In Vivo Protein – Protein Interaction Assays: Beyond Proteins**

Hening Lin and Virginia W. Cornish*

In 1989 Fields and Song introduced the "Yeast Two-Hybrid Assay", which provides a straightforward method for detecting protein – protein interactions in vivo.[1] Up until the development of the two-hybrid methodology, protein binding had been detected using traditional biochemical techniques such as co-immunoprecipitation, affinity chromatography, and photoaffinity labeling.[2] There are three significant advantages to this in vivo assay which led almost immediately to its widespread use: first, it is technically straightforward and can be carried out rapidly; second, the sequence of the two interacting proteins can be read off directly from the DNA sequence of the plasmids encoding them; and third, it does not depend on the identity of the interacting proteins, and so is general.

The two-hybrid assay was based on the observations that eukaryotic transcriptional activators can be dissected into two functionally independent domains, a DNA-binding domain (DBD) and a transcription activation domain (AD), and that hybrid transcriptional activators can be generated by mixing and matching these two domains. It seems that the DNA-binding domain only needs to bring the activation domain into the proximity of the transcription start site; this suggests that the linkage between the DNA-binding and activation domains can be manipulated without disrupting activity. Thus, the linkage in the two-hybrid assay is the noncovalent bond between the two interacting proteins. As outlined in Figure 1, the yeast two-hybrid system consists of two protein chimeras and a reporter gene downstream from the binding site for the transcriptional activator. If the two proteins of interest (X and Y) interact, they effectively dimerize the DNA-binding protein chimera (DBD-X) and the transcription activation protein chimera (AD-Y). Dimerization of the DNA-binding domain and the transcription activation domain helps to recruit the transcription machinery to a promoter adjacent to the binding site for the transcriptional activator, thereby activating transcription of the reporter gene.

[*] Prof. V. W. Cornish, H. Lin
Department of Chemistry
Columbia University
New York, NY 10027 (USA)
Fax: (+1) 212-932-1289
E-mail: vc114@columbia.edu

[**] We would like to thank Tony Siu, Dr. Charles Cho, and the members of our lab for their helpful comments as we were preparing this manuscript.

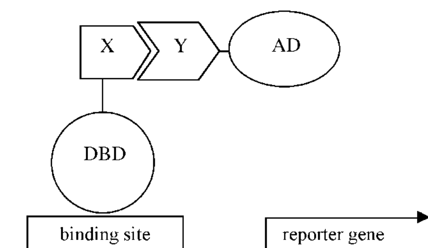


Figure 1. The yeast two-hybrid assay. DBD – X is a chimeric protein consisting of a DNA-binding domain (DBD) fused to protein X, and AD – Y is a second chimeric protein consisting of a transcription activation domain (AD) and protein Y. If protein X and Y bind to one another with reasonable affinity, the AD will be brought into the proximity of the reporter gene and will activate its transcription.

The assay was demonstrated initially using two yeast proteins known to be physically associated in vivo.[1] The yeast SNF1 protein, a serine – threonine protein kinase, was fused to the GAL4 DNA-binding domain, and the SNF1 activator protein SNF4 was fused to the GAL4 activation domain. A GAL4 binding sequence was placed upstream of a β-galactosidase reporter gene. Plasmids encoding the protein fusions and the reporter gene were introduced into yeast, and β-galactosidase synthesis levels were quantified using standard biochemical assays. Control experiments established that neither the DBD and AD domains on their own nor the individual protein chimeras induced β-galactosidase synthesis above background levels. β-galactosidase synthesis levels were increased 200-fold when the DBD – SNF1 and SNF4 – AD fusion proteins were introduced together. By comparison, the direct DBD – AD fusion protein activated β-galactosidase synthesis levels 4000-fold.

