
An avidin-like domain that does not bind biotin is adopted for oligomerization by the extracellular mosaic protein fibropellin

ITAI YANAI,^{1,4,5} YONG YU,^{2,3,4} XIAHUI ZHU,^{3,6} CHARLES R. CANTOR,^{2,3} AND ZHIPING WENG^{1,3}

¹Bioinformatics Program, ²Center for Advanced Biotechnology, and ³Biomedical Engineering Department, Boston University, Boston, Massachusetts 02215, USA

(RECEIVED May 27, 2004; FINAL REVISION September 29, 2004; ACCEPTED October 9, 2004)

Abstract

The protein avidin found in egg white seems optimized for binding the small vitamin biotin as a stable homotetramer. Indeed, along with its streptavidin ortholog in the bacterium *Streptomyces avidinii*, this protein shows the strongest known noncovalent bond of a protein with a small ligand. A third known member of the avidin family, as similar to avidin as is streptavidin, is found at the C-terminal ends of the multidomain fibropellin proteins found in sea urchin. The fibropellins form a layer known as the apical lamina that surrounds the sea urchin embryo throughout development. Based upon the structure of avidin, we deduced a structural model for the avidin-like domain of the fibropellins and found that computational modeling predicts a lack of biotin binding and the preservation of tetramerization. To test this prediction we expressed and purified the fibropellin avidin-like domain and found it indeed to be a homotetramer incapable of binding biotin. Several lines of evidence suggest that the avidin-like domain causes the entire fibropellin protein to tetramerize. We suggest that the presence of the avidin-like domain serves a structural (tetrameric form) rather than functional (biotin-binding) role and may therefore be a molecular instance of exaptation—the modification of an existing function toward a new function. Finally, based upon the oligomerization of the avidin-like domain, we propose a model for the overall structure of the apical lamina.

Keywords: avidin-like domain; biotin; fibropellin; streptavidin; avidin

The initial steps in the development of an embryo require an enclosure to keep the differentiating blastomeres together. In the sea urchin, *Strongylocentrotus purpuratus*, this enclosure, known as the hyaline layer, is secreted by the fertilized egg and tightly surrounds the embryo until metamorphosis (Cameron and Holland 1985; Spiegel et al. 1989;

Burke et al. 1998). The hyaline layer is a complex matrix that serves as a substrate to which cells can adhere to and migrate upon during specific periods of development. It is itself composed of distinct layers (Hall and Vacquier 1982). In 1 M glycine solution, hyalin in the outer hyaline layer dissolves and leaves behind the apical lamina, which is a fibrous matrix composed predominantly of three glycoproteins, the fibropellins (Hall and Vacquier 1982).

The fibropellins are mosaics of three domains: CUB (Bork and Beckmann 1993), EGF (Hursh et al. 1987), and surprisingly, an avidin-like domain (Hunt and Barker 1989). They are present as two genes in the sea urchin genome, SpEGF I and SpEGF III (Bisgrove et al. 1991; Bisgrove and Raff 1993). SpEGF I produces two splice variants, Ia and Ib (Bisgrove et al. 1991). All three fibropellin proteins share

Reprint requests to: Zhiping Weng, Biomedical Engineering Department, 44 Cummington Street, Boston University, Boston, MA 02215, USA; e-mail: zhiping@bu.edu; fax: (617) 353-6766.

⁴These authors contributed equally to this work.

Present addresses: ⁵Department of Molecular and Cellular Biology (MCB), Harvard University, Cambridge, MA 02138, USA; ⁶AstraZeneca R&D Boston, Waltham, MA 02451, USA.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.04898705>.

the same domain architecture differing only in the number of EGF repeats. The domain motif is EGFICUBIEGF[8, 13, or 21 repeats]AVIDIN-LIKE (Fig. 1). Avidin forms stable homotetramers, and is famous for “rapid and almost irreversible binding” of biotin and anything that may be attached to it (Green 1990). To date, this protein has only been observed in the *Aves* lineage (birds), the bacterium *Streptomyces avidini*, where it is known as streptavidin, and as a domain in the fibropellin proteins. The affinity of avidin and streptavidin for biotin has become an invaluable tool for biotechnology and is very well studied; however, the present work constitutes the first experimental study of the third member of the avidin family, the avidin-like domain of fibropellins, henceforth *edin* (Egf, avidIN).

