glycosyl-enzyme intermediate faster than water, resulting in an apparent (but misnamed) "activation" of the enzyme (Fig.2).

Use of activated esters rather than amino acid

Table. Release of *p*-nitrophenol from *p*-nitrophenyl acetate catalysed by  $\alpha$ -chymotrypsin in the presence of various alcohols.

Alcohol (0.68 M)	Relative rate of <i>p</i> -nitrophenol release
None	1.0
Methanol	2.0
Ethanol	1.0
1-Propanol	2.0
2-Propanol	0.55
1-Butanol	6.5
2-Methyl-1-propanol	2.8
2-Butanol	0.76
1,1-Dimethylethanol	0.3
1-Pentanol <sup>a</sup>	12.0
1-Hexanol <sup>b</sup>	15.0

<sup>a</sup> 0.048 M

<sup>b</sup> 0.048 M, but substrate incompletely dissolved.

C.E. McDonald and A.K. Balls, *J. Biol. Chem.*, 1956, 221, 993.

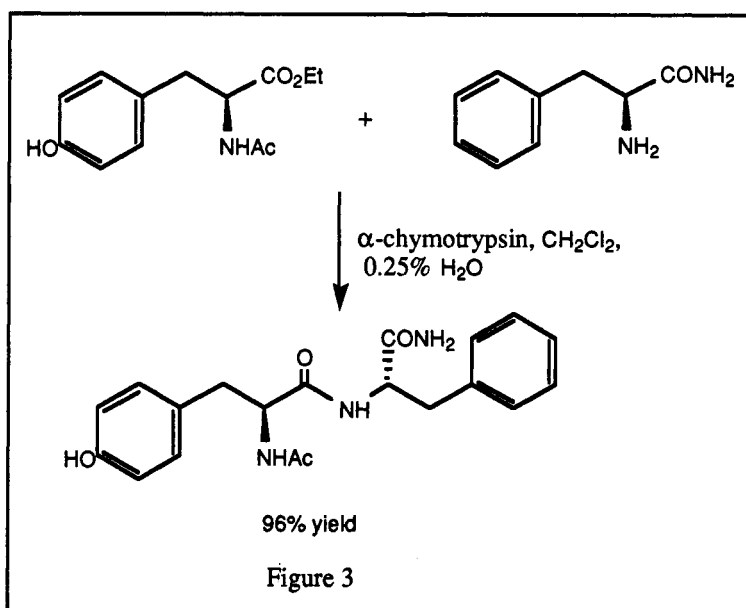
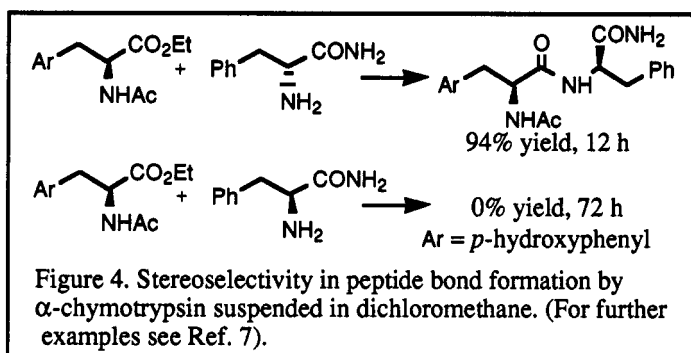


