Tracking evolution of myoglobin stability in cetaceans using experimentally calibrated computational methods that account for generic protein relaxation

Holm, Jeppe; Dasmeh, Pouria; Kepp, Kasper Planeta

Published in:
Biochimica et Biophysica Acta - Proteins and Proteomics

Link to article, DOI:
[10.1016/j.bbapap.2016.04.004](https://doi.org/10.1016/j.bbapap.2016.04.004)

Publication date:
2016

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Holm, J., Dasmeh, P., & Kepp, K. P. (2016). Tracking evolution of myoglobin stability in cetaceans using experimentally calibrated computational methods that account for generic protein relaxation. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1864(7), 825-834. <https://doi.org/10.1016/j.bbapap.2016.04.004>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Tracking evolution of myoglobin stability in cetaceans using experimentally calibrated computational methods that account for generic protein relaxation

Jeppe Holm¹, Pouria Dasmeh², and Kasper P. Kepp^{*1}

¹ *Technical University of Denmark, DTU Chemistry, DK-2800 Kongens Lyngby, Denmark.*

² *Harvard University, Department of Chemistry and Chemical Biology, Cambridge, Massachusetts, USA.*

* Corresponding Author E-mail: kpj@kemi.dtu.dk Phone: +045 45 25 24 09

Abstract

The evolution of cetaceans (whales, dolphins, and porpoises) from land to water is one of the most spectacular events in mammal evolution. It has been suggested that selection for higher myoglobin stability (ΔG of folding) allowed whales to conquer the deep-diving niche. The stability of multi-site protein variants, including ancient proteins, is however hard to describe theoretically. From a compilation of experimental $\Delta\Delta G$ vs. ΔG we first find that protein substitutions are subject to large generic protein relaxation effects. Using this discovery, we develop a simple two-parameter model that predicts multi-site $\Delta\Delta G$ as accurately as standard methods do for single-site mutations and reproduces trends in contemporary myoglobin stabilities. We then apply this new method to the study of the evolution of Mb stability in cetaceans: With both methods the main change in stability (about 1 kcal/mol) occurred very early, and stability was later relaxed in dolphins and porpoises, but was further increased in the sperm whales. This suggests that single proteins can affect whole organism evolution and indicates a role of Mb stability in the evolution of cetaceans. Transition to the deep-diving niche probably occurred already in the ancestor of contemporary baleen and toothed whales. In summary, we have discovered generic stability relaxation effects in proteins that, when incorporated into a simple model, improves the description of multi-site protein variants.

Key words: myoglobin; whales; cetaceans; evolution; protein stability; protein misfolding

1. Introduction

A major challenge in biology is to understand, quantify, and predict the molecular forces of evolution producing the diverse biological structure that we observe around us. Solving this challenge would enable us to rationalize natural evolution phenomena at the single-protein level. In this quest, proteins that infer direct advantages to the organism phenotype are particularly relevant, as the mapping from structure and function to organism and macroevolution then becomes feasible.

One example of single protein evolution is myoglobin (Mb) of deep-diving cetaceans (whales, dolphins, and porpoises)[1][2][3]. Most marine mammals originate from ancestors re-entering the oceans [4]. The ancestor of contemporary cetaceans probably re-entered the ocean approximately 50 million years ago (Mya) in a transition to aquatic life style [5][6]. The associated adaptations and specializations are among the most remarkable in mammalian biology: Various limbs were discarded to increase aquatic streamlining, by inactivation of genes coding for redundant body parts, and tails were evolved for propulsion [5], followed by one of the most amazing divergences of general feeding strategy occurring around 35 Mya leading to the contemporary suborders of toothed and baleen whales (*Odontoceti* and *Mysticeti*) [6][7], one using massive filter-feeding of plankton, the other using advanced echo-location techniques to hunt prey at substantial depths. Dive durations and –depths vary greatly within *Cetacea* [8]: Humpback whales and harbor porpoises dive approximately 3–7 minutes to ~60 meters, whereas specialization towards deep dives has produced the largest mammals known, with e.g. large sperm whales routinely diving for more than an hour at a kilometer’s depth [9][10].

Selection pressure to increase dive lengths has thus partly dictated the evolution of whales: They acquired increased blood volume, efficient modes of locomotion such as stroke-and-glide swimming, and ~10–20 times higher Mb concentrations (C_{Mb}) [9][11]. Many cetaceans depend heavily

on Mb in skeletal muscles with up to half of all stored O₂ located there, vs. only 15% in humans [9]. During routine dives where oxygen consumption is fairly constant, dive lengths consequently scale nearly linearly with C_{Mb} [9][12]. From this, it has been argued that the fitness of whales who evolved diving foraging strategies can be considered proportional to dive duration and thus, C_{Mb}, providing a direct coupling from single-protein properties to organism fitness [12].

Figure 1. Generic protein relaxation affects stability changes. A) Conserved sites around heme group of myoglobin, contributing to O₂ affinity (made with Pymol based on structure 1BZ6); B) Average experimental $\Delta\Delta G$ upon mutation as a function of the total ΔG of the wild type for different proteins (Supporting Information, Table S1).

Sites near the O₂-binding heme group of Mb are evolutionarily conserved across mammals (Figure 1A), and their mutation substantially impairs O₂-affinity [13][14]. The O₂ affinity is under strong purifying selection due to the delicate need for fast, reversible O₂ binding to heme [15] and the

co-evolved O₂-saturation curves of the ingenious hemoglobin-myoglobin transport/storage partnership [16]. If O₂ affinity is changed even marginally, animal dive lengths are drastically reduced, explaining why such mutants are not found in nature, providing a tangible example of single-protein evolution directly affecting organism fitness [12]. Thus, as has been argued, the only feasible way to increase O₂-storage was to increase C_{Mb} [1].

Such increased cellular concentration of a protein, however, does not come without a cost [17]. Selection pressures act on abundant proteins to reduce their rate of evolution [18]. Three reasons have been suggested for this universal tendency, one being constraints on the cellular translation efficiency to minimize errors [17], another being the toxic burden of misfolded protein copies [19][20][21], and a third uniting these, the increased maintenance energy of proteostasis at both RNA and protein levels, which reduces available reproductive energy, and by inference, fitness [22].

