Molecular BioSystems

Cite this: DOI: 10.1039/c2mb05492b

www.rsc.org/molecularbiosystems

PAPER

Engineering a large protein by combined rational and random approaches: stabilizing the *Clostridium thermocellum* cellobiose phosphorylase†

Xinhao Ye,*[a] Chenming Zhang[a] and Y.-H. Percival Zhang*[abc]

Received 7th December 2011, Accepted 22nd March 2012
DOI: 10.1039/c2mb05492b

The *Clostridium thermocellum* cellobiose phosphorylase (*Ct*CBP) is a large protein consisting of 812 amino acids and has great potential in the production of sugar phosphates, novel glycosides, and biofuels. It is relatively stable at 50 °C, but is rapidly inactivated at 70 °C. To stabilize *Ct*CBP at elevated temperatures, two protein-engineering approaches were applied, *i.e.* site-directed mutagenesis based on structure-guided homology analysis and random mutagenesis at various mutation rates. The former chose substitutions by comparison of the protein sequences of CBP homologs, utilized structural information to identify key amino acid residues responsible for enhanced stability, and then created a few variants accurately. The latter constructed large libraries of random mutants at different mutagenesis frequencies. A novel combinational selection/screening strategy was employed to quickly isolate thermostability-enhanced and active variants. Several stability-enhanced mutants were obtained by both methods. Manually combining the stabilizing mutations identified from both rational and random approaches led to the best mutant (CM3) with the halftime of inactivation at 70 °C extended from 8.3 to 24.6 min. The temperature optimum of CM3 was increased from 60 to 80 °C. These results suggested that a combination of rational design and random mutagenesis could have a solid basis for engineering large proteins.

Introduction

Development of thermostable and hyper-thermostable enzymes is a very active research area because it not only broadens the industrial applicability of enzymes, but also furthers our understanding of protein structure–function relationship.[1–3] Numerous methods have been employed to improve enzyme thermostabilization, including both non-biologically- and biologically-based methods.[4] Non-biological methods use additives (*e.g.*, surfactant, salt, reducing agents, *etc.*) or/and immobilize enzymes on a solid support to guard against denaturation.[5,6] Biological methods aim to isolate enzyme variants from thermophiles or to enhance enzyme stability by protein engineering.[7,8]

Protein engineering offers numerous promising advantages over the other methods,[9,10] often by two distinctive approaches: rational design and directed evolution.[11,12] Directed evolution is an effective way to evolve targeted biocatalysts with desired performance.[13,14] It is independent of the knowledge of three-dimensional structures of proteins but does often require high-throughput screening and/or effective selection systems. The success of directed evolution is restricted by sequences explored, biases in mutagenesis methods, as well as the degeneracy of genetic codes.[15] For example, beneficial mutations can be accumulated and selected or screened from iterative rounds of directed evolution.[16,17] On the other hand, thermostable enzyme variants can also be rationally designed by introducing hydrogen bonds, strengthening salt or disulfide bridges, improving core packing, optimizing surface charges, increasing rigidity with preferred residue substitutions, or stabilizing α-helix, β-turns or flexible termini or loops.[18,19] Rational protein design could be less labor intensive than directed evolution, but requires extensive knowledge of protein structure and function.[4,20,21] Among numerous protein design approaches, a semi-rational approach is to design mutants by comparison of the amino acid sequences of homologous enzymes.[22,23] Statistical analysis extracts recurring amino acid replacement trends, from which the respective consensus amino acids are presumed to contribute more than average to protein stability than the non-consensus residues.[1] This 'consensus concept' has been used to produce many new thermostable proteins.[24,25] However, this approach may be hampered by noises accompanied with random genetic drift.[15,26] In view of the pros and cons of protein engineering

[a] *Biological Systems Engineering Department, Virginia Tech, Blacksburg, Virginia 24061, USA. E-mail: ypzhang@vt.edu, xhye@vt.edu; Fax: +1 540 231-3199; Tel: +1 540 231-7414*
[b] *Institute for Critical Technology and Applied Science (ICTAS), Virginia Tech, Blacksburg, Virginia 24061, USA*
[c] *DOE Bioenergy Science Center, Oak Ridge, Tennessee 37831, USA*
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c2mb05492b

methods, a combination of directed evolution and rational design represents a more effective route for improving the properties and functions of enzymes.[12,27] An impressive example of a successful combination of directed evolution and rational design is the improvement of stability of *Coprinus cinereus* heme peroxidase.[28] Other examples include increasing the thermal stability of *Bacillus subtilis* 3-isopropylmalate dehydrogenase by a combination of *in vivo* mutagenesis, *in vitro* evolution, and rational design,[29,30] and creating a thermostable penicillin G acylase and a thermostable glucose dehydrogenase by a structure-guided consensus approach.[4,25]

Cellobiose phosphorylase (CBP, EC 2.4.1.20) catalyses phosphorolysis of cellobiose into glucose 1-phosphate and glucose. It belongs to glycoside hydrolase (GH) family 94 and plays an important role in the energy-efficient metabolism of long-chain polysaccharides.[31,32] CBP has been applied for hetero-oligosaccharide synthesis,[33] sugar 1-phosphate production,[34] and enzymatic hydrogen production.[35] The *C. thermocellum* CBP (*Ct*CBP) consists of 812 amino acid residues. It is relatively stable at 50 °C, but the halftime of inactivation is less than 10 min at 70 °C.[36–38]

The aim of our work is to create a *Ct*CBP variant with improved thermal stability and comparable activity to the wild-type enzyme. The hyperthermostable and active variant is of essence to increase the biohydrogen production rate by cell-free Synthetic Pathway Biotransformation (SyPaB) at elevated temperatures,[35] as well as to decrease enzyme costs of enzymatic hydrogen production[39] and glycoside synthesis.[40,41] Unlike the previous studies in that glycoside phosphorylases were stabilized *via* random mutagenesis over part of the enzymes,[42,43] we created the *Ct*CBP mutants by two parallel approaches: site-directed mutagenesis based on a structure-guided homology approach and random mutagenesis of the whole *ctcbp* gene *via* error-prone PCR. The best mutant based on a combination of both approaches significantly extended the inactivation halftime of *Ct*CBP and had a three-fold increase of the specific activity at 80 °C.

Materials and methods

Materials

All chemicals were reagent-grade, purchased from Sigma (St. Louis, MO, USA), unless otherwise noted. *Clostridium thermocellum* ATCC 27405 genomic DNA was a gift from Dr Jonathan Mielenz at the Oak Ridge National Laboratory (Oak Ridge, TN). The Luria-Bertani (LB) medium and M9/cellobiose medium were prepared as described elsewhere.[44] Regenerated amorphous cellulose (RAC) was prepared from Avicel after water slurrying, cellulose dissolution in concentrated H_3PO_4, and regeneration in water.[45]

Bacterial strains and plasmids

Escherichia coli JM109 was used for cloning and mutant library construction; *E. coli* Rosetta BL21 (DE3) was employed for CBP production. The *cbp* gene was amplified from genomic DNA of *C. thermocellum* using primers P1 and P2 (Table S1, ESI†). The DNA fragments were digested by *Pst*I and *Bam*HI, and then ligated into the digested plasmid pUC19 (New England Labs, Ipswich, MA, USA) to give the plasmid pUCB. For protein characterization, the wild-type and mutant *cbp* genes were amplified from the plasmid pUCB with the primers P2 and P3 (Table S1, ESI†). The PCR product was digested by *Bam*HI and *Xho*I, and ligated with a *Bam*HI/*Xho*I digested pCIG vector[46] for the plasmid pCIB. All plasmid sequences described herein were verified by DNA sequencing (MCLab, San Francisco, CA).

