ELSEVIER

Journal of Molecular Catalysis B: Enzymatic 21 (2003) 3–8

JOURNAL OF MOLECULAR CATALYSIS B: ENZYMATIC

www.elsevier.com/locate/molcatb

Rational engineering of the stability and the catalytic performance of enzymes

Sigrid Gåseidnes [a], Bjørnar Synstad [a], Jens E. Nielsen [b], Vincent G.H. Eijsink [a,*]

[a] *Department of Chemistry and Biotechnology, Agricultural University of Norway, P.O. Box 5040, 1432 Ås, Norway*

[b] *Howard Hughes Medical Institute & Department of Chemistry and Biochemistry, Mail code 0365, University of California, San Diego, La Jolla, CA 92093, USA*

Abstract

Rational engineering of enzyme properties has an enormous potential but is hampered by limitations in our understanding of the structural determinants of these properties. Recent examples of efforts in rational engineering illustrate the current state of the art.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Protease; Autolysis; Chitinase; Electrostatics; Unfolding

1. Introduction

The applicability of naturally occurring enzymes in industrial processes is often hampered by the fact that these enzymes primarily are adapted to nature and not to industry. Thus, successful application of an enzyme in an industrial process is often preceded by efforts in enzyme engineering that are aimed at tailoring specific enzyme properties. Unfortunately, rational design of improved enzyme variants is rather complicated as illustrated by the fact that there are a lot of carefully designed mutant enzymes around that just do not work (and that usually do not appear in the literature). The complexity of mutant design is even more clearly illustrated by recent results from directed evolution studies: enzyme variants with new properties have been obtained, but many mutations selected in these variants are not obvious from a theoretical point of view [1–4]. Despite apparent complexities, impressive examples of rational enzyme design do exist, some of which are discussed below.

2. Stability

Many enzymes are not sufficiently stable under process conditions, which makes stability one of the most often engineered protein properties. Mutational studies of reversible thermal unfolding of small proteins have provided quite some fundamental insight in structural factors that determine protein stability [5–8]. These studies have established several important strategies for protein stabilization, such as the introduction of disulfide bridges [9], helix capping and optimization [10–13] and “entropic stabilization” (Gly → Xxx; Xxx → Pro; [14]). In addition, there exist numerous attempts to derive “rules” for protein stabilization from sequence comparisons of naturally occurring homologous proteins with varying stabilities (see [15,16] for recent examples). Such statistical rules are

* Corresponding author. Tel.: +47-64-949472; fax: +47-64-947720.
E-mail address: vincent.eijsink@ikb.nlh.no (V.G.H. Eijsink).

1381-1177/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S1381-1177(02)00136-4

often lacking a theoretical explanation and examples of successful application of such rules are scarce.

Despite the existence of some experimentally verified rules [5–8], rational improvement of the stability of an enzyme in an industrial setting is by no means well-established. One major problem relates to the fact that protein inactivation (by temperature or other denaturing factors) is likely to result from (undefined) irreversible processes that are governed by local, as opposed to global unfolding (e.g. [17–20]). If this is the case, the main problem for the protein engineer is to find "weak spots" in the protein that are involved in the local unfolding processes that determine the rate of the irreversible process (usually aggregation or proteolysis).

The importance of local processes is well-illustrated by the design of a hyperstable variant of the thermolysin-like protease from *Bacillus stearothermophilus* (TLP-ste) [18,21,22]. In the initial phase of the design process, it was noticed that most rationally designed mutations had remarkably small effects on thermal stability [23]. On the basis of these and other results it was suggested that thermal inactivation of TLP-ste is governed by local unfolding processes that precede the irreversible step (autolysis ([23,24])). Further mutagenesis studies revealed that mutations in one particular region in the N-terminal domain had profound effects on stability [18], suggesting that this region was the weak spot in TLP-ste (Fig. 1). Once this weak spot had been identified, rational design of stabilizing mutations in TLP-ste became remarkably successful [24] and hyperstable variants of TLP-ste could be engineered [21]. The weak spot probably plays a prominent role in the early steps of TLP-ste unfolding [20]; in fact the term "unfolding region" has been used to describe this stability-determining part of the protein [25]. The extent of unfolding that precedes autolysis is not known but it is likely that larger parts of the N-terminal domain of the protein are involved [22,23].