Since the initial paper by Fields and Song, there have been significant technical improvements in the method, and it is now an integral tool in biochemistry and genetics laboratories. For example, as a first step towards testing their hypothesis that the cyclin-dependent kinase Cdc20 is the target for the spindle checkpoint in budding yeast, Murray and co-workers used the yeast two-hybrid assay to determine if any of the proteins known to be involved in the spindle checkpoint physically interact with Cdc20.[3] By facilitating the discovery of cascades of interacting proteins—in this case, the spindle checkpoint—the yeast two-hybrid assay helps researchers put together entire biochemical pathways and begin to understand how these proteins function together inside a cell. A

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2001 1433-7851/01/4005-0871 $ 17.50+.50/0

recent twist is to use automation techniques to screen entire genomes.[4] Every open-reading frame that encodes a protein—there are approximately 6000 in budding yeast—is fused both to the DBD and to the AD, and the two fusion libraries are screened against one another.

Now in widespread use, the two-hybrid assay has been used to identify thousands of new protein–protein interactions.[4, 5] In response to the utility of this approach, several laboratories have begun to develop transcription-based assays that can be carried out in bacteria, or protein–protein interaction assays based on alternative readouts such as enzyme complementation or fluorescence resonance energy transfer (FRET). Moreover, the assay is being extended to the detection of other interactions. There are one-hybrid assays for detecting DNA–protein interactions, and three-hybrid assays for detecting RNA–protein and small molecule–protein interactions. A few of these advances are highlighted in this article.

In Vivo Protein–Protein Interaction Assays

While similarities in transcription among eukaryotic organisms allow the yeast two-hybrid assay to be transferred readily to mammalian cell lines, new transcription-based assays must be developed for bacteria. There are several potential advantages to working in bacteria. Generally, molecular biology techniques have been optimized in *E. coli*. Since the transformation efficiency of *E. coli* is several orders of magnitude greater than that of yeast, a larger number of protein variants can be screened. The rapid doubling time of *E. coli* would decrease the time required for the selection experiments. For several years, only the yeast two-hybrid assay was in use. Recently, however, bacterial protein–protein interaction assays have begun to be reported.[6–9]

One approach to developing transcription-based assays in bacteria has taken advantage of the fact that many bacterial repressors and activators are dimeric proteins with structurally distinct DNA-binding and dimerization domains. Hu et al.[6] demonstrated that the C-terminal dimerization domain of the λ-repressor could be replaced with the leucine zipper dimerization domain from the yeast transcriptional activator GCN4. The chimeric protein is stable and can functionally replace the λ-repressor in vivo, so providing immunity to superinfection by the λ-bacteriophage or efficient repression of an artificial λ-promoter and β-galactosidase reporter gene. There are other strategies for detecting protein–protein interactions in bacteria as well.[6–9]

Alternatively, there is interest in developing assays for detecting protein–protein interactions that are not based on transcriptional activators or repressors. While the information that is known about the biochemical mechanism of transcription in eukaryotes suggests that two-hybrid assays should be able to detect weak interactions and be relatively insensitive to conformation, it may be that other types of assays are more effective. Several alternative approaches have been devised. Generally, these approaches rely either on the induced interaction of two complementary fragments of a protein to reconstitute enzymatic activity or fluorescence resonance energy transfer.[9–13]

One such assay is based on the synthesis of adenosine 3′,5′-cyclic monophosphate (cAMP) by adenylate cyclase (Figure 2).[9] The adenylate cyclase from *B. pertussis* can be split

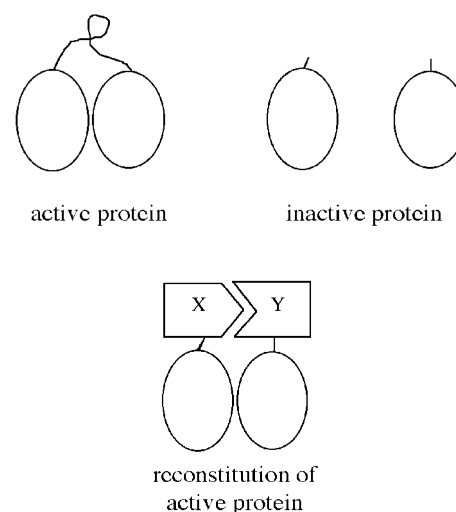


Figure 2. Protein complementation assay. A protein that carries out a detectable function is cleaved into two halves that are inactive when no longer connected. The two protein halves are fused to protein X and protein Y, respectively. If proteins X and Y bind to one another the protein's function is reconstituted.

