Subsite Mapping of Enzymes. Studies on *Bacillus subtilis* Amylase*

John A. Thoma,† Charles Brothers,‡ and Joseph Spradlin‡

ABSTRACT: The bond ruptured by a lytic enzyme hydrolyzing a biopolymer reflects the positioning of the substrate on the specificity site. For a given substrate, selective positioning and bond rupture gives rise to a discrete product distribution characteristic of a given substrate. Since the product variation reflects microscopic enzyme–substrate dissociation constants, a quantitative analysis of product pattern can give a useful subsite profile of the specificity center. A representational model based upon these concepts is presented and used for the interpretation of the cleavage patterns of *Bacillus subtilis* amylase acting on α-methylmaltodextrin oligosaccharides. Substrates range in size from 3 to 12 glucose residues. Analysis of the data leads to the conclusion that the enzyme specificity site spans nine residues and that the subsites exhibit a wide range of monomer affinity. The predictive value of the model is also demonstrated.

The interactions of depolymerizing enzymes with their substrates and substrate analogs has led various authors to conclude that the enzyme specificity sites are composed of tandem subsites geometrically complementary to the monomer units of the polymer substrate (Abramowitz *et al.*, 1967; Chipman *et al.*, 1967; Cuatrecasas *et al.*, 1968; Kazuyuki and Oka, 1968; Robyt and French, 1963, 1970; Schechter and Berger, 1966, 1967). Various authors have speculated that the specificity sites of the different enzymes may span as few as three or as many as nine monomer residues. The most carefully examined enzyme, hen egg-white lysozyme, is reported to interact with six monomer units (Blake *et al.*, 1967; Chipman *et al.*, 1968). Some insight into the number, substrate affinity, and heterogeneity of the subsites can be obtained by measuring inhibition constants with substrate analogs (Chipman *et al.*, 1967; Cuatrecasas *et al.*, 1968; Karush, 1962; Rupley, 1967; Rupley and Gates, 1967).

The product patterns characteristic of depolymerizing enzymes are also capable of yielding similar information. The chain length and product concentrations depend upon how the substrate is positioned across the catalytic center (Robyt and French, 1963; Greenwood and Milne, 1968a) and the relative concentration of the different possible complexes or binding modes (Scheme I). The population density of these modes will be dictated by the relative affinities of the subsites for the monomer residues. The population distribution of the complexes can be expected to shift systematically with the substrate size. Since the product patterns are governed by the binding mode population, it is apparent that the action pattern of an enzyme conveys quantitative information about the nature and number of subsites. The purpose of this paper is to illustrate how product distributions are valuable for outlining some of the quantitative features of an enzyme's binding site. The rationale for this procedure is discussed in the next section.

Model

The model is based upon the supposition that the specificity site on the enzyme is composed of a number of tandem subsites which are structurally complementary to a monomer residue. As a first approximation it is assumed that the unitary free energy (Karush, 1962) of subsite association with the monomer residue is independent of the presence or absence of contiguous interactions. In this instance, the relative population density for a specific binding mode of an *n*-mer will be proportional to the sum of subsite energy lost when *n* contiguous sites are occupied.

Scheme I illustrates how the relative population density of the various productive binding modes govern the initial distribution of products.

For example, hydrolysis of the complexes VIII 4 and IX 4 leads to Glc_3Me[1] and Glc_2Me, respectively, so that the ratio, $Glc_3Me:Glc_2Me$, measures the apparent ratio, [$\widetilde{\text{IX}}$ 4]:[$\widetilde{\text{VIII}}$ 4]. Since both complexes extend across sites VI and VII the partitioning of substrate between modes VIII 4 and IX 4 should be exclusively controlled by the difference in apparent unitary free energy (Karush, 1962) for the association of a glucosyl residue with sites V and IX. The following derivation shows how this apparent free-energy difference can be computed from the product distribution ratios.

The Michaelis equation for the model represented by Scheme I can be written as

$$E + A_n \underset{k_{-1,i,n}}{\overset{k_{1,i,n}}{\rightleftharpoons}} E_iA_n \xrightarrow{k_{2,i,n}} E + P_{i,m} \qquad (1)$$

* From the University of Arkansas, Department of Chemistry, Fayetteville, Arkansas 72701. *Received October 30, 1969.* This work was supported by Agricultural Research Service, U. S. Department of Agriculture, Grant No. 12-14-100-9116(71), administered by the Northern Utilization Research and Development Division; by grants from the U. S. Department of Public Health (AM 11666); and by Corn Industries Research Foundation.

† To whom inquires should be addressed.

‡ Present address: Union Starch, Granite City, Ill.

[1] Abbreviation used is Glc_nMe, α-methyl maltooligosaccharide of *n* units.

where i corresponds to the position occupied by the "reducing" glucosyl residue, n and m represent, respectively, substrate and product chain length and E and A represent, respectively, free enzyme and substrate concentrations.