The stability of apo-Mb in cetaceans is 1–3 kcal/mol higher than in related terrestrial mammals [23], suggesting that Mb stability could be positively selected. By directly mapping chemical properties of protein orthologs onto the species phylogeny (we refer to this as “protein-explicit phylogeny”), it was found that substitutions that increase Mb stability tend to be positively selected in cetaceans [1], and not due to coincidental or nearly-neutral evolution. The increased stability protects against unwanted side-effects of the increased C_{Mb}: A 10-fold higher total C_{Mb} would lead to a 10-fold increase in unfolded Mb ([U]) at steady-state if no change in stability occurred. However, if the protein had evolved to be 1.5 kcal/mol more stable, the increase in [U] would be effectively nullified. Thus, a conspicuous relationship between experimentally observed elevations in Mb stability and concentration can be rationalized and argues for selection against misfolded protein copies *via* protein stability [1].

Heterologous expression levels of Mb orthologs are also proportional to their respective ApoMb stabilities but depend little or even inversely on heme affinity, because high ApoMb stability increases the life time and thus total copy number of holoMb formed upon heme incorporation[24], which is a specific event on the folding pathway of almost completely folded ApoMb[25]. This leads to an apparent correlation between heterologous expression levels and Mb concentrations measured in the real host mammals; in both cases, this arises from the stability of the Mb ortholog[1][24]. However, the dependency on stability arises for different reasons: In one case due to protein synthesis rates, in the other, as argued previously[1], due to selection pressure on steady state concentrations of whales.

Later, a similar study found the net charge to be significantly higher in Mbs of diving mammals [26]. Although the sites that increased charge were not shown to be selected rather than occurring coincidentally, increased net charge occurred in multiple aquatic clades. Net charge is one of the factors that reduce protein aggregation propensity [21] and thus supports a mechanism of selection against misfolded protein copies in diving mammals as first suggested by Dasmeh et al.[1]

A main challenge of protein-explicit phylogeny is the accuracy of the computed properties, which is challenged by the many protein variants in the phylogeny that differ in multiple sites, leading to potential error propagation. Computational methods have been designed to quantify the change in folding free energy ($\Delta\Delta G$) upon single-site mutation [27][28], yet their advantage mainly manifests when many variants with multiple differing sites are compared, because experimental screening becomes impossible [29]. There is thus a major need to extend computational protein stability estimation to multi-site variations, as attempted only in a few cases [29][30].

In this paper, it is shown that multi-site variations can be studied with an accuracy that approaches that of single-site mutations, if the methods are calibrated to account for protein relaxation

effects. Interestingly, these relaxation effects are generic, which is of fundamental interest and relevant to the design of improved stability predictors. The two calibrated methods (Popmusic 2.1 [31] and I-Mutant 2.0 [32]) were then used to trace the stability of Mb through the evolutionary history of cetaceans. Both the applied methods provide results in good agreement with the experimentally known stability trends of cetacean vs. terrestrial Mbs and recapitulate the higher stability of sperm whales vs. other cetaceans. Both methods find that most stability increase occurred in the early evolution of *Cetacea* but that stability was later relaxed in dolphins and porpoises.

2. Methods

2.1. Experimental data sets

To benchmark computational estimates of the stability changes ($\Delta\Delta G$) of multi-site variants in the Mb phylogeny, we used three experimental data sets.

The first data set is a compilation from the ProTherm database [33] of $\Delta\Delta G$ of single-site mutations where also the total stability (ΔG) of the wild-type (WT) protein is reported. Departure from optimal pH (pH_{opt}) strongly scales down both ΔG , as is well-known, but also $\Delta\Delta G$, which is usually used for benchmarking and design of computational $\Delta\Delta G$ estimators [27][28]. This is a major problem and shows that only nearly-neutral pH data should be applied in the design and benchmarking of methods. This data set thus only included those data that were reported at nearly neutral pH, as collected in Supporting Information, Table S1. This data set is used in this work to identify a generic relationship between the starting stability of a protein and the associated impact of typical mutations.

For the main subject of this paper, the study of Mb stability during evolution, we used two data sets specifically for Mb: one data set for multi-site alanine mutations by Lin et al.[34], and one for the

stabilities of several cetaceans and terrestrial mammals by Scott et al.[23] Unfolding equilibrium constants were recalculated into $\Delta\Delta G$ using $-RT\ln K$ and all stabilities were reported as relative to that of the wild type sperm whale Mb ΔG value.

2.2. Applied crystal structures and structure insensitivity of computed results

All computations were carried out on PDB files as specified by the ProTherm entries for the non-Mb data sets. For the Mb data sets, we tested several PDB structures since it has previously been reported that some stability estimators are sensitive to structure input [35]. Previous work on single-site mutations in Mb found that I-Mutant 2.0 [32] and Popmusic 2.1 [31] are accurate for Mb and at the same time display low structure sensitivity [36]. Yet, to confirm low structure sensitivity also for the present work's mutations, we studied the performance of the methods for four near-atomic resolution crystal structures with all 153 residues included: 1BZ6 and 1BZP [37], 2EKT [38], and 2ZSN [39]: 1BZ6 represents Mb with H₂O bound as distal ligand (aqua-Mb) at a resolution of 1.20 Å in a monoclinic crystal; 1BZP is deoxy-Mb at 1.15 Å in a monoclinic crystal; 2EKT is aqua-Mb at 1.10 Å in a hexagonal crystal; and 2ZSN is CO-Mb at 1.21 Å in a monoclinic crystal. For comparison, we also included the method CUPSAT [40] in the account of structure sensitivity.

2.3. Computational tools

The two methods I-Mutant 2.0 [32] and Popmusic 2.1 [31] were used for specific stability calculations, as they have been shown in previous benchmarks to perform most accurate for Mb mutations [36]. I-Mutant is a support vector machine method using an input vector of 43 parameters that account for T,

pH, each amino acid in a specific local environment, and solvent accessibility; the method was calibrated using data for 2087 single-point mutations [32][41]. Popmusic 2.1 uses parameterized energy functions depending on amino acid volume and solvent accessibility, and the method was parameterized to reproduce $\Delta\Delta G$ for 2648 single-point mutations [31].

All calculations reflect a fundamental change in free energy of folding caused by mutation i , i.e.

$$\Delta\Delta G_i = \Delta G_i(\text{mutant}) - \Delta G(WT) \quad (1)$$

Here, a negative number implies a stabilization effect of the amino acid change (mutation or substitution). Popmusic 2.1 uses this sign convention, whereas I-Mutant 2.0 uses the reverse sign convention; hence numbers computed with the latter method were multiplied by -1 .

The ancestral states were inferred using the maximum likelihood method with a Dayhoff matrix-based model [42] as implemented in the MEGA5 software [43].

