

The Evolutionary Portrait of Metazoan NAD Salvage

João Carneiro^{1,2,3}, Sara Duarte-Pereira^{1,3}, Luísa Azevedo¹, L. Filipe C. Castro³, Paulo Aguiar⁴, Irina S. Moreira⁵, António Amorim^{1,2}, Raquel M. Silva^{1*}

1 IPATIMUP-Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal, **2** Faculty of Sciences, University of Porto, Porto, Portugal, **3** Interdisciplinary Centre for Marine and Environmental Research (CIIMAR), CIMAR Associate Laboratory, University of Porto, Porto, Portugal, **4** CMUP - Centro de Matemática da Universidade do Porto, Porto, Portugal, **5** REQUIMTE - Rede de Química e Tecnologia, Faculty of Sciences, University of Porto, Porto, Portugal

Abstract

Nicotinamide Adenine Dinucleotide (NAD) levels are essential for cellular homeostasis and survival. Main sources of intracellular NAD are the salvage pathways from nicotinamide, where Nicotinamide phosphoribosyltransferases (NAMPTs) and Nicotinamidases (PNCs) have a key role. NAMPTs and PNCs are important in aging, infection and disease conditions such as diabetes and cancer. These enzymes have been considered redundant since either one or the other exists in each individual genome. The co-occurrence of *NAMPT* and *PNC* was only recently detected in invertebrates though no structural or functional characterization exists for them. Here, using expression and evolutionary analysis combined with homology modeling and protein-ligand docking, we show that both genes are expressed simultaneously in key species of major invertebrate branches and emphasize sequence and structural conservation patterns in metazoan NAMPT and PNC homologues. The results anticipate that NAMPTs and PNCs are simultaneously active, raising the possibility that NAD salvage pathways are not redundant as both are maintained to fulfill the requirement for NAD production in some species.

Citation: Carneiro J, Duarte-Pereira S, Azevedo L, Castro LFC, Aguiar P, et al. (2013) The Evolutionary Portrait of Metazoan NAD Salvage. PLoS ONE 8(5): e64674. doi:10.1371/journal.pone.0064674

Editor: Andrew R. Dalby, University of Westminster, United Kingdom

Received: February 3, 2013; **Accepted:** April 16, 2013; **Published:** May 28, 2013

Copyright: © 2013 Carneiro et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Portuguese Foundation for Science and Technology (PTDC/BIA-PRO/099888/2008). IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Education and Science, and is partially supported by FCT. LA and RMS are supported by FCT Ciência2007 (Hiring of PhDs for the SCTN - financed by POPH - QREN - Typology 4.2 - Promoting Scientific Employment, co-financed by MES national funding and The European Social Fund). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: raquelsilva@ua.pt

☛ These authors contributed equally to this work.

Introduction

Nicotinamide Adenine Dinucleotide (NAD) is an essential molecule to cells. As a cofactor in redox reactions, NAD regulates the metabolism and energy production and, as a substrate for NAD-consuming enzymes such as poly(ADP-ribose) polymerases (PARPs) and sirtuins, NAD is involved in DNA repair, transcriptional silencing and cell survival [1]. To maintain adequate NAD levels, several routes are used for NAD synthesis that depend on distinct precursors: *de novo* pathways synthesize NAD from tryptophan or aspartic acid whereas salvage pathways recycle NAD from nicotinamide (Nam), nicotinic acid (Na) and their ribosides [2–4].

The nicotinamide salvage pathway is the major source of intracellular NAD in humans [5,6] and is also required for growth in several microorganisms [7–10]. NAD salvage from Nam is a two- or four-step reaction, in which the rate-limiting enzymes and functional homologues are, respectively, nicotinamide phosphoribosyltransferases (NAMPTs) and nicotinamidases (PNCs) [11–13]. In humans, *NAMPT* is widely studied due to its involvement in inflammation and disease such as cancer [14,15]. In contrast, humans lack nicotinamidase but expression of the *Drosophila* Pnc protects human neuronal cells from death originated by oxidative stress [16]. Moreover, an increased Pnc1 and sirtuin activity confers protection to proteotoxic stress in yeast and *C. elegans* [17,18]. The yeast Pnc1 is a biomarker of stress and a regulator of sirtuin activity [11,18], and thus, most studies in yeast and

invertebrates have focused in the link between these enzymes and aging [16,19]. Notwithstanding, despite their importance to major cellular processes, there is a poor functional characterization of nicotinamidases [20,21] and their role in infection has been less explored [7,8,22].

NAMPTs and PNCs act as regulators of enzymes from similar pathways, coordinating the overall metabolism and stress responses [23]. Moreover, both are pharmacologically relevant. NAMPT inhibitors are being used in clinical trials as anti-cancer agents [24–27] and nicotinamidases are attractive targets to the development of drugs for infectious diseases and anti-parasitic therapies [7,8,22,28–30].

NAMPTs and PNCs do not co-occur in all organisms and, until very recently, lineages with both *NAMPT* and *PNC* had been only found in bacteria and algae [30–32]. *NAMPT* was thought to be absent from invertebrates but the discovery that *NAMPT* homologues are present in several invertebrate species and that some species have both *NAMPT* and *PNC* homologues [33] challenged the classical view that these enzymes are redundant and mutually exclusive [1], emphasizing the need for studies characterizing the structural and functional properties of these enzymes.

Motivated by the lack of information for *NAMPT* and *PNC* homologues in relevant invertebrate species, which would render the biological meaning of simultaneous *versus* unique occurrence of these proteins more evident, we carried out an integrated study to

establish gene expression, amino acid conservation and structural comparisons. We provide experimental evidence that both genes are expressed simultaneously in key invertebrate species. In addition, evolutionary conserved patterns at the amino acid sequence and at the structural levels were detected. Also, using homology modeling and protein-ligand docking, we identify the amino acids that bind Nam in the active sites of invertebrate NAMPTs and PNCs. Taken together, the results suggest that invertebrate NAMPTs and PNCs are concurrently functional and, thus, that both NAD salvage pathways might not be redundant.

Results

Expression of invertebrate NAMPTs and PNCs

NAMPT homologues have been previously found in the vibriophage KVP40 [34], bacteria [10,32], and the unicellular green algae *Chlamydomonas reinhardtii* [31], motivating the search for NAMPT homologues in invertebrates, some of which simultaneously have PNC sequences [33] (Table S1). No recognizable NAMPT homologue has been detected so far in representative species of the phyla Arthropoda or Nematoda, although NAMPT and PNC were found in more basal lineages such as the choanoflagellate *Monosiga brevicollis* and the sea anemone *N. vectensis* (Figure 1). Such phylogenetic distribution is consistent with a scenario where both genes were present in the Metazoan ancestor and were selectively lost in specific lineages, as evidenced by the different patterns in protostomes. Namely, both genes were found in lophotrochozoans that includes mollusks (*Lottia gigantea*) and annelids (*C. teleta* and *Helobdella robusta*), and the absence of NAMPT was observed in ecdysozoans such as nematodes and arthropods. In deuterostomes, which comprises chordates, hemichordates and echinoderms, both genes were likely present in early lineages, which is supported by the evidence from the extant *B. floridae*, *Saccoglossus kowaleskii* and *S. purpuratus* species, but NAMPT was secondarily lost in the urochordate *Ciona intestinalis* while PNC was lost in vertebrates (Figure S1). RT-PCR of selected species showed that both NAMPT and PNC genes are expressed in the adult forms of *Branchiostoma floridae* (Cephalochordata), *Strongylocentrotus purpuratus* (Echinodermata), *Capitella teleta* (Annelida) and *Nematostella vectensis* (Cnidaria) (Figure 1). In addition, available EST (Expressed Sequence Tag) data indicates that NAMPT and PNC genes are also co-expressed during developmental stages (Table S2), suggesting a widespread usage of both Nam salvage pathways across Metazoans.

