

2 The DAPIN family: a novel domain links apoptotic and interferon response proteins

Die folgende Erklärung legt die individuellen Beiträge der Co-Autoren zum nachfolgenden Manuskript dar. Von den Co-Autoren unterschriebene Fassungen dieser Erklärung liegen dieser Dissertation bei. Die Erklärungen sollen belegen, dass die erzielten Resultate im wesentlichen auf meiner Arbeit beruhen und ihre Verwendung innerhalb dieser Dissertation gerechtfertigt ist.

The DAPIN manuscript: declaration about contributions of authors

Hereby I, co-author of the manuscript „*The DAPIN family: a novel domain links apoptotic and interferon response proteins*“ which appeared in *Trends in Biochemical Sciences (2001) vol.26, no.2, pp.83-85* declare my contributions to the results and to the preparation of the manuscript. Furthermore, I agree that the statements below describe the contributions of the other co-authors correctly.

1) Eike Staub

- conducted the protein sequence analysis,
- discovered and characterised the domain,
- predicted the death domain-like structure,
- inferred the function of the domain,
- wrote the text and prepared the figures for the manuscript,
- served as the corresponding author during the review process.

2) Edgar Dahl

- raised the interest in the sequence analysis of the ASC protein by his initial observation of the differential expression of the ASC transcript in breast tumors,
- contributed to final stages of manuscript preparation by helpful comments on the style of the manuscript.

3) André Rosenthal

- was the supervisor of the project,
- contributed to final stages of manuscript preparation by helpful comments on the style of the manuscript.

The DAPIN family: a novel domain links apoptotic and interferon response proteins

Eike Staub, Edgar Dahl and André Rosenthal

We report the discovery of a protein domain, hereafter referred to as DAPIN, in diverse vertebrate and viral proteins that is associated with tumor biology, apoptosis and inflammation. Based on a secondary structure prediction, we suggest an all- α fold for DAPIN, which is also adopted by apoptotic protein domains of the CARD, death domain and death effector domain type.

Using an EST data-mining approach to search for genes that are differentially expressed between normal and cancerous breast tissue¹, we identified a gene which was recently named ASC (after apoptosis-associated speck-like protein containing a CARD). This name was given because the protein precipitates with monoclonal antibodies that target apoptotic speck-like bodies². The ASC protein sequence contains a C-terminal caspase recruitment domain (CARD), which is an adapter domain in apoptotic proteins (e.g. RAIDD, ICH-1, Ced-3, ICE) that binds to other CARDS via homophilic interactions³. The presence of a CARD suggests that ASC is involved in apoptosis.

The N-terminal sequence of ASC is similar to that of the Mediterranean fever protein pyrin. Mutations in the pyrin-coding *MEFV* gene are the cause for familial Mediterranean fever (FMF), an autosomal recessive disease characterized mainly by recurrent attacks of fever and serositis^{4,5}. Eleven of 16 mutations that contribute to the disease phenotype are located in the B30.2 domain⁶, with none in

the N-terminal region⁷. Pyrin localizes to the perinuclear cytoplasm and interacts with a putative Golgi transport protein⁸.

Using the similarity between ASC and pyrin at the N-terminus as a starting point, we identified a further 11 proteins with similar N-termini by an iterative process that involved construction of multiple alignments with CLUSTAL X, building of hidden Markov models (HMM) and protein database scanning^{9,10}. The final manually edited alignment (Fig. 1) resulted in a HMM that detected each family member with an expectation value E of $<1 \times 10^{-38}$. We named this new protein domain DAPIN (domain in apoptosis and interferon response).

The DAPIN seems to be unique to vertebrates or vertebrate-specific viruses, as no similarity to any known or predicted protein from fly, worm, yeast or bacteria could be detected. The 3' ends of the first coding exons of the pyrin, IFI 16 (interferon inducible gene 16)^{4,11} and *MNDA* genes (www.ensembl.org; Gene ID, ENSG00000073840) are consistent with the C-terminal boundaries of the alignment. Secondary structure prediction with JNet (Ref. 12) suggests an all- α structure with five α -helices. Thus, DAPIN resembles apoptotic adapter domains, the CARD, the death domain and the death effector domain, which all fold into six α -helices¹³⁻¹⁵. This similarity might indicate a common evolutionary origin for the four domains. The combination of DAPIN and CARD in two

different proteins (Fig. 2) might be the product of domain duplication and divergent evolution.

The best-characterized DAPIN family members originate from gene clusters of interferon-inducible genes on human chromosome 1q22 and the syntenic mouse region. They code for the HIN-200 family of hematopoietic interferon-inducible nuclear proteins^{16,17}. All proteins translocate to the nucleus after induction. Single family members differ in the dependence of their induction on distinct interferon subtypes and in their expression pattern among the different hematopoietic cell lineages. This family includes AIM2 (absent in melanoma 2), a possible tumor suppressor that was discovered in a screen for differentially expressed genes between malignant and benign melanoma cells¹⁸.

