AIMS Biophysics

AIMS Biophysics, 13(1): 21−28.
DOI: 10.3934/biophy.2026002
Received: 07 January 2025
Revised: 30 January 2026
Accepted: 03 February 2026
Published: 05 February 2026

http://www.aimspress.com/journal/biophysics

Review

Secondary structure of proteins in the context of molecular interactions

Yoshito Abe*

School of Pharmacy at Fukuoka, International University of Health and Welfare, 137-1 Enoki-zu, Okawa Fukuoka 831-8501, Japan

* **Correspondence:** Email: y_abe@ihwg.jp; Tel: +81944326128; Fax: +81-944-89-2001.

Abstract: Protein secondary structure has long been viewed as a local property encoded in the amino acid sequence. While classical definitions of α-helices and β-sheets and established experimental and computational methods have provided a solid framework for structural analysis, growing evidence indicates that secondary structure is strongly influenced by intra- and intermolecular interactions. Such context dependence poses significant challenges for conventional sequence-based prediction approaches. In this review, I discuss the limitations of traditional secondary-structure prediction methods and highlight recent advances enabled by deep learning–based structure prediction, with a particular focus on AlphaFold. Using oligomeric proteins and polytopic membrane proteins as representative examples, I illustrate how interaction-dependent secondary structures can be revealed by integrating computational predictions with experimental data. These observations emphasize the importance of considering molecular interaction contexts to achieve a comprehensive understanding of protein secondary structure and its functional roles.

Keywords: secondary structure; oligomerization; polytopic membrane protein; protein-protein interaction; protein folding

1. Introduction

Protein secondary structure denotes recurrent local conformational motifs of the polypeptide backbone and constitutes a fundamental level of protein structural organization. The canonical α-helix and β-sheet were formulated by Pauling and colleagues in 1951 [1,2] and were later substantiated by X-ray crystallographic analyses of myoglobin and hen egg white lysozyme [3,4]. The α-helix is stabilized by intramolecular hydrogen bonds between the backbone amide (NH) group of residue i and

the carbonyl (CO) group of residue i+4, giving rise to a right-handed helical conformation with side chains oriented outward from the helical axis. In contrast, β-sheets are formed through interstrand hydrogen-bonding networks, producing extended sheet-like architectures in which side chains alternate on opposite sides of the sheet plane.

Side-chain interactions emerging from these secondary-structure frameworks play essential roles in both intramolecular folding and intermolecular association. The secondary structures are routinely categorized based on backbone dihedral angles (φ and ψ), as represented in the Ramachandran plot [5]. In addition, residues deposited in the Protein Data Bank (PDB) are assigned to one of eight secondary-structure states according to hydrogen-bonding patterns defined by the Dictionary of Secondary Structure of Proteins (DSSP), derived from X-ray crystallographic and cryo-electron microscopy data [6]. This formalism has become an indispensable standard for describing protein structural features. Experimental characterization of secondary structure at residue-level resolution is achievable using NMR-based techniques, including chemical shift analysis [7], residual dipolar couplings [8], and TALOS-derived torsion angle prediction [9]. In contrast, ensemble-averaged secondary-structure content is commonly evaluated using circular dichroism (CD) and infrared (IR) spectroscopy. Among these methods, CD spectroscopy remains particularly advantageous for solution-state analysis, providing rapid and quantitative estimates of α-helical and β-sheet content based on characteristic spectral signatures [10].

For proteins recalcitrant to experimental structure determination, sequence-based secondary-structure prediction methods have been intensively developed since the 1970s. The pioneering Chou–Fasman method exploited empirically derived amino acid propensities based on the limited structural data available at that time [11,12]. Although conceptually simple, this approach achieved prediction accuracies of approximately 70% for soluble globular proteins and laid the foundation for subsequent methodological advances. Later generations of predictors incorporated expanded structural databases, statistical learning frameworks, and experimentally informed constraints, resulting in incremental but consistent improvements in predictive performance [13−15]. Despite these advances, accurate secondary-structure prediction remains problematic for proteins lacking close homologs in existing databases, particularly oligomeric assemblies with undefined intermolecular interfaces and polytopic membrane proteins. Secondary structure has traditionally been treated as a predominantly local property encoded in the primary sequence. However, a growing body of evidence indicates that secondary-structure formation and stability are strongly influenced by intra- and intermolecular interactions. In this review, I critically examine the limitations of conventional secondary-structure prediction paradigms and discuss emerging perspectives informed by interaction-dependent structural behavior.

