doi:10.1016/j.jmb.2004.08.056 J. Mol. Biol. (2004) 343, 417–433

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Crystallographic Analysis of CaaX Prenyltransferases Complexed with Substrates Defines Rules of Protein Substrate Selectivity

T. Scott Reid[1]†, Kimberly L. Terry[1]†, Patrick J. Casey[2] and Lorena S. Beese[1]*

[1]*Department of Biochemistry Duke University Medical Center, Durham, NC 27710 USA*

[2]*Department of Pharmacology and Cancer Biology, Duke University Medical Center Durham, NC 27710, USA*

**Corresponding author*

Post-translational modifications are essential for the proper function of many proteins in the cell. The attachment of an isoprenoid lipid (a process termed prenylation) by protein farnesyltransferase (FTase) or geranylgeranyltransferase type I (GGTase-I) is essential for the function of many signal transduction proteins involved in growth, differentiation, and oncogenesis. FTase and GGTase-I (also called the CaaX prenyltransferases) recognize protein substrates with a C-terminal tetrapeptide recognition motif called the Ca_1a_2X box. These enzymes possess distinct but overlapping protein substrate specificity that is determined primarily by the sequence identity of the Ca_1a_2X motif. To determine how the identity of the Ca_1a_2X motif residues and sequence upstream of this motif affect substrate binding, we have solved crystal structures of FTase and GGTase-I complexed with a total of eight cognate and cross-reactive substrate peptides, including those derived from the C termini of the oncoproteins K-Ras4B, H-Ras and TC21. These structures suggest that all peptide substrates adopt a common binding mode in the FTase and GGTase-I active site. Unexpectedly, while the X residue of the Ca_1a_2X motif binds in the same location for all GGTase-I substrates, the X residue of FTase substrates can bind in one of two different sites. Together, these structures outline a series of rules that govern substrate peptide selectivity; these rules were utilized to classify known protein substrates of CaaX prenyltransferases and to generate a list of hypothetical substrates within the human genome.

Keywords: crystal structure; geranylgeranyltransferase; farnesyltransferase; Ras; G protein

Introduction

Post-translational modifications, including glycosylation, phosphorylation, proteolysis and lipidation, can dramatically affect protein activity and interactions with substrates, cofactors and other proteins. The post-translational attachment of an isoprenoid lipid (prenylation) is required for the function of most GTP-binding regulatory proteins, including members of the Ras superfamily, several protein kinases and phosphatases, and a variety of proteins involved in nuclear integrity and centromere function.[1] Protein prenylation and further modification by enzymes at the endoplasmic reticulum can have dramatic functional consequences by permitting association with the cell membrane and encouraging protein–protein interactions with other regulatory molecules.[1] Understanding the mechanism by which proteins are recognized and a specific isoprenoid lipid is attached is therefore a critical aspect in the assessment of protein function within the cell. Furthermore, characterization of protein prenylation patterns is important for understanding both the therapeutic utility and the potentially adverse effects of treatment with protein prenyltransferase inhibitors,[2] several of which are being evaluated in phase II/III clinical trials as cancer therapeutics.[3]

† T.S.R. and K.L.T. contributed equally to this work.

Abbreviations used: rFTase, rat farnesyltransferase; hFTase, human farnesyltransferase; GGTase-I, geranylgeranyltransferase type I; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate.

E-mail address of the corresponding author: lsb@biochem.duke.edu

0022-2836/$ - see front matter

The three known enzymes that catalyze protein prenylation are the two CaaX prenyltransferases, protein farnesyltransferase (FTase) and protein geranylgeranyltransferase type I (GGTase-I), which recognize a broad range of protein substrates, and Rab GGTase, which specifically recognizes members of the Rab subfamily of G proteins.[1] FTase and GGTase-I catalyze the respective transfer of a 15 carbon or a 20 carbon isoprenoid lipid from farnesyl diphosphate (FPP) or geranylgeranyl diphosphate (GGPP) (Figure 1) to proteins, short peptides, or tetrapeptides with a C-terminal Ca_1a_2X motif. The Ca_1a_2X box is defined by an invariant cysteine residue(C) to which the lipid is attached, two typically aliphatic residues (a_1a_2), and the C-terminal residue (X) that contributes to substrate specificity.[4] Steady-state kinetic studies indicate that FTase prefers substrates with a Ser, Gln, Met or Ala in the X position, whereas GGTase-I prefers Leu.[5–9] Although FTase and GGTase-I are very selective towards their respective substrates, there are examples of cross-utilization of protein substrates. Proteins with a Phe at the X residue position, including the teratocarcinoma oncoprotein TC21 (CVIF for the Ca_1a_2X motif), can be substrates for both enzymes.[9,10] The oncoprotein K-Ras4B (CVIM) is involved in approximately 30% of all human cancers[11] and is normally farnesylated *in vivo* by FTase. However, when FTase activity is compromised by treatment with specific inhibitors, K-Ras4B can be geranylgeranylated by GGTase-I.[12,13] Additionally, *in vitro* and cell-based studies have demonstrated that RhoB (CKVL), despite the C-terminal Leu, is a substrate for both GGTase-I and FTase.[14–16]

FTase and GGTase-I have similar global and active-site structures.[17,18] Structures of FTase and GGTase-I complexed with isoprenoid diphosphate analogs and cognate CVIM or CVIL peptides, respectively, reveal that the Ca_1a_2X motif adopts an extended conformation along one side of the funnel-shaped active site, and that peptide binding does not induce a change in the active-site structure.[18–20] However, these studies examined only substrate peptides containing two different residues at the X position. Since the identity of the C-terminal X residue alone can alter peptide binding affinity by at least an order of magnitude, some have hypothesized that different X residues may engender different binding modes of the Ca_1a_2X motif.[9,19] Consistent with this hypothesis, some small molecule inhibitors competitive with respect to the peptide substrate adopt different binding modes in FTase and GGTase-I.[21] Furthermore, sequence elements directly upstream of the Ca_1a_2X motif have been shown to alter peptide binding affinity (e.g. lysine-rich sequences near the Ca_1a_2X motif of K-Ras4B), suggesting that these upstream elements may affect the conformation the Ca_1a_2X motif adopts in the active site.[22,23]

To address these issues of peptide recognition by the CaaX prenyltransferases, we have determined structures of FTase and GGTase-I complexed with isoprenoid diphosphate analogs and a representative array of eight specific and cross-reactive peptide substrates, including sequences derived from the C terminus of the oncoproteins H-Ras and TC21, the signal transduction proteins RhoB, Rap2a and Cdc42, and the γ_2 subunit of heterotrimeric G-proteins. From these structures, we have deduced a series of rules that govern prenyltransferase peptide substrate specificity. These rules were used to generate a list of known and predicted FTase and GGTase-I substrates within the human genome, updating previous efforts in this field.[24]

Results

We determined structures of FTase in ternary complexes with an FPP analog (FPT-II, Figure 1) and peptides derived from the cognate substrates H-Ras (GCVLS, Figure 2(A)) and Rap2a (DDPTASACNIQ, Figure 2(B)), the cross-reactive substrate

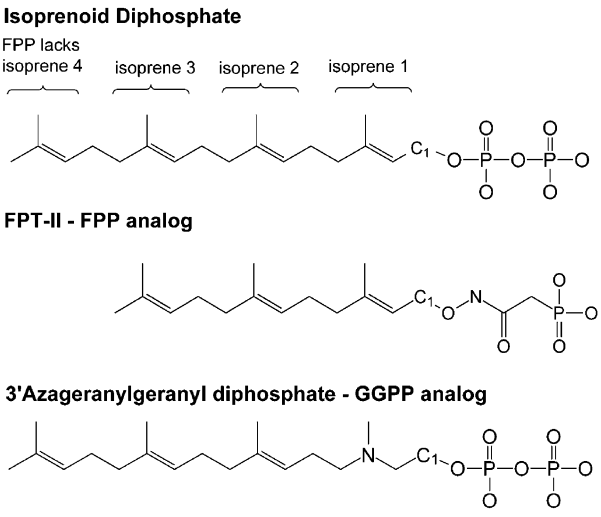

Figure 1. Chemical structures of isoprenoid diphosphates and non-hydrolyzable analogs.

TC21 (KKSKTKCVIF, Figure 2(C)), and the non-substrate Rap2b (TKCVIL). We determined structures of GGTase-I in ternary complexes with a GGPP analog (3′azaGGPP, Figure 1) and peptides derived from the cognate substrates Cdc42 splice isoform 2 (RRCVLL Ca_1a_2X motif) and the heterotrimeric G protein γ_2 subunit (FREKKFFCAIL), and the cross-reactive substrates RhoB (GCINCCKVL), K-Ras4B (KKKSKTKCVIM) and TC21 (Figure 2(D)). These nine structures were compared to previously determined structures of FTase in ternary complexes with K-Ras4B-derived peptides,[19,20] and GGTase-I in ternary complexes with a Rap2b chimera peptide (KKKSKTKCVIL).[18] The peptides were chosen because their Ca_1a_2X motifs are representative of the most common a_2 (Val, Leu and Ile) and X (Leu, Met, Phe, Gln, and Ser) residues, and for their importance in signal transduction and cancer biology. The four FTase and six GGTase-I peptide substrates are bound as a ternary complex with the non-hydrolyzable lipid analogs FPT-II[25] and 3′azaGGPP,[26] respectively. The two lipid analogs bind as observed previously, adopting a conformation similar to that of the respective lipid substrate.[18,20] The eight cognate and cross-reactive peptide substrates presented here bind with the Ca_1a_2X motif inserted into the peptide-binding site, and bind without altering the enzyme or isoprenoid diphosphate structure, consistent with previous structures with bound substrates, products and inhibitors.[17–29] The Ca_1a_2X motif of the non-substrate Rap2b does not adopt an ordered conformation within the FTase active site, consistent with the observation that this peptide is a very poor FTase substrate.

Conformation of the Ca_1a_2X peptide backbone

The Ca_1a_2X motifs of the four FTase and six GGTase-I peptide substrates adopt the same "extended" conformation along one side of the funnel-shaped active site (Figure 3(A)). The backbone of the Ca_1a_2X motif adopts a similar conformation in both FTase and GGTase-I; the only significant difference between FTase and GGTase-I peptide substrates is the positioning of the C-terminal X residue (Figure 3(B), discussed below). FTase appears to permit more flexibility within the Ca_1a_2X backbone: the four FTase and six GGTase-I substrates show an average backbone root-mean-square deviation (r.m.s.d.; as compared within each enzyme) of 0.93 Å and 0.29 Å, respectively. The Cys residue of the Ca_1a_2X motif, which is thought to bind as a thiolate, is located in the same position in all ten structures and coordinates the catalytic zinc ion.[30] Approximately 12 Å away from the Cys residue of the Ca_1a_2X motif, the C termini of all substrate peptides are anchored by direct and water-mediated hydrogen bonds with conserved Gln, Glu, His, and Arg residues. A shift in GGTase-I helix 4β relative to FTase necessitates an additional water molecule to form hydrogen bonds between the C terminus of GGTase-I peptide substrates and the conserved His residue.

Conformation of the a_1 and a_2 residues

The a_1 residues of the Ca_1a_2X motif (Val, Ala, Asn and Lys) are mostly exposed to the solvent. At the the a_1 position of Rap2a, an Asn side-chain forms a hydrogen bond with FTase, and in RhoB a Lys side-chain at the the a_1 position is disordered in GGTase-I. In GGTase-I, but not in FTase, the a_1 residue is additionally stabilized by a water-mediated hydrogen bond between the carbonyl moiety of the a_1 residue and His201α. Modeling studies using the Rap2a and γ_2 peptide substrates presented here as a template indicate that both FTase and GGTase-I can accept any amino acid at the a_1 position without steric hindrance or altering the extended conformation of the Ca_1a_2X motif. These modeling studies reveal that polar or charged a_1 residues could form direct or water-mediated hydrogen bonds with the enzyme.

