Published in final edited form as:
J Mol Biol. 2005 December 9; 354(4): 743–750. doi:10.1016/j.jmb.2005.10.018.

Correlation of Gene and Protein Structures in the FXYD Family Proteins

Carla M. Franzin, **Jinghua Yu**, **Khang Thai**, **Jungyuen Choi**, and **Francesca M. Marassi***
The Burnham Institute, 10901 North Torrey Pines Road La Jolla, CA 92037, USA

Abstract

The FXYD family proteins are auxiliary subunits of the Na,K-ATPase, expressed primarily in tissues that specialize in fluid or solute transport, or that are electrically excitable. These proteins range in size from about 60 to 160 amino acid residues, and share a core homology of 35 amino acid residues in and around a single transmembrane segment. Despite their relatively small sizes, they are all encoded by genes with six to nine small exons. We show that the helical secondary structures of three FXYD family members, FXYD1, FXYD3, and FXYD4, determined in micelles by NMR spectroscopy, reflect the structures of their corresponding genes. The coincidence of helical regions, and connecting segments, with the positions of intron–exon junctions in the genes, support the hypothesis that the FXYD proteins may have been assembled from discrete structural modules through exon shuffling.

Keywords

FXYD; Na,K-ATPase; NMR structure; membrane protein; gene

The FXYD family proteins are tissue-specific and physiological-state-specific auxiliary subunits of the Na, K-ATPase, the primary enzyme responsible for maintaining the equilibrium of Na and K ion concentrations across animal cell membranes.[1-3] All of the FXYD genes are expressed in the early stages of fetal life, with prevalence in tissues that specialize in fluid or solute transport, or that are electrically excitable. In keeping with their distribution in tissues, several FXYD family members have been shown to regulate ion transport by binding and modulating the activity of Na, K-ATPase molecules.[2-11] In addition, certain FXYD family members can induce ionic currents in *Xenopus* oocytes, or in phospholipid bilayers, although the direct formation of ion channels has not been demonstrated *in vivo*.[12-16]

Seven FXYD family members have been identified in mammals. FXYD1 (PLM; phospholemman) is the principal substrate of hormone-stimulated phosphorylation by cAMP-dependent protein kinase A and C in sarcolemma heart.[17] FXYD2 (gamma), unique in the family for having two alternative splice variants (FXYD2a and FXYD2b), and FXYD4 (CHIF; channel-inducing factor; corticosteroid hormone induced factor), are each expressed in distinct, specialized segments of the kidney, with unique expression patterns that help explain the physiological differences in Na,K-ATPase activity among the nephron segments.[5,6,13,18,19] Two family members, FXYD3 (Mat8; mammary tumor protein 8 kDa) and FXYD5 (RIC; resembles ion channel; dyshaderin), are expressed in cancers and play a role in tumor progression.[20-23] FXYD6 and FXYD7 are expressed exclusively in the brain,

*Corresponding author fmarassi@burnham.org.

and are believed to play a role in neuron excitability during postnatal development and in the adult brain.[8,24]

Membrane proteins belonging to what is now recognized as the FXYD family were first discovered, in association with the Na,K-ATPase,[4] or as a major substrate of hormone-stimulated phosphorylation in sarcolemma.[25] The subsequent sequencing and identification of additional proteins, with similar structural and functional features, pointed to the existence of a family of homologous membrane proteins involved in ion transport regulation.[5,13,17,20,21] The complete gene family was established recently by expressed sequence tags (EST) database analysis, in a detailed genetic study that has provided complete cDNA sequences, protein signature sequences, and expression patterns for the genes.[1] A surprising finding of this study was that the proteins, despite their relatively small sizes, ranging from about 60 to 160 amino acid residues, are all encoded by genes with six to nine small exons. In contrast, phospho-lamban, another small, single-spanning, membrane protein that regulates the Ca-ATPase of cardiac tissue, and has no sequence homology to the FXYD proteins, is encoded by a single exon.[26] Here, we report that the secondary structures of the FXYD proteins reflect the structures of their corresponding genes, suggesting that they might have been assembled from modules through exon shuffling.

The FXYD family members share a core homology of 35 invariant and conserved amino acid residues, in and around a single transmembrane segment (Figure 1). The short signature motif PFXYD (Pro, Phe, X, Tyr, Asp), from which the family takes its name, is invariant in all seven known mammalian examples, and identical in those of other vertebrates, except for Pro, and residue X is typically Tyr, but can also be substituted by Thr, Glu, or His. In all the proteins, conserved basic residues flank the transmembrane domain, the extracellular N termini are acidic, and the cytoplasmic C termini are basic; however, outside the signature homology motif there is little sequence conservation among the family members. The presence of two conserved Gly residues in the transmembrane domain of all FXYD proteins (G30 and G41 in FXYD2a), suggests that it is involved in specific intramembrane helix–helix interactions. Indeed, in a family with primary hypomagnesemia, a disease characterized by renal or intestinal Mg loss, the dominant-negative mutation, G41R, in FXYD2a causes misrouting of the protein by inhibiting its association with the Na,K-ATPase, suggesting that G41 is important for mediating helix–helix interactions between the FXYD transmembrane helix and the Na,K-ATPase.[27]

The first step in our structural studies of the FXYD family proteins was to find a detergent that would be suitable for high-resolution NMR spectroscopy. We examined the $^{1}H/^{15}N$ heteronuclear single quantum correlation (HSQC) spectra of the proteins in the best-characterized micelle-forming detergents (DHPC, di-heptanoyl-phosphocholine; DPC, dodecylphosphocholine; LPPG, lyso-palmitoylphosphoglycerol; OG, octyl-glucopyranoside; SDS, sodium dodecylsulfate), at various conditions (protein, detergent and salt concentration; pH; temperature). The highest quality spectra were obtained for the proteins in SDS micelles, at 40°C (Figure 2). Although SDS is widely assumed to be a universal protein-denaturing detergent because of its common use in protein electrophoresis, many hydrophobic membrane proteins actually retain their structures in SDS micelles.[28] Furthermore, since all functional studies of Na,K-ATPase have been done on enzymes purified in the presence of SDS, and the non-covalent associations of α, β and FXYD subunits are maintained through the SDS purification process,[29-31] we reasoned that this detergent would also be a good choice for FXYD structural studies.

