Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 1029–1033, February 1991
Biochemistry

FKB1 encodes a nonessential FK 506-binding protein in *Saccharomyces cerevisiae* and contains regions suggesting homology to the cyclophilins

GREG WIEDERRECHT, LEONARDO BRIZUELA, KEITH ELLISTON, NOLAN H. SIGAL, AND JOHN J. SIEKIERKA*

Merck Sharp & Dohme Research Laboratories, P.O. Box 2000, Rahway, NJ 07065

Communicated by Edward M. Scolnick, October 29, 1990

ABSTRACT **FK 506, a powerful immunosuppressant that blocks allograft rejection by preventing T-cell activation, binds to an 11-kDa protein called the FK 506-binding protein (FKBP). Like cyclophilin, a cytosolic protein that binds another immunosuppressant, cyclosporin A, FKBP possesses peptidylprolyl cis–trans isomerase activity. We have isolated a genomic clone encoding the yeast FKBP (*FKB1*). The gene encodes a protein of 114 amino acids having a calculated M_r of 12,158. Disruption of the gene shows that *FKB1* is not essential for growth. A search of translated nucleic acid data bases revealed bacterial FKBP homologs in *Neisseria meningiditis* and *Pseudomonas aeruginosa*. Comparison of the conserved amino acids in FKBP homologs with the conserved amino acids in the cyclophilins has revealed a region of similarity that we speculate to be a homologous domain related to the functional similarities of the two proteins.**

FK 506, a neutral macrolide produced by *Streptomyces tsukabaensis*, is an important new immunosuppressive drug with 10–100 times the potency of another immunosuppressant, the cyclic peptide cyclosporin A (CsA) (1). Both drugs inhibit T-lymphocyte activation by affecting the signal-transduction events leading to transcription of the same set of early-phase lymphokine genes (2). Although their effects upon T-cell activation are indistinguishable, FK 506 and CsA bind to distinct, abundant (0.1–0.4% of total cellular protein), cytosolic receptors termed the FK 506-binding protein (FKBP; M_r 11,000) and cyclophilin (M_r 17,000), respectively (3–5). FKBP and cyclophilin are not lymphoid-specific proteins but are widely distributed in tissues and throughout the phyla (6, 7). Both proteins are members of a class of enzymes termed peptidylprolyl isomerases (PPIases) that catalyze isomerization between the cis and trans forms of the Xaa-Pro bond in peptides and proteins (4, 5, 8, 9). The PPIase activities of FKBP and cyclophilin are inhibited by FK 506 and CsA, respectively (4, 5). FKBPs have been purified from human T lymphocytes, calf thymus, and *Saccharomyces cerevisiae* and all exhibit PPIase activity that is inhibited by FK 506, immunological cross-reactivity, and equivalent masses (7).

Protein sequencing of the *S. cerevisiae*, bovine, and human FKBPs revealed few amino acid differences between the human and bovine proteins and demonstrated that the yeast protein is homologous to the mammalian FKBPs (7). The high degree of amino acid conservation in FKBP from these divergent species as well as the abundance of the protein in all three organisms indicates that FKBP has an important role in cellular physiology. Comparison of cyclophilin clones from such phylogenically diverse species as human, rat, *Neurospora crassa*, and *S. cerevisiae* has also demonstrated excellent conservation at the amino acid level (10, 11). It is intriguing that two chemically unrelated ligands, FK 506 and CsA, affect expression of the same set of lymphokine genes and bind with nanomolar affinity to distinct cytosolic enzymes, both of which have PPIase activity. However, similarities between the amino acid sequences of FKBP and cyclophilin are not conspicuous at first glance. Understanding the roles that the FKBP and cyclophilin may perform during the process of T-cell activation will help reveal the mechanism of action of FK 506 and CsA.

The powerful genetic and molecular approaches available in *S. cerevisiae* make it an especially attractive organism in which to undertake a detailed molecular analysis of FKBP. Furthermore, a comprehension of the function of FKBP in yeast may provide insights to its function in T cells. We report the cloning and sequencing of the gene encoding the yeast FKBP, *FKB1*.† *FKB1* is a nonessential gene present once per haploid genome.

By computer analysis, we compared the derived amino acid sequences from the human (refs. 12 and 13; G.W., N.H.S., and J.J.S., unpublished work), *S. cerevisiae*, and *N. crassa* (14) FKBP-encoding genes. Their evolutionary conservation with each other and their similarities to two unidentified open reading frames (URFs) found in the genomes of *Neisseria meningitidis* and *Pseudomonas aeruginosa* are discussed. These URFs may encode bacterial homologs of the eukaryotic FKBPs. FKBP and cyclophilin are functionally related because both are PPIases and because both bind drugs that prevent T-cell activation. With these similarities in mind, we have compared the conserved amino acids in six FKBP homologs with the conserved amino acids in nine cyclophilin homologs. A region of similarity between FKBP and cyclophilin is suggested. We speculate that this homologous domain is important for one of the shared activities of these two proteins.

