Library Biosynthesis

Genetically Selected Cyclic-Peptide Inhibitors of AICAR Transformylase Homodimerization**

*Ali Tavassoli and Stephen J. Benkovic**

The de novo purine biosynthetic pathway is used by virtually all organisms for the production of purine nucleotides. The final two steps of this pathway (Scheme 1) are catalyzed by aminoimidazole-4-carboxamide ribonucleotide transformylase/inosine monophosphate cyclohydrolase (AICAR Tfase/IMPCH), the two activities of a highly conserved 64 kDa bifunctional protein (ATIC) that possesses two distinct domains.[1] The C-terminal AICAR Tfase domain (residues 200–593) catalyzes the transfer of a formyl group from N_{10}-formyltetrahydrofolate (10-f-thf) to AICAR. The N-terminal IMPCH domain (residues 1–199) catalyzes the final step of the pathway.[2]

Cancer cells rely heavily on the de novo pathway for purine biosynthesis.[3] Thus, the inhibition of enzymes in this pathway is an attractive approach for the development of anticancer agents. As well as their potential uses in the treatment of malignant diseases, ATIC inhibitors are used in the treatment of inflammatory diseases, such as rheumethoid arthritis.[4]

The AICAR Tfase activity of ATIC is dependent on its homodimerization, whereas the IMPCH activity is not. The recently reported crystal structure shows ATIC as a dimer with an interface of ≈ 5000 Å^2.[2] There is much potential for the development of a new generation of therapeutic agents that act by inhibiting protein–protein interactions.[5] We chose genetic selection as the means to identify small molecules that specifically inhibit ATIC homodimerization and thereby inhibit AICAR Tfase activity. The selected approach utilizes whole cells as reporters of a designated intracellular event (interruption of a protein–protein interaction) by correlating host growth to the desired functional property of a small molecule. An advantage of this method is the selection of library members in vivo, which allows both affinity and selectivity to be assayed simultaneously.

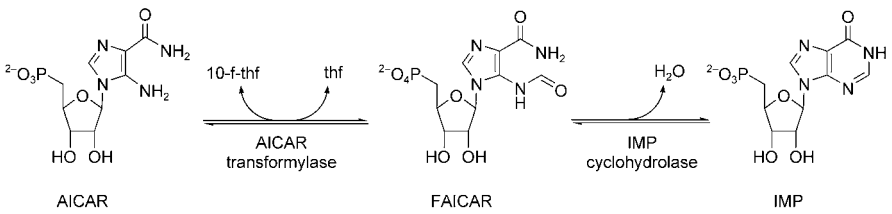
Specifically, the combination of our split intein-mediated circular ligation of peptides and proteins (SICLOPPS) technology (Scheme 2)[6] with a bacterial reverse two-hybrid system (RTHS) provides a method with the above characteristics for the systematic identification of small-molecule

[*] Dr. A. Tavassoli, Prof. S. J. Benkovic
Department of Chemistry
The Pennsylvania State University
University Park, PA 16802 (USA)
Fax: (+1) 814-865-2973
E-mail: sjb1@psu.edu

[**] We thank A. R. Horswill, T. A. Naumann, S. W. Nelson, S. N. Savinov, and M. E. Webb for helpful discussions. This work was funded by National Institutes of Health grant GM24129-28.
AICAR = aminoimidazole-4-carboxamide ribonucleotide.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

  DOI: 10.1002/ange.200500417



Scheme 1. Final two steps of the de novo purine biosynthetic pathway catalyzed by ATIC.

inhibitors of protein–protein interactions.[7] SICLOPPS allows the intracellular synthesis of libraries containing up to 10^8 cyclic peptides,[6] which is several orders of magnitude larger than those possible by conventional synthetic methods. The cyclization of peptides confers in vivo stability through the resulting resistance to degradation by proteases.[8]

Our bacterial RTHS[7] is based on the bacteriophage regulatory system,[9] which links the disruption of the fusion-protein homodimer to the expression of three reporter genes (Figure 1). HIS3[10] (imidazole glycerol phosphate dehydratase) and Kan^R (aminoglycoside 3′-phosphotransferase for kanamycin resistance) are two chemically tunable, conditionally selective reporter genes. The third reporter gene, LacZ (β-galactosidase), is used to quantify the protein–protein interaction through β-galactosidase assays.