2.4. Calibrated multi-site protein stability model

To compute the $\Delta\Delta G_N$ of an N -site mutant and of protein orthologs and ancestors, where multiple sites differ from the structural template, we first considered the simplest model where the relative free energy equals the summed contribution of all individual amino acids changes:

$$\Delta\Delta G_N = \sum_i^N \Delta\Delta G_i \quad (2)$$

Here, all changes are calculated relative to a template wild-type structure.

High correlation coefficients were observed between summed single-mutation $\Delta\Delta G$ values and corresponding full triple- and quadruple-mutation $\Delta\Delta G$ values ($R = 0.72$ and 0.66 , respectively) [44].

These correlation coefficients are comparable to those of computational methods applied only to the prediction of experimental single-site mutations [27][28]. Thus, the additive approximation is a valid approach since full, specific protein relaxation effects are beyond current computational capacities.

However, numerical values from Equation (2) are not accurate because they overestimate the impact of each change in a multi-site variant (see below). Therefore, we investigated simple approaches to solve this issue and found that the multi-site method needs to correct for the number of changed sites N , since the error grows with this number. The simplest such approach is to compute the mean of the single-site mutation $\Delta\Delta G$ values constituting the multi-site variant:

$$\Delta\Delta G_N = \frac{\sum_i^N \Delta\Delta G_i}{N} \quad (3)$$

From a quick estimate of the $\Delta\Delta G$ values of both multi-site and single-site mutants reported in the ProTherm database, this seems much more reasonable than Equation (2), since unsigned multi-site $\Delta\Delta G$ values are generally not much larger than unsigned single-site $\Delta\Delta G$ values. This is apparent from the pioneering work on exhaustive mutations in staphylococcal nuclease [44], where the $\Delta\Delta G$ values of quadruple-site mutations were much smaller than the sums of the single-site $\Delta\Delta G$ values, despite the high trend correlation discussed above.

As a compromise between these two extremes, we thus devised the simplest possible model that maintains the physical trend accuracy already provided by the combined single-site mutations but scales the total numerical value of $\Delta\Delta G$ according to a compromise between the sum and average:

$$\Delta\Delta G_N = \frac{\sum_i^N \Delta\Delta G_i + 1/N \sum_i^N \Delta\Delta G_i}{2z} = \frac{\sum_i^N \Delta\Delta G_i (1 + \frac{1}{N})}{2z} \quad (4)$$

In this work, we assumed that z is a constant; however, z may also be optimized as dependent on N , and equation (4) assumes equal weight of the two terms, which could easily be changed in future work. As

shown below, this simple expression works remarkably well once calibrated to minimize mean absolute errors between experimental and computed data.

For the final computed estimates of the Mb phylogenies, an additional calibration was done to make sure that the values were as accurate as possible for the specific protein of interest, Mb. These method-specific calibrated values were computed as:

$$\Delta\Delta G_{N,Calibrated} = a \Delta\Delta G_N + b \quad (5)$$

As seen from Equation (4), the a constant includes z , and thus only two free parameters were optimized. We then computed the estimated total stability of any Mb of any sequence (ΔG_{Mammal}) using the atomic-resolution sperm whale structure 1BZ6 as template with the equation:

$$\Delta G_{Mammal} = \Delta\Delta G_{N,Calibrated} + \Delta G_{Sperm\ whale} \quad (6)$$

The experimental value from Scott et al. of $\Delta G_{Sperm\ whale} = -7.17$ kcal/mol was used [23].

3. Results and discussion

3.1. Generic protein relaxation effects upon mutation

The typical $\Delta\Delta G$ of a random single-site mutation is 1 kcal/mol towards destabilization, based on the data available in ProTherm [45]. Yet, by inspection of ProTherm, the corresponding values for multi-site mutants are not that much larger than for typical single-site mutations, and certainly not of a magnitude that resembles the sum of single-site $\Delta\Delta G$ values. From this observation, it is apparent that energy relaxation compensates the first-order destabilizing effect of a mutation once several mutations are introduced.

Protein homologs and orthologs usually exhibit relatively similar stabilities, as also seen for Mb variants in this work, despite differing in many sites. However, since these real proteins differ from the typical mutants in ProTherm by being the result of natural evolution, selection pressure on stability could contribute to this “compression” of protein variants stabilities. The intrinsic relaxation effect originates in the fundamental electrostatic interactions between introduced amino acids and remaining protein and is thus qualitatively different from the probabilistic selection dynamics that may produce marginal protein stability by mutation-selection balance because most mutations are destabilizing [46]; the latter selection effect would occur even if $\Delta\Delta G$ did not depend on ΔG . However, the relative magnitudes of these two effects are not well-known.

A fundamental anti-correlation between the introduced magnitude of $\Delta\Delta G$ and the stability of the wild type protein ΔG was noted and discussed by Serohijos et al., who found that this anti-correlation leads to more abundant proteins evolving slowly, if misfolded protein copies are selected against [18][47]: the copy number of misfolded proteins grows with loss of stability according to two-

state unfolding thermodynamics [20] and the fitness is proportional to this copy number times a cost constant c per misfolded copy defined heuristically by Drummond et al.[17] and derived from protein turnover kinetics in an energy framework by Kepp and Dasmeh [22].

To investigate this relationship in more detail, we collected data sets from ProTherm that included both the $\Delta\Delta G$ values of single-site mutations and the specifically determined total stability ΔG of the WT involved in the experiment. Some data were for the same protein at different pH with very different ΔG values, since most proteins have a pseudo-Gaussian stability surface centered at an optimal pH value, pH_{opt} . However, we saw that pH also has a large effect on $\Delta\Delta G$. This is important because in current computational stability estimators, $\Delta\Delta G$ values of mutations were parameterized against experimental $\Delta\Delta G$ values at variable pH values. Some of the anti-correlation of $\Delta\Delta G$ and ΔG noted previously [20] could thus be due to pH effects and not due to intrinsic protein relaxation.

Therefore, we sorted the data for single-site mutations to only include the pH range from 5.5 to 8.5, with most data at pH 7 (Supporting Information, Table S1). The average $\Delta\Delta G$ of single-site mutations for each protein was plotted against the total ΔG of the protein (Figure 1B) giving a correlation coefficient of linear regression $R = 0.85$ ($R^2 \sim 0.73$) and $p = 4.6 \times 10^{-5}$. This correlation is much more significant than previously, because it averages out individual protein effects and reduces noise from experimental protocols. The anti-correlation shows that more stable proteins tend to have more destabilizing mutations already *in vitro*. The reason is that physical interactions tend to scale in the same way with external variables, and ΔG and $\Delta\Delta G$ are made up by such similar interactions. From our analysis, the slope was -0.3 (in the work by Serohijos et al., it was -0.13 [20]). Thus Figure 1B quantifies, in our opinion, a real generic, physical relaxation effect of proteins distinct from and present without any evolutionary dynamics. A correlation does not prove cause, and there could be other reasons for such a correlation. Yet, the most likely alternative, experimental bias in choices of mutants,

is not supported by the fact that the wild type proteins and well as the mutants differ substantially in stability, the general difficulty of *a priori* predicting the properties of protein mutants before choosing them, and the averaging over many mutations in each site in these instances. Furthermore and most importantly, multi-site mutations have stability changes that numerically resemble those of single-site mutated sites, speaking strongly for a relaxation effect that would not be explained by experimental bias.

