Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 17 (2007) 5548–5551

Evaluation of protein farnesyltransferase substrate specificity using synthetic peptide libraries

Amanda J. Krzysiak,[a,†] Sarah A. Scott,[a,†] Katherine A. Hicks,[b] Carol A. Fierke[b] and Richard A. Gibbs[a,*]

[a]*Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, 575 Stadium Mall Drive, West Lafayette, IN 47907, USA*
[b]*Departments of Chemistry and Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA*

Received 19 July 2007; revised 9 August 2007; accepted 13 August 2007
Available online 16 August 2007

Abstract—Farnesylation, catalyzed by protein farnesyltransferase (FTase), is an important post-translational modification guiding cellular localization. Recently predictive models for identifying FTase substrates have been reported. Here we evaluate these models through screening of dansylated-GCaaS peptides, which also provides new insights into the protein substrate selectivity of FTase.

Farnesylation is a post-translation modification that tags proteins with a farnesyl (C_{15}) isoprenoid supplied by farnesyl diphosphate (FPP) for the purpose of locating the protein to cellular membranes. These lipidated protein substrates are modified by farnesyltransferase (FTase) on their C-terminal FTase recognition sequence called the CaaX box. When it was found that farnesylation occurs on oncogenic Ras proteins,[1] FTase became a chemotherapeutic drug target.[2] Although initially developed based on a simple paradigm where they would target Ras-driven tumors, farnesyltransferase inhibitors (FTIs) have proven to work via a complex mechanism, and their activity is now attributed to the perturbation of a number of cellular proteins.[2,3]

The complex and unexpected biology observed with FTIs has made a precise definition of the set of farnesylated proteins in a human cell critically important. It is not known how many proteins in the cell are farnesylated or what are the critical targets of FTIs. Early biochemical studies of Brown and Goldstein[4] and the Merck group[5] demonstrated that tetrapeptides bearing a cysteine, two amino acids, and the appropriate X residue are farnesylated and serve as the minimum substrate for FTase recognition. Recent modeling studies have provided predictions for FTase Ca_1a_2X box specificity and thus its protein substrates.[6,7] These models are only predictive and require additional investigation[8,9] to determine cellular protein farnesylation. Using traditional biological tools (radiolabeling and/or Western blot analysis), it would be time-consuming to confirm the cellular farnesylation of these hypothetical FTase substrates. Therefore, a screening approach to validate that FTase accepts and modifies the minimal substrate Ca_1a_2X boxes of a select group of these proteins would be useful (Fig. 1).

As part of our laboratory's investigation into FTase specificity, we have synthesized a library of Dansyl-GCa_1a_2S pentapeptides representing FTase substrate candidates. The sequences were identified from a Swiss-prot database search for carboxyl-terminal Ca_1a_2S boxes. Sequences were chosen to represent (a) biologically important farnesylated proteins, and (b) interesting and diverse 'aa' amino acid sequences. In view of the current interest in models to define FTase substrate specificity, we now report the substrate ability of these Ca_1a_2S peptides as an experimental test of these models.

The 27 member Dansyl-GCa_1a_2S-OH library was synthesized on Wang resin primarily in an automated fashion using a standard Fmoc peptide chemistry, with HBTU/HOBt coupling and piperidine/DMF deprotection (Supplementary data). The resin-bound CaaS tetrapeptide was capped by coupling with Dansyl-Gly-OH, followed by cleavage from the resin and side-chain

Keywords: Farnesyl; Protein modification; Protein farnesyltransferase; Peptide synthesis; Fluorescence assay.
* Corresponding author. E-mail: rag@pharmacy.purdue.edu
† These authors contributed equally to this work.

