ELSEVIER

Chemistry & Biology 8 (2001) 891–898

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Novel enzyme activities and functional plasticity revealed by recombining highly homologous enzymes

SunAi Raillard [a], Anke Krebber [a], Yonghong Chen [a], Jon E. Ness [a], Ericka Bermudez [a], Rossana Trinidad [a], Rachel Fullem [a], Christopher Davis [a], Mark Welch [a], Jennifer Seffernick [b], Lawrence P. Wackett [b], Willem P.C. Stemmer [a], Jeremy Minshull [a,] *

[a] *Maxygen, Inc., 515 Galveston Drive, Redwood City, CA 94063, USA*

[b] *Department of Biochemistry, Molecular Biology and Biophysics and Biological Process Technology Institute, University of Minnesota, St. Paul, MN 55108-6106, USA*

Received 27 February 2001; revisions requested 26 April 2001; revisions received 12 June 2001; accepted 26 June 2001

First published online 27 July 2001

Abstract

Background: Directed evolution by DNA shuffling has been used to modify physical and catalytic properties of biological systems. We have shuffled two highly homologous triazine hydrolases and conducted an exploration of the substrate specificities of the resulting enzymes to acquire a better understanding of the possible distributions of novel functions in sequence space.

Results: Both parental enzymes and a library of 1600 variant triazine hydrolases were screened against a synthetic library of 15 triazines. The shuffled library contained enzymes with up to 150-fold greater transformation rates than either parent. It also contained enzymes that hydrolyzed five of eight triazines that were not substrates for either starting enzyme.

Conclusions: Permutation of nine amino acid differences resulted in a set of enzymes with surprisingly diverse patterns of reactions catalyzed. The functional richness of this small area of sequence space may aid our understanding of both natural and artificial evolution. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Enzyme activity; Functional plasticity; Homologous enzymes; Triazine hydrolases

1. Introduction

The unpredictability with which protein function alters with sequence has led to the development of directed evolution as a laboratory process that mimics some of the features of natural evolution in order to obtain proteins, metabolic pathways, viruses, cells or organisms with desired properties. Variants are typically produced by mutation and/or recombination, and the libraries thus produced are screened for their ability to perform a specific function. Recursive application of this procedure has been used successfully to modify physical and catalytic properties of enzymes such as pH optima [1], thermotolerance [2], solvent stability [3], stereoselectivity [4], catalytic activity and substrate specificity [5], as well as toxicity resistance mechanisms in bacteria [6] and the host range and stability of viruses [7,8].

Directed evolution has so far been used almost exclusively as a tool for engineering proteins; to date little predictive insight has been gained from directed evolution experiments. The effects of mutations are often interpreted in terms of their locations in and possible effects upon a protein's crystal structure [4,9,10] which may provide some information regarding the structure–function relationship for a specific protein but is seldom used to modify a protein further.

Directed evolution may be a direct way to provide information about the distribution of function in sequence space. Although usually applied to produce a single specific phenotype, recombination-based methods such as DNA shuffling [11] could produce an enormous amount of sequence and functional diversity [1,12]. We wished to test whether DNA shuffling could be used to generate

* Corresponding author.
E-mail address: jeremy.minshull@maxygen.com (J. Minshull).

1074-5521/01/$ – see front matter © 2001 Elsevier Science Ltd. All rights reserved.
PII: S1074-5521(01)00061-8

information about the potential catalytic activities, and in particular the substrate specificities, accessible in a small defined region of sequence space. Here we ask two specific questions: (1) can correlations between sequences and functional information be used predictively without structural information and (2) how densely do different functionalities appear in real (rather than modeled) sequence space.

We chose as a test system the closely related but distinct enzymes AtzA and TriA, which initiate hydrolysis of *s*-triazines by dechlorination or deamination respectively [13–15]. *s*-Triazines have recently appeared in the environment because of their uses as herbicides, pesticides and in plastics [16,17]. Degradation of *s*-triazines by microbes was observed relatively rapidly after their introduction [18]. AtzA was isolated from an atrazine-degrading *Pseudomonas* species found at an atrazine spill-site [19]. Atrazine is a herbicide that has been used extensively for several decades in the USA for control of broad-leaved weeds in maize, sugar cane and sorghum; *atzA* genes have been found in many geographically distinct sites where atrazine is used [20]. Melamine is used as a plasticizer in the production of products as diverse as fibreboard and tableware. Although not spread into the environment deliberately, melamine spill-sites also exist and it is from municipal sewage that a melamine-degrading *Pseudomonas* strain and the *triA* gene were obtained [15,21].

AtzA and TriA catalyze dechlorination and deamination reactions respectively (Fig. 1), with little overlap in substrate preference despite differing by only nine amino acids out of a total of 475 [15,22]. It is likely that *s*-triazine degradation pathways have evolved very recently, probably in response to the environmental appearance of *s*-triazines. The small difference in sequences between AtzA and TriA strongly suggest a very recent common ancestor. Ellington and colleagues have recently shown that laboratory evolution of a β-glucuronidase to a β-galactosidase proceeds through generalist intermediates with broader substrate specificities [23]. We might therefore expect to see intermediates between the naturally evolved TriA and AtzA enzymes that also possess broader substrate specificities. Here we describe an exploration of the functional plasticity of the region of sequence space around AtzA and TriA, permutating the nine amino acid differences by DNA shuffling and screening the chimeric library for variants that hydrolyze other *s*-triazines.