The conservation expression is given by

$$[E_0] = [E] + \sum_i [E_i A_n] \tag{2}$$

and the rate of formation of a specific product is

$$\frac{d[P_{i,m}]}{dt} = k_{2,i,n}[E_i A_n] \tag{3}$$

Combining eq 1, 2, and 3 the Michaelis equation is found to be

$$\frac{d[P_{i,m}]}{dt[E_0]} = \frac{k_{2,i,n}[A]_n/K_{i,n}}{1 + [A_n]\sum_i (1/K_{i,n})} \tag{4}$$

The overall initial velocity, $\bar{v}_0$, is given as

$$\frac{\bar{v}_0}{[E_0]} = \sum_j \frac{d[P_{i,m}]}{dt[E_0]} = \frac{\sum_j k_{2,j,n}[A]_n/K_{j,n}}{1 + [A_n]\sum_i (1/K_{i,n})} \tag{5}$$

where $j = i$ for all productive complexes. In the denominator of eq 5, i is summed over all productive complexes and those nonproductive complexes which interfere with productive binding. $K_{i,n}$ is interpreted as a microscopic Michaelis constant characteristic of binding in the ith mode and takes the form, $k_{2,i,n}/(k_{-1,i,n} + k_{2,i,n})$. The product distribution ratio for a pair of adjacent modes is

$$\frac{[\dot{P}_{i,m}]}{[\dot{P}_{j,m+1}]} = \frac{k_{2,i,n}/K_{i,n}}{k_{2,j,n}/K_{j,n}} = \frac{\tilde{K}_{i,n}}{\tilde{K}_{j,n}} \tag{6}$$

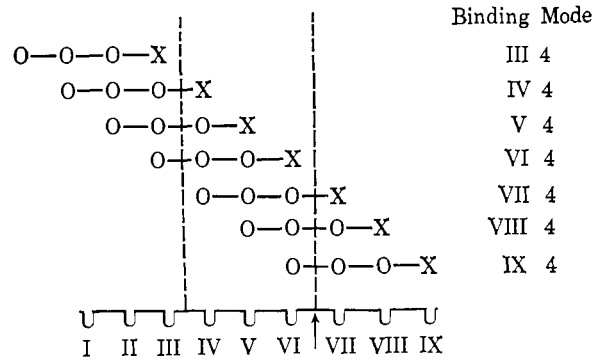
where $\sim$ implies an apparent function and $j = i + 1$ and $[\dot{P}]$ is $d[P]/dt$. Thus $\tilde{K}_{i,n}$ and $\tilde{K}_{j,n}$ may be interpreted as apparent association constant for the ith and jth complexes in so far as the microscopic Michaelis constant approaches the true dissociation constant and the catalytic coefficients are the same. Equation 6 predicts that product ratios may be a function of the rate coefficients characteristic of hydrolysis of a given binding mode as well as the microscopic Michaelis constant associated with that mode of binding. It is possible then that product ratios and subsite interaction energies will be influenced by rate phenomena. In the limiting case when $k_{2,i} = k_{2,j}$ and $k_{-1} \gg k_2$, the product distribution ratio will correspond to the real population of complexes of the i and j binding modes. However, in any case, the apparent population of contiguous modes can be computed from the observed product ratios.

Since

$$-RT \ln \tilde{K}_{i,n} = \sum_{i}^{i-n+1} \tilde{\Delta G}_{u,i} \tag{7}$$

where i may range in value from 1 to the number of subsites $+ (n - 1)$ and $\tilde{\Delta G}_{u,i}$ is the apparent unitary free energy of interaction of the ith site.

SCHEME I[a]



[a] Illustration showing various possible binding modes of α-methyl maltotetraoside on *Bacillus subtilis* amylase. ↑, position of catalytic site; —∪—, monomer subsite; ——, α-1→4 bond; X, α-methyl D-glucopyranoside "residue"; O, glucopyranoside "residue." Binding modes are designated by Roman numeral specifying subsite occupied by the "reducing" end of the substrate while the Arabic numeral designates substrate size. Modes VII 4 to IX 4 are productive complexes while modes I 4 to VI 4 are nonproductive complexes. Dotted lines are visual aids.

Combining eq 6 and 7 gives

$$RT \ln \frac{[\dot{P}_{i,m}]}{[\dot{P}_{j,m+1}]} = RT \ln \frac{\tilde{K}_{i,n}}{\tilde{K}_{j,n}} = \sum_{j}^{j-n+1} \tilde{\Delta G}_{u,j} - \sum_{i}^{i-n+1} \tilde{\Delta G}_{u,i} \tag{8}$$

Which simplifies to

$$RT \ln \frac{[\dot{P}_{i,m}]}{[\dot{P}_{j,m+1}]} = \tilde{\Delta G}_{u,j} - \tilde{\Delta G}_{u,j-n} \tag{9}$$

The significance of 9 can be appreciated by reference to Scheme I. This equation predicts that when $n = 4$, the apparent subsite association affinities of positions V and IX, govern the substrate distribution between binding modes VIII 4 and IX 4. For example, if $\tilde{\Delta G}_{u,V}$ is more negative than $\tilde{\Delta G}_{u,IX}$, the concentration of complex VIII 4 will dominate that of complex IX 4 and Glc_2Me will dominate Glc_3Me in the reaction digest. The product ratio Glc_2Me:Glc_3Me substituted into eq 9 gives the differences in apparent unitary free energy for subsite association between sites IX and V.