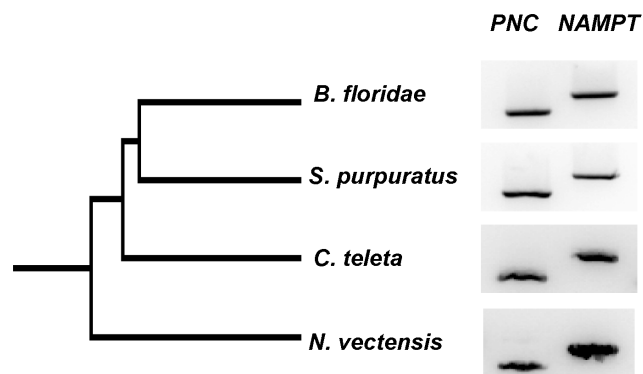


Figure 1. NAMPT and PNC homologues are co-expressed in invertebrates. RT-PCR analysis shows that NAMPT and PNC are simultaneously expressed in *Branchiostoma floridae*, *Strongylocentrotus purpuratus*, *Capitella teleta* and *Nematostella vectensis*. doi:10.1371/journal.pone.0064674.g001

Evolutionary divergence of NAMPTs and PNCs

We have further characterized the evolutionary divergence of NAMPT and PNC homologues, measured as protein distances calculated from amino acid sequence alignments (Figure 2). The resulting matrix (Figure 2) showed that NAMPT is conserved, even when large evolutionary distances are considered. For example, the divergence between the human and cnidarian (*N. vectensis*) NAMPT homologues is about 50%, as much as when compared with amphioxus (*B. floridae*). Among invertebrates the sequences showing the smallest divergence are from *N. vectensis* and *C. teleta* (31.2%). Conversely, PNC sequences are highly divergent even in closely related species, as shown for the annelids *C. teleta* and *H. robusta*, or the basal chordates *B. floridae* and *C. intestinalis*. Curiously, the smallest divergence between PNC sequences was found for *C. teleta* and *B. floridae* (51.3%). This trend was also evident when we plotted protein distances taking implicitly in consideration the evolutionary divergence time between each pair of species studied (Movie S1 and Table S3). Analyses of protein distances (pd) indicated that NAMPT homologues are considerably more conserved ($pd = 0.447 \pm 0.116$) than PNC ($pd = 0.842 \pm 0.151$) (mean \pm std), which is remarkable for species spanning over 1000 million years of divergence (Table S3). For PNC proteins, in addition to the larger values, no correlation with evolutionary distance was observed, while NAMPT distances were smaller and increased consistently with the evolutionary distance (ed) between species. The Kendall rank correlation coefficient was used to measure the dependence between pd and ed , showing no relevant dependence between both quantities for PNC ($\tau = -0.052$). However, for NAMPT both quantities vary consistently ($\tau = 0.413$).

Motif conservation in NAMPTs and PNCs

We next used the previously constructed amino acid sequence alignments dataset to search for conserved motifs in NAMPT and PNC homologues. In line with the aforementioned results, analyses of NAMPT sequences (Figure 3A) revealed conserved amino acid motifs surrounding catalytic residues [24,25,35–37] Tyr18, Phe193, Asp219, His247, Asp279, Asp313, corresponding to the boxed amino acids in Figure 3. As well, Asp16 and Arg311, Gly353 and Asp354, and Gly384 that bind nicotinamide, ribose or phosphate, respectively, are preserved and the additional NMN interacting residues Arg196 and Gly383 in rat NAMPT [25] are present in all sequences analyzed. The amino acid stretches that represent the dimer interface are also conserved in invertebrate NAMPTs (Figure 3A and Figure S2), as previously shown for vertebrates [25].

Similar analyses on PNC homologues showed that, while overall amino acid sequence identity is low (Figure 3B), motifs surrounding metal-binding and catalytic residues (boxed amino acids) show up. Indeed, all PNC sequences have conserved residues that coordinate the metal ion (corresponding to *Saccharomyces cerevisiae* Asp51, His53 and His94) and the catalytic triad (*S. cerevisiae* Asp8, Lys122 and Cys167). The characteristic *cis*-peptide bond that has been identified in available nicotinamidase/pyrazinamidase structures also corresponds to conserved residues present in these species, namely Val-Ala in *Pyrococcus horikoshii*, *S. cerevisiae*, *Leishmania infantum* and *C. intestinalis* [7,38,39], Ile-Ala in *Mycobacterium tuberculosis*, *Acinetobacter baumannii*, *H. robusta* and *B. floridae* [40,41], or Val-Leu in *Streptococcus pneumoniae* [42], and are preceded by a conserved glycine that has a role in catalysis [38,40,41]. Additionally, mutations that lead to *M. tuberculosis* loss of pyrazinamidase activity have defined residues that delineate the active site scaffold [38], corresponding to *S. cerevisiae* Glu10, Asp12, Phe13, Leu20, His57, Trp91, Gly123, Tyr131, Ser132, Val162,

	Hs	Mm	Dr	Bf	Ci	Sp	Ct	Hr	Lg	Dm	Ce	Nv	Ta	Sc	Mb
<i>H. sapiens</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>M. musculus</i>	0,035	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>D. rerio</i>	0,119	0,114	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>B. floridae</i>	0,483	0,498	0,494	-	0,618	0,607	0,513	0,855	0,944	0,664	0,841	0,533	-	0,898	0,749
<i>C. intestinalis</i>	-	-	-	-	-	0,687	0,749	0,898	0,991	0,749	0,898	0,629	-	0,869	0,749
<i>S. purpuratus</i>	0,509	0,509	0,505	0,341	-	-	0,736	0,787	0,898	0,787	0,944	0,736	-	1,007	0,724
<i>C. teleta</i>	0,532	0,525	0,513	0,374	-	0,367	-	0,913	0,944	0,585	0,913	0,675	-	1,058	0,787
<i>H. robusta</i>	0,589	0,58	0,597	0,436	-	0,494	0,443	-	0,841	1,007	1,007	0,828	-	1,04	0,814
<i>L. gigantea</i>	0,564	0,556	0,552	0,371	-	0,394	0,335	0,454	-	1,075	1,007	0,855	-	1,04	0,991
<i>D. melanogaster</i>	-	-	-	-	-	-	-	-	-	-	0,991	0,675	-	1,007	0,652
<i>C. elegans</i>	-	-	-	-	-	-	-	-	-	-	-	0,801	-	1,075	0,959
<i>N. vectensis</i>	0,505	0,502	0,49	0,325	-	0,397	0,312	0,502	0,364	-	-	-	-	0,869	0,749
<i>T. adhaerens</i>	0,548	0,56	0,544	0,377	-	0,443	0,367	0,505	0,415	-	-	0,377	-	-	-
<i>S. cerevisiae</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,093
<i>M. brevicollis</i>	0,576	0,576	0,556	0,367	-	0,45	0,415	0,505	0,461	-	-	0,408	0,454	-	-

Figure 2. Evolutionary divergence between NAMPT and PNC homologues. The estimates of evolutionary divergence were calculated as amino acid substitutions per site between NAMPT (green) and PNC (orange) sequences for several species. Hs, *Homo sapiens*; Mm, *Mus musculus*; Dr, *Danio rerio*; Bf, *Branchiostoma floridae*; Ci, *Ciona intestinalis*; Sp, *Strongylocentrotus purpuratus*; Ct, *Capitella teleta*; Hr, *Helobdella robusta*; Lg, *Lottia gigantea*; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*; Nv, *Nematostella vectensis*; Ta, *Trichoplax adhaerens*; Sc, *Saccharomyces cerevisiae*; Mb, *Monosiga brevicollis*. doi:10.1371/journal.pone.0064674.g002

Ala163, Tyr166 and Thr171, and most of them are conserved in all invertebrate PNC sequences as well (Figure 3B and Figure S3).