Each of the three FXYD proteins that we studied adopts a unique structure in SDS, reflected in the HSQC spectra shown in Figure 2, and the 1.5 ppm dispersion of the amide ^{1}H chemical shift frequency is typical of native helical membrane proteins in micelles.

Furthermore, the HSQC spectra obtained in $^{2}H_{2}O$, revealed a core region in each protein, with amide protons that exchange very slowly with the surrounding aqueous solvent (Figure 3(b)). These regions, spanning V22 to S37 for FXYD1, I23 to S38 for FXYD3, and I23 to L37 for FXYD4, match the hydrophobic transmembrane segments of the proteins identified by the hydropathy plots (Figure 3(a)), and reflect the strong intra-molecular hydrogen bonds of transmembrane helices in the low dielectric environment of the membrane or, in this case, the micelle interior.

To determine the secondary structures of the three proteins, we relied primarily on $^{1}H–^{15}N$ residual dipolar couplings (RDCs), measured from weakly aligned samples and analyzed in terms of dipolar waves.[32] These were supplemented with chemical shifts, analyzed in terms of chemical shift indices (CSI),[33] and with TALOS.[34] The RDCs measured in solution NMR experiments provide direct orientation restraints relative to an intra-molecular reference frame,[35-37] and are very useful for evaluating helical residues, helix regularity, and the relative orientations of helical segments.[38,39] This is particularly useful for membrane proteins where structure determination is usually hindered by the shortage of reliable long-range nuclear Overhauser effects (NOEs) as distance restraints.

In the transmembrane helix of each of the three FXYD proteins, the fits of experimental RDCs to a sinusoid with the signature α-helical periodicity of 3.6 residues demonstrate that the transmembrane helices are close to ideal (Figure 3(c)). In each protein, the boundaries of the transmembrane helix were identified from the changes in sinusoid periodicity, and correlate well with $^{13}C^{\alpha}$ CSI profiles (Figure 3(d)), as well as with dihedral angles obtained from TALOS analysis. In all three proteins, the transmembrane helix is preceded by a short helical stretch of approximately five residues, beginning at the aspartate residue of the FXYD signature sequence. This helix-break-helix motif spans exactly the entire length of exon 4 in FXYD1, exon 6 in FXYD3, and exon 5 in FXYD4, and is followed by a third short helix, spanning the stop-transfer signal sequence of each protein, and encoded by the next exon (exon 5 in FXYD1; exon 7 in FXYD3; exon 6 in FXYD4). In all the three proteins, these three helices, plus the preceding PFXYD signature motif, constitute a common core module, which spans the entire 35 residue FXYD protein homology region, over three exons.

The heteronuclear $^{1}H–^{15}N$ NOE measurements are excellent indicators of local backbone motions in proteins, and have been used to characterize the dynamics of helical membrane proteins in micelles (reviewed by Opella & Marassi[40]). Throughout the core helical modules of FXYD1, FXYD3, and FXYD4, all residues have similar positive values of the $^{1}H–^{15}N$ NOE, reflecting similar values of the rotational correlation time, and indicating that the three helices are rigidly connected (Figure 3(e)). Lower negative values of the heteronuclear NOE, reflecting additional backbone motions, are present in residues near the N and C termini, and at the boundaries of the core modules, marked by the junction between exons 5 and 6 in FXYD1, 7 and 8 in FXYD3, and 6 and 7 in FXYD4.

Resonance line-widths also reflect local backbone dynamics, and correlate with resonance intensities, which can be easily measured as peak heights in the HSQC spectra of proteins. Large-amplitude backbone motions that are rapid compared to the overall reorientation rate of the protein reduce the line-widths, and increase the resonance intensities. All of the resonances from amino acids within the core helical modules of the three proteins have similar peak intensities that plateau at minimum values (Figure 3(f)), indicating that the three helices are rigidly connected. In contrast, residues at the core module boundaries, and at the terminal regions of the proteins have greater intensities reflecting increased dynamics in these regions.

Beyond the core helical module, FXYD1 has a fourth helix, spanning from F60 to S68 near the C terminus, which is well defined by the RDCs, as well as by CSI and TALOS analyses of the chemical shifts. This helix contains the S63 and S68 consensus sites for phosphorylation by PKA and PKC, and it is separated from the core module by a flexible linker, with significantly smaller values of the $^{1}H–^{15}N$ NOE, and significantly greater peak intensities. The position of the flexible linker coincides with the junction between exons 5 and 6, and the fourth helix ends at the end of exon 6.

The addition of $MnCl_2$ to FXYD1 in micelles caused significant line broadening and disappearance of the $^{1}H/^{15}N$ HSQC peaks from amino acids in helix 1, in the flexible connecting segment between the helices 3 and 4, and in the N and C-terminal regions of FXYD1. In contrast, the peaks from amino acids in helix 2 (the transmembrane helix), in helix 3, and in helix 4, were largely unaffected. The paramagnetic electrons in the Mn^{2+} induce distance-dependent broadening of peaks from protein sites that are solvent-exposed, while residues in the hydrophobic interior of the micelle, as those in the transmembrane helix, are mainly unaffected. This result suggests that helices 3 and 4 are tightly associated with the micelle, and is consistent with the ^{15}N chemical shift solid-state NMR spectra of FXYD1 in lipid bilayers, which indicate the presence of a helical segment associated with the membrane surface.[41]

FXYD3 also appears to form a fourth helix at the start of exon 8, although it is very short and not as well defined by the smaller values of the RDCs. A short and somewhat more flexible stretch of amino acids is evident from the small, but significant, increase in peak intensities at the junction between exons 7 and 8 connecting helix 3 to helix 4. FXYD4, on the other hand, has a very well defined fourth helix, which begins at the start of exon 7 and spans its entire length, before reaching the mobile C terminus, which coincides with exon 8.