Results and Discussion

Computational analysis indicates that edin is a tetramer and does not bind biotin

Although sharing only 33% identity (46% similarity) at the protein sequence level, avidin and streptavidin have an identical β -barrel fold, with only a 1.15 Å root-mean-square deviation. Edin is 33% and 28% identical to avidin and streptavidin, respectively. Based upon the known X-ray structures of avidin, we have built a homology model of edin (Peitsch et al. 1995; Szustakowski and Weng 2000; Wang et al. 2000). Figure 2A shows the multiple sequence alignment of avidin, streptavidin, and the edin domains of SpEGFI and SpEGFIII, constructed based upon the structural alignment of avidin and streptavidin. Edin I and III are 80% identical, and all residues discussed in this paper are conserved between the two.

Overall the biotin-binding pocket is well conserved among the three proteins, with 10 of the 16 key residues in avidin and streptavidin present in edin (Fig. 2A). Nonetheless, the sequence alignment indicates that edin cannot bind biotin. A key component of biotin binding is the set of four Trp residues that surround the ligand. Two of these Trps are absent in edin. One of them is found as Arg65, and is likely to form a salt bridge with Glu35. Glu35 is mutated from Thr38 in avidin where the residue is not involved in biotin

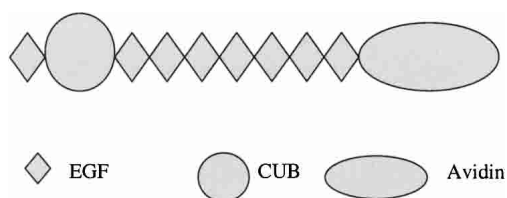


Figure 1. Domain architecture of fibropellin. All three fibropellin proteins share the same domain architecture differing only in the number of EGF repeats.

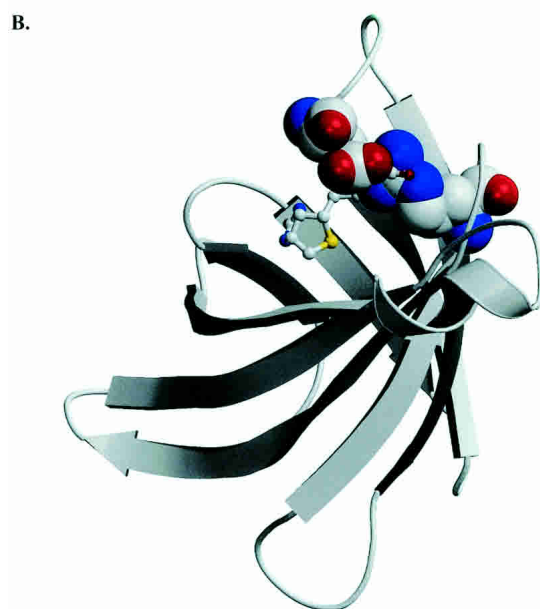
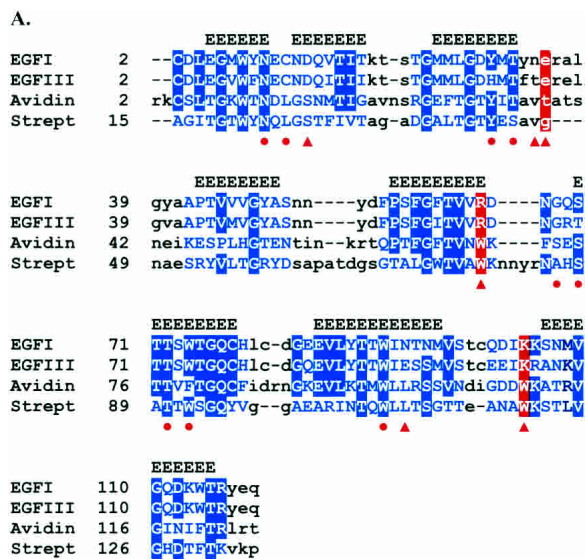


Figure 2. (A) Structure-guided multiple sequence alignment of the fibropellins (SpEGFI and SpEGFIII) avidin-like domain, avidin, and streptavidin. The *top* line of each aligned block relates secondary structure information (E, β -strand). The *bottom* line identifies those residues that are within 4.5 Å to the bound biotin in both streptavidin and avidin (●, conserved with the fibropellins; ▲, nonconserved). Only the backbone of the chain is proximate to biotin in the position including V37 of avidin. (B) Structural model for the avidin-like domain of fibropellin. The binding site is blocked by glutamic acid 35 (*left*) and arginine 65 (*right*), shown space-filled. Biotin is shown in “ball-and-stick” mode. The figure was prepared using Molscript (Kraulis 1991).

binding (Fig. 2A). Arg65 and Glu35 sit at the lip of the binding site and the salt bridge could block the binding site for biotin (Fig. 2B).