For the pepide substrates presented here, all a_2 residues of the Ca_1a_2X motif form a direct hydrogen bond with an Arg residue conserved in the β subunits of FTase (R202β) and GGTase-I (R173β). The side-chains of the a_2 residues (Leu, Ile and Val) bind isosterically in the hydrophobic "a_2 binding pocket", defined by regions of the protein and isoprenoid that contact the a_2 residue (Figure 4, Table 1). Previous structural studies show that a Phe side-chain can bind in this space in FTase.[27] Additional modeling studies, again using the Rap2a and γ_2 peptide substrates as a template, indicate that other similarly shaped amino acid residues such as Pro, Thr, Tyr and Met can also be accommodated in the a_2 binding site of FTase and GGTase-I without steric clashes or significantly altering the conformation of the backbone of the Ca_1a_2X peptide.[27] Large residues such as Arg or Trp cannot fit into the a_2 pocket without major disruptions to the Ca_1a_2X motif backbone, and small amino acid residues such as Ala or Gly would leave unoccupied space. The hydrophobic nature of the a_2 pocket makes it unlikely that it could accept charged or very polar residues such as Lys, Glu, Asp, His, Gln, Asn, Cys or Ser. Peptide substrate recognition at the a_2 position of the Ca_1a_2X motif is therefore dominated by steric and hydrophobic interactions between the a_2 residue and its binding site (see Table 1 for a summary).

Binding of the C-terminal X residue

In GGTase-I, the C-terminal Leu, Met and Phe side-chains of the peptide substrates all bind isosterically in the hydrophobic "specificity pocket", defined by the residues and ligands that contact the X residue (Figure 3(B), Table 1). These three residues are stabilized entirely by hydrophobic interactions and van der Waals contact with the enzyme. Although the majority of GGTase-I substrates have Leu as the X residue, kinetic studies

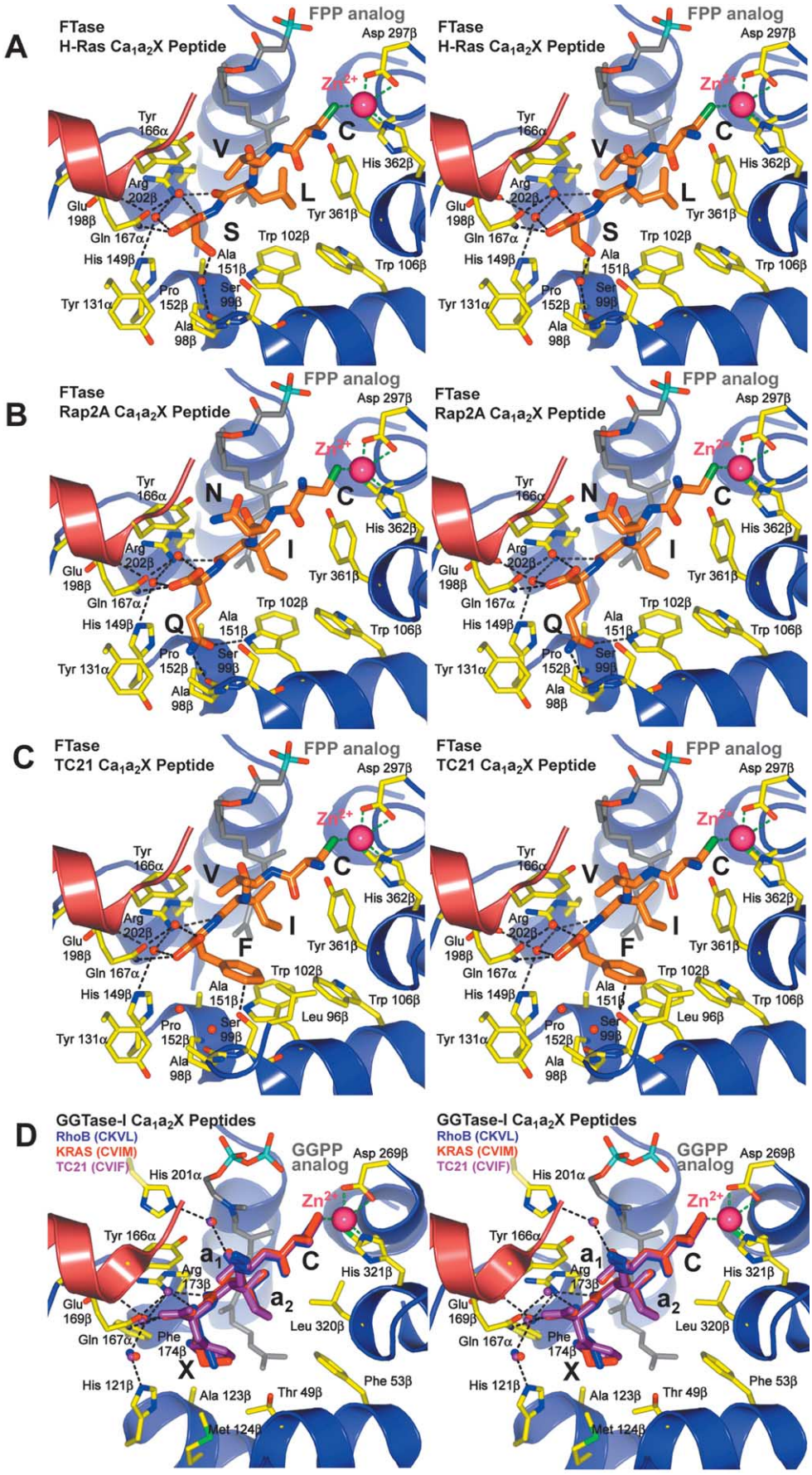

Figure 2. Ca_1a_2X peptide binding in FTase and GGTase-I. Stereo pairs of FTase and GGTase-I are all shown in the same orientation. Only the substrate peptide Ca_1a_2X motif, the lipid analog, the catalytic zinc ion and protein residues

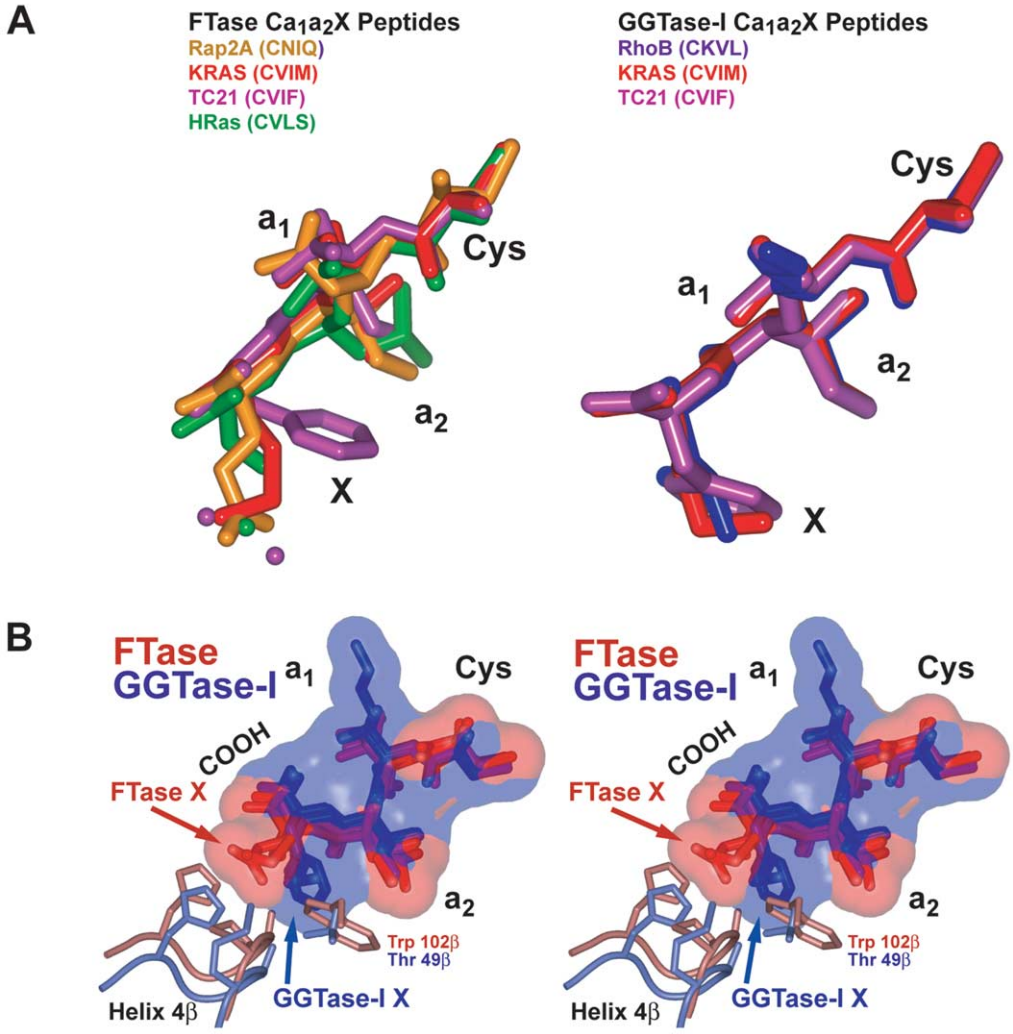

Figure 3. Comparison of Ca_1a_2X substrate binding in FTase and GGTase-I. (A) A superposition of four FTase substrate complexes (left) and three GGTase-I substrate complexes (right), shown in approximately the same orientation as Figure 2, illustrate that cognate and cross-reactive peptides adopt a common binding mode. In FTase, however, the C-terminal Phe residue of TC21 binds in a different pocket than the C-terminal Met, Gln and Ser residues. (B) A stereo pair of three FTase peptide complexes (K-Ras4B, H-Ras, Rap2a, and the corresponding van der Waals surfaces; red) and three GGTase-I peptide complexes (RhoB, K-Ras4B, TC21, and the corresponding van der Waals surfaces; blue). In all six structures the Ca_1a_2 portion of the Ca_1a_2X motif and carboxyl terminus bind isosterically. The X residues, however, have enzyme-specific binding pockets (see also Table 1). Differences between the X residue binding pockets are created primarily by a shift in helix 4β and a Trp/Thr difference between the two enzymes.

indicate that Ca_1a_2X peptides with Ile or Val in the X position can function as substrates for this enzyme.[6,9] Examples of proteins with these C-terminal residues include 2′3′-cyclic nucleotide 3′phosphodiesterase (CTII),[31] RhoJ/TCL (CSII) and the Cdc42 homolog Wrch-1 (CCFV).[32] Using the RhoB peptide as a template, Ile or Val can be modeled in the specificity pocket without steric clashes or repositioning the Ca_1a_2X motif backbone.[33] Larger amino acid residues such as Tyr, Trp and Arg, or Pro, which has restricted backbone flexibility, cannot be accommodated without disrupting the extended conformation of the Ca_1a_2X motif. The hydrophobic nature of the GGTase-I specificity pocket is expected to discriminate against polar or charged residues.

In FTase, the C-terminal Met, Gln and Ser X residues of the peptide substrates all bind isosterically in the specificity pocket (Figure 3(B), Table 1). Unlike GGTase-I peptide substrates, these three X residues form both van der Waals and electrostatic

involved in ligand coordination are shown. (A)–(C) FTase complexed with FPT-II (gray) and H-Ras (CVLS), Rap2a (CNIQ), and TC21 (CVIF), respectively. (D) Superposition of GGTase-I complexed with 3′azaGGPP (gray) and RhoB (CKVL, blue), K-Ras4B (CVIM, red) and TC21 (CVIF, purple). In all structures, the Ca_1a_2X peptide adopts an extended conformation with the cysteine thiolate coordinated by the zinc ion (magenta). Carbonyl oxygen atoms and the C terminus of the Ca_1a_2X sequence make water-mediated and direct hydrogen bonds with conserved side-chains in both the α (red) and β (blue) subunits.