Genetic architecture conservation of NAMPT homologues

Given the degree of conservation of both proteins, and taking into account the divergence times of over 1000 million years between the species considered here, we have investigated the conservation at the gene structure and genome organization levels. *NAMPT* retained microsynteny in chordates, as indicated by the conserved gene order between *H. sapiens*, *M. musculus*, *D. rerio* and *B. floridae*, and also showed macrosynteny conservation in some lineages, namely between *Trichoplax adhaerens* and either *H. sapiens*, *N. vectensis* or *M. brevicollis* (Figure 4 and Figures S4, S5). For *PNC* homologues, no syntenic regions were found. Although recent studies point to a higher level of microsynteny conservation in metazoans than previously estimated [43], these evidences are challenging in some lineages due to poor genome annotation and breakdown in small scaffolds. At the level of exon-intron structure, *NAMPT* is more homogeneous than *PNC*, considering the number and size of exons, and total gene length (Figures S6, S7). The exception is observed in *N. vectensis*, where *NAMPT* has many small exons spanning 14 Kb in the genome, while *PNC* has only two exons in less than 2 Kb. Using the information on conserved motifs and gene structure, we were able to successfully identify *NAMPT* and *PNC* homologues as well as predict the corresponding gene structures in the hemichordate *S. kowaleskii*, a phylogenetic informative species (Figure S8).

Secondary structure conservation of PNC homologues

Nicotinamidase sequences are poorly conserved even in closely related species (Figures 2 and 3). Yet, considering some structures determined for archaea (*P. horikoshii*, PDB id: 1IM5), bacteria (*A. baumannii*, PDB id: 2WTA) and fungi (*S. cerevisiae*, PDB id: 2H0R), sharing only 30% protein identity (Figure 5A), the 3D structures are perfectly superimposable (Figure 5B). Such structural conservation is observed across the three domains of life, as all PNC enzymes share a similar core fold (Figure S9), with a potential

increase in complexity of the enzyme that is active as a monomer in *P. horikoshii* [38], dimer in *A. baumannii* [40] and heptamer in *S. cerevisiae* [39]. Thus, we have aligned PNC sequences based on secondary structure predictions and determined that invertebrate PNCs also show structural conservation (Figure 6). The regions corresponding to alpha-helices (red) and beta-sheets (yellow) are conserved at the structural level, even if the amino acids are not (Figure 6A). For instance, the alpha-helices of regions I, II and III comprise different amino acids while displaying a similar fold. To illustrate this, region II is shown in detail for *P. horikoshii*, *A. baumannii* and *S. cerevisiae* (Figure 6B).

Modeling and docking analyses of invertebrate NAMPTs and PNCs

To gain insight into the structures of invertebrate NAMPTs and PNCs, we have performed homology modeling and protein-ligand docking. To overcome limitations in the interpretation of results, we have used several templates to generate the models (Table S4). The LIGPLOT program was used to generate schematic diagrams between ligand (Nicotinamide, NCA) and receptor (NAMPT and PNC), which are shown in Figure 7. The prediction accuracy redocking test performed for the NAMPT (PDB 2E5D from *H. sapiens*) and PNC (PDB 3R2J from *L. infantum*), were in agreement with the ligand-receptor conformation in these X-ray structures. We obtained a similar active site ligand-receptor interaction for both NAMPT and PNC, which insure that the docking approach was accurate enough to be applied to the various molecular systems.

In NAMPT protein active site, all species, except *N. vectensis*, maintained most of the ligand-receptor interactions when compared with the structure of human NAMPT (Figure 7A). The homologous NAMPT of *B. floridae* has a hydrogen bond network that stabilizes the active site with two H-bonds between the side-chain of Arg-293 and the oxygen atom of the ligand. A similar bonding network can be observed in the human protein (PDB 2E5D) where Asp-219 binds to the nitrogen atom of the substrate (NCA). Hydrophobic interactions are similar when compared with

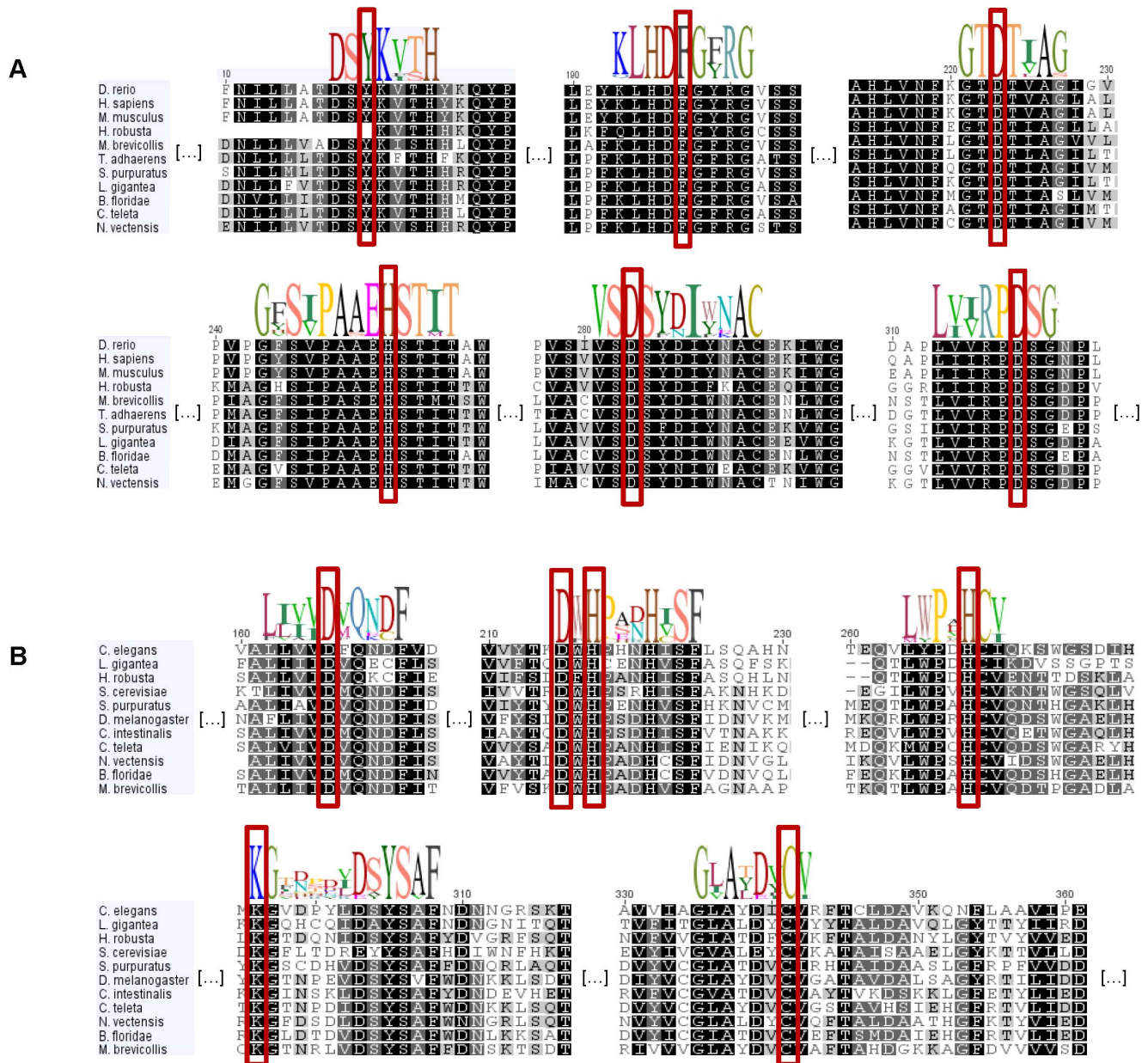


Figure 3. Amino acid motifs found in NAMPT and PNC homologues. The conserved amino acid motifs surrounding the active site residues (boxed) are shown as logos and displayed above the aligned sequences. NAMPT conservation is highlighted by the large blocks of identical amino acids that are found in the species analyzed (A). In PNC homologues, although the overall amino acid identity is low, the presence of conserved motifs is still detected throughout the species analyzed that range wide evolutionary distances (B).
doi:10.1371/journal.pone.0064674.g003

the human active site. In *C. teleta*, H-bond interactions between Arg-300 and NCA oxygen moiety and between Asp-209 and Asp-16 to both NCA nitrogens preserve the NCA conformation in the active site. Two hydrophobic interactions in *C. teleta* (Tyr-18 and Phe-183) with ligand atoms are not seen. In *N. vectensis* no H-bond interaction is present, but the most important hydrophobic interactions, Phe-182(B), Arg-298(B) and Tyr-17(A), are preserved. The H-bond interaction network of *S. purpuratus* shows that Asp-210(B) H-bond is maintained. Two other H-bonds, Tyr-19(A) and Glu-235(B), and hydrophobic interactions of the NCA ligand to Phe-184 (B) and Ala-233 (B) are also present. Globally, the NAMPT binding modes obtained by docking for the species analyzed shared the critical hydrophobic and hydrogen bonding

interactions and, if not (e.g. *N. vectensis*), the conformational status of NCA was maintained.