The absence of another key Trp residue in edin further suggests its lack of biotin binding function. Avidin and

Table 1. Interaction energies

	Monomer/monomer interface		Dimer/dimer interface	
	ΔE_{elec}	ΔG_{desolv}	ΔE_{elec}	ΔG_{desolv}
Avidin	-14.3	-9.6	-1.4	-12.1
Avidin W110K	-13.7	-9.6	-2.3	-0.9
Edin	0.8	-37.2	-14.4 (-13.9)	8.5 (6.3)

All energies are in kcal/mol. Electrostatic interaction energies (ΔE_{elec}) were computed using CHARMM (Brooks et al. 1983), with a dielectric constant $4r$ (where r indicates the distance between two atoms). We use the Atomic Contact Energy (ACE) (Zhang et al. 1997) to estimate desolvation. The total desolvation score of a protein complex (ΔG_{desolv}) is simply the sum of the ACE scores of atom pairs within a distance cutoff of 6 Å. The energy values in parentheses for edin were computed after side chain conformational searches using SCWRL (Dunbrack and Cohen 1997). We have extensively demonstrated the accuracies of these two energy terms in the context of protein docking (Chen and Weng 2002, 2003; Chen et al. 2003a,b,c).

streptavidin benefit from tetramerization as each monomer aids in the biotin binding of another monomer by way of a surface loop containing a Trp (number 110 in avidin and 120 in streptavidin) which folds over the trapped biotin. Surprisingly, in edin this Trp is replaced by a Lys. The Lys points to a nonpolar region of biotin, thus making the burial of its charge extremely unfavorable. Laitinen et al. (1999) used this Lys substitution in edin as the conceptual model for mutating avidin and streptavidin. With this mutation, the biotin binding affinities of avidin and streptavidin decreased by seven orders of magnitude, and interestingly, both mutant proteins form stable dimers in solution (Laitinen et al. 1999).

A molecular explanation for the oligomerization of edin can also be gleaned from the homology model (Guex and Peitsch 1997) using the crystal structure of avidin (Nardone et al. 1998) as the template, where the tetramer is a dimer of dimers. Energy calculations (Table 1) indicate that edin's monomer/monomer interface is predominately stabilized by desolvation energy, compared to that of avidin, which has both electrostatic and desolvation contributions. The overall interaction energy (sum of electrostatic and desolvation energies) for edin is considerably stronger than for avidin. The interaction between the dimers assumes different mechanisms between edin and avidin. Table 1 indicates that for avidin, there is almost no electrostatic contribution while desolvation is strong (-12.1 kcal/mol). In agreement with the findings by Laitinen et al. (1999), this strong desolvation vanishes entirely upon the mutation of Trp110 to Lys (-0.9 kcal/mol), while electrostatic is almost unaffected. As a result, avidin with the Trp110 → Lys mutation is a dimer, instead of a tetramer. The dimer-dimer interaction for edin has unfavorable desolvation (8.5 kcal/mol). To partially account for the inaccuracy of the homology model, we used a side chain optimization algorithm SCWRL (Dunbrack and Cohen 1997) to search the side chain conformations in the

dimer-dimer interface. The desolvation energy improved slightly to 6.3 kcal/mol. This is not surprising given Lys104 in edin (equivalent of Trp110 in avidin). Unlike avidin, however, the electrostatic is highly favorable at the dimer interface of edin (-14.4 kcal/mol). This indicates that edin is a tetramer, with charged residues forming a network of salt bridges that strengthens the tetramerization.

Experimental analyses indicate that edin is a tetramer that does not bind biotin

Construction, expression, and purification of edin

Beginning with a cDNA clone of the SpEGFIII gene, we subcloned a DNA fragment encoding edin, the avidin-like domain of the fibropellin protein that consists of residues V442 to A570 (Bisgrove and Raff 1993) into an expression vector pET3a. Edin was overexpressed in *Escherichia coli* BL21(DE3)(pLysS) using the T7 expression system (Studier et al. 1990), in the form of insoluble inclusion bodies (Fig. 3). The protein was folded in a buffer that contains L-arginine, and reduced and oxidized glutathione. Edin was purified to homogeneity by gel filtration chromatography (Fig. 4), with a yield of approximately 20 mg/L of culture. When edin was folded in the presence of 10 μM biotin, the yield increased by 30%. Purified edin remains stable in 20 mM TrisHCl (pH 8.0), 50 mM NaCl after storage at 4°C for over 1 mo.