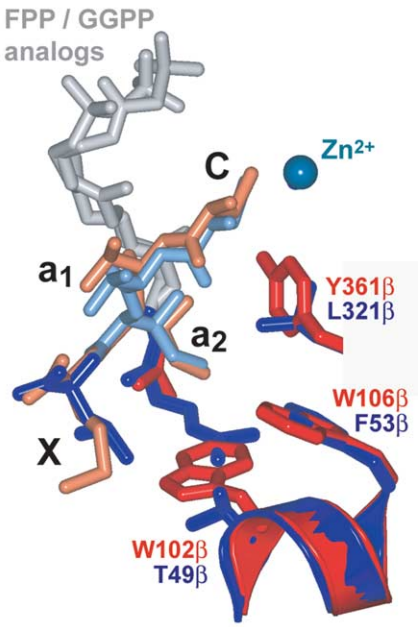

Figure 4. Comparison of the a_2 binding site in FTase and GGTase-I. Superposition of FTase and GGTase-I shows residues that interact with the a_2 residue of the Ca_1a_2X motif. Regions of the FTase ternary complex forming the a_2 binding site (enzyme residues Trp102β, Trp106β and Tyr361β and isoprene 3) are colored red. Corresponding regions of the GGTase-I ternary complex (enzyme residues Thr49β, Phe53β and Leu321β, isoprenes 3 and 4, and substrate peptide X residue) are colored blue. Portions of the FPP and GGPP analogs (gray) and of the CVIM (pink) and CVIL (light blue) peptides not forming the a_2 site are shown.

interactions with the enzyme. The C-terminal Met (K-Ras4B) is oriented such that the thioether can accept a weak hydrogen bond from the Ser 99β hydroxyl group (sulfur-to-oxygen distance 3.2 Å).[34] The C-terminal Gln (Rap2a) donates a hydrogen bond to the 98β carbonyl oxygen atom and an adjacent buried water molecule, and accepts a hydrogen bond from Trp102β $N_{\varepsilon 1}$. The smaller Ser (H-Ras) is accompanied by a water molecule hydrogen bonded to the Ser hydroxyl group and the adjacent Ala 98β carbonyl oxygen atom. A comparison of the binding of these three substrates reveals the FTase specificity pocket possesses two distinct subsites: one that can donate a hydrogen bond, and one that can accept a hydrogen bond (Figure 5).

A surprising result was obtained when the binding of the TC21 peptide substrate in FTase was examined. Although the FTase specificity pocket is shaped incorrectly to accommodate the C-terminal Phe of the TC21 substrate, this peptide binds to FTase by placing the side-chain of the X residue in an alternative hydrophobic binding site (Figure 2(C), Table 1), and two solvent molecules are now observed in the specificity pocket. The phenyl ring of the Phe is stabilized by face-on-face aromatic stacking interactions with Trp102β,[35,36] and a weak hydrogen bond with the Ser 99β hydroxyl group (carbon-to-oxygen distance of 2.96 Å).[34] In addition to accommodating Ca_1a_2X substrates terminating in a Phe residue, this secondary binding site could be utilized by atypical Ca_1a_2X peptides (see below).

Although the majority of known mammalian FTase substrates have Met, Ser or Gln X-residues (Figure 7(C)), kinetic studies indicate that Thr, Cys or Ala can function as substrates.[6,9] Examples of

Table 1. Comparison of Ca_1a_2X binding sites in FTase and GGTase-I

Residue		FTase	GGTase-I
Upstream residues	Binding site Accepts	Enzyme surface (variable) No apparent restrictions	Enzyme surface (variable) No apparent restrictions
c	Binding site Accepts	Coordinates catalytic zinc Cys only	Coordinates catalytic zinc Cys only
a_1	Binding site Accepts	Solvent accessible H-bonding to enzyme possible No apparent restrictions	Solvent accessible H-bonding to enzyme possible No apparent restrictions
a_2	Binding site Accepts	Trp 102β, Trp 106β, Tyr 361β, FPP isoprene 3 Val, Ile, Leu, Phe, Tyr, Thr, Met preference for Ile, Val	Thr 49β, Phe 53β, Leu 320β, GGPP isoprene 3-4, X residue Val, Ile, Leu, Phe, Tyr, Thr, Met preference for Ile, Leu
x	Binding site 1 "specificity pocket" Accepts Binding site 2 Accepts	Tyr 131α, Ala 98β, Ser 99β, Trp102β, His 149β, Ala 151β Pro 152β Met, Gln—polar interactions Ser, Ala, Thr, Cys—with buried water Leu 96β, Ser 99β, Trp 102β, Trp 106β, Ala 151β, 3rd FPP isoprene, a_2 residue Phe, possibly Leu, Asn, or His	Thr 49β, His 121β, Ala 123β, Phe174β, 4th GGPP isoprene, a_2 residue Leu, Ile, Val, Phe Not observed Not observed

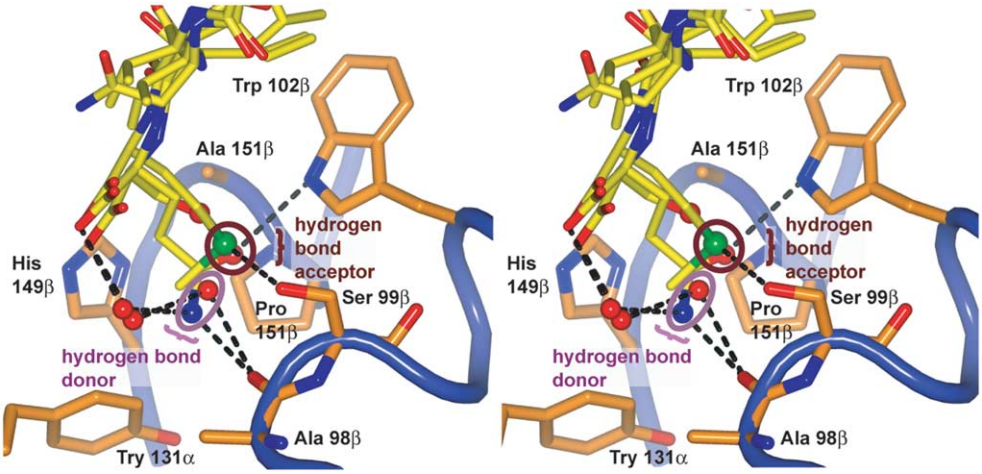

Figure 5. FTase C-terminal X specificity pocket. Stereo pair shows a superposition of K-Ras4B (CVIM), Rap2a (CNIQ) and H-Ras (CVLS) peptides bound to FTase. Only the Ca_1a_2X motif of substrate peptides and the residues in FTase that interact with the X-residue of the Ca_1a_2X peptide, the "specificity pocket", are shown. A comparison of the binding of these three substrates reveals the FTase specificity pocket possesses two distinct subsites: one that can donate a hydrogen bond, and one that can accept a hydrogen bond. The S^{δ} atom of Met (K-Ras4B; green sphere) and $O^{\varepsilon 1}$ atom of Gln (Rap2a; red sphere) are hydrogen bond acceptors (circled in brown). The C-terminal $N^{\varepsilon 2}$ atom of Gln (blue sphere), and the water molecule that accompanies the Ser of H-Ras (red sphere, see the text), are both hydrogen bond donors (circled in purple). The C-terminal residue of FTase substrates must complement the shape of the specificity pocket, and must place hydrogen bond donor or acceptor atoms in specific locations.

proteins with these C-terminal residues include the human proteins RhoQ (CLIT),[37] the prostacyclin receptor (CSLC),[38] and cerebral protein-5 (CVLA).[39] Using the H-Ras structure as a template, Cys, Ala or Thr can be modeled in the specificity pocket without steric clashes or repositioning the Ca_1a_2X motif backbone. As with Ser, these three X-residues would be accompanied by a buried water molecule that fills unoccupied space in the specificity pocket. The FTase specificity pocket is shaped incorrectly to accommodate other residues with a van der Waals shape similar to that of Gln or Met, including His, Asn, Leu, Phe and Ile. Larger residues such as Tyr, Trp and Arg, or Pro, which has restricted backbone flexibility, cannot be accommodated without disrupting the extended conformation of the Ca_1a_2X motif. Although Leu or the polar residues Asn and His (at pH 7.5) cannot be modeled in the FTase specificity pocket without steric clashes, Ca_1a_2X peptides with these residues in the X position (e.g. peptides derived from RhoB, with a C-terminal Leu) can function as FTase substrates.[6,9,14,16] Using the TC21 peptide as a template, a C-terminal Asn, His or Leu can be placed in the secondary binding site where Phe binds without steric clashes or distorting the Ca_1a_2X backbone. As with TC21, Ca_1a_2X sequences terminating in these three residues would be predicted to bind with two water molecules occupying the specificity pocket. Furthermore, although the charged residues Glu or Asp could bind in a manner similar to that of Gln or Asn, respectively, there are no residues available to stabilize the negative charges of these two residues, so such binding would be expected to be of low affinity.

Effect of peptide sequence upstream of the Ca_1a_2X motif on binding

In FTase, the upstream sequences of Rap2a (DDPTASA, Figure 6(A) and (B)) and K-Ras4B (KKKSKTK) bind along the rim of the peptide-binding site, stabilized primarily by direct and water-mediated hydrogen bonds with the surface of the enzyme. Although these two upstream sequences each adopt significantly different conformations, the Ca_1a_2X motifs of the two peptides adopt the same extended conformation discussed above (Figure 6(C)). In all GGTase-I substrate complexes, the sequences upstream of the Ca_1a_2X motif (KKKSKTK for K-Ras4B, FREKKFF for γ_6 and GCINC for RhoB) are partly disordered, with electron density disappearing two or three residues beyond the Cys residue of the Ca_1a_2X motif. Regardless of the upstream sequence, all Ca_1a_2X substrates adopt the same conformation in the GGTase-I active site. Indeed, these results show that the Ca_1a_2X motif of RhoB (CKVL) binds identically to GGTase-I whether the upstream sequence is native (GCINC), or contains the polylysine sequence derived from K-Ras4B (KKSKTK).

