

Research article

Open Access

The TROVE module: A common element in Telomerase, Ro and Vault ribonucleoproteins

Alex Bateman*¹ and Valerie Kickhoefer²

Address: ¹The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK and ²Department of Biological Chemistry and the Jonsson Comprehensive Cancer Center, The David Geffen School of Medicine, University of California, Los Angeles, California 90095-1737, USA

Email: Alex Bateman* - agb@sanger.ac.uk; Valerie Kickhoefer - VKick@mednet.ucla.edu

* Corresponding author

Published: 16 October 2003

Received: 26 March 2003

BMC Bioinformatics 2003, 4:49

Accepted: 16 October 2003

This article is available from: <http://www.biomedcentral.com/1471-2105/4/49>

© 2003 Bateman and Kickhoefer; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Ribonucleoproteins carry out a variety of important tasks in the cell. In this study we show that a number of these contain a novel module, that we speculate mediates RNA-binding.

Results: The TROVE module – Telomerase, Ro and Vault module – is found in TEP1 and Ro60 the protein components of three ribonucleoprotein particles. This novel module, consisting of one or more domains, may be involved in binding the RNA components of the three RNPs, which are telomerase RNA, Y RNA and vault RNA. A second conserved region in these proteins is shown to be a member of the vWA domain family. The vWA domain in TEP1 is closely related to the previously recognised vWA domain in VPARP a second component of the vault particle. This vWA domain may mediate interactions between these vault components or bind as yet unidentified components of the RNPs.

Conclusions: This work suggests that a number of ribonucleoprotein components use a common RNA-binding module. The TROVE module is also found in bacterial ribonucleoproteins suggesting an ancient origin for these ribonucleoproteins.

Background

Many important cellular components are ribonucleoprotein (RNP) complexes, such as the spliceosome and ribosome that have key roles in gene regulation and translation. The telomerase RNP is a reverse transcriptase that maintains the telomeric repeats of eukaryotic chromosomes. Telomerase is composed of two proteins, the functionally essential reverse transcriptase TERT and the non-essential TEP1 (also known as TP1 or TLP1) as well as the telomerase RNA. TEP1 [1,2] is also found to be a component of the enigmatic vault RNP [3]. The vault is a

huge structure (13 Md) of unknown function. The vault RNP is mainly composed of the major vault protein MVP, but also contains smaller amounts of TEP1 and VPARP as well as the vault RNA. Although predominately cytoplasmic, a portion of vaults are found associated with nuclear pores [4]. Vaults have been suggested to be involved in multidrug resistance, nucleo-cytoplasmic transport, and formation of RNPs [5]. While investigating the components of these RNPs an interesting protein similarity was noticed.