It should be noted that $[\dot{P}_{i,m}]$:$[\dot{P}_{j,m+1}]$ is only independent of substrate concentration when $[A_n] \to 0$. At high substrate concentrations the binding of a second substrate in a nonproductive mode may seriously alter the population densities of the productive modes. For example, a second substrate whose reducing end occupied site V will inhibit production of Glc_1Me and Glc_2Me.

Unfortunately no information can be deduced about the relative affinities of sites VI and VII, the subsites adjacent to the catalytic amino acids. Since all the productive complexes occupy both these positions their interaction with substrates will not shift population densities of productive modes.

The apparent interaction energy for site VIII may be evaluated by comparing the product ratio, $Glc_1Me:Glc_2Me$, for an $(n + 1)$-mer to the product ratio, $Glc_2Me:Glc_3Me$, for an n-mer. As an example, the latter ratio for $n = 5$ compares subsite IV and IX while the former ratio for $n = 4$ compares subsite IV to VIII. Subtracting the two appropriate free energies yields the interaction energy of VIII with respect to IX, the reference subsite. The apparent unitary free energies of subsites with indexes of X or larger and those to the left of subsite I may be computed in an analogous fashion.

In addition to estimating relative subsite–monomer energetics, product patterns may also aid in delineating the span of the enzyme's specificity center. When the substrate becomes large enough to extend beyond the binding center the unitary free energy of all overlapping monomer units will be zero. Since these groups do not interact with the enzyme they should not influence the action pattern and the ratio, $[\dot{P}_m]:[\dot{P}_{m+1}]$, should become constant. As an illustration, the model predicts that $Glc_2Me:Glc_3Me$ will become constant for $n \geq 9$ if there are no chain-length-dependent conformational effects or other factors regulating product frequencies.

For convenience we have arbitrarily set $\widetilde{\Delta G}_{u,IX} = 0$ designating this site as the reference. If $Glc_2Me:Glc_3Me$ is substituted into eq 9 for any n-mer then $\widetilde{\Delta G}_{u,IX-n}$ is computed directly. Thus the ratio, $Glc_2Me:Glc_3Me$, resulting from Glc_4Me hydrolysis gives $\widetilde{\Delta G}_{u,V}$.

Experimental Section

Cyclodextrin Transglucosidase. This enzyme was isolated from a culture of *Bacillus macerans* kindly supplied by J. Robyt. The *B. macerans* was cultured at 37° for 3–4 days in a medium recommended by J. Robyt. The medium contained 10 g of finely diced potatoes, 2 g of $CaCO_3$, 1.5 g of yeast extract, 1.0 g of Bacto tryptone, 0.5 g of urea, 0.1 g of barbituric acid, and 100 ml of distilled water. During the growth period the medium was vigorously aerated.

The enzyme was prepared by centrifuging the culture medium, and concentrating to one-third the original volume by ultrafiltration with a Diaflow membrane, UM-1 (Aminco) with a 10^4 mol wt cutoff. The concentrate which possessed about 50 units of activity (Thoma *et al.*, 1965) was placed in a 55–58° bath and held at that temperature for 10 min. The solution was centrifuged and allowed to incubate overnight at room temperature with a small amount of DNase and RNase. Aliquots were then prepared by passage though a Sephadex G-50 (fine) column previously equilibrated with buffer.

Substrates. A homologous series of nonreducing saccharides was prepared by incubating α-methyl D-glucoside with highly purified cyclohexaamylose (French *et al.*, 1963) with cyclodextrin transglucosidase at pH 4.8 in 0.01 M sodium acetate buffer. The digest contained [α-^{14}C]methyl D-glucoside (15.4 mg, 100 μCi; Calbiochem), cyclohexaamylose (60 mg), and enzyme (5 units; 1.1 ml) and was held at 37° for 24 hr. The enzyme was destroyed by gently heating the digest to boiling.

The solution was then streaked across the short end of a sheet of Whatman No. 3MM paper (14 × 17 in.). The sheet was placed in a tank equilibrated with the irrigating solvent of H_2O–EtOH (95%)–CH_3NO_2 (26:36:38, v/v) (Thoma and French, 1957). A single ascent was capable of resolving the sugars of chain length 1–9. Four ascents were employed for resolution of the sugars of chain length 10–12. The sugars were detected by autoradiographic techniques using a 14 × 17 in. Kodak No-Screen X-Ray Film employing an 18–24-hr exposure.

Using the developed X-ray film as a guide, the individual oligosaccharides were cut from the chromatogram and washed from the strips with distilled water. The compounds are generally removed in the first 10–15 drops. The compounds were stored in the freezer until further use.

Action of Bacillus subtilis Amylase. Crystalline *B. subtilis* lot 20086 was purchased from Nagase and Co. Ltd., Osaka, Japan (U. S. distributor, Enzyme Development Corp., N. Y.).

