

## Central Lancashire Online Knowledge (CLoK)

Title	Genome mining, in silico validation and phase selection of a novel aldo-keto reductase from <i>Candida glabrata</i> for biotransformation
Type	Article
URL	<a href="https://clock.uclan.ac.uk/18674/">https://clock.uclan.ac.uk/18674/</a>
DOI	<a href="https://doi.org/10.1080/21655979.2017.1342911">https://doi.org/10.1080/21655979.2017.1342911</a>
Date	2018
Citation	Basak, Souvik, Sahoo, Nanda Gopal and Pavanagam, Angayar (2018) Genome mining, in silico validation and phase selection of a novel aldo-keto reductase from <i>Candida glabrata</i> for biotransformation. <i>Bioengineered</i> , 9 (1). pp. 186-195. ISSN 2165-5979
Creators	Basak, Souvik, Sahoo, Nanda Gopal and Pavanagam, Angayar

It is advisable to refer to the publisher's version if you intend to cite from the work.  
<https://doi.org/10.1080/21655979.2017.1342911>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

To link to this article: <https://doi.org/10.1080/21655979.2017.1342911>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Accepted author version posted online: 23 Jun 2017.  
Published online: 13 Jul 2017.



Submit your article to this journal 



Article views: 238



[View Crossmark data](#)

ADDENDUM



## Genome mining, *in silico* validation and phase selection of a novel aldo-keto reductase from *Candida glabrata* for biotransformation

Souvik Basak<sup>a</sup>, Nanda Gopal Sahoo<sup>b</sup>, and Angayar K. Pavanasam<sup>c</sup>

<sup>a</sup>Dr. B.C. Roy College of Pharmacy & Allied Health Sciences, Durgapur, WB, India; <sup>b</sup>Nanoscience and Nanotechnology Centre, Department of Chemistry, Kumaun University, Nainital, Uttarakhand, India; <sup>c</sup>International College of Engineering and Management (University of Central Lancashire, UK Affiliation), Muscat, Oman

### ABSTRACT

Previously, we published cloning, overexpression, characterization and subsequent exploitation of a carbonyl reductase (*cr*) gene, belonging to general family aldo-keto reductase from *Candida glabrata* CBS138 to convert keto ester (COBE) to a chiral alcohol (ethyl-4-chloro-3-hydroxybutanoate or CHBE). Exploiting global transcription factor CRP, rDNA and transporter engineering, we have improved batch production of CHBE by trinomial bioengineering. Herein, we present the exploration of *cr* gene in *Candida glabrata* CBS138 through genome mining approach, *in silico* validation of its activity and selection of its biocatalytic phase. For exploration of the gene under investigation, 3 template genes were chosen namely *Saccharomyces cerevisiae* YDR541c, YGL157w and YOL151w. The CR showed significant homology match, overlapping of substrate binding site and NADPH binding site with the template proteins. The binding affinity of COBE toward CR (−4.6 Kcal/ mol) was found higher than that of the template proteins (−3.5 to −4.5 Kcal/ mol). Biphasic biocatalysis with cofactor regeneration improved product titer 4~5 times better than monophasic biotransformation. Currently we are working on DNA Shuffling as a next level of strain engineering and we demonstrate this approach herein as a future strategy of biochemical engineering.

### ARTICLE HISTORY

Received 3 March 2017

Revised 25 May 2017

Accepted 11 June 2017

### KEYWORDS





binding site; cofactor regeneration; docking; DNA shuffling; homology match

## Introduction

In previous studies, many enzymes of Aldo-keto Reductase (AKR) and Carbonyl reductase (CR) family were cloned, characterized and used in the asymmetric synthesis of (S)-CHBE<sup>1–4</sup> and (R) CHBE.<sup>5</sup> AKR & CR find extensive applications in pharmaceutical industry. For example, they are used as key chiral intermediates in the enantioselective synthesis of slagenins B and C, they serve as 2 potential compounds against murine leukemia, they are also used for synthesis of HMG-CoA reductase inhibitors (hypolipidemic agents) and can be converted into 1,4 dihydropyridine type blocker (antihypertensive agents).<sup>1–5</sup> Similarly, they are also used for conversion of ethyl-4-chloro-3-oxo-butanoate (COBE) to optically active ethyl-4-chloro-3-hydroxybutanoate (CHBE) as CHBE serve as a versatile precursor for pharmacologically valuable products.<sup>1–5</sup> Although attempts have been made to

augment bioconversion by either genetically manipulating the biocatalytic system with cofactor regeneration or fabricating the reaction media with single or multiple solvents, productivity has often faced short-fall due to obtaining higher reaction rate only at small substrate concentration (5~230 mM) thus eliciting a limiting batch output within the reactor.<sup>6–9</sup> Hence, establishment of a biocatalysis system with industrial competence has always been a prime search through the years which has driven researchers to find out newer proteins with improved activity and higher productivity.

In this context, in our previous publication,<sup>10</sup> we have reported the cloning, expression and purification of a carbonyl reductase from *Candida glabrata* CBS 138. We have also introduced the recombinant gene into globally engineered strain constructed by manipulating global transcription factor CRP. The improved

**CONTACT** Souvik Basak  [souvik\\_basak1@yahoo.com](mailto:souvik_basak1@yahoo.com)  Dr. B.C. Roy College of Pharmacy & Allied Health Sciences, Bidhan Nagar, Durgapur-713206, West Bengal, India; Nanda Gopal Sahoo  [ngsahoo@yahoo.co.in](mailto:ngsahoo@yahoo.co.in)  Nanoscience and Nanotechnology Centre, Department of Chemistry, Kumaun University, Nainital, Uttarakhand, India.

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

tolerance by the host cell against organic phase, which had been added as an essential component of the bio-transformation, led the output of the process shoot outstandingly higher. However, detailed methodology of the gene exploration together with optimization of the biocatalysis medium had not been described in details in the previous report. Thus, exploration of the gene through bioinformatics guided approach and its validation through docking studies has been presented in this work. In addition, this work also focuses on enhancing cell phenotype through DNA shuffling. Preliminary data obtained from DNA shuffling appears to be promising with better product titer and hence it is envisaged to extend our research on improvement of microbial cell factory through the above mentioned techniques.