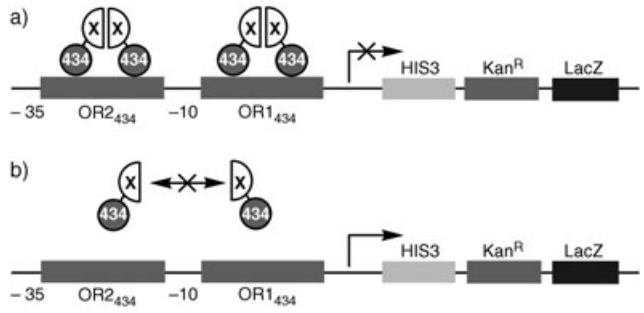
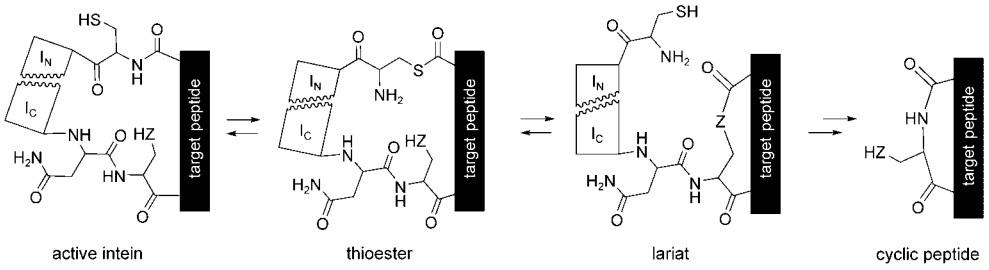


Figure 1. a) The protein fusions (X = target protein) associate (by binding to the 434 operator OR1 and OR2) to form a functional repressor of the promoter that directs expression of the reporter genes, thus inhibiting growth on minimal media. b) A small molecule inhibits the protein–protein interaction, thus restoring growth by the induction of HIS3 and Kan^R.

ATIC was cloned as a fusion with the DNA-binding domain of bacteriophage 434 repressor and placed under the control of an isopropyl-β-D-thiogalactoside (IPTG) inducible promoter. The fusion constructs showed IPTG-dependent repression of the reporter genes on selective media, thus confirming the formation of a functional repressor (Figure 1 a). To improve the selection conditions, a new RTHS strain was constructed by integrating the ATIC fusion onto the chromosome. The level of IPTG that leads to optimal repression was determined to be 50 μM by β-galactosidase assays.

The first SICLOPPS library transformed into the selection strain encoded a hexapeptide with five random residues and a cysteine nucleophile. Approximately 10^7 transformants were plated onto histidine-free minimal media supplemented with arabinose (inducer for SICLOPPS), 3-amino-1,2,4-triazole (3-AT, competitive inhibitor of the HIS3 product), and kanamycin at a density of 10^6 per plate (100×15 mm). The plates were incubated until colonies were readily visible (approximately 1 in 10^5). A second library, which encoded an octapeptide with five random residues and an invariable SGW motif, was also tested (Scheme 2). Around 200 colonies were picked and screened for arabinose-dependent growth advantage and IPTG-dependent inhibition of growth to eliminate false positive results. The expected phenotype was further confirmed by isolating and retransforming the selected SICLOPPS plasmids into the selection strain. The 14 remaining cyclic peptides were then ranked for activity by spotting serial dilutions of the corresponding cells onto selective media; the conferred growth advantage was compared at each dilution level (Figure 2).

To assess the in vivo target specificity of the selected cyclic peptides, a new RTHS strain containing a 434 repressor DNA-binding domain fusion with the *Saccharomyces cerevisiae* GCN4 leucine zipper (LZ) on its chromosome was constructed. The SICLOPPS plasmids of the active selectants were transformed into the LZ RTHS strain and ranked by drop spotting. ATIC-specific cyclic-peptide inhibitors were expected to be inactive in the LZ strain (identical to the ATIC RTHS strain except for the LZ homodimer). Five of the 14 selectants conferred a growth advantage (arabinose-dependent) on the LZ RTHS strain and were therefore discarded.