These *in vitro* mutations are introduced, not fixated by evolution, and thus, they say nothing about a tendency of a stable protein to better accept destabilizing mutations during natural evolution. However, the relaxation effects contribute to $\Delta\Delta G$ and thus indirectly, to the fixation of mutations under selection pressure for maintaining stability [20].

Chen and Stites [44] showed that most of the multi-site $\Delta\Delta G$ is predictable from the local physical effects of individual sites, whereas correlation terms are dominated by pair-wise interactions and are smaller than the direct physical effect of $\Delta\Delta G$ for the involved single mutations ($R = 0.72$ for linear regression of summed single-site $\Delta\Delta G$ according to Equation 2, vs. true triple-site mutant $\Delta\Delta G$). Thus, generic scaling of interactions make $\Delta\Delta G$ proportional to ΔG , and relaxation within proteins reduces the impact of single-site mutations but keep the most important qualitative feature of the mutation in terms of destabilization or stabilization. Together, this shows that interactions within proteins are subject to uniform scaling and protein relaxation.

Computational predictors towards unspecified $\Delta\Delta G$ values will make an error when neglecting this important relaxation effect, and some of the lack of universality of current methods may be due to the neglect of ΔG in $\Delta\Delta G$ estimates. The inclusion of this relaxation effect will thus also help us

provide more numerically accurate estimates of protein stabilities where many sites differ, relevant to the present work.

3.2. Multi-site $\Delta\Delta G$ are well predicted by the combined single-site $\Delta\Delta G$ values

After identifying the generic scaling of interactions causing correlation between $\Delta\Delta G$ and ΔG and protein relaxation effects on introduced mutations, we investigated whether multi-site mutations can be accurately described by single-site mutations if one includes such scaling. Specific, local correlations between amino acids are responsible for the missing part of the correlation between $\Delta\Delta G$ and ΔG . These correlations, which may e.g. lead to same-gene epistasis during evolution [48], may be modeled separately by other tools and are beyond the scope of the present work, which concerns the generic scaling effect evident from Figure 1.

First, we wanted to ensure that the choice of crystal structure does not affect significantly the results. Recent work has shown that stabilities, and indeed other structure-derived protein properties, can be very sensitive to the structure input, not only from MD simulation where this relates to the problem of sampling [35], but also in the experimental crystal structures, which differ in terms of physical conditions, resolution, and protein state [49]. We therefore tested the sensitivity of $\Delta\Delta G$ values to the use of four different near-atomic-resolution Mb crystal structures, 1BZ6, 1BZP, 2EKT, and 2ZSN. 1BZ6 and 2EKT both have water as distal ligand but differ in crystal symmetry type (1BZ6 is monoclinic whereas 2EKT is hexagonal). 1BZ6 is the deoxy version of monoclinic 1BZP, whereas 2ZSN is a carbon monoxide-bound Mb also in monoclinic symmetry.

The results of this analysis are shown in Supporting Information, Table S2, showing the computed $\Delta\Delta G$ values for single-site mutations involved in the transition from sperm whale to pygmy

sperm whale Mb, and from sperm whale to pig Mb. The mean $\Delta\Delta G$ values and standard deviations for the involved mutations are shown for each method in the bottom of Table S2. It can be seen that I-Mutant 2.0 and Popmusic 2.1 are very insensitive to crystal structure and distal ligand change, with maximum changes in mean and standard deviations of 0.05 kcal/mol, much less than the error of the methods. In contrast, CUPSAT displays large structure sensitivity, with mean values varying by up to 1.11 kcal/mol (from -0.50 kcal/mol with 2EKT to 0.61 kcal/mol for 1BZ6) and standard deviations differing by up to 0.27 kcal/mol. Thus, the results described in the following work using calibrated I-Mutant 2.0 and Popmusic 2.1 methods are not dependent on the choice of Mb structure in the near-atomic resolution limit.

Figure 2 shows the predicted $\Delta\Delta G$ of multi-site mutants and mammalian Mb orthologs relative to sperm whale Mb, computed either by Popmusic (Figure 2A) or I-Mutant (Figure 2B) using the simple additive approximation of Equation 2. The significant correlations demonstrate that a major part of the physics of multi-site mutations is described by the properties of the individual sites, consistent with the finding from experimental mutagenesis studies by Chen and Stites [44]. However, because of the generic relaxation effects discussed above, the numerical accuracy is limited. We therefore used these data to optimize the parameters a and b of Equation (5) by linear regression. For Popmusic, this gave the parameters $a = 0.77$ and $b = -0.43$, whereas for I-Mutant, $a = 0.54$ and $b = -0.54$. Because of the scale factor $2z = 4$ that may be modeled as N -dependent but is a constant in this work, the effective calculation expression for one multi-site mutation is:

$$\Delta\Delta G_{Calibrated} = a \frac{\sum_i^N \Delta\Delta G_i (1 + \frac{1}{N})}{4} + b \quad (7)$$

The scatter plots of $\Delta\Delta G_N$ giving rise to these parameters are shown in Figure 2C and 2D. From comparison of Figures 2A and 2B vs. 2C and 2D, the effect of the N -quenching on bringing the

numerical stability effects into a realistic range is evident. Numerical values for the computed numbers can be found in Supporting Information, Table S3 and Table S4.

It is notable that our Mb calculations include the hemes of the functional holoproteins present in the real animals, whereas experimental stabilities were obtained for the apoMbs. This could affect the agreement between experimental and calculated data as well. Yet, it is still meaningful to compare our methods with available apoMb data rather than not benchmarking it at all: Although the heme affinity affects stability[50], the differential effect of heme among orthologs is small compared to the direct effect of amino acid substitutions, i.e. the heme is identical in all orthologs and is more constant in its effect except for correlations with substituted sites close to the heme pocket. This was further confirmed by Kepp[36] who showed that I-Mutant and Popmusic (but not other, more structure-sensitive methods) produce most of the correlation with experimental apoMb data with heme present in the crystal structures and better than expected for typical proteins without cofactors.