The structural feature common to all members of the HIN-200 family is the presence of one or two copies of a 200-amino-acid domain^{16,17,19} (HIN-200-aa in Fig. 2). Furthermore, a high degree of amino acid similarity in the N-terminal DAPIN region was described for the HIN-200 family members, except for interferon-inducible protein 202 (IFI 202). Attention was drawn to an 'imperfect' Leu zipper motif¹⁹ (compare the alignment position 72-94 of proteins 1-6, Fig. 1) with Leu residues being partly replaced by Val, Ile and Met. Our secondary structure prediction suggests two α -helices in this sequence region and therefore contradicts the assumption of a Leu-zipper

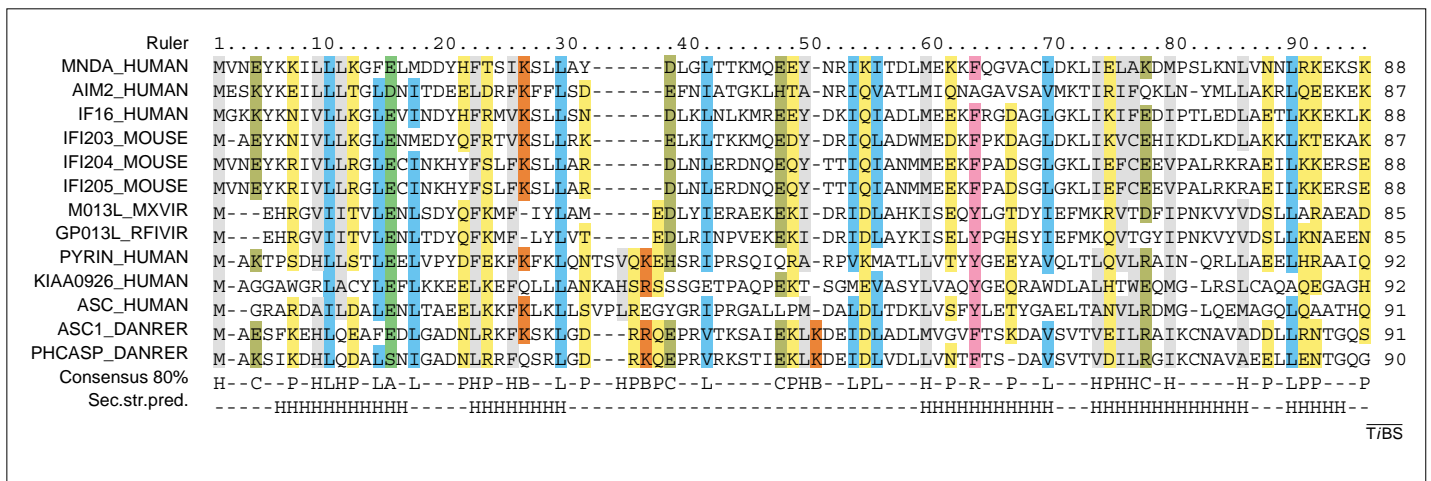


Fig. 1. Multiple alignment of DAPIN (domain in apoptosis and interferon response) regions in 13 family members. The alignment was constructed using the CLUSTAL X program with subsequent manual editing. The amino terminal Met was not used for hidden Markov model (HMM) construction. The numbers on the far right-hand side indicate the position of the DAPIN residue in the sequence that is closest to the C-terminus. The amino acid residues are colored according to an 80% consensus: P, polar (DEHKNQRST; yellow); C, charged (DEHKR; brown); A, negative (ED; green); B, positive (HRK; red); R, aromatic (FHVY; purple); L, aliphatic (ILV; blue); H, hydrophobic (ACFGHILMVVWY; grey). Secondary structure prediction (sec. str. pred.): H, α -helix; E, β -strand. The protein names consist of gene and organism identifier (DANRER, danio ferio; MXVIR, myxoma virus; MND_A, myeloid nuclear differentiation antigen; RFIVIR, rabbit fibroma virus; PHCASP, pyrin-homolog caspase). GenBank identifiers: MND_A_HUMAN, 730038; AIM2_HUMAN, 2558942; IFI16_HUMAN, 184569; IFI203_MOUSE, 6016336; IFI204_MOUSE, 124489; IFI205_MOUSE, 2833215; M013L_MXVIR, 6523868; GP013L_RFIVIR, 6578691; PYRIN_HUMAN, 4557743; KIAA0926_HUMAN, 4589484; ASC_HUMAN, 6482372; ASC1_DANRER, 7673624; PHCASP_DANRER, 7673640.

conformation. Deletion mutant studies have shown that homodimerization of MND_A depends on amino acids 52–82 in the DAPIN region²⁰. IFI 16 is a transcriptional repressor and its first 159 amino acids are sufficient to bind DNA, either directly or indirectly²¹. The abundance of basic amino acids in the DAPIN region seems to be a special feature of the HIN-200 subfamily and might

facilitate direct or indirect DNA binding. However, we do not assume a DNA-binding function for the DAPIN domain based on the predicted secondary structure similarity to other apoptotic adapter domains and on the MND_A dimerization studies.

