Searching for the Protein Targets of Bioactive Molecules

Christopher Chidley, Hirohito Haruki, Miriam Grønlund Pedersen, Cindy Fellay, Simone Moser, and Kai Johnsson*

Abstract: The identification of all protein targets of a given drug or bioactive molecule within the human body is a prerequisite for an understanding of its beneficial and deleterious activities. Current approaches to reveal protein targets often fail to reveal physiologically relevant interactions. Here we review a recently introduced yeast-based approach for the identification of the binding partners of small molecules. We discuss the advantages and limitations of the approach using the clinically approved drug sulfasalazine as an example.

Keywords: Sulfasalazine · Target deconvolution · Tetrahydrobiopterin · Yeast three-hybrid screening

Target Deconvolution of Drugs and Bioactive Small Molecules

Most approved pharmaceutical drugs are small molecules that are usually either derived from natural products or are entirely synthetic. The great majority of these small molecule drugs elicit their therapeutic activity by interfering with the function of one or more proteins by physically binding to them.[1,2] Interfering with multiple protein targets can either enable a drug to be therapeutically used for multiple diseases or increase the efficacy of a drug to a particular disease if several of the targets are involved in pathways relevant to that disease.[3] However, harmful side effects can also arise from binding to multiple protein targets in addition to the protein target responsible for the efficacy of the drug. The identification of all protein targets of a small molecule drug, *i.e.* target deconvolution, therefore provides the basis for understanding its beneficial or deleterious actions. The discovery of unknown molecular targets of small molecule drugs could be useful for both compensating harmful side effects by prescribing adjunct therapies or by dose adjustments, and for repurposing drugs for new therapeutic uses. Target deconvolution of bioactive molecules has also become of greater importance in academia, mainly due to the advent of chemical genetics in which libraries of compounds are routinely screened for molecules capable of generating a certain cellular phenotype of interest. In these experiments, the identification of the target protein(s) is crucial as it opens the path towards the discovery of new biological insights. In conclusion, target deconvolution is an important aspect of research involving bioactive small molecules.

Current Approaches for Target Deconvolution of Drugs and other Bioactive Molecules

As outlined above, the identification of the binding partners of bioactive molecules or drugs is the focus of intensive research in academia and industry. In the following we will list some of the methods that have been successfully used for target deconvolution in the past years.[4,5] One approach towards target deconvolution is the analysis of gene expression after exposure of cells to a bioactive molecule, even if gene expression signatures of cells upon drug exposure are no direct proof for a physical interaction between a drug and a protein. Conceptually similar to the analysis of the transcriptome, protein or metabolite levels can be measured through proteomics strategies.[6–8] Another approach that can be used for the identification of proteins or pathways that are (again directly or indirectly) affected by bioactive molecules takes advantage of the ease with which the model organism yeast can be manipulated.[9–11] Here, the growth rate of different yeast strains in which a defined gene is deleted or overexpressed is measured in the presence of a drug. This approach requires that the molecule of interest binds both the human and yeast protein. Furthermore, different computational or cheminformatics-based approaches have shown great potential for identifying drug–protein interactions.[5,12,13] A recent example is the use of a computational chemical similarity approach to identify new targets for old drugs. So far the most successful approach for the direct identification of binding partners of bioactive molecules is chemical proteomics.[3] Chemical proteomics combines small molecule affinity chromatography for isolation of binding proteins, modern high-resolution mass spectrometry (MS) for identification of binding proteins, and statistics or bioinformatics for subsequent analysis of the data generated by MS. This approach benefits not only from the tremendous technological developments in MS during recent years but also from the experience that the community has gathered through the characterization of protein–protein interactions using affinity purification and MS (*i.e.* classical proteomics). The study of the binding proteins of the BCR-ABL inhibitors dasatinib, nilotinib and imatinib or the identification of the protein cereblon as a primary target of thalidomide teratogenicity are illustrative examples of the power of modern chemical proteomics.[14,15] Nevertheless, the use of chemical proteomics remains difficult for the identification of low-abundance target proteins or for proteins with low solubility or stability in cell extracts.