Figure 2. Experimental vs. computed $\Delta\Delta G$ relative to WT myoglobin for the full myoglobin data set (orthologs and multi-site mutants). A) Computed with Popmusic 2.1. B) Computed with I-Mutant 2.0. C) Calibrated Popmusic 2.1. D) Calibrated I-Mutant 2.0.

3.3. The evolution of Mb stability in cetaceans using protein-explicit phylogeny

We now apply our new method that incorporates generic protein relaxation in a simple way to orthologs and ancestors of cetacean Mbs. Using the species tree of the Cetacean clade as described in recent work [7], the sequences of ancestors were previously reconstructed [1]. From these sequences, we computed the estimated stabilities of the ancestor proteins that resided within extinct cetaceans

using the calibrated methods (Figure 2C and D). The numerical data for these computations can be found in Supporting Information, Table S5. Figure 3 shows these stabilities estimated using the calibrated Popmusic method, whereas Figure 4 shows the same tree with numbers from the calibrated I-Mutant method. The use of two different methods both calibrated against experimental data serves to increase the confidence in the stability estimates.

For a single mutation, computed $\Delta\Delta G$ values are in general not very significant (this is generally true for theoretical estimates of single mutants with any known method). Our method has the same accuracy for multi-site substitutions as standard methods have for single-site substitutions, which is a step forward, but this error is still ~ 1 kcal/mol as estimated by Potapov et al. in their independent benchmark[28], and as seen in our data. Therefore, stability estimation tools should only be applied strictly to *batches* of mutants, both in protein engineering and as done in this study. Then the general drifts in stability among many substitutions become statistically significant. For a group of mutations, the group mean becomes increasingly accurate with N , and so does the confidence in the direction, if any, of these. This effect is seen in the correlation coefficient of our calibrated Popmusic method, which predicts experimental ortholog stabilities with a p -value of 0.033 (Supporting Information, Table S6, and Figure S1). Thus, while some changes are insignificant, a few changes are significant, to be discussed below.

First, we find that the major transition from Ancestor A to B (resembling the transition to Cetacea from the common ancestor of hippos and cetaceans) was accompanied by an estimated 1 kcal/mol increase in Mb stability (-1.34 kcal/mol for Popmusic, -0.74 kcal/mol for I-Mutant, average -1.04 kcal/mol). Both methods find this to a high extent and suggest that Ancestor B had a stability of approximately -7 kcal/mol, approximately as contemporary sperm whale (-7.2 kcal/mol (notice that this number is the experimental number used as reference to compute all other ΔG values)). As seen in

Figure 5, Z-score analysis finds this change in stability from ancestor A to B to be statistically significant compared to other branches of the phylogenetic tree with p -values of 4.4×10^{-70} and 3.1×10^{-25} for the changes computed with Popmusic (Figure 5A) and I-Mutant (Figure 5B), respectively. The overall stability drift that we obtain resembles in magnitude the experimental data by Scott et al. and much larger than the structure-sensitivity discussed above, i.e. the result will not change if another structure input was used.

Both methods also found that sperm whales experienced additional stabilization of Mb by 0.14–0.17 kcal/mol during subsequent evolution, whereas other branches experienced reduced stability by up to 1 kcal/mol. Except for the two transitions leading to Minke and humpback whales, the two methods agree on the sign of all transitions within the clade. The loss of stability in dolphins vs. pilot and melon headed whales is also a consistent feature of both methods. All these tendencies furthermore agree qualitatively with the calculations by Dasmeh et al. using a third method, FoldX [51], making the sign changes and approximate magnitudes more significant. Third, the particular stability of sperm whale Mb vs. other cetacean Mbs is consistent with experimental acid-denaturation estimates of contemporary cetacean Mb stabilities[52]. Thus, why most minor changes in the trees are insignificant, the two major drift events discussed above are significant from p -values vs. experimental data for orthologs, Z-score analysis, and the fact that both methods agree on magnitude and sign.

Figure 3. Evolution of myoglobin stability in cetaceans according to calibrated Popmusic method.

The tree is a simplified version of the full tree that includes 130 mammalian myoglobin sequences.

Figure 4. Evolution of myoglobin stability in cetaceans according to calibrated I-Mutant method.

The tree is a simplified version of the full tree that includes 130 mammalian myoglobin sequences.

Figure 5. Z-scores for the change in myoglobin stability model for all branches in the cetacean phylogenetic tree. A) Calculated using the data from the calibrated Popmusic model; B) from the calibrated I-Mutant 2.0. The first bar corresponds to the ancestral transition A to B.

Figure 6. Key substitutions of ancestor myoglobins. A) Myoglobin ancestor A; B) Myoglobin ancestor B. Residues are colored according to positive charge (blue), negative charge (red), polar (purple), hydrophobic (green), or other/special (orange). The conserved proximal and distal histidines are shown in yellow.

3.4. Individual amino acid changes important to the evolution of cetacean Mb

The substitutions defining the transition to the common ancestor of *Cetacea*, relative to its terrestrial counterparts (from ancestor A to B), are the charged residue changes E27D, K118R, and N140K, and the neutral substitutions G1V, G15A, V28I, V101I, G129A, S132, and M142I. Together, these substitutions increased protein stability by approximately 1 kcal/mol, as estimated by both methods. The homology models of ancestors A and B are shown in Figure 6A and 6B, respectively. It is highly significant that all these neutral substitutions are associated with increased hydrophobicity and the likely associated destabilization of the unfolded protein chain. These residues form the core of the chemical changes that accompanied the transition from land to water between Ancestor A and B. A few

of these sites later relaxed back: V1 returned to G in the transition to dolphins and porpoises, a transition that was accompanied by the distinct Q152H substitution found in dolphins and porpoises but not in their closest-relative toothed whales and baleen whales. It is also notable that the Mb sequence of pygmy hippopotamus (the closest living relative of cetaceans) differs from Ancestor A only in two positions, by having E27 and I142 like Ancestor B.

The most remarkable and repeated tendency in the evolution of cetacean Mb can be seen to be that of valine-to-isoleucine substitutions, starting with V28I and V101I in the transition to the common cetacean ancestor, and followed by V13I in the evolution of baleen whales and V21I in the sperm whales. All of these lead to increased hydrophobicity and expected increased folding stability due to reduced stability of the unfolded state. The variable site 121 is occupied by S or G in the terrestrial Artiodactyla relatives but by A in dolphins, porpoises, and Minke whales, and G and S in other cetaceans.