Another DAPIN protein, the large predicted human KIAA0926 protein, has

been presented as a member of the novel NACHT (after NAIP, CIIA, HET-E and TP1 proteins) NTPase family²², which suggests a function in apoptosis or activation of major histocompatibility complex (MHC)-class II transcription. In addition, KIAA0926 contains a CARD on its C-terminus. The domain architecture and size of this protein support the assumption that KIAA0926 plays an important role in the control of apoptosis, possibly as a docking platform for other proteins.

Two open-reading frames, named M013L and GP013L from the genomes of rabbit viruses, myxoma and fibroma virus^{23,24}, encode highly similar proteins that consist almost entirely of the DAPIN domain. This supports the idea that the DAPIN exon is a functionally independent module. Myxoma virus infection causes a rapid systemic infection in European rabbits and is almost always lethal, whereas fibroma virus causes a benign fibroma at the site of invasion. In adult hosts an adaptive immune response to fibroma virus is able to cause tumor shrinking and virus elimination. M013L and GP013L are novel candidates for viral proteins that tackle the host immune response. We hypothesize that this is achieved by the formation of nonfunctional heterodimers with cellular DAPIN proteins, thereby interfering with programmed cell death in response to virus infection.

Recently, two zebrafish genes that encode other DAPIN members have been discovered²⁵. One is a predicted

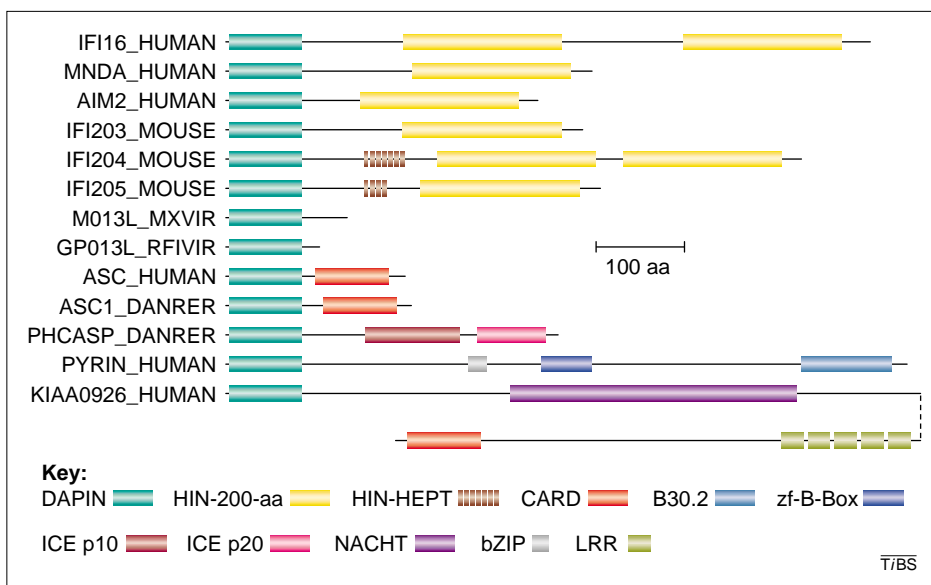


Fig. 2. Domain architecture of the DAPIN proteins (scale is approximate). Myeloid nuclear differentiation antigen (MND_A), AIM2 (absent in melanoma 2), IFI 16, IFI 203, IFI 204 and IFI 205 (where IFI is the interferon inducible protein) belong to the HIN-200 family. The domain key is given underneath. Abbreviations: ASC, apoptosis-associated speck-like protein containing a CARD; B30.2, named after the exon B30.2, identified in a search for coding sequences in the major histocompatibility complex (MHC)-class I region; bZIP, basic leucine zipper domain; CARD, caspase recruitment domain; DAPIN, domain in apoptosis and interferon response; HIN-200aa, characteristic 200-amino-acid domain for HIN-200 family (hematopoietic interferon-inducible proteins with a 200-amino-acid repeat); HIN-HEPT, heptamer repeats from HIN-200 proteins IFI 204 and IFI 205; ICE p10/ICE p20, interleukin-1 β -converting enzyme p10 or p20 domain; LRR, Leu-rich repeats; NACHT, after NAIP, CIIA, HET-E and TP1 proteins; zf-B-Box, B-box zinc finger domain.

ortholog of the human ASC protein; the other is an ICE (interleukin-1 β -converting enzyme)-like protease. The combination of the DAPIN with ICE-like protease p10 and p20 domains is consistent with the assumption of an adapter function for DAPIN that complements the protease effector function in this novel caspase.