2. Secondary-structure prediction at interaction interfaces

Protein tertiary structure arises from specific interactions among secondary-structure elements, and perturbation of these interactions can induce pronounced secondary-structure rearrangements. A paradigmatic example is provided by amyloidogenic proteins, which undergo conversion from native conformations to β-sheet-rich fibrillar assemblies [16]. Likewise, so-called chameleon sequences exhibit context-dependent structural plasticity, adopting either α-helical or β-sheet conformations depending on their molecular environment [17]. In addition, numerous intrinsically disordered proteins acquire stable secondary structures upon binding to cognate partners [18]. These phenomena

underscore fundamental limitations in predicting secondary structure solely from amino acid sequence. The prediction of interactions among secondary-structure elements poses an even greater challenge, particularly for proteins whose three-dimensional architectures remain unresolved. Homology modeling approaches, such as SWISS-MODEL [19], can provide useful approximations when sufficient sequence similarity exists; however, their reliability deteriorates rapidly in low-homology regimes. The advent of AlphaFold2 in November 2020 represented a major conceptual shift in protein structure prediction [20]. AlphaFold2 integrates co-evolutionary signals extracted from multiple sequence alignments with Transformer-based deep learning architectures to infer residue–residue geometric constraints, which are subsequently assembled into three-dimensional structural models. While the complete workflow includes iterative refinement and physics-based relaxation, the central innovation lies in the deep integration of evolutionary and geometric information. AlphaFold-Multimer extends this framework to the prediction of protein–protein interactions within oligomeric complexes [21].

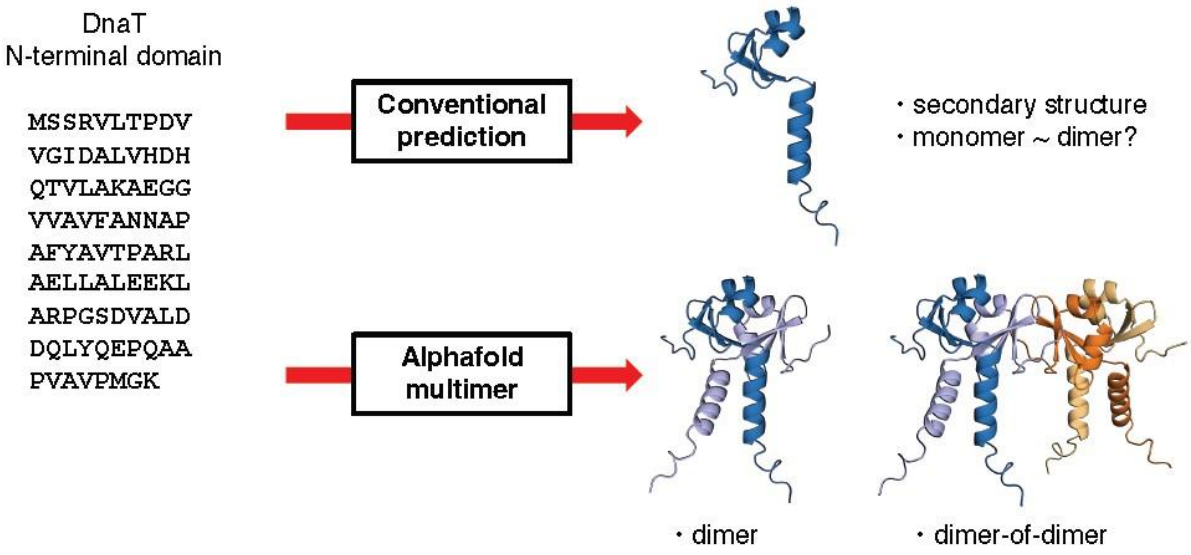

Figure 1. Prediction of the oligomeric structure of the DnaT N-terminal domain.

Using this approach, together with co-workers, I investigated the oligomeric organization of DnaT, a replication restart protein from *Escherichia coli*. DnaT comprises N-terminal and C-terminal domains linked by a flexible segment [22−24]. The C-terminal domain binds single-stranded DNA and forms a helical assembly, as established by X-ray crystallography [25], whereas the oligomeric arrangement of the N-terminal domain had remained unresolved. Biochemical analyses, including size-exclusion chromatography and fluorescence spectroscopy, suggested that the N-terminal domain forms trimers or tetramers [26,27]. Structural characterization, however, was impeded by low-resolution crystallographic data and extensive NMR line broadening due to oligomerization [22,23]. On the basis of earlier fold and secondary-structure predictions, the domain was proposed to adopt an antitoxin-like β–α–β–β–α topology, although direct experimental validation was not achievable. Given that antitoxin proteins frequently assemble as homodimers, AlphaFold-Multimer was first applied to predict a dimeric structure of the DnaT N-terminal domain [22]. Interface-directed mutagenesis demonstrated that substitutions at predicted α-helical and β-sheet contact surfaces destabilized oligomer formation, lending support to the dimeric model. Consistent with experimental indications

of higher-order assembly, a tetrameric model was subsequently generated. This model revealed a distinct β-sheet-mediated interface absent in the dimer, resulting in a dimer-of-dimers architecture (Figure 1). Mutations targeting this interface substantially impaired oligomerization. Collectively, these results illustrate that AlphaFold-based approaches can provide mechanistic insights into oligomerization processes that are inaccessible to conventional secondary-structure prediction methods. This approach is used not only for DnaT but also for other proteins, the KCTD Family, and the Zika virus NS4A protein, to predict oligomerization sites [28,29]. Recently, it has also been utilized as part of proteomics [30]. Nevertheless, the limited capacity of current models to capture conformational heterogeneity and transient assemblies necessitates careful experimental corroboration.

3. Secondary-structure prediction of polytopic membrane proteins

Membrane proteins constitute a structurally heterogeneous class of biomolecules whose conformations are strongly influenced by the lipid bilayer environment. Monotopic membrane proteins are excluded from the present discussion. Bitopic membrane proteins traverse the bilayer once, with transmembrane segments stabilized primarily by hydrophobic interactions with lipid acyl chains, a context that strongly favors α-helical backbone hydrogen bonding. The hydrophobic core of biological membranes is approximately 20 Å thick, whereas α-helices extend by 1.5 Å per residue. Consequently, contiguous stretches of 20 hydrophobic residues have been widely used as indicators of transmembrane helices. This principle forms the basis of the hydropathy plot method introduced by Kyte and Doolittle, which remains a cornerstone of transmembrane segment prediction [31]. Subsequent refinements incorporating amphipathic character and hidden Markov models have further enhanced predictive accuracy [32,33]. However, such approaches implicitly assume that transmembrane segments are lipid-exposed, an assumption that is frequently violated in polytopic membrane proteins. In these systems, transmembrane helices often engage in extensive helix–helix interactions, and stabilization may involve polar contacts or noncanonical secondary-structure elements.

