BIOINFORMATICS DISCOVERY NOTE

Vol. 29 no. 19 2013, pages 2371–2374
doi:10.1093/bioinformatics/btt412

Sequence analysis

Advance Access publication July 18, 2013

Tiki, at the head of a new superfamily of enzymes

Luis Sanchez-Pulido* and Chris P. Ponting

MRC Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX1 3QX, UK

Associate Editor: Alfonso Valencia

ABSTRACT

Summary: Tiki proteins appear to antagonize Wnt signalling pathway by acting as Wnt proteases, thereby affecting Wnt solubility by its amino-terminal cleavage. Tiki1 protease activity was shown to be metal ion-dependent and was inhibited by chelating agents and thus was tentatively proposed to be a metalloprotease. Nevertheless, Tiki proteins exhibit no detectable sequence similarity to previously described metalloproteases, but instead have been reported as being homologues of TraB proteins (Pfam ID: PF01963), a widely distributed family of unknown function and structure. Here, we show that Tiki proteins are members of a new superfamily of domains contained not just in TraB proteins, but also in erythromycin esterase (Pfam ID: PF05139), DUF399 (domain of unknown function 399; Pfam ID: PF04187) and MARTX toxins that contribute to host invasion and pathogenesis by bacteria. We establish the core fold of this enzymatic domain and its catalytic residues.

Contact: luis.sanchezpulido@dpag.ox.ac.uk

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on February 11, 2013; revised on June 11, 2013; accepted on July 15, 2013

1 INTRODUCTION

Wnt signalling is a key developmental pathway controlling both embryonic development and tissue maintenance by regulation of stem cell proliferation in adults. Aberrant activation of the Wnt signalling pathway can lead to uncontrolled cell proliferation. Wnt-related diseases are extensive in number and include neurodegenerative disease, inflammatory bowel disease, cleft lip/palate disorders and various cancers (Herr *et al.*, 2012). The first Wnt gene was discovered >30 years ago, and despite extensive research into this signalling pathway, it has remained unclear how Wnt morphogen levels and activities are regulated. In large part, this is because of the vast numbers of Wnt isoforms, receptors, co-receptors and different secreted antagonists encoded in animal genomes (Bouldin and Kimelman, 2012; Vincent and Magee, 2012; Willert and Nusse, 2012).

Our understanding of Wnt signalling has recently been advanced substantially by the characterization of Tiki1, a gene which when over-expressed in *Xenopus* embryos results in head enlargement; Tiki refers to large-headed humanoid carvings in Polynesian cultures. Tiki1, and its paralogue Tiki2, antagonize Wnt activity upstream of its receptors (Zhang *et al.*, 2012).

*To whom correspondence should be addressed.

2 RESULTS AND DISCUSSION

2.1 Computational protein sequence analysis. The TIKI domain.

Initially, we performed a sequence similarity search using the human Tiki1 protein as query using BLAST (Altschul *et al.*, 1997). A multiple sequence alignment of Tiki homologous sequences was generated with T-Coffee (Notredame *et al.*, 2000) and used as input to make a profile using HMMER2 (Eddy, 1996). Iterative profile versus sequence searches against the UniRef50 protein sequence database (Wu *et al.*, 2006) identified a conserved region that more clearly defines the family of Tiki homologous proteins (corresponding to amino acids 52–331 of human Tiki1). The resulting Tiki alignment is extended relative to the TraB Pfam family (PF01963) (Punta *et al.*, 2012), which subsequently provided greater power to identify remote homologies.

We then compared sequence conservation profiles from different protein families using HHpred (Söding *et al.*, 2005) against Pfam (Punta *et al.*, 2012) and pdb70 profile databases (Söding *et al.*, 2005). Our first finding was that the conserved region in the Tiki/TraB family is similar in sequence to domain of unknown function 399 (DUF399) proteins (Bateman *et al.*, 2010; Chan *et al.*, 2006) and to a family of erythromycin esterases (Morar *et al.*, 2012) (Fig. 1) (E-value: 1.7×10^{-3}; Supplementary Fig. S1). Both families have structurally characterized members: for the DUF399 family, ChaN [protein data bank (PDB)]: 2G5G) (Chan *et al.*, 2006) was identified as a heme interacting protein, although no enzymatic function for this protein was suggested. The erythromycin esterase family has two structurally characterized members: BcR136 (PDB: 2QGM) and BcR135 (PDB: 2RAD & 3B55) (Northeast Structural Genomics Consortium, unpublished). A detailed functional characterization, including active site identification and structural interpretation, was recently published for the erythromycin esterase EreB protein (Morar *et al.*, 2012). The structural similarity between DUF399 and erythromycin esterase families has already been described by Morar *et al.* (2012) (Supplementary Fig. S2).