### Discovery of CR protein from *Candida glabrata* CBS138

Three open reading frames namely *S. cerevisiae* YDR541c, YGL157w and YOL151w, reported for encoding aldo-keto reductases,<sup>11</sup> were subjected to Protein BLAST (BLASTP) along with 2 crystallographically elucidated proteins of the same class namely aldehyde reductase 2 from *Sporobolomyces salmonicolor* AKU4429 (PDB ID: 1UJM) and an aldehyde reductase from *Sporidiobolus salmonicolor* (PDB ID: 1Y1P). The BLASTP results from all attempts revealed the repetitive hypothetical protein from *Candida glabrata* CBS138 (Protein ID: XP\_445913.1). The BLAST scores revealed that the yeast proteins bear a plausible 55% to 62% identity to the target protein together with a staggering 98% sequence coverage with the latter. In contrast, 1UJM and 1Y1P possessed only 30% identity with the CR protein, however their 95% sequence swap with the target protein led us to a reliable approximation that XP\_445913.1 from *Candida glabrata* CBS138 might belong to the same family as *S. cerevisiae* YDR541c, YGL157w, YOL151w together with 1UJM and 1Y1P. Multiple sequence alignment of the structures showed that they preserve a high degree of homology match with each other (Fig. 1) with a comprehensive amount of conserved amino acids at most of the positions.

### Homology modeling

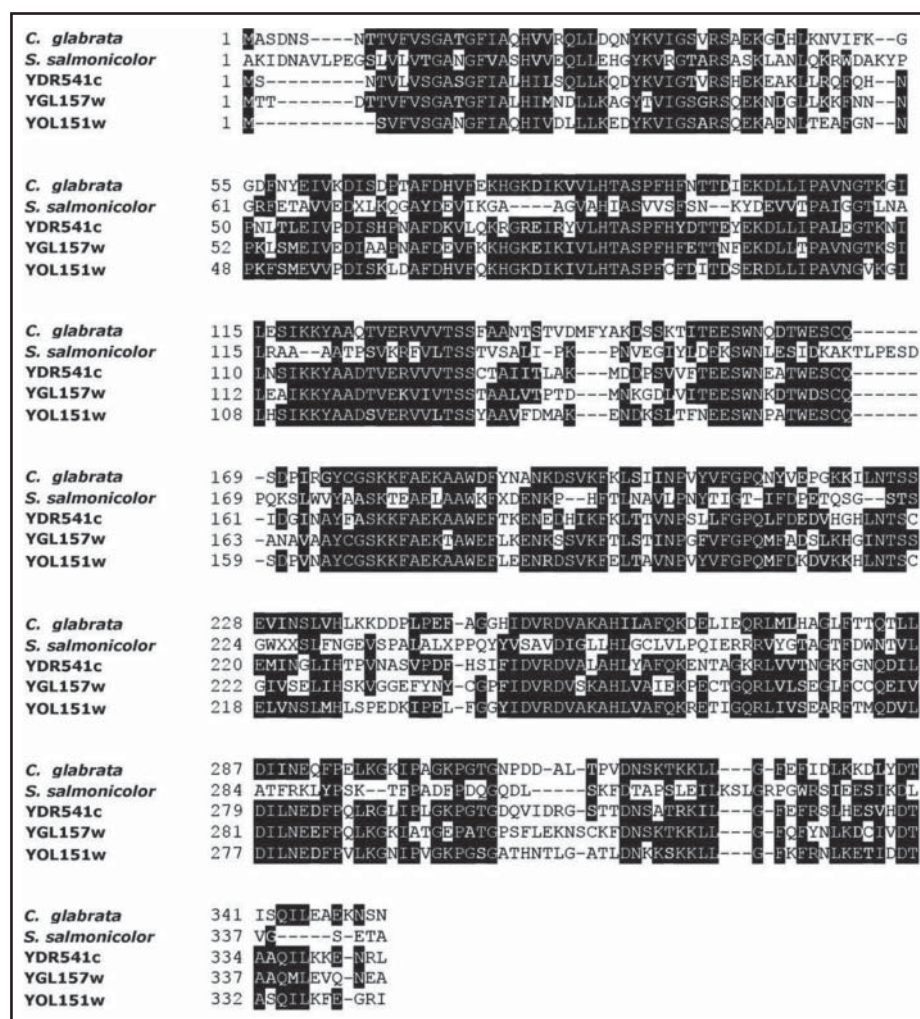
For further investigations of XP\_445913.1 from *C. glabrata* CBS138, homology model of the same was

constructed using 5 template proteins as used for BLAST. In this process, primary homology models have been created for *S. cerevisiae* YDR541c, YGL157w and YOL151w since crystallographic structure has not been elucidated till date for these proteins. Homology modeling of all the 4 proteins were acquiesced using MODELLER9.12 ([www.scilab.org/modeler](http://www.scilab.org/modeler)). The overall folding of the homology model structure was same (RMSD 0.466) as analyzed by swiss pdb.

A further insight and analogical comparison of the modeled structure of XP\_445913.1 with other proteins exhibited that the conserved catalytic residues such as S134, Y175 and K179 were similar to the crystal structure of carbonyl reductase from *Sporobolomyces salmonicolor* and other yeast proteins (Fig. 2). In addition, similarities have been obtained for a lot of other amino acids too such as amino acids spanning hydrophobic channel of the 2 proteins, such as Phe 94→Phe 97, Trp 226→Val229, Pro241→Ala238, Ileu172→Leu174. However, for XP\_445913.1, Q168 and E244 make the beginning of the channel little less hydrophobic than the crystal structure (1UJM) where there is Pro170 and Leu 241 in the equivalent positions (Table 1).

For estimation of modeling parameters, several parameters of our model has been performed such as ANOLEA (Atomic Non Local Environment),<sup>12</sup> QMEAN (Qualitative Model Energy Analysis),<sup>13</sup> GROMOS (Groningen Molecular Simulation Computer Simulation Package).<sup>14</sup> While ANOLEA calculates knowledge based distance dependent mean force potential, QMEAN evaluates the quality of the model based on certain scoring function and GROMOS is the force field based on molecular simulation. PROCHECK<sup>15</sup> has been used for Ramachandran Plot. The modeled protein quality check with ANOLEA, QMEAN and GROMOS have been provided in Figure 3. The model quality assessment has been performed using the Swiss-model workspace.<sup>16</sup>

Ramachandran plot has been performed to assess the model quality by analyzing the favored, allowed and generously allowed perturbations of residue-residue interaction (Fig. 4). From the plot, it can be observed that most of the residues have been clustered in  $\alpha$  and  $\beta$  regions with a very few outliers suggesting that most of the interactions are favored folding interactions. Total quality of the model is assessed based on the reliability model and was found to be QMEAN6 of 0.528. The coloring residues plot with respect to errors (Fig. 5A) and normalized QMEAN6 plot with respect



**Figure 1.** Multiple sequence alignment of XP\_445913.1 from *Candida glabrata* CBS138 with other reported Aldo-keto Reductase group of proteins namely one aldose reductase from *Sporobolomyces salmonicolor* and 3 *Saccharomyces cerevisiae* derived genes encoding carbonyl reductases (YDR541c, YGL157w, YOL151w). The conserved domains are highlighted as black and gray boxes

to query protein residues (Fig. 5B) have been provided to demonstrate the CR model accuracy. The pseudo energy plot of the contributing terms (Fig. 5C) has also been provided with their *z*-scores (with respect to the scores obtained from high-resolution structures in this protein subset). The scores obtained from high-resolution structures solved by X-ray crystallography has been taken as baseline score.

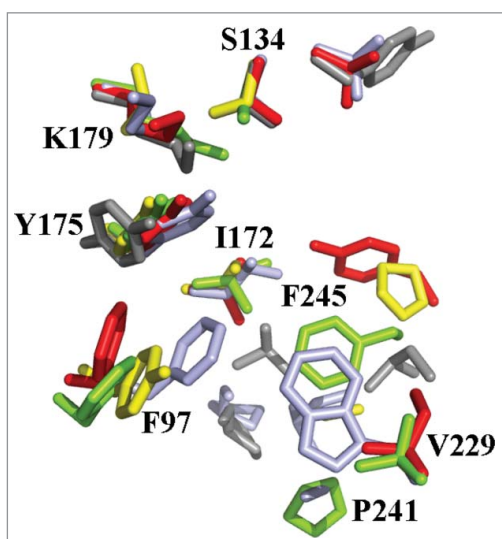
### Docking studies

The docking studies were performed in AutoDockVina to evaluate the binding affinity of the substrate with the enzyme. The binding affinity of the target enzyme was compared with that of aforementioned standard AKR proteins to elucidate the enzyme potential as aldose reductase against COBE. The energy minimized

structure of COBE was prepared and converted to PDBQT format through MGL Tools 1.5.6. The modeled protein structures were freed from water molecules and inbound ligand by Discovery Studio 3.5 Visualizer. All the bonds and torsional angles of the ligand were allowed to rotate freely. *C. glabrata* CR was found to have more binding affinity (−4.6 Kcal/mol) compared with other same family of proteins (binding affinity ranging from −3.5 to −4.5 Kcal/mol) (Table 2). A binding site analysis displayed that the substrate can fit nicely into the hydrophobic cavity of the enzyme and the amino acids especially T111 and Y175 can form hydrogen bonds with the substrate carbonyl oxygen atom (Fig. 6).

NADPH dependence for *C. glabrata* CBS138 CR was interpreted by aligning the crystallographic structure of *S. salmonicolor* AKR protein together with that of *C. glabrata* and followed by comparing the cofactor





**Figure 2.** Positions of important amino acids for COBE binding shown in stick representation by pyMOL. Overlap of crystal structure from *S. salmonicolor* (light blue) and homology model structure from *C. glabrata* (green), *S. cerevisiae* YDR541C (yellow), *S. cerevisiae* YGL157W (red), *S. cerevisiae* YOL151W (gray).

attachment domain between the 2 proteins (Fig. 7). The cofactor domain mapping of the target protein together with comparing it with that of *S. salmonicolor* (1UJM) and *S. cerevisiae* (YDR541c, YGL157w and YOL151w) AKR proteins has been accomplished using standard alignment and labeling tools of PyMOL ([www.pymol.org](http://www.pymol.org)).

Thus homology model construction, alignment of important amino acids residues, analyzing substrate binding site, exploring cofactor domain and docking score comparisons with other standard ALR/ CR group of enzymes suggested that XP\_445913.1 should belong to the same family of enzymes as the standards and thus we treated XP\_445913.1 as CR (Carbonyl Reductase) group of protein from *Candida glabrata* CBS138. These findings encouraged us to try and explore its actual biocatalytic potential in realistic experimental condition. Thus we cloned the gene in heterologous vector, overexpressed and subsequently purified the protein, characterized through kinetic studies, calibrated

through pH and temperature and finally exploited it in actual bioconversion through trinomial bioengineering as described in our earlier report.