0960-894X/$ - see front matter
doi:10.1016/j.bmcl.2007.08.024

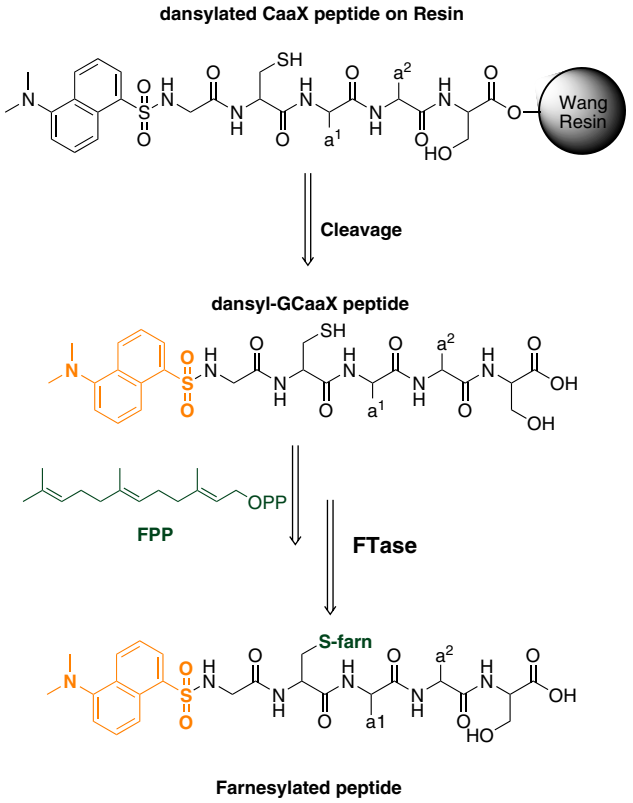

Figure 1. Analysis of Dansyl-GCaaX peptides through screening.

deprotection (90% CF_3COOH, 5% iPr_3SiH, and 5% H_2O). The library was successfully synthesized with yields for each member of the library ranging from 70 to 100 mg (66–95% overall yield). The purity of each of the peptide sequences was >70%, as confirmed by RP-HPLC analysis. The identity of all peptides was also confirmed by ESI-MS.

The FTase substrate activity of dansylated pentapeptide CaaX boxes can be measured through a fluorescence-based assay,[10] in a 96-well plate format.[11] Briefly, 3 μM dansylated-CaaX peptide and 1–9 μM FPP are combined, farnesylation is initiated by addition of recombinant mammalian FTase (0.05 μM), and the increase in fluorescence intensity is measured at 485 and 535 nm emission (Supplementary data). To confirm the farnesylation of the dansylated-CaaX peptides, HPLC analysis was performed for each of the dn-GCaaS peptide reactions.[11]

The results from the screening of the dn-GCaaS peptides are summarized in Table 1. The peptides are presented in descending order of reactivity. Of the 27 peptides screened, 24 were found to be substrates for FTase by fluorescence screening and, in 20 cases, by HPLC analysis (Supplementary data). These peptides did vary widely (~150-fold) in their ability to be farnesylated by FTase. Of the 27 dn-GCaaS boxes screened, 13 were either known substrates, or hypothesized FTase substrates according to the structural analysis-derived 'Beese model' developed by Reid, Casey, and Beese.[6] All of the CaaS sequences representing known farnesylated proteins are substrates in our in vitro system, providing support for its use in evaluating FTase substrate selectivity. Note that 11 CaaS sequences whose farnesylation was neither known nor hypothesized were also substrates for FTase. Seven of the 11 sequences are relatively poor substrates, exhibiting less than ten percent of the fluorescence change of dn-GCVLS, and in three cases no product was detected by HPLC. However, four of the sequences—CRQS, CVHS, CKSS, and CFSS—are well accepted by FTase, and are comparable in reactivity to known and hypothesized CaaS boxes. Three naturally-occurring CaaS sequences were not observable as substrates under our experimental conditions: CLRS, CIRS, and CFNS.