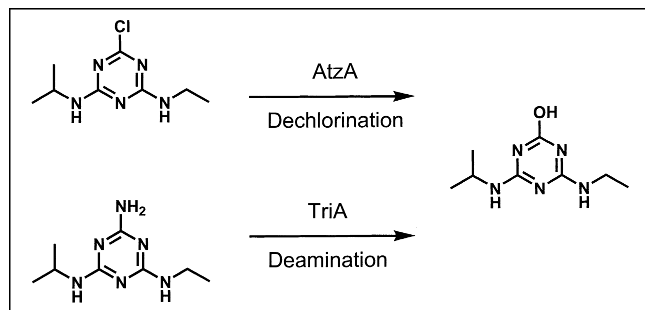

Fig. 1. Hydrolysis of atrazine and aminoatrazine by AtzA and TriA.

R2 \ R1	HN–ethyl	HN–isopropyl	morpholine
NH_2	Aminoatrazine	Aminopropazine	Aminomorphazine
Cl	Atrazine	Propazine	Morphazine
$O-CH_3$	Atratone	Prometon	Morphaton
$HN-CH_3$	N-Methylamino atrazine	N-Methylamino propazine	N-Methylamino morphazine
$S-CH_3$	Ametryn	Prometryn	Morphatryn

Fig. 2. Triazine library design to allow a structured exploration of chemical space. We fixed one of the side chains as isopropylamine in all compounds tested. We varied the second side chain (R1) to confer increased bulkiness and the leaving group (R2) to cover a 5×10^4-fold range of leaving group effectiveness in nucleophilic aromatic substitution reactions.

2. Triazine library design

We designed a family of 15 triazine compounds, differing both in their leaving group chemistry and in their steric hindrance (Fig. 2). We fixed one of the side chains as isopropylamine in all compounds tested. We varied the second side chain (R1) to confer increased bulkiness, altering R1 from ethylamine to isopropylamine to morpholine (the morpholine moiety was chosen to mimic the bulkiness of a cyclohexylamine group while maintaining solubility of the substrate in aqueous buffer). We varied the leaving group (R2) from a primary amine to chloride, methoxy, methylamine and methylthio moieties. These substituents cover a 5×10^4-fold range of leaving group effectiveness in nucleophilic aromatic substitution reactions [24].

3. Chimeric triazine hydrolases have improved and novel substrate specificities

We constructed a library of chimeric enzymes by DNA shuffling *triA* with *atzA*. We then screened 1600 members of the shuffled library, tested as crude bacterial lysates, against each of the 15 triazine substrates. The shuffled library contained enzymes that hydrolyzed five of the eight substrates that were not hydrolyzed by either of the parental enzymes. As shown in Fig. 3, enzymes hydrolyzing three of the recalcitrant substrates (prometon, prometryn and *N*-methylaminopropazine) were identified with hydrolysis rates respectively 180-, 20- and 90-fold above the minimum rates required for detection in our assay. No activity was detectable for any of these substrates with

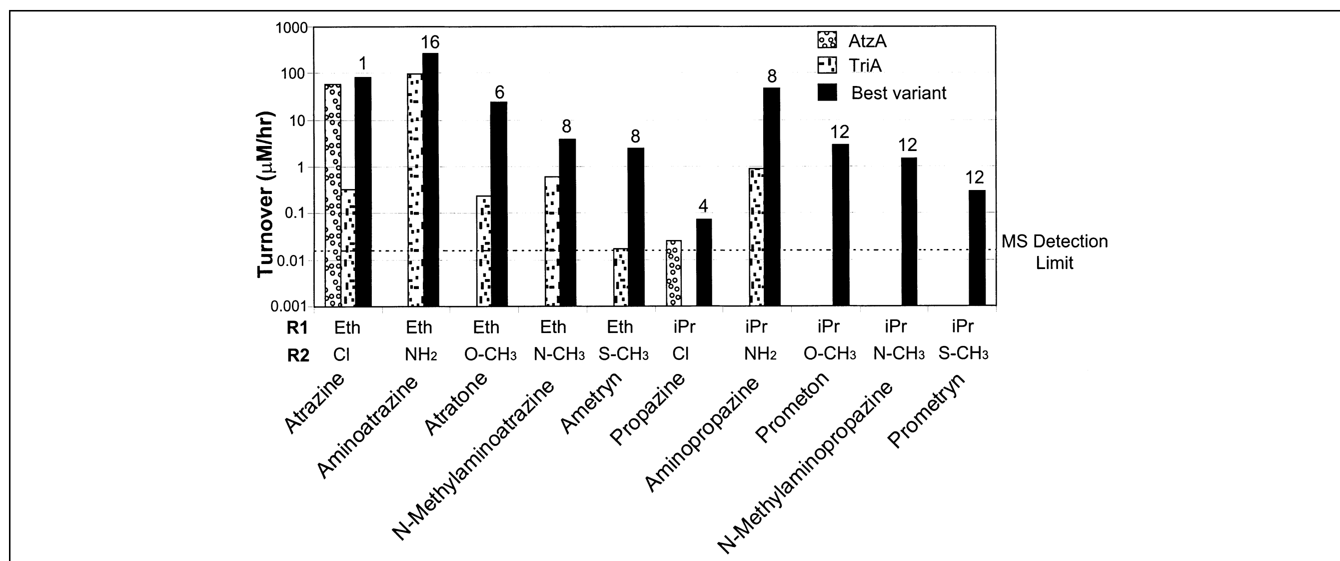

Fig. 3. Turnover rates of the best chimeric triazine hydrolases acting on different substrates. Turnover rates in μM/h for reactions containing approximately 40 ng of protein from a cell lysate in a 100 μl reaction. Rates for AtzA (open circles) and TriA (solid squares) with each of the 10 substrates with either ethyl or isopropyl R1 groups. The filled bar for each substrate represents the best performing member of the original library towards that substrate; the variant number is indicated on the bar. Errors calculated by linear regression were less than 10%; error bars are not marked because they do not show up on the logarithmic scale used.