### Optimization of bioconversion

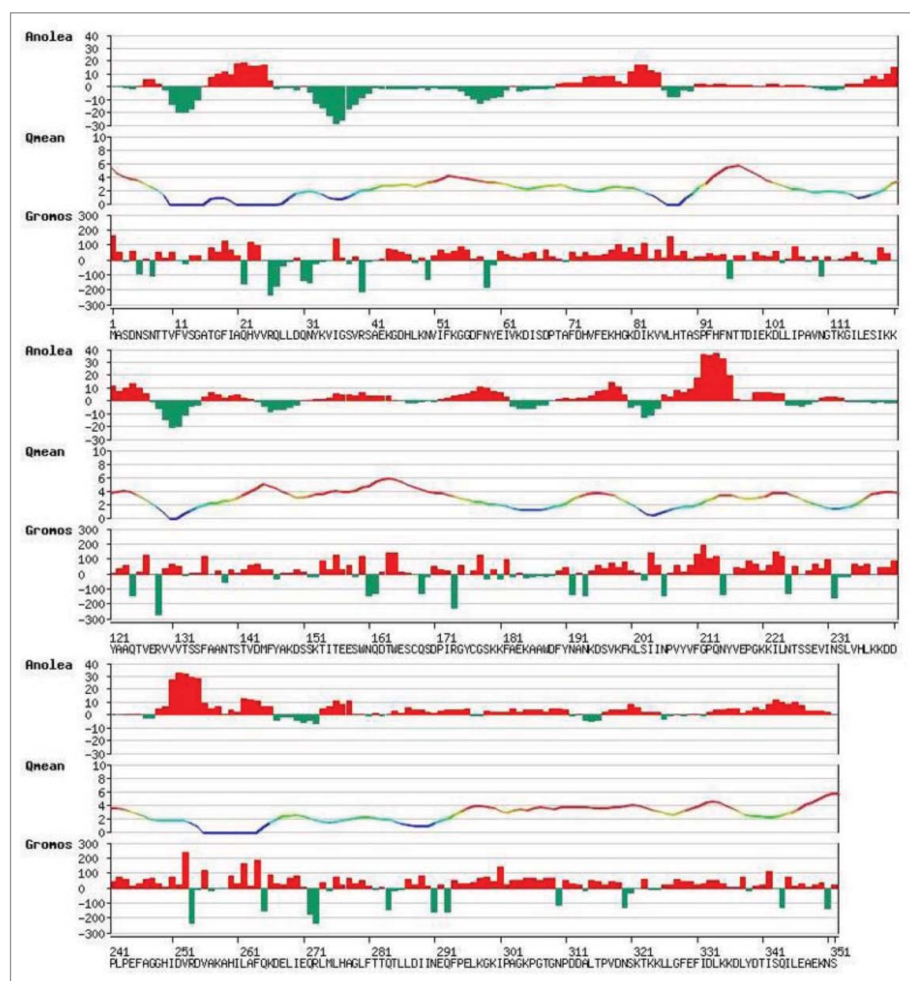
The optimization of the reaction system was accomplished by analyzing the outputs with any 2 variables of the 3 key factors controlling the bioconversion such as monophasic system, biphasic system and NADPH as cofactor. Apparently, best yields were obtained when biphasic reaction system was used together with cofactor regeneration as reported by other researchers.<sup>1-5</sup> As mentioned earlier,<sup>10</sup> gene encoding Glucose dehydrogenase (GDH) from *Bacillus subtilis* has been cloned in the recombinant microbial cell factory together with CR from *Candida glabrata* CBS 138 to follow cofactor regeneration of NADPH from oxidized NADP<sup>+</sup> via exogenously added substrate glucose. When whole cells over-expressing the CR and GDH proteins (cofactor regenerating) were used with previously reported conditions<sup>10</sup> within biphasic reaction system, the substrate being in butyl acetate phase while recombinant host in the buffer phase, the product formation improved rapidly up to 4h and obtained steady-state within 6h (88.3% bioconversion from COBE to CHBE, Fig. 8A). In contrast, without cofactor regeneration even in biphasic system, only 5% bioconversion was achieved together with having a longer reaction time ~8–10 h (Fig. 8A). However, a cofactor regenerating system within single buffer phase (0.1 M Potassium phosphate buffer, pH 7.5) produced a bioconversion of 20.80% necessitating the contribution of cofactor in enhancing the product yield. Figure 8B depicts a comparative portfolio of bioconversion under 3 aforementioned conditions.

### DNA shuffling- Library formation and mutant selection for future host construction

We used error-prone PCR (ep-PCR) in our previous study<sup>10</sup> to construct highly stress tolerant mutant as

**Table 1.** Comparison of binding site.

	Catalytic residues					Hydrophobic channel				
<i>C. glabrata</i>	S133	Y 175	K 179	F 94	I 172	F 245	V229	P 241	—	
<i>S. salmonicolor</i>	S134	Y 177	K 181	F 97	L 174	L 241	W 226	A 238	P170	
<i>S. cerevisiae</i> YDR541C	S129	Y 167	K 171	Y 89	I 164	P 235	—	—	—	
<i>S. cerevisiae</i> YGL157W	S131	Y 169	K 173	F 91	V 166	Y 239	I223	—	—	
<i>S. cerevisiae</i> YOL151W	S127	Y 165	K 169	—	V 162	I 232	—	—	P161	



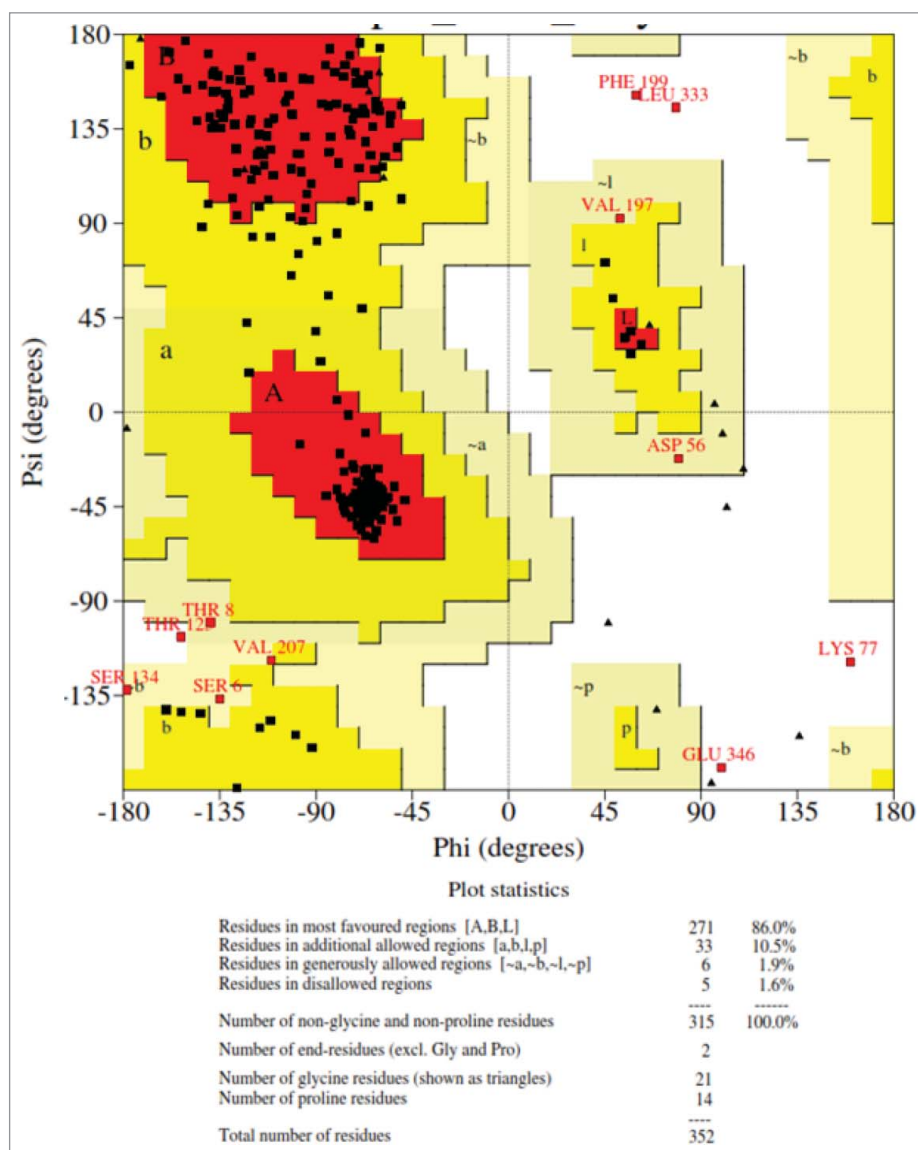
**Figure 3.** ANOLEA, QMEAN and GROMOS plot of modelled XP\_445913.1 from *Candida glabrata* CBS138.