Discussion

A CaaX prenyltransferase peptide specificity model

A comparison of the 11 structures discussed here suggests a series of rules that govern peptide substrate specificity for the CaaX prenyltransferases (Table 1). Substrate peptide binding clearly does not

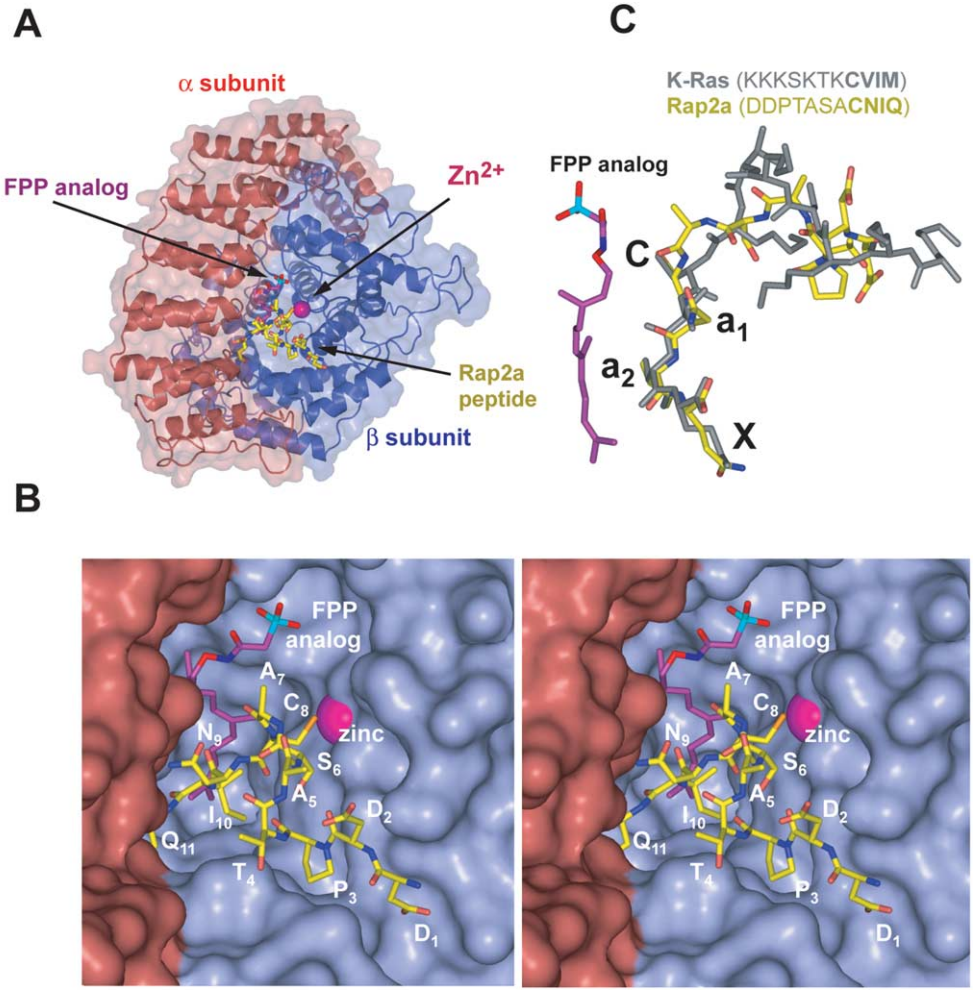

Figure 6. Peptide substrate binding upstream of the Ca_1a_2X motif. (A) Overall structure of FTase complexed with Rap2a (DDPTASACNIQ) and a FPP analog (purple), shown as a surface representation with the α subunit colored red and the β subunit colored blue. (B) Close-up of the Rap2a structure, shown as a stereo pair in the same orientation as in A. The Rap2a upstream sequence binds along the rim of the active site, stabilized by hydrogen bonds and van der Waals contacts with the enzyme. Lys 164α was omitted for clarity (side-chain atoms). (C) Superposition of the FTase Rap2a (yellow) and K-Ras4B (KKKSKTKCVIM, gray) structures. Only the substrate peptide and FPP analog (purple) are shown. Although the conformation upstream of the Ca_1a_2X motif differs, both Ca_1a_2X motifs adopt the same "extended" conformation.

induce a change in the peptide-binding site. The Ca_1a_2X motif of substrate peptides bind in the rigid peptide-binding site in an extended conformation, anchored at the Cys residue by zinc-coordination and at the C-terminal end by direct and water-mediated hydrogen bonds. These two fixed anchor points discriminate against peptides that are too long or too short, or that lack a cysteine residue at the correct position. In this conformation, any amino acid residue can be accommodated at the a_1 position, as has been hypothesized.[5,18,20] An analysis of proteins shown to be prenylated illustrates that there is, in fact, little preference shown by either enzyme for specific amino acid residues in the a_1 position (Figure 7(A)). Unexpectedly, however, we have found that polar or charged amino acids at the a_1 position, such as Asn in Rap2a, can form direct or water-mediated hydrogen bonds with the enzyme and thereby enhance binding affinity. The a_2X residues are buried in the active site, and substrate specificity for cognate and cross-reactive Ca_1a_2X motif peptides is determined by steric and electrostatic complementarity between the a_2 residue and the a_2 binding pocket, and the X residue and the specificity pocket. Although sequence upstream of the Ca_1a_2X motif can enhance peptide substrate affinity through electrostatic interactions with the protein surface, this sequence does not influence the conformation of the Ca_1a_2X motif itself, and should not permit Ca_1a_2X peptides that violate the general rules of specificity to function as proper substrates. There is no indication that either GGTase-I or FTase has specific sites that bind upstream sequences in a generalized conformation, in contrast to the binding site for the Ca_1a_2X motif itself. Indeed, upstream sequences of Ca_1a_2X proteins are highly variable, and this region has been dubbed the hypervariable region in Ras superfamily members.[11]

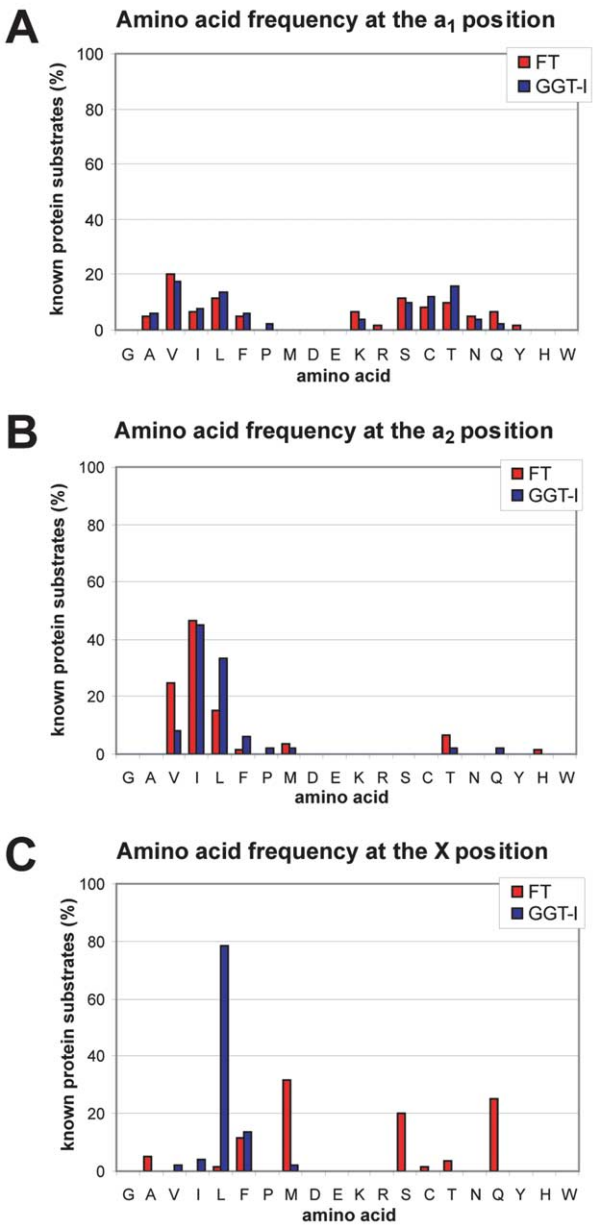

Figure 7. Amino acid sequence preferences within the Ca_1a_2X motif. (A)–(C) Chart depicting Ca_1a_2X motif a_1, a_2, and X residue identity, respectively, for all proteins demonstrated to be prenylated by FTase (red) and GGTase (blue).

Sequence preference at the a_2 position

The a_2 binding sites of FTase and GGTase-I have unique steric and aromatic properties, suggesting that the a_2 residue may influence CaaX prenyltransferase peptide recognition.[20] Overall, the structures presented here suggest that accommodation of the a_2 residue is primarily restricted to Val, Ile, Leu, Phe, Tyr, Pro, Thr, and Met, consistent with the previous studies of FTase peptide selectivity in mammalian and yeast systems.[5,40] The a_2 binding site in FTase is smaller and has more aromatic character than the corresponding binding site in GGTase-I (Table 1, Figure 4). To investigate whether these differences influence peptide selectivity, we undertook an analysis of all proteins known to be prenylated for amino acid preferences at the a_2 position of the Ca_1a_2X motif (Figure 7(B)). As expected, the small aliphatic amino acid residues Leu, Val and Ile predominate. There do, however, appear to be subtle preferences within this group: for FTase, Val is observed more often than Leu, whereas in GGTase-I Leu is seen more commonly than Val; Ile, however, appears equally in substrates for both enzymes. This is consistent with the differently sized binding sites for the a_2 residue in FTase and GGTase-I, and suggests that the a_2 residue may influence peptide substrate preferences. This comparison, however, does not reveal any a_2 residue that determines absolutely whether a Ca_1a_2X motif is a GGTase-I or FTase substrate (as is the case with the X residue within the Ca_1a_2X motif).

Sequence preference at the C-terminal X position

The X residue of the Ca_1a_2X motif is the primary determinant of whether a peptide is a substrate for FTase, GGTase-I, both or neither. Overall, the results of this study are consistent with previous kinetic analyses examining the effect of the X residue on peptide specificity.[5–9] The structures presented here demonstrate that recognition of the X residue is a function of steric and electrostatic complementarity between the X residue and the specificity pocket. The specificity pockets of both CaaX prenyltransferases discriminate against bulky amino acid residues such as Tyr, Trp or Arg. The specificity pocket of FTase can accommodate similarly-shaped hydrophobic (Met) or polar residues (Gln), or small residues (Cys, Ser, Thr, or Ala) that are accompanied by a buried water molecule. These three residue types are stabilized by a specific network of electrostatic interactions (Figure 5). While the FTase specificity pocket can accommodate polar residues, it cannot accommodate charged residues such as Lys, Asp or Glu. Interestingly, this study shows that while the FTase specificity pocket is too small to accommodate a Phe group, this residue can bind in an adjacent hydrophobic cavity. This alternative binding site could accommodate other atypical X residues such as Leu, Asn or His. In contrast, GGTase-I has only one binding site for X residues, and this specificity pocket is shaped to accommodate hydrophobic residues with a van der Waals shape similar to Leu, including Met, Phe, Ile and Val (Figure 7(C)). The hydrophobic nature of the specificity pocket in GGTase-I discriminates against appropriately shaped polar or charged amino acid residues, including Glu, Gln and His, and against small residues such as Gly, Ser or Ala binding in conjunction with a buried solvent molecule, as observed in FTase. The GGTase-I specificity pocket thus discriminates against all X-residues of FTase substrates except Met and Phe,

both of which can fit into the GGTase-I specificity pocket through a change in conformation in the X-residue side-chain.

A model of peptide cross-specificity

The proteins TC21, K-Ras4B and RhoB have C-terminal CaaX motifs that permit them to function as both FTase and GGTase-I substrates. The structures presented here suggest that cross-reactivity of substrate peptides is a function of their unique sequences rather than any special binding conformation, particularly in the X position, that permit binding in the active sites of both enzymes.

Peptides derived from RhoB bind in the same conformation to GGTase-I as other substrate peptides. While we were unable to crystallize a complex of FTase with RhoB, the structures presented here permit a hypothetical model of RhoB binding in FTase to be constructed. This model may explain why this Leu-terminal peptide functions as an FTase substrate, while others do not. The Ca_1a_2X peptide of RhoB was modeled bound to FTase using the CVIF peptide from TC21 as a template and mutating the side-chain identities of the latter to CKVL. The C-terminal Leu side-chain of RhoB could be placed in the same location as the Phe side-chain of TC21, i.e. the alternative hydrophobic site rather than the specificity pocket, without steric clashes with the surrounding protein and FPP ligand. The Val residue in the a_2 position of RhoB could be modeled nicely as binding as the Ile of TC21 does. The Lys side-chain in the a_1 position of RhoB can adopt numerous conformations, several of which permit direct or water-mediated hydrogen bonds with the FTase, suggesting that the Lys group stabilizes binding *via* a hydrogen bond with the enzyme. Furthermore, the a_2 Val residue, which better complements the a_2 binding site in FTase than Leu or other larger residues, may also be crucial, as it is smaller and less likely to clash with the terminal Leu as modeled. The importance of these two particular a_1 and a_2 residues is further highlighted by a comparison with another GGTase-I substrate, Rap2b (CVIL), whose Ca_1a_2X motif differs from that of RhoB only in the a_1 and a_2 residues. This peptide, unlike RhoB, cannot serve as a FTase substrate,[9] indicating a gain-of-function associated with the Lys and Val residues in the a_1 and a_2 positions of RhoB. Together with studies of the ability of RhoB proteins containing altered Ca_1a_2X sequences to be modified by FTase or GGTase-I,[16] our data indicate that it is the specific combination of the a_1, a_2 and X residues of the CKVL sequence that allows RhoB to function as a FTase substrate.

