

Enantioselective Enzymes by Computational Design and In Silico Screening**

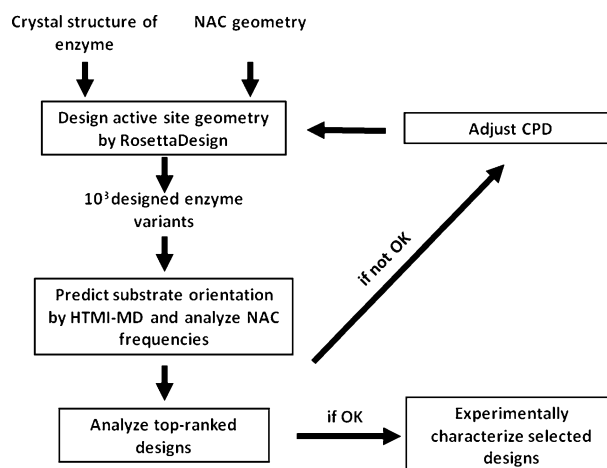
Hein J. Wijma, Robert J. Floor, Sinisa Bjelic, Siewert J. Marrink, David Baker, and Dick B. Janssen*

Abstract: Computational enzyme design holds great promise for providing new biocatalysts for synthetic chemistry. A strategy to design small mutant libraries of complementary enantioselective epoxide hydrolase variants for the production of highly enantioenriched (*S,S*)-diols and (*R,R*)-diols is developed. Key features of this strategy (CASCO, catalytic selectivity by computational design) are the design of mutations that favor binding of the substrate in a predefined orientation, the introduction of steric hindrance to prevent unwanted substrate binding modes, and ranking of designs by high-throughput molecular dynamics simulations. Using this strategy we obtained highly stereoselective mutants of limonene epoxide hydrolase after experimental screening of only 37 variants. The results indicate that computational methods can replace a substantial amount of laboratory work when developing enantioselective enzymes.

Recent successes with de novo computational protein design (CPD)^[1] and redesign of enzyme active sites^[2] suggest that computational methods can be used for controlling enzyme selectivity, which is an important goal in industrial biocatalyst development. Enantioselectivity engineering is often pursued by directed evolution, in most cases with the use of large mutant libraries.^[3] Although important successes have been reported,^[3] further development could benefit from improved predictability of mutant properties and a higher abundance of beneficial substitutions in mutant libraries. We have recently explored the use of computational methods for the design of small mutant libraries harboring thermostable enzyme variants.^[4] In the work reported herein, we demonstrate the use

of computational library design for engineering enzyme stereoselectivity.

The computational strategy that we explored is to place a substrate in the enzyme active site in an orientation required for the desired enantioselective conversion and to redesign the active site to stabilize that geometry. Protein sequences forming predefined substrate-binding sites or sites with shape complementarity to transition-state models can be generated by CPD methods such as implemented in RosettaDesign^[5] after which laboratory screening is carried out to identify improved variants.^[6] In the approach that we termed CASCO (Scheme 1), high-throughput molecular dynamics simulations



Scheme 1. The developed CASCO framework for redesign of catalytic selectivity by integration of computational enzyme design and HTMI-MD. The computational design includes the introduction of steric hindrance to prevent unwanted substrate binding modes.

that predict relative reactivity and thereby enantioselectivity replace most of the experimental screening. The strategy was tested with a challenging case of biocatalytic relevance: the design of a pair of enantiocomplementary epoxide hydrolases for the enantioselective transformation of cyclopentene oxide (**1a**) to yield either (*R,R*)- or (*S,S*)-cyclopentane-1,2-diol (**1b**) (Scheme 2). As the stereochemical outcome of this asymmetric conversion is controlled by regioselectivity of water attack, accurate placement of the substrate in the active site is essential. This is highly challenging because epoxide **1a** lacks bulky or polar groups that could position the substrate in a defined way by directional interactions such as hydrogen bonds or electrostatic attraction.

Asymmetric enzymatic synthesis of diol **1b** requires an active site geometry that binds substrate **1a** in a reactive

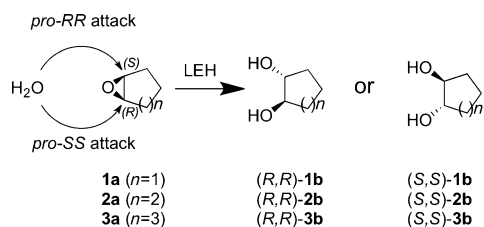
[*] Dr. H. J. Wijma,^[†] R. J. Floor,^[†] Prof. S. J. Marrink, Prof. D. B. Janssen
Groningen Biomolecular Sciences and Biotechnology Institute
University of Groningen
Nijenborgh 4, 9747 AG Groningen (The Netherlands)
E-mail: d.b.janssen@rug.nl

Prof. S. J. Marrink
Zernike Institute for Advanced Materials, University of Groningen
Nijenborgh 7, 9747 AG Groningen (The Netherlands)
Dr. S. Bjelic, Prof. D. Baker
Department of Biochemistry, University of Washington
Seattle, WA 98105 (USA)

[†] These authors contributed equally to this work.

