PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Human FEZ1 has characteristics of a natively unfolded protein and dimerizes in solution

Daniel C. F. Lanza,[1,2] Julio C. Silva,[1,3] Eliana M. Assmann,[1] Alexandre J. C. Quaresma,[1,2] Gustavo C. Bressan,[1,2] Iris L. Torriani,[1,3] and Jörg Kobarg[1,2]*

[1] Laboratório Nacional de Luz Síncrotron, Campinas, SP, Brasil
[2] Instituto de Biologia, Universidade Estadual de Campinas, Campinas, SP, Brasil
[3] Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas, Campinas, SP, Brasil

ABSTRACT

The fasciculation and elongation protein Zeta 1 (FEZ1) is the mammalian orthologue of the *Caenorhabditis elegans* protein UNC-76, which is necessary for axon growth. Human FEZ1 interacts with Protein Kinase C (PKC) and several regulatory proteins involved in functions ranging from microtubule associated transport to transcriptional regulation. Theoretical prediction, circular dichroism, fluorescence spectroscopy, and limited proteolysis of recombinant FEZ1 suggest that it contains disordered regions, especially in its N-terminal region, and that it may belong to the group of natively unfolded proteins. Small angle X-ray scattering experiments indicated a mainly disordered conformation, proved that FEZ1 is a dimer of elongated shape and provided overall dimensional parameters for the protein. *In vitro* pull down experiments confirmed these results and demonstrated that dimerization involves the N-terminus. *Ab-initio* 3D low resolution models of the full-length conformation of the dimeric constructs 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) were obtained. Furthermore, we performed *in vitro* phosphorylation assays of FEZ1 with PKC. The phosphorylation occurred mainly in its C-terminal region, and does not cause any significant conformational changes, but nonetheless inhibited its interaction with the FEZ1 interacting domain of the protein CLASP2 *in vitro*. The C terminus of FEZ1 has been reported to bind to several interacting proteins. This suggests that FEZ1 binding and transport function of interacting proteins may be subject to regulation by phosphorylation.

Proteins 2009; 74:104–121.

Key words: SAXS; spectroscopy; protein-protein interactions; microtubular transport; circular dichroism; limited proteolysis; axonal transport.

INTRODUCTION

The fasciculation and elongation protein Zeta-1 (FEZ1) was initially identified as an orthologue of the *Caenorhabditis elegans* protein UNC-76, which is necessary for the formation and extension of the worms axons.[1] In rats it was later shown that FEZ1 mRNA is abundantly expressed in early stages of the developing brain at the onset of neurogenesis.[2]

In a yeast two-hybrid assay the FEZ1 protein was identified as a prey interacting with the N-terminal domain of PKCζ.[3] Co-expression of PKCζ and FEZ1 cells alters the sub-cellular localization of the latter in COS-7 and increases the rate of differentiation to neurons in PC12 cells. Several other independent yeast two-hybrid studies employing different, and at first sight, seemingly unrelated proteins as baits, all resulted in the identification of FEZ1 as a prey. These include the nuclear protein DISC1 (Disrupted-In-Schizophrenia 1), the murine E4B U-box-type ubiquitin ligase and the agnoprotein of the human polyoma JC virus.[4–6] In the latter study a functional association of FEZ1 with microtubule was demonstrated. By co-precipitation studies it was demonstrated that FEZ1 binds to the microtubule and that the viral agnoprotein can block this interaction. Based on this activity the agnoprotein is furthermore capable of inhibiting neurite outgrowth in PC12 cells.

The involvement of FEZ1 in transport processes has clear implications for its role in neuronal differentiation and axon growth. Several recent articles explored the potential role of FEZ1 in microtubule associated transport. In rat hippocampal neurons it was

Abbreviations: CC region, coiled-coil region; CLASP2, clip-associated protein; DISC1, disrupted-in-schizophenia 1; GST, glutathione-S-transferase; FEZ1/2, fasciculation and elongation factor 1/2; KIF3A, kinesin family member 3A; NEK1, NIMA related kinase 1; PKC, protein kinase C.

Grant sponsors: Fundação de Amparo à Pesquisa do Estado São Paulo (FAPESP), projects SmolbNet, CEPID, grant number: 05/00235-1, Conselho Nacional de Pesquisa e Desenvolvimento (CNPq), the LNLS.
*Correspondence to: Jörg Kobarg, Laboratório Nacional de Luz Síncrotron, Centro de Biologia Molecular Estrutural, Rua Giuseppe Máximo Scolfaro 10.000, C.P. 6192, 13084-971 Campinas - SP, Brasil. E-mail: jkobarg@lnls.br
Received 29 October 2007; Revised 18 April 2008; Accepted 28 April 2008
Published online 9 July 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22135

shown that inhibition of endogenous FEZ1 protein expression by iRNA-mediated gene silencing, leads to inhibition of axonal polarization and slower transport rates of mitochondria along microtubule.[7] Another recent study identified FEZ1 as a Kinesin heavy chain interacting protein and demonstrates further that FEZ1 collaborates with JIP1 (JNK-interacting protein) in order to activate the molecular motor protein Kinesin-1.[8]

After FEZ1 frequent detection as an interacting prey protein, its C-terminal region was recently explored also as a bait in yeast two-hybrid assays in order to identify additional interacting proteins and to get new clues on the possible cellular roles of FEZ1.[9,10] A total of 36 new interacting proteins were identified that can be grouped as follows: associated to microtubular transport (including CLASP2), neuronal cell development (including FEZ1 itself, RAB3, KIBRA, HNOEL-iso, *tachykinin,* and PTPRS), apoptose (Programmed Cell Death 7), mitochondrial function (including MAGMAS, KIAA1387, TOMM20), proteins with extracellular functions (Transthyretrin, P4HB, Parathyroid hormone, COL9A2), and the regulation of transcription or other nuclear functions (including DRAP1, SAP30L, BAF60a, HTAT-SF1, TFII2F).

Proteins that interact with more than 30 other protein are frequently called hubs and the chance that they are essential for the cells functions is three times larger than that of proteins with fewer functional links.[11] Interestingly, an analysis revealed that hub proteins have been described to contain a high surface charge content, and disordered domains, which allow them to interact with a multitude of proteins.[12]

Here we describe theoretical predictions and results from spectroscopic studies and limited proteolysis experiments, all of which suggest that FEZ1 could belong to the class of natively unfolded proteins. It is already known that proteins with multiple functions have frequently a disordered intrinsic structure.[13–18] It is well known that high resolution structures of proteins with a low degree of compactness are hard to obtain and small angle X-ray scattering (SAXS) is the most adequate technique to obtain dimensional parameters and low resolution 3D conformational models of very large or partially unstructured molecules. This technique was applied to unstructured proteins such as synucleins, prothymosine alpha and P53.[19–21] The SAXS analysis in solution of 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) confirmed both the dimerization of FEZ1 as well as its open extended shape, compatible with a largely unfolded protein. Finally, we show that FEZ1 interaction with the microtubule proteins CLASP2 is inhibited by its phosphorylation *in vitro.* Together, our results suggest that FEZ1 presents a dimeric conformation and that its phophorylation status may regulate its interaction with other regulatory proteins or transported cargoes.

METHODS

FEZ1 secondary structure prediction and sequence analysis

Prediction of disorder in FEZ1 was performed using the program PONDR (Prediction of Natural Disordered Regions, http://www.pondr.com) using the default predictor VL-XT.[22,23] This predictor integrates the three neural networks: the VL1 predictor and N- and C-terminal predictors, which use the disordered regions identified from missing electron density in X-ray crystallography and nuclear magnetic resonance (NMR) studies. Access to PONDR is provided by Molecular Kinetics Indianapolis (www.molecularkinetics.com). Further disorder predictions were performed using the Fold Index software (http://bip.weizmann.ac.il/fldbin/findex) that predicts if a given protein sequence is intrinsically unfolded implementing the algorithm of Uversky and coworkers, which is based on the average residue hydrophobicity and net charge of the sequence.[24–26] Additional analyses were performed using 12 prediction programs: DisEMBL TM,[27] DRIPPRED,[28] DISpro,[29] GlobPlot,[30] IUPred,[31] PreLink,[32] RoNN,[33] SPRITZ,[34] FoldUnfold,[35,36] VL2,[37] VL3H,[38] VSL2,[39] available in a Database of protein disorder DISPROT (www.disprot.org/predictors.php). For coiled coil predictions we used the COILS (http://www.ch.embnet.org/software/COILS_form.html) and Multicoil (http://groups.csail.mit.edu/cb/multicoil/cgi-bin/multicoil.cgi) programs.[40,41]

Plasmid constructions

To express full-length FEZ1(1-392) or truncated FEZ1(1-227) fused to a 6xHis-tag, the corresponding nucleotide sequences were amplified by PCR and inserted into a modified version of the bacterial expression vector pET28a (Novagen/EMD Biosciences, San Diego, CA) as described.[10] pET-TEV-28a has as the main difference the substitution of the thrombin cleavage site which follows the 6xHis-tag by a TEV protease cleavage site. The fusion protein constructs of both FEZ1 protein versions have the initiating Methionine encoded by the vector, followed by an additional 31 amino acids: Gly, 6xHis-tag, 7 amino acids from the TEV cleavage site and 16 poly-linker encoded amino acids. In the case of both fusion protein constructs of FEZ1 the natural first Met residue was left out to prevent its usage as an additional translation initiation site. These protein constructs have been called 6xHis-FEZ1(1-392) or FEZ1(1-392) and 6xHis-FEZ1(1-227) or FEZ1(1-227), respectively (throughout the article). We also cloned the C-terminus of FEZ1(221-392) into pET28a, but it only showed little expression and was almost exclusively found in the insoluble fraction of the bacterial lysate. This precluded further comparative studies with the full-length and N-terminal constructions, both of which expressed very well and were very soluble.

For expression of different constructs coding indicated protein fragments fused to GST, the corresponding nucleotide sequences were cloned into a modified vector pET28a-GST that codifies GST protein upstream of the protein to be inserted. The cDNAs encoding the proteins or protein fragments identified to interact with FEZ1 in the yeast two-hybrid assay were sub-cloned from the pACT2 vector (Clontech) into bacterial expression vector pGEX-4T-2 (GE Healthcare, Waukesha, WI) as described.[10] This way full length SAP30L(1-183) and DRAP1(1-205), as well as the interacting protein fragments CLASP2(1046-1251), RAI 14 isoform(720-983), KIBRA(869-1113) and SMC3(881-1217), all fused to GST were expressed as described below.

Protein expression and purification

Soluble FEZ1 (complete 1-392, or deletion 1-227), fused to an N-terminal 6xHis-tag like described above, was purified from 1 L of culture of *E. coli* BL21 (DE3) cells that were induced for 2.5 h to protein expression at 30°C using 0.5 m*M* isopropyl 1-thio-β-D galactopyranoside. 6xHis-FEZ1(1-392) or 6xHis-FEZ1(1-227) proteins used in this study were purified using a HiTrap chelating column in an ÄKTA^{TM} FPLC^{TM} (GE Healthcare) as follows. Cells were harvested by centrifugation at 4500*g* for 10 min, and the cell pellet was resuspended and incubated for 30 min with 10 volumes of lysis buffer (137 m*M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4, 1.8 m*M* KH_2PO_4, pH 7.4, 1 mg/mL lysozyme, 1 m*M* phenylmethylsulfonyl fluoride, and 0.05 mg/mL DNase). After three cycles of sonication, soluble and insoluble fractions were separated by centrifugation at 28,500*g* for 30 min at 4°C. The cleared supernatant was then loaded onto a HiTrap chelating column (GE Healthcare) preequilibrated with lysis buffer (lacking lysozyme and DNase), followed by extensive wash of the column with the same buffer. Bound proteins were eluted in a gradient of 0–100% of elution buffer (137 m*M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4, 1.8 m*M* KH_2PO4, 1 m*M* phenylmethylsulfonyl fluoride, and 500 m*M* imidazole, pH 7.4). Aliquots of each eluted fraction obtained were analyzed by SDS-PAGE, and peak fractions containing FEZ1 fusion protein were dialyzed with PBS buffer (137 m*M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). The proteins fused to GST were induced for expression and *E. coli* cells were lysed and processed as described above for 6xHis fusion proteins. The resulting cleared supernatant was incubated with glutathione-Uniflow resin (Clontech) and used for *in vitro* binding assays.

