Protein Engineering vol.15 no.9 pp.727–737, 2002

Prediction of the structure of human Janus kinase 2 (JAK2) comprising JAK homology domains 1 through 7

Fabrizio Giordanetto[1] and Romano T.Kroemer[1,2,3]

[1]Department of Chemistry, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, UK

[2]Present address: Molecular Modelling and Design, Discovery Research Oncology, Pharmacia, Viale Pasteur 10, 20014 Nerviano (MI), Italy

[3]To whom correspondence should be addressed.
E-mail: romano.kroemer@pharmacia.com

A theoretical model of human Janus kinase 2 (JAK2) comprising all seven Janus homology domains is presented. The model was generated by application of homology modelling approaches. The three-dimensional structure contains, starting from the N-terminus, FERM (4.1, ezrin, radixin, moesin), SH2 (Src homology region 2), tyrosine kinase-like, and tyrosine kinase domains. The predicted inter-domain orientation in JAK2 is discussed and the currently existing mutational data for Janus kinases are evaluated. Structural details of the SH2 and the FERM domains are presented. The predictions indicate that the SH2 domain is not fully functional. A number of hydrophobic amino acids of the FERM domain that are predicted to be involved in the constitutive association with the cytokine receptors are highlighted. The model gives new insights into the structure–function relationship of this important protein, and areas that could be investigated by mutation studies are highlighted.

Keywords: homology modelling/JAK2/
receptor-associated tyrosine kinase/signal transduction

Introduction

The Janus kinase or just another kinase (JAK) protein family (Ihle, 1994; Ihle *et al.*, 1994) is involved in the first cytoplasmic steps of the signal transduction cascade mediated by a variety of cytokines and hormones (Ihle, 1995; Leonard and O'Shea, 1998). Currently, four mammalian members of this group are known, JAK1, JAK2, JAK3 and TYK2 (Firmbachkraft *et al.*, 1990; Wilks *et al.*, 1991; Harpur *et al.*, 1992; Rane and Reddy, 1994) plus another member encoded by the *Drosophila* hopscotch gene (Binari and Perrimon, 1994). The main function of these proteins is to act as tyrosine kinases. Since many type I and II cytokine receptors lack a protein tyrosine kinase domain, they rely on JAKs to initiate the cytoplasmic signal transduction cascade. Ligand binding to their extracellular domain induces oligomerization of the receptors, which then activates the receptor associated JAK proteins at the inside of the cell membrane. Activation is accomplished through trans-phosphorylation of the JAK proteins, which subsequently phosphorylate tyrosine residues along the receptor chains they are associated with. These phosphotyrosine residues are a target for a variety of SH2-domain-containing transducer proteins. Among those are the signal transducers and activators of transcription (STAT proteins) which, after binding to the receptor chains, are phosphorylated by the JAK proteins. Phosphorylation enables them to dimerize and to translocate into the nucleus where they alter the expression of cytokine-regulated genes. This whole process has become known as the JAK–STAT pathway (Kazuroni and Leonard, 2000).

The amino acid sequence of the known members of the JAK family is characterized by the presence of seven highly conserved domains referred to as JAK homology domains (JH) (Harpur *et al.*, 1992) ranging from 1 to 7. The C-terminal domain (JH1) contains the tyrosine kinase function. The next domain along the sequence (JH2) is also known as the tyrosine kinase-like domain, as its sequence shows high homology to functional kinases, but it does not possess any catalytic activity (Wilks *et al.*, 1991). Its function is not yet well established, although there is some evidence for a regulatory role on the JH1 domain, thus modulating the catalytic activity (Luo *et al.*, 1997).

The N-terminal portion of the JAKs (spanning JH7 to JH3) is important for receptor association and non-catalytic activity (Frank *et al.*, 1994). Biochemical studies using JAK segments, expressed in bacteria and in mammalian cells, have shown that the N-terminus, including JH7 and part of JH6, is crucial for receptor association (Velazquez *et al.*, 1995; Zhao *et al.*, 1995; Chen *et al.*, 1997; Kohlhuber *et al.*, 1997). For example, the JH7-6 segment of JAK2 fused to JAK1 was sufficient for binding to the R2 subunit of the interferon (IFN)-gamma receptor (Kohlhuber *et al.*, 1997). A study of the interaction between JAK3 and the gamma common (γ_c) receptor chain indicated that the first 193 amino acid residues of JAK3 are sufficient for receptor binding (Chen *et al.*, 1997). However, 60 more amino acids were needed in a JAK3/JAK2 chimera to reconstitute interleukin (IL)-2 signalling (Cacalano *et al.*, 1999). In TYK2 the first 221 amino acids expressed *in vitro* bound efficiently to the IFNαR1 subunit of the IFN-α/β receptor, but *in vivo* additional JH domains were necessary for the *in vivo* assembly of TYK2 and IFNαR1 (Richter *et al.*, 1998; Yan *et al.*, 1998). For binding to granulocyte-macrophage colony stimulating factor (GM-CSF)-β_c receptor, amino acids 1–294 of JAK2 were essential (Quelle *et al.*, 1994; Zhao *et al.*, 1995).

The N-terminal region of JAK2 appears therefore crucial for the interaction with common type I/II cytokine receptors. The counterpart on the receptor has been the object of several studies. Considerable sequence homology has been detected within the membrane-proximal regions of the hematopoietic growth factor receptors (Fukanaga *et al.*, 1991; Murakami *et al.*, 1991; Taniguchi, 1995). This area includes an eight amino acid segment that is rather conserved among JAK2 interacting partners. The segment contains between two and four proline residues and is known as the proline-rich or box1 motif (Murakami *et al.*, 1991). Cytokine receptors also display a second conserved stretch of amino acids known as 'box2' in their cytoplasmic regions but its role is not well established. Evidence, however, exists for a particular role of the box1 motif in JAK2 recognition and association. Studies carried out

on growth hormone receptor (GHR) mutants by VanderKuur *et al.* (VanderKuur *et al.*, 1994) demonstrated the importance of an 85 amino acid region containing the box1 motif for the stable binding of GHR to JAK2: two mutant receptor chains—one lacking the entire proline-rich fragment and the other containing alanine for proline substitutions inside box1—did not immunoprecipitate with wild-type JAK2. Tanner *et al.* (Tanner *et al.*, 1995) extended this type of mutational analysis to other cytokine receptor box1 domains. Mutants of erythropoietin receptor (EPOR) and IL-6 receptor-associated transducer gp130 that lacked the proline rich sequence (PGIPSP and PNVPDP, respectively) were unable to precipitate with JAK2. Similar results were obtained by Zhao *et al.* (Zhao *et al.*, 1995) for the GM-CSF receptor beta chain. In their study, the receptor chain depended on amino acid segment 458–495, which includes the box1 motif, for the binding and consequent activation of JAK2. Bach *et al.* (Bach *et al.*, 1996) performed alanine-scanning mutagenesis of the cytoplasmic domain of the IFN-γ receptor beta chain. They identified the sequence $P^{263}PSIP^{267}$ as essential for a constitutive association between IFN-γ receptor and JAK2. The prolactin receptor (PRLR) box1 was also found important for the interaction with JAK2. Lebrun *et al.* (Lebrun *et al.*, 1995), employing several PRLR mutants, showed that amino acids 267–274 were required for stable interaction with JAK2. Since this fragment contains the proline-rich motif, they implicated this region as necessary for signal transduction.

A recent study indicated that the N-terminal amino acids from JH7 until the middle of the JH4 domain show significant homology to the FERM domain (Girault *et al.*, 1998), which was first described in the band 4.1 protein from erythrocytes (Leto and Marchesi, 1984). This protein cross-links the actin cytoskeleton to the erythrocyte membrane through a C-terminal actin-binding domain and the N-terminal FERM domain (Anderson and Marchesi, 1985). A variety of proteins, including the JAKs, were subsequently found to contain a FERM domain. These proteins include the tumour suppressor merlin (Rouleau *et al.*, 1993), the cell–cell contact protein talin (Rees *et al.*, 1990), the unconventional myosin VIIA (Chen *et al.*, 1996), the ezrin–radixin–moesin (ERM) proteins (Tsukita *et al.*, 1997), some protein tyrosine phosphatases (Banville *et al.*, 1994) and another group of tyrosine kinases, the focal adhesion kinases (FAKs) (Schaller *et al.*, 1992). Notably, the FERM domains of ezrin, moesin and radixin are involved in a stable association with the cytoplasmic regions of CD44, CD43 and ICAM-2 glycoproteins (Yonemura *et al.*, 1998). Therefore, the detection of such a domain in the N-terminus of the JAK proteins is consistent with the finding that the corresponding amino acids are crucial for receptor association.

It has been noted that the second half of JH4 plus the whole of the JH3 domain bears some weak similarity to a Src homology region 2 (SH2) domain (Bernards, 1991). More recently, this analysis has been corroborated by additional studies (Bork and Gibson, 1996; Kampa and Burnside, 2000). These findings are consistent with the fact that many protein tyrosine kinases possess an SH2 domain adjacent to the kinase domain. In the case of the JAK proteins, however, the adjacent domain, JH2, is a non-functional kinase.