[**] This work was supported by the European Union 7th framework projects Metaexplore (KBBE-2007-3-3-05, 222625) and Kyrobio (KBBE-2011-5, 289646), as well as by BE-Basic and by NWO (Netherlands Organization for Scientific Research) through an ECHO grant.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201411415>.



I116F/W, and F134W/Y could sterically hinder pro-RR attack, while the mutations L35F/W, L74F/W, M78F/W, I80F/W, and L103F/W could prevent pro-SS attack. The forced introduction of these mutations was necessary, as these bulky substitutions only occurred in low frequencies amongst the initial designs (Supporting Information, Figure S1). Algorithms to automatically carry out such tasks are still under development, and will aid future design projects.^[15] After selecting a substitution that enforces steric hindrance, the other ten residues were computationally redesigned in separate design calculations, which resulted in 945 pro-RR and 931 pro-SS designs with improved sequence diversity (Supporting Information, Figures S1, S2, Table S1).

When these second-generation designs with enforced steric hindrance were subjected to HTMI-MD simulations, a much larger fraction of the variants exhibited the desired enantioselectivity (Figure 1 B; Supporting Information, Figures S3, S4). Of the initial designs, 1.6 % had an ee^{pred} of over 90 %, whereas after the introduction of steric hindrance by mutations L103F or L103W, over 20 % of the designs was calculated to have a high ee^{pred} for pro-RR attack (Supporting Information, Table S1). Similarly, substitutions L35F/W, M78W, and I80F/W increased the fraction of the enantiocomplementary designs with an ee^{pred} for pro-SS attack from 10 to 20 % of the designs (Supporting Information, Table S1).

To evaluate the design, we expressed and tested the 10 designs with the highest predicted pro-RR selectivity as well as 27 pro-SS designs. For this a recently engineered thermostable variant of LEH was used^[4b] because thermostability increases the probability that mutations are tolerated without formation of misfolded protein.^[16] Conversion assays with the purified mutant enzymes with analysis of diol produced from cyclopentene oxide **1a** by chiral GC revealed that 28 of the 33 active variants had the enantioselectivity they were designed for (Supporting Information, Table S2). Of these, nine were highly selective, producing the desired diol with an ee of over 75 %. The most enantioselective variants were pro-RR-8 which contains five mutations and produces (*R,R*)-**1b** with an ee of 85 % and pro-SS-16, which has seven mutations and forms (*S,S*)-**1b** with 90 % ee (Figures 2 and 3; Supporting Information, Figure S5). Both these variants have a bulky residue (L103F and L35W, respectively) that was introduced to prevent a substrate binding mode that would allow nucleophilic attack by water at the unwanted position, and the ee value demonstrates that in these variants 92 % and 94 %, respectively, of the nucleophilic attack events occurred on the correct carbon atom (see the Supporting Information, Equation S2). These results show that there is a good

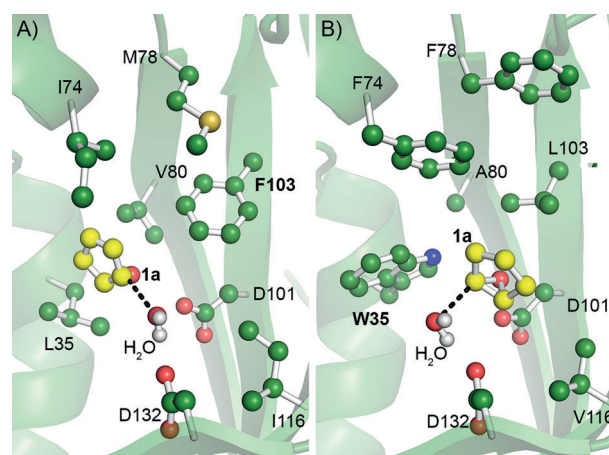


Figure 2. Predicted structures of the active sites of limonene epoxide hydrolase variants pro-RR-8 (A) and pro-SS-16 (B). The docked substrate **1a** is shown in yellow. The water molecule performing the nucleophilic attack is also shown. Residues that were introduced to prevent substrate binding in the undesired pose are labeled in bold.