MATERIALS AND METHODS

Design of Oligonucleotide Probes and Screening of a Yeast Library. Degenerate, oligonucleotide probes were synthesized on a MilliGen Biosearch cyclone model 8400 DNA synthesizer. The oligonucleotide sequences as well as the yeast FKBP amino acid sequences they were derived from are shown below:

```
ASP GLY ALA THR PHE PRO LYS THR GLY ASP
CTA CCI CGI TGI AAA GGI TTT TGI CCI CT
  G               G       C
and
ASN ILE GLY VAL GLY GLN VAL ILE LYS GLY
TTA TAI CCI CAI CCI GTT CAI TAI TTT CC
  G                   C           C
```

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: FKBP, FK 506-binding protein; URF, unidentified open reading frame; PPIase, peptidylprolyl cis–trans isomerase; CsA, cyclosporin A.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M57967).

The yeast library was screened according to the methods described by Itoh *et al.* (15). DNA inserts from positive phage were subcloned into pUC19 and sequenced by the dideoxynucleotide chain-termination method from denatured plasmid miniprep DNA.

Yeast Strains and Methods. Yeast were grown on YPD media (16). The diploid strain, YFK016, used to construct the gene disruption is the product of a cross between haploid strains YPH54 and YPH98 (17). Yeast protein extracts were prepared from 6 ml of early stationary cultures grown in YPD. Cells were disrupted by glass beads in 0.3 ml of buffer containing 150 mM Tris (pH 7.5), 10 mM $MgCl_2$, 1 mM dithiothreitol, 10% (vol/vol) glycerol, 1 mM phenylmethylsulfonyl fluoride, benzamidine at 3 μg/ml, aprotinin at 1 μg/ml, and leupeptin at 1 μg/ml. Insoluble material was removed by centrifugation for 15 min in a microcentrifuge. The supernatant was assayed using the [^{3}H]FK 506-binding assay as described (4).

Sequence Analysis. Sequence analysis of the various FKBPs and cyclophilins was performed using the University of Wisconsin Genetics Computer Group package (18). Percent similarity, percent identity, and pairwise homologies were assessed using the program GAP, molecular weights were calculated using the program PEPSORT, and isoelectric points were calculated using the program ISOELECTRIC. Data base searches were conducted against GenBank release 63 (19), European Molecular Biology Laboratory (EMBL) release 23 (20), and Protein Identification Resource release 25 (21) by using the Pearson and Lipman programs FASTA and TFASTA (22) as implemented in the University of Wisconsin Genetics Computer Group package. Multiple sequence alignments were constructed using the programs CLUSTAL (23) and TREEALIGN (24). Similarity measurements were calculated using the Gribskov and Burgess scoring table (25), which is a normalized version Dayhoff's "log-odds" amino acid comparison matrix (26). Amino acids were considered similar when their score was 0.5 or above. URFs were analyzed for protein-coding potential by using the program TESTCODE (27) as implemented in the University of Wisconsin Genetics Computer Group package.

RESULTS AND DISCUSSION

Isolation of the Yeast FKBP-Encoding Gene. The amino acid sequence of the *S. cerevisiae* FKBP (7) was used to design two nonoverlapping, inosine-containing oligonucleotide probes for screening a yeast genomic DNA library. Sixteen thousand recombinant phage were screened, and five plaques hybridized to both oligonucleotide probes through several rounds of screening. All five phage contained a 2-kilobase (kb) insert when digested with *Eco*RI. Sequencing of ≈1 kb of the 2-kb insert (Fig. 1) revealed an open reading frame encoding a 114-amino acid protein of M_r 12,150, in agreement with the molecular weight of the purified yeast FKBP as determined by SDS/PAGE (7).

The authenticity of the clone is verified by the fact that the amino acid sequence is entirely consistent with the 77 amino acids obtained from peptide sequencing (7). Compatible with the observation that the FKBP is an abundant protein are the facts that the initiator methionine is in an excellent context for translation initiation, with an adenine at −3 and a thymidine at +6, a consensus residue among highly expressed yeast genes (28). Furthermore, the codon bias index of *FKB1* was calculated to be 0.71 [the codon bias index for yeast cyclophilin is 0.77 (10)], a value demonstrating highly biased codon representation and predicting a high level of mRNA in the cytoplasm (29). Promoter analysis reveals several potential TATA boxes, at −227, −218, −204, −169, and −32, suggesting that, like most yeast genes, *FKB1* has multiple transcription start sites.

