Appl Biochem Biotechnol (2008) 147:107–117
DOI 10.1007/s12010-007-8087-x

Development of Activity-based Cost Functions for Cellulase, Invertase, and Other Enzymes

Chris C. Stowers · Elizabeth M. Ferguson · Robert D. Tanner

Received: 11 May 2007 / Accepted: 18 October 2007 /
Published online: 13 December 2007

Abstract As enzyme chemistry plays an increasingly important role in the chemical industry, cost analysis of these enzymes becomes a necessity. In this paper, we examine the aspects that affect the cost of enzymes based upon enzyme activity. The basis for this study stems from a previously developed objective function that quantifies the tradeoffs in enzyme purification via the foam fractionation process (Cherry et al., Braz J Chem Eng 17:233–238, 2000). A generalized cost function is developed from our results that could be used to aid in both industrial and lab scale chemical processing. The generalized cost function shows several nonobvious results that could lead to significant savings. Additionally, the parameters involved in the operation and scaling up of enzyme processing could be optimized to minimize costs. We show that there are typically three regimes in the enzyme cost analysis function: the low activity prelinear region, the moderate activity linear region, and high activity power-law region. The overall form of the cost analysis function appears to robustly fit the power law form.

Keywords Cost function · Cellulase · Invertase · Foam fractionation · Separations cost

Introduction

Since the rapid development of the chemical industry in the 1940s, the chemical industry has morphed from a commodity chemical market into a specialty chemical market where bio-related products and processes have become increasingly important. Enzymes are now commonly used as catalysts to produce products such as proteins, sugars, and lipids as well as processing tools to enhance more traditional products, such as paper pulp. In fact,

C. C. Stowers · R. D. Tanner (✉)
Department of Chemical Engineering, Vanderbilt University, Nashville, TN 37235, USA
e-mail: Robert.D.Tanner@Vanderbilt.edu

E. M. Ferguson
Pediatrics Department/Judith B Ratner MD Office, George Washington University,
2150 Pennsylvania Ave NW, Washington, DC 20016, USA

enzyme industry sales are expected to increase to 2.2 billion dollars by 2010 [4]. It has been predicted that the kinetics of cost reduction for industrial enzymes will control the extent of how the enzyme market will grow [5]. Thus, cost analysis of enzymes is becoming an important factor for engineers and scientists dealing with mainstream bio-products.

The price of enzymes is often controlled by the method of production and more importantly the purification method [6]. Many enzymes can be isolated from common plants and microorganisms while others must be harvested from animals, particularly mammals such as pigs, rabbits, and even humans. Generally, enzymes isolated from common plants are less costly than those isolated from animals. Likewise, enzymes that require little purification are less costly. Some enzymes require additional processing beyond traditional standard processing, which includes foam fractionation, salting out, and liquid–liquid extraction [7]. These additional processing steps can become extremely costly and can lead to increased market cost of the enzyme. These processing steps often include standard gel chromatography and/or high performance/pressure liquid chromatography (HPLC).

In this study, we will explore whether an objective function, relating enzyme activity to cost, can be developed to establish the cost function of cellulase and other enzymes. We will determine the cost increase in crude cellulase (in terms of activity per mass) after processing used to enhance the purity and concentration of this protein. We shall start with a more generalized objective function comprised of measurable purification process responses to create our cost model and reduce that model to the specific cost function for this study. A previously developed objective function quantifies the tradeoff between maximizing the enzyme concentration in a separation process such as a foam fractionation process and minimizing the loss of enzyme mass and enzyme activity in that process [1] is shown below.

$$\Phi = (\mathrm{AR})^{a}(\mathrm{MR})^{b}(\mathrm{ER})^{c} \tag{1}$$

where AR=activity recovery=Afoam/Ai
MR=mass recovery=Mfoam/Mi
ER=enrichment recovery=Cfoam/Ci

Typically, a, b and c are positive in the case where Φ is a generalized, desired performance factor. On the other hand, as we shall see in the results section, when MR is replaced by purchase mass and Φ is replaced by \$/mg, b becomes negative. Here, we explore whether a generalized relationship between cost, activity, and recovered mass alone can characterize a cost function, and if so, determine the appropriate coefficients.

This approach to the development of cost-based processing can result in the lowering of product cost at each processing step by selecting the control variables which can maximize the respective Φ or minimize the cost-based functions at each step. In a foam-fractionation process, in particular, these control parameters are typically the pH and the foaming-gas superficial velocity [1]. In this approach, it is convenient to define Φ as a generalized value (or price) function. In this study, we define price as the price published in leading biochemical catalogs used in this study. Additional processing to generate higher purity enzyme will result in incremental increases in the value of Φ, which can be represented as the first derivative of the objective function with respect to activity. In particular, in this study, we shall compare the catalog values to a parameter fitted model for industrial enzymes to determine whether there are general trends and quantitative similarities between classes of enzymes. Our goal is to develop an objective function based on market values, making the model most useful from the purchasing (consumer) point of view. However, the

model could also be used from a manufacturing point of view to determine whether the production of certain products would be profitable by determining the economic value of a potentially new product before investment in production. In the initial part of this study, we will assume the second and third terms of the objective function model (Eq. 1), Φ, are constant and can be lumped into a new parameter γ. This permits the model to be simplified to the following equation, where the activity term is defined as activity per unit mass, as is often expressed in enzyme sales catalogs.

$$\Phi = \gamma(\mathrm{A})^{\mathrm{a}} \tag{2}$$

With this framework, our underlying assumption for the remainder of the work is that for a given enzyme, the processing cost is captured by the power law framework and by difference production (fermentation) costs are included within the constant γ. This is a good assumption since the majority of industrial enzymes are produced in a microbial environment or directly harvested from plants. These enzymes, such as cellulases, are often processed directly for activity [2, 3], which requires increasingly larger costs to reach higher activity. Enzymes generally used for medical applications, which are produced in a mammalian cell line or from animals, may tend to deviate from this assumption. Additionally, the cost of specialty enzymes, which are produced in only small quantities for R&D purposes, may be dictated by investment cost. We shall focus on industrial enzymes here.

The cost of storage for enzymes can also influence the cost of these chemicals at the customer level. It can be readily established that significant savings (greater than 25%) can be achieved by simply buying in bulk from an enzyme supplier. However, bulk storage costs of these enzymes are not negligible considering that most enzymes must be kept at −20°C. The additional cost of storage resulting from bulk purchases needs to be considered by engineers and scientists in industry. Thus, we present a slightly modified cost function, Φ, that accounts for the cost effects of buying in bulk quantity and the resulting enzyme storage.

$$\Phi = \gamma(A)^{a}\delta(M)^{b}(1 + \alpha * t) \tag{3}$$

where Φ=enzyme cost in $/mg
M=mass purchase amount
t=storage time in weeks
a and b are dimensionless constants
γ, δ, and α have appropriate dimensions.

Two key observations can be made regarding this improved cost function. First, the cost per unit mass is directly proportional to activity and inversely proportional to the amount of mass purchased. This observation is fairly trivial knowing that if a customer wants a higher activity per unit mass product, it will cost more and the more purchased will warrant a decrease in price per unit. Secondly, the parameter b must be negative knowing that the price per unit should decrease as more product is purchased. The larger the absolute value of b, the more discount a consumer will receive for buying in bulk. Thus, the parameter b and the storage cost function (third term) must be balanced to attain the most cost-effective purchase amount for a given purchase. α has the units of 1/week (inverse time) and the parameters γ and δ have units dependent on the values of the coefficients a and b and the units of activity and mass are expressed, respectively. Of course, activity and the mass purchase amounts could be expressed in dimensionless form to generalize Eq. 3. It is also important to notice that the parameters γ and δ can be lumped into one single parameter. However, for our purposes, it is easier to now look at the two parameters separately so that

we can investigate the proposed model terms individually. The first term can be investigated as an independent activity dependent cost function, while the second term can be investigated to determine the effects of mass purchase amount on the cost function. This framework will allow us to empirically determine whether the simple models shown in Eqs. 2 and 3 can accurately relate enzyme activity and purchase amount to cost for industrial enzymes.