Edin is a folded homotetramer

Circular dichroism (CD) spectra were used to provide information of edin's conformation in terms of secondary

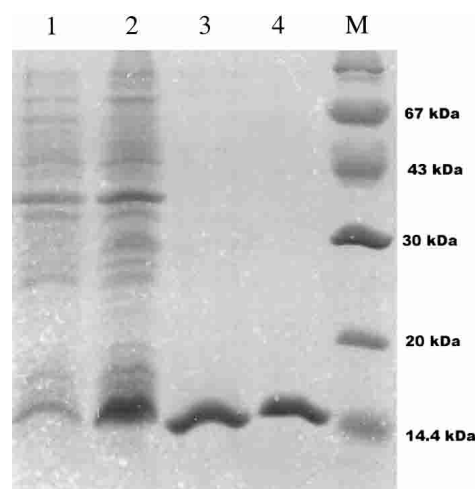


Figure 3. Expression and purification of edin in *E. coli* carrying the pET3a-edin expression vector. Samples boiled in the presence of β -mercaptoethanol were loaded onto a 12% SDS-PAGE. Lane 1, total cell proteins of BL21(DE3)(pLysS)(pET3a-edin), uninduced; lane 2, total cell proteins of BL21(DE3)(pLysS)(pET3a-edin), induced with 0.4 mM IPTG; lane 3, purified inclusion bodies of edin; lane 4, purified edin, lane M, molecular mass standard proteins.

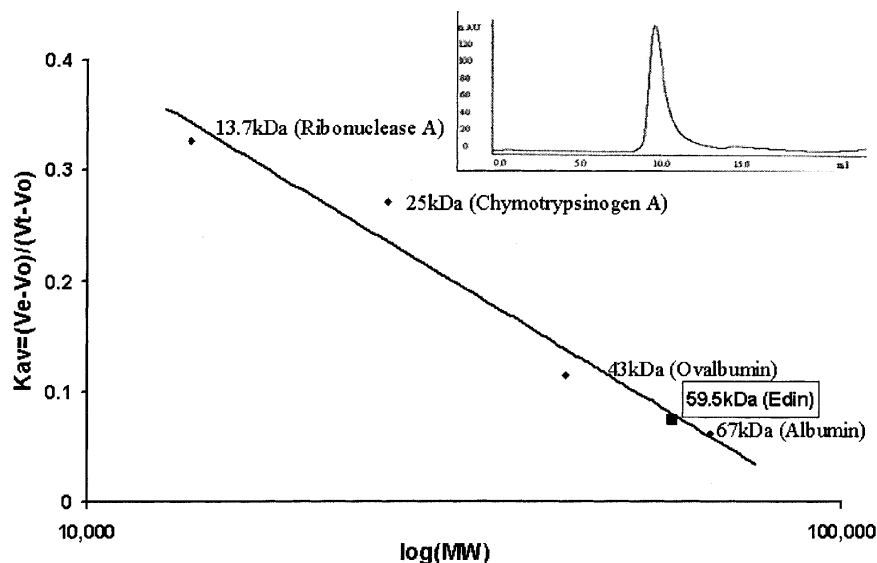


Figure 4. Edin is a tetramer. The molecular weight of edin was determined by FPLC gel filtration using a Pharmacia Superdex 75 HR 10/30 column. The column was calibrated with a gel filtration LMW calibration kit (Pharmacia). Edin shows a molecular weight of 59.5 kDa, indicating that edin is a tetramer. The *inset* shows that edin was observed as a well-defined peak.

structures (Fig. 5). Edin folded with and without biotin have almost identical CD spectra. The CD spectra showed that edin had a minimum at 211 nm ($-6344^{\circ}\text{cm}^2\text{dmol}^{-1}$), and crossover at about 199 nm and 223 nm, indicative of a β -sheet structure (Fig. 5). The positive peak at about 191 nm and a negative peak at 211 nm also indicate a high degree of β -sheet content. These results demonstrate that edin is folded with a predominantly β -sheet structure, in agreement with the prediction of the secondary structure in Figure 2B and the homology modeling.