Circular dichroism and fluorescence spectroscopy

Circular dichroism (CD) measurements were done using a Jasco J-810 spectropolarimeter with a temperature of 25°C, controlled by a Peltier-type system PFD 425S. 6xHis-FEZ1 proteins were re-suspended in 20 m*M* Tris-HCl, pH 7.5 with 10 m*M* $MgCl_2$. Data were collected at a scanning rate of 50 nm/min with a spectral bandwidth of 1 nm using a 0.1 mm path length cell. In the TFE assay the sample concentrations were 8.1 μ*M* for 6xHis-FEZ1 (1-392) and 24.2 μM for 6xHis-FEZ1(1-227) and for the phosphorylation assays the sample concentration was 3.7 μ*M* for 6xHis-FEZ1(1-392) and 3.56 μ*M* for 6xHis-FEZ1(1-227) in 20 m*M* Tris-HCl and 10 m*M* $MgCl_2$ buffer, pH 7.5. All buffers used were of analytical grade and were filtered before use to avoid light scattering by small particles. Graphics were generated by Origin 7.5 software. Fluorescence measurements were performed at 25°C with a FP-6500 spectrofluorimeter (Jasco, Inc., Easton, MD). Intrinsic fluorescence spectra were taken between 310 and 400 nm, with excitation at 283 nm. Buffer contributions were subtracted from the raw fluorescence data to give corrected spectra.

SAXS experiments

The two protein samples were submitted to DLS (dynamic light scattering) analysis before the SAXS experiments. DLS data of two samples (6xHis-FEZ1, 1-392 and 6xHis-FEZ1, 1-227) showed narrow, single peaks and predicted masses and percentage of poly-dispersivity indicative of solutions of dimers in a monodisperse solution. The SAXS experiments were performed at the D02A-SAXS2 beamline of the Laboratório Nacional de Luz Síncrotron (LNLS, Campinas, Brazil). Measurements were performed with a monochromatic X-ray beam with a wavelength of $\lambda = 1.488$ Å. The X-ray patterns were recorded using a two-dimensional position-sensitive MARCCD detector. The sample-to-detector distances were set at 1374.4 mm and 1788.8 mm, resulting in a scattering vector range of $0.009\ Å^{-1} < q < 0.25\ Å^{-1}$, where q is the magnitude of the $\boldsymbol{q}$-vector defined by $q = (4\pi/\lambda)\sin\theta$ (2θ is the scattering angle). The samples of 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) in PBS buffer, were centrifuged for 30 min in an ultracentrifuge, at 356.000*g*, at 4°C to remove any aggregates or particles and then placed on ice. For SAXS measurements protein samples were placed in a 1-mm path length cell with mica windows, temperature-controlled ($T = 20$°C) via water circulation.[42] Three successive frames of 300 s each were recorded for each sample. The buffer scattering data were recorded before and after the sample scattering data for 300 s each and thereafter averaged. The scattering curves were individually corrected for the detector response and scaled by the incident beam intensity and the samples absorption. The average buffer scattering was subtracted from the corresponding sample scattering. The resulting curve was carefully inspected to check for possible radiation-induced damage, but no such effects were observed. The scattering patterns were recorded at

two different concentrations for each sample: 4.99 mg/mL and 2.57 mg/mL for 6xHis-FEZ1(1-392) and 9.17 mg/mL and 2.57 mg/mL for 6xHis-FEZ1(1-227). After scaling for concentration, the scattering data of both samples were analyzed to investigate concentration dependence. This effect was not observed. A 4.59 mg/mL bovine serum albumin (BSA, 66 kDa) solution, in the same buffer of the samples, was used as a standard sample to determine the molecular masses of the 6xHis-FEZ1 proteins. The molecular mass of each 6xHis-FEZ1 sample was inferred from the ratio of the extrapolated value of the intensity at the origin $I(0)$.[43]

SAXS data analysis

The radius of gyration was first evaluated using the Guinier approximation.[44–46] It is worth noting that in the case of an unstructured protein, the Guinier approximation holds true on a very restricted q range corresponding to $q < 1/R_g$. Such a very narrow range contains a small number of experimental points, thereby limiting the accuracy of the R_g determination. In that case, a more accurate determination of the radius of gyration can be obtained by Debye's equation,[47–51] which adequately describes the scattering in the domain $q < 1.4/R_g$ by:

$$\frac{I(q)}{I(0)} = \frac{2}{x^2}(x - 1 + e^{-x}) \qquad (1)$$

where $x = q^2R_g^2$.

Moreover, R_g was also evaluated from the pair distance distribution function $p(r)$ which was calculated using the indirect transform package GNOM.[52] The $p(r)$ function represents the histogram of all distances within a molecule and provides the maximum dimension D_{max} of the molecule where its value reaches zero.[22,23]

To analyze the natively unfolded conformation of 6xHis-FEZ1(1-227) and 6xHis-FEZ1(1-392) chains, a particularly useful representation of the SAXS intensity was used: the so-called Kratky Plot ($q^2I(q)$ vs. q). This plot is usually applied to study unstructured proteins because it provides information about the compactness of a molecule. For compact and structured proteins, the Kratky representation presents a bell-shaped plot with a well-defined maximum, because the scattering function satisfies Porod's law for large q values (i.e. $I(q)$ is proportional to $1/q^4$). Conversely, the Kratky representation of an ideal Gaussian chain,[47] (equivalent to a random coil in the case of an infinitely thin chain) presents a plateau at high q-values because $I(q)$ is proportional to $1/q^2$. Finally, in the case of a persistence length chain model, where short-range interactions between adjacent chain segments produce stiffness of the chain, the Kratky plot also displays a plateau over a specific q-range followed by a monotonic increase at still higher q-values.[50,51]

Ab initio modeling

Although the "shape" of naturally unfolded proteins is per se hard to define, we restored the low resolution average conformations of 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) from the experimental SAXS curves by the following *ab initio* approach. To obtain an overall view from the conformational space occupied by the proteins, the program DAMMIN,[53] was used. In this approach, each protein was represented as an assembly of densely packed spherical beads (dummy atoms) of radius $r_0 \ll D_{max}$ inside a sphere of diameter D_{max}, which was directly determined from the scattering data using the GNOM routine. Using simulated annealing, the program DAMMIN starts from a random configuration of beads and searches for a configuration that fits the experimental scattering pattern. Ten calculations were performed. The normalized spatial discrepancies (NSD) were evaluated using the DAMAVER suite,[54] and the most typical model (with lowest NSD value) were considered as close descriptions of the possible conformations of the molecule.

Limited proteolysis

Limited proteolysis with thrombin was performed by incubating 1 μg of thrombin (Sigma) with 100 μg of purified 6xHis-FEZ1(1-392) or 6xHis-FEZ1(1-227) in PBS buffer at 30°C. This corresponds to molar ratios of 1/55 and 1/83 protease/protein for the two different forms, respectively. Aliquots were collected in time intervals and reactions were stopped by adding sample buffer (50 m*M* Tris-HCl, pH 6.8, 2 m*M* EDTA, 1% sodium dodecyl sulfate (w/v), 1% β-mercaptoetanol, 8% glycerol, 0.025% bromophenol blue (w/v), 2 m*M* PMSF) and boiling for 5 min. The cleaved products were analyzed by SDS-PAGE on a 12.5% acryl amide gel. For proteolysis with proteinase K we utilized 100 ng of proteinase K for 5 μg of recombinant 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) as described.[55] For endogenous FEZ1 proteolysis, we used 400 ng of proteinase K which were added to the lysates of Hek293 cells.[55]

In vitro phosphorylation

For *in vitro* phosphorylation 5 μg of purified 6xHis-FEZ1(1-392) or (1-227) were incubated in 25 μL of the reaction mixture (20 m*M* Tris, 10 m*M* $MgCl_2$, 50 nm PMA; 20 m*M* ATP, pH 7.5) with 50 ng of PKC-Pan (a mixture of α, β, γ with lesser amounts of ζ and δ isoforms, purified from rat brain) or recombinant isoforms PKC α, δ, ζ (Calbiochem). After addition of 55 kBq of [γ-^{32}P]ATP (~220 TBq/mmol), the mixture was incubated at 30°C for 30 min. Samples were analyzed by SDS PAGE (10% acryl amide) and subsequent autoradiography. For spectroscopic analysis and protein–protein pull-down experiments nonradioactively labeled phospho-

6xHis-FEZ1 protein samples were prepared accordingly by leaving out the radioactive ATP. In case of the pull-down experiments FEZ1 was previously phosphorylated by PKC-Pan, and in the CD and fluorescence experiments with PKC-ζ.

In vitro binding assays and western blotting

Purified 6xHis-FEZ1 proteins were allowed to bind to 25 μL of Ni-NTA sepharose in PBS for 1 h at 4°C. After incubation, the beads containing bound recombinant proteins were washed three times with PBS at 4°C. GST or GST-fusion proteins were added to the supernatants of the protein coupled beads and incubated in 0.1 mL of PBS for 1 h at 4°C to allow protein–protein interactions to occur. The beads were then washed three times with 0.5 mL of PBS, followed by 10 washes with 1 mL NP40 buffer (1% NP-40, 0.15*M* NaCl, 50 m*M* Tris-HCl, pH 7.2, 2 m*M* EDTA), and three additional washes with 1 mL of PBS. Resin-bound proteins were run out by 10% acrylamid SDS-PAGE and transferred to PVDF membranes by electro blotting. After saturation with a solution of 5% dry milk in TBS-T (0.15*M* NaCl, 20 m*M* Tris-HCl, 0.05% Tween 20, pH 7.2), membranes were incubated either with a mouse anti-4xHis monoclonal antibody (1:5000, Qiagen) or with a mouse monoclonal anti-GST antibody 5.3.3 (hybridoma supernatant 1:5) for 1 h.[10] Alternatively, for detection of endogenous FEZ1 from Hek293 cells, we used a specific polyclonal rabbit antiserum that had been generated by four subsequent immunizations of rabbits with 1 mg 6xHis-FEZ1(131-392) fusion protein. After three washes with TBS-T (0.15*M* NaCl, 20 m*M* Tris-HCl, 0.05% Tween 20, pH 7.2), the membranes were incubated with the secondary horseradish peroxidase conjugated anti-rabbit IgG antibody (1:5000; Santa Cruz Biotechnology) for 1 h and washed again three times with TBS. The membranes were developed by chemi-luminescence using the reagent Luminol (Santa Cruz Biotechnology) for detection of His-tagged or GST fusion proteins. Interaction assays with proteins that interacted with FEZ1 in the two hybrid assay were performed as described.[10]

RESULTS

FEZ1 secondary structure prediction

A bioinformatics analysis of the amino acid sequence of human FEZ1 with the computer programs Fold Index and PONDR predicted a large contend of potentially unstructured regions [Fig. 1(A)]. PONDR predicts disordered regions throughout the whole sequence of FEZ1 but especially in the N-terminal and central regions of the protein sequence. Only at the C-terminus a short structured region was predicted. These predictions are at large in agreement with those obtained by the Fold index software. Additional analysis by 13 different predicting softwares at large confirmed these data that FEZ1 presents basically two unfolded regions with great probability in the N-terminal (regions between amino acids, 1-69 and 110-229) and a higher probability of structuring at the C-terminal (data not shown).

Using the software COILS and a window size of 28, we identified one region with a high probability to form coiled coils in the region between amino acids 230–265 (∼96% probability). The software MULTICOILS identified (window 28) more or less the same region (233–261, probability ∼54%). One additional region identified only by COILS but of high probability is located further C-terminal (278–306, probability ∼63%) and three regions of low probability at the N-terminus of FEZ1. The MultiCoil software furthermore predicted that the coiled coil region has a higher probability to form dimer than trimer coils.