For FXYD1, FXYD3, and FXYD4, solid-state NMR studies of uniformly and selectively ^{15}N-labeled proteins in lipid bilayers, indicate the presence of transmembrane and C-terminal helical segments similar to those found in micelles.[41] For FXYD4, the emerging data indicate that the transmembrane helix exhibits the helix-break-helix motif that is observed in micelles, hinting that the structures in bilayers are similar to those in SDS micelles. This has been observed for other membrane polypeptides studied by both methods.[40]

Many proteins are composed of discrete domains that are characterized by their ability to fold independently and to confer specific biological functions. These modular proteins are often generated by exon shuffling, where copies of individual exons are recruited from different unrelated genes, to create new combinations of exons.[42] Because the modules can be easily moved around, they have played a dominant role in protein evolution. Each module tends to be encoded by one, or a combination of exons, that begin and end in the same splice frame, and the module organization in the proteins is strongly correlated with the intron–exon structure of their corresponding genes. According to the rules of exon shuffling, introns may interrupt the reading frame of a gene between two consecutive codons (phase 0 introns), between the first and second nucleotide of a codon (phase 1 introns), or between the second and the third nucleotide (phase 2 introns). Exons are then classified into symmetric (0-0; 1-1; 2-2) or asymmetric (0-1; 0-2; 1-0; 1-2; 2-0; 2-1) classes, depending on the phase of the flanking introns.[43,44] Symmetric exons, or exon sets, are the only ones that can be inserted into introns of the same phase, undergo tandem duplication, or be deleted without disturbing the reading frame.

To examine the relationship between gene structure and protein structure in the FXYD family, we mapped the splice frame diagrams of the seven mammalian genes onto the

organization of the structured modules determined by NMR, and of the protein homology domains (Figure 4). The gene data, which had been determined by EST analysis and experimentally,[1,45-47] were obtained from the NCBI databank (National Center for Biotechnology Information†). The majority of mobile modules are encoded by symmetric exons, or exon sets, of class 1-1, while only a few class 0-0 and class 2-2 domains are known.[43,44] Similarly, in the FXYD family proteins, many of the intron–exon junctions are conserved across the family, and the majority of introns (64%) have phase 1, and separate symmetric exons, or exon sets, of class 1-1.

In all the mammalian FXYD proteins, the FXYD signature motif (Figure 4, red boxes) is encoded by a symmetric exon of class 1-1. It is followed by the transmembrane domain (Figure 4, green boxes), also encoded by a class 1-1 exon, which forms a helix-break-helix structure that is conserved in FXYD1, FXYD3, and FXYD4, with helix breaks that exactly match the intron–exon junctions in these three family members. The stop-transfer signal (Figure 4, white boxes) is encoded by an asymmetric class 1-2 exon, which is interrupted by a phase 2 intron, and in FXYD1, FXYD3, and FXYD4, it forms a short helix followed by a more mobile linker. In these three proteins, the neighboring exon belongs to class 2-1, and encodes a fourth helix. Taken together, this asymmetric pair begins and ends in the same splice frame as the preceding sequence, and forms a helix-link-helix structure that is a good candidate for a mobile module.

The large number of exons in the genes of the relatively small FXYD proteins is intriguing and has been suggested to reflect modular gene assembly.[1,45] The presence of conserved modules in different FXYD family members suggests that they are derived from a common ancestor gene, from which FXYD5 appears to have been the first to diverge.[3] The coincidence of intron–exon junctions with helical structures and flexible connecting segments, in the three family members, FXYD1, FXYD3, and FXYD4, further supports this hypothesis, and indicates that the proteins in this family may have been assembled from discrete structural modules through exon shuffling. The multiple exon organization of the FXYD genes could serve to confer high levels of structural and functional diversity among the family members.

Acknowledgments

This research was supported by a grant from the National Institutes of Health (R01 CA082864). The NMR studies utilized the Burnham Institute NMR Facility, supported by a grant from the National Institutes of Health (P30 CA030199).

References

1. Sweadner KJ, Rael E. The FXYD gene family of small ion transport regulators or channels: cDNA sequence, protein signature sequence, and expression. Genomics 2000;68:41–56. [PubMed: 10950925]
2. Crambert G, Geering K. FXYD proteins: new tissue-specific regulators of the ubiquitous Na, K-ATPase. Sci STKE 2003;2003:RE1. [PubMed: 12538882]
3. Garty H, Karlish SJ. Role of FXYD proteins in ion transport. Annu Rev Physiol. 2005 In the press.
4. Forbush B 3rd, Kaplan JH, Hoffman JF. Characterization of a new photoaffinity derivative of ouabain: labeling of the large polypeptide and of a proteolipid component of the Na, K-ATPase. Biochemistry 1978;17:3667–3676. [PubMed: 210802]
5. Mercer RW, Biemesderfer D, Bliss DP Jr, Collins JH, Forbush B 3rd. Molecular cloning and immunological characterization of the gamma polypeptide, a small protein associated with the Na, K-ATPase. J Cell Biol 1993;121:579–586. [PubMed: 8387529]