Applying FPLC gel filtration column, edin appeared as a well-defined peak with an apparent molecular weight of about 59.5 kDa (Fig. 4). SDS-PAGE analysis showed that purified edin had a molecular weight of about 14.6 kDa (Fig. 3). Purified edin was also verified by MALDI-TOF-

MS. The mass of edin was 14,609 Da by singly charged molecular ion giving an error of -4 Da relative to the theoretical value of 14,613 Da. These results clearly indicate that edin is a homotetramer.

Edin does not bind biotin

Our computation analysis predicted that edin is not able to bind biotin. Experimentally, two methods were employed to test edin's biotin binding capability. In the first method, Ultrafree-MC centrifugal filtration units with a 10 kDa cut-off were used to separate the free ^3H -biotin from the ^3H -biotin possibly complexed with edin. Approximately, 680 pmol of edin subunits was incubate with 740 pmol or 74 pmol of D-[8,9- ^3H]biotin for overnight at 4°C . No difference of radioactivity was observed between the starting mixture and the filtrate fraction, indicating that edin does not bind biotin. A titration technique was further used, in which a fixed ^3H -biotin concentration of $0.37\ \mu\text{M}$ or $0.74\ \mu\text{M}$ were mixed with edin at different concentrations ranging from 0.1 – $100\ \mu\text{g}$ ($68\ \text{nM}$ – $68\ \mu\text{M}$). The mixtures were incubated for overnight at 4°C . No biotin binding was observed with different edin concentrations. This confirms that edin does not bind biotin. In the second method, Dispo-Equilibrium Biodialyzers (10 kDa cutoff) were used. Approximately 340 pmol of subunits of edin in $50\ \mu\text{L}$ of PBS (pH 7.4) in one chamber was dialyzed against another chamber with about 370 pmol or 37 pmol of D-[8,9- ^3H]biotin in $50\ \mu\text{L}$ of PBS. The two chambers reached equilibrium after incubation for 24 h at 4°C . The samples from both chambers were measured for radioactivity with liquid scintillation counting. No difference of radioactivity between

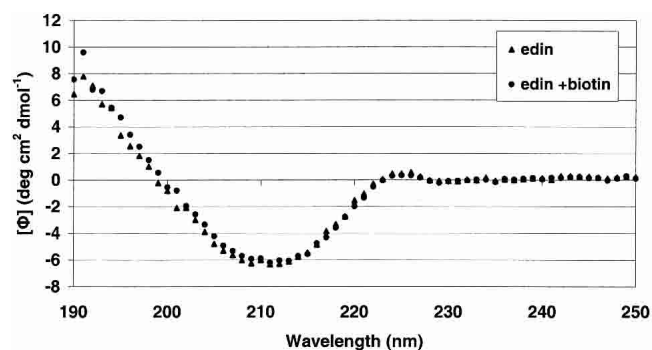


Figure 5. Edin is a folded β -sheet protein. Circular dichroism (CD) spectra of edin (with and without biotin) show that edin has a high content of β -sheet secondary structure.

two chambers was observed. These results also indicate that edin does not bind biotin at the conditions tested.

A model of the fibropellin matrix in apical lamina

Although the present work is limited to the edin domain, it is reasonable to extrapolate that the edin causes the entire fibropellin protein to tetramerize. Several lines of arguments support this notion. First, we have shown by both computational and experimental analyses that edin forms a stable tetramer. In addition, the monomer–monomer interaction energy is much stronger than in avidin (Table 1). Second, extracellular proteins are typically multimeric (Engel 1991). An example is the family of extracellular matrix proteins that might form trimeric or hexameric molecules. Tenascins contain an oligomerization region made of heptad repeats that allow the formation of trimeric molecules that are stabilized by disulfide bonds (Engel and Kammerer 2000). Third, the location of edin as the last domain of a large multidomain protein also suggests a role as an oligomerization interface.

Since the structures of homologs of the fibropellin components (CUB, EGF, and avidin) have been solved, a rough model for the entire protein can be proposed. We suggest a rod-like structure with edin at one end and the CUB domain at the other. This configuration assumes the structural independence of the composing domains. However, the functioning of the EGF domains as spacers has been long recognized (Appella et al. 1988; Engel 1989; Davis 1990) and is exemplified in such extracellular proteins as laminin, thrombospondin, and tenascin (Engel and Kammerer 2000).

