

Using bioinformatics to investigate a novel Cyclin A/Cdk 2-interacting protein

David Khoo and Brian D. Dynlacht, PhD

Department of Pathology, New York University School of Medicine

Introduction

The Cyclin A/Cdk2 complex is critical for progression through the S-phase as well as entry into the M-phase of the eukaryotic cell cycle. However, much remains to be understood of the mechanism by which this complex carries out its duties, as identification of its substrates and regulators has only just begun. Towards this end, we identified CAIP, a novel Cyclin A/ Cdk2-interacting protein by affinity purification and peptide sequencing. Although the cDNA of this 160kDA protein was already known, its function had not been determined. Here we employ various bioinformatics tools to facilitate our investigation of this protein.

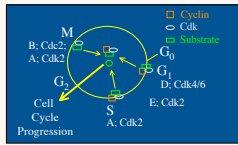


Fig. 1: Progress through the cell cycle is facilitated by the various cyclin/cdk complexes; Cyclin A/cdk2 is important in the S- and M-phases, but little is known about its substrates and regulators

Materials and methods

•All protein and DNA analysis was performed using the *Homo Sapiens* ZNF291 gene sequence (zinc-finger protein 291 - accession #AF242528), which is identical to CAIP.

•BLASTP search for cross-species homology: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein> and follow BLINK for ZNF291.

•Protein motif/domain prediction:

- GCG suite via SeqWeb
- ProDOM

•Protein multiple alignments of cross-species homologs:

- ProDOM
- UCSC Genome Browser

•Comparison of hydropathic profiles between two select homologous proteins was restricted to regions of homology. Gaps were removed from the BLAST alignment prior to application of the Kyte-Doolittle algorithm:

http://bioinformatics.weizmann.ac.il/hydroph/cmp_hydp.html

Results

BLAST alignment

CAIP cDNAs and truncations thereof were found in a wide range of genomes, including those of the human, mouse, rat, chicken, dog and cow, indicating a high degree of sequence conservation among higher eukaryotes, despite the lack of an assigned gene function in any of these organisms. No equivalent was found in lower eukaryotes such as *S.cerevisiae* or *C.Elegans*.

Motif/Domain prediction

Use of various motif and domain-prediction tools on the CAIP protein yielded multiple results (see Figure 2). The prediction of a transmembrane domain by the Nbest algorithm in the TransMem GCG tool was substantiated by the mostly hydrophobic surface of the helical wheel of the putative transmembrane sequence (Fig. 3).

The KKXX ER retention motif, coiled-coil domains, zinc finger motif, Cyclin A RXL binding motif and transmembrane domain are all found in mouse, rat, human and chicken cDNA, although the length of the respective regions vary from species to species. Notably, a ProDom search revealed local homology of CAIP with the N- and C-termini of a drosophila protein of unknown function (Uniprot accession # Q9VJ35).

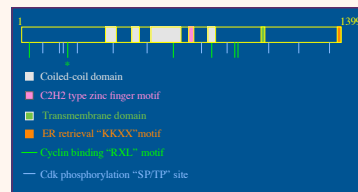
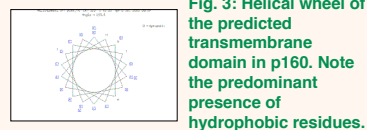


Fig. 2: Summary of predicted motifs and domains in CAIP. Although multiple Cyclin-binding motifs are present, only one (*) is essential for interaction with Cyclin A.



Distantly-related sequences

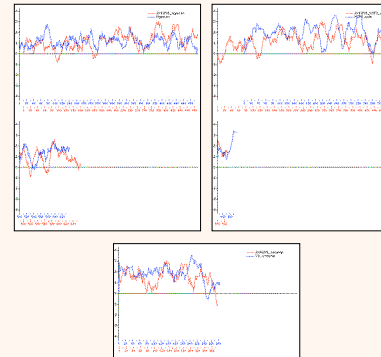


Fig. 4: Use of hydropathy plots to evaluate homology of proteins that share weak similarity with CAIP. Clockwise from left: Human Myosin heavy chain, non-muscle; Human Translation initiation factor IF2; Frog Inner Centromere protein.

Several sequences of low-scoring similarity with E values of less than 0.05 were detected in the Blast alignment (via BLINK), raising the possibility of a distant but non-random evolutionary relationship in the form of a paralog or an ortholog. Hydropathy plots (Fig 4) provide a rough approximate of the secondary structure (hydrophobicity and hydrophilicity) of each protein. By comparing hydropathy plots, evolutionarily related proteins that have diverged in amino acid sequence but have retained a common secondary structure may be detected, allowing a function to be inferred for CAIP.

Conclusions

The absence of orthologs or paralogs of high sequence similarity to CAIP underscores its uniqueness as a Cyclin A/ Cdk2 interacting protein. The fact that it is found predominantly in higher eukaryotes suggests a relatively late evolutionary appearance as a functional protein, which may be consistent with greater cell cycle regulatory needs in more developed organisms.

Immunofluorescence experiments have shown that p160 is localized mainly in the endoplasmic reticulum, suggesting that the highly conserved KKXX ER retention motif is indeed functional.

There is also evidence that overexpression of CAIP can influence the localization of Cyclin A, suggesting a regulatory role for CAIP, in which its endomembrane localization allows it to retain Cyclin A outside the nucleus, thus inhibiting its activity.

The presence of coiled coil motifs suggests that CAIP can interact with other proteins in addition to Cyclin A and cdk2, and it would be interesting to characterize these other CAIP. Might the large size of CAIP indicate a role as a scaffolding protein? If so, the significant evidence for a transmembrane domain would suggest that this multimeric complex is membrane anchored. At the same time, the distinct similarities between the hydropathy plots of CAIP and non-muscle myosin and human translation initiation factor 2 offer tantalizing alternatives for the role of this enigmatic protein.

Acknowledgments

The author wishes to thank William Tsang, PhD and Alex Blais, PhD, for invaluable input in the development of this project, as well as Brian Dynlacht, PhD for his continuing mentorship and support.

For further information

Please contact khoo01@med.nyu.edu.