From Ancestor A to the closest living relative of cetaceans, the semi-aquatic hippopotamus, the following substitutions were found to have occurred (using the pygmy hippo Mb sequence): K35T, D45E, H49N, K57R, S59C, D61N, A85E, and Q92H. Of these, D45E is a mutation in a site that functions as distal gate-keeper for oxygen binding to heme, and it has substantial impact on the dive length, as seen in previous estimates of dive lengths based on O₂ binding affinities using an integrated Krogh model of muscle O₂ storage in diving mammals [12]. Site 35 was previously found to be positively selected in the transition to deep-diving cetaceans [1].

3.5. Concluding remarks and evolutionary implications

Three main findings are reported in this work: First, we identify strong correlations between stability effects of mutations vs. the stability of each respective wild type protein, and a damping effect of multi-site mutant stabilities vs. the sum of single-site mutant stability changes. From this, we conclude that proteins are subject to generic stability relaxation effects that should be accounted for in future studies of protein stability. Second, we develop a simple model that implements this generic relaxation to predict stability changes of multi-site variations with the same accuracy as for single-site variations by damping with the number of changed sites N . Third, we apply the new method to the case study of cetacean myoglobin evolution, and show how such a method can be useful.

Current thoughts on cetacean evolution supports the notion of a land-based ancestor of a clade residing within the artiodactyls [5], with the closest living terrestrial relative being the semi-aquatic hippopotamus [6], although specific locations within *Cetacea* remain debated [6][53]. *Pakicetus*, a wolf-sized, hoofed, land-based animal has been affiliated with *Cetacea* [5]: It has skull features that partly resemble whales, but its skeleton shows that it was fully land-based, and it drank fresh water [54]. Its upwards pointing eyes differ from any cetacean and suggests a semi-aquatic submergence strategy resembling that of crocodiles, consistent with the finding of fossils within previous rivers and coastal areas [54]. Then followed the increasingly aquatic ambulocetids and remintonocetids, which lost dependence on fresh water and had laterally positioned eyes, but still had no nasal drift towards the blowhole of modern cetaceans, and then the protocetidae, where nasal drift was initiated. The oldest known obligate aquatic cetaceans are the basilosaurids and dorudontids at about 35–40 Mya [5].

Mb is a protein with a typical rate of evolution in general (~1 amino acid substitution per site per one billion years). From this average rate, we would expect about 5 substitutions to have occurred

from Ancestor B at 35 Mya to any extant species. The evolution of Mb has been of this speed for the baleen whales but roughly three times faster in the toothed whales, as estimated from the analysis by Dasmeh et al. (~15 substitutions from Ancestor B to extant toothed whales) [1], and other researchers have confirmed the accelerated Mb evolution in cetaceans [3]. The increased concentration and stability [23], and possibly net charge [26], of cetacean Mbs are the main phenotypes that seem to have been selected relative to terrestrial relatives, and are thus likely responsible for the observable elevation of substitution rates, providing a direct example of positive selection in a mammal.

Both our calibrated methods suggest that the enhanced stability of Mbs in cetaceans is largely due to major adaptations occurring in the early transition from Ancestor A to B, a period that spans from the divergence of cetaceans from their closest relatives and to the common ancestor of current cetaceans. With the two methods, 1.3 and 0.7 kcal/mol were gained in this period. If Mb stability is a main adapted feature of deep-diving cetaceans, the finding that Mb stability increased in the transition somewhere between Pakicetus and the ancestor of contemporary toothed and baleen whales (B) agrees with the scenario that this ancestor and the closest living relatives, hippos, share a semi-aquatic life-style (as distinct from a purely terrestrial life style) and that the transition to the marine niche occurred after Pakicetidae, the first known likely ancestor of whales [7].

Furthermore, the high stability of Mb of the common ancestor B of contemporary cetaceans would imply that it was already fully aquatic. This is consistent with the morphology of basilosaurids and dorudontids that lived around the time of Ancestor B (35–40 Mya) and were possibly the first obligate aquatic cetaceans the size of large contemporary whales (Ancestor B may be a dorudontid, although this remains unknown) [5].

The 0.3–1 kcal/mol reduced Mb stability that we compute for dolphins, porpoises and baleen whales was complemented by analysis of evolutionary relaxation. We compared the evolutionary rate of Mb in these clades with that of sperm whales using the RELAX method [55]. In brief, RELAX tests the hypothesis of evolutionary rate relaxation in selected branches of a phylogenetic tree, compared with reference branches. A K-value is devised to evaluate whether selective strength was relaxed ($K < 1$) or intensified ($K > 1$). The results are shown in Figure 7B for the comparison of the clades in Figure 7A (numerical results can be found in Table S7): The estimated K was 0.25 against the null model in which $K = 1$ ($p = 0.028$). Dolphins and beaked whales (blue bars) converge towards the whale rates (red bars) and towards 1. Thus, the selection pressure towards stabilization of Mb shown previously [1] seems to have been relaxed in some lineages, based on both evolutionary rate relaxation analysis and actual computed Mb stabilities. After the move to the marine environment, several specializations occurred, e.g. filter feeding that affords large size without the need for deep diving, the transition to fresh-water habitats by river dolphins, and the radiation of dolphins [7]. As our data indicate, these groups have experienced relaxed Mb stability during their recent evolution.

Figure 7. Evolutionary rates of myoglobin evolution. A) Evolutionary rate relaxation analysis (RELAX) of dolphins, porpoises, and beaked whales, shown in blue as test branches, and sperm

whales, shown in red as reference branches; B) Proportion of sites (%) with evolutionary rates ω indicating relaxation behavior.

Supporting information

Table S1 shows the data set of stability changes upon mutation vs. the stability of the wild type protein used to correlate ΔG and $\Delta\Delta G$; Table S2 provides computed $\Delta\Delta G$ values used to assess the structure sensitivity of the applied protocols; Table S3 describes the computed and experimental $\Delta\Delta G$ values for the data set by Lin et al.[34]; Table S4 shows the computed and experimental $\Delta\Delta G$ values for the data set by Scott et al.[23]; Table S5 shows the $\Delta\Delta G$ values computed for the relevant ancient and contemporary Mb proteins of the cetacean clade; Table S6 and Figure S1 show the linear trend correlation of our best calibrated method vs. experimental ortholog data; Table S7 shows the numerical results for the evolutionary relaxation test.

Acknowledgements

This work was supported by the Danish Center for Scientific Computing (grant ID: 2012-02-23).