Circular dichroism spectropolarimetry

To address the prediction of a largely disordered structure for FEZ1 experimentally we performed CD spectropolarimetry experiments. The spectrum of 6xHis-FEZ1(1-392) has an pronounced minimum at 205 nm as well as in the regions 208 and 222 nm [Fig. 1(B)]. 6xHis-FEZ1(1-227) has a pronounced minimum at 205 nm region. Disordered proteins or unfolded proteins are characterized by minima in their far UV CD spectra around 200 and 205 nm.[56] Therefore the spectra of full-length and even so more that of 6xHis-FEZ1(1-227) show clear characteristics of largely disordered proteins. This is in agreement with the theoretical secondary structure prediction by 12 softwares on the NPS@ server (http://npsa-pbil.ibcp.fr), which indicated an average of 48.7% of alpha helix, 41.13% of random coil and 4.73% of ambiguous regions in the amino acid sequence of FEZ1. Interestingly, the CD spectrum of full-length 6xHis-FEZ1 and 6xHis-FEZ1(1-227) showed a change in profile upon addition of 5% of TFE [Fig. 1(B,C)]. The change in the shape of the spectrum of full length FEZ1 suggests that there is a visible global gain of mainly α-helical content. This could suggest that FEZ1, although intrinsically disordered, can gain structure not only upon addition of TFE but also upon encountering with interacting proteins. The low signal intensity in this experiment may be explained by a certain aggregation of FEZ1 at higher concentrations, although we did not observe any precipitation during the experiments. Additional CD analyses, however, revealed a gain of signal in lower concentrations of FEZ1 (data not shown).

SAXS measurements

The corrected and normalized experimental SAXS curves for 6xHis-FEZ1(1-227) and 6xHis-FEZ1(1-392)

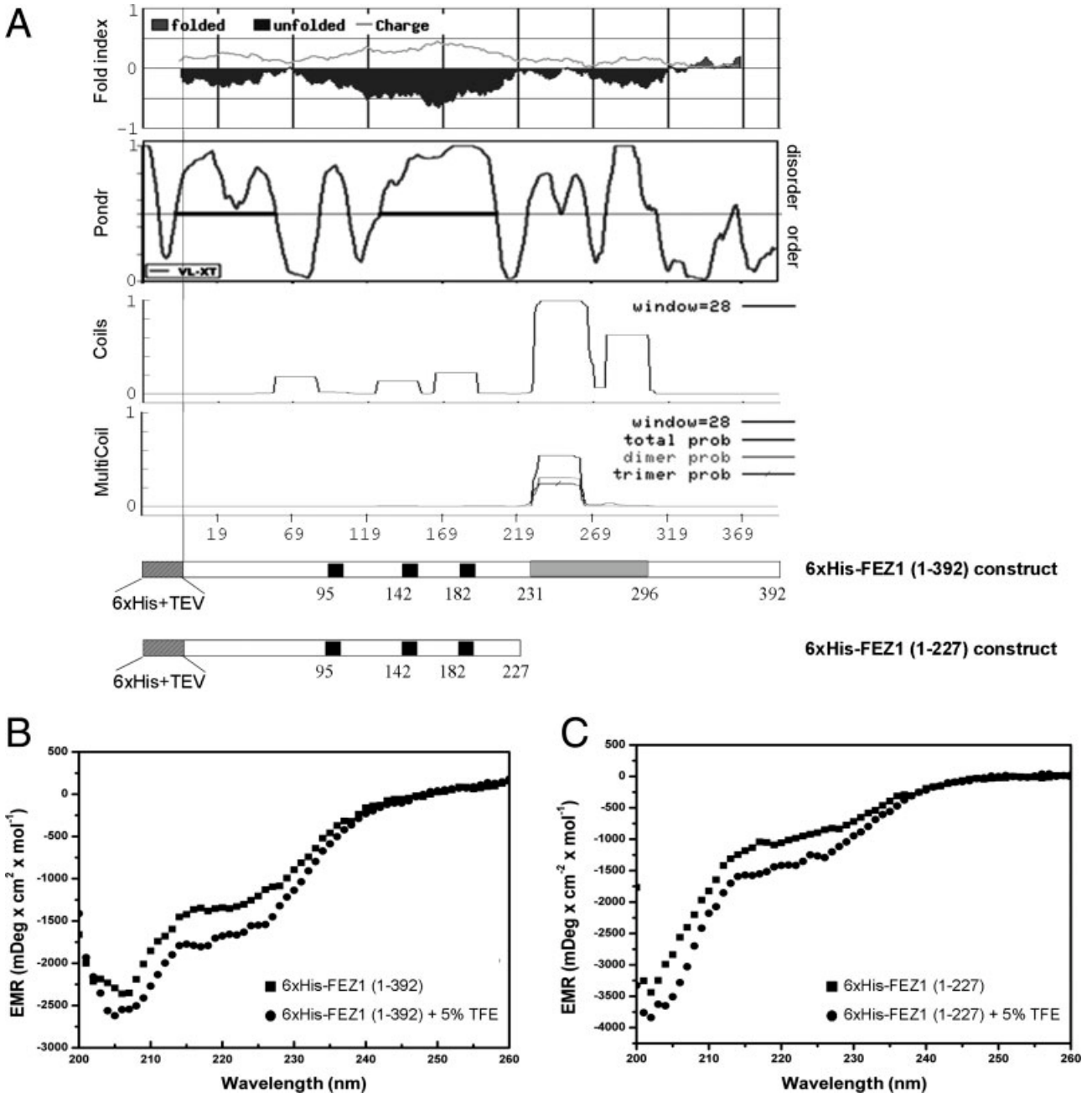


Figure 1

FEZ1 has features of a natively unfolded protein. (**A**) Unstructured and coiled-coil regions were predicted in the amino acid sequence of the fusion protein 6xHis-FEZ1(1-392) expressed from the plasmid peT28a. The y axes represent structure probability scores in the "Fold index" and "PONDR" softwares or coiled-coil probability in the "Coils" and "MultiCoil" programs. The x axis represents the amino acid sequence of human FEZ1. The fusion protein constructs of both FEZ1 protein versions have the initiating Methionine encoded by the vector, followed by an additional 31 amino acids: Gly, 6xHis-tag, 7 amino acids from the TEV cleavage site and 16 poly-linker encoded amino acids. In the case of both fusion protein constructs of FEZ1 the natural first Met residue was left out, to prevent usage of an additional translation initiation site. Schematic representations of the full-length human FEZ1(1-392) and the N-terminal fragment FEZ1(1-227) constructs are shown. At its N-terminal region FEZ1 contains three motifs rich in glutamic acid residues (black boxes). At its C-terminal region it contains an extensive coiled- coil region (grey box) of up to 65 amino acids. The indicated amino acid numberings refer to the native complete FEZ1 amino acid sequence. (**B**) Circular dichroism (CD) measurements indicate a secondary structure composition consisting mainly of random coil regions and alpha helices in 6xHis-FEZ1(1-392). A small but significant gain in alpha helix content is observed upon addition of 5% TFE in the CD spectrum of 6xHis-FEZ(1-392). (**C**) 6xHis-FEZ1(1-227) shows predominantly random coil content. Small change in the alpha-helical signal of 6xHis-FEZ1(1-227) was observe upon addition of 5% TFE.

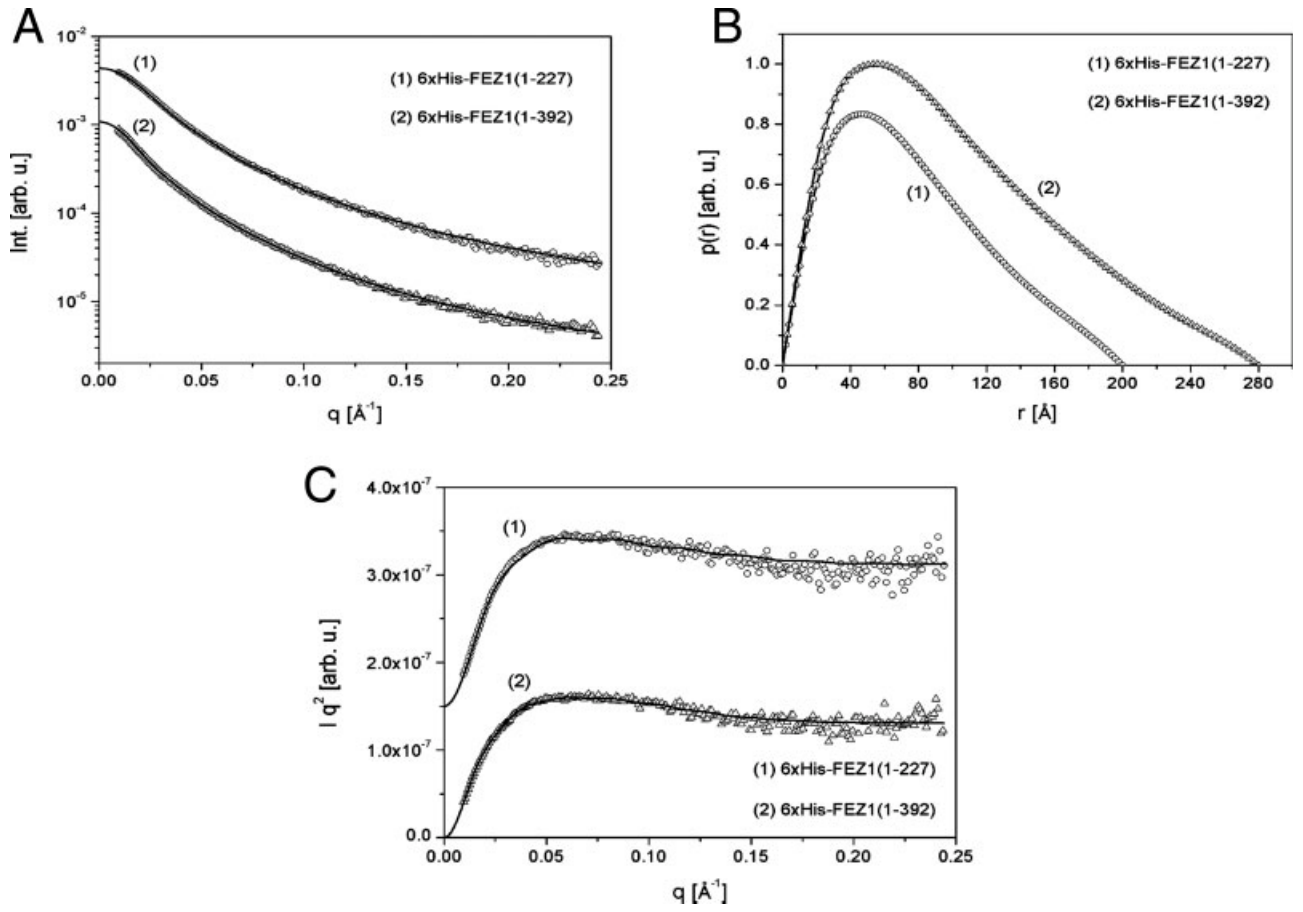


Figure 2

Experimental SAXS curves for human 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) in solution and the results of the fitting procedures. (**A**) Experimental scattering curve of 6xHis-FEZ1(1-227) (open circles) and 6xHis-FEZ1(1-392) (open triangles) and the theoretical fitting (solid lines) of data by using the program GNOM. (**B**) Pair distance distribution functions p(r). (**C**) Kratky Plots. In (A) and (C) the curve of 6xHis-FEZ1(1-392) was dislocated by a constant factor to allow better visualization and comparison.