As a check for a contaminating glucanosyltransferase the enzyme was chromatographed according to the procedure of Pazur and Okada (1968) with a 150-μl gradient. The enzyme was dissolved at pH 5.5 in 0.04 M sodium acetate buffer. Hydrolysis was initiated by adding 100 ml of enzyme to 50 μl of substrate. The final substrate concentrations were $5 \pm 2.5 \times 10^{-4}$ M. At appropriate time intervals 10-μl aliquots were withdrawn and pipetted onto Whatman No. 3MM paper and dried with a hair dryer. The hot air drying was adequate to inactivate the enzyme. In alternate experiments the 10-μl digest was added to an equal volume of concentrated NH_4OH on a paraffin plate to inactivate the enzyme. The action pattern was unaffected by either of these treatments. Spots of the original oligosaccharide mixture were placed at regular intervals across the chromatogram to serve as markers. Additional samples were also spotted at zero time to furnish the corrections necessary for self-radioautolysis.

At the termination of the reaction the paper sheets were chromatographed as described above and the spots located by radioautography. The radioactive compounds located directly by darkening on the X-ray film were cut from the chromatogram folded singly and placed in scintillation vials. A scintillator solution (5 ml) was added which was prepared by dissolving 4 g of 2,5-diphenylazole and 50 mg of 1,4-bis[2-(5phenyloxazolyl)]benzene in 1 l. of reagent grade toluene. Counting was done with a Beckman LS-100 liquid scintillation system and direct data read-out module.

Results

The crystalline enzyme used in these studies was homogeneous by ultracentrifugal analysis, disc electrophoresis, and gel permeation chromatography on Sephadex G-200 (L. Li and J. A. Thoma, 1965, unpublished results).

In order to check for the presence of transferase activity in our crystalline enzyme preparation, the enzyme was submitted to ion-exchange chromatography (Pazur and Okada, 1968) which cleanly separates a transferase component from the hydrolytic component in a crude preparation. Only a single protein peak appeared in the elution profile superimposable with the hydrolytic activity. About 0.2–0.5% contamination would have been detected.

Glc_6Me and Glc_4Me were hydrolyzed with the chromatographed enzyme at a substrate concentration two orders of magnitude lower than that originally used. Within experimental error no significant change in action pattern was discovered (see Table I). On the basis of these observations it can safely be concluded that if any contaminating transferase

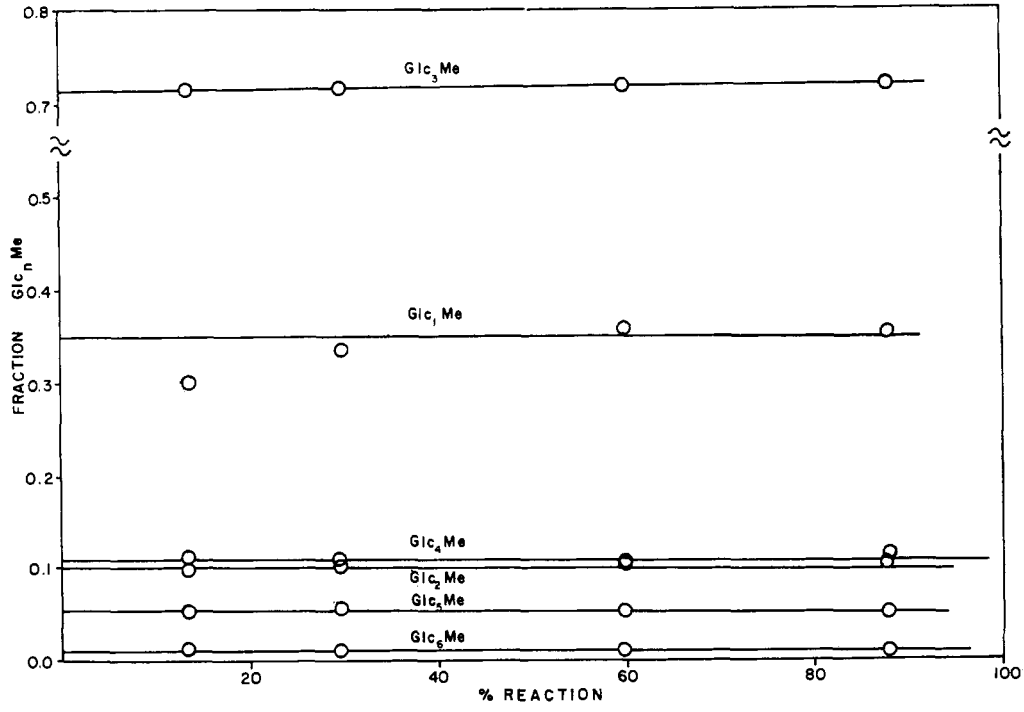


FIGURE 1: Typical time course of hydrolysis of substrate by *B. subtilis* amylase. Substrate, Glc_1Me; product curves, are labeled individually. Glc_1Me plotted on a tenfold-expanded axis. The precision of Glc_1Me data is unusually good for such a low yield.