Materials and Methods

The enzyme market was investigated for the following enzymes: cellulase, invertase, collagenase, papain, alpha-amylase, and elastase. The market was surveyed by determining the cost per mg of enzyme from Fisher Scientific, Sigma-Aldrich, Carolina Biochemical, Worthington Biochemical, and Elastin Products Company [8–12]. If these suppliers sold a particular enzyme in different allotment sizes, the cost per mg was averaged over all allotments to determine an average price per mg from that particular supplier. This averaging was introduced here to reduce the market scattering of the data to enable us to more clearly observe the trend of enzyme activity on cost. Data were also collected for invertase from an available industrial market report [13]. Only enzymes that were sold in allotments in terms of mass were considered (those sold in terms of enzyme units were ignored). The activity per mg of each product was recorded along with the respective cost per mg. When the supplier gave an activity range, the lower limit of the range was selected to be the finite activity used in this study. Only enzymes that had reported activities with similar units were compared. A complete list of the data collected is shown in Table 1. The activity units between different enzymes varied because the enzyme activity was generally expressed as the amount of substrate utilized per unit time; however, the amount of substrate and time often varied between enzymes. A complete list of activities for enzymes used in this study is also shown in Table 1.

The data for price per mg for each enzyme were then plotted versus the corresponding activity per mg, and subsequent least-squares regression analysis was then performed. When more data were available for the lower activity purity range, the high purity data were weighted such that the number of high and low purity data points had equivalent power in the regressions, meaning each data set consisted of the same number of data points in the upper and lower purity range. If this normalization is not performed, the low purity data points will not allow the regression function to capture the dynamics associated with the high purity enzyme. Since data cannot be homogenously sampled over the entire activity domain, some form of normalization is necessary to fit the data over the entire activity domain. Performing a regression on the raw data, simply provides an excellent fit to the more numerous low range activity data points, but completely misses the scarce high-activity-range data. For the alpha amylase data, the regression technique was modified to compensate for an excessive number of data points available in the low purity regime. These data points were so low in the activity domain that they were not allowing the regression function to capture the dynamics associated with the high purity enzyme. In other words, the numerous data points for amylase in the low activity range were controlling the dynamics of the regression in the high activity range such that the regression did not fit the moderate to high activity data points. Thus, the three lowest data points were not included in the regression to provide the best qualitative fit to the data over the entire activity domain. All regression analysis was performed with Microsoft Excel Version XP using the built-in least squares algorithm. Several regressions were attempted including

Table 1 Data collected for several enzymes from different suppliers.

Enzyme	Activity (U/mg)	Price ($/mg)	Supplier	Activity unit definition
Invertase	20	0.00005	13	1 U hydrolyzes 1 μmol of saccharose per min at pH=4.65 and *T*=25 °C
	100	0.01018	10	
	200	0.05700	10	
	300	0.50600	10	
Papain	1.5	0.00072	10	1 U hydrolyzes 1 μmol of *N*-benzoyl-L-arganine-ethyl ester per min at pH=6.2 and *T*=25 °C
	3	0.00097	10	
	12	0.64983	10	
Elastase	3	1.04333	11	1 U cleaves 1 μmol of *N*-succinyl-L-alanyl-L-alanyl-L-alanine-*p*-nitroanilide per min at pH=8.0 and *T*=25 °C
	8	4.00000	11	
	18	1910.0	12	
	18	1928.7	9	
Alpha Amylase	1.5	0.00023	10	1 U liberates 1 mg of maltose from starch in 3 min at pH=6.9 and *T*=20°C
	20	0.09240	10	
	30	0.00250	10	
	150	0.41700	10	
	380	0.16800	10	
	500	8.28000	10	
	1,000	97.200	10	
Cellulase	1	0.00223	10	1 U releases 0.01 mg of glucose per hour from microcrystalline cellulose at pH=5.0 and *T*=37 °C
	6	0.00500	9	
	25	0.02650	11	
	45	0.10800	11	
	50	0.91000	8	

power law, exponential, polynomial, and linear fits. Power law regressions seemed to fit the best across all enzymes. This regression expression also passes through point (0,0). A similar algorithm was used to determine the enzyme cost as a function of purchase amount. The data for pricing and mass were normalized to cost per mg and inverse mass to suit the regression type limitations of Microsoft Excel.

Results

We found that there appears to be a generalized model for enzyme cost per mg in terms of enzyme activity. From the analyzed data, this simple two-parameter model follows a power law trend as depicted in the previously described Eq. 2. Here, it is found that a is approximately 3.7±0.6 and tends to vary only slightly from enzyme to enzyme. The parameter a describes the separation (purification) cost of these enzymes, since an increase in activity results in a power law growth in the cost function (the rate at which is controlled by the value of the parameter a). γ is dependent on the method of enzyme production and can vary greatly between enzymes, as shown in Table 2. Indeed, higher production cost, such as mammalian cell culture, will result in a larger value for γ, whereas lower production cost, such as observed in enzymes isolated from plants, generally results in a lower value for γ. This can be seen in Table 2 above, which shows elastase, which was isolated from a pig pancreas has a γ value of 0.005, whereas cellulase isolated from a fungus has a γ value of 3.00E-08. The generalized behavior of the enzymes quantified by γ and a is depicted in Figs. 1, 2, 3, 4 and 5.

