Disease-associated variants in PYP AF1 and NOD2 result in similar alterations of conserved sequence

Mario Albrecht1,∗, Thomas Lengauer1 and Stefan Schreiber2

1Max-Planck-Institute for Informatics, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany and 2First Department of Medicine, Christian-Albrechts-Universität Kiel, Schittenhelmstr. 12, 24105 Kiel, Germany

Received on February 5, 2003; accepted on May 7, 2003

ABSTRACT
Sequence variations in the gene products PYP AF1/CIAS1 and NOD2/CARD15 have been associated with several auto-inflammatory diseases that, although clinically different, share a similar inflammatory pathophysiology. A multiple sequence alignment of homologous proteins demonstrates that some of the missense variants are located in highly conserved regions of the NTPase domain and possibly impair NTP-hydrolysis. Intriguingly, one of the variations, which is found identically in PYP AF1 and NOD2, is located at the same alignment position. Our findings suggest that evolutionary gene duplication can give rise to disease families because variants affect conserved sequence in a similar fashion.

Contact: mario.albrecht@mpi-sb.mpg.de

INTRODUCTION
Recently, several missense variants in the PYPAF1 (CIAS1) and NOD2 (CARD15) genes have been identified to be independently associated with clinically distinct diseases, which, however, have shared pathophysiological characteristics of autoimmune inflammatory processes. Genetic variations in the PYPAF1 gene have been linked to chronic infantile neurological cutaneous and articular syndrome (CINCA) [also known as neonatal-onset multisystem inflammatory disease (NOMID)], familial cold autoinflammatory syndrome (FCAS) [also called familial cold urticaria (FCU)], and Muckle–Wells syndrome (MWS) (Hoffman et al., 2001, 2003; Aganna et al., 2002; Dodé et al., 2002; Feldmann et al., 2002). Single nucleotide polymorphisms in the NOD2 gene have been found to confer susceptibility to Blau syndrome (BS) (Micieli-Richard et al., 2001; Wang et al., 2002) and Crohn’s disease (CD), one of the two main types of chronic inflammatory bowel disease (Hampe et al., 2001; Hugot et al., 2001; Ogura et al., 2001; Lesage et al., 2002). FCAS, MWS and BS are rare autosomal-dominant disorders.

CINCA shows joint manifestations with recurrent fever and inflammation, skin rash, chronic meningitis, progressive visual defect and perceptive deafness with increasing age. FCAS and MWS are characterized by recurrent episodes of fever, urticarial rash, arthralgia, myalgias and conjunctivitis. Symptoms of FCAS are triggered by exposure to cold, whereas a progressive sensorineural hearing loss of late onset occurs in MWS. The main complication of MWS is secondary renal amyloidosis, which is sometimes associated with FCAS and also seen in CD. The observed symptoms of BS are skin rash with camptodactyly, granulomatous arthritis, uveitis and occasionally amyloidosis.

The expressed PYPAF1 and NOD2 proteins, with as yet unknown biological function, are members of a protein family putatively involved in the regulation of apoptosis and inflammatory reactions due to the recognition of bacterial antigens (Koonin and Aravind, 2000; Harton et al., 2002; Grenier et al., 2002). These proteins share a similar domain architecture that contains N-terminal homologous PYRIN (named after the protein pyrin) or caspase-activating recruitment domains (CARDs), one centrally located NACHT NTPase domain and one C-terminal leucine-rich repeat (LRR) domain. In the following, the comparative analysis of the genetic variations with respect to their structural impact on the protein level will give important insights into disease mechanisms.

MATERIALS AND METHODS
Protein sequences were retrieved from the SWISS-PROT/TrEMBL (SPTrEMBL) database (Bairoch and Apweiler, 2000). Their accession numbers are given in the caption of Figure 1. The structure of the β-subunit of the bovine F1-ATPase was obtained from the PDB (Berman et al., 2000), using the PDB identifier 1bmf, chain F. Its secondary structure assignment was taken from the DSSP database (Kabsch and Sander, 1983). The multiple sequence alignment was constructed by CLUSTAL W (Thompson et al., 1994; Higgins et al., 1996) and improved by minor manual modifications. To predict the secondary structure of proteins with unknown
Fig. 1. Multiple sequence alignment of the nucleotide-binding sites in NACHT NTPase domains and in the β-subunit of the F₁-ATPase. (a) Region with Walker A motif. (b) Region with Walker B motif. (c) Region with cluster of sequence variations. The secondary structure of bovine F₁-ATPase (bF₁beta) and the corresponding prediction by the PSIPRED server for PYPAF1 are depicted in the upper part, and for NOD2 and CIITA in the lower part of the alignment (α-helices are represented by curled lines, β-strands by arrows). The alignment columns with strictly conserved residues are highlighted in dark gray boxes, those in which more than 60% of the residues are physico-chemically equivalent are shown in light gray boxes. Dashed lines above the alignment indicate the positions of the Walker A and B motifs. Missense variants associated with the autoinflammatory diseases CINCA, FCAS, MWS, BS and sequence variations observed with CD patients are annotated for PYPAF1 and NOD2. An asterisk marks the conserved mutation arginine to tryptophan (R→W) in PYPAF1 and NOD2. A triangle points to the conserved aspartate that participates in the coordination of the magnesium ion chelating the nucleotide. Solid circles denote the other residues of the F₁-ATPase involved in binding to the Mg-nucleotide. The SPTrEMBL accession numbers of the selected PYRIN/CARD-, NACHT- and LRR-containing proteins are as follows (note that all human proteins PYPAF1–7, DEFCAP and MATER contain a N-terminal PYRIN domain, while this domain seems to replaced by another uncharacterized domain in mouse MATER; all other proteins carry N-terminal CARD domains; however, some splicing variants of CIITA do not have this domain): human PYPAF1/CIAS1/CRYOPYRIN/NALP3, Q96P20; human PYPAF2/NALP2/NBS1/PAN1, Q9NX02; human PYPAF3/NALP7/NOD12, Q8WX94; human PYPAF4/NALP4/PAN2/RNHI, Q9MN2; human PYPAF5/NALP6/PAN3, P59044; human PYPAF6/NALP11/NOD17, P59045; human PYPAF7/MONARCH-1/NALP12/PAN6, P59046; human MATER/PYPAF8/NALP5, P59047; mouse MATER, Q9JLR2; human DEFCAP/CARD7/NALP1/NAC, Q9CO00; human CLAN/CARD12/IPAF, Q9NP44; human CIITA/MHC2TA, P33076; mouse CIITA/MHC2TA, Q9TPP1; human NOD1/CARD4, Q9Y23; human NOD2/CARD15/IBD1, Q9HC29.

RESULTS AND DISCUSSION

We compiled a list of proteins with PYRIN/CARD, NACHT and LRR domains and assembled a multiple sequence alignment of their NACHT domains with focus on the nucleotide-binding site (NBS) and the positions of the known sequence variants (Fig. 1). The NTPase domain of the ATP-hydrolyzing β-subunit of the F₁-ATPase was included because its crystal structure is known (Abrahams et al., 1994) in contrast to that of the NACHT domain. We annotated the sequence alignment with the known secondary structures of the F₁-ATPase and the corresponding predictions for the NACHT domains.

Generally, the numerous members of the NTPase family are characterized by their exact arrangement of at least four parallel β-strands forming a central pleated β-sheet with few flanking α-helices, though the overall folding shape of the NBS is preserved in all families (Ogura and Wilkinson, 2001). Therefore, it is not surprising that the consistently predicted secondary structures of the NACHT domains are slightly different from the F₁-ATPase structure. However, the sequence alignment indicates a conserved structural core of NTP-binding residues near the C-termini of the β-strands. This observation is typical of most NTPases and confirmed by experimentally determined structures (Ogura and Wilkinson, 2001).
In the multiple sequence alignment (Fig. 1), we examined the residues of the F1-ATPase that are known to participate in binding to the Mg-nucleotide (Abrahams et al., 1994). NTPases normally contain two well-characterized consecutive NBS signatures, the Walker A and B motifs. The Walker A motif contains the P-loop whose residues bind to the polyphosphate moiety of NTP. Specific amino acids of the Walker B motif are also implicated in NTP-binding with the conserved aspartate particularly involved in Mg2+ coordination.

We located the mutations Arg260Trp (778 CGG → TGG) and Arg334Trp (1000 CGG → TGG), underlying both FCAS and MWS in PYPAF1 and BS in NOD2, respectively, in the sequence alignment near the Walker A motif (Fig. 1a). The respective variant amino acids in PYPAF1 and NOD2 align with each other within a conserved region of four neighboring residues close to the C-terminus of the preceding β-strand. As observed in the aligned F1-ATPase and similarly in other NTPases, the variant arginine, the preceding glutamate and...
the glutamate three residues downstream are involved in binding to the Mg-nucleotide (Abrahams et al., 1994; Muneyuki et al., 2000). The replacement of the negatively charged base arginine by the amphiphatic tryptophan results in a different physico-chemical environment that may impair binding and hydrolysis of NTP.

The finding of an identical mutation at analogous sequence positions in two homologous proteins must be considered extremely improbable to occur by chance. Thus it points to a close relationship of the diseases. Additional genetic variations in PYPAF1 and NOD2 associated with CINCA and MWS, or found with CD patients, support this link as they are often observed to cluster in the vicinity of each other in the alignment, see Figure 1c for an exemplary cluster located inside the structural core of NTPases near the C-terminus of a β-strand. Several other residue changes observed may also disturb NTP-hydrolysis because they are located close to the P-loop of the Walker A motif (Fig. 1a) and the NBS in the Walker B motif (Fig. 1b), where we further identified the identical mutation D303N in PYPAF1 as associated both with CINCA and MWS.

CONCLUSIONS

In summary, based on a multiple sequence alignment, we have observed clusters of missense variants of the homologous disease gene products PYPAF1 and NOD2 in highly conserved and functionally relevant sequence regions of the NACHT NTPase domain. Some of the identified variations possibly impair NTP-hydrolysis as judged from the comparison with the NBS in the structure of the F1-ATPase. In particular, the variation Arg → Trp in both homologs is associated with three clinically distinct diseases and is located at the same alignment position that is known to function in binding the nucleotide to the F1-ATPase.

Our results suggest a disease process with shared similarities in the granulomatous diseases CINCA, FCAS, MWS, BS and CD. This finding strongly supports the hypothesis that different autoinflammatory diseases are controlled by regulatory pathways that are closely related because the primary variants affect conserved sequences in gene families. Specifically, gene duplications could permit the development of disease families. It remains to be seen whether this mechanism could explain species differences in disease development. A limited diversity of primary pathways, as has been suggested from a comparison of linkage analyses (Becker et al., 1998), could explain the remarkable overlap of clinical symptoms and pathophysiological features of many disorders.

ACKNOWLEDGEMENTS

Part of this research has been funded by the Deutsche Forschungsgemeinschaft (DFG) under contract no. LE 491/14-1 and by the National Genome Research Network (NGFN).

REFERENCES


