



Oxygen channels and fractal wave–particle duality in the evolution of myoglobin and neuroglobin



Vedant Sachdeva, James C. Phillips*

Department of Physics and Astronomy, Rutgers University, Piscataway, NJ, 08854, United States

HIGHLIGHTS

- Amino acid sequences alone describe many functional differences between globins.
- Dynamics involve allosteric balance, quantified by hydrodynamic level sets.
- Differences between the two oxidation channels are seen in hydropathic waves.

ARTICLE INFO

Article history:

Received 10 March 2016

Received in revised form 9 May 2016

Available online 15 July 2016

Keywords:

Globin

Allometric

Hydropathic

Level set

Scaling

ABSTRACT

The evolution of terrestrial and aquatic globins is dominated by changes in two proximate–distal His channels, here monitored quantitatively by hydropathic waves. These waves reveal allometric functional features inaccessible to single amino acid stereochemical contact models, and even very large all-atom Newtonian simulations. The evolutionary differences between these features between myoglobin and neuroglobin are related to the two oxidation channels through hydropathic wave analysis, which identifies subtle interspecies functional differences inaccessible to traditional size and metabolic scaling studies. Level set analysis involves dynamic synchronization of allometric interactions across entire globins. Amino acid sequences alone show functional differences between species, which reflect basic metabolic differences (for instance, between temperate and tropical fresh water fish, or differing escape strategies of mice and rabbits).

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Here we discuss the evolution of myoglobin (Mb) and neuroglobin (Ngb) in functional terms, connecting hydropathic allometric globular properties and oxidation kinetics. Our analysis assumes that globins have evolved from aquatic predecessors to take nearly optimal advantage of the 28 times larger oxygen concentration available terrestrially. Mb stores O₂ in tissues, was the first protein whose structure was determined [1], and is the best understood globin. Ngb is concentrated in the neural network and the retina, and it exhibits instructive differences from Mb, and substantial sequence differences (similar overall length, but only 24% identity, 39% similarity). Features of Mb and Ngb evolution are largely hidden from structural studies, which generally show little difference between human and chicken protein backbones. Phylogenetic trees explore evolution in terms of single-site amino acid (aa) identities or similarities; more complex overall evolutionary scenarios are poorly described by such one-dimensional models [2].

It was first suggested that short-range hydrogen bond donation by the neutral NεH tautomer of distal His(E7) regulates Mb oxygen affinity, a small energy change [3]. Later structural studies showed that ligand entry into Mb occurs by swinging

* Corresponding author.

E-mail address: jcphillips8@comcast.net (J.C. Phillips).

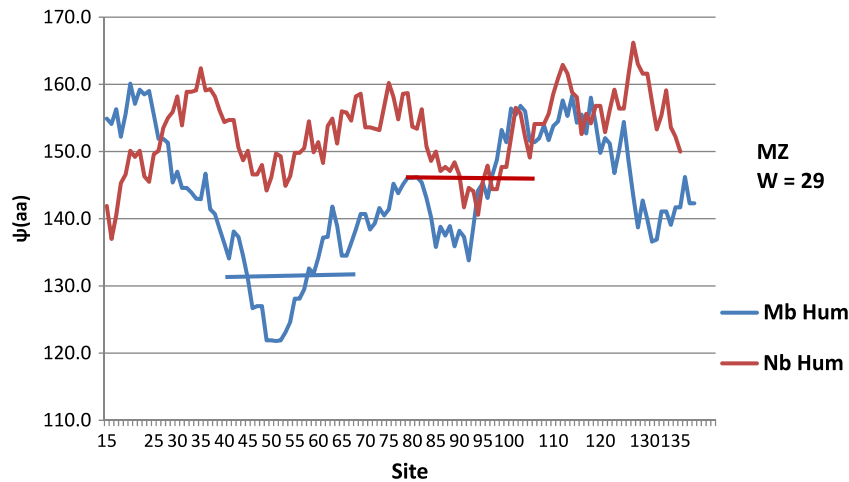


Fig. 1. The ψ (aa, $W = 29$) hydropathic profiles for human Myoglobin and Neuroglobin. Also shown is the ψ (aa) averages for the elastic region spanning 29 amino acids between the Mb heme contact points 42–71 discussed in the text. The much lower elastic blue value for Mb enables the E8 His (64) gate to open and function as a small gas molecule channel. The much larger red elastic value for Ngb stiffens this region, so that the functional channel is switched from Mb. In Ngb the channel goes from E8 (64) to the C-terminal ends of the E and H helices [7,8]. Note that in Mb the deepest hydrophilic minimum is centered near the distal His 64, while in Ngb the deepest minimum is centered near the proximate His 96. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

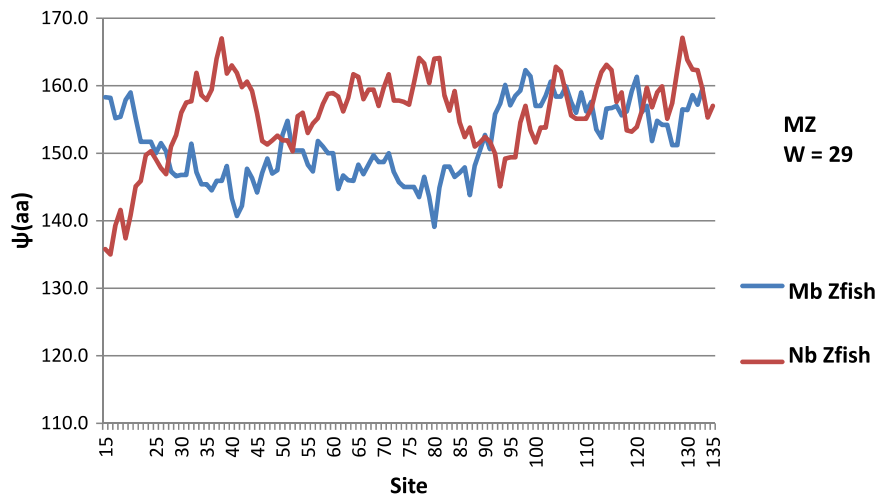


Fig. 2. The zebrafish profiles here should be compared to the human profiles of Fig. 1. The largest difference is the upward (hydrophobic) shift of the Mb hinge near 50 from human to zebrafish. The Ngb profile is little changed from fish to human in the 50–110 central region, which is shifted downwards in human by only 4 (about 10% of the overall range) in ψ (aa, $W = 29$).

the distal His E7 imidazole gate near the solvent edge of the Fe-binding porphyrin ring [4], a medium-range interaction (a still smaller energy) that also plays a key role in discrimination between O_2 and CO binding [5]. While the separation of the proximate and distal His in Euclidean space on opposite sides of the porphyrin ring is small, their separation tangentially along globin chains looping around the porphyrin ring is large, and in Mb and Hemoglobin (α , β) it is usually 29 aa. This fixed chain spacing is not so surprising for Mb ($\sim 75\%$ aa conservation between human and chicken), or for Hb ($\sim 70\%$ conservation), yet their functions (storage or transport) are different. Moreover, the recent discovery of Ngb, which appears to resemble Mb [6], has shown a dis-his Ngb spacing of 32 aa in all species sequenced so far. This could be related to a different oxidation channel opposite to the standard Mb channel, as shown by detailed structural comparisons ([7], esp. Figs. 2 and 3, and [8]). Here we are involved with still smaller energies.