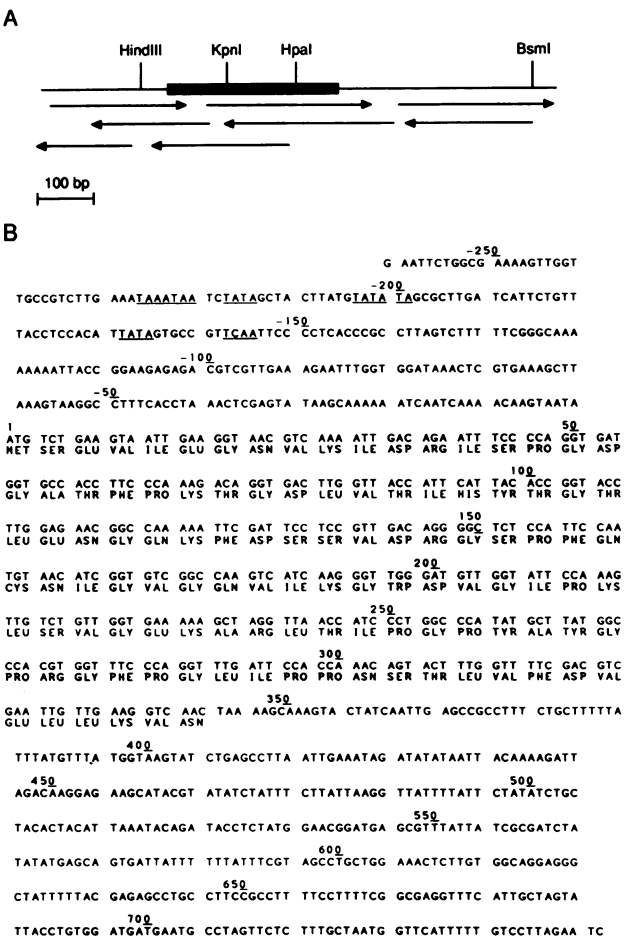

FIG. 1. (*A*) Sequencing strategy for the yeast FKBP-encoding gene. The thick line represents the open reading frame. (*B*) Nucleotide sequence of the *S. cerevisiae* FKBP-encoding gene and the predicted amino acid sequence. The sequence shown is 1009 bases of the 2-kb genomic *Eco*RI fragment containing *FKB1*. Potential TATA boxes are underlined. The predicted amino acid sequence is shown below the nucleotide sequence.

To determine whether the FKBP is essential in yeast, a disruption mutant of *FKB1* was constructed by one-step gene replacement (30). The 2-kb *Eco*RI genomic fragment containing *FKB1* was first subcloned into pUC19 (Fig. 2*A1*). A 1.6-kb *Sma* I–*Hpa* I fragment containing the *URA3* selectable marker was cloned in the antisense orientation relative to *FKB1* into the unique *Bst*EII site (Fig. 2*A2*). The *Xho* I fragment, now containing the disrupted *FKB1* gene, was excised from the vector and used to transform the *ura3/ura3* diploid yeast strain YFK016 to uracil prototrophy. The phenotype of the disruption mutant was examined in haploid by sporulation of six independent transformants and dissection of the resulting tetrads. In all asci examined, two of the spores were uracil auxotrophs, and two of the spores were uracil prototrophs. In all instances, all spores were viable. Southern analysis of the Ura$^+$ transformants confirmed that the proper physical changes had taken place in the genome (Fig. 2*B*). The undisrupted diploid strain carries the *FKB1* gene on a single 2-kb *Eco*RI fragment (lane 1). *Eco*RI digestion of genomic DNA prepared from the Ura$^+$ diploid transformant demonstrated that one chromosome contains the wild-type *FKB1* gene, whereas the other chromosome contains the disrupted copy (Fig. 2*B*, lane 2). The disrupted haploid grows at the same rate and reaches the same final cell density at saturation as the wild-type haploid strain.