Table 2 A list of fitted a and γ values and regression R^2 values for the simplified two parameter objective function model given in Eq. 2.

Parameter values			
	a	γ	R^2
Papain	3.48	8.00E-05	0.92
Elastase	4.32	0.005	0.95
Invertase	3.29	3.00E-09	0.99
Cellulase	4.18	3.00E-08	0.77
Alpha Amylase	3.45	3.00E-09	0.81

The data used to compute these parameters and generate the subsequent plots were collected from [6–10].

The equation relating the bulk price per mg is:

$$\Phi = \delta(M)^b \tag{4}$$

where b=−0.2 and δ=\$2.049 per $mg^{0.8}$ for papain (Figure not shown).

The negative parameter value for b, in Eq. 4, depicts the savings achieved by purchasing in bulk. Significant savings can be achieved when buying bulk enzymes, but this gain may be offset by storage costs. We have not quantified the tradeoffs associated with buying bulk and storage requirements (in Eq. 3). We expect that the parameters for capital costs of storage will vary greatly from facility to facility.

Figure 6 further elaborates on the enzyme cost relationship by segregating the enzyme activity domain into three regimes characterized by the method of separation.

In Fig. 6, the generalized function used was $y=4\times10^{-6}\times x^{3.7}$, where x is the activity per mg, and y is the cost per mg of the generalized enzyme. The parameter values used, γ=4×