JAK2 plays an important role in the development of leukaemia. Three leukaemic patients displayed similar chromosomal translocations. These genetic abnormalities expressed fusion proteins between TEL, a member of the ETS transcription factor family, and JAK2 (Lacronique *et al.*, 1997; Peeters *et al.*, 1997). These fusion products possessed constitutive kinase activity and they caused leukaemic effects *in vitro* (Lacronique *et al.*, 1997). The use of a specific JAK2 inhibitor was fundamental in blocking leukaemic cell growth in both *in vitro* and *in vivo* studies (Meydan *et al.*, 1996). Therefore, structural studies of JAK2 would offer a better understanding of its role and function. Unfortunately, no experimentally derived three-dimensional structure is currently available for the JAK2 protein. A theoretical model for the kinase domain (JH1) plus kinase-like domain (JH2) of JAK2 has been generated recently (Lindauer *et al.*, 2001). Using this structure as a starting point the N-terminal region of JAK2 was also modelled, and here the entire predicted three-dimensional structure of human JAK2, comprising all seven JAK homology domains, is presented.

Methods

The amino acids sequences of JAK family members were retrieved from the SWISSPROT database (Bairoch and Apweiler, 2000), using the DBGET retrieval system (Migimatsu and Fujibuchi, 1996): last accessed May 8, 2002. Multiple sequence alignment was performed employing CLUSTALW (Thompson *et al.*, 1994) and Dialign (Morgenstern *et al.*, 1998). Amino acid conservation was calculated using phylogenetic relationships (Clamp, 1998): last accessed May 8, 2002.

Secondary structure predictions were carried out using PSIpred V2.0 (Jones, 1999b) and PredictProtein (Rost and Sander, 1993, 1994; Rost *et al.*, 1994).

Fold recognition was performed using a variety of programs, such as 123D (Alexandrov *et al.*, 1996), TopLign (Thiele *et al.*, 1999), GenThreader and Threader (Jones, 1999a) and UCLA-DOE (Fischer and Eisenberg, 1996).

Three-dimensional structure coordinate files were obtained from the Protein Data Bank (Berman *et al.*, 2000). Search and evaluation of structural neighbours and structural comparison was performed with 3DALI (Holm and Sander, 1993), FSSP (Holm and Sander, 1996), SCOP (Murzin *et al.*, 1995), VAST (Gibrat *et al.*, 1996), CATH (Orengo *et al.*, 1997) and CE (Shindyalov and Bourne, 1998).

The three-dimensional models were generated by application of restraint-based homology modelling methods implemented in the program MODELLER (Šali and Blundell, 1993). Loop fragments were chosen according to the best score for homology and r.m.s. fit for the anchor residues, as generated by loop search algorithms (Jones and Thirup, 1986; Claessens *et al.*, 1989).

The relative orientation between the FERM domain and the SH2–JH2–JH1 part of the protein was calculated using the program 3D_Dock (Aloy *et al.*, 1998). Different angular deviations (20°, 15°, 12°, 9°), electrostatic filtering for the total space sampling, and an angular deviation ranging from 3° to 9° for each Euler angle in the refinement step were applied. The protein interface obtained was then optimized using the Multidock program (Jackson *et al.*, 1998).

Molecular properties such as electrostatic potentials were calculated using DelPhi (Nicholls and Honig, 1991) and visualized with GRASP (Nicholls *et al.*, 1991). Non-covalent interactions on the FERM domain were evaluated using the 'dry' probe within the GRID package (Goodford, 1985). The same program was employed for the assessment of surface property complementarity at the predicted interfaces. In this case, a cationic (Na^+), an anionic (Cl^-) alongside with the

hydrophobic (dry) probe were employed. A phosphate dianion probe (PO_4^{2-}) served for the simulation of the interaction between the SH2 domain and the phosphate group of a phosphotyrosine. Residue solvent-accessible areas were obtained with Naccess v2.1.1 (Hubbard and Thornton, 1992): last accessed May 8, 2002. Surface curvatures were calculated and displayed with GRASP (Nicholls *et al.*, 1991).

The final structure was subjected to 2000 steps of energy minimization using the Amber force field (Pearlman *et al.*, 1995). Subsequently a 20 ps molecular dynamics simulation at 300°C (step length 0.001 ps) involving only the side chain atoms was carried out. The structure was then subjected to another 2000 steps of energy minimization. Structural evaluation of the overall model was accomplished using the program Procheck (Laskowski *et al.*, 1993).

The structural models for the peptides corresponding to the GHR, EPOR and IFN-γ receptor box1 fragments were built with the Biopolymer module of SYBYL (SYBYL® 6.7.1, Tripos Inc., St Louis, MO, USA) selecting a random conformation. The peptides were subsequently minimized with the AMBER force field (Pearlman *et al.*, 1995; 2000 steps), solvated by 10 Å of water in each direction of a cubic box and submitted to molecular dynamics simulation of 80 ps at 300 K, applying the SHAKE algorithm. The conformations of the peptides obtained in this way were used in rigid docking simulations with the program 3D_Dock (Aloy *et al.*, 1998). The peptides were docked to the lobe B of the FERM domain of JAK2, employing an angular increment of 12°, applying an electrostatic and a pair potential filtering scheme.

The best complexes, according to surface complementarity, electrostatic and pair potential scores, served as starting points of subsequent molecular dynamic simulations. The complexes were solvated by 10 Å of water in each direction of a cubic box. The backbone atoms of the FERM domain of JAK2 were constrained, whereas the box1 peptides were left fully flexible during the whole process. MD was performed for 100 ps at 300 K. The resulting geometries were then energy minimized with AMBER.

Subsequent docking simulations were carried out using the program Autodock v3.0 (Morris *et al.*, 1998) with the macromolecule fixed and the peptides being flexible. The peptides corresponding to the GHR, EPOR and IFN-γ receptor box1 motifs were restrained only in the torsion around the peptide bonds. Affinity maps for all the atom types present, as well as an electrostatic map, were computed on a cubic grid of 120 points in each direction with a grid spacing of 0.5 Å. The search was carried out with the Lamarckian genetic algorithm, a population of 500 individuals, and a mutation rate of 0.03. Evaluation of the results was done by sorting the different complexes with respect to the predicted binding energy. A cluster analysis based on r.m.s.d. values, with reference to the starting geometry, was subsequently performed.

A different docking study was carried out in parallel to the procedure described above. The algorithm implemented in the QXP program (McMartin and Bohacek, 1997) allows for fully flexibility of the ligand (the receptor chain fragments in this case) and simultaneous flexibility of the receptor (the side chains of the JAK2 FERM domain in this case). Each docking run included 15 000 steps of Monte Carlo perturbation, subsequent fast searching, and final energy minimization. The results were evaluated in terms of total estimated binding energy.

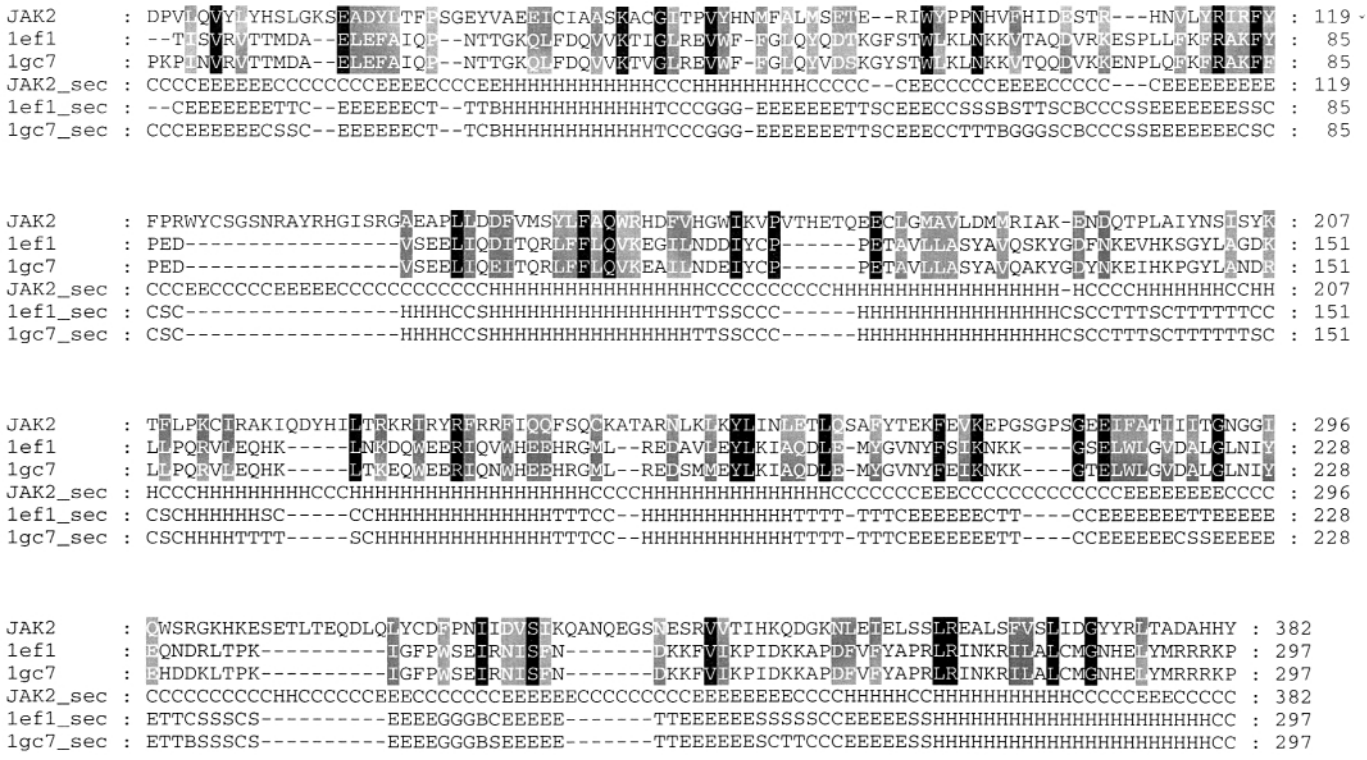

Fig. 1. Alignment for the FERM domain of JAK2 (D36–Y382). 1ef1 and 1gc7 are the PDB identifiers for the three-dimensional structures of human moesin (Pearson *et al.*, 2000) and radixin (Hamada *et al.*, 2000) FERM domains. Secondary structure prediction for JAK2 generated with PSIpred. (H, helix; E, extended strand; C, coil). Secondary structure assignments for the template structures are indicated as follows: H, helix; E, extended strand; C, coil; B, isolated beta-bridge; G, 3/10 helix; S, bend; T, hydrogen-bonded turn. Black shading indicates amino acid identities while grey shading represents conserved physico-chemical properties.