A second oligomerization property of the fibropellin may occur through the dimerization of CUB domains. A crystal structure of a dimer of two distinct CUB domains, PSP-I and PSP-II, is available (Romero et al. 1997). Furthermore, the CUB domain's role in mediating a complex has been directly observed (Busby and Ingham 1990).

Given these two oligomerization properties, the tetramerization of edin and the dimerization of the CUB domain, it is immediately clear how a matrix of fibropellins might emerge. A grid-like network can form where intersections are composed of an avidin-like tetramer and edges are formed by a CUB dimer (Fig. 6). Such a network can become three dimensional when several layers are superimposed and some interprotein links connect between the layers at the expense of a perfect grid.

Concluding remarks

The present work represents the first study of the molecular properties of the avidin-like domain of the sea urchin fibropellin protein. Edin has been indirectly studied by Laitinen et al. (1999), who used sequence comparisons between streptavidin/avidin and edin to direct a Trp → Lys mutation

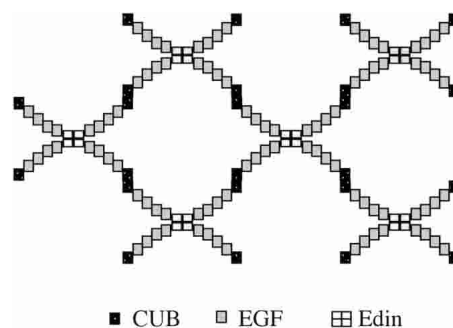


Figure 6. A model for the molecular structure of the apical lamina.

in avidin and streptavidin. The mutation caused a significant reduction in the biotin binding capability and resulted in avidin and streptavidin dimers as opposed to tetramers. Based on these results the authors predicted that edin is dimeric and it functions in fibropellin solely for its dimerization property. This notion is partly consistent with the results presented here. We have shown experimentally that edin possesses no ability to bind biotin and a computational modeling of the structure offers a molecular explanation.

Contrary to the prediction by Laitinen and colleagues, however, we have shown that edin forms a tetramer. Interestingly, both the monomer–monomer and the dimer–dimer interactions follow distinct mechanisms between edin and avidin. Several lines of evidence support the notion that the entire fibropellin protein tetramerizes. Thus, it is reasonably concluded that the avidin domain has been “exapted” (Gould and Vrba 1982) from a small (~120 amino acid) single-domain protein that binds biotin to an oligomerizing role as a domain in a large multidomain protein. Such a case highlights the context-dependence of the function of a domain in terms of the entire protein. Fibropellin's involvement in the formation of the fibrous meshwork can be further characterized if we could produce a complete fibropellin protein.

Materials and methods

Construction of expression vector

The SpEGF III cDNA clone (Bisgrove et al. 1991; Bisgrove and Raff 1993) in pBluescript (a kind gift from Dr. Rudolf Raff) was used as a starting material. The DNA encoding for the avidin-like domain, amino acid 420–570, was amplified with PCR. The PCR product was digested with NdeI and BamHI, purified, and cloned into the NdeI/BamHI-treated pET3a. The cloned DNA was sequenced and the resultant vector, pET3a-edin, was transformed into an *E. coli* strain BL21(DE3)(pLysS).

Expression of edin in E. coli

The *E. coli* strain BL21(DE3)(pLysS) carrying an expression vector, pET3a-edin, was grown at 37°C with shaking in Luria-Bertani

medium supplemented with 100 $\mu\text{g}/\text{mL}$ ampicillin and 25 $\mu\text{g}/\text{mL}$ chloramphenicol. When the OD_{600} of the culture reached 0.5–0.6, isopropyl b-D-thiogalactopyranoside was added to a final concentration of 0.4 mM to induce the expression of the T7 RNA polymerase gene placed under the control of the *lacUV5* promoter. Four hours after IPTG induction and incubation at 37°C with shaking, cells were harvested by centrifugation at 8000g for 10 min.

Isolation and purification of edin

Cell pellets were resuspended in 50 mM Tris-HCl (pH 8.0), 25% (W/V) sucrose, 1 mM EDTA, 0.1% (w/v) NaAzide, and 10 mM DTT. After the cells were treated with three cycles of freeze-thaw, the pellets were sonicated and centrifuged three times in 50 mM Tris-HCl (pH 8.0), 0.5% Triton-X100, 100 mM NaCl, 1 mM EDTA, 0.1% NaAