# Proto-*vav* and gene expression

SIR — Two recent papers[1,2] describe the *vav* proto-oncogene product as containing 'leucine-zipper' (LZ) and 'zinc-finger' (ZF) motifs, two nuclear localization signals and a 'helix–loop–helix' (HLH) domain. We were surprised to find, however, that none of these transcription-factor and DNA-binding motifs attributed to proto-*vav* is detectable when objective or quantitative criteria are applied.

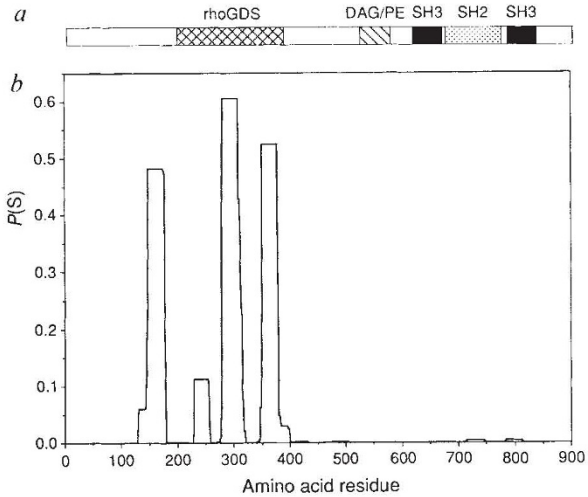
We examined the proto-*vav* amino-acid sequence as follows. Initial database searches (see figure legend) confirmed the presence of two SH3 and one SH2 domains but failed to provide evidence of sequence similarity to any ZF, HLH or LZ proteins. In contrast, the relationships between *max*, *myc* and other HLH/LZ proteins were easily detectable and highly significant ($P<10^{-7}$) under the same search conditions. Considering perhaps that similarities between proto-*vav* and other HLH/LZ and ZF sequences might be obscured by the more significant search results, we performed subsequent analyses with an edited query sequence encompassing residues 1–613 of proto-*vav* but excluding the C–terminal SH3–SH2–SH3 domains. A search of the general-purpose sequence databases revealed significant local similarities between proto-*vav* (residues 513–561) and the diacylglycerol/phorbol ester (DAG/PE) binding domains of protein kinase C — a fact recognized by others (for example, ref. 3) and documented in the PROSITE database (see below). Proto-*vav* also shows some borderline similarities to the *dbl* oncogene product[4] and some insignificant local similarities to intermediate-filament proteins (see below).

Directed searches of the Zinc Finger[5] and Transcription Factor[6] databases provided no evidence for even marginal similarity between *vav* and any of the 1,313 ZF domains and 811 transcription-factor domains present in these collections. Finally, the complete proto-*vav* sequence was scanned for the presence of any of the 605 motifs in the PROSITE database[7]. Proto-*vav* does not meet the criteria for nuclear localization signals and the 'zinc-finger' region corresponds to the DAG/PE-binding motif mentioned above (PROSITE entry PS00479). Statistical methods confirmed the presence of an acidic domain (residues 130–174) in proto-*vav*, but charge clusters associated with DNA-binding proteins and transcription factors are usually basic in nature[8].

Adams *et al.*[9] have just published a revision of the mouse proto-*vav* sequence affecting 32 codons specifying residues 325–355 of the protein. When this revised sequence is used in database searching as described above, the results are essentially the same as before with the profound exception that the sequence now shows extremely significant ($P\sim 2\times10^{-10}$) similarity to a family of proteins (including human Dbl, Bcr and yeast Cdc24) that act as guanine nucleotide dissociation stimulators for the *rho/rac* family of *ras*-like small GTPases (see ref. 10). This strongly confirms the results of Adams *et al.*[9] and supports their conclusion that *vav* functions in a signal transduction pathway involved in cytoskeletal organization.