into two functionally complementary fragments, T18 and T25, which allows protein dimerization to be assayed based on dimerization of T18 and T25 and reconstitution of adenylate cyclase activity. Adenylate cyclase is essential for the synthesis of cAMP. cAMP activates the catabolite activator protein (CAP), and the activated cAMP/CAP complex induces transcription of several genes, including the lac operon. Thus, it is possible to screen for adenylate cyclase activity using β-galactosidase plate assays or selections for growth on lactose. Ladant and co-workers[9] have demonstrated that dimerization of the leucine zipper domain of GCN4, the *N*-terminal domain of tyrosyl-tRNA synthase, and the yeast splicing factors Prp11 and Prp12 can all be detected using this assay. Similar assays using reconstitution of dihydrofolate reductase (DHFR) or β-galactosidase have been described by Remy and Michnick,[12] and Blau and co-workers,[11] respectively. It will be interesting to see the relative merits of different assays, and there may be ways to exploit assays that show conformational dependence or other supposed weaknesses.

DNA– and RNA–Protein Interactions

Early on, it was realized that, just as the yeast two-hybrid assay could be used to detect protein–protein interactions, transcriptional activators could be used directly, in a "one-hybrid" assay, to detect DNA–protein interactions (Figure 3 A). DNA-binding proteins that bind to a given target DNA sequence could be isolated from cDNA libraries encoding all of the proteins expressed in a given organism or specific cell type. Alternatively, the optimal or naturally occuring recognition sequences for a given regulatory protein could be determined. For example, with such an approach Wang and Reed cloned Olf-1, a transcriptional activator

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2001 1433-7851/01/4005-0872 $ 17.50+.50/0

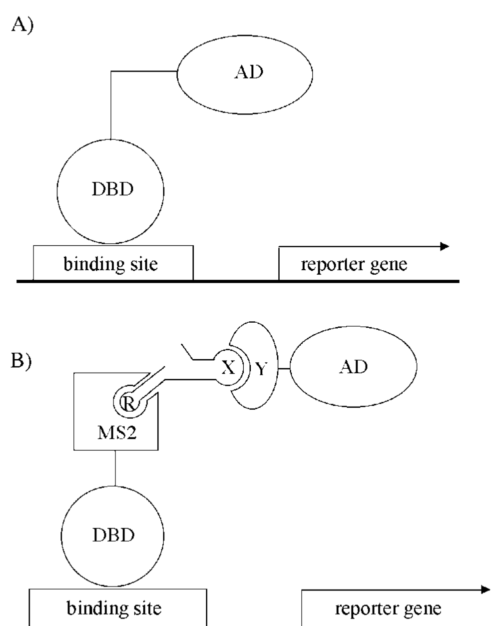


Figure 3. One- and three-hybrid assays. A) In the one-hybrid assay, the transcription activation domain (AD) is fused directly to the DNA-binding domain (DBD). This assay can be used to identify either a DBD that binds to a specific DNA sequence or the binding site for a given DBD. B) Compared with the yeast two-hybrid system, the three-hybrid system used to detect RNA–protein interactions has one additional component, a hybrid RNA molecule. One half of the hybrid RNA is a known RNA (R) that can bind the MS2 coat protein (MS2) with high affinity and serves as an anchor. The other half is RNA X, whose interaction with protein Y is being tested.

believed to be the critical switch for the coordinated expression of olfactory-specific genes.[14] A factor had been identified previously that was expressed in nuclear extracts from the nasal epithelium but not other tissues. The difficulty was in cloning this factor and determining its sequence. An olfactory cDNA library was subcloned into a GAL4 AD plasmid to produce millions of plasmids, each expressing a different olfactory protein fused to the GAL4 AD. These plasmids were then introduced into a yeast strain that allowed the Olf-1–GAL4 fusion protein to be selected from the millions of other olfactory protein fusions by a growth selection for transcription of an essential histidine biosynthesis gene. With an entirely different purpose, Pabo and coworkers recently used a bacterial one-hybrid selection system, not for studying natural DNA–protein interactions, but to engineer a zinc finger DNA-binding protein that binds to a designed target DNA sequence.[8]