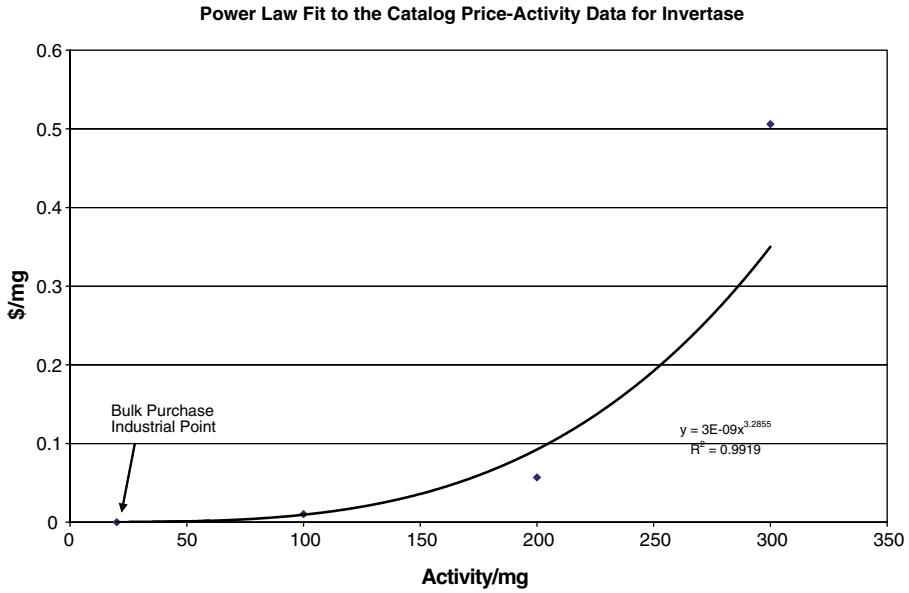

Fig. 1 Power law fit to the catalog price-activity data for invertase

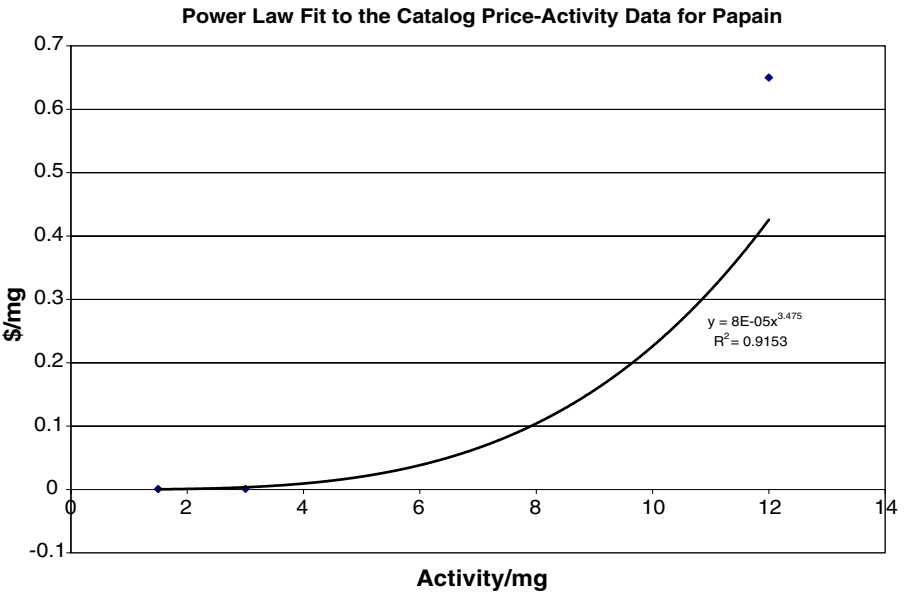

Fig. 2 Power law fit to the catalog price-activity data for papain

10^{-4} and a=3.7 were chosen arbitrarily, but lie within the realm of the parameters fitted to the collected data (see Table 2).

We now try to further interpret the possible meaning, in terms of separation cost, of Eq. 3. To do this analysis, we break up the enzyme activity domain into three parts: the low activity-prelinear region, the moderate activity linear region, and the high activity exponential region. We do this based on a general consensus seen within all the collected