TABLE I: Experimental and Theoretical Effect of Chain Length on Product Distribution of Radiolabeled Substrate.[a]

Substrate Chain Length	Per Cent Product Produced[b,c]													
	Chain Length of Products													
	1		2		3		4		5		6		7	
	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd	Obsd	Calcd
3	72		28											
		66		34										
	74[e]		26											
4	22		66		12									
		21		68		10.8								
	21[f]		70		9									
5	5	4	47	48	48	48								
6	37		13		50		Trace[d]							
		37.6		13.7		47.5		1.2						
	34[f]		16		50		1							
7	23	27	69	65	8.2	7.3	Trace[d]	0.6						
8	3	2.7	56	56	41	41	Trace[d]	1.2						
9	1	0.65	18	13	79	84	2	1.5	Trace[d]	0.03				
10	1[d]	0.57	11	12	76	75	9	12	2	1.4				
11	0.35	0.5	10	11	72	67	11	10.5	6	10.5	1	1.2		
12	*d*	0.46	10	9.5	63	60	10	9.5	9	9.5	6	9.5	3	1.1

[a] These products are the radiolabeled α-methyl maltooligosaccharides corresponding to the "reducing" end of the parent substrate after hydrolysis. The fate of the maltooligosaccharides corresponding to the "nonreducing" end of the parent substrate was not followed. [b] Computed values based on column 4 in Table II. [c] On the basis of several duplicate experiments it is estimated that the relative (not absolute) precision of the initial product is $\pm 5\%$ for the predominant components and 10–50% for the minor components. [d] Too much experimental error to accurately evaluate. [e] The values reported in this row were run at substrate concentration $\sim 1 \times 10^{-6}$ M. [f] The values reported in these rows were run at substrate concentration $\sim 5 \times 10^{-6}$ M.

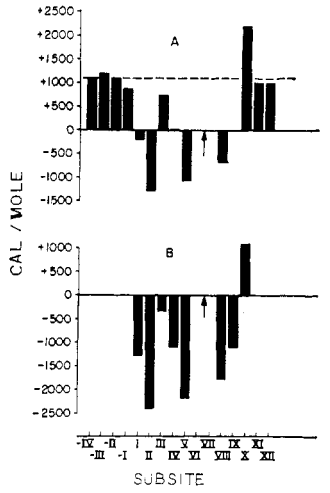


FIGURE 2: Histogram of subsite interaction energy with anhydroglucopyranoside on *B. subtilis* amylase. (A) Site IX used as thermodynamic reference; (B) imaginary sites used as thermodynamic reference. Data plotted from Table II.

is present it does not manifest itself. Since transfer or condensation reactions catalyzed by *B. subtilis* amylase are second order in substrate, the failure to observe any significant change in action patterns with 100-fold range of substrate concentration (Robyt and French, 1969) is interpreted to mean that this enzyme acts exclusively as a hydrolylase under our reaction conditions.

The substrate levels employed for this investigation were $\sim 5 \times 10^{-4}$ M or $\bar{K}_m \pm 0.5\ \bar{K}_m$ of Glc_6Me–$Glc_{12}Me$ (L. Li and J. A. Thoma, unpublished observations). However experiments to test for transferase or condensation were run at substrate levels of 5×10^{-6} to 1×10^{-6} M. The product patterns were examined as a function of chain length from $n = 3$ to 12 inclusive. Initial distributions were measured by extrapolating plots of fractions of a specified product *vs.* the extent of the reaction to zero extent of reaction. The fraction of n-mer is defined as $[Glc_nMe]/\Sigma[Glc_nMe]$. Hydrolysis was generally terminated after 90–99% reaction of the original substrate. Plots of fraction of Glc_nMe extent of the reaction invariably yielded straight lines with essentially zero slope. Typical data are presented in Figure 1. The reliability of the data is inversly proportional to the per cent yield of a product. In the terminal stages of these reactions (>95%) slight curvature of the plots for large substrates was sometimes noted and is probably the result of secondary attack on products by the enzyme.

The product patterns as a function of substrate chain length from $n = 3$ to 12 are recorded in Table I. Although Glc_2Me and Glc_3Me predominate as products for all sugars hydrolyzed, larger oligosaccharides appear in increasing quantities as the chain length of the substrate is increased (Table I). We conclude that *B. subtilis* amylase is a true endoenzyme with a high preference to cleave near the "reducing" end of the substrate.

The relative affinity of the subsites for a monomer group computed from eq 9 are recorded in column 3 of Table II and Figure 2A using either subsite IX as the reference and in column 4 of Table II and Figure 2B using the average of

TABLE II: Apparent Unitary Free Energy of Subsites.

Substrate Size	Subsite	Thermodynamic Reference Site: Subsite IX[a]	Thermodynamic Reference Site: Imaginary Subsites[b]
12	−IV	+1100	
11	−III	+1200	
10	−II	+1100	
9	−I	+ 880	
8	I	− 190	−1290
7	II	−1300	−2400
6	III	+ 740	− 360
5	IV	+ 10	−1100
4	V	−1100	−2200
	VI, VII	Experimentally inaccessible, see text	
c	VIII	− 700	−1800
c	IV	Reference	−1100
c	X	+2200	+1100
c	XI	+1000	
c	XII	+1000	

[a] Computed as $RT \ln (Glc_3Me/Glc_2Me)_n$, where n is substrate size. [b] Value of −1100 was subtracted from sites I to X and the imaginary sites set to zero. [c] From Table III.

subsites −I, −IV, and X–XII as the reference point (see Discussion).