are displayed in Figure 2(A), together with the individual GNOM curve fittings (solid lines). The $p(r)$ functions resulting from these calculations are shown in Figure 2(B). The maximum dimension values obtained for the molecules were 200 Å for 6xHis-FEZ1(1-227) and 280 Å for 6xHis-FEZ1(1-392). The R_g values derived from Guinier's law, Debye's law [see Eq. (1) above] and the $p(r)$ functions were respectively 62 ± 2 Å, 62.9 ± 0.3 Å, and 61.6 ± 0.6 Å for 6xHis-FEZ1(1-227), and 80 ± 2 Å, 84.6 ± 0.2 Å, and 83 ± 1 Å for 6xHis-FEZ1(1-392). The most reliable R_g value for each molecule is that obtained from the $p(r)$ function because it takes into account the complete experimental curve. However, all these values are in close agreement. It is interesting to compare the hydrodynamic radii estimated by DLS: 40 Å for 6xHis-FEZ1(1-227) and 47 Å for 6xHis-FEZ1(1-392) with those radii of gyration determined by SAXS (~62 and ~83 Å, respectively). The resulting ratios R_g/R_h are 1.55 and 1.77, respectively for the two constructs of FEZ1. Interestingly, ratios of this magnitude are frequently observed for proteins with extended shape. R_g/R_h ratios are reported to vary from 0.78 for homogeneous spheres, up to values nearing 2 for extended coils and prolate ellipsoids.[57]

Using BSA as reference sample, the molecular masses for FEZ1 proteins were obtained from SAXS results by comparison with the BSA scattering data. The values of ~60 kDa for 6xHis-FEZ1(1-227) and ~95 kDa for 6xHis-FEZ1(1-392) indicate that both the full length protein and the N-terminal construct exist in a dimeric state in solution, since the values obtained are approximately twice the theoretically values (calculated from the primary sequence using ProtParam)[58]: 29.7 kDa for 6xHis-FEZ1(1-227) and 48.6 kDa for 6xHis-FEZ1(1-392). These overall parameters suggest that both molecules have relatively elongated shapes, which is also evidenced by the

corresponding asymmetric and characteristic shapes of the p(r) functions. The Kratky plots of the two proteins are shown in Figure 2(C). Both representations are similar and they display a plateau for $q > 0.15$ Å^{-1}. The absence of a maximum clearly indicates that both the 6xHis-FEZ1(1-227) and 6xHis-FEZ1(1-392) present flexible chains and do not adopt compact conformations. These data suggest that both the full-length protein and the N-terminal construct are dimeric molecules with a largely open conformation.

Ab initio shape determination of human FEZ1

The low resolution *ab-initio* shapes of human 6xHis-FEZ1(1-227) and 6xHis-FEZ1(1-392) were restored from the experimental data using the approach described in "Experimental procedures". The low resolution structure models were derived from the experimental data without imposing any symmetry constrains for the dimers. As already mentioned in a previous section, in the absence of a unique solution for the *ab initio* calculations, several runs were performed for each protein with similar results. The obtained models from *ab-initio* calculations for 6xHis-FEZ1(1-227) and 6xHis-FEZ1(1-392) are shown in Figure 3(A,B), respectively. The NSD values for the set of 10 DAMMIN models ranged from 0.87 to 1.00. The low resolution of this representation only shows that the conformations of both proteins are extended. Inspection of the shapes obtained for 6xHis-FEZ1(1-227) [Fig. 3(A)] and 6xHis-FEZ1(1-392) [Fig. 3(B)], reveals a close resemblance of the conformations obtained for both proteins. The ratio between the DAMIN predicted volumes of 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) equals 1.52. This value is fairly close to ratio 1.65 of the number of amino acids of the two protein constructs.

Limited proteolysis of FEZ1

Natively unfolded proteins are due to their intrinsic flexibility more susceptible to proteolytical cleavage in comparison to more structured globular proteins which are more resistant.[59] Therefore, we submitted FEZ1 to proteolytical cleavage by thrombin and proteinase K [Fig. 4(A–D)]. We observed that when we incubated both 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) with thrombin a stable fragment of about 36 kDa appeared at 15 min and showed no further degradation up to 48 h. Other fragments are observed, which disappear until 48 h of digestion, indicating that they are intermediate products of cleavage. It is important to mention that both FEZ1 proteins, probably due to their high content of charged amino acids, show an anomalous mobility in SDS-PAGE: 6xHis-FEZ1(1-392) runs like a protein of about 70 kDa, although predicted to have a molecular mass of 48.6 kDa. 6xHis-FEZ1(1-227) runs like a protein of about 45–50 kDa, inspite a theoretical prediction of 29.7 kDa. The aberrant electroporetic mobility is a typical feature of intrinsic unstructured proteins and was observed for other proteins with intrinsically disordered regions and charged regions. Proteolysis with proteinase K generated similar results, giving rise to a prominent band of about 40 kDa in the case of both types of proteins [Fig. 4(C,D)]. However, this band is only stable up to 10 min, after which time point we can no longer detect 6xHis-FEZ1 in the SDS-PAGE. Next, we analyzed the same samples by anti-4xHis Western blot and were able to detect the same stable ~40 kDa band [Fig. 4(E,F)]. This indicates that both the full length and the N-terminal construct of 6xHis-FEZ1(1-227) still maintained their 6xHis-containing N-terminus and suggests that proteolysis occurs from the direction of the C-terminus inwards to the central region of the protein.

Finally, we submitted also endogenous FEZ1, from a Hek 293 cell extract, to limited proteolysis [Fig. 4(G)]. By Western blot, we found again a stable band of about 40 kDa, which maintained stability up to 30 min of incubation with proteinase K. Interestingly, this band was already present before incubation with proteinase K, suggesting that FEZ1 may suffer cleavage or proteolytical degradation either already *in vivo* or immediately after lysate preparation. Incubation of the control protein BSA for up to 40 min *in vitro* with proteinase K [Fig. 4(H)] did not show any signs of degradation. This demonstrates the capacity of the performed reactions to distinguish between globular (BSA) and open/unfolded proteins (FEZ1).

In summary, these results suggest that FEZ1 behaves like a natively unfolded protein with an open and flexible conformation which is susceptible to a significant amount of proteolysis under conditions that did not result in any cleavage in the globular control protein BSA. On the other hand we observe the accumulation of a relatively stable N-terminal fragment of about 40 kDa, which still contains the 6xHis-tag, and which is probably stabilized by the formation of the dimer (see Fig. 3).

Phosphorylation of FEZ1 by PKC at its C-termius *in vitro*

It was already known that FEZ1 interacts with PKCζ and that FEZ1 is phosphorylated.[3] A prediction of phosphorylated amino acids using the NetPhos software (http://www.cbs.dtu.dk/services/NetPhos/), showed that FEZ1 has possible Ser and Thr phosphorylation sites with varying probability of phosphorylation throughout its whole amino acid sequence (not shown). Using the program NetPhosK (http://www.cbs.dtu.dk/services/NetPhosK/) and defining a minimum score of 50% for PKC phosphorylation sites, we were able to identify six possible PKC sites (score ranging from 50.3–85.5), all of which are located in the C-terminal region of FEZ1 [Fig. 5(B,C)], right after the first predicted coiled-coil region

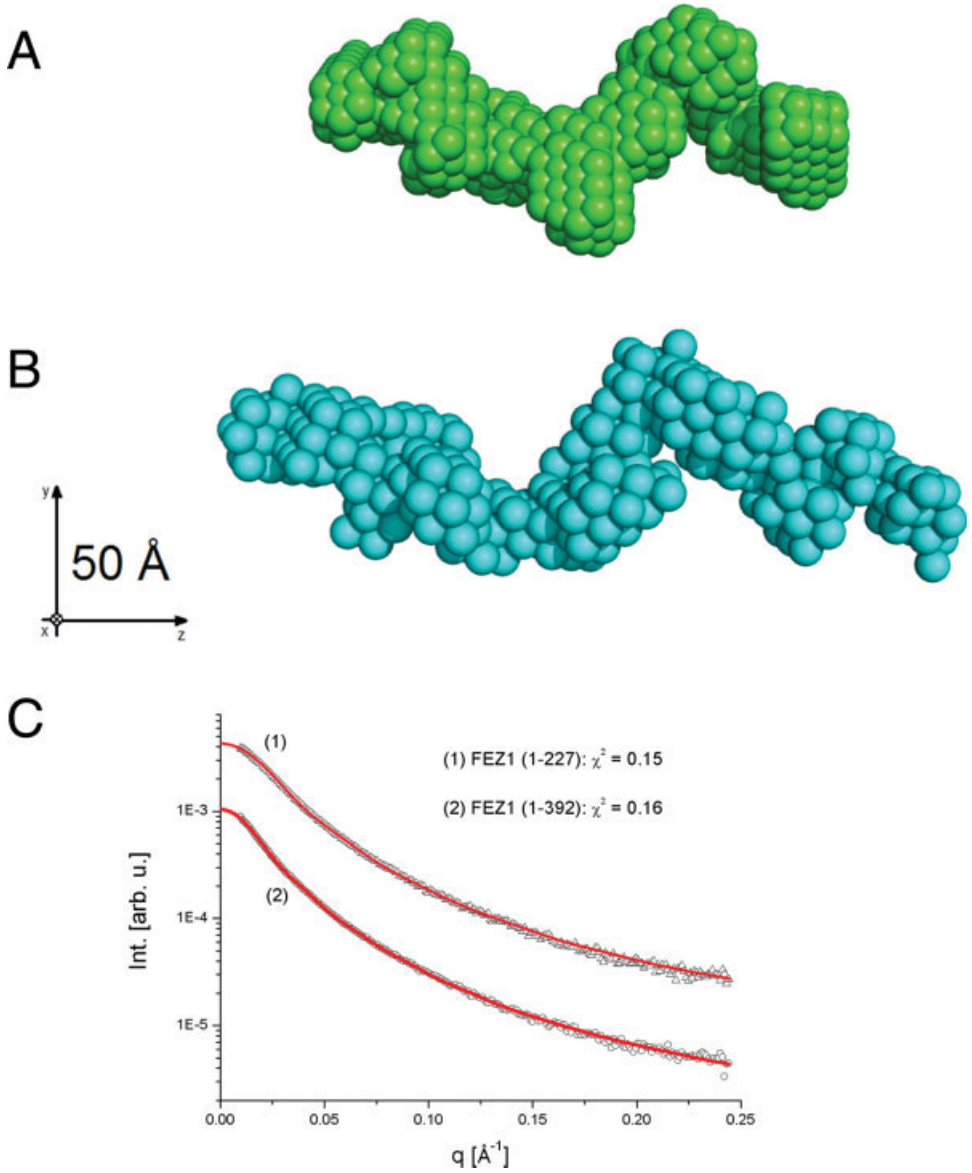


Figure 3

Low resolution *ab initio* models derived from the SAXS data. (**A**) Selected view of the 6xHis-FEZ1(1-227) dimer model (green). (**B**) Selected view of the 6xHis-FEZ1(1-392) dimer model (blue). The models were displayed by the program PyMOL.[67] (**C**) Curve fitting evaluated from the most typical model to the experimental data using DAMMIN.

(>aa 270). Two of them also scored high in phosphorylation probability in the NetPhos program [S326 and T331, Fig. 5(A)].

Based on these predictions we next employed both 6xHis-FEZ1(1-392) and 6xHis-FEZ1(1-227) in *in vitro* phosphorylation assays with recombinant PKC isoforms α, δ *e* ζ as well as with PKC-Pan [Fig. 5(D)]. 6xHis-FEZ1(1-227) as well as the control protein GST suffered some phosphorylation, but only by PKC-Pan suggesting that this phosphorylation may be rather unspecific. FEZ1(1-392) however was strongly phosphorylated by all tested PKC isoforms, especially by PKC Pan and PKC δ. The fact that the full length but not the N-terminal region of FEZ1 was phosphorylated under the same conditions, suggests strongly that the C-terminal region of FEZ1 could be the main target of phosphorylation by PKCs, as predicted theoretically [Fig. 5(A–C)]. The N-terminal of FEZ1(1-227) in contrast has only few predicted sites of phosphorylation by PKC and the observed phosphorylation by PKC Pan might be unspecific, as also indicated by the much weaker intensity of the band when compared to that of full length FEZ1.