our biocatalysis host and thus achieved significant improvement of product yield with elongation of cell fitness during the biphasic reaction. However we are still in search for further improvement of cell phenotype as we believe that cell fitness can be a key factor to tower the product yield during such bioreaction. Many such strategies are under investigation, among them DNA shuffling has emerged as a successful tool to improve cell phenotype in multitude of conditions.<sup>17-20</sup>

DNA shuffling was performed with a little modification of the procedure as described by Stemmer, 1994<sup>21</sup> taking whole *crp* operons of 3 ep-PCR mutants (M1 ~ M3) as templates. The acquisition of M1~ M3 and their amino acid mutations have been already described in our previous publication.<sup>10</sup> 4  $\mu$ g of total template DNA was used for DNA shuffling. The total template DNA was digested by DNase-I at 15°C for 3 mins and subsequently 50–200 bp DNA fragments had been recovered from gel electrophoresis for

further process. Afterwards, DNA fragments were subjected to undergo PCR without primer and finally chimeric *crp* was recovered by amplification with forward and reverse *crp* primer, the sequences as discussed in our previous publication.

The DNA shuffling library was finally constructed by cloning the chimeric *crp* into pACYC Duet-I plasmid with *Kpn* I and *Bam* HI and subsequent introduction into *E. coli* DH5 $\alpha$  as discussed earlier.<sup>10</sup> The introduction of chimeric *crp*-pACYC Duet-I conjugate into *E. coli* DH5 $\alpha$  through electroporation yielded a variant library in the order of  $10^4 \sim 10^5$ . DNA shuffling mutants were grown in LBGMg medium (Bacto tryptone 10 g/L, Yeast extract 1 g/L, NaCl 10 g/L, Glucose 2 g/L, MgSO<sub>4</sub> 10 mM) and the winner was selected by subjecting the mutant library in 3 rounds of selection with 0.23~0.25% (v/v) toluene. 50 mL High Density Polypropylene centrifuge tubes (BD Biosciences, USA) with parafilm sealing have been used for culturing the variant with proper oxygenation (37°C, 200 rpm).



**Figure 4.** Ramachandran Plot of modelled XP\_445913.1 from *Candida glabrata* CBS138.

Proper dilution of the culture from each round was plated onto LBGMg-agar plates to isolate the colonies. Selected clones from the third round were taken, the plasmid DNA was isolated by mini-preparation with QIAGEN plasmid isolation kit (Qiagen, USA). The shuffled mutations were verified by DNA sequencing and relevant plasmids were re-introduced into fresh *E. coli* DH5 $\alpha$  background to create fresh variants. The fresh variants were challenged in LBGMg medium under 0.4% – 0.5% (v/v) toluene pressure to select the variant with best growth profile.

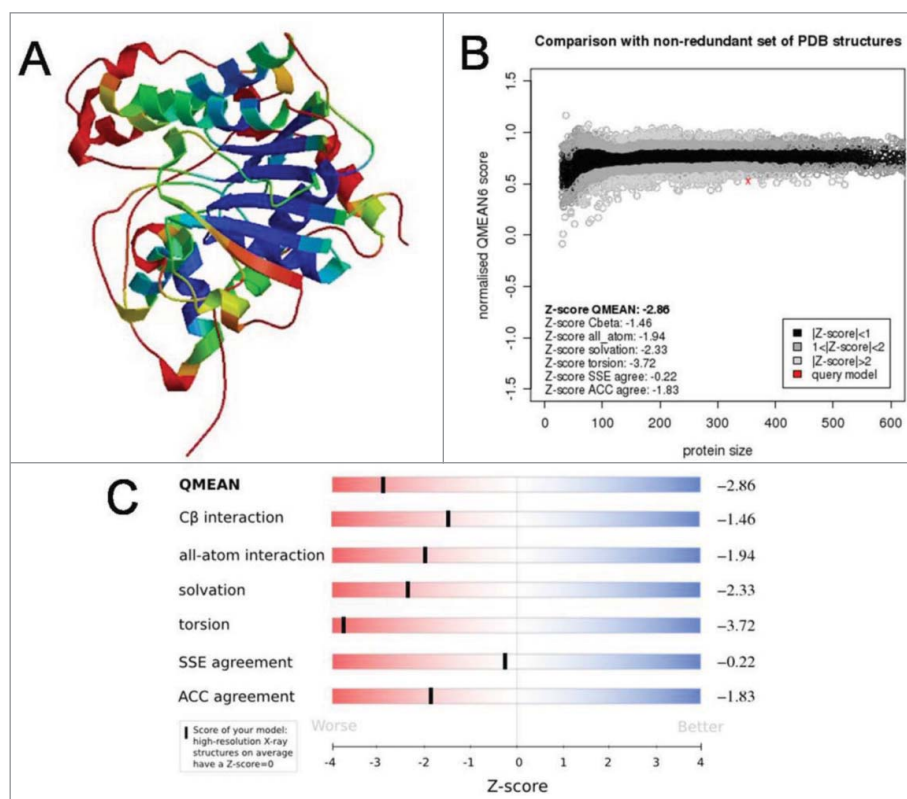
One DNA shuffling mutant (DSM) was isolated and sequenced to reveal amino acid substitutions such as T127N F136I T208N. The DSM revealed better growth profile against higher concentrations of

toluene (0.4% – 0.5% (v/v)). Under 0.40% toluene, the DSM reached to OD<sub>600</sub> ~2.5 in 24 h where other mutants' growth (M1~ M3) remained within OD 2.0 ~2.3 (Fig. 9A). In 0.50% toluene, where other reached a saturation OD ~2.0 in 22 h, DSM reached a OD ~2.4 in the same time (Fig. 9B).

