Subsite Mapping of Enzymes

APPLICATION OF THE DEPOLYMERASE COMPUTER MODEL TO TWO α -AMYLASES

By JIMMY D. ALLEN and JOHN A. THOMA

Department of Chemistry, University of Arkansas, Fayetteville, AR 72701, U.S.A.

(Received 9 March 1976)

In the preceding paper (Allen & Thoma, 1976) we developed a depolymerase computer model, which uses a minimization routine to establish a subsite map for a depolymerase. In the present paper we show how the model is applied to experimental data for two α -amylases. Michaelis parameters and bond-cleavage frequencies for substrates of chain lengths up to twelve glucosyl units have been reported for *Bacillus amyloliquefaciens*, and a subsite map has been proposed for this enzyme [Thoma *et al.* (1971) *J. Biol. Chem.* **246**, 5621–5635]. By applying the computer model to the experimental data, we have arrived at a ten-subsite map. We find that a significant improvement in this map is achieved by allowing the hydrolytic rate coefficient to vary as a function of the number of occupied subsites comprising the enzyme-binding region. The bond-cleavage frequencies for *Aspergillus oryzae* α -amylase are reported for chain lengths of three to twelve. By using the computer model in conjunction with these bond-cleavage frequencies, the enzyme is found to have eight subsites. A partial subsite map is arrived at, but the entire binding region cannot be mapped because Michaelis parameters are complicated by transglycosylation reactions. The hydrolytic rate coefficients for this enzyme are not constant.

In the preceding paper (Allen & Thoma, 1976) we summarized the theory of subsite mapping of depolymerases and showed how the theory can be applied to experimentally accessible parameters to (1) ascertain the number of subsites on the binding region and the position of the catalytic site within these subsites, (2) establish the subsite-substrate-monomer-unit binding energies and (3) assess the values of the hydrolytic rate coefficients. A depolymerase computer model using a minimization routine was developed to convert the experimental Michaelis parameters and bond-cleavage frequencies into a subsite map. The applicability of the computer model to subsite mapping was established through the use of simulated data. In the present paper we will apply the model to experimental data for a bacterial (*Bacillus amyloliquefaciens*) α -amylase and to a fungal (*Aspergillus oryzae*) α -amylase. [1,4- α -D-Glucan glucanohydrolase (EC 3.2.1.1) *A. oryzae* α -amylase is frequently referred to as 'Taka-amylase A'.]

Thoma *et al.* (1970, 1971) have proposed from a bond-cleavage-frequency analysis that *B. amyloliquefaciens* α -amylase is composed of ten subsites, and have predicted the binding energies of the subsites from bond-cleavage frequencies and Michaelis parameters. The inadequacies of the techniques used by Thoma *et al.* (1970, 1971) were thoroughly explored in the preceding paper (Allen & Thoma,

1976). Because of these inadequacies, we have re-examined the subsite map of this enzyme by using a more valid statistical approach. The existence of the acceleration factor proposed by Thoma *et al.* (1971) is critically examined.

Nitta *et al.* (1971) have proposed a partial subsite map for *Aspergillus oryzae* α -amylase. They estimated the number of subsites comprising the binding region of the enzyme as seven from a plot of $\log \bar{V}$ against substrate chain length. However, we have shown that the results of this type of analysis are difficult to interpret, and that a plot of $\log \bar{V}$ against chain length can only set a lower limit on the number of subsites (Thoma & Allen, 1976).

Nitta *et al.* (1971) positioned the catalytic site within the subsites from qualitative bond-cleavage frequencies (Okada *et al.*, 1969) for reducing-end labelled oligosaccharides of chain length four to six. But, as we have shown by using simulated data, the bond-cleavage frequencies of oligosaccharides smaller than the specificity site of a depolymerase are poor probes to determine the position of the catalytic site (Thoma & Allen, 1976). Nitta *et al.* (1971) used Michaelis parameters for oligosaccharides of chain lengths of two to seven to predict the binding energies of selected subsites. Unfortunately these parameters are complicated by transglycosylation reactions (Allen, 1975a) and are therefore not suitable

for measuring unimolecular hydrolysis. Consequently the only remaining experimentally measurable parameters which are useful in determining the binding energies for *A. oryzae* α -amylase, are quantitative bond-cleavage frequencies determined under unimolecular conditions (Allen, 1975a). We have evaluated bond-cleavage frequencies for chain lengths of three to twelve and use this data to re-evaluate the subsite map for this enzyme.

In the preceding paper (Allen & Thoma, 1976) we established the best procedure for subsite mapping by using the computer-minimization model. In the present paper we apply the procedure to *B. amyloliquefaciens* α -amylase and, as far as applicable, to *A. oryzae* α -amylase. We then examine the goodness-of-fit of the subsite model to the experimental data and ascertain the improvement of the fit caused by allowing the acceleration factor to vary.