The identification of a common DAPIN domain links a well-characterized family of nuclear interferon-inducible proteins to other proteins with putative functions in apoptosis and tumor biology, viral infection and inflammation. Future research on DAPIN proteins should reveal the physiological and biochemical function of this domain and the small viral DAPIN proteins might be especially helpful for this task.

Acknowledgements

We would like to thank an anonymous reviewer for his / her helpful suggestions.

References

- Schmitt, A.O. *et al.* (1999) Exhaustive mining of EST libraries for genes differentially expressed in normal and tumor tissues. *Nucleic Acids Res.* 27, 4251–4260
- Masumoto, J. *et al.* (1999) ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells. *J. Biol. Chem.* 274, 33835–33838
- Hofmann, K. *et al.* (1997) The CARD domain: a new apoptotic signalling motif. *Trends Biochem. Sci.* 22, 155–156
- The French FMF Consortium (1997) A candidate gene for familial Mediterranean fever. *Nat. Genet.* 17, 25–31
- The International FMF Consortium (1997) Ancient missense mutations in a new member of the RoRet gene family are likely to cause familial Mediterranean fever. *Cell* 90, 797–807
- Vernet, C. *et al.* (1993) Evolutionary study of multigenic families mapping close to the human MHC class I region. *J. Mol. Evol.* 37, 600–612
- Mulley, J.C. (1999) The genetic basis for periodic fever. *Am. J. Hum. Genet* 64, 939–942
- Chen, X. *et al.* (2000) The familial mediterranean fever protein interacts and colocalizes with a putative golgi transporter. *Proc. Soc. Exp. Biol. Med.* 224, 32–40
- Jeanmougin, F. *et al.* (1998) Multiple sequence alignment with CLUSTAL X. *Trends Biochem. Sci.* 23, 403–405
- Eddy, S.R. (1998) Profile hidden Markov models. *Bioinformatics* 14, 755–763
- Trapani, J.A. *et al.* (1994) Genomic organization of IFI16, an interferon-inducible gene whose expression is associated with human myeloid cell differentiation: correlation of predicted protein domains with exon organization. *Immunogenetics* 40, 415–424
- Cuff, J.A. *et al.* (2000) Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins* 40, 502–511
- Chou, J.J. *et al.* (1998) Solution structure of the RAIDD CARD and model for CARD/CARD interaction in caspase-2 and caspase-9 recruitment. *Cell* 94, 171–180
- Eberstadt, M. *et al.* (1998) NMR structure and mutagenesis of the FADD (Mort1) death-effector domain. *Nature* 392, 941–945
- Huang, B. *et al.* (1996) NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain. *Nature* 384, 638–641
- Dawson, M.J. *et al.* (1996) HIN-200: a novel family of IFN-inducible nuclear proteins expressed in leukocytes. *J. Leukoc. Biol.* 60, 310–316
- Landolfo, S. *et al.* (1998) The I β 200 genes: an emerging family of IFN-inducible genes. *Biochimie* 80, 721–728
- DeYoung, K.L. *et al.* (1997) Cloning a novel member of the human interferon-inducible gene family associated with control of tumorigenicity in a model of human melanoma. *Oncogene* 15, 453–457
- Johnstone, R.W. *et al.* (1999) Transcription and growth regulatory functions of the HIN-200 family of proteins. *Mol. Cell Biol.* 19, 5833–5838
- Xie, J. *et al.* (1997) MNDA dimerizes through a complex motif involving an N-terminal basic region. *FEBS Lett.* 408, 151–155
- Johnstone, R.W. *et al.* (1998) The human interferon-inducible protein, IFI 16, is a repressor of transcription. *J. Biol. Chem.* 273, 17172–17177
- Koonin, E.V. *et al.* (2000) The NACHT family – a new group of predicted NTPases implicated in apoptosis and MHC transcription activation. *Trends Biochem. Sci.* 25, 223–224
- Cameron, C. *et al.* (1999) The complete DNA sequence of myxoma virus. *Virology* 264, 298–318
- Willer, D.O. *et al.* (1999) The complete genome sequence of Shope (rabbit) fibroma virus. *Virology* 264, 319–343
- Inohara, N. *et al.* (2000) Genes with homology to mammalian apoptosis regulators identified in zebrafish. *Cell Death Differ.* 7, 509–510

E. Staub*

E. Dahl

A. Rosenthal

metaGen Gesellschaft für Genomforschung mbH, 14195 Berlin, Germany.

*e-mail: eike.staub@metagen.de