† www.ncbi.nlm.nih.gov

6. Arystarkhova E, Wetzel RK, Asinovski NK, Sweadner KJ. The gamma subunit modulates Na(+) and K(+) affinity of the renal Na,K-ATPase. J Biol Chem 1999;274:33183–33185. [PubMed: 10559186]
7. Mahmmoud YA, Vorum H, Cornelius F. Identification of a phospholemman-like protein from shark rectal glands. Evidence for indirect regulation of Na,K-ATPase by protein kinase *c via* a novel member of the FXYDY family. J Biol Chem 2000;275:35969–35977. [PubMed: 10961995]
8. Beguin P, Crambert G, Monnet-Tschudi F, Uldry M, Horisberger JD, Garty H, Geering K. FXYD7 is a brain-specific regulator of Na,K-ATPase alpha 1-beta isozymes. EMBO J 2002;21:3264–3273. [PubMed: 12093728]
9. Crambert G, Fuzesi M, Garty H, Karlish S, Geering K. Phospholemman (FXYD1) associates with Na,K-ATPase and regulates its transport properties. Proc Natl Acad Sci USA 2002;99:11476–11481. [PubMed: 12169672]
10. Garty H, Lindzen M, Scanzano R, Aizman R, Fuzesi M, Goldshleger R, et al. A functional interaction between CHIF and Na-K-ATPase: implication for regulation by FXYD proteins. Am J Physiol Renal Physiol 2002;283:F607–F615. [PubMed: 12217851]
11. Crambert G, Li C, Claeys D, Geering K. FXYD3 (Mat-8), a new regulator of Na,K-ATPase. Mol Biol Cell 2005;16:2363–2371. [PubMed: 15743908]
12. Moorman JR, Palmer CJ, John JE 3rd, Durieux ME, Jones LR. Phospholemman expression induces a hyperpolarization-activated chloride current in *Xenopus* oocytes. J Biol Chem 1992;267:14551–14554. [PubMed: 1378834]
13. Attali B, Latter H, Rachamim N, Garty H. A corticosteroid-induced gene expressing an “IsK-like” K+ channel activity in *Xenopus* oocytes. Proc Natl Acad Sci USA 1995;92:6092–6096. [PubMed: 7597086]
14. Moorman JR, Ackerman SJ, Kowdley GC, Griffin MP, Mounsey JP, Chen Z, et al. Unitary anion currents through phospholemman channel molecules. Nature 1995;377:737–740. [PubMed: 7477264]
15. Morrison BW, Moorman JR, Kowdley GC, Kobayashi YM, Jones LR, Leder P. Mat-8, a novel phospholemman-like protein expressed in human breast tumors, induces a chloride conductance in *Xenopus* oocytes. J Biol Chem 1995;270:2176–2182. [PubMed: 7836447]
16. Minor NT, Sha Q, Nichols CG, Mercer RW. The gamma subunit of the Na,K-ATPase induces cation channel activity. Proc Natl Acad Sci USA 1998;95:6521–6525. [PubMed: 9600999]
17. Palmer CJ, Scott BT, Jones LR. Purification and complete sequence determination of the major plasma membrane substrate for cAMPdependent protein kinase and protein kinase C in myocardium. J Biol Chem 1991;266:11126–11130. [PubMed: 1710217]
18. Shi H, Levy-Holzman R, Cluzeaud F, Farman N, Garty H. Membrane topology and immunolocalization of CHIF in kidney and intestine. Am J Physiol Renal Physiol 2001;280:F505–F512. [PubMed: 11181413]
19. Wetzel RK, Sweadner KJ. Immunocytochemical localization of Na-K-ATPase alpha- and gamma-subunits in rat kidney. Am J Physiol Renal Physiol 2001;281:F531–F545. [PubMed: 11502602]
20. Morrison BW, Leder P. neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors. Oncogene 1994;9:3417–3426. [PubMed: 7970700]
21. Fu X, Kamps MP. E2a-Pbx1 induces aberrant expression of tissue-specific and developmentally regulated genes when expressed in NIH 3T3 fibroblasts. Mol Cell Biol 1997;17:1503–1512. [PubMed: 9032278]
22. Ino Y, Gotoh M, Sakamoto M, Tsukagoshi K, Hirohashi S. Dysadherin, a cancer-associated cell membrane glycoprotein, down-regulates E-cadherin and promotes metastasis. Proc Natl Acad Sci USA 2002;99:365–370. [PubMed: 11756660]
23. Kayed H, Kleeff J, Kolb A, Ketterer K, Keleg S, Felix K. FXYD3 is overexpressed in pancreatic ductal adenocarcinoma and influences pancreatic cancer cell growth. Int J Cancer 2005;118:43–54. [PubMed: 16003754]
24. Kadowaki K, Sugimoto K, Yamaguchi F, Song T, Watanabe Y, Singh K, Tokuda M. Phosphohippolin expression in the rat central nervous system. Brain Res Mol Brain Res 2004;125:105–112. [PubMed: 15193427]

25. Walaas O, Walaas E, Lystad E, Alertsen AR, Horn RS. The effect of insulin and guanosine nucleotides on protein phosphorylations by sarcolemma membranes from skeletal muscle. Mol Cell Endocrinol 1979;16:45–55. [PubMed: 227762]
26. Simmerman HK, Jones LR. Phospholamban: protein structure, mechanism of action, and role in cardiac function. Physiol Rev 1998;78:921–947. [PubMed: 9790566]
27. Meij IC, Koenderink JB, van Bokhoven H, Assink KF, Groenestege WT, de Pont JJ, et al. Dominant isolated renal magnesium loss is caused by misrouting of the $Na^{(+)},K^{(+)}$-ATPase gamma-subunit. Nature Genet 2000;26:265–266. [PubMed: 11062458]
28. Tanford C, Reynolds JA. Characterization of membrane proteins in detergent solutions. Biochim Biophys Acta 1976;457:133–170. [PubMed: 135582]
29. Jorgensen PL. Purification of Na+, K+-ATPase: enzyme sources, preparative problems, and preparation from mammalian kidney. Methods Enzymol 1988;156:29–43. [PubMed: 2835612]
30. Maunsbach AB, Skriver E, Jorgensen PL. Analysis of Na+,K+-ATPase by electron microscopy. Methods Enzymol 1988;156:430–441. [PubMed: 2835625]
31. Ivanov AV, Gable ME, Askari A. Interaction of SDS with Na^+/K^+-ATPase: SDSsolubilized enzyme retains partial structure and function. J Biol Chem 2004;279:29832–29840. [PubMed: 15123703]
32. Mesleh MF, Opella SJ. Dipolar waves as NMR maps of helices in proteins. J Magn Reson 2003;163:288–299. [PubMed: 12914844]
33. Wishart DS, Sykes BD. The ^{13}C chemicalshift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. J Biomol NMR 1994;4:171–180. [PubMed: 8019132]
34. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 1999;13:289–302. [PubMed: 10212987]
35. Bax A, Kontaxis G, Tjandra N. Dipolar couplings in macromolecular structure determination. Methods Enzymol 2001;339:127–174. [PubMed: 11462810]
36. Prestegard JH, Kishore AI. Partial alignment of biomolecules: an aid to NMR characterization. Curr Opin Chem Biol 2001;5:584–590. [PubMed: 11578934]
37. Lee S, Mesleh MF, Opella SJ. Structure and dynamics of a membrane protein in micelles from three solution NMR experiments. J Biomol NMR 2003;26:327–334. [PubMed: 12815259]
38. Mesleh MF, Veglia G, DeSilva TM, Marassi FM, Opella SJ. Dipolar waves as NMR maps of protein structure. J Am Chem Soc 2002;124:4206–4207. [PubMed: 11960438]
39. Valafar H, Prestegard JH. REDCAT: a residual dipolar coupling analysis tool. J Magn Reson 2004;167:228–241. [PubMed: 15040978]
40. Opella SJ, Marassi FM. Structure determination of membrane proteins by NMR spectroscopy. Chem Rev 2004;104:3587–3606. [PubMed: 15303829]
41. Franzin, CM.; Marassi, FM. NMR structure determination of proteins in bilayer lipid membranes: the FXYD family proteins. In: Tien, HT.; Ottova-Leitmannova, A., editors. Advances in Planar Lipid Bilayers and Liposomes. Vol. 2. Elsevier; Amsterdam: 2005. p. 77-93.
42. Gilbert W. Why genes in pieces? Nature 1978;271:501. [PubMed: 622185]
43. Patthy L. Intron-dependent evolution: preferred types of exons and introns. FEBS Letters 1987;214:1–7. [PubMed: 3552723]
44. Kolkman JA, Stemmer WP. Directed evolution of proteins by exon shuffling. Nature Biotechnol 2001;19:423–428. [PubMed: 11329010]
45. Sweadner KJ, Wetzel RK, Arystarkhova E. Genomic organization of the human FXYD2 gene encoding the gamma subunit of the Na, K-ATPase. Biochem Biophys Res Commun 2000;279:196–201. [PubMed: 11112438]
46. Bogaev RC, Jia LG, Kobayashi YM, Palmer CJ, Mounsey JP, Moorman JR, et al. Gene structure and expression of phospholemman in mouse. Gene 2001;271:69–79. [PubMed: 11410367]
47. Aizman R, Asher C, Fuzesi M, Latter H, Lonai P, Karlish SJ, Garty H. Generation and phenotypic analysis of CHIF knockout mice. Am J Physiol Renal Physiol 2002;283:F569–F577. [PubMed: 12167609]

48. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 2004;5:113. [PubMed: 15318951]
49. Clamp M, Cuff J, Searle SM, Barton GJ. The Jalview Java alignment editor. Bioinformatics 2004;20:426–427. [PubMed: 14960472]
50. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 1997;10:1–6. [PubMed: 9051728]
51. Crowell KJ, Franzin CM, Koltay A, Lee S, Lucchese AM, Snyder BC, Marassi FM. Expression and characterization of the FXYD ion transport regulators for NMR structural studies in lipid micelles and lipid bilayers. Biochim Biophys Acta 2003;1645:15–21. [PubMed: 12535606]
52. Thai K, Choi J, Franzin CM, Marassi FM. Bcl-XL as a fusion protein for the high-level expression of membrane-associated proteins. Protein Sci 2005;14:948–955. [PubMed: 15741345]
53. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–132. [PubMed: 7108955]
54. Cavanagh, J. Protein NMR Spectroscopy: Principles and Practice. Academic Press; San Diego: 1996.
55. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293. [PubMed: 8520220]
56. Goddard, TD.; Kneller, DG. SPARKY 3. University of California; San Francisco: 2004.
57. Mori S, Abeygunawardana C, Johnson MO, Vanzijl PCM. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. J Magn Reson B 1995;108:94–98. [PubMed: 7627436]
58. Ikura M, Kay LE, Bax A. A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of proteins: heteronuclear tripleresonance three-dimensional NMR spectroscopy. Application to calmodulin. Biochemistry 1990;29:4659–4667. [PubMed: 2372549]
59. Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog Nucl Magn Reson Spectrosc 1999;34:93–158.
60. Grzesiek S, Bax A. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. J Am Chem Soc 1992;114:6291–6293.
61. Farrow NA, Zhang O, Forman-Kay JD, Kay LE. A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical exchange rates of systems in slow equilibrium. J Biomol NMR 1994;4:727–734. [PubMed: 7919956]
62. Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S. Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. J Biomol NMR 2000;18:303–309. [PubMed: 11200524]
63. Ishii Y, Markus MA, Tycko R. Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. J Biomol NMR 2001;21:141–151. [PubMed: 11727977]
64. Chou JJ, Gaemers S, Howder B, Louis JM, Bax A. A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. J Biomol NMR 2001;21:377–382. [PubMed: 11824758]
65. Ottiger M, Delaglio F, Bax A. Measurement of *J* and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson 1998;131:373–378. [PubMed: 9571116]
66. Ding K, Gronenborn AM. Sensitivityenhanced 2D IPAP, TROSY-anti-TROSY, and E.COSY experiments: alternatives for measuring dipolar ^{15}N–^{1}HN couplings. J Magn Reson 2003;163:208–214. [PubMed: 12914836]
67. Mesleh MF, Lee S, Veglia G, Thiriot DS, Marassi FM, Opella SJ. Dipolar waves map the structure and topology of helices in membrane proteins. J Am Chem Soc 2003;125:8928–8935. [PubMed: 12862490]

Abbreviations used

NOE	nuclear Overhauser effect
HSQC	heteronuclear single quantum coherence
RDC	residual dipolar coupling
CSI	chemical shift index

