Protein Science (1995), *4*:1851–1859. Cambridge University Press. Printed in the USA.

Evidence from sequence information that the interleukin-1 receptor is a transmembrane GTPase

THOMAS P. HOPP

Protein Research Laboratories, Inc., San Diego, California 92126

(RECEIVED April 11, 1995; ACCEPTED June 15, 1995)

Abstract

Evidence is presented that the cytoplasmic domain of the type I interleukin-1 receptor (IL-1R) may be a GTPase. This domain conserves segments of hydrophobic amino acids that suggest a structural relatedness to the ras proto-oncogene protein and other members of the GTPase superfamily, despite a lack of significant detectable sequence homology. When the hydrophobic segments of the IL-1R were aligned with similar segments of the GTPases, it became apparent that the IL-1Rs possess a number of conserved amino acids that represent plausible functional residues for base-specific binding of GTP, magnesium chelation, and phosphate ester hydrolysis. Furthermore, a segment of five contiguous residues was found that is identical between ras and the IL-1R, and which is positioned to form part of the guanine base binding pocket. If this model is correct, then the IL-1Rs possess a highly conserved effector protein binding region, but one that is entirely unrelated to the effector regions of other superfamily members. Therefore, if the IL-1R is indeed a GTPase, then its activation function may be directed to as-yet unrecognized effector target proteins, as part of a unique cellular signal transduction pathway.

Keywords: GTPase; homology; interleukin-1 receptor; signaling

The type I interleukin-1 receptor (IL-1R) is remarkable in that its extracellular IL-1 binding region belongs to the immunoglobulin (Ig) superfamily, whereas its cytoplasmic region has not been matched to any of the structural groups known to exist among other receptors, whether they be Ig superfamily members or not. Among Ig superfamily receptors, the cytoplasmic, non-Ig-like regions have been identified as being members of a diverse collection of functional types. These include protein tyrosine kinases, for example, the cytoplasmic portions of the platelet-derived growth factor receptor, the fibroblast growth factor receptor, the colony stimulating factor 1 receptor, and others (Williams & Barclay, 1988). Furthermore, this group appears to be extremely ancient, as is evidenced by the recent discovery of Ig/tyrosine kinase receptors in animals as primitive as sponges (Schaecke et al., 1994). Many Ig superfamily receptors do not contain tyrosine kinase domains, however, and several have been found to possess cytoplasmic domains belonging to the family of protein tyrosine phosphatases (Krueger & Saito, 1992) or to possess short cytoplasmic tails that do not exhibit intrinsic activities but which are required for binding to other cytoplasmic functional proteins such as α-actinin (Carpen et al., 1992), actin (Letourneau & Shattuck, 1989), and spectrin (Pollenberg et al., 1987). Other Ig superfamily members, such as the membrane forms of Igs themselves, possess cytoplasmic domains of only one or several amino acids and rely on other transmembrane proteins for signal transduction (Keegan & Paul, 1992).

From the foregoing, it is clear that the IL-1R belongs to a heterogeneous family that diversified through a process of domain shuffling (Patthy, 1985) to yield a variety of cytoplasmic functionalities all of which are attached to Ig superfamily receptor domains. The converse appears to be true as well, because the IL-1R is also a member of a small family of receptors that share homology in their cytoplasmic domains. The *Drosophila* developmental receptor Toll has been shown to possess a cytoplasmic domain that is similar to that of the IL-1R, but an extracellular portion that is a member of the leucine-rich repeat family, which is entirely unrelated to the Ig superfamily (Kobe & Deisenhofer, 1993). In addition, a mammalian nonreceptor, cytoplasmic, myeloid-differentiation-mediating protein, MyD88, is a further homologue of the IL-1R cytoplasmic domain (Hultmark, 1994). Thus, the IL-1R shares superfamily relationships both in its extracellular and intracellular portions, albeit to entirely different groups of proteins. These complex evolutionary relationships place the IL-1R within a spectrum of molecules whose functions have evolved through a process of mixing and matching functional domains to achieve a variety of activities. Given the wide range of domain shuffling already observed among these receptors, it would not be surprising to find that the IL-1R possesses a cytoplasmic signaling domain with a functionality entirely different from the ones already identified.

Reprint requests to: Thomas P. Hopp, Protein Research Laboratories, Inc., 10606-8 Camino Ruiz, Suite 281, San Diego, California 92126; e-mail: tomhopp@aol.com.

The search for a signaling function for the IL-1R is confounded by difficulty in deciphering the complex mixture of signals sent after IL-1 binds to its receptor. Upon IL-1 binding, cells undergo a general stimulation that may include, among other effects, the augmentation of cell proliferation and the initiation or suppression of gene expression for several proteins. Cells such as osteoblasts, fibroblasts, mesangial cells, thymic epithelial cells, and platelet stem cells respond to IL-1 by proliferating, whereas other cell types such as hepatocytes, chondrocytes, endothelial cells, and adrenal cortex cells do not directly proliferate because of IL-1 but rather undergo profound changes in cellular activity and secretory behavior. These changes are mediated by a number of cytoplasmic secondary signaling molecules, and at least two nuclear factors that have been shown to be IL-1 inducible: NF-kB (Shirakawa & Mizel, 1989) and AP-1 (Muegge et al., 1989). Controversy has arisen recently regarding IL-1 signal transduction mechanisms, partly because of the heterogeneity of the responses between different cell types, and partly because of the accumulation of reports regarding different signaling processes that may or may not be influenced by IL-1 in different cell types (reviewed in Dinarello, 1991). Against this formidable background, several groups have made progress in beginning to unravel the signaling functions of the IL-1R. Experiments wherein full-length IL-1R or truncated forms with part or all of the cytoplasmic domain deleted were expressed in cells demonstrated several key features of the IL-1R. (1) The full IL-1R is capable of introducing IL-1 responsiveness to cells without the need for other co-transfected proteins (Curtis et al., 1989), indicating that a fully functional IL-1R is contained within the single IL-1R chain. (2) Deletion of part or all of the cytoplasmic domain blocks IL-1-dependent bioactivity, confirming that the cytoplasmic domain has a signaling function. Interestingly, several inactivating truncations did not eliminate IL-1/IL-1R translocation to the nucleus, suggesting that nuclear translocation is not necessary for at least part of the signaling mechanism.

Several recent reports have indicated that cell membranes bearing the IL-1R bind to GTP analogs in an IL-1-dependent manner, but that the toxin susceptibility patterns of the IL-1-activated GTPase do not match any known G protein (Saklatvala et al., 1990). Papers have implicated G protein-mediated activation of adenyl cyclase (Mizel, 1990), activation of a novel protein-serine kinase (Saklatvala et al., 1990), and evidence also has been offered that the IL-1R may be coupled to a G protein that activates phospholipase C (O'Neill et al., 1990) or D (Bursten & Harris, 1994) to produce diacylglycerols or phosphatidic acids as further messengers. In these experiments, it was found that IL-1R-mediated activation was extremely rapid, triggering measurable GTP binding within 1 min (O'Neill et al., 1990). However, no evidence was provided that directly implicated the IL-1R in this GTP binding activity. Despite a wide variety of experimental approaches, no clear picture of the signal transduction mechanism has yet emerged (Dinarello, 1991; Kopp & Ghosh, 1995).

The known IL-1R family has grown to four members (human, rat, mouse, and chicken), and within this group the cytoplasmic domain retains considerable sequence identity, suggesting that a conserved catalytic or binding function may exist. Despite the lack of apparent homology to cytoplasmic domains of known function, the IL-1R cytoplasmic domain possesses features such as segments of hydrophobic amino acids, which suggest that it may fold into a globular form (Hopp, 1989). This observation, along with the presence of a number of conserved charged and hydrophilic amino acids, suggests that it may possess an enzymatic function or play some role other than just as a cytoplasmic tail.

In light of the background presented above, I sought to draw comparisons of the IL-1R sequence with other proteins involved in cellular signal transduction, in order to attempt to identify a signaling mechanism for IL-1. Several observations regarding the cytoplasmic domain of the type I IL-1R suggested a possible structural relationship to members of the GTPase superfamily. (1) The calculated molecular weight of the IL-1R cytoplasmic domain is 24,603 Da, making it similar in size to the ras-type GTPases, which have molecular weights of around 21,000 Da. (2) In the cytoplasmic domain sequence, there exist several stretches of three to six contiguous hydrophobic and β-branched amino acids (V, I, L, T, F) that correlate with deep hydrophobic valleys of hydrophilicity profiles. These are reminiscent of the core β-strands that traverse the interiors of the GTPases. (3) The amino acid composition of the cytoplasmic domain has the characteristic amino acid distribution for an α/β-type folded structure (Nakashima et al., 1986), again similar to most GTPases. (4) Secondary structure predictions (Chou & Fasman, 1978; Garnier et al., 1978) suggest a mixture of α-helical and β-stranded secondary structures, some of which might correspond to those seen in the crystallographically determined 3D structures of all GTPases described to date. (5) Several strong amphipathic helix signals are detectable in the IL-1R cytoplasmic domain sequence, as is also true for the GTPases.

The observations listed above prompted a comparison of the sequences of different GTPases with the cytoplasmic domain of IL-1R. When the sequences were aligned and gaps were inserted by roughly matching the largest hydrophobic regions between IL-1R and the GTPases, it became apparent that many of the critical residues necessary for GTP binding and hydrolysis were conserved between IL-1R and the GTPases. This report describes the sequence alignment process, discusses the conserved residues and their potential roles in GTP binding and hydrolysis, and considers the implications of the proposed receptor GTPase activity in the pathway of IL-1 signaling.

Results

Efforts to detect sequence homology

All efforts carried out initially to detect significant sequence homology between IL-1Rs and GTPases were unsuccessful. When whole IL-1R cytoplasmic domains were run against the entire ENTREZ sequence database, the only statistically significant matches that were found were with the IL-1Rs themselves and the closely related ST2L protein. In all other cases, the number of identical residues was small enough that random matches were likely to be occurring, leading to unacceptably low alignment scores (Needleman & Wunsch, 1970). Subsequent to the alignment process, which uncovered the sequence relationships between the IL-1R and the GTPases (described below), homology searching was revisited, this time using a segment alignment approach (Takio et al., 1984). Homology was sought between the GTPases used in the alignment (below) and portions of the IL-1Rs, after breaking the receptor cytoplasmic domains down into four consecutive 50-residue segments for comparisons

(omitting the C-terminal 13–17 residues, which lie outside the putative GTPase domain). This led to several matches with ras and other proteins that are consistent with the alignments but, just as often, led to irrelevant matches that have no relationship to the alignments ultimately chosen. The only segment that consistently lined up with the same region of the GTPases was the fourth 50-residue segment of the IL-1R, which contains the five-contiguous-residue match, SAKTR, with members of the ras subgroup.

Matching hydrophobic regions

Figure 1 shows the technique used to align the IL-1R cytoplasmic region with the ras protein sequence. As has been demonstrated before, the deepest valley regions of Hopp and Woods hydrophilicity profiles correlate with those secondary structure elements that are packed into the cores of globular proteins (Hopp, 1989). In order to determine a possible structural alignment of ras with the IL-1R cytoplasmic region in the absence of demonstrable sequence homology, it was necessary to search for possible matching core secondary structure elements based on the hydrophilicity profiles alone. This was done by first aligning the two sequences at the points indicated by the filled arrow in Figure 1. As can be seen, this area represents the deepest valley region for both profiles and was therefore assumed to be a common structural element, in this case, corresponding to one of the central core β-strands known to exist in ras. Proceeding outward from this point, other valleys were aligned by allowing gaps in the ras sequence where necessary. As can be seen in Figure 1, this process allowed a rough alignment of the two proteins based only on the hydrophilicity profiles. Where appropriate, each aligned sequence segment was then shifted one or several residues to either side in order to line up any identical residues. During this process, the five-residue identity with ras, SAKTR, near the C-terminus was first identified and matched.

Superfamily member alignments

In order to refine the IL-1R/ras alignment and produce a sequence alignment for the other superfamily members (Fig. 2), the following alignments were taken from the literature: transducin α-chain (Gt) aligned with elongation factor-TU (EFTU) and ras (Lambright et al., 1994); Gt with ADP ribosylation factor-1 (ARF1), ras, rho, and ral (Sewell & Kahn, 1988); ras with ras 1, ral A, and rad (Cohen et al., 1994); ADP ribosylation factor domain-containing protein-1 (ARD1) and GTP binding protein (GBP) with ARF1 and ARF6 (Mishima et al., 1993); IL-1R with Toll and MyD88 (Hultmark, 1994); and ST2L with IL-1R (Yanagisawa et al., 1993). These published alignments were preserved as the overall alignment in Figure 2 was produced but were adjusted by adding gaps where necessary (usually where gaps existed in the other alignments) or enlarging gaps, to produce the alignment shown. The alignment between the IL-1R/ST2L/Toll group, on the one hand, and the GTPase group, on the other hand, was accomplished by matching each group with the previously aligned sequences of the IL-1R and ras. The initial alignment achieved by these steps was then refined manually to increase the number of identical residue matches without adding additional gaps (for instance, by moving residues from one side of a gap to the other).

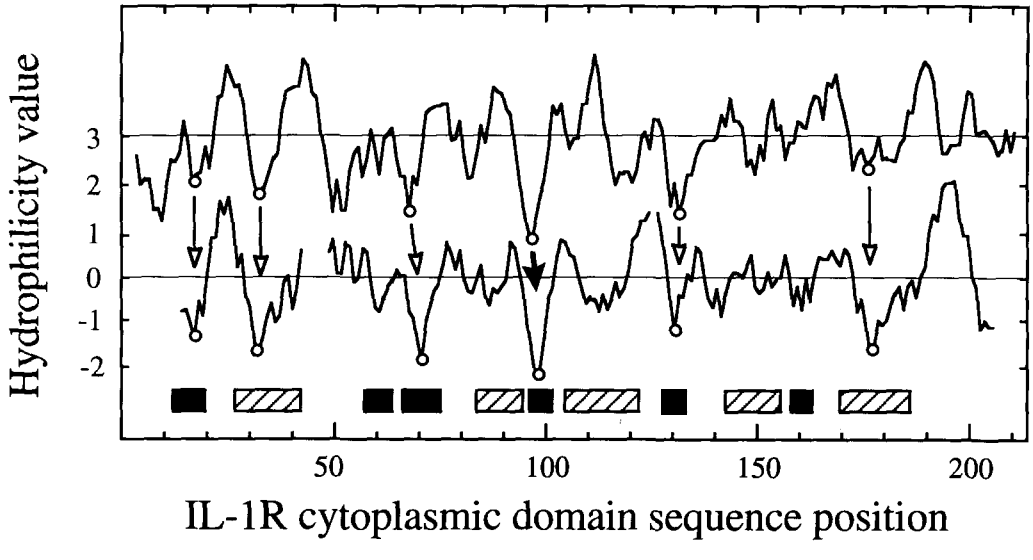

Fig. 1. Hydrophilicity profile comparison between ras and IL-1R cytoplasmic domain. A profile was generated for each protein by the hydrophilicity plotting method of Hopp and Woods (Hopp & Woods, 1981; Hopp, 1989) using the amino acid "acrophilicity" scale (Hopp, 1985). For the purposes of this display, the profile for ras was kept at its original position, centered near the zero point of the y axis (horizontal line), whereas the profile for the human IL-1R was elevated by 3 hydrophilicity units, so that it is arbitrarily centered near the $+3$ y-axis value. This facilitates comparison of the hydrophobic valley regions of the profiles. Bars beneath the lower profile indicate the known secondary structures of ras: gray bars, β-strands; hatched bars, α-helices. Note that all of the deepest valleys correspond to individual secondary structure elements in ras. Circles on the profiles indicate the valley bottom positions chosen for use in matching the IL-1R sequence to the ras sequence. The pair of valleys joined by the filled arrow were matched first, then the other valley pairs were matched, working outward from this point. Gaps were opened in the ras profile to facilitate the production of this figure.