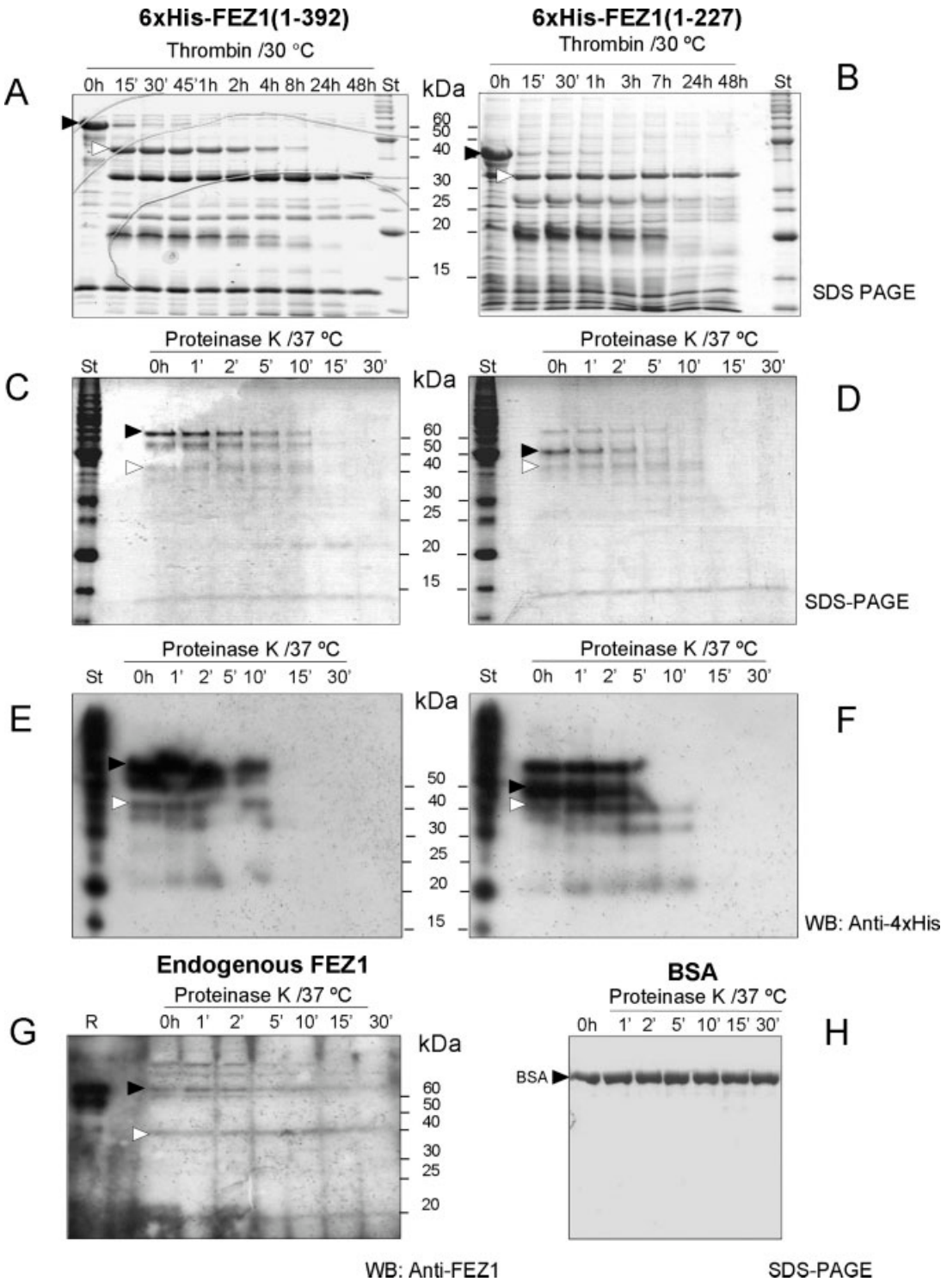


Figure 4

Limited proteolysis of 6xHis-FEZ1(1-392), 6xHis-FEZ1(1-227) and endogenous, cellular FEZ1 results in a common stable N-terminal fragment of ~40 kDa. (**A–D**) After incubation for the indicated times at the indicated temperatures of 6xHis-FEZ1(1-392) or 6xHis-FEZ1(1-227) with thrombin or proteinase K, samples were analyzed by SDS-PAGE (12.5% acryl amid). (**E–F**) anti-4xHis Western blot analysis of the proteolysis profile of the FEZ1 proteins obtained after digestion with proteinase K. (**G**) Proteolysis of endogenous FEZ1 from Hek 293 lysate, with proteinase K, followed by Western blot (anti FEZ1). (**H**) Control SDS-PAGE of proteinase K digested BSA. St: Protein molecular weight standard; R- recombinant 6xHis-FEZ1(1-392). The black arrows indicate undigested FEZ1 (A-G) or undigested BSA (H) and the white arrows the stable fragment from the N-terminus of FEZ1.

A

NetPhos

Serine predictions

Pos	Context	Score
301	LSLQSSRIE	0.305
326	NILQSGIRQ	0.735
334	QTFGSSGTD	0.334

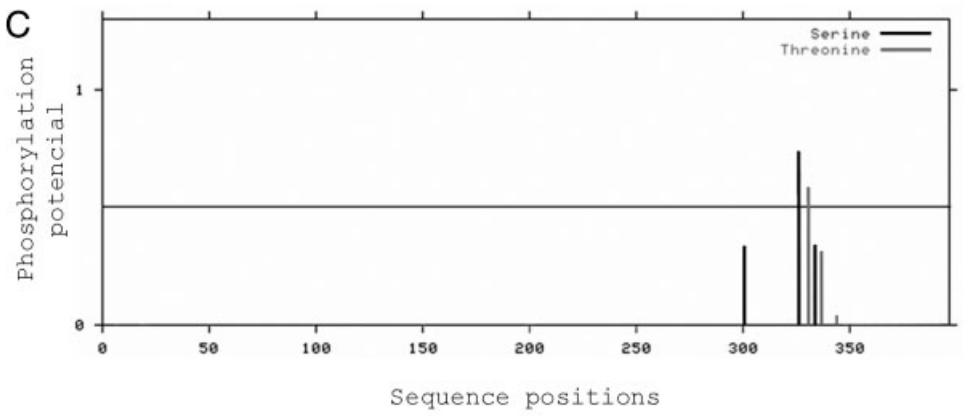
Threonine predictions

Pos	Context	Score
331	GIRQTFGSS	0.585
337	GSSGTDKQY	0.307
344	QYLNTVIPY	0.038

B

NetPhosK with ESS filtering

Pos:	AA:	Kinase:	Score:
301	S	PKC	0.713
326	S	PKC	0.852
331	T	PKC	0.855
334	S	PKC	0.503
337	T	PKC	0.845
344	T	PKC	0.777



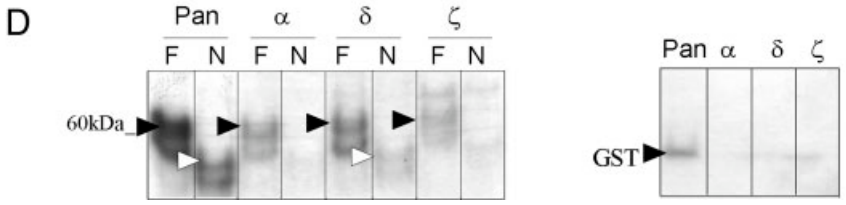


Figure 5

PKCα, δ and ζ phosphorylate 6xHis-FEZ1 mainly in its C-terminal region *in vitro*. (**A,B**) Commons predicted phosphorylation sites in human FEZ1 amino acid sequence as identified by NetPhos (A) and NetPhosK programs (B). (**C**) Graphic plot of the phoshorylation site prediction in FEZ1 as predicted by NetPhos. The peaks indicate the localization in the amino acid sequence (*x* axis) and the probability (*y* axis) of the phosphorylation. Only common PKC specific sites also predicted by NetPhosK are shown. (D) *In vitro* phosphorylation by different isoforms of PKC of 6xHis-FEZ1(1-392) (=F), 6xHis-FEZ1(1-227) (=N) and GST control protein. The indicated proteins have been submitted to *in vitro* phosphorylation by indicated PKC isoform, then analyzed by SDS-PAGE and autoradiography. The autoradiography is shown. The black arrows indicate phosphorylated 6xHis-FEZ1(1-392) (left panel) or GST control (right panel). White arrows indicate phosphorylated 6xHis-FEZ1(1-227).

These data in summary suggest that FEZ1 C-terminus is the main target of phosphorylation by PKCs. The fact that the C-terminal region also interacts with all the FEZ1 interacting proteins,[10] indicates that it may represent a critical region for phosphorylation dependent regulation of a subset of the interactions.

FEZ1 dimerization does not depend on its C-terminus or phosphorylation status

Next, we performed *in vitro* FEZ1/FEZ1 pull down assays to test if the phosphorylation status of FEZ1 influences its dimerization and to map protein regions involved in the dimerization process (see Fig. 6). In a first round of experiments we coupled full length or 6xHis-FEZ1(1-227) in both its apo- or phospho-forms (the latter previously phosphorylated by PKC *in vitro*) to Ni-NTA sepharose beads. A subsequent incubation was performed with either apo- or phospho GST-FEZ(1-392) or GST-FEZ(1-227) [Fig. 6(A,B)]. The results show clearly that the phosphorylation status of neither the beads-coupled FEZ1 nor that of GST-FEZ1 in the supernatant had any influence on the dimerization of FEZ1. Furthermore, we can also observe

