v1: 19 November 2024

Review Article

The CCN Family of Proteins: A Critical Approach to the Multi-Modular Structure of the CCN Domains

Preprinted: 21 October 2024
Peer-approved: 19 November 2024

Qeios, Vol. 6 (2024)
ISSN: 2632-3834

Bernard Perbal[1]

1. International CCN Society, Nice, France

The CCN family of proteins is composed of six members (CCN1-CCN6) sharing a tetra-modular organization and a striking conservation of their primary structure. The CCN acronym was originally assigned in 1993 by P. Bork to three newly discovered factors (originally called CTGF, CYR61, and NOV), which he proposed to constitute a new family of proteins on the basis of their common physical features. Six years later, three other proteins (Wisp1-3), sharing the same tetramodular organization, joined the family (figure 1). The HUGO-recognized acronyms for the CCN proteins were officialized in 2018[1].

The CCN family turned out to contain positive and negative regulators of cell proliferation and differentiation, with pro- and anti-tumorigenic activities. A significant amount of work has been performed to identify the participation of the constitutive modules in these biological features. The aim of this review is to briefly examine the potential roles assigned to the constitutive modules of CCN proteins and propose a critical view of the structural basis for their interactions and functions.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

The general picture

Soon after the discovery of the multi-modular organization of the CCN proteins, questions were raised regarding the possible evolutionary biological advantages resulting from the reunion of five constitutive exons in a single polypeptide[2].

The main question, which is still open, was to determine "*the participation of each module in the function of the full-length protein. Either the activities of each CCN module add up or they confer on the whole protein specific functions that might substitute or add to the function of the individual modules*" Perbal[2]. The case of CCN proteins is not unique.

There are many examples of evolutionary functional domains clustering, leading to a coordinated regulation of the elementary activities encoded by different genes contributing to a final biological unified function. One famous example that we presented in a recent review[3] is provided by the evolution of the pyrimidine biosynthetic pathway from bacteria to eukaryotes.

Briefly, the bacterial *de novo* pyrimidine biosynthetic pathway requires the sequential participation of six independent enzymatic reactions performed by physically separate genes (pyrA to PyrF) that catalyze the biosynthesis of uridine monophosphate from glutamate and carbonic acid. In mammals, the first three steps of the biosynthetic pathway are performed by a single polypeptide multi-enzymatic hexamer of 243 kD (CAD) harboring the three enzymatic activities leading to the formation of DHO[4] (figure 2).

 doi.org/10.32388/K901XY.2

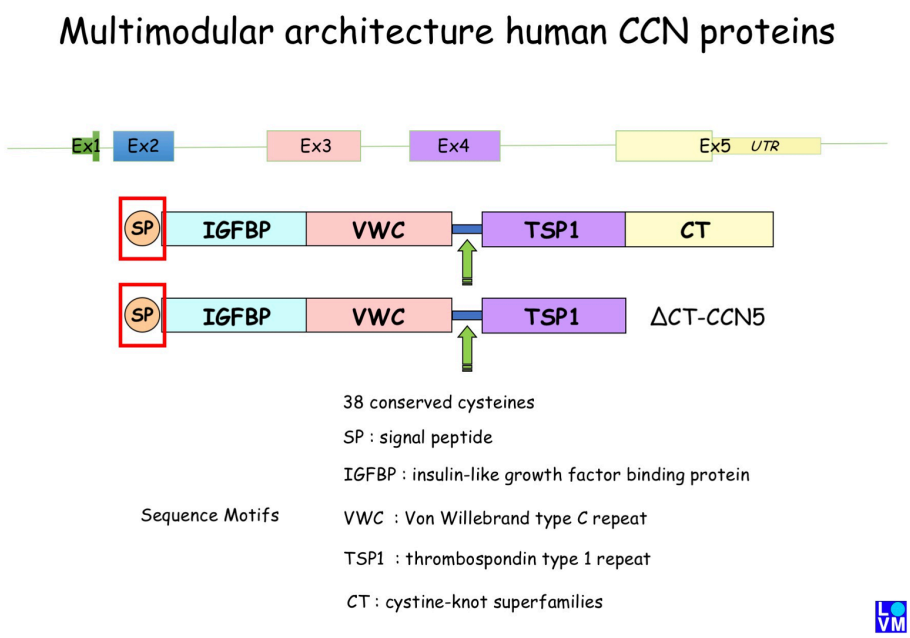
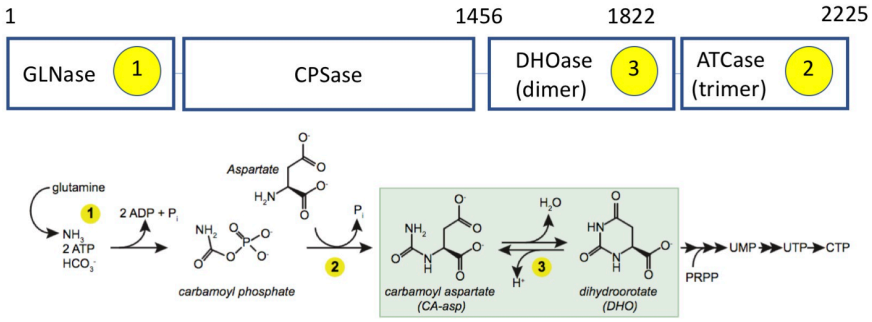


Figure 1.

CAD: Carbamoyl-phosphate synthetase 2, Aspartate transcarbamoylase, and Dihydroorotase



Three first steps of *de novo* biosynthesis of pyrimidines catalyzed by CAD. Glutamine-dependent carbamyl phosphate synthetase (GLN-CPSase)+ aspartate transcarbamylase (ATCase) + dihydroorotase (DHOase) activities

After Grande-García et al. 2014

Figure 2.

The last two steps in the biosynthesis of UDP are under the control of bifunctional proteins containing both the OPRT (Orotate-phosphoribosyltransferase) and OMPDC (orotidine decarboxylase) domains[5][6].

doi.org/10.32388/K901XY.2

Functional analysis of the modules

The phylogenetic reunion of the CCN modules on one single polypeptide might serve the purpose of grouping in a unique transcriptional unit several structural modules having different functions in order to permit a topographical synchronization. A common theme in evolution.

Many reviews have already addressed the biological activities of individual CCN domains. Unfortunately, a few aspects were overlooked. Not trendy, questioned some established "dogma," or were simply ignored because of a lack of referencing.

It is neither the place nor the aim of this review to come back to these questions. We do hope that they will make a comeback when the CCN community will finally consider more thoroughly the points P. Bork expressed over 30 years ago in his seminal paper about the shuffling of exons supplying the ideas of: «*the molecular basis for a complex functional network allowing multiple regulation in and between nearly all tissues. It combines the most different functional complexes like, for example, the antibody framework, inflammation processes or the hematopoietic system...*» and «*coordination of various functions within the world of extracellular proteins*»[7]

In my opinion, a few pending questions regarding the participation of the individual modules' biological features remain to be addressed.

For example, can we draw simple conclusions from the evolutionary conservation of the number of modules and their relative positions in the full-length CCN proteins?

Are we approaching the problems in a timely fashion when we still focus on one particular isolated function at a time of global spatial biology and whole genome sequencing, which are widely used in the studies of several normal and pathological systems?

Does it make sense to consider any particular CCN protein as a unique tool or study object when data that has accumulated over the past decades supports physical and biological positive and negative interactions between different members of the CCN family, within the CCN group, and out of the group with the large families of signaling factors with which they undoubtedly interact?

As previously pointed out[8], "*Altogether, the presence of these four modules in a single protein is not just a curiosity that is worth mentioning. It likely provides the cell with another « multi sensing station » permitting coordinated interactions with key signaling pathways involved in the regulation of cell behavior*", and as Bork also stated in his article « *coordination of various functions within the world of extracellular proteins*»[7].

Are we right to consider only the activities of the isolated modules under conditions that cannot represent their participation in the full-length CCN proteins?

