Comparison and Visualization of Coiled-Coil Predictions

Dominic Simm, Klas Hatje and Martin Kollmar

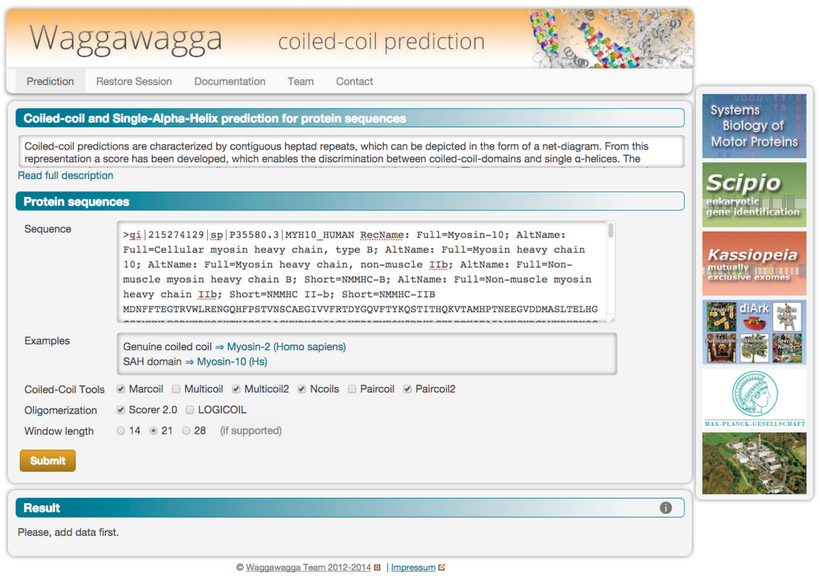
Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany

Group Systems Biology of Motor Proteins

MAX-PLANCK-GESELLSCHAFT

Introduction

Coiled-coils belong to the most common structural motives for proteins. The sequences of coiled-coils are typically characterized by contiguous heptad repeats $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$ with hydrophobic residues in '*a*' and '*d*' positions and the remaining residues mainly charged thus favouring alpha-helical formation. In most cases, the individual alpha-helices are not very stable but become stabilized by wrapping around each other. Thus left-handed coiled-coils are formed, in which the hydrophobic residues are buried in the centre of the molecule. Coiled-coils can either be parallel, anti-parallel, homodimers, heterodimers, trimers, or any other oligomer. Several programs exist to predict coiled-coil regions like Marcoil [1], Multicoil [2] or Paircoil2 [3]. However, all these prediction programs are biased towards highly charged sequences. Charged residues are even found, although rarely, in '*a*' and '*d*' positions. Although predicted as homodimeric coiled-coils, many of these sequences indeed form stable single-alpha-helices (SAH) instead.



Conclusion

Here, we present a new web service for comparing and visualizing coiled-coil predictions, which we called WaggaWagga. The user can run among others Marcoil [1], Multicoil [2] and Paircoil2 [3] on the query sequence, and use Scorer 2.0 [4] and LOGICOIL [5] for oligomerisation prediction. The query sequence is visualized as helical wheel-diagram of parallel or anti-parallel homodimers, or parallel homotrimers, and as net diagram. With sliders the user can interactively move through the sequence automatically updating the visualizations.