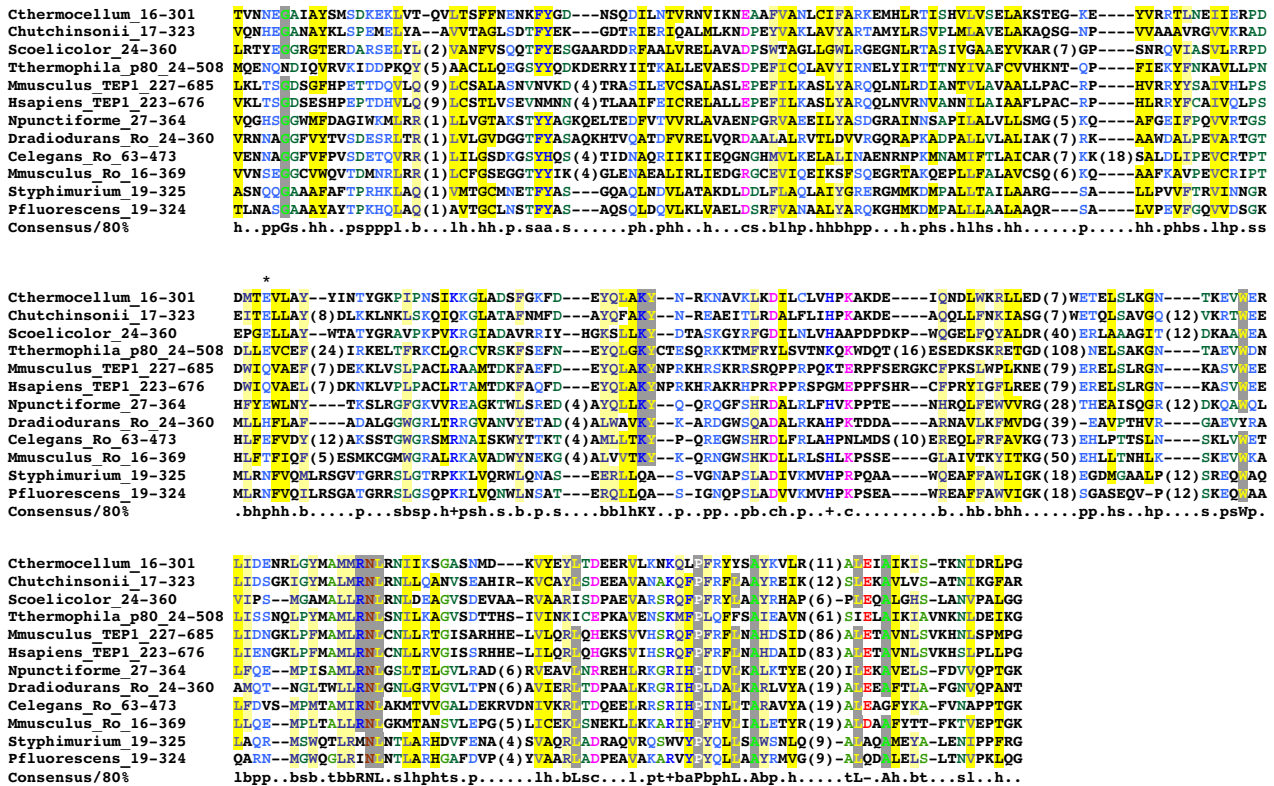


Figure 1

An alignment of TROVE modules. The alignment was generated using MAFFT [15]. The alignment has been coloured using Chroma with the default colouring scheme [16]. The 4th position in the RNP-I motif proposed by van Horn et al [12] in marked with an asterisk. The Swiss-Prot or GenBank accession numbers for the proteins in the alignment are as follows: Cthermocellum (ZP_00060193), Scoelicolor (Q9X9W7), Tthermophila_p80 (Q94818), Mmusculus_TEP1 (P97499), Hsapiens_TEP1 (Q99973), Npunctiforme (ZP_00108461), Dradiodurans_Ro (Q9RUW8), Celegans_Ro (Q27274), Mmusculus_Ro (O08848), Styphimurium (Q8ZLH8), Pfluorescens (ZP_00086137).

Results and Discussion

The complete sequence of the *Tetrahymena thermophilus* telomerase p80 component (Swiss:Q94818), a homologue of TEP1, was used to seed a PSI-blast (Position Specific Iterated-Blast) search at NCBI using the default inclusion threshold [6]. The search identified the vertebrate telomerase/vault component TEP1 and uncharacterised bacterial sequences from *Clostridium thermocellum*, *Cytophaga hutchinsonii*, *Streptomyces coelicolor* and *Salmonella typhimurium* in the first round. The second round of searching identified vertebrate homologues of the Ro60 ribonucleoprotein, with E-values as low as 6×10^{-12} , as well as three further bacterial sequences from *Pseudomonas fluorescens*, *Nostoc punctiforme* and *Deinococcus radiodurans*. The sequence of *D. radiodurans* has been previously identified as a homologue of the vertebrate Ro60 protein [7]. A short region of similarity between p80 and Ro60 was

noted previously, but the biological significance of this observation was not discussed further [2].