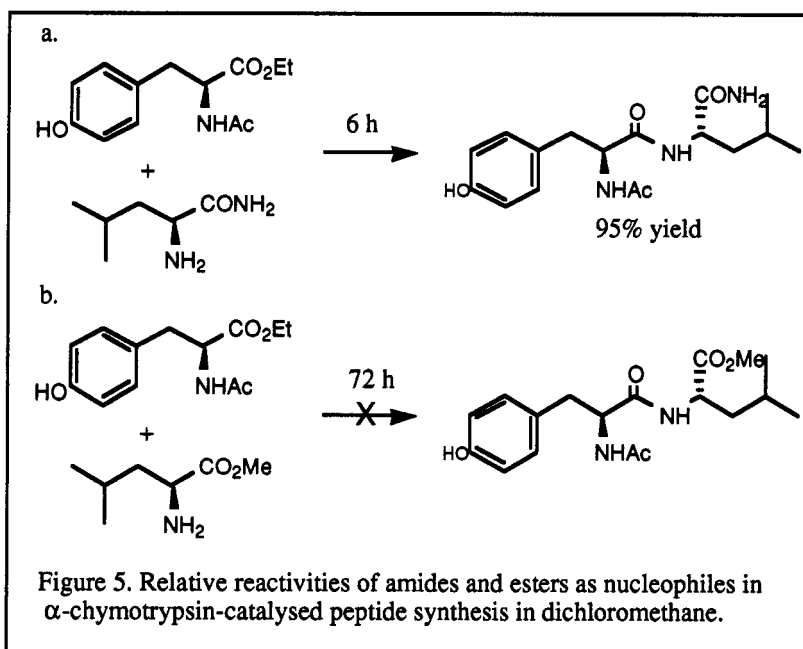
Figure 3



vantage of minimal (i.e. undetectable by n.m.r. or t.l.c.) racemisation. An example is shown in Fig. 3.

The system can be used for the synthesis of numerous peptides.<sup>7</sup> However, as with many other systems,<sup>8</sup> the environmental change from water to dichloromethane suspension is accompanied by significant changes in catalytic properties. Some of these are illustrated in Fig. 4. This shows that although the system retains its strict L-specificity with

regard to the acyl donor, in dichloromethane it is much more accommodating with respect to the configuration of the nucleophile. According to the nomenclature of Schechter and Berger<sup>9</sup>, the nucleophile (the P<sub>1</sub> residue) is bound at the S<sub>1</sub>' subsite. In water, the rate of reaction with D-amino acid derivatives is very slow, whereas in the dichloromethane suspension, peptide formation is relatively rapid, implying a significant change in conformation, affecting the S<sub>1</sub>' subsite. At the same time, it was found that the size of the residue that could be accommodated at the S<sub>1</sub> subsite (encompassing the hydrophobic pocket that confers on  $\alpha$ -chymotrypsin its specificity for aromatic amino acids) is much more restricted than when the enzyme operates in an aqueous environment. It has been suggested that in media of low dielectric constant, coulombic interactions between charged or polar groups become stronger and that this leads to greater rigidity in the protein structure.<sup>10</sup> Whether it is appropriate to use a concept based on macroscopic systems at the protein-molecular level is debatable. In the present case, although the altered properties in the organic suspension cannot be used to support the argument for greater rigidity, it is clear that in this environment, marked conformational changes are induced. To what extent the altered properties are also attributable to penetration of solvent molecules into the interior of the protein cannot be decided with certainty, although various lines of evidence, including the observation that certain solvents can inhibit competitively the enzymatic reaction, make solvent penetration a distinct possibility.<sup>11</sup>

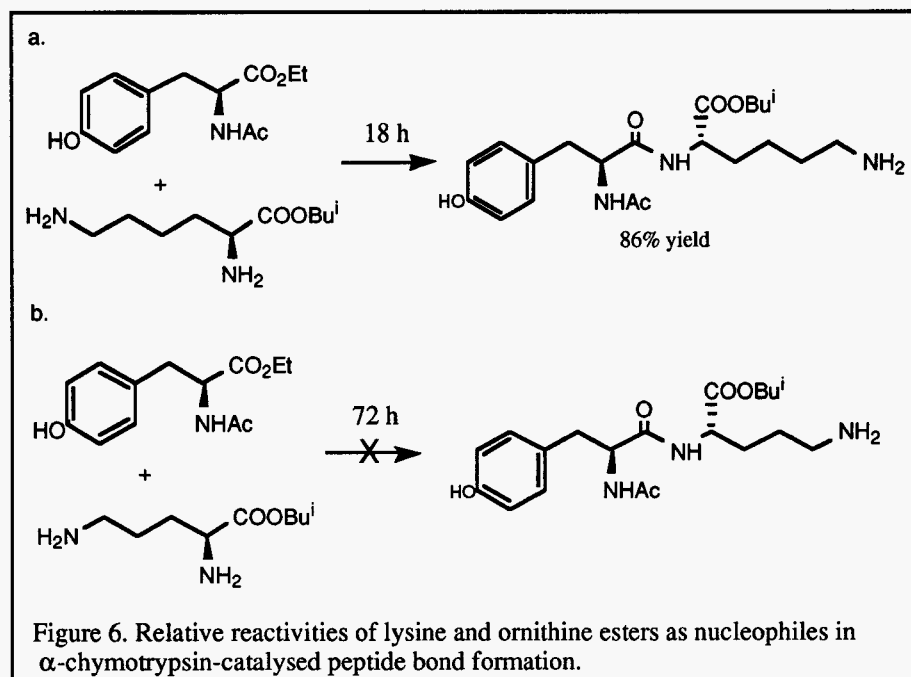


Molecular modelling was used to investigate certain properties of the  $\alpha$ -chymotrypsin system, not particularly related to the organic environment. For example, it was noted that amino acid amides were much better substrates than corresponding esters (Fig. 5). Similar observations have been made for comparable substrates in an aqueous environment. Modelling of the oxyanion intermediate (Fig. 7) revealed the formation of a strong hydrogen bond between the amide amino group of the P<sub>1</sub>' component and the oxygen atom of the backbone carbonyl group of Phe41. The possibility of forming such a hydrogen bond is clearly lost on re-

placing the amide group by an ester function.

An initially unexpected observation was that lysine isobutyl ester functioned as an excellent nucleophile in the acylation reaction (Fig. 6a). Modelling of the oxyanion intermediate revealed that the lysine side chain

could be accommodated in a strainless, fully antiperiplanar conformation in a groove leading up to Cys58, to the backbone carbonyl group of which the  $\epsilon$ -amino group formed a strong hydrogen bond (Fig. 7). This observation is particularly interesting in the light of the conclusion, based on a statistical analysis of proteolysis data, that  $\alpha$ -chymotrypsin displays a specificity for peptide bonds that have a lysine or arginine residue in the P'<sub>1</sub> position.<sup>12</sup>



The modelling studies suggested that if the side chain were shortened, as in ornithine derivatives, the hydrogen bond to Cys58 could not be formed and an inferior substrate would result. In the event (Fig. 6b), the

isobutyl ester of ornithine proved to be quite ineffectual as a substrate. The modelling studies also provided an explanation for the P<sub>1</sub> specificity for L-amino acids. Modelling of the oxyanion formed from *N*-acetyl-D-tyrosylalanine methylamide showed that although this substrate could be accommodated, it was with a penalty with respect to torsional strain in the D-tyrosyl backbone of some 8 kcal mol<sup>-1</sup>, which would be prohibitive as far as participation in the catalytic reaction is concerned. The success of this approach to the modelling of  $\alpha$ -chymotrypsin-catalysed reactions (which encompassed a considerably wider range of substrates than those described here)<sup>7</sup> suggests that a similar approach to the modelling of lipase-catalysed reactions would be equally successful, using the X-ray crystallographic data already in existence for three such enzymes.<sup>13</sup>

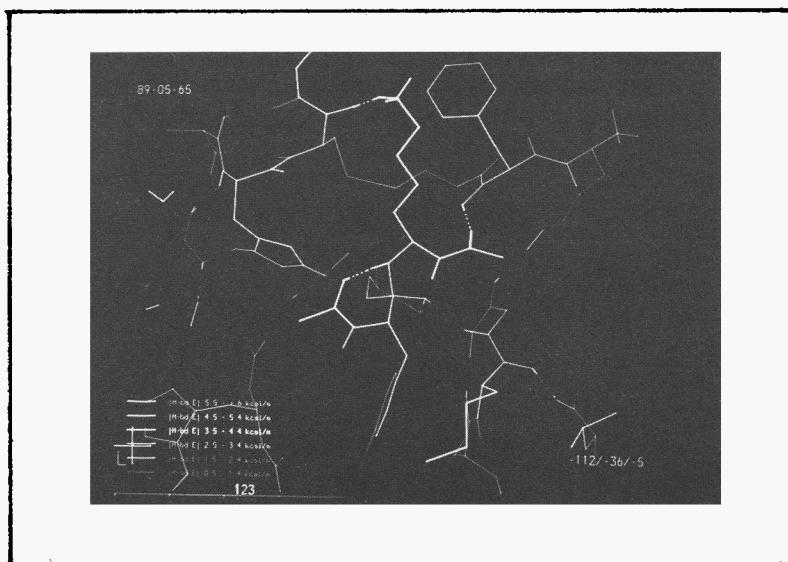


Figure 7. Molecular modelling of the oxyanion formed following binding of *N*-acetyllysylalaninemethylamide to  $\alpha$ -chymotrypsin. The figure illustrates both the hydrogen bond formed between the  $\epsilon$ -amino group of the lysyl component with the backbone carbonyl group of Cys58 and the hydrogen bond formed between the alanyl residue bound at the P'<sub>1</sub> subsite with the backbone carbonyl group of Phe41.

*Please note:* A colour version of Figure 7 is printed on prelim page iv.

## GLYCOSIDASE-CATALYSED SYNTHESIS OF OLIGOSACCHARIDES

The importance of carbohydrates in molecular recognition processes, many of which are of profound medical importance, has stimulated renewed interest in oligosaccharide synthesis. By comparison with peptide and oligonucleotide synthesis, the current state of oligosaccharide synthesis is not equal to the demands of biological and medical sciences for bioactive oligosaccharides, the simplest of which cannot be synthesised without considerable effort. This is not at all a reflection on the inadequacy of research in oligosaccharide synthesis. Rather, it is a reflection of the great complexity of oligosaccharide structures. The development of new, efficient methods for oligosaccharide synthesis represents the major challenge to organic chemistry over the next decade. It is in this context that biotransformations have much to offer. Glycosidase-catalysed glycoside formation in particular, provides an efficient method by which glycosidic bonds can be generated with no requirement for protection-deprotection sequences and with total control of configuration at the newly formed anomeric centre. Glycosyl donors such as natural substrates (e.g. lactose with  $\beta$ -galactosidase) or activated substrates such as *o*- and *p*-nitrophenyl glycosides, can be used to generate a glycosyl-enzyme intermediate that can be trapped by an added nucleophile in competition with water. When such activated substrates are used with, for example,  $\beta$ -galactosidase, nucleophilic attack on the glycosyl enzyme intermediate (cf Fig. 1) becomes rate determining and an apparent "activation" of the enzyme results, as noted above (Fig. 2). Although water is present at 55 M and the nucleophile N at, typically, 100 mM, capture of the intermediate by the nucleophile competes effectively with capture by water. With  $\beta$ -galactosidase, this has been attributed to the hydrophobic character of the nucleophile binding region at the active site.<sup>3</sup> Yields based on glycosyl donor (the nucleophilic acceptor is usually present in excess) are typically 10–30% (but may be much higher, as shown below). However, such yields compare favourably with overall yields in purely chemical procedures and the enzymatic reactions are considerably more efficient with respect to the expenditure of time *per* coupling step. The penalty for this efficiency is that glycosyl transfer is not regioselective. More than one transfer product may be obtained and our understanding of the nature of the catalytic sites of glycosidases at present is such that the regioselectivity of transfer cannot be predicted except on the basis of empirical observations. The trade-off for the efficiency and simplicity of glycosidase-catalysed oligosaccharide synthesis is therefore the necessity to apply separation procedures for the isolation of the target product.

Although not predictable, regioselectivity of glycosyl transfer can be modulated, as demonstrated by Nilsson,<sup>14</sup> by control of the configuration at the anomeric centre in the acceptor sugar, a process for which the term "anomeric control" has been coined.<sup>14</sup> This we applied to the synthesis of 1,4-linked GalNAc–GluNAc disaccharides using an *N*-acetylgalactosaminidase from *Aspergillus oryzae*. With *p*-nitrophenyl *N*-acetyl- $\beta$ -galactosaminide as glycosyl donor and methyl  $\beta$ -glucopyranoside as acceptor, corresponding 1,3- and 1,4-linked disaccharides were produced, with the former preponderating (Fig. 8). However, with methyl  $\alpha$ -glucopyranoside as acceptor, the 1,4-linked product predominated and the minor product was found to be the 1,6-linked disaccharide (Fig. 9).<sup>15</sup> These re-