Experimental

Enzyme

Three-times-recrystallized *Aspergillus oryzae* α -amylase was a generous gift from Y. Nitta (Laboratory of Biophysical Chemistry, College of Agricultural Chemistry, University of Osaka Prefecture, Sakai, Japan). The enzyme was purified on a Sephadex A-50 anion-exchange column (Toda & Akabori, 1963) by using conditions described previously (Allen, 1975a). The purified enzyme preparation proved to be a single component by polyacrylamide-gel electrophoresis and ultracentrifugation (Allen, 1975a).

Maltodextrin oligosaccharides

A series of reducing-end-labelled maltodextrin oligosaccharides was prepared by the action of cyclodextrin glucanotransferase (EC 2.4.1.19) on [^{14}C]glucose (Pazur, 1955) by using conditions described previously (Allen, 1975a). The radioactively labelled oligosaccharides were separated on Whatman 3MM chromatography paper by using a solvent system of water/ethanol (95%, v/v)/nitromethane (13:18:19, by vol.) (Thoma & French, 1957) with multiple ascents (French & Wild, 1953) at room temperature (20–23°C). The chromatograms were subjected to radioautography, and the oligosaccharides were eluted from the paper with water (Allen, 1975).

Total carbohydrate was determined by a microadaptation of the phenol/sulphuric acid assay (Dubois *et al.*, 1956; Allen, 1975a). (Radioactivity counting was performed on 2.4 cm filter-paper discs (Geiger & Wright, 1960) in the system described below.

Oligosaccharide hydrolysis by *A. oryzae* amylase

Since *A. oryzae* amylase has been shown to have significant transglycosylase activity, the hydrolyses for this study were carried out by using experimental conditions where these complicating reactions have been shown to be minimal (Allen, 1975a). The initial substrate concentrations in μM were: G_3 ,* 65.0; G_4 , 100.0; G_5 , 39.0; G_6 , 16.0; G_7 , 8.3; G_8 , 9.1; G_9 , 6.1; G_{10} , 6.2; G_{11} , 8.8; G_{12} , 9.6.

The hydrolyses were carried out in 0.05M-pyridine/acetic acid buffer, pH 5.3, at 25°C (Allen, 1975a). A reaction was initiated by the addition of *A. oryzae* amylase to give an enzyme concentration which would result in about 50% hydrolysis of original substrate in 1 h. Typically, 15 samples were withdrawn during the course of hydrolysis. The sample size was adjusted to give sufficient radioactivity in the products so that about 0.5% hydrolysis at a single bond in an oligosaccharide was detectable.

The sample was added to twice its volume of 15M- NH_3 in a depression on a wax plate and immediately spotted on Whatman 3MM chromatography paper and dried with hot air. For large samples, where it was necessary to concentrate the sample before spotting, NH_3 denaturation was unsatisfactory since *A. oryzae* amylase is able to regain its activity as NH_3 is driven off (Takagi & Toda, 1962; Fischer & deMontmollin, 1951). Therefore when concentration was necessary, heat denaturation was used. The samples were heated to 100°C for 10 min, which was sufficient to denature the enzyme, and concentrated *in vacuo* before spotting.

Radioactively labelled oligosaccharide standards were spotted on the chromatograms, and the chromatograms were developed in the solvent system described above. Enough ascents were used (two ascents for chain length of three; up to eight ascents for a chain length of twelve) to give sufficient resolution for the oligosaccharides under investigation. The chromatograms were subjected to radioautography. The resulting X-ray film was stapled to the chromatogram and used as a guide for cutting out the areas containing radioactive products. When a particular product oligosaccharide was not present in sufficient quantities to darken the film, the standards were used as a guide in cutting out the chromatogram.

The cut-out sections of the chromatogram were placed in the bottom of scintillation vials and 5 ml of scintillation fluid, prepared as previously described (Allen, 1975a), was added. The sample radioactivity was counted on a Beckman LS-100 liquid-scintillation counter. The sample radioactivity was counted to 0.3% error (based on 2σ statistics) or for a maximum of 5 min.

* G_n is a malto-oligosaccharide of n glucopyranoside units.

Bond-cleavage-frequency analysis

Operationally the bond-cleavage frequency for a particular bond in an oligosaccharide is the fraction of the labelled product resulting from the cleavage of that bond of the total labelled products. For simple unimolecular hydrolysis (Allen, 1975a) of an oligosaccharide, the bond-cleavage frequency represents the relative rate of hydrolysis of a bond in an oligosaccharide.