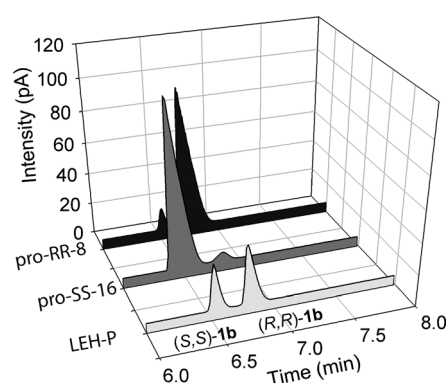


Figure 3. Chiral-phase gas chromatography of diols **1b** produced by LEH variants. The plot displays FID detector signal intensity versus the retention time on a Hydrodex b-TBDAC column (Aurora Borealis, The Netherlands).

correlation between the predicted and the experimentally observed enantioselectivities.

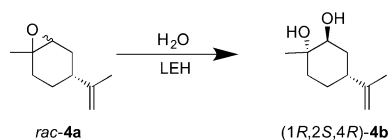
The two most selective variants (pro-RR-8 and pro-SS-16) were characterized in more detail. Both variants were expressed at 50 mg protein per liter of culture broth, which is comparable to the yield of the parent LEH-P.^[4b] The two variants maintained good catalytic activity for the hydrolysis of **1a**, albeit with lower rates than variant LEH-P (Table 1).

Table 1: Computationally predicted and experimentally observed enantioselectivities of the best limonene epoxide hydrolase variants for the hydrolysis of cyclopentene oxide (**1a**).^[a]

| Variant | Computational predictions Active site mutations ^a | Experimental analysis | | | |
|------------|---|-----------------------|----------------|-------------------------------------|--------------------------------|
| | | NAC pro-RR [%] | NAC pro-SS [%] | k_{obs} [s ⁻¹] | Major product ee [%] |
| WT (LEH-P) | none | | | 0.026 ± 0.0006 | (<i>R,R</i>)- 1b 23.9 |
| pro-RR-8 | M32L/L74I/I80V/ L103F /F139L | 5.95 | 0.004 | 0.0048 ± 0.0003 | (<i>R,R</i>)- 1b 85.5 |
| pro-SS-16 | M32L/ L35W /L74F/M78F/I80A/I116V/F139L | 0.69 | 48.1 | 0.0031 ± 0.0003 | (<i>S,S</i>)- 1b 90.2 |

[a] NAC percentages obtained from 10 independently initialized 100 ps MD simulations. The mutations enforcing steric hindrance are in bold. Data on the other improved variants are given in the Supporting Information, Table S2.

Variants pro-RR-8 and pro-SS-16 had apparent melting temperatures of 69.5°C and 66.5°C, respectively, compared to 68.0°C for LEH-P, showing that the thermostability which was introduced earlier was maintained (Supporting Information, Table S3). Variants pro-RR-8 and pro-SS-16 exhibited also enantioselectivity with the larger alicyclic epoxides **2a** and **3a** (Scheme 2). For **2** and **3**, pro-RR-8 was moderately (*R,R*) selective while pro-SS-16 was highly (*S,S*)-selective (Supporting Information, Table S3). The natural substrate limonene epoxide **4a** (Scheme 3) was also converted with high enantioselectivity (Supporting Information, Table S3).



Scheme 3. Hydrolysis of limonene epoxide *rac*-**4a** by LEH.

The enantioselectivity of mutants pro-RR-8 and pro-SS-16 is similar to that of two variants that were discovered by directed evolution;^[3f] mutants H173 and H178 produced (*R,R*)- and (*S,S*)-cyclopentane-1,2-diol with an *ee* of 80% and 93%, respectively. The variants pro-RR-8 and pro-SS-16 each have only one substitution in common with the corresponding directed evolution mutants (L74I in pro-RR-8 and H173, I116V in pro-SS-16 and H178). These particular substitutions do not have a large influence on the shape of the active site (Figure 2). Whereas the outcomes of directed evolution and computational approach are similar in terms of enzyme performance, they differ in the sequences of the enzyme that are produced. More importantly, the approaches differ in experimental screening effort, which encompassed 4700 variants examined by chiral GC for directed evolution^[3f] but only 37 mutants in case of the CASCO-based library described herein. Thus, computational library design and computational screening can replace most of the laboratory screening of a directed evolution project.