The tale of TSP1 and nuclear CCN addressing

Recently, new TSP1 advances were presented during the 12th Workshop on the CCN family of Genes[9].

Previous studies[10] proposed that CCN proteins are preproteins that need to be processed to become active. This suggestion was in full agreement with previous data from our laboratory, which demonstrated the existence of truncated CCN protein variant forms that were detected in cell culture medium and cellular extracts from CCN-producing cells[2][11], and the demonstration that CCN3 was subjected to proteolytic post-translational modification and that the CT-Terminal module was responsible for the anti-proliferative activity of CCN3[12][13].

This year, the group reported[9] that the TSP1 domain of both CCN3 and CCN5 can counteract the principal profibrotic activity of TGFbeta 1 in fibroblasts. The TGFbeta 1-stimulated fibroblast to myofibroblast differentiation was partially reversed by CCN5-TSP1.

Previous work from the group of Park[14] had demonstrated that the CCN5 TSP1 domain alone is not sufficient for the antifibrotic function of the full-length CCN5. They also showed that, in addition to the TSP1 domain, either the IGFBP or the VWC domain is also required, raising the possibility that new functions may be enhanced or created by the modification of the modules' environment, independent of the required post-translational modification that still needs to be demonstrated in vivo.

These observations suggest that the integration of TSP1 within a physiological context can have its biological properties modulated, provided that the flanking modules are present. This is in favor of

the biological properties of the constitutive CCN modules needing to be carefully evaluated in the *in vivo* context.

Furthermore, in their work, Park et al. established that the TSP1 domain of CCN5 is essential and sufficient to support the nuclear localization of the full-length protein[14].

Initially discovered in tumor cells[15], the nuclear localisation of CCN3 was not fully accepted by the scientific community, even though the nuclear localisation of CCN2 was demonstrated two years later[16].

The use of clones expressing various parts of CCN3 permitted the identification of the CT domain as the one responsible for the translocation of the full-length CCN3 and showed that amino-truncated proteins lacking the signal peptide and one or two N-terminal domains were translocated to the nucleus, where the CCN proteins were shown to colocalize with the HSV1-ICP4 transcription factor and physically interact with the rpb7 subunit of RNA polymerase II. Even though these observations were reinforced by the immunogold localization of CCN3 in the nucleus[17], the nuclear localization of CCN3 was not considered.

The situation might evolve in the right direction thanks to the Park publication. Observations confirmed that the nuclear translocation of CCN5 was critical for its anti-fibrotic activity and that the TSP1 domain of CCN5 was essential for the endocytosis and nuclear translocation of CCN5.

Additional observations by the Park Group that should be taken into account by others in the field demonstrate that TSP1-VWC, but not IGFBP-TSP1, could reverse the transdifferentiation of myofibroblasts to fibroblasts, leading us to the last part of this communication.

The hinge that separates the two duplexes of amino and carboxy-proximal CCN domains has been recognized as potential protease targets. Interestingly, the digestion of CCN2 (CTGF) by plasmin generated two N- and C-terminal paired domains migrating with an apparent MW of 20 kDa, showing different biological activities[18]. The carboxy-proximal bifold part was shown to induce cell proliferation and DNA synthesis in concert with EGF, while the amino-terminal bipartite part stimulated differentiation and collagen synthesis in concert with IGF. These authors suggested that CCN2 "*could be viewed as two cytokines that have been linked into a single gene product*", in agreement with our above proposal that during the evolutionary processes, the four constitutive CCN exons were brought together to provide a "coordinated" level of expression, leading to a balanced production of each CCN component.

Structural features guiding the modular organization of the CCN proteins?

We have seen above that the four distinctive CCN modules are organized in a very conserved way among vertebrates. Whether this arrangement originated from the splicing of an ancient single polycistronic transcriptional unit giving rise to four polypeptide chains is not supported by the chromosomal distinct localisation of the CCN genes. It was therefore proposed that exon shuffling had been responsible for the constitution of the CCN proteins as we see them in vertebrates. Interestingly, the "CCN-like" modules mapped in ancient living species show an arrangement that is very similar to their organization in "modern" CCN proteins, in spite of the reduced number.