In summary, a variety of different methods for identifying the protein targets of bioactive molecules exist. However, as testified by (i) the large number of drugs with unknown mechanism of action and unexplained side effects and (ii) the ac-

Correspondence: Prof. Dr. K. Johnsson
Institute of Chemical Sciences and Engineering
NCCR Chemical Biology
École Polytechnique Fédérale de Lausanne
CH-1015 Lausanne
Tel.: +41 21 693 93 56
E-mail: kai.johnsson@epfl.ch

knowledged difficulty to identify the targets of bioactive compounds identified in chemical genetics experiments, there still is an urgent need for new approaches to identify small molecule–protein interactions.

A Yeast Three-hybrid System for Identifying the Protein Targets of Small Molecules

One potentially attractive alternative for target deconvolution of bioactive molecules is the so-called yeast three-hybrid system. The yeast three-hybrid system is an adapted version of the yeast two-hybrid system.[16] In the two-hybrid system the interaction of two proteins leads to the reconstitution of an active transcription factor and the subsequent transcription of a reporter gene (Fig. 1a). This can then be exploited for the identification of binding partners of a given protein by selecting genes out of cDNA or genomic libraries. In the yeast three-hybrid system, the reconstitution of an active transcription factor is linked to the interaction of a small molecule with a protein (Fig. 1b). For this, the small molecule of interest needs to be derivatized with a specific ligand that permits its binding to one of the partners of the original two-hybrid system. The approach has a number of attractive features: (i) interactions are detected in intact eukaryotic cells; (ii) the abundance of proteins with low expression levels is increased through expression in yeast; (iii) cDNA libraries from various human tissues and other organisms are commercially available; (iv) the identification of ligand-binding domains of proteins is possible; and (v) it requires no specialized equipment. However, it was found experimentally that initial yeast three-hybrid experiments suffered from limited sensitivity and presence of high levels of false positives. As a consequence, only very few successful examples of yeast three-hybrid-based target deconvolutions have been published so far.[16,17]

In the past few years we have developed a robust yeast three-hybrid system for the selection of small molecule–protein interactions that overcomes the shortcomings mentioned above.[18] Our system uses SNAP-tag as an anchor protein and O^6-benzylguanine (BG) derivatives of the bioactive molecules of interest.[19] SNAP-tag is a self-labeling protein tag that can be covalently labeled with BG derivatives in living cells (Fig. 2a).[20] In the SNAP-tag-based three-hybrid system the molecule of interest is first coupled to BG and then covalently linked to a SNAP-tag fusion protein in yeast, thereby coupling the interaction of the bioactive molecule with a protein to transcription of a gene (Fig. 2b). To
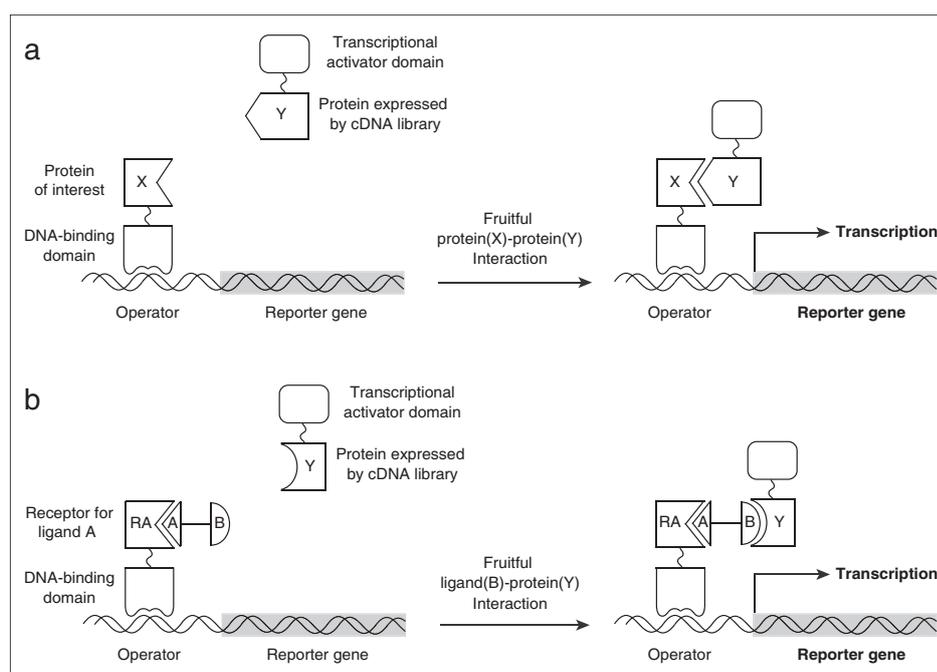

