THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 279, No. 9, Issue of February 27, pp. 8441–8451, 2004
Printed in U.S.A.

NMR Solution Structure of the Focal Adhesion Targeting Domain of Focal Adhesion Kinase in Complex with a Paxillin LD Peptide

EVIDENCE FOR A TWO-SITE BINDING MODEL*[S]

Received for publication, September 4, 2003, and in revised form, November 18, 2003
Published, JBC Papers in Press, December 7, 2003, DOI 10.1074/jbc.M309808200

Guanghua Gao‡, Kirk C. Prutzman‡, Michelle L. King§, Danielle M. Scheswohl§, Eugene F. DeRose¶, Robert E. London¶, Michael D. Schaller§‖‡‡§§, and Sharon L. Campbell‡‖¶¶**

*From the ‡Department of Biochemistry and Biophysics, the §Department of Cell and Developmental Biology, ‖Lineberger Comprehensive Cancer Center, **Comprehensive Center for Inflammatory Disorders, and ‡‡Carolina Cardiovascular Biology Center, University of North Carolina, Chapel Hill, North Carolina 27599 and ¶Laboratory of Structural Biology, NIEHS, National Institutes of Health, Research Triangle Park, North Carolina 27709*

Focal adhesion kinase (FAK) is a non-receptor tyrosine kinase that is regulated by integrins. Upon activation, FAK generates signals that modulate crucial cell functions, including cell proliferation, migration, and survival. The C-terminal focal adhesion targeting (FAT) sequence mediates localization of FAK to discrete regions in the cell called focal adhesions. Several binding partners for the FAT domain of FAK have been identified, including paxillin. We have determined the solution structure of the avian FAT domain in complex with a peptide mimicking the LD2 motif of paxillin by NMR spectroscopy. The FAT domain retains a similar fold to that found in the unliganded form when complexed to the paxillin-derived LD2 peptide, an antiparallel four-helix bundle. However, noticeable conformational changes were observed upon the LD2 peptide binding, especially the position of helix 4. Multiple lines of evidence, including the results obtained from isothermal titration calorimetry, intermolecular nuclear Overhauser effects, mutagenesis, and protection from paramagnetic line broadening, support the existence of two distinct paxillin-binding sites on the opposite faces of the FAT domain. The structure of the FAT domain-LD2 complex was modeled using the program HADDOCK based on our solution structure of the LD2-bound FAT domain and mutagenesis data. Our model of the FAT domain-LD2 complex provides insight into the molecular basis of FAK-paxillin binding interactions, which will aid in understanding the role of paxillin in FAK targeting and signaling.

Focal adhesion kinase (FAK),[1] is an essential protein tyrosine kinase as *fak*$^{-/-}$ mice exhibit embryonic lethality (1). FAK signaling is regulated by the interactions of transmembrane receptors, integrins, with their extracellular matrix ligands (2–4). Integrin signaling via FAK controls several critical biological processes including cell spreading, cell migration, and cell survival. In addition to its role in the control of these functions in normal cells, FAK may also function in the pathology of human disease, *e.g.* cancer (5–10). FAK is overexpressed in a number of tumors, including cancers of the breast and prostate. Furthermore, experimental overexpression of FAK can promote some phenotypes of cancer cells, including growth in soft agar and invasion (11–14). Attenuation of endogenous FAK signaling in some cancer cells can impair anchorage-independent growth, motility, and invasion (15–18). These investigations demonstrate that FAK is fundamentally important in normal development and that aberrant FAK signaling may contribute to carcinogenesis.

FAK discretely co-localizes with integrins at focal adhesions, sites of close contact with the extracellular matrix (2, 3). The sequences within FAK responsible for localization have been mapped to ~140 residues near the C terminus, a region defined as the focal adhesion targeting (FAT) sequence or domain (19, 20). Two focal adhesion-associated proteins interact with FAK via the FAT domain, paxillin and talin (19, 21, 22). The precise mechanism of focal adhesion targeting has not been elucidated. Evidence both for and against paxillin and talin as the sole mediators of FAK localization has been presented in the literature (19, 21, 23–25). This discrepancy may be resolved by postulating multiple mechanisms for localization, *i.e.* a paxillin-dependent and a talin-dependent mechanism. Proper subcellular localization of FAK is essential for integrin-dependent regulation and for transmission of downstream signals (24, 26). Furthermore, the C-terminal domain of FAK has been used as a potent dominant negative mutant to inhibit signaling via endogenous FAK (27). Thus, this region of FAK is important for biochemical and biological signaling and is potentially a target for therapeutic intervention.

Paxillin is also an essential protein as the *paxillin*$^{-/-}$ mice are embryonically lethal (28). Paxillin has been implicated in

* This work was supported by National Institutes of Health Grants CA90901 (to M. D. S.), DE13079 (to M. D. S.), and HL45100 (to M. D. S. and S. L. C.).

[S] The on-line version of this article (available at http://www.jbc.org) contains Figs. 1 and 2.

The atomic coordinates and structure factors (code 1qvx) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§§ To whom correspondence may be addressed: Dept. of Cell and Developmental Biology, CB 7090, University of North Carolina, Chapel Hill, NC 27599. Tel.: 919-966-0391; Fax: 919-966-1856; E-mail: crispy4@med.unc.edu.

¶¶ To whom correspondence may be addressed: Dept. of Biochemistry and Biophysics, CB 7260, University of North Carolina, Chapel Hill, NC 27599. Tel.: 919-966-7139; Fax: 919-966-2852; E-mail: campbesl@med.unc.edu.

[1] The abbreviations used are: FAK, focal adhesion kinase; FAT, focal adhesion targeting; ITC, isothermal titration calorimetry; NOE, nuclear Overhauser effect; NOSEY, nuclear Overhauser effect spectroscopy; GST, glutathione *S*-transferase; r.m.s.d., root mean square deviation; PDB, protein data bank.

controlling a subset of the biological processes under the control of FAK, including cell spreading and cell migration (28, 29). In several different scenarios, paxillin has been implicated in the control of motility. The $\alpha_4\beta_1$ integrin promotes cell motility, and paxillin was identified as a direct binding partner of the cytoplasmic tail of the α_4 subunit (30). A mutant of α_4 that is defective for paxillin binding fails to promote cell motility. This interaction is also required for $\alpha_4\beta_1$-dependent tyrosine phosphorylation of FAK and Pyk2, and these two tyrosine kinases were implicated in regulating motility downstream of $\alpha_4\beta_1$ (30, 31). Tyrosine phosphorylation of paxillin may play a role in regulating cell motility. A mutant of paxillin completely defective for tyrosine phosphorylation functions as a dominant negative mutant and in one system promotes motility, whereas in a second system it inhibits motility (32, 33). Although the discrepancy between these studies remains to be completely resolved, the evidence supports a role for tyrosine phosphorylation in the regulation of motility by paxillin. FAK and Pyk2 can regulate tyrosine phosphorylation of paxillin, directly or indirectly. Furthermore, the association of FAK with paxillin may be a mechanism to promote maximal phosphorylation of paxillin (34). These findings demonstrate the important interplay between FAK/Pyk2 and paxillin that appears to control tyrosine phosphorylation and the regulation of cell motility.

Paxillin is an adaptor protein with docking sites for a variety of binding partners (35). The C-terminal half of paxillin contains four LIM domains, which are domains that function in mediating protein-protein interactions. The third LIM domain is vital for paxillin function as it contains the major focal adhesion localization determinant within the molecule (36). The binding partner promoting localization has not been identified. The N-terminal half of paxillin contains sites of tyrosine phosphorylation and five copies of a leucine-rich peptide motif called the LD motif, with a consensus sequence LD*X*LL*XX*L (35, 37). This peptide sequence also functions in mediating protein-protein interactions. Many paxillin-binding partners associate via these LD motifs including FAK, vinculin, PKL, actopaxin, ILK, and the E6 oncoprotein of papillomavirus (35, 37). The binding of FAK to paxillin is mediated by two LD motifs, called LD2 and LD4 (36). Although LD motif-mediated interactions appear important for the function of paxillin, neither the molecular and structural basis for these interactions nor the binding specificity have been elucidated.

Recent structural analysis of the FAT domain of FAK has demonstrated that this domain forms an anti-parallel four-helix bundle (38–40). Two hydrophobic patches on opposite faces of the FAT domain were initially proposed as docking sites for LD motifs of paxillin (39). Hydrophobic patch 1 (HP1) was at the interface of helix 2 and 3, and HP2 was at the interface of helix 1 and 4. In this model, the LD motifs were assumed to be α-helical in nature, and two LD motifs were proposed to bind to the FAT domain, one to each hydrophobic patch (39). In contrast, Arold *et al.* (38) proposed a single paxillin-binding site at the interface between helix 2 and 3 (HP1) based upon their crystal structure along with some previously reported mutagenesis data. A subsequent solution structure of FAT in complex with an LD motif peptide fundamentally contradicted the original model (40). The LD2-docking site on FAT was identified using a paramagnetically labeled peptide, as no interactions could be detected between the peptide and the FAT domain by NOE-based approaches. Although evidence that the LD motif peptide was α-helical in solution was presented, other aspects of the interaction were different. Their results indicated that the FAT domain contains a single LD motif-binding site at HP2 (40). Furthermore, the peptide docked in the opposite orientation relative to the model of Hayashi *et al.* (39), which was based on the crystal structure of the FAT domain. Although these studies have provided important insight into the structure of the FAT domain, there are a number of unresolved issues, including the details of the interaction with paxillin.

Here we report the solution structure of the avian FAT domain in complex with a paxillin LD2 peptide solved by NMR spectroscopy. Our study demonstrates that two sites containing either HP1 or HP2, which are on the opposite faces of the FAT domain, function as paxillin LD motif-binding sites with similar binding affinities, yet the thermodynamic basis for binding to the two sites is fundamentally different. Our experimentally derived models of FAT domain-LD2 complexes provide important insight into the structural basis of paxillin recognition.

EXPERIMENTAL PROCEDURES

Expression and Purification of the FAT of FAK and Peptide Synthesis—The FAT sequence, containing residues 920–1053 of the avian FAK plus a 12-amino acid N-terminal linker, was expressed and purified as previously described (34).[2] A 27-amino acid residue LD2 peptide (^{133}MTSTSLGSNLSELDRLLLELNAVQHNP^{159}) was chemically synthesized at the University of North Carolina microprotein sequencing and peptide synthesis facility. A [^{15}N]leucine-labeled LD2 peptide and a shorter LD2 peptide lacking 8 N-terminal residues (residues 141–159) were also synthesized. Protein and peptide concentrations were measured using the BCA assay (Pierce) because the results lie within 5% of the determinations by amino acid composition analysis.

