Bioorganic & Medicinal Chemistry Letters xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Cell-free expression with the toxic amino acid canavanine

Emanuel G. Worst [a,†], Matthias P. Exner [b,†], Alessandro De Simone [b], Marc Schenkelberger [a], Vincent Noireaux [c], Nediljko Budisa [b], Albrecht Ott [a,*]

[a] Universität des Saarlandes, Department of Experimental Physics, Campus Saarbruecken, 66123 Saarbruecken, Germany
[b] Technische Universität Berlin, Institute of Chemistry, Mueller-Breslau-Strasse 10, 10623 Berlin, Germany
[c] University of Minnesota, School of Physics and Astronomy, Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:
Received 28 May 2015
Revised 12 June 2015
Accepted 13 June 2015
Available online xxxx

Keywords:
Arginine
Antimetabolite
Canavanine
Cell-free expression
Noncanonical amino acid

ABSTRACT

Canavanine is a naturally occurring noncanonical amino acid, which is analogous to arginine. It is a potent antimetabolite and natural allelochemic agent, capable of affecting or blocking regulatory and catalytic reactions that involve arginine. Incorporated into proteins at arginine positions, canavanine can be detrimental to protein stability and functional integrity. Although incorporation of canavanine into proteins has long been documented, due to its toxicity, expression in *Escherichia coli* and other common hosts remains a considerable challenge. Here, we present a simple, cell-free expression system with markedly improved performance compared to heterologous expression. The cell-free expression system does not require any tuning besides substitution of arginine by canavanine. We show that our technique enables highly efficient protein expression in small volumes with arginine being fully replaced by canavanine for functional and structural studies.

L-Arginine (Arg) is composed of a hydrophobic chain and a positively charged guanidinium group. The residue is generally located on the surface of a protein. It is often critical to catalytic processes, and it also participates in the formation of salt bridges. L-Canavanine (Can, Fig. 1A) is the oxyguanidino-analog of Arg. It was discovered first in raw seeds of *Canavalia ensiformis*[1] and found in several other plant species and families.[2] Its natural function is chemical defense, that is, Can serves as an allelochemic agent that deters the feeding activity of plant-eating insects and other herbivores.[3] This inspired a lot of studies on its antimetabolic, antibacterial, antifungal and antiviral properties.[4] Can is capable of selectively inhibiting inducible nitric oxide synthase, which prevents the overproduction of nitrite oxide.[5] It has been proposed as anti-cancer agent[6,7] and stimulator of autoimmune diseases in humans.[8]

Already in 1950s and 1960s, chemical, immunological and microbiological studies established Can incorporation into cellular proteins[9] both, in vivo and in vitro.[10–13] Indeed K_M for Arg is 45 μM whereas K_i for Can is 50 μM in yeast arginyl-tRNA-synthetases.[14] The reduced basicity of Can relative to arginine leads to distorted proteins and enzyme inactivation, exerting toxic effects in many species.[2,3,11] Most noncanonical amino acids that have been incorporated into proteins so far are normally limited to growth inhibitors, more or less well tolerated by the metabolism of living cells.[15] Conversely, the presence of Can in the living cell leads to immediate death.[15] Not surprisingly, only one instance of a protein with confirmed full replacement of arginines by canavanine has been reported so far, using an elaborate single-protein production system.[16]

The problem of high toxicity can be circumvented by the use of sophisticated in vitro expression systems. Such cell-free systems play an increasing role in a number of biotechnology and proteomic applications.[17] They produce recombinant proteins rather efficiently. High yields often rely on the presence of the T7 or T3 RNA polymerase and promoters. However, these elements entail bacteriophage-like transcription, limiting the functionality of the extract compared to the endogenous *Escherichia coli* transcription and translation machinery. *E. coli* based cell-free systems that work with the endogenous core RNA polymerase and the housekeeping sigma factor 70 overcome these restraints.[18]

In this work, we use a cell-free expression system[19] based on the endogenous *E. coli* transcription system that expresses proteins as efficiently as in current bacteriophage systems (0.5–1 mg/ml of recombinant protein) and enables complex procedures due to its native-like composition. The work presented here is one of a few examples of an in vitro system for supplementation-based incorporation of noncanonical amino acids. The rarer amino acid tryptophan was successfully substituted by an analog, at all positions