The human erythrocyte anion exchanger Band 3 exemplifies these limitations. Early hydropathy-based analyses predicted 14 transmembrane helices, as shown in Figure 2A [34], a model that conflicted with biochemical and functional data [35−37]. The X-ray crystal structure of Band 3, determined in 2015 [38], revealed substantial deviations from these predictions, particularly in transmembrane helices 3 and 10. These transmembrane helices are embedded in the lipid bilayer in the 14 transmembrane helices model, but are largely shielded from the lipid environment in the crystal structure (Figures 2A and 2B). They participate in interhelical interactions and incorporate β-sheet elements that expand the repertoire of permissible membrane-embedded structures. Comparing the secondary structures of the crystal structure and the prediction, as shown in Figure 2C, for transmembrane helices 2 and 9, which interact with lipids, the positions of the transmembrane helices predicted by hydropathic analysis and those observed in the crystal structure are largely consistent. However, for transmembrane helices 3 and 10, the positions differ between the prediction and the crystal structure. Band 3 is organized into a core domain and a gate domain, whose relative motions enable anion translocation [39,40]. Transmembrane segments 3 and 10 reside at the domain interface and, together with conserved anion-binding residues, constitute a structural motif critical for conformational switching. Comparable noncanonical membrane-embedded architectures are observed in other polytopic membrane proteins, including the NPA motif of aquaporins [41] and the selectivity filter of potassium channels [42], both of which form atypical conduits rather than lipid-facing helices.

These examples indicate that deviations from hydropathy-based secondary-structure predictions often reflect functionally essential architectural features rather than predictive failures per se.

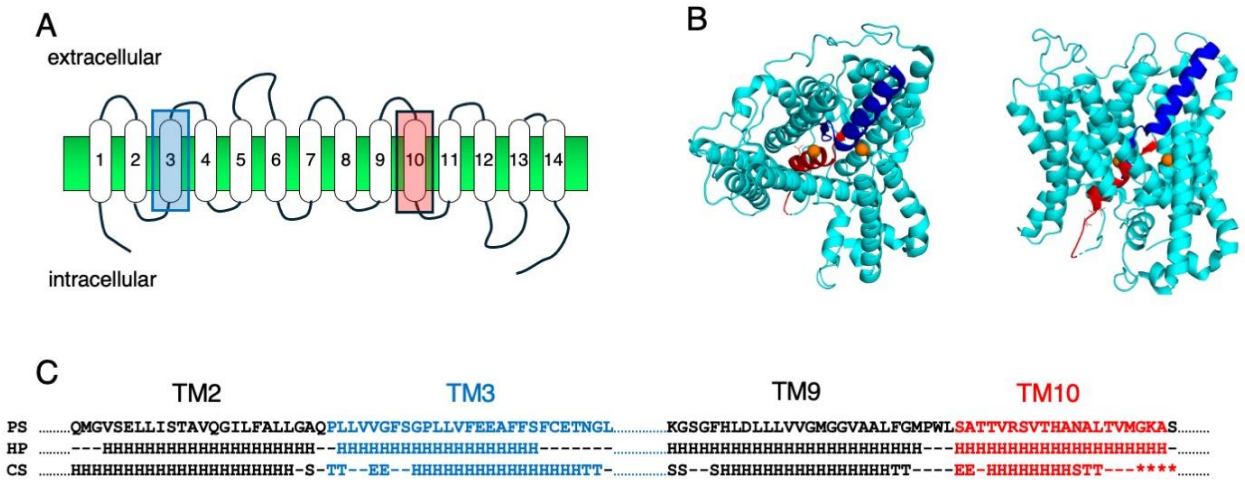

Figure 2. Structure of the transmembrane region of the human erythrocyte Band 3 protein. TM3 and TM10 are colored blue and red, respectively. A. Hydropathic prediction model. B. Views of the crystal structure from the extracellular side (left) and the membrane side (right). Anion binding sites are indicated by orange spheres. C. Comparison of secondary structures in transmembrane segments between hydropathic predictions (HP) based on primary structure (PS) and crystal structure (CS).

4. Concluding remarks

Secondary structure has long been viewed as a predominantly sequence-encoded and locally determined property. However, evidence discussed in this review demonstrates that secondary-structure formation, stability, and functional relevance are profoundly shaped by intra- and intermolecular interactions. The integration of AlphaFold-derived predictions with rigorous experimental validation promises to refine our understanding of interaction-dependent secondary structures that have historically resisted reliable prediction. Such advances are expected to contribute not only to structural biology but also to the rational engineering of proteins with novel architectures and functions.