Isothermal Titration Calorimetry (ITC)—Protein solutions were dialyzed thoroughly against ITC buffer (50 mM HEPES, 100 mM NaCl, and 2 mM NaN_3 at pH 7.0) and degassed prior to ITC measurements. For experiments at different temperatures, the pH was corrected to ensure constant pH. Titrations were conducted at the University of North Carolina Macromolecular Interactions Facility using a Microcal VP-ITC. The FAT domain or FAT domain variants (~50 μM) were placed in the calorimetric cell and titrated with ~1 mM LD2 peptide in the ITC buffer. The heat of association was obtained from the difference between the integrated heat of reaction and the corresponding heat of dilution. Data were analyzed using the ITC version of Origin 5.0 with embedded calorimetric fitting routines, with ΔH (the change in binding enthalpy, cal/mol) and K_a (the binding constant, M^{-1}) as adjustable parameters. An ITC study was also conducted on a FAK homolog, Pyk2.

Mapping the Binding Interface from Line Broadening by Gd-EDTA Paramagnetic Probe—Gd(III)-EDTA, a paramagnetic relaxation broadening reagent, was titrated into samples of uniformly ^{15}N-enriched free and LD2-bound ^{15}N-enriched FAT domain samples, and the effects of line broadening were compared as the difference in normalized height of amide HSQC peaks (42). Two dimensional ^1H_N-^{15}N HSQC spectra of 0.2 mM ^{15}N-labeled FAT domain samples, free or in complex with the LD2 peptide (molar ratio 1:8), were collected in the presence of a range of concentrations of 1:1 Gd(III)-EDTA (0.6, 1.2, 1.8, 2.4, 3.0, 6.0, and 9.0 mM). Line broadening was measured as a ratio of the height of a backbone amide peak at a given concentration of Gd(III)-EDTA to the height in the absence of the paramagnetic probe. Normalized heights of NH peaks obtained from the HSQC spectrum of the free FAT domain in the presence of 3 mM Gd(III)-EDTA were subtracted from normalized heights of NH peaks taken from the HSQC spectrum of the FAT domain-LD2 complex in the presence of 3 mM Gd(III)-EDTA.

Interactions with Full-length Paxillin—GST fusion proteins were induced as described above. After immobilization on glutathione beads, the fusion proteins were used to test paxillin binding. CE cells were lysed in Tx-RIPA buffer (50 mM Tris, pH 7.3, 150 mM NaCl, 1% Triton X-100, 0.5% deoxycholate) (34), and the protein concentration was determined by using the BCA assay. Lysates were pre-cleared with 50 μg of GST per 1 mg of lysate for 1 h at 4 °C. Then 500 μg of lysate was incubated with 6 μg of GST fusion proteins for 1 h at 4 °C. The beads were washed twice with Tx-RIPA and twice with phosphate-buffered saline, and bound proteins were eluted by boiling in Laemmli's sample buffer. The samples were analyzed by Western blotting using a paxillin monoclonal antibody (BD Transduction Laboratories, Lexington, KY).

[2] K. C. Prutzman, G. Gao, M. L. King, V. V. Iyer, G. A. Mueller, M. D. Schaller, and S. L. Campbell, submitted for publication.

For expression in 293 cells, wild type FAK and mutants were subcloned into the pcDNA3 expression vector (Clontech). Cells were transfected using Lipofectamine according to the manufacturer's recommendations (Invitrogen). Seventy two hours post-transfection, cells were lysed in Tx-RIPA buffer and protein concentrations determined using the BCA assay. FAK was immunoprecipitated from 1 mg of lysate using the BC4 polyclonal antiserum (26, 43). Immune complexes were washed twice with Tx-RIPA buffer and twice with phosphate-buffered saline, and bound proteins were eluted with sample buffer. Bound paxillin was detected by Western blotting by using a paxillin monoclonal antibody. The blot was then stripped and re-probed for FAK as a loading control.

NMR Samples—Purified proteins were exchanged into NMR buffer (20 mM deuterated Tris maleate (Isotec, Miamisburg, OH), 50 mM NaCl, 0.1 mM EDTA, 0.1% NaN_3, and 10% D_2O, pH 6.0), using centrifugal filter devices with a 10,000-dalton molecular weight cut-off BioMax membrane (Millipore, Bedford, MA). Lyophilized LD2 peptide (protein/peptide molar ratio 1:8) was added into diluted protein solution (~1.5 mg/ml). NMR samples of the FAT domain-LD2 complex were then concentrated to desired concentrations using an Amicon stirred ultrafiltration cell (model 8003) equipped with a 1,000-dalton molecular weight cut-off YM1 ultrafiltration membrane (Millipore, Bedford, MA). 1 μM D-Phe-Pro-Arg-chloromethylketone (Calbiochem) and 0.5 mg/ml pefabloc SC (Roche Applied Science) were used in all NMR samples to inhibit residual thrombin remaining from GST cleavage during purification.

NMR Spectroscopy—All NMR experiments were carried out at 310 K. ^{1}H chemical shifts were referenced relative to an external 2,2-dimethyl-2-silapentane-5-sulfonate sample, whereas ^{13}C and ^{15}N chemical shifts were referenced indirectly to 2,2-dimethyl-2-silapentane-5-sulfonate as described (44). Spectra were processed with NMRPipe program version 2.1 (45) and analyzed by NMRView program version 5.0.4 (46). A three-dimensional simultaneous ^{13}C/^{15}N-edited NOESY spectrum (75-ms mixing time) was collected on a Varian Inova 800 MHz spectrometer (47). To identify intermolecular contacts between the FAT domain and LD2 peptide, a pair of three-dimensional simultaneous ^{13}C/^{15}N edited NOESY spectra (100 ms mixing time) were also acquired under identical conditions on a Varian Inova 600 MHz spectrometer, with or without F1 decoupling during acquisition. All other spectra were collected on a Varian Inova 600-MHz spectrometer. Backbone sequential assignments were obtained manually from the following series of experiments on a 0.6 mM uniformly ^{13}C/^{15}N-labeled FAT protein with 4.8 mM unlabeled LD2 peptide: two-dimensional ^1H_N-^{15}N TROSY HSQC (see representative spectrum in Supplemental Material Fig. 1), two-dimensional ^{1}H-^{13}C HSQC, three-dimensional HNCA (48), three-dimensional HN(CO)CA (49), three-dimensional HNCACB (48, 50), three-dimensional CBCA(CO)NH (51), and three-dimensional HNCO (52). Two dimensional ^1H_N-^{15}N TROSY HSQC spectra were also collected on a 0.2 mM selectively [^{15}N]Leu-labeled protein in complex with LD2 peptide, and a 0.2 mM [^{15}N]Lys(−)-labeled protein in complex with LD2 peptide (protein/peptide molar ratio 1:8). These two spectra provided starting points for the assignments. Side chain assignments were achieved using the following experiments: three-dimensional HCCH-COSY (53), three-dimensional HCCH-TOCSY (54), three-dimensional H(CCO)-TOCSY-NNH, and three-dimensional (H)C(CO)-TOCSY-NNH (55, 56). Aromatic side chain assignments were obtained from two-dimensional (Hβ)Cβ(CγCδCϵ)Hϵ and two-dimensional (Hβ)Cβ(CγCδ)Hδ experiments (57). ^1H_N-^{15}N residual dipolar coupling values were measured using 7.5 mg/ml Pf1 phage as aligning medium (ASLA Ltd., Riga, Latvia). The protein concentration was 0.25 mM, and peptide concentration was 2 mM. No significant spectral difference was observed in two-dimensional ^1H_N-^{15}N HSQC spectra with and without phage. Two dimensional spectra were collected using gNhsqc_IPAP pulse sequence from ProteinPack (Varian) on the complex samples with and without Pf1 phage (58).

Structure Calculations—Results of the CSI predictions and the TALOS program were analyzed for identification of secondary structure elements (59–61). The crystal structure (Protein Data Bank access code 1k40) from Hayashi *et al.* (39) served as starting structure for calculations after minor modifications, because of small differences in the primary sequence between mouse and the avian FAT domain. The program Insight II (Accelrys, San Diego, CA) was used to append the 12-amino acid residue linker to the N terminus of the avian FAT domain and to substitute the following amino acid residues: T972S, I973L, A975V, G996A, S1000N, L1046I, G1047S, and T1049S. Dihedral angle constraints were derived from the Cα chemical shifts by using the program TALOS, which matches sequence and chemical shifts with a data base of known structures and chemical shifts (61). Using the criteria of TALOS program, only "good" matches of dihedral angles were used for restraints, and errors were set as $\pm 2\sigma$ or at least $\pm 20°$. The hydrogen bonds were assigned based on manually assigned NOEs and the secondary structure predictions by TALOS. In total, 212 hydrogen bond and 92 dihedral angle constraints were used in structure calculations. RDC values (−5.9 to 9 Hz) were fit to our preliminary NMR structure of the FAT domain-LD2 complex using the RDCA program (62), and a rhombicity value of 0.5075 was obtained. Uncertainties of $^1D_{NH}$ RDC were estimated to be 1.5 Hz.

The program CNS version 1.1 with ARIA version 1.2 module was used to calculate the structures of the avian FAT domain as described below (63–66). The structural restraints included NOE, hydrogen bond, dihedral angle, and RDC restraints. The NOE distance restraints were derived from a three-dimensional simultaneous ^{13}C/^{15}N-edited NOESY peak lists with 75-ms mixing time. The peak lists were generated automatically by NMRView and edited manually to remove any obvious water and apodization artifacts. The peak lists were unassigned and uncalibrated with respect to distance. To calibrate the NOE distances, the rotational correlation time of the complex, τ_c, was set to 10 ns and was based on τ_c values obtained on proteins of this size at similar temperature. The default parameters for ARIA produced structures with poor convergence. Several parameters were optimized, including the ambiguous cut-off (ρ), the violation tolerance (v_{tol}), and maximum number of ambiguities per peak (n_{max}), to provide better noise discrimination. The following scheme led to the best convergence over nine iterations: ρ = (0.95, 0.95, 0.95, 0.94, 0.93, 0.92, 0.91, 0.90, and 0.80), v_{tol} = (5.0, 2.0, 1.0, 0.5, 0.25, 1.0, 0.1, 0.1, and 0.1), n_{max} = 5 for all iterations. An ensemble of 25 calculated NMR structures of the avian FAT domain was selected for further analysis.