Fig. 1. Basic principle of the yeast two- and three-hybrid system. (a) Identification of protein–protein interactions using the yeast two-hybrid system. In yeast cells, the protein of interest (X) is expressed as a fusion with a DNA-binding domain and a cDNA library clone (Y) is expressed as a fusion to a transcriptional activator domain. The interaction of protein X with a protein Y induces the activation of transcription of a reporter gene by bringing the two fused domains into close proximity. Fruitful protein–protein interactions can be selected using auxotrophic markers or screened using colorimetric markers. (b) Identification of small molecule–protein interactions using the yeast three-hybrid system. In yeast cells, the receptor for a specific ligand (RA) is expressed as a fusion with a DNA-binding domain and a cDNA library clone is expressed as a fusion to a transcriptional activator domain. A synthetic hybrid molecule composed of the specific ligand (A) linked to the small molecule of interest (B) is used to anchor B onto the DNA-binding domain. The interaction of B with a protein (Y) induces the activation of transcription of a reporter gene by bringing the two fused domains into close proximity. Fruitful small molecule–protein interactions can be selected using auxotrophic markers or screened using colorimetric markers.
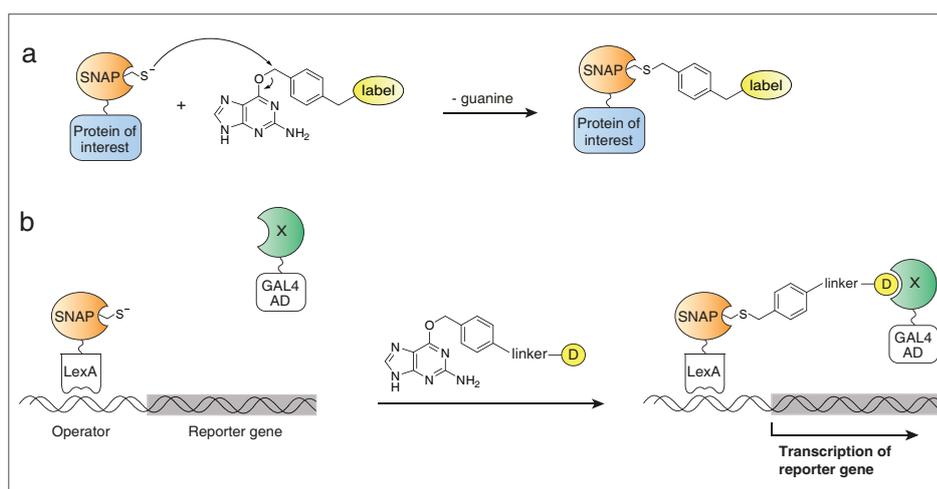

Fig. 2. The SNAP-tag-based yeast three-hybrid system. (a) Representation of the SNAP-tag labeling technique to label fusion proteins with small molecules of interest inside living cells. The protein of interest is expressed as a fusion protein with SNAP-tag and the small molecule of interest (label) is chemically synthesized as an O^6-benzylguanine (BG) derivative. Inside living cells, an active cysteine residue of SNAP-tag reacts with BG derivatives, resulting in the transfer of the label to the fusion protein. (b) Scheme of the SNAP-tag-based yeast three-hybrid system. Two fusion proteins consisting of a DNA-binding domain (LexA) fused to SNAP-tag and of a transcriptional activator domain (GAL4AD) fused to a target protein (or fused to a cDNA library clone) (X) are expressed in a reporter yeast strain. SNAP-tag reacts in living yeast cells with a BG derivative of the small molecule drug of interest (D), which results in the covalent anchoring of the molecule to LexA. A fruitful interaction of the small molecule with a target protein is detected by specific activation of a reporter gene.