The radioactivity that corresponds to a particular oligosaccharide on a chromatogram contains sample background radioactivity. This background radioactivity, which is present before the action of the enzyme, is due to contaminating oligosaccharides and autodegradation products. Sample background may be accounted for by spotting a sample at zero time, i.e. before the enzyme has hydrolysed substrate, and subtracting the radioactivity of this zero-time determination from all subsequent samples of the digest (Thoma *et al.*, 1970). However, it has been shown that this procedure propagates the sample background error throughout the measured values (Allen, 1975b), and a better way of manipulating the data is by plotting

$$\frac{\sum_{j=1}^{n-1} G_j}{\sum_{j=1}^n G_j} \text{ against } \frac{G_i}{\sum_{j=1}^n G_j}$$

where G_j is the measured radioactivity (c.p.m.) for an oligosaccharide of chain length j , and G_i is the radioactivity for a product of chain length i , which is formed when a substrate oligosaccharide of chain length n is hydrolysed. The slope of such a plot for each labelled product equals the bond-cleavage frequency for the bond that when cleaved yields the plotted product. This procedure makes a zero-time determination unnecessary and, consequently, conserves radiolabelled substrates. A typical bond-cleavage-frequency plot is shown in Fig. 2.

The counting data were analysed as a routine by a computer program that transformed the radioactivity values into the functions to be plotted. A computer program to perform a least-squares fit allowing for differential weighting due to error in both x and y values (Christian *et al.*, 1974) was generously supplied by E. H. Lane, University of Oklahoma, Norman, OK 73069, U.S.A. The weighting factors were calculated by the computer program based on the radioactive-counting error (Friedlander *et al.*, 1964) and a standard propagation error analysis (Beers, 1953). The computer program automatically plotted the data and the least-squares line on a CalComp plotter. Because of the experimental error, bond-cleavage frequencies will not necessarily add to unity.

Curvature in this type of plot is diagnostic of a

change in product ratios. This change may be due to a shift in mechanism as substrate is depleted (Allen, 1975) or may be due to secondary attack (i.e. degradation of the products of the original bond cleavage of substrate oligosaccharide). Experimental conditions used in this work were such that the only operative mechanism was unimolecular hydrolysis. However, in the later stages of hydrolysis, curvature was often evident, due to secondary attack on products. The degree of hydrolysis of substrate at which secondary attack becomes significant depends on the relative rates of hydrolysis of products and substrate oligosaccharide. Since maltose is a poor substrate (Nitta *et al.*, 1968) maltotriose could be carried to a high degree of hydrolysis with no curvature evident in the plot; however, G_{12} -oligosaccharide-hydrolysis plots exhibited curvature at about 50% hydrolysis. In this analysis, any points exhibiting curvature in the later stages of hydrolysis were not used to determine bond-cleavage frequencies.

Results and Discussion*Bacillus amyloliquefaciens*

Evaluation of the number of subsites and the position of the catalytic site. As discussed in the preceding paper (Allen & Thoma, 1976), Thoma *et al.* (1970) found that the binding region of *B. amyloliquefaciens* is comprised of ten subsites with the catalytic site located between subsites VI and VII. By using the technique outlined in the preceding paper (Allen & Thoma, 1976), we can use the minimization model to test the validity of this prediction.

The results of the analysis using minimization with bond-cleavage frequencies are given in Table 1. The ten subsites predicted by Thoma *et al.* (1970) are optimized in minimization number 3 and give $Q_{\text{min.,b.c.f.}} = 0.2$. [A complete listing of symbols used is given in Table 1 of the preceding paper (Allen & Thoma, 1976).] Minimizations numbers 1 and 2 test the effect of removing an end subsite from the binding region of this ten-subsite map. Removing subsite number X results in a 64-fold decrease in fit, and removing subsite number I gives a 26-fold decrease in fit as evidenced by $Q_{\text{min.,b.c.f.}}$ values. Hence, at least ten subsites are required to explain the bond-cleavage frequencies.

In minimization number 4, we see the effect of adding a subsite to the right of the binding region. The $Q_{\text{min.,b.c.f.}}$ shows no improvement, and subsite number XI is optimized at 390 J/mol. We showed in the preceding paper (Allen & Thoma, 1976) that experimental scatter can cause apparent subsite energies of ± 630 J/mol; hence, the binding energy of subsite XI is not significantly greater than zero and not indicative of a real subsite on the enzyme. In minimization number 5 an additional subsite is added to the left of the binding region of the ten-sub-

Table 1. Evaluation of the number of subsites on *B. amyloliquefaciens* α -amylase

The experimental bond-cleavage frequencies measured by Thoma *et al.* (1970) and given in Table 4 were used to optimize the binding energy of selected subsites (designated by X), while the remaining subsite binding energies were constrained at zero.