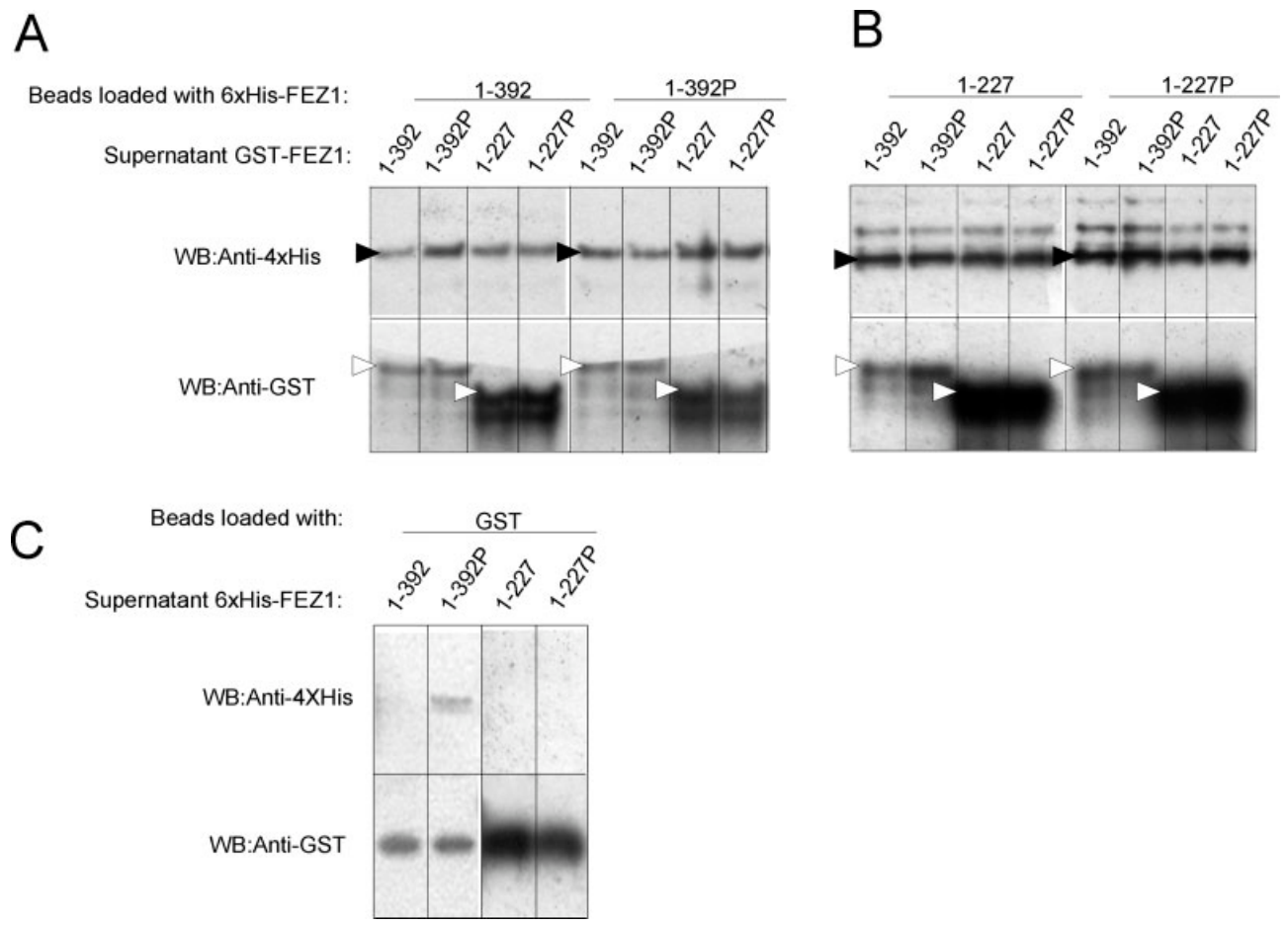


Figure 6

FEZ1 dimerization does not depend on its C-terminal region and is not influenced by its phosphorylation status. (**A,B**) *In vitro* pull-down assay between apo- or phospho-6xHis- or GST-FEZ1(1-392) (A) or FEZ1(1-227) constructs (B). 6xHis-FEZ1 fusion constructs were coupled to Ni-NTA sepharose beads, and the similar constructs fused to GST (as indicated in the figure) added to the supernatant for interaction (P, Phospho protein). After high stringency washes using NP-40 buffer, the aliquots were separated on two separate SDS-PAGE and transferred to a PVDF membranes for Western blot analysis. Blots were developed using either the monoclonal mouse anti-5xHis tag antibody or anti-GST monoclonal antibody 5.3.3. The black arrows indicate the positions of the 6xHis-proteins and the white arrows the GST fusion proteins. (C) GST was loaded onto glutathione sepharose beads. After washing beads were incubated with indicated 6xHis-FEZ1 apo- or phospho-proteins and washed stringently again. After separation on two separate SDS-PAGE the blots were developed as above.

that the dimerization occurs predominantly through the N-terminal region of FEZ1. In a second round of control experiments we coupled GST to glutathione-sepharose beads and subsequently incubated them with apo- or phospho-6xHis-FEZ1(1-227) or -FEZ1(1-392) [Fig. 6(C)]. No significant interaction was observed this time, except a weak but unspecific interaction between GST and phospho-6xHis-FEZ1(1-392). This demonstrates the specificity of the dimerization experiments described above. In summary, these data suggest, that FEZ1 dimerizes via its N-terminal domain and that in the dimer the outwards pointing C-terminal regions, which contain the coiled-coil regions, are free to interact with other proteins. Furthermore, the C-terminals of the FEZ1 dimer are accessible to phosphorylation by PKC, which may regulate some of the interactions.

FEZ1 phosphorylation does not promote conformational changes

The *in vitro* phosphorylation of neither 6xHis-FEZ1(1-227) nor 6xHis-FEZ1(1-392) promotes any significant changes in the secondary structure when monitored by circular dichroism spectroscopy [Fig. 7(A,B)]. Three Trp residues which are located in the N-terminal region of FEZ1 can be used as fluorescence probes to assay any changes in tertiary structure upon phosphorylation. However, no significant differences were observed in the fluorescence emission spectra of apo- and phospho-6xHis-FEZ1(1-227) or -FEZ1(1-392) [Fig. 7(C,D)]. Together these data indicate that the phosphorylation of FEZ1 does not promote any spectroscopically detectable changes in its conformation. This is in agreement with the observations that dimerization

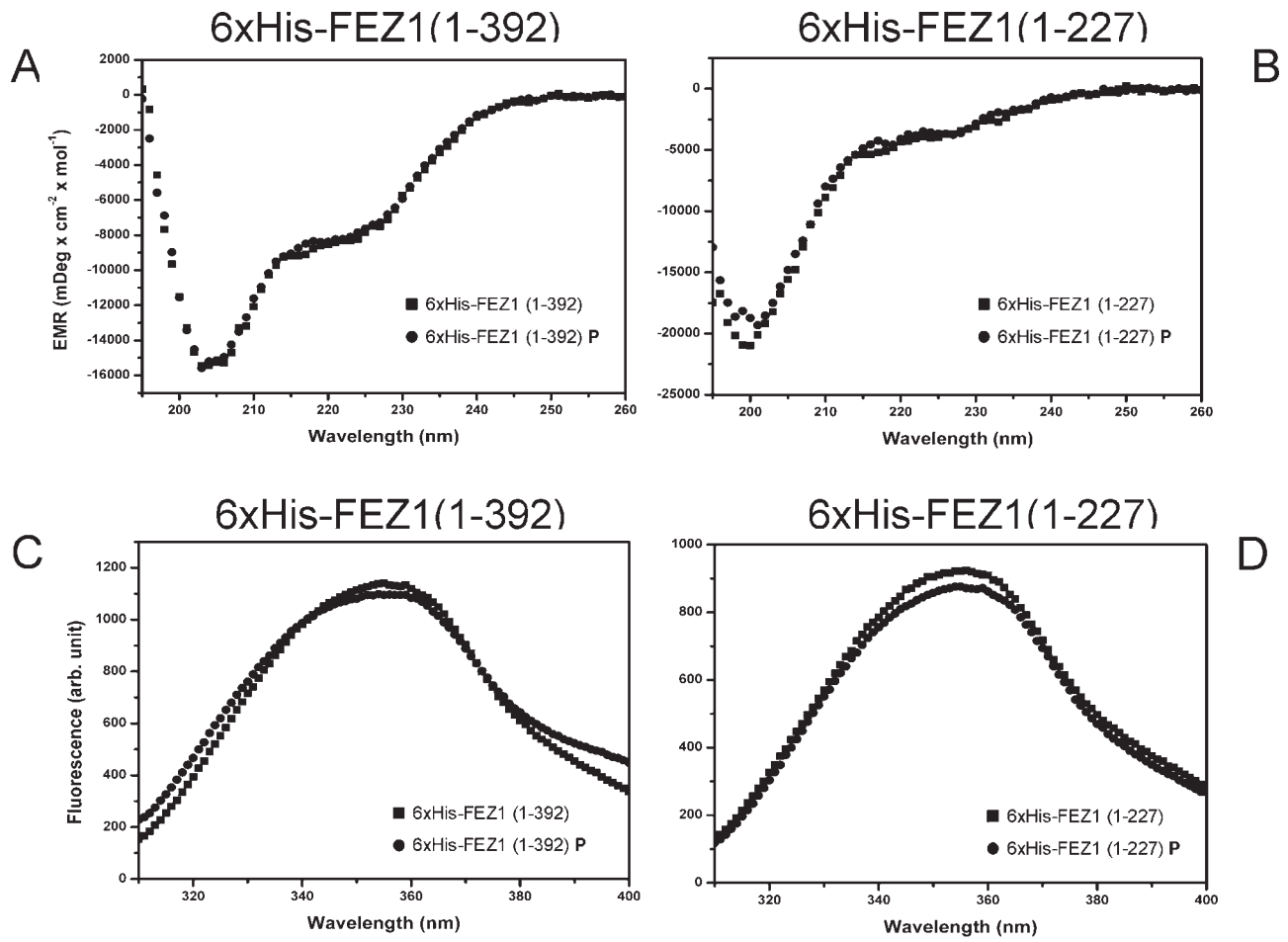


Figure 7

In vitro phosphorylation by PKC-ζ does not promote detectable conformational changes in FEZ1. (**A**) Circular dichroism experiment of 6xHis-FEZ1(1-392). Residual molar ellipticity of apo- and phospho (P) 6xHis-FEZ1(1-392) was measured from 195 to 260 nm in 20 m*M* Tris-HCl buffer, pH 7.5 with 10 m*M* $MgCl_2$ at 25°C, using a Jasco J-810 spectropolarimeter. (**B**) Circular dichroism experiment of apo- and phospho-6xHis-FEZ1(1-227). (**C**) Fluorescence spectroscopy experiment of apo- and phospho-6xHis-FEZ1(1-392). Samples were excited at 283 nm and fluorescence emission was recorded from 310 to 400 nm. (**D**) Fluorescence spectroscopy experiment of apo- and phospho-6xHis-FEZ1(1-227).

occurs at the N-terminus and phosphorylation predominantly at the C-terminus, which is not involved in the dimerization and structure maintenance but whose main function may be, as our data suggest, to interact with other proteins.

FEZ1 phosphorylation inhibits its interaction with CLASP2 *in vitro*

To test whether the phosphorylation status of FEZ1 can influence its interaction with other interacting proteins, we performed *in vitro* pull down assays between apo- and phospho-6xHis-FEZ1(1-392) and a subset of those proteins or protein fragments we had previously described to interact with FEZ1 (see Fig. 8).[10] Among the six GST fusion proteins tested, CLASP2(1046-1251) ceased to interact with FEZ1(1-392) when the latter had been phosphorylated by PKC *in vitro*. The observed lack of interaction with CLASP2 is specific, since none of the other five proteins showed a decreased interaction after phosphorylation of FEZ1 and GST even showed a weak but unspecific binding to phospho- but not apo-FEZ1. CLASP2 is a protein of important intracellular transport functions and localizes to the growing tip of microtubule.[60] A multiprotein transport complex associated to microtubule and consisting of FEZ1/JIP1/Kinesin has also been described.[8] Together, this suggests that FEZ1 may be involved in microtubular transport processes and that the interaction with CLASP2 could be regulated by phosphorylation.

DISCUSSION

FEZ1 is an orthologue of the *C. elegans* protein UNC-76 and both proteins have been shown to be of crucial