make this system suitable for screenings, it needed to be optimized by engineering the yeast strain through deletion of multi-drug export pumps and by modifying the conditions of selection. Model selections showed that using these engineered strains, interactions with dissociation constants in the low micromolar range can be detected. This sensitivity is about 100 times better than what has been reported for initial yeast three-hybrid systems and is well-suited for detection of biologically relevant interactions. In order to reduce false positive levels, the reporter yeast strain used for three-hybrid selections was further engineered by incorporation of a new reporter gene. The incorporated reporter gene URA3 allows (i) positive selection on medium lacking uracil and (ii) negative selection on plates containing 5-fluoroorotic acid. The overall selection strategy is based on a negative selection step against unspecific interactions in the absence of BG derivative prior to the positive three-hybrid selection. This modification decreases the rate of false positives by a factor 10–100.

The workflow of our yeast three-hybrid system is depicted in Fig. 3. There are two key steps of the procedure that should be noted: A simple respotting step can be used for the elimination of the remaining false positives. This step consists in respotting each colony on selective plates containing the drug derivative and on selective plates not containing the drug derivative. Specific three-hybrid interactions are characterized by growth which is dependent on the presence of the drug derivative. This step combined with the negative selection step reduces the false positive rate to very low levels. Furthermore, the versatility of SNAP-tag allows the use of the same drug derivatives in yeast selections and in validation experiments (Fig. 3b). The validation of the hits is routinely performed by affinity chromatography: recombinant SNAP-tag covalently labeled with a drug derivative can be immobilized on beads and the specific binding of the hit protein to the immobilized drug can be detected by Western blot. Competition experiments with underivatized drug furthermore confirm the relevance of the identified interaction. These straightforward pull-down experiments are important for an independent confirmation of potential interactions.

Sulfasalazine and Tetrahydrobiopterin Biosynthesis

In the following we discuss the target deconvolution of the anti-inflammatory drug sulfasalazine as a representative example of our approach. Sulfasalazine is widely used against ulcerative colitis and rheumatoid arthritis. Introduced in 1942,
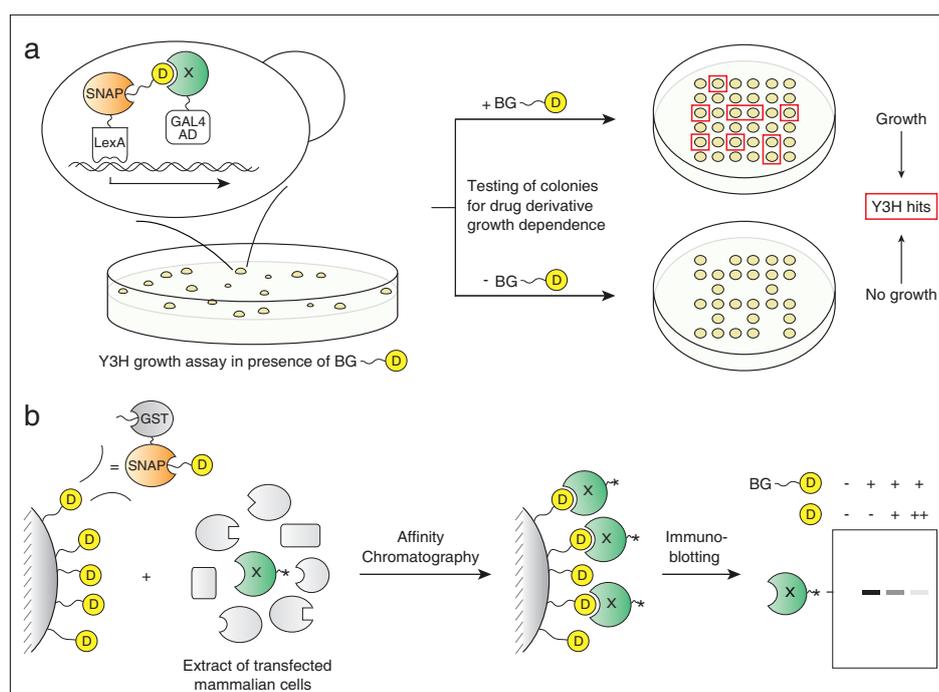

