



Cite this: *Chem. Commun.*, 2016, 52, 5053

Received 29th February 2016,
Accepted 11th March 2016

DOI: 10.1039/c6cc01823h

www.rsc.org/chemcomm

Enzyme classification using complex dynamic hemithioacetal systems†

Yan Zhang,^a H. Surangi N. Jayawardena,^b Mingdi Yan^{*ab} and Olof Ramström^{*a}

A complex dynamic hemithioacetal system was generated for the evaluation of lipase reactivities in organic media. In combination with pattern recognition methodology, twelve different lipases were successfully classified into four distinct groups following their reaction selectivities and reactivities. A probe lipase was further categorized using the training matrix with predicted reactivity.

Enzyme catalysis has over the last decades been increasingly explored for organic synthesis due to favorable properties such as environmental benignity, high efficiency and selectivity.^{1–9} A multitude of different enzymes have been applied, among which lipases have drawn special attention, in part due to their performances in organic solvents coupled to high stereoselectivities.^{10–13} Applications of lipases have also been recently expanded owing to new findings on enzyme modification and promiscuity, where improved or additional activities have been illustrated.^{14–17} However, lipases from individual sources or families may not behave the same in regard to substrates or reaction types, and in order to distinguish the specificities among different lipases, an efficient method for lipase screening and grouping would be useful, accelerating further applications in synthesis.

Classification of enzyme performances in principle requires collection and analysis of sizable data sets, involving interpretation of arrays of reaction results. This process can however be facilitated using two methodologies: dynamic chemistry and pattern recognition. Pattern recognition has been demonstrated to be a powerful technique with broad applications in classification and identification, for example in retrosynthetic analysis and protein differentiation.^{18–20} Methods such as principal component analysis (PCA), linear discriminant analysis (LDA), and

partial least square (PLS) analysis have thus been developed for the extraction of relevant data features, generating classification patterns which can be used to predict unknown species.^{21–23} In conjunction to these techniques, dynamic systemic resolution (DSR), a concept based on constitutional dynamic chemistry (CDC), can be efficiently applied to generating and probing complex systems.^{24–28} Typical CDC systems involve reversible (covalent) reactions, making all components mutually interchange under thermodynamic control. Applied selection pressures may then exert rearrangements of the systems, causing amplification of optimal system species.^{29–51}

In DSR systems, kinetic steps are coupled to the reversible reactions, resulting in irreversible resolution processes.^{43,52–61} For example, kinetic enzymatic transformations have been used as external selection pathways, enabling rapid substrate screening. Therefore, the combination of DSR and pattern recognition would result in a powerful platform for enzyme screening and classification in reference to the enzyme reactivities and substrate selectivities. Herein, this dynamic diversity has been demonstrated and further applied to distinguish lipase selectivities. Dynamic hemithioacetal systems were thus generated and further subjected to biocatalyzed transesterification, and the resulting information subsequently analyzed by pattern recognition methodology resulting in discrete classification of lipase reactivity (Fig. 1).

Hemithioacetal chemistry shows attractive reversibility features, and can be applied to generate dynamic systems in both aqueous and organic media.^{43,57,59} We have for example

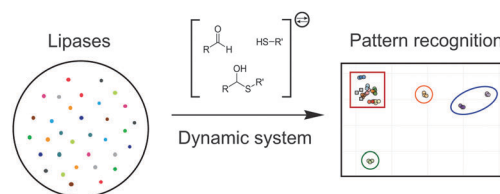


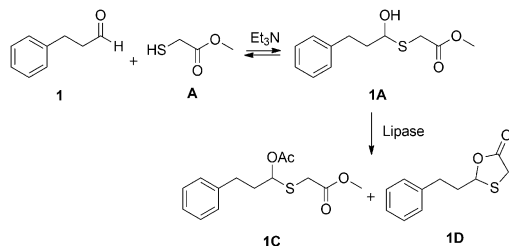
Fig. 1 Lipase classification through dynamic hemithioacetal systems in combination with pattern recognition methodology.

^a Department of Chemistry, KTH – Royal Institute of Technology, Teknikringen 30, 10044 Stockholm, Sweden. E-mail: ramstrom@kth.se

^b Department of Chemistry, University of Massachusetts Lowell, 1 University Ave., Lowell, MA 01854, USA. E-mail: Mingdi_Yan@uml.edu

† Electronic supplementary information (ESI) available: NMR spectra, linear discriminant analysis. See DOI: 10.1039/c6cc01823h





Scheme 1 Lipase-catalyzed acetylation/lactonization of dynamic hemithioacetal system.

previously demonstrated the selective lipase-catalyzed cyclization to 1,3-oxathiolan-5-ones using sulfanylacetates.⁵⁷ When additional acyl donors are added into these systems, linear (non-cyclized) acylated products can potentially be produced in addition to cyclized oxathiolanones. This leads to increased diversity of the system, thereby enabling additional information of the lipase catalysis process. In order to evaluate the potential of forming both intermolecularly acetylated and intramolecularly lactonized products, test reactions between 3-phenylpropanal and methyl 2-sulfanylacetate in the presence of phenyl acetate were first performed under the catalytic action of two different lipases: PSIm from *Burkholderia* (*Pseudomonas*) *cepacia* and CALB from *Candida antarctica* (Scheme 1). Interestingly, with triethylamine (Et₃N) as base,⁵⁷ both types of products were detected in the ¹H-NMR spectra of the reaction mixtures at room temperature, but the ratio between these two products varied according to the lipases used. While PSIm gave predominantly the intermolecular, 'linear' acetylated product (1C:1D = 6.4:1), CALB mostly yielded the cyclized intramolecular product (1C:1D = 1:2.7). These results thereby confirmed the potential of using these dynamic systems for lipase reactivity profiling.

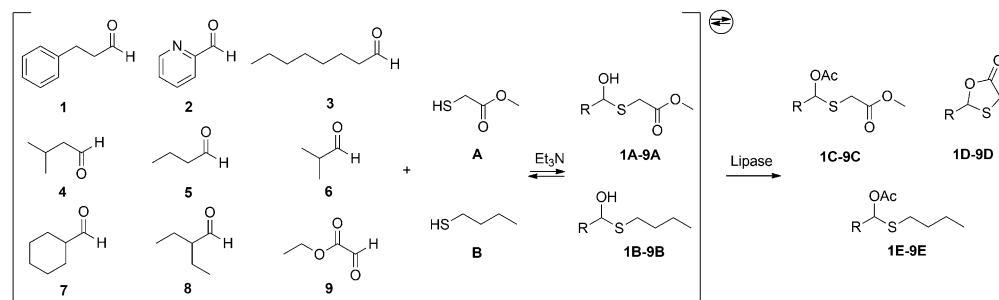
A dynamic hemithioacetal system with enhanced complexity was subsequently established. Nine aldehydes (1–9) of different structure were incorporated, including both aromatic and aliphatic compounds, structures containing short and longer chains, and linear and cyclized functional groups (Scheme 2). Besides methyl 2-sulfanylacetate, butanethiol was also included in the system, only able to form linear acetylated products. Under the same conditions as in the test reactions, twelve commonly used and commercially available lipases were chosen

to challenge the system as external selectors, irreversibly transforming the intermediates 1A–9A and 1B–9B to the respective preferred products. Thus, lipases from *Aspergillus niger* (ANL), *Candida antarctica* (CALA and CALB), *Candida rugosa* (CRL), *Mucor javanicus* (MJL), *Mucor miehei* (MML), *Pseudomonas fluorescens* (PFL), porcine pancreas (PPL), *Burkholderia cepacia* (PSCI and PSIm), *Rhizopus arrhizus* (RAL), and *Rhizopus niveus* (RNL), were evaluated. In principle, under these conditions, 36 different intermediates of two types and 54 possible products of three different types can be generated, forming a dynamic system with the complexity of 101 different compounds.

¹H-NMR was used to monitor the reaction progress in the system, which showed the typical intermediate and product peaks around 4–7 ppm. As expected, this complex dynamic system produced largely overlapping signals, difficult to resolve by conventional means. Pattern recognition methodology was then adopted to analyze the results. Among the data classification techniques, LDA is excellent at processing two or more classes of objects, and is in addition simple and robust.⁶² The principle of LDA analyses is based on a training matrix created by a collection of independent (explanatory) and dependent variables; where the dependent variables (lipases) are classified according to the values generated by the explanatory variables (NMR signal intensities). LDA treats the array response as the linear combination of discriminant functions or canonical scores.^{63,64}

The relative signal integrations using Et₃N as the internal standard were first extracted from the ¹H-NMR spectra, with every 0.1 ppm in the range of 4.6–6.7 ppm, which were further employed as 21 explanatory variables in the LDA training matrix (Table S1, ESI†). The dependent variables to generate the training matrix were the twelve lipases with their repeated triplicate experimental data. For the best separation achieved in Fig. 2, the lipases were classified into groups on the factor plot where the canonical factors F1 and F2 contained variation of 73.93% and 21.49%, respectively.

From the pattern generated from the training matrix, the twelve different lipases could be classified into several distinct areas in the analysis window: three discrete areas with CALB, PFL, and PSCI/PSIm, and an extended region grouping all remaining lipases (Fig. 2). Validation tests revealed some cross-correlation between PSCI and PSIm, as could be expected, and strong correlation between the lipases of the extended



Scheme 2 Dynamic hemithioacetal system with enhanced complexity used for lipase evaluation.



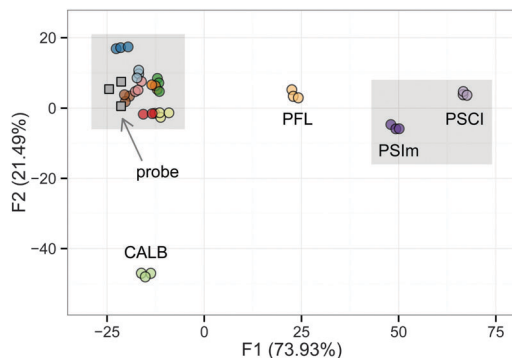


Fig. 2 Factor plot obtained from the LDA analysis of lipases tested with the dynamic hemithioacetal system; (●) ANL, (●) CALA, (●) CALB, (●) CRL, (●) MJL, (●) MML, (●) PFL, (●) PPL, (●) PSCI, (●) PSIm, (●) RAL, (●) RNL, (■) probe.

group. The overall cross-validation score was thus relatively low (0.39, Table S7, ESI[†]). However, upon grouping of the lipases in the extended region, ANL, CALA, CRL, MJL, MML, PPL, RAL, RNL, the validation result was considerably higher (0.97, Table S7, ESI[†]).

Furthermore, the correlation between the pattern and the enzyme activities was also studied. CALB, CRL, PFL and PSIm, were selected as representative enzymes of the different groups and evaluated in reaction tests with 3-phenylpropanal. These studies revealed clear differences between the respective enzymes. CALB, PFL and PSIm were comparatively more reactive, whereas the distributions of the formed products were opposite: CALB showed preference for the cyclized product (1D), while PSIm was selective for the linear species (1C). For PFL, the resulting ratio between the linear acetylated- and cyclized product was 29:1, displaying strong resemblance to the PSIm/PSCI group. On the other hand, CRL yielded only trace amount of cyclized product, which demonstrated the inactive behavior of this group of lipases toward intermolecular acetylations under these reaction conditions.

Information from the grouping pattern of all the lipases shown above was then applied to test a probe lipase. Thus, a lipase from *Penicillium camembertii* (PCL) was introduced as the testing lipase into the complex system under the same conditions as for all other lipases. As illustrated in Fig. 2, this probe lipase was positioned in the region of the extended group.

In order to examine the correlation between the found pattern and the activity of PCL, the test reaction between 3-phenylpropanal and methyl 2-sulfanylacetate with phenyl acetate as acetyl donor was challenged also by PCL catalysis. The resulting pattern demonstrated the similarity to the extended lipase region.

In conclusion, complex dynamic hemithioacetal systems were generated and coupled to enzymatic transformation, providing two types of potential products, intermolecular linear acetylated compounds and intramolecular cyclized structures. Pattern recognition methodology was further employed to evaluate the selectivities and reactivities of lipases based on the signal integrations in the ¹H-NMR spectra of the reaction mixture.

Thus, twelve different lipases were classified into four groups in the analysis window, which showed good correlation to their selectivities between the two types of products. Lipase PCL as a probe lipase from the training set was tested with the same dynamic system, giving major recognition into an extended lipase group, which was consistent with the enzyme reactivity.

This work was in part supported by the Swedish Research Council, the National Institutes of Health (R01GM080295 and 2R15GM066279 to MY), and the Royal Institute of Technology. YZ thanks the China Scholarship Council for a special scholarship award.