Docking Protocol—Model 3 from the family of 25 accepted structures of the FAT domain in the LD2 peptide bound state was selected as a representative set of coordinates. The flexible N-terminal 12-residue linker was removed. The structured region of LD2 peptide (^{138}LGSNLSELDRLLLEL^{152}) in the FAT domain-LD2 complex was determined from the analysis of a two-dimensional ^1H_N-^{15}N HSQC spectrum of a complex of unlabeled FAT domain and selectively [^{15}N]leucine-labeled LD2 peptide (see HSQC spectra in Supplemental Material Fig. 2). Upon binding, the amide peaks of 6 of 7 leucine residues were clearly broadened. We assume the amide peak at 8.3 and 120.4 ppm, which are the narrowest and most intense leucine NH peaks that do not shift in the spectra of both free and bound states, correspond to leucine residues at the ends of the peptide sequence. Furthermore, consistent with the assumptions made in the previous studies (39, 40), we assume the peptide residues 138–152 adopt a helical conformation when complexed to FAT given its strong propensity to form α-helix (39, 40). The initial peptide coordinates were generated in the Discover module of the Insight II program using standard helical conformation, followed by 1,000 steps of molecular dynamics and 3,000 steps of energy minimization in the CHARM force field. The docking process was carried out using the HADDOCK (high ambiguity driven docking) program (67). Docking of the LD2 peptide onto helix 1, 4 (HP2) and helix 2, 3 (HP1) sites of the FAT domain were conducted separately. The side chains of the LD2 peptide were allowed to move to allow for conformational rearrangements. Because both chemical shift perturbation and paramagnetic probe mapping were not able to define precisely the interaction sites, only mutagenesis data were used to drive the docking process. According to established criteria of the HADDOCK program, the "active" residues are those that have been shown by mutations to abolish or significantly perturb complex formation and are solvent-exposed. The passive residues correspond to the residues that are surface neighbors of the active residues and also have a high solvent accessibility (67). FAT residue Ile-937 in helix 1, Glu-949, Lys-956, and Arg-963 in helix 2, Asn-992 and Lys-1003 in helix 3 (24, 39), and Lys-1019, His-1026, and Lys-1033 in helix 4 are therefore considered to be active residues according to our ITC results and previous mutagenesis studies. For the LD2 peptide, only the critical Asp-146 residue is identified as an active residue. Four neighboring LD2 residues, Leu-142, Ser-143, Leu-149, and Leu-150, are considered to be passive residues. Relative solvent accessibility of all these residues is larger than 70% as calculated with NACCESS (68). The default HADDOCK parameters were used except only 200 initial complex structures were generated, and the best 25 solutions in terms of intermolecular energies were then refined in water (default values were 1,500 and 200, respectively) (67).

RESULTS AND DISCUSSION

Characterization of FAT Domain Binding to LD2 Peptide by ITC—To characterize the association of paxillin with FAK, the interaction of synthetic peptides mimicking the LD2 motif of

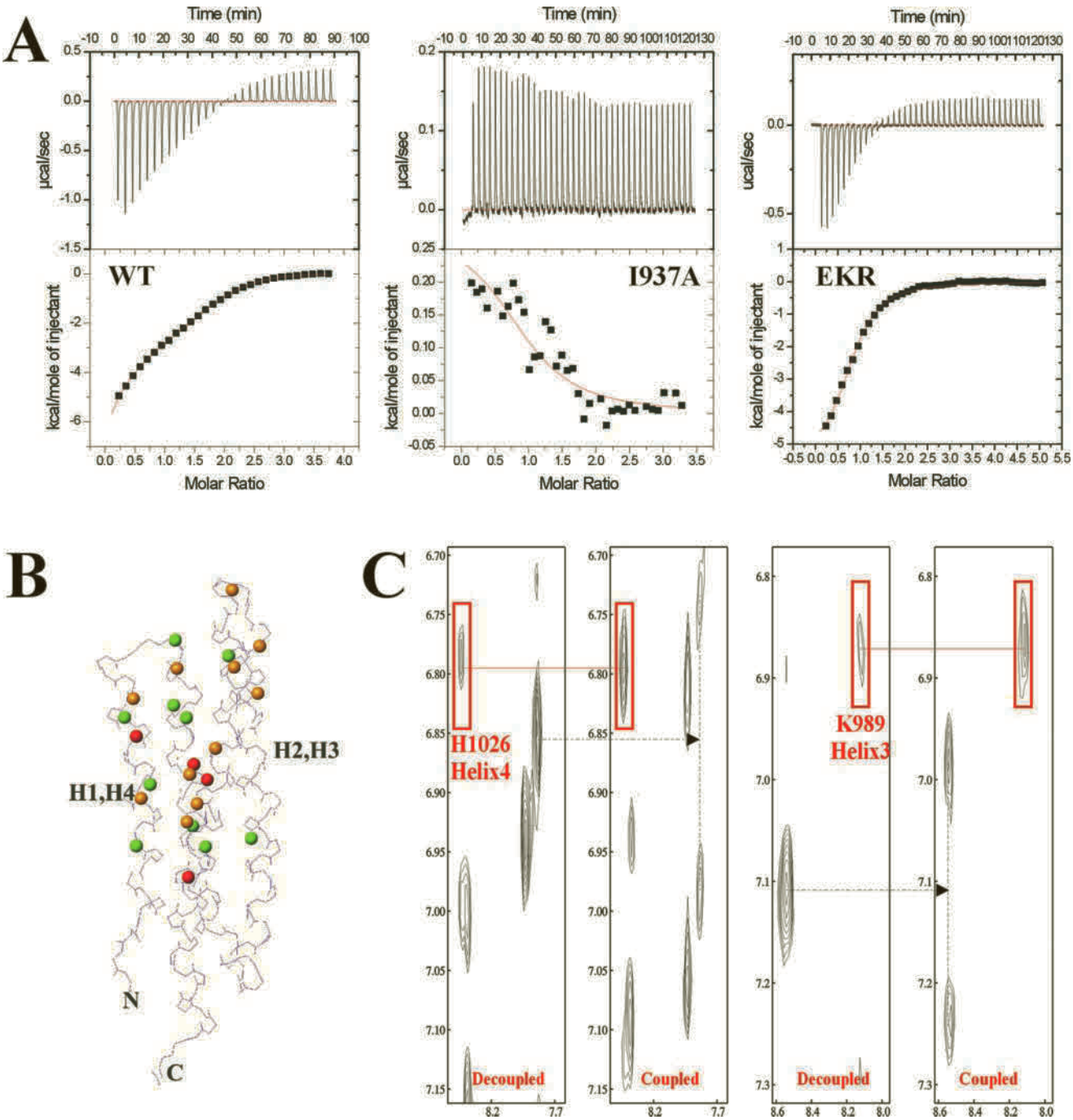

FIG. 1. *A*, calorimetric titrations of wild type FAT, single mutant I937A (helix 1), and triple mutant E949A/K956A/R963A (helix 2) with LD2 peptide in HEPES buffer at pH 7.0, 299 K. *Top panels* show raw data with integration base line (*red*). *Bottom panels* show data after peak integration, subtraction of blank titrations, and concentration normalization. The *red solid lines* in the *bottom panels* represent the fit to a two binding site model for wild type FAT and a single binding site model for I937A and EKR mutant, respectively. Thermodynamic parameters are summarized in Table I. *B*, LD2 peptide contact sites of FAT were mapped by paramagnetic relaxation broadening (see "Experimental Procedures" for details). Amide protons experiencing protection from 3.0 mM Gd(III)-EDTA are drawn as colored balls in the presence of LD2 peptide (*green*) and a shorter LD2 peptide lacking N-terminal 8 residues (*gold*). Amide protons protected with both peptides are shown as *red balls*. *C*, two-dimensional ^{1}H-^{1}H planes taken from a pair of three-dimensional simultaneous $^{13}C/^{15}N$ edited NOESY spectra with and without F1 decoupling during acquisition. Intermolecular NOEs between labeled FAT and unlabeled LD2 peptide are singlets in both spectra, whereas intramolecular NOEs for labeled FAT are doublets in the coupled spectrum. The *boxed peaks* correspond to intermolecular NOEs between the amide protons of residue His-1026 (helix 4) and Lys-989 (helix 3) to the peptide, respectively.

paxillin with the purified FAT domain of FAK was examined by ITC. Thermodynamic parameters were calculated from Equation 1.

$$\Delta G = -RT\ln K_a = \Delta H - T\Delta S \qquad \text{(Eq. 1)}$$

Representative titrations are shown in Fig. 1*A*, and the thermodynamic parameters of binding are listed in Table I. The experimental data best fit a two-site model for the wild type FAT domain, with fixed stoichiometry (n) of 1 for both sites in order to more accurately fit K_a and ΔH by decreasing the number of variables. The χ^2 value for the goodness of fit of the data to the two-site binding model was 536, compared with 10^5 if the data were fit to a one-site model. The ITC results also demonstrate that the two sites are energetically different, *i.e.*

TABLE I
Thermodynamic parameters of the interaction between FAT variants or Pyk2 and LD2 peptide

	Helix 1,4 site			Helix 2,3 site		
	K_d	ΔH	$-T\cdot\Delta S$	K_d	ΔH	$-T\cdot\Delta S$
	μM	*kcal/mol*		μM	*kcal/mol*	
Wild type FAT	9.0 ± 0.5	−7.2 ± 0.2	0.3 ± 0.2	11.5 ± 0.5	1.1 ± 0.2	−7.8 ± 0.2
I937A				13.6 ± 4.9	0.3 ± 0.1	−6.9 ± 0.3
E949A/K956A/R963A	4.4 ± 0.3	−5.5 ± 0.2	−1.8 ± 0.2			
Pyk2	45 ± 4	−10 ± 1	4 ± 1	21 ± 1	0.3 ± 0.4	−6.7 ± 0.4

one being largely enthalpically driven and the other mainly entropically driven. Both sites have binding affinities of ~10 μM. However, it is important to point out that usually in multiparameter fitting there could be strong correlations between certain variables that are being determined, which can introduce substantial uncertainties in the actual values. So the real uncertainties of the thermodynamic parameters should be much larger than the error ranges given in Table I. Nevertheless, the ITC data from the two FAT mutants as discussed below strongly support the interpretation of the presence of two energetically distinct LD2-binding sites within the FAT domain. ITC studies were also performed at 308 K. The change in heat capacity of binding (ΔC_p) was determined using Equation 2.

$$\Delta C_p = (\Delta H_{T2} - \Delta H_{T1})/(T_2 - T_1) \quad \text{(Eq. 2)}$$

The heat capacity of binding, ΔC_p, was estimated to be −0.3 kcal·mol^{-1}·K^{-1} and −0.4 kcal·mol^{-1}·K^{-1} for the two binding sites. The large negative values of ΔC_p indicate significant hydrophobic interactions are involved in binding of the LD2 peptide at both binding sites. Interestingly, the close family member of FAK, Pyk2, possesses very similar thermodynamic characteristics of LD2 binding to that of the FAT domain (Table I). Pyk2 and FAK share a similar structural organization with a tyrosine kinase domain flanked by noncatalytic domains at both N and C termini. Moreover, the FAT domain and the C-terminal domain of Pyk2 share ~60% sequence identity, and constitutive association of Pyk2 with paxillin has been demonstrated (69–71). Our ITC results support the hypothesis that the C-terminal domain of Pyk2 and the FAT domain are structurally related and may share similar global folding and paxillin binding properties.