Structure-guided homology analysis

Calculation of the CBP consensus sequence was performed as follows. The amino acid sequence of *Ct*CBP was firstly blasted in GenBank by PSI-BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi); 55 unique CBP sequences were identified, then sorted by the optimal growth temperatures of the microorganism where the CBPs were originated, and finally aligned together by the program ClustalW from BioEdit (Carlsbad, CA, USA). Nine CBP sequences, which exhibited the highest identities with *Ct*CBP (56.4–73.9%), were chosen for homology analysis (Fig. S1, ESI†). Among them, *Ia*CBP is from *Ignisphaera aggregans* (optimal growth temperature: 95 °C), *Tn*CBP is from *Thermotoga neapolitana* (optimal growth temperature: 90 °C), *Tm*CBP is from *Thermotoga maritime* (optimal growth temperature: 80 °C), *Dt*CBP is from *Dictyoglomus thermophilum* (optimal growth temperature: 78 °C), *Csa*CBP is from *Caldicellulosiruptor saccharolyticus* (optimal growth temperature: 70 °C), *Cst*CBP is from *Clostridium stercorarium* (optimal growth temperature: 65 °C), *Cu*CBP is from *Cellulomonas uda* (optimal growth temperature: 50 °C), *Bf*CBP is from *Butyrivibrio fibrisolvens* (optimal growth temperature: 37 °C), and *Cg*CBP is from *Cellvibrio gilvus* (optimal growth temperature: 30 °C) (Table 1). Following sequence alignment, the candidates for substitution were selected by applying the following criteria: (1) at the certain residue the substitutions were obviously related with the change of growth temperatures; (2) the substitutions preferred the consensus amino acid from thermophiles and hyperthermophiles;[1] (3) the substitutions were not located in stabilization centers, or were capable to expand the stabilization centers (Table S2, ESI†);[47,48] and (4) the substitutions did not contradict any structural knowledge with respect to stability enhancement.

Table 1 Putative stabilizing mutations identified by structure-guided homology analysis

CBP ID	Origin	Growth temp./°C	Amino acids at the potential residues[a]					
			130–131	201	292	411	423	781
*Ia*CBP	*I. aggregans*	95	WW	*P*[b]	L	*G*[b]	*S*[b]	*K*[b]
*Tn*CBP	*T. neapolitana*	90	HH	S	*K*[b]	G	S	K
*Tm*CBP	*T. maritime*	80	*HY*[b]	P	K	G	S	K
*Dt*CBP	*D. thermophilum*	78	HY	P	K	S	S	K
*Csa*CBP	*C. saccharolyticus*	70	HY	P	K	S	S	K
*Cst*CBP	*C. stercorarium*	65	NQ	P	I	S	A	N
*Ct*CBP	*C. thermocellum*	60	**QK**[c]	**K**[c]	**N**[c]	**S**[c]	**A**[c]	**A**[c]
*Cu*CBP	*C. uda*	50	QK	Q	A	S	G	G
*Bf*CBP	*B. fibrisolvens*	37	TK	K	Q	S	G	A
*Cg*CBP	*C. gilvus*	30	QK	R	A	S	G	A

[a] The residues were numbered based on the protein sequence of *Ct*CBP. [b] The chosen amino acids (also italicized) were introduced to *Ct*CBP to replace their counterparts. [c] The residues of *Ct*CBP (also marked in bold) were chosen to mutate.

For example, if the mutation was found in a helix, a helix-stabilizer was not changed to a helix-destabilizer.[25]

Site-directed mutagenesis

Site-directed mutagenesis was performed with the Phusion site-directed mutagenesis kit (New England Labs, Ipswich, MA, USA) according to the manufacturer's instructions. Each reaction (50 μL) contained 100 pg of template plasmid pCIB and 25 pmol of each mutagenic primer. The primers were commercially phosphorylated at the 5′ end and listed in Table S1 (ESI†). Thermal cycling was performed in Eppendorf Mastercycler pro Thermal cyclers (Hauppauge, NY) with the following cycling conditions: 98 °C for 30 s, 25 cycles of 98 °C for 10 s, 54–61 °C for 20 s, and 72 °C for 5 min, and finally 72 °C for 10 min. The PCR products were purified by the Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA), self-ligated by the NEB quick ligase kit, and then transformed into *E. coli* Rosetta BL21 (DE3).

Construction of random mutant libraries

Random mutagenesis of the *Ctcbp* gene was performed in a reaction mix containing 0.2 ng μL^{-1} plasmid pUCB as the template, 10 mM Tris–HCl (pH 8.3), 0.2 mM dATP, 0.2 mM dGTP, 1 mM dCTP, 1 mM dTTP, 0.4 μM each of the primers P4 and P5 (Table S1, ESI†), 30 mM KCl, $MgCl_2$, $MnCl_2$, and 2.5 U NEB Taq polymerase. Three mutagenic libraries with different mutation frequencies were made by varying concentration of Mg^{2+} and Mn^{2+}, *i.e.* 5 mM $MgCl_2$ for the library (Library L) with a low mutation frequency, 5 mM $MgCl_2$ plus 0.3 mM $MnCl_2$ for the one (Library O) with an estimated optimal mutation frequency (Fig. S5, ESI†), and 7 mM $MgCl_2$ with 0.5 mM $MnCl_2$ for the library (Library H) with a high mutation frequency.[49] The cycling scheme was 94 °C for 2 min, 30 cycles of 94 °C for 1 min, 54 °C for 1 min, and 72 °C for 3 min, followed by 72 °C for 10 min. The 2852 bp PCR products, including the entire *cbp* gene, 244 bp of an upstream sequence, and 175 bp of a downstream sequence, were purified using the Zymoclean DNA Recovery Kit. The resulting DNA fragment was then digested by *Pst*I and *Bam*HI, ligated with the plasmid pUC19, and transformed into *E. coli* JM109.

Selection and screening

Putative *Ct*CBP mutants were isolated by a combinatorial selection/screening approach, similar to our previous work for beta-glucosidase.[44] The whole process consisted of two steps involving the selection for mutants with retaining CBP activity and subsequent screening for improved thermostability. All transformants were first spread on solid agar plates (selection plates) with 1.5% agar, M9 minimal medium with 0.4% cellobiose, 75 μg mL^{-1} ampicillin, and 0.1 mM IPTG. Meanwhile, a small fraction of the transformed cells from each library was grown on a LB plate containing 100 μg mL^{-1} ampicillin, which would be used as the reference plate. All plates were incubated at 30 °C until the colonies could be visualized easily. Next, the colonies were transferred to nylon membranes, subjected to heat treatment at 80 °C for 10 min, and eventually lysed *in situ* on the nylon membranes, resulting in the release of intracellular CBP. The membranes with released CBP were overlaid on the screening plates that contained M9 minimal medium with 1% cellobiose, 100 μg mL^{-1} ampicillin, 0.5% agar, and an indicator strain *E. coli* JM109/pUC19. Since the indicator strain was able to utilize glucose and glucose 1-phosphate but not cellobiose, after heat treatment the survived CBP variants could digest cellobiose and then support the growth of the indicator strain. Therefore, the colonies (white spots) in the screening plates would help identify the clones expressing thermostable CBP mutants. Colonies on the reference plates were counted to estimate the library sizes, and the mutation rate of each library was determined by sequencing, in both directions, the *Ctcbp* gene from 6–10 randomly picked clones. The selection power was calculated from the ratio of the number of colonies on the M9-cellobiose plates, where only the transformants expressing active CBP can grow, to that on the LB reference plates, where all transformants can grow.