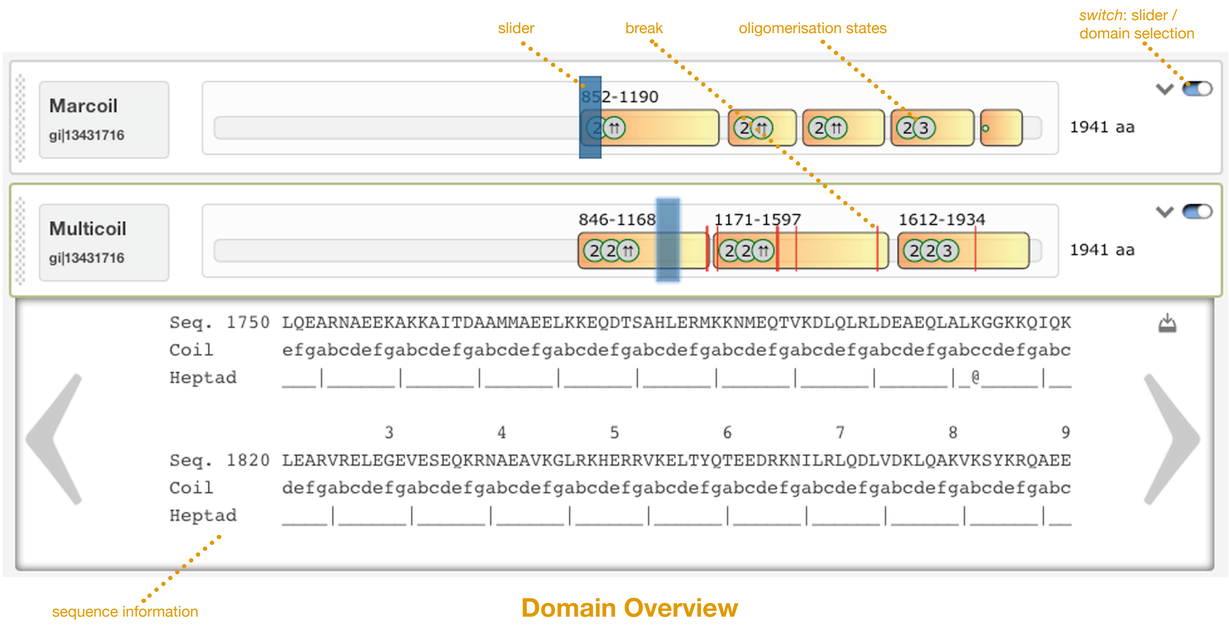
WaggaWagga is available at
http://waggawagga.motorprotein.de
or scan the QR code.

Visualizing Coiled-Coil Predictions

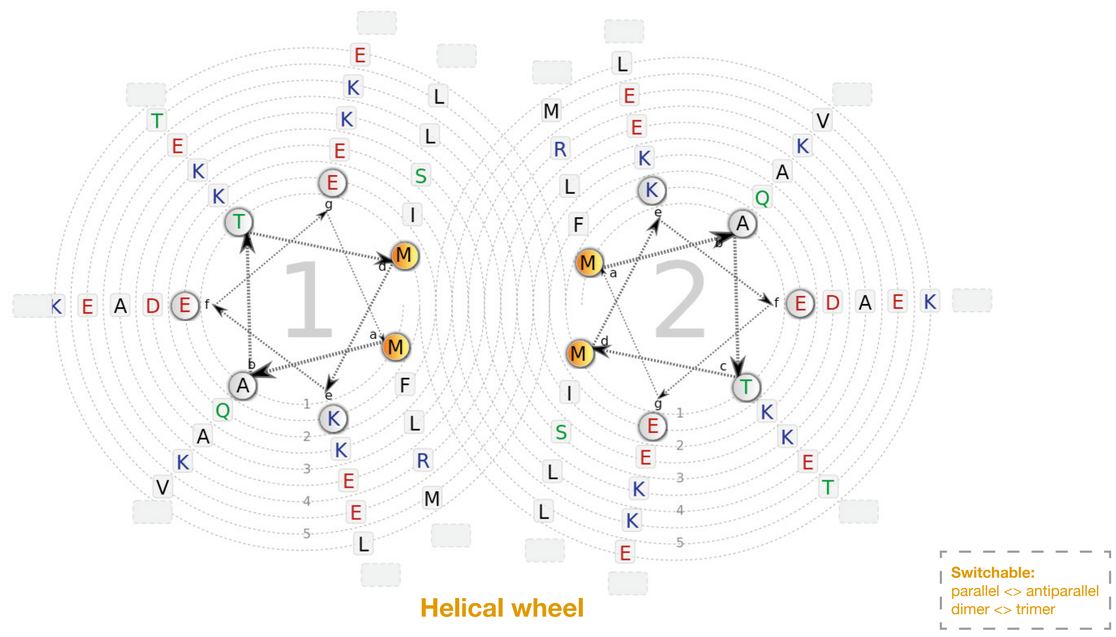
WaggaWagga presents the coiled-coil predictions in four different result sections. First, an overview diagram of the predicted coiled-coil regions for each chosen prediction tool is displayed to the user. This part represents simultaneously a navigation along the query sequence through a free movable slider. Changing the position of the slider causes instantaniously an update of all subsequent visualizations. By deactivating the slider, the predicted regions can also be accessed directly via clicking the highlighted domains. Each region diagram gives further information about the oligomerisation states and comprises a detailed view with sequence information like positions, the heptad repeats and breaks.

