Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Chemical Biology

Design of computational retrobiosynthesis tools for the design of *de novo* synthetic pathways

Noushin Hadadi[1,2] and Vassily Hatzimanikatis[1,2]

Designing putative metabolic pathways is of great interest in synthetic biology. Retrobiosynthesis is a discipline that involves the design, evaluation, and optimization of *de novo* biosynthetic pathways for the production of high-value compounds and drugs from renewable resources and natural or engineered enzymes. The best candidate pathways are then engineered within a metabolic network of microorganisms that serve as synthetic platforms for synthetic biology. The complexity of biological chemistry and metabolism requires computational approaches to explore the full possibilities of engineering synthetic pathways towards target compounds. Herein, we discuss recent developments in the design of computational tools for retrosynthetic biochemistry and outline the workflow and design elements for such tools.

Addresses
[1] Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland
[2] Swiss Institute of Bioinformatics (SIB), CH-1015, Switzerland

Corresponding author: Hatzimanikatis, Vassily (vassily.hatzimanikatis@epfl.ch)

Current Opinion in Chemical Biology 2015, **28**:99–104

This review comes from a themed issue on **Synthetic biology**

Edited by **Michelle CY Chang** and **Huimin Zhao**

For a complete overview see the Issue and the Editorial

Available online 11th July 2015

http://dx.doi.org/10.1016/j.cbpa.2015.06.025

Introduction

One of the principal aims of synthetic biology and metabolic engineering is to design and build microbial factories for the sustainable fabrication of high-value compounds and industrial chemicals [1–4]. To create efficient microbial factories and broaden the range of biosynthetic pathways for the production of both natural and non-natural compounds, it is necessary to go beyond natural pathways by exploring the chemistry and synthetic capabilities of biological systems [1,5]. The *de novo* design of pathways is crucial for exploiting the incredible natural diversity of enzymatic transformations.

Retrobiosynthesis, a promising approach for *de novo* pathway design, is inspired by the retro-evolution hypothesis that was first proposed in 1945 by Norman Horowitz [6,7] and has its origins in retrosynthetic organic chemistry. Retrosynthetic design starts by defining a target molecule of interest to produce and then 'walks' backwards through the known chemical transformation rules to modify the target molecule and identify potential precursors and reactions [8,9].

This basic concept of walking backwards from a molecule and using the biotransformation rules to reconstruct biochemical pathways is also used: (i) to find novel pathways for the biodegradation of pollutants [10,11] to generate hypothetical pathways for metabolites and lipids that are found in metabolomics and lipidomics studies, but have an unknown metabolism [12•]. Although retrosynthesis and retrobiosynthesis are molecular *design* methods, the term retrobiosynthesis was also initially used to describe the *analysis* of experimental ^{13}C labeling data for identifying biosynthetic routes [13].

In retrobiosynthesis, the aim is to produce a target molecule through enzymatic biotransformation steps that occur in a metabolic pathway of microorganisms. This analysis results in *de novo* pathways that connect the target molecule to either a cellular metabolite or a biochemical feedstock using natural or engineered enzymes.

Before a *de novo* pathway can be built in the laboratory and integrated in a microorganism, it should first be designed and evaluated. Although intuition and manual design can assist in postulating novel pathways, these are not sufficient to guarantee the generation of *all* potentials and to select the most efficient ones [1,2,14–19]. Hence, computational prediction tools are indispensible for retrobiosynthesis analysis, not only for assisting with generating novel hypotheses but also for screening for the most efficient pathways. Computational frameworks result in the extensive generation of all possible *de novo* biosynthetic pathways to allow for the exploration of the entire realm of feasible biotransformations in a given cell [11,18–23,24••,25•,26].

The combinatorial explosion is the most important risk associated with computational approaches, as these methods generate compounds and reactions that may or may not actually occur in nature. Therefore, the next crucial step is to screen the generated biosynthetic pathways through feasibility studies. Various techniques can be used to prune the *de novo* generated pathways and select the most promising ones. In the next sections, we discuss

the general workflow and essential design elements to be integrated into the development of a sound framework for retrobiosynthetic studies, and we compare different available tools based on the consideration of these elements.