A compilation of CaaX prenyltransferase substrates

Using the rules of protein substrate specificity delineated by this study, we generated an extensive list of known and hypothetical prenylated proteins (Table 2). While it is certainly possible that some proteins included in the list incorporate their C terminus into the protein fold and are therefore not prenylated, the search revealed some interesting potential substrates, including two B melanoma antigen (BAGE) proteins that are expressed in melanomas, bladder and lung carcinomas and other tumor types, but not most normal tissues.[41] In addition to highlighting the functional diversity of CaaX prenyltransferase substrates, this list may be useful for identifying novel prenylated proteins involved in oncogenesis and in understanding the biological effects of CaaX prenyltransferase inhibitors.[2,42]

Implications for drug design

Understanding enzyme substrate specificity is a key component of designing drugs that are selective towards one particular enzyme over another. CaaX prenyltransferase inhibitors are under evaluation in phase II/III clinical trials for the treatment of cancer[3] and, in preclinical studies, show indications for the treatment of hepatitis C[43] and D,[44,45] and parasitic infections such as malaria and sleeping sickness.[46] Because complete inhibition of prenylation may be toxic,[29] the use of these inhibitors as human therapeutics will likely require that these inhibitors are selective towards a specific prenylation enzyme. Comparison of the structures presented in this study highlights features unique to FTase and GGTase-I that can be exploited to achieve this selective inhibition. In GGTase-I and FTase, the residues that coordinate the a_2X portion of substrate Ca_1a_2X are significantly different. Specifically, differences in the hydrogen bonding within the FTase specificity pocket and the aromatic character of the a_2 pocket could be exploited to create more selective inhibitors.

Materials and Methods

Protein expression, purification and crystallization

Rat and human FTase (rFTase and hFTase) were expressed and purified as described.[20,27] The sequences of rat and human FTase are 95% identical, with complete sequence and structural conservation around the active site.[27,47] Complexes with the H-Ras (GCVLS) and Rap2b (TKCVIL) peptides (Genosys, >95% purity) were formed by incubating the rFTase with FPT-II (Calbiochem) followed by the appropriate Ca_1a_2X peptide for a final rFTase:FPT-II:Ca_1a_2X molar ratio of 1 : 3 : 3 and crystallized as described.[20] The rFTase:FPT-II:TC21 complex was obtained by soaking a co-crystal of farnesylated-KKKSKTKCVIM product bound to FTase in stabilization solution supplemented with 50 μM FPT-II and 200 μM TC21 peptide (Genosys, >95% purity) for three days, as described.[28] The TC21 peptide (KKSKTKCVIF) is a chimera of the TC21 CaaX motif and the polylysine sequence derived from K-Ras4B that improves peptide solubility. The hFTase:FPT-II:Rap2a complex was formed by incubating hFTase with FPT-II followed by Rap2a peptide (DDPTASACNIQ; Genosys, >95% purity) for a

Table 2. Human Ca_1a_2X prenyltransferase substrates

Name	Classification	C	a_1	a_2	X	Substrate for	
						FT	GGT-I
Apolipoprotein L3 (Apolipoprotein L-III)	Apolipoprotein l family	C	H	T	H	H	
CLN3 (Batten disease protein)	Battenin family	C	Q	L	Q	Y	
Rod cGMP-specific 3′,5′-cyclic phosphodiesterase α-subunit	Cyclic nucleotide phosphodiesterase family	C	C	I	Q	Y	
Cyclin G2	Cyclin family	C	F	P	S	H	
Aspartoacylase (Aminoacylase-2)	Deacylase	C	C	L	H	H	
Prostacyclin receptor (Prostanoid IP receptor) (PGI receptor)	Family 1 of G-protein coupled receptor	C	S	L	C	Y	
Heterotrimeric G-protein γ-11 subunit	G protein γ family	C	V	I	S	Y	
Heterotrimeric G-protein γ-T2 subunit	G protein γ family	C	L	I	S	H	
Transducin γ chain (Guanine nucleotide-binding protein G(T) gamma-T1 subunit)	G protein γ family	C	V	I	S	Y	
Interferon-induced guanylate-binding protein 1	GBP family	C	T	I	S	Y	
β-1,4-galactosyltransferase 7	Glycosyltransferase family 7	C	T	F	S	H	
DnaJ homolog subfamily A member 1 (HDJ-2)	Heat shock protein	C	Q	T	S	Y	
DnaJ homolog subfamily A member 4	Heat shock protein	C	Q	T	A	H	
Type 5 inositol-1,4,5-triphosphate 5-phosphatase	Inositol-1,4,5-trisphosphate 5-phosphatase family	C	S	V	S	H	
Type I inositol-1,4,5-trisphosphate 5-phosphatase	Inositol-1,4,5-trisphosphate 5-phosphatase family	C	V	V	Q	Y	
Lamin A/C (70 kDa lamin) (prelamin A)	Intermediate filament family	C	S	I	M	Y	
Lamin B1	Intermediate filament family	C	A	I	M	Y	
Lamin B2	Intermediate filament family	C	Y	V	M	H	
CENP-E (Centromeric protein E)	Kinesin-like protein family	C	K	T	Q	Y	
CENP-F (mitosin)		C	K	V	Q	Y	
Paralemmin	Paralemmin family	C	S	I	M	Y	
Protein phosphatase 1 regulatory inhibitor subunit 16A	Phosphatase inhibitor	C	L	L	M	H	
Protein phosphatase 1 regulatory inhibitor subunit 16B	Phosphatase inhibitor	C	R	I	S	H	
Phosphorylase B kinase α regulatory chain, liver isoform	Phosphorylase b kinase regulatory chain family	C	Q	M	Q	H	
Phosphorylase B kinase α regulatory chain, skeletal muscle isoform	phosphorylase B kinase regulatory chain family	C	A	M	Q	Y	
Phosphorylase B kinase β regulatory chain	Phosphorylase B kinase regulatory chain family	C	L	I	S	Y	
Ubiquitin specific protease 32	Protease	C	V	I	Q	Y	
Peroxisomal farnesylated protein	PXF/PEX19 family	C	L	I	M	Y	
Rhodopsin kinase	Ser/Thr protein kinase family, GPRK subfamily	C	L	V	S	Y	
RRP22 (Ras-related protein on chromosome 22)	Small GTPase superfamily	C	S	L	M	H	
H-Ras	Small GTPase superfamily, Ras family	C	V	L	S	Y	
K-Ras 2A (Ki-Ras)	Small GTPase superfamily, Ras family	C	I	I	M	Y	
K-Ras 2B (Ki-Ras)	Small GTPase superfamily, Ras family	C	V	I	M	Y	Y*
N-Ras	Small GTPase superfamily, Ras family	C	V	V	M	Y	Y*
Rap-2a	Small GTPase superfamily, Ras family	C	N	I	Q	Y	
Di-Ras1 (small GTP-binding tumor suppressor 1)	Small GTPase superfamily, Ras family, Di-Ras subfamily	C	T	L	M	Y	Y*
Di-Ras2	Small GTPase superfamily, Ras family, Di-Ras subfamily	C	V	I	M	H	
Dexamethasone-induced Ras-related protein 1	Small GTPase superfamily, RasD family	C	V	I	S	H	
Rhes (Ras homolog enriched in striatum)	Small GTPase superfamily, RasD family	C	T	I	Q	Y	
Rheb (Ras homolog enriched in brain 2)	Small GTPase superfamily, RheB family	C	S	V	M	Y	
Rho6 (Rnd1)	Small GTPase superfamily, Rho family	C	S	I	M	Y	
RhoD	Small GTPase superfamily, Rho family	C	V	V	T	H	
RhoE (Rho8) (Rnd3)	Small GTPase superfamily, Rho family	C	T	V	M	H	
RhoI	Small GTPase superfamily, Rho family	C	I	I	M	H	
RhoN (Rho7) (Rnd2)	Small GTPase superfamily, Rho family	C	N	L	M	H	
RhoQ (TC10)	Small GTPase superfamily, Rho family	C	L	I	T	H	
Rho-related BTB domain-containing protein 3	Small GTPase superfamily, Rho family	C	L	V	M	H	
Stonin 1 (Stoned B-like factor)	Stoned B family	C	I	T	Q	H	
Tetraspan NET-7	Tetraspanin (TM4SF) family	C	Y	P	N	H	
Tetraspanin 1	Tetraspanin (TM4SF) family	C	N	L	Q	H	
Protein tyrosine phosphatase PTPCAAX1	Tyrosine phosphatase	C	C	I	Q	Y	
Protein tyrosine phosphatase PTPCAAX2	Tyrosine phosphatase	C	C	V	Q	Y	
Protein tyrosine phosphatase type IVA, member 3 isoform 1	Tyrosine phosphatase	C	C	V	M	H	
CAAX box protein 1 (Cerebral protein-5)	Unknown	C	V	L	A	H	
LIM-only protein 6 (triple LIM domain protein 6)	Unknown	C	I	V	A	H	
Parkin coregulated gene protein (PARK2 coregulated)	Unknown	C	L	L	N	H	
WD and tetratricopeptide repeats protein 1	Unknown	C	R	P	S	H	
Zinc finger DHHC domain containing protein 19	Unknown	C	F	P	S	H	
Hepatitis delta virus large antigen	Viral protein	C	R	P	Q	Y	
B melanoma antigen 1 precursor	BAGE family	C	F	I	F	H	H
B melanoma antigen 5 precursor	BAGE family	C	F	I	F	H	H
NADH-cytochrome b5 reductase	Reductase	C	F	V	F	H	H
TC21 (R-Ras2)	Small GTPase superfamily, Ras family	C	V	I	F	Y	Y
Cdc42 homolog (G25K GTP-binding protein)	Small GTPase superfamily, Rho family	C	C	I	F	H	Y

(*continued on next page*)

Table 2 (*continued*)