* Corresponding author. Tel.: +49 681 302 68550; fax: +49 681 302 68554.
E-mail address: albrecht.ott@physik.uni-saarland.de (A. Ott).
† Authors contributed equally.

http://dx.doi.org/10.1016/j.bmcl.2015.06.045
0960-894X/

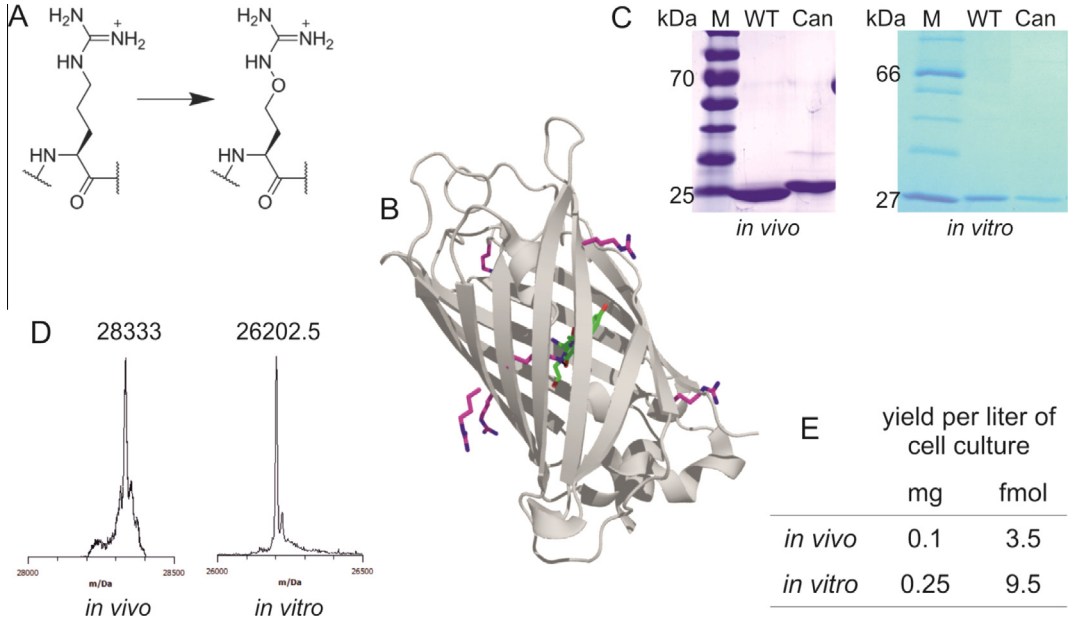

Figure 1. Comparison of in vivo and in vitro expression of green fluorescent protein (GFP) with Canavanine (Can). (A) Chemical structures of arginine and canavanine. (B) Distribution of arginine residues (highlighted in magenta) in a GFP protein model (structure from pdb id 2YOG). All residues are located in β-sheets. Five side chains face outwards, and only one is in contact with the fluorophore, localized in the barrel structure. EGFP has an additional arginine residue at position 2. (C) Gel of purified GFP model proteins. Left: In vivo expressed EGFP(WT) and EGFP(Can) after affinity and size exclusion chromatography; right: In vitro expressed dEGFP(WT) and dEGFP(Can) after affinity chromatography. (D) Confirmation of full incorporation of Can by ESI-MS; expected masses are 28333 Da (in vivo) and 26204 Da (in vitro). The mass difference of 1.5 Da for dEGFP(Can) is within the error of spectrum deconvolution. (E) Typical yields of EGFP(Can) (in vivo) and dEGFP(Can) (in vitro). In vitro yield is a prediction based on cell-free expression using 90 μL of cell extract derived from approximately 50 mL cell culture.

of model proteins, using a specialized trp-depleted cell extract.[20] The incorporation of canavanine (and several other noncanonical amino acids) was achieved with a highly diafiltrated cell extract containing T7 polymerase.[21] To the best of our knowledge, this is the first publication that reports the extract-based cell-free incorporation of canavanine into model proteins at all arginine positions with a native *E. coli* system, enabling incorporation into a wide range of target proteins. Additionally, we attempted incorporation of canavanine into model proteins in vivo using a standard supplementation pressure incorporation protocol based on depletion of arginine in auxotrophic cells and subsequent addition of canavanine as a surrogate amino acids.