Next we also analyzed the conformational changes of PNC active and catalytic sites (flexible residues) in the four species (Figure 7B). In the *B. floridae* PNC, Phe-22, Trp-110, Val-182 and Cys-183 hydrophobic interactions contribute to the binding status of NCA. The three hydrogen bonding interactions (His-113 to NCA oxygen atom, Asp-62 to NCA nitrogen atom and His-113 to Asp-62) sustain the conformational position of the ligand. The Zn^{2+} keeps the strong binding to the ligand that was also present in *L. infantum* PNC (PDB 3R2J). In *C. teleta* the Trp-110 (hydrophobic interaction) and Tyr-147 (H-bond interaction) are the residues from the active site that play an important role in the ligand-

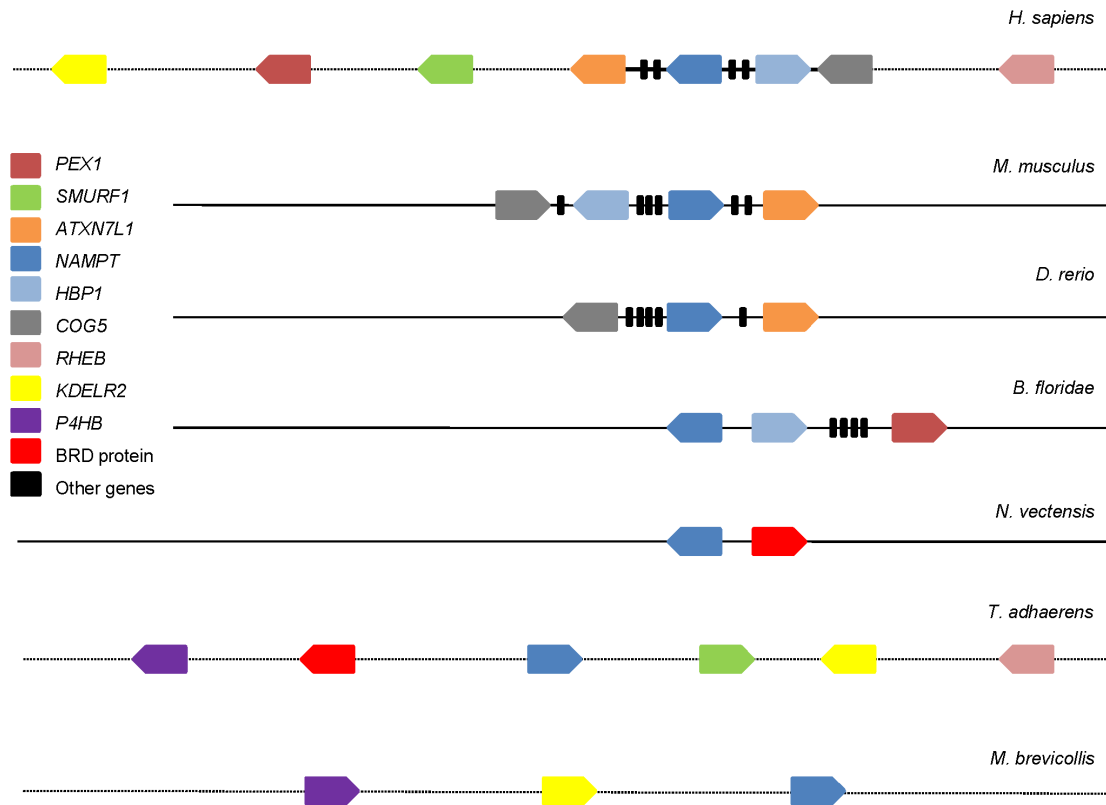


Figure 4. Syntenic organization of NAMPT homologues. Gene order and organization are represented for several lineages, and show conservation of microsynteny in chordates. *H. sapiens* chromosome 7, *M. musculus* chromosome 12, *D. rerio* chromosome 4, *B. floridae* scaffold 633, *N. vectensis* scaffold 360, *T. adhaerens* scaffold 2 and *M. brevicollis* scaffold 7 are displayed and dots indicate intervals containing multiple genes (>4). doi:10.1371/journal.pone.0064674.g004

receptor interaction. As in *B. floridae*, the His-113 has a hydrogen bond connection with NCA. Two other H-bond interactions not present in *L. infantum* PNC (Ser-70 and Lys-108) appear to be important to ligand binding. The interaction between Zn^{2+} and ligand is maintained. Although no significant changes in ligand conformation were observed, hydrogen bonds in *N. vectensis* were not predicted. When compared to 3R2J, hydrophobic interactions Cys-21, Trp-99, Ala-169 and Cys-173 are kept for the active site residues. Hydrophobic interactions for the catalytic residue Cys-173 are present, as well as a newly arisen Phe-11 interaction with the ligand. In *S. purpuratus*, hydrophobic contacts between Tyr-106, Trp-143, Ala-175 and the ligand are retained. Catalytic site residues Asp-17 and Cys-179 also bind to the ligand through an H-bond and a hydrophobic interaction. Two unique hydrogen bonds (Asp-57 and Leu-174) and hydrophobic contacts (His-109 and Phe-22) arise in the ligand-protein interaction. It can also be noticed that a conserved histidine (His-113 in *B. floridae* and *C. teleta*, and His-109 in *S. purpuratus*) maintains the interaction with the ligand.

Discussion

Nicotinamide phosphoribosyltransferases (NAMPTs) and nicotinamidases (PNCs) are the main NAD salvage enzymes and, until recently, were thought to occur in distinct lineages. Our data show that several Metazoan species have predicted homologues of both enzymes and that both genes are simultaneously expressed in *B. floridae*, *S. purpuratus*, *C. teleta* and *N. vectensis*. The distribution of NAMPT and PNC homologues points to the presence of both genes in early eukaryote evolution with selective gene loss and

retention in different animal lineages. Interestingly, loss of either one of the genes was predominantly found in fast evolving lineages, namely *D. melanogaster*, *C. elegans* and *C. intestinalis*, while slow-evolving species such as *B. floridae* retained both [44]. This is also reflected in genome architecture, with conserved *NAMPT* microsynteny in vertebrates and *B. floridae*.

We also highlight different conservation patterns in NAMPT and PNC homologues, at the protein amino acid sequence and at the 3D structural level. NAMPT sequences are highly conserved, as evidenced by small evolutionary divergences between species and long stretches of identical amino acids surrounding important catalytic and structural positions. As dimerization is required for NAMPT activity [25], in addition to active site residues that interact with the substrates and reaction products, amino acids that constitute the dimer interface are also conserved. For PNCs the sequence identity is lower, yet, critical amino acids are conserved and the overall fold is maintained in all the three domains of life. These are unifying features of nicotinamidases, even though there is a diversity of catalytic mechanisms described, with some exceptions concerning metal binding and metal ion coordination [7,20,21,29,41,42].