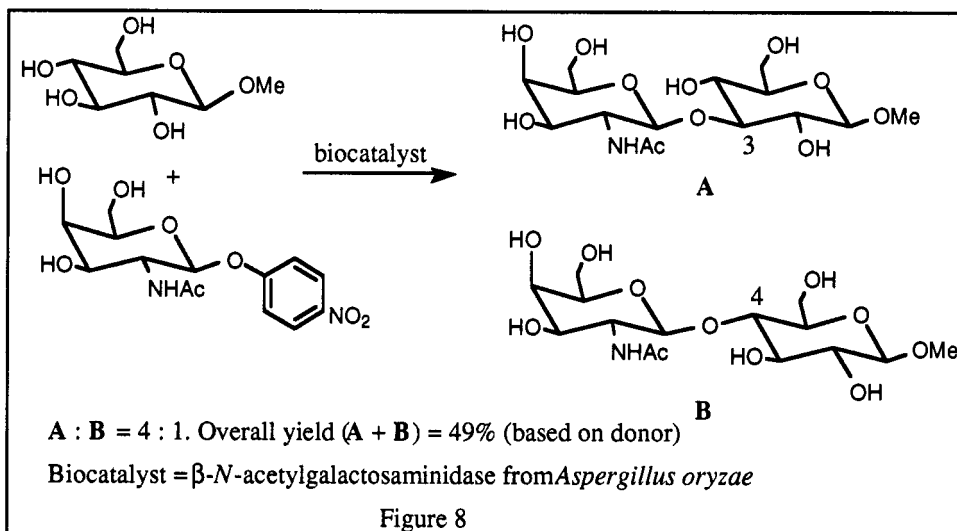


Figure 8

sults suggested that the  $\alpha$ -configuration in the acceptor favoured transfer to the 4-position. It was predictable, therefore, that synthesis of the target compound GalNAc( $\beta$ 1,4)GluNAc would best be carried out using methyl  $\alpha$ -*N*-acetylglucosaminide as acceptor. In the event, transfer was found to be highly efficient

using this substrate (Fig. 10) and highly selective with respect to transfer to the 4-position.<sup>16</sup>

These observations clearly show that glycosidase-catalysed oligosaccharide synthesis has much to offer. When integrated with chemical methods, a powerful combination results that increases the efficiency with which bioactive oligosaccharides can be prepared.