Most interestingly, the VWC-TSP1 domains pair is found in all *Branchiostoma floridae*, *Ciona intestinalis*, *Drosophila melanogaster*, and Crassostera Gigas, while the IGFBP domain is found only in *Ciona intestinalis* and the CT domain in *Crassostera gigas*[19] (figure 3).

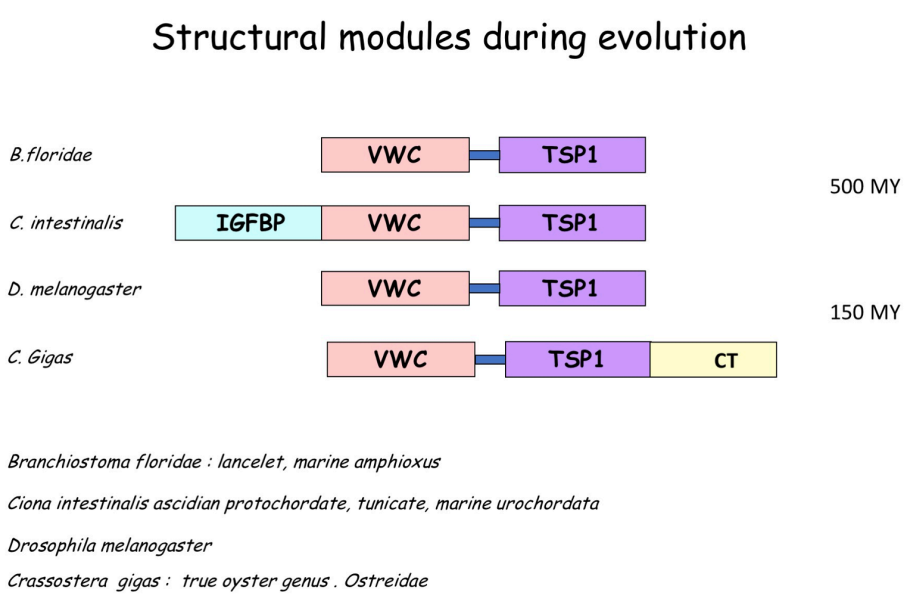


Figure 3.
After Hu et al.[19]

The evolutionary conserved synteny of the VWC-TSP1 pair and the occasional appearance of the IGFBP and CT domains may suggest that specific structural features guide the relative position of the VWC and TSP1, whose functions may be interdependent.

It is remarkable that during evolution, the order of the constitutive domains is conserved. Even if mutational events affecting the organization of these four domains may have been counter-selected for functional reasons[20], it is unlikely that the conserved scheme is the result of fortuity, based on the unique functions that they confer on the full-length proteins.

Also, the variant CCN species, lacking modules, which have been identified in normal and pathological conditions, showed the same arrangement for the conserved modules[21][13].

As I suggested in several CCN workshops, it would be interesting to determine whether experimental domain shuffling would affect or alter the biological properties of CCN proteins. For example, the addition of a CCN-CT domain to CCN5 would recast its functions and modify the range of its interactions with other regulatory factors. The same approach could address the importance of the various cysteine residues in the maintenance of the biological properties of the four modules.

It is well established that CCN proteins are involved in a wide variety of interactions with regulatory factors[22][3] (figure 4). The conserved organisation and order of the CCN domains may be critical to their physical interaction and biochemical partnership with other biological factors. The unique tertiary structure of each domain, which is determined by the presence of conserved disulfide bonds, is of considerable importance to promote physical proximity and interactions. A great number of potential partners have been published.