Protein production and purification

Wild-type CBP and its thermostable mutants were produced from the expression vectors in *E. coli*, firstly as fusion protein CBM-intein-CBP. The *E. coli* strain was grown in LB medium at 37 °C with a rotary shaking rate of 250 rpm. When the absorbance of the culture (A_{600}) reached 0.8, IPTG was added to a final concentration of 0.25 mM and the growth temperature was decreased to 16 °C. After overnight culture, the *E. coli* cells were harvested by centrifugation and resuspended in a 50 mM HEPES buffer (pH 7.2). The cell suspension was sonicated, and the cell debris was removed by centrifugation (10 000*g* for 20 min at 4 °C). Then the fusion protein in the supernatant was bound to regenerated amorphous cellulose (RAC). After pH adjustment for intein self-cleavage, the cleaved CBP protein was obtained in the supernatant.[46] Finally, the protein solution was dialyzed with a 50 mM HEPES buffer (pH 7.2) and concentrated with a protein concentration kit from Millipore (Billerica, MA, USA) with a molecular weight cut-off of 50 000. The concentration of purified protein was determined by the Bradford method with bovine serum albumin (BSA) as the standard. Enzyme purity was checked by SDS-PAGE.

Enzyme assay

All enzymatic reactions were conducted in 5 mL glass culture tubes (12 × 75 mm, Fisher Scientific). The enzyme activities were assayed at 40 °C in 50 mM HEPES buffer (pH 7.2) containing 30 mM cellobiose, 10 mM glucose 1-phosphate, 1 mM DTT, and 1 mM Mg^{2+} unless otherwise noted. Kinetic parameters (K_m and k_{cat}) were determined by non-linear regression of Michaelis–Menten data *via* CurveExpert v1.4 (Hixson, TN, USA). The kinetics were examined at 40 °C in a 50 mM HEPES buffer (pH 7.2) containing 1 mM Mg^{2+}, 1 mM DTT, and various substrate concentrations between 0.2 and 5 times of their respective K_m values.[50] Enzyme concentrations were set at 5.0 mg L^{-1} for most assays. The reactions were stopped by placing the tubes in a boiling water bath for 10 min. The product phosphate was measured by the mild pH phosphate assay as described elsewhere.[51] One unit of cellobiose phosphorylase was defined as the amount of enzyme that generates one μmol of phosphate per min.

Temperature effects and thermostability

The effects of temperature were examined by measuring the enzyme activities at various temperatures under standard assay conditions. Halftime of thermal inactivation was determined by incubating the enzymes (0.05 mg mL^{-1}) in 50 mM HEPES buffer (pH 7.2) at 70 °C for different time intervals, as described before.[44] After the incubation, samples were chilled on ice immediately. Then the residual activity was assayed at 40 °C as described above. The activation energy (E_a) was determined from the slope of the Arrhenius plot.

Homology modeling

The homology model was built over the assumption that structure is much more conserved than sequence during the evolution and the modeling structure can reflect the structure changes of the mutants. The structures of *Ct*CBP variants were estimated online by ESyPred3D[52] using the crystal structure of *Ct*CBP (PDB: 3qde) as template.[53,54] Before structural analysis, every model was verified for consistency with known protein folds and allowed ϕ and ψ angles by Insight II (Accelrys Inc., San Diego, CA, USA), both with default settings. For each selected protein, the structure was visualized by VMD (UIUC, Urbana-Champaign, IL, USA), and the protein contact map was built by the Contact Map plugin of VMD 1.8.6.[55]

Results

Site-directed mutagenesis based on structure-guided homology analysis

The amino acid sequence of *C. thermocellum* cellobiose phosphorylase (*Ct*CBP) was compared with nine homologous proteins for identifying the amino acid candidates for site-directed substitutions. The ten CBPs having more than 50% sequence identities were chosen from different species with diverse growth temperatures (Fig. S1, ESI†). For example, *Ia*CBP, *Tn*CBP, along with *Tm*CBP are from hyperthermophiles (growth optimum > 80 °C); *Dt*CBP, *Csa*CBP, *Cst*CBP, *Cu*CBP, as well as *Ct*CBP are from thermophiles (growth optimum > 45 °C); *Bf*CBP and *Cg*CBP are from mesophiles (growth optimum > 25 °C).

Seven mutations were identified by the structure-guided homology analysis (Table 1), and individually introduced to the wild-type *Ctcbp* gene, except Q130H and K131Y. Q130 and K131 interact with each other, and both function as the stabilization center (Table S2, ESI†). They are surrounded by aromatic amino acids, *e.g.* Y71, F72, Y103, F119, F120, Y258, and F260, as illustrated in the contact map (Fig. S2, ESI†) and in Fig. 1A and B. The mutations, Q130H together with K131Y, create new π–π interactions, resulting in an extension of the inactivation halftime at 70 °C from 8.34 to 14.1 min. K201P, N292K, and S411G are all located in the surface loops. K201P was chosen because the pyrolidine ring of Pro could restrict the number of conformations and eliminate entropy gain in the denatured state.[56] Replacement of Asn 292 by Lys might increase the number of ion pairs,[57] while substitution of Ser 411 by Gly was because (hyper)thermophilic proteins statistically contain more Gly and less Ser.[58] As a consequence, the mutation S411G improved the protein stability (Fig. 1C and D, and Table 2), whereas K201P and N292K reduced

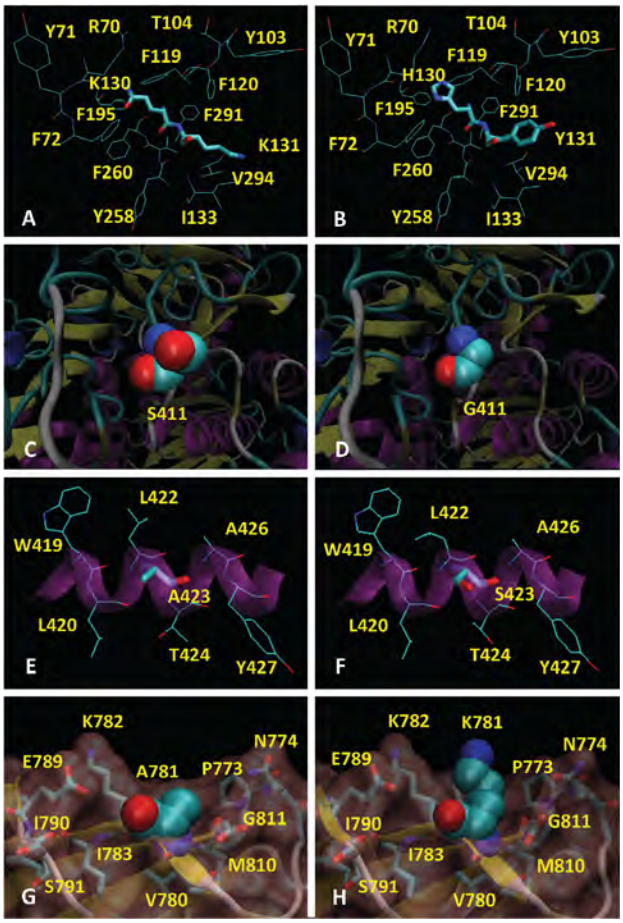

Fig. 1 Structural interpretation of the residues that contribute the differences in thermostability between *Ct*CBP and its mutants. (A) The residues surround Q130 and K131 in *Ct*CBP; (B) the residues surround H130 and Y131 in the mutant Q130H/K131Y; (C) surface portion around S411 in *Ct*CBP; (D) surface portion around G411 in the mutant S411G; (E) an α-helix containing A423 in *Ct*CBP; (F) an α-helix containing G423 in the mutant A423G; (G) a structural cavity that surrounds A781 in *Ct*CBP; (H) a structural cavity that surrounds K781 in the mutant A781K.

the thermostability of *Ct*CBP (data not shown). The fifth site-directed mutant A423S was designed to stabilize the α-helix by capping the interactions among W419, S423, and Y427 (Fig. 1E and F). It practically increased the halftime to 11.3 min (Table 2). The most striking effect on protein stability occurred upon the mutation of A781. The A781K mutant offered a better packed C-terminal domain, built the interaction between K781 and N774, and pushed K782 toward E789 (Fig. 1G and H). All in all, it reduced loop flexibility in the terminus and enhanced the halftime to 15.3 min (Table 2).

As the individual mutations were combined in successive rounds of site-directed mutagenesis, they were found to be approximately multiplicative in their effects on both improved thermostability and reduced activity. The thermal stability effects were cumulative. The final combined mutant I (CM1) from the modified structure-guided homology approach (Q130H, K131Y, S411G, A423S, and A781K) lengthened the halftime at 70 °C by more than 2-fold to 17.7 min (Table 2).

Table 2 Characterization of the thermostability-enhanced mutants

Enzyme	$T_{1/2}$[a]/min	Kinetics		
		k_{cat}/s^{-1}	K_m/mM	k_{cat}/K_m
Wild-type	8.3 ± 0.2	3.40 ± 0.11	16.5 ± 0.58	0.206
Q130H and K131Y	14.1 ± 0.3	4.50 ± 0.14	18.8 ± 0.24	0.239
S411G	13.6 ± 0.3	1.32 ± 0.18	6.94 ± 0.20	0.189
A423S	11.3 ± 0.4	2.96 ± 0.38	10.3 ± 0.45	0.287
A781K	15.3 ± 0.2	2.90 ± 0.25	11.0 ± 0.19	0.263
CM1[b]	17.7 ± 0.4	2.13 ± 0.16	15.9 ± 0.33	0.134
M52[c]	16.6 ± 0.8	0.85 ± 0.21	15.3 ± 0.42	0.056
M52m[d]	15.3 ± 0.3	3.12 ± 0.15	19.0 ± 0.61	0.164
CM2[e]	17.4 ± 0.4	1.44 ± 0.21	7.92 ± 0.41	0.182
CM3[f]	24.6 ± 0.3	1.52 ± 0.23	14.7 ± 0.34	0.103

[a] $T_{1/2}$ denotes the halftime of thermal inactivation at 70 °C. [b] CM1 represents the mutant with combined mutations I (Q130H, K131Y, S411G, A423S, A781K). [c] M52 is a positive mutant obtained from moderate-frequency random mutagenesis, which contained five mutations, *i.e.*, R48K, K142R, R189L, A423S, and V526A. [d] M52m is an alternative mutant to M52 with a reverse mutation L189R. So it held four mutations, such as R48K, K142R, A423S, and V526A. [e] CM2 represents the mutant with combined mutations II (R48K, K142R, A423S, V526A, and A781K). [f] CM3 represents the mutant with combined mutation III (R48R, Q130H, K131Y, K142R, S411G, A423S, V526A, and A781K).

Random mutagenesis with different mutation frequencies

Three *Ctcdp* mutant libraries were generated with different mutation rates ranging from 0.14 ± 0.02% in Library L, to 0.28 ± 0.07% in Library O, and to 1.02 ± 0.44% in Library H. Three libraries had sizes of *ca.* 1.0×10^4. The mutations generated included all possible transitions and transversions with errors biased toward AT to GC changes.

Selection was designed based on the hypothesis that introduction of a heterologous cellobiose phosphorylase enables non-cellobiose-utilizing *E. coli* to grow in the M9 synthetic medium containing cellobiose as the sole carbon source (M9-CB plate).[32,59] On LB plates both *E. coli* JM109/pUC19 (a negative control) and *E. coli* JM109/pUCB that produced CBP grew well (Fig. 2A), while on the M9-CB plates only *E. coli* JM109/pUCB grew (Fig. 2B). These results validated the selection hypothesis. Since the selection approach ensured that only the transformants with active CBP expressed survived, the active CBP mutants were easily isolated from the mutant libraries (Fig. 2C). The difference of colony numbers in selection plates (M9-CB plates) against reference plates (LB plates) inferred that the selection powers were 64 ± 2%, 45 ± 3%, and 12 ± 2% for Library L, Library O, and Library H, respectively.

The following screening method was designed to further identify the enhanced thermostability CBP mutants. The colonies on the selection plates were imprinted on nylon membranes, and lysed *in situ* on membranes by heat. The heat treatment also deactivated the negative mutants and the wild type. Hence only more thermostable mutants remained active to degrade cellobiose and support the growth of the indicator strain *E. coli* JM109/pUC19 in the screening M9 plates containing cellobiose as a sole carbon source (Fig. 2D).

Four putative CBP mutants were found from Library L, one was identified from Library O, but none was isolated from Library H. All of the five putative mutants were purified and characterized. As a result, only one (namely M52) from Library O exhibited a considerably enhanced thermostability with the inactivation halftime at 70 °C extended to 16.6 min (Table 2). However, the turnover number (k_{cat}) and the catalytic efficiency (k_{cat}/K_m) of M52 decreased to 0.85 s^{-1} and 0.056 mM^{-1} s^{-1} from 3.40 s^{-1} and 0.206 mM^{-1} s^{-1} of the wild type, respectively (Table 2).

Identification of the beneficial mutations

The mutant M52 contains 7 mutations in the DNA sequence, five of which resulted in amino acid substitutions (Table 3). Consensus analysis indicated that R189 was strictly conserved among the 10 CBP homologies so that the mutation R189L may be the main reason for decrease in activity. Replacement of leucine back to the conserved arginine restored the activity of M52, but decreased its thermal stability slightly, as shown as M52m in Table 2.

It is interesting to note that the stabilizing mutation A423S identified by rational design was also found in M52. Reverse mutation R142K of the M52m and single mutation K142R of the wild type slightly reduced the halftime of inactivation at 70 °C to 12.4 min and 11.3 min, respectively (Fig. 3). K142R was confirmed to be a stabilizing mutation. Either reverse mutation K48R or A526V reduced the inactivation halftime of M52m. However, single mutation R48K and V526A cast deleterious effects on the thermostability of *Ct*CBP, shortening the halftime to 8.0 min and 7.9 min, respectively (Fig. 3). The results suggested that the combination of R48K and V526A improved the thermostability of M52m, even though introduction of R48K or V526A, individually, had negative impacts on the stability of *Ct*CBP.