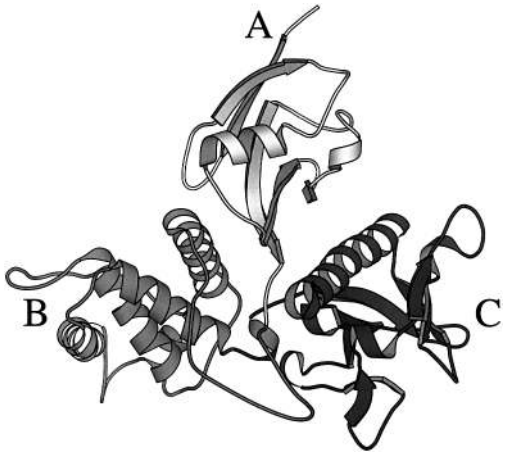

Fig. 2. Three-dimensional model of the FERM domain (D36–Y382) of JAK2. A, B and C indicate the three lobes according to the notation employed by Hamada *et al.* (Hamada *et al.*, 2000). Figure drawn with Molscript (Kraulis, 1991).

Results

FERM domain

The region comprising domains JH7 to the middle of JH4 was investigated with several fold recognition programs. GenThreader and Threader identified the three-dimensional structures of the FERM domains of radixin and moesin (Pearson *et al.*, 2000 and Hamada *et al.*, 2000, respectively) as suitable templates for the JAK2 amino acid segment, as indicated by maximum score values and reliability factors. The derived alignment is presented in Figure 1. The two template structures are very similar both in amino acid composition and in three-dimensional fold, as demonstrated by an 89% amino acid identity score and an r.m.s.d. of 0.406 Å using all the α carbon atoms when superimposed. Their alignment with respect to the JAK2 sequence is therefore identical. The first part of JAK2 (D36–R122) aligns well with the three-dimensional templates. Secondary structure prediction matched all the secondary features with the exception of the third β-strand of the FERM domain. This segment was predicted to contain an α-helix and a coil region. Therefore, an alternative alignment was generated, matching the F^{85}ALMSET^{91} segment of JAK2 with the F^{45}GLQYQD^{51} β-strand of moesin and the corresponding F^{45}GLQYVD^{51} segment on human radixin as shown in Figure 1. Using this alignment it was possible to conserve the physico-chemical properties of the amino acids (F^{45}, L^{47} and Y^{49}) that participate in a hydrophobic core, as observed in the crystal structures. The alignment of JAK2 with radixin and moesin displays a rather long insertion comprising residues W123 to G139. Secondary structure predictions indicated the possible presence of a β-strand (R130–H134) in this segment, but the confidence values were in a range between 10 and 80%. Therefore, this region was modelled employing loop-searching techniques (Jones and Thirup, 1986; Claessens *et al.*, 1989).

For amino acids A140 to Q347 of JAK2 a good alignment with the template structures could be generated. Several amino acid identities were found in this region and in many cases the physico-chemical properties were conserved. Secondary structure predictions for JAK2 were in good agreement with the features of the template structures.

The last part of the FERM domains of moesin and radixin contains a β-strand, a coil and a final long α-helix. Secondary structure predictions matched exactly the beginning of the α-helix, where two amino acid identities are found (L359 and R360).

According to the alignment, the JAK2 FERM domain resembles the usual architecture of a FERM pattern, consisting of three major lobes, A, B and C, which are building up an overall trigonal shape structure (Hamada *et al.*, 2000), as presented in Figure 2. Lobe A (residues D35–P121) folds into an α+β structure containing one α-helix and a five-stranded mixed β-sheet (Hamada *et al.*, 2000; Pearson *et al.*, 2000). Lobe B, comprising residues D147 to L258, is an all α-structure, including four longer helices and one shorter helix. Lobe C, consisting of residues from E268 to Y382, exhibits a seven-stranded β-sandwich core with one α-helix at the C-terminal end. Two large loops are found in the predicted model. The first one is located between lobes A and B and ranges from W123 to G139 (17 residues). The second one connects the two β-strands before S306 and after Q315 (10 amino acids). Since no reliable secondary structure prediction could be made for these two loops, they were modelled using standard loop searches (Jones and Thirup, 1986; Claessens *et al.*, 1989).

Different experimental studies (Frank *et al.*, 1995; Lebrun *et al.*, 1995; Tanner *et al.*, 1995; Zhao *et al.*, 1995; Kohlhuber *et al.*, 1997) indicated that the FERM domain of JAK2 is responsible for the association with the 'box1' motif of different cytokine receptors. This conserved amino acid segment displayed overall hydrophobic characteristics. Therefore, we investigated the presence of an adequate counterpart on the predicted JAK2 FERM domain. The interaction energies between the hydrophobic 'dry' probe and the FERM model were calculated using the program GRID. The 'dry' probe estimates the hydrophobic energy as the sum of water entropy contributions and Lennard-Jones components minus H-bond interaction energies. The most favourable interaction energies were found surrounding three hydrophobic amino acids (M181, F236 and F240) that formed a pocket on the JAK2 FERM domain (Figure 3A). M181 is located on the second α-helix of lobe B, whereas F236 and F240 reside on the fourth α-helix.

In order to evaluate the prediction of the hydrophobic pocket, the same interaction potential was calculated for the well known hydrophobic pocket of carbonic anhydrase II (Hakansson *et al.*, 1992) and the results were compared. Although the two pockets were different in sizes and in number of residues constituting the pocket, with the one belonging to carbonic anhydrase being larger, the calculated interaction energy values per unit surface area were comparable.

In order to verify more thoroughly the predictions for the interaction between the FERM domain of JAK2 and the box1 of different cytokine receptors, different simulations were carried out. In the alignment of the box1 motifs, belonging to cytokine receptors known to interact constitutively with JAK2, three different patterns of prolines can be found (Figure 4). Different peptides, corresponding to the relevant parts of the GHR, EPOR and IFN-γ receptors and accounting for the different proline patterns among the box1 motifs of cytokine receptors were docked to the FERM domain of JAK2. The docking studies were carried out as described in the Methods section.

More specifically, two different approaches were employed. The first type of procedure (A in Table I) consisted of a series of consecutive steps, starting with rigid docking (Aloy *et al.*, 1998), followed by molecular dynamics simulations of the

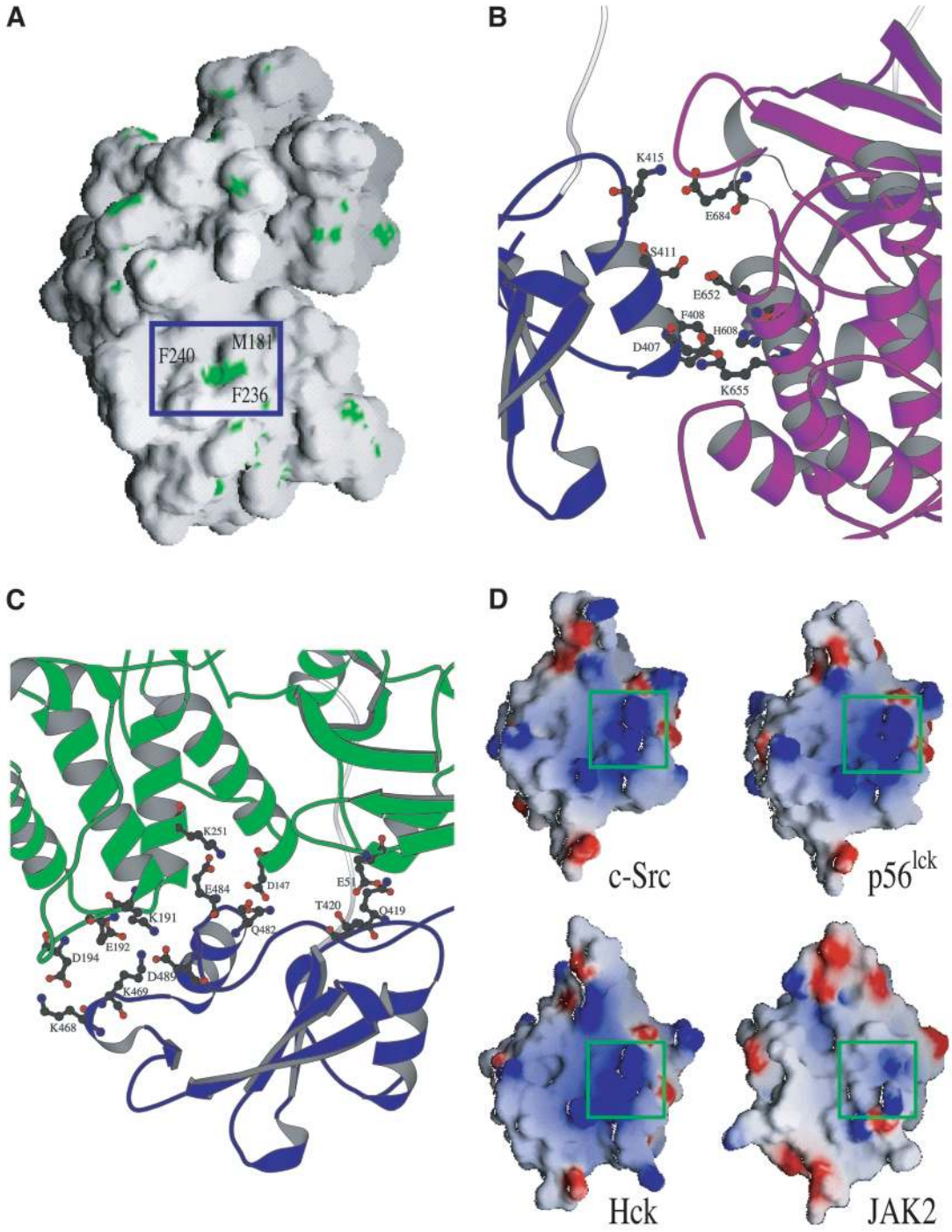

Fig. 3. (**A**) Hydrophobic pocket of the predicted FERM domain. The square highlights the area with the most favourable hydrophobic potential (–2.785 kcal/mol, compared to the next best value of –1.025 kcal/mol). (**B**) Close-up of the interface between the SH2 (blue) and the JH2 domain (purple). Figure drawn with Molscript (Kraulis, 1991). (**C**) Close-up of the SH2 (blue)–FERM domain (green) interface. Interacting amino acid side chains are indicated. Figure prepared with Molscript. (**D**) Molecular electrostatic potential of the SH2 domains of JAK2 and three templates structures. Potentials were calculated with DelPhi and displayed with GRASP (Nicholls *et al.*, 1991). Green boxes indicate the phosphotyrosine-binding pocket.

```
GHR      :  IKMLILPPVPVPKIKGIDP  :  311
PRLR     :  MVTCIFPPVPGPKIKGFDA  :  281
IFN-γ    :  IKYWFHTPPSIPLQIEEYL  :  295
gp130    :  IKKHIWPNVPDPSKSHIAQ  :  665
EPOR     :  LKQKIWPGIPSPESEFEGL  :  296
GM-CSFR  :  RKNPLWPSVPDPAHSSLGS  :  672
```

Fig. 4. Alignment of the box1 segments, including flanking residues, of receptors known to interact with JAK2. Black shading indicates the prolines in the box1 motif, whereas grey boxes highlight conserved physico-chemical properties.

resulting complexes, and finishing with flexible ligand docking (Morris *et al.*, 1998). The second, parallel, approach (B in Table I) included a docking simulation with flexible treatment of both the FERM domain side chains and the whole cytokine receptor peptides, employing QXP (McMartin and Bohacek, 1997). Using these two approaches it was possible to identify specific amino acids involved in hydrophobic interactions between the FERM domain and the box1 peptides. Most of these amino acids were identified as important by both docking procedures, as shown in Table I. The key hydrophobic residues

on the JAK2 FERM domain were the ones already identified with GRID (M181, F236 and F240). Also, I223 was engaged in hydrophobic contacts in all the complexes. Important contact residues on the GHR peptide were found in the box1 stretch (L296, L298, P299, P300 and V301), whereas aromatic amino acids (W283, F293), alongside with a box1 proline (P284), were found to be crucial for the EPOR peptide. In the FERM–IFN-γ complex only P288 of the cytokine receptor peptide was implied in hydrophobic interactions by both approaches.

SH2-like domain

The fold recognition programs used in this study predicted a few high scoring templates for the sequence spanning the second half of JH4 and the whole JH3 region. These templates are proteins closely related to the JAK family, both in terms of biological function and structure, and show an SH2 region next to the catalytic domain. Alignment with these proteins, as shown in Figure 5, comprises the SH2 domain of the human p56^{lck} kinase (1lkk; Tong *et al.*, 1996), the SH2 domain belonging to the tyrosine kinase c-Src (1fmk; Xu *et al.*, 1997) and the SH2 domain of the hematopoietic cell kinase (1qcf; Schindler *et al.*, 1999). The c-Src and Hck structures included not only the respective SH2 domains but also the catalytic domains. Figure 5 also contains the alignment of the start of their kinase sequences with the initial part of the kinase-like domain (JH2) of JAK2. The JH3-4 segment of JAK2 aligns very well with the corresponding template structures. Secondary structure prediction for JAK2 matches exactly the secondary structure features of the known structures. The predicted structure of the SH2 domain (or SH2-like domain, see below) comprising L393–P500 displays the usual SH2 fold, which consists of a central four-stranded β-sheet being flanked at both sides by an α-helix. Only one small insertion had to be modelled (E484–V486 when compared to the Hck tyrosine kinase structure), after the last α-helix of the SH2 fold.

Table I. Residues predicted to be involved in hydrophobic interactions in the peptide–FERM complexes according to docking protocol A only, protocol B only, and common to both approaches (A and B)

Complex	Docking protocol		
	A	B	A and B
GHR & FERM	L298, V301	L296, P299	P300
	V170, F236		M181, I223, F240
EPOR & FERM		I286, P287	W283, P284, F293
			M181, I223, F236, F240
IFN-γ & FERM	F281, P287	H282, P284, I291	P288
	F236		M181, I223, F240

SH2–JH2 interface

The interface between the SH2 domain and the tyrosine kinase-like domain (JH2) of JAK2 was generated using the structures of c-Src tyrosine kinase (Xu *et al.* 1997) and of Hck tyrosine kinase (Schindler *et al.*, 1999) as templates. Both of these structures contain an SH2 domain followed by a tyrosine kinase domain.

The SH2 model and the JH2 domain of Lindauer's model (2001) were fitted onto the corresponding regions of the two template structures, using the α carbon atoms of M406–K413, L122–C427, E448–K456 and L174–Y481 for the SH2 domain and L545–Q553, K558–R564 and I647–E666 for the JH2 region. Using c-Src kinase as a template, r.m.s.d. values of 0.906 Å for the SH2 domain and of 1.726 Å for the kinase-like region were obtained. The same procedure using the hematopoietic cell kinase yielded r.m.s.d. values of 1.082 and 2.018 Å, respectively. Therefore, the JAK2/c-Src tyrosine

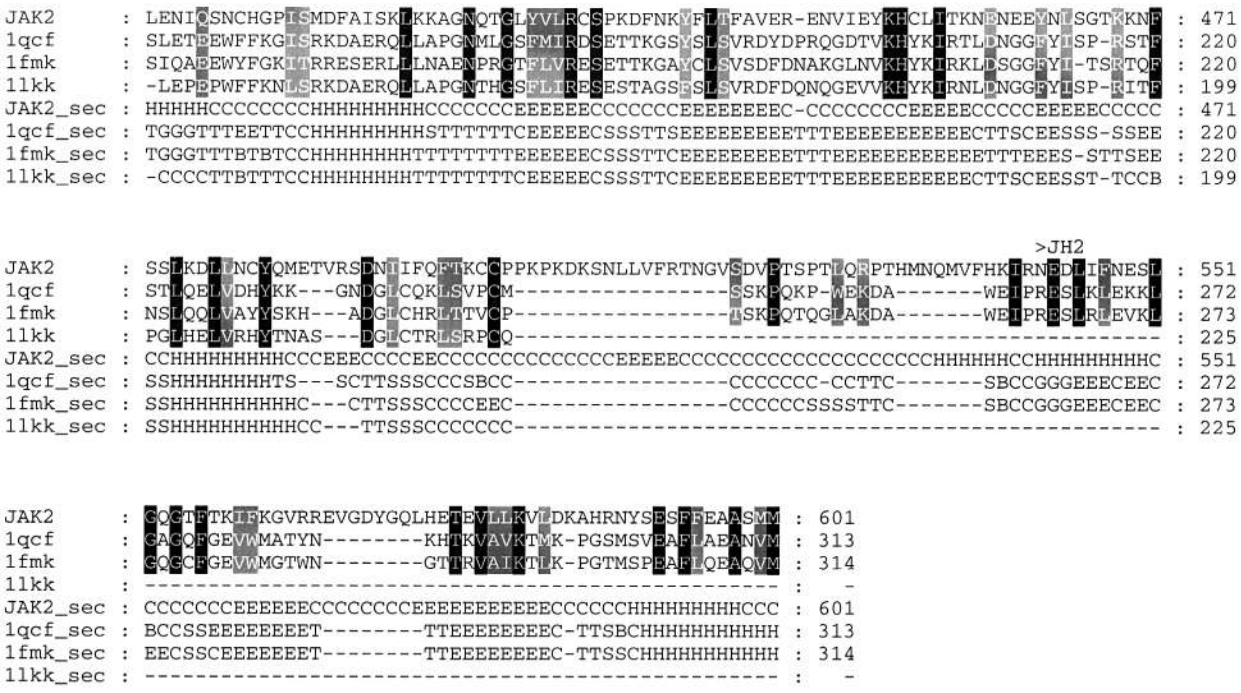

Fig. 5. Sequence–structure alignment of the JAK2 SH2 domain (L393–P500). 1qcf, 1fmk and 1lkk are the PDB accession numbers for the SH2 domains of hematopoietic cell kinase (Schindler *et al.*, 1999), c-Src tyrosine kinase (Xu *et al.*, 1997) and p56^{lck} tyrosine kinase (Tong *et al.*, 1996). Grey shadings as in Figure 1.

kinase superimposition was used as a model for the SH2–JH2 interface.