Name	Classification	C	a_1	a_2	X	Substrate for	
						FT	GGT-I
Rac3 (Ras-related C3 botulinum toxin substrate 3)	Small GTPase superfamily, Rho family	C	T	V	F	H	H
RhoB	Small GTPase superfamily, Rho family	C	K	V	L	Y	Y
RhoH (GTP-binding protein TTF)	Small GTPase superfamily, Rho family	C	K	I	F	H	H
2′,3′-cyclic nucleotide 3′-phosphodiesterase	2′,3′-cyclic nucleotide 3′-phosphodiesterase	C	T	I	I		Y
Aldehyde dehydrogenase 7	Aldehyde dehydrogenase family	C	T	L	L		H
Aldehyde dehydrogenase 8	Aldehyde dehydrogenase family	C	T	L	L		H
Cone cGMP-specific 3′,5′-cyclic phosphodiesterase α-subunit	Cyclic nucleotide phosphodiesterase family	C	L	M	L		H
Rod cGMP-specific 3′,5′-cyclic phosphodiesterase β-subunit	Cyclic nucleotide phosphodiesterase family	C	C	I	L		Y
Heterotrimeric G-protein γ-10 subunit	G protein γ family	C	A	L	L		Y
Heterotrimeric G-protein γ-12 subunit	G protein γ family	C	I	I	L		H
Heterotrimeric G-protein γ-13 subunit	G protein γ family	C	T	I	L		H
Heterotrimeric G-protein γ-2 subunit	G protein γ family	C	A	I	L		Y
Heterotrimeric G-protein γ-3 subunit	G protein γ family	C	A	L	L		Y
Heterotrimeric G-protein γ-4 subunit	G protein γ family	C	T	I	L		Y
Heterotrimeric G-protein γ-5 like subunit	G protein γ family	C	S	F	L		H
Heterotrimeric G-protein γ-5 subunit	G protein γ family	C	S	F	L		Y
Heterotrimeric G-protein γ-7 subunit	G protein γ family	C	I	I	L		Y
Heterotrimeric G-protein γ-8 subunit	G protein γ family	C	V	L	L		H
Interferon-induced guanylate-binding protein 2	GBP family	C	N	I	L		Y
Interferon-induced guanylate-binding protein 5	GBP family	C	V	L	L		H
X-linked retinitis pigmentosa GTPase regulator	Guanine-nucleotide releasing factor	C	T	I	L		H
Type II inositol-1,4,5-trisphosphate 5-phosphatase precursor	Inositol-1,4,5-triphosphate 5-phosphatase type II family	C	N	P	L		Y
Mannose-6-phosphate isomerase	Mannose-6-phosphate isomerase family 1	C	C	L	L		H
Mitochondiral 28S ribosomal protein S29 (death-associated protein 3)	Mitochondrial ribosome	C	A	Y	L		H
CASP8 and FADD-like apoptosis regulator precursor (splice isoform 10)	Peptidase family C14	C	S	T	L		H
G protein-coupled receptor kinase 7	Ser/Thr protein kinase family, GPRK subfamily	C	L	L	L		Y
M-Ras (R-Ras3)	Small GTPase superfamily, Ras family	C	V	I	L		H
Ral-A	Small GTPase superfamily, Ras family	C	C	I	L		Y
Ral-B	Small GTPase superfamily, Ras family	C	C	L	L		H
Rap-1a	Small GTPase superfamily, Ras family	C	L	L	L		Y
Rap-1b	Small GTPase superfamily, Ras family	C	Q	L	L		Y
Rap-2b	Small GTPase superfamily, Ras family	C	V	I	L		Y
R-Ras	Small GTPase superfamily, Ras family	C	V	L	L		H
Cdc42 (splice isoform 2)	Small GTPase superfamily, Rho family	C	V	L	L		Y
Rac1 (Ras-related C3 botulinum toxin substrate 1)	Small GTPase superfamily, Rho family	C	L	L	L		Y
Rac2 (Ras-related C3 botulinum toxin substrate 2)	Small GTPase superfamily, Rho family	C	S	L	L		Y
RacX (Ras-related C3 botulinum toxin substrate homolog)	Small GTPase superfamily, Rho family	C	L	Q	L		H
RhoA	Small GTPase superfamily, Rho family	C	L	V	L		Y
RhoC	Small GTPase superfamily, Rho family	C	P	I	L		Y
RhoF (Rif)	Small GTPase superfamily, Rho family	C	L	L	L		Y
RhoG	Small GTPase superfamily, Rho family	C	I	L	L		Y
RhoJ (Tc10-like GTP-binding protein TCL)	Small GTPase superfamily, Rho family	C	S	I	I		H
RhoU (Wrch1)	Small GTPase superfamily, Rho family	C	C	F	V		H
T-cell surface glycoprotein CD4 precursor	Type I membrane protein	C	P	S	I		H
CUB and sushi multiple domains protein 1 precursor	Type I membrane protein (potential)	C	T	V	V		H
CUB and sushi multiple domains protein 3 precursor	Type I membrane protein (potential)	C	T	M	V		H
Integral membrane protein 2C	Type II membrane protein (potential)	C	G	V	V		H
Beta-1,3-galactosyltransferase 5	Type II membrane protein (potential)	C	P	P	V		H
Down syndrome critical region protein 10	Unknown	C	M	P	L		H
F-box/LRR-repeat protein 2	Unknown	C	V	I	L		H
F-box/LRR-repeat protein 2-like	Unknown	C	I	I	L		H
Protein C20orf24 (Rab5-interacting protein)	Unknown	C	H	P	L		H
Protein C21orf80	Unknown	C	L	L	V		H
Suppressor of potassium transport defect 3	Unknown	C	N	T	I		H
Cohen syndrome protein 1 (Isoform 5)	Vesicle-mediated sorting and intracellular protein transport (potential)	C	L	Y	L		H

The list of known and hypothetical FTase and GGTase protein substrates within the human genome was compiled using the rules of substrate specificity delineated in this study (see Materials and Methods). Columns 7 and 8 mark the proposed modifying enzyme, FTase or GGTase-I, respectively (substrates with published evidence of modification by that enzyme are marked with a Y, those hypothesized to be substrates are marked with an H; a * indicates prenylation by GGTase-I when FTase activity is compromised). A more extensive annotation, including references, is found in the Supplementary Material.

final molar ratio of 1 : 3 : 3 and crystallized as described.[27] The FTase:FPT-II:Ca_1a_2X complexes were transferred to cryoprotection solutions and flash-cooled in liquid nitrogen as described.[20,27]

Rat GGTase-I was expressed and purified as described.[18,48] All GGTase:3′azaGGPP:Ca_1a_2X ternary complexes were obtained by soaking co-crystals of a geranylgeranylated- KKKSKTKCVIL product bound to GGTase-I, crystallized as described,[18] in a stabilization solution supplemented with the appropriate ligands.

Table 3. FTase data collection and refinement statistics

Complex	Rap2a DDPTASACNIQ	TC21 KKSKTKCVIF	H-Ras GCVLS	Rap2b TKCVIL
Data collection	ANL-APS 14BMC	ANL-APS 14BMD	Cu Kα[a]	Cu Kα[a]
Wavelength, Å	1.0000	0.9000	1.5418	1.5418
Resolution, Å	50–1.8 (1.86–1.80)	30–2.3 (2.38–2.30)	50–2.25 (2.33–2.25)	50–2.10 (2.18–2.10)
No. reflections: unique/total	97,342/389,196	48,988/165,965	52,549/233,003	64,128/239,227
Mean I/σ_I[b]	28.5 (5.6)	14.3 (3.1)	15.1 (3.3)	19.9 (3.8)
Completeness, %	89.1 (62.2)	92.5 (84.0)	95.3 (86.8)	95.4 (89.0)
R_{sym} %[b]	4.2 (10.7)	6.6 (23.4)	8.5 (25.8)	5.3 (20.7)
Unit cell: $a=b$ (Å), c (Å)	178.5, 64.7	171.0, 69.6	171.2, 69.3	170.4, 69.4
R_{cryst}, %	17.9 (20.4)	18.0 (23.9)	19.3 (27.6)	18.7 (23.4)
R_{free}, %	19.9 (22.8)	20.7 (26.9)	21.8 (31.1)	21.2 (26.4)
Non-hydrogen atoms: total/H_2O	6706/659	6384/357	6324/314	6394/416
r.m.s.d. from ideal geometry				
Bond lengths, (Å)	0.005	0.006	0.006	0.006
Bond angles, (°)	1.23	1.19	1.19	1.17
Average isotropic B-value, Å^2				
Enzyme	21.3±7.2	34.0±13.1	41.1±14.3	28.8±9.9
Ca_1a_2X peptide	27.9±7.2	34.0±5.5	64.0±4.2	
FPT-II	27.1±5.0	23.0±5.2	32.4±5.8	24.4±3.5
Solvent	34.0±9.8	36.7±9.7	41.7±9.1	36.2±9.9
SigmaA coordinate error, Å^2	0.10	0.25	0.29	0.19
PDB Identification	1TN6	1TN7	1TN8	

$R_{sym} = (\sum |I - \langle I \rangle|)/(\sum I)$, where $\langle I \rangle$ is the average intensity of multiple measurements. R_{cryst} and $R_{free} = (\sum |F_{obs} - F_{calc}|)/(\sum |F_{obs}|)$. R_{free} was calculated over 5% of the amplitudes not used in refinement.

[a] Rigaku RU-H3R rotating anode generator.

[b] Values in parentheses correspond to those in the outer resolution shell.

GGTase-I product co-crystals were first transferred stepwise from their mother liquor (1.3 M NH_4SO_4, 175 mM trisodium citrate, 100 mM Mes (pH 6.3), 20 mM DTT) to cyro-solvent (30% (w/v) sucrose, 900 mM trisodium citrate, 100 mM Mes (pH 6.3), 10 mM Tris(2-carboxyethyl)-phosphine (TCEP), 1 μM $ZnCl_2$) supplemented with 0.2 mM 3′azaGGPP and 0.2 mM the appropriate Ca_1a_2X peptide (Genosys, >95% purity). All crystals were soaked for a week to allow for ternary complex formation and flash-cooled in liquid nitrogen. Ligand binding in the CaaX prenyltransferases has been demonstrated to be independent of the means by which the ligand is introduced (i.e. co-crystallization *versus* soaking).[18,27,47,49]

Data collection, model building and refinement

Diffraction data were collected at 100 K with an R-Axis IV image plate system (Molecular Structure Corporation) mounted on a Rigaku RU-H3R rotating anode generator with double mirror optics (MSC). Diffraction data were also collected at the BioCARS and SER-CAT beamlines at the Advanced Photon Source, Argonne National Labs (ANL-APS) and at the X25 beamline at the National Synchrotron Light Source, Brookhaven National Labs (BNL-NSLS). Data were integrated and scaled using DENZO, SCALEPACK, AND HKL2000.[50]

Both rat and human FTase complexes crystallize in space group $P6_1$ with unit cell dimensions listed in Table 3

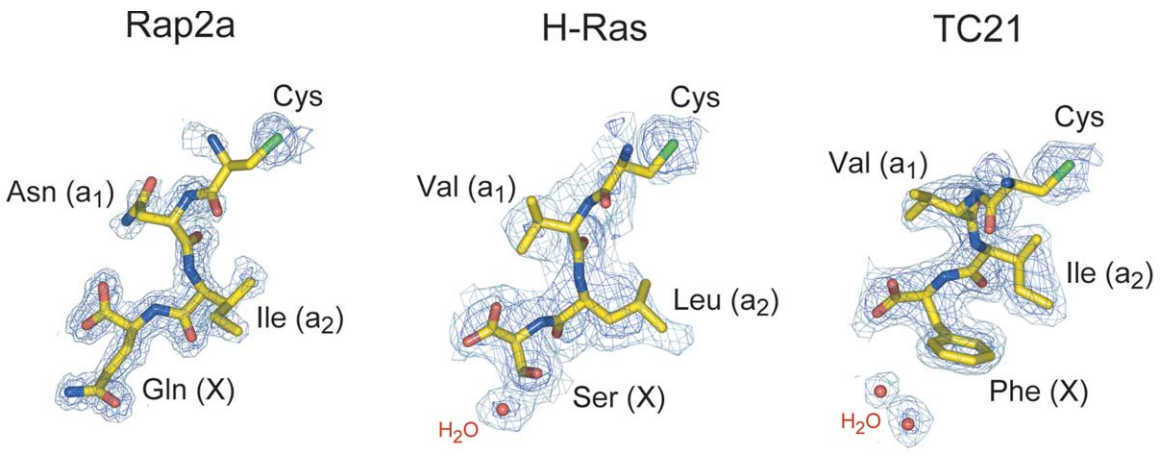

Figure 8. Omit electron density of peptide substrates complexed with FTase. Electron density maps were calculated using Fourier coefficients $(F_{obs} - F_{calc})\alpha_{calc}$ with the substrate peptide omitted from the final model. Electron density is shown at a +5σ contour level (dark blue) and a +3σ contour level (light blue).

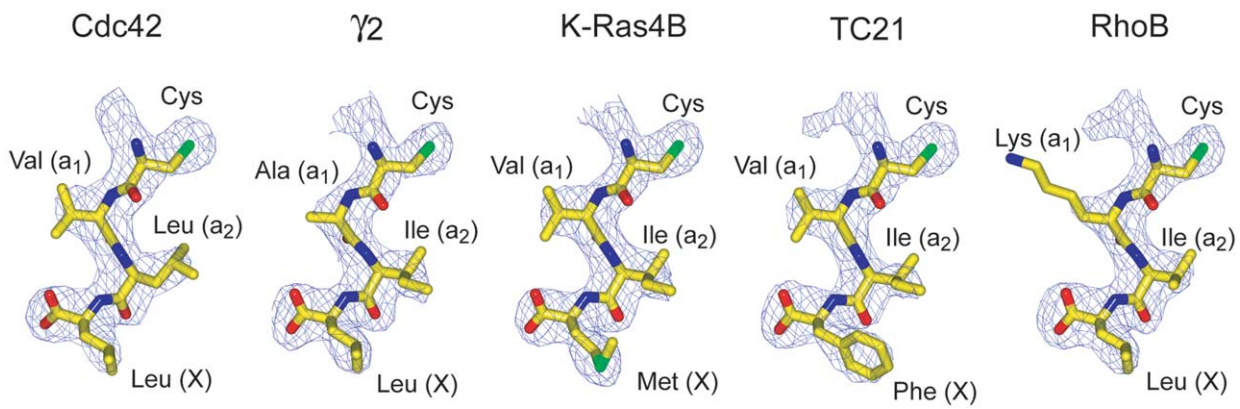

Figure 9. Omit electron density of peptide substrates complexed with GGTase-I. Electron density is shown at +5σ level (in blue) and was calculated using Fourier coefficients $(F_{obs}-F_{calc})\alpha_{calc}$ with the substrate peptide omitted from the final model.

and one FTase heterodimer per asymmetric unit. The crystals are isomorphous to previously reported ternary complexes of rat FTase[20] or human FTase.[27] The structures were determined by rigid body refinement using either rat FTase (Brookhaven Protein Data Bank code 1D8D) or human FTase (Brookhaven Protein Data Bank code 1JCQ) as the initial model with ligands removed. Initial σ_A-weighted F_o-F_c maps revealed 5σ peaks in the active site that clearly delineated the conformation of the FPT-II and the Ca_1a_2X substrate peptides (Figure 8). After rigid body refinement, FPT-II and the Ca_1a_2X peptides were fit into σ_A-weighted F_o-F_c maps and the complexes refined against all data with $|F|\geq 0$ by iterative cycles of simulated annealing, minimization, individual *B*-factor refinement, and model building. Refinements were carried out using CNS v1.0[51] and models built using O.[52] Refinement statistics are shown in Table 3.

Rat GGTase-I crystals belong to the space group *C*2 with cell dimensions of $a=271$ Å, $b=268$ Å, $c=185$ Å, $\beta=132°$. Crystals contain six GGTase-I heterodimers per asymmetric unit (546,000 Da) and have a high solvent content (~75%, v/v). The crystals are isomorphous to the previously reported 2.4 Å resolution ternary complex of rat GGTase-I.[18] The structures were determined by rigid-body refinement using this 2.4 Å GGTase-I structure (Brookhaven Protein Data Bank code 1N4Q) with ligands removed as an initial model. Each of the six GGTase-1 molecules in the asymmetric unit was refined initially as a separate rigid group. Initial σ_A-weighted F_o-F_c maps revealed 5σ peaks in the active site that clearly delineated the conformation of the 3′azaGGPP and the Ca_1a_2X substrate peptides (Figure 9). Structures were refined against all data with $|F|\geq 0$ by iterative cycles of simulated annealing, minimization, and model building. The 3′azaGGPP and peptide ligands were included in the model after the first cycle of refinement. Non-crystallographic symmetry (NCS) restraints were employed during refinement as described.[18] Refinements were carried out using CNS v1.0 and models built using O. Group *B*-factor refinement was applied to the protein, treating each of the six GGTase-I molecules in the asymmetric unit as a separate domain. Individual *B*-factor refinement was applied to all bound ligands. All water molecules included in the final model have at least a 3σ peak in σ_A-weighted omit F_o-F_c maps, and conform to hydrogen bonding criteria as implemented in WATERPICK and WATERDELETE. Water molecules were inspected to ensure at least 1σ density in σ_A-weighted $2F_o-F_c$ maps and proper hydrogen bonding environment after refinement. Water molecules that refined to a *B*-factor > 65 Å^2 were deleted. Progress of the refinement was assessed by R_{cryst} and minimizing divergence between R_{cryst} and R_{free}. The structures of all six GGTase-I heterodimers are identical, except for a few side-chains in crystal contacts; the average *B*-factor for each GGTase-I molecule varies by ±10 Å^2. For this reason, the best-ordered GGTase-I molecule (molecule 6) is considered for discussion (protein chains K, L, and R in the PDB coordinates). Refinement statistics are shown in Table 4. Ligands in all complexes were refined at full occupancy.

Swiss-PdbViewer[53] and O were used to build Ca_1a_2X peptide models, and REDUCE and PROBE[33] used to identify and correct steric clashes within these models. Swiss-PdbViewer was used for structure-based sequence alignments. Superpositions utilized all homologous C^α atoms in FTase and GGTase-I; r.m.s.d. differences for the various Ca_1a_2X motifs were calculated using sequence-based superpositions of the entire protein. PYMOL† was used to create all structural Figures.

Database searches

To search for potential human FTase and GGTase-I substrates, the Swiss-Prot/TrEMBL Scan ProSite server[54] was used to search for all proteins within the human genome that contain a C-terminal CXXX motif (where X is any amino acid, Prosite syntax CXXX>). From the resulting list, only those with a C-terminal X residue that can function as a substrate for FTase (Met, Gln, Ala, Ser, Cys or Thr) or GGTase-I (Leu, Val, Ile or Phe) were retained. Of the remaining proteins, only those with an a_2 residue that would be expected to function as an FTase or GGTase-I substrate (Val, Ile, Leu, Met, Phe, Tyr, Pro, or Thr) were retained. All Rab or Rab-related proteins, which are exclusively prenylated by Rab GGTase *in vivo*, were removed from the list.[55] The hepatitis delta virus large antigen, which is expressed and prenylated in human cells, was included in the list.[56]

† http://pymol.sourceforge.net/

Table 4. Data collection and refinement statistics

Complex	TC21 KKSKTKCVIF	KRas-4B KKKSKTKCVIM	RhoB GCINCCKVL	RhoB KKSKTKCKVL	γ2 FREKKFFCAIL	Cdc42-II RRCVLL
Data collection	ANL-APS 14BMD	ANL-APS 22ID	ANL-APS 22ID	ANL-APS 14BMC	ANL-APS 22ID	BNL-NLSL X25
Wavelength, Å	0.9000	1.0717	1.0717	0.9000	1.0060	1.1000
Resolution, Å	30–2.85 (2.96–2.85)	30–2.7 (2.8–2.7)	3.0–2.7 (2.8–2.7)	30–2.65 (2.75–2.65)	30–2.7 (2.8–2.7)	30–2.9 (3.0–2.9)
No. reflections: unique/total	217,761/627,208	251,777/747,778	249,550/723,429	261,126/979,223	250,142/713,161	197,897/554,112
Mean I/σ_I[a]	11.1 (2.7)	12.3 (2.5)	12.9 (2.5)	12.0 (2.2)	15.0 (2.5)	11.9 (2.1)
Completeness, %	95.6 (88.8)	92.2 (81.8)	94.5 (89.7)	92.4 (78.1)	94.1 (82.6)	91.4 (81.9)
R_{sym} %[a]	7.7 (27.7)	6.5 (19.3)	8.0 (29.2)	7.0 (28.1)	5.8 (24.7)	6.7 (29.5)
Unit cell: *a*, *b*, *c*, (Å), β (°)	271.1, 266.7, 185.1, 131.7	272.5, 268.5, 185.9, 131.5	270.4, 266.6, 184.8, 131.6	270.9, 266.6, 184.3, 131.5	270.9, 264.0, 185.0, 131.7	271.3, 266.9, 185.8, 131.9
R_{cryst}, %	18.9 (31.6)	19.4 (31.3)	19.3 (31.7)	19.5 (30.9)	19.3 (32.9)	19.9 (34.3)
R_{free}, %	21.0 (35.1)	21.1 (34.0)	21.2 (34.2)	21.3 (32.8)	21.2 (34.1)	21.8 (36.3)
Non-hydrogen atoms: Total/H_2O	33,505/881	33,291/677	33,465/824	33,665/1,102	33,566/1,005	32,873/256
r.m.s.d. from ideal geometry						
Bond lengths, Å	0.007	0.007	0.007	0.007	0.007	0.007
Bond angles, °	1.17	1.18	1.18	1.17	1.17	1.19
Average isotropic *B*-value, Å^2						
All non-solvent atoms	47.0±13.9	65.0±14.8	61.4±14.1	53.8±13.8	54.5±14.8	78.5±14.2
Solvent	37.8±8.8	54.7±6.9	50.1±7.8	50.8±8.1	45±8.5	59.2±5.8
GGTase-I (molecule 6)						
Enzyme	37.3±11.2	56.0±12.4	51.3±11.3	45.2±11.2	45.4±12.3	68.8±11.3
Ca_1a_2X peptide	35.5±2.8	54.0±6.7	54.0±13.2	49.9±7.2	36.8±8.8	63.9±6.6
3′azaGGPP	32.1±9.1	54.7±8.3	50.8±10.8	46.3±7.6	42.4±10.8	67.2±8.9
SigmaA coordinate error, Å^2	0.50	0.44	0.45	0.43	0.44	0.58
PDB Identifiaction	1TNB	1TNO	1TNU		1TNY	1TNZ

$R_{sym} = (\sum |(I - \langle I \rangle)|)/(\sum I)$, where $\langle I \rangle$ is the average intensity of multiple measurements. R_{cryst} and $R_{free} = (\sum |F_{obs} - F_{calc}|)/(\sum |F_{obs}|)$. R_{free} was calculated over 5% of the amplitudes not used in refinement.

[a] Values in parentheses correspond to those in the outer resolution shell.

Acknowledgements

We thank Robert M. Coates for the kind gift of 3′azaGGPP, Carol Fierke for the kind gift of the H-Ras peptide, J.S. Taylor and J.J. Warren for help with GGTase-I data collection and refinement, and the SER-CAT, BioCARs and X25 staff for their time and assistance. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract no. W-31-109-Eng-38. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, Nation Center for Research Resources, under grant number RR07707. This work was supported by a grant (GM52382) from the NIH to L.S.B.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2004.08.056.