Selecting for RNA–protein interactions is less straightforward because RNA–protein fusions cannot be generated directly in vivo and because routine biochemical assays that turn RNA-binding events into an amplifiable signal are not available. This difficulty was circumvented by adding a third component to the two-hybrid system and making a "three-hybrid" assay (Figure 3B).[15, 16] The third component is a hybrid RNA molecule where one half is a well-studied RNA molecule that binds to a known protein with high affinity and the other half is the RNA molecule of interest whose protein binding partner is in question. In total, then, the three-hybrid system consists of two protein chimeras, one RNA chimera, and a reporter gene. The hybrid RNA molecule bridges the DNA-binding and activation domain fusion proteins and activates transcription of a reporter gene. Wickens and coworkers first built this three-hybrid system[15] and then used it to identify a regulatory protein from *C. elegans* that binds to the 3′-untranslated region of the fem-3 gene and mediates the sperm/oocyte switch in hermaphrodites.[16]

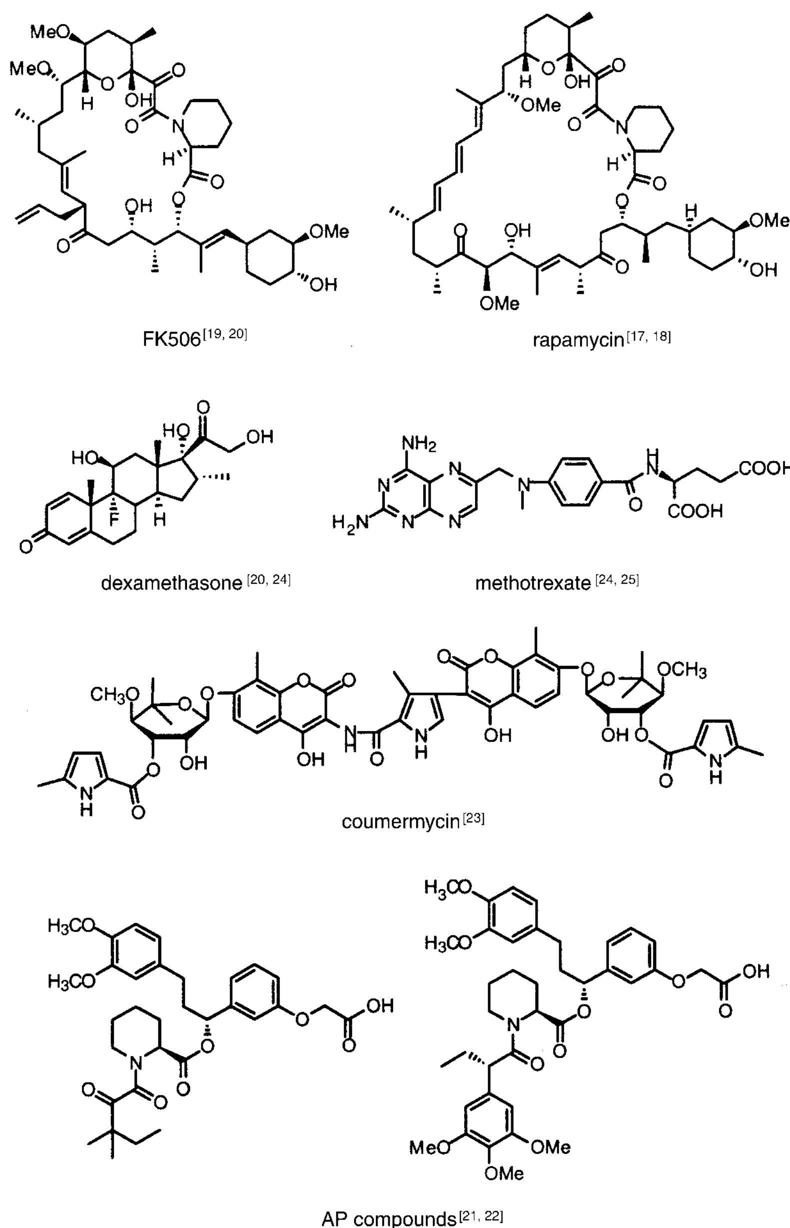
Small Molecule–Protein Interactions

Just as an RNA molecule can be introduced to mediate the interaction between the DNA-binding and activation domains, so can a small molecule. These small molecules have often been termed chemical inducers of dimerization (CIDs). To date, both monomeric and dimeric small molecules have been employed. Stan et al.[17] showed that the pharmacophore rapamycin could dimerize two fusion proteins, the DBD targets of rapamycin (TOR2) and the FK506/rapamycin binding protein (FKBP12)-AD, and could activate transcription of a reporter gene. This result was consistent with previous biochemical data suggesting that rapamycin acts by dimerizing TOR2 and FKBP12 in vivo. Furthermore, they used the transcription assay to show that the interaction requires a conserved Ser residue in TOR2, as a Ser 1975 $\rightarrow$Arg mutant of TOR2 failed to activate transcription of the reporter gene. Concurrently, a mammalian homologue of yeast TOR was identified by Chiu et al. from a mouse cDNA library using a similar transcription-based assay,[18] which demonstrated that this modified two-hybrid system could also be used in large-scale screening and selections.