Scheme 2. An expressed fusion protein folds to form an active intein, which undergoes a series of rearrangements to generate a cyclic peptide. In this case the target cyclic peptide contains a series of randomly encoded amino acids to give libraries of $\approx 10^8$ members. Library 1: Z = S, target peptide = $CX_1X_2X_3X_4X_5$; library 2: Z = O, target peptide = $SGWX_1X_2X_3X_4X_5$ (X_n = random amino acid).

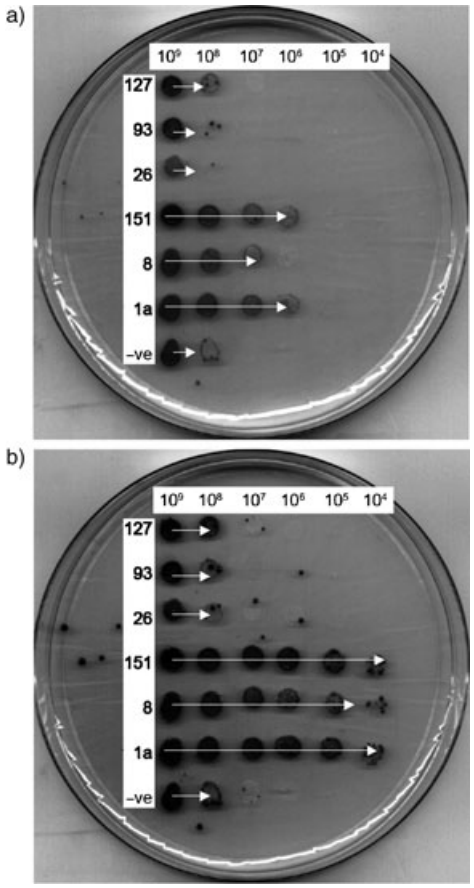


Figure 2. Determination of cyclic-peptide activity (rows) by drop spotting serial dilutions (2.5 μL of $\approx 10^n$ cells mL^{-1}) onto selective media containing arabinose and observing growth after 2 days (a) and after 4 days (b). In this case peptides **1a**, **8**, and **151** are active. Growth was not observed past the first column in the absence of arabinose (data not shown). −ve = control strain containing a blank SICLOPPS plasmid that does not express a cyclic peptide.

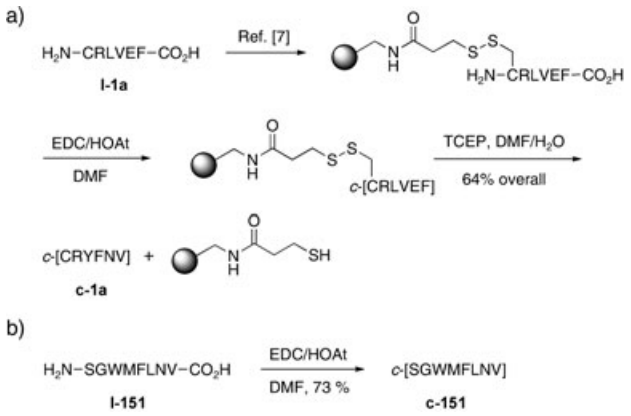
An inherent advantage of using genetically encoded libraries is the relative ease with which the structure of the active members can be determined (in contrast to deciphering synthetically derived libraries). Thus, DNA sequencing of the variable inserts present on the selected SICLOPPS plasmids readily revealed the amino acid sequence of the ATIC-specific cyclic-peptide inhibitors (Table 1).

There is considerable sequence homology in the genetically selected peptides. In general, arginine is favored in position 1, followed in position 2 by an aromatic amino acid (tyrosine or phenylalanine) in the more active, or an aliphatic amino acid (isoleucine, leucine, or valine) in the less active, cyclic peptides. The third random position is mainly occupied by leucine or phenylalanine. The three most active inhibitors contain an amino acid with an amide side chain (asparagine or glutamine) in position 4. The fifth amino acid is mostly valine or leucine.

Table 1: Sequence of the selected cyclic peptides in order of biological activity.