A number of previously characterized mutants of the FAT domain have been examined for paxillin binding. These include I937E and I999E, which contain substitutions for conserved hydrophobic residues in helix 1 and 3, respectively, and E949A/K956A/R963A and H1026A/K1033A, each containing substitutions for two conserved positively charged residues (Lys-956 and Arg-963) in helix 2 and 4, respectively. These mutants were capable of binding paxillin *in vitro*, albeit qualitatively less well than wild type FAK (24, 39). For ITC characterization, these mutations were engineered into the GST-FAT sequence construct, with the exception that alanine substitutions were made for Ile-937 and Ile-999, instead of the more dramatic glutamic acid substitution. Each mutant was expressed, purified, and analyzed for LD2 peptide binding by ITC (Fig. 1*A*). In contrast to the data for the wild type FAT domain (Fig. 1*A*, *left panel*), the I937A mutant data fit well only to a one binding site model. From the thermodynamic parameters shown in Table I, it is obvious that I937A mutation abolishes the LD2 binding activity of the helix 1, 4 site and has little effect on the helix 2, 3 binding site. The I999A mutant exhibited LD2 binding activity similar to wild type FAT (data not shown). The ITC binding data for the E949A/K956A/R963A mutant also fits well only to a one binding site model. This mutation disrupts the LD2 binding activity of helix 2, 3 site, leaving the helix 1, 4 site almost intact. In contrast, the H1026A/K1033A mutant decreases the LD2 binding affinity of the helix 1, 4 site by just 2-fold. A second mutation, K1019A/H1026A/K1033A, has a significant effect on the binding affinity of the helix 1, 4 site and decreases the binding affinity of this site by ~10-fold (data not shown).

Mapping the Paxillin-binding Sites on FAT Using Gadolinium(III)-EDTA as a Paramagnetic Line Broadening Reagent—Attempts to identify the interaction sites of the LD2 peptide within the FAT domain by chemical shift mapping failed because addition of the peptide induced dramatic chemical shift changes for the majority of the residues of the FAT domain, consistent with previous observations (40). As an alternative strategy, Gd(III)-EDTA was used as a paramagnetic probe to map the binding interface (42). Relative amide peak heights were estimated as a ratio of the height of a particular peak in the two-dimensional ^1H_N-^{15}N HSQC spectrum with paramagnetic probe present to the HN peak height with no paramagnetic probe present. In particular, relative heights for the free FAT domain with 3 mM Gd(III)-EDTA present were subtracted from relative heights for the FAT domain bound to LD2 with 3 mM Gd(III)-EDTA present. The threshold for amide protons considered to be significantly protected was set to 2σ, the standard deviation of all subtracted values. Amide proton resonances associated with the FAT domain residues that were protected from broadening in the FAT domain-LD2 complex are localized to both helix 1, 4 and helix 2, 3 faces of the FAT domain (Fig. 1*B*). These results support the ITC data and provide further evidence for the two-site binding model of the FAT domain-LD2 complex.

Observation of Intermolecular NOEs between the LD2 Peptide and the FAT Domain—In order to observe intermolecular NOEs between protein and ligand, an isotope-labeled, isotope-filtered NOESY experiment was implemented. This technique makes use of ^{13}C and/or ^{15}N labeling in one of the binding components, *i.e.* the FAT domain in this case while leaving its binding partner unlabeled (*i.e.* LD2 peptide). However, no NOE cross-peaks were observed due to exchange broadening of resonances associated with the LD2 peptide in the FAT domain-LD2 complex, consistent with earlier observations (40). To increase the sensitivity associated with detection of intermolecular NOEs, we collected a pair of three-dimensional simultaneous $^{13}C/^{15}N$-edited NOESY spectra of a $^{13}C/^{15}N$ double-labeled FAT domain/unlabeled LD2 peptide complex, with and without F1 decoupling during acquisition. All intramolecular NOE peaks from the labeled FAT domain should appear as doublets in the spectrum without F1 decoupling due to $^1J_{NH}$ or $^1J_{CH}$ coupling, whereas intermolecular NOE peaks involving protons from unlabeled peptide appear as singlets in both spectra (72). Comparison of coupled and decoupled ^{15}N-edited spectra revealed ~10 intermolecular cross-peaks. Although the resulting information could not be applied as NOE constraints during structure calculations due to the inability to assign exchange broadened peaks associated with the LD2 peptide in the complex, we have been able to clearly prove the existence of two LD2 peptide-binding sites on the FAT domain based on the

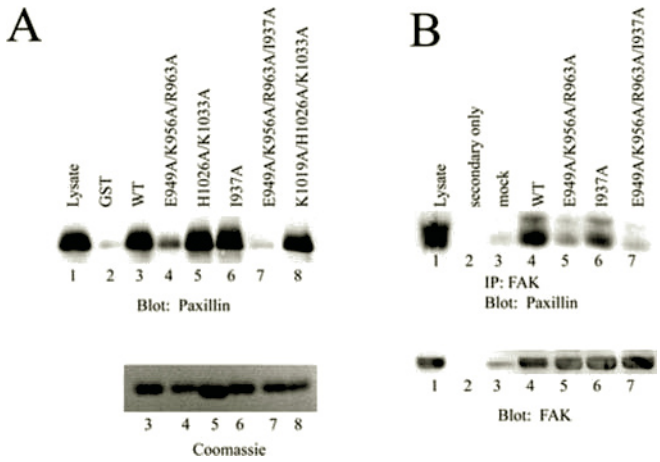

FIG. 2. **Association of FAK mutants with full-length paxillin *in vitro* and *in vivo*.** *A*, CE lysate (500 μg) was precleared with 25 μg of GST and then incubated with 6 μg of GST (*lane 2*) or GST-FAT domain constructs: wild type (*WT*) (*lane 3*), E949A/K956A/R963A (*lane 4*), H1026A/K1033A (*lane 5*), I937A (*lane 6*), E949A/K956A/R963A/I937A (*lane 7*), or K1019A/H1026A/K1033A (*lane 8*). After washing, bound proteins were eluted in Laemmli's sample buffer, and bound paxillin was detected by Western blotting (*top panel*). As a control, 40 μg of lysate was directly loaded (*lane 1*). Comparable amounts of each fusion protein were used in the analysis, as determined by SDS-PAGE and Coomassie Blue staining (*bottom panel*). *B*, wild type FAK (*lanes 2* and *4*), E949A/K956A/R963A (*lane 5*), I937A (*lane 6*) or E949A/K956A/R963A/I937A (*lane 7*) were transiently expressed in 293 cells. Control, mock-transfected cells (*lane 3*), and FAK-expressing cells were lysed and FAK immunoprecipitated (*IP*). Bound paxillin was detected by Western blotting (*top panel*). As a control for nonspecific paxillin binding, protein A Sepharose alone was used as a negative co-immunoprecipitation control (*lane 2*). Lysate (25 μg) was used as a control (*lane 1*). The blot was stripped and reprobed for FAK to verify comparable recovery of FAK variants (*bottom panel*).

available FAT domain assignments using this data. Two planes taken from each of three-dimensional simultaneous $^{13}C/^{15}N$ edited NOESY spectra, with or without F1 decoupling, are shown in Fig. 1*C*. Peaks in the boxes that remain as singlets in the coupled spectra are NOE peaks between the labeled protein and the unlabeled peptide. In contrast, those peaks appearing as doublets in the coupled spectra are NOE peaks within the labeled FAT domain (72). According to our FAT domain assignments, the boxed peaks correspond to close contacts between the amide protons of residue His-1026 in helix 4 and Lys-989 in helix 3 and the LD2 peptides. These two residues (His-1026 and Lys-989) are on the opposite sides of the molecule, providing additional evidence for the presence of two peptide-binding sites in the FAT domain.

Interactions between the FAT Domain and Full-length Paxillin, Both in Vitro and in Vivo—The results of the peptide-binding experiments demonstrate that a synthetic peptide mimicking the LD2 motif of paxillin can dock to two distinct sites on the FAT domain. To explore the role of the two paxillin docking sites in mediating interactions of FAK with paxillin, mutants were expressed as GST-FAT domain fusion proteins, and their capacity to associate with paxillin was examined *in vitro*. Fusion proteins were immobilized on glutathione beads and then incubated with CE cell lysates. Bound paxillin was detected by Western blotting. As shown in Fig. 2*A*, paxillin bound to the wild type FAT domain of FAK but very poorly associated with GST alone. The E949A/K956A/R963A mutant exhibited reduced binding to paxillin compared with wild type, although the mutant still exhibited significant binding relative to the GST control. The I937A mutant bound to paxillin similar to the wild type protein. This was surprising, because the I937A mutant completely abolished binding of the LD2 peptide to the helix 1, 4 binding site. Two other mutants, H1026A/K1033A and K1019A/H1026A/K1033A, that retained LD2 peptide binding to the helix 1, 4 binding site, but with reduced affinity, were also analyzed. Both mutants bound to paxillin similar to wild type. However, mutation of the helix 1, 4 binding site (I937A) in combination with mutation of the helix 2, 3 binding site (E949A/K956A/R963A) was required to ablate paxillin binding activity of the GST-FAT domain, as demonstrated with the E949A/K956A/R963A/I937A mutant.