The screening data presented in Table 1 provide a qualitative ordering of the substrate ability of the CaaS sequences, and this provides information on the peptide substrate preference of FTase. In general, the reactivity order agrees with the previous observations and predictions that the FTase a_1 site is tolerant of many amino acid side chains, while the a_2 site has a preference for the larger aliphatic residues as well as methionine, phenylalanine, and tyrosine. We have now demonstrated that glutamine, serine, and histidine are also accepted at a_2. The a_1 position was previously not believed to play a major role in the selectivity for FTase, but our data suggest that the amino acid at a_1 may modulate reactivity, as indicated by the fourfold variation in the change in fluorescence of the five peptides of the general sequence dn-GCxIS (Table 1). The CRIS and CKIS peptides are the most effective substrates among the 27 examined, suggesting a preference for a positively-charged side chain in the a_1 position. This is not completely unexpected as the Beese model predicts that solution hydrogen bonding of the a_1 side chain to the solvent exposed enzyme is possible.

As with much else in the protein prenylation field, the discovery of farnesylation on Ras in the late 1980s drove the CaaX box meme, as a sequence where the central residues are aliphatic (H-Ras: CVLS; N-Ras: CVVM; K-Ras4B: CVIM).[2] However, it became quickly apparent that this is not strictly the case. A study of inhibitors based on the K-Ras CVIM tetrapeptide found significant flexibility in the CxIM series, but a strong preference for aliphatic residues with CVxM.[4] An early site-directed mutagenesis study from the Merck group also demonstrated that certain charged residues (lysine, but not glutamate) are accepted at the a_1 position, but not a_2.[5] Subsequently, a combinatorial screening approach defined the following optimal tetrapeptide FTase substrates: CKQQ and CKQM.[12] Based on extensive structural studies, Beese and coworkers[6] developed an FTase CaaX selectivity model, where a_1 is flexible and a_2 is limited to hydrophobic residues, and determined that 61 human proteins are substrates for FTase. Previously, individual CaaX peptides derived from FTase substrate candidates have been investigated as FTase substrates,[2] but here we report for the first time a more comprehensive investigation of the ability of a diverse set of human CaaX sequences to act as substrates. Our data suggest that CaaS boxes CRQS,

Table 1. Comparison of CaaX substrate predictions

CaaX	Protein	Rel. fluor. increase[a]	Beese prediction[c]	PRENbase[d]	PRENbase XCVLS[e]
CRIS	Protein phosphatase 1 reg (inhibitor) subunit 16B	886	H	+	
CKIS	Rab40A	574	H*	++	
CQTS	DNAJ	508	K	+	
CVLS	H-Ras	388	K	++	
CVIS	Transducin gamma subunit	331	K	++	
CTIS	Guanylate binding protein 1 interferon-inducible	325	K	++	
CFPS	CCNG2: cyclin G2	280	H	−−	−
CLIS	Phosphorylase kinase B	218	K	++	
CLVS	G protein-coupled receptor kinase 1	217	K	++	
CSVS	Inositol polyphosphate-5-phosphatase	180	H	+	
CRPS	WDTC1	171	H	+	
CTFS	Xylosyl protein beta 1,4-galactosyltransferase	166	H	−−	+
CRQS	Unknown protein fragment (Q15693)	121	NH	−	++
CVHS	Galectin-12	100	NH	−−	+
CKSS	Unknown fragment (q12814)	82	NH	−−	−
CFSS	Collagen type V alpha 3 subunit	60	NH	−	++
CDMS	Topo I binding Arg/Ser rich	42	H*	−−	+
CAKS	Rab38	33[b]	NH	−−	++
CEGS	Protein with 8 zinc finger domains	28	NH	−−	++
CQKS	Cytidine and deoxycytidylate deaminase	26[b]	NH	−	++
CVES	Unknowm fragment (Q29856)	20	NH	−−	−−
CPAS	Ribosomal protein L12	19	NH	−−	+
CGAS	Kinesin family member 22	6	NH	−−	+
CAES	G protein-coupled receptor 41	6[b]	NH	−−	++
CIRS	Rab3 interacting protein 1	0	NH	−	++
CLRS	Globoside alpha-1,3-*N*-acetylgalactosaminyltransferase 1	0	NH	−−	+
CFNS	Unknown Fragment (Q14922)	0	NH	−−	−

[a] Relative fluorescence increase was determined using a Perkin-Elmer Fusion plate reader (E_x 335/E_m 535; 30 min values reported).
[b] In these cases, the fluorimetric indication of farnesylated peptide product was not conclusively confirmed by HPLC analysis.
[c] Predictions derived from the FTase substrate model reported by Reid et al. K, known substrate; H, hypothesized substrate; H*, not reported as a substrate by Reid et al., but consistent with their rules for substrate ability; NH, not hypothesized to be a substrate.
[d] PRENbase prediction for the sequence XXXXXXXXCaaS, as described in the text.
[e] PRENbase prediction for the sequence XXXXXXXXCVLS, as described in the text.

CVHS, CKSS, and CFSS are farnesylated and proteins bearing these C-terminal sequences may be farnesylated in the cell (vide infra). These dnGCa_1a_2X peptides exhibit a_1 and a_2 residues that differ from the commonly accepted Ca_1a_2X box model, where a_1 and a_2 are aliphatic residues. While it was clear from earlier studies that there are few substrate constraints on the a_1 residue, it is surprising, both from known prenylated proteins and from the structure of the FTase a_2 binding pocket, that certain polar residues are accepted in the a_2 position.

Maurer-Stroh et al.[13] developed a database tool (PRENbase) using sequence conservation across species and other information to predict human proteins that are likely FTase substrates. PRENbase evaluates the 12 C-terminal residues of a protein and provides a score from ++ (prenylation is very likely) to −− (prenylation is very unlikely). These values are reported in Table 1, along with scores where the CaaX box of the C-terminal sequence was mutated in silico to CVLS, as a control for the effects of the sequence upstream of the CaaS box on FTase recognition and CaaS accessibility. The PRENbase analyses of our 27 selected CaaS proteins indicated that ten of the proteins are likely farnesylation candidates, including the six known farnesylated proteins in the set. Thus, PRENbase is significantly more conservative than the Beese model. While some of this disparity is due to the analysis of the eight residues 'upstream' of the CaaX box, there is also an element of conservative analysis of the CaaS motif itself. Note that the xylosyl transferase sequence (CTFS), that is not predicted to be a substrate, is predicted to be one when the CTFS CaaX box is 'mutated' to CVLS. The CTFS sequence is acceptable in the Beese model, and a good substrate in our experimental Ca_1a_2S screen.

Another intriguing divergence between the models involves cyclin G2. The Beese model predicts that its CaaX box CFPS is a substrate, while it is disfavored as a substrate by PRENbase (although this is even maintained with the CVLS 'mutant'). Our in vitro screen indicates that CFPS is a very good FTase substrate. Our data, combined with the PRENbase analysis of the CVLS 'mutants' to provide information of the suitability of the 'upstream' sequence, indicate that an uncharacterized gene product (Q15693) and two unexpected proteins—galectin-12 and collagen type five alpha three subunit—are candidates for farnesylation. The farnesylation of a subunit of the extracellular protein collagen seems unlikely, but galectin-12 is an intriguing, newly reported intracellular protein, whose lipid modification status has not yet been investigated.[14]

The screening of the dn-GCaaS library of peptides has resulted in the identification of 23 substrate CaaX sequences. Our results essentially validate the model for

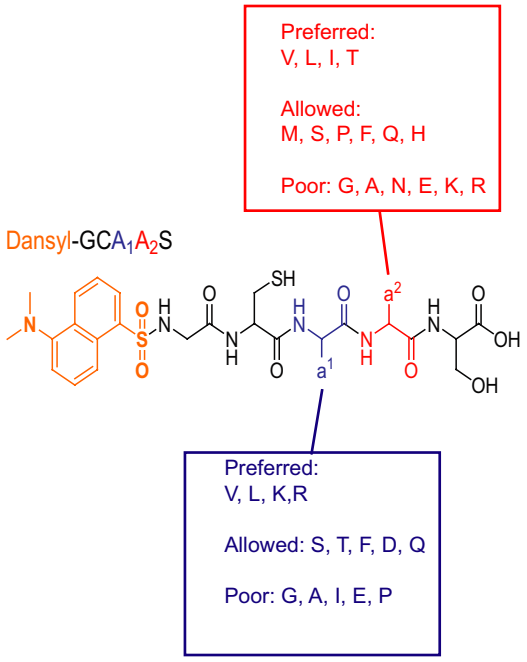

Figure 2. Summary of FTase amino acid specificity.

FTase substrate prediction published by Reid et al., but also indicate that additional a_2 residues, including Gln, Ser, and His, are well-tolerated by FTase. Furthermore, these data shed light on the importance of both a_1 and a_2 positions' residues on determining substrate ability for peptide sequences. In particular, our results highlight the preference for Lys and Arg at the a_1 position (see summary in Fig. 2). With the capability to perform peptide synthesis and rapid biochemical evaluation, CaaX screening provides an important tool for identifying FTase substrate proteins. Together with the previously reported structurally-based model database-derived sequence analysis (e.g., PRENbase), biochemical peptide screening will provide a valuable guide to future proteomic and other biological searches for farnesylated proteins.

Acknowledgments

This work was supported by NIH Grants CA78819 (R.A.G.), GM40602 (C.A.F.), and P30CA21368 (Purdue Cancer Center), and by Purdue Research Foundation fellowships (A.J.K. and S.A.S.).

Supplementary data

Experimental procedures for solid phase peptide synthesis and the spectrofluorimetric fluorescence assay, and HPLC traces for reaction of each member of the dnGCaaS library with FPP in the presence of FTase. Supplementary data associated with this article can be found in the online version. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2007.08.024.

References and notes

1. Konstantinopoulos, P. A.; Karamouzis, M. V.; Papavassiliou, A. G. *Nat. Rev. Drug Disc.* **2007**, *6*, 541.
2. Basso, A. D.; Kirschmeier, P.; Bishop, W. R. *J. Lipid Res.* **2006**, *47*, 15.
3. Sebti, S. M.; Der, C. J. *Nat. Rev. Cancer* **2003**, *3*, 945.
4. Reiss, Y.; Stradley, S.; Gierasch, L.; Brown, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 732.
5. Moores, S. L.; Schaber, M. D.; Mosser, S. D.; Rands, E.; O'Hara, M. B.; Garsky, V. M.; Marshall, M. S.; Pompliano, D. L.; Gibbs, J. B. *J. Biol. Chem.* **1991**, *266*, 14603.
6. Reid, T. S.; Terry, K. L.; Casey, P. J.; Beese, L. S. *J. Mol. Biol.* **2004**, *343*, 417.
7. Maurer-Stroh, S.; Eisenhaber, F. *Genome Biol.* **2005**, *6*, R55.
8. Kho, Y.; Kim, S. C.; Jiang, C.; Barma, D.; Kwon, S. W.; Chemg, J.; Jaunbergs, J.; Weinbaum, C.; Tamanoi, F.; Falck, J.; Zhao, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12479.
9. Suzuki, T.; Ito, M.; Ezure, T.; Shikata, M.; Ando, E.; Utsumi, T.; Tsunasawa, S.; Nishimura, O. *Proteomics* **2007**, *7*, 1942.
10. Cassidy, P. B.; Dolence, J. M.; Poulter, C. D. *Methods Enzymol.* **1995**, *250*, 30.
11. Krzysiak, A. J.; Rawat, D. S.; Scott, S. A.; Pais, J. E.; Handley, M.; Harrison, M. L.; Fierke, C. A.; Gibbs, R. A. *ACS Chem. Biol.* **2007**, *2*, 385.
12. Boutin, J. A.; Marande, W.; Petit, L.; Loynel, A.; Desmet, C.; Canet, E.; Fauchere, J. L. *Cell. Signalling* **1999**, *11*, 59.
13. Maurer-Stroh, S.; Koranda, M.; Benetka, W.; Schneider, G.; Sirota, F. L.; Eisenhaber, F. *PLoS Comput. Biol.* **2007**, *3*, e66.
14. Yang, R. Y.; Liu, F. T. *Cell Mol. Life Sci.* **2003**, *60*, 267.