Dimeric CIDs have been used much in the same way as hybrid RNA molecules, with one half of the molecule serving as an anchor and the other half being the compound in question. The first dimeric CIDs were dimers of the immunosuppresant FK506.[19] Licitra and Liu[20] built what they called a "yeast three-hybrid assay" that employs two fusion proteins: the glucocorticoid receptor (GR) fused to the DNA-binding protein LexA, and FKBP12 fused to the B42 transcription activation domain. The two fusion proteins are bridged by a heterodimeric dexamethasone–FK506 molecule. Dexamethasone binds to GR with a low nanomolar dissociation constant, as does FK506 to FKBP12. The key to the success of these systems is most likely the ligand–receptor pairs, hence one major area of development is new CID pairs. Researchers at ARIAD Gene Therapeutics have developed several CIDs based on purely synthetic analogues of FK506.[21, 22] Coumermycin, which is a naturally occurring asymmetric homodimer, has also been used as a CID.[23] We recently reported a dexamethasone–methotrexate CID that efficiently dimerizes GR and DHFR in the yeast three-hybrid assay.[24] The main advantage to this system is that both dexamethasone and methotrexate are readily available and that the heterodimeric derivative can be prepared readily. In addition the low picomolar affinity of methotrexate for the *E. coli* DHFR probably contributes to the efficacy of this CID. A methotrexate homodimer is being developed for use in the bacterial λ-repressor system.[25]

In their simplest application, CIDs can be used in combination with transcription-based dimerization assays as diffusible inducers of gene transcription. There are also many

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2001 1433-7851/01/4005-0873 $ 17.50+.50/0



other potential applications. Licitra and Liu[20] suggested that these systems be used to identify the cellular targets of known pharmacophores. In a proof of principle experiment using the immunosuppressant FK506, they showed that a heterodimeric dexamethasone–FK506 molecule could be used to select the natural target of FK506, FKBP12, from a Jurkat cDNA library fused to the B42 activation domain. The yeast three-hybrid assay can also serve as a tool for engineering proteins with new binding specificities. A variant of the FKBP12–rapamycin-binding (FRB) domain of the FKBP12–rapamycin-associated protein (FRAP) that is selective for a rapamycin analogue was isolated from a library of FRB mutants by Liberles et al. using a three-hybrid assay carried out in a mammalian cell line.[26] The FRB mutants were designed to create a new pocket in their binding sites to accommodate an additional substituent installed on rapamycin to block binding to wild-type FRB. The most selective mutant was identified by screening several FRB mutants in the three-hybrid assay.