doi.org/10.32388/K901XY.2

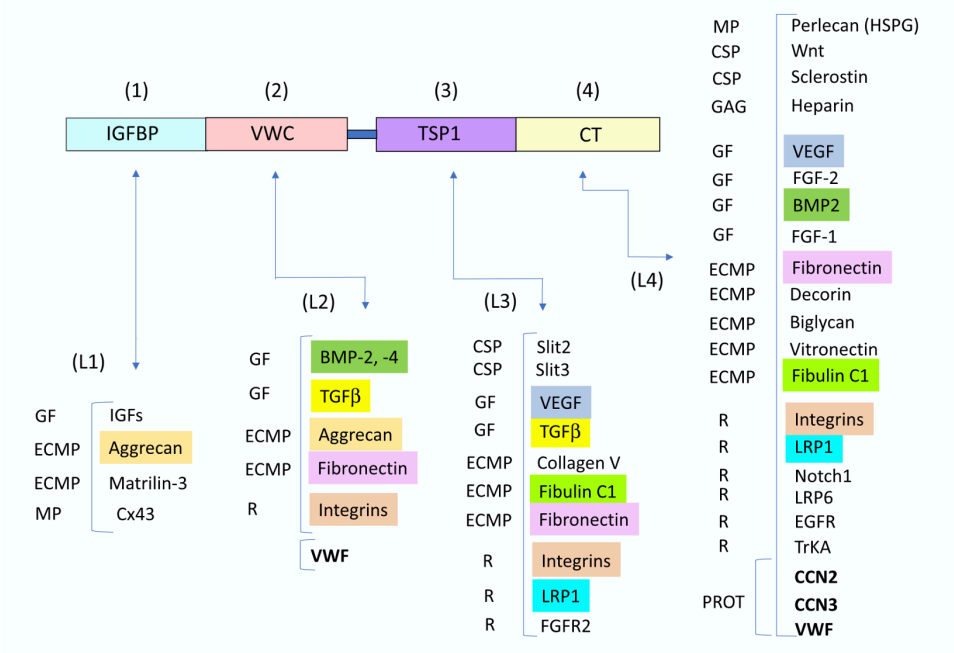


Figure 4.
After Perbal et al.[3]

The spatiotemporal regulatory model that was proposed several years ago[2] was meant to clarify the way to approach this particular question. The model does not only rely on the experimentally established bioavailability of the different partners and their accessibility; it also requires that the interactions between the targets are made possible by the physical occupancy or vacancy of the association active site. However, no study addressed the problem of potential dual occupancy for the same binding regions or sites. Furthermore, most binding studies were performed on isolated domains without taking into account the state of the target in the whole protein. This is a critical and necessary point to evaluate the likelihood of ligands-targets combinations.

In other words, significant binding studies must be performed on full-length native CCN proteins in precisely defined contexts.

The main conclusion emerging from all these considerations is pointing out the inadequateness of studies performed on isolated domains.

Conclusion

As previously advocated, future approaches should be comprehensive and based on spatial biology methodologies, which will be mandatory to reach successful biomedical applications based on a deep understanding of the unique multi-domain biology of the CCN proteins.

Statements and Declarations

Author Contributions

The sole author confirms responsibility for the following: study conception and design, literature analysis and interpretation, and manuscript preparation.

Acknowledgements

Sincere thanks are due to Prof. Herman Yeger for the critical review of the manuscript and to Annick Perbal for support and suggestions.