either wild type enzyme, so the activity above the detection limit represents the minimum improvement in activity obtained for these substrates. Although Fig. 3 shows that variant 12 was the most active enzyme in the hydrolysis of each of these substrates, Table 1 shows that we obtained multiple enzymes capable of hydrolyzing these compounds. Two other recalcitrant substrates (morphazine, and aminomorphazine) were also hydrolyzed by members of the enzyme library, but the absence of a hydroxymorphazine standard prevented us from obtaining quantitative kinetic data.

For each of the seven triazines for which the parental enzymes showed some activity (atrazine, atratone, ametryn, *N*-methylaminoatrazine, aminoatrazine, propazine and aminopropazine), the library of shuffled enzymes also yielded multiple recombinant enzymes with increased transformation rates up to 150-fold greater than the best parental enzyme (Fig. 3).

4. Triazine hydrolase sequence space is densely packed with functional variation

To investigate whether the information generated by DNA shuffling and high-throughput mass spectrometry could be used predictively we analyzed the sequences and activities of the best 18 chimeras, 7 randomly chosen weak positives and 5 inactive variants (Table 1), and characterized their activities against the 15 triazines.

Comparison of the sequences revealed that AtzA-derived asparagine at position 328 and serine at position 331 are found in enzymes that are capable of only dechlorination (variants 1–5, 19 and 20), while TriA-derived aspartate at position 328 is found with either serine or cysteine at position 331 in enzymes catalyzing any reaction other than dechlorination.

Also notable is a comparison between the activities of variants 10, 11 and 12 against substrates with bulky side chains and methoxy, methylamino or methylthio leaving groups. The only difference between these three variants is decreasing size of the amino acid at position 92: from leucine to valine to alanine (the alanine in variant 12 is a fortuitous randomly introduced amino acid change resulting from mutagenesis in the DNA shuffling process). As the size of amino acid 92 decreases from leucine to alanine, activity against substrates containing large R1 side chains increases.

In almost all variants sequenced positions 217 and 219 were maintained as the AtzA-derived threonines. We do not believe that this is a result of a failure to shuffle these amino acids due to their close proximity for several reasons. Firstly, amino acids at positions 253 and 255 are also separated by only one amino acid, and yet they are seen to be uncoupled in 3 of the 18 best variants sequenced (variants 2, 8 and 18). Secondly, we observed the TriA-derived isoleucine at position 217 in combination with the AtzA-derived threonine in position 219 in one of the randomly selected weak positives (variant 25). This provides direct evidence of recombination between these amino acids. Thirdly the TriA-derived 217I and 219P only appear together in two inactive variants (26 and 27). We therefore believe that their absence from the successful chimeras is because of their lower fitness either alone or in combination.

The information from the sequence and activity comparison allows us to make broad generalizations concern-

