



Structure, evolution and expression of collagen XXVIII: Lessons from the zebrafish

Jan M. Gebauer^{1,†}, Birgit Kobbe², Mats Paulsson^{1,2,3} and Raimund Wagener^{1,2}

1 - Center for Biochemistry, Medical Faculty, University of Cologne, D-50931 Cologne, Germany

2 - Center for Molecular Medicine, University of Cologne, D-50931 Cologne, Germany

3 - Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, D-50931 Cologne, Germany

Correspondence to Raimund Wagener: Institute for Biochemistry II, Medical Faculty, University of Cologne, Joseph-Stelzmann-Str. 52, D-50931 Cologne, Germany. [†] Current address: Institute of Biochemistry, University of Cologne, D-50674 Cologne, Germany. raimund.wagener@uni-koeln.de

<http://dx.doi.org/10.1016/j.matbio.2015.07.001>

Edited by R. Iozzo

Abstract

Collagen XXVIII is the last discovered member of the collagen superfamily and thus has been only sparsely investigated. We studied collagen XXVIII in zebrafish to gain insight into its structure, evolution and expression. In contrast to human and mouse, the zebrafish genome contains four collagen XXVIII genes, *col28a1a* and *-b*, and *col28a2a* and *-b*. Genomic context and phylogenetic analysis revealed that the *a2* branch was lost during evolution of mammals, whereas the duplication of the *a1* and *a2* branches results from the whole genome duplication in the teleost lineage. Sequence analysis revealed conservation of domain structure and the unique imperfections in the triple helical domain. Two major forms of collagen XXVIII were identified, Col28a1b in adult and Col28a2a in 3–5 dpf zebrafish. Composite agarose/polyacrylamide gel electrophoresis revealed that both these chains mainly form dimers of trimers, although Col28a1b appears to be more polydisperse. Homodimers are abundant, although it is possible that complexes consisting of Col28a2a and Col28a1a or *-a2b* occur. Peptide mass fingerprint analysis revealed that the C-terminal Kunitz domain is often proteolytically processed. In contrast to murine collagen XXVIII, the zebrafish orthologs are widely expressed and not only present in the nervous system. They are differentially expressed in the liver, thymus, muscle, intestine and skin. Altogether our results point to a unique nature of collagen XXVIII within the collagen family.

© 2015 Elsevier B.V. All rights reserved.

Introduction

The collagen superfamily is a heterogeneous group of proteins with many *in vivo* functions. Its members play important roles in tissue formation, stability and maintenance (e.g. cell adhesion, migration, differentiation, and wound healing), for review see [1]. To date 28 different collagens are known which are encoded by 46 different genes. The protein family can be subdivided into fibril, beaded filament, anchoring fibril, and network forming collagens. Structurally, collagens are defined by their Gly-X-Y repeats where X and Y are often a proline or hydroxyproline residue leading to the formation of left-handed polyproline II helices that assemble into twisted, tightly packed right-handed triple helices. However, many collagens also contain non-

collagenous domains and depending on the individual domain structure they can form very different supra-molecular assemblies.

Information about collagen XXVIII is sparse, knockout or knockdown models are not available and it is not clear if this collagen belongs to one of the four collagen subgroups or if it is unique. The human and murine proteins consist of an N-terminal VWA (von Willebrand factor type A) domain followed by a 528 amino acid residue collagenous domain, which has altogether 16 uniformly arranged, very short imperfections (G1G and G4G) in the Gly-X-Y repeat, unique among the known collagens [2]. The C-terminus is made up by a second VWA domain followed by a unique domain and a domain related to the BPTI/Kunitz family of serine protease inhibitors

za2a 1 MAVMAHSLSWCVCLGLSLVCAQ...DFYEERTGPT.....RVKTRGPTASNLQKNGDAILED...SLELAFLLDSSSETAKDN.HQEEKKTMDDIE
 za2b 1 ~MMLRLCVRWLVLSSAAQVLAQCKDFDDEDEDPEFVEDNRNRKRK.PITSNVIPIAN...KDENCNLELAFLVDSSESADKN.HGEEKSEFVTDIVN
 za1a 1 ~~~~~~QCCAVMCLILLI.....CCSARSONRRKNQKPDNNV...NTNGVKTLFCFVELAFLVDSSEKAT.MLFEQRQREBFLRFST
 za1b 1 ~~~~~~NESVLLVCVSWLL.....VSAARCNRROGETP.NNL...TTK..QNACSLVAFILDSSESAGKLLFSROKEFVRFSR
 ma1 1 ~~~~~~MRRRDVAFCLLLLPAFTMQAVYGQK..KGKPNILARIN..DFQDAICFTDVFILDSSES...KILFLDQKDFVDSISE

za2a 89 GLOSTRIDTRKTSWRAALLOYSSHVITEOTLKQKGTENFKSSAPMAYIGHGTYTTYAIINMTKIFVEESSPERIKIALLTDGFFHPRNPDIASAMA
 za2b 94 HTPNIRLQTCQGNFRFALLQYSSHVITEQSFKDWGTPSEQSRVASIPIGHGTYTTYAIINLTRYIEESGPGTVKVALMYGGASHPNPNPDIFASALA
 za1a 175 DAKIHDIHMEVIGLPGSSRDQYGRLRSTIASAPCOYLFLSLTDPLDDKLFREISELANRACPQPKSCICEKGERGLPGNPGKVGADGDPGPKGSR
 za1b 74 RLMEVQV.SGWHIRTRIALYQSSSVHINQHFNDWQDLVFLDQEDASYGICGTYSTYAIASNATLFTREISGQSVVSLMTDGDHPRNPDIIMVVA
 ma1 77 KIFQITPGRSLKYDIKLAALQSSSVQIDPLSSWKDRTKQKVKSNLIGQSTSYAIASNATLLKREGRKDGKVALMLTMDGIDHPNPDVQISIE

VWA1

za2a 189 DAKNQCKVKEFTIGLRTANPVNAANLRLSSTPASRELYNLQDITNVMEKVITQTAQLANDGCPLSQKCVCEKGERGPSGPAGKKGRPGEDGTGPAKQKQK
 za2b 194 DAKNOGKIFRIVGLISAA...NMEKQLLASAPASRYVHNQDKGVVDKIITREITKVDEVCFFSPKCTCEKGRGPSGPAGKKGRPGEDGSPGPKGQK
 za1a 175 DAKIHDIHMEVIGLPGSSRDQYGRLRSTIASAPCOYLFLSLTDPLDDKLFREISELANRACPQPKSCICEKGERGLPGNPGKVGADGDPGPKGSR
 za1b 173 EAKSHNTKIFAILGLSMRAMSNS.AKLRAVASSPACQVHSLDRGLERLLOQETTAQKQDCPRPLVLCCKQKQGLPGAPGKQDQVGGAPGAKGSR
 ma1 177 DAKILGISFITVGLSTV...VNBKLRLLSGDESNEFVLLSDPTLVDRIOERGVLFERKEH.KICECEKGEPPGPGPPGTHNPGIKGERGPKGNP

* *

za2a 289 GEAGLSGLPQRETEGEPGYKGEGERGECCTPGIKGDRGPEGPGVGTQSRGLOGLPGFQGLGPGGAQCKKGERGHAGLPGLQGDGSLGPGPKGMGF
 za2b 290 GEGSSGAPGRDTEGATGYKGEKGERGECCTPGVKGDRGPEGPGVIRGSRGLOGLPGFQGLGPGGAQCKKGERGHAGLPGLQGDGSLGPGPKGMGF
 za1a 275 GQPLNCRPGLGILRGAPGLKGEKGGKGECCASGLKGDQGPNGSGPPRGRPGDKLGTGMPGDTGPEGPVSGKGERGPGSGTGPDEGVGFPGAKGKGS
 za1b 272 GEPGAPGLPGINGEPGAPGKGDQKQDQCNCGPPGCKGKGGEGPPGPRGRPGQGLKQSPGDQSGQPPGPKDRGHTCASGLPGDITGVGFPGPKGVKN
 ma1 272 GDA.QKSETGERPGVCPGYKGDKGERGECCKPGKGDQKPEGPGVGPGRGTQGLGPPGDFGPKGFQCNKGEPPGPGYGFPGATGICQGVKGERGQ

* *

za2a 389 QGRPGPPPGPGLGEPGLPGPGQPGVQEGKPGQEGFPGPKGDRGLEPKKPRGQGLGLKGDQKGFPGPPGLPGPVGLGPGVGTQGEKGVGEPGRPGGGRG
 za2b 390 QGRMGPPGLGEPGLPGPGQPGQSGPKGPPGEGFPGPKGRGLEPGPRGPGATAIKGDQKGFPGSPGLPGPVGLGPGVGTQGEKGVGEPGRPGGGRG
 za1a 375 QGRPGTGPVIGEPGPGPEGTQGVQCNQCFPEGGLPGKGRDRFEGPKGVIRGPPGSSIKGDQKNTGERGLPGGLGQGVGFPGKGDGIPGPPGRG
 za1b 372 PGRAGPGLGQVIGEPGLPGPEGLQGNPQAEAGEGLPGEKGRDQYAGETGTRGPEGYVVKAKGAGPAGPGSPGLGMPGGRGTQGEKGVGEPVPPGQGRG
 ma1 371 EGRMGAPGPGIIGEPGPGPRGPGGAPGERGLGEGFPGPKGKQSGEGHIGQSLGGLSTIKGDQKGLGPGVPGQSPAGLEGLSGQGGGTQGPSPPGPGQ