The proximate–distal His channel aa spacing in Mb has oscillated from 32 (ultraprimitive vertebrates, lamprey and hogfish) to 25 (amphibian bullfrog), and finally settled at 29 in Mb and Hb α and β , across many species, including even primitive pufferfish Mb and lungfish Hb. The mechanics of channel opening and closing involve long-range interactions (network strain), not just medium range channel unblocking near His 64. Only recently has it become possible to show the existence of a single Mb oxidation channel spanning so many aa side groups, and involving long diffusion times, with a single escape pathway, using Newtonian mechanical models [9–11].

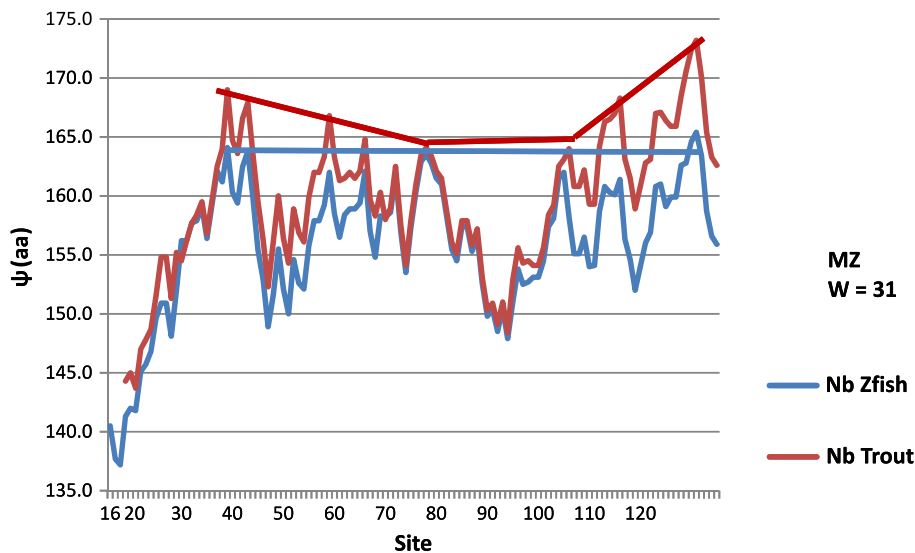


Fig. 3. Comparison of the hydroprofiles of tropical and temperate freshwater fish. The central region involving the heme and the distal His 64—proximate His 96 channel is conserved, while the remaining regions are parabolically more hydrophobic for temperate trout, and nearly level for tropical Zebrafish. Presumably the parabolically more hydrophobic regions are necessary because of larger temperature fluctuations in temperate regions. Note also the unmarked hydrophilic minima at 47 and 94. These also divide the proteins into three nearly equal regions, with the Zebrafish minima again being more nearly level.

2. Methods

For the reader's convenience we summarize here the main features of our recently developed bioinformatic scaling method [12]. The entire approach is post-Newtonian, and initially it may appear counter-intuitive. The central starting point is the implicit information contained in amino acid sequences in NCBI and their related structures in the PDB. Quite generally, one suspects that buried in these data are several key features of Nature's solutions for specific proteins to Levinthal's paradox (how proteins can function when they involve many exponentially complex, small energy differences) [13]. Newtonian all-atom simulation models that involve motions of thousands of atoms are not only extremely cpu intensive (typically months to reach a few results), but they do not address the evolutionary nature of protein sequence-structure-function relations, or provide tools for extracting functional information hidden in NCBI and PDB.

The first break in addressing this problem was the landmark discovery of a uniquely effective hydropathicity scale [14]. Competing effects of hydrophobic and hydrophilic segments of a given protein have long been known to be the primary driving force behind the folding of protein chains into protein globules. There are secondary effects associated with longitudinal hydrogen bonding (α helices) and transverse hydrogen bonding (β strands), and even weaker charge effects, but in most proteins the dominant physico-chemical factor in an overall property such as aggregation [12,15] is hydrophobic interactions. This new scale focuses on the hydrogen bonds in the interface between the water film and the protein. It is based on the solvent-accessible areas of amino acid van der Waals surfaces partitioned in overlapping regions according to the Voronoi rules. It has fundamental value because it describes globular protein sequence-structure-function relations in the context of a thermodynamic model of structural transitions dominated by the rearrangement energies of the protein–water interfacial hydrogen bond network.

The second step is unexpected, but the justification of it has become apparent from studying many examples. The first popular application of hydropathicity scales Ψ was to the study of all- α heptad transmembrane opsins, whose seven internal transmembrane segments are predominantly hydrophobic, with a typical length around 20 amino acids. One can carry this description to the similarly 20 amino acid thick layers adjacent to cell membranes where proteins can interact most effectively, as they are temporarily confined to a narrower space. Some readers will recognize a similarity to heterogeneous catalysis, with the membrane here serving as catalytic substrate. This picture is a generalization of enzyme catalysis (proteins with organic molecules) to all protein–protein interactions and evolution in vivo. It also helps to resolve Levinthal's paradox [13] on a functional level wider than protein folding, a larger puzzle than the big bang of cosmology.

In this membrane-based surface space proteins can both interact on a 20 amino acid length scale, or evolve via exchange of modular elements of ~ 20 amino acids. This is a post facto explanation for the remarkable discovery made by Ref. [14]. Instead of restricting the calculation of solvent-accessible amino acid areas to entire folded proteins from the PDB, [14] studied those areas associated with protein fragments, centered on specific amino acids, of length L , $9 \leq L \leq 35$, an interval itself centered on the typical membrane thickness 21. (These fragments are sometimes called modules in models of protein evolution [16]). These solvent-accessible areas were then found to be linear in a log–log plot against L , over the same range, for all 20 amino acids. This collinearity is indicative of special properties of self-organized (nearly perfect) protein networks

called self-similarity, and has been conjectured to be the key to a wide range of other nearly perfect networks. The amino-acid specific parameters of the model are universal fractals [17,18].