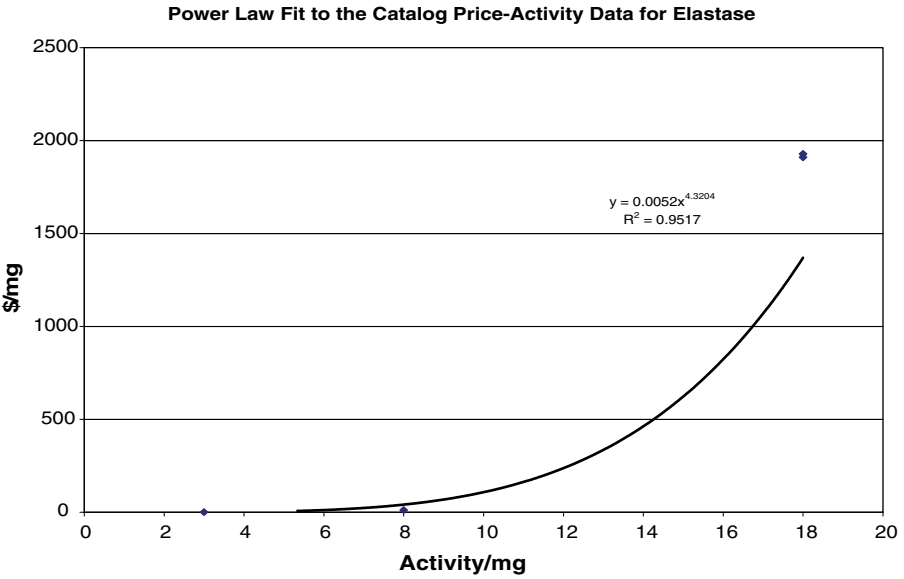

Fig. 3 Power law fit to the catalog price-activity data for elastase

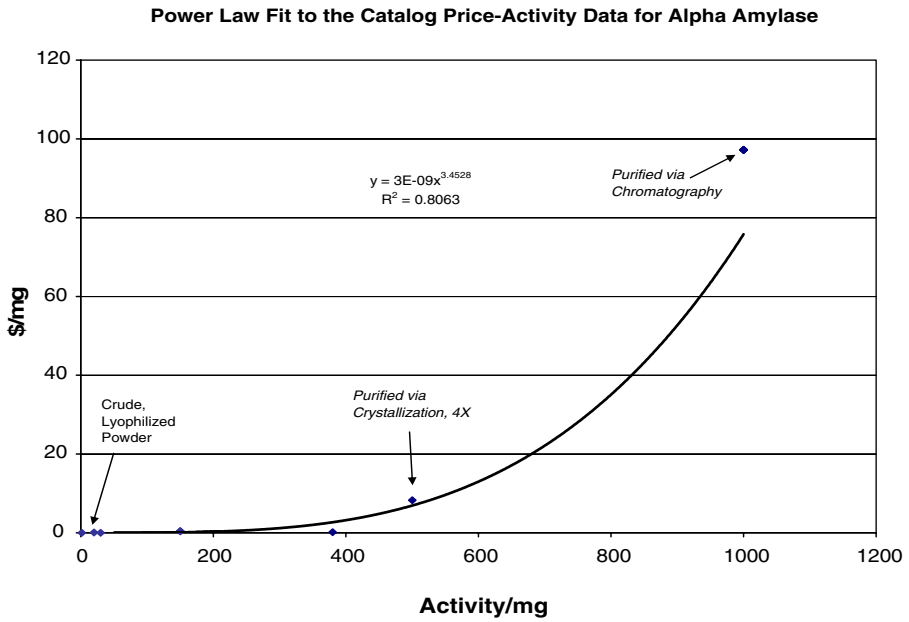

Fig. 4 Power law fit to the catalog pice-activity data for alpha amylase

data that the activity (and price) of each enzyme correlates with the separation techniques used. This correlation can be most easily seen for actual amylase data in Fig. 4.

The initial low activity prelinear region, is characterized by little change in cost per mg as activity per mg is increased. This region appears to be characterized by low purity