Conclusion

The two-hybrid assay has already been used to detect thousands of new protein interactions, so what remains? One open question is whether other dimerization assays might be even more powerful than the two-hybrid assay—able to pick up lower affinity interactions or less sensitive to changes in conformation. It will be interesting to see how new in vivo dimerization assays fare over the next few years. There is still considerable room for optimization of the RNA or small-molecule ligand/receptor handles in the yeast three-hybrid assays to increase their sensitivity. The real question, however, is how this powerful assay can be exploited. While thousands of new protein interactions have been detected, the field is still wide open for chemists interested in understanding and manipulating biological interactions.[8, 26] Even more intriguing, perhaps, is the possibility that this assay can be used as more than just a readout for binding interactions.[27, 28]

[1] S. Fields, O. Song, *Nature* **1989**, *340*, 245–246.
[2] E. M. Phizicky, S. Fields, *Microbiol. Rev.* **1995**, *59*, 94–123.
[3] L. H. Hwang, L. F. Lau, D. L. Smith, C. A. Mistrot, K. G. Hardwick, E. S. Hwang, A. Amon, A. W. Murray, *Science* **1998**, *279*, 1041–1044.
[4] P. Uetz, L. Giot, G. Cagney, T. Mansfield, R. Judson, J. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J. Rothberg, *Nature* **2000**, *403*, 623–627.
[5] These interacting proteins can now be searched online at http://portal.curagen.com.
[6] J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, *Science* **1990**, *250*, 1400–1403.
[7] S. Dove, J. Joung, A. Hochschild, *Nature* **1997**, *386*, 627–630.
[8] J. K. Joung, E. I. Ramm, C. O. Pabo, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7382–7387.
[9] G. Karimova, J. Pidoux, A. Ullmann, D. Ladant, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5752–5756.
[10] S. Adams, A. Harootunian, Y. Buechler, S. Taylor, R. Tsien, *Nature* **1991**, *349*, 694–697.
[11] F. Rossi, C. Charlton, H. Blau, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8405–8410.
[12] I. Remy, S. Michnick, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5394–5399.
[13] N. Johnsson, A. Varshavsky, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10340–10344.
[14] M. M. Wang, R. R. Reed, *Nature* **1993**, *364*, 121–126.
[15] D. SenGupta, B. Zhang, B. Kraemer, P. Pochart, S. Fields, M. Wickens, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8496–8501.
[16] B. Zhang, M. Gallegos, A. Puoti, E. Durkin, S. Fields, J. Kimble, M. P. Wickens, *Nature* **1997**, *390*, 477–484.
[17] R. Stan, M. M. McLaughlin, R. Cafferkey, R. K. Johnson, M. Rosenberg, G. P. Livi, *J. Biol. Chem.* **1994**, *269*, 32027–32030.
[18] M. I. Chiu, H. Katz, V. Berlin, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12574–12578.
[19] D. Spencer, T. Wandless, S. Schreiber, G. Crabtree, *Science* **1993**, *262*, 1019–1024.
[20] E. Licitra, J. Liu, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12817–12821.

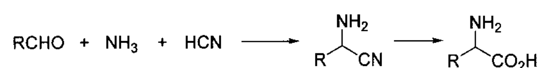
 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2001 1433-7851/01/4005-0874 $ 17.50+.50/0

[21] J. Amara, T. Clackson, V. Rivera, T. Guo, T. Keenan, S. Natesan, R. Pollock, W. Yang, N. Courage, D. Holt, M. Gilman, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10618–10623.
[22] T. Clackson, W. Yang, L. W. Rozamus, M. Hatada, J. F. Amara, C. T. Rollins, L. F. Stevenson, S. R. Magar, S. A. Wood, N. L. Courage, X. Lu, F. Cerasoli, M. Gilman, D. A. Holt, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10437–10442.
[23] M. A. Farrar, J. Alberola-lla, R. M. Perlmutter, *Nature* **1996**, *383*, 178–181.
[24] H. Lin, W. M. Abida, R. Sauer, V. Cornish, *J. Am. Chem. Soc.* **2000**, *122*, 4247–4248.
[25] S. J. Kopytek, R. F. Standaert, J. C. Dyer, J. C. Hu, *Chem. Biol.* **2000**, *7*, 313–321.
[26] S. D. Liberles, S. T. Diver, D. J. Austin, S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 7825–7830.
[27] S. Firestine, F. Salinas, A. Nixon, S. Baker, S. Benkovic, *Nat. Biotechnol.* **2000**, *18*, 544–547.
[28] R. Briesewitz, G. Ray, T. Wandless, G. Crabtree, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1953–1958.