Our hydrophobic analysis shows that small-gas ligand (O_2 , CO, NO, ...) entry and exit globin allometric dynamics can be dominated by longitudinal hydrophobic strain at a single chain length, the spacing between proximate receptor–distal gate His. Compared to enriched multi-dimensional Newtonian all-atom simulation models [9–11], our hypermodern approach based only on one-dimensional optimized hydrophobic amino acid square wave profiles may appear simplistic. However, there are compensating advantages, as the economical methods used here have been applied to many other proteins with many different functions [19,20]. As we will see, evolution has employed similar optimizing tools on many proteins, which are recognizable from the very large amino acid sequence data base alone, with minimal support from known secondary α helix and β strand structural data.

The motivation for constructing allometric hydrophobic profiles is also based on thermodynamic concepts concerning the differential aqueous sculpting of globular protein surfaces near a critical point [12]. It can be compared to elastometric treatments of hinge-bending conformational transition pathways [21–23]. Euclidean structural models are static, but some dynamical features can be inferred by mutagenesis experiments that replace larger aa by smaller aa [8]. Note that mutagenic experiments involve non-WT proteins, whereas our bioinformatic scale involves only the differential surface geometry of WT proteins, a much smaller space with strong functional constraints.

Because globins perform the key elementary biochemical metabolic functions and are very well studied for many species, one can use them to extend and refine some previous correlations between hydrophobic extrema and species functionality, especially in connection with rapid escape from predators. Previously these largely referred to leveling of extrema as a mechanism to accelerate conformational motions by synchronizing chain segments that are well separated along the chain, but may be close in Euclidean space because of chain folding [19]. Here evolution has conserved the distal–proximate His spacing over many species, so that the same constructs can be used repeatedly, producing results that are nearly parameter-free and strongly over-determined.

3. Results: Two channels

Mb, Hb α and β , and Ngb globin sequence identities (positives) are less than 25(40)%. [7] presents a structure-based sequence alignment showing 75% α helices and heme contact points (their Fig. 1), as well as views of superposed Mb and Ngb numbered structures (their Fig. 2(A)). Apparently the heme stabilizes a common Mb–Ngb globin fold, while the oxidation channels of Mb and Ngb differ. These superposed Euclidean structures [7] can be compared to the human Mb and Ngb hydrophobic profiles shown here in Fig. 1.

The Mb channel, which exits past the distal His E7 gate, is opposite to the Ngb “apolar” tunnel, which is analogous to the exit past His100 in a mini-hemoglobin, as shown in Figs 1A and B of Ref. [8]. Experimental data, as well as the recent “milestone” calculation for Ngb [9], agree that each of these opposing channels is unique. Evolutionary optimization should favor a more efficient single channel, and our hydrophobic results in Fig. 1 show a simple elastic explanation for the existence of two opposing channels in Mb and Ngb. Relative to Mb, the Ngb profile shows a small softening near the beginning of the G chain, but the largest change by far is the Mb opening of a deep hydrophilic elastic hinge near 50, which releases the His E7 (64) imidazole ring gate for an exit channel.

Stated differently, the Ngb hinge near 50 is much less hydrophilic than the deep Mb valley. The soft Mb hinge is centered on the 17 aa sequence 45DKFKHLKSEDEMKASED61, which is strongly hydrophilic, as 13 aa are K, D or E. Using our rescaled version [15] of the MZ bioinformatic scale [20], which has an average hydrophobicity ~ 155 , the average for the soft Mb 45–61 sequence is 115 (strongly hydrophilic), compared to 45NCRQFSSPEDLSPE60 = 132 (only moderately hydrophilic; only 4 aa are R, D or E) for Ngb. These elastic sequences fill the chains between the mainly hydrophobic heme contact points (42, 44) and (64, 67, 68, 71) [7]. The full range of the contact points is again $W = 71 - 42 = 29$! According to Hooke’s law (1660, 1678), the average elasticity of this range is a rough estimate of the extent to which the distal His E7 imidazole ring gate (64) can open to create a channel entrance and exit.

Overall the human hydrophobic profiles shown in Fig. 1 contain three parts, hydrophobic near the N- and C- terminals, and an active central part 50–110 centered between distal and proximate His. In addition to the soft hinge near 50 the central region in Mb is bounded by a weaker second hinge above the proximate His F8 (96). The overall globular structure is stabilized by weakly hydrophobic shoulders near the C- and N- terminals. In Ngb, the softest region has switched to be centered on the proximate His F8 (96), and it seems very likely that this is another indication that the exit channel of Ngb is different from that of Mb.

The largest evolutionary changes occur between aquatic and terrestrial species, so Fig. 2 shows Mb and Ngb hydrophobic profiles of zebrafish, a tropical freshwater fish. It is immediately obvious that the largest difference is the upward (hydrophobic) shift of the Mb hinge near 50 from human to zebrafish. The Ngb profile is little changed from fish to human in the 50–110 central region: for example, the correlation there of (human, zfish) for ψ (aa, $W = 29$) is 0.62, while the same correlation for Mb is 0.41. Ngb (Zfish) exhibits a large number of nearly level hydrophobic extrema, and the main change in Ngb (human) is a small hydrophilic shift of the active central region, presumably increasing its activity.

Hydrophobic profiles are sensitive to function and environment, even for Ngb in freshwater fish, for example, zebrafish (tropical) compared to trout (temperate). As shown in Fig. 3, the N- and C- terminal hydrophobic stabilization trout shoulders

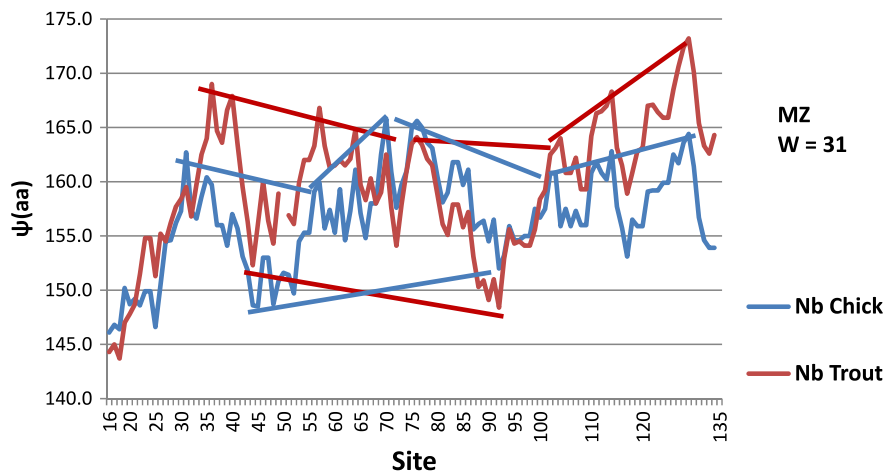


Fig. 4. Ngb chicken is unusual (see text).