Homology modeling and protein-ligand docking indicates that active site residues and interactions of invertebrate NAMPTs with the substrate, nicotinamide, are similar to what is described for vertebrate NAMPTs [24,25,35–37]. In invertebrate PNCs, most interactions are maintained while additional hydrogen bonds and hydrophobic contacts were found. These new interactions might derive from complementary amino acid changes as a result of

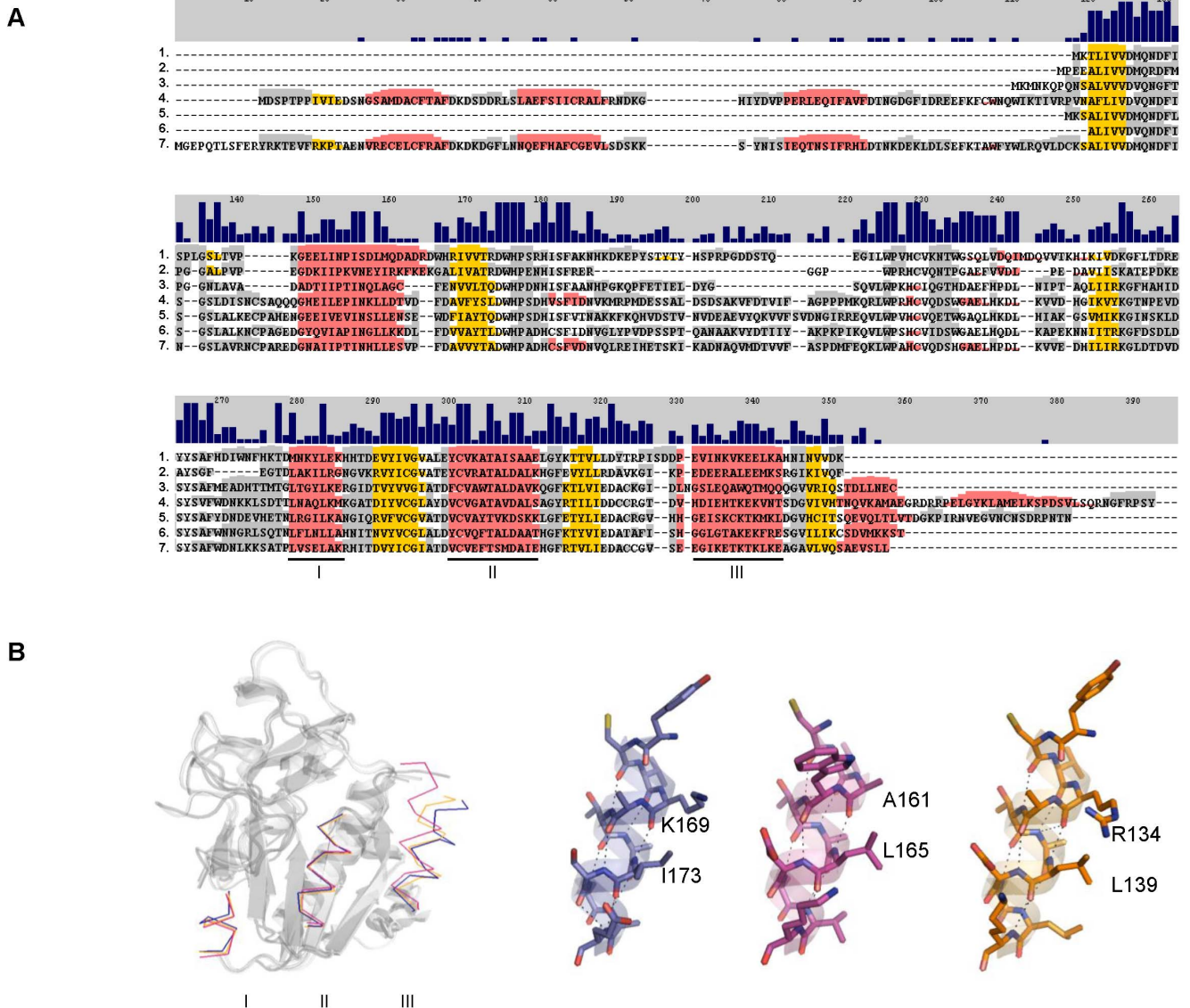


Figure 6. Predicted secondary structure of PNC homologues. (A) Aligned amino acid sequences of representative PNC homologues are displayed in function of the secondary structure. Alpha-helices are shown in red, beta-sheets are in yellow and grey represents coiled regions. Regions of structural conservation are highlighted in color even when the primary sequences are not conserved as demonstrated by the graphic bars above the sequences. 1, *Saccharomyces cerevisiae*; 2, *Pyrococcus horikoshii*; 3, *Acinetobacter baumannii*; 4, *Drosophila melanogaster*; 5, *Ciona intestinalis*; 6, *Nematostella vectensis*; 7, *Branchiostoma floridae*. (B) Alpha-helices I, II and III formed by groups of unrelated amino acids are structurally equivalent as shown by the 3D superimposition. In blue, *S. cerevisiae*; in pink, *A. baumannii*; and in yellow, *P. horikoshii*. doi:10.1371/journal.pone.0064674.g006

Molecular homology modeling

Prime [52] was used to search homologous proteins in NCBI PDB database (<http://www.rcsb.org/pdb/home/home.do>) for PNC and NAMPT. PDB templates (Table S4) were selected considering lowest e-values ($<1 \times 10^{-6}$), and structures without many missing residues (gaps $<20\%$). PNC and NAMPT sequences for the species *B. floridae*, *C. teleta*, *S. purpuratus* and *N. vectensis* were used to generate the alignments with homologue proteins. For secondary structure prediction the third-party program SSpro [53,54] was used and then all the templates re-aligned to the query sequence. The resulting alignment was used to build the protein models. The LigPrep [55] interactive optimizer (protein preparation wizard) with neutral pH was used to optimize the protein model. Finally hydrogens were added, bond order was assigned

and selenomethionines were converted to methionines for the generated models.

Molecular docking simulations

The 3D-structures of ligands were obtained from the PDB structures. The protein-ligand complexes were prepared with AutoDockTools [56,57]: hydrogen atoms were added for each protein and Kollman united atom charges assigned. Hydrogens were also added to the ligand (NCA) and charges were calculated by the Gasteiger-Marsili method. The rotatable bonds in the ligands were assigned with AutoDockTools. The Zn atom of PNC was assigned a charge of +2. AutoDock4.2 [56] was used to perform protein-ligand docking calculations. To insure the accuracy of our methodological approach we first have done