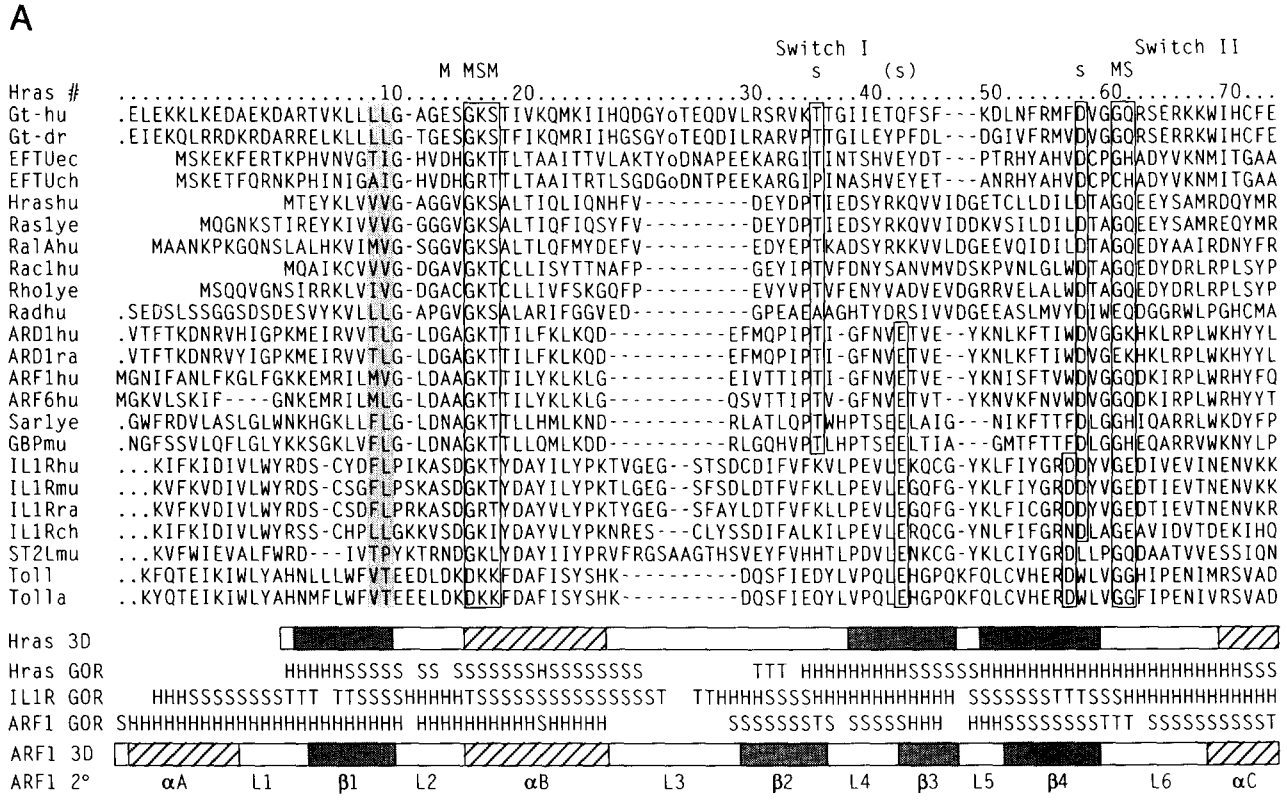

Fig. 2. Sequence alignment of the IL-1R/GTPase superfamily. Sequences were aligned in subgroups and then combined into a single alignment as described in the text. The key at the end of the figure explains the letters and symbols used. Important conserved functional residues are boxed. Conserved hydrophobic clusters corresponding to the core β-strands are shaded. The IL-1R sequences begin with the first lysine residue on the cytoplasmic side of the membrane-spanning region and continue through the C-terminus. Identities of proteins are indicated at the left: Gt-hu, human transducin α-chain; Gt-dr, *Drosophila melanogaster* transducin α-chain; EFTUec, *Escherichia coli* elongation factor TU; EFTUch, *Chlamydia trachomatis* elongation factor TU; Hrashu, human H-ras; Ras1ye, *Saccharomyces cerevisiae* ras 1; RalAhu, human ral A; Rac1hu, human rac 1; Rho1ye, *S. cerevisiae* rho 1; Radhu, human rad; ARD1hu, human ARF domain-containing protein; ARD1ra, rat ARF domain-containing protein; ARF1hu, human ADP ribosylation factor 1; ARF6hu, human ADP ribosylation factor 6; Sar1ye, *S. cerevisiae* sar 1; GBPmu, mouse GTP binding protein; IL1Rhu, -mu, -ra, -ch, human, mouse, rat, and chicken IL-1R; ST2Lmu, mouse ste2 gene long product; Toll, *D. melanogaster* Toll protein; Tolla, *D. pseudoobscura* Toll protein. The MyD88 protein was omitted because no single clear alignment could be made in the N-terminal region. (*Continues on facing page.*)

Homology

After the alignment was completed, the percent identity with the IL-1R was calculated for each protein, relative to the total number of residues of that protein used in the alignment. Segments that had been omitted from the alignment were not counted when determining the total residues for percentages. These omitted segments included the C-terminal tails of ras and related proteins, the large multihelix insertion found in transducin chains, and an extended N-terminus in MyD88. Evidence of homology within the IL-1R/ST2L/Toll group was substantial, as expected, and mouse, rat, and chicken receptors showed 77, 74, and 59% identity in their cytoplasmic domains. ST2L retains 33% identity with IL-1R, whereas Toll and MyD88 are only distantly related, with 21 and 16% identity to IL-1R, although their homology to each other is more evident. The argument for homology between the IL-1R and the GTPases, on the other hand, ranges from marginal in most cases to insignificant in others. The highest identity seen was 15% (for both ARF6 and GBP), whereas other values ranged down to a low of 8%, for EFTU. Homology in this range is not normally considered significant (Needleman & Wunsch, 1970).

Secondary structures

At the bottom of Figure 2, both predicted and known secondary structures of GTPases are shown. The crystallographically determined α-helices and β-strands of ras and ARF1 are displayed on the bars beneath the sequences. A comparison of these to each other and to the predictions displayed between them demonstrates the similarities as well as variations between the known and predicted secondary structures. The predicted and known secondary structures for ras and ARF1 agree in 50–60% of the positions, as is expected from previous assessments of prediction accuracy (Chou & Fasman, 1978; Garnier et al., 1978), and it is reasonable to assume that such a relationship is true for the IL-1R predictions as well. Because of this limited accuracy, no attempt is made in this report to specify the extent of the secondary structures of the IL-1Rs. It is worth noting, however,

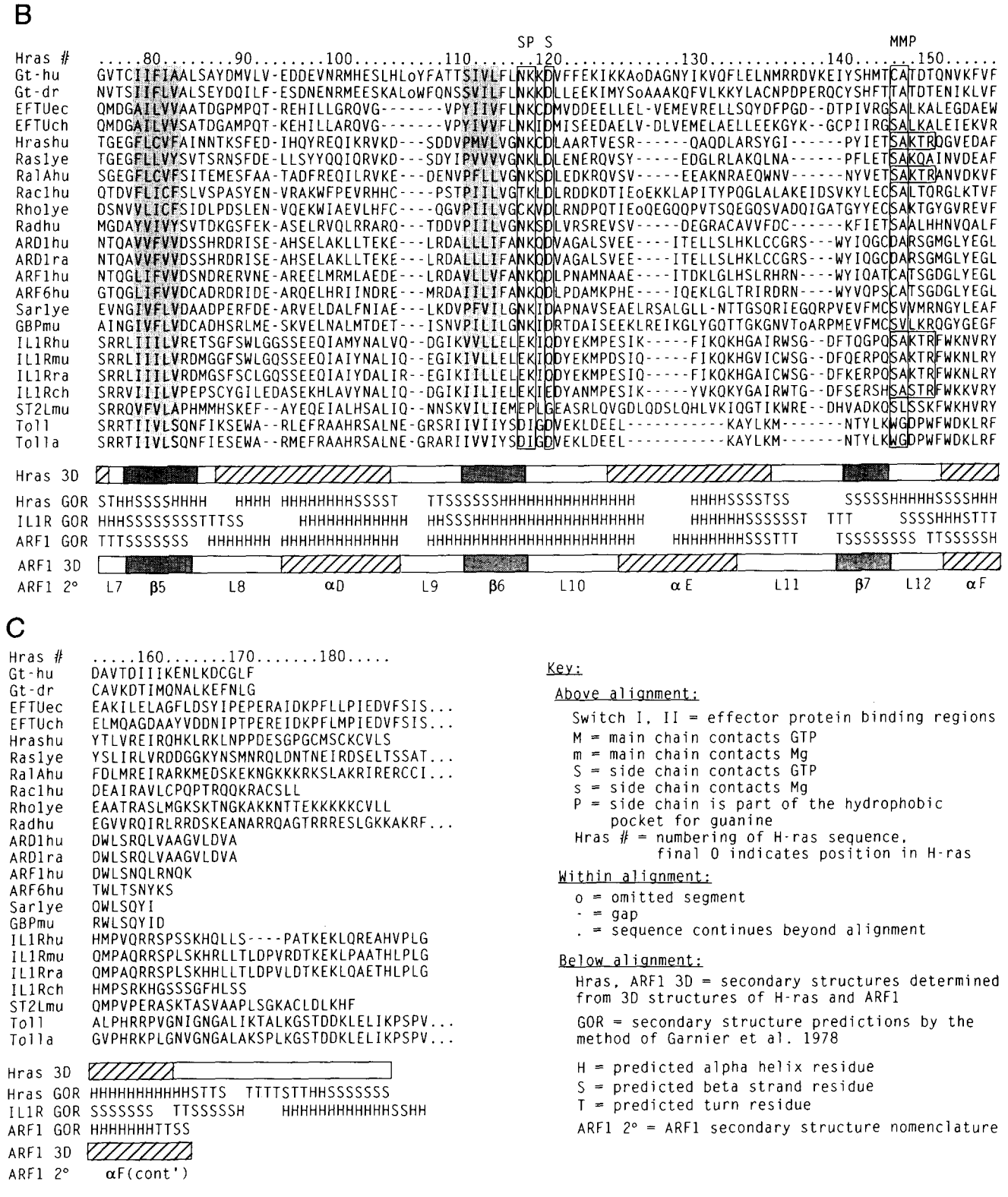

Fig. 2. *Continued.*

that the three clearly conserved hydrophobic segments (gray shaded residues in Fig. 2), known to be core β-strands in ras and ARF1, are associated with segments of the human IL-1R that predict as β-strands.

Key residues

Given the low conservation of sequence homology mentioned above, the findings for the key binding and catalytic residues

of the GTPases are striking. The IL-1Rs all retain the majority of these residues or replace them with plausible alternatives that are themselves retained among the IL-1R group (this is not true for the Toll/MyD88 proteins, which may not have retained GTPase functions, as discussed later). Figure 2 indicates these functional residues at the top of the alignment by letters that note main-chain (M) or side-chain (S) interactions with nucleotide or magnesium (refer to key at the end of the figure). The following paragraphs relate the present findings to the recent description of the consensus residues (Pennington, 1994) that exist in most or all of the GTPases and that are known to be involved in GTP binding and hydrolysis.

GK(S/T) (Ras positions 15–17)

This group includes the lysine (rarely, arginine) residue that contacts the β-phosphate of the nucleotide via its side chain, and two residues that have main-chain contacts with GTP phosphate oxygens. Three IL-1Rs retain the lysine, and one, rat IL-1R, possesses the other permissible residue, arginine, in this triplet. All IL-1Rs retain the glycine and three retain the threonine, with chicken IL-1R possessing the conservative Thr–Ile substitution. All members of the IL-1R/Toll/MyD88 family conserve the central Lys/Arg, but Toll and MyD88 each make nonconservative replacements of the other two residues. These changes at this critical site suggest that Toll and MyD88 have lost their GTPase function. This is not entirely unexpected because other members of the GTPase superfamily have lost this function as well (Lee et al., 1995). It is worth noting that, in the case of both Toll and MyD88, this triplet and adjacent loop residues (to the left in the sequence) have undergone changes that add negatively charged side chains to this region. Because this is the triphosphate contact region, it is tempting to speculate that the added negative charges somehow compensate for the four negative charges normally provided by GTP when it binds. If this is the case, then it might be reasonable to assume that Toll and MyD88 represent constitutively active members of this superfamily, even in the absence of bound GTP. It is not, however, the purpose of this paper to propose or support such a concept for Toll or MyD88, and it is appropriate to distinguish this question from the main point of this discussion, which is limited to the IL-1Rs. As mentioned above, the conservation of the GKT triplet among the IL-1Rs is consistent with a retention of GTPase activity. In fact, given that the lysine has been shown to play a catalytic role, through electrostatic interactions with the β-phosphate, this conservation is strong evidence for a GTPase activity among the IL-1Rs.

T (ras position 35) or E (ras position 42)

Upon binding GTP, this region of the GTPase moves in order to "switch" to the signaling conformation of the protein, which then can bind to effector proteins via the switch I and II segments of polypeptide chain. This motion brings threonine 35 of ras into contact with the magnesium ion. Switch regions vary among the GTPases depending on the particular effector proteins they are supposed to bind to. Interestingly, a comparison of the 3D structures of ras and ARF1 (Amor et al., 1994) shows that the magnesium ligand from this region has changed from threonine 35 in ras to a glutamic acid further down the chain (corresponding to ras position 42) in ARF1. It would appear that IL-1R resembles ARF1 in this region, because no conserved threonine exists, but a conserved glutamic acid is indeed available at this position matching ras amino acid 42 (ARF1 position 54). This glutamic acid exists in all IL-1Rs and all other IL-1R/Toll/MyD88 family members as well. This implies that all family members retain a magnesium binding function.

DXXG(Q/H) (ras positions 57–61)

The aspartic acid in this segment is another magnesium ligand and the side chain of the glutamine or histidine has a critical catalytic function in that it contacts the leaving phosphate in the transition state, abstracting a proton (Sondek et al., 1994). A glutamic acid occurs and is conserved here in IL-1R. Interestingly, mutation data from ras studies have shown that a Gln to Glu substitution at this position retained full GTPase activity, whereas all substitutions of other amino acids caused a great reduction of activity (Frech et al., 1994). Thus, the IL-1Rs retain an acceptable catalytic residue as well as the magnesium-chelating aspartic acid and the conformationally important glycine residue. The Gln/His/Glu residue has, however, been lost in the Toll/MyD88 subgroup, again consistent with loss of GTPase activity.