The linker between the SH2 and the kinase domains of both c-Src kinase and Hck tyrosine kinase (T248–R264 and S247–R264, respectively) is short when compared to the corresponding part of JAK2 (P501–N542). Secondary structure prediction indicated a β-strand (L509–R513) and an α-helix (Q534–K539) in this region of JAK2, but with low confidence values. Therefore, this loop was modelled employing loop-searching techniques. The loop chosen, according to the above-mentioned criteria, did not correspond to the secondary structure predictions, but as this loop is not involved in any inter-domain interactions it was selected as a suitable candidate.

In the predicted interface the first α-helix of the SH2 domain (M406–K414) interacts with an α-helix (I647–E666) and two turn fragments (S605–V610, I682–R687) of the JH2 (kinase-like) domain. The model of this interface displays several interactions. In particular, two ionic pairs (K415 with E684, and D409 with K655) and a hydrogen bond between the side chains of S411 and E652 can be found (Figure 3B). The electrostatic potential indicates electrostatic complementarity of the two interacting protein fragments. A positively charged zone is located around the side chain of K415, the area around S411 is neutral, while a negatively charged area surrounds D407. At the other side of the interface a complementary negative cluster is located on the side chains of E684 and E652, and a positively charged area is centred on K655. Similar results were obtained by application of the program GRID. The interaction energies of two charged probes (Na^+ as well as Cl^-) indicated favourable interaction energies between the amino acids. The hydrophobic probe revealed positive hydrophobic interactions corresponding to the side chains of F408 and H608 of the SH2 and the JH2 domains, respectively. These two amino acids are closely packed in the predicted interface and would therefore contribute to the stability of the interface.

Interface with the FERM domain

For generating the relative orientation of the FERM domain with respect to the remainder of the molecule no template structures were available. Therefore, a model for the inter-domain orientation was generated using a protein–protein docking approach (Aloy *et al.*, 1998).

The FERM domain of JAK2 was docked to different parts of the remainder of the JAK2 protein, namely the SH2 domain alone, the SH2 plus JH2 complex, and the whole SH2–JH2–JH1 structure. This was done in order to evaluate whether the shape and size of the FERM docking partner would affect the results. Different angular increments (20°, 15°, 12° and 9°) were used and an electrostatic filter was applied for the solutions. In order to investigate whether some of the larger loops (W123–G139 and S306–Q315 in the FERM domain, P501–N542 between SH2 and JH2, and S207–F832 between JH2 and JH1) affect the results, the corresponding fragments were docked also without these loop segments being present.

A 10-residue segment (L383–V392) connects the FERM and SH2 domains. The maximum possible length of this segment (32 Å in an extended conformation) provided therefore a distance filter for evaluating the solutions of the docking runs. Those docked orientations that satisfied this distance filter were further optimized using a finer angular increment (3°) during the subsequent rotational search. Interestingly, the resulting complexes with the highest values of surface complementarity displayed the same overall orientation, independent of the conditions of the docking procedure. The best docked orientation showed a surface complementarity correlation value of 61 and an electrostatic correlation value of –28.6, as calculated by fast-Fourier transform. Other orientations, which satisfied the distance filter described above, displayed lower surface complementarity scores (57 in the best case) and were rejected because they showed only slightly negative electrostatic scores (–4.6 for the best one, where positive values represent unfavourable electrostatic interactions).

The parts of the FERM domain that are predicted to form the interface with the SH2 domain are as follows. The so-called A lobe (using the notation of Hamada *et al.*, 2000) participates with a coil segment along with the second β-strand (L2967–L55) and the end of the α-helix (A70–G76). Parts of the B lobe that are involved include: the initial part of the first α-helix (D146–Y152), the end of the second α-helix plus the adjacent coil segment (M187–Q195), and the end of the third α-helix, the start of the fourth α-helix (S241–K251) as well as the segment connecting these two helices. The SH2 domain contributes to this interface with the linker between the first α-helix and the first β-strand (G417–G421), the second α-helix plus flanking residues (T467–D489) and the last part of the SH2 domain (F495–C499). The interface is shown in Figure 3C, and it contains several interacting pairs of amino acids. The side chains of D489/K191, K468/D194 and E484/K251 are forming salt bridges. Hydrogen bonds are formed between Q482 and D147 as well as between K469 and E192. The side chain of E51 is forming two hydrogen bonds, one with the side chain of Q419 and the other with the main chain NH group of T420.

An analysis of the electrostatic properties mapped onto the protein surfaces revealed a deep positively charged pocket corresponding to K191 and K251, whereas negatively charged areas were found around E51, D147, D194 and E197. On the other side of the interface a complementary cluster of negatively charged zones surrounded E484 and D489, while neutral areas were found around the side chains of Q419 and Q482. The same probe set as the one used for the SH2–JH2 interface indicated good interaction energies in the predicted SH2–FERM interface. In addition to a number of stabilizing electrostatic interactions, several amino acids displayed a favourable interaction with the 'dry' probe of GRID. In particular F148, I189 and A247 belonging to the FERM domain and M483 alongside with T496 showed the best interaction energies. These amino acids were found in close contact in the proposed interface. The overall predicted structure of JAK2 is shown in Figure 6.

Discussion

FERM domain

Hamada *et al.* (Hamada *et al.*, 2000) found that lobe C of radixin's FERM domain resembles the classical architecture of a phosphotyrosine-binding domain (PTB). In particular the structure was very similar to the PTB domain of the human insulin receptor substrate-1 (IRS-1) (Dhe-Paganon *et al.*, 1999). The amino acid comparison between radixin FERM's lobe C and the PTB of IRS-1 showed that two basic residues, required for the co-ordination of phosphotyrosine with IRS-1, were conserved in the radixin FERM structure (R246 and K262 of human radixin). The proposed structure of Jak2 contains I324

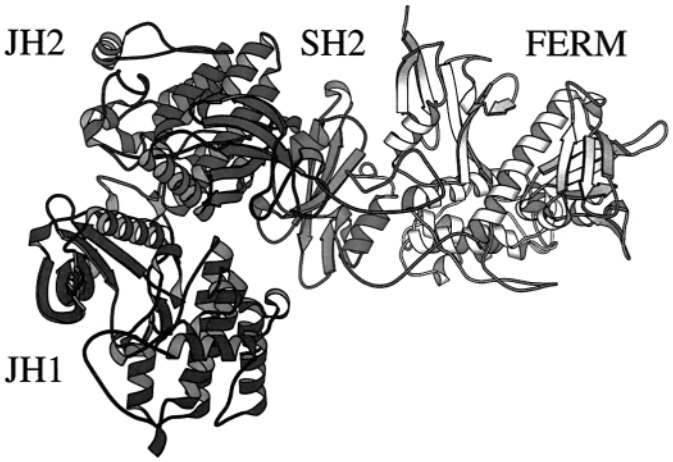

Fig. 6. Overall model of JAK2. Figure drawn with Molscript (Kraulis, 1991).

and Q347 at those positions. Our data therefore suggest that the FERM domain of JAK2 does not display a functional phosphotyrosine-binding site. Although the corresponding amino acid segment aligns very well with the template, the possibility of different alignments at those positions was evaluated. However, a suitable alternative alignment giving rise to a PTB could not be determined.

In the FERM domain of radixin Y146 is one of the residues that undergo phosphorylation by the tyrosine kinase protein v-Src (Takeda *et al.*, 1995). This position corresponds to N202 in the predicted model of JAK2. The adjacent residue in JAK2 (Y201) remains fully exposed to the solvent and could be a target for phosphorylation. Nevertheless, thus far there is no experimental evidence for the presence or importance of phosphorylation sites within the N-terminal domain of JAK2.

Several tyrosine residues present in the FERM region of JAK2 were systematically mutated to phenylalanine by Kohlhuber *et al.* (Kohlhuber *et al.*, 1997) in order to investigate the effects on STAT1 activation and subsequent signalling. None of these mutations showed a significant effect. In the model, two of the mutated residues (Y119 and Y254) are almost completely buried and they are involved in hydrophobic contacts with Y266 and M186, respectively. Therefore, these two tyrosines are not likely to represent suitable phosphorylation sites. Kohlhuber *et al.* (Kohlhuber *et al.*, 1997) also indicated the presence of a potential tyrosine-phosphorylating site in the sequence S^{367}LIDGYYRLTAD^{378}, similar to the one recognized by the SH2 domain of hematopoietic cell phosphatase (Songyang *et al.*, 1994). This amino acid segment was found at the end of the last α-helix of the JAK2 FERM domain model. The tyrosine at position 372 was mutated by Kohlhuber *et al.* (Kohlhuber *et al.*, 1997) into phenylalanine without detectable effects. In our model, Y372 is completely buried and it interacts with F118. Interestingly, Y373, which has not been subjected to mutational analysis, is fully exposed in our predicted structure. Moreover, a number of tyrosines (at positions 42, 44, 54, 62, 81, 201, 206, 221 and 231) are solvent exposed according to our model. Therefore, we propose these amino acids as suitable targets for mutational studies, in order to comprehensively evaluate the presence of tyrosine phosphorylation sites in the JAK2 FERM domain.