Use of generative-AI tools declaration

I used AI tools for English editing and proofreading.

Conflict of interest

The author declares no conflicts of interest.

References

1. Pauling L, Corey RB, Branson HR (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* 37: 205–211. https://doi.org/10.1073/pnas.37.4.205
2. Pauling L, Corey RB (1951) The pleated sheet, a new layer configuration of polypeptide chains, Proc. *Natl Acad Sci USA* 37: 251–256. https://doi.org/10.1073/pnas.37.5.251
3. Kendrew JC, Bodo G, Dintzis HM (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181: 662–666. https://doi.org/10.1038/181662a0
4. Blake CC, Koenig DF, Mair GA (1965) Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature* 206: 757–761. https://doi.org/10.1038/206757a0
5. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7: 95−99. https://doi.org/10.1016/S0022-2836(63)80023-6
6. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22: 2577–2637. https://doi.org/10.1002/bip.360221211.
7. Wishart DS, Sykes BD, Richards FM (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* 222: 311–333. https://doi.org/10.1016/0022-2836(91)90214-q
8. Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* 278: 1111−1114. https://doi.org/10.1126/science.278.5340.1111
9. Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Bio NMR* 13: 289–302. https://doi.org/10.1023/a:1008392405740
10. Abe M, Abe Y, Ohkuri T (2013) Mechanism for retardation of amyloid fibril formation by sugars in Vλ6 protein. *Protein Sci* 22: 467–474. https://doi.org/10.1002/pro.2228
11. Chou PY, Fasman GD (1974) Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. *Biochemistry* 13: 211–222. https://doi.org/10.1021/bi00699a001
12. Chou PY, Fasman, GD (1974) Prediction of protein conformation. *Biochemistry* 13: 222–245. https://doi.org/10.1021/bi00699a002
13. Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* 120: 97–120. https://doi.org/10.1016/0022-2836(78)90297-8
14. Cuff JA, Clamp ME, Siddiqui AS (1998) Jpred: A consensus secondary structure prediction server. *Bioinformatics* 14: 892–893. https://doi.org/10.1093/bioinformatics/14.10.892
15. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292: 195–202. https://doi.org/10.1006/jmbi.1999.3091
16. Mishima T, Ohkuri T, Monji A (2009) Residual structures in the acid-unfolded states of Vlambda6 proteins affect amyloid fibrillation. *J Mol Biol* 392: 1033–1043. https://doi.org/10.1016/j.jmb.2009.07.078