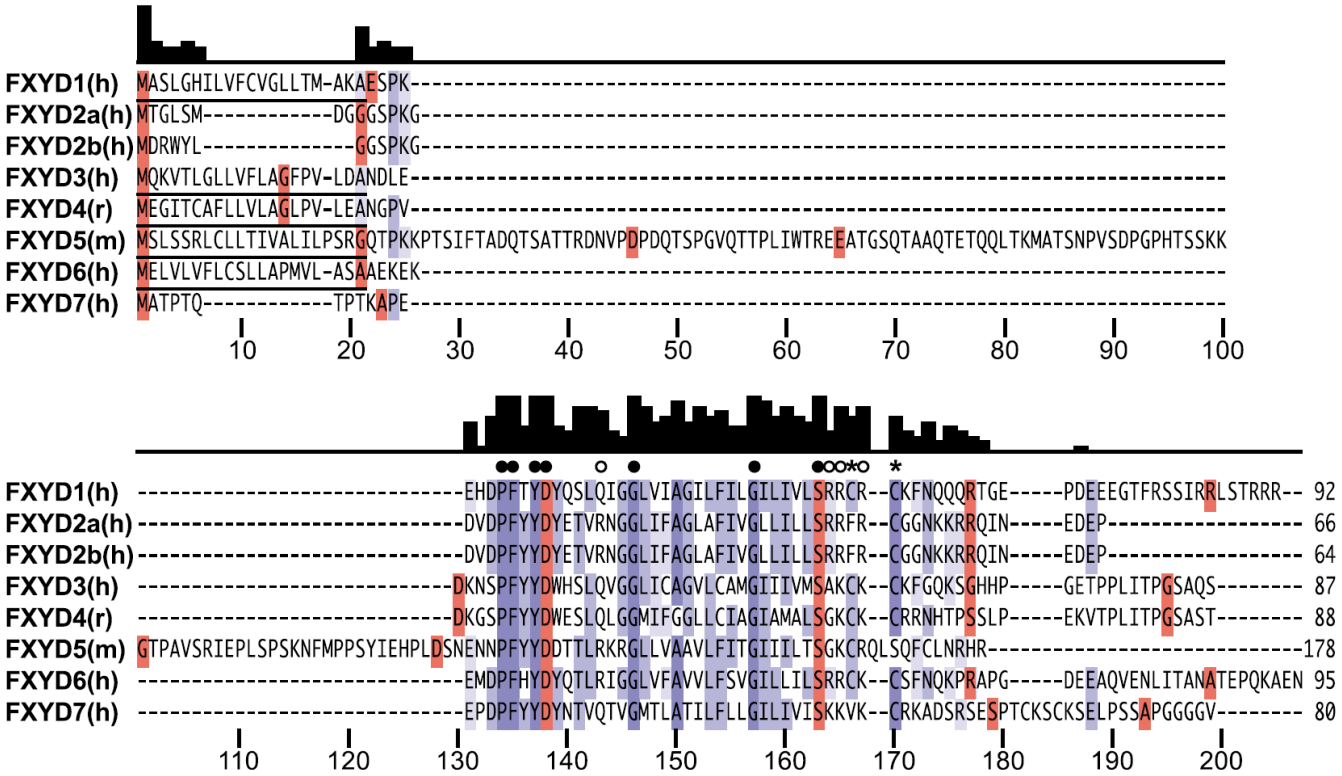

Figure 1.

Multiple sequence alignment of the seven mammalian FXYD family proteins. The alignment was produced with MUSCLE,[48] and rendered with Jalview.[49] Signal peptides, denoted by underlined sequence, were predicted with SignalP.[50] The intensity of blue color reflects percentage amino acid identity. Amino acid residues in red mark the starts of coding exons. Directly above the sequences, invariant residues are marked by filled circles, conserved positively charged residues by open circles, and conserved cysteine residues by asterisks. FXYD2 has two alternative splice variants (FXYD2a, FXYD2b). The bar graph above the alignment plots sequence conservation as a function of residue number. Amino acid numbers below the alignment serve as markers but do not reflect the actual residue numbers in the sequences. The NCBI gene and protein accession numbers are: NM_005031 and NP_068702 (human FXYD1); NM_001680 and NP_067614 (human FXYD2a); NM_021603 and NP_001671 (human FXYD2b); NM_005971 and NP_005962 (human FXYD3); NM_022388 and NP_071783 (rat FXYD4); NM_008761 and NP_032787 (mouse FXYD5); NM_022003 and NP_071286 (human FXYD6); NM_022006 and NP_071289 (human FXYD7).

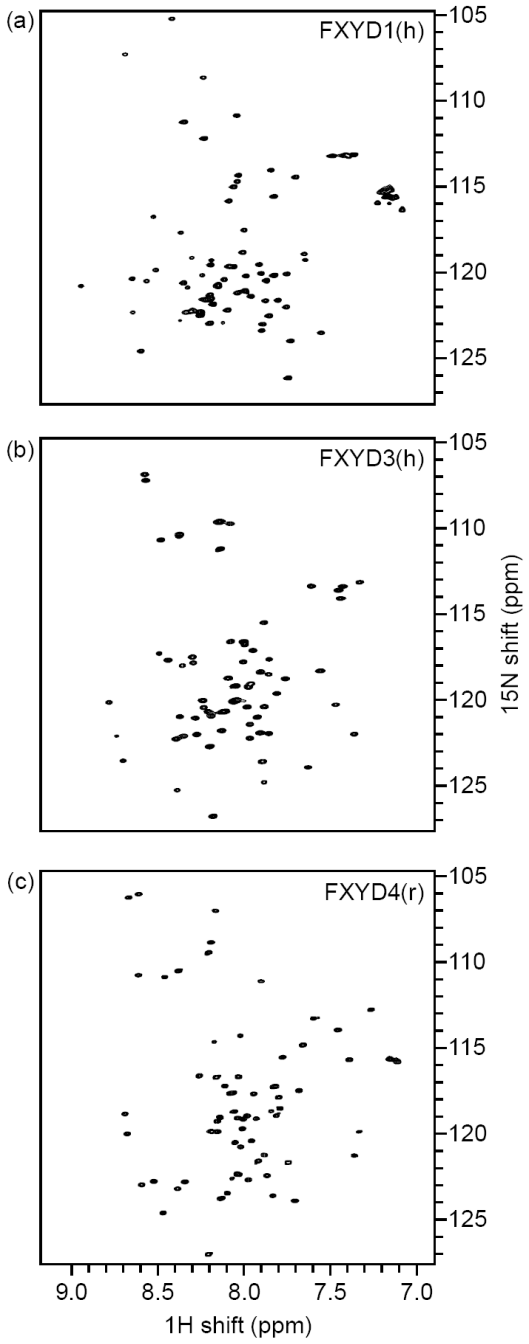

Figure 2.

Two-dimensional $^{1}H/^{15}N$ HSQC spectra of uniformly ^{15}N-labeled FXYD1, FXYD3, and FXYD4, in SDS micelles. Protein expression and purification were as described.[51,52] Isotopically ^{15}N-labeled amino acids and salts, $^{2}H_2O$, and ^{13}C-labeled glucose were from Cambridge Isotopes (www.isotope.com). The samples for solution NMR were prepared by dissolving each protein in SDS-containing buffer (500 mM SDS, 20 mM sodium phosphate (pH 5), 10 mM DTT, 1 mM sodium azide, in 90% H_2O, 10% $^{2}H_2O$).