From our experience in developing the retrobiosynthesis framework BNICE.ch and the analysis of other available tools, we propose a retrobiosynthetic workflow that includes three main steps, and each step requires the implementation of certain technical design elements (Figure 1).

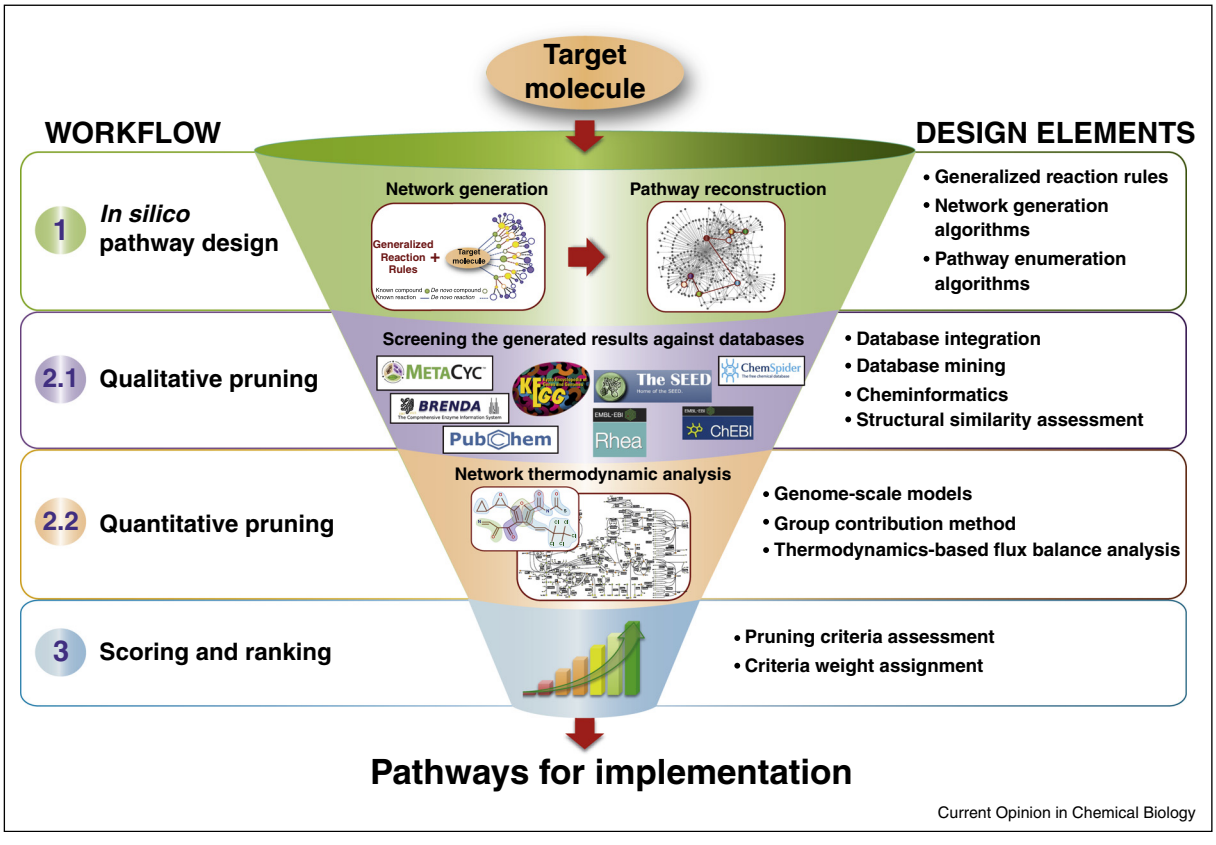
In silico pathway design

The most common *in silico* pathway prediction tools offer the enumeration of pathways in two ways: either they effectively combine known reactions from databases that lead to the production of a desired compound from different organisms (heterologous pathways) [23,27–29] or they construct *de novo* pathways that include not only known reactions but also hypothetical steps whose corresponding enzymes might not actually exist in nature [11,18,19,22,24••,25•,30]. A comprehensive algorithm for the *in silico* prediction of *de novo* pathways is a significant driver for the success of retrobiosynthetic analyses, and a variety of such tools have been developed in the past decade (Table 1).