## Discussion and conclusion

The CR from *Candida glabrata* was identified through BLASTP guided sequence search method. It is known that if there is a significant sequence match between 2 proteins, then it is highly likely that there are similarities in their functions. In this research work, it was found that there was a 55% ~62% identity match





**Figure 5.** Different model validation parameters of model XP\_445913.1 from *Candida glabrata* CBS138. A. Coloring residue plot with errors (blue depicting most reliable residues while red suggests potentially unreliable residues). B. Pseudo-energy plot of contributing terms with their z-scores. C. QMEAN6 plot with query residues.

between CR and 3 reported AKR family of proteins YDR151c, YGL157w and YOL151w.<sup>11</sup> This suggests that the target protein might have an AKR like activity. The overlapping of active site between the target and the template proteins also imply similar tertiary structure of the proteins.

Since the CR has been structurally elusive together with YDR151c, YGL157w and YOL151w, all the proteins have been modeled to churn any similarities between their structures as well as the active sites. The models were first validated then the active site comparison revealed that similar hydrophobic channels spanned the target and template protein residues including structurally elucidated proteins 1UJM and 1Y1P. Thus it is a plausible assumption the substrate COBE may bind with the proteins in the same way,

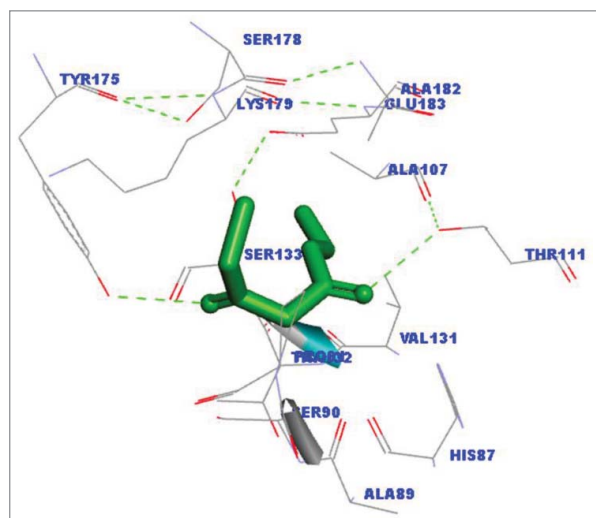
however the specificity and affinity may change depending upon the hydrophobicity of the channel. Furthermore, an *in silico* docking of COBE with the proteins revealed that COBE has even more binding affinity toward CR than other template proteins. This really encouraged us to try the CR from *Candida glabrata* in realistic bioconversion.

One critical bottleneck for enzymatic bioconversion is the hydrophilicity of the enzyme since enzymes usually are dissolved in buffered system for prolonging its stability and activity. Thus, a prior investigation of CR in this sort through GRAVY (grand average of hydrophobicity) analysis (<http://web.expasy.org/cgi-bin/protparam/protparam>) revealed a negative gravity index as  $-0.359$ . This indicated protein's hydrophilicity<sup>22</sup> thus advocating its competence in buffer mediated bioconversion. Then we challenged the enzyme in actual bioconversion and optimized it through whole cell biotransformation.

It is important to note that biphasic biotransformation provides better output in this reaction. The reason being that improved solubility of COBE in the organic phase (compared with aqueous) enhances the

**Table 2.** Docking scores of proteins.

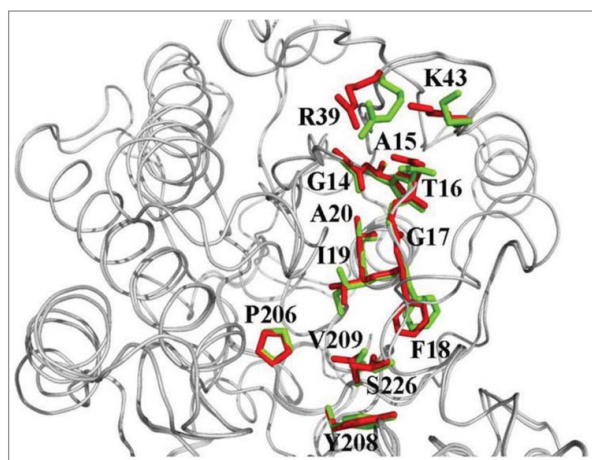
Protein	Binding affinity (Kcal/mol)
<i>C. glabrata</i> XP_445913.1	-4.6
<i>S. cerevisiae</i> YDR541c	-3.9
<i>S. cerevisiae</i> YGL157w	-3.5
<i>S. cerevisiae</i> YGL157w	-4.5
AKR protein from <i>S. salmonicolor</i>	-3.7



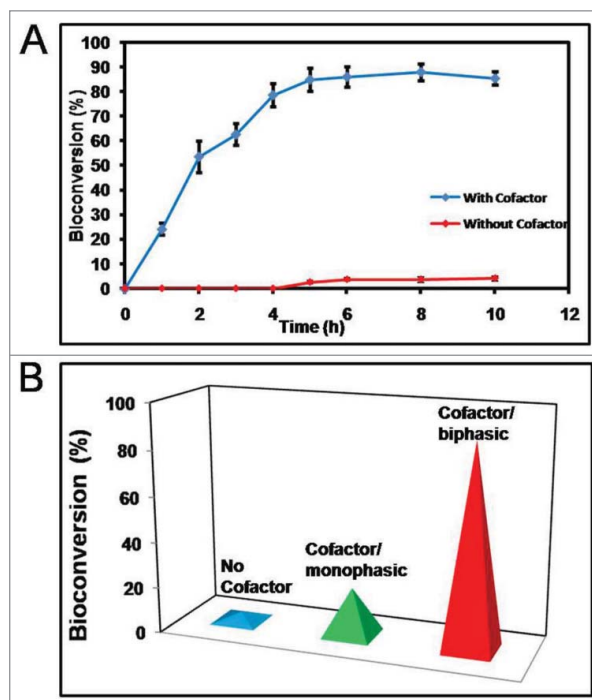
**Figure 6.** Binding site analysis of COBE in CR from *Candida glabrata* CBS138. The binding site analysis has been done using the standard protocol in Discovery Studio Visualizer 3.5. The dotted bond represents hydrogen bonding.

substrate carriage to the host cells. Also, COBE undergoes partial hydrolysis in aqueous phase which retards the rate of reaction in the monophasic biotransformation (data not shown). Cofactors are required to supply the hydrogen for the biotransformation.<sup>1-5,23</sup>

Harboring the aforementioned highly efficient recombinant enzyme for improved biotransformation, we have tried to overcome secondary challenges in biphasic biocatalysis such as tolerance of the host cell toward organic phase and improved substrate uptake inside the host cell reported in our previous publication.<sup>10</sup> We have demonstrated that error prone PCR has been undertaken to improve cell phenotype

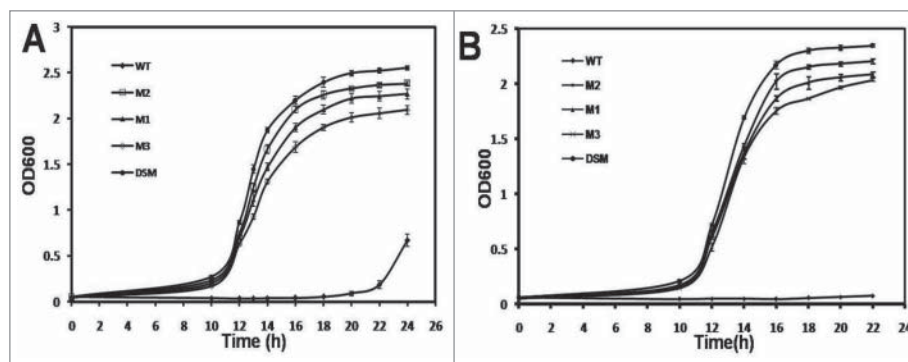


**Figure 7.** Positions of important amino acids for NADPH binding for *C. glabrata* (green) and *S. salmonicolor* (red).



**Figure 8.** (R)-CHBE production profile in biotransformation reaction. A. Yield profile with or without cofactor regeneration B. Comparative study of bioconversion under 3 reaction schemes The data points are the average ( $\pm$  standard deviation) of 3 independent observations ( $n = 3$ ).

during biocatalysis. Also, cell phenotype has been improved by DNA shuffling. For DNA shuffling we have used whole *crp* operon from the mutants because the *crp* operon contains 3 parts: the transcription factor binding site (Transcription Factor B involving 2 CRP-cAMP binding sites, 4 FISbinding sites, and 3 *crp* native promoters), the *crp* gene, and a specifically designed *rrnB* terminator.<sup>24,25</sup> The resultant *in vitro* recombination yielded a super-performing mutant (DSM) which showed improved performance than the mothers against extremely hydrophobic toluene. Toluene has been selected as the challenging solvent because a “winner” against such extremely hydrophobic solvent would have good chance to survive other organic solvents too. LBGMg medium has been a choice of medium for survival of colonies against organic solvent.<sup>26</sup> The DSM combines mutations at 3 regions namely T127 (in the cAMP binding pocket- $\alpha$  helix that stabilizes cAMP- CRP binding),<sup>27</sup> F136 (stabilizing interdomain hinge)<sup>28</sup> and T208 (in the DNA binding domain).<sup>27</sup> We propose that implementation of DNA Shuffling mutant in future biotransformation might improve the product titer even better and might gain other applications as well.



**Figure 9.** Tolerance profile of DNA shuffling mutant (DSM) with respect to parent mutants. The tolerance has been analyzed by growth of the mutant against hydrophobic solvent such as A. 0.40% (v/v) toluene B. 0.50% (v/v) toluene

## Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

## Acknowledgment

The authors are highly thankful to School of Chemical and Biomedical Engineering (SCBE), Nanyang Technological University (NTU), Singapore for providing the laboratory infrastructure to carry out biomolecular experiments. The authors are also grateful to Dr. Jiang Rongrong, Asst/P. SCBE, NTU, Singapore for her help and suggestions regarding the study.