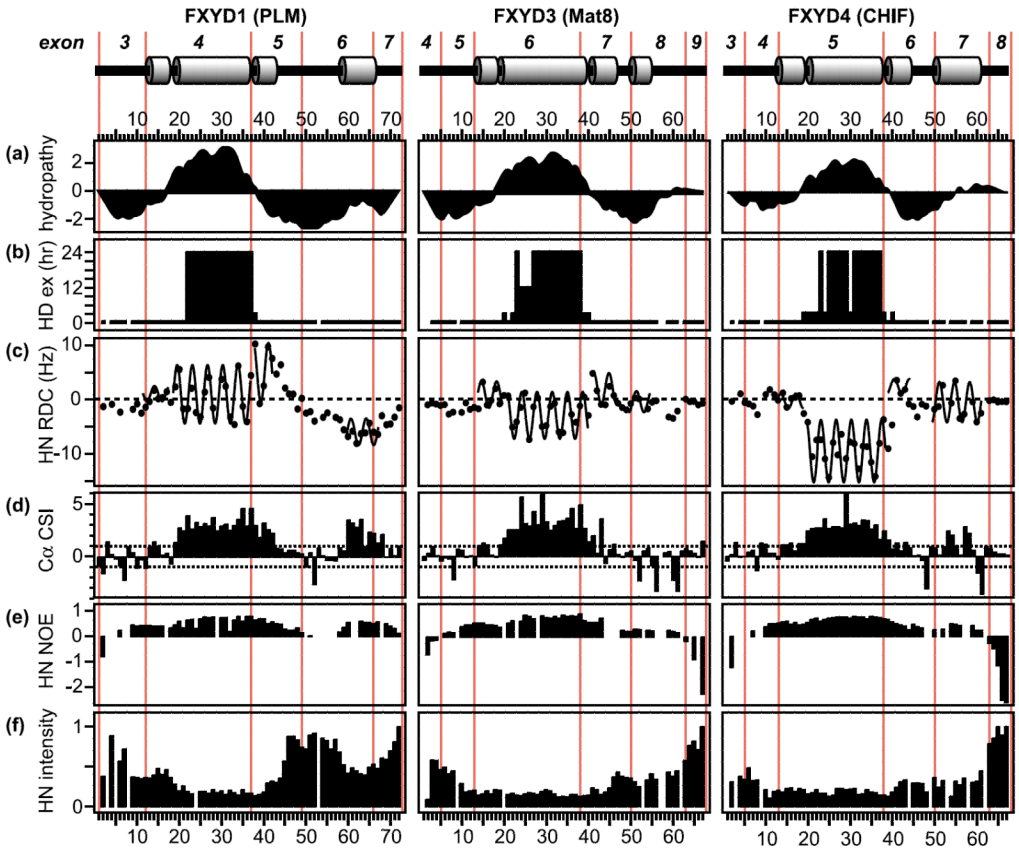

Figure 3.

Summary of NMR restraints, plotted as a function of residue number, for FXYD1, FXYD3, and FXYD4. The protein secondary structures are shown at the top of the Figure. Exon numbers are written above the structures, and the positions of intron–exon junctions are marked by red lines. The residue numbers for the three proteins begin after the signal sequences. (a) Kyte–Doolittle hydropathy plot.[53] (b) Amide hydrogen/deuterium exchange profiles, with exchange rates classified as rapid (less than 1 h), short (less than 3 h), medium (less than 12 h), or long (more than 24 h). (c) Values of $^{1}H–^{15}N$ RDCs with the characteristic periodicity in the α-helical regions of the proteins fitted sinusoids. (d) $^{13}C^{\alpha}$ chemical shift index. (e) $^{1}H/^{15}N$ heteronuclear NOEs. (f) Normalized $^{1}H/^{15}N$ HSQC peak intensities. Positions that are left blank correspond to proline residues, or overlapped resonances. NMR experiments were performed on a Bruker AVANCE 600 MHz spectrometer using a triple-resonance $^{1}H/^{13}C/^{15}N$ probe equipped with three-axis pulsed field gradients (www.bruker-biospin.com). All NMR experiments were performed at 40°C using a 1 s recycle delay. The chemical shifts are referenced to the $^{1}H_2O$ resonance, set to its expected position of 4.5999 ppm at 40°C.[54] The NMR data were processed using NMRPipe,[55] and the spectra were assigned and analyzed using SPARKY.[56] Chemical shifts were analyzed in terms of chemical shift indices[33] and with TALOS[34] to verify the protein residues in helical segments. The standard fHSQC experiment was used for isotropic samples with 1024 points in *t*2 and 256 in *t*1.[57] Backbone resonance assignments were made using a standard HNCA experiment with constant time evolution for ^{15}N, and solvent suppression was accomplished with a water flip-back pulse after the original $^{1}H–^{15}N$ magnetization transfer.[58-60] The spectra from selectively ^{15}N-labeled protein samples were necessary to resolve assignment ambiguities due to the extensive overlap among the C^{α} resonances that is typical of helical membrane proteins in micelles. $^{1}H–^{15}N$ heteronuclear