One of the key design elements of the BNICE.ch tool is a database of 'biochemical transformation rules' that mimic the functions of enzymes and serve as *in silico* enzymatic reactions. As there are a large number of characterized enzymes, one can organize those that employ similar reaction mechanisms into '*generalized enzymatic reaction rules*' [19,20]. The concept of generalized reaction rules has been adopted by several other similar methods [18,22,24••,25•,26]. When acting on a molecule, the generalized reaction rules recognize the biologically reactive sites of a molecule and apply the biotransformation, whereby atoms and bonds rearrange to form a product. Therefore, a generalized rule is capable of acting upon a wide range of substrates in addition to specific native substrates. This leads to the identification of candidate sequences for designing enzymes with broad or altered substrate specificities.

Repeating this process iteratively using a 'network generation algorithm' results in the generation of a biochemical network of all theoretically possible compounds and reactions, including those that have no known experimental

Figure 1



General workflow and design elements for a computational retrobiosynthesis tool.

Table 1

Available retrobiosynthesis tools and their characteristics

Tools	Generalized reaction rules	*De novo* reaction	Gibbs free energy of formation and reaction	Network thermodynamics	Protein sequence identification based on enzyme promiscuity and reaction similarity	Enzyme docking	Host organism specificity	Pathways scoring and ranking	Tool development and applications
BNICE.ch	✓	✓	✓	✓	✓	✓		✓	[12•,19–21,31,32••]
DESHARKY	✓				✓		✓	✓	[23,33]
ReBiT http://www.retro-biosynthesis.com	✓	✓	✓					✓	[26]
Method developed with Cho *et al.*	✓	✓	✓					✓	[18]
RetroPath http://www.issb.genopole.fr/~faulon/retropath.php	✓	✓	✓		✓		✓	✓	[14,22,34,35•,36,37]
SimPheny	✓	✓	✓		✓		✓	✓	[24••]
GEM-Path	✓	✓	✓		✓		✓	✓	[25•]

counterpart (*de novo* compounds and reactions). The next required design element is a 'pathway reconstruction algorithm' that constructs all possible pathways from a given substrate to a target molecule. These algorithms perform either a graph-based search in the network or use optimization-based methods to identify possible pathways from potential substrates for the synthesis of a target compound in the generated metabolic network [31,38,39].

Pruning the generated data

A retrobiosynthetic analysis risks a combinatorial explosion in two ways. First, in the network generation process, the actions of generalized reaction rules on the target compound results in the generation of all possible compounds and reactions, which may or may not actually occur in nature and exponentially increase in every iteration of the network generation algorithm. Second, because of the combinatorial nature of the pathway enumeration step, an enormous number of pathways from a substrate to the same target compound are generated. Thus, the very important next step is the evaluation of the proposed compounds, reactions, and pathways and the selection of the most feasible enzymes, reactions, and pathways to be tested in the laboratory. Pruning analysis is performed using two strategies:

(1) Qualitative pruning of generated data;
(2) Quantitative pruning of generated pathways.

Qualitative pruning of generated results

Qualitative pruning of the generated pathways is the process of surveying which fraction of the obtained information is already *known* or *novel* and asking how similar is the novel information compared with the known data, i.e., the metabolites, reactions and pathways in the databases. These databases are biological, such as KEGG [40] and Metacyc [41], and chemical such as PubChem [42] and ChEBI [43]. Qualitative pruning in general is independent of the organism of choice and is done by comparing the metabolites and reactions in the synthetic pathways with the entries in existing databases. By screening through existing databases, not only can we differentiate between known and novel knowledge, but we can also directly capture available biochemical properties for the compounds and reactions. One such property, as implemented in RetroPath for the qualitative pruning of *de novo* pathways, is the toxicity of known reactants and products of reactions [14,36].

Qualitative pruning in the network generation step

In the network generation algorithm, screening against databases is most commonly carried out after pathway reconstruction. In BNICE.ch, we have also introduced the notion of supervised network generation through the adaptable search space in the *de novo* pathway prediction process. The adoptable search space allows searching within a domain of metabolites and reactions that are predefined as a parameter, and the supervision can be applied for the generated compounds or reactions, or both, leading to the following features:

- Selection of the compound search space:
  where in each iteration, we keep only those compounds that are part of a biological or chemical database, or both (vs keeping all known and novel compounds in each iteration).
- Selection of the reaction search space:

where in each iteration, we allow only known KEGG reactions or reactions that are part of a specific database (vs keeping all known and novel reactions in each iteration).