Notes and references

- 1 A. Cuetos, M. García-Ramos, E.-M. Fischereder, A. Díaz-Rodríguez, G. Grogan, V. Gotor, W. Kroutil and I. Lavandera, *Angew. Chem., Int. Ed.*, 2016, **55**, 3144–3147.
- 2 Y.-L. Du, R. Singh, L. M. Alkhalaf, E. Kuatsjah, H.-Y. He, L. D. Eltis and K. S. Ryan, *Nat. Chem. Biol.*, 2016, **12**, 194–199.
- 3 J. L. Porter, R. A. Rusli and D. L. Ollis, *ChemBioChem*, 2016, **17**, 197–203.
- 4 J.-b. Wang and M. T. Reetz, *Nat. Chem.*, 2015, **7**, 948–949.
- 5 B. M. Nestl, S. C. Hammer, B. A. Nebel and B. Hauer, *Angew. Chem., Int. Ed.*, 2014, **53**, 3070–3095.
- 6 N. J. Turner and E. O'Reilly, *Nat. Chem. Biol.*, 2013, **9**, 285–288.
- 7 A. Lumbroso, M. L. Cooke and B. Breit, *Angew. Chem., Int. Ed.*, 2013, **52**, 1890–1932.
- 8 L. Hu, F. Schaufelberger, Y. Zhang and O. Ramström, *Chem. Commun.*, 2013, **49**, 10376–10378.
- 9 *Enzyme Catalysis in Organic Synthesis*, ed. K. Drauz, H. Gröger and O. May, John Wiley & Sons, 2012.
- 10 E. A. Manoel, J. M. Robert, M. C. C. Pinto, A. C. O. Machado, M. D. Besteti, M. A. Z. Coelho, A. B. C. Simas, R. Fernandez-Lafuente, J. C. Pinto and D. M. G. Freire, *RSC Adv.*, 2016, **6**, 4043–4052.
- 11 Y. Kitamoto, Y. Kuruma, K. Suzuki and T. Hattori, *J. Org. Chem.*, 2015, **80**, 521–527.
- 12 A. S. de Miranda, L. S. Miranda and R. O. de Souza, *Biotechnol. Adv.*, 2015, **33**, 372–393.
- 13 P. Adlercreutz, *Chem. Soc. Rev.*, 2013, **42**, 6406–6436.
- 14 E. Stavila, G. O. R. Alberda van Ekenstein, A. J. J. Woortman and K. Loos, *Biomacromolecules*, 2014, **15**, 234–241.
- 15 M. Sakulsombat, P. Vongvilai and O. Ramström, *Chem. – Eur. J.*, 2014, **20**, 11322–11325.
- 16 M. S. Humble and P. Berglund, *Eur. J. Org. Chem.*, 2011, 3391–3401.
- 17 P. Vongvilai, M. Linder, M. Sakulsombat, M. Svedendahl Humble, P. Berglund, T. Brinck and O. Ramström, *Angew. Chem., Int. Ed.*, 2011, **50**, 6592–6595.
- 18 X. Bi, D. Li and Z. Liu, *Anal. Chem.*, 2015, **87**, 4442–4447.
- 19 H. S. N. Jayawardena, X. Wang and M. Yan, *Anal. Chem.*, 2013, **85**, 10277–10281.
- 20 S. Nieto, C. Cativiela and E. P. Urriolabeitia, *New J. Chem.*, 2012, **36**, 566–569.
- 21 S. G. Elci, D. F. Moyano, S. Rana, G. Y. Tonga, R. L. Phillips, U. H. F. Bunz and V. M. Rotello, *Chem. Sci.*, 2013, **4**, 2076–2080.
- 22 S. Stewart, M. A. Ivy and E. V. Anslyn, *Chem. Soc. Rev.*, 2014, **43**, 70–84.
- 23 A. Buryak and K. Severin, *J. Am. Chem. Soc.*, 2005, **127**, 3700–3701.
- 24 J.-M. Lehn, *Angew. Chem., Int. Ed.*, 2015, **54**, 3276–3289.
- 25 Y. Zhang, L. Hu and O. Ramström, *Supramolecular Systems in Biomedical Fields*, The Royal Society of Chemistry, 2013, pp. 397–418.
- 26 *Constitutional Dynamic Chemistry*, ed. M. Barboiu, Springer Verlag, Berlin, Heidelberg, 2012.
- 27 *Dynamic Combinatorial Chemistry: In Drug Discovery, Bioorganic Chemistry, and Materials Science*, ed. B. L. Miller, John Wiley & Sons, Hoboken, NJ, 2010.
- 28 *Dynamic Combinatorial Chemistry*, ed. J. N. H. Reek and S. Otto, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010.
- 29 Y. Zhang and M. Barboiu, *Chem. Rev.*, 2016, **116**, 809–834.
- 30 T. Kosikova, H. Mackenzie and D. Philp, *Chem. – Eur. J.*, 2016, **22**, 1831–1839.



- 31 A. Takahashi, T. Ohishi, R. Goseki and H. Otsuka, *Polymer*, 2016, **82**, 319–326.
- 32 G. A. Barcan, X. Zhang and R. M. Waymouth, *J. Am. Chem. Soc.*, 2015, **137**, 5650–5653.
- 33 Y. Zhang and M. Barboiu, *Chem. Commun.*, 2015, **51**, 15925–15927.
- 34 F. Schaufelberger and O. Ramström, *Chem. – Eur. J.*, 2015, **21**, 12735–12740.
- 35 F. Schaufelberger, L. Hu and O. Ramström, *Chem. – Eur. J.*, 2015, **21**, 9776–9783.
- 36 S. A. Boer and D. R. Turner, *Chem. Commun.*, 2015, **51**, 17375–17378.
- 37 G. Li, W. Zheng, Z. Chen, Y. Zhou, Y. Liu, J. Yang, Y. Huang and X. Li, *Chem. Sci.*, 2015, **6**, 7097–7104.
- 38 Y. Zhou, Y. Yuan, L. You and E. V. Anslyn, *Chem. – Eur. J.*, 2015, **21**, 8207–8213.
- 39 M. Mondal and A. K. Hirsch, *Chem. Soc. Rev.*, 2015, **44**, 2455–2488.
- 40 R.-C. Brachvogel and M. von Delius, *Chem. Sci.*, 2015, **6**, 1399–1403.
- 41 D. B. Rasale, S. Biswas, M. Konda and A. K. Das, *RSC Adv.*, 2015, **5**, 1529–1537.
- 42 C. W. Hsu and O. S. Miljanic, *Angew. Chem., Int. Ed.*, 2015, **54**, 2219–2222.
- 43 Y. Zhang and O. Ramström, *Chem. – Eur. J.*, 2014, **20**, 3288–3291.
- 44 E. Burda and J. Rademann, *Nat. Commun.*, 2014, **5**, 5170.
- 45 A. Wilson, G. Gasparini and S. Matile, *Chem. Soc. Rev.*, 2014, **43**, 1948–1962.
- 46 A. Herrmann, *Chem. Soc. Rev.*, 2014, **43**, 1899–1933.
- 47 M. C. Misuraca, E. Moulin, Y. Ruff and N. Giuseppone, *New J. Chem.*, 2014, **38**, 3336–3349.
- 48 S. Ulrich and P. Dumy, *Chem. Commun.*, 2014, **50**, 5810–5825.
- 49 A. G. Santana, E. Jimenez-Moreno, A. M. Gomez, F. Corzana, C. Gonzalez, G. Jimenez-Oses, J. Jimenez-Barbero and J. L. Asensio, *J. Am. Chem. Soc.*, 2013, **135**, 3347–3350.
- 50 M.-K. Chung, K. Severin, S. J. Lee, M. L. Waters and M. R. Gagné, *Chem. Sci.*, 2011, **2**, 744.
- 51 M. Capela, N. J. Mosey, L. Xing, R. Wang and A. Petitjean, *Chemistry*, 2011, **17**, 4598–4612.
- 52 L. Hu, Y. Zhang and O. Ramström, *Sci. Rep.*, 2015, **5**, 11065.
- 53 Y. Zhang, P. Vongvilai, M. Sakulsombat, A. Fischer and O. Ramström, *Adv. Synth. Catal.*, 2014, **356**, 987–992.
- 54 L. Hu, Y. Zhang and O. Ramström, *Org. Biomol. Chem.*, 2014, **12**, 3572–3575.
- 55 L. Hu and O. Ramstrom, *Chem. Commun.*, 2014, **50**, 3792–3794.
- 56 Y. Zhang, L. Hu and O. Ramström, *Chem. Commun.*, 2013, **49**, 1805–1807.
- 57 M. Sakulsombat, Y. Zhang and O. Ramström, *Chem. – Eur. J.*, 2012, **18**, 6129–6132.
- 58 Y. Zhang, M. Angelin, R. Larsson, A. Albers, A. Simons and O. Ramström, *Chem. Commun.*, 2010, **46**, 8457–8459.
- 59 R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero and O. Ramström, *Angew. Chem., Int. Ed.*, 2010, **49**, 589–593.
- 60 P. Vongvilai and O. Ramström, *J. Am. Chem. Soc.*, 2009, **131**, 14419–14425.
- 61 P. Vongvilai, M. Angelin, R. Larsson and O. Ramström, *Angew. Chem., Int. Ed.*, 2007, **46**, 948–950.
- 62 P. C. Jurs, G. A. Bakken and H. E. McClelland, *Chem. Rev.*, 2000, **100**, 2649–2678.
- 63 J. W. McFarland and D. J. Gans, *J. Med. Chem.*, 1987, **30**, 46–49.
- 64 D. Coomans, D. L. Massart and L. Kaufman, *Anal. Chim. Acta*, 1979, **112**, 97–122.