Girault *et al.* (Girault *et al.*, 1999) were the first to propose an alignment between JAK/FAK tyrosine kinases and FERM members. In particular, they used, among others, the sequence of JAK2 from mouse and moesin from *Drosophila melanogaster.* Their use of hydrophobic cluster analysis indicated that there are six highly conserved residues among the JAK, FAK and ERM proteins. Girault assumed that the FERM domain consisted of a duplication of two 140-residue domains. Later, however, a three-lobed nature of moesin FERM domain was revealed (Pearson *et al.*, 2000). The proposed model of the JAK2 FERM domain preserves three of the six conserved residues indicated by Girault. The first of these three residues, W95, is involved in hydrophobic contact with the R117 side chain carbon atoms, which in turn is involved in a network of hydrophobic interactions with F118, L363 and Y372. These interactions appear to stabilize the relative orientation of lobes A and C. The second residue, Q156, is hydrogen-bonded to R115 and is therefore involved in an important interaction between lobes A and B. These findings underscore the importance of these two amino acids in preserving and maintaining the overall spatial disposition of the JAK2 FERM domain. The third residue, G292, is adjacent to a turn that connects two β-strands forming a β-sheet in lobe C of the FERM domain and is packed against the side chain of F320. The crystal structures of the templates employed in the current study display similar types of interactions for the amino acids mentioned above. In particular, the tryptophan residue (W58 for moesin and radixin) is closely packed against the carbon atoms of the side chain of moesin and radixin K83. The glutamine (Q105 in human moesin and radixin FERM domain) is forming a hydrogen bond with the backbone amide group of A82 in the structures of moesin and radixin. The glycine that occurs at position 224 for moesin and radixin is in close contact with the side chain of W242 in both the structures. It could therefore be inferred that any side chain substituting the glycine H-α would destabilize the fold in this region.

A single amino acid substitution (G341E) in the *drosophila* hop^{tum-l} JAK kinase was found to cause leukaemia-like hematopoietic defects (Luo *et al.*, 1995). According to the present sequence alignment, the corresponding amino acid in JAK2 is N321. This residue is located in a linker segment between two β-strands on the surface of lobe C in the FERM domain. In the model the side chain of N321 is solvent accessible. This suggests a possible role for this mutation in the interaction between JAK2 and other protein partners.

An intensively studied mutation in the N-terminal region of JAK3 concerns Y100. The Y100C mutation was detected in a patient suffering from autosomal severe combined immunodeficiency (Macchi *et al.*, 1995). Y100, together with L98 and I102, are thought to be essential for interaction of JAK3 with the γ_c chain of the IL-2 receptor (Cacalano *et al.*, 1999). Sequence alignment of JAK2 with the other members of the JAK family, revealed that these residues are well conserved. In the present model of JAK2 the corresponding amino acids (Y114, V112 and I116) are not solvent exposed, but they interact with A86 and M88, L43 and H45, Y152 and L153, respectively, forming a prolonged hydrophobic network. This would therefore indicate that these residues are not involved in direct interaction with the receptor, but that they perform a crucial role in maintaining the proper fold for receptor interaction.

The FERM domains of ezrin, radixin and moesin are known to interact with the cytoplasmic regions of CD44, CD43 and ICAM-2 (Tsukita *et al.*, 1994). Intracellular segments of these receptors that are rich in basic amino acids and are strongly conserved were identified as the specific binding sites for the

FERM domains of ezrin, radixin and moesin (Yonemura *et al.*, 1998). Hamada *et al.* (Hamada *et al.*, 2000) proposed a cluster of acidic amino acids as interaction partners on radixin's FERM domain. However, in the FERM domain of JAK2 only two out of 17 acidic residues are retained. This is due to the fact that binding partners for the N-terminal of JAK2 differ significantly from the ones that interact with ERM proteins.

The FERM domain of JAK2 is thought to interact with the cytoplasmic regions of trans-membrane cytokine receptors. These receptors contain a conserved structural feature called 'box1', which is also referred to as the proline-rich motif, as it contains between two and four proline residues. Mutation and binding studies for the predicted FERM domain have indicated that this region is responsible for recognition and association of JAK2 with cytokine receptors. As the box1 motif of the cytokine receptors is very hydrophobic in nature, the surface of the model of the FERM domain was explored for hydrophobic zones. The region with the highest hydrophobic potential was identified near amino acids M181, F236 and F240 (Figure 3A). These three amino acids are well conserved among JAK members. For JAK1 and TYK2 they are identical, in JAK3 they are replaced by other hydrophobic amino acids. This is a further indication that these residues play an important role in JAK proteins. Another hydrophobic amino acid close to this hydrophobic cluster is I223. According to the predictions this residue could also be involved in hydrophobic interactions with the 'box1' motif.

The crystal structure of human moesin (Pearson *et al.*, 2000) was determined as a complex between the FERM domain and the C-terminal tail of human moesin. One of the interfaces in the structure was stabilized by the hydrophobic interactions between L124 and W175 of the FERM domain and V518, H521 and L525 of the C-terminal tail. Interestingly, L124 and W175 correspond to M181 and F236 in the current FERM domain of JAK2. These residues would therefore be interesting targets for site-directed mutagenesis experiments, in order to confirm their role in JAK2 receptor interactions.

SH2-like domain

Comparing the model of the putative SH2 domain in JAK2 with SH2 domains of other proteins, significant differences can be observed. In the crystal structure of the complex between the phosphotyrosine peptide and the SH2 domain of $p56^{lck}$, R134, R154 and S158 are interacting with the phosphate group (Tong *et al.*, 1996). In addition, the aromatic ring of the phosphotyrosine side chain displays amino aromatic with the guanidinium group of R134. The most conserved basic residue among the SH2 domains (Schaffhausen, 1995) corresponds to R154 in $p56^{lck}$. The equivalent residue in the model of JAK2 is R426 and could be involved in electrostatic interactions with a negatively charged phosphate group. However, R134 of $p56^{lck}$ is replaced by M406 in JAK2, thereby abolishing the putative interaction with a phosphate group. K430 of JAK2 corresponds to S158 of $p56^{lck}$, and its longer side chain would make it difficult to interact in the same manner with the phosphotyrosine than S158 in $p56^{lck}$. Another difference between the SH2 domains of JAK2 and of $p56^{lck}$ concerns the lining of the phosphotyrosine-binding pocket. E157 of $p56^{lck}$ is replaced by P429 in JAK2. In the model this proline appears to restrict access to the pocket. The flexible side chain of E157 points into the open space, while the rigid side chain of P429 on JAK2 is constrained to point into the binding pocket, thereby restricting the access space for the approaching phosphotyrosine. Figure 3D displays a comparison between the electrostatic potentials projected onto the molecular surfaces of $p56^{lck}$ tyrosine kinase, c-Src kinase, hematopoietic cell kinase, and the predicted model of the SH2 domain of JAK2. Overall the structures appear very similar in shape and their electrostatic properties appear rather well conserved. However, in the three template structures the region around the phosphotyrosine-binding pocket is more positively charged than in JAK2, which can be explained by the mutations in the SH2 domain of JAK2. Analysis of the structures with GRID using the phosphate dianion probe indicated that all four structures displayed the most negative interaction energy on a grid point close to the arginine residue (R426 for JAK2) that is invariant across the SH2 domains. Nevertheless, the three template structures displayed significantly better interaction energies with the probe compared to JAK2 (–34.2 for $p56^{lck}$, –33.0 for Hck, –30.8 for c-Src, and –22.1 for JAK2; all values in kcal/mol).

There appears to be overall sequence homology between an SH2 domain and the amino acids of JAK2 located in the JH3 and JH4 domains. However, analysis of the corresponding model structure and of the interaction potentials suggests less favourable interactions between a phosphotyrosine and the SH2 domain of JAK2, when compared with functional SH2 domains, despite the fact that the most important residue (R426) is conserved. These data are in line with recent experimental studies of Kohlhuber *et al.* (Kohlhuber *et al.*, 1997), where the R426A mutant had no detectable effect on STAT1 activation in transient transfections. According to the present model this can be explained by a general change in the properties of this domain, which alters its function from phosphotyrosine binding to an—as of yet—unknown role.

The SH2 domains of many tyrosine kinases play an important role in modulating catalytic activity by interacting with phosphotyrosines within the C-terminal tail. Therefore, a similar role for the putative SH2 domain of JAK was investigated. According to the predicted model, an intramolecular interaction between the SH2 domain and any tyrosine in the JH1 domain is unlikely to occur since it would require a significant conformational rearrangement.

Considering that the JAK proteins are thought to come in close contact during trans-activation, an alternative way of JAK2 self-regulation through intermolecular SH2–JH1 interactions involving two JAK2 proteins was explored as well. However, the present model did not allow for generation of a dimer involving the putative SH2 domain of one protein and either Y1007 or Y1008 of the activation loop of JH1 in the other protein. Formation of such a dimer was prevented by major steric overlap of the one JH2 domain with parts of the other protein. Analysis of the model of the SH2 domain casts doubts on whether it is fully functional. Therefore, it would be unable to interact also with other phosphotyrosine-containing proteins. This is also supported by the absence of any experimental data indicating that it is involved in interactions of this type. Therefore, the SH2 domain of JAK2 may represent another example of conservation of three-dimensional fold without conservation of function (Holm and Sander, 1995; Kiefer *et al.*, 1997).