NKXD (ras positions 116–119)

This motif occurs immediately following one of the extended stretches of hydrophobic amino acids (shaded in Fig. 2) that correspond to the core β-strands of the domain. This region is critical for specific recognition of the guanine base, with the side chains of the asparagine and aspartic acid hydrogen bonding to the amino groups of the base. These roles are preserved in the IL-1Rs, although the amide and acid functions have switched between the glutamic acid and glutamine at the corresponding positions, and chicken IL-1R utilizes glutamic acid for both of these functions. The lysine, which forms part of the hydrophobic pocket for the guanine with its side-chain methylenes, is conserved in every GTPase, and in all IL-1Rs as well. Although it is possible that the Asn to Glu and Asp to Gln changes could result in unacceptable steric changes in the guanine binding pocket, it is noteworthy that this segment of chain exists as a flexible loop in the known GTPase 3D structures, and differences between GTPases in the exact binding geometry and energetics have been noted (Weng et al., 1994). This in turn suggests that slight rearrangements in main-chain configuration may allow the Glu and Gln to bind to the guanine effectively.

(C/S)AX (ras positions 145–147)

This segment forms main-chain hydrogen bonds to the nucleotide from its first two residues and utilizes the side chain of its third residue as part of the hydrophobic pocket for the guanine base. All of the IL-1Rs conserve both of the first two residues, and the third, lysine, matches the lysine found at this position in ras. In fact, when two adjacent residues on the right side are included, this segment contains the five-residue exact match to ras, mentioned above. Such a match, occurring in a key functional region, strongly implies that the IL-1Rs do indeed share a common ancestry with ras, within the GTPase superfamily.

Other observations

No further residue matches were found in the C-terminal region. However, all of the key conserved residues of the GTPase superfamily have been accounted for. Figure 3 shows examples of amphipathic helices found in IL-1R. One of these (Fig. 3A) occurs in the C-terminal region of the IL-1R and matches an amphipathic stretch that occurs in ras and corresponds to the C-terminal helix found in all GTPase domains.

It is worth noting that the phosphorylatable threonine residue of the IL-1R (Bird et al., 1991), at ras position 176, but not corresponding to any ras residue, probably lies outside the GTPase domain region of the IL-1R, in a C-terminal tail. This phosphorylation occurs in an IL-1-dependent manner, although nothing is known of its effect on IL-1R function. This region shows no amphipathicity and no substantial hydrophobicity, also consistent with a nonpacked hydrophilic tail function.

Discussion

The most significant conclusion to be drawn from the findings reported here is that the GTPase activity known to be associated with the IL-1R may in fact be a function of the IL-1R itself. This could explain the extreme rapidity of the GTP binding response to IL-1 and also the inability of investigators to identify a separate, IL-1R-associated GTPase. Notably, investigations of ARF1 demonstrated that, due to the need for ARF1 to dimerize in order to be active, it was extremely difficult to extract the active form of the GTPase from cells (Amor et al., 1994) and the same may be true for the IL-1R. If the concept of the IL-1R as a GTPase is born out in the future by experimental testing, then this will be the first example of a GTPase known to be fused to a member of the Ig superfamily, and the first GTPase that exists in the form of a transmembrane receptor protein. Most members of the GTPase superfamily are either soluble monomers or are anchored to membranes via lipids attached to their N- or C-termini.

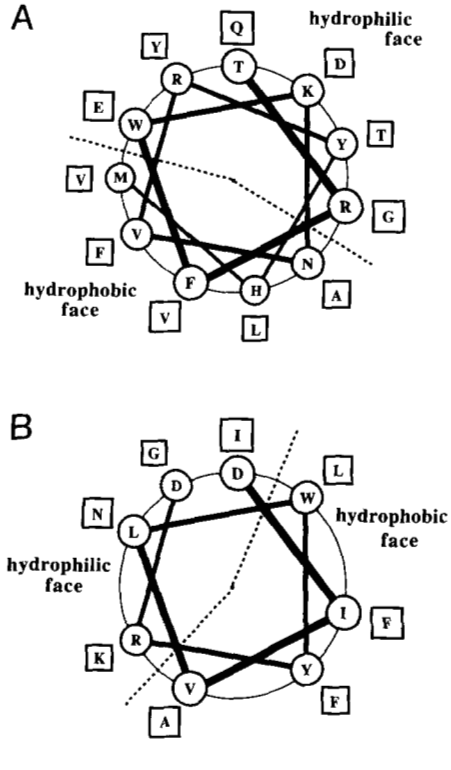

Fig. 3. Examples of amphipathic helices in IL-1R, ras, and ARF1. The putative IL-1R helices are indicated by the circled residues (single letter amino acid code) arranged around the large, inner helix circles. **A:** The C-terminal helix of ras is depicted by the outer circle of squares and the corresponding region of IL-1R is shown in the inner circle. Each protein shows a clear amphipathicity in this region. **B:** Squares represent the N-terminal helix of ARF1 and circles show the putative amphipathic helix at the N-terminus of the IL-1R cytoplasmic domain.