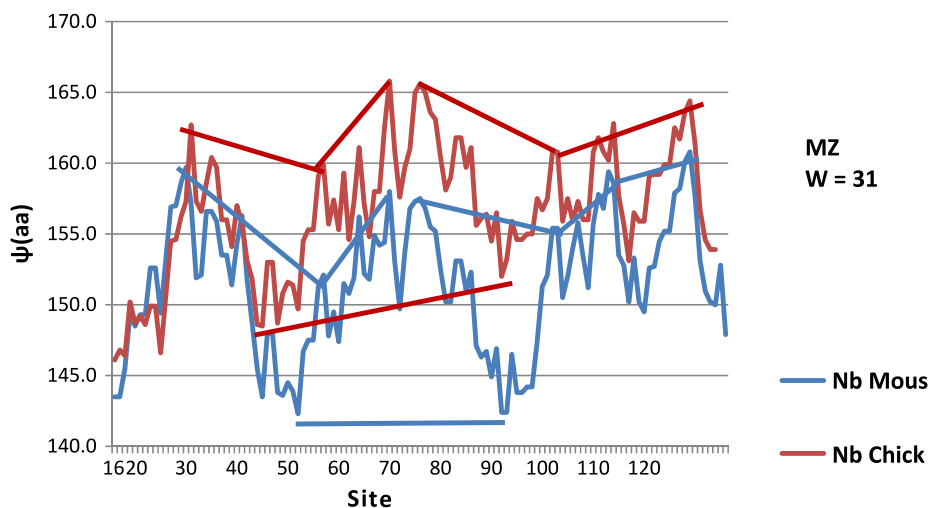


Fig. 5. Comparison of Ngb for chicken and mouse with the MZ scale.

(buttresses) are increased, so that the hydrophobic extrema change from flat (Zfish) to parabolic (trout). Meanwhile the central region, involving the oxidation channel, is almost unchanged.

4. Neuroglobin evolution

There are key changes from aquatic to terrestrial Ngb, specifically trout to chicken (Fig. 4). The ordering of the hydrophilic extrema reverses, but only slightly, which is probably not enough to shift the exit channel to distal His 64 in chicken. A new terrestrial feature emerges, as the hydrophobic maxima develop a W pattern, with the new central peak near 80. This is the Euclidean apex of the chain segments leading to distal His 64 and proximate His 96 on opposite heme sides [7]. Stiffening the central apex stabilizes the network contacting the heme. Studies of escape channels marked by pockets that accommodate Xe suggest that the apical chain region halfway between distal His 64 and proximate His 96 is the only region common to Mb and Ngb small molecule channels [24]. It is possible that mouse Ngb uses both distal His 64 and proximate His 96 channels for faster oxidation.

The chief Ngb change from chicken to mouse (Fig. 5) is a large mouse hydrophilic shift (35% of the profile width near the center hydrophobic maximum, and 40% at the proximate His F8 hydrophilic minimum). Making the overall Ngb profile more hydrophilic enables mouse to act more quickly, which is appropriate for a small rodent prey. Note the leveling of the two hydrophilic extrema. This facilitates synchronization of conformational transitions (Figure 3 of Ref. [19]), desirable for immediate mouse response to danger. The visual systems of primitive chickens were already optimized for perceiving rapid movement [25].

In earlier studies [12] of hydroprofiles based on the bioinformatic MZ scale [14,15], it was noted that because the MZ scale has both a thermodynamic and an evolutionary basis in terms of self-organized criticality, its results are generally more

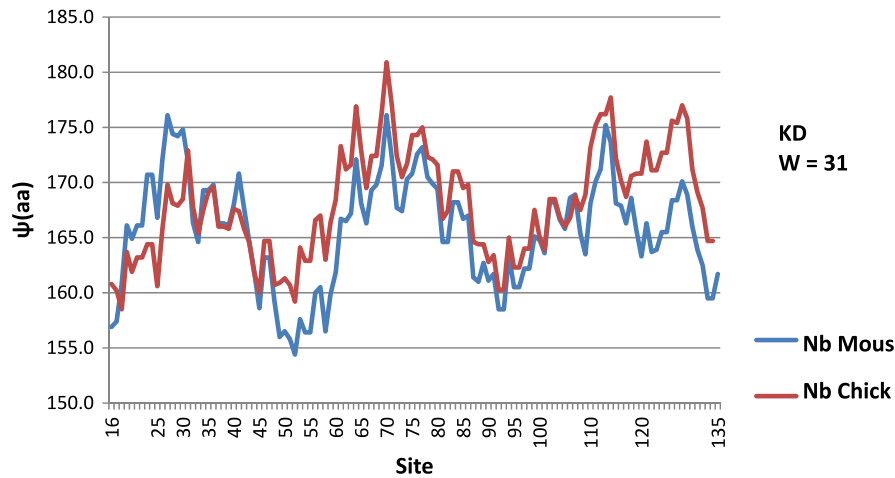


Fig. 6. Comparison of Ngb for chicken and mouse with the KD scale. The large differences seen with the MZ scale in Fig. 5 have become small or disappeared.

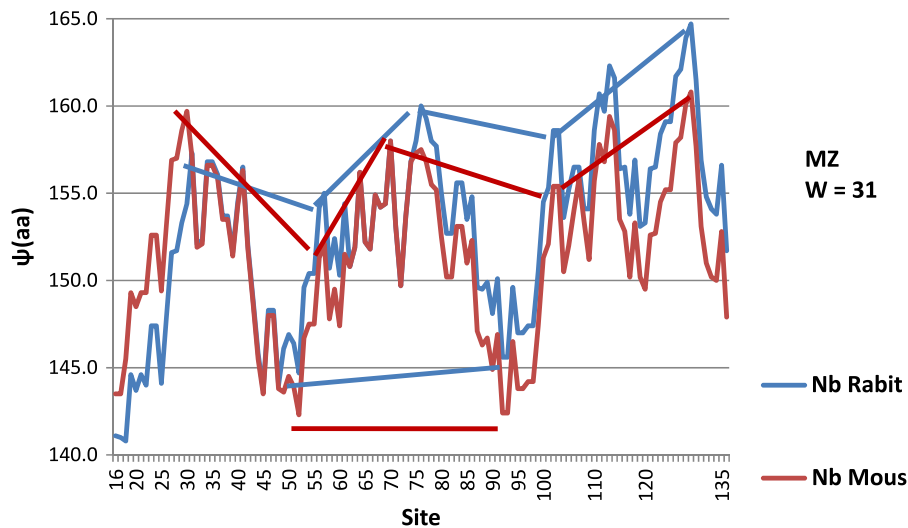


Fig. 7. Rabbit Ngb is more hydrophobic than mouse Ngb, and its hydrophilic hinges are less nearly level, reflecting a delayed escape strategy [27].

informative than those obtained with the standard KD scale, based on water–air complete chain unfolding [26]. Because the large differences between Ngb chicken and mouse shown in Fig. 5 with the MZ scale are so informative, Fig. 6 shows the same profiles calculated with the KD scale. Mouse is still more hydrophilic, but most of the differences have decreased sharply: for instance, the 40% MZ difference at the proximate His F8 hydrophilic minimum is $< 10\%$ with the KD scale. Also the KD hydrophilic extrema are not level, which disrupts conformational synchronization. The improvements using the MZ scale compared to the KD scale are an internal consistency check on the relevance of hydrophobic extrema.