References

- [1] P. Dasmeh, A.W.R. Serohijos, K.P. Kepp, E.I. Shakhnovich, Positively Selected Sites in Cetacean Myoglobins Contribute to Protein Stability, *PLoS Comput. Biol.* 9 (2013) e1002929.
- [2] S. Mirceta, K.L. Campbell, M. Berenbrink, Molecular evolution of myoglobin in small diving mammals, *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.* 153 (2009) S98–S99. doi:10.1016/j.cbpa.2009.04.117.
- [3] M.F. Nery, J.I. Arroyo, J.C. Opazo, Accelerated evolutionary rate of the myoglobin gene in

- long-diving whales, *J. Mol. Evol.* 76 (2013) 380–387. doi:10.1007/s00239-013-9572-1.
- [4] M.R. McGowen, J. Gatesy, D.E. Wildman, Molecular evolution tracks macroevolutionary transitions in Cetacea, *Trends Ecol. Evol.* 29 (2014) 336–346. doi:10.1016/j.tree.2014.04.001.
 - [5] J.G.M. Thewissen, L.N. Cooper, J.C. George, S. Bajpai, From Land to Water: the Origin of Whales, Dolphins, and Porpoises, *Evol. Educ. Outreach.* 2 (2009) 272–288. doi:10.1007/s12052-009-0135-2.
 - [6] M.D. Uhen, Evolution of marine mammals: Back to the sea after 300 million years, *Anat. Rec. Adv. Integr. Anat. Evol. Biol.* 290 (2007) 514–522. doi:10.1002/ar.20545.
 - [7] J. Gatesy, J.H. Geisler, J. Chang, C. Buell, A. Berta, R.W. Meredith, et al., A phylogenetic blueprint for a modern whale, *Mol. Phylogenet. Evol.* 66 (2013) 479–506.
 - [8] B.S. Stewart, Diving Behavior, *Encycl. Mar. Mamm.* (2009) 321–327.
 - [9] S.R. Noren, T.M. Williams, Body size and skeletal muscle myoglobin of cetaceans: adaptations for maximizing dive duration., *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* 126 (2000) 181–191. doi:10.1016/S1095-6433(00)00182-3.
 - [10] S.H. Montgomery, J.H. Geisler, M.R. McGowen, C. Fox, L. Marino, J. Gatesy, the Evolutionary History of Cetacean Brain and Body Size, *Evolution* (N. Y). 67 (2013) 3339–3353. doi:10.1111/evo.12197.
 - [11] R.W. Davis, T.M. Williams, The marine mammal dive response is exercise modulated to maximize aerobic dive duration, *J. Comp. Physiol. A.* 198 (2012) 583–591. doi:10.1007/s00359-012-0731-4.
 - [12] P. Dasmeh, R.W. Davis, K.P. Kepp, Aerobic dive limits of seals with mutant myoglobin using combined thermochemical and physiological data, *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.* 164 (2013) 119–128. doi:10.1016/j.cbpa.2012.10.010.
 - [13] B.J. Smagghe, J. a. Hoy, R. Percifield, S. Kundu, M.S. Hargrove, G. Sarath, et al., Review: Correlations between oxygen affinity and sequence classifications of plant hemoglobins, *Biopolymers.* 91 (2009) 1083–1096.
 - [14] F. Draghi, A.E. Miele, C. Travaglini-Allocatelli, B. Vallone, M. Brunori, Q.H. Gibson, et al., Controlling Ligand Binding in Myoglobin by Mutagenesis, *J. Biol. Chem.* 277 (2002) 7509–7519.
 - [15] K.P. Jensen, U. Ryde, How O₂ Binds to Heme: REASONS FOR RAPID BINDING AND SPIN INVERSION, *J. Biol. Chem.* 279 (2004) 14561–14569.
 - [16] P. Dasmeh, K.P. Kepp, Bridging the gap between chemistry, physiology, and evolution: Quantifying the functionality of sperm whale myoglobin mutants, *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.* 161 (2012) 9–17. doi:10.1016/j.cbpa.2011.07.027.
 - [17] D.A. Drummond, C.O. Wilke, Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution., *Cell.* 134 (2008) 341–352. doi:10.1016/j.cell.2008.05.042.
 - [18] D.A. Drummond, J.D. Bloom, C. Adami, C.O. Wilke, F.H. Arnold, Why highly expressed proteins evolve slowly., *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 14338–14343.

doi:10.1073/pnas.0504070102.

- [19] J.D. Bloom, D.A. Drummond, F.H. Arnold, C.O. Wilke, Structural determinants of the rate of protein evolution in yeast., *Mol. Biol. Evol.* 23 (2006) 1751–1761. doi:10.1093/molbev/msl040.
- [20] A.W.R. Serohijos, Z. Rimas, E.I. Shakhnovich, Protein biophysics explains why highly abundant proteins evolve slowly., *Cell Rep.* 2 (2012) 249–256. doi:10.1016/j.celrep.2012.06.022.
- [21] F. Chiti, C.M. Dobson, Protein misfolding, functional amyloid, and human disease., *Annu. Rev. Biochem.* 75 (2006) 333–366. doi:10.1146/annurev.biochem.75.101304.123901.
- [22] K.P. Kepp, P. Dasmeh, A model of proteostatic energy cost and its use in analysis of proteome trends and sequence evolution., *PLoS One.* 9 (2014) e90504. doi:10.1371/journal.pone.0090504.
- [23] E.E. Scott, E. V. Paster, J.S. Olson, The stabilities of mammalian apomyoglobins vary over a 600-fold range and can be enhanced by comparative mutagenesis, *J. Biol. Chem.* 275 (2000) 27129–27136. doi:10.1074/jbc.M000452200.
- [24] P.P. Samuel, L.P. Smith, G.N.J. Phillips, J.S. Olson, Apoglobin Stability Is the Major Factor Governing both Cell-free and in Vivo Expression of Holomyoglobin., *J. Biol. Chem.* 290 (2015) 23479–23495. doi:10.1074/jbc.M115.672204.
- [25] P. Dasmeh, K.P. Kepp, Unfolding Simulations of Holomyoglobin from Four Mammals: Identification of Intermediates and β -Sheet Formation from Partially Unfolded States, *PLoS One.* 8 (2013) e80308. <http://dx.plos.org/10.1371/journal.pone.0080308>.
- [26] S. Mirceta, A. V. Signore, J.M. Burns, A.R. Cossins, K.L. Campbell, M. Berenbrink, Evolution of mammalian diving capacity traced by myoglobin net surface charge, *Science.* 340 (2013) 1234192. doi:10.1126/science.1234192.
- [27] S. Khan, M. Vihinen, Performance of protein stability predictors., *Hum. Mutat.* 31 (2010) 675–684. doi:10.1002/humu.21242.
- [28] V. Potapov, M. Cohen, G. Schreiber, Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details., *Protein Eng. Des. Sel.* 22 (2009) 553–560. doi:10.1093/protein/gzp030.
- [29] N.J. Christensen, K.P. Kepp, Stability Mechanisms of Laccase Isoforms using a Modified FoldX Protocol Applicable to Widely Different Proteins, *J. Chem. Theory Comput.* 9 (2013) 3210–3223. doi:10.1021/ct4002152.
- [30] J. Tian, N. Wu, X. Chu, Y. Fan, Predicting changes in protein thermostability brought about by single- or multi-site mutations., *BMC Bioinformatics.* 11 (2010) 370. doi:10.1186/1471-2105-11-370.
- [31] Y. Dehouck, J.M. Kwasigroch, D. Gilis, M. Rooman, PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality., *BMC Bioinformatics.* 12 (2011) 151. doi:10.1186/1471-2105-12-151.
- [32] E. Capriotti, P. Fariselli, R. Casadio, I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure, *Nucleic Acids Res.* 33 (2005) W306–W310.
- [33] M.M. Gromiha, Prediction of protein stability upon point mutations., *Biochem. Soc. Trans.* 35