Histogram 2B was used to compute the theoretical product distribution *vs.* substrate length. Theoretical and experimental values for end-labeled substrates are compared in Table I. The agreement is generally within the bounds of experimental error.

Table III is a check for internal consistency of the model. If the rate coefficients for hydrolysis of the different complexes do not vary widely with chain length it would be expected that the apparent unitary free energy of a given site would be independent of substrate size and mode of binding. The data in Table III are consistent with the assumption that subsite interactions tend to be independent of substrate size.

Discussion

It is now generally recognized that the action pattern of an amylase is characteristic of its origin (Greenwood and Milne, 1968a,b) and the experimental conditions of hydrolysis (Robyt and French, 1967). The quantitative explanation of these observations in terms of molecular events is only partially complete. Two proposals, repetitive attack and a series of productive enzyme–substrate complexes (Scheme I), have been proffered as the basic explanation of these phenomena.

The former concept was devised by French to account for the production of small oligosaccharides in the early phases of amylose hydrolysis (Robyt and French, 1967, and references therein). He envisions that during a single enzyme substrate encounter several bond ruptures may occur before the complex dissociates and that the number of hydrolytic events is

TABLE III: Apparent Unitary Free Energy of Subsite VIII and Imaginary Subsites in Calories per Mole and Substrate Sizes Used for Measurement.

Substrates	$-\tilde{\Delta G}_{VIII}$[a]	Substrates	$+\tilde{\Delta G}_{X}$[b]	Substrates	$+\tilde{\Delta G}_{XI}$[c]	Substrates	$+\tilde{\Delta G}_{XII}$[d]
3 and 4	520						
4 and 5	670						
5 and 6	600						
5 and 7	750						
7 and 8	840						
8 and 9	860	8 and 9	2000				
9 and 10	570	9 and 10	2200	8 and 10	700		
10 and 11	*e*	10 and 11	2300	9 and 11	1200	8 and 11	900
	820	11 and 12	2300	10 and 12	1200	9 and 12	1100
Average standard deviation	700 ± 130		2200 ± 140		1000 ± 300		1000 ± 140

[a] Computed as $-RT[\ln (Glc_2Me/Glc_3Me)_n - \ln (Glc_1Me/Glc_2Me)_{n-1}]$. [b] Computed as $RT[\ln (Glc_3Me/Glc_4Me)_n - \ln (Glc_2Me/Glc_3Me)_{n-1}]$. [c] Computed as $RT[\ln (Glc_4Me/Glc_5Me)_n - \ln (Glc_2Me/Glc_3Me)_{n-2}]$. [d] Computed as $RT[\ln (Glc_5Me/Glc_6Me)_n - \ln (Glc_2Me/Glc_3Me)_{n-3}]$. [e] This value excluded in average because of excessive scatter in Glc_1Me data from $Glc_{10}Me$ hydrolysis.

proportional to k_{cat}/k_{dissoc}. With this reaction scheme the enzyme appears to nibble the end of a large substrate producing small oligosaccharides after the first random attack.

Although the number of ruptures per effective encounter is basically a kinetic phenomenon, the size of the products released will probably be influenced by monomer–subsite interactions. After the first cleavage rearrangement of the enzyme–substrate complex will yield a series of complexes whose relative concentrations partially depend upon the subsite–monomer interaction energies (Scheme I). In the absence of repetitive attack product concentrations will normally be expected to parallel the concentration of the enzyme–substrate complexes from which they arise. Product concentrations will then be governed by the microscopic Michaelis constants characteristic of each binding mode. In this manuscript we demonstrate that the product patterns yield an apparent thermodynamic outline of the binding site.

For this purpose the end-labeled α-methylmaltodextrins were chosen as substrates. They have two advantages. First, the Michaelis parameters for nonreducing substrates are more easily evaluated by standard reducing sugar assays than for reducing substrates. This will facilitate comparison of measured Michaelis constants with those calculated from product patterns. Second, only end-labeled substrates unambiguously and directly furnish relative rates of bond hydrolysis. For example cleavage of mode VII 3 yields radioactive Glc_1Me and maltose while rupture of mode VIII 3 yields radioactive Glc_2Me and glucose. The relative amounts of the radiolabeled products is a direct measure of the effective rates of bond rupture. By way of contrast, hydrolysis of unlabeled maltotriose yields equimolar quantities of glucose and maltose regardless of the relative frequencies of bond cleavage and yield no information about the population of productive complexes.

The action pattern of *B. subtilis* amylase reported here is essentially the same as that reported in the more qualitative studies by Robyt and French (1963) on the reducing maltooligosaccharides except for Glc_6Me. Although they employed slightly different reaction conditions we can offer no satisfying explanation for the divergence in results of Glc_6Me hydrolysis. It may be pertinent though that Welker and Campbell (1967) report some variations in action patterns of the amylases from different strains of *B. subtilis*.