Rabbit is much larger than mouse, and can afford a more sophisticated delayed escape strategy. The behavior of the hare is characterized by specific standing and flushing distances [27]. The speediest escape strategy of mouse is refined to include delayed escape by rabbit, as reflected by tilted hydrophilic extrema in Fig. 7. It also lives more than twice as long, and benefits from a more hydrophobic Ngb profile, which is what Fig. 7 also shows. Humans are much larger and live much longer, but the differences between mammals are small (approaching optimized criticality), see Fig. 8. In the proximate His 96 region Human Ngb closely resembles Mouse Ngb, while Human is more hydrophobic and more stable in the N- and C-terminal shoulders, consistent with greater longevity.

In most species' Ngb, there is a Cys disulfide bond between the C and D helices, with a variable Cys–Cys spacing of 6–9 aa (for example, human 45NCRQFSSPEDCLSSPE60). This disulfide bond must regulate dynamics of the CD loop [28]. While the disulfide bond is present in trout, chicken, rabbit and human, it is strikingly absent in mouse, where the dis–prox His tangential chain spacing is fully conserved. This is consistent with the extremally hydrophilic Ngb profile of mouse in the central region between dis and prox His, consistent with the mouse's need for rapid escape.

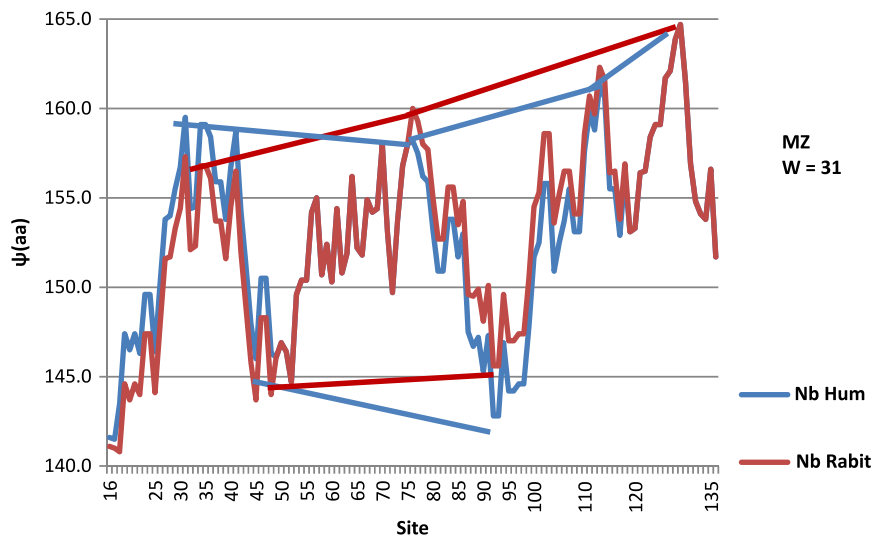


Fig. 8. Compared to rabbit Ngb, human Ngb is stabilized near the N terminal, and more hydrophilic at the proximate His 96 site.

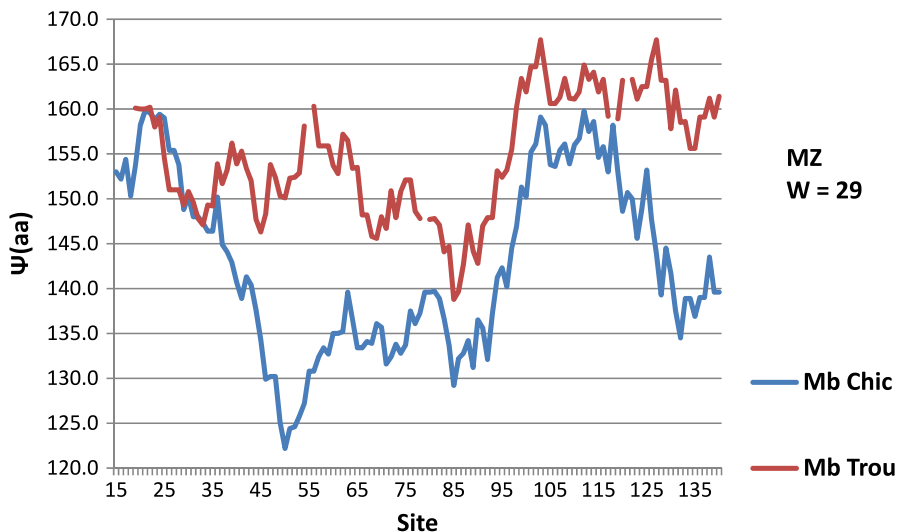


Fig. 9. Hydrophobicity profiles of Mb trout and chicken. Trout is much more hydrophobic, and the elastically soft hydrophilic region near 50 has become hydroneutral.

5. Myoglobin evolution

Mb is one of the most studied proteins [1,3,4], and the emerging evidence, particularly from fish species, indicates that Mb fulfills a broad array of physiological functions in a wider range of different tissues [29]. However, almost all Mb structural studies have focused on sperm whale or human Mb. Studies of tuna have shown that the D helix centered on 55 has been disordered [30]. As we saw in discussing Fig. 1, the key to the flexibility of the His 64 gate is the hydrophilic softness of the chain region between the heme contact points (42, 44) and (64, 67, 68, 71) [7]. Helix D stabilizes this chain region, and when this region becomes hydroneutral (Fig. 9), it is possible that in the absence of helix D the distal His E7 may still function as a gate for a channel similar to the terrestrial oxidation channel, but probably more complex. At present the Mb aquatic oxidation channel is unknown.

For mouse we concluded that Ngb is especially hydrophilic because mouse must recognize danger quickly (Ngb is concentrated in retina and brain [31]). The most striking new feature, after chicken, of Mb in mouse (Fig. 10) is the appearance of the apical hydrophobic peak at 80, which we suggested above stabilizes the network contacting both sides of the heme. Such stabilization could increase signal transmission and release oxygen more rapidly to tissues.