If the stability of a protein depends on global unfolding, the effects of independent mutations on stability would normally be additive [20,26]. If, on the other

Fig. 1. C-α trace of TLP-ste. The locations of three stabilizing point mutations in the unfolding region are illustrated by the side chains of the introduced residues (Phe63, Pro65, Pro69). An engineered disulfide bridge (Gly8- > Cys; Asn60- > Cys) that alone stabilized TLP-ste by 16.7 °C [24] is drawn as a dotted line and indicated by an arrow. The bold line represents the C-α trace between residues 56 and 69. The dotted part of the trace indicates the region where mutations were introduced to create TLP-ste variants with two unfolding regions [20].

hand, local (partial) unfolding processes are important, stabilizing mutations will be non-additive unless they are in the principal unfolding region. To illustrate this Vriend et al. [20] created four pseudo wild-types of TLP-ste in which a second unfolding region had been created by mutations in the C-terminal domain of the enzyme (Fig. 1). The distances between the two regions were in the order of 30–40 Å meaning that direct interactions between the two regions could be excluded. Subsequently, the effects of stabilizing one or both of the two weak regions were studied. The results illustrated two things.

(1) *The "enough is enough" effect*. In one variant the second unfolding region was created by three independent mutations that together destabilized the wild-type by 7.1 °C. Wild-type stability could be retained by each combination of two of the three reverse mutations (and up to 7.0 °C could be obtained by only one of them). This illustrates that once a region of a protein has become so stable that its unfolding no longer contributes to the thermal inactivation process, new mutations in this region have no effect.
(2) *Overadditivity*. Combining stabilizing mutations in each of the two unfolding regions gave large synergistic effects. For example, in one case, stabilizing mutations in the two regions that yielded +2.3 and +4.1 °C individually yielded +14.6 °C when combined. This illustrates that the individual effects of stabilizing mutations in each of the regions are limited by the fact that there is a second weak region.

The fact that local unfolding is rate-limiting in irreversible inactivation has important implications for the rational design of stabilizing mutants. Mutation strategies should aim at finding the "weak spot" of the protein, and then use established strategies for protein stabilization. When all weak spots of the protein are eliminated a more globally oriented mutation strategy may be adopted.

Finding the weak spot or unfolding region in a protein obviously is not a simple task and will most likely need to be based on random approaches. Clearly, the combination of rational design with random mutagenesis, directed evolution and high throughput screening of mutants may be beneficial in this respect (see also, e.g. [27,28]).

Fortunately, recent efforts in enzyme engineering have also yielded insights that enlighten the task. First of all, it is now well-established that only a limited number of well-designed mutations is needed to achieve drastic improvements in stability properties [21,29–32]. Furthermore, opposed to naturally occurring thermostable enzymes, engineered stable enzymes seem to keep their activity at low temperature (e.g. [21,29–32]). This difference between naturally occurring and engineered enzymes is probably due to the fact that selection pressures in nature are rather different from those that we apply in the laboratory. The latter nicely illustrates that the protein engineer has access to what is physically possible and not only to what is biologically relevant.

Increases in computational power and in the number of well-characterized protein variants have spurred the development of algorithms capable of predicting stabilizing mutations in small proteins that undergo reversible unfolding (see e.g. [33–35]). These methods are likely to be most successful when used for calculating stabilizing mutations in the interior of the protein. When it comes to calculating mutations on the surface or in flexible parts of the protein, the rate of success is smaller [34–36]. As with all computational methods based on energy-calculations, there are limitations imposed by the size of the protein and by possible shortcomings in the force-field. These computational methods may not be generally applicable to the stabilization of industrial enzymes, where stability is likely to be determined by (undefined) partial unfolding processes followed by e.g. aggregation. On the other hand, impressive methods for "automated" protein stabilization are evolving [33] which may find their way to industry.

3. Catalysis

Enzymes catalyze an enormous variety of reactions in complex catalytic centers using mechanisms that are often poorly understood. One major complication concerns the role of structural motions (involving a few residues or complete loops/domains; e.g. [37–39]), which clearly are important but which are difficult to rationalize/engineer. Also catalytic mechanisms differ from enzyme to enzyme, making it intrinsically difficult to devise a set of more or less "general"

strategies that may be tried in efforts to engineer catalysis.