(2007) 1569–1573.

- [34] L. Lin, R.J. Pinker, G.N. Phillips, N.R. Kallenbach, Stabilization of myoglobin by multiple alanine substitutions in helical positions., *Protein Sci.* 3 (1994) 1430–1435. doi:10.1002/pro.5560030909.
- [35] N.J. Christensen, K.P. Kepp, Accurate stabilities of laccase mutants predicted with a modified FoldX protocol., *J. Chem. Inf. Model.* 52 (2012) 3028–3042. doi:10.1021/ci300398z.
- [36] K.P. Kepp, Towards a “Golden Standard” for computing globin stability: Stability and structure sensitivity of myoglobin mutants, *Biochim. Biophys. Acta - Proteins Proteomics.* 1854 (2015) 1239–1248. doi:10.1016/j.bbapap.2015.06.002.
- [37] G.S. Kachalova, a N. Popov, H.D. Bartunik, A steric mechanism for inhibition of CO binding to heme proteins, *Sci. (New York, NY).* 284 (1999) 473–476. doi:10.1126/science.284.5413.473.
- [38] K. Harada, M. Makino, H. Sugimoto, S. Hirota, T. Matsuo, Y. Shiro, et al., Structure and ligand binding properties of myoglobins reconstituted with monodepropionated heme: functional role of each heme propionate side chain., *Biochemistry.* 46 (2007) 9406–9416. doi:10.1021/bi7007068.
- [39] A. Tomita, T. Sato, K. Ichianagi, S. Nozawa, H. Ichikawa, M. Chollet, et al., Visualizing breathing motion of internal cavities in concert with ligand migration in myoglobin, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 2612–2616. doi:0807774106 [pii]r10.1073/pnas.0807774106.
- [40] V. Parthiban, M.M. Gromiha, D. Schomburg, CUPSAT: prediction of protein stability upon point mutations., *Nucleic Acids Res.* 34 (2006) W239–W242. doi:10.1093/nar/gkl190.
- [41] E. Capriotti, P. Fariselli, R. Casadio, A neural-network-based method for predicting protein stability changes upon single point mutations., *Bioinformatics.* 20 Suppl 1 (2004) i63–i68. doi:10.1093/bioinformatics/bth928.
- [42] R. Schwarz, M. Dayhoff, Matrices for detecting distant relationships, in: M. Dayhoff (Ed.), *Atlas Protein Seq., National Biomedical Research Foundation*, 1979: pp. 353–358.
- [43] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.* 28 (2011) 2731–2739. doi:DOI 10.1093/molbev/msr121.
- [44] J. Chen, W.E. Stites, Higher-Order Packing Interactions in Triple and Quadruple Mutants of, *Biochemistry.* 40 (2001) 14012–14019. doi:10.1021/bi011269d.
- [45] N. Tokuriki, F. Stricher, J. Schymkowitz, L. Serrano, D.S. Tawfik, The stability effects of protein mutations appear to be universally distributed., *J. Mol. Biol.* 369 (2007) 1318–1332. doi:10.1016/j.jmb.2007.03.069.
- [46] R.A. Goldstein, The evolution and evolutionary consequences of marginal thermostability in proteins., *Proteins.* 79 (2011) 1396–1407. doi:10.1002/prot.22964.
- [47] D.A. Drummond, C.O. Wilke, The evolutionary consequences of erroneous protein synthesis., *Nat. Rev. Genet.* 10 (2009) 715–724. doi:10.1038/nrg2662.
- [48] D.D. Pollock, G. Thiltgen, R.A. Goldstein, PNAS Plus: Amino acid coevolution induces an evolutionary Stokes shift, *Proc. Natl. Acad. Sci.* 109 (2012) E1352–E1359.

doi:10.1073/pnas.1120084109.

- [49] K. Kepp, Computing Stability Effects of Mutations in Human Superoxide Dismutase 1, *J. Phys. Chem. B.* 118 (2014) 1799–1812.
- [50] D.S. Culbertson, J.S. Olson, Role of Heme in the Unfolding and Assembly of Myoglobin, *Biochemistry.* 49 (2010) 6052–6063. doi:10.1021/bi1006942.
- [51] J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, L. Serrano, The FoldX web server: an online force field, *Nucleic Acids Res.* 33 (2005) W382–W388. doi:10.1093/nar/gki387.
- [52] M.A. Flanagan, B. Garcia-Moreno, S.H. Friend, R.J. Feldmann, H. Scouloudi, F.R. Gurd, Contributions of individual amino acid residues to the structural stability of cetacean myoglobins, *Biochemistry* 22 (1983) 6027–6037.
- [53] J.H. Geisler, M.R. McGowen, G. Yang, J. Gatesy, A supermatrix analysis of genomic, morphological, and paleontological data from crown Cetacea., *BMC Evol. Biol.* 11 (2011) 112. doi:10.1186/1471-2148-11-112.
- [54] J.G. Thewissen, E.M. Williams, L.J. Roe, S.T. Hussain, Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls., *Nature.* 413 (2001) 277–281. doi:10.1038/35095005.
- [55] J.O. Wertheim, B. Murrell, M.D. Smith, S.L. Kosakovsky Pond, K. Scheffler, RELAX: Detecting Relaxed Selection in a Phylogenetic Framework., *Mol. Biol. Evol.* 32 (2014) 1–13. doi:10.1093/molbev/msu400.