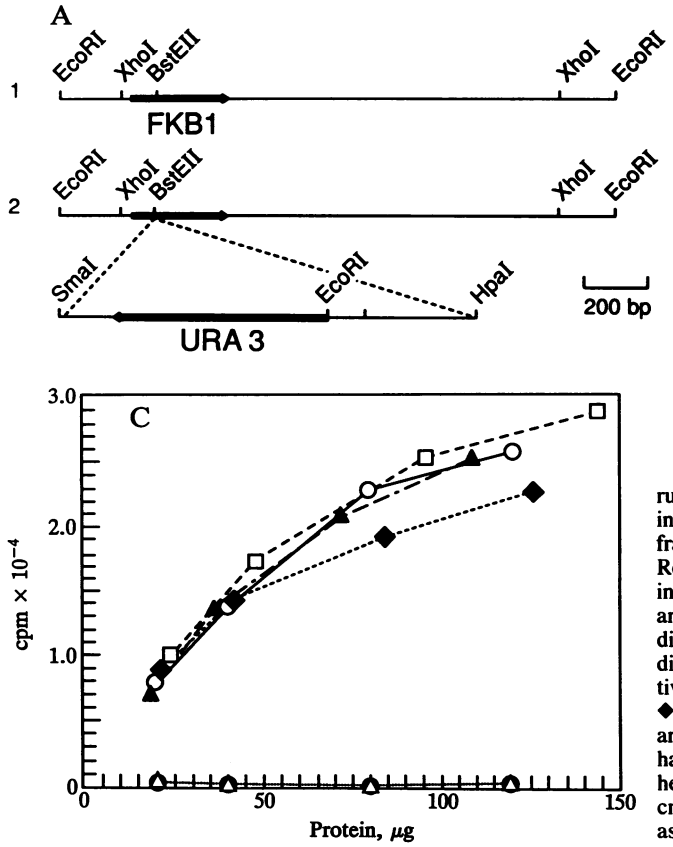

FIG. 2. (*A*) Graphic representation depicting the disruption mutant, *fkb1*, made by inserting the *URA3* gene into *FKB1*. The location of *FKB1* within the 2-kb *Eco*RI fragment and relevant restriction sites are shown in *1*. Relevant restriction sites within *URA3* and its insertion into the *Bst*EII site of *FKB1* are depicted in *2*. (*B*) Southern analysis of DNA (10 μg) isolated from the nondisrupted diploid strain YFK016 (lane 1) and from the *FKB1*-disrupted heterozygote (lane 2). (*C*) FK 506-binding activity of the parental haploids YPH54 and YPH98 (□ and ◆, respectively), of the nondisrupted haploid products (○ and ▲) and of the disrupted haploid products (● and △). The haploid products resulted from tetrad dissection of the heterozygous, disrupted diploid. Increased amounts of crude extract prepared from each of the strains was assayed for FK 506 binding.

Extracts were prepared from the nondisrupted haploid products and from the disrupted haploid products and assayed for [^{3}H]FK 506 binding activity (Fig. 2*C*). As a control, extracts were prepared from the parental haploid strains mated to produce the diploid strain used for the disruption. The parental strains and the nondisrupted Ura$^-$ haploid products of the disrupted diploid contained equivalent amounts of binding activity. The disrupted Ura$^+$ haploid products contained no detectable FK 506-binding activity. The genetic linkage of uracil prototrophy to the lack of FK 506-binding activity corroborates that *FKB1* has been disrupted and further authenticates the correctness of the clone. These results have been confirmed with a deletion mutant of *FKB1* (data not shown). The inactivation experiment and the Southern analysis described above suggest that *FKB1* is a single-copy gene that is not essential for growth. The results also indicate that the FKBP is the major, if not the sole, FKBP in yeast cells.

Comparison of Eukaryotic FKBPs and Identification of Possible Prokaryotic FKBP Homologs. The human (refs. 12 and 13; G.W., N.H.S., and J.J.S., unpublished results), bovine (7), *S. cerevisiae*, and *Neurospora* FKBPs (14) have similar masses, and the human, bovine, and *Neurospora* FKBPs have similar pI values. However, the yeast protein is more acidic than the human and *Neurospora* FKBPs (7). Alignment of the amino acid sequences derived from the human, *Neurospora*, and yeast FKBP clones and from the bovine FKBP peptide sequence is shown in Fig. 3*A*. The human FKBP shares 57% identity and 75% similarity with the yeast FKBP and 44% identity and 64% similarity with the *Neurospora* FKBP. The alignment focuses on the conserved regions of these proteins that begins between 20 and 30 residues into their respective sequences. Within the 84 amino acids of bovine amino acid sequence available (7), there is 95% identity with the human FKBP.