To further explore the role of the two paxillin-binding sites of the FAT domain in the interaction with paxillin *in vivo*, wild type full-length FAK or FAK mutants were transiently expressed in 293 cells. FAK was immunoprecipitated and co-immunoprecipitated paxillin was detected by Western blotting (Fig. 2*B*). A small amount of paxillin was co-immunoprecipitated with endogenous FAK. A dramatic increase in the amount of associated paxillin was observed in cells expressing exogenous wild type FAK, due to the elevated level of FAK expression. The E949A/K956A/R963A mutant exhibited reduced binding to paxillin relative to wild type FAK, although more paxillin was immunoprecipitated than with endogenous FAK, demonstrating that the mutant retained some paxillin binding activity. I937A exhibited a small reduction in binding activity, but clearly associated with paxillin substantially better than E949A/K956A/R963A. Again, mutation of both paxillin-binding sites in the E949A/K956A/R963A/I937A mutant virtually abolished paxillin binding. These results suggest that the two paxillin docking sites of the FAT domain may not contribute equally to binding full-length paxillin. Disruption of the helix 2, 3 binding site results in a more dramatic reduction in paxillin binding relative to disruption of the helix 1, 4 binding site. However, simultaneous disruption of both binding sites is required to eliminate paxillin binding.

Overall Structure of the FAT Domain—Following well established assignment procedures, 98.0% of the backbone resonances and 94.6% of the side chain resonances were assigned for the structured regions of the FAT domain (residues 920–1053) bound to the LD2 peptide. We have solved the NMR solution structure of the FAT domain in complex with the LD2 peptide at moderately high resolution. Twenty five of the 100 NMR-derived structures with the lowest energies and the smallest numbers of violations against experimental restraints in the structural calculation were selected for further analysis. The quality of the NMR structures is evident in the statistics shown in Table II. Approximately 92% of the residues from 920–1053 appear in the most favored regions of a Ramachandran plot. For backbone atoms of residues in the secondary structures, the average r.m.s.d. to the mean structure is 0.67 ± 0.11 Å. For all heavy atoms of residues 924–1050, the average r.m.s.d. is 0.85 ± 0.16 Å. The FAT domain is organized into a "right-turn" four-helix bundle (Fig. 3). The helices are relatively straight with different lengths: helix 1 is the shortest, whereas helix 4 is the longest among the four helices. All the helices are nearly parallel except helix 4. The helical axis of helix 4 is about 10° from the elongated axis of the helix bundle (Fig. 3, *A* and *B*), which is different from the available structures of either the free FAT domain or the FAT domain bound to a shorter LD2 peptide (Fig. 3*D* and Table III). Notably, the helix bundle has a well defined hydrophobic core shown by the ordered *green side chains* in Fig. 3*A*. The helix bundle is likely stabilized by extensive hydrophobic interactions among these residues from all four helices. In fact, a large number of inter-helical NOEs were observed not only between neighboring helices (helix 1 to helix 2 and helix 3 to helix 4) but also between diagonal helices (helix 1 to helix 3 and helix 2 to helix 4). The hydrophobic core of the helix bundle is indeed highly conserved across species (39). However, the hydrophobic core is inter-

TABLE II
Structural statistics for the family of 25 accepted structures

Distance restraints	
All	2844
H-CH	1257
H-NH	1375
Unambiguous	1362
Ambiguous	1270
Hydrogen bonds	212
Dihedral angle restraints	
All	184
Φ	92
Ψ	92
Dipolar coupling restraints	
$^1D_{NH}$	50
Ramachandran plot	
Residues in most favored regions	91.9%
Residues in additional allowed regions	5.3%
Residues in generally allowed regions	1.9%
Residues in disallowed regions	0.8%
Average violations per structure	
Distance restraints (>0.5 Å)	2.2
Dihedral angle restraints (>5°)	4.4
Mean r.m.s.d. from experimental restraints	
Distance (Å)	0.05 ± 0.01
Dihedral angle (°)	1.39 ± 0.20
Residual dipolar coupling (Hz)	0.96 ± 0.17
Mean r.m.s.d. from idealized geometry	
Bonds (Å)	0.0049 ± 0.0002
Angles (°)	0.72 ± 0.02
Improper (°)	2.10 ± 0.10
R.m.s.d. to the mean structure (Å)	
All residues excluding the termini (924–1050)	
Backbone atoms	0.85 ± 0.16
All heavy atoms	1.69 ± 0.19
Secondary structures (926–941, 950–976, 983–1006, 1016–1048)	
Backbone atoms	0.67 ± 0.11
All heavy atoms	1.57 ± 0.19

rupted by a cluster of four methionines shown in *magenta*, which creates a distinct cavity at the top of the bundle (Fig. 3*A*). Loop 1 connecting helix 1 and helix 2 and loop 3 connecting helix 3 and 4 at the top of the bundle also have larger r.m.s.d. values than the r.m.s.d. value of loop 2 at the bottom of the bundle that connects helix 2 and 3 (Fig. 3*C*). Finally, as mentioned above, the distance between helix 1 and 4 is slightly wider at the top than at the bottom of the bundle because of the ~10° tilt angle of helix 4 with the rest of the helices. Taken together, these structural features indicate the region around the Met cluster of the helix bundle is less compact compared with the hydrophobic core lying underneath. Interestingly, three of four Tyr residues within FAT domain also appear at the top of the bundle (Fig. 3*A*). Among them, Tyr-951 is buried between loop 1 and 3, immediately above the Met cluster. Two other conserved Tyr residues (Tyr-1008 and Tyr-1017) are solvent-exposed. They are located at the two ends of loop 3, and their aromatic side chains are arranged in a "ring-stacking" position (Fig. 3*A*) and thus may play a role in stabilizing the top end of the helix bundle.

Previous DALI data base searches revealed that despite the lack of significant sequence similarity, the FAT domain is structurally related to the vinculin tail domain (PDB access code 1qkr), domains within α-catenin (PDB access code 1h6g), and apolipoprotein E3 (PDB access code 1nfn) (38–40). All of these structures include a four-helix bundle with a simple up-down-up-down right-handed topology (39). By inspection of the PDB coordinates, we found that the vinculin tail domain and a subset of apolipoproteins also have Met clusters within the helix bundle. The helix bundles of both proteins are proposed to "unfurl" to promote interaction with lipid substrates using the exposed hydrophobic sites (73–75). NMR and x-ray crystallographic studies of the free FAT domain revealed intrinsic flexibility of the helices in the helix bundle of the FAT domain, especially helix 1, which is in intermediate exchange on the NMR time scale and possesses enough conformational flexibility to "domain exchange" with helix 1 of another FAT domain molecule to form a domain swapped dimer in the crystalline state (38).[2] The less compact cavity containing the methionine cluster, in addition to a strained polyproline-rich hinge region at the top of FAT domain bundle, may be responsible for the capability of the FAT domain to "open up." However, the plasticity of the helix bundle, particularly the helix 1, is largely decreased by the association with the LD2 peptides in our solution structure of the complex, as evidenced by perfectly well defined conformation of all four helices, and by apparent lack of the line broadening of NMR resonances seen in the spectra of the free FAT domain (data not shown). In addition to common structural and perhaps dynamic properties, functional similarities between the FAT domain and the vinculin tail are of particular interest, as both FAK and vinculin are localized to focal adhesions through their FAT domain-like 4-helix bundle tail domains. Like the FAT domain, the vinculin tail binds to LD2 and LD4 motifs of paxillin. However, it is unclear whether the FAT domain and the vinculin tail bind paxillin using a similar binding epitope.

Surface Properties of the FAT Domain—The most noticeable feature of the surface of the FAT domain is the presence of two large hydrophobic patches on the opposite faces of the molecule termed HP1 and HP2 (Fig. 4, *A* and *B*) by Hayashi *et al.* (39). HP1 is located at the interface between helix 2 and 3, whereas HP2 is located at the interface between helix 1 and 4. However, comparison of recent structures of the FAT domain, both free and LD2 bound, show two noticeable differences in the hydrophobic patches. 1) HP1 traverses much of the helix 2, 3 surface from top to bottom in our solution structure of the FAT domain bound by the LD2 peptide, in contrast to HP1 in the crystal structure and our recent solution structure of the free FAT domain (38, 39),[2] where the HP1 patch resides in the center of helix 2, 3 interface. This is presumably due to the subtle conformational changes of the helix 2 and 3. 2) The hydrophobic area of HP2 becomes larger upon binding the LD2 peptide because of the slight split between helix 1 and 4 and/or the partial exposure of the Met cluster. From our thermodynamic studies, it is very clear that the binding between the FAT domain and the LD2 peptide involves large scale hydrophobic interactions. This information, when combined with our paramagnetic mapping experiments, indicates that HP1 and HP2 represent the binding sites for the hydrophobic faces of the LD2 peptide. Indeed HP2 and especially HP1 are two surfaces that have a great deal of sequence conservation among FAT domains (38, 39). The electrostatic surfaces shown in Fig. 4, *C* and *D,* represent views illustrating the same surfaces as in Fig. 4, *A* and *B*. Areas of positively charged residues are present at positions indicated in Fig. 4, *C* and *D,* immediately adjacent to the two hydrophobic patches, including conserved residues Lys-934 in helix 1, Lys-956, Arg-963 in helix 2, Lys-1003 in helix 3, and Lys-1019, His-1026, and Lys-1033 in helix 4. Considering the acidic nature of LD2 peptide and its strong propensity to form an amphipathic α-helical conformation, the conservation of these distinct electropositive spots and adjacent large hydrophobic patches strongly suggest that the interfaces between helix 1 and 4, and helix 2 and 3 are the binding sites for the LD2 peptide. Analysis of the surface formed by helix 1 and 2 and the surface formed by helix 3 and 4 reveals different

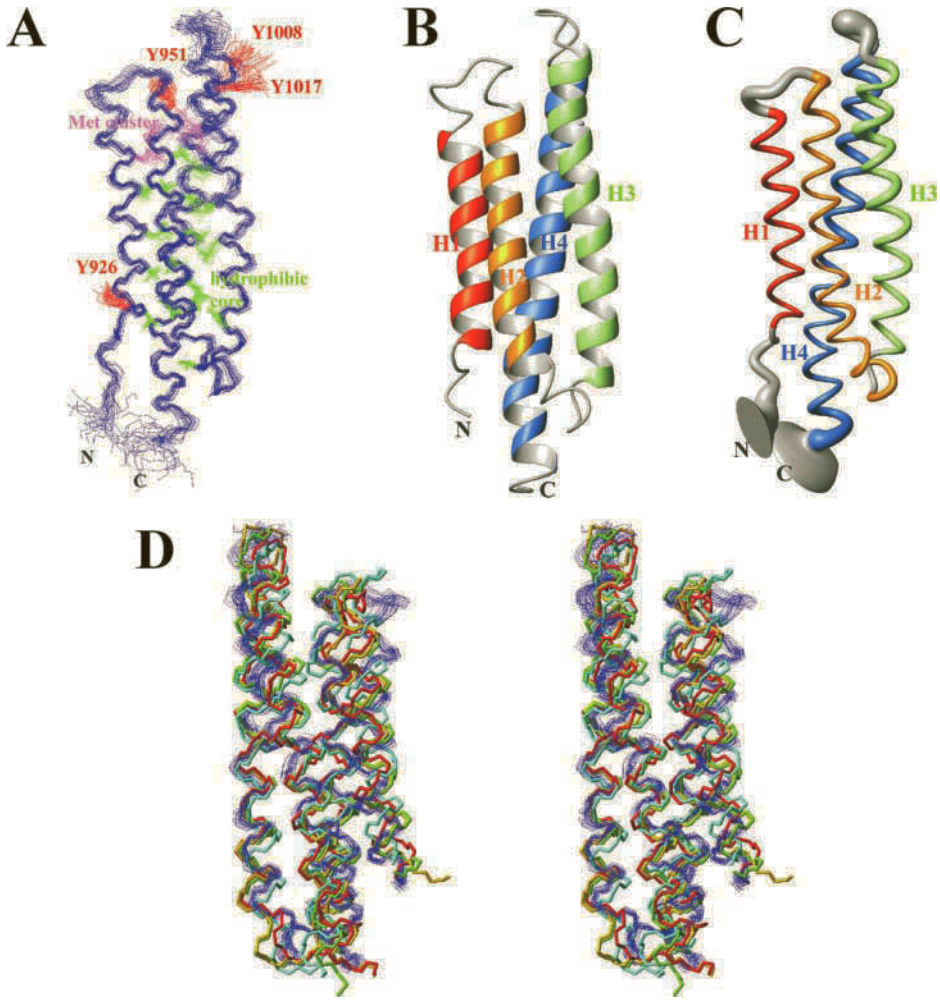

FIG. 3. **Solution structure of FAT bound by LD2 peptide (133–159).** *A*, ensemble of 25 structures (residues 920–1053). Side chain heavy atoms of well defined residues in the hydrophobic core of the helical bundle are colored *green*. Side chain heavy atoms of four Tyr residues are highlighted in *red*, with the residue numbers labeled. The Met cluster is shown in *magenta*. *B*, ribbon plot of the backbone structure of FAT in complex with LD2 peptide in an orientation similar to *A*. Representative model 3 is drawn with four α-helices colored. *C*, the 25 solution structures of FAT are represented with a backbone tube of thickness proportional to the backbone r.m.s.d., again in an orientation similar to *A*. Four α-helices are colored in the same manner as in *B*. Loops and termini are colored *gray*. *D*, superposition of 25 NMR models of avian FAT (1qvx, *blue*) bound by LD2 peptide with the following structures: representative avian FAT NMR structure (1ktm, model 1, *red*) in complex with a similar LD2 peptide; representative free avian FAT NMR structure[2] (1pv3, model 1, *cyan*); free mouse FAT x-ray structure (1k40, *green*); and free human FAT x-ray structure (1k05, *yellow*). The backbone atoms (N, Cα, C′) of FAT residues 926–941, 950–976, 983–1006, and 1016–1047 were superimposed on the backbone atoms of mouse or human FAT residues 925–940, 949–975, 982–1005, and 1015–1046. The view in *D* is a rotation of the view of *A*, nearly 180° about the long axis of helical bundle. All figures were generated by MolMol (76).

TABLE III
Comparison of structure of FAT models

Comparison of structure of FAT models was based on pairwise superposition of backbone carbon and nitrogen atoms of residues in secondary structure (926–941, 950–976, 983–1006, and 1016–1047). Fitting was performed using the McLachlan algorithm (41) as implemented in the program ProFit (/www.bioinf.org.uk/software/profit).

Pairwise r.m.s.d.	FAT bound by LD2, NMR, 1ktm	Free FAT, NMR, 1pv3[a]	Free FAT, x-ray, 1k40	Free FAT, x-ray, 1k05[b]
FAT bound by LD2, NMR, 1qvx	2.40 ± 0.08	2.54 ± 0.07	2.04 ± 0.08	1.83 ± 0.06
FAT bound by LD2, NMR, 1ktm		2.12 ± 0.07	1.47 ± 0.06	1.36 ± 0.04
Free FAT, NMR, 1pv3[a]	2.12 ± 0.07		1.64	1.91
Free FAT, X-ray, 1k40	1.47 ± 0.06	1.64		0.94

[a] Using model 1, judged most representative by its authors.
[b] A chain of 1k05.

properties than the surfaces containing HP1 and HP2: 1) the absence of the large hydrophobic area; and 2) the presence of dominantly electronegative patches in the center of the surfaces (Fig. 4, *E* and *F*). These distinct features suggest it is impossible for the helix 1, 2 and helix 3, 4 surfaces to interact with paxillin LD motifs in a similar fashion. However, it is possible these two surfaces provide additional protein-protein interaction sites for other FAT domain ligands, such as talin. Our solution structure of the FAT domain in complex with the LD2 peptide also explains the widespread chemical shift changes shown in the ^{1}H-^{15}N HSQC spectra for virtually all residues upon peptide binding. These changes were due to the following: 1) the subtle conformational changes of helices that were observed between the free and bound states of the FAT domain, especially helix 4; and 2) the involvement of all four helices in peptide binding, altering the chemical environments experienced by almost all residues in the FAT domain.

Structure of FAT Domain Bound by LD2 Peptide Accounts for Available Mutagenesis Data—Two studies have characterized several mutations in the FAT domain in an effort to better understand FAT domain function (23, 24), particularly the relationship between paxillin binding and focal adhesion targeting. In the first study (23), point mutations were introduced into either helix 1 or 4 of the FAT domain within regions of sequence homology with vinculin (23). However, the majority of mutants that abrogated paxillin binding are associated with residues involved in helix packing and are likely to result in structural alterations of the FAT domain (V928G, L931R, V935A, L1034S, L1035A, V1037D corresponding to V929G, L932R, V936A, L1035S, L1036A, V1038D in the avian FAT domain). The majority of mutations that had no effect on paxillin binding or focal adhesion targeting involved solvent-exposed residues on either the helix 1, 2 face (K923E corresponding to K924E in the avian FAT domain), helix 1, 4 face

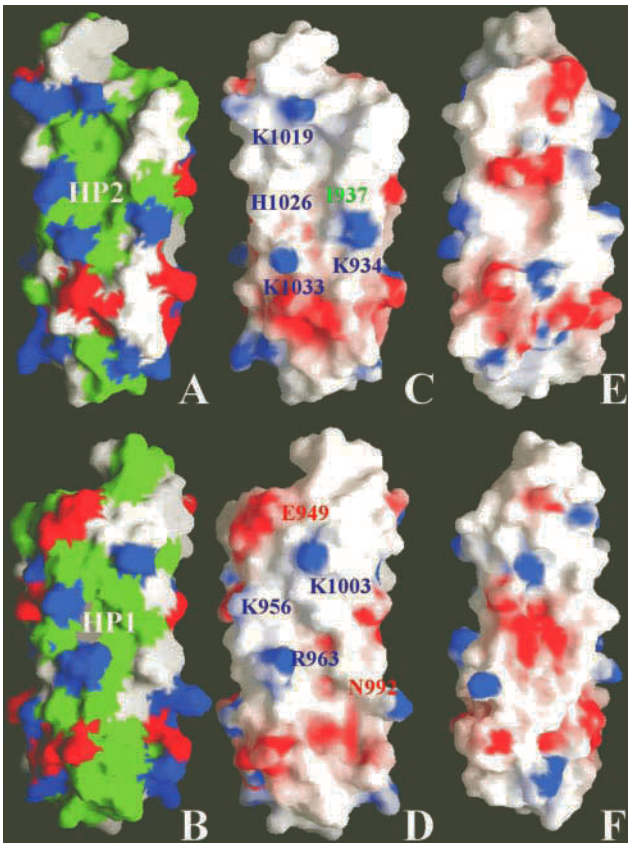

FIG. 4. **Surface of solution structure of FAT bound to LD2.** Surface of helix 1,4 (*A*) and 2,3 (*B*) are color-coded for hydropathy and charge (*green*, hydrophobic; *white*, hydrophilic; *red*, Asp and Glu; *blue*, Lys, Arg and His). *C–F* are electrostatic potential surfaces of FAT bound to LD2, colored from −10 (*red*) to +10 (blue) kT e^{-1}. The surfaces in *C* and *D* are the same as those in *A* and *B*. The views in *E* and *F* involve a 90° rotation, relative to *C* and *D*, respectively, about the long axis of the helix bundle. They represent surfaces of helix 1,2 (*E*) and helix 3,4 (*F*). These molecular surfaces, derived from representative NMR model 3, were generated using GRASP (77).

(T929S/R, K933E, D1036H, L1043R corresponding to T930S/R, K934E, D1037H, L1044R in the avian FAT domain), or the helix 3, 4 face (N1033D, Q1040E/G/K corresponding to N1034D, Q1041E/G/K in the avian FAT domain). Two of the mutants that abrogated paxillin binding to the FAT domain are solvent-exposed residues located near the end of helix 4 (D1039A, R1042G corresponding to D1040A, R1043G in the avian FAT domain), which is distant from the proposed helix 1, 4 binding site. Based upon either our models of LD motif-FAT domain interactions or that reported by Liu *et al.* (40), it is unclear how these mutations perturb paxillin binding.

In the second study, a set of avian FAT domain mutants targeted to helices 2, 3, and 4 was engineered (24). Two mutants (L995A/L998A and L1028A/L1035A) abrogated both paxillin binding as well as focal adhesion targeting. Both mutants, however, targeted residues located in the hydrophobic core of the FAT domain and may perturb the structure of the FAT domain, which could account for disruption of paxillin binding and resulting FAK signaling deficiencies. The E985A/K989A mutant retains paxillin binding and focal adhesion targeting properties similar to those of the wild type FAT domain. Glu-985 and Lys-989 are not in the immediate vicinity of the proposed binding site; therefore, this mutant would be predicted to have a minimal effect on paxillin binding. Two additional mutants were made that reduced paxillin binding and still targeted FAK to focal adhesions. The E949A/K956A/R963A mutant removes significant positive charge from the hydrophobic patch of the helix 2, 3 paxillin-binding site, whereas the H1026A/K1033A mutant similarly removes positive charge near the hydrophobic patch of the helix 1, 4 binding site. As will be discussed below, these electropositive residues are important for the FAT-LD2 interaction.

Modeling of FAT-LD2 Complex Using Mutagenesis Data by HADDOCK Program—NMR efforts failed to experimentally extract intermolecular NOEs between the FAT domain and the LD2 peptide in the FAT domain-LD2 complex, or intramolecular LD2 peptide NOEs in the FAT domain-LD2 complex for structure calculations, in both a previous study (40) and our present NMR study. Consistent with these observations, we have observed that resonances associated with the LD2 peptide are in intermediate exchange on the NMR time scale and are consequently broadened beyond detection in the FAT domain-LD2 complex. Hayashi *et al.* first generated a model for these interactions by manually docking the LD2 peptide onto their free FAT domain crystal structure followed by energy minimization (39). In the previous NMR study, Liu *et al.* (40) were able to extract a few intermolecular distance restraints between the FAT domain and the LD2 peptide by site-directed paramagnetic spin labeling of the peptide, and then used these restraints to model their structure of the free LD2 peptide onto the solution structure of the FAT domain in the LD2-bound state (40). Although Hayashi *et al.* (39) and Liu *et al.* (40) both modeled the LD2 peptide in a helical conformation when it binds to the FAT domain, their resultant models are fundamentally different as follows. 1) Hayashi *et al.* (39) identified two surfaces containing either HP1 or HP2 as the LD2-binding sites, whereas Liu *et al.* (40) only found one LD2-binding site at the interface of helix 1 and 4, corresponding to the surface containing HP2 in the model by Hayashi. 2) The LD2 peptide docked at the interface of helix 1 and 4 in these two models are in opposite orientations. There is no question that the manual modeling approach is subjective in nature. However, the paramagnetic spin-labeling technique is also not error-free. Extra caution has to be taken during the analysis of line broadening data involving the spin-labeling technique. 1) It is very possible that the Cys mutation of the LD2 peptide residues followed by chemical modification will alter the affinity of the LD2 peptide for the FAT domain. 2) The presence of excessive concentrations of spin-labeled peptides in solution may lead to nonspecific binding interactions with the protein, therefore a very careful "background" calibration is necessary.

A recently introduced approach, using the HADDOCK program (67), provided us with an alternative method to dock the LD2 peptide onto FAT and has been proven to be very successful in the present study. By using this approach, we introduced available mutagenesis data as ambiguous interaction restraints to drive the docking process. An ambiguous interaction restraint is defined as an ambiguous distance between all residues shown to be involved in the interaction. Another advantage of the application of HADDOCK over the two techniques described above is that it accommodates conformational rearrangements during the docking process by allowing side chain rearrangements, which is crucial for the FAT domain-LD2 complex. The final five models with the lowest intermolecular energy after the docking process have very low backbone r.m.s.d. for both the helix 1, 4 and helix 2, 3 sites: 0.65 ± 0.12 Å for LD2 peptide docked onto helix 1, 4 site with an intermolecular energy of −220 ± 12 kcal·mol^{-1}, and 0.43 ± 0.07 Å for LD2 peptide docked onto helix 2, 3 site with an intermolecular energy of −232 ± 10 kcal·mol^{-1}. Fig. 5 represents a view of the helix 1, 4 (Fig. 5*A*) and helix 2, 3 site (Fig. 5*B*). The orientation of the LD2 peptide at the helix 1, 4 binding site is similar to that reported previously by Liu *et al.* (40). Two general simi-

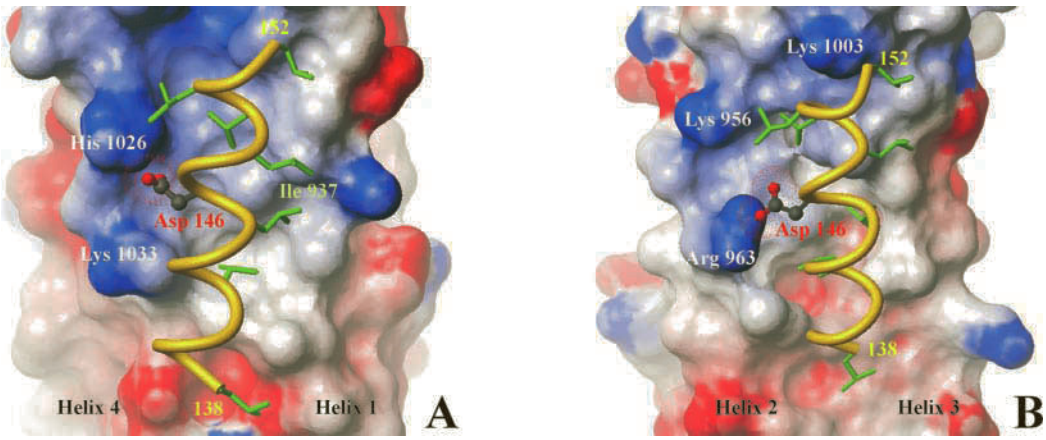

FIG. 5. **NMR-derived models of the FAT-paxillin complex using the HADDOCK program.** Proposed model of LD2 peptide (*gold*) in helical conformation docked onto the surface of helix 1, 4 (*A*) and helix 2, 3 (*B*). Electrostatic potential surfaces of FAT were generated by MolMol (76) using default parameters. The side chain of the critical LD2 residue, Asp-146, is shown as a *ball-and-stick* representation. Side chains of seven LD2 leucine residues are colored in *green*. FAT residues identified as active residues during the docking process, as described in the "Experimental Procedures," are indicated.

larities between the helix 1, 4 and helix 2, 3 binding sites are as follows. 1) The critical Asp-146 residue on the LD2 peptide makes salt bridges to the conserved basic residues of the FAT domain (His-1026 and Lys-1033 within the helix 1, 4 site; Arg-963 and/or Lys-956 within the helix 2, 3 site) in close proximity to the hydrophobic patches HP1 and HP2. 2) The hydrophobic side chains from the LD2 peptide, especially the leucine residues shown in *green*, fit extremely well into the hydrophobic grooves of HP1 and HP2. Our docking models suggest that the structured regions of the LD2 peptide bind to the helix 1, 4 and helix 2, 3 sites in a similar manner. However, the ITC results clearly indicate the thermodynamic basis governing the protein-peptide interactions are fundamentally different despite the similar affinities; the binding between helix 1, 4 site and the LD2 peptide is mainly enthalpically driven, whereas the binding between helix 2, 3 site and the LD2 peptide is essentially entropically driven (Table I). We postulate, based on these results, that the relatively flexible N- and/or C-terminal residues may also interact with the FAT domain. In fact, ~2–3-fold decrease in the affinity was observed for both FAT-binding sites when a shorter LD2 peptide lacking the highly conserved N-terminal 8 residues was used during the ITC titrations (data not shown). One important difference between our current model of the FAT domain-LD2 complex and that reported by Liu *et al.* (40), besides the presence of two LD2-binding sites *versus* one binding site, is that there appears to be a slight split at the top of the helix bundle, between helix 1 and 4 in our structure in the LD2 peptide-bound state. Discrepancies between these structures may be due to the following differences in the complexes under investigation. 1) Different LD2 peptide sequences were employed (LD2 residues 139–162 by Liu *et al.* (40) *versus* 133–159 in this study). 2) Different protein/peptide molar ratios were used for the NMR studies of the complex (1:5 by Liu *et al.* (40) *versus* 1:8 in this study). However, our ITC and paramagnetic relaxation broadening results acquired on a shorter LD2 peptide lacking 8 N-terminal residues (LD2 residues 141–159) indicated a similar mode of binding to the FAT domain, albeit with a modest decrease in the affinity for each of the two binding sites (data not shown). Hence, the underlying reasons for the differences between these two independent NMR studies are unclear.

Biological Implications for Multiple Paxillin LD Motif-binding Sites on the FAT Domain—FAK and paxillin are focal adhesion-associated, phosphotyrosine-containing proteins that physically interact (34). The binding of FAK to paxillin is mediated by LD2 and LD4 motifs within the N-terminal half of paxillin (34, 36). Interestingly, a larger fragment of paxillin encompassing both LD2 and LD4 motifs binds 5–10-fold more strongly than each individual LD motif peptide, and the critical Asp mutations in both LD motifs of the larger fragment are required to dramatically impair FAK binding *in vitro* (34). The reason for two FAK-binding sites on paxillin and two paxillin-binding sites on FAK is unclear. Multiple docking sites may simply be a mechanism to increase the avidity of the interaction. Alternatively, multiple docking sites may allow the assembly of higher order complexes containing a number of FAK and paxillin molecules. Formation of such complexes may function in the regulation of FAK or paxillin signaling. In addition, the functional significance of two binding sites is unclear, *i.e.* is paxillin bound to HP1 functionally redundant with paxillin bound to HP2?

The role of distinct paxillin-binding sites *in vivo* is open to speculation. Whereas previous mutational analyses, albeit with most of the mutations now predicted to perturb structure, demonstrate that paxillin binding is dispensable for localization of FAK, disruption of paxillin binding may have more subtle effects upon localization. Analysis of a paxillin mutant defective for FAK binding reveals reduced levels of phosphotyrosine relative to wild type paxillin (34). These results suggested that paxillin must bind FAK for maximal paxillin phosphorylation in response to cell adhesion. Hence, FAK may function to direct tyrosine phosphorylation of paxillin. Several analyses have also suggested that disruption of paxillin function may impair tyrosine phosphorylation of FAK and Pyk2 (28, 31). These findings suggest that paxillin association could also potentially regulate FAK signaling. Given the additional complexity of the FAK-paxillin interaction, *i.e.* presence of multiple binding sites, the question of functional redundancy between specific FAK-paxillin complexes must again be raised. Although these intriguing questions remain to be addressed, the mutants developed in this study will be invaluable in further studies designed to further explore these issues.

Accession Numbers—The resonance assignments have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under BMRB accession number 5924. The ensemble of 25 NMR models has been deposited under PDB accession code 1qvx.

Acknowledgments—We thank G. Mueller for assistance with ARIA program and helpful discussions. We thank B. Temple for help with the Insight II program, and A. Tripathy and T. Walker for help with the ITC experiments. We thank G. Mueller and A. Lee for critically reading this manuscript. Duke University and Bruker Biospin Corp. provided Inova 800 MHz and DMX 600 NMR time with R. Venters and C. Anklin assisting.