These features can also be implemented in every network generation algorithm for the efficient organization of results based on the knowledge that exists in databases and to address the risk of combinatorial explosion.

Qualitative pruning in the pathway enumeration step

The complications that arise from the huge number of generated pathways resulting in the existing pathway enumeration methods has been previously discussed [14,39,44] and solutions have been proposed to enumerate 'a set of viable pathways' based on predefined criteria, rather than all possible pathways.

In BNICE.ch, we have also implemented the notion of supervised pathway enumeration to evaluate pathways based on our knowledge of compounds and reactions in databases. For example, we can enumerate only pathways with a prespecified percentage of their steps existing in biological databases as known enzymatic reactions.

*Protein sequence identification for **de novo** reactions*

A compelling aspect of the interactive analysis with databases is the structural similarity comparison of substrates and products of *de novo* reactions with the substrates and products of known reactions. The results of such a comparison could be quantified using different chemoinformatics metrics, such as 'compound fingerprint comparison' using the 'Tanimoto distance' and assigning to novel reactions a similarity score with respect to the existing reactions. Using such a metric, one can identify gene and protein sequences for the *de novo* steps of a pathway based on their structural similarities to known reactions. The enzymes encoded by those genes might be able to catalyze novel reactions but with very low activity, or they might perform very similar catalysis reactions, as they will belong to the same 3rd level in the Enzyme Commission (EC) classification system. Therefore, one must use evolution-based protein engineering and computational protein design [32••] to obtain sequences and enzymes for the experimental implementation of novel pathways with significant performance [7,45].

Quantitative pruning of generated pathways

Once we enumerate *de novo* pathways of interest and screen them against databases, the next step is to perform a feasibility analysis to determine the fitness and performance of individual pathways and to quantitatively prune the proposed pathways down to a set of the most biologically feasible ones. Quantitative pruning is generally context-dependent for the chassis organism. Different metrics can be applied to evaluate the likelihood of an *in silico*-designed pathway being proficiently implemented in an organism.

One crucial metric is the thermodynamics of the reaction steps and consequently the synthetic pathway to allow us to discard those pathways that are energetically unfavorable. To perform such a thermodynamics analysis, we developed a Group Contribution Method to estimate the Gibbs free energy for metabolites and consequently for reactions [46]. This method has been used in several frameworks to estimate the thermodynamics feasibility of the *in silico* generated synthetic pathways [18,24••,25•]. Furthermore, in BNICE.ch, we apply constraint-based modeling by incorporating the synthetic pathways one at a time into the genome scale model of chosen organism and performing Thermodynamics-based Flux Balance Analysis (TFBA) [47,48]. This additional step allows us to adjust the estimated Gibbs free energy based on the metabolite concentration, ionic strength, and pH to get closer to *in vivo* conditions. By performing a TFBA analysis, we guarantee that the obtained pathways are feasible with respect to mass balance (stoichiometrically), we assess the network thermodynamic feasibility of the generated pathways and we eventually quantify their overall effects on the metabolic profile of the organism by calculating the energetic cost and changes in the biomass yield for each molecule of the generated product [49,50]. One of the most important outcomes of TFBA for biotechnological applications is also the pruning and ranking of pathways based on the maximum production yield of the target molecule from each individual synthetic pathway. Other practical aspects have been also used for the quantitative pruning of *de novo* pathways, such as enzyme kinetics and gene compatibility [35•,36,51].

Scoring and ranking the biosynthetic pathways

By reconciling the metrics obtained in the qualitative and quantitative pruning strategies, one can define a scoring and ranking feature that combines and scales different factors and assigns an overall score for the prioritization of *in silico* generated pathways. Such a score gives the capability of pinpointing the best candidate synthetic pathways that are most likely to produce a desired target molecule and can be implemented in the metabolic network of the chassis organism. Additionally, one can rank the scores for a certain criterion as the primary ranking, and then perform a secondary ranking based on another criterion. For instance, choosing pathways with a maximum (or economically feasible) yield, and from those pathways choosing those with a minimum number of novel reactions, since their implementation will involve a smaller number of engineering enzyme steps.

Conclusions and perspectives

Computational retrobiosynthesis tools feature *de novo* pathway design and prioritization for synthetic biology and metabolic engineering studies.

Pruning approaches are crucial to avoid the risk of combinatorial explosion in the enumeration of *de novo* pathways. Here, we systematically classified the established methods and proposed strategies for pruning the generated *de novo* pathways.

One should be careful when applying certain criteria used for pruning the obtained data, recognizing that this is a multi-objective problem and different applications might give different weights to different criteria. Moreover, some of these criteria depend upon current technologies, and although some pathways can be currently ruled as infeasible, new technologies could enable their realization in the future.

Acknowledgments

The research in this study was supported by the LipidX RTD project of the Swiss initiative in Systems Biology (SystemsX), funded by the Swiss National Science Foundation (SNSF), and by funds from Ecole Polytechnique Fédéral de Lausanne (EPFL).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Shin JH, Kim HU, Kim DI, Lee SY: **Production of bulk chemicals via novel metabolic pathways in microorganisms**. *Biotechnol Adv* 2013, **31**:925-935.

2. Planson AG, Carbonell P, Grigoras I, Faulon JL: **A retrosynthetic biology approach to therapeutics: from conception to delivery**. *Curr Opin Biotechnol* 2012, **23**:948-956.

3. Medema MH, van Raaphorst R, Takano E, Breitling R: **Computational tools for the synthetic design of biochemical pathways**. *Nat Rev Microbiol* 2012, **10**:191-202.

4. Stephanopoulos G: **Synthetic biology and metabolic engineering**. *Acs Syn Biol* 2012, **1**:514-525.

5. Curran KA, Alper HS: **Expanding the chemical palate of cells by combining systems biology and metabolic engineering**. *Metab Eng* 2012, **14**:289-297.

6. Metzenberg RL: **Norman Harold Horowitz, 1915–2005**. *Genetics* 2005, **171**:1445-1448.

7. Bachmann BO: **Biosynthesis: is it time to go retro?** *Nat Chem Biol* 2010, **6**:390-393.

8. Corey EJ: **The logic of chemical synthesis — multistep synthesis of complex carbogenic molecules**. *Angew Chem-Int Ed Engl* 1991, **30**:455-465.

9. Law J, Zsoldos Z, Simon A, Reid D, Liu Y, Khew SY, Johnson AP, Major S, Wade RA, Ando HY: **Route designer: a retrosynthetic analysis tool utilizing automated retrosynthetic rule generation**. *J Chem Information Modeling* 2009, **49**:593-602.

10. Finley SD, Broadbelt LJ, Hatzimanikatis V: **Computational framework for predictive biodegradation**. *Biotechnol Bioeng* 2009, **104**:1086-1097.

11. Ellis LB, Roe D, Wackett LP: **The University of Minnesota Biocatalysis/Biodegradation Database: the first decade**. *Nucleic Acids Res* 2006, **34**:D517-D521.

12. • Hadadi N, Soh KC, Seijo M, Zisaki A, Guan XL, Wenk MR, Hatzimanikatis V: **A computational framework for integration of lipidomics data into metabolic pathways**. *Metab Eng* 2014, **23**:1-8.
In this study, the retrobiosynthesis concept is used to integrate lipid species with unknown metabolic pathways into biological networks based on pathways for their biosynthesis and biodegradation that were identified using BNICE.ch.

13. Bacher A, Rieder C, Eichinger D, Arigoni D, Fuchs G, Eisenreich W: **Elucidation of novel biosynthetic pathways and metabolite flux patterns by retrobiosynthetic NMR analysis**. *Fems Microbiol Rev* 1998, **22**:567-598.

14. Carbonell P, Fichera D, Pandit SB, Faulon JL: **Enumerating metabolic pathways for the production of heterologous target chemicals in chassis organisms**. *Bmc Syst Biol* 2012:6.

15. Lee JW, Na D, Park JM, Lee J, Choi S, Lee SY: **Systems metabolic engineering of microorganisms for natural and non-natural chemicals**. *Nat Chem Biol* 2012, **8**:536-546.