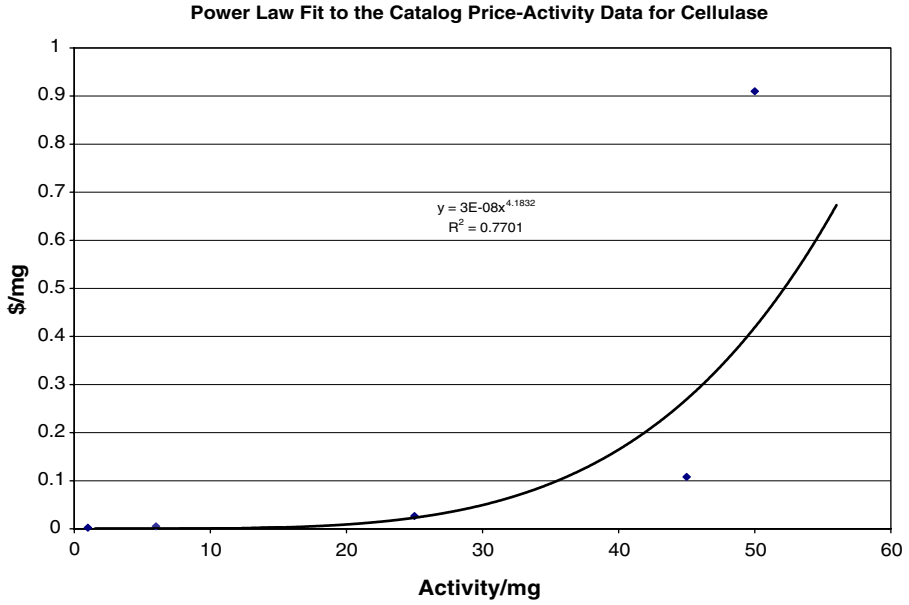

Fig. 5 Power law fit to the catalog price-activity data for cellulase

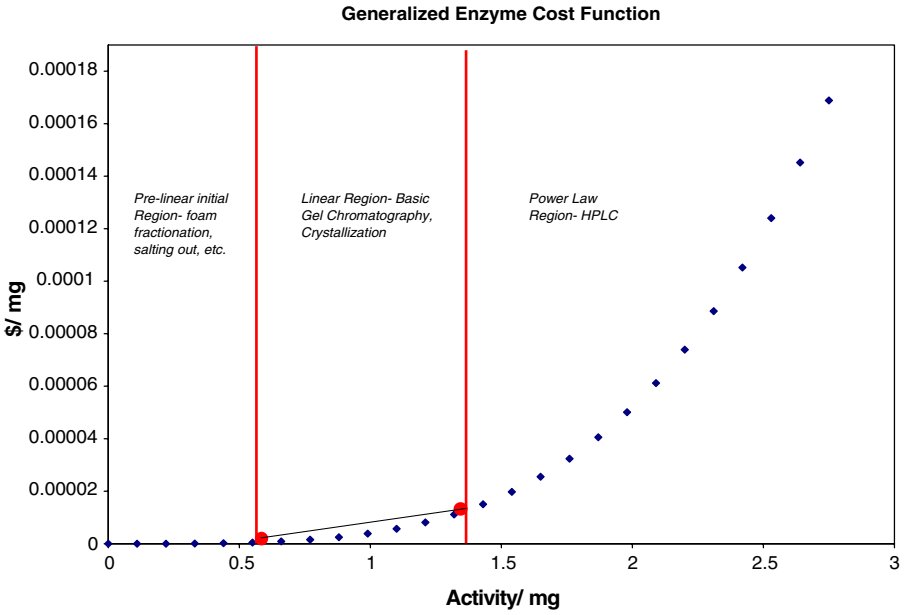

Fig. 6 Generalized enzyme cost function

enzymes harvested from plants (or other readily available sources such as fermentation broths) and purified via traditional low cost processes such as foam fractionation or salting out. These purification processes can be extended at little cost to generate increased enzyme activity. Since the additional processing within this region requires little additional capital cost and only modest operating costs, the change in enzyme cost per mg may be marginal.

The second region, the moderate activity region (linear), is characterized by nearly a linear change, over the narrow band of the $4\times10^{-6}\times x^{3.7}$ function, in cost per mg as activity is increased. This region is generally characterized by moderately pure enzymes that have been purified by more costly processing than the prelinear region. It appears that the majority of enzymes considered in this domain are purified via gel chromatography. The capital and operating costs of this processing are typically higher than the costs associated with foam fractionation and salting out processes [7]. Additional processing, for additional enzyme activity in this middle region where there is moderate activity, requires larger chromatography columns or larger affinity gels whose costs tend to increase linearly in price.

The third domain is characterized by the observable (empirical) power law growth in enzyme cost per mg with respect to activity. This domain extends to the highest purity of the desired enzyme and is generally comprised of enzymes that require expensive processing. The overall processing is thus generally a combination of salting out, foam fractionation, and multistage gel chromatography, but often includes a final purification step, which is generally HPLC. The capital and operating costs for HPLC are significantly higher than that of the other mentioned purification processes. Additional purification through HPLC generally requires running a sample multiple times on additional columns. During each step of HPLC, a large fraction of the enzyme is lost due to the inefficiencies of the process, making the increased cost per mg behave like a power law function when several steps are used in series.

The derivative of the generalized cost function shows the incremental change in cost as activity is increased. This derivative function, represented as $y=1.5\times10^{-5}\times x^{2.7}$ leads to some key observations. For example, the power law exponent is still larger than one. This observation holds for every enzyme analyzed in this study, which increases our confidence in the analysis of these enzymes. Even with a power greater than one, the cost will only change minimally with activities smaller than unity. Above unity, there will be a small region where price will increase only marginally as a function of activity. However, regardless of the value of the preexponential parameter, at some threshold where the activity per mass is greater than unity, the enzyme cost will begin increasing significantly and exponentially as a function of enzyme activity.

Conclusions

The significant conclusion and most interesting observation of this study is that there appears to be one simple function that estimates the cost of an enzyme in $ per mg as a function of activity, $\Phi=\gamma\,(A)^a$, where a appears to be a generalizable exponent of the order 3.7. The generalizable function appears to fit best over the low to moderate activity range, but often under-predicts enzyme cost in the high activity range. This implies that there may not be a simple two-parameter model that accurately captures the dynamics of enzyme activity related to cost over the entire activity domain. Additional data in the moderate to high activity range are needed to further characterize the limitations of Eq. 2. However, the function developed in this analysis is still useful as an estimation tool where little or no real data may be known.

The generalizable function can furthermore be analyzed in terms of three characteristic domains describing the enzyme cost with respect to activity. These domains can be useful when purchasing enzymes for industrial or laboratory purposes. For example, if a low purity enzyme is being purchased, it is logical to purchase the enzyme at that purity equivalent to the end of the prelinear phase of the generalized enzyme curve, denoted by the critical point located near an activity per mass ratio of 0.5 on the generalized enzyme cost function plot (Fig. 6), giving the buyer the most amount of active enzyme at a very low cost. A complimentary strategy for high purity enzymes would be to choose the activity of the upper end of the moderate activity linear range (the lower end of the high activity exponential domain). If processing cost is dependent on enzyme impurities, the cost of processing can be compared to the cost of increasing enzyme activity, by using Eq. 2 in a manner to minimize overall costs. Since the parameter a generally ranges between 3 and 4 between enzymes, a twofold increase in activity is generally associated with an eightfold increase in cost.