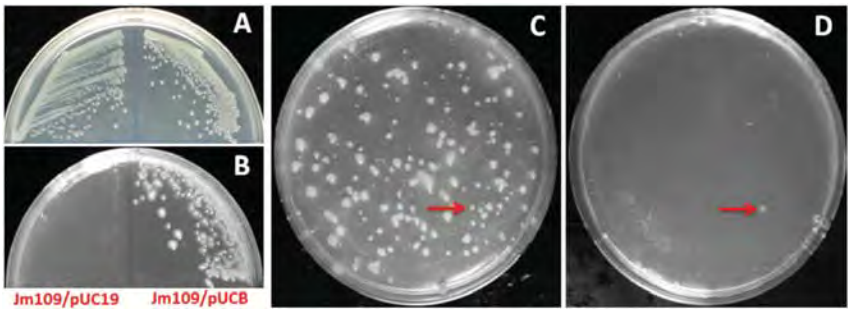

Fig. 2 The combinatorial selection/screening strategy. *E. coli* strain JM109/pUCB and the control JM109/pUC19 were streaked on the plates with LB medium (A) or M9 minimal medium plus 0.4% cellobiose (B). (C) A selection plate with the colonies expressing active cellobiose phosphorylase; (D) the screening plate corresponding to (C). The red arrows, pointing out the growth of indicator strain (JM109/pUC19) in the selection plate, highlighted the putative mutants in the selection plate.

Table 3 Consensus analysis of the mutations found in M52

DNA mutation	Amino acid mutation	Consensus analysis[a]
C9T	—	—
G143A	R48K	50% R, 50% K
A425G	K142R	60% K, 30% R, 10% T
G566T	R189L	100% R, 0% L
C681T	—	—
G1267T	A423S	50% S, 30% G, 20% A
T1577C	V526A	60% V, 40% A

[a] Consensus analysis was conducted based on the homology of the ten CBPs listed in Table 1.

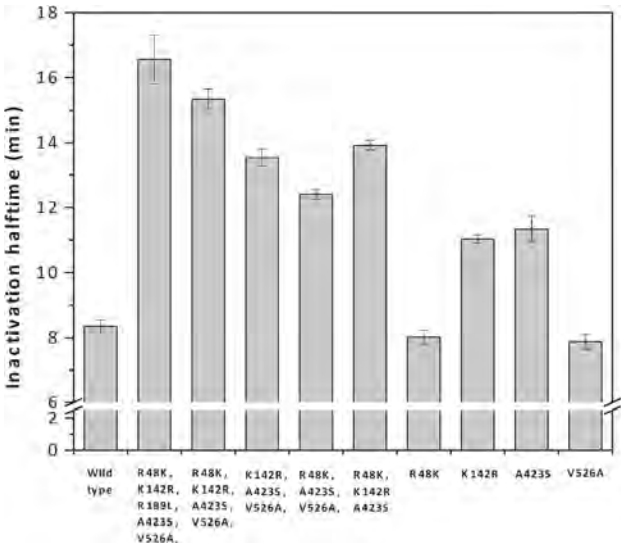

Fig. 3 Inactivation halftime (at 70 °C) for the mutants in relation to M52. By directed evolution, a thermostable mutant M52 was identified that contained 5 amino acid mutations, including R48K, K142R, R189L, A423S, and V526A. Replacement of L189 back to the strictly conserved arginine helped the enzyme restore its activity. The other mutations were also reversed in order to verify their effects. The results showed K142R and A423S improved the thermostability, while R48K and V526A individually had little impact on the stability of M52. But the combination of R48K and V526A improved the thermostability. It is a good example demonstrating the presence of synergistic epistasis as well as the unique advantages of optimal random mutagenesis.

Combining the stabilizing mutations from rational design and directed evolution

To further improve the thermal stability, the substitution A781K, with nearly two-fold increase of inactivation halftime, was firstly introduced into M52m, generating a new mutant II (CM2, as shown in Fig. 4). The inactivation halftime of CM2 was thus extended to 17.4 min (Table 2). The incremental effects were more significant as we incorporated all the stabilizing mutations identified by rational design to the M52m. The final mutant with combined mutations III (CM3, with R48K, Q130H, K131Y, K142R, S411G, A423S, V526A, and A781K) had inactivation halftime at 70 °C of 24.6 min, approximately three-fold longer than the wild type (Table 2). In contrast to the wild type, the temperature optimum of CM3 was raised from 60 to 80 °C (Fig. 5), with the activation energy (E_a) decreased from 59.2 to 31.7 kJ mol^{-1}. The catalytic efficiency was 0.103 mM^{-1} s^{-1}, nearly half of the wild type at 40 °C (Table 2). However, the specific activity of CM3 was 3.43 U mg^{-1}, 3-fold higher than the wild type (1.03 U mg^{-1}) at 80 °C (Fig. 5).

Discussion

Engineering *Ct*CBP was challenging because it is a large dimeric enzyme. Since the effort to construct combinatorial libraries increased substantially for large proteins and a large number of variants would tend to be inactive, directed evolution of *Ct*CBP had a low probability of success for whole sequence mutagenesis, especially in a traditional manner.[60,61] Therefore, De Groeve *et al.*[34] converted *Cellulomonas uda* cellobiose phosphorylase to lactose phosphorylase by generating random mutations in the region from T216 to V757 (542 out of the total 822 residues), while Yamamoto *et al.*[43] improved the thermostability of *Thermoanaerobacter brockii* kojibiose phosphorylase by running a random mutagenesis restricted to the residues between S269 and T700 (432 out of the total 775 residues). Additionally, *Ct*CBP itself is a thermostable protein.[36] The rational attempts to develop enhanced thermostability variants were also risky because even the mechanisms that stabilize *Ct*CBP under modest conditions (< 50 °C) remained unclear.

In this study, we fulfilled the task by two complementary approaches: structure-guided homology analysis and random mutagenesis with different mutation rates (Fig. 4). For the structure-guided homology analysis, the homologous set included ten diverse CBPs (to reduce biases), had high sequence identity (to minimize random drift), and adapted to very disparate temperatures (to maximize signal).[26] In comparison with previous homology-based protein design that set the vote weight of each species identical and focused on replacing poorly conserved residues at a given position with the most representative (or the consensus) type of residue,[18,62] our method favored the residues where substitutions were closely related with growth temperature and preferred the consensus residue among the thermophilic and hyperthermophilic counterparts. Then structural information was utilized to reduce the number of residues to be mutated for stabilization. Consequently, of the six mutants, four (~67%) rendered more stable effects, and the effects were additive. The best mutant A781K involved three base-pair changes in DNA (from GCT to AAA or AAG), which was often inaccessible by regular error-prone PCR. Two unsuccessful mutations, K201P and N292K, were located in the protein surface. Despite extensive investigation of protein stability enhanced by surface mutations,[22,57,63] no general rule was available to predict the stability of proteins upon surface mutations. The incomplete understanding of underlying mechanisms accounts for the failed rational attempts, and further points out a major drawback that restricts the application of rational design.