Activity rank	Name	Peptide sequence
1	**1a**	R Y F N V C
1	**151**	M F L N V S G W
2	**8**	R I L Q L C
2	**4**	R F F I C C
3	**6**	T V L M F C
3	**15**	S M M V L C
3	**5**	R I L V L C
3	**26**	P V L L L C
3	**25**	M L L I V C

As the AICAR Tfase activity of ATIC is dependent on its dimerization, disruption of the homodimer can be monitered in vitro by AICAR Tfase assays. The two most active cyclic peptides (**1a** and **151**) were synthesized chemically for in vitro testing. The synthesis of cyclic peptide **1a** involved the immobilization of the corresponding linear sequence on a modified amino polyethylene glycol acrylamide copolymer (PEGA) resin by forming a disulfide bond with its cysteine side chain.[7] The immobilized peptide was then cyclized and cleaved from the PEGA resin (Scheme 3a). The cyclization of linear peptide **151** was carried out in *N*,*N*-dimethylformamide (DMF) at high dilution to favor monomolecular cyclization (Scheme 3b).



Scheme 3. Synthesis of cyclic-peptide inhibitors **c-1a** and **c-151**. In both cases the linear peptides are cyclized by treatment with 1-ethyl-3-(3′-dimethylaminopropyl)carbodiimide (EDC) and 1-hydroxy-7-azabenzotriazole (HOAt) in DMF. Peptide **c-1a** is cleaved from the resin with tris(2-carboxyethyl)phosphane (TCEP).

The cyclic peptides were purified by reversed-phase chromatography. The chemical nature of the peptides was confirmed by comparison with samples prepared biologically (SICLOPPS) by reversed-phase HPLC and electrospray-ionization mass spectrometry. Cyclic peptides **c-1a** and **c-151**, as well as their linear counterparts **l-1a** and **l-151**, were assayed against AICAR Tfase. The peptides were assumed to compete with f-10-thf, which binds to ATIC and stabilizes its dimerization. From the measured k_{cat} value of the enzyme ($1.1\ s^{-1}$) and K_m value of f-10-thf (33.9 μM), the equation for competitive inhibition was used to determine the K_i values (Figure 3).

Peptide **c-1a** was found to have a K_i value of 17 ± 4.2 μM, whereas its linear counterpart **l-1a** has a K_i value of 142 ± 22.5 μM. Inhibitor **c-151** has a K_i value of 59 ± 6.8 μM, and again the linear peptide **l-151** is less active with a K_i value of

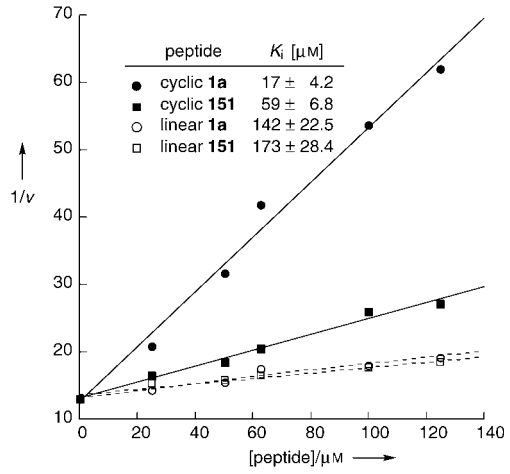


Figure 3. The K_i values of cyclic-peptide inhibitors **1a** and **151** and their linear counterparts as determined by assuming competitive inhibition with respect to 10-f-thf.

173 ± 28.4 μM. That both cyclic peptides are several times more potent than their linear counterparts confirms the superior activity of the genetically selected cyclic epitopes and demonstrates the inherent entropic benefit of a constrained scaffold. The cyclic peptides were also assayed against IMPCH and showed no inhibitory effects. IMPCH activity is not dependent on enzyme dimerization, which suggests that the compounds act by inhibiting ATIC dimerization.

