Molecular Phylogenetics and Evolution 105 (2016) 160–165

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Evolution under pressure and the adaptation of visual pigment compressibility in deep-sea environments

Megan L. Porter [a,1], Nicholas W. Roberts [b,*,1], Julian C. Partridge [c]

[a] Department of Biology, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA
[b] School of Biological Sciences, University of Bristol, Tyndall Avenue, Bristol BS8 1TQ, UK
[c] School of Animal Biology and the Oceans Institute, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

ARTICLE INFO

Article history:
Received 18 April 2016
Revised 22 July 2016
Accepted 11 August 2016
Available online 12 August 2016

Keywords:
Opsin
Evolution
Deep-sea
Compressibility
Teleost
Cephalopod

ABSTRACT

Understanding the link between how proteins function in animals that live in extreme environments and selection on specific properties of amino acids has proved extremely challenging. Here we present the discovery of how the compressibility of opsin proteins in two evolutionarily distinct animal groups, teleosts and cephalopods, appears to be adapted to the high-pressure environment of the deep-sea. We report how in both groups, opsins in deeper living species are calculated to be less compressible. This is largely due to a common set of amino acid sites (bovRH# 159, 196, 213, 275) undergoing positive destabilizing selection in six of the twelve amino acid physiochemical properties that determine protein compressibility. This suggests a common evolutionary mechanism to reduce the adiabatic compressibility of opsin proteins. Intriguingly, the sites under selection are on the proteins' outer faces at locations known to be involved in opsin-opsin dimer interactions.

1. Introduction

The world's oceans cover approximately 72% of the Earth's surface and, at an average depth of 3800 m, the deep-sea represents the major portion of the planet's biosphere (Somero, 1992). As depth increases, so does the hydrostatic pressure, from 0.1 MPa at the surface to over 100 MPa in the deepest oceanic trenches: a thousand-fold range. Life, including macrofauna, is found at all depths, an observation that belies the extent to which pressure can affect biological function. Increasing pressure affects fundamental cellular processes such as enzyme activity, action potential propagation, synaptic transmission, and overall protein function and regulation (Campenot, 1975; Gross and Jaenicke, 1994; Somero, 1992).

In marine animals two types of changes are known to occur that maintain protein function at elevated hydrostatic pressures: extrinsic (changes in cellular and/or membrane composition) and intrinsic (amino acid substitutions leading to modifications of protein structure). An extrinsic change common to many deep-sea animals is that of homeoviscous adaptation to maintain membrane fluidity (Behan et al., 1992; Eguchi et al., 1994; Hazel and Williams, 1990; Somero, 1992). Another is the use of trimethylamine oxide (TMAO) to stabilize proteins (Yancey et al., 2014; Yancey and Siebenaller, 1999, 2015). Intrinsic amino acid changes that increase protein stability are also documented in several marine fish. A number of cytosolic proteins have been investigated in both shallow and deep living species (Stefanni et al., 2014), including lactate dehydrogenase (Brindley et al., 2008; Campenot, 1975; Gross and Jaenicke, 1994; Hennessey and Siebenaller, 1987; Somero, 1992), α-actin (Morita, 2010; Swezey and Somero, 1982; Wakai et al., 2014), and myosin heavy chain proteins (Morita, 2010). These studies have demonstrated that proteins from deep-sea species are functionally less susceptible to increased pressure than those from shallow living species. However, although such studies have suggested that specific amino acid changes may contribute to protein functional adaptation due to increased hydrostatic pressure, very few have made the link between protein molecular evolution (i.e. changes in amino acid residues), specific amino acid physicochemical properties, and the bulk protein property (i.e. compressibility) that is under natural selection (Wakai et al., 2014).

Transmembrane proteins, and in particular G-protein coupled receptors (GPCRs), are a diverse and extremely important class of proteins involved in cell signaling (Bockaert and Pin, 1999). Several studies have demonstrated that GPCRs are extremely susceptible to perturbation by high pressures (Siebenaller and Garrett, 2002; Siebenaller and Murray, 1999; Somero, 1992). The GPCR (i.e. opsin protein) based visual systems in many deep-sea organisms implies that these proteins have evolved to function under the increased

* Corresponding author.
E-mail address: nicholas.roberts@bristol.ac.uk (N.W. Roberts).
[1] Both authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ympev.2016.08.007
1055-7903/

hydrostatic pressures of the deep sea. Studying GPCR function relative to increased pressure is notoriously difficult, as removing the GPCR proteins from the membrane for manipulation alters the overall physiological function of the protein. Therefore, to investigate potential intrinsic changes that have evolved in GPCRs in high-pressure, deep-sea environments, we obtained gene sequences of the transmembrane GPCR opsin from two independent taxonomic groups of marine animals: 128 species of teleost fish and 65 species of cephalopods, representing animals living at a wide range of depths. In fish we investigated the C-type rod opsin, Rh1, and in cephalopods the R-type opsin, representing the two main evolutionarily distinct groups of opsin proteins used in animal vision (Porter et al., 2012). Furthermore, each opsin amino acid sequence provides a convenient analytical way to calculate the adiabatic compressibility of the protein by using well-tested empirical relationships based on the individual physicochemical properties of each constituent amino acid (Gromiha and Ponnuswamy, 1993, and Table 1 therein) (Table S1). Using these taxonomically diverse and independently evolved sets of sequences, we reconstructed the phylogenetic history of each set of opsins, calculated the compressibility for each protein, and used comparative evolutionary methods to look at the evolution of opsin protein compressibility in relation to living depth in order to test the hypothesis that species living at deeper depths have less compressible opsin proteins.

2. Materials and methods

2.1. Taxon sampling and phylogenetic reconstruction

Opsin sequence data were collected from GenBank for exclusively marine species from two major taxonomic groups: Rh1 sequences from 128 species of fish within the taxon Teleostei and rhodopsin sequences from 65 species of cephalopods (see Supplementary Data for full lists of species and accession numbers). Within both the fish and cephalopod datasets, multiple independent opsin lineages were represented that included shallow living, mid-water, and deep-sea species (Figs. S1, S2, Table S1). Each DNA dataset was translated to amino acid sequence and aligned using the MAFFT v6 online server (Katoh and Toh, 2008). For the fish dataset, this resulted in an alignment of 281 amino acids, spanning from the beginning of transmembrane helix I to the end of transmembrane helix VII, and a cephalopod opsin alignment of 213 amino acids from the beginning of transmembrane helix IV to past the cytoplasmic helix VIII in the C-terminal tail.

The aligned datasets were used to reconstruct maximum likelihood phylogenies (Fig. S1 (fish opsin) and S2 (cephalopod opsin)) using Randomized Axelerated Maximum Likelihood (RAxML) v7.2.8 with rapid bootstrapping (100 iterations) as implemented on the Cyberinfrastructure for Phylogenetic Research (CIPRES) Portal v3.1 (Miller et al., 2010; Stamatakis, 2006). Based on preliminary phylogenetic analyses including non-teleost sarcopterygian and agnathan marine fish (e.g. *Latimeria* and *Petromyzon*, *respectively*), the opsin from *Clupea harengus* (EU492243) was determined to be the basal marine teleost sequence, and therefore was used to root the phylogeny. Based on preliminary phylogenetic analyses including non-cephalopod molluscs (e.g. the bivalve scallop, *Mizuhopecten yessoensis* AB006454), the opsin from *Vampyroteuthis infernalis* (AY545563) was determined to be the basal cephalopod sequence, and therefore was used to root the phylogeny.

2.2. Depth and compressibility calculations

For each species in the opsin datasets, median depth was calculated from catch depth records obtained from the Ocean Biogeographic Information System (OBIS) website (UNESCO) (Table S1). Species with less than five catch records were removed from the dataset. For the remaining species, all records were used to calculate a median catch depth. For the analyses of positive selection, species with median catch depths of 500 m or deeper were considered to be deep-sea species. This was chosen as a conservative measure of depth, given the biased nature of catch depth data, which are skewed towards shallow depths.

Adiabatic compressibility (β_s) is the relative change in the volume of the system per unit adiabatic change in its pressure, and can be estimated for proteins based solely on amino acid sequences (Gromiha and Ponnuswamy, 1993). Briefly, the twelve amino acid properties with the highest correlations to experimentally determined protein compressibility were used to compute a regression equation (Gromiha and Ponnuswamy, 1993, Eq. (5)):

$$\begin{aligned}\beta_s = {} & 56.5064 + 0.3057K^0 + 33.3672P_t - 47.4189P_c \\ & - 102.1064F + 14.0347H_t + 1.7694M_w - 0.7816B_t \\ & - 0.5990\mu - 11.3579P_\alpha + 5.1392R_\alpha + 29.1553\alpha_n \\ & - 2.1465V^0, \end{aligned} \quad (1)$$

where the amino acid physicochemical properties are:

K^0 = compressibility
P_t = turn tendency
P_c = coil tendency
F = mean r.m.s. fluctuational displacement
H_t = thermodynamic transfer hydrophobicity
M_w = molecular weight
B_t = bulkiness
μ = refractive index
P_α = α-helix tendency
R_a = solvent accessible reduction ratio
α_n = power to be at the N-terminus of an α helix
V^0 = partial specific volume.

Using this empirical relationship developed from experimental data, protein adiabatic compressibility can be calculated from its amino acid sequence alone; for each specific property, the total number of each amino acid type, multiplied by that specific property value, is summed across all 20 amino acids. To automate these calculations, a program was written in LabView 2010 (National Instruments) to calculate β_s for a given amino acid sequence (available upon request). To ensure that calculations were comparable across species with different portions of the opsin protein sequence available, we conducted tests comparing the calculated compressibility of an entire opsin protein to compressibility of datasets with the N- and C-terminal tails not included or including just the transmembrane or just the loop regions. These tests showed that the trend of less compressible opsin proteins in deeper living species was detected in all dataset partitions, although strongest in the transmembrane portions of the protein. Therefore, we proceeded to use our program to calculate the compressibility of every opsin amino acid sequence in both aligned datasets generated for phylogeny reconstruction (see below), representing ~48% of the total cephalopod and ~80% of the complete fish rhodopsin proteins.

2.3. Phylogenetic comparative analyses

For both squid and cephalopod datasets, we investigated the relationship between calculated opsin protein compressibility and mean catch depth for all species, using phylogenetic generalized least squares (PGLS), which accounts for the shared evolutionary history of genes. PGLS regressions were run in the 'caper' package of the software program 'R' with ultrametric phylogenies