The second section, the helical wheel diagram, illustrates the possible interacting residues of two helix-sequences when forming a coiled-coil in the depicted manner. The turns of the potential coiled-coil go into the plane, each dashed circle symbolizes one step. Important are here the positions '*a*' and '*d*', which have have to be hydrophobic to build a coiled-coil.

The third section, the helical net diagram, is a further visualization, that helps to determine on the basis of the predictions, whether the regarded region is a coiled-coil or only a single-alpha-helix domain.



Domain Overview



Helical wheel

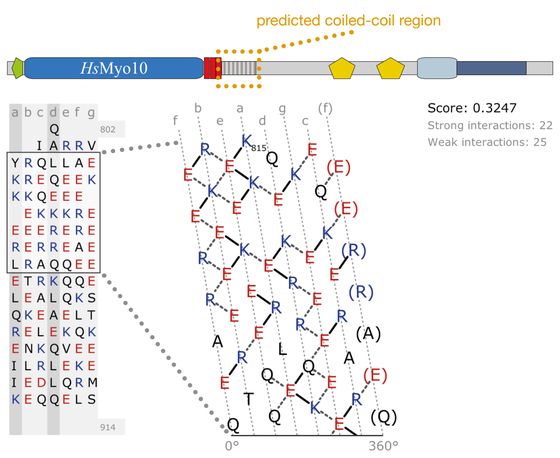
Discriminating SAHs and Coiled-Coils

SAH domains are formed, when a high number of both positive and negative charged residues are present in a protein sequence and the arrangement of these residues follows a specific pattern. In this case, intrahelical interactions are formed, which stabilize a SAH domain.

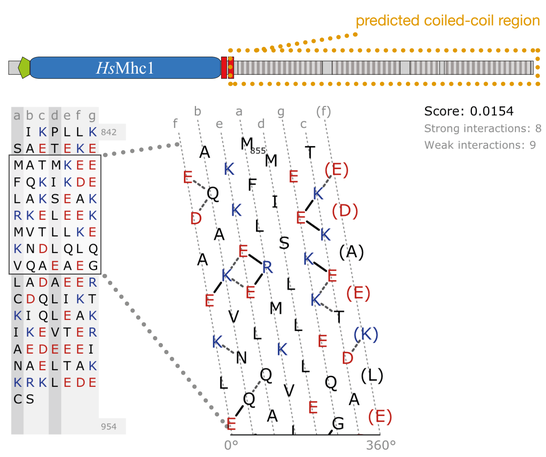
Waggawagga has the helical net visualization available with an integrated discriminating score, that enables the user to determine, based on the predictions, between stable SAHs and Coiled-Coil regions.

The scoring algorithm uses for its prediction between these two classes the characteristic difference in the interaction structure in the net representation.

The web service provides two examples, one is a non-muscle myosin heavy chain, which builds a SAH domain and the other a myosin heavy chain from a skeletal muscle, which forms a genuine coiled-coil.