Minimization no.	Subsite index no.	...	Catalytic site											Q _{min., b.c.f.}	
			-I	I	II	III	IV	V	VI ↓ VII	VIII	IX	X	XI		
1				X	X	X	X	X			X	X			12.9
2					X	X	X	X			X	X	X		5.2
3				X	X	X	X	X			X	X	X		0.2
4				X	X	X	X	X			X	X	X	X	0.2
5			X	X	X	X	X	X			X	X	X		0.1
6			X	X	X	X	X	X			X	X	X	X	0.1

Table 2. Optimized subsite maps for *B. amyloliquefaciens* α -amylase

The subsite binding energies were optimized by using the data of Thoma *et al.* (1970, 1971). ----- shows the position of the catalytic site.

Subsite no.	Apparent subsite binding energies (kJ/mol)			
	ΔG_a constrained at zero			ΔG_a allowed to vary†
	$Q_{\text{b.c.f.}}$ minimized	$Q_{K_m} + Q_V + Q_{K_{\text{int.}}}$ minimized	Q_{total} minimized	Q_{total} minimized
I	-4.85		-3.72	-4.48
II	-9.83		-8.87	-10.21
III	-1.00		-0.54	-0.67
IV	-4.35		-3.72	-4.23
V	-9.67		-8.16	-9.54
VI	*	25.48	20.79	13.81

VII	*	-9.58	-12.05	-14.39
VIII	-7.28		-6.28	-7.20
IX	-4.27		-3.68	-4.02
X	5.23		-6.02	5.27
ΔG_a (kJ/mol)	0	0	0	1.55
Normalized sum of residual error squared				
$Q_{\text{min.}}$				
b.c.f.	0.2	0.2	0.6	0.3
K_m	†	50.4	24.7	9.9
$\bar{P}$	†	31.3	6.6	6.0
$K_{\text{int.}}$	†	0.003	0.3	0.02
Total		11.3	4.7	2.4

* Subsites adjacent to the catalytic site cannot be estimated from bond-cleavage frequencies.

† Michaelis parameters and $K_{\text{int.}}$ cannot be computed without a complete subsite map.

‡ The binding energies listed are apparent energies and contain the acceleration factor of 1.55 kJ/mol.

site enzyme. In this case, the improvement of the fit is twofold over minimization number 3. Subsite number -I is predicted as -920 J/mol. Since end subsites of up to ± 630 J/mol were artifacts of experimental scatter, and we can only predict subsite energies with ± 840 J/mol due to the experimental scatter, we surmise that this optimized subsite is not a real subsite. We cannot state conclusively that the additional subsite is not warranted, due to the im-

provement in $Q_{\text{min., b.c.f.}}$; however, it is seen below that minimization number 3 predicts bond-cleavage-frequency data within experimental error. We propose that the ten subsites originally predicted by Thoma *et al.* (1970) are correct. The optimized binding energies from minimization number 3 are given in Table 2 (column 1) and are within 420 J/mol of those predicted by Thoma *et al.* (1971) from a less rigorous bond-cleavage-frequency analysis.

Evaluation of the binding energies of subsites VI and VII. By using the optimized binding energies of subsites I–V and VIII–X established by bond-cleavage frequencies (Table 2, column 1), the binding energies of subsites VI and VII were optimized by using experimental Michaelis parameters and K_{int} . (Thoma *et al.*, 1971). $Q_{min.}$ was achieved when subsites VI and VII were equal to 25.48 and –9.58 kJ/mol respectively. The $Q_{min.}$ values are given in Table 2, column 2, with $Q_{min.,total} = 11.3$ after this second stage of optimization. No additional (i.e. local) minima (Allen & Thoma, 1976) were found, even though a wide range of initialized energies (–42 to 42 kJ/mol) for subsites VI and VII were tested.

The Michaelis parameters computed by this optimized map are compared with the experimental data in Table 3, columns 2 and 5. The ratios of the experimental to calculated parameters for this map with ΔG_a constrained at zero obviously do not approach a random distribution about a ratio of unity. There is a trend in the ratios, first increasing and then decreasing. When the error is not randomly distributed, a poor model is indicated (Bártfai & Mannervik, 1972).

To achieve the best possible fit to the data, without allowing the hydrolytic coefficients to vary, all of the subsites were optimized by using all of the experimental data to obtain $Q_{min.,total}$. The optimized subsite map is given in Table 2, column 3. To improve the fit with Michaelis parameters (Q_{K_m} , Q_V), the bond-cleavage-frequency fit is decreased, as evidenced

by a threefold increase in $Q_{min.,b.c.f.}$. $Q_{min.,total}$ has been improved from 11.3 to 4.7.