SH2–JH2 interface

In the present model of JAK2 the SH2 domain is in contact with both the JH2 and FERM domains. Two structures out of three SH2 domain-containing proteins, which were identified

Table II. Residues pairs showing significant interactions at the predicted interfaces

Interface	Polar interactions	Hydrophobic interactions
SH2–JH2	D407–K655	F408–H608
	S411–E653	
	K415–E685	
FERM–SH2	D489–K191	T496–F148
	K468–D194	M483–I189, A247
	E484–K251	
	Q482–D147	
	K469–E192	
	E51–Q419, T420	

by fold recognition methods as possible templates for the SH2 domain of JAK2 SH2, displayed a kinase domain adjacent to the SH2 domain. These two structures, c-Src tyrosine kinase (Xu *et al.*, 1997) and hematopoietic cell kinase (Hck; Schindler *et al.*, 1999) perform a similar role to the JAKs in signal transduction. In both these proteins the first α-helix of the SH2 domain (R155–A165 for the c-Src tyrosine kinase and R155–P165 for the Hck protein) contributes to this interface. Several parts of the kinase are involved in the interface. These include the linker between the first α-helix and the fourth β-strand (H319–L322 for both structures), the third α-helix (L360–S372 and L360–E372) and the turn between the sixth and seventh β-strand (E396–C400 for c-Src and A396–C400 for Hck). Both template structures contain two conserved pairs of charged amino acids interacting at their interfaces (R160–D365 and E157–K321).

Figure 3B displays a close-up of the corresponding interface of JAK2 with the most important interactions. The interactions are also listed in Table II. In the model one of these interaction pairs corresponds to S411 and E652, whose side chains are H-bonded to each other. F408 and H608 represent the other pair, which is displaying hydrophobic contacts. In addition to these two conserved pairs, the proposed model contains other interacting amino acid pairs. In particular, two salt-bridges between K415 and E684, as well as D407 and K655, appear to be stabilizing the SH2–JH2 interface. All the amino acids found to interact in the predicted interface are rather well conserved among JAK members, which lends support to the prediction of a structural role of these amino acids.

FERM–SH2 interface

The orientation of the FERM domain with respect to the remainder of the protein was obtained by application of a rigid docking method (Aloy *et al.*, 1998). The best structure resulting from this procedure was found in contact with the SH2 domain alone. Three ion pairs, D489–K191, K468–D194 and E484–K251 contribute to the stability of the FERM–SH2 interface (Figure 3C, Table II). Additionally, four hydrogen bonds are present in this interface. Two of them are present in the Q482–D147 and K469–E192 interactions, whereas the side chain of E51 is bonding to the main chain amide group of T420 as well as the side chain amide group of Q419.

The interface is also characterized by good electrostatic complementarity between the two domains, as indicated by polar probes (Na^+ and Cl^-) in GRID. Using the 'dry' probe, a cluster of favourable hydrophobic interactions could be identified in the interface. The corresponding faces were located around F148, I189 and A247 of the FERM domain and near M483 and T496 of the SH2 domain. In the predicted interface, parts of the phenyl ring of F148 are closely packed against the side chain of T496. Also, the side chains of A247, I189 and M483 are engaged in contacts. Although there is no direct experimental evidence for this interface, its prediction is supported by many stabilizing interactions and good overall complementarity. It would be therefore interesting to generate mutations at the predicted interface in order to test the predictions.

Conclusions

The overall model of JAK2 displays a multidomain architecture. It comprises, from the C- to the N-terminus: the JH1 conserved block as the catalytic core and JH2 as an activity-regulating subunit, the SH2 domain and the FERM domain. The SH2 region of JAK2 retains the overall fold and the conserved arginine of the common SH2 domain. However, our predictions, in agreement with currently available experimental results (Kohlhuber *et al.*, 1997), suggest a silent, or yet undefined, role for it. In this way, JAK2 could be added to the class of those proteins that show an already known fold in the structure but with a different or abolished function.

The FERM domain is known as essential for the stable interaction of JAK2 with the cytokine receptors. Employing different approaches, a group of hydrophobic amino acids are predicted as crucial for JAK2–receptor association. Those residues appear to be well conserved among JAK members and they serve a similar function in the FERM domain of human moesin. These predictions are consistent with experimental results that map onto the JH7–6 region the part responsible for the stable interaction of JAK2 with the cytokine receptors.

The proposed model could help to investigate further the role of the SH2 domain and the receptor–FERM interaction via targeted mutagenesis studies. In order to study the FERM–receptor interactions, primary targets for mutation would be the hydrophobic cluster comprising M181, F236 and F239. The predictions indicate that the SH2 domain is a non-functional domain, which has been supported experimentally by a R426A mutant, but further mutation studies, using the proposed model as a basis, could shed further light on the role of this domain.

Additionally, the model will serve as a basis for the modelling of interactions of JAK2 with other proteins, thereby contributing to new insights into the signal transduction pathway it is involved in.

The co-ordinates of the Jak2 model are available upon request from the authors via e-mail. Also, they can be downloaded from http://www.chem.qmul.ac.uk/ccs/JAK2.pdb.gz

Acknowledgements

The authors gratefully acknowledge support from the Nuffield Foundation and from the National Foundation for Cancer Research (USA) for this work.

References

Alexandrov,N.N., Nussinov,R. and Zimmer,R.M. (1996) In Hunter,L. and Klein,T. (eds), *Biocomputing: Proceedings of the Pacific Symposium.* World Scientific Publishing Co., Singapore, pp. 53–72.

Aloy,P., Moont,G., Gabb,H.A., Querol,E., Aviles F.X. and Sternberg,M.J.E. (1998) *Proteins*, **33**, 535–549.

Anderson,R.A. and Marchesi,V.T. (1985) *Nature*, **318**, 295–298.

Bach,E.A., Tanner,J.W., Marsters,S., Shkenazi,A., Aguet,M., Shaw,A.S. and Schreiber,R.D. (1996) *Mol. Cell. Biol.*, **16**, 3214–3221.

Bairoch,A. and Apweiler,R. (2000) *Nucleic Acids Res.*, **28**, 45–48.