Table 1
A comparison of enzyme sequences with their function

R1 R2	Eth Cl	Eth NH_2	Eth OCH_3	Eth NCH_3	Eth SCH_3	iPr Cl	iPr NH_2	iPr OCH_3	iPr NCH_3	iPr SCH_3	Mor Cl	Mor NH_2	Mor OCH_3	Mor NCH_3	Mor SCH_3	84	92	125	217	219	253	255	328	331	Addnl Mut
AtzA	59	0	0	0	0	0.025	0	0	0	0	0	0	0	0	0	F	V	E	T	T	I	G	N	S	
TriA	0.3	94	0.3	0.4	0.03	0	0.9	0	0	0	0	0	0	0	0	L	L	D	I	P	L	W	D	C	
1	81	0	0	0	0	0.01	0	0	0	0	0	0	0	0	0	F	V	E	T	T	L	W	N	S	
2	51	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0	F	V	E	T	T	L	G	N	S	E251G
3	42	0	0	0	0	0.002	0	0	0	0	0	0	0	0	0	F	V	D	T	T	L	W	N	S	
4	40	0	0	0	0	0.08	0	0	0	0	0.004	0	0	0	0	L	V	E	T	T	L	W	N	S	
5	21	0	0	0	0	0.04	0	0	0	0	0	0	0	0	0	L	V	E	T	T	L	W	N	S	S182P
6	14	220	24	1.0	0.7	0	0.9	0.2	0	0	0	0	0	0	0	F	L	E	T	T	I	G	D	C	
7	13	220	19	1.7	0.5	0	5.2	1.0	0.06	0	0	0	0	0	0	F	V	E	T	T	I	G	D	C	
8	12	260	21	9.0	3.0	0.003	19	1.2	0.1	0	0	0.1	0	0	0	L	L	E	T	T	L	G	D	C	T386I and M465I
9	11	130	12	3.3	2.4	0	5.3	1.0	0.03	0	0	0	0	0	0	L	L	E	T	T	L	W	D	S	
10	10	210	16	8.0	1.8	0	12	0.8	0.05	0	0	0	0	0	0	L	L	E	T	T	L	W	D	C	
11	4.7	160	8.2	6.0	0.4	0	17	1.4	0.4	0	0	0	0	0	0	L	V	E	T	T	L	W	D	C	
12	11	140	12	7.7	1.2	0.05	12	3.1	1.4	0.3	0.002	1.7	0	0	0	L	A	E	T	T	L	W	D	C	
13	8.8	78	8.8	1.6	0.3	0.005	10	1.4	0.2	0	0	0	0	0	0	L	V	E	T	T	I	G	D	S	
14	3.5	100	5.9	2.6	0.3	0	16	0.9	0.2	0	0	0	0	0	0	L	V	E	T	T	I	G	D	C	
15	4.2	87	6.5	3.1	0.4	0.004	13	1.1	0.3	0	0	0.08	0	0	0	L	V	E	T	T	I	G	D	C	D30G and M160V
16	8.5	340	15	2.3	1.4	0	0.8	0.07	0	0	0	0	0	0	0	L	L	D	T	T	L	W	D	S	
17	8.8	190	13	5.8	1.8	0	13	0.6	0.03	0	0	0	0	0	0	L	L	E	T	T	I	G	D	C	
18	3.5	300	11	1.8	0.7	0	2.1	0.03	0	0	0	0	0	0	0	L	L	D	T	T	I	W	D	C	S473G
Weakly active chimeras																									
19	37	0	0	0	0	0.002	0	0	0	0	0	0	0	0	0	F	V	D	T	T	I	G	N	S	
20	17	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0	L	V	E	T	T	I	G	N	S	V189A
21	10	250	14	1.8	0.2	0	4.7	0.7	0.05	0	0	0	0	0	0	F	V	E	T	T	L	W	D	C	I393V
22	9.0	140	15	0.7	0.3	0	3.7	0.6	0.03	0	0	0	0	0	0	F	V	E	T	T	I	G	D	C	Q453R
23	5.9	50	6.1	0.8	0.2	0.002	7.3	1.0	0.10	0	0	0	0	0	0	L	V	E	T	T	I	G	D	S	R271H
24	5.3	110	13	0.4	0.3	0	0	0	0	0	0	0	0	0	0	F	L	D	T	T	I	G	D	C	
25	3.0	60	4.2	1.2	0.3	0	2.1	0.09	0.03	0	0	0	0	0	0	L	V	D	I	T	L	W	D	S	
Inactive chimeras																									
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	F	V	D	I	P	L	W	D	C	
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	F	L	E	I	P	L	W	D	C	
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	F	V	E	T	T	L	W	N	S	K514E and L863P
Site-directed variants																									
29	5.0	130	7.4	1.1	1.6	0	11	1.4	0.7	0.2	0	0.7	0	0	0	L	G	E	T	T	L	W	D	C	
30	33	0	0	0	0	0.2	0	0	0	0	ND	0	0	0	0	L	A	E	T	T	L	W	N	S	
31	13	0	0	0	0	0.04	0	0	0	0	ND	0	0	0	0	L	G	E	T	T	L	W	N	S	
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	L	A	E	T	T	L	W	N	C	
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	L	G	E	T	T	L	W	N	C	

Reaction rates (μM/h) were measured at 22°C for reactions containing approximately 40 ng of protein from a cell lysate in a 100 μl reaction. Measurements were made at single timepoints, different reaction times were used for each substrate so that products were measured in the linear range of the reactions: aminoatrazine, 15 min; atrazine and aminopropazine, 30 min; atratone and ametryn, 1 h; *N*-methylaminoatrazine, 3 h; all other substrates 20 h. ND: not done. Values for substrates in which R1 = morpholino are approximations, the standard curve for hydroxypropazine was used as hydroxymorphazine was unobtainable. Observing the effect of the fortuitous mutation 92A, we also constructed a version of variant 12 with 92G (variant 29), versions of variant 4 containing 92A and 92G (variants 30 and 31), and versions of variants 12 and 29 containing 328N (variants 32 and 33). Amino acids in each of the nine variant positions are indicated, as are additional mutations introduced by the shuffling process.

ing structure and function relationships. These include: amino acids at positions 84 and 92 affect the bulk of side chain R1 that can be accepted, AtzA-derived amino acids are favored at positions125, 217 and 219, enzymes containing 253L and 255W from TriA tend to be more active against a broader range of substrates than those with the AtzA-derived amino acids, and 328D is required for all hydrolytic reactions other than dechlorination. Nevertheless it is clear from results shown in Table 1 that these generalizations represent a significant oversimplification, the precise effect of any one amino acid change depends on its context and the states of the other eight variable amino acids.

5. Comparisons of enzyme sequences with activities can provide predictive information

Having obtained correlations between sequence and function, one of our aims was to see whether this information could be used to engineer enzymes with predictable substrate specificities. While the 1600 variants screened represent about three times the potential number of recombinants ($2^9 = 512$), it probably does not represent a complete screening of the library. Although every variant sequenced had undergone at least one and as many as four recombination events, recombination does not occur equally between each varying amino acid position: there are three pairs of amino acids separated by only one or two residues and these pairs tend to remain linked together. In addition about a third of the variants sequenced had acquired additional amino acid changes as a result of the shuffling procedure. If it is possible to derive function–sequence correlations from the screened variants, this may substitute for exhaustive screening of a library.