Optimization of acceleration factor. When all of the subsites and the acceleration factor are allowed to vary to obtain optimum energies as dictated by $Q_{min.,total}$, the acceleration factor is optimized to 1.55 kJ/mol. The final map is shown in Fig. 1 and the $Q_{min.}$ values are given in Table 2, column 4. By allowing an acceleration factor, $Q_{min.,total}$ has been decreased from 4.7 to 2.4. The F test shows that the improvement of fit is significant at the 99.5%-confidence level (Pearson & Hartley, 1970; Sokal & Rohlf, 1969). We showed in the preceding paper (Allen & Thoma, 1976) that the error in the acceleration factor due to experimental scatter is ± 21 J/mol. Therefore the predicted value of 1.55 kJ/mol is not likely due to experimental error. We conclude, from the F test and the magnitude of ΔG_a that the acceleration factor reflects an actual physical phenomenon and is not an artifact of data processing.

The Michaelis parameters calculated from the optimized map are compared with experimental data in Table 3, columns 3 and 6, and the computed and experimental bond-cleavage frequencies are presented in Table 4. It is obvious that the experimental data for substrates of chain length one to twelve is closely predicted.

In the preceding paper (Allen & Thoma, 1976) we showed that there are three contributing factors to $Q_{min.}$: $q_{req.}$ (minimization inadequacies), $q_{var.}$ (experimental variance), and q_{bias} (model inadequacies). We further showed that $q_{req.}$ was significant so that

Table 3. Comparison of experimental and computed Michaelis parameters for *B. amyloliquefaciens* α -amylase

The experimental Michaelis parameters (Thoma, 1971) are compared with those computed from optimized subsite maps where the acceleration factor was constrained at zero ($\Delta G_a = 0$ kJ/mol) and where the acceleration factor was optimized ($\Delta G_a = 1.55$ kJ/mol).

Substrate chain length	K_m or K_1^* (M)			V (relative)†		
	Experimental†	Experimental/calculated		Experimental	Experimental/calculated	
		$\Delta G_a = 0$ kJ/mol	$\Delta G_a = 1.55$ kJ/mol		$\Delta G_a = 0$ kJ/mol	$\Delta G_a = 1.55$ kJ/mol
1	6.0×10^{-1}	2.0	2.7	—	—	—
2	2.2×10^{-2}	0.7	0.7	—	—	—
3	1.9×10^{-2}	2.3	1.6	2.4×10^{-4}	3.1	3.7
4	1.7×10^{-2}	8.0	0.7	2.0×10^{-3}	5.2	0.9
5	1.9×10^{-2}	53.7	2.2	6.3×10^{-3}	11.8	0.9
6	9.9×10^{-3}	28.0	1.3	3.4×10^{-2}	11.6	1.0
7	5.2×10^{-3}	14.8	1.3	2.5×10^{-1}	8.3	1.4
8	1.5×10^{-3}	4.5	1.2	4.1×10^{-1}	1.6	0.8
9	8.8×10^{-4}	3.0	1.3	1.0	1.2	1.1
10	5.6×10^{-4}	1.9	0.8	1.0	1.1	1.1
11	4.9×10^{-4}	1.7	0.8	1.0	1.1	1.0
12	8.6×10^{-4}	3.1	1.4	1.0	1.0	1.0

* The value for glucose is an inhibition constant; all others are K_m values.

† Normalized to a substrate chain length of 12, maximum velocity.

Fig. 1. Subsite map for *B. amyloliquefaciens* α -amylase

The number of subsites and the position of the catalytic site were determined from a bond-cleavage-frequency minimization as outlined in the text. The binding energies and ΔG_a were determined by minimization of $Q_{\min.,total}$. The arrow shows the position of the catalytic amino acids. The open bars depict the apparent binding energies of subsites I-V and VIII-X obtained from bond-cleavage-frequency analysis. The solid bars are more positive by 1.55 kJ/mol (the acceleration factor) and show the actual subsite binding energies.

$Q_{\min.} \approx q_{var.} + q_{bias.}$ For bond-cleavage frequencies, $q_{var.}$ was estimated to be about 0.9 and for *B. amyloliquefaciens* we have $Q_{\min.,b.c.f.} = 0.3$; hence, bond-cleavage frequencies are predicted correctly within experimental error.

A reasonable upper limit for $q_{var.}$ for K_m and V was shown to be 2.1 and 1.4 respectively (Allen & Thoma, 1976). We see from Table 2 that the minimized values for Q_{K_m} and Q_V are 9.9 and 6.0 respectively. Therefore the model cannot predict K_m and V within experimental error. The q_{bias} for K_m is about 7.8, or four times $q_{var.}$ for K_m , and q_{bias} for V is about 4.6, or three times $q_{var.}$ for V . This bias is probably due to the approximation that $\Delta G_{a,1}$ is equal for subsites on the enzyme. It is unlikely that each subsite in the binding region contributes equally to the hydrolytic coefficient; hence, the equality of $\Delta G_{a,1}$ can only be an approximation. The computer model shows that this approximation does significantly improve the fit; consequently we surmise that the acceleration factor does reflect the fact that binding influences catalysis, but each subsite probably does not contribute equally.

Aspergillus oryzae α -amylase

We have previously shown (Allen, 1975a) that *A. oryzae* α -amylase has significant transglycosylase activity, hence the bond-cleavage frequencies reported here were determined under conditions where complicating multimolecular reactions are minimal. A typical bond-cleavage-frequency-analysis plot is shown in Fig. 2 and the bond-cleavage frequencies for chain lengths of three to twelve are given in Table 5.

We have also found that *A. oryzae* α -amylase is a repetitive-attack enzyme (i.e. cleaves more than one glycosidic bond per encounter) (J. D. Allen & J. A. Thoma, unpublished work); however, as with pig pancreatic α -amylase (Robyt & French, 1970), the polarity of multiple attack is towards the non-reducing end of the substrate. Since the substrates used in this study were labelled in the reducing end, the measured bond-cleavage frequencies will not be influenced by a subsequent attack on the unlabelled portion of the substrate molecule.

Evaluation of the number of subsites and the position of the catalytic site. An examination of the experimental bond-cleavage frequencies in Table 5 reveals a trend that allows a qualitative estimate of the number of subsites and the position of the catalytic site comprising the binding region of *A. oryzae* amylase. It was established in the preceding paper (Allen & Thoma, 1976) that when a substrate of chain length n becomes larger than the binding region composed of l subsites (i.e. $n > l$) that there are $n-l+1$ bond-cleavage frequencies in the n -mer that are identical. As indicated in Table 5 by underlining, for substrates larger than malto-octaose there are bond-cleavage frequencies for each substrate which are approximately constant. For example, for $n=9$, the underlined bond-cleavage frequencies are 0.28 ± 0.02 and for $n=10$ the underlined bond-cleavage frequencies are 0.22 ± 0.01 . These data are indicative of an eight-subsite enzyme with the catalytic site located between subsites III and IV as shown in Fig. 3. The binding modes for $n=8, 9$ and 10 where all of the subsites are occupied are shown.

In Table 6, it is shown how this intuitive approach to the number of subsites and the position of the catalytic site is placed on a more statistically sound footing by use of the minimization model. In each minimization selected subsite binding energies were optimized while the remaining subsites were constrained at a binding energy of zero. It can be seen that removing either subsite VIII (minimization 1) or subsite I (minimization 2) results in a much poorer fit than the eight-subsite map as evidenced by a 39- and 15-fold increase in $Q_{\min.,b.c.f.}$ respectively. Hence at least an eight-subsite map is necessary to accommodate the bond-cleavage-frequency data. When the minimization routine was allowed to

Notations used: O, D-glucopyranoside unit, Φ , 'reducing' radioactively labelled glucopyranoside unit [Thoma *et al.* (1970) used the α -methyl malto-oligosaccharide series]; —, α -(1 $\rightarrow$ 4) glycosidic bond. The experimental bond-cleavage frequencies are from Thoma *et al.* (1970). The calculated bond-cleavage frequencies were computed from the optimized map shown in Fig. 1.

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

□, Glucose; ○, maltose; △, maltotriose; ■, maltotetraose; ●, maltohexaose; G_i, malto-oligosaccharide of *i* glucosyl units. The radioactivity due to each sugar in a sample/the total radioactivity in each sample is plotted against the radioactivity in product sugars in a sample/the total radioactivity in each sample. The slope of each line is the bond-cleavage frequency for the substrate bond, which results in the plotted product. A complete description of the analysis is in the Experimental section.

Table 6. Evaluation of the number of subsites on *A. oryzae* α -amylase

The experimental bond-cleavage frequencies for *A. oryzae* amylase given in Table 5 were used to optimize the binding frequency of selected subsites (designated by X), while the remaining subsite binding energies were constrained at zero.

Minimization no.	Subsite index no.	...	-I	I	II	Catalytic site		V	VI	VII	VIII	IX	$Q_{\text{min.,b.c.f.}}$
						III	IV						
1				X	X			X	X	X			3.9
2					X			X	X	X	X		1.5
3				X	X			X	X	X	X		0.1
4				X	X			X	X	X	X	X	0.1
5			X	X	X			X	X	X	X		0.1
6			X	X	X			X	X	X	X	X	0.1

Fig. 4. Apparent subsite-binding energies for *A. oryzae* α -amylase

The number of subsites and the position of the catalytic site were determined by a bond-cleavage-frequency minimization. The apparent binding energies for subsites I, II, and V–VIII were evaluated by optimizing these binding energies to give the best fit to bond-cleavage-frequency data as established by $Q_{\text{min.,b.c.f.}}$. $\uparrow$ is the position of the catalytic amino acids. The binding energies of subsites III and IV and the acceleration factor cannot be evaluated, because of complicating transglycosylation reactions (Allen, 1975a).

are considerably lower (150–9.5 mM) than for the oligosaccharides. In the concentration range represented by K_m values for the larger substrates, bimolecular reactions are significant; hence, we can use these K_m values to approximate $K_{\text{int.}}$, the dissociation constant for a positional isomer where all subsites are occupied. The values of K_m were plotted against $n-l+1$ (Allen & Thoma, 1976) and fit to a hyperbola by using a weighted least-squares analysis (Cleland, 1967). The value of $K_{\text{int.}}$ is 1.59 mM so that the sum of the binding energies is 26.0 kJ/mol [see the preceding paper, Allen & Thoma (1976)]. The sum of the apparent binding energies for subsites I, II and V–VIII shown in Fig. 4 is –50.7 kJ/mol. Hence the sum of the binding energies for subsites III and IV is 24.7 kJ/mol; or, more likely as we shall see, the apparent binding energies contain an acceleration factor, $\Delta G_{a,i}$, and must be made more positive by this factor to yield true binding energies.

Evaluation of the acceleration factor. Because of a lack of K_m and $\bar{V}$ as a function of chain length, we cannot evaluate the acceleration factor; but we can show that the apparent binding energies shown in Fig. 4 must incorporate an acceleration factor. Suetsugu *et al.* (1968) have reported that glucose and maltose are competitive inhibitors for *A. oryzae* α -amylase when using *p*-nitrophenyl α -maltoside as a substrate. The reported K_i values for glucose and maltose are 78 and 95 mM respectively. An examination of Fig. 4 reveals that binding at subsite II alone predicts a dissociation constant of 0.28 mM, so that the apparent binding energies predict about 300-fold too tight binding. This is diagnostic that the apparent binding energies contain a rate-coefficient term and are too low by the value $\Delta G_{a,i}$. Unfortunately we do not presently have the means to estimate the value of this acceleration factor, because of the inaccessibility of Michaelis parameters.

We have shown how the depolymerase computer model can be successfully applied to the subsite mapping of two α -amylases. For *B. amyloliquefaciens* α -amylase the complete subsite map is evaluated, and

the acceleration factor is estimated as 1.55 kJ/mol. For *A. oryzae* α -amylase, a partial subsite map is obtained, and it is shown that the apparent subsite-binding energies necessarily have an acceleration factor incorporated into them.

The map for *B. amyloliquefaciens* α -amylase in Fig. 1 has apparent subsite-binding energies that are within 710 J/mol of those predicted by Thoma *et al.* (1971) with the exception of subsites VI and VII, which vary by about 2.5 kJ/mol from those predicted by Thoma *et al.* (1971). We find an optimized acceleration factor of 1.55 kJ/mol as compared to 1.88 kJ/mol found by Thoma *et al.* (1971). The map shows a strongly unfavourable interaction at subsite VI, which is analogous to the unfavourable interaction of 12.5–25 kJ/mol at subsite D (the subsite immediately to the left of the catalytic site) in lysozyme (Chipman & Sharon, 1969). This unfavourable interaction energy is consistent with the X-ray structure of lysozyme, where it is proposed that the glycosyl unit interacting with this subsite is distorted from the stable chair configuration to the half-chair. This distortion subsite therefore helps the reaction proceed toward the transition state.

The *B. amyloliquefaciens* map as well as the *A. oryzae* map have another subsite, which interacts unfavourably with the substrate monomer unit. Subsite X in Fig. 1 and subsite VII in Fig. 4 have positive free energies of binding and will be referred to as 'barrier subsites'. For the *A. oryzae* map this unfavourable interaction is manifested in the bond-cleavage frequencies as an unfavourable positional isomer, which results in a lower bond-cleavage frequency. For example, consider the bond-cleavage frequencies in Table 5 for malto-octaose and the diagrammatic representation in Fig. 3. When malto-octaose binds in binding mode VI,8 the barrier subsite is avoided, and the bond-cleavage frequency is 0.27, but when malto-octaose binds in binding mode VII,8 the barrier subsite is encountered and the bond-cleavage frequency is depressed to 0.11. In this positional isomer formed here, all subsites are occupied, binding mode VIII,8, the unfavourable barrier-subsite energy of 2.13 kJ/mol is more than offset by the favourable interaction of subsite VIII at –3.22 kJ/mol to result in a net binding energy for interaction at these two subsites of –1.09 kJ/mol. Consequently the bond-cleavage frequency resulting from binding mode VIII,8 is 0.39, higher than for binding modes VI,8 or VII,8. For *A. oryzae* α -amylase the binding mode where all subsites are occupied results in the highest bond-cleavage frequency. A high bond-cleavage frequency such as this has been used to predict the number of subsites on pig pancreatic α -amylase (Robyt & French, 1970). The hazards of using this criterion can be demonstrated with the *B. amyloliquefaciens* α -amylase data.