Unexpectedly, in the subsequent profile-versus-profile comparison analysis, we then also found sequence similarity to close homologues of the *Pasteurella multocida* Toxin (PMT, also known as the Dermonecrotic Toxin) (E-value: 6.9×10^{-2}; Supplementary Fig. S1); this sequence similarity was restricted to the C2-2 domain whose tertiary structure is known in PMT (Kitadokoro *et al.*, 2007). The C2-2 domain-containing family is made up of multidomain protein toxins such as PMT and MARTX proteins from *Vibrio cholerae*, *Vibrio vulnificus* and

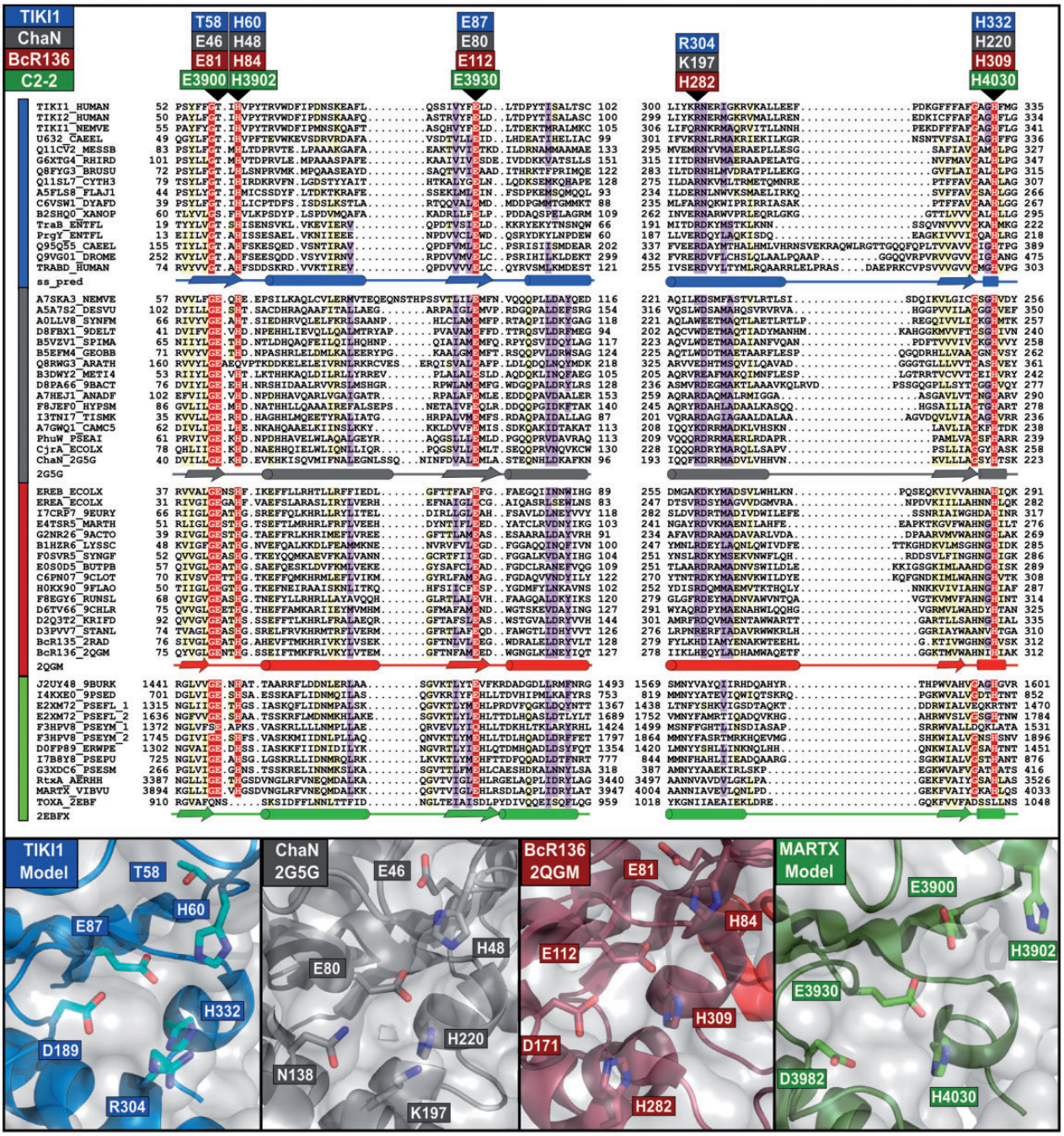

Fig. 1. Multiple sequence alignment of representative TIKI domain-containing families. Multiple sequence alignments for each family (Tiki/TraB, erythromycin esterase, DUF399 and Dermonecrotic toxins/MARTX) were generated with the program T-Coffee (Notredame *et al.*, 2000) using default parameters, slightly refined manually and visualized with the Belvu program (Sonnhammer and Hollich, 2005). The final superfamily alignment was generated using a combination of profile-to-profile comparisons (Söding *et al.*, 2005) and sequence alignments derived from structural superpositions for those families whose structure is known (Holm and Sander, 1995). Active site residues are labelled and coloured according to reference protein sequences: Tiki1, ChaN (DUF399 family), erythromycin esterase family member BcR136 (PDB: 2QGM, Northeast Structural Genomics Consortium, unpublished) and *V.vulnificus* MARTX in blue, grey, red and green, respectively. Families of TIKI domain-containing proteins are indicated by coloured bars to the left of the alignment: TIKI/TraB, DUF399, erythromycin esterase, C2-2 Dermonecrotin Toxin in blue, grey, red and green, respectively. Secondary structure predictions (Jones, 1999) were performed for the TIKI/TraB family (show in ss_pred lane), these are consistent with X-ray determined secondary structures of the DUF399 family member ChaN (PDB: 2G5G) (Chan *et al.*, 2006), erythromycin esterase BcR136 (PDB: 2QGM) (Northeast Structural Genomics Consortium, unpublished) and *P.multocida* C2-2 domain (PDB: 2EBF chain X) (Kitadokoro *et al.*, 2007). Alpha-helices and beta-strands are indicated by cylinders and arrows, respectively. The amino acid colouring scheme indicates average BLOSUM62 scores (which are correlated with amino acid conservation) for each alignment column: red (greater than 2), violet (between 2 and 0.8) and light yellow (between 0.8 and 0.2). The boundaries of the two most conserved regions in the TIKI domain are indicated by the residue positions on each side. The numbering