When using *FKB1* to query the translated nucleotide data bases, EMBL and GenBank, high-score homologies to URFs in *N. meningitidis* and *P. aeruginosa* were discovered. The translated URFs are aligned in their regions of homology with the eukaryotic FKBPs in Fig. 3*A*. The homologous URF in *N. meningitidis* is near two tandemly arranged, silent, truncated pilin genes (37). The translated URF encodes a protein with a calculated M_r of 9700 that is 43% and 47% identical and 65% and 71% similar to the human and yeast FKBPs, respectively. The *N. meningitidis* URF encodes an initiator methionine preceded by stop codons in all three reading frames and is in a good context for translation initiation, having the preferred adenine at position −3 (28) and a potential ribosome binding site from −7 to −12. A TESTCODE analysis (27) further confirmed that the *N. meningitidis* URF encodes a functional protein. In an *in vitro*-coupled transcription–translation reaction with a clone containing the URF as template, a protein in the appropriate size range was well-expressed (37). The URF in *P. aeruginosa* (38) maps near a region encoding regulatory proteins involved in alginate biosynthesis and translates into a potential 23-kDa polypeptide. The C-terminal domain of the translated URF (i.e., the region that overlaps with known FKBPs) is 36% and 32% identical and 55% and 56% similar to the human and yeast FKBPs, respectively. The TESTCODE measurement of codon bias indicates that this URF also has high protein-coding potential.

FKBP and Cyclophilin May Share a Homologous Domain. FKBP and cyclophilin share at least two activities. Both are receptors for immunosuppressive drugs that prevent expression of the same set of lymphokine genes in T cells, and both are small, abundant, and ubiquitous proteins having PPIase activity. We have examined the two proteins for consensus amino acid sequences that might correspond to their similar activities. Multiple sequence alignment of the eukaryotic FKBPs as well as their bacterial homologs is shown in Fig. 3*A*. The amino acid sequences of nine eukaryotic cyclophilins or cyclophilin homologs (8, 10, 11, 31–36) are aligned in Fig. 3*B*. The N-terminal sequences of both the FKBPs and the

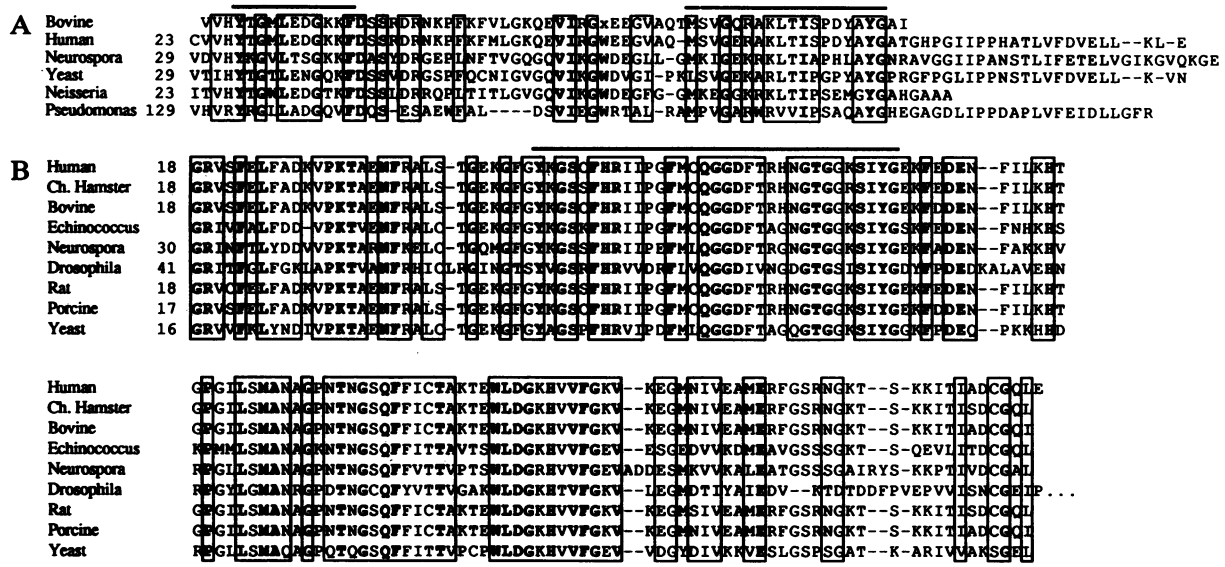

FIG. 3. (*A*) Amino acid sequence (in single-letter code) alignment of the human, bovine, *Neurospora*, and *S. cerevisiae* FKBPs and the potential FKBP homologs in *N. meningitidis* and *P. aeruginosa*. Similar residues are enclosed in boxes, where identities in all six FKBPs are represented by boldface type and conserved residues in all six sequences are depicted in regular type. Gaps required to align the sequences are represented by dashes. Numbers indicate the amino acid number of the first residue shown. The bovine sequence is unnumbered because the N-terminal sequence is not known for certain. The bold lines are those portions of the FKBPs homologous to the cyclophilins. (*B*) Amino acid sequence alignment of the human (31), Chinese (Ch.) hamster (32), bovine (33), *Echinococcus* (34), *Neurospora* (35), rat (11), porcine (8), *Drosophila* (36), and *S. cerevisiae* (10) cyclophilins or cyclophilin homologs. Explanation of numbers, dashes, and of boxed and boldface residues are described for *A*. The *Echinococcus* sequence is unnumbered because the cDNA sequence was incomplete. The bold line describes that region of the cyclophilins homologous to the FKBPs.