We chose fluorescent proteins as model proteins to gain a direct visual control of protein expression. For in vivo expressions, we use enhanced green fluorescent protein (EGFP) with an N-terminal hexahistidine-tag (and a total of 7 arginine residues), while in vitro a truncated variant of EGFP (dEGFP, 6 arginine residues) is expressed (see Fig. 1B for structure and Supplementary Fig. S1 for sequences).

Canavanine was successfully incorporated into model proteins in vivo and in vitro (Fig. 1C). The replacement of arginine was complete in both cases and no arginine incorporation was observed, as confirmed by mass spectrometry (Fig. 1D). EGFP (7 arginines replaced) was expressed in vivo with a yield of just 0.1 mg/L.

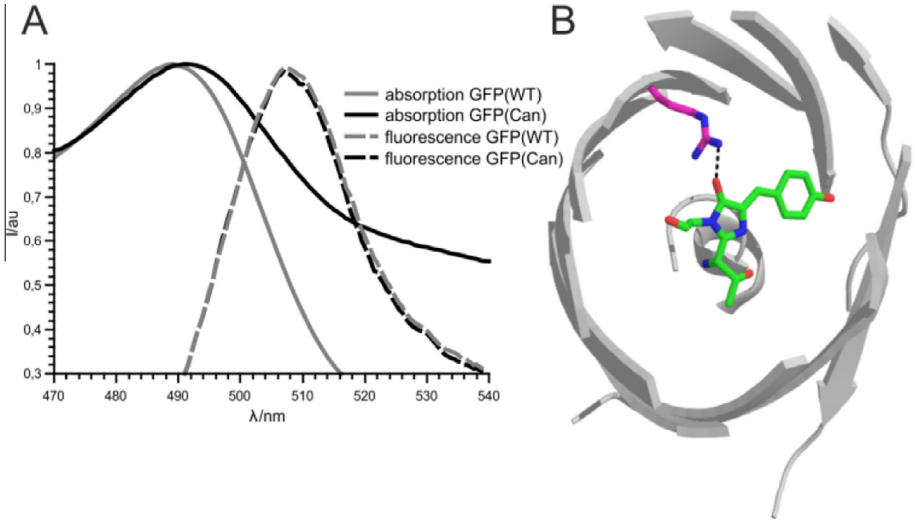

Figure 2. Spectroscopic properties of dEGFP(Can). (A) Absorption and fluorescence spectra of dEGFP(WT) and dEGFP(Can). Five spectra were accumulated for each sample and normalized. After the incorporation of canavanine, the absorption maximum is slightly red-shifted (491 nm vs 489 nm) with unchanged emission maximum at 508 nm. (B) Crystal structure of EGFP (top view, from PDB ID 2Y0G) showing a polar contact between the fluorophore and an arginine/canavanine sidechain. The fluorophore is highlighted in green, the arginine/canavanine position is highlighted in magenta.

Please cite this article in press as: Worst, E. G.; et al. *Bioorg. Med. Chem. Lett.* (2015), http://dx.doi.org/10.1016/j.bmcl.2015.06.045

dEGFP (6 arginines replaced) was expressed at 0.25 mg/L using our cell-free expression system. Yields are summarized in Figure 1E. While the expression yields cannot be directly compared due to the fact that different model proteins were used, these values show that cell-free expression is a viable alternative to fermentation for production of proteins containing canavanine at all arginine positions.

Furthermore the reduced expression volume of cell-free reactions (270 μL vs 4 L) greatly simplifies protein purification procedures. Interestingly, despite not being specifically depleted or purified, the cell extract does not retain enough arginine to promote translation of protein without addition of either arginine or canavanine, as shown in Supplementary Figure S2. An additional advantage is the enhanced stability of Can during cell-free expression.

In several batches of in vivo expressions, we also found proteins with significantly lower mass than expected (Supplementary Fig. S3). We speculate that this is due to canavanine cleavage to homoserine and hydroxyguanidine, a well-described pathway.[22,23] However, the corresponding protein mass is 5 Da above what is expected from this type of degradation pathway. This mass difference can most probably be attributed to deconvolution errors resulting from high background noise. In contrast, we never detected the lower mass protein in cell-free expressions.

Can incorporation leaves the spectroscopical properties of dEGFP virtually unchanged (Fig. 2A), however, its fluorescence efficiency is reduced. This may be due to the presence of the canavanine side chain in close vicinity to the fluorophore (see Fig. 2B).

In conclusion, we present the application of an easy to use cell-free expression system as a viable strategy to produce proteins with canavanine replacing arginine at all positions. Other amino acid substitutions could potentially be applied to this methodology. Our method enables the efficient production of target canavanine proteins while avoiding the limitations of in vivo systems, such as canavanine toxicity, or the strong dependency on mRNA sequence in strategies based on the single protein production (SPP) method.[16] The general applicability and usability make our method a potent tool for further research in canavanine toxicity and therapeutic application.

Acknowledgments

M.P.E. and N.B. acknowledge financial support from the FP7 EU-funded METACODE and SYNPEPTIDE Consortia. E.G.W., M.S. and A.O. acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) within the collaborative research center SFB 1027.

Supplementary data

Supplementary data (experimental details and supplementary information) associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmcl.2015.06.045.

References and notes

1. Richmond, M. H. *Bacteriol. Rev.* **1962**, *26*, 398.
2. Rosenthal, G. A. Q. *Rev. Biol.* **1977**, *52*, 155.
3. Rosenthal, G. a.; Dahlman, D. L. *J. Agric. Food Chem.* **1991**, *39*, 987.
4. Thomas, D. A.; Rosenthal, G. A.; Gold, D. V.; Dickey, K. *Cancer Res.* **1986**, *46*, 2898.
5. Liaudet, L.; Feihl, F.; Rosselet, A.; Markert, M.; Hurni, J. M.; Perret, C. *Clin. Sci. (Lond.)* **1996**, *90*, 369.
6. Bence, A. K.; Crooks, P. a. *J. Enzyme Inhib. Med. Chem.* **2003**, *18*, 383.
7. Bence, A. K.; Worthen, D. R.; Adams, V. R.; Crooks, P. *Drugs* **2002**, *13*, 313.
8. Akaogi, J.; Barker, T.; Kuroda, Y.; Nacionales, D. C.; Yamasaki, Y.; Stevens, B. R.; Reeves, W. H.; Satoh, M. *Autoimmun. Rev.* **2006**, *5*, 429.
9. Richmond, M. H. *Biochem. J.* **1959**, *73*, 261.
10. Kruse, P. F.; White, P. B.; Carter, H. A. *Cancer Res.* **1959**, *19*, 122.
11. Schachtele, C. F.; Rogers, P. *J. Mol. Biol.* **1965**, *14*, 474.
12. Murphy, E. C.; Kopchick, J. J.; Watson, K. F.; Arlinghaus, R. B. *Cell* **1978**, *13*, 359.
13. Murphy, E. C.; Arlinghaus, R. B. *J. Virol.* **1980**, *33*, 954.
14. Cavarelli, G. E. J. *Aminoacyl-tRNA Synthetases*; Landes Biosciences: Austin, 2003.
15. Budisa, N. *Angew. Chem., Int. Ed.* **2004**, *43*, 6426.
16. Ishida, Y.; Park, J. H.; Mao, L.; Yamaguchi, Y.; Inouye, M. *J. Biol. Chem.* **2013**, *288*, 7564.
17. He, M. *New Biotechnol.* **2008**, *25*, 126.
18. Zubay, G. *Annu. Rev. Genet.* **1973**, *7*, 267.
19. Shin, J.; Noireaux, V. *J. Biol. Eng.* **2010**, *4*, 8.
20. Singh-Blom, A.; Hughes, R. a.; Ellington, A. D. *J. Biotechnol.* **2014**, *178*, 12.
21. Oh, S.-J.; Lee, K.-H.; Kim, H.-C.; Catherine, C.; Yun, H.; Kim, D.-M. *Bioprocess. Eng.* **2014**, *19*, 426.
22. Kihara, H.; Prescott, J. M.; Snell, E. E. *J. Biol. Chem.* **1955**, *217*, 497.
23. Kalyankar, G. D.; Ikawa, M.; Snell, E. E. *J. Biol. Chem.* **1958**, *233*, 1175.