(continued)

Aeromonas hydrophila (Jeong and Satchell, 2012; Kwak *et al.*, 2011; Lee *et al.*, 2007, 2008, 2013; Lo *et al.*, 2011; Suarez *et al.*, 2012), as well as bacterial effector proteins from *Pseudomonas syringae* and *Erwinia amylovora* (Mansfield *et al.*, 2012). MARTX toxins, the most experimentally studied group of this family, are large proteins characterized by the presence of RtxA repeats (Pfam ID: PF07634) at the N and C termini, essential for translocation of the central portion of the toxin across the eukaryotic plasma membrane. After autoproteolysis of the holotoxin, at the host cell cytosol, the effector domains located in the central region of the toxin are released, which possess critical functions in toxinosis, and promote colonization and evasion of host immune system (Geissler *et al.*, 2010; Roig *et al.*, 2011; Satchell, 2011). The C2-2 domain is one such effector domain whose function in pathogenesis has remained unknown.

The unexpected finding of sequence similarity between erythromycin esterases and close homologues of PMT, prompted us to compare their known protein tertiary structures, which subsequently were revealed as being similar (PDB codes 2QGM and 2EBF; DALI Z-score = 7, root-mean-square-deviation of superimposed alpha-carbon atoms = 3.6 Å over 126 aligned residues with 13% identity; Supplementary Figs. S1 and S2) (Holm and Sander, 1995). The shared core domain contains four beta-strands and five alpha-helices, precisely where the catalytic residues described in the esterase structure are located (Supplementary Fig. S2) (Morar *et al.*, 2012).

Three reasons likely explain why these remote homologous relationships between C2-2 domains of bacterial toxins, erythromycin esterase/DUF399 and Tiki/TraB families had hitherto escaped detection: (i) the insertion of a non-conserved helical region that interrupts this domain (Fig. 1 and Supplementary Fig. S3); (ii) the low degrees of sequence identity among the different members of the superfamily; and (iii) the absence of the putative catalytic residues in the C2-2 domain of the structurally characterized PMT (see TOXA_2EBF sequence in Fig. 1), which are otherwise conserved in many members of this family of bacterial toxins (see representative members shown in green in Fig. 1 and Supplementary Fig. S3).

Reciprocal profile versus profile similarity searches produced convergent results, which is strongly indicative of homology (Supplementary Fig. S1). No further homologues were identified in subsequent sequence similarity searches. Additionally, a secondary structure prediction was performed for the TIKI/TraB family (ss_pred lane in Fig. 1 and Supplementary Fig. S3) (Jones, 1999), showing good agreement with the known secondary structure of different members of the superfamily (lanes: 2C5G, 2QGM and 2EBFX in Fig. 1 and Supplementary Fig. S3). The described similarities in sequence and structure provide strong evidence that TIKI/TraB, erythromycin esterase, DUF399 and C2-2 domain-containing proteins families form a new superfamily of enzymes. We term the domain common to all of these proteins 'TIKI', as it lies at the head of this large and diverse superfamily.

2.2 Reaction mechanism and active site conservation

TIKI superfamily proteins appear to catalyze different hydrolytic (esterase or protease) reactions and are characterized by a number of conserved amino acids (Fig. 1 and Supplementary Fig. S3) that form part of the recently described erythromycin esterase active site (Morar *et al.*, 2012). Morar *et al.* showed that the *Escherichia coli* erythromycin esterase EreB activity is highly dependent on two amino acids (H46 and E74 in EreB; H84 and E112 in Bcr136), which correspond to human Tiki1 (H60 and E87). It is notable that these two amino acids, together with structurally adjacent basic residues (R304 and H332 in Tiki1), are highly conserved across TIKI domain proteins (Fig. 1 and Supplementary Fig. S3). Consequently, we propose that enzymatically active TIKI domain proteins adopt the catalytic mechanism of the erythromycin esterase family wherein a conserved His–Glu pair activates a water molecule, which acts as a nucleophile to hydrolyse a carbonyl group, which is common to both ester and peptide bonds (Morar *et al.*, 2012).

The proposed metal ion-dependence described for Tiki1 (Zhang *et al.*, 2012) and EreA (erythromycin esterase family) may not be extended to all TIKI superfamily members because the reaction mechanism is metal ion-independent and resistant to chelating agents in EreB and Bcr136 members of the erythromycin esterase family (Morar *et al.*, 2012).

3 CONCLUSION

Our analyses demonstrate that human TIKI proteins lie at the head of a widely phyletically distributed protein superfamily that includes not only the TraB family but also erythromycin esterase, DUF399 and MARTX families. Furthermore, our identification of likely active site residues in TIKI and MARTX proteins should assist experiments that investigate the function of these critical proteins in different human pathologies and encourage the development of drugs that target the identified TIKI superfamily conserved active centre of enzymes, with key roles in antibiotic resistance (erythromycin esterase family), bacterial toxinosis (MARTX and Dermonecrotic toxin families) and the Wnt signalling pathway (TIKI family). During completion of this work, Bazan *et al.* (2013) reported similar results that focused on the description of the TIKI/TraB subfamily. Our work complements these results by presenting a more extensive description and annotation of sequence conservation in the predicted active site across TIKI subfamilies.

Fig. 1. Continued
after the protein name indicates the domain-repeat number when more than one TIKI domain is detected in the sequence (in E2XM72 and F3HPV8 bacterial toxin proteins). Sequences are named according to their UniProt identification or common name (details provided in Supplementary Fig. S3). Below the alignment are shown known and predicted 3D structures of representative members of the TIKI superfamily. Highly conserved residues in TIKI domains that form part of their predicted active centres are labelled and side chains shown using sticks. The human Tiki1 and *V.vulnificus* MARTX TIKI domains' structural models were created using Modeller (Sali and Blundell, 1993) based on the DUF399 family member ChaN (PDB: 2G5G) (Chan *et al.*, 2006) and erythromycin esterase BcR136 (PDB: 2QGM) (Northeast Structural Genomics Consortium, unpublished) structures. Models and structures are presented using Pymol (http://www.pymol.org)

Funding: Medical Research Council UK.

Conflict of Interest: none declared.

REFERENCES

Altschul,S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.

Bateman,A. *et al.* (2010) DUFs: families in search of function. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, **66**, 1148–1152.

Bazan,J.F. *et al.* (2013) The TIKI/TraB/PrgY family: a common protease fold for cell signaling from bacteria to metazoa? *Dev. Cell*, **13**, 225–227.