The receptor shares more features with ARFs than with other members of the superfamily. (1) The N-terminus of the IL-1R cytoplasmic domain is amphipathic (Fig. 3B), suggesting that it may be a helix similar to the N-terminal segment of ARF1. Furthermore, there are no proline or glycine residues among the first 16 cytoplasmic amino acids of any IL-1R. This suggests that the transmembrane helix may continue unbroken into this N-terminal region of the cytoplasmic domain. (2) ARF1 is the only GTPase known to form homodimers. It does so primarily via the N-terminal helix and the β-strand, β2 (Fig. 2, bottom), which hydrogen bonds to its counterpart in the opposite monomer. The N-terminal region is a conserved sequence among all of the IL-1Rs, as is the region corresponding to the β2-strand. This is consistent with the known fact that IL-1Rs must dimerize to become active in signaling (Dinarello, 1991) and suggests that the IL-1R may dimerize in an orientation that is similar to that of the ARFs. (3) The conserved glutamic acid in the switch I region suggests ARF-like binding of magnesium in this segment of IL-1R. (4) It has been noted that a "positive patch" of spatially adjacent basic residues occurs in ARF1 in the region thought to interact with the membrane (Amor et al., 1994). Such a patch would also exist in IL-1R if it folds into a 3D structure approximating that of ARF1. Highly conserved basic residues in this patch would include the two lysines in the N-terminal helix, the lysines and arginines of the conserved KKSRR sequence corresponding to ras positions 72–76, and several arginines in the C-terminal helix corresponding to ARF1 helix αF.

As mentioned above, Toll and MyD88 are members of the IL-1R family that fail to conserve many of the key binding and catalytic residues of the GTPases. This suggests that these proteins, which are only distantly related to the IL-1R group, have lost their GTPase activity. It is worth noting that these proteins each have inserted multiple negatively charged residues into the phosphate binding region (ras positions 10–15) and a tryptophan in the (C/S)AX segment that contacts the guanine base. It is conceivable that these amino acids fill the triphosphate and guanine binding pockets to yield constitutively active signaling proteins. This would suit their roles in embryonic development and differentiation, which are fundamentally different (and longer-acting) than the role of IL-1R in immune and inflammatory responsiveness.

Although the preceding information suggests that IL-1R and ARFs may be similar in their 3D structures, relationship to the membrane, and dimerization potential, there is no reason to conclude that the receptor is indeed an ADP ribosylation factor. In fact, there is nothing else evident in the sequences to indicate what the specific function of the receptor might be. Knowing that IL-1 binding to IL-1R-bearing cells leads to the activation of a number of kinases, adenyl cyclase, ras itself, sphingomy-

elinase, and a variety of other signaling mediators, it seems premature to place the IL-1R into any subcategory of the GTPase superfamily, nor would it be wise to exclude it from any.

This paper represents one of only a few reports thus far that use hydropathic profiles to make arguments concerning structural and functional properties of proteins (Grandori & Carey, 1994; Lee et al., 1995). This technique may see wider use in the future as more sequence data become available without known 3D structures or obvious homologies. At present, however, it remains unclear just how informative hydropathic plot comparisons are in general, so that caution should be exercised in their interpretation. In the present case, the additional evidence offered by active site residue conservation makes a stronger argument for GTPase relatedness. Nevertheless, this relationship should not be considered proven until experimental data are obtained to demonstrate a definite GTPase function for the IL-1R.

Methods

Homology searches

MacVector implementations of the ALIGN protein library search algorithms (Lipman & Pearson, 1985), with PAM250, PAM250s, and residue-identity search tables were used in efforts to detect homology between IL-1Rs, GTPases, and other proteins. The procedures were carried out on the entire ENTREZ sequence database (release 11.0) and on selected whole proteins without statistically significant results, other than with IL-1R/ST2L family members. As an alternative, attempts were made using another approach (Takio et al., 1984) wherein sequences are broken down into segments for searching.

Hydrophilicity profiles

The hydrophilicity profiling method of Hopp and Woods was used as described (Hopp, 1989). This procedure uses a moving average of hydrophilic/hydrophobic values assigned to the 20 amino acids in order to produce profiles of the type shown. An averaging window of six amino acids has been demonstrated to be optimal for structure prediction purposes (Hopp & Woods, 1981; Hopp, 1985) and was used here. The amino acid values used were taken from the amino acid "acrophilicity" scale (Hopp, 1985). Among hydrophilic/phobic scales, the acrophilicity scale has been shown to have a particularly strong relationship with protein secondary structure elements, especially α-helices and β-strands that pack into the cores of globular protein domains (Hopp, 1986). The program (HYDRO_A) is available from the author and is included in the Electronic Appendix to this paper.

Secondary structure predictions

The predicted secondary structure assignments shown at the bottom of Figure 2 were determined by the method of Garnier and coworkers (Garnier et al., 1978) as implemented in the PCgene protein analysis software (Intelligenetics).

Percent identity calculations

In this paper, the percent identity between any two aligned proteins is expressed as the number of identical residues divided by the total number of residues compared, multiplied by 100. Portions of proteins that extended beyond the ends of the IL-1R cytoplasmic domain were not included in comparisons and were therefore not counted as nonmatched amino acids.

Supplementary material in the Electronic Appendix

A Macintosh version of the HYDRO_A program is included in the Electronic Appendix to this report. This program uses FASTA format sequence files (e.g., those obtained from the ENTREZ database) to produce hydrophilicity profiles for proteins, together with a data file containing the numeric output.

Acknowledgment

This work was carried out without the benefit of grant support or other funding.

References

Amor JC, Harrison DH, Kahn RA, Ringe D. 1994. Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature 372*:704–708.

Bird TA, Woodward A, Jackson JL, Dower SK, Sims JE. 1991. Phorbol ester induces phosphorylation of the 80 kilodalton murine interleukin 1 receptor at a single threonine residue. *Biochem Biophys Res Commun 177*:61–67.

Bursten SL, Harris WE. 1994. Interleukin-1 stimulates phosphatidic acid mediated phospholipase D activity in human mesangial cells. *Am J Physiol 266*:C1093–C1104.

Carpen O, Pallai P, Staunton DE, Springer TA. 1992. Association of intercellular adhesion molecule-1 (ICAM-1) with actin-containing cytoskeleton and α-actinin. *J Cell Biol 118*:1223–1234.

Chou PY, Fasman GD. 1978. Prediction of the secondary structure of proteins from their amino acid sequences. *Adv Enzymol 47*:45–148.

Cohen L, Mohr R, Chen YY, Huang M, Kato R, Dorin D, Tamanoi F, Goga A, Afar D, Rosenberg N, Witte O. 1994. Transcriptional activation of a ras-like gene (kir) by oncogenic tyrosine kinases. *Proc Natl Acad Sci USA 91*:12448–12452.