A comparison between the activities of variants 10, 11 and 12 shows that as the size of amino acid 92 decreases from leucine to valine to alanine (92A occurs in variant 12 as a fortuitous mutation), activity increases against most substrates with an isopropylamine R1 side chain. Variants 10, 11 and 12 have the 328D 331C pair from TriA. Variant 4, the most active variant against propazine, is identical to variant 11 except that it contains the dechlorination-specific 328N 331S amino acids from AtzA. By analogy with variant 12, we reasoned that the isopropylamine R1 side chain of propazine would be better accommodated by incorporating the 92A substitution into variant 4. We therefore constructed and tested a version of variant 4 with alanine at position 92 (variant 30). Consistent with our prediction derived by comparing sequence and function we found that this change does increase the activity of the enzyme towards propazine. We also noted that activity of variant 30 towards atrazine was slightly less than that of variant 4, showing that this increase in activity is not simply the result of an undetected increase in expression of the enzyme.

We also tested whether decreasing the size of this residue further would continue the trend. We constructed versions of variants 12 and 30 with glycine at position 92 (variants 29 and 31 respectively) and tested their activities. Table 1 shows that these variants have the same patterns of activity as their 92A parents, but in both cases the 92G substitution is somewhat less active than the 92A variant. This may be because glycine, as well as being less bulky than alanine, allows more conformational flexibility: glycine is also known to be disfavored in specific protein structures like helices which could influence the overall three-dimensional structure of the mutant enzyme [25].

We further noted that while aspartate at residue 328 appears with either cysteine or serine at residue 331, we only recovered chimeras with 328N paired with 331S, its original partner in AtzA. To test whether this apparent preference reflected a functional requirement we engineered two variants (32 and 33) containing 328N 331C. Both were completely inactive against all substrates, consistent with their absence from the functional clones sequenced from the chimeric library. From these results we conclude that predictive information can be gained through DNA shuffling and subsequent analysis of sequences and functions of library members.

The correlations between sequence and function that we have obtained here have been limited by the size of the library and the number of variants sequenced. Because recombination produced systematically varied sequences, we have been able to obtain predictive information even from this small dataset. We anticipate that by increasing the number of parents shuffled, the number of progeny sequenced and functionally characterized, and the sophistication of the tools used to analyze the correlations, sequence and function data may allow the design of variants with specific functionalities. Minimally it should allow the construction of smaller more functional libraries, thereby significantly reducing the number of variants that must be screened to obtain the desired characteristics.

6. Can evolution select for functionally plastic proteins?

Two features of evolution contribute to the rapidity with which new functions can be generated from the enormous number of possible amino acid combinations. The first is that new proteins evolve from proteins which have been previously selected for their ability to form stable, functional structures. The second is that during the course of evolution, sequence-searching mechanisms have themselves evolved, so that organisms surviving today are those which have been able to continually adapt and reinvent themselves throughout hundreds of millions of years of changing selective pressures.

Microorganisms capable of catabolizing new xenobiotic compounds, including *s*-triazines, often appear only a few decades after the chemical is first released into the environ-