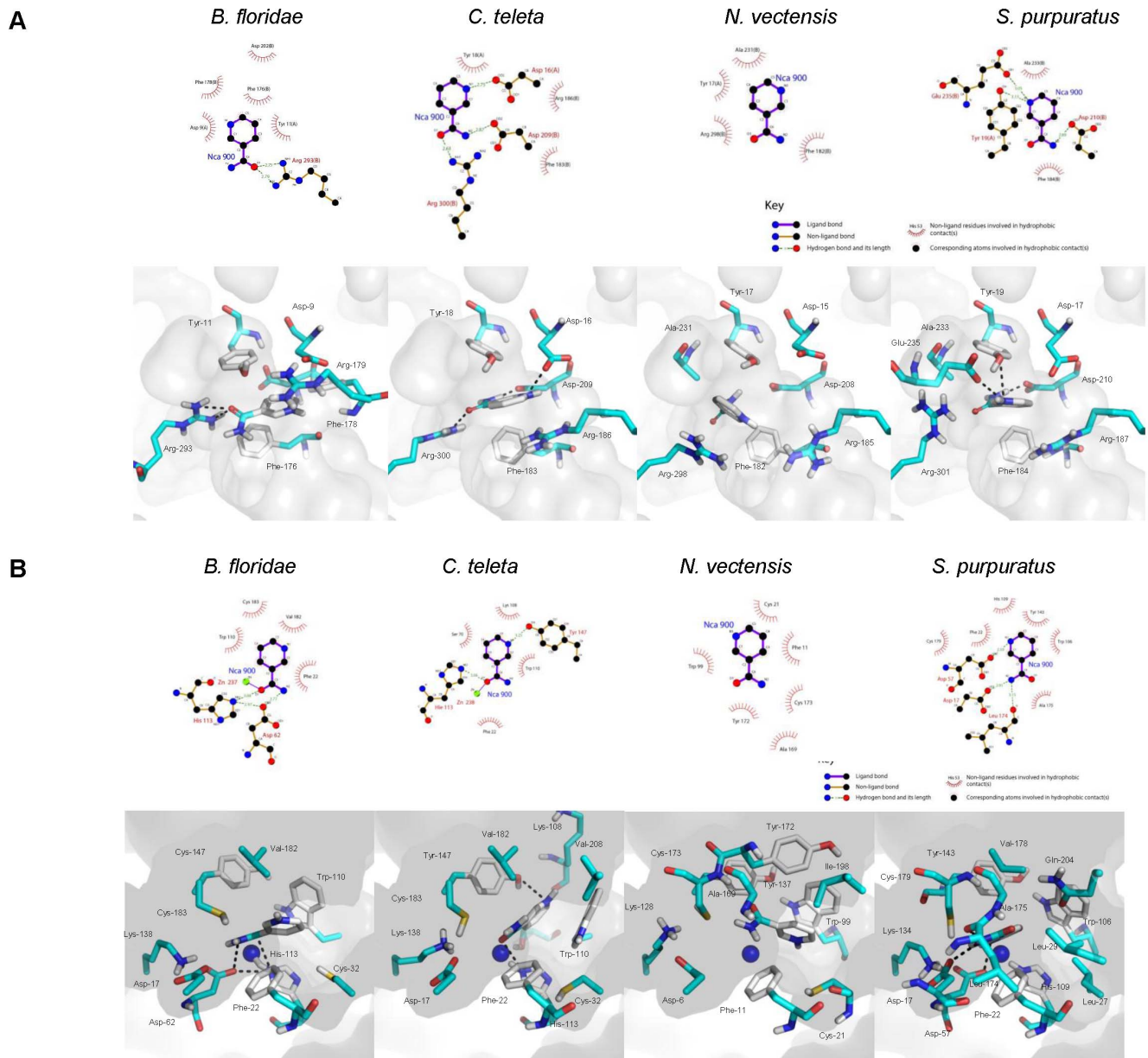


Figure 7. Hydrophobic interactions and hydrogen bonding network between the ligand (NCA) and the various receptors. NAMPT (A) and PNC (B) representations in LIGPLOT (upper panels) and PyMOL (lower panels) representations are shown for *Branchiostoma floridae*, *Capitella teleta*, *Nematostella vectensis*, and *Strongylocentrotus purpuratus*. The major binding determinants are represented in cyan stick. The Zn^{2+} atom is in blue vdW representation.

doi:10.1371/journal.pone.0064674.g007

redocking of the two most recently available X-ray structures (NAMPT PDB 2E5D and PNC PDB 3R2J) and then applied it to the various predicted protein-ligand systems. Various grid sizes were tested using as structural criteria the similarity between our docked results and the X-ray structure of *H. sapiens* NAMPT (2E5D) and *L. infantum* PNC (3R2J). We have selected a cubic grid box of $30 \times 25 \times 40$ Å for NAMPT and $35 \times 35 \times 40$ Å for PNC, centered on the C2–C5 ligand atoms distance mean with a grid spacing of 0.375 Å as shown in Table S4.

We considered the binding pockets described in the literature [7,36] (also shown in Table S5) to perform the flexible protein-ligand docking. The corresponding residues in the homology

alignment are described in Table S5. We performed the docking simulations using 100 independent Lamarckian genetic algorithm (LGA) runs, with the population size set to 200, the number of energy evaluations set to 10 000 000 and the maximum number of generations set to 27 000. All other parameters were used as default [58,59]. The results were analysed clustering together the conformations within a RMSD of 2 Å. The cluster with lower energy and with a conformation similar to the X-ray structure of NAMPT (PDB id: 2E5D) and PNC (PDB id: 3R2J) was selected for each species.

H-bonds and hydrophobic interactions for ligand-receptor molecules

Interactions between the ligand (NCA) and receptors (NAMPT and PNC) were calculated using LIGPLOT [60]. The hydrogen bonds were calculated using geometrical criteria [61] of protein-ligand complex (The used criteria is: H–A distance $<2.7 \text{ \AA}$, D–A distance $<3.3 \text{ \AA}$, D–H–A angle $>90^\circ$, D–A–AA angle $>90^\circ$ and H–A–AA angle $>90^\circ$, where A is the hydrogen acceptor, D is the hydrogen donor, AA is the atom attached to the hydrogen acceptor, and H an atom of hydrogen). LIGPLOT also calculates non-covalent bond interactions (hydrophobic interactions) by applying a simple cut-off of 3.9 \AA . LIGPLOT diagrams were generated for each species. PyMOL [62] was used to generate the 3D images.

Expression analysis

B. floridae (whole organism), *C. teleta* (whole organism), *S. purpuratus* (gonad) and *N. vectensis* (whole organism) samples were obtained from Ocean Genome Legacy (OGL Accession ID numbers S13045, S13061, S13034 and S13115, respectively) [63]. RNA was extracted with the Illustra TriplePrep kit (GE Healthcare) and genomic DNA was removed from RNA preparations with an additional DNase treatment using DNase I, RNase-free (Fermentas, Thermo Fisher Scientific Inc.), according to the manufacturer's procedure. Complementary DNA (cDNA) was synthesized from $1 \mu\text{g}$ of total RNA using the RETROscrip[®] First Strand Synthesis Kit (Ambion) with oligo-dT primers according to the manufacturer's instructions. Reverse-transcription PCR reactions were prepared using Hot-StarTaq[®] Master Mix Kit (Qiagen) with $2 \mu\text{l}$ of the synthesized cDNA in a $10 \mu\text{l}$ final volume. Q solution was included in the reaction (10%) in NAMPT amplification in *B. floridae* and *S. purpuratus*. PNC and NAMPT were amplified with species-specific primers described in Table S6, with a final concentration of $0.2 \mu\text{M}$. Thermocycling conditions were as follows: initial denaturation at 95°C for 15 min, 40 cycles at 95°C for 30 sec, variable annealing temperatures ranging from 52°C to 62°C (Table S6) for 1 min 30 sec, and 72°C for 1 min, and a final extension step of 10 min at 72°C . All amplification products were visualized on 1.5% agarose gels and were confirmed by sequencing. For that, PCR products were purified with ExoSAP-IT (USB Corporation) by incubation at 37°C for 15 min, followed by enzyme inactivation for 15 min at 85°C . The resulting purified fragments were sequenced using an ABI Big Dye Terminator Cycle Sequencing Ready Reaction kit v 3.1 (Applied Biosystems) and analyzed in an ABI PRISM 3130xl (Applied Biosystems).

Expressed Sequence Tag (EST) information was retrieved from available databases for *B. floridae* [64], *S. purpuratus* [65] and *N. vectensis* [66] and is detailed in Table S2.

Supporting Information

Movie S1 Evolutionary divergence between NAMPT and PNC homologues. Protein distances were plotted for each pair of species arranged accordingly to their respective divergence time. This plot shows that NAMPT is highly conserved across large evolutionary distances, while PNC is less conserved even in closely related species. Notice that in addition to being highly conserved, protein distances and evolutionary distances are correlated in NAMPT (quantified by the Kendall coefficient of 0.413), as opposed to PNC (where the Kendall coefficient was -0.052). (MP4)

Figure S1 Distribution of NAMPT and PNC homologues across the tree of life. The presence or absence of NAMPT and PNC sequences is indicated on the columns by a plus or a minus.

Protostomes are divided in ecdysozoans and lophotrochozoans (green and blue boxes of the tree, respectively), while Deuterostomes are represented in the red box.

(TIF)

Figure S2 Alignment of the amino acid sequences from NAMPT homologues. Catalytic residues are marked with red dots and residues that bind nicotinamide, ribose, phosphate or NMN are highlighted in blue.