Random mutagenesis was conducted over the entire *Ctcbp* gene (2433 bp plus 419 bp pre- and pro sequences) to maximize the evolutionary search. A novel combinatorial selection/screening approach was established to meet the challenges of substantially expanded fitness landscapes.[16] The selection was designed for the mutants with adequate CBP activity, while the subsequent screening was developed for the mutants with improved thermostability. The whole process allowed fast and easy identification of the thermostability-enhanced active mutants. Meanwhile, unlike the traditional directed evolution strategy

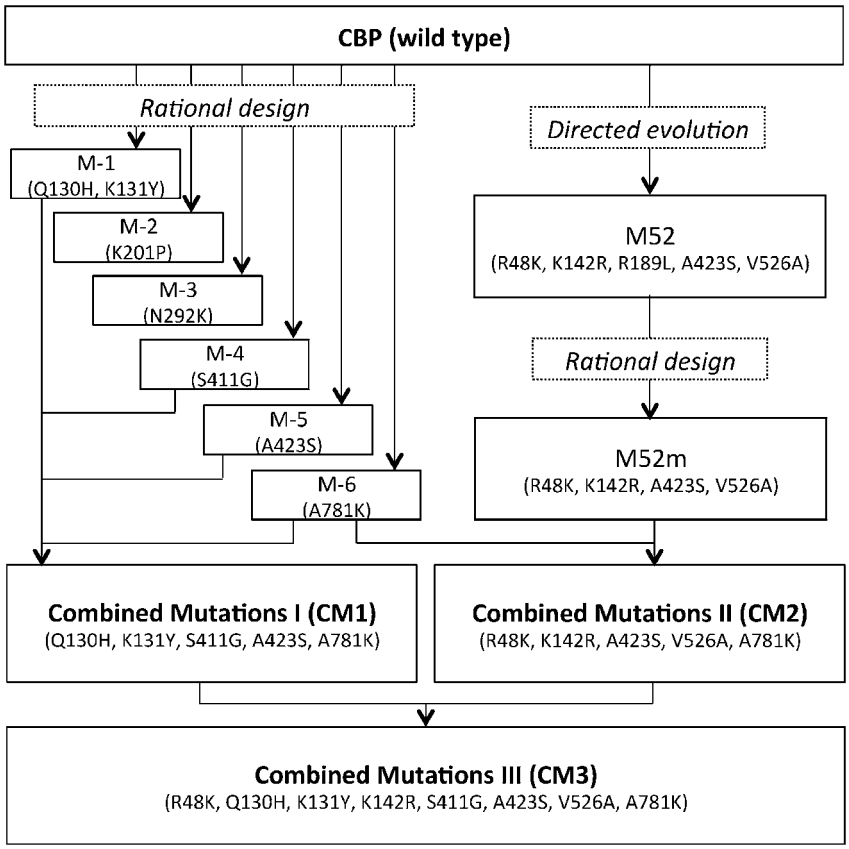

Fig. 4 Flowchart illustrating the development of thermostability-enhanced active *Ct*CBP mutants by a combination of rational design and directed evolution.

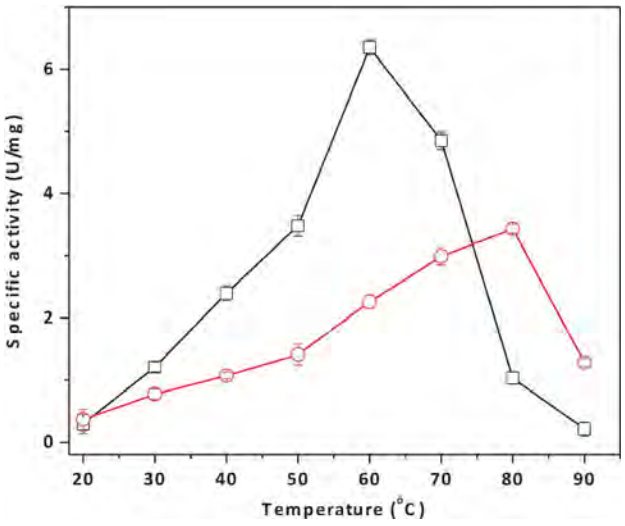

Fig. 5 Temperature optimum of *Ct*CBP and the mutant CM3. *Ct*CBP: black rectangles, CM3: red circles.

'*climbing fitness peaks one amino acid at a time*',[64] the mutant libraries were constructed at various mutagenesis frequencies here. Low mutagenesis frequency offered Library L a high probability of functional sequences and a low probability of beneficial mutations, whereas high mutagenesis frequency brought out a high probability of lethal mutations with a high probability of unique sequences to Library H.[65] Although the unique sequences of these two libraries were difficult to assess, such a theory resulted in the lowest selection power (63.7%) to Library L and the highest selection power (12.1%) to Library H. It was also believed an optimal mutation rate practically existed that balances diversity and retention of the function.[66–68] Following Sun's and Drummond's model,[65,69] we predicted the optimal mutation rate of this work was *ca.* 0.25% (Fig. S5, ESI†), by which Library O was developed. Four putative mutants from Library L and one from Library O were identified *via* selection and screening, whereas none was found from Library H possibly owing to the small number of surviving mutants (~1200) screened. Characterization of the purified enzymes demonstrated that the mutants from Library L did not have the properties of interest improved, while the one from Library O (M52) gained enhanced thermostability but significantly lost the catalytic activity. The false positive mutants screened by the *in vivo* approach did not lead to desired mutants *in vitro*, suggesting big differences between *in vivo* and *in vitro*.[70] In particular, the evolved improvement *in vivo* may stem from the increase in expression level, polymerase folding or stability that are specific to the context of the cytoplasm, rather than in the *in vitro* features.[71,72]

The mutant M52 contained five amino acid substitutions, *i.e.* R48K, K142R, R189L, A423S, and V526A. Consensus analysis suggested that it was mainly deactivated by R189L (Table 3). Reverse mutation L189R then significantly restored the enzyme activity. A423S was a stabilizing mutation, which had been proven by rational design. Replacing K142 with Arg could increase the salt bridges in the protein surface and

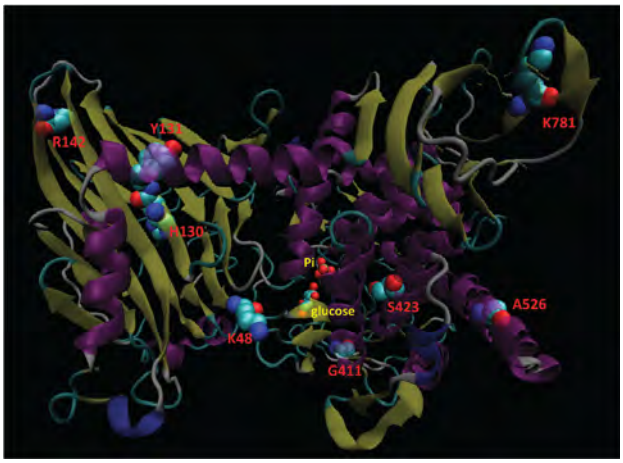

Fig. 6 Overview of the beneficial mutations identified from rational and random attempts. The ligands glucose and Pi were located according to the crystal structure of *Cellvibrio gilvus* CBP (PDB: 3AFJ), which share 63% sequence identity and 98% structural identity with *Ct*CBP.

thus strengthened the stability. R48, located in the CBM10 domain and near the entrance of active sites (Fig. 6), may affect the formation of the enzyme–substrate complex. V526 was located in an α-helix, sterically close to S423 (Fig. 6). The substitution R48K and V526A individually destabilized *Ct*CBP, whereas the synergistic mutations led to a modest increase in the thermostability. The above results were difficult to explain on structural ground, but clearly suggested that the combination of R48K, V526A, may together with A423S have a cooperative effect in stabilizing the *Ct*CBP. Since such kind of mutational epistasis was not implemented by the traditional directed evolution strategies or by current rational attempts,[16,65] it set a good example demonstrating the unique advantages of optimal random mutagenesis.

After all, several enhanced thermostability *Ct*CBP mutants were generated by rational design and directed evolution. The success of each method depended upon the level of understanding of the protein structure and function, or the effectiveness of the selection/screening scheme over large mutant libraries with the proper mutation frequencies. Although either rational design or directed evolution could be effective, a combination of both strategies represented the most successful route to engineer an enhanced thermostability active *Ct*CBP. The final combination of stabilizing mutations (CM3) identified from rational design and random mutagenesis extended the inactivation halftime at 70 °C to 24.6 min, three-fold higher than the wild type and more than two-fold higher than any mutant obtained from the above methods alone. In general, the mutants with enhanced stability displayed smaller k_{cat} and lower enzyme efficiency (k_{cat}/K_m) (Table 2), because high activity required a high flexibility of the protein to undergo conformational changes that cast a concomitant negative impact on thermostability.[60] Of interest, the mutant CB3 retained the specific activity three-fold higher than the wild type at 80 °C. The thermostability-enhanced active enzyme would work as a building block for cell-free SyPaB and has the potential for large-scale enzymatic biohydrogen production and glycoside synthesis under elevated temperatures.[35,41]

Conclusion

Recent advances in protein engineering expedited the development of robust enzymes tailored for industrial applications.[11,14] However, the knowledge of engineering a large protein (> 500 amino acid residues) was still very limited since the space of all possibilities was too large (and expensive) to exhaustively investigate.[73] Even though the extremozymes have attracted much attention,[8] there are few reports on enhancing the stability of thermostable enzymes because of the high risk of failure and the small space for further improvement. In this work, we developed two methods to fulfill these challenges. The rational method depended on homology analysis of ten CBP homologs from diverse hosts with a wide range of growth temperatures, by which the signal (stabilizing mutations) was maximized and the noise (ineffective mutations) was reduced. Meanwhile, structural information was facilitated to select the stabilizing mutations as well as to ensure success rates. As a result, four of six (~67%) were positive. Directed evolution of *Ct*CBP was performed at different mutation rates. In contrast to the traditional directed evolution strategies, high frequency mutagenesis, in an adaptive walk (3–25 mutations per *Ctcbp* gene), searched a larger sequence space that was then screened by a novel high-throughput selection/screening method. One enhanced thermostability mutant was identified with an apparent manifestation of mutational epistasis. Combining the best mutations from the rational and random attempts generated a *Ct*CBP mutant (CM3) with a three-fold increase in the inactivation halftime at 70 °C. The temperature optimum was raised from 60 to 80 °C. It was three times more active than the wild type at 80 °C. These results clearly demonstrated that a hybrid approach of rational design and directed evolution enabled to engineer large proteins and held great potential for creating extremozymes.

Acknowledgements

This work was not possible without support from the Biological Systems Engineering Department of Virginia Tech, the Air Force Office of Scientific Research (FA9550-08-1-0145), the USDA Biodesign and Bioprocess Center, and DOE BESC to YPZ. The authors appreciated the constructive suggestions from the editor and the reviewers.

References

1 M. Lehmann and M. Wyss, *Curr. Opin. Biotechnol.*, 2001, **12**, 371–375.
2 S. Kumar, C.-J. Tsai and R. Nussinov, *Protein Eng.*, 2000, **13**, 179–191.
3 W. Liu, X. Z. Zhang, Z. Zhang and Y. H. P. Zhang, *Appl. Environ. Microbiol.*, 2010, **76**, 4914–4917.
4 K. M. Polizzi, J. F. Chaparro-Riggers, E. Vazquez-Figueroa and A. S. Bommarius, *Biotechnol. J.*, 2006, **1**, 531–536.
5 X. Ye, J. Rollin and Y.-H. P. Zhang, *J. Mol. Catal. B: Enzym.*, 2010, **65**, 110–116.
6 S. Myung, X.-Z. Zhang and Y.-H. P. Zhang, *Biotechnol. Prog.*, 2011, **27**, 969–975.
7 G. D. Haki and S. K. Rakshit, *Bioresour. Technol.*, 2003, **89**, 17–34.
8 D. W. Hough and M. J. Danson, *Curr. Opin. Chem. Biol.*, 1999, **3**, 39–46.
9 F. Wen, N. U. Nair and H. Zhao, *Curr. Opin. Biotechnol.*, 2009, **20**, 412–419.
10 D. Böttcher and U. T. Bornscheuer, *Curr. Opin. Microbiol.*, 2010, **13**, 274–282.
11 R. Chen, *Trends Biotechnol.*, 2001, **19**, 13–14.
12 U. T. Bornscheuer and M. Pohl, *Curr. Opin. Chem. Biol.*, 2001, **5**, 137–143.