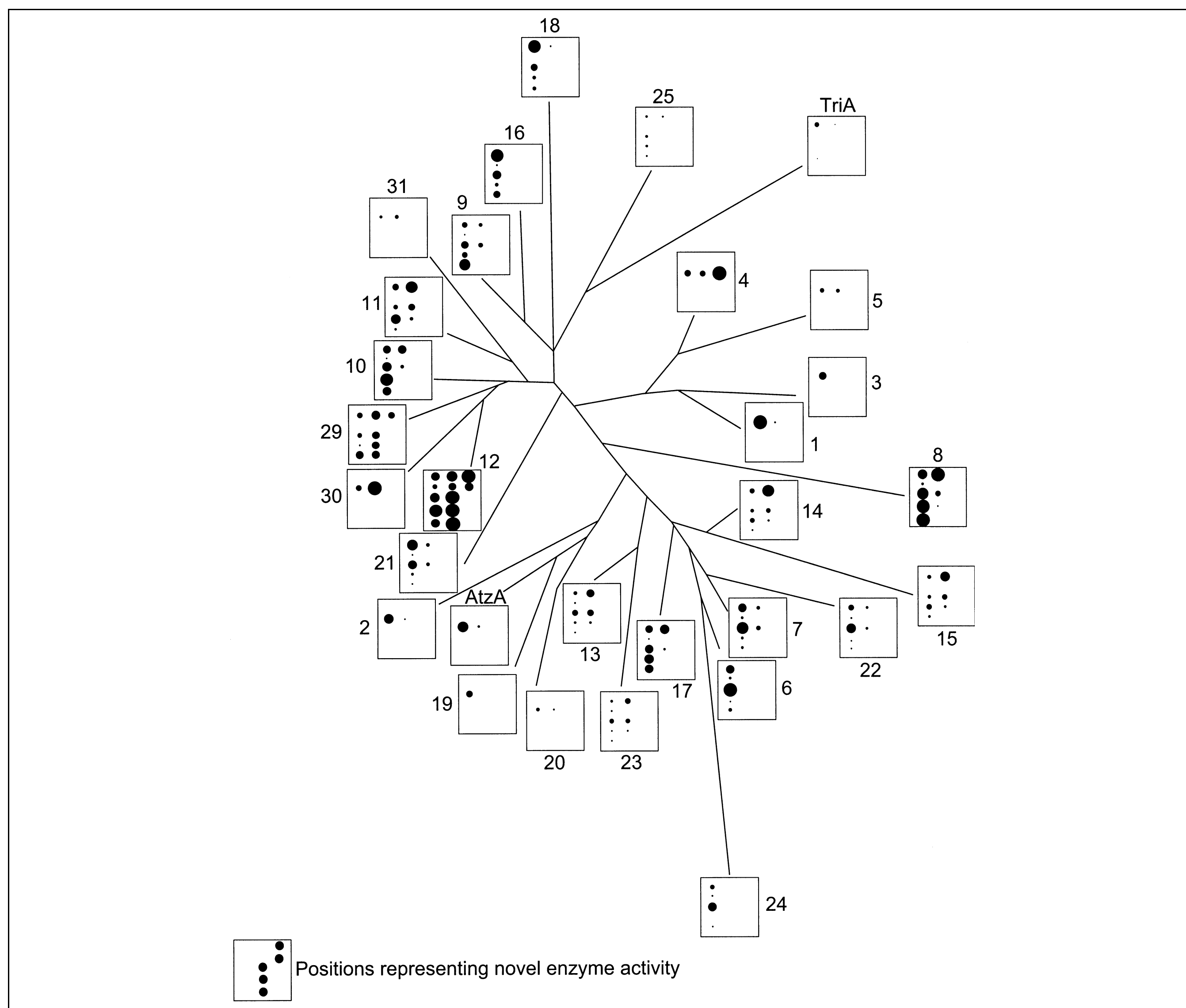

Fig. 4. Distribution of functional activity in sequence space. Variants in Table 1 are represented on a distance matrix tree constructed using DARWIN (cbrg.inf.ethz.ch/Darwinshome.html), in which variants are ordered to reflect the extent of differences between them. Activities towards each of the 15 substrates are indicated by circles arranged in the 3×5 grid shown in Fig. 2. For each substrate the circle area is proportional to the activity as shown in Table 1. We have compensated for the 100-fold range in maximal activity towards different substrates by normalizing circle areas for each substrate to the maximum activity obtained towards that substrate. Untested substrates are indicated by horizontal dashes.

ment. Here we have shown that there is a great richness of substrate specificities in the small region of sequence space accessed by permutating the two highly homologous enzymes AtzA and TriA (Fig. 4). The accessibility of functional variation in a biological system is a measure of its evolutionary potential [26,27], thus AtzA and TriA, both of which have probably been recruited very recently for their current function, are highly evolvable enzymes. While evolution cannot directly select for functions that will be required in the future, we propose that the changing availability of compounds that can serve as microbial nutrients could provide a selective advantage for catabolic enzymes that retain the ability to rapidly change substrate specificity. Consistent with this hypothesis we note that the substrate specificity determining large subunit of the catabolic enzyme biphenyl dioxygenase also appears to be highly evolvable [28,29].

Not all enzymes share the degree of plasticity we find in the triazine hydrolases: directed evolution of β-galactosidase and thymidine kinase required 4–7 rounds of shuffling, screening of over 10 000 variants per round to obtain 30–60-fold increases in specific activity [5,30]. Here we screened 10-fold fewer variants, performed only a single round of shuffling and obtained enzymes with activities two orders of magnitude greater than the best parent for four substrates, and about 20-fold greater than the best parent for another three substrates. The ability to identify highly evolvable protein sequences would be a very power-

ful protein engineering tool. Such proteins might be identified by a better understanding of their natural roles and the selective pressures upon them. AtzA and TriA possess close sequence similarity coupled with functional diversity: these may be good criteria by which to identify 'functional plasticity islands' in sequence space, representing other natural proteins that may have acquired permutability during evolution, thus readily yielding new functionalities via recombination.

7. Methods

7.1. Molecular biology

Genes for AtzA (from *Pseudomonas* sp. ADP, GenBank accession number P72156) and TriA (from *Acidovorax avenea sitrulli* NRRL B-12227, formerly *Pseudomonas*, GenBank accession number AAG41202) were cloned under control of the *lac* promoter. Genes were shuffled as previously described [31] and transformed into *Escherichia coli* TG1. Site-directed mutagenesis was performed by overlap extension [32] and DNA sequences verified by sequencing both strands on an ABI 377 sequencer.

7.2. Growth, expression and reaction conditions

7.2.1. High-throughput screening

Transformants were plated on LB-agar plates containing 1% glucose, picked into microtiter plates grown and induced in 1 mM IPTG. Cells were harvested by centrifugation, washed twice with 150 µl 10 mM ammonium acetate buffer pH 6.8. The cell pellet was resuspended in 150 µl ammonium acetate. Cells were diluted 5-fold into a reaction mix containing substrate at 250 µM in ammonium acetate buffer (125 µM for propazine and morphazine because of their relative insolubility). Atrazine, atratone, ametryn, aminomethylatrazine, aminoatrazine and propazine were reacted for 20 h at 22°C. Prometryn, prometon, *N*-methylaminopropazine, aminopropazine, morphazine, morphaton, morphatryn, *N*-methylaminomorphazine and aminomorphazine turnover were reacted for 20 h at 37°C. Reactions were stopped by addition of 100 µl methanol. Cells were removed by centrifugation and 10 µl supernatant was analyzed by mass spectrometry. Initial hits were retested as described above, but with varied reaction times from 10 min to 30 h at 22°C depending on the substrate being tested. Total ion current was converted to product concentration by comparison with standard curves.

7.2.2. Turnover rate determination

Nine clones were chosen to determine specific turnover rates for all 15 substrates. Cells were grown in 3 ml cultures and normalized to $A_{600} = 3.0$. Induction, washing and reaction conditions were as described above. Reaction samples were withdrawn for eight different timepoints between 5 min and 28 h, stopped by addition of methanol and analyzed by mass spectrometry. A dilution series of known concentrations of hydroxyatrazine or hydroxypropazine were mixed with negative control cells and used for calibration of hydrolysis products from compounds with ethyl or isopropyl R1 groups respectively. Precise quantitative turnover rates for morphazine and aminomorphazine could not be obtained due to the lack of a hydroxymorphazine standard for calibration.

Expression of the triazine-hydrolyzing enzymes was observed by running uninduced and induced samples from the cultures on a denaturing SDS–polyacrylamide gel and staining with Coomassie blue. We detected no variation in expression levels between the different variants. Comparison of the induced protein levels with quantified standards allowed us to calculate that 20 µl of cells $A_{600} = 3.0$ contained approximately 40 ng of enzyme. Turnover rates (µM/h for an enzyme concentration of 400 ng/ml) were determined for each substrate.

7.3. Chemical syntheses

Atrazine, propazine, atratone, prometon, ametryn, prometryn, hydroxyatrazine, hydroxypropazine and aminoatrazine were purchased from Cresent Chemical Co. or Riedel-de-Haen Co. The other library members were synthesized using the following general protocols.

7.3.1. N-Isopropyl-dichlorotriazine

To a stirred solution of cyanuric chloride (1.845 g, 10 mmol) in anhydrous dichloromethane (100 ml) at −5°C, was added Hunig's base (1.92 ml, 11 mmol) followed by dropwise addition of isopropylamine (860 µl, 10 mmol). The solution was stirred for 1 h and the reaction was quenched by pouring into 10% citric acid (100 ml). The organic phase was separated and washed with 10% citric acid, sat. brine and dried. The solvent was removed to yield a clear oil (2.07 g, quant).

7.3.2. Addition of the second side chain (R1)

To a stirred solution of *N*-isopropyl-dichlorotriazine (1.03 g, 5 mmol) in anhydrous dichloromethane (50 ml) at 0°C, was added Hunig's base (960 µl, 5.5 mmol) followed by dropwise addition of the nucleophile (morpholine, isopropylamine, 5 mmol). The solution was stirred at 0°C for 1 h and the reaction was allowed to warm to ambient temperature overnight. The reaction was quenched by pouring into 10% citric acid (50 ml). The organic phase was separated and washed with 10% citric acid, sat. brine and dried. The solvent was removed to yield a white powder, which was recrystallized from ethyl acetate (1.1 g, 86%).

7.3.3. Addition of the third side chain (R2)

To a stirred solution of disubstituted chlorotriazine (0.5 mmol) in anhydrous aminomethyl pyrrolidone (5 ml) at 20°C, was added Hunig's base (960 µl, 5.5 mmol) and the nucleophile (5 mmol). The solution was warmed to 80°C and stirred until reaction was complete (1–3 h). The reaction was quenched by pouring into 10% citric acid (50 ml). The organic phase was separated and washed with 10% citric acid, sat. brine and dried. The solvent was removed to yield a white powder, which was purified by flash column chromatography and recrystallized from ethyl acetate when possible (quant). In the case where R2 = amino group, sodium azide was the nucleophile. The azidotriazine was then reduced with borohydride on Amberlite in methanol to give the desired aminotriazine. The identity of the chromatographically pure library members was confirmed by mass spectrometry.

7.4. Triazine hydrolysis assays

All mass spectrometry data were recorded on a Micromass

Quattro LC (Micromass, Beverly, MA, USA) equipped with an electrospray ionization source. The mobile phase (methanol) was introduced at a constant flow rate of 300 μl/min. Samples (10 μl) were drawn from a 96-well microtiter plate and analyzed in flow-injection mode at a speed of one sample every 40 s.

To detect for all 15 substrates and their products, we used the fixed side chain (isopropylamine group) as a mass spectroscopy neutral loss detection marker. The neutral loss MS/MS spectra were obtained in the triple-quadrupole neutral loss scanning mode by searching for mass losses of 42 (C_3H_6) in the mass range of 130–300 amu. The pressure in quadrupole 2 (collision cell) normally was set at 5×10^{-4} mbar. The collision energy was adjusted to 30 eV. The photomultiplier was operated at 650 V.

Acknowledgements

We thank Yan Chen and Kim Perry for technical assistance, Robin Emig, Rob Pak, Claes Gustafsson, Sridhar Govindarajan, Troy Obero and Walker Lutringer for automation and data analysis support, and Stephen Raillard, Brian Shoichet, Mike Sadowsky and Russell Howard for helpful discussions, insights and comments on the manuscript. Jennifer Seffernick was supported by NIH training Grant GM08347.