Ro60 is the protein component of the Ro RNP complex, that also contains a Y RNA. The region of similarity between Ro60 and TEP1 was over 800 amino acids in length. Protein domains range from 30 to 500 amino acids in length, therefore the region of similarity between Ro60 and TEP1 was too long to be a single domain, so smaller regions of these proteins were investigated. A PSI-blast search with the C-terminal residues 514 to 719 of p80 revealed matches to known vWA proteins; indicating that the C-terminal region in these proteins is a vWA domain. This new search also identified VPARP a poly-ADP-ribose polymerase associated with the vault complex [8] (see also <http://www.vaults.arc.ucla.edu/>) as containing a closely related vWA domain that was noted previ-

P10155 (538 aa) Ro60 protein (Human) & Q94818 (719 aa) p80 (Tetrahymena)



Q99973 (2627 aa) TEP1 protein, telomerase (Human)



Q9UKK3 (1724 aa) vault polyA polymerase (Human)



Figure 2

A schematic view of the domain architectures of TROVE module containing proteins. Domains shown with Pfam [17] or SMART [18] accessions: WD40 (PF00400), BRCT (PF00533), PARP (PF00644), VIT (SM0609), vWA (SM0327, PF00092), TEP1_N (PF05386). The MVPint domain is the MVP interaction domain [8,9].

ously [9]. It is somewhat surprising that two components of the vault have a highly related vWA domain. It has been suggested that the vWA domain in VPARP binds to a metal ion and might be involved in complex assembly [10]. However, the region of the vWA domain is unlikely to be a site of major vault protein (MVP) binding [9] so perhaps it could be involved in an interaction between TEP1 and VPARP, or bind an as yet unidentified transient component.

The amino-terminal 500 amino acids of p80 were found to be restricted to TEP1, Ro60 and other uncharacterised bacterial proteins using PSI-blast as above. A multiple sequence alignment of this region is shown in Figure 1. This region is large ranging from 286 residues in the *C. thermocellem* homologue to 485 residues in the p80 protein from *Tetrahymena thermophila* and so may not correspond to a single protein domain. Therefore we call this evolutionary conserved region a module. The longer members of this family have multiple long insertions that are not found in the shorter homologues. This region is named the TROVE module after Telomerase, Ro and Vault ribonucleoproteins in which it is found. The alignment of the TROVE module contains a few absolutely conserved residues. None of these conserved residues are the polar types of amino acids found in active sites, so it seems unlikely this region has an enzymatic function. *Tetrahymena* p80 is known to bind telomerase RNA [11], so the RNA-binding activity must reside in either the TROVE or vWA domains of p80. Given the known functions of vWA

domains it is likely that the RNA-binding function resides within the TROVE module.

An RNA-binding RRM domain has been proposed in the Ro60 proteins from human, frog and worm [12] due to the presence of the two classic RNP-1 and RNP-2 motifs [13]. This domain would lie within the proposed TROVE module. Although there are interesting similarities with the RNP RNA-binding motifs, examination of the broader TROVE alignment indicates that the 4th RNP-1 position (marked with an asterisk in Figure 1), is a conserved polar residue. However, in known RRM domains this position is a buried beta-sheet anchor residue, and consistently non-polar. In addition the spacing of the proposed RNP-1 and RNP-2 motifs in human and frog would be one of the shortest observed separations in RRM domains, and is inconsistent with the known structures of RRMs. Based on this sequence analysis it seems that the presence of an RRM although plausible based on Ro60 function is unlikely.

Common domains are often found in proteins involved in related cellular processes. For example the PAZ domain is found in Dicer and Piwi proteins that are involved in post-transcriptional gene silencing [14] and are both part of the RISC complex. The discovery of the TROVE module in three RNPs is intriguing and suggests that these three RNPs might be involved in inter-related processes.

What is the function of the bacterial TROVE containing proteins? The *Deinococcus* homologue is known to be