Bouldin,C.M. and Kimelman,D. (2012) Taking a bite out of Wnts. *Cell Res.*, **22**, 1621–1623.

Chan,A.C. *et al.* (2006) Cofacial heme binding is linked to dimerization by a bacterial heme transport protein. *J. Mol. Biol.*, **362**, 1108–1119.

Eddy,S.R. (1996) Hidden Markov models. *Curr. Opin. Struct. Biol.*, **6**, 361–365.

Geissler,B. *et al.* (2010) Identification of a conserved membrane localization domain within numerous large bacterial protein toxins. *Proc. Natl Acad. Sci. USA*, **107**, 5581–5586.

Herr,P. *et al.* (2012) WNT secretion and signalling in human disease. *Trends Mol. Med.*, **18**, 483–493.

Holm,L. and Sander,C. (1995) Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.*, **20**, 478–480.

Jeong,H.G. and Satchell,K.J. (2012) Additive function of *Vibrio vulnificus* MARTX(Vv) and VvhA cytolysins promotes rapid growth and epithelial tissue necrosis during intestinal infection. *PLoS Pathog.*, **8**, e1002581.

Jones,D.T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.*, **292**, 195–202.

Kitadokoro,K. *et al.* (2007) Crystal structures reveal a thiol protease-like catalytic triad in the C-terminal region of *Pasteurella multocida* toxin. *Proc. Natl Acad. Sci. USA*, **104**, 5139–5144.

Kwak,J.S. *et al.* (2011) *Vibrio vulnificus* rtxA1 gene recombination generates toxin variants with altered potency during intestinal infection. *Proc. Natl Acad. Sci. USA*, **108**, 1645–1650.

Lee,J.H. *et al.* (2007) Identification and characterization of the *Vibrio vulnificus* rtxA essential for cytotoxicity in vitro and virulence in mice. *J. Microbiol.*, **45**, 146–152.

Lee,B.C. *et al.* (2008) *Vibrio vulnificus* RTX toxin plays an important role in the apoptotic death of human intestinal epithelial cells exposed to *Vibrio vulnificus*. *Microbes Infect.*, **10**, 1504–1513.

Lee,C.T. *et al.* (2013) MARTX of *Vibrio vulnificus* biotype 2 is a virulence and survival factor. *Environ. Microbiol.*, **15**, 419–432.

Lo,H.R. *et al.* (2011) RTX toxin enhances the survival of *Vibrio vulnificus* during infection by protecting the organism from phagocytosis. *J. Infect. Dis.*, **203**, 1866–1874.

Mansfield,J. *et al.* (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.*, **13**, 614–629.

Morar,M. *et al.* (2012) Mechanism and diversity of the erythromycin esterase family of enzymes. *Biochemistry*, **51**, 1740–1751.

Notredame,C. *et al.* (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.*, **302**, 205–217.

Punta,M. *et al.* (2012) The Pfam protein families database. *Nucleic Acids Res.*, **40**, 290–301.

Roig,F.J. *et al.* (2011) Domain organization and evolution of multifunctional autoprocessing repeats-in-toxin (MARTX) toxin in *Vibrio vulnificus*. *Appl. Environ. Microbiol.*, **77**, 657–668.

Sali,A. and Blundell,T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, **234**, 779–815.

Satchell,K.J. (2011) Structure and function of MARTX toxins and other large repetitive RTX proteins. *Annu. Rev. Microbiol.*, **65**, 71–90.

Söding,J. *et al.* (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, **33**, W244–W248.

Sonnhammer,E.L. and Hollich,V. (2005) Scoredist: a simple and robust protein sequence distance estimator. *BMC Bioinformatics*, **6**, 108.

Suarez,G. *et al.* (2012) Actin cross-linking domain of *Aeromonas hydrophila* repeat in toxin A (RtxA) induces host cell rounding and apoptosis. *Gene*, **506**, 369–376.

Vincent,J.P. and Magee,T. (2012) Tiki casts a spell on Wnt. *Cell*, **149**, 1426–1427.

Willert,K. and Nusse,R. (2012) Wnt proteins. *Cold Spring Harb. Perspect. Biol.*, **4**, a007864.

Wu,C.H. *et al.* (2006) The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.*, **34**, D187–D191.

Zhang,X. *et al.* (2012) Tiki1 is required for head formation via Wnt cleavage-oxidation and inactivation. *Cell*, **149**, 1565–1577.