For *B. amyloliquefaciens* (Fig. 1) the barrier subsite is on the end of the binding region. Therefore there is no compensating favourable subsite as is the case of *A. oryzae* α -amylase. The bond-cleavage frequencies in Table 4 show the consequences of the barrier subsite. For maltodecaose the largest bond-cleavage frequency is not a result of all subsites being filled but is a result of avoidance of the barrier subsite VII. As a result of the unfavourable interactions at subsite VII, cleavage near the non-reducing end of the substrate is very low (Table 4).

B. amyloliquefaciens, then is approaching the properties of an exo-amylase. It is obvious that, as binding at the barrier subsite becomes more unfavourable, hydrolysis beyond the third glucosyl unit from the reducing end will become less frequent until it is essentially non-existent. Thus there is no sharp division between exo- and endo-depolymerases, but a continuum with enzymes such as *A. oryzae* α -amylase at the one end of the scale, β -amylase and glucoamylase at the other with *B. amyloliquefaciens* α -amylase lying somewhere in between.

This work was supported by grants from the National Science Foundation and the University of Arkansas Computing Center.

References

- Allen, J. D. (1975a) Ph.D. Thesis, University of Arkansas
- Allen, J. D. (1975b) *Carbohydr. Res.* **39**, 312–315
- Allen, J. D. & Thoma, J. A. (1976) *Biochem. J.* **159**, 105–120
- Bárfai, T. & Mannervik, B. (1972) in *Analysis and Simulation of Biochemical Systems* (Hemker, H. C. & Hess, B., eds.), pp. 198–209, Elsevier, New York
- Beers, Y. (1953) *Introduction to the Theory of Error*, pp. 26–35, Addison-Wesley, Cambridge, MA
- Chipman, D. M. & Sharon, N. (1969) *Science* **165**, 454–465
- Christian, S. D., Lane, E. H. & Garland, F. (1974) *J. Chem. Educ.* **51**, 475–476
- Cleland, W. W. (1967) *Adv. Enzymol.* **29**, 1–32
- Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. (1956) *Anal. Chem.* **28**, 350–356
- Fischer, E. H. & deMontmollin, R. (1951) *Helv. Chim. Acta* **34**, 1987–1994
- French, D. & Wild, G. M. (1953) *J. Am. Chem. Soc.* **75**, 2612–2616
- Friedlander, G., Kennedy, J. W. & Miller, J. M. (1964) *Nuclear and Radiochemistry*, 2nd edn., pp. 166–190, John Wiley and Sons, New York
- Geiger, J. W. & Wright, L. D. (1960) *Biochem. Biophys. Res. Commun.* **2**, 282–286
- Nitta, Y., Hiromi, K. & Ono, S. (1968) *J. Biochem. (Tokyo)* **63**, 632–636
- Nitta, Y., Mizushima, M., Hiromi, K. & Ono, S. (1971) *J. Biochem. (Tokyo)* **69**, 567–576

- Okada, S., Kitahata, S., Higashihara, M. & Fukumoto, J. (1969) *Agr. Biol. Chem.* **33**, 900–906
- Pazur, J. H. (1955) *J. Am. Chem. Soc.* **77**, 1015–1017
- Pearson, E. S. & Hartley, H. O. (eds.) (1970) *Biometrika Tables for Statisticians*, vol. 1, 3rd edn., pp. 169–175, Cambridge University Press, London
- Robyt, J. F. & French, D. (1970) *J. Biol. Chem.* **245**, 3917–3926
- Sokal, R. R. & Rohlf, F. J. (1969) *Biometry*, pp. 175–203, W. H. Freeman, San Francisco
- Suetsugu, N., Hiromi, K., Takagi, M. & Ono, S. (1968) *J. Biochem. (Tokyo)* **64**, 619–624
- Takagi, T. & Toda, H. (1962) *J. Biochem. (Tokyo)* **52**, 16–27
- Thoma, J. A. & Allen, J. D. (1976) *Carbohydr. Res.* **48**, 105–124
- Thoma, J. A. & French, D. (1957) *Anal. Chem.* **29**, 1645–1648
- Thoma, J. A., Brothers, C. & Spradlin, J. (1970) *Biochemistry* **9**, 1768–1775
- Thoma, J. A., Rao, G. V. K., Brothers, C., Spradlin, J. & Li, L. H. (1971) *J. Biol. Chem.* **246**, 5621–5635
- Toda, H. & Akabori, S. (1963) *J. Biochem. (Tokyo)* **53**, 102–110