(TIF)

Figure S3 Alignment of the amino acid sequences from PNC homologues. Catalytic residues are marked with red dots and residues that bind zinc are highlighted in blue. Additional residues of the active site are shown in green.

(TIF)

Figure S4 Vertebrate NAMPT synteny. Conserved synteny blocks detected between the Human, Mouse and Zebrafish genomes. Input data was automatically retrieved from Ensembl release 64 using CHSminer. Corresponding chromosomes are indicated.

(TIF)

Figure S5 Invertebrate NAMPT synteny. Synteny blocks were retrieved from JGI and represent *T. adhaerens* versus *H. sapiens* (A), *M. brevicollis* (B), and *N. vectensis* (C) genomes. Corresponding scaffolds and chromosomes are indicated.

(TIF)

Figure S6 NAMPT gene structure. Exon-intron predictions were performed in Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/>).

(TIF)

Figure S7 PNC gene structure. Exon-intron predictions were performed in Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/>).

(TIF)

Figure S8 NAMPT (A) and PNC (B) amino acid alignments indicating exon size and number. *Saccoglossus kowalewskii* (Sk) protein sequences were manually predicted, based on the conserved motifs identified in this paper. Exon predictions were performed in Genescan (<http://genes.mit.edu/GENSCAN.html>).

(TIF)

Figure S9 PNC alignments of all the templates available. (A) Alignments of the amino acid sequences. (B) Structural alignments of 2H0R (blue), 1IM5 (yellow), 2WTA (pink), 3R2J (grey), 3PL1 (purple) and 3O90 (green). (C) RMSD scores of the structural alignments. All structures are superimposed on the right.

(TIF)

Table S1 Sequences of NAMPT and PNC orthologues.

The table lists the species and the associated protein identifier and source used in this work.

(XLSX)

Table S2 EST sequences of NAMPT and PNC. The table lists the species with available EST data and the associated identifier and source.

(XLSX)

Table S3 MATLAB input data. For each pair of species, mean divergence times are shown in million years as well as protein distances calculated from NAMPT and PNC amino acid alignments. -100 was the value assigned for the cases where one of the proteins is absent from the species pair.

(XLSX)

Table S4 PDBs used as reference and grid parameters for the docking calculations.

(XLSX)

Table S5 Residue positions for each species that correspond to the reference residues (NAMPT PDB id 2E5D and PNC PDB id 3R2J) of the alignment used to model the proteins.

(XLSX)

Table S6 Oligonucleotide sequences used for amplification of NAMPT and PNC from *B. floridae*, *S. purpuratus*, *C. teleta* and *N. vectensis*.

(XLSX)

References

1. Revollo JR, Grimm AA, Imai S (2007) The regulation of nicotinamide adenine dinucleotide biosynthesis by Nampt/PBEF/visfatin in mammals. *Curr Opin Gastroenterol* 23: 164–170.
2. Belenky P, Christensen KC, Gazzaniga F, Pletnev AA, Brenner C (2009) Nicotinamide riboside and nicotinic acid riboside salvage in fungi and mammals. Quantitative basis for Urh1 and purine nucleoside phosphorylase function in NAD⁺ metabolism. *J Biol Chem* 284: 158–164.
3. Bogan KL, Brenner C (2008) Nicotinic acid, nicotinamide, and nicotinamide riboside: a molecular evaluation of NAD⁺ precursor vitamins in human nutrition. *Annu Rev Nutr* 28: 115–130.
4. Belenky P, Bogan KL, Brenner C (2007) NAD⁺ metabolism in health and disease. *Trends Biochem Sci* 32: 12–19.
5. Hara N, Yamada K, Shibata T, Osago H, Hashimoto T, et al. (2007) Elevation of cellular NAD levels by nicotinic acid and involvement of nicotinic acid phosphoribosyltransferase in human cells. *J Biol Chem* 282: 24574–24582.
6. Magni G, Amici A, Emanuelli M, Orsomando G, Raffaelli N, et al. (2004) Enzymology of NAD⁺ homeostasis in man. *Cell Mol Life Sci* 61: 19–34.
7. Gazanion E, Garcia D, Silvestre R, Gerard C, Guichou JF, et al. (2011) The Leishmania nicotinamidase is essential for NAD⁺ production and parasite proliferation. *Mol Microbiol* 82: 21–38.
8. Jewett MW, Jain S, Linowski AK, Sarkar A, Rosa PA (2011) Molecular characterization of the *Borrelia burgdorferi* in vivo-essential protein PncA. *Microbiology* 157: 2831–2840.
9. Li YF, Bao WG (2007) Why do some yeast species require niacin for growth? Different modes of NAD synthesis. *FEMS Yeast Res* 7: 657–664.
10. Martin PR, Shea RJ, Mulks MH (2001) Identification of a plasmid-encoded gene from *Haemophilus ducreyi* which confers NAD independence. *J Bacteriol* 183: 1168–1174.
11. Anderson RM, Bitterman KJ, Wood JG, Medvedik O, Sinclair DA (2003) Nicotinamide and PNC1 govern lifespan extension by calorie restriction in *Saccharomyces cerevisiae*. *Nature* 423: 181–185.
12. Gallo CM, Smith DL Jr, Smith JS (2004) Nicotinamide clearance by Pnc1 directly regulates Sir2-mediated silencing and longevity. *Mol Cell Biol* 24: 1301–1312.
13. Revollo JR, Grimm AA, Imai S (2004) The NAD biosynthesis pathway mediated by nicotinamide phosphoribosyltransferase regulates Sir2 activity in mammalian cells. *J Biol Chem* 279: 50754–50763.
14. Galli M, Van Gool F, Rongvaux A, Andris F, Leo O (2010) The nicotinamide phosphoribosyltransferase: a molecular link between metabolism, inflammation, and cancer. *Cancer Res* 70: 8–11.
15. Mesko B, Poliska S, Szegedi A, Szekanecz Z, Palatka K, et al. (2010) Peripheral blood gene expression patterns discriminate among chronic inflammatory diseases and healthy controls and identify novel targets. *BMC Med Genomics* 3: 15.
16. Balan V, Miller GS, Kaplun L, Balan K, Chong ZZ, et al. (2008) Life span extension and neuronal cell protection by *Drosophila* nicotinamidase. *J Biol Chem* 283: 27810–27819.
17. Burnett C, Valentini S, Cabreiro F, Goss M, Somogyvari M, et al. (2011) Absence of effects of Sir2 overexpression on lifespan in *C. elegans* and *Drosophila*. *Nature* 477: 482–485.
18. Silva RM, Duarte IC, Paredes JA, Lima-Costa T, Perrot M, et al. (2009) The yeast PNC1 longevity gene is up-regulated by mRNA mistranslation. *PLoS One* 4: e5212.
19. van der Horst A, Schavemaker JM, Pellis-van Berkel W, Burgering BM (2007) The *Caenorhabditis elegans* nicotinamidase PNC-1 enhances survival. *Mech Ageing Dev* 128: 346–349.
20. French JB, Cen Y, Vrablik TL, Xu P, Allen E, et al. (2010) Characterization of nicotinamidases: steady state kinetic parameters, classwide inhibition by nicotinaldehydes, and catalytic mechanism. *Biochemistry* 49: 10421–10439.
21. Zhang JL, Zheng QC, Li ZQ, Zhang HX (2012) Molecular Dynamics Simulations Suggest Ligand's Binding to Nicotinamidase/Pyrazinamidase. *PLoS One* 7: e39546.
22. Domergue R, Castano I, De Las Penas A, Zupancic M, Lockett V, et al. (2005) Nicotinic acid limitation regulates silencing of *Candida* adhesins during UTLI. *Science* 308: 866–870.
23. Garten A, Petzold S, Korner A, Imai SI, Kiess W (2009) Nampt: linking NAD biology, metabolism and cancer. *Trends Endocrinol Metab* 20: 130–138.
24. Khan JA, Tao X, Tong L (2006) Molecular basis for the inhibition of human NMPRTase, a novel target for anticancer agents. *Nat Struct Mol Biol* 13: 582–588.
25. Kim MK, Lee JH, Kim H, Park SJ, Kim SH, et al. (2006) Crystal structure of visfatin/pre-B cell colony-enhancing factor 1/nicotinamide phosphoribosyltransferase, free and in complex with the anti-cancer agent FK-866. *J Mol Biol* 362: 66–77.
26. Olesen UH, Petersen JG, Garten A, Kiess W, Yoshino J, et al. (2010) Target enzyme mutations are the molecular basis for resistance towards pharmacological inhibition of nicotinamide phosphoribosyltransferase. *BMC Cancer* 10: 677.
27. Zhang LY, Liu LY, Qie LL, Ling KN, Xu LH, et al. (2012) Anti-proliferation effect of APO866 on C6 glioblastoma cells by inhibiting nicotinamide phosphoribosyltransferase. *Eur J Pharmacol* 674: 163–170.
28. Ma B, Pan SJ, Zupancic ML, Cormack BP (2007) Assimilation of NAD(+) precursors in *Candida glabrata*. *Mol Microbiol* 66: 14–25.
29. Seiner DR, Hegde SS, Blanchard JS (2010) Kinetics and inhibition of nicotinamidase from *Mycobacterium tuberculosis*. *Biochemistry* 49: 9613–9619.
30. Sorci L, Blaby I, De Ingeniis J, Gerdes S, Raffaelli N, et al. (2010) Genomics-driven reconstruction of acinetobacter NAD metabolism: insights for antibacterial target selection. *J Biol Chem* 285: 39490–39499.
31. Lin H, Kwan AL, Dutcher SK (2010) Synthesizing and salvaging NAD: lessons learned from *Chlamydomonas reinhardtii*. *PLoS Genet* 6(9): e1001105.
32. Gazzaniga F, Stebbins R, Chang SZ, McPeck MA, Brenner C (2009) Microbial NAD metabolism: lessons from comparative genomics. *Microbiol Mol Biol Rev* 73: 529–541.
33. Gossman TI, Ziegler M, Puntervoll P, de Figueiredo LF, Schuster S, et al. (2012) NAD(+) biosynthesis and salvage - a phylogenetic perspective. *FEBS J* 279: 3355–3363.
34. Miller ES, Heidelberg JF, Eisen JA, Nelson WC, Durkin AS, et al. (2003) Complete genome sequence of the broad-host-range vibriophage KVP40: comparative genomics of a T4-related bacteriophage. *J Bacteriol* 185: 5220–5233.
35. Burgos ES, Ho MC, Almo SC, Schramm VL (2009) A phosphoenzyme mimic, overlapping catalytic sites and reaction coordinate motion for human NAMPT. *Proc Natl Acad Sci U S A* 106: 13748–13753.
36. Takahashi R, Nakamura S, Nakazawa T, Minoura K, Yoshida T, et al. (2010) Structure and reaction mechanism of human nicotinamide phosphoribosyltransferase. *J Biochem* 147: 95–107.
37. Wang T, Zhang X, Bheda P, Revollo JR, Imai S, et al. (2006) Structure of Nampt/PBEF/visfatin, a mammalian NAD⁺ biosynthetic enzyme. *Nat Struct Mol Biol* 13: 661–662.
38. Du X, Wang W, Kim R, Yakota H, Nguyen H, et al. (2001) Crystal structure and mechanism of catalysis of a pyrazinamidase from *Pyrococcus horikoshii*. *Biochemistry* 40: 14166–14172.
39. Hu G, Taylor AB, McAlister-Henn L, Hart PJ (2007) Crystal structure of the yeast nicotinamidase Pnc1p. *Arch Biochem Biophys* 461: 66–75.
40. Fyfe PK, Rao VA, Zemla A, Cameron S, Hunter WN (2009) Specificity and mechanism of *Acinetobacter baumannii* nicotinamidase: implications for activation of the front-line tuberculosis drug pyrazinamide. *Angew Chem Int Ed Engl* 48: 9176–9179.
41. Petrella S, Gelus-Ziental N, Maudry A, Laurans C, Boudjelloul R, et al. (2011) Crystal structure of the pyrazinamidase of *Mycobacterium tuberculosis*: insights into natural and acquired resistance to pyrazinamide. *PLoS One* 6: e15785.
42. French JB, Cen Y, Sauve AA, Ealick SE (2010) High-resolution crystal structures of *Streptococcus pneumoniae* nicotinamidase with trapped intermediates provide insights into the catalytic mechanism and inhibition by aldehydes. *Biochemistry* 49: 8803–8812.