Recent Developments in Catalytic Asymmetric Strecker-Type Reactions

Larry Yet*

The Strecker amino acid synthesis, which involves treatment of aldehydes with ammonia and hydrogen cyanide (or equivalents) followed by hydrolysis of the intermediate α-aminonitriles to provide α-amino acids (Scheme 1), was first

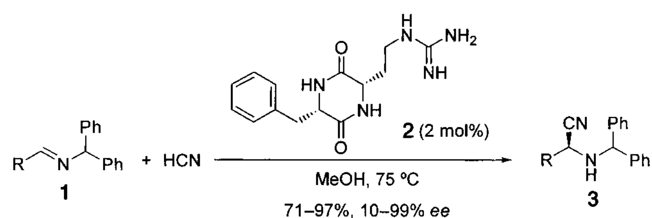


Scheme 1. Classical Strecker synthesis of α-amino acids.

reported in 1850.[1] This method has been applied on an industrial scale toward the synthesis of racemic α-amino acids, but more recently interest in nonproteinogenic α-amino acids in a variety of scientific disciplines has prompted intense activity in the asymmetric syntheses of α-amino acids.[2] The catalytic asymmetric Strecker-type reaction offers one of the most direct and viable methods for the asymmetric synthesis of α-amino acid derivatives. It is the purpose of this Highlight to disclose recent developments in this emerging field of importance.

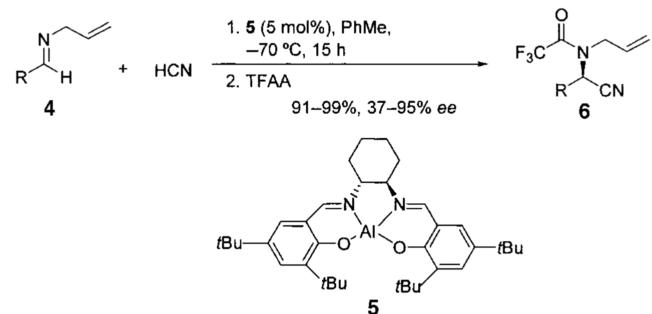
Lipton and co-workers investigated the viability of the asymmetric Strecker amino acid synthesis in which they utilized cyclic guanidine dipeptide **2** in the reaction of *N*-benzhydrylimines **1** with hydrogen cyanide to give *N*-benzhydryl-α-aminonitriles **3** (Scheme 2).[3] *N*-Benzhydrylimines **1**, derived from aromatic aldehydes, gave products **3** in generally high enantiomeric excess. However, electron-deficient 3-nitro, 3-pyridyl, and aliphatic aldehyde derivatives afforded racemic products.

[*] Dr. L. Yet
Albany Molecular Research, Inc.
21 Corporate Circle, P.O. Box 15098
Albany, NY 12212-5098 (USA)
Fax: (+1) 518-257-2021
E-mail: larryy@albmolecular.com



Scheme 2. Asymmetric Strecker synthesis with cyclic dipeptide **2** (Lipton and co-workers).

Sigman and Jacobsen reported the first example of a metal-catalyzed enantioselective Strecker-type reaction using a chiral Al^{III}–salen complex (salen = *N*,*N'*-bis(salicylidene)-ethylenediamine dianion).[4] A variety of *N*-allylimines **4** were evaluated in the reaction catalyzed by complex **5** to give products **6**, which were isolated as trifluoroacetamides in good yields and moderate-to-excellent enantioselectivities (Scheme 3). Substituted arylimines **4** were the best substrates,



Scheme 3. Asymmetric Strecker synthesis with chiral Al^{III}–salen catalyst **5** (Sigman and Jacobsen). TFAA = trifluoroacetic anhydride.

while alkyl-substituted imines afforded products with considerably lower *ee* values. Jacobsen and co-workers also reported that non-metal Schiff base catalysts **8** and **9** proved to be effective in the Strecker reaction of imines **7** with hydrogen

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2001 1433-7851/01/4005-0875 $ 17.50+.50/0