In conclusion, we developed a strategy for redesign of catalytic selectivity by computation (Scheme 1). The combined use of computational design and molecular dynamics screening resulted in the rapid discovery of epoxide hydrolases that produced diols with high enantiomeric excess. We anticipate that computational library design will increase reliability and cost efficiency of protein engineering efforts aimed at controlling product profiles of enzymes relevant for applied biocatalysis.

Received: November 26, 2014
Published online: January 30, 2015

Keywords: biocatalysis · computational design · epoxide hydrolase · screening · stereoselectivity

[1] a) S. Bjelic, L. G. Nivón, N. Çelebi-Ölçüm, G. Kiss, C. F. Rosewall, H. M. Lovick, E. L. Ingalls, J. L. Gallaher, J. Seetharaman, S. Lew, G. T. Montelione, J. F. Hunt, F. E. Michael, K. N.

- Houk, D. Baker, *ACS Chem. Biol.* **2013**, *8*, 749; b) R. Blomberg, H. Kries, D. M. Pinkas, P. R. E. Mittl, M. G. Grutter, H. K. Privett, S. L. Mayo, D. Hilvert, *Nature* **2013**, *503*, 418; c) L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Röthlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas, D. Hilvert, K. N. Houk, B. L. Stoddard, D. Baker, *Science* **2008**, *319*, 1387; d) J. B. Siegel, A. Zanghellini, H. M. Lovick, G. Kiss, A. R. Lambert, J. L. S. Clair, J. L. Gallaher, D. Hilvert, M. H. Gelb, B. L. Stoddard, K. N. Houk, F. E. Michael, D. Baker, *Science* **2010**, *329*, 309; e) D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik, D. Baker, *Nature* **2008**, *453*, 190.
- [2] a) S. R. Gordon, E. J. Stanley, S. Wolf, A. Toland, S. J. Wu, D. Hadidi, J. H. Mills, D. Baker, I. S. Pultz, J. B. Siegel, *J. Am. Chem. Soc.* **2012**, *134*, 20513; b) M. J. Grisewood, N. P. Gifford, R. J. Pantazes, Y. Li, P. C. Cirino, M. J. Janik, C. D. Maranas, *PLoS One* **2013**, *8*, e75358; c) S. D. Khare, Y. Kipnis, P. Greisen, Jr., R. Takeuchi, Y. Ashani, M. Goldsmith, Y. Song, J. L. Gallaher, I. Silman, H. Leader, J. L. Sussman, B. L. Stoddard, D. S. Tawfik, D. Baker, *Nat. Chem. Biol.* **2012**, *8*, 294; d) L. Liu, P. Murphy, D. Baker, S. Lutz, *Chem. Commun.* **2010**, *46*, 8803; e) D. P. Nannemann, K. W. Kaufmann, J. Meiler, B. O. Bachmann, *Protein Eng. Des. Sel.* **2010**, *23*, 607; f) X.-D. Kong, S. Yuan, L. Li, S. Chen, J.-H. Xu, J. Zhou, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15717; g) X.-D. Kong, Q. Ma, J. Zhou, B.-B. Zeng, J.-H. Xu, *Angew. Chem. Int. Ed.* **2014**, *53*, 6641; *Angew. Chem.* **2014**, *126*, 6759.
- [3] a) M. Alexeeva, A. Enright, M. J. Dawson, M. Mahmoudian, N. J. Turner, *Angew. Chem. Int. Ed.* **2002**, *41*, 3177; *Angew. Chem.* **2002**, *114*, 3309; b) U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185; c) M. T. Reetz, *Angew. Chem. Int. Ed.* **2013**, *52*, 2658; *Angew. Chem.* **2013**, *125*, 2720; d) N. J. Turner, *Nat. Chem. Biol.* **2009**, *5*, 567; e) G. J. Williams, T. Woodhall, L. M. Farnsworth, A. Nelson, A. Berry, *J. Am. Chem. Soc.* **2006**, *128*, 16238; f) H. Zheng, M. T. Reetz, *J. Am. Chem. Soc.* **2010**, *132*, 15744.
- [4] a) R. J. Floor, H. J. Wijma, D. I. Colpa, A. Ramos-Silva, P. A. Jekel, W. Szymański, B. L. Feringa, S. J. Marrink, D. B. Janssen, *ChemBioChem* **2014**, *15*, 1660; b) H. J. Wijma, R. J. Floor, P. A. Jekel, D. Baker, S. J. Marrink, D. B. Janssen, *Protein Eng. Des. Sel.* **2014**, *27*, 49.
- [5] a) D. Hilvert, *Annu. Rev. Biochem.* **2013**, *82*, 447; b) G. Kiss, N. Çelebi-Ölçüm, R. Moretti, D. Baker, K. N. Houk, *Angew. Chem. Int. Ed.* **2013**, *52*, 5700; *Angew. Chem.* **2013**, *125*, 5810; c) W. S. Mak, J. B. Siegel, *Curr. Opin. Struct. Biol.* **2014**, *27*, 87; d) F. Richter, A. Leaver-Fay, S. D. Khare, S. Bjelic, D. Baker, *PLoS One* **2011**, *6*, e19230; e) H. J. Wijma, D. B. Janssen, *FEBS J.* **2013**, *280*, 2948.
- [6] a) M. M. Y. Chen, C. D. Snow, C. L. Vizcarra, S. L. Mayo, F. H. Arnold, *Protein Eng. Des. Sel.* **2012**, *25*, 171; b) S. M. Lippow, T. S. Moon, S. Basu, S.-H. Yoon, X. Li, B. A. Chapman, K. Robison, D. Lipovšek, K. L. J. Prather, *Chem. Biol.* **2010**, *17*, 1306.
- [7] Pro-RR attack or hydrolysis is defined as the reaction (or binding mode) that yields (*R,R*)-diol; pro-SS attack gives the (*S,S*)-diol.
- [8] a) M. Arand, B. M. Hallberg, J. Y. Zou, T. Bergfors, F. Oesch, M. J. Van der Werf, J. A. M. de Bont, T. A. Jones, S. L. Mowbray, *EMBO J.* **2003**, *22*, 2583; b) K. H. Hopmann, B. M. Hallberg, F. Himo, *J. Am. Chem. Soc.* **2005**, *127*, 14339; c) M. E. S. Lind, F. Himo, *Angew. Chem. Int. Ed.* **2013**, *52*, 4563; *Angew. Chem.* **2013**, *125*, 4661; d) M. J. Van der Werf, R. V. A. Orru, K. M. Overkamp, H. J. Swarts, I. Osprian, A. Steinreiber, J. A. M. de Bont, K. Faber, *Appl. Microbiol. Biotechnol.* **1999**, *52*, 380.
- [9] a) A. Leaver-Fay, M. J. O'Meara, M. Tyka, R. Jacak, Y. Song, E. H. Kellogg, J. Thompson, I. W. Davis, R. A. Pache, S. Lyskov, J. J. Gray, T. Kortemme, J. S. Richardson, J. J. Havranek, J.