The nature of the inhibition of the most active peptide, **c-1a**, was verified by progress-curve analysis.[11–14] The progress curves were fitted to a model in which the inhibitor binds a single protomer of ATIC, thereby preventing dimerization, as well as to the standard inhibition models (noncompetitive, uncompetitive, mixed, and competitive), by using DynaFit.[11] The data for peptide **c-1a** best fitted the nonstandard model (inhibition of enzyme dimerization) with respect to 10-f-thf and the noncompetitive inhibition model with respect to AICAR (see Supporting Information). This result is consistent with both the ordered binding observed for the enzyme and stabilization of the catalytic dimer by 10-f-thf.[15,16] Furthermore, the K_i value of **c-1a** obtained by this method (18 ± 8.6 μM) matches closely that obtained by assuming competitive inhibition with f-10-thf (17 ± 4.2 μM). The collective kinetic data confirm that cyclic peptide **1a** acts by inhibiting the dimerization of ATIC (as also indicated by the in vivo studies). We are currently attempting to evolve more-potent inhibitors by using second-generation SICLOPPS libraries (based on the selected sequences) and peptidomimetics.[17]

In summary, we have demonstrated a novel approach for the genetic selection of cyclic-peptide inhibitors of AICAR Tfase by combining the RTHS and SICLOPPS technologies. Nine cyclic peptides were selected from an intracellular library of 10^7 members and confirmed to function by selective disruption of the ATIC homodimer in vivo and in vitro. These compounds represent a striking structural departure from the antifolate-based inhibitors generally targeted against this enzyme.[18] The reported methodology allows rapid identification of small-molecule inhibitors of protein–protein interactions, and is thus a powerful and novel approach to drug discovery.

Received: February 3, 2005
Published online: April 13, 2005

Keywords: bioorganic chemistry · enzymes · inhibitors · library biosynthesis · peptides

[1] L. Ni, K. Guan, H. Zalkin, J. E. Dixon, *Gene* **1991**, *106*, 197.
[2] S. E. Greasley, P. Horton, J. Ramcharan, G. P. Beardsley, S. J. Benkovic, I. A. Wilson, *Nat. Struct. Biol.* **2001**, *8*, 402.
[3] R. C. Jackson, R. J. Harkrader in *Nucleosides and Cancer Treatment* (Eds.: M. H. N. Tattersall, R. M. Fox), Academic Press, Sydney, **1981**, p. 18.
[4] P. Gadangi, M. Longaker, D. Naime, R. I. Levin, P. A. Recht, M. C. Montesinos, M. T. Buckley, G. Carlin, B. N. Cronstein, *J. Immunol.* **1996**, *156*, 1937.
[5] R. Zutshi, M. Brickner, J. Chmielewski, *Curr. Opin. Chem. Biol.* **1998**, *2*, 62.
[6] C. P. Scott, E. Abel-Santos, A. D. Jones, S. J. Benkovic, *Chem. Biol.* **2001**, *8*, 801.
[7] A. R. Horswill, S. N. Savinov, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15591.
[8] Y. Q. Tang, J. Yuan, G. Osapay, K. Osapay, D. Tran, C. J. Miller, A. J. Ouellette, M. E. Selsted, *Science* **1999**, *286*, 498.
[9] J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, *Science* **1990**, *250*, 1400.
[10] J. K. Joung, E. I. Ramm, C. O. Pabo, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7382.
[11] P. Kuzmic, *Anal. Biochem.* **1996**, *237*, 260.
[12] S. R. Stone, H. J. Fromm, *Biochemistry* **1980**, *19*, 620.
[13] M. Bauer, H. Griengl, W. Steiner, *Biotechnol. Bioeng.* **1999**, *62*, 20.
[14] C. Frieden, *Trends Biochem. Sci.* **1993**, *18*, 58.
[15] J. M. Vergis, K. G. Bulock, K. G. Fleming, G. P. Beardsley, *J. Biol. Chem.* **2001**, *276*, 7727.
[16] K. G. Bulock, G. P. Beardsley, K. S. Anderson, *J. Biol. Chem.* **2002**, *277*, 22168.
[17] S. A. Andronati, T. L. Karaseva, A. A. Krysko, *Curr. Med. Chem.* **2004**, *11*, 1183.
[18] C. G. Cheong, D. W. Wolan, S. E. Greasley, P. A. Horton, G. P. Beardsley, I. A. Wilson, *J. Biol. Chem.* **2004**, *279*, 18034.