cyclophilins are less well conserved, and so the alignments shown start between 16 and 41 residues into the respective proteins (except for the *P. aeruginosa* FKBP, where the alignment begins at amino acid 129). Only amino acids identical or similar in all six FKBPs or in all nine cyclophilins are enclosed by boxes in Fig. 3. Of special note are the facts that the translated URFs of *Neisseria* and *Pseudomonas* are conserved in the same regions as the other FKBPs, further corroboration that they encode prokaryotic FKBP homologs.

Examination of the aligned sequences in Fig. 3 reveals that while the cyclophilins are extremely well conserved throughout, the conservation throughout the six FKBPs is somewhat

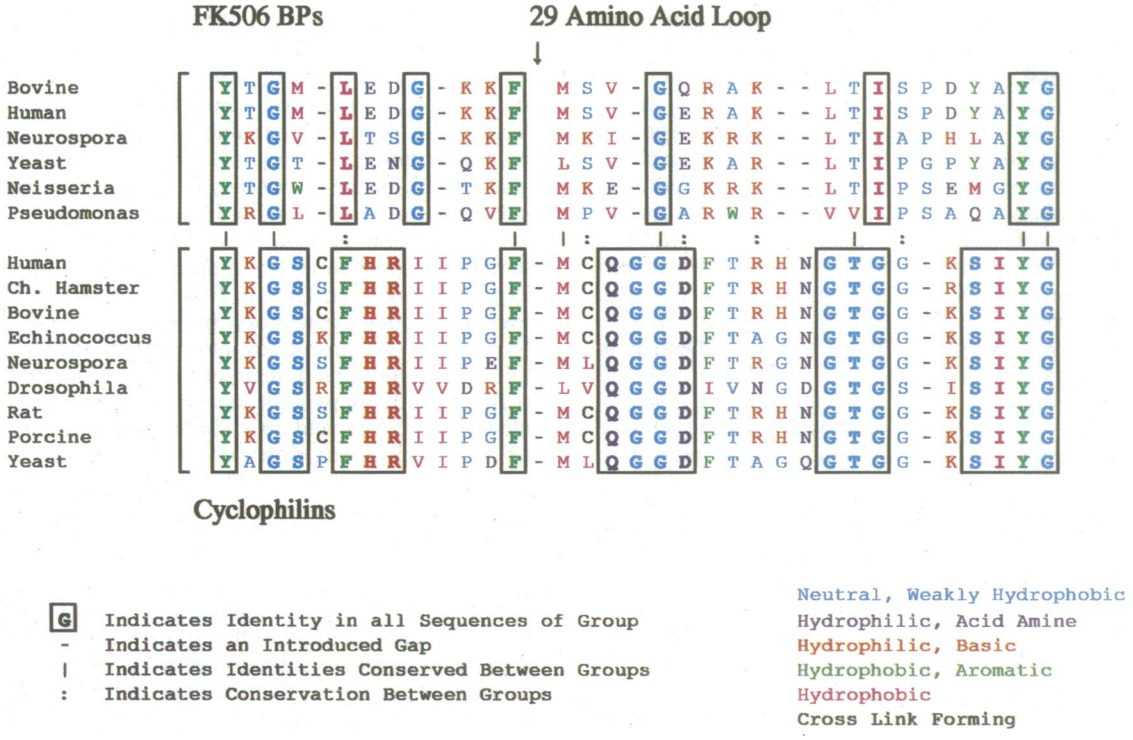

FIG. 4. Alignment of a homologous region in the FKBPs (above) with the cyclophilins (below). Explanation of boxed residues, boldfaced residues, and colors of the various amino acids is shown below the figure. Identities (|) and similarities (:) between the two groups are from the consensus sequences derived from each group. Amino acids were determined similar according to the normalized (25) Dayhoff matrix (26).