## References

- [1] Wang Q, Shen L, Ye T, Cao D, Chen R, Pei X, Xie T, Li Y, Gong W, Yin X. Overexpression and characterization of a novel (S)-specific extended short-chain dehydrogenase/reductase from *Candida parapsilosis*. *Bioresour Technol* 2012; 123:690-4; PMID:22939596; <https://doi.org/10.1016/j.biortech.2012.07.060>
- [2] Cao H, Mi L, Ye Q, Zang G, Yan M, Wang Y, Zhang Y, Li X, Xu L, Xiong J, et al. Purification and characterization of a novel NADH-dependent carbonyl reductase from *Pichia stipitis* involved in biosynthesis of optically pure Ethyl (S)-4-Chloro-3-hydroxybutanoate. *Bioresour Technol* 2011; 102:1733-9; PMID:20933386; <https://doi.org/10.1016/j.biortech.2010.08.072>
- [3] Ye Q, Cao H, Mi L, Yan M, Wang Y, He Q, Li J, Xu L, Chen Y, Xiong J, et al. Biosynthesis of (S)-4-chloro-3-hydroxybutanoate ethyl using *Escherichia coli* co-expressing a novel NADH-dependent carbonyl reductase and a glucose dehydrogenase. *Bioresour Technol* 2010; 101:8911-4; PMID:20630744; <https://doi.org/10.1016/j.biortech.2010.06.098>
- [4] You ZY, Liu ZQ, Zheng YG. Characterization of a newly synthesized carbonyl reductase and construction of a biocatalytic process for the synthesis of Ethyl (S)-4-Chloro-3-hydroxybutanoate with high space-time yield. *Appl Microbiol Biotechnol* 2014; 98:1671-80; <https://doi.org/10.1007/s00253-013-5042-3>
- [5] Liu Y, Xu Z, Jing K, Jiang X, Lin J, Wang F, Cen P. Asymmetric reduction of ethyl 4-chloro-3-oxobutanoate to ethyl (R)-4-chloro-3-hydroxybutanoate with two co-existing, recombinant *Escherichia coli* strains. *Biotechnol Lett* 2005; 27:119-25; PMID:15703875; <https://doi.org/10.1007/s10529-004-7336-0>
- [6] Chen XH, Wei P, Wang XT, Zong MH, Lou WY. A novel carbonyl reductase with anti-Prelog stereospecificity from *Acetobacter sp.* CCTCC M209061: purification and characterization. *PLoS One* 2014; 9:e94543; PMID:24740089; <https://doi.org/10.1371/journal.pone.0094543>
- [7] Wang L-J, Li CX, Ni Y, Zhang J, Liu X, Xu JH. Highly efficient synthesis of chiral alcohols with a novel NADH-dependent reductase from *Streptomyces coelicolor*. *Bioreour Technol* 2011; 102:7023-8; PMID:21570826; <https://doi.org/10.1016/j.biortech.2011.04.046>
- [8] Yoon SA, Kim HK. Development of a bioconversion system using *Saccharomyces cerevisiae* Reductase YOR120W and *Bacillus subtilis* glucose dehydrogenase for chiral alcohol synthesis. *J Microbiol Biotechnol* 2013; 23:1395-402; PMID:23770562; <https://doi.org/10.4014/jmb.1305.05030>
- [9] Lee DH, Lee YJ, Ryu YW, Seo JH. Molecular cloning and biochemical characterization of a novel erythrose reductase from *Candida magnoliae* JH110. *Microb Cell Fact* 2010; 9:43; <https://doi.org/10.1186/1475-2859-9-43>
- [10] Basak S, Ghosh SK, Punetha VD, Aphale AN, Patra PK, Sahoo NG. An experimental modeling of trinomial bioengineering- crp, rDNA, and transporter engineering within single cell factory for maximizing two-phase bioreduction. *Int J Biol Macromol* 2017; 95:818-25; PMID:27923567; <https://doi.org/10.1016/j.ijbiomac.2016.12.001>
- [11] Kaluzna IA, Matsuda T, Sewell AK, Stewart JD. Systematic investigation of *Saccharomyces cerevisiae* enzymes catalyzing carbonyl reductions. *J Am Chem Soc* 2004; 126:12827-32; PMID:15469278; <https://doi.org/10.1021/ja0469479>
- [12] Melo F, Feytmans E. 1998. Assessing protein structures with a non-local atomic interaction energy. *J Mol Biol* 1998; 277:1141-52.

- [13] Benkert P, Tosatto SC, Schomburg D. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* 2008; 71:261-77; <https://doi.org/10.1002/prot.21715>
- [14] Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, Heinz TN, Kastenholz MA, Kräutler V, Oostenbrink C, et al. The GROMOS software for biomolecular simulation: GROMOS05. *J Comput Chem* 2005; 26:1719-51; PMID:16211540; <https://doi.org/10.1002/jcc.20303>
- [15] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 1993; 26:283-91; <https://doi.org/10.1107/S0021889892009944>
- [16] Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 2006; 22:195-201; PMID:16301204; <https://doi.org/10.1093/bioinformatics/bti770>
- [17] Zhang H, Chong H, Ching CB, Jiang R. Random mutagenesis of global transcription factor cAMP receptor protein for improved osmotolerance. *Biotechnol Bioeng* 2012; 109:1165-72; PMID:22179860; <https://doi.org/10.1002/bit.24411>
- [18] Zhang H, Chong H, Ching CB, Song H, Jiang R. Engineering global transcription factor cyclic AMP receptor protein of *Escherichia coli* for improved 1-butanol tolerance. *Appl Microbiol Biotechnol* 2012; 94:1107-17; PMID:22466954; <https://doi.org/10.1007/s00253-012-4012-5>
- [19] Cramer A, Raillard SA, Bermudez E, Stemmer WP. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 1998; 391:288-91; PMID:9440693; <https://doi.org/10.1038/34663>
- [20] Gao X, Zhao H, Zhang G, He K, Jin Y. Genome shuffling of *Clostridium acetobutylicum* CICC 8012 for improved production of acetone-butanol-ethanol (ABE). *Curr Microbiol* 2012; 65:128-32; PMID:22562601; <https://doi.org/10.1007/s00284-012-0134-3>
- [21] Stemmer WP. DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Proc Natl Acad Sci USA* 1994; 91:10747-51; PMID:7938023; <https://doi.org/10.1073/pnas.91.22.10747>
- [22] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982; 157:105-32; [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)
- [23] He YC, Tao ZC, Zhang X, Yang ZX, Xu JH. Highly efficient synthesis of ethyl (S)-4-chloro-3-hydroxybutanoate and its derivatives by a robust NADH-dependent reductase from *E. coli* CCZU-K14. *Bioresource Technol* 2014; 161:461-4; <https://doi.org/10.1016/j.biortech.2014.03.133>
- [24] Brosius J, Dull TJ, Sleeter DD, Noller HF. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J Mol Biol* 1981; 148:107-27; [https://doi.org/10.1016/0022-2836\(81\)90508-8](https://doi.org/10.1016/0022-2836(81)90508-8)
- [25] Okamoto K, Freundlich M. Mechanism for the autogenous control of the *crp* operon: Transcriptional inhibition by a divergent RNA transcript. *Proc Natl Acad Sci USA* 1986; 83:5000-4; <https://doi.org/10.1073/pnas.83.14.5000>
- [26] Basak S, Song H, Jiang R. Error-prone PCR of global transcription factor cyclic AMP receptor protein forenhanced organic solvent (toluene) tolerance. *Process Biochem* 2012; 47:2152-8; <https://doi.org/10.1016/j.procbio.2012.08.006>
- [27] Won HS, Lee YS, Lee SH, Lee BJ. Structural overview on the allosteric activation of cyclic AMP receptor protein. *Biochim Biophys Acta* 2009; 1794:1299-308; PMID:19439203; <https://doi.org/10.1016/j.bbapap.2009.04.015>
- [28] Fas BA, Tutar Y, Haliloglu T. Dynamic Fluctuations Provide the Basis of a Conformational Switch Mechanism in Apo Cyclic AMP Receptor Protein. *PLoS Comput Biol* 2013; 9:e1003141; PMID:23874183; <https://doi.org/10.1371/journal.pcbi.1003141>