The final stage of Mb evolution from mouse to human is shown in Fig. 11. As with Ngb, mouse is an extremum between chicken and mammals, especially near the apical site 80 heme stabilizing hydrophobic peak. The overall differences are

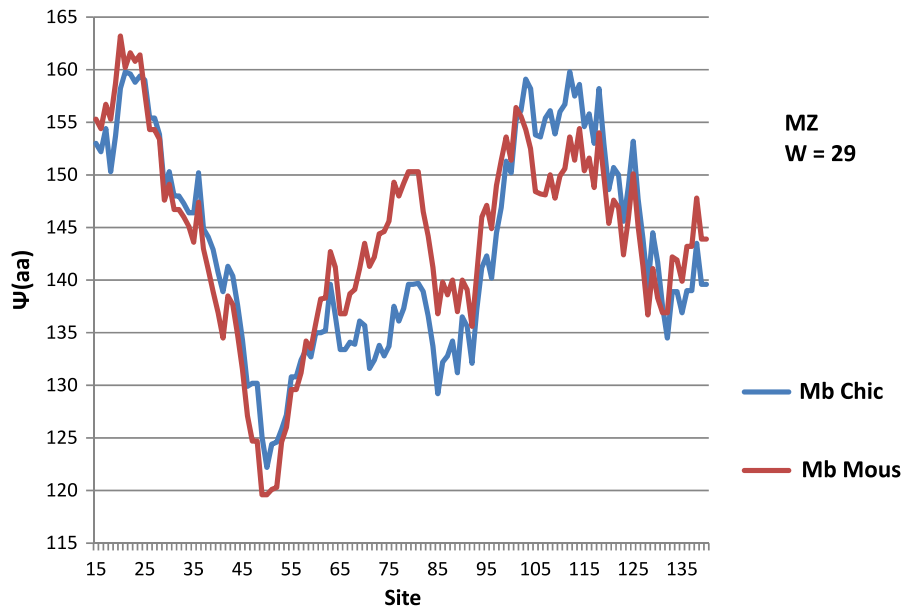


Fig. 10. The largest change here is the mouse apical hydroneutral peak near site 80.

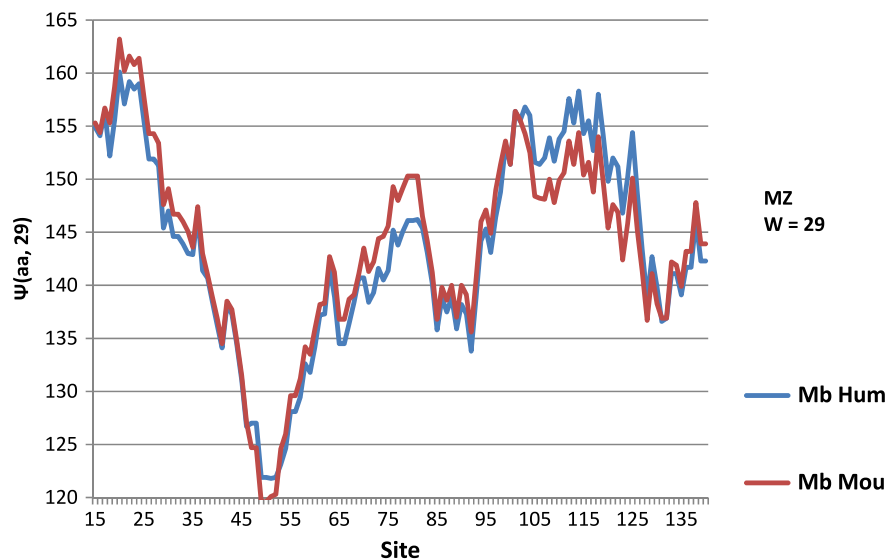


Fig. 11. Mb has stabilized from rodents to mammals.

smaller with Mb than with Ngb, which was to be expected, as Ngb is finely tuned to function optimally in neural networks, whereas Mb has many functions in all tissues.

We saw in Figs. 5–7 how the extremal properties of mouse Ngb appeared as much larger differences with the MZ conformational (second-order) scale, than with the KD scale. In Fig. 12 we show the Mb profiles of Fig. 11 (MZ scale), now with the KD scale. With Mb profiles obtained with the two scales are much closer, and the differences are now about 10% larger with the KD scale. At first this seems surprising, but it could be merely a reflection of the fine tuning of Ngb for neural networks, and the wide range of functions of Mb, which are better described using the first-order KD scale suited to protein unfolding. In the network glass and polymer literature these heterogeneous first-order effects [32] are sometimes described as random, which is simplistic [33]. The homogeneous limit of dynamical network stretched exponential relaxation is far from random, and involves topological fractals similar to the MZ scaling fractals [34].

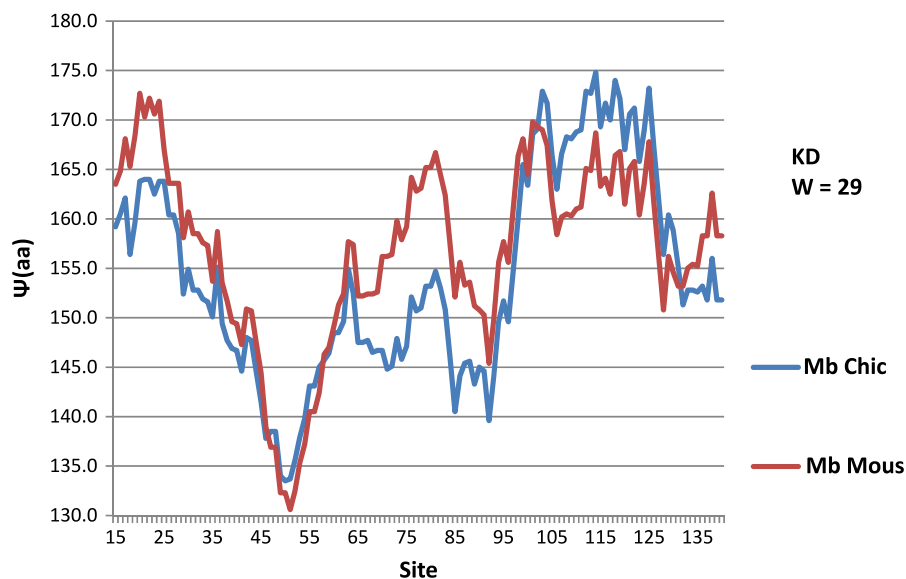


Fig. 12. Here the profiles are evaluated using the first-order KD scale, to be compared with Figs. 10, 5 and 6.

6. Discussion

Both the structure and function of a protein chain in its native state are determined by its amino acid sequence [35]. Mutagenic and structural studies have established many Euclidean short-range features of Mb and Ngb, including the existence of different oxidation channels [1,4,7–11].