16. Martin CH, Nielsen DR, Solomon KV, Prather KLJ: **Synthetic metabolism: engineering biology at the protein and pathway scales**. *Chem Biol* 2009, **16**:277-286.

17. Long MR, Ong WK, Reed JL: **Computational methods in metabolic engineering for strain design**. *Curr Opin Biotechnol* 2015, **34C**:135-141.

18. Cho A, Yun H, Park JH, Lee SY, Park S: **Prediction of novel synthetic pathways for the production of desired chemicals**. *Bmc Syst Biol* 2010:4.

19. Hatzimanikatis V, Li CH, Ionita JA, Henry CS, Jankowski MD, Broadbelt LJ: **Exploring the diversity of complex metabolic networks**. *Bioinformatics* 2005, **21**:1603-1609.

20. Hatzimanikatis V, Li CH, Ionita JA, Broadbelt LJ: **Metabolic networks: enzyme function and metabolite structure**. *Curr Opin Struct Biol* 2004, **14**:300-306.

21. Soh KC, Hatzimanikatis V: **DREAMS of metabolism**. *Trends Biotechnol* 2010, **28**:501-508.

22. Carbonell P, Planson AG, Fichera D, Faulon JL: **A retrosynthetic biology approach to metabolic pathway design for therapeutic production**. *Bmc Syst Biol* 2011:5.

23. Rodrigo G, Carrera J, Prather KJ, Jaramillo A: **DESHARKY: automatic design of metabolic pathways for optimal cell growth**. *Bioinformatics* 2008, **24**:2554-2556.

24. •• Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R *et al.*: **Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol**. *Nat Chem Biol* 2011, **7**:445-452.
The authors introduce a *de novo* pathway approach for the production of 1,4-butanediol in *E. coli*. They also used strain design algorithms to engineer a strain of *E. coli* for the production of up to 18 g/L of 1,4-butanediol. This is the first case of experimental design and implementation of a novel pathway discovered from computational retrobiosynthesis.

25. • Campodonico MA, Andrews BA, Asenjo JA, Palsson BO, Feist AM: **Generation of an atlas for commodity chemical production in *Escherichia coli* and a novel pathway prediction algorithm, GEM-Path**. *Metab Eng* 2014, **25**:140-158.
The authors performed a retrobiosynthesis analysis for the design and analysis of *de novo* pathways in *E. coli*, and they proposed 245 unique synthetic pathways for 20 commodity chemicals. In the process of pathway enumeration, the proposed algorithm filters out infeasible pathways by associating them with the genome-scale model of *E. coli* and analyzing their cellular integrity.

26. Prather KL, Martin CH: **De novo biosynthetic pathways: rational design of microbial chemical factories**. *Curr Opin Biotechnol* 2008, **19**:468-474.

27. Chou CH, Chang WC, Chiu CM, Huang CC, Huang HD: **FMM: a web server for metabolic pathway reconstruction and comparative analysis**. *Nucleic Acids Res* 2009, **37**:W129-W134.

28. Faust K, Dupont P, Callut J, van Helden J: **Pathway discovery in metabolic networks by subgraph extraction**. *Bioinformatics* 2010, **26**:1211-1218.

29. Dale JM, Popescu L, Karp PD: **Machine learning methods for metabolic pathway prediction**. *BMC Bioinformatics* 2010, **11**:15.

30. Saurabh Srivastava JK, Stephi Hamilton, Paul Ruan, Jeff Tsui J, Christopher Anderson: **Pathway Synthesis using the Act Ontology**. *4th International Workshop on Bio-Design Automation (IWBDA)*. 2012.

31. Henry CS, Broadbelt LJ, Hatzimanikatis V: **Discovery and analysis of novel metabolic pathways for the biosynthesis of industrial chemicals: 3-hydroxypropanoate**. *Biotechnol Bioeng* 2010, **106**:462-473.

32. •• Brunk E, Neri M, Tavernelli I, Hatzimanikatis V, Rothlisberger U: **Integrating computational methods to retrofit enzymes to synthetic pathways**. *Biotechnol Bioeng* 2012, **109**:572-582.
Using a molecular dynamic modeling approach, the authors used 3D information regarding enzymes to perform a quantitative assessment of the binding affinity and catalytic power of an enzyme with a non-native substrate. This approach demonstrates for the first time how one can prune *de novo* pathways based on the structures of substrates and enzymes.

33. Martin CH, Nielsen DR, Solomon KV, Prather KL: **Synthetic metabolism: engineering biology at the protein and pathway scales**. *Chem Biol* 2009, **16**:277-286.

34. Feher T, Planson AG, Carbonell P, Fernandez-Castane A, Grigoras I, Dariy E, Perret A, Faulon JL: **Validation of RetroPath, a computer-aided design tool for metabolic pathway engineering**. *Biotechnol J* 2014, **9**:1446-1457.

35. • Carbonell P, Parutto P, Herisson J, Pandit SB, Faulon JL: **XTMS: pathway design in an eXTended metabolic space**. *Nucl Acids Res* 2014, **42**:W389-W394.
This study illustrates how chemoinformatics methods can be used in retrobiosynthesis. The authors used the molecular signatures of compounds as a filtering criterion in the pathway enumeration process. The maximum feasible yield of the products, toxicity, and enzyme efficiency are also applied as crucial design criteria.

36. Planson AG, Carbonell P, Paillard E, Pollet N, Faulon JL: **Compound toxicity screening and structure–activity relationship modeling in *Escherichia coli***. *Biotechnol Bioeng* 2012, **109**:846-850.

37. Carbonell P, Planson AG, Faulon JL: **Retrosynthetic design of heterologous pathways**. *Methods Mol Biol* 2013, **985**:149-173.

38. Mavrovouniotis M, Stephanopoulos G, Stephanopoulos G: **Synthesis of biochemical production routes**. *Comput Chem Eng* 1992, **16**:605-619.

39. Yousofshahi M, Lee K, Hassoun S: **Probabilistic pathway construction**. *Metab Eng* 2011, **13**:435-444.

40. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes**. *Nucl Acids Res* 2000, **28**:27-30.

41. Karp PD, Riley M, Saier M, Paulsen IT, Paley SM, Pellegrini-Toole A: **The EcoCyc and MetaCyc databases**. *Nucl Acids Res* 2000, **28**:56-59.

42. Engel T: **The structural- and bioassay database PubChem**. *Nachrichten Aus Der Chemie* 2007, **55**:521-524.

43. Degtyarenko K, De Matos P, Ennis M, Hastings J, Zbinden M, McNaught A, Alcantara R, Darsow M, Guedj M, Ashburner M: **ChEBI: a database and ontology for chemical entities of biological interest**. *Nucl Acids Res* 2008, **36**:D344-D350.

44. Pertusi DA, Stine AE, Broadbelt LJ, Tyo KE: **Efficient searching and annotation of metabolic networks using chemical similarity**. *Bioinformatics* 2015, **31**:1016-1024.

45. Obexer R, Studer S, Giger L, Pinkas DM, Grutter MG, Baker D, Hilvert D: **Active site plasticity of a computationally designed retroaldolase enzyme**. *Chemcatchem* 2014, **6**:1043-1050.

46. Jankowski MD, Henry CS, Broadbelt LJ, Hatzimanikatis V: **Group contribution method for thermodynamic analysis of complex metabolic networks**. *Biophys J* 2008, **95**:1487-1499.

47. Finley SD, Broadbelt LJ, Hatzimanikatis V: **In silico feasibility of novel biodegradation pathways for 1,2,4-trichlorobenzene**. *Bmc Systems Biol* 2010:4.

48. Henry CS, Broadbelt LJ, Hatzimanikatis V: **Thermodynamics-based metabolic flux analysis**. *Biophys J* 2007, **92**:1792-1805.

49. Kumar VS, Maranas CD: **GrowMatch: an automated method for reconciling in silico/in vivo growth predictions**. *PLoS Comp Biol* 2009:5.

50. Gonzalez-Lergier J, Broadbelt LJ, Hatzimanikatis V: **Analysis of the maximum theoretical yield for the synthesis of erythromycin precursors in *Escherichia coli***. *Biotechnol Bioeng* 2006, **95**:638-644.

51. Carbonell P, Lecointre G, Faulon JL: **Origins of specificity and promiscuity in metabolic networks**. *J Biol Chem* 2011, **286**:43994-44004.