Fig. 3. Workflow of SNAP-tag-based yeast three-hybrid screening. (a) Yeast cells pre-transformed with the bait (LexA-SNAP) and with the cDNA library (GAL4-X) are plated on medium selective for reporter gene activation and in presence of a BG-drug conjugate. Colonies that grow on the selection plates are tested for drug derivative growth dependence and isolated hits are identified by DNA sequencing. (b) Proteins identified in (a) are expressed in mammalian cells with an epitope tag (represented as*) and drug binding is verified through a pull-down assay. For the latter assay, the drug is immobilized *via* a SNAP-GST fusion protein on beads.

it was generated by linking the antibiotic sulfapyridine with 5-aminosalicylic acid (mesalamine) through an azo bond.[21] Sulfasalazine is incompletely absorbed from the small intestine and carried to the colon where it is split by gut bacteria at the azo linkage, releasing 5-aminosalicylic acid and sulfapyridine (Scheme 1a).[22] It is believed that sulfasalazine serves as a vehicle to deliver its possible active components to the colon in higher concentrations than could be achieved by oral administration of either one alone.[23] With respect to the pharmacokinetics of the two molecules released, 5-aminosalicylic acid is believed to act locally in the colon, as it is poorly absorbed and rapidly eliminated.[24] In contrast, sulfapyridine is mostly absorbed and may act both locally, during mucosal absorption, and systemically.[23,25] In ulcerative colitis, it is believed that the therapeutic effect of sulfasalazine is, at least partially, due to the topical release of 5-aminosalicylic acid in the colon.[26] Sulfasalazine and mesalamine differ however, in their therapeutic properties. For example, whilst both drugs show similar efficacy in inducing remission of ulcerative colitis,[27] sulfasalazine shows superior efficacy in maintaining that remission.[28] Such differences between sulfasalazine and mesalamine might be due to their respective pharmacokinetics; alternatively, sulfasalazine or sulfapyridine might have additional targets.[29,30] In rheumatoid ar-

thritis, it was shown that sulfapyridine, and not 5-aminosalicylic acid, had a therapeutic effect similar to that of sulfasalazine.[31,32] The mechanism by which sulfasalazine and sulfapyridine affect the inflammation that characterizes rheumatoid arthritis is unclear.[33]

The medical importance of sulfasalazine combined with its obscure mechanism of action made it an attractive candidate for target deconvolution using our three-hybrid system. To subject sulfasalazine to three-hybrid screening, the drug needed to be derivatized first with BG. Based on practical considerations and the observation that sulfapyridine on its own also has activity against rheumatoid arthritis we decided to derivatize sulfasalazine *via* its carboxyl group (Scheme 1b). It should be noted that the chemical derivatization of a bioactive molecule might affect its binding to target proteins. This is an inherent limitation of all target deconvolution approaches that require drug derivatization. To minimize that risk, a bioactive molecule should ideally be derivatized at different positions and each derivative should be subjected to a screening. In the SNAP-tag-based yeast three-hybrid screening of sulfasalazine, three hits from a kidney cDNA library were identified as being dependent on BG-sulfasalazine for growth on selective medium. All three hits encoded the enzyme sepiapterin reductase (SPR, 261
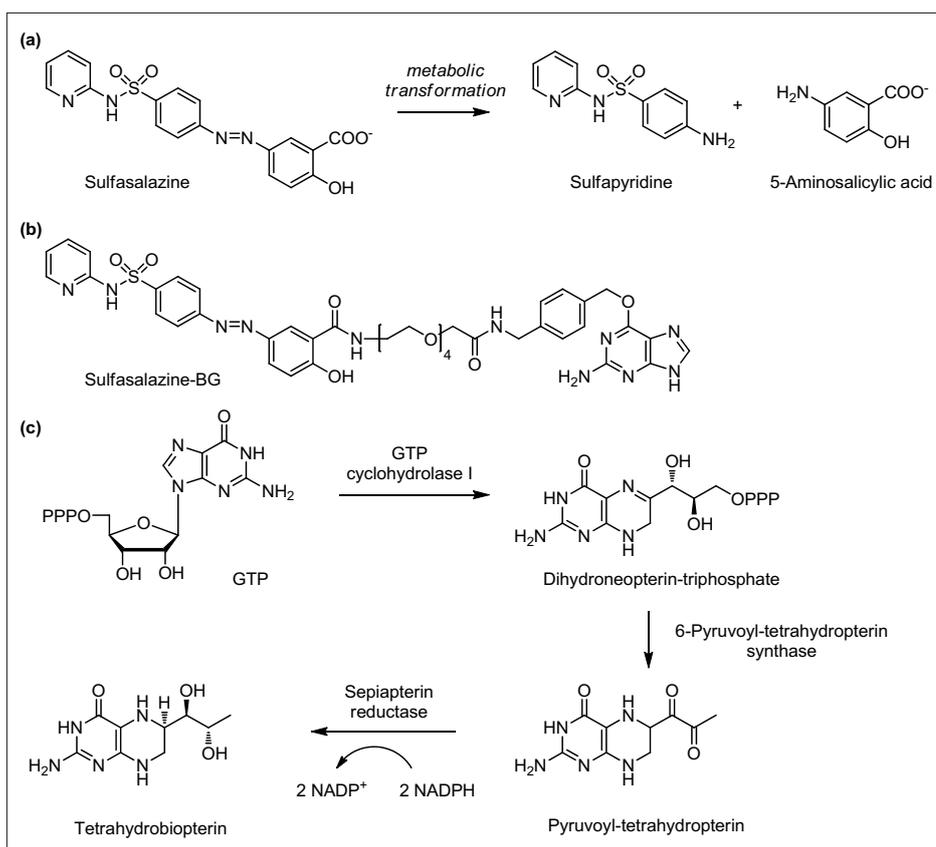
Scheme 1. Sulfasalazine and sepiapterin reductase (SPR). (a) Metabolism of sulfasalazine to sulfapyridine and aminosalicylic acid. (b) Structure of sulfasalazine and the BG derivative used for three hybrid screening. (c) Biosynthesis of tetrahydrobiopterin.

residues), two hits were full-length and the third hit had a very small N-terminal truncation (Δ1-4). The isolation of different variants of the same gene is a strong indication that a real interaction was identified. Indeed, the interaction of sulfasalazine with SPR was subsequently verified in a pull-down assay. What is the biological role of SPR? SPR catalyzes the NADPH-dependent reduction of 6-pyruvoyl tetrahydropterin to tetrahydrobiopterin (BH4), which is the final step in the biosynthesis of BH4 (Scheme 1c).[34] BH4 is utilized as a cofactor by a number of enzymes. These include: (i) hydroxylases that are involved in the biosynthesis of tyrosine, dopamine, and serotonin; (ii) glyceryl ether monooxygenases; (iii) all nitric oxide synthases (NOS).

A drug–protein interaction identified through a yeast three-hybrid screen does not establish that the binding of the molecule affects the activity of the protein. We therefore investigated if sulfasalazine inhibits the enzymatic activity of SPR. Using recombinant SPR, we could demonstrate that sulfasalazine and its metabolites sulfapyridine and mesalamine are all inhibitors of SPR: sulfasalazine, the parent compound, being the most potent inhibitor (IC_{50} = 25 nM); sulfapyridine being a potent inhibitor (IC_{50} = 2.3 μM); and finally mesalamine being a weak SPR inhibitor (IC_{50} = 680

μM). Could an inhibition of SPR explain the therapeutic activity of sulfasalazine and its metabolites? Increased nitric oxide (NO) levels and increased expression of inducible NOS (iNOS) have been associated with rheumatoid arthritis and ulcerative colitis, respectively. Furthermore, there is evidence in the literature that (i) inhibition of BH4 biosynthesis reduces the production of NO in various inflammatory models[35,36] and (ii) selective iNOS inhibitors reduced inflammation both in rheumatoid arthritis[37] and ulcerative colitis[38] models. Based on these data and our findings, we therefore propose that sulfasalazine and sulfapyridine inhibit iNOS activity indirectly through depletion of the cofactor and that this inhibition contributes to the anti-inflammatory activity of the drug. To test this hypothesis, we next measured the effect of sulfasalazine and its metabolites on intracellular, total biopterin levels in a cellular assay. At physiological relevant concentrations, sulfasalazine and sulfapyridine effectively reduced intracellular biopterin levels. These findings provide additional confirmation that the inhibition of BH4 biosynthesis by sulfasalazine and its metabolites plays an important role in the mechanism of action of the drug. A specific inhibition of BH4 biosynthesis also suggests new therapeutic applications for this drug. For example, investigations