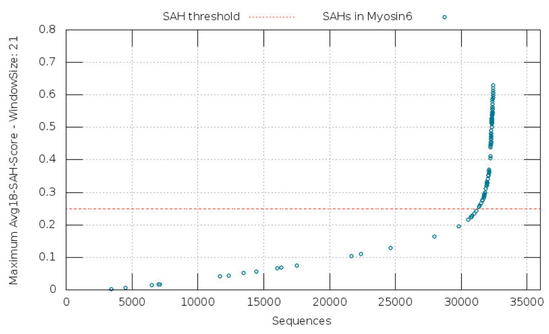
Helical net - SAH

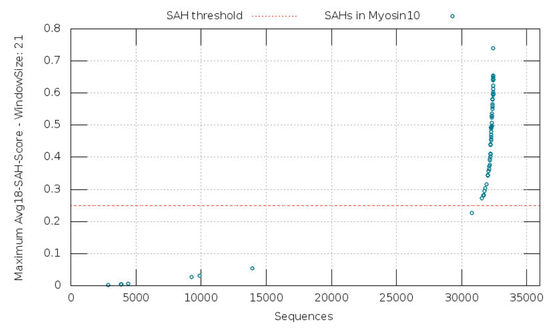


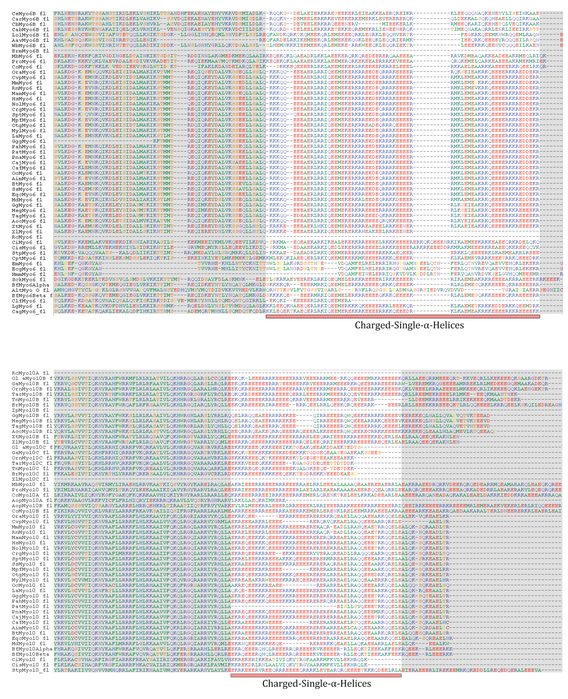
Helical net - Coiled-Coil

SAH-Score Evaluation

For an evaluation of the SAH-score, an analysis of a protein sequence dataset, composed of 45 Myosin classes, has been made. Therefore all of the sequences were scored by the SAH-scoring algorithm and the results were manually compared for a verification to the Multiple Sequence Alignments of the single Myosin classes.







References

[1] M Delorenzi, and T Speed. (2002) An HMM model for coiled–coil domains and a comparison with PSSM-based predictions. *Bioinformatics*, **18**: 617-625.

[2] P S Kim, B Berger, and E Wolf. (1997) MultiCoil: A program for predicting two-and three-stranded coiled coils. *Protein Science* **6**: 1179–1189.

[3] A V McDonnell, T Jiang, and B Berger. (2006) Paircoil2: Improved prediction of coiled coils from sequence. *Bioinformatics*, **22**: 356–358.

[4] C T Armstrong, T L Vincent, P J Green, and D N Woolfson. (2011) SCORER 2.0: An algorithm for distinguishing parallel dimeric and trimeric coiled-coil sequences. *Bioinformatics*, **27**: 1908–1914.

[5] T L Vincent, P J Green, and D N Woolfson. (2013) LOGICOIL—Multi-state prediction of coiled-coil oligomeric state. *Bioinformatics*, **29**: 69–76.

[6] M Peckham and P J Knight. (2009) When a predicted coiled coil is really a single α-helix, in myosins and other proteins. *Soft Matter*, **5**: 2493–2503.

Implementation

- WaggaWagga is implemented as a Ruby on Rails 3.1 web application using Ruby 2.0. Data is stored for fast delivery in the Key-Value-Store Redis.
- BioRuby handles the FASTA file format and garantuees a structured data treatment.
- The prediction is realized through the provision of Marcoil [1], Multicoil [2], Multicoil2, Ncoils, Paircoil and Paircoil2 [3] as well as PrOCoil, Scorer 2.0 [4] and LOGICOIL [5].
- Visualizations are generated either in pure Ruby or with the help of gnuplot and Inkscape.