more localized. Closer scrutiny of these conserved regions within the FKBPs reveals two domains that, when combined, are conserved with respect to a single domain in the cyclophilins. These domains are marked by heavy lines above the sequences in Fig. 3 and are aligned with one another in Fig. 4. The consensus amino acid sequence that is generated from the conserved cyclophilin domain shown in Fig. 4 is as follows:

YKGSX**F**H**R**I**I**PG**F**M**C**QGGD**F**T**R**H**N**GTG**G**KSI**Y**G,

where residues in boldface type represent absolute identity in all nine cyclophilin sequences; an underlined residue represents an amino acid that predominates at that position and also means that all amino acids at that specific position are chemically similar according to the Dayhoff matrix. Residues that are neither in boldface nor underlined represent amino acids that predominate at that position but where there is at least one nonsimilar amino acid. X represents no conservation. Of particular note is the fact that the methionine residue is present in all eight of the true cyclophilins and is replaced by a leucine (a conserved change) in the *Drosophila ninaA* sequence.

The consensus amino acid sequence generated from the combined FKBP domains in Fig. 4 is the following:

YTGX**L**E**D**GKK**F**\29 amino acids**M**SVGERA**K**L**T**IPPXX**A**Y**G**

The residues are as described for the cyclophilin consensus sequence above, except that an underlined residue means that the identical amino acid is present in at least half of the FKBPs at that particular position and that all residues at that position are similar. As in the cyclophilins, the methionine is present in all but one of the FKBPs, where it is replaced by a leucine. The threonine in the C-terminal portion of the sequence is present in five of six FKBPs. The number of amino acids between the phenylalanine and methionine in FKBP is well conserved between the various species—30 in the bovine protein, 29 in the human, yeast, *Neurospora*, and *Neisseria* proteins, and 25 in the *Pseudomonas* protein. When the 29 amino acids between the phenylalanine and methionine are removed from the FKBP consensus sequence, the FKBP and cyclophilin consensus sequences can be aligned as follows:

```
Cyclophilin YKGSXFHRIIPGFMCQGGDFTRHNGTGGK-SIYG
                | |   :       ||:  |:  :   | :    ||
FKBP        YTGX-LEDG-KKFMSV-GERAK--LTIPPXXAYG,
```

where dashes represent introduced gaps to align the sequences, vertical bars mark identity between residues, and colons represent a conserved change between the sequences. Of the 29 amino acids of the FKBP consensus that are aligned, there is 27.5% identity and 46% similarity with cyclophilin. The significance of this homology is made even more compelling by the fact that relatively rarer amino acids (phenylalanine, tyrosine, and methionine) can be aligned. We theorize that one or more of the conserved activities of FKBP and cyclophilin resides within this site. Possible candidates for such conserved activities are a common PPIase domain or a binding site for interaction with another common polypeptide involved in signal transduction. Because cyclophilin and FKBP act through the formation of the same transition-state intermediate between the cis and trans rotamer conformations of proline (39, 40), the likelihood that the consensus sequences represent the PPIase domain is high. The lack of even greater homology may be due to the observation that FKBP and cyclophilin have distinct substrate specificities [related to the amino acid preceding the proline (40)], thereby suggesting sites of unique interactions within a common structural motif. To account for the observation that analogs of FK 506 and CsA can show a discrepancy between immunosuppressive and PPIase inhibitory activity, we have recently postulated that cyclophilin and FKBP may complex with other polypeptides to exert their immunosuppressive activity (41). Assuming a common site of interaction between FKBP and cyclophilin, the conserved domain could represent the binding site for such a protein.

We thank Drs. K. Bostian, R. Goldberg, D. Linemeyer, E. O'Neill, S. Parent, S. Streicher, and M. Tocci for their helpful comments on this manuscript.