NOE measurements were made using difference experiments with and without 3 s of saturation of the ^{1}H resonances between scans.[61] Hydrogen exchange experiments were performed by dissolving the lyophilized protein in SDS buffer with 100% $^{2}H_2O$, and then acquiring HSQC spectra at 0.5 h intervals. The HSQC spectra of FXYD1 in the presence of $MnCl_2$ were obtained with 0.8 mM uniformly ^{15}N-labeled FXYD1, after adding $MnCl_2$, dissolved in the SDS-containing NMR buffer, to final concentrations of 0.5 mM, 0.8 mM, and 1.6 mM. For the measurement of RDCs, the protein-SDS micelles were weakly aligned in 7% polyacrylamide gels, using either vertical compression[62,63] or expansion.[64] The $^{1}H–^{15}N$ RDCs were measured using a sensitivity-enhanced $^{1}H–^{15}N$IPAP experiment modified for suppression of the NH_2 signals from the acrylamide in the gel.[63,65,66] The contribution to the residual dipolar coupling splitting from the isotropic scalar coupling was determined by performing the same experiment on an isotropic micelle sample, and subtracting the value of the isotropic *J*-coupling obtained from that measured for the weakly aligned gel sample. The RDCs were analyzed using MATLAB scripts as described.[32,38,67] Helical regions were identified by applying a sliding window algorithm to fit the experimental RDCs. The RDCs within a five-residue window were fit to a sinusoid of periodicity 3.6, and the RMSD between the sinusoid and the data were plotted as a function of residue number. Continuous stretches of amino acids with low RMSD (less than the experimental error of 1.5 Hz) were identified as helices and fitted to a single sinusoid. Higher RMSDs were generally interpreted as deviations from ideality, including kinks, curvature, and loops. This analysis relating the orientation of the helix to the amplitude, average value, and phase of the sinusoid is an initial step toward structure determination, and can determine the relative orientations of helices to within four degenerate solutions.[38,39]

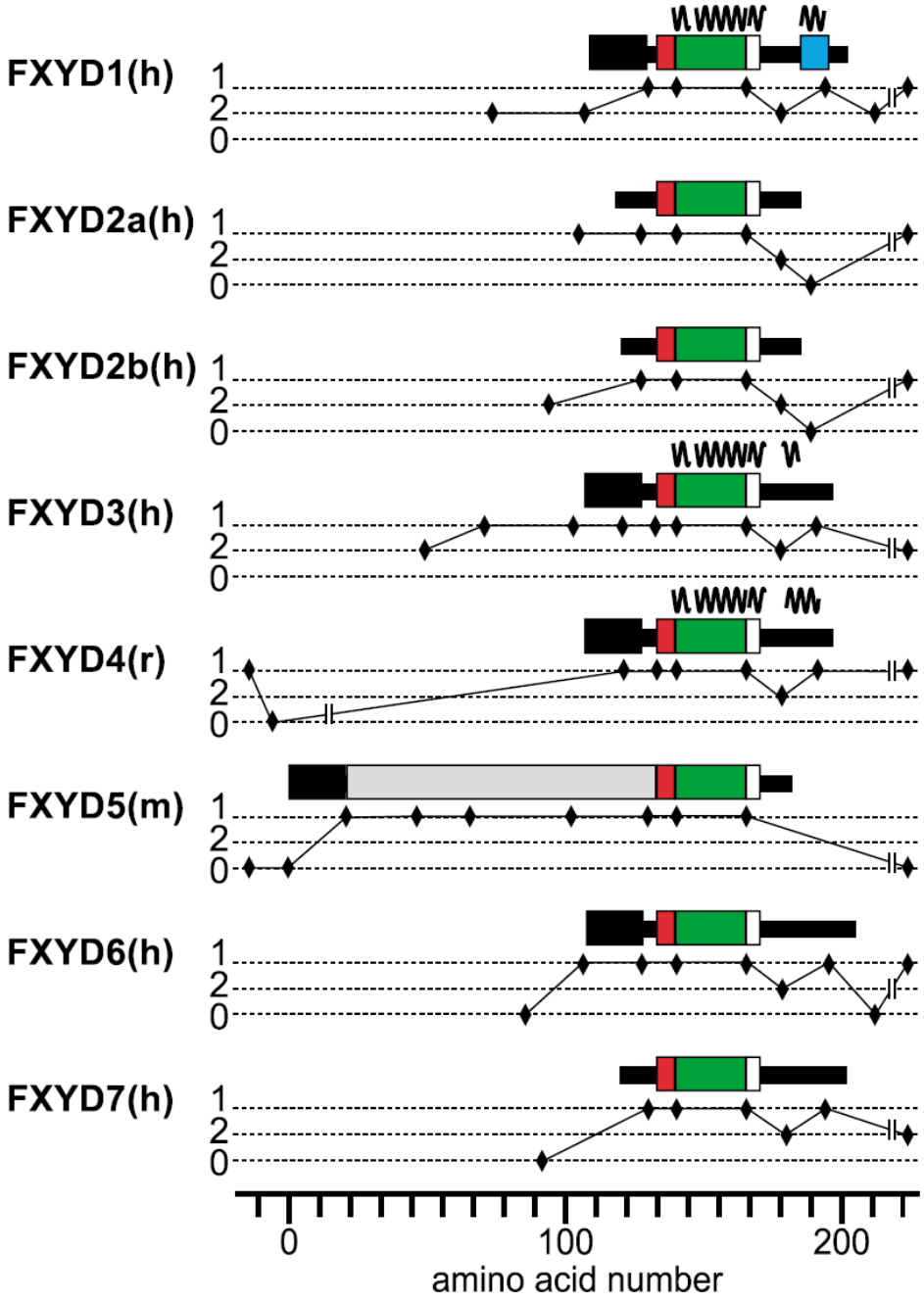

Figure 4.

Splice frame diagrams and organization of structured modules in the seven mammalian FXYD family proteins. Protein sequences are designated by thick black lines or colored boxes. The FXYD signature motif is in red, the transmembrane domain in green, the stop-transfer signal in white, and the predicted signal sequences in black. In FXYD1, the phosphorylation domain is in blue. In FXYD5, regions with homology to O-linked glycosylated proteins are in gray. Helical regions, specific to the proteins FXYD1, FXYD3, and FXYD4, are marked by sinusoids. Exons are designated by thin black lines below each protein, and the positions of intron–exon junctions are marked by diamonds. The vertical axis indicates the phase of the intervening introns, which depends on the intron position

relative to the reading frame of the gene. Introns can interrupt the gene reading frame between consecutive codons (phase 0), between the first and second nucleotide of a codon (phase 1), or between the second and third nucleotide of a codon (phase 2). The gene data[1,45-47] were obtained from the NCBI (National Center for Biotechnology Information; www.ncbi.nlm.nih.gov). The gene and protein accession numbers are listed in the legend to Figure 1. The sizes of domains and structured modules are to scale. Amino acid numbers below the alignment serve as markers but do not reflect the actual residue numbers in the sequences.