on cancer patients with a reduced-function haplotype in BH4 biosynthesis[39] suggest that sulfasalazine might be effective in delaying pain in cancer patients. The observation that sulfasalazine blocks the development of tactile allodynia in diabetic rats supports the hypothesis that sulfasalazine could be used for alleviation of chronic pain.[40] Another interesting aspect of our findings on the mechanism of action of sulfasalazine concerns some of the side effects of the drug. Side effects of sulfasalazine include anorexia, headache, nausea and vomiting as well as mental or mood changes. It is possible that such side effects may at least partially be caused by a change in the concentration of neurotransmitters dependent on BH4 for their biosynthesis. As patients with deficiency in SPR are responsive to treatment with L-dopa,[41] it is reasonable to speculate that these side effects could be attenuated by adjunct therapy with molecules such as L-dopa that restore a correct balance of neurotransmitters.

Conclusions

Our results on sulfasalazine and other clinically approved drugs[18] demonstrate that the SNAP-tag-based yeast three-hybrid system is a powerful tool for target deconvolution of bioactive small molecules. It is experimentally simple and the availability of cDNA libraries from different tissues and organisms makes it suitable for the analysis of molecules with activities against various organisms. We consider the approach as complementary to other approaches for target deconvolution, in particular since the derivatized bioactive molecules used for three-hybrid screenings can also be used for chemical proteomics. The large number of bioactive small molecules with obscure mechanisms of action certainly creates ample opportunities for the three-hybrid system to prove its utility.

[1] J. P. Overington, B. Al-Lazikani, A. L. Hopkins, *Nat. Rev. Drug Discov.* **2006**, *5*, 993.
[2] P. Imming, C. Sinning, A. Meyer, *Nat. Rev. Drug Discov.* **2006**, *5*, 821.
[3] U. Rix, G. Superti-Furga, *Nat. Chem. Biol.* **2009**, *5*, 616.
[4] G. C. Terstappen, C. Schlupen, R. Raggiaschi, G. Gaviraghi, *Nat. Rev. Drug Discov.* **2007**, *6*, 891.
[5] J. N. Chan, C. Nislow, A. Emili, *Trends Pharmacol. Sci.* **2010**, *31*, 82.

[6] A. A. Cohen, N. Geva-Zatorsky, E. Eden, M. Frenkel-Morgenstern, I. Issaeva, A. Sigal, R. Milo, C. Cohen-Saidon, Y. Liron, Z. Kam, L. Cohen, T. Danon, N. Perzov, U. Alon, *Science* **2008**, *322*, 1511.

[7] S. M. Watkins, J. B. German, *Curr. Opin. Mol. Ther.* **2002**, *4*, 224.

[8] N. Vinayavekhin, E. A. Homan, A. Saghatelian, *ACS Chem. Biol.* **2010**, *5*, 91.

[9] G. Giaever, D. D. Shoemaker, T. W. Jones, H. Liang, E. A. Winzeler, A. Astromoff, R. W. Davis, *Nat. Genet.* **1999**, *21*, 278.

[10] S. Hoon, R. P. St Onge, G. Giaever, C. Nislow, *Trends Pharmacol. Sci.* **2008**, *29*, 499.

[11] G. Giaever, P. Flaherty, J. Kumm, M. Proctor, C. Nislow, D. F. Jaramillo, A. M. Chu, M. I. Jordan, A. P. Arkin, R. W. Davis, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 793.

[12] P. Kolb, R. S. Ferreira, J. J. Irwin, B. K. Shoichet, *Curr. Opin. Biotechnol.* **2009**, *20*, 429.

[13] C. McInnes, *Curr. Opin. Chem. Biol.* **2007**, *11*, 494.

[14] T. Ito, H. Ando, T. Suzuki, T. Ogura, K. Hotta, Y. Imamura, Y. Yamaguchi, H. Handa, *Science* **2010**, *327*, 1345.

[15] U. Rix, O. Hantschel, G. Durnberger, L. L. Remsing Rix, M. Planyavsky, N. V. Fernbach, I. Kaupe, K. L. Bennett, P. Valent, J. Colinge, T. Kocher, G. Superti-Furga, *Blood* **2007**, *110*, 4055.

