

# Wenxiang 3.0: Evolutionary Visualization of $\alpha$ , $\pi$ , and 3/10 Helices

John R Jungck<sup>1</sup> and Metehan Cebeci

Delaware Biotechnology Institute, University of Delaware, Newark, DE, USA.

Evolutionary Bioinformatics  
Volume 18: 1–3  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/11769343221101014



**ABSTRACT:** Wenxiang diagrams illustrate protein helices as spirals on a plane and thus have the advantage over helical wheels of being planar graphs. Wenxiang 3.0 extends the original version by adding 3 major features: (1) individual amino acid residues can be colored according to their evolutionary conservation in comparative multiple sequence alignments using CONSURF encoding; (2)  $\alpha$ ,  $\pi$ , and 3/10 helices can be illustrated by overlaying arcs representative of the pitches of these helices; and, (3) the physico-chemical properties of amino acids residues in the protein sequence can be re-presented by colored geometric shapes.

**KEYWORDS:** Wenxiang diagrams,  $\alpha$ ,  $\pi$ , and 3/10 helices, evolutionary conservation

**RECEIVED:** January 10, 2022. **ACCEPTED:** April 28, 2022.

**TYPE:** Software or Database Review

**FUNDING:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work has been supported by the Delaware INBRE program, supported by a grant from the National Institute of General Medical Sciences—NIGMS (P20 GM103446) from the National Institutes of Health and the State of Delaware and the Summer Undergraduate Research Program of the University of Delaware.

**DECLARATION OF CONFLICTING INTERESTS:** The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**CORRESPONDING AUTHOR:** John R Jungck, Delaware Biotechnology Institute, University of Delaware, 15 Innovation Way, Newark, DE 19716, USA. Email: jungck@udel.edu

## Introduction

Wenxiang diagrams<sup>1</sup> illustrate protein helices as spirals on a plane (Figure 1) and thus have the advantage over helical wheels<sup>2</sup> of being planar graphs. Wenxiang 3.0 extends the original version by adding 3 major features: (1) individual amino acid residues can be colored according to their evolutionary conservation in comparative multiple sequence alignments using CONSURF<sup>3</sup> encoding (Figure 1a and b); (2)  $\alpha$ ,  $\pi$ , and 3/10 helices can be illustrated by overlaying arcs representative of the pitches of these helices ((Figure 1b and c); and, (3) the physicochemical properties of amino acids residues in the protein sequence can be represented by colored geometric shapes (Figure 1c).

Besides helical wheels, 3 additional planar visualization tools of alpha helices in protein secondary structure include helixvis,<sup>4</sup> HELIQUEST,<sup>5</sup> and helical net-diagrams.<sup>6</sup> We examined all 4 approaches to develop 5 different design criteria that we felt would be particularly helpful in featuring different aspects of the structure, function, and evolution of helices in proteins. In order to find relevant  $\alpha$ ,  $\pi$ , and 3/10 helical regions in protein sequences, we used STRIDE.<sup>7</sup>

## Software Features

Our 5 different reasons for developing Wenxiang 3.0 are: First,  $\pi$  and 3/10 helices are not currently represented in any helix planar visualization tool of protein secondary structure that we could find. Kumar and Bansal<sup>8</sup> review the importance of  $\pi$  helices and Cooley et al<sup>9</sup> and Khrustalev et al<sup>10</sup> review the importance of 3/10 helices. Ren et al<sup>11</sup> stress that large winding and unwinding of helices occur in membrane proteins such as rhodopsin and these result in transformations among these 3 configurations. Thus, a tool for analyzing potential helical interactions of a single amino acid sequence can provide helpful insights. As educators, we have extensively used origami

models of  $\alpha$ ,  $\pi$ , and 3/10 helices<sup>12</sup> and 3D printed models of helices (Herman et al<sup>13</sup> and Davenport et al<sup>14</sup>). Yet our students still struggled with understanding the differences between these 3 types of helices and the importance of using evolutionary conservation to comprehend which residues or residue types were important to sustaining structure and function of proteins.

Second, in order to understand the different pitches of  $\alpha$ ,  $\pi$ , and 3/10 helices we wanted to overlay arcs representative of whether hydrogen bonding was occurring between either every third, fourth, or fifth residue in the pitches of these helices. These are in close correspondence to the various pitches which are usually reported to be pitches of 3 residues per turn for 3/10 helices<sup>20</sup>, 3.6 for  $\alpha$  helices, and 4.4 for  $\pi$  helices. While there is not a one-to-one correspondence between our integer values with measured values fractional values around these 3 integers, we believe our visualization of arcs representing different helical pitches has heuristic value in inferring vertical interactions between different amino acid residues involved in helical structures that contribute to their stability. We are particularly interested in identifying weak non-covalent potential interactions including hydrophobic interactions, salt bridges, van der Waals forces, etc. that have been evolutionarily conserved.

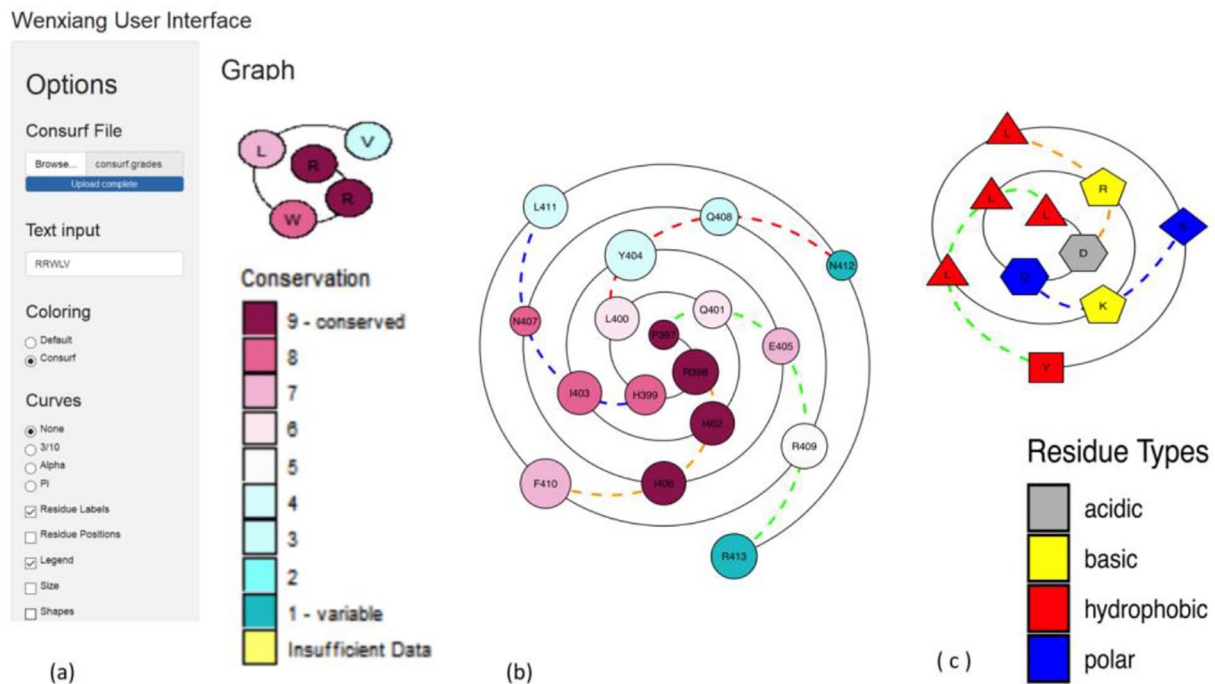
Third, residue size (small, medium, large, very large) was based upon the Voronoi polyhedral volumes (in cubic Angstroms) of amino acid residues which were divided into these bins based upon breaks in properties.

Fourth, chemical properties residues (acidic, basic, polar, hydrophobic) could be represented by different geometries and colors.

Fifth, and finally, we adopted a numerical identification of successive residues in the helical region of the amino acid sequence.



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).



**Figure 1.** (a) Wenxiang 3.0 interface wherein the sequence of a 3/10 helix from 1GPB (glycogen phosphorylase B from rabbit (*Oryctolagus cuniculus*)) and its notation of evolutionary conservation using the key from CONSURF applied to a correlated multiple sequence alignment. (b) Arcs representing hydrogen bonding interactions for an alpha helix in 1GPB (glycogen phosphorylase B) with sequential residue numbers on each amino acid residue in the peptide and again with evolutionary conservation coloring. (c) Arcs representing hydrogen bonding interactions for a Pi helix in 1GPB (glycogen phosphorylase B) with coloring of physicochemical properties (acidic, basic, hydrophobic, polar) and geometry of size (triangular=small, rhombic=medium, pentagonal=large, and hexagonal=very large).

### Use of the software

For generating visualizations in Wenxiang 3.0 like Figure 1a and b, we recommend that users identify proteins with published 3D structures determined by X-ray crystallography and that they have generated multiple sequence alignments of sequences similar to that of the published 3D structure. If a user is only interested in visualizing the physicochemical interactions and the arcs of  $\alpha$ ,  $\pi$ , and 3/10 helices of an individual protein such as in Figure 1c, the user can simply enter an amino acid sequence (our current limit is 18 residues long). Specific choices of whether to implement: curves for which of the 3 helices, coloring of degrees of evolutionary conservation, and chemical properties of different amino acid residues are simple check boxes on the user interface (Figure 1a).

### Future directions

The decision of which features were most important to include was difficult. The number of physicochemical categories of amino acids varies enormously from 2 (H: hydrophobic/hydrophilic) to the 4 shown herein to Dayhoff et al's<sup>15</sup> 7 and Knight et al's<sup>16</sup> 8. Since there are so many different multiple sequence alignment software packages, we will be implementing some data wrangling routines for importing from ones identified by our users. In terms of 3D representations of evolutionary

conservation, 3DPatch<sup>17</sup> has the advantage that it “does not require the user to have prior knowledge of any 3D structures matching the query sequence.” However, because it does not have as many features as CONSURF and depends upon entropy information measures rather than section, we decided to use CONSURF in our current version of wenxiang 3.0. While we have used STRIDE to find examples of  $\alpha$ ,  $\pi$ , and 3/10 helices, it would be helpful to have a general-purpose tool to find various helical forms more generally. Le and Nguyen<sup>18</sup> and Tng et al<sup>19</sup> have used deep learning techniques to explore proteins that span membranes in mitochondria such as members of the electron transport chain. Their use of measures of specificity, sensitivity, and accuracy as well as visualization with sequence logos may provide useful extensions for identifying  $\alpha$ ,  $\pi$ , and 3/10 helices.

### Acknowledgements

The authors are grateful to Professor Jacqueline Fajardo for introducing them to Wenxiang diagrams as well Pi and 3/10 helices.

### Authors' Contributions

Conception and design: John R. Jungck. Programming: Metehan Cebici. Manuscript writing: John R. Jungck. Final approval of manuscript: Both authors.

**ORCID iD**John R Jungck  <https://orcid.org/0000-0003-4621-5333>**Data Availability**

Wenxiang 3.0 is encoded in R and is available as an open source software package. <https://qubeshub.org/publications/2893/1>

**REFERENCES**

1. Chou K-C, Lin W-Z, Xiao X. Wenxiang: a web-server for drawing wenxiang diagrams. *Nat Sci*. 2011;03:862-865.
2. Schiffer M, Edmundson AB. Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophys J*. 1967;7:121-135.
3. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*. 2016;44:W344-W350.
4. Wadhwa R, Subramanian V, Stevens-Truss R. Visualizing alpha-helical peptides in R with helixvis. *J Open Source Softw*. 2018;3:1008.
5. Gautier R, Douguet D, Antonny B, Drin G. HELIQUEST: A web server to screen sequences with specific alpha-helical properties. *Bioinformatics*. 2008;24:2101-2102.
6. Dunnill P. The use of helical net-diagrams to represent protein structures. *Biophys J*. 1968;8:865-875.
7. Heinig M, Frishman D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res*. 2004;32:W500-W502.
8. Kumar P, Bansal M. Dissecting  $\pi$ -helices: sequence, structure and function. *FEBS J*. 2015;282:4415-4432.
9. Cooley RB, Arp DJ, Karplus PA. Evolutionary origin of a secondary structure:  $\pi$ -helices as cryptic but widespread insertional variations of  $\alpha$ -helices that enhance protein functionality. *J Mol Biol*. 2010;404:232-246.
10. Khrustalev VV, Barkovsky EV, Khrustaleva TA. The influence of flanking secondary structures on amino acid content and typical lengths of 3/10 helices. *Int J Proteom*. 2014;2014:360230.
11. Ren Z, Ren PX, Balusu R, Yang X. Transmembrane helices tilt, bend, slide, torque, and unwind between functional states of rhodopsin. *Sci Rep*. 2016;6:34129-34213.
12. Pikaart M. The turn of the screw: an exercise in protein secondary structure. *Biochem Mol Biol Educ*. 2011;39:221-225.
13. Herman T, Morris J, Colton S, et al. Tactile teaching: exploring protein structure/function using physical models. *Biochem Mol Biol Educ*. 2006;34:247-254.
14. Davenport J, Pique M, Getzoff E, Huntoon J, Gardner A, Olson A. A self-assisting protein folding model for teaching structural molecular biology. *Structure*. 2017;25:671-678.
15. Dayhoff MO, Eck RV, Park CM. A model of evolutionary change in proteins. In: MO Dayhoff, ed. *Atlas of Protein Sequence and Structure*. Vol. 5. National Association for Biomedical Research; 1972.
16. Knight RD, Freeland SJ, Landweber LF. Selection, history and chemistry: the three faces of the genetic code. *Trends Biochem Sci*. 1999;24:241-247.
17. Jakubec D, Vondrášek J, Finn RD. 3DPatch: fast 3D structure visualization with residue conservation. *Bioinformatics*. 2019;35:332-334.
18. Le NQ, Nguyen BP. Prediction of FMN binding sites in electron transport chains based on 2-D CNN and PSSM Profiles. *IEEE/ACM Trans Comput Biol Bioinform*. 2021;18:2189-2197.
19. Tng SS, Le NQK, Yeh HY, Chua MCH. Improved prediction model of protein lysine crotonylation sites using bidirectional recurrent neural networks. *J Proteome Res*. 2022;21:265-273.
20. Enkhbayar P, Hikichi K, Osaki M, Kretsinger RH, Matsushima N. 3/10-helices in proteins are parahelices. *Proteins Struct Funct Bioinf*. 2006;64:691-699.