17. Minor Jr DL, Kim PS (1996) Context-dependent secondary structure formation of a designed protein sequence. *Nature* 380: 730–734. https://doi.org/10.1038/380730a0
18. Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* 293: 321–331. https://doi.org/10.1006/jmbi.1999.3110
19. Waterhouse A, Bertoni M, Bienert S (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46: W296–W303. https://doi.org/10.1093/nar/gky427
20. Jumper J, Evans R, Pritzel A (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589. https://doi.org/10.1038/s41586-021-03819-2
21. Evans R, O'Neill M, Pritzel A (2021) Protein complex prediction with AlphaFold-Multimer. bioRxiv
22. Inoue S, Ikeda Y, Fujiyama S (2023) Oligomeric state of the N-terminal domain of DnaT for replication restart in *Escherichia coli*. *Biochim Biophys Acta Proteins Proteom* 1871: 140929. https://doi.org/10.1016/j.bbapap.2023.140929
23. Fujiyama S, Abe Y, Tani J (2014) Structure and mechanism of the primosome protein DnaT-functional structures for homotrimerization, dissociation of ssDNA from the PriB-ssDNA complex, and formation of the DnaT-ssDNA complex. *FEBS J* 281: 5356–5370. https://doi.org/10.1111/febs.13080
24. Abe Y, Ikeda Y, Fujiyama S (2020) A structural model of the PriB-DnaT complex in *Escherichia coli* replication restart. *FEBS Lett* 595: 341–350. https://doi.org/10.1002/1873-3468.14020
25. Liu Z, Chen P, Wang X (2014) Crystal structure of DnaT84–153-dT10 ssDNA complex reveals a novel single-stranded DNA binding mode. *Nucleic Acids Res* 42: 9470–9483. https://doi.org/10.1093/nar/gku633
26. Huang YH, Huang CY (2013) The N-terminal domain of DnaT, a primosomal DNA replication protein, is crucial for PriB binding and self-trimerization. *Biochem Biophys Res Commun* 442: 147–152. https://doi.org/10.1016/j.bbrc.2013.11.069
27. Szymanski MR, Jezewska MJ, Bujalowski W (2013) The *Escherichia coli* primosomal DnaT protein exists in solution as a monomer-trimer equilibrium system. *Biochemistry* 52: 1845–1857. https://doi.org/10.1021/bi301568w
28. Esposito L, Balasco N, Vitagliano L (2022) Alphafold predictions provide insights into the structural features of the functional oligomers of all members of the KCTD family. *Int J Mol Sci* 23: 13346. https://doi.org/10.3390/ijms232113346
29. Surya W, Liu Y, Torres J (2023) The cytoplasmic N-terminal tail of Zika virus NS4A protein forms oligomers in the absence of detergent or lipids. *Sci Rep* 13: 137360. https://doi.org/10.1038/s41598-023-34621-x
30. Schweke H, Pacesa M, Levin T (2024) An atlas of protein homo-oligomerization across domains of life. *Cell* 187: 999−1010. https://doi.org/10.1016/j.cell.2024.01.022
31. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105–132. https://doi.org/10.1016/0022-2836(82)90515-0
32. Hirokawa T, Boon-Chieng S, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14: 378–379. https://doi.org/10.1093/bioinformatics/14.4.378

33. Krogh A, Larsson B, von Heijne G (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567–580. https://doi.org/10.1006/jmbi.2000.4315
34. Wood PG (1992) The anion exchange proteins: homology and secondary structure. *Prog Cell Res* 2: 325–352. https://doi.org/10.1016/B978-0-444-89547-9.50037-0
35. Hamasaki N, Okubo K, Kuma H (1997) Proteolytic cleavage sites of band 3 protein in alkali-treated membranes: fidelity of hydropathy prediction for band 3 protein. *J Biochem* 122: 577–585. https://doi.org/10.1093/oxfordjournals.jbchem.a021792
36. Kuma H, Abe Y, Askin D (1997) Molecular basis and functional consequences of the dominant effects of the mutant band 3 on the structure of normal band 3 in Southeast Asian ovalocytosis. *Biochemistry* 41: 3311–3320. https://doi.org/10.1021/bi011678+
37. Hamasaki N, Abe Y, Tanner MJ (2002) Flexible regions within the membrane-embedded portions of polytopic membrane proteins. *Biochemistry* 41: 3852–3854. https://doi.org/10.1021/bi011918l
38. Arakawa T, Kobayashi-Yurugi T, Alguel Y (2015) Crystal structure of the anion exchanger domain of human erythrocyte band 3. *Science* 350: 680–684. https://doi.org/10.1126/science.aaa4335
39. Zhekova HR, Jiang J, Wang W (2022) Cryo EM structures of anion exchanger 1 capture multiple states of inward- and outward-facing conformations. *Commun Biol* 5: 1372. https://doi.org/10.1038/s42003-022-04306-8
40. Su CC, Zhang Z, Lyu M (2024) Cryo-EM structures of the human band 3 transporter indicate a transport mechanism involving the coupled movement of chloride and bicarbonate ions. *Plos Biol* 22: e3002719. https://doi.org/10.1371/journal.pbio.3002719
41. Sui H, Han BG, Lee JK Structural basis of water-specific transport through the AQP1 water channel. *Nature* 414: 872–878. https://doi.org/10.1038/414872a
42. Doyle DA, Cabral JM, Pfuetzner RA (1998) The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science* 280: 69–77. https://doi.org/10.1126/science.280.5360.69

AIMS Press