Hydropathic scaling provides a different and complimentary picture, involving long-range (allometric) elastic interactions at the molecular level, which connects amino acid sequences directly to function, as well as elucidating previously hidden structural features. The most striking feature of all Mb and Ngb profiles is their tripartite structure, which is barely visible in aquatic profiles (especially tropical), but is well developed and closely related to function in terrestrial species. The central peak is associated with the porphyrin ring, while the side extrema facilitate allometric synchronized motion of N and C wings. This produces symmetric patterns, especially for mouse, which it requires for its escape from predators. On a more detailed level, profiles identify the elastically flexible region near 50 in Mb, which enables His E7 imidazole ring to function as the Perutz gate in Mb. In Ngb the hydrophilic region near 50 stiffens, and it is no longer the minimum. That minimum has shifted to near the proximate His at 96, and this shift is accompanied by the appearance of an oxidation channel on the opposite side of the heme. Thus tripartite hydropathic profiles alone strongly suggest different escape channels, from opposite sides of the heme, in Mb and Ngb.

The extreme elasticity of the Mb hinge near 50 should be compared to that of several more elastic regions: in the 230–260 region of the A4 (amyloid precursor) 770 aa protein of the neural network, and near the center of 413 aa Aspein, which is responsible for directed formation of calcite in the shell of the pearl oyster—an ultrasoft buffer supporting the growing ultrahard oyster shell [36,37].

Returning to Figs. 1 and 2, one notices that Mb has evolved from aquatic to terrestrial species much more than Ngb. At first, this seems surprising, but aquatic (terrestrial) species are cold- (warm-) blooded, and ligand exchange kinetics in respiratory tissues of aquatic species may be quite different from that in terrestrial species [38]. By contrast, the neural and retinal networks of all vertebrates differ largely in size [36].

The leveling of hydropathic extrema emphasized in Figs. 3–8 has been discussed by differential geometers in terms of level sets as applied to the diffusive flow of liquids, in at least 700 papers [39,40]. The synchronization comparisons made here are driven in terrestrial species by primary biological forces (prey vs. predator).

7. Conclusions

The systematics of Mb and Ngb evolution can be described accurately using two hydropathic scales [14,26], with the second-order MZ scale being much more accurate for Ngb, and the first-order KD scale being slightly more accurate for Mb. The evolution of the competing enzymatic and lytic functions of lysozyme c, and its role in suppressing amyloid aggregation, are described much more accurately by the MZ scale [12]. This scale represents the quantitative realization, using the tools of thermodynamics and differential geometry, of the general theory of dominant conformational motions slaved by the hydration shell and the bulk solvent [41]. Specific soft modes are seen in Mb–O₂ compared with those in Mb–CO [42]. The quantitative power of bioinformatic scales is especially large for the globins, because they perform relatively simple metabolic functions, compared to lysozyme c, which performs at least three functions.

It is interesting to compare the present molecular scaling treatment of the evolution of allometric interactions with macroscopic models relating vertebrate metabolism to surface area [43]. It appears that these vary significantly between fish, amphibians, reptiles, birds and mammals [44]. There is also substantial scatter in visual response rates when these are scaled by size or metabolism [45]. Here we have seen significant differences even between temperate and tropical fish (Fig. 3). We have easily connected molecular differences between mouse Ngb and other species Ngb to the mouse's predator escape strategy.

Scaling, improved by inclusion of multiple individual factors, is likely to be an increasingly important tool in medical contexts [46,47]. At the molecular level it has already provided a detailed description of viral sequence evolution [48], which may well be more accurate in selecting suitable vaccine targets than antigenic clustering [49,50]. The accuracy of scaling in describing the functional evolution of Ngb suggests that it could be usefully employed in metabolomics when its database is larger [51]. Because of the complexity of proteins and their interactions, economy, simplicity and universality are critical features of theoretical biophysical tools.

Methods Although bioinformatic scaling is novel, the calculations described in this paper used only an EXCEL Macro built by Niels Voorhoeve and refined by Douglass C. Allan. Our basic assumption that proteins are nearly perfect and are near a self-organized critical point has been proved in a recent study 7678 NMR data of proteins with 10 different structures [52]. The conclusion is that all studied proteins were thermodynamically critical, in the sense that the effective correlation length spanned the entire chain.