collagenous

za2a 489 PPGGAPGPKGEOGLPGEMCTPGERGCEPEPAKGEPPGSSGLAGVPLPGEDGAPGQKGPGLPGFRGPEGAQCTGTQGEKGDQGRGIRGLPGPPGISGP
 za2b 490 PPGGAPGPKGDOGLPGESGAPGERGAGEPAKGEPPGAGLSGLPLPGEDGAPGQKGPGLGATGLRGAEGAAQVGTQGEKGDQGRGIRGLAGSPGAGP
 za1a 475 PPGVGLVPGKQDQGFPGPGPGESEGPQGERGTEGPGKPGEPVPGAPGIPGPGEDGSEVGEKGMGLPGKGERGQEGSPGKIPGKGDGIPGPKGDRG
 za1b 472 NPCTGTGPKGEOGPGPNPAGPGRGTGEPGPKGEPGLRGLAGDPGIPGQDGAQSGKGEIIGLPGAAGPPGPPGRLPGEKGRDGERGSRQPGPTGAGA
 ma1 471 PPGQSPGPKGEPVQMGPTGPRGPMGLSVQGPKEPGTVGLPGQPGVPGEDGASCKKBAGLPGTRGPEGMPEKSPGPKGDECKKSKKNQGRGQGFPG

* *

za2a 589 SCAKGEPPGTGRLGMPGLGRLASPKGDLGAPGPPGTGEGTGLGGLPGKGRDRNGEPHGESGPKGDGPGPVGLPGLPLGPEGPEGVGVPKPGKDVG
 za2b 590 SSKSGEPGTGROGLPGLGRLASPKGDLGPPGPGSPVGFEGGLTGLPKGRDRGPGLEGLTGLKGDGYPGPTGLPGPGLPGEPOCGVGPKPGKDVG
 za1a 575 PGAKGPPSGMGMGLPGPSGRGLPGAKGEPGAGPPGPHVGEHGTGTPGKGRDRGSPGPPVPGPMKGDGYPGPGLPGLPGLTGPGECCGTPGPKGDRG
 za1b 572 TCGKGNPGNVPPGVPGLGGLPGAKGDLGVPVPPGAGEGLGLAGPKGLIGLPGSGPPGLKGEYVGPYPGPPGLPLGLTGEIGPEGLPLPGPKGERG
 ma1 571 ECPKGEPCVMPGPMGASIFGPSKPGDRGPPGMPGLKGEGLPWRGPKQAQCPRGVPCAPGLKGDGYPGVAGPRGLPGPPGPMGLRGVSDTCAKGEPP

* *

za2a 689 FRGLPGLPFPFEGSLOQKQENTGRPGPPGPHGPPGEGIQGPKGDQAGPVTGPRGPSGELPGAQGDRCLOGERGSKGGKGMGDPGPTPLAGQGLKGE
 za2b 690 FRGLPFPFPFEGSLOQKQENTGRPGPPGPHGPPGEGIQGPKGDQAGPVTGPRGPSGELPGAQGDRCLOGERGSKGGKGMGDPGPTPLAGQGLKGE
 za1a 675 SPGVGTGSPGEGTGLPGPKGAVGPPGAGPGLSGPPGEGIQGPKGDGPGGLTGPGRGPGGLPGKGEKGRDGLSGERGRKGERGMPQGTQSKGAPGPKGDRG
 za1b 672 PQGSPGAGATGLGQMGPKGSLGLTGVPVPPGLGEGIQGKGEFGYQGLQGRGLTGLGLPGQKGRDRCFOGAQGRKGEPPGROGEPGLTGPPGNAGKKE
 ma1 671 VRGPPGSGFRGLSTQGPKGDTCQKGLFPFPFEGYCSOGIKGEGOGPGFPSPKGTGGLSLPGQKGEHGRGDRGVGRKGEKGTGEPGSPGKQGLQPKGD