13 J. R. Cherry and A. L. Fidantsef, *Curr. Opin. Biotechnol.*, 2003, **14**, 438–443.
14 C. Schmidt-Dannert and F. H. Arnold, *Trends Biotechnol.*, 1999, **17**, 135–136.
15 S. Lutz, *Curr. Opin. Biotechnol.*, 2010, **21**, 734–743.
16 P. A. Romero and F. H. Arnold, *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, 866–876.
17 L. Giver, A. Gershenson, P.-O. Freskgard and F. H. Arnold, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 12809–12813.
18 M. Lehmann, D. Kostrewa, M. Wyss, R. Brugger, A. D'Arcy, L. Pasamontes and A. P. G. M. van Loon, *Protein Eng.*, 2000, **13**, 49–57.
19 R. Jaenicke, H. Schurig, N. Beaucamp and R. Ostendorp, in *Advances in Protein Chemistry*, ed. D. S. E. Frederic, M. Richards and S. K. Peter, Academic Press, 1996, vol. 48, pp. 181–269.
20 Y. H. P. Zhang, M. E. Himmel and J. R. Mielenz, *Biotechnol. Adv.*, 2006, **24**, 452–481.
21 X. Ye, Z. Zhu, C. Zhang and Y. H. Zhang, *Appl. Microbiol. Biotechnol.*, 2011, **92**, 551–560.
22 D. Perl, U. Mueller, U. Heinemann and F. X. Schmid, *Nat. Struct. Biol.*, 2000, **7**, 380–383.
23 B. Van den Burg, G. Vriend, O. R. Veltman, G. Venema and V. G. H. Eijsink, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 2056–2060.
24 B. Steipe, B. Schiller, A. Plückthun and S. Steinbacher, *J. Mol. Biol.*, 1994, **240**, 188–192.
25 E. Vázquez-Figueroa, J. Chaparro-Riggers and A. S. Bommarius, *ChemBioChem*, 2007, **8**, 2295–2301.
26 P. J. Haney, J. H. Badger, G. L. Buldak, C. I. Reich, C. R. Woese and G. J. Olsen, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 3578–3583.
27 X. Ye, J. Chu, Y. Zhuang and S. Zhang, *Front. Biosci.*, 2005, **10**, 961–965.
28 J. R. Cherry, M. H. Lamsa, P. Schneider, J. Vind, A. Svendsen, A. Jones and A. H. Pedersen, *Nat. Biotech.*, 1999, **17**, 379–384.
29 S. Akanuma, A. Yamagishi, N. Tanaka and T. Oshima, *Eur. J. Biochem.*, 1999, **260**, 499–504.
30 S. Akanuma, A. Yamagishi, T. Oshima and N. Tanaka, *Protein Sci.*, 1998, **7**, 698–705.
31 J. Lou, K. Dawson and H. Strobel, *Appl. Environ. Microbiol.*, 1996, **62**, 1770–1773.
32 Y. H. P. Zhang and L. R. Lynd, *Appl. Environ. Microbiol.*, 2004, **70**, 1563–1569.
33 A. Percy, H. Ono and K. Hayashi, *Carbohydr. Res.*, 1998, **308**, 423–429.
34 M. R. M. De Groeve, M. De Baere, L. Hoflack, T. Desmet, E. J. Vandamme and W. Soetaert, *Protein Eng., Des. Sel.*, 2009, **22**, 393–399.
35 X. Ye, Y. Wang, R. C. Hopkins, M. W. W. Adams, B. R. Evans, J. R. Mielenz and Y. H. P. Zhang, *ChemSusChem*, 2009, **2**, 149–152.
36 K. Tanaka, T. Kawaguchi, Y. Imada, T. Ooi and M. Arai, *J. Ferment. Bioeng.*, 1995, **79**, 212–216.
37 Y.-K. Kim, M. Kitaoka, M. Krishnareddy, Y. Mori and K. Hayashi, *J. Biochem.*, 2002, **132**, 197–203.
38 J. K. Alexander, *J. Biol. Chem.*, 1968, **243**, 2899–2904.
39 Y. H. P. Zhang, *Biotechnol. Bioeng.*, 2010, **105**, 663–677.
40 H. Nakai, M. A. Hachem, B. O. Petersen, Y. Westphal, K. Mannerstedt, M. J. Baumann, A. Dilokpimol, H. A. Schols, J. Ø. Duus and B. Svensson, *Biochimie*, 2010, **92**, 1818–1826.
41 M. R. M. De Groeve, T. Desmet and W. Soetaert, *J. Biotechnol.*, 2011, **156**, 253–260.
42 F. Kazutoshi, I. Masae, Y. Michiyo, T. Takeshi and K. Takashi, *J. Appl. Glycosci.*, 2006, **53**, 91–97.
43 T. Yamamoto, K. Mukai, H. Yamashita, M. Kubota, S. Fukuda, M. Kurimoto and Y. Tsujisaka, *J. Biosci. Bioeng.*, 2005, **100**, 212–215.
44 W. Liu, J. Hong, D. R. Bevan and Y.-H. P. Zhang, *Biotechnol. Bioeng.*, 2009, **103**, 1087–1094.
45 Y. H. P. Zhang, J. Cui, L. R. Lynd and L. R. Kuang, *Biomacromolecules*, 2006, **7**, 644–648.
46 J. Hong, Y. Wang, X. Ye and Y.-H. P. Zhang, *J. Chromatogr., A*, 2008, **1194**, 150–154.
47 Z. Dosztányi, C. Magyar, G. Tusnády and I. Simon, *Bioinformatics*, 2003, **19**, 899–900.
48 Z. Dosztányi, A. Fiser and I. Simon, *J. Mol. Biol.*, 1997, **272**, 597–612.
49 *Directed Evolution Library Creation: Methods and Protocols*, ed. F. H. Arnold and G. Georgiou, Humana Press, 2003.
50 Y. H. P. Zhang, J. Hong and X. Ye, in *Biofuels: Methods and Protocols*, ed. J. R. Mielenz, Humana Press Inc., 2009, vol. 581, pp. 213–231.
51 S. Saheki, A. Takeda and T. Shimazu, *Anal. Biochem.*, 1985, **148**, 277–281.
52 C. Lambert, N. Léonard, X. De Bolle and E. Depiereux, *Bioinformatics*, 2002, **18**, 1250–1256.
53 M. Hidaka, M. Kitaoka, K. Hayashi, T. Wakagi, H. Shoun and S. Fushinobu, *Biochem. J.*, 2006, **398**, 37–43.
54 C. M. Bianchetti, N. L. Elsen, B. G. Fox and G. N. Phillips, Jr., *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, 2011, **67**, 1345–1349.
55 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
56 E. Frare, P. P. de Laureto, E. Scaramella, F. Tonello, O. Marin, R. Deana and A. Fontana, *Protein Eng., Des. Sel.*, 2005, **18**, 487–495.
57 K.-B. Wong, C.-F. Lee, S.-H. Chan, T.-Y. Leung, Y. W. Chen and M. Bycroft, *Protein Sci.*, 2003, **12**, 1483–1495.
58 S. Kumar and R. Nussinov, *Cell. Mol. Life Sci.*, 2001, **58**, 1216–1233.
59 R. Sekar, H.-D. Shin and R. Chen, *Appl. Environ. Microbiol.*, 2012, **78**, 1611–1614.
60 J. D. Bloom and F. H. Arnold, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 9995–10000.
61 K. B. Zeldovich, P. Chen and E. I. Shakhnovich, *Proc. Natl. Acad Sci. U. S. A.*, 2007, **104**, 16152–16157.
62 X. Jiang, J. Kowalski and J. W. Kelly, *Protein Sci.*, 2001, **10**, 1454–1465.
63 R. A. Goldstein, *Protein Sci.*, 2007, **16**, 1887–1895.
64 C. A. Tracewell and F. H. Arnold, *Curr. Opin. Chem. Biol.*, 2009, **13**, 3–9.
65 D. A. Drummond, B. L. Iverson, G. Georgiou and F. H. Arnold, *J. Mol. Biol.*, 2005, **350**, 806–816.
66 T. Miura and P. Sonigo, *J. Theor. Biol.*, 2001, **209**, 497–502.
67 P. S. Daugherty, G. Chen, B. L. Iverson and G. Georgiou, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 2029–2034.
68 M. Zaccolo and E. Gherardi, *J. Mol. Biol.*, 1999, **285**, 775–783.
69 F. Sun, *J. Comput. Biol.*, 1995, **2**, 63–86.
70 Y.-H. P. Zhang, *Biotechnol. Adv.*, 2011, **29**, 715–725.
71 T. Bulter, M. Alcalde, V. Sieber, P. Meinhold, C. Schlachtbauer and F. H. Arnold, *Appl. Environ. Microbiol.*, 2003, **69**, 987–995.
72 K. M. Esvelt, J. C. Carlson and D. R. Liu, *Nature*, 2011, **472**, 499–503.
73 H. Zhao, K. Chockalingam and Z. Chen, *Curr. Opin. Biotechnol.*, 2002, **13**, 104–110.