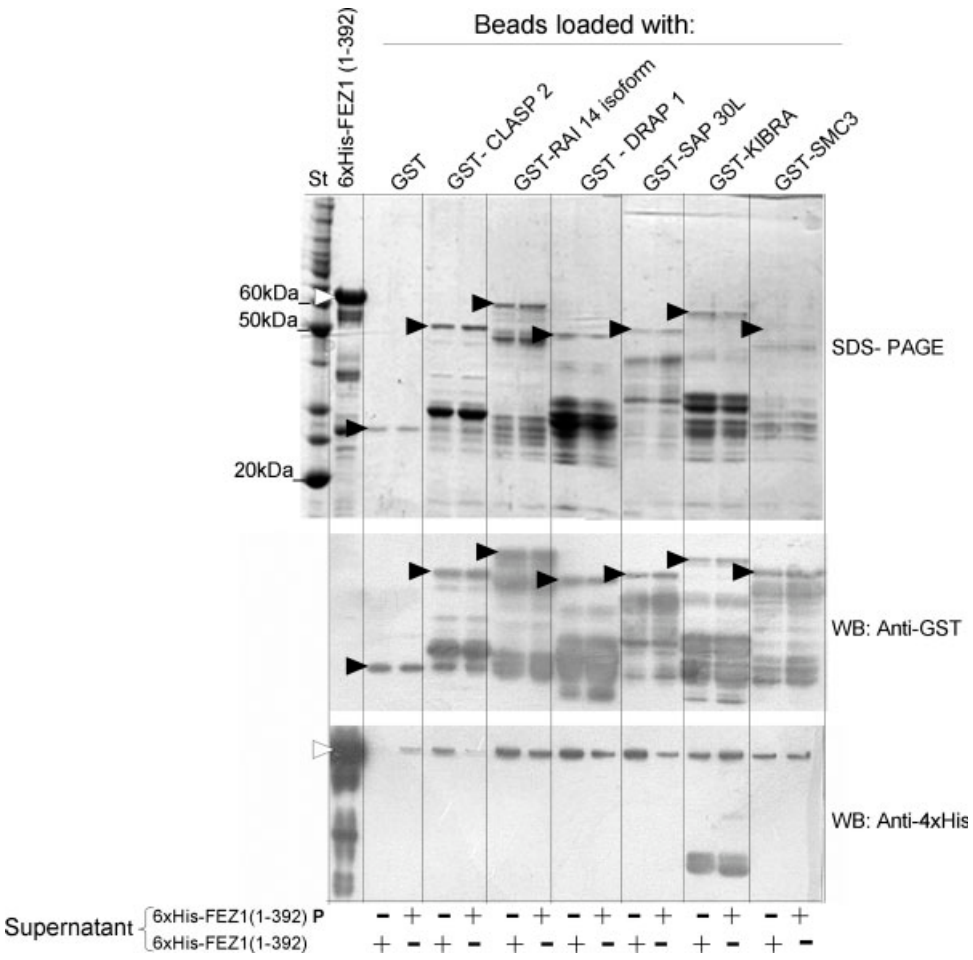


Figure 8

Phosphorylation of 6xHis-FEZ1(1-392) by PKC-ζ inhibits its interaction with the FEZ1 interaction domain of protein CLASP2 *in vitro*. *In vitro* pull-down assays between apo-(=FEZ1) or phospho-6xHis-FEZ1(1-392) (=FEZ1P) and selected proteins or protein fragments fused to GST that have been previously identified to interact with FEZ1.[10] GST, GST-CLASP2(1046-1251), GST-RAI14-isoform(720-983), GST-DRAP1(1-205), GST-SAP30L(1-183), GST-KIBRA(869-113), and GST-SMC3(881-1217) were loaded onto glutathione-sepharose beads. After washing the loaded beads were incubated with purified apo-6xHis-FEZ1(1-392) or phospho-6xHis-FEZ1(1-392), previously phosphorylated *in vitro* by PKC-ζ. After high stringency washes, samples were analyzed by Coomassie-stained SDS-PAGE and Western blotting. Blots were developed with monoclonal mouse anti-GST antibody 5.3.3 or anti-4xHis monoclonal antibody. Black arrows indicate the positions of the GST fusion proteins and the white arrows the position of FEZ1. ST, molecular weight standard proteins.

importance in neuronal differentiation and axonal outgrowth and elongation.[1,3,5] Recent experimental data pin point FEZ1 main activity in the course of axonal outgrowth to a microtubule associated cargo transport function,[7] since the inhibition of FEZ1 protein expression via iRNA-mediated interference leads to a retardation of the anterograde mitochondrial transport along the neurites of hippocampal neurons. Furthermore, the mitochondria in theses cells are elongated and the axon formation and neuronal polarization are inhibited, possibly in consequence of a lack of mitochondrial anterograde transport and an associated lack of metabolic energy supply in the growing axon tip. Further recent molecular evidence support this vision by associating FEZ1 physically and functionally to the heavy chain of the molecular motor protein Kinesin-1.[8] It was demonstrated that FEZ1 cooperates in conjunction with JIP-1 to activate the molecular motor activity of Kinesin-1.

Here we report a 3D low resolution structural model of the FEZ1 dimer, which we obtained by studying both the full length 6xHis-FEZ1 as well as a deletion construct representing amino acids 6xHis-FEZ1(1-227) of its N-terminus by SAXS measurements and *ab initio* modeling.

FEZ1 C-terminal region is an important protein–protein docking domain where all of the interacting proteins identified and tested up to now interact.[9,10] If we assume that the two C-termini of the full length FEZ1 dimer point outwards, away from the center of the dimer, which seem to involve its N-terminus [compare Fig. 3(A) and 3(B)], this may suggest that the interacting

proteins could bind at the C-termini for transport and that FEZ1 may function as a bivalent and dimeric adaptor protein.

Should this be true, we may begin to understand why FEZ1 binds to so many different proteins. Some or most of the over 42 cellular or viral proteins identified so far to interact with FEZ1 may actually represent transported cargo molecules.[9,10] The coiled-coil region in the C-terminal of FEZ1 may serve as a promiscuous docking domain for a large number of proteins, at least 27 of which actually show also a high probability of coiled-coil regions in their sequences, including those regions that were identified to interact with FEZ1.

In our previous yeast two-hybrid screen, where we used the C-terminal region (aa 221-392) of FEZ1 as a bait, we identified among other 14 proteins, FEZ1 itself.[10] Here, we demonstrated that the dimerization of FEZ1 occurs in its N-terminal region and that the C-terminus of FEZ1 is likely to be responsible for binding and possibly transport of several other proteins.[10] This may suggest that FEZ1 itself may in addition to its adaptor role also be a cargo molecule to be transported. More experiments need to be performed in order to test if a C-terminally mediated "transport" of FEZ1 by FEZ1 does actually happen in human cells. Another possibility is a regulatory role of this mode of FEZ1-FEZ1 binding. In this context it is worth mentioning that the super-expression of FEZ1 in fibroblast cells conferred resistance to retroviral infection to these cells.[61] The authors demonstrated that FEZ1 mRNA super expression prevented viral nuclear entry, thereby blocking the viral life cycle. Knowing that FEZ1 may function as a transport adaptor we can now speculate that excess FEZ1 can bind to the viral capsid or capsid associated cellular proteins, thereby preventing the virus to interact with the microtubular motor transport machinery. A second possibility is that at the highly elevated concentration, FEZ1 binds in a regulatory fashion to the C-terminal protein interaction domain of FEZ1, as it was observed *in vitro*,[10] and thereby inhibits the interaction of FEZ1 with its cargo proteins. One of the possible candidate proteins that can be a binding target/cargo of FEZ1 on the macromolecular complex of the HIV viral capsid and the associated viral and human proteins is HTATSF1 (HIV TAT specific factor 1), which has been in deed identified as an interacting protein in a yeast two-hybrid system assay.[9]

FEZ1 mediated transport functions may furthermore not be limited to viral components,[6,61] and proteins in general,[9,10] but can be possibly extended to organelles or other larger cellular components, including chromosomes. Recent studies demonstrated that FEZ1 may be involved in transport of mitochondria,[7] binds to KIF3,[6] and KIF5,[62] and possibly even binds to chromosomes or other nuclear components mediated via interactions with a ternary complex of KIF3A/SMAP/HCAP.[63]

Besides the dimeric SAXS structure of FEZ1 shows a long and extended conformation, which is in agreement with our theoretical prediction as well as with the spectroscopic and limited proteolysis analyses. Altogether, these data suggest that FEZ1 may belong to the group of natively unfolded proteins. The Kratky representation of the SAXS curves is typical of proteins with an elevated degree of flexibility and elevated random coil content. It is worth pointing out that the main function of the FEZ1 N-terminal domain may be to provide a platform for dimerization. Through dimerization and by means of the observed arrangement in the model, the C-termini remain apart from each other and are free for interaction with the binding protein partners, which may or may not represent cargo proteins in transport processes.

We may speculate that FEZ1 dimerizes in its N-terminal region, since both 6xHis-FEZ1(1-227) and 6xHis-FEZ1(1-392) show the same rough dimer arrangement (see Fig. 3), involving a N-terminal region which is heavily hydrated (data not shown). Possibly, this region contains one or more of the three poly-Glu motifs found in the middle of FEZ1 N-terminal region [Fig. 1(A), amino acids 95-194). The comparison of both models may also suggest a possible exposition for the two C-terminal regions in a FEZ1(1-392) dimer away from the heavily hydrated dimerization region located at the center of the full length FEZ1 dimer. This may be functionally important, bearing in mind the previous data that the C-termini of FEZ1 represents essentially a docking domain involved in all protein–protein interactions described so far.[10] Therefore the main function of the FEZ1 N-terminal domain may be to provide a platform for dimerization. Through such a mode of dimerization, the C-termini would remain apart from each other and hence free for interaction with the binding protein partners, which may or may not represent cargo proteins in transport processes.

FEZ2 is a human paralogue of FEZ1 with ubiquitous expression, whereas FEZ1 is predominantly found in neuronal cells. Based on this it is tempting to speculate that not only FEZ1-FEZ1 but also FEZ2-FEZ2 dimerization may occur in a given cell. A theoretically possible FEZ1-FEZ2 heterodimer seems rather unlikely, firstly because both proteins are usually not expressed in the same cells (except for neuronal cells), and secondly because the N-terminal region which we found here to be crucial for the dimerization of FEZ1 shows very little sequence conservation between FEZ1 and FEZ2 (28% identity and 60% similarity of the amino acid sequences). Clearly, more experiments are necessary to address these open but important issues.

The rather unfolded nature of FEZ1 fits well with its hub position in protein interaction networks, where it has been found to bind to at least 42 different proteins to date. In fact flexibility and high surface charge have been predicted and found to be important properties for

hub proteins, which have the need to interact with multiple proteins.[12] Furthermore, multiple functions have been described for various natively unfolded proteins, similar to FEZ1, which has been implicated in transcriptional regulation, intracellular transport and neuronal cell development.[13,64,65]

The majority of these functions are regulated by protein phosphorylation, which in this case may not regulate the protein activity and binding to other proteins through conformational changes, since natively unfolded proteins tend to be rather structurally disordered. Regulation of interaction may occur rather by simple electrostatic or steric blockage of a region important in contacting a specific interacting protein. Indeed our studies of the *in vitro* phosphorylation of FEZ1 by PKC suggest that this could be the case also for FEZ1. We found that the overall secondary and tertiary structures of FEZ1 are basically not affected by its phosphorylation through PKC. The additional findings that the phophorylation is almost absent in the N-terminal region of FEZ1 suggested that the C-terminal is the main site for phosphorylation, *in vitro*. This is further supported by the fact that dimerization, which involved the N-terminal region of FEZ1, is not at all affected by phosphorylation.

The C-terminus represents on the other hand the main protein docking region of FEZ1, suggesting that phophorylation could be an important event for the regulation of interaction with other proteins. We indeed found that interaction with the microtubular tip protein CLASP2 *in vitro* is practically blocked after phosphorylation of FEZ1. The interaction of FEZ1 with CLASP2 may be of crucial importance to understand the mechanism underlying our recent observation that the over-expression of GFP-FEZ1 in Hek293 cells causes the microtubular dependent formation of multi-lobulated nuclei, a hall mark of aggressive leukemia cells.[66] Further studies must address the importance of FEZ1 phosphorylation *in vivo* in this process and the elucidation of the mechanism that causes the multi-lobulation of the nuclei will be important for our understanding of the steps that lead to development of these cancer cells. It is however tempting to speculate that FEZ1 bivalency and its characteristic as a hub protein, able to interact with many proteins, especially the microtubular components including CLASP2, KIF3A, and tubulin, may prove to be essential to explain the observed phenotype.

In summary, our data suggest that human FEZ1 has several features of natively unfolded proteins. Its dimer configuration in solution, may suggest that it could act as a bivalent transport adaptor protein. The two C-terminal regions of the dimer are pointing outwards and may serve as docking domains for cargos and/or to make contact with microtubules via CLASP2, KIF3A or other connecting proteins, yet to be identified. Future high resolution structural studies of protein complexes may reveal the details of the previously described interactions with cargo proteins or microtubule associated proteins.[9,10] Functional cellular studies should also be performed to determine FEZ1 exact role in transport processes and the exact mode by which it connects to microtubules.

ACKNOWLEDGMENTS

The authors thank Maria Eugenia R. Camargo, Zildene D. Correa and Adriana C. Alves for technical assistance. This work was performed at the: Laboratório Nacional de Luz Síncrotron, Campinas, SP, Brasil.