[16] E. J. Licitra J. O. Liu, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12817.

[17] F. Becker, K. Murthi, C. Smith, J. Come, N. Costa-Roldan, C. Kaufmann, U. Hanke, C. Degenhart, S. Baumann, W. Wallner, A. Huber, S. Dedier, S. Dill, D. Kinsman, M. Hediger, N. Bockovich, S. Meier-Ewert, A. F. Kluge, N. Kley, *Chem. Biol.* **2004**, *11*, 211.

[18] C. Chidley, H. Haruki, M. G. Pedersen, E. Muller, K. Johnsson, *Nat. Chem. Bio.* **2011**, *7*, 375.

[19] S. Gendreizig, M. Kindermann, K. Johnsson, *J. Am. Chem. Soc.* **2003**, *125*, 14970.

[20] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86.

[21] N. Svartz, *Acta Medica Scandinavica* **1942**, *110*, 577.

[22] R. Caprilli, M. Cesarini, E. Angelucci, G. Frieri, *J. Crohns & Colitis* **2009**, *3*, 149.

[23] K. M. Das, R. Dubin, *Clin. Pharmacokinet.* **1976**, *1*, 406.

[24] L. A. Christensen, J. Fallingborg, K. Abildgaard, B. A. Jacobsen, G. Sanchez, S. H. Hansen, S. Bondesen, E. F. Hvidberg, S. N. Rasmussen, *Aliment Pharmacol. Ther.* **1990**, *4*, 523.

[25] A. K. Azadkhan, S. C. Truelove, J. K. Aronson, *Br. J. Clin. Pharmacol.* **1982**, *13*, 523.

[26] A. K. Azad Khan, J. Piris, S. C. Truelove, *Lancet* **1977**, *2*, 892.

[27] L. Sutherland, J. K. Macdonald, *Cochrane Database Syst. Rev*. **2006**, CD000543.

[28] L. Sutherland, J. K. Macdonald, *Cochrane Database Syst. Rev.* **2006**, CD000544.

[29] J. E. Baggott, S. L. Morgan, T. Ha, W. H. Vaughn, R. J. Hine, *Biochem. J.* **1992**, *282 Pt 1*, 197.

[30] C. Wahl, S. Liptay, G. Adler, R. M. Schmid, *J. Clin. Invest.* **1998**, *101*, 1163.

[31] T. Pullar, J. A. Hunter H. A. Capell, *Br. Med. J. (Clin. Res. Ed.)* **1985**, *290*, 1535.

[32] V. C. Neumann, A. J. Taggart, P. Le Gallez, C. Astbury, J. Hill, H. A. Bird, *J. Rheumatol.* **1986**, *13*, 285.

[33] B. N. Cronstein, *Br. J. Rheumatol.* **1995**, *34 Suppl 2*, 30.

[34] B. Thony, G. Auerbach, N. Blau, *Biochem. J.* **2000**, *347 Pt 1*, 1.

[35] S. S. Gross, R. Levi, *J. Biol. Chem.* **1992**, *267*, 25722.

[36] A. J. Bune, M. P. Brand, S. J. Heales, J. K. Shergill, R. Cammack, H. T. Cook, *Biochem. Biophys. Res. Commun.* **1996**, *220*, 13.

[37] J. R. Connor, P. T. Manning, S. L. Settle, W. M. Moore, G. M. Jerome, R. K. Webber, F. S. Tjoeng, M. G. Currie, *Eur. J. Pharmacol.* **1995**, *273*, 15.

[38] E. Kankuri, K. Vaali, R. G. Knowles, M. Lahde, R. Korpela, H. Vapaatalo, E. Moilanen, *J. Pharmacol. Exp. Ther*. **2001**, *298*, 1128.

[39] J. Lotsch, P. Klepstad, A. Doehring, O. Dale, *Pain* **2010**, *148*, 103.

[40] L. N. Berti-Mattera, T. S. Kern, R. E. Siegel, I. Nemet, R. Mitchell, *Diabetes* **2008**, *57*, 2801.

[41] N. Blau, L. Bonafe, B. Thony, *Mol. Genet. Metab.* **2001**, *74*, 172.