In order to ascribe the observed action pattern of an enzyme to heterogeneity of the subsites it is necessary to exclude other factors as contributing to the action pattern. The action pattern of depolymerizing enzymes will be influenced by a wide range of phenomena. These include the number and heterogeneity of the subsites, differences in the hydrolysis rates of the various productive modes of a given substrate, differences in substrate structure including branching and monomer heterogeneity, repetitive attack, conformationally dependent chain-length effects and transferase activity catalyzed by the hydrolyase or by contaminating enzymes. High levels of substrates and products may stimulate reversion or condensation reactions or cause shifts in the population of the productive modes of the complexes by forming ternary complexes. Other experimental variables such as pH or temperature may also induce shifts in enzyme–substrate populations.

Repetitive attack or multiple cleavages per encounter of an enzyme with substrate (Robyt and French, 1967) occur when $k_2 \geq k_{-1}$. This factor can be excluded as an influence on the action pattern of *B. subtilis* amylase because plots of blue values (iodine stain) *vs.* reducing value is independent of pH (L. Li and J. A. Thoma, unpublished observations). The slope of this plot is regarded as an index of repetitive attack (Robyt and French, 1967) and should depend upon pH when $k_2 \geq k_{-1}$. Since the degree of multiple attack drops to zero at pH values distant from the pH optimum (k_2 becomes $\ll k_{-1}$) this complicating feature of the action pattern can be eliminated for any enzyme by running the reactions at a suitable pH.

Chain-length effects can be eliminated as a controlling factor because a considerable literature exists to indicate that the conformation of the maltodextrins is probably independent of chain length (for pertinent references, see Foster, 1965, and Erlander *et al.*, 1968).

The observation that the product ratios are independent of concentration over the range 1–0.01 $\bar{K}_m$ rules out reversions, condensations, transfer reactions, or substrate concentration effects on the populations of productive modes as influencing factors in the action of the amylase (Robyt and French, 1969). Furthermore, the enzyme appears to be homogeneous by a number of physiochemical techniques (L. Li and J. A. Thoma, unpublished observations).

Since our substrates are homologous linear polymers and other variables have been excluded, the effect of chain length on product variation will be interpreted in terms of subsite heterogeneity and variability in rate coefficients. Since the effective concentration of a given mode is the quotient of a rate coefficient and a microscopic Michaelis constant (see Model) rate phenomena may be disguised in apparent subsite association constants. However, if the rate coefficients for hydrolysis are the same, this factor will cancel when subsite interactions are evaluated according to eq 9. If the rates are unequal then the apparent free energies will be influenced by the factor $RT \ln (k_{2,i}/k_{2,j})$.

The product patterns recorded in Table I offer a means to check for this effect and for the internal consistency of the model. If either $k_{2,i}$, the microscopic hydrolytic coefficient, is the same for all complexes or if the rate of formation of a given length product is independent of substrate size then the unitary free energy of subsite VIII and the imaginary subsites X–XII will appear to be independent of substrate size. On the other hand, if the microscopic hydrolytic coefficients, $k_{2,i}$, change markedly from one complex to another, computed subsite interactions will radically fluctuate as a function of substrate length. When $k_{2,i}$ is constant, the apparent unitary free energy at subsite i, $\widetilde{\Delta G}_{u,i}$, will probably reflect the real monomer–subsite interaction energy and the true population densities of the different modes of binding. When $k_{2,i}$ is variable $\widetilde{\Delta G}_{u,i}$ gives only an effective or apparent measure of population densities because $\widetilde{\Delta G}_{u,i}$ is a combination of a thermodynamic and kinetic term. The method for calculating the free energy of interaction for subsites VIII, X–XII is outlined in the Model section and the results of these computations are listed in Table III. The computed subsite energies using different pairs of sugars are constant within experimental error and consistent with the proposal that to a first approximation k_2 is independent either of the binding mode or less stringently of the product size. The internal consistency of the representational model is also validated.

The data for the binding energies listed in Table II (column 3) and plotted in Figure 2A reveal a very interesting feature of the active site of *B. subtilis* amylase. The free energies of association of subsites −I to −IV and XI and XII are all approximately 1.1 kcal larger than the reference site, IX. The obvious interpretation of these data is that residues occupying these "sites" actually extend beyond the enzyme. When these imaginary sites are used as the reference ($\widetilde{\Delta G}_u = 0$) then subsites I to IX must be assigned a unitary free energy of 1.1 kcal more negative than those assigned in Table II (column 3) and Figure 2A. The revised assignments are shown in Table II (column 4) and Figure 2B. This histogram suggests that topologically the binding site spans nine monomer units. If the site interacts with an extended oligosaccharide chain, it would span approximately 45 Å. Figure 2B also reveals a marked degree of heterogeneity among the subsites. Similar observations have been observed for hen egg-white lysozyme (Blake *et al.*, 1967) and have been proposed for other enzymes based on more qualitative investigations (Abramowitz *et al.*, 1967; Robyt and French, 1963; Schechter and Berger, 1967).

The peculiar action pattern of *B. subtilis* amylase is readily rationalized in terms of the subsite histogram in Figure 2B. The marked preference for attack near the reducing end is caused by the positive unitary energy of introducing a monomer residue into site X. The unfavorable energetics is most likely the consequence of some steric impediment to binding. This obstruction tends to enhance the population of complexes in which the whole substrate molecule lies completely to the left of subsite X. Hydrolysis of this population then furnishes large amounts of Glc_1Me, Glc_2Me, and Glc_3Me. However as the substrate size becomes large with respect to the binding site the number of internal segments which can completely occupy the site (I–X) increases. The concentration of these complexes is proportional to chain length (Thoma, 1966); so that the frequency of internal rupture increases with substrate size and the relative proportions of Glc_1Me, Glc_2Me, and Glc_3Me will progressively decrease. Eventually with very long polymers the initial attack will appear to be random. This trend is already perceptible in Table I. This phenomenon is undoubtedly one factor which causes shifts from an apparently random attack on a substrate of long chain length to a highly specific attack on substrates of short chain length (Greenwood and Milne, 1968b). A continual shift in action pattern will also occur during the hydrolysis of any large substrate as the average chain length decreases, and end effects begin to dominate the attack specificity.

If the model is an accurate representation of reality and if the subsite assignments are correct then the model should be capable of predicting the action pattern as a function of chain length. In order to test the predictive value of the model, a program was written to compute the relative concentrations of the productive complexes from the subsite association energies. Relative product yields were then set equal to the relative concentrations of the productive complexes for each substrate. Thus the fraction of Glc_3Me from Glc_4Me hydrolysis was equated to the fraction (see Scheme I)

$$Glc_3Me = \frac{[\text{mode IX 4}]}{[\text{mode IX 4}] + [\text{mode VIII 4}] + [\text{mode VII 4}]} \quad (10)$$

Replacement of [mode IX 4] in the numerator by [mode VIII 4] and [mode VII 4] gives the fraction of Glc_2Me and Glc_1Me, respectively.

Similar computations were performed for each substrate (n = 3–12) and the results recorded in Table I. The input for this program is the histogram in Figure 2B and the assumption was made that sites assigned negative indexes and XI or greater are imaginary, *i.e.*, that $\Delta G_u = 0$. The agreement between the computed and experimental data is a striking illustration of the practical merits of the representational model. In principle the computer program can predict product distributions accurately for even larger substrates as well

as the time course of the digestion of a large substrate. In order to compute the time course of hydrolysis for a large substrate it is only necessary to know the $\tilde{K}_m$ as a function of substrate size. Investigations are currently in progress to experimentally test the proposal. A more extensive description of the program will be published later.

On the basis of this investigation we conclude that *B. subtilis* amylase has a binding site that topologically spans nine glucose residues, that the catalytic center is located between subsites VI and VII, that a steric impediment at subsite X profoundly affects the action pattern of the enzyme, and that the subsite map generated from product distribution patterns of small substrates quantitatively predicts the action pattern on larger oligosaccharides.

Acknowledgments

We are grateful to Allen Jennings who provided valuable technical assistance and Dr. J. Robyt who furnished the *B. macerans* culture and growth conditions.

References

Abramowitz, N., Schechter, I., and Berger, A. (1967), *Biochem. Biophys. Res. Commun. 29*, 862.

Blake, C. C. F., Mair, G. A., North, A. C. J., Phillips, D. C., and Sarma, V. R. (1967), *Proc. Roy. Soc. (London) B167*, 365.

Chipman, D. M., Grisaro, V., and Sharon, N. (1967), *J. Biol. Chem. 242*, 4388.

Cuatrecasas, P., Wilchek, M., and Anfinsen, C. B. (1968), *Science 162*, 1491.

Erlander, S., Purvinas, R. M., and Griffin, H. L. (1968), *Cereal Chem. 45*, 140.

Foster, J. (1965), *in* Starch: Chemistry and Technology, Vol. I, Whistler, R., and Paschall, E., Ed., New York, N. Y., Academic, p 349.

French, D., Pulley, A. O., and Whelan, W. J. (1963), *Staerke 10*, 349.

Greenwood, C. T., and Milne, E. A. (1968a), *Staerke 20*, 139.

Greenwood, C. T., and Milne, E. A. (1968b), *Advan. Carbohydrate Chem. 23*, 281.

Karush, F. (1962), *Advan. Immunol. 2*, 1.

Kazuyuki, M., and Oka, T. (1968), *Biochem. Biophys. Res. Commun. 30*, 625.

Pazur, J., and Okada, S. (1968), *J. Biol. Chem. 243*, 4732.

Robyt, J., and French, D. (1963), *Arch. Biochem. Biophys. 100*, 451.

Robyt, J., and French, D. (1967), *Arch. Biochem. Biophys. 122*, 8.

Robyt, J., and French, D. (1970), *J. Biol. Chem.* (in press).

Rupley, J. A. (1967), *Proc. Roy. Soc. (London) B167*, 416.

Rupley, J. A., and Gates, V. (1967), *Proc. Natl. Acad. Sci. U. S. 57*, 496.

Schechter, I., and Berger, A. (1966), *Biochemistry 5*, 3371.

Schechter, I., and Berger, A. (1967), *Biochem. Biophys. Res. Commun. 27*, 157.

Thoma, J. A. (1966), *Biochemistry 5*, 1365.

Thoma, J. A., *et al.* (1965), *Anal. Biochem. 13*, 367.

Thoma, J. A., and French, D. (1957), *Anal. Chem. 29*, 1645.

Welker, N. E., and Campbell, L. L. (1967), *Biochemistry 6*, 3681.