REFERENCES

1. Bloom L, Horvitz HR. The *Caenorhabditis elegans* gene *unc*-76 and its human homologs define a new gene family involved in axonal outgrowth and fasciculation. Proc Natl Acad Sci USA 1997; 94:3414–3419.
2. Fujita T, Ikuta J, Hamada J, Okajima T, Tatematsu K, Tanizawa K, Kuroda S. Identification of a tissue-non-specific homologue of axonal fasciculation and elongation protein zeta-1. Biochem Biophys Res Commun 2004;313:738–744.
3. Kuroda S, Nakagawa N, Tokunaga C, Tatematsu K, Tanizawa K. Mammalian homologue of the *Caenorhabditis elegans* UNC-76 protein involved in axonal outgrowth is a protein kinase C zeta-interacting protein. J Cell Biol 1999;114:403–411.
4. Honda A, Miyoshi K, Baba K, Taniguchi M, Koyama Y, Kuroda S, Katayama T, Tohyama M. Expression of fasciculation and elongation protein zeta-1 (FEZ1) in the developing rat brain. Brain Res Mol Brain Res 2004;122:89–92.
5. Okumura F, Hatakeyama S, Matsumoto M, Kamura T, Nakayama K. Functional regulation of FEZ1 by the U-box-type Ubiquitin ligase E4B contribuyes to neuritogenesis. J Biol Chem 2004; 279:53533–53543.
6. Suzuki T, Okada Y, Semba S, Orba Y, Yamanouchi S, Endo S, Tanaka S, Fujita T, Kuroda S, Nagashima K, Sawa H. Identification of FEZ1 as a protein that interacts with JC virus agnoprotein and microtubules. J Biol Chem 2005;280:24948–24956.
7. Ikuta J, Maturana A, Fujita T, Okajima T, Tatematsu Kenji, Tanizawa K, Kuroda S. Fasciculation and elongation protein zeta-1 (FEZ1) participates in the polarization of hippocampal neuron by controlling the mitochondrial motility. Biochem Biophys Res Commun 2007;353:127–132.
8. Blasius TL, Cai D, Jih GT, Toret CP, Verhey KJ. Two binding partners cooperate to activate the molecular motor Kinesin-1. J Cell Biol 2007;176:11–17.
9. Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, Koeppen S, Timm J, Mintzlaff S, Abraham C, Bock N, Kietzmann S, Goedde A, Tokso ZE, Droege A, Krobitsch S, Korn B, Birchmeier W, Lehrach H, Wanker EE. A human protein-protein interaction network: a resource for annotating the proteome. Cell 2005;122:957–968.
10. Assmann EM, Alborghetti MR, Camargo MER, Kobarg J. FEZ1 dimerization and interaction with transcription regulatory proteins involves its coiled-coil region. J Biol Chem 2006;281:9869–9881.
11. Jeong H, Mason SP, Barabasi AL, Oltvai ZN. Lethality and centrality in protein networks Nature 2001;411:41–42.
12. Patil A, Nakamuraa H. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. FEBS Lett 2006;580:2041–2045.
13. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J 2005;272:5129–5148.

14. Ekman D, Light S, Bjorklund AK, Elofsson A. What properties characterize the hub proteins of the protein-protein interactrion network of *Saccharomyces cerevisiae*? Genome Biol 2006,7:R45.
15. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM. Intrinsic disorder is a common fetature of hub proteins from four eukaryotic interactomes. PLoS Comput Biol 2006,2:e100.
16. Dosztanyi Z, Chen J, Dunker AK, Simon I, Tompa P. Disorder and sequence repeats in hub proteins and their implications for network evolution. J Proteome Res 2006;5:2985–2995.
17. Singh GP, Dash D, Intrinsic disorder in yeast transcriptional regulatory network. Proteins 2007;68:602–605.
18. Tompa P. The interplay between structure and function in intrinsically unstructured proteins. FEBS Lett 2005;579:3346–3354.
19. Uversky VN, Fink AL. Conformational behavior of human alpha-synuclein is modulated by familial Parkinson's disease point mutations A30P and A53T. Neurotoxicology 2002;23:553–567.
20. Uversky VN, Gillespie JR, Millett IS, Khodyakova AV, Vasiliev AM, Chernovskaya TV, Vasilenko RN, Kozlovskaya GD, Dolgikh DA, Fink AL, Doniach S, Abramov VM. Natively unfolded human prothymosin alpha adopts partially folded collapsed conformation at acidic pH. Biochemistry 1999;38:15009–15016.
21. Tidow H, Melero R, Mylonas E, Freund SMV, Grossmann JG, Carazo JM, Svergun DI, Valle M, Fersht A. Quaternary structures of tumor surppressor p53 and a specific p53-DNA complex. Proc Natl Acad Sci USA 2007;104:12324–12329.
22. Romero P, Obradovic Z, Dunker AK. Sequence data analysis for long disordered regions prediction in the calcineurin family. Genome Inf Ser 1997;8:110–124.
23. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. Proteins 2001;42:38–48.
24. Uversky VN, Gillespie JR, Fin AL. Why are natively unfolded proteins unstructured under the physiological conditions? Proteins: Struct Funct Genet 2000;42:415–427.
25. Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, Dunker AK. Comparing and combining predictors of mostly disordered proteins. Biochemistry 2005;44:1989–1200.
26. Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg E, Man O, Beckmann JS, Silman I, Joel L, Sussman JL. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics 2005;21:3435–3438.
27. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein disorder prediction: implications for structural proteomics. Structure 2003;11:1453–1459.
28. MacCallum RM. Order/disorder prediction with self organizing maps. Online paper. Available at: http://www.forcasp.org/paper2127.html. CASP 6 meeting.
29. Cheng J, Sweredoski M, Baldi P. Accurate prediction of protein disordered regions by mining protein structure data. Data Mining Knowl Discov 2005;11:213–222.
30. Linding R, Russell RB, Neduva V, Gibson TJ. Glob plot: exploring protein sequences for globularity and disorder. Nucleic Acids Res 2003;31:3701–3708.
31. Dosztanyi Z, Csizmok V, Tompa P, Simon I. IU Pred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 2005;21:3433–3434.
32. Coeytaux K, Poupon A. Prediction of unfolded segments in a protein sequence based on amino acid composition. Bioinformatics 2005;21:1891–1900.
33. Yang ZR, Thomson R, McNeil P, Esnouf RM. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics 2005;21:3369–3376.
34. Vullo A, Bortolami O, Pollastri G, Tosatto S. Spritz: a server for the prediction of intrinsically disordered regions in protein sequences using kernel machines Nucleic Acids Res 2006;34:W164–W168.
35. Garbuzynskiy SO, Lobanov MY, Galzitskaya OV. To be folded or to be unfolded? Protein Sci 2004;13:2871–2877.
36. Galzitskaya OV, Garbuzynskiy SO, Lobanov MY. Prediction of natively unfolded regions in protein chain. Mol Biol (Mosk) 2006; 40:341–348.
37. Vucetic S, Brown CJ, Dunker AK, Obradovic Z. Flavors of protein disorder. Proteins 2003;52:573–584.
38. Obradovic Z, Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK. Predicting intrinsic disorder from amino acid sequence. Proteins 2003;53:566–572.
39. Obradovic Z, Peng K, Vucetic S, Radivojac P, Dunker AK. Exploiting heterogeneous sequence properties improves prediction of protein disorder. Proteins 2005;7:176–182.
40. Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences. Science 1991;252:1162–1164.
41. Wolf E, Kim PS, Berger B. MultiCoil: a program for predicting two and three-stranded coiled coils. Protein Sci 1997;6:1179–1189.
42. Cavalcanti LP, Torriani IL, Plivelic TS, Oliveira CLP, Kellermann G, Neuenschwander. Two new sealed sample cells for small angle X-ray scattering from macromolecules in solution and complex fluids using synchrotron radiation. R Rev Sci Inst 2004;75:4541–4546.
43. Orthaber D, Bergmann A, Glatter O. J Appl Crystallogr 2000;33: 218–255.
44. Guinier A, Fournet G. Small angle scattering of X-rays, translated by Walker CB and Yudowitch KL. New York: Wiley; 1955, pp 5–78.
45. Feigin LA, Svergun DI. Structure analysis by small angle X-ray and neutron scattering. New York: Plenum Press; 1987, pp 59–104.
46. Glatter O, Kratky O. Small angle X-ray scattering. New York: Academic Press; 1982, pp 17–50.
47. Kirste RG, Oberthür RC. Small angle X-ray scattering. In: Glatter O, Kratky O, editors. London, New York: Academic Press; 1982, pp 387–431.
48. Debye PJ. Molecular-weight determination by light scattering. Phys Colloid Chem 1947;51:18–32.
49. Calmettes P, Durand D, Desmadril M, Minard P, Receveur V, Smith JC. How random is a highly denatured protein? Biophys Chem 1994;53:105–114.
50. Perez J, Vachette P, Russo D, Desmadril M, Durand D. Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering. J Mol Biol 2001;308:721–743.
51. Moncoq K, Broutin I, Craescu CT, Vachette P, Ducruix A, Durand D. SAXS study of the PIR domain from the Grb14 molecular adaptor: a natively unfolded protein with a transient structure primer? Biophys J 2004;87:4056–4064.
52. Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Cryst 1992;25:495–503.
53. Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 1999;76:2879–2886.
54. Volkov VV, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. J Appl Cryst 2003;36:860–864.
55. Denning DP, Uversky V, Patel SS, Fink AL, Rexach M. The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein. J Biol Chem 2002;36:33447–33455.
56. Uversky VN. Natively unfolded proteins: a point where biology waits for physics. Protein Sci 2002;11:739–756.
57. Rizos AK, Tsikalas I, Morikis D, Galanakis P, Spyroulias GA, Krambovitis E. Characterization of the interaction between peptides derived from the gp120/V3 domain of HIV-1 and the amino terminal of the chemokine receptor CCR5 by NMR spectroscopy and light scattering. J Non-Cryst Solids 2006;352:4451–4458.
58. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The proteomics protocols handbook. Totowa, NJ: Humana Press; 2005, pp 571–607.

59. Fontana A, Polverino de Laureto P, Spolaore B, Frare E, Picotti P, Zambonin M. Probing protein structure by limited proteolysis. Acta Biochim Pol 2004;51:299–321.
60. Mimory-Kiyouse Y, Grigoriev I, Lansberger G, Sasaki H, Matsui C, Severin F, Galjart N, Grosveld F, Vorobjev I, Tsukita S, Akhmanova A. CLASP1 and CLASP2 bind to EB1 and regulate microtubule plus-end dynamics at the cell cortex. J Cell Biol 2005;168:141–153.
61. Naghavi MH, Hatziioannou T, Gao G, Goff SP. Overexpression of fasciculation and elongation protein ζ-1 (FEZ1) induces a post-entry block to retroviruses in cultured cells. Genes Dev 2005;19:1105–1115.
62. Fujita T, Maturana AD, Ikuta J, Hamada J, Walchli S, Suzuki T, Sawa H, Wooten MW, Okajima T, Tatematsu K, Tanizawa K, Kuroda S. Axonal guidance protein FEZ1 associates with tubulin and kinesin motor protein to transport mitochondria in neurites of NGF-stimulated PC12 cells. Biochem Biophys Res Commun 2007; 361:605–610.
63. Shimizu K, Shirataki H, Honda T, Minami S, Takai Y. Complex formation of SMAP/KAP3, a KIF3A/B ATPase motor-associated protein, with a human chromosome-associated polypeptide. J Biol Chem 1998;273:6591–6594.
64. Dafforn TR, Smith CJI. Natively unfolded domains in endocytosis: hooks, lines and linkers. EMBO reports 2004;5:1046–1052.
65. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life J Mol Biol 2004;337:635–645.
66. Lanza DCF, Trindade DM, Assmann EM, Kobarg J. Over-expression of GFP-FEZ1 causes generation of multi-lobulated nuclei mediated by microtubules in HEK293 cells. Exp Cell Res 2008, doi: 10.1016/j.yexcr, 2008.02.012
67. DeLano WL. The PyMOL molecular graphics system, DeLano Scientific, Palo Alto, CA, USA, 2002.