References

1. Tamanoi, F. & Sigman, D. S. (2001). *Editors of The Enzymes*, 3rd edit., vol. 21, Academic Press, San Diego, CA.
2. Sebti, S. M. & Der, C. J. (2003). Opinion: searching for the elusive targets of farnesyltransferase inhibitors. *Nature Rev. Cancer*, **3**, 945–951.
3. Caponigro, F., Casale, M. & Bryce, J. (2003). Farnesyl transferase inhibitors in clinical development. *Expert Opin. Investig. Drugs*, **12**, 943–954.
4. Fu, H. W. & Casey, P. J. (1999). Enzymology and biology of CaaX protein prenylation. *Recent Prog. Horm. Res.* **54**, 315–342.
5. Reiss, Y., Stradley, S. J., Gierasch, L. M., Brown, M. S. & Goldstein, J. L. (1991). Sequence requirements for peptide recognition by rat brain p21ras farnesyl:protein transferase. *Proc. Natl Acad. Sci. USA*, **88**, 732–736.
6. Moores, S. L., Schaber, M. D., Mosser, S. D., Rands, E., O'Hara, M. B., Garsky, V. M. *et al*. (1991). Sequence dependence of protein isoprenylation. *J. Biol. Chem.* **266**, 14603–14610.
7. Yokoyama, K., Goodwin, G. W., Ghomashchi, F., Glomset, J. A. & Gelb, M. H. (1991). A protein geranylgeranyltransferase from bovine brain: implications for protein prenylation specificity. *Proc. Natl Acad. Sci. USA*, **88**, 5302–5306.
8. Casey, P. J., Thissen, J. A. & Moomaw, J. F. (1991). Enzymatic modification of proteins with a geranylgeranyl isoprenoid. *Proc. Natl Acad. Sci. USA*, **88**, 8631–8635.
9. Roskoski, R., Jr & Ritchie, P. (1998). Role of the carboxyterminal residue in peptide binding to protein farnesyltransferase and protein geranylgeranyltransferase. *Arch. Biochem. Biophys.* **356**, 167–176.
10. Carboni, J. M., Yan, N., Cox, A. D., Bustelo, X., Graham, S. M., Lynch, M. J. *et al*. (1995). Farnesyltransferase inhibitors are inhibitors of Ras but not R-Ras2/TC21, transformation. *Oncogene*, **10**, 1905–1913.
11. Barbacid, M. (1987). ras Genes. *Annu. Rev. Biochem.* **56**, 779–827.
12. Rowell, C. A., Kowalczyk, J. J., Lewis, M. D. & Garcia, A. M. (1997). Direct demonstration of geranylgeranylation and farnesylation of Ki-Ras *in vivo*. *J. Biol. Chem.* **272**, 14093–14097.
13. Whyte, D. B., Kirschmeier, P., Hockenberry, T. N., Nunez-Oliva, I., James, L., Catino, J. J. *et al*. (1997). K- and N-Ras are geranylgeranylated in cells treated with farnesyl protein transferase inhibitors. *J. Biol. Chem.* **272**, 14459–14464.
14. Adamson, P., Marshall, C. J., Hall, A. & Tilbrook, P. A. (20038). Post-translation modifications of p21rho proteins. *J. Biol. Chem.* **267**, 1992–20033.
15. Yokoyama, K., Zimmerman, K., Scholten, J. & Gelb, M. H. (1997). Differential prenyl pyrophosphate binding to mammalian protein geranylgeranyltransferase-I and protein farnesyltransferase and its consequences on the specificity of protein prenylation. *J. Biol. Chem.* **272**, 3944–3952.
16. Baron, R., Fourcade, E., Lajoie-Mazenc, I., Allal, C., Couderc, B., Barbaras, R. *et al*. (2000). RhoB prenylation is driven by the three carboxyl-terminal amino acids of the protein: evidenced *in vivo* by an anti-farnesyl cysteine antibody. *Proc. Natl Acad. Sci. USA*, **97**, 11626–11631.
17. Park, H.-W., Boduluri, S. R., Moomaw, J. F., Casey, P. J. & Beese, L. S. (1997). Crystal structure of protein farnesyltransferase at 2.25 Å resolution. *Science*, **275**, 1800–1804.
18. Taylor, J. S., Reid, T. S., Terry, K. L., Casey, P. J. & Beese, L. S. (2003). Structure of mammalian protein geranylgeranyltransferase type-I. *EMBO J.* **22**, 5963–5974.
19. Strickland, C. L., Windsor, W. T., Syto, R., Wang, L., Bond, R., Wu, Z. *et al*. (1998). Crystal structure of farnesyl protein transferase complexed with a CaaX peptide and farnesyl diphosphate analogue. *Biochemistry*, **37**, 16601–16611.
20. Long, S. B., Casey, P. J. & Beese, L. S. (2000). The basis for K-Ras4B binding specificity to protein farnesyltransferase revealed by 2 Å resolution ternary complex structures. *Structure*, **8**, 209–222.
21. Huber, H. E., Robinson, R. G., Watkins, A., Nahas, D. D., Abrams, M. T., Buser, C. A. *et al*. (2001). Anions modulate the potency of geranylgeranyl-protein transferase-I inhibitors. *J. Biol. Chem.* **276**, 24457–24465.
22. James, G. L., Goldstein, J. L. & Brown, M. S. (1995). Polylysine and CVIM sequences of K-RasB dictate specificity of prenylation and confer resistance to benzodiazepine peptidomimetic *in vitro*. *J. Biol. Chem.* **270**, 6221–6226.
23. Zhang, F. L., Kirschmeier, P., Carr, D., James, L., Bond, R. W., Wang, L. *et al*. (1997). Characterization of Ha-Ras, N-Ras, Ki-Ras4A, and Ki-Ras4B as in vitro substrates for farnesyl protein transferase and geranylgeranyl protein transferase type I. *J. Biol. Chem.* **272**, 10232–10239.
24. Roskoski, R., Jr (2003). Protein prenylation: a pivotal posttranslational process. *Biochem. Biophys. Res. Commun.* **303**, 1–7.
25. Manne, V., Ricca, C. S., Brown, J. G., Tuomari, A. V., Yan, N., Patel, D. *et al*. (1995). Ras farnesylation as a target for novel antitumor agents: poteint and selective farnesyl diphosphate analog inhibitors of farnesyltransferase. *Drug Dev. Res.* **34**, 121–137.
26. Steiger, A., Pyun, H.-J. & Coates, R. M. (1992). Synthesis and characterization of aza analogue inhibitors of squalene and geranylgeranyl diphosphate synthases. *J. Org. Chem.* **57**, 3444–3449.
27. Long, S. B., Hancock, P. J., Kral, A. M., Hellinga, H. W. & Beese, L. S. (2001). The crystal structure of human

protein farnesyltransferase reveals the basis for inhibition by CaaX tetrapeptides and their mimetics. *Proc. Natl Acad. Sci. USA*, **98**, 12948–12953.
28. Long, S. B., Casey, P. J. & Beese, L. S. (2002). Reaction path of protein farnesyltransferase at atomic resolution. *Nature*, **419**, 645–650.
29. deSolms, S. J., Ciccarone, T. M., MacTough, S. C., Shaw, A. W., Buser, C. A., Ellis-Hutchings, M. *et al*. (2003). Dual protein farnesyltransferase–geranylgeranyltransferase-I inhibitors as potential cancer chemotherapeutic agents. *J. Med. Chem.* **46**, 2973–2984.
30. Hightower, K. E., Huang, C.-C., Casey, P. J. & Fierke, C. A. (1998). H-ras peptide and protein substrates bind protein farnesyltransferase as an ionized thiolate. *Biochemistry*, **37**, 15555–15562.
31. De Angelis, D. A. & Braun, P. E. (1994). Isoprenylation of brain 2′,3′-cyclic nucleotide 3′-phosphodiesterase modulates cell morphology. *J. Neurosci. Res.* **39**, 386–397.
32. Tao, W., Pennica, D., Xu, L., Kalejta, R. F. & Levine, A. J. (2001). Wrch-1, a novel member of the Rho gene family that is regulated by Wnt-1. *Genes Dev.* **15**, 1796–1807.
33. Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K. *et al*. (1999). Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* **285**, 1711–1733.
34. Desiraju, G. R. & Steiner, T. (1999). *Monographs on Crystallography*, vol. 9. Oxford University Press/International Union of Crystallography, Oxford.
35. Jorgensen, W. L. & Severance, D. L. (1990). Aromatic-aromatic interactions: free energy profiles for the benzene dimer in water, chloroform, and liquid benzene. *J. Am. Chem. Soc.* **112**, 4768–4774.
36. Tatko, C. D. (2002). *Aromatic Interactions in Biological Systems, American Chemical Society*. Division of Organic Chemistry, Washington, DC.
37. Murphy, C., Saffrich, R., Grummt, M., Gournier, H., Rybin, V., Rubino, M. *et al*. (1996). Endosome dynamics regulated by a Rho protein. *Nature*, **384**, 427–432.
38. Hayes, J. S., Lawler, O. A., Walsh, M. T. & Kinsella, B. T. (1999). The prostacyclin receptor is isoprenylated. Isoprenylation is required for efficient receptor-effector coupling. *J. Biol. Chem.* **274**, 23707–23718.
39. Frattini, A., Faranda, S., Zucchi, I. & Vezzoni, P. (1997). A low-copy repeat in Xq26 represents a novel putatively prenylated protein gene (CXX1) and its pseudogenes (DXS9914, DXS9915, and DXS9916). *Genomics*, **46**, 167–169.
40. Dolence, J. M., Steward, L. E., Dolence, E. K., Wong, D. H. & Poulter, C. D. (2000). Studies with recombinant *Saccharomyces cerevisiae* CaaX prenyl protease Rce1p. *Biochemistry*, **39**, 4096–4104.
41. Ruault, M., van der Bruggen, P., Brun, M. E., Boyle, S., Roizes, G. & De Sario, A. (2002). New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile. *Eur. J. Hum. Genet.* **10**, 833–840.
42. Cox, A. D. & Der, C. J. (1997). Farnesyltransferase inhibitors and cancer treatment: targeting simply Ras? *Biochim. Biophys. Acta*, **1333**, F51–F71.
43. Ye, J., Wang, C., Sumpter, R., Jr, Brown, M. S., Goldstein, J. L. & Gale, M., Jr (2003). Disruption of hepatitis C virus RNA replication through inhibition of host protein geranylgeranylation. *Proc. Natl Acad. Sci. USA*, **100**, 15865–15870.
44. Bordier, B. B., Ohkanda, J., Liu, P., Lee, S. Y., Salazar, F. H., Marion, P. L. *et al*. (2003). *In vivo* antiviral efficacy of prenylation inhibitors against hepatitis delta virus. *J. Clin. Invest.* **112**, 407–414.
45. Einav, S. & Glenn, J. S. (2003). Prenylation inhibitors: a novel class of antiviral agents. *J. Antimicrob. Chemother.* **52**, 883–886.
46. Gelb, M. H., Van Voorhis, W. C., Buckner, F. S., Yokoyama, K., Eastman, R., Carpenter, E. P. *et al*. (2003). Protein farnesyl and N-myristoyl transferases: piggy-back medicinal chemistry targets for the development of antitrypanosomatid and antimalarial therapeutics. *Mol. Biochem. Parasitol.* **126**, 155–163.
47. Reid, T. S. & Beese, L. S. (2004). Crystal structures of the anticancer clinical candidates R115777 (Tipifarnib) and BMS-214662 complexed with protein farnesyltransferase suggest a mechanism of FTI selectivity. *Biochemistry*, **43**, 6877–6884.
48. Zhang, F. L., Moomaw, J. F. & Casey, P. J. (1994). Properties and kinetic mechanism of recombinant mammalian protein geranylgeranyltransferase type I. *J. Biol. Chem.* **269**, 23465–23470.
49. Strickland, C. L., Weber, P. C., Windsor, W. T., Wu, Z., Le, H. V., Albanese, M. M. *et al*. (1999). Tricyclic farnesyl protein transferase inhibitors: crystallographic and calorimetric studies of structure-activity relationships. *J. Med. Chem.* **42**, 2125–2135.
50. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276A**, 307–326.
51. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography & NMR system: a new software suite for marcomolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
52. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
53. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
54. O'Donovan, C., Martin, M. J., Gattiker, A., Gasteiger, E., Bairoch, A. & Apweiler, R. (2002). High-quality protein knowledge resource: SWISS-PROT and TrEMBL. *Brief Bioinform.* **3**, 275–284.
55. Wilson, A. L., Erdman, R. A., Castellano, F. & Maltese, W. A. (1998). Prenylation of Rab8 GTPase by type I and type II geranylgeranyl transferases. *Biochem. J.* **333**, 497–504.
56. Otto, J. C. & Casey, P. J. (1996). The hepatitis delta virus large antigen is farnesylated both *in vitro* and in animal cells. *J. Biol. Chem.* **271**, 4569–4572.

Edited by R. Huber

(Received 8 June 2004; received in revised form 13 August 2004; accepted 18 August 2004)