References

1. ^Perbal B, Tweedie S, Bruford E (2018). "The official unified nomenclature adopted by the HGNC calls f or the use of the acronyms, CCN1-6, and discontinuation in the use of CYR61, CTGF, NOV and WISP 1-3 r espectively." J Cell Commun Signal. **12**(4):625–629.
2. a, b, c, dPerbal B (2001). "NOV (nephroblastoma overexpressed) and the CCN family of genes: structural and functional issues." Mol Pathol. **54**:57–79.
3. a, b, cPerbal B, Perbal M, Perbal A (2023). "Cooperation is the key: the CCN biological system as a gate to high complex protein superfamilies' signaling." J Cell Commun Signal. **17**(2):233–253.
4. ^Moreno-Morcillo M, Grande-García A, Ruiz-Ramos A, Del Caño-Ochoa F, Boskovic J, Ramón-Maiques S (2017). "Structural Insight into the Core of CAD, the Multifunctional Protein Leading De Novo Pyrimi dine Biosynthesis." Structure. **25**(6):912–923.e5.
5. ^Grande-García A, Lallous N, Díaz-Tejada C, Ramón-Maiques S (2014). "Structure, functional character ization, and evolution of the dihydroorotase domain of human CAD." Structure. **22**(2):185–98.
6. ^Wittmann JG, Heinrich D, Gasow K, Frey A, Diederichsen U, Rudolph MG (2008). "Structures of the hu man orotidine-5'-monophosphate decarboxylase support a covalent mechanism and provide a frame work for drug design." Structure. **16**:82–92.
7. a, bBork P (1993). "The modular architecture of a new family of growth regulators related to connective tissue growth factor." FEBS Lett. **327**:125–30.
8. ^Perbal B (2018). "The concept of the CCN protein family revisited: a centralized coordination networ k." J Cell Commun Signal. **12**(1):3–12.
9. a, bAttramadal H, Weiskirchen R, Perbal B (2024). "Report on the 12th international workshop on the C CN family of genes, Oslo, June 20–23, 2024." J Cell Commun Signal. **18**(3). doi:10.1002/ccs3.12049.
10. ^Kaasbøll OJ, Gadicherla AK, Wang JH, Monsen VT, Hagelin EMV, Dong MQ, Attramadal H (2018). "Con nective tissue growth factor (CCN2) is a matricellular preproprotein controlled by proteolytic activatio n." J Biol Chem. **293**(46):17953–17970.
11. ^Kyurkchiev S, Yeger H, Bleau AM, Perbal B (2004). "Potential cellular conformations of the CCN3(NO V) protein." Cell Commun Signal. **2**(1):9.
12. ^Planque N, Long Li C, Saule S, Bleau AM, Perbal B (2006). "Nuclear addressing provides a clue for the t ransforming activity of amino-truncated CCN3 proteins." J Cell Biochem. **99**(1):105–16.
13. a, bLazar N, Manara C, Navarro S, Bleau AM, Llombart-Bosch A, Scotlandi K, Planque N, Perbal B (200 7). "Domain-specific CCN3 antibodies as unique tools for structural and functional studies." J Cell Com mun Signal. **1**(2):91–102.
14. a, bSong MH, Jo Y, Kim YK, Kook H, Jeong D, Park WJ (2022). "The TSP-1 domain of the matricellular pro tein CCN5 is essential for its nuclear localization and anti-fibrotic function." PLoS One. **17**(4):e0267629.
15. ^Perbal B (1999). "Nuclear localisation of NOVH protein: a potential role for NOV in the regulation of ge ne expression." Mol Pathol. **52**(2):84–91.
16. ^Wahab NA, Brinkman H, Mason RM (2001). "Uptake and intracellular transport of the connective tiss ue growth factor: a potential mode of action." Biochem J. **359**(Pt 1):89–97.
17. ^Thomopoulos GN, Kyurkchiev S, Perbal B (2001). "Immunocytochemical localization of NOVH protein and ultrastructural characteristics of NCI-H295R cells." J Submicrosc Cytol Pathol. **33**(3):251–60.
18. ^Grotendorst GR, Duncan MR (2005). "Individual domains of connective tissue growth factor regulate f ibroblast proliferation and myofibroblast differentiation." FASEB J. **19**(7):729–38.
19. a, bHu K, Tao Y, Li J, Liu Z, Zhu X, Wang Z (2019). "A comparative genomic and phylogenetic analysis of t he origin and evolution of the CCN gene family." Biomed Res Int. **19**:8620878.
20. ^Gilbert W (1978). "Why genes in pieces?" Nature. **271**:9.
21. ^Perbal B (2004). "CCN proteins: multifunctional signalling regulators." Lancet. **363**(9402):62–4. doi:10.1016/S0140-6736(03)15172-0.
22. ^Zaykov V, Chaqour B (2021). "The CCN2/CTGF interactome: an approach to understanding the versati lity of CCN2/CTGF molecular activities." J Cell Commun Signal. **15**:567–580.

Declarations

Funding: No specific funding was received for this work.
Potential competing interests: No potential competing interests to declare.