Because of the local similarities between proto-*vav* and some intermediate-filament proteins, we determined the propensity of the proto-*vav* sequence to adopt a coiled-coil conformation. We found three regions (residues 148–178, 281–311, 351–379) with coiled-coil probabilities of about 50% (*b* in the figure). Note that these regions do not include the putative leucine-zipper motif (residues 70–91). A summary of some objective and quantitative sequence attributes of proto-*vav* is shown in *a* in the figure. None of our findings supports the hypothesis[1,2] that proto-*vav* might have DNA-binding or transcription-factor activities.

**Mark S. Boguski**
*National Center for Biotechnology Information,*
*National Library of Medicine,*
*National Institutes of Health,*
*Bethesda,*
*Maryland 20894, USA*
**Amos Bairoch**
*Department of Medical Biochemistry,*
*University of Geneva,*
*1211 Geneva 4, Switzerland*
**T. K. Attwood**
*Department of Biochemistry and Molecular Biology,*
*University of Leeds,*
*Leeds LS2 9JT, UK*
**George S. Michaels**
*Division of Computer Research and Technology,*
*National Institutes of Health,*
*Bethesda,*
*Maryland 20892, USA*




Computable sequence attributes of the *vav* proto-oncogene product. The complete human proto–*vav* sequence was constructed by merging 5′ data[11] with the SWISS-PROT entry for *vav* (accession no. p15498), excluding the Tn*5*-derived residues at the amino terminus of the latter sequence. Proto-*vav* was then searched against a merged, non-redundant collection of sequences derived from PIR 31.0, SWISS–PROT 21.0 and translated GenBank 71.0 using the BLASTP program[12] with default parameters. Searches against the Zinc-Finger and Transcription-Factor databases were conducted with FASTA[13] and BLASTP[12], respectively. This analysis was repeated with the revised mouse proto-*vav* sequence[3,9]. *a*, Schematic summary of significant search results as described in the text. The sizes and locations of the rhoGDS, DAG/PE and SH3, SH2 and SH3 domains are drawn to scale and correspond to residues 198–385, 516–564, 617–665, 671–665 and 789–837, respectively, in the mouse proto-*vav* sequence. *b*, The proto-*vav* sequence was also analysed for the presence of potential coiled coil regions using the method of Lupas *et al.*[14] with a window of 28 residues. The probabilities of segments forming coiled coils are plotted as a function of their location in the linear sequence. The regions showing some coiled-coil potential correspond to the rhoGDS domain (previously described as the leucine-rich domain[3]) but do not include the putative leucine-zipper region. *P*(S) is the probability of forming a coiled coil of score S defined in ref. 14.

1. Bustelo, X. R., Ledbetter, J. A. & Barbacid, M. *Nature* **356**, 69–71 (1992).
2. Margolis, B. *et al. Nature* **356**, 71–74 (1992).
3. Coppola, J., Bryant, S., Koda, T., Conway, D. & Barbacid, M. *Cell Growth Diff.* **2**, 95–105 (1991).
4. Galland, F., Katzav, S. & Birnbaum, D. *Oncogene* **7**, 585–587 (1992).
5. Jacobs, G. & Michaels, G. *New Biol.* **2**, 583–584 (1990).
6. Ghosh, D. *Trends biochem. Sci.* **16**, 445–447 (1991).
7. Bairoch, A. *Nucleic Acids Res.* **19**, 2241–2245 (1991).
8. Brendel, V. & Karlin, S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5698–5702 (1989).
9. Adams, J. M., Houston, H., Allen, J., Lints, T. & Harvey, R. *Oncogene* **7**, 611–618 (1992).
10. Boguski, M. S. & McCormick, F. *Neuron* (in the press).
11. Katzav, S., Martin-Zanca, D. & Barbacid, M. *EMBO J.* **8**, 2283–2290 (1989).
12. Atlschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. molec. Biol.* **215**, 403–410 (1990).
13. Pearson, W.R. & Lipman, D.J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444–2448 (1988).
14. Lupas, A., Van Dyke, M. & Stock, J. *Science* **252**, 1162–1164 (1991).