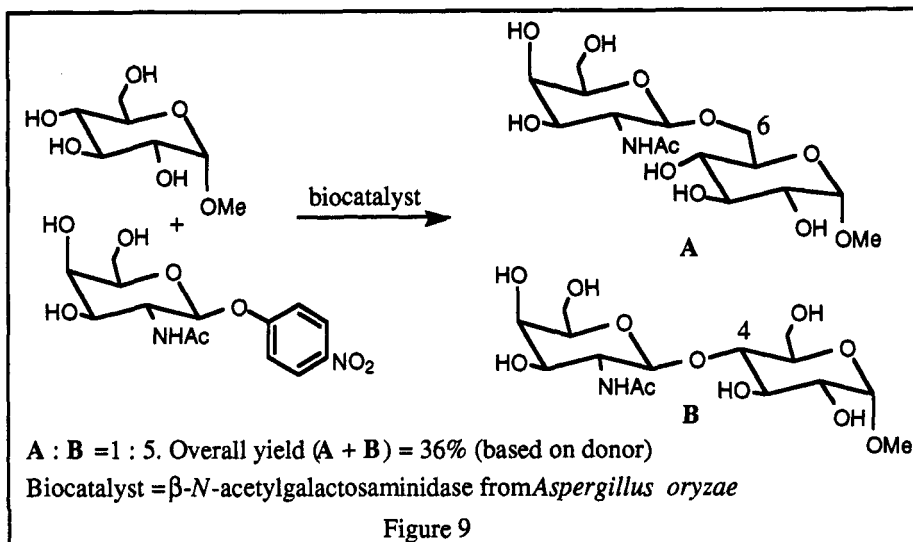


Figure 9

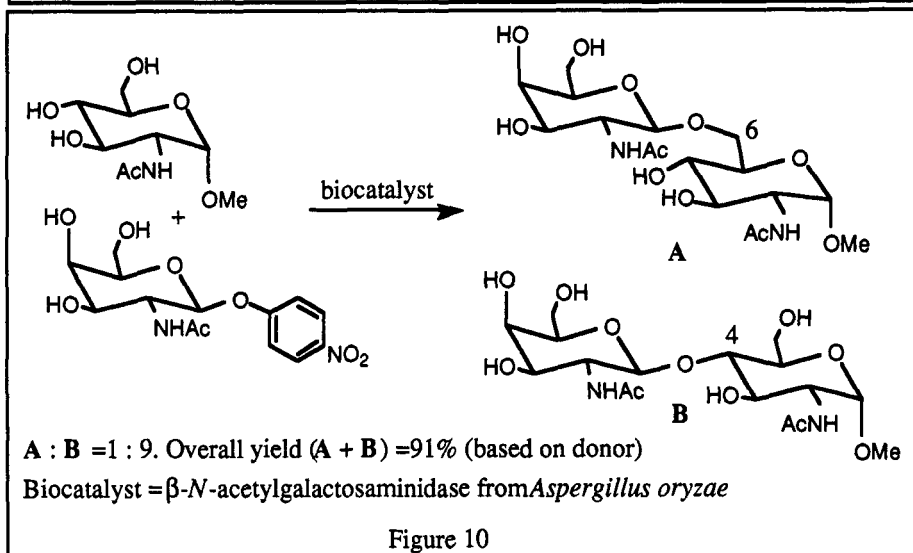


Figure 10

### Acknowledgements

The molecular modelling studies were carried out in collaboration with Dr K. Müller, Hoffmann LaRoche Ltd., Basel. Financial support was provided by the Science and Engineering Research Council, Unilever Research and Rhone Poulenc.

### REFERENCES

- W. Kullmann, "Enzymatic Peptide Synthesis", CRC Press Inc., Boca Raton, Florida, 1987.
- D.G. Drueckhammer, W.J. Hennen, R.L. Pedersen, C.F. Barbas, C.M. Gautheron, T. Krach and C-H. Wong, *Synthesis*, 1991, 499-525.
- K. Wallenfells and R. Weil., in "The Enzymes", 3rd edition, Vol.7 (P.D. Boyer, ed), Academic Press, 1972, pp 617-663.
- R.J. Heath, B.E.P. Swoboda and D.H.G. Crout, unpublished data.
- J. Fatsrez and A.R. Fersht, *Biochemistry*, 1973, 12, 2025-2034.
- F.R. Dastoli and N.A. Musta, *Arch. Biochem. Biophys.*, 1966, 115, 44-47.
- J-M. Ricca and D.H.G. Crout, *J. Chem. Soc., Perkin Trans 1*, 1989, 2126-2127.
- A. Zaks and M. Klibanov, *J. Biol. Chem.*, 1988, 263, 3194-3201.
- I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun.*, 1967, 27, 157-162.
- P.A. Fitzpatrick and A.M. Klibanov, *J. Am. Chem. Soc.*, 1991, 113, 3166-3171.
- G.E. Clement and M.L. Bender, *Biochemistry*, 1963, 2, 836-843; J.L. Miles, E. Morey, F. Crain, S. Gross, J. San Julian and W.J. Canady, *J. Biol. Chem.*, 1962, 237, 1319-1322.
- B. Keil, *Protein Sequences Data Anal.*, 1987, 1, 13-20.
- L. Brady *et al.*, *Nature*, 1990, 343, 767-770; F.K. Winkler, A. D'Arcy and W. Hunziker, *Nature*, 1990, 343, 771-774.
- K.G.I. Nilsson, *Trends Biotechnol.*, 1988, 6, 256-264
- D.H.G. Crout, O.W. Howarth, S. Singh, B.E.P. Swoboda, P. Critchley and W.T. Gibson, *J. Chem. Soc., Chem. Commun.*, 1991, 1550-1551.
- D.H.G. Crout, S. Singh, B.E.P. Swoboda, P. Critchley and W.T. Gibson, *J. Chem. Soc., Chem. Commun.*, in the press.