References

[1] J.E. Ness, M. Welch, L. Giver, M. Bueno, J.R. Cherry, T.V. Borchert, W.P. Stemmer et al., DNA shuffling of subgenomic sequences of subtilisin, Nat. Biotechnol. 17 (1999) 893–896.
[2] H. Zhao, F.H. Arnold, Directed evolution converts subtilisin E into a functional equivalent of thermitase, Protein Eng. 12 (1999) 47–53.
[3] J.C. Moore, F.H. Arnold, Directed evolution of a *para*-nitrobenzyl esterase for aqueous-organic solvents, Nat. Biotechnol. 14 (1996) 458–467.
[4] K. Liebeton, A. Zonta, K. Schimossek, M. Nardini, D. Lang, B.W. Dijkstra, M.T. Reetz et al., Directed evolution of an enantioselective lipase, Chem. Biol. 7 (2000) 709–718.
[5] J.-H. Zhang, G. Dawes, W.P.C. Stemmer, Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening, Proc. Natl. Acad. Sci. USA 94 (1997) 4505–4509.
[6] A. Crameri, G. Dawes, E. Rodriguez, S. Silver, W.P.C. Stemmer, Molecular evolution of an arsenate detoxification pathway by DNA shuffling, Nat. Biotechnol. 15 (1997) 436–438.
[7] N. Soong, L. Nomura, K. Pekrun, M. Reed, L. Sheppard, G. Dawes, W.P.C. Stemmer, Molecular breeding of viruses, Nat. Genet. 4 (2000) 436–439.
[8] S.K. Powell, M.A. Kaloss, A. Pinkstaff, R. McKee, I. Burimski, M. Pensiero, E. Otto et al., Breeding of retroviruses by DNA shuffling for improved stability and processing yields, Nat. Biotechnol. 18 (2000) 1279–1282.
[9] B. Spiller, A. Gershenson, F.H. Arnold, R.C. Stevens, A structural view of evolutionary divergence, Proc. Natl. Acad. Sci. USA 96 (1999) 12305–12310.
[10] M.C. Orencia, M.A. Hanson, R.C. Stevens, Structural analysis of affinity matured antibodies and laboratory-evolved enzymes, Adv. Protein Chem. 55 (2000) 227–259.
[11] W.P. Stemmer, Rapid evolution of a protein in vitro by DNA shuffling, Nature 370 (1994) 389–391.
[12] J.E. Ness, S.B. delCardayre, J. Minshull, W.P.C. Stemmer, Molecular breeding: the natural approach to protein design, Adv. Protein Chem. 55 (2000) 261–292.
[13] K. Jutzi, A.M. Cook, R. Hutter, The degradative pathway of the *s*-triazine melamine. The steps to ring cleavage, Biochem. J. 208 (1982) 679–684.
[14] M.L. de Souza, M.J. Sadowsky, L.P. Wackett, Atrazine chlorohydrolase from *Pseudomonas* sp. strain ADP: gene sequence, enzyme purification and protein characterization, J. Bacteriol. 178 (1996) 4894–4900.
[15] J.L. Seffernick, M.L. de Souza, M.J. Sadowsky, L.P. Wackett, Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different, J. Bacteriol. 183 (2001) 2405–2410.
[16] H.O. Esser, G. Dupois, E. Ebert, G.J. Marco, C. Vogel, *s*-Triazines, in: P.C. Kearney, D.D. Kaufman, Herbicides: Chemistry, Degradation and Mode of Action, Vol. 1, 2nd edn., Chapter 2, Marcel Dekker, New York, 1975.
[17] F.A. Gunther, J.D. Gunther, The Triazine Herbicides, Springer-Verlag, New York, 1970.
[18] D.D. Kaufman, P.C. Kearney, Microbial degradation of *s*-triazine herbicides, Residue Rev. 32 (1970) 235–265.
[19] R.T. Mandelbaum, D.L. Allan, L.P. Wackett, Isolation and characterization of a *Pseudomonas* sp. that mineralizes the *s*-triazine herbicide atrazine, Appl. Environ. Microbiol. 61 (1995) 1451–1457.
[20] M.L. de Souza, J. Seffernick, B. Martinez, M.J. Sadowsky, L.P. Wackett, The atrazine catabolism genes atzABC are widespread and highly conserved, J. Bacteriol. 180 (1998) 1951–1954.
[21] A.M. Cook, R. Hutter, *s*-Triazines as nitrogen sources for bacteria, J. Agric. Food Chem. 29 (1981) 1135–1143.
[22] J.L. Seffernick, G. Johnson, M.J. Sadowsky, L.P. Wackett, Substrate specificity of atrazine chlorohydrolase and atrazine-catabolizing bacteria, Appl. Environ. Microbiol. 66 (2000) 4247–4252.
[23] I. Matsumura, A.D. Ellington, In vitro evolution of β-glucuronidase into a β-galactosidase proceeds through non-specific intermediates, J. Mol. Biol. 305 (2001) 331–339.
[24] J. March, Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, John Wiley and Sons, New York, 1992.
[25] B.W. Matthews, H. Nicholson, W.J. Becktel, Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, Proc. Natl. Acad. Sci. USA 84 (1987) 6663–6667.
[26] C.L. Burch, L. Chao, Evolvability of an RNA virus is determined by its mutational neighbourhood, Nature 406 (2000) 625–628.
[27] B.G. Hall, Toward an understanding of evolutionary potential, FEMS Microbiol. Lett. 178 (1999) 1–6.
[28] T. Kumamaru, H. Suenaga, M. Mitsuoka, T. Watanabe, K. Furukawa, Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase, Nat. Biotechnol. 16 (1998) 663–666.
[29] F. Bruhlmann, W. Chen, Tuning biphenyl dioxygenase for extended substrate specificity, Biotechnol. Bioeng. 63 (1999) 544–551.
[30] F.C. Christians, L. Scapozza, A. Crameri, G. Folkers, W.P.C. Stemmer, Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling, Nat. Biotechnol. 17 (1999) 259–264.
[31] W.P. Stemmer, DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution, Proc. Natl. Acad. Sci. USA 91 (1994) 10747–10751.
[32] R.M. Horton, In vitro recombination and mutagenesis of DNA, in: B.A. White (Ed.), PCR Protocols: Current Methods and Applications, Vol. 15, Humana, Totowa, NJ, 1993.