References

- [1] J.C. Kendrew, G. Bodo, H.M. Dintzis, et al., A three-dimensional model of the myoglobin molecule obtained by X-ray analysis, *Nature* 181 (1958) 662–666.
- [2] D.H. Huson, D. Bryant, Application of phylogenetic networks in evolutionary studies, *Mol. Biol. Evol.* 23 (2006) 254–267.
- [3] L. Pauling, Nature of the iron–oxygen bond in oxyhaemoglobin, *Nature* 203 (1964) 182–183.
- [4] M.F. Perutz, F.S. Mathews, An x-ray study of azide methaemoglobin, *J. Mol. Biol.* 21 (1966) 199–202.
- [5] I. Birukou, R.L. Schweers, J.S. Olson, Distal histidine stabilizes bound O₂ and acts as a gate for ligand entry in both subunits of adult human hemoglobin, *J. Biol. Chem.* 285 (2010) 8840–8854.
- [6] T. Burmester, T. Hankeln, What is the function of neuroglobin? *J. Exp. Biol.* 212 (2009) 1423–1428.
- [7] A. Pesce, S. Dewilde, M. Nardini, et al., Human brain neuroglobin structure reveals a distinct mode of controlling oxygen affinity, *Structure* 11 (2003) 1087–1095.
- [8] M.D. Salter, G.C. Blouin, J. Soman, et al., Determination of ligand pathways in globins: apolar tunnels versus polar gates, *J. Biol. Chem.* 287 (2012) 33163–33178.
- [9] T.-Q. Yu, M. Lapelosa, E. Vanden-Eijnden, et al., Entry, internal diffusion, and exit in myoglobin from transition-path theory simulations, *J. Am. Chem. Soc.* 137 (2015) 3041–3050.
- [10] H. Na, G. Song, Quantitative delineation of how breathing motions open ligand migration channels in myoglobin and its mutants, *Proteins- Struct. Funct. Bioinform.* 83 (2015) 757–770.
- [11] J.P. Arcon, P. Rosi, A.A. Petruk, et al., Molecular mechanism of myoglobin autoxidation: Insights from computer simulations, *J. Phys. Chem. B* 119 (2015) 1802–1813.
- [12] J.C. Phillips, Fractals and self-organized criticality in proteins, *Phys. A* 415 (2014) 440–448.
- [13] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, *Nat. Struct. Biol.* 4 (1997) 10–19.
- [14] M.A. Moret, G.F. Zebende, Amino acid hydrophobicity and accessible surface area, *Phys. Rev. E* 75 (2007) 011920.
- [15] J.C. Phillips, Scaling and self-organized criticality in proteins: Lysozyme c, *Phys. Rev. E* 80 (2009) 051916.
- [16] H. Lipson, J.B. Pollack, N.P. Suh, On the origin of modular variation, *Evolution* 56 (2002) 1549–1556.
- [17] B.B. Mandelbrot, Is nature fractal? *Science* 279 (1998) 783.
- [18] P. Bak, K. Sneppen, Punctuated equilibrium and self-organized criticality in a simple model of evolution, *Phys. Rev. Lett.* 71 (1993) 4083–4086.
- [19] J.C. Phillips, Fractals and self-organized criticality in anti-inflammatory drugs, *Phys. A* 415 (2014) 538–543.
- [20] J.C. Phillips, Similarity is not enough: Tipping points of ebola zaire mortalities, *Physica A* 427 (2015) 277–281.
- [21] N. Sinha, S. Kumar, R. Nussinov, Interdomain interactions in hinge-bending transitions, *Structure* 9 (2001) 1165–1181.
- [22] L. Yang, G. Song, R.L. Jernigan, How well can we understand large-scale protein motions using normal modes of elastic network models? *Biophys. J.* 93 (2007) 920–929.
- [23] L. Cardamone, A. Laio, V. Torre, R. Shahapure, A. DeSimone, Cytoskeletal actin networks in motile cells are critically self-organized systems synchronized by mechanical interactions, *Proc. Natl. Acad. Sci. USA* 108 (2011) 13978–13983.
- [24] A. Tomita, T. Sato, K. Ichianagi, et al., Visualizing breathing motion of internal cavities in concert with ligand migration, *Proc. Natl. Acad. Sci. USA* 106 (2009) 2612–2616.
- [25] D. Rubene, O. Hastad, R. Tauson, et al., The presence of UV wavelengths improves the temporal resolution of the avian visual system, *J. Exp. Biol.* 213 (2010) 3357–3363.
- [26] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* 157 (1982) 105–132.
- [27] S. Focardi, M. Rizzotto, Optimal strategies and complexity: A theoretical analysis of the anti-predatory behavior of the hare, *Bull. Math. Biol.* 61 (1999) 829–847.
- [28] B.G. Guimaraes, D. Hamdane, C. Lechauve, et al., The crystal structure of wild-type human brain neuroglobin reveals flexibility of the disulfide bond that regulates oxygen affinity, *Acta Crystallogr. D* 70 (2014) 1005–1014.
- [29] S. Helbo, R.E. Weber, A. Fago, Expression patterns and adaptive functional diversity of vertebrate myoglobins, *Biochim. Biophys. Acta, Proteins Proteomics* 38 (2013) 1311–1322.
- [30] G.I. BirNgbbaum, S.V. Evans, M. Przybylska, et al., 1.70 Angstrom resolution structure of myoglobin from yellowfin tuna – an example of a myoglobin lacking the D-helix, *Acta Crystallogr. D* 50 (1994) 283–289.
- [31] M. Schmidt, A. Giessl, T. Laufs, et al., How does the eye breathe? Evidence for neuroglobin-mediated oxygen supply in the mammalian retina, *J. Biol. Chem.* 278 (2003) 1932–1935.
- [32] G. Adam, J.H. Gibbs, On temperature dependence of cooperative relaxation properties in glass-forming liquids, *J. Chem. Phys.* 43 (1965) 1398.
- [33] T.R. Kirkpatrick, D. Thirumalai, Colloquium: Random first order transition theory concepts in biology and physics, *Rev. Modern Phys.* 87 (2015) 183–209.
- [34] R.C. Welch, J.R. Smith, M. Potuzak, et al., Dynamics of glass relaxation at room temperature, *Phys. Rev. Lett.* 110 (2013) 265901.
- [35] C. Anfinsen, The formation and stabilization of protein structure, *Biochem. J.* 128 (1972) 737–749.
- [36] J.C. Phillips, Thermodynamic description of beta Amyloid formation using physico-chemical scales and fractal bioinformatic scales, *ACS Chem. Neurosci.* 6 (2015) 745–750.
- [37] J.C. Phillips, Fractal Scaling of Cortical Matter, Amyloid Fragmentation and Plaque Formation across Rodents and Primates, 2014. [arXiv1410.1419](https://arxiv.org/abs/1410.1419).

- [38] L. Sundin, M.L. Burleson, A.R. Sanchez, et al., Respiratory chemoreceptor function in vertebrates - comparative and evolutionary aspects, *Integr. Comp. Biol.* 47 (2007) 592–600.
- [39] L.C. Evans, J. Spruck, Motion of level sets by mean curvature.1, *J. Differential Geom.* 33 (1991) 635–681.
- [40] J.A. Sethian, Analysis and applications of the Voronoi Implicit Interface method, *J. Comput. Phys.* 231 (2012) 6051–6085.
- [41] H. Frauenfelder, G. Chen, Joel Berendzen, et al., A unified model of protein dynamics, *Proc. Natl. Acad. Sci. USA* 106 (2009) 5129–5134.
- [42] K.N. Woods, Using THz time-scale infrared spectroscopy to examine the role of collective, thermal fluctuations in the formation of myoglobin allosteric communication pathways and ligand specificity, *Soft Matter* 10 (2014) 4387–4402.
- [43] G.B. West, J.H. Brown, B.J. Enquist, The fourth dimension of life: Fractal geometry and allometric scaling of organisms, *Science* 284 (1999) 1677–1679.
- [44] C.R. White, N.F. Phillips, R.S. Seymour, The scaling and temperature dependence of vertebrate metabolism, *Biol. Lett.* 2 (2006) 125–127.
- [45] K. Healy, L. McNally, G.D. Ruxton, et al., Metabolic rate and body size are linked with perception of temporal information, *Anim. Behav.* 216 (2013) 685–696.
- [46] F.E. Dewey, D. Rosenthal, Dj Murphy, et al., Does size matter? Clinical applications of scaling cardiac size and function for body size, *Circulation* 117 (2008) 2279–2287.
- [47] D.A. Bluemke, R. Kronmal, J.A. Lima, et al., The relationship of left ventricular mass and geometry to incident cardiovascular events, *J. Am. Coll. Cardiol.* 52 (2008) 2148–2155.
- [48] J.C. Phillips, Punctuated evolution of influenza virus neuraminidase (A/H1N1) under opposing migration and vaccination pressures, *BioMed. Res. Int.* 2014 (2014) 907381.
- [49] J.M. Fonville, S.H. Wilks, S.L. James, et al., Antibody landscapes after influenza virus infection or vaccination, *Science* 346 (2014) 996–1000.
- [50] X. Li, J.C. Phillips, M.W. Deem, Early recognition of emerging flu strain clusters, 2015. [arXiv1512.08442](https://arxiv.org/abs/1512.08442).
- [51] D.B. Kell, *Metabolomics and systems biology: making sense of the soup*, *Curr. Opin. Microbiol.* 7 (2004) 296–307.
- [52] Q.-Y. Tang, Y.Y. Zhang, J. Wang, et al. Critical fluctuations in proteins native states, 2016. [arXiv1601.03420](https://arxiv.org/abs/1601.03420).