Curtis BM, Gallis B, Overell RW, McMahan CJ, DeRoos P, Ireland R, Eisenman J, Dower SK, Sims JE. 1989. T-cell interleukin 1 receptor cDNA expressed in CHO cells regulates functional responses to interleukin 1. *Proc Natl Acad Sci USA 86*:3045–3049.

Dinarello CA. 1991. Interleukin-1 and interleukin-1 antagonism. *Blood 77*:1627–1652.

Frech M, Darden TA, Pedersen LG, Foley CK, Charifson PS, Anderson MW, Wittinghofer A. 1994. Role of glutamine-61 in the hydrolysis of GTP by p21H-ras: An experimental and theoretical study. *Biochemistry 33*:3237–3244.

Garnier J, Osguthorp DJ, Robson B. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol 120*:97–120.

Grandori R, Carey J. 1994. Six new candidate members of the α/β twisted open-sheet family detected by sequence similarity to flavodoxin. *Protein Sci 3*:2185–2193.

Hopp TP. 1985. Protein surface analysis: Methods for identifying antigenic determinants and other interaction sites. *J Immunol Methods 88*:1–18.

Hopp TP. 1986. Protein interaction sites predictable by hydrophilicity analysis. *Protides Biol Fluids 34*:59–62.

Hopp TP. 1989. The use of hydrophilicity plotting procedures to identify protein antigenic segments and other interaction sites. *Methods Enzymol 178*:571–585.

Hopp TP, Woods KR. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci USA 78*:3824–3828.

Hultmark D. 1994. Macrophage differentiation marker MyD88 is a member of the Toll/IL-1 receptor family. *Biochem Biophys Res Commun 199*:144–146.

Keegan AD, Paul WE. 1992. Multichain immune recognition receptors: Similarities in structure and signaling pathways. *Immunol Today 13*:63–68.

Kobe B, Deisenhofer J. 1993. Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature 366*:751–756.

Kopp EB, Ghosh S. 1995. NF-kB and rel proteins in innate immunity. *Adv Immunol 58*:1–22.

Krueger NX, Saito H. 1992. A human transmembrane protein-tyrosine-

phosphatase, PTP zeta, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrases. *Proc Natl Acad Sci USA 89*:7417–7421.

Lambright DG, Noel JP, Hamm HE, Sigler PB. 1994. Structural determinants for activation of a heterotrimeric G protein. *Nature 369*:621–628.

Lee JO, Rieu P, Arnaout MA, Liddington R. 1995. Crystal structure of the A domain from the α subunit of integrin CR3. *Cell 80*:631–638.

Letourneau PC, Shattuck TA. 1989. Distribution and possible interactions of actin-associated proteins and cell adhesion molecules of nerve growth cones. *Development 105*:505–519.

Lipman DJ, Pearson WR. 1985. Rapid and sensitive protein similarity searches. *Science 227*:1435–1440.

Mishima K, Tsuchiya M, Nightingale MS, Moss J, Vaughan M. 1993. ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-terminal ADP-ribosylation factor domain. *J Biol Chem 268*:8801–8807.

Mizel SB. 1990. Cyclic AMP and interleukin-1 signal transduction. *Immunol Today 11*:390–391.

Muegge K, Williams TM, Kant J, Karin M, Chiu R, Schmidt A, Siebenlist U, Young HA, Durum SK. 1989. Interleukin-1 costimulatory activity on the interleukin-2 promoter via AP-1. *Science 246*:249.

Nakashima H, Nishikawa K, Ooi T. 1986. The folding type of a protein is relevant to the amino acid composition. *J Biochem 99*:153–162.

Needleman SB, Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol 48*:443–453.

O'Neill LAJ, Bird TA, Saklatvala J. 1990. Interleukin-1 signal transduction. *Immunol Today 11*:392–394.

Patthy L. 1985. Evolution of the proteases of blood coagulation and fibrinolysis by assembly from modules. *Cell 41*:657–663.

Pennington SR. 1994. GTP-binding proteins 1: Heterotrimeric G proteins. *Protein Profile 1*:169–234.

Pollenberg E, Burridge K, Krebs K, Goodman S, Schachner M. 1987. The 180 kD component of the neural cell adhesion molecule N-CAM is involved in cell–cell contacts and cytoskeleton–membrane interactions. *Cell Tissue Res 250*:227–236.

Saklatvala J, Bird TA, Kaur P, O'Neill LAJ. 1990. IL-1 signal transduction: Evidence of activation of G protein and protein kinase. *Prog Clin Biol Res 349*:285–295.

Schaecke H, Mueller WEG, Gamulin V, Rinkevich B. 1994. The Ig superfamily includes members from the lowest invertebrates to the highest vertebrates. *Immunol Today 15*:497–498.

Sewell JL, Kahn RA. 1988. Sequences of bovine and yeast ADP-ribosylation factor and comparison to other GTP-binding proteins. *Proc Natl Acad Sci USA 85*:4620–4624.

Shirakawa F, Mizel SB. 1989. In vitro activation and nuclear translocation of NF-kB catalyzed by cAMP derived protein kinase and protein kinase C. *Mol Cell Biol 9*:2424.

Sondek J, Lambright DG, Noel JP, Hamm HE, Sigler PB. 1994. GTPase mechanism of G proteins from the crystal structure of transducin a-GDP-AlF4. *Nature 372*:276–279.

Takio K, Wade RD, Smith SD, Krebs EG, Walsh KA, Titani K. 1984. Guanosine cyclic 3′,5′-phosphate dependent protein kinase, a chimeric protein homologous with two separate protein families. *Biochemistry 23*:4207–4218.

Weng G, Chen CX, Balogh-Nair V, Callender R, Manor D. 1994. Hydrogen bond interactions of G proteins with the guanine ring moiety of guanine nucleotides. *Protein Sci 3*:22–29.

Williams AF, Barclay AN. 1988. The immunoglobulin superfamily—Domains for cell surface recognition. *Annu Rev Immunol 6*:381–405.

Yanagisawa K, Takagi T, Tsukamoto T, Tetsuka T, Tominaga SI. 1993. Presence of a novel primary response gene ST2L, encoding a product highly similar to the interleukin 1 receptor type 1. *FEBS Lett 318*:83–87.