Banville,D., Ahmad,S., Stocco,R. and Shen,S.H. (1994) *J. Biol. Chem*., **269**, 22320–22327.
Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Beth,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.C. (2000) *Nucleic Acids Res.*, **28**, 235–242.
Bernards,A. (1991) *Oncogene*, **6**, 1185–1187.
Binari,R. and Perrimon,N. (1994) *Genes Dev*., **8**, 300–312.
Bork,P. and Gibson,T.J. (1996) *Methods Enzymol*., **266**, 162–184.
Cacalano,N.A., Migone,T., Bazan,F., Hanson,E.P., Chen,M., Candotti,F., O'Shea,J.J. and Johnston,J.A. (1999) *EMBO J*., **18**, 1549–1558.
Chen,Z.Y., Hasson,T., Kelley,P.M., Schwender,B.J., Schwartz,M.F., Ramakrishnan,M., Kimberling,W.J., Mooseker,M.S. and Corey,D.P. (1996) *Genomics*, **36**, 440–448.
Chen,M., Cheng,A., Chen,Y., Hymel,A., Hanson,E.P., Kimmel,L., Minami,Y., Taniguchi,T., Changelian,P.S. and O'Shea,J.J. (1997) *Proc. Natl Acad. USA*, **94**, 6910–6915.
Claessens,M., Van Cutsem,E., Lasters,I. and Wodak,S. (1989) *Protein Eng*., **2**, 335–345.
Clamp,M. (1998) *Jalview*. http://www.ebi.ac.uk/~michele/jalview/
Dhe-Paganon,S., Ottinger,E.A., Nolte,R.T. and Shoelson,S.E. (1999) *Proc. Natl Acad. USA*, **96**, 8378–8383.
Firmbachkraft,I., Byers,M., Shows,T., Dallafavera,R. and Krolewski,J.J. (1990) *Oncogene*, **5**, 1329–1336.
Fischer,D. and Eisenberg,D. (1996) *Protein Sci.*, **5**, 947–955.
Frank,S.J., Gilliland,G., Fraft,A.S. and Arnold,C.S. (1994) *Endocrinology*, **135**, 2228–2239.
Frank,S.J., Yi,W.S., Zhao,Y.M., Goldsmith,J.F., Gilliland,G., Jiang,J., Sakai,I. and Kraft,A.S. (1995) *J. Biol. Chem.*, **270**, 14776–14785.
Gibrat,J.F., Madej,T. and Bryant,S.H. (1996) *Curr. Opin. Struct. Biol*., **6**, 377–385.
Girault,J.A., Labesse,G., Mornon,J. and Callebaut,I. (1998) *Mol. Med*., **4**, 751–769.
Girault,J.A., Labesse,G., Mornon,J. and Callebaut,I. (1999) *Trends Biochem. Sci*., **24**, 54–57.
Goodford,P.J. (1985) *J. Med. Chem*., **28**, 849–857.
Hakansson,K., Carlsson,M., Svensson,L.A. and Liljas,A. (1992) *J. Mol. Biol.*, **227**, 1192–1204.
Hamada,K., Shimizu,T., Matsui,T., Tsukita,S., Tsukita,S. and Hakoshima,T. (2000) *EMBO J*., **19**, 4449–4462.
Harpur,A.G., Andres,A.C., Ziemiecki,A., Aston,R.R. and Wilks,A.F. (1992) *Oncogene*, **7**, 1347–1353.
Holm,L. and Sander,C. (1993) *J. Mol. Biol*., **233**, 123–128.
Holm,L. and Sander,C. (1995) *EMBO J*., **14**, 1287–1293.
Holm,L. and Sander,C. (1996) *Science*, **273**, 595–602.
Hubbard,S. and Thornton,J. (1992) *NACCESS 2.1.1*. http://wolf.bms.umist.ac.uk/naccess/
Ihle,J.N. (1994) *Proc. Soc. Exp. Biol. Med*., **206**, 268–272.
Ihle,J.N. (1995) *Nature*, **377**, 591–594.
Ihle,J.N., Witthuhn,B.A., Quelle,F.W. and Silvennoinen,O. (1994) *Trends Biol. Sci.*, **19**, 227–227.
Jackson,R.M., Gabb,H.A. and Sternberg,M.J.E. (1998) *J. Mol. Biol*., **276**, 265–285.
Jones,D.T. (1999a) *J. Mol. Biol*., **287**, 797–815.
Jones,D.T. (1999b) *J. Mol. Biol*., **292**, 195–202.
Jones,T.A. and Thirup,S. (1986) *EMBO J*., **5**, 819–822.
Kampa,D. and Burnside,J. (2000) *Biochem. Biophys. Res. Commun*., **278**, 175–182.
Kazuroni,I. and Leonard,W.J. (2000) *Mol. Immunol*., **37**, 1–11.
Kiefer,J.R., Mao,C., Hansen,C.J., Basehore,S.L., Hogrefe,H.H., Braman,J.C. and Beese,L.S. (1997) *Structure*, **5**, 95–108.
Kohlhuber,F., Rogers,N.C., Watling,D., Feng,J., Guschin,D., Briscoe,J., Witthuhn,B.A., Kotenko,S.V., Pestka,S., Stark,G.R., Ihle,J.N. and Kerr,I.M. (1997) *Mol. Cell. Biol.*, **17**, 695–706.
Kraulis,P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.
Lacronique,V., Boureux,A., Della Valle,V., Poirel,H., Tran Quang,C., Mauchauffe,M., Berthou,C., Lessard,M., Berger,R., Ghysdael,J. and Bernard,O.A. (1997) *Science*, **278**, 1309–1312.
Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) *J. Appl. Crystallogr.*, **26**, 283–291.
Lebrun,J.J., Ali,S., Ullrich,A. and Kelly,P.A. (1995) *J. Biol. Chem*., **270**, 10664–10670.
Leonard,W.J. and O'Shea,J.J. (1998) *Annu. Rev. Immunol*., **16**, 293–322.
Leto,T.L. and Marchesi,V.T. (1984) *J. Mol. Biol*., **259**, 4603–4608.
Lindauer,K., Loerting,T., Liedl,K.R. and Kroemer,R.T. (2001) *Protein Eng*., **14**, 27–37.
Luo,H., Hanratty,W.P. and Dearolf,C.R. (1995) *EMBO J*., **14**, 1412–1420.
Luo,H., Rose,P., Barber,D., Hanratti,W.P., Lee,S., Roberts,T.M., Dandrea,A.D. and Dearolf,C.R. (1997) *Mol. Cell. Biol*., **17**, 1562–1571.
Macchi,P., Villa,A., Gillani,S., Sacca,M., Prattini,A., Porta,F., Ugazio,A.G., Johnston,J.A., Candotti,F., O'Shea,J.J., Vezzoni,P. and Notarangalo,L.D. (1995) *Nature*, **377**, 65–68.
McMartin,C. and Bohacek,R. (1997) *J. Comp. Aid. Mol. Des*., **11**, 333–344.
Meydan,N., Grünberger,T., Dadi,H., Shahar,M., Arpaia,E., Lapidot,Z., Leeder,S., Freedman,M., Cohen,A., Gazit,A., Levitzki,A. and Roifman,C.M. (1996) *Nature*, **379**, 645–648.
Migimatsu,H. and Fujibuchi,W. (1996) *Version 2 of DBGET*. http://www.genome.ad.jp/dbget/dbget_manual.html
Morgenstern,B., Frech,K., Dress,A. and Werner,T. (1998) *Bioinformatics*, **14**, 290–294.
Morris,G.M., Goodsell,D.S., Halliday,R.S., Huey,R., Hart,W.E., Belew,R.K. and Olson,A.J. (1998) *J. Comput. Chem.*, **19**, 1639–1662.
Murakami,M., Narazaki,M., Hibi,M., Yawata,H., Yasukawa,K., Hamaguchi,M., Taga,T. and Kishimoto,T. (1991) *Proc. Natl Acad. USA*, **88**, 11349–11353.
Murzin,A.G., Brenner,S.E., Hubbard,T. and Chothia,C. (1995) *J. Mol. Biol*., **247**, 536–540.
Nicholls,A. and Honig,B. (1991) *J. Comput. Chem*., **12**, 435–445.
Nicholls,A., Sharp,K. and Honig,B. (1991) *Proteins*, **11**, 281–287.
Orengo,C.A., Michie,A.D., Jones,S., Jones,D.T., Swindells,M.B. and Thornton,J.M. (1997) *Structure*, **5**, 1093–1108.
Pearlman,D.A., Case,D.A., Caldwell,J.W., Ross,W.S., Cheatham,T.E., DeBolt,S., Ferguson,D., Seibel,G. and Kollman,P. (1995) *Comput. Chem. Comm*., **91**, 1–41.
Pearson,M.A., Reczek,D., Bretscher,A. and Karplus,P.A. (2000) *Cell*, **101**, 259–270.
Peeters,P., Raynaud,S.D., Cools,J., Wlodarska,I., Grosgeorge,J., Philip,P., Monpoux,F., Van Rompaey,L., Baens,M., Van den Berghe,H. and Marynen,P. (1997) *Blood*, **90**, 2535–2540.
Quelle,F.W., Sato,N., Witthuhn,B.A., Inhorn,R.C., Eder,M., Miyajima,A., Griffin,J.D. and Ihle,J.N. (1994) *Mol. Cell. Biol*., **14**, 4335–4341.
Rane,S.G. and Reddy,E.P. (1994) *Oncogene*, **9**, 2415–2423.
Rees,D.J.G., Ades,S.A., Singer,S.J. and Hynes,R.O. (1990) *Nature*, **347**, 685–689.
Richter,M.F., Duménil,G., Uzé,G., Fellous,M. and Pellegrini,S. (1998) *J. Biol. Chem.*, **273**, 24723–24729.
Rost,B. and Sander,C. (1993) *J. Mol. Biol.*, **232**, 584–599.
Rost,B. and Sander C. (1994) *Proteins: Struct. Funct. Genet.*, **19**, 55–72.
Rost,B., Sander,C. and Schneider,R. (1994) *CABIOS*, **10**, 53–60.
Rouleau,G.A. *et al.* (1993) *Nature*, **363**, 515–521.
Šali,A. and Blundell,T.L. (1993) *J. Mol. Biol*., **234**, 779–815.
Schaffhausen,B. (1995) *Biochem. Biophys. Acta*, **1242**, 61–75.
Schaller,M.D., Borgman,C.A., Cobb,B.S., Vines,R.R., Reynolds,A.B. and Parsons,J.T. (1992) *Proc. Natl Acad. USA*, **89**, 5192–5196.
Schindler,T., Sicheri,F., Pico,A., Gazit,A., Levitzki,A. and Kuriyan,J. (1999) *Mol. Cell*, **3**, 639–647.
Shindyalov,I.N. and Bourne,P.E. (1998) *Protein Eng*., **11**, 739–747.
Songyang,Z. *et al.* (1994) *Mol. Cell. Biol*., **14**, 2777–2785.
Takeda,H., Nagafuchi,A., Yonemura,S., Tsukita,S., Behrens,J., Birchmeier,W. and Tsukita,S. (1995) *J. Cell. Biol*., **131**, 1839–1847.
Taniguchi,T. (1995) *Science*, **268**, 251–255.
Tanner,J.W., Chen,W., Young,R.L., Longmore,G.D. and Shaw,A.S. (1995) *J. Biol. Chem*., **270**, 6523–6530.
Thiele,R., Zimmer,R. and Lengauer,T. (1999) *J. Mol. Biol*., **290**, 757–779.
Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) *Nucleic Acids Res*., **22**, 4673–4680.
Tong,L., Warren,T.C., King,J., Betageri,R., Rose,J. and Jakes,S. (1996) *J. Mol. Biol*., **256**, 601–610.
Tsukita,S., Oishi,K., Sato,N., Sagara,J., Kawai,A. and Tsukita,S. (1994) *J. Cell. Biol.*, **126**, 391–401.
Tsukita,S., Yonemura,S. and Tsukita,S. (1997) *Curr. Opin. Cell. Biol*., **9**, 70–75.
VanderKuur,J.A., Wang,X.Y., Zhang,L.Y., Campbell,G.S., Allevato,G., Billestrup,N., Norstedt,G. and Carter-Su,C. (1994) *J. Biol. Chem*., **269**, 21709–21717.
Velazquez,L., Mogensen,K.E., Barbieri,G., Fellous,M., Uzé,G. and Pellegrini,S. (1995) *J. Biol. Chem*., **270**, 3327–3334.
Wilks,A.F, Harpur,A.G., Kurban,R.R., Ralph,S.J., Zurcher,G. and Ziemiecki,A. (1991) *Mol. Cell. Biol*., **11**, 2057–2065.
Xu,W.Q., Harrison,S.C. and Eck,M.J. (1997) *Nature*, **385**, 595–602.
Yan,H., Piazza,F., Krishnan,K., Pine,R. and Krolewski,J.J. (1998) *J. Biol. Chem.*, **273**, 4046–4051.
Yonemura,S., Hirao,M., Doi,Y., Takahashi,N., Kondo,T., Tsukita,S. and Tsukita,S. (1998) *J. Cell. Biol*., **140**, 885–895.
Zhao,Y.M., Wagner,F., Frank,S.J. and Kraft,A.S. (1995) *J. Biol. Chem*., **170**, 13814–13818.

Received August 10, 2001; revised May 14, 2002; accepted June 10, 2002