There are numerous examples in the literature that illustrate the delicacies of catalysis and the sensitivity of catalysis to minor changes in active site structure (e.g. [40,41]). To underline this, Mesecar et al. [40] introduced the term "orbital steering" after having observed that small structural perturbations in the active site of isocitrate hydrogenase had large effects on catalytic activity. The observed effects were suggested to be due to perturbations no larger than small changes in the positioning of overlapping orbitals in enzyme and substrate. The results of studies aimed at converting trypsin into chymotrypsin by (semi-) rational design provide another beautiful illustration of the complexity of engineering catalysis. From a structural point of view, this conversion seemed rather straightforward in the sense that it seemed attainable by a few mutations in the specificity-determining S_1 subsite. It turned out, however, that several "non-predictable" mutations quite far away from the S_1 subsite were needed to obtain significant chymotrypsin activity ([41,42], and references therein).

By combining mutagenesis, structural studies, enzymology and computational methods McIntosh et al. [43–45] recently investigated electrostatic interactions involved in catalysis in a xylanase from *Bacillus circulans*. They show that the pK_a values of catalytic glutamate residues depend on a variety of interactions involving most of the residues in the active site. Some of these interactions change during the catalytic cycle, which results in pK_a cycling during catalysis [43]. Most importantly, in their most recent paper, Joshi et al. conclude that attempts to engineer the catalytic performance of the enzyme should probably focus on mutations at positions somewhat outside the active site. Such mutations would leave the structure of the active site unchanged but could still have considerable effects on catalysis via longer range electrostatic interactions [44].

Whereas the importance of electrostatic interactions for catalysis is well-established [46], the rationalization of these interactions still is a major challenge. It is possible to treat local short-range electrostatic forces using quantum mechanical techniques [47]. However, quantum mechanical calculations (QMC) are limited to treating only 100 atoms in a single calculation, and it is, therefore, not possible to use QMC to describe long-range electrostatic forces in enzymes. Consequently, researchers have been forced to use less detailed methods such as finite-difference solvers of the Poisson–Boltzmann equation (PBE). PBE solvers [48–50] have been used extensively to analyze electrostatic phenomena in proteins. These methods do, however, suffer from a number of inherent complexities. A more severe limitation is the description of protein flexibility by the use of a single (or in some cases two) protein dielectric constant(s). A single dielectric constant is clearly not able to accurately describe the dielectric response in an enzyme active site, and it is also inherently problematic to describe microscopic electrostatic effects using a macroscopic quantity as the dielectric constant (see Sham et al. [51] for an excellent discussion on protein relaxation and dielectric constants in protein electrostatic calculations).

The importance of long-range electrostatic interactions has been confirmed by mutagenesis studies in a number of enzymes, e.g. subtilisin [52], thermolysin-like proteases [53], and chitinases [54]. However, there are no convincing examples of fully *rational* engineering of catalysis via manipulation of long-range electrostatic interactions. The delicacy of catalysis that makes rational engineering difficult is also a potential goldmine to the protein scientist. It is becoming increasingly clear that enzymes evolve in nature by relatively minor modifications of active-site structures [1,55]. This is a challenge because it inclines that small changes give large effects, but it also makes redesign of an enzyme to get different catalytic specificities possible without large adjustments of the enzyme structure. This is well-illustrated by recent studies on redesign of enzymatic activities of $(\beta/\alpha)_8$ barrel proteins [56]. The $(\beta/\alpha)_8$ motif is unique in that it occurs in roughly ten percent of all enzymes, representing a broad spectrum of enzymatic activities. Further studies aimed at rational redesign of a variety of $(\beta/\alpha)_8$ enzymes may provide important clues for rational engineering of enzymes in general.

4. Conclusions

New techniques for creating large and diverse libraries of protein variants and robotized high-throughput screening for desired properties provide

important new tools for the enzyme engineer [1]. Although these new strategies (often referred to as "directed evolution") have yielded impressive results, they have also revealed the continued importance of rational design. The combination of rational design and random approaches is probably the most powerful and effective [1,27,57]. Random approaches will be particularly important in cases where there is insufficient structural information available as well as for the engineering of parameters that are difficult to rationalize. Rational design will remain of utmost importance when it comes to engineering properties that are difficult to screen for (e.g. enantioselectivity; [28]). It will also be used to create an initial set of mutants as a starting point for "fine-tuning" by directed evolution methods. Methods for rational design of effective site-directed mutations and, eventually, complete active enzymes will profit from the results of ongoing massive efforts in protein X-ray crystallography (often referred to as "structural genomics"; [58]). A drastic increase in the protein structure database will also help in further unraveling of the structural determinants of enzyme stability and catalytic performance.

References

[1] F.H. Arnold, Nature 409 (2001) 253.
[2] K. Liebeton, A. Zonta, K. Schimossek, M. Nardini, D. Lang, B.W. Dijkstra, M.T. Reetz, K.E. Jaeger, Chem. Biol. 7 (2000) 709.
[3] S. Oue, A. Okamoto, T. Yano, H. Kagamiyama, J. Biol. Chem. 274 (1999) 2344.
[4] B. Spiller, A. Gershenson, F.H. Arnold, R.C. Stevens, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 12305.
[5] B.W. Matthews, Annu. Rev. Biochem. 62 (1993) 139.
[6] B.W. Matthews, Adv. Protein Chem. 46 (1995) 249.
[7] A. Shaw, R. Bott, Curr. Opin. Struct. Biol. 6 (1996) 546.
[8] A.R. Fersht, L. Serrano, Curr. Opin. Struct. Biol. 3 (1993) 75.
[9] M. Matsumura, G. Signor, B.W. Matthews, Nature 342 (1989) 291.
[10] M. Blaber, X.J. Zhang, B.W. Matthews, Science 260 (1993) 1637.
[11] H. Nicholson, W.J. Becktel, B.W. Matthews, Nature 336 (1988) 651.
[12] L. Serrano, A.R. Fersht, Nature 342 (1989) 296.
[13] L. Serrano, J.L. Neira, J. Sancho, A.R. Fersht, Nature 356 (1992) 453.
[14] B.W. Matthews, H. Nicholson, W.J. Becktel, Proc. Natl. Acad. Sci. U.S.A. 84 (1987) 6663.
[15] A.M. Facchiano, G. Colonna, R. Ragone, Protein Eng. 11 (1998) 753.
[16] P.J. Haney, J.H. Badger, G.L. Buldak, C.I. Reich, C.R. Woese, G.J. Olsen, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 3578.
[17] B.A. Chrunyk, R. Wetzel, Protein Eng. 6 (1993) 733.
[18] V.G.H. Eijsink, O.R. Veltman, W. Aukema, G. Vriend, G. Venema, Nat. Struct. Biol. 2 (1995) 374.
[19] J.M. Finke, M. Roy, B.H. Zimm, P.A. Jennings, Biochemistry 39 (2000) 575.
[20] G. Vriend, H.J. Berendsen, B. van den Burg, G. Venema, V.G.H. Eijsink, J. Biol. Chem. 273 (1998) 35074.
[21] B. Van den Burg, G. Vriend, O.R. Veltman, G. Venema, V.G.H. Eijsink, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 2056.
[22] V.G.H. Eijsink, G. Vriend, B. Van den Burg, Biocatal. Biotransform. 19 (2001) 443.
[23] G. Vriend, V.G.H. Eijsink, J. Comput. Aided Mol. Des. 7 (1993) 367.
[24] J. Mansfeld, G. Vriend, B.W. Dijkstra, O.R. Veltman, B. Van den Burg, G. Venema, R. Ulbrich-Hofmann, V.G.H. Eijsink, J. Biol. Chem. 272 (1997) 11152.
[25] J. Mansfeld, G. Vriend, B. Van den Burg, V.G. Eijsink, R. Ulbrich-Hofmann, Biochemistry 38 (1999) 8240.
[26] J.A. Wells, Biochemistry 29 (1990) 8509.
[27] R. Chen, Trends Biotechnol. 19 (2001) 13.
[28] U.T. Bornscheuer, M. Pohl, Curr. Opin. Chem. Biol. 5 (2001) 137.
[29] H. Zhao, F.H. Arnold, Protein Eng. 12 (1999) 47.
[30] J.C. Williams, J.P. Zeelen, G. Neubauer, G. Vriend, J. Backmann, P.A. Michels, A.M. Lambeir, R.K. Wierenga, Protein Eng. 12 (1999) 243.
[31] L. Giver, A. Gershenson, P.O. Freskgard, F.H. Arnold, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 12809.
[32] K. Miyazaki, P.L. Wintrode, R.A. Grayling, D.N. Rubingh, F.H. Arnold, J. Mol. Biol. 297 (2000) 1015.
[33] S.M. Malakauskas, S.L. Mayo, Nat. Struct. Biol. 5 (1998) 470.
[34] J. Funahashi, K. Takano, K. Yutani, Protein Eng. 14 (2001) 127.
[35] D. Gilis, M. Rooman, J. Mol. Biol. 272 (1997) 276.
[36] D. Gilis, M. Rooman, Protein Eng. 13 (2000) 849.
[37] A. Kohen, J.P. Klinman, Chem. Biol. 6 (1999) R191.
[38] M.J. Osborne, J. Schnell, S.J. Benkovic, H.J. Dyson, P.E. Wright, Biochemistry 40 (2001) 9846.
[39] O.R. Veltman, V.G.H. Eijsink, G. Vriend, A. de Kreij, G. Venema, B. Van den Burg, Biochemistry 37 (1998) 5305.
[40] A.D. Mesecar, B.L. Stoddard, D.E. Koshland Jr, Science 277 (1997) 202.
[41] J.J. Perona, L. Hedstrom, W.J. Rutter, R.J. Fletterick, Biochemistry 34 (1995) 1489.
[42] L. Hedstrom, L. Szilagyi, W.J. Rutter, Science 255 (1992) 1249.
[43] L.P. McIntosh, G. Hand, P.E. Johnson, M.D. Joshi, M. Korner, L.A. Plesniak, L. Ziser, W.W. Wakarchuk, S.G. Withers, Biochemistry 35 (1996) 9958.
[44] M.D. Joshi, G. Sidhu, J.E. Nielsen, G.D. Brayer, S.G. Withers, L.P. McIntosh, Biochemistry 40 (2001) 10115.
[45] M.D. Joshi, G. Sidhu, I. Pot, G.D. Brayer, S.G. Withers, L.P. McIntosh, J. Mol. Biol. 299 (2000) 255.
[46] B. Honig, A. Nicholls, Science 268 (1995) 1144.