* *

za2a 789 QGLTREIITIKLKEICCGCKCKERPMLVLFVIDSSSVGPNFELIKDFVAILDVRTVGRNATRIGLILYSLEVHLEFNILARYMTKQDVKKAVRRMPY
 za2b 790 SGLTRDEIIRMIREICCGCKCKERPMLVLFVIDSSSVGPNFELIKDFVAILDVRTVGRNATRIGLILYSLEVHLEFNILARYMTKQDVKKAVRRMPY
 za1a 775 PLGRREEVITLREICCGKCKRTSPLELVFVIDSSSVGPNFELIKDFVAILDVRTVGRNATRIGLILYSLEVHLEFNILARYMTKQDVKKAVRRMPY
 za1b 772 PGLTREIITIKLKEICCGCKCKRTSPLELVFVIDSSSVGPNFELIKDFVAILDVRTVGRNATRIGLILYSLEVHLEFNILARYMTKQDVKKAVRRMPY
 ma1 771 LGLTKEEIIKLIKEICCGCKCKETPLELVFVIDSSSVGPNFELIKDFVAILDVRTVGRNATRIGLILYSLEVHLEFNILARYMTKQDVKKAVRRMPY

* *

za2a 889 MGGTYTGTAIRKATQBAFFSARNQVRKVAIVITDQOTDKREPVLKIDAVREAQVANIEMAYAGIVNTSDPTQAEFLREINLTASDPDSEHMLIDDYNT
 za2b 890 IEGTYTGSIAHNATHAFYSARTGVKKVAIVITDQOTDKREPVLKIDAVREAQVANIEMAYAGIVNTSDPTQAEFLREINLTASDPDSEHMLIDDYNT
 za1a 875 LGGETGTSAIRATQLFQAARPGVRKVAIVITDGLADNRDAVSLKDAEAGHAGTETFFVGVVNSDSQAEFKNEINIATSDPDENYVYLTDDFLK
 za1b 872 IEGTYTGSIRKANEMFAFARPGVRKVAIVITDQOTDKRDVTKLEDAVREAHLNITMFAAGVNOQSDPIVDDFKQELKSTASPTTEHMSVSDERM
 ma1 871 LGGTYTATALLQAN.DMEKEARPGVKVAIVITDQOTDSRDKKLABVVKDANDSNVLEFVGVVKKDDPNLEIFHEKMNLIAT..DAEHVYQDFDFT

VWA2

za2a 989 LP.ALESKLVNQCEDENCALYFNRIITN...GLNGNGFGYNGQEDISSYGNIVYGSRFQESLDRQSHTRGRGHSLPLPISFG.PLKPVQENDSDG
 za2b 990 LP.ALESKLVNQCEDENSRALFNRIINDFENGYGNGLRINGNGGHNFDTSYKGSRENG.TFSATNVITERGRGTFEASTAGP..DIKQEVERVP..
 za1a 974 LH.ALESKLLNHICEHNGKVF.....SSSGKLIHPFGDPVDRLEPPSTDYFITEGKDDTPIDFTFPQQPEDY
 za1b 971 LH.ELQSKLLQKVCENIDGST.....SSKPGAFESGVPFPYD...PESDIVGT...IREVYLEREDT.....
 ma1 968 LQDTLKQKLSKKICEDFDSYI.....QVFGSPSPPEFGVSEREVSVS.....PKPAKESKSFNV

unique

za2a 1082 EDLDIKTQTKSGTGVVVKNTVOAQPGTSSSSSSSTVSLSSSVSTHSTSVNVDTKRRPVVTKESAPLDR...CQLSFEVQGSRCNVYVIRWYYVQKQ
 za2b 1084KTPDIRA.....PAVTGTSIPSKHSSGTTSSSSVNTYSTSMQFIAAPRPPPKKEVVLDP...CKLLLDQGPCREYVIRWYYVQQA
 za1a 1045 DDHIDHSYFDGSVNEPWWTETE...DGNKLRQSSISVVGPPPINWHVQ...ESSGPKTLQPPHSDNSLKHVCGQGLDGPDCREYVIRWYYVQQA
 za1b 1027ILNL.....E..DFSRR.....TTDPPLTFQDFTADE...RCLQPLDGPDCREYVIRWYYVQQA
 ma1 1027 RGQNEETESVLTETAGLAIPTPEATNTLEPLSSREGVETRI.....ENNLQSEKSL..YKDPREEALKFGEGGVVIRWYYVQQA

za2a 1178 NSCAQFWYGGGDGNDNRHTEEECKTTCVLSSTV
 za2b 1165 NSCAQFWYGGCGENNRHTEEECKTTCVVV~
 za1a 1140 NSCAQFWYGGCGNSNRHTEEDICKSTCVQT~
 za1b 1079 NSCAQFWYGGCKKNQDFSELTCKTTCVRL~
 ma1 1111 NSCAQFWYSGCNSGNRHFSEKRETCIKQ~
 ~ ~ ~ ~ ~ kunitz ~ ~ ~ ~ ~

Download English Version:

<https://daneshyari.com/en/article/8455176>

Download Persian Version:

<https://daneshyari.com/article/8455176>

[Daneshyari.com](https://daneshyari.com)