1. Kino, T., Hatanake, H., Miyata, S., Inamura, N., Nishiyama, M., Yajima, T., Goto, T., Okuhara, M., Kohsaka, M., Aoki, H. & Ochai, T. (1987) *J. Antibiot. (Tokyo)* **40**, 1256–1265.
2. Tocci, M., Matkovich, D., Collier, K., Kwok, P., Dumont, F., Lin, S., DeGudicibus, S., Siekierka, J. J., Chin, F. & Hutchinson, N. I. (1989) *J. Immunol.* **143**, 718–726.
3. Handschumacher, R., Harding, M., Rice, J., Drugge, R. & Speicher, D. (1984) *Science* **226**, 544–547.
4. Siekierka, J., Hung, S., Poe, M., Lin, C. & Sigal, N. (1989) *Nature (London)* **341**, 755–757.
5. Harding, M. W., Galat, A., Uehling, D. E. & Schreiber, S. L. (1989) *Nature (London)* **341**, 758–760.
6. Koletsky, A., Harding, M. & Handschumacher, R. (1986) *J. Immunol.* **137**, 1054–1059.
7. Siekierka, J., Wiederrecht, G., Greulich, H., Boulton, D., Hung, H., Cryan, J. & Sigal, N. (1990) *J. Biol. Chem.* **265**, 21011–21015.
8. Takahashi, N., Hayano, T. & Suzuki, M. (1989) *Nature (London)* **337**, 473–475.
9. Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, Th. & Schmid, F. (1989) *Nature (London)* **337**, 476–478.
10. Haendler, B., Keller, R., Hiestand, P., Kocher, H., Wegmann, G. & Movva, M. (1989) *Gene* **83**, 39–46.
11. Danielson, P., Forss-Petter, S., Brow, M., Calavetta, L., Douglass, F., Milner, R. & Sutcliffe, J. (1988) *DNA* **7**, 261–267.
12. Maki, N., Sekiguchi, G., Nishimaki, J., Miwa, K., Hayano, T., Takahashi, N. & Suzuki, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5440–5443.
13. Standaert, R., Galat, A., Verdine, G. & Schreiber, S. (1990) *Nature (London)* **346**, 671–674.
14. Tropschug, M., Wachter, E., Mayer, S., Schonbrunner, E. & Schmid, F. (1990) *Nature (London)* **346**, 674–677.
15. Itoh, N., Tanaka, N., Mihashi, S. & Yamashina, I. (1986) *J. Biol. Chem.* **262**, 3132–3135.
16. Sherman, F., Fink, G. & Hicks, J. (1979) *Methods in Yeast Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
17. Sikorski, R. & Hieter, P. (1989) *Genetics* **122**, 19–27.
18. Deveraux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
19. Bilofsky, H., Burks, C., Fickett, J., Goad, W., Lewitter, F., Rindone, W., Swindell, C. & Tung, C. (1986) *Nucleic Acids Res.* **14**, 1–4.
20. Hamm, G. & Cameron, G. (1986) *Nucleic Acids Res.* **14**, 5–10.
21. George, D., Barker, W. & Hunt, L. (1986) *Nucleic Acids Res.* **14**, 11–16.
22. Pearson, W. & Lipman, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
23. Higgins, D. & Sharp, P. M. (1988) *Gene* **73**, 237–244.
24. Hein, J. (1989) *Methods Enzymol.* **183**, 626–644.
25. Gribskov, M. & Burgess, R. (1986) *Nucleic Acids Res.* **14**, 6745–6763.
26. Schwartz, R. & Dayhoff, M. (1979) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), pp. 353–358.
27. Fickett, J. (1982) *Nucleic Acids Res.* **10**, 5303–5318.
28. Dobson, M., Tuite, M., Roberts, N., Kingsman, A., Kingsman, S., Perkins, R., Conroy, S., Dunbar, B. & Fothergill, L. (1982) *Nucleic Acids Res.* **10**, 2625–2637.
29. Bennetzen, J. & Hall, B. (1982) *J. Biol. Chem.* **257**, 3026–3031.
30. Rothstein, R. (1983) *Methods Enzymol.* **101**, 202–211.
31. Haendler, B., Gofer-Warbinek, R. & Hofer, E. (1987) *EMBO J.* **6**, 947–950.
32. Bergsma, D. & Sylvester, D. (1990) *Nucleic Acids Res.* **18**, 200.
33. Harding, M., Handschumacher, R. & Speicher, D. (1986) *J. Biol. Chem.* **261**, 8547–8555.
34. Lightowlers, M., Haralambous, A. & Rickard, M. (1989) *Mol. Biochem. Parasitol.* **36**, 287–290.
35. Tropschug, M. (1990) *Nucleic Acids Res.* **18**, 190.
36. Schneuwly, S., Shortridge, R., Larrivee, D., Ono, T., Ozaki, M. & Pak, W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5390–5394.
37. Perry, A., Nicolson, I. & Saunders, J. (1988) *J. Bacteriol.* **170**, 1691–1697.
38. Kato, J., Chu, L., Kitano, K., DeVault, J., Kimbara, K., Chakrabarty, A. & Misra, T. (1989) *Gene* **84**, 31–38.
39. Harrison, R. & Stein, R. (1990) *Biochemistry* **29**, 1684–1689.
40. Harrison, R. & Stein, R. (1990) *Biochemistry* **29**, 3813–3816.
41. Sigal, N., Siekierka, J. & Dumont, F. (1990) *Biochem. Pharmacol.* **40**, 2201–2208.