[47] R.A. Friesner, M.D. Beachy, Curr. Opin. Struct. Biol. 8 (1998) 257.
[48] A. Nicholls, B. Honig, J. Comput. Chem. 12 (1991) 435.
[49] A. Nicholls, K.A. Sharp, B. Honig, Proteins 11 (1991) 281.
[50] J.D. Madura, J.M. Briggs, R.C. Wade, M.E. Davis, B.A. Luty, A. Ilin, J. Antosiewicz, M.K. Gilson, B. Bagheri, L.R. Scott, J.A. McCammon, Comp. Phys. Comm. 91 (1995) 57.
[51] Y.Y. Sham, I. Muegge, A. Warshel, Biophys. J. 74 (1998) 1744.
[52] S.E. Jackson, A.R. Fersht, Biochemistry 32 (1993) 13909.
[53] A. De Kreij, B. Van den Burg, G. Venema, G. Vriend, V.G.H. Eijsink, J.E. Nielsen, J. Biol. Chem. 277 (2002) 15432.
[54] D.M.F. Van Aalten, D. Komander, B. Synstad, S. Gåseidnes, M.G. Peter, V.G.H. Eijsink, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 8979.
[55] P.C. Babbitt, J.A. Gerlt, Adv. Protein Chem. 55 (2000) 1.
[56] C. Jürgens, A. Strom, D. Wegener, S. Hettwer, M. Wilmanns, R. Sterner, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 9925 (as explained at L270).
[57] J.R. Cherry, M.H. Lamsa, P. Schneider, J. Vind, A. Svendsen, A. Jones, A.H. Pedersen, Nat. Biotechnol. 17 (1999) 379.
[58] E. Abola, P. Kuhn, T. Earnest, R.C. Stevens, Nat. Struct. Biol. 7 (2000) 973.