Since the parameters determined within this study were empirically computed from current market prices, these parameters are likely to change over long time-scales given the dynamic nature of any economic market. The enzymes used for this analysis are generally widely used proteins that, therefore, have fairly stable, well-developed markets. Enzymes used in small quantities for less industrial purposes may not follow similar trends since market fluctuations can be significant. This allows industrial enzymes to be best suited for this type of study since their market will be less dynamic than other enzymes that are still within market development. However, economies of scale will still play a factor, especially in enzymes that are expected to see a wide increase in use and production scale-up, such as cellulase. These changes will likely result in a scaling factor that will linearly decrease the value of γ. This observation is drawn from analyzing scale-up cost for typical bioprocessing

operations involved in enzyme purification that show linear cost reductions as scale increases [14]. Similarly, this observation can be made by looking at typical equipment scale-up cost. One of the most common protein purification methods is liquid chromatography [15]. Chromatography columns tend to scale up linearly since the capital cost for the column is minimal compared to the high resin cost, which is required in directly proportional amounts to column throughput. Scale-up of enzyme production is generally sublinear, but the data presented in this analysis demonstrates that purification costs tend to dominate overall product cost [15]. Regardless of changes in scaling due to market growth and scale-up, the functional forms determined for the relationship of enzyme activity and purchase amount to price should hold. The independent variables within Eq. 2 were shown to be directly related to the process variables for enzyme production. Although the model under-predicts enzyme cost for the high activity range, Eq. 2 still provides a useful tool for engineers to estimate enzyme cost based on activity.

References

1. Cherry, J., Ko, S., Grainger, R., Prokop, A., & Tanner, R. (2000). Developing an objective function to characterize the tradeoffs in salting out and the foam and droplet fractionation processes. *Brazilian Journal of Chemical Engineering, 17*, 233–238.
2. Zhang, Q., Lo, C.-M. & Ju, L.-K. (2007). Factors affecting foaming behavior in Cellulase fermentation by *Trichoderma reesei* Rut C-30. *Bioresource Technology, 98*(4), 753–760.
3. Lo, C.-M., Zhang, Q., Lee, P., & Ju, L-K. (2005). Cellulase production by *Trichoderma reesei* using Sawdust Hydrolysate. *Applied Biochemistry and Biotechnology, 121–124*, 561–573.
4. Graff, G., "Pharma Market Spur Increased Enzyme Demand" 3/1/2007 http://www.purchasing.com/article/CA6419082.html.
5. Poulsen, P. (1987). Trends in industrial applications of enzymes. Annals of the New York Academy of Sciences. *Enzyme Engineering, 501*, 413–419.
6. Nilesh; A., Kamat, M., & Arvind, L. (2004). Expanded bed affinity purification of bacterial α-amylase and Cellulase on composite substrate analogue-Cellulose matrices. *Process Biochemistry, 39*, 565–570.
7. Zhang, Q., Lo, C.-M., & Ju, L.-K. (2006). Affinity foam fractionation of Trichoderma Cellulase. *Applied Biochemistry and Biotechnology, 129–132*, 1051–1065.
8. Carolina Biochemical Corporation. http://www.carolina.com. (catalog information on enzymes). 4/1/07.
9. Fisher Scientific. http://new.fishersci.com. (catalog information on enzymes). 4/1/07.
10. Sigma-Aldrich Corporation. http://wwwsigmaaldrich.com. (catalog information on enzymes). 4/1/07.
11. Worthington Biochemical Corporation. http://www.worthington-biochem.com. (catalog information on enzymes). 4/1/07.
12. Elastin Products Company, Inc. http://www.elastin.com (catalogue information on enzymes) 4/1/07.
13. "Enzymatic Production of Invert Sugar", Ensymm Consulting for Biotechnology (2007). http://www.ensymm.com/pdf/ensymmProjectstudyreportInversugarproduction.pdf.
14. Cacciuttolo, M., & Arunakumari, A., " Scale-Up Considerations for Biotechnology-Derived Products" http://biomedical.rutgers.edu/doc/Scale%20Up%20of%20Biotechnology%20Products.pdf.
15. Blanch, H., & Clark, D. (1997). *Biochemical Engineering*. New York, NY: Marcel Dekker, pp. 678–682.