Protist, Vol. 161, 434–451, July 2010
http://www.elsevier.de/protis
Published online date 27 January 2010

Protist

ORIGINAL PAPER

Secondary Structure Models for the Internal Transcribed Spacer (ITS) Region 1 from Symbiotic Dinoflagellates

Daniel J. Thornhill[1], and Jenna B. Lord

Department of Biology, Bowdoin College, 6500 College Station, Brunswick, ME 04011, USA

Submitted August 22, 2009; Accepted November 21, 2009
Monitoring Editor: Michael Melkonian

Ribosomal genes and their spacers have been extensively utilized to examine the biodiversity and phylogenetics of protists. Among these, the internal transcribed spacer regions 1 and 2 (ITS1 and ITS2) are known to form secondary structures that are critically important for proper processing of the pre-rRNA into mature ribosomes. Although the secondary structure of ITS2 has been widely investigated, considerably less is known about ITS1 and its secondary structure. Here, secondary structures of the ITS1 were modeled for 46 ITS "types" from *Symbiodinium*, a diverse dinoflagellate genus that forms symbioses with many protists and metazoans, using comparative phylogenetic and minimum free energy approaches. The predicted ITS1 secondary structures for each *Symbiodinium* "type" were highly stable ($\Delta G = -46.40$ to -85.30 kcal mol^{-1} at 37 °C) and consisted of an open loop with five helices separated by single-stranded regions. Several structural characteristics were conserved within monophyletic sub-groups, providing additional support for the predicted structures and the relationships within this genus. Finally, the structures were applied to identify potential pseudogenes from five *Symbiodinium* ITS1 datasets. Consequently, ITS1 secondary structures are useful in understanding the biology and phylogenetics, as well as recognizing and excluding questionable sequences from datasets, of protists such as *Symbiodinium*.

Key words: coral reef; compensatory base change (CBC); endosymbiont; ITS1; ribosome; rRNA; pre-rRNA processing; pseudogene; *Symbiodinium*.

Introduction

Ribosomes are essential components of all cells due to their integral function in protein biosynthesis. In eukaryotes, the RNA component of these complexes is partially encoded as tandemly repeated rDNA operons consisting of small subunit (18S) and large subunit (5.8S and 28S) rRNA genes separated by external and internal transcribed spacer regions (ETS and ITS) (Maroteaux et al. 1985; Perry 1976). During ribosome production, the entire operon is transcribed as a single pre-rRNA molecule by RNA polymerase I and the spacer regions are subsequently removed during rRNA maturation (Perry 1976). Although the transcribed spacer regions do not become part of the functional ribosome, they do form secondary structures that are critical for appropriate processing of the transcript (Côté and Peculis 2001; Liu and Schardl 1994; Michot et al. 1999; Musters et al. 1990; van Nues et al. 1995) and mutations that modify these secondary structures are oftentimes not tolerated (Abeyranthne et al.

[1]Corresponding author; fax +1 207 725 3405
e-mail thornhill.dan@gmail.com (D.J. Thornhill).

doi:10.1016/j.protis.2009.11.004

2002; Abeyranthne and Nazar 2005; Beltrame and Tollervey 1992; Beltrame et al. 1994; Lalev and Nazar 1998; Lalev et al. 2000; Musters et al. 1990; van der Sande et al. 1992; van Nues et al. 1994). In yeast for instance, changes to the secondary structure of internal transcribed spacer region 2 (ITS2) result in dramatically decreased production of ribosomal subunits and cellular growth rates (van Nues et al. 1995).

Of the two internal transcribed spacer regions, the secondary structure of ITS2 has been well-explored in many protists and other eukaryotes. Based on this work, it is now recognized that the ITS2 region from a wide-range of taxa forms a secondary structure consisting of a closed loop with four double-stranded helices separated by single-stranded regions (e.g., Coleman 2007; Goertzen et al. 2003; Gottschling and Plötner 2004; Hunter et al. 2007; Kocot and Santos 2009; Mai and Coleman 1997; Schultz et al. 2005). In contrast, considerably less is known about the secondary structure of the other internal transcribed spacer, region 1 (ITS1). This is somewhat surprising because ITS1 secondary structure is also a critical component of ribosome maturation (e.g., Abeyranthne and Nazar 2005; Lalev et al. 2000; Musters et al. 1990; van Nues et al. 1994). In the eukaryotic taxa examined thus far, the ITS1 secondary structure is variable from one taxonomic group to another, but typically forms an open loop containing multiple double-stranded helices (Abeyranthne and Nazar 2005; Campbell et al. 2005; Goertzen et al. 2003; Gottschling et al. 2001; Itskovich et al. 2008; Lalev and Nazar 1998; Tippery and Les 2008; von der Schulenburg et al. 2001; Yeh et al. 1990). The secondary structural arrangement of the ITS1 is notably under-explored among protists, with only a handful of studies addressing this topic (e.g., Beiggi and Piercey-Normore 2007; Coleman et al. 1998; De Jonckheere 1998; Gottschling and Plötner 2004; Ki and Han 2007). Here, we contribute to this area through the proposal of ITS1 secondary structures for members of the dinoflagellate genus *Symbiodinium*.

Symbiodinium is a diverse group of photosynthetic dinoflagellates that forms symbioses with various protists and invertebrate metazoans, including reef-building corals. These symbiotic relationships are responsible, in large part, for the high rates of primary productivity and calcification in coral reef ecosystems (reviewed in Barnes and Chalker 1990; Coffroth and Santos 2005; Trench 1993). Previous studies using various molecular markers, including rRNA genes, have revealed that *Symbiodinium* consists of eight genetically divergent Clades (designated A to H; Pochon et al. 2006) that can be further distinguished into sub-cladal "types" according to differences in molecular genetics, physiology, biochemistry, ecology, and host-specificity (Coffroth and Santos 2005; Trench 1993 and references within). Many recent *Symbiodinium* diversity studies have specifically targeted the ITS regions (ITS1 and ITS2) because the variability of these genes enables sub-cladal "types" to be differentiated in an ecologically meaningful context (e.g., Baillie et al. 2000; Goulet et al. 2008; LaJeunesse 2002; Magalon et al. 2007; Pawlowski et al. 2001; Pochon et al. 2001; Sampayo et al. 2007; Thornhill et al. 2008; van Oppen 2004; van Oppen et al. 2001).

Given the above, a considerable amount of *Symbiodinium* ITS1 data are available for exploring secondary structural evolution in this important protistan group. Currently, a single ITS1 structural model is available for only one lineage of *Symbiodinium* Clade C (i.e., isolate 1591, Gottschling and Plötner 2004). In this study, we expand upon this previous work by elucidating the ITS1 secondary structures of 46 *Symbiodinium* "types" representing Clades A through H. In addition to potentially improving knowledge of the basic biology, phylogenetics, and ecology of these endosymbionts, the secondary structural models presented here form a comparative framework for exploring ITS1 structure and evolution in other protists.

Results

Secondary Structural Characteristics of *Symbiodinium* ITS1

Secondary structures of the ITS1 from 46 *Symbiodinium* ITS1 "types" in the currently recognized sub-generic Clades A-H were modeled via comparative phylogenetic and minimum free energy approaches. These ITS1 sequences varied between 198 and 247 nucleotides in length[2] and had a % GC content ranging between

[2]Sequence length determinations were based upon the results of Gottschling and Plötner (2004) and Thornhill et al. (2007). Although the ITS1 sequence lengths reported here may differ slightly from these sequences' GenBank annotation, this length variation has no notable impact on the results and discussion presented here because sequence endpoints predominantly occur in the single-stranded regions of the structures (Fig. 1, Table 2).