- Snoeyink, D. Baker, B. Kuhlman, *Methods Enzymol.* **2013**, 523, 109; b) C. A. Voigt, D. B. Gordon, S. L. Mayo, *J. Mol. Biol.* **2000**, 299, 789.
- [10] a) O. Khersonsky, D. Röthlisberger, A. M. Wollacott, P. Murphy, O. Dym, S. Albeck, G. Kiss, K. N. Houk, D. Baker, D. S. Tawfik, *J. Mol. Biol.* **2011**, 407, 391; b) H. K. Privett, G. Kiss, T. M. Lee, R. Blomberg, R. A. Chica, L. M. Thomas, D. Hilvert, K. N. Houk, S. L. Mayo, *Proc. Natl. Acad. Sci. USA* **2012**, 109, 3790; c) J. Z. Ruscio, J. E. Kohn, K. A. Ball, T. Head-Gordon, *J. Am. Chem. Soc.* **2009**, 131, 14111; d) L. Wang, E. A. Althoff, J. Bolduc, L. Jiang, J. Moody, J. K. Lassila, L. Giger, D. Hilvert, B. Stoddard, D. Baker, *J. Mol. Biol.* **2012**, 415, 615.
- [11] H. J. Wijma, S. J. Marrink, D. B. Janssen, *J. Chem. Inf. Model.* **2014**, 54, 2079.
- [12] S. Hur, T. C. Bruice, *Proc. Natl. Acad. Sci. USA* **2003**, 100, 12015.
- [13] A. Westerbeeck, W. Szymański, H. J. Wijma, S. J. Marrink, B. L. Feringa, D. B. Janssen, *Adv. Synth. Catal.* **2011**, 353, 931.
- [14] B. Schiøtt, T. C. Bruice, *J. Am. Chem. Soc.* **2002**, 124, 14558.
- [15] A. Leaver-Fay, R. Jacak, P. B. Stranges, B. Kuhlman, *PLoS One* **2011**, 6, e20937.
- [16] N. Tokuriki, F. Stricher, L. Serrano, D. S. Tawfik, *PLoS Comput. Biol.* **2008**, 4, e1000002.