Addendum—After we submitted this paper, the crystal structures of the human FAT domain in complex with a shorter paxillin LD2 or LD4 peptide were reported (Hoellerer, M. K., Noble, M. E., Labesse, G., Campbell, I. D., Werner, J. M., and Arold, S. T. (2003) *Structure* **11,** 1207–1217). This paper confirms the presence of two LD-binding sites within the FAT domain, and the structures are in good agreement with our models demonstrating the utility of the HADDOCK/NMR approach for modeling FAT-LD2 binding interactions.

REFERENCES

1. Ilic, D., Furuta, Y., Kanazawa, S., Takeda, N., Sobue, K., Nakatsuji, N., Nomura, S., Fujimoto, J., Okada, M., and Yamamoto, T. (1995) *Nature* **377,** 539–544
2. Schlaepfer, D. D., Hauck, C. R., and Sieg, D. J. (1999) *Prog. Biophys. Mol. Biol.* **71,** 435–478
3. Schaller, M. D. (2001) *Biochim. Biophys. Acta* **1540,** 1–21
4. Parsons, J. T. (2003) *J. Cell Sci.* **116,** 1409–1416
5. Cance, W. G., and Liu, E. T. (1995) *Breast Cancer Res. Treat.* **35,** 105–114
6. Hecker, T. P., and Gladson, C. L. (2003) *Front. Biosci.* **8,** S705–S714
7. McLean, G. W., Avizienyte, E., and Frame, M. C. (2003) *Expert Opin. Pharmacother.* **4,** 227–234
8. Jones, R. J., Brunton, V. G., and Frame, M. C. (2000) *Eur. J. Cancer* **36,** 1595–1606
9. Kornberg, L. J. (1998) *Head Neck* **20,** 745–752
10. Gabarra-Niecko, V., Schaller, M., and Dunty, J. (2003) *Cancer Metastasis Rev.* **22,** 359–374
11. Schneider, G. B., Kurago, Z., Zaharias, R., Gruman, L. M., Schaller, M. D., and Hendrix, M. J. (2002) *Cancer* (*Phila.*) **95,** 2508–2515
12. Renshaw, M. W., Price, L. S., and Schwartz, M. A. (1999) *J. Cell Biol.* **147,** 611–618
13. Hsia, D. A., Mitra, S. K., Hauck, C. R., Streblow, D. N., Nelson, J. A., Ilic, D., Huang, S., Li, E., Nemerow, G. R., Leng, J., Spencer, K. S., Cheresh, D. A., and Schlaepfer, D. D. (2003) *J. Cell Biol.* **160,** 753–767
14. Frisch, S. M., Vuori, K., Ruoslahti, E., and Chan-Hui, P. Y. (1996) *J. Cell Biol.* **134,** 793–799
15. Jones, G., Machado, J., Jr., and Merlo, A. (2001) *Cancer Res.* **61,** 4978–4981
16. Leyton, J., Garcia-Marin, L. J., Tapia, J. A., Jensen, R. T., and Moody, T. W. (2001) *Cancer Lett.* **162,** 87–95
17. Hauck, C. R., Sieg, D. J., Hsia, D. A., Loftus, J. C., Gaarde, W. A., Monia, B. P., and Schlaepfer, D. D. (2001) *Cancer Res.* **61,** 7079–7090
18. Li, X., Regezi, J., Ross, F. P., Blystone, S., Ilic, D., Leong, S. P., and Ramos, D. M. (2001) *J. Cell Sci.* **114,** 2665–2672
19. Hildebrand, J. D., Schaller, M. D., and Parsons, J. T. (1995) *Mol. Biol. Cell* **6,** 637–647
20. Hildebrand, J. D., Schaller, M. D., and Parsons, J. T. (1993) *J. Cell Biol.* **123,** 993–1005
21. Chen, H. C., Appeddu, P. A., Parsons, J. T., Hildebrand, J. D., Schaller, M. D., and Guan, J. L. (1995) *J. Biol. Chem.* **270,** 16995–16999
22. Turner, C. E., and Miller, J. T. (1994) *J. Cell Sci.* **107,** 1583–1591
23. Tachibana, K., Sato, T., D'Avirro, N., and Morimoto, C. (1995) *J. Exp. Med.* **182,** 1089–1099
24. Cooley, M. A., Broome, J. M., Ohngemach, C., Romer, L. H., and Schaller, M. D. (2000) *Mol. Biol. Cell* **11,** 3247–3263
25. Zheng, C., Xing, Z., Bian, Z. C., Guo, C., Akbay, A., Warner, L., and Guan, J. L. (1998) *J. Biol. Chem.* **273,** 2384–2389
26. Shen, Y., and Schaller, M. D. (1999) *Mol. Biol. Cell* **10,** 2507–2518
27. Richardson, A., and Parsons, T. (1996) *Nature* **380,** 538–540
28. Hagel, M., George, E. L., Kim, A., Tamimi, R., Opitz, S. L., Turner, C. E., Imamoto, A., and Thomas, S. M. (2002) *Mol. Cell. Biol.* **22,** 901–915
29. Wade, R., Bohl, J., and Vande Pol, S. (2002) *Oncogene* **21,** 96–107
30. Liu, S., Thomas, S. M., Woodside, D. G., Rose, D. M., Kiosses, W. B., Pfaff, M., and Ginsberg, M. H. (1999) *Nature* **402,** 676–681
31. Rose, D. M., Liu, S., Woodside, D. G., Han, J., Schlaepfer, D. D., and Ginsberg, M. H. (2003) *J. Immunol.* **170,** 5912–5918
32. Yano, H., Uchida, H., Iwasaki, T., Mukai, M., Akedo, H., Nakamura, K., Hashimoto, S., and Sabe, H. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9076–9081
33. Petit, V., Boyer, B., Lentz, D., Turner, C. E., Thiery, J. P., and Valles, A. M. (2000) *J. Cell Biol.* **148,** 957–970
34. Thomas, J. W., Cooley, M. A., Broome, J. M., Salgia, R., Griffin, J. D., Lombardo, C. R., and Schaller, M. D. (1999) *J. Biol. Chem.* **274,** 36684–36692
35. Schaller, M. D. (2001) *Oncogene* **20,** 6459–6472
36. Brown, M. C., Perrotta, J. A., and Turner, C. E. (1996) *J. Cell Biol.* **135,** 1109–1123
37. Tumbarello, D. A., Brown, M. C., and Turner, C. E. (2002) *FEBS Lett.* **513,** 114–118
38. Arold, S. T., Hoellerer, M. K., and Noble, M. E. (2002) *Structure* **10,** 319–327
39. Hayashi, I., Vuori, K., and Liddington, R. C. (2002) *Nat. Struct. Biol.* **9,** 101–106
40. Liu, G., Guibao, C. D., and Zheng, J. (2002) *Mol. Cell. Biol.* **22,** 2751–2760
41. McLachlan, A. D. (1982) *Acta. Cryst. Sect. A* **38,** 871–873
42. Arumugam, S., Hemme, C. L., Yoshida, N., Suzuki, K., Nagase, H., Berjanskii, M., Wu, B., and Van Doren, S. R. (1998) *Biochemistry* **37,** 9650–9657
43. Schaller, M. D., Borgman, C. A., Cobb, B. S., Vines, R. R., Reynolds, A. B., and Parsons, J. T. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 5192–5196
44. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR* **6,** 135–140
45. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
46. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR* **4,** 603–614
47. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Formankay, J. D., and Kay, L. E. (1994) *J. Magn. Reson.* **103,** 197–201
48. Ikura, M., Kay, L. E., and Bax, A. (1990) *Biochemistry* **29,** 4659–4667
49. Matsuo, H., Kupce, E., Li, H., and Wagner, G. (1996) *J. Magn. Reson.* **113,** 91–96
50. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson.* **101,** 201–205
51. Yamazaki, T., Lee, W., Arrowsmith, C., Muhandiram, D. R., and Kay, L. E. (1994) *J. Am. Chem. Soc.* **116,** 11655–11666
52. Kay, L. E., Xu, G., and Yamazaki, T. (1994) *J. Magn. Reson.* **109,** 129–133
53. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) *J. Magn. Reson.* **88,** 425–431
54. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Formankay, J. D. (1993) *J. Magn. Reson.* **101,** 333–337
55. Lyons, B. A., Tashiro, M., Cedergren, L., Nilsson, B., and Montelione, G. T. (1993) *Biochemistry* **32,** 7839–7845
56. Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) *J. Biomol. NMR* **3,** 225–231
57. Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) *J. Am. Chem. Soc.* **115,** 11054–11055
58. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* **131,** 373–378
59. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry* **31,** 1647–1651
60. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* **4,** 171–180
61. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
62. Yang, D., Venters, R., Mueller, G. A., Choy, W. Y., and Kay, L. E. (1999) *J. Biomol. NMR* **14,** 333–343
63. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR* **13,** 51–59
64. Nilges, M. (1995) *J. Mol. Biol.* **245,** 645–660
65. Nilges, M. (1997) *Folding Des.* **2,** S53–S57
66. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **54,** 905–921
67. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) *J. Am. Chem. Soc.* **125,** 1731–1737
68. Hubbard, S. J., and Thornton, J. M. (1993) "*NACCESS*" *Computer Program*, Version 2.1.1, University College London, London, UK
69. Schaller, M. D., and Sasaki, T. (1997) *J. Biol. Chem.* **272,** 25319–25325
70. Salgia, R., Avraham, S., Pisick, E., Li, J. L., Raja, S., Greenfield, E. A., Sattler, M., Avraham, H., and Griffin, J. D. (1996) *J. Biol. Chem.* **271,** 31222–31226
71. Li, X., and Earp, H. S. (1997) *J. Biol. Chem.* **272,** 14341–14348
72. Boetzel, R., Czisch, M., Kaptein, R., Hemmings, A. M., James, R., Kleanthous, C., and Moore, G. R. (2000) *Protein Sci.* **9,** 1709–1718
73. Wilson, C., Wardell, M. R., Weisgraber, K. H., Mahley, R. W., and Agard, D. A. (1991) *Science* **252,** 1817–1822
74. Narayanaswami, V., and Ryan, R. O. (2000) *Biochim. Biophys. Acta* **1483,** 15–36
75. Bakolitsa, C., de Pereda, J. M., Bagshaw, C. R., Critchley, D. R., and Liddington, R. C. (1999) *Cell* **99,** 603–613
76. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics* **14,** 51–55
77. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296