Acknowledgments

We are thankful to Fátima Gil from Aquário Vasco da Gama (Lisboa, Portugal), Ana Lopes Ribeiro from INEB (Porto, Portugal) and Timery S. DeBoer from Ocean Genome Legacy (Ipswich MA, USA) for providing samples.

Author Contributions

Conceived and designed the experiments: RMS. Performed the experiments: JC SDP IM RMS. Analyzed the data: LA LFCC RMS. Contributed reagents/materials/analysis tools: PA AA. Wrote the paper: JC SDP LA LFCC PA IM AA RMS.

43. Irimia M, Tena JJ, Alexis MS, Fernandez-Minan A, Maeso I, et al. (2012) Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. *Genome Research* 22:2356–2367.
44. Louis A, Roest Crollius H, Robinson-Rechavi M (2012) How much does the amphioxus genome represent the ancestor of chordates? *Brief Funct Genomics* 11: 89–95.
45. Azevedo L, Carneiro J, van Asch B, Moleirinho A, Pereira F, et al. (2009) Epistatic interactions modulate the evolution of mammalian mitochondrial respiratory complex components. *BMC Genomics* 10: 266.
46. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.
47. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
48. Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. Edited in *Evolving Genes and Proteins* by V Bryson and HJ Vogel Academic Press, New York: 97–166.
49. Drummond A, Ashton B, Buxton S, Cheung M, Cooper A, et al. (2011) Geneious. v5.5 ed: <http://www.geneious.com>.
50. Hedges SB, Dudley J, Kumar S (2006) TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22: 2971–2972.
51. Wang Z, Ding GH, Yu ZH, Liu L, Li YX (2009) CHSMiner: a GUI tool to identify chromosomal homologous segments. *Algorithms for Molecular Biology* 4: 2.
52. Schrödinger LLC (2009) Prime, version 2.1, New York, NY.
53. Baldi P, Brunak S, Frasconi P, Soda G, Pollastri G (1999) Exploiting the past and the future in protein secondary structure prediction. *Bioinformatics* 15: 937–946.
54. Pollastri G, Przybylski D, Rost B, Baldi P (2002) Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins* 47: 228–235.
55. Schrödinger LLC (2011) LigPrep, version 2.5, New York, NY.
56. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 30: 2785–2791.
57. Sanner MF (1999) Python: a programming language for software integration and development. *J Mol Graph Model* 17: 57–61.
58. Pohorille A, Jarzynski C, Chipot C (2010) Good practices in free-energy calculations. *J Phys Chem B* 114: 10235–10253.
59. Shirts MR, Pitera JW, Swope WC, Pande VS (2003) Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *Journal of Chemical Physics* 119: 5740–5761.
60. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8: 127–134.
61. McDonald IK, Thornton JM (1994) Satisfying hydrogen bonding potential in proteins. *J Mol Biol* 238: 777–793.
62. Schrödinger LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1 (<http://pymol.org/>).
63. Ocean Genome Legacy, Ocean Genome Resource database. Published on the Web at: www.oglf.org/Catalog.htm; accessed September 2011.
64. Yu JK, Wang MC, Shin IT, Kohara Y, Holland LZ, et al. (2008) A cDNA resource for the cephalochordate amphioxus *Branchiostoma floridae*. *Dev Genes Evol* 218: 723–727.
65. Tu Q, Cameron RA, Worley KC, Gibbs RA, Davidson EH (2012) Gene structure in the sea urchin *Strongylocentrotus purpuratus* based on transcriptome analysis. *Genome Res* 22: 2079–2087.
66. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, et al. (2007) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 35: D5–12.