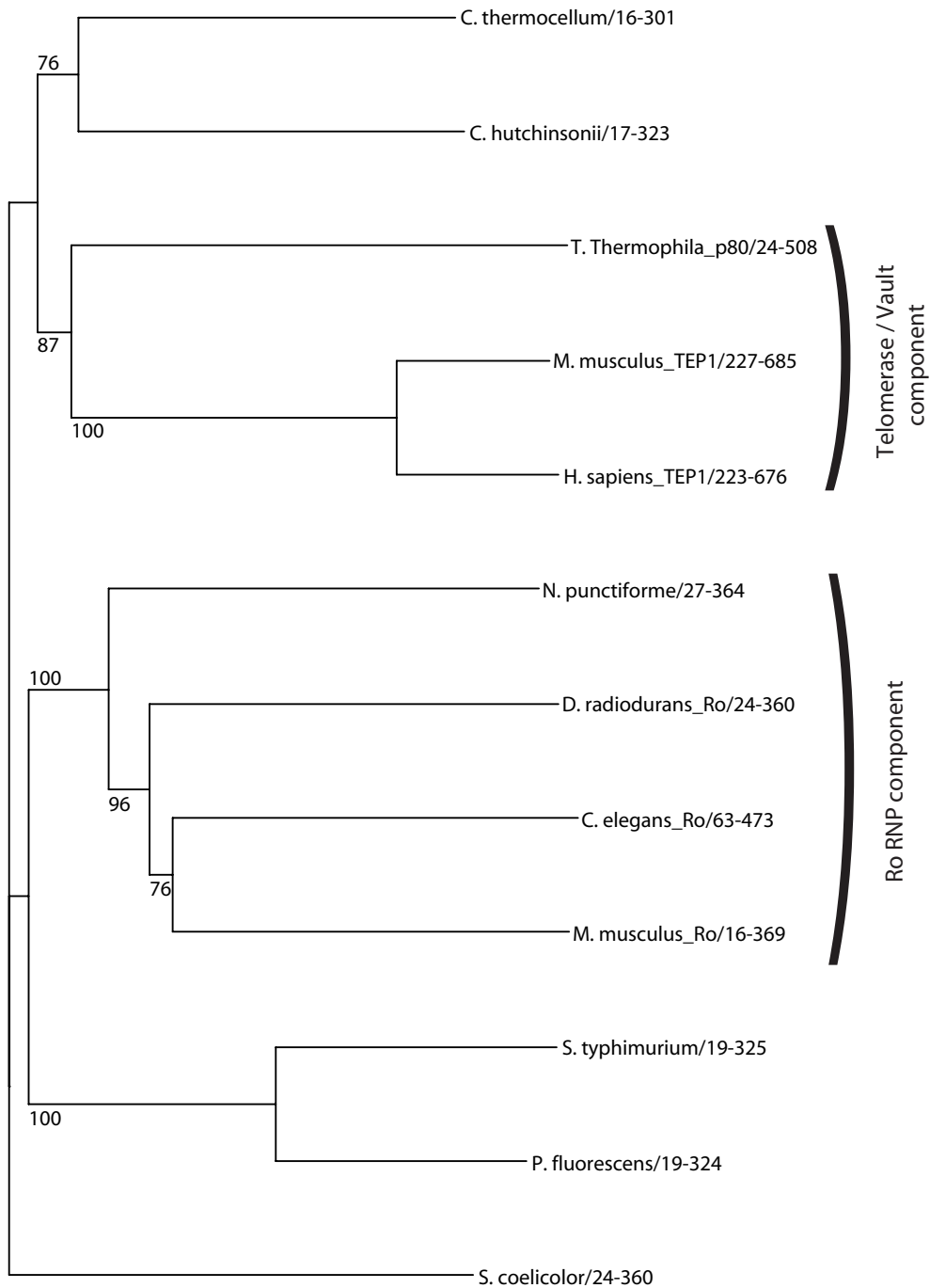


Figure 3

A tree constructed using the neighbour-joining algorithm implemented in the QuickTree program [19]. 500 bootstrap replicates were used and values over 75% are shown.

part of a Ro-like RNP that even contains a Y-like RNA molecule [7]. Building a phylogenetic tree of the TROVE module alignment, see figure 3, shows that the *Deinococcus* homologue does indeed cluster with the known Ro60 proteins, as does the *Nostoc punctiforme* homologue. It seems likely that *Nostoc punctiforme* also contains a Ro RNP. The other bacterial homologues cannot be attributed to either the TEP1-like or Ro60-like subfamilies. So we cannot assign any function to these proteins except that they may be part of an as yet unidentified RNP complex. Given the wide but patchy distribution of the TROVE module containing proteins we suggest that they are an ancient RNA-binding component of RNP complexes.

Author's Contributions

AB carried out sequence analysis and produced figures. AB and VK authored the manuscript.

Acknowledgements

We would like to thank Ewan Birney for providing an expert opinion on the provenance of the RRM in the Ro60 proteins. AB is supported by the Wellcome Trust. We would also like to thank Mike Poderycki for useful comments on the manuscript.

References

- Nakayama J, Saito M, Nakamura H, Matsuura A and Ishikawa F: **TLP1: a gene encoding a protein component of mammalian telomerase is a novel member of WD repeats family.** *Cell* 1997, **88**:875-84.
- Harrington L, McPhail T, Mar V, Zhou W, Oulton R, Bass MB, Arruda I and Robinson MO: **A mammalian telomerase-associated protein.** *Science* 1997, **275**:973-7.
- Kickhoefer VA, Stephen AG, Harrington L, Robinson MO and Rome LH: **Vaults and telomerase share a common subunit, TEP1.** *J Biol Chem* 1999, **274**:32712-7.
- Chugani DC, Rome LH and Kedersha NL: **Evidence that vault ribonucleoprotein particles localize to the nuclear pore complex.** *J Cell Sci* 1993, **106**:23-9.
- Suprenant KA: **Vault ribonucleoprotein particles: sarcophagi, gondolas, or safety deposit boxes?** *Biochemistry* 2002, **41**:14447-54.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucl Acids Res* 1997, **25**:3389-3402.
- Chen X, Quinn AM and Wolin SL: **Ro ribonucleoproteins contribute to the resistance of *Deinococcus radiodurans* to ultraviolet irradiation.** *Genes Dev* 2000, **14**:777-82.
- Kickhoefer VA, Siva AC, Kedersha NL, Inman EM, Ruland C, Streuli M and Rome LH: **The 193-kD vault protein, VPARP, is a novel poly(ADP-ribose) polymerase.** *J Cell Biol* 1999, **146**:917-28.
- van Zon A, Mossink MH, Schoester M, Scheffer GL, Scheper RJ, Sonneveld P and Wiemer EA: **Structural domains of vault proteins: a role for the coiled coil domain in vault assembly.** *Biochem Biophys Res Commun* 2002, **291**:535-41.
- Whittaker CA and Hynes RO: **Distribution and evolution of von willebrand/integrin a domains: widely dispersed domains with roles in cell adhesion and elsewhere.** *Mol Biol Cell* 2002, **13**:3369-87.
- Collins K, Kobayashi R and Greider CW: **Purification of Tetrahymena telomerase and cloning of genes encoding the two protein components of the enzyme.** *Cell* 1995, **81**:677-86.
- Van Horn DJ, Eisenberg D, O'Brien CA and Wolin SL: **Caenorhabditis elegans embryos contain only one major species of Ro RNP.** *Rna* 1995, **1**:293-303.
- Birney E, Kumar S and Krainer AR: **Analysis of the RNA-recognition motif and RS and RGG domains: conservation in meta-zoan pre-mRNA splicing factors.** *Nucleic Acids Res* 1993, **21**:5803-16.
- Cerutti L, Mian N and Bateman A: **Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain.** *Trends Biochem Sci* 2000, **25**:481-482.
- Katoh K, Misawa K, Kuma K and Miyata T: **MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.** *Nucleic Acids Res* 2002, **30**:3059-66.
- Goodstadt L and Ponting CP: **CHROMA: consensus-based colouring of multiple alignments for publication.** *Bioinformatics* 2001, **17**:845-6.
- Bateman A, Birney E, Cerruti L, Durbin R, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M and Sonnhammer ELL: **The Pfam protein families database.** *Nucleic Acids Res* 2002, **30**:276-280.
- Letunic I, Goodstadt L, Dickens NJ, Doerks T, Schultz J, Mott R, Ciccarelli F, Copley RR, Ponting CP and Bork P: **Recent improvements to the SMART domain-based sequence annotation resource.** *Nucleic Acids Res* 2002, **30**:242-4.
- Howe K, Bateman A and Durbin R: **QuickTree: building huge Neighbour-Joining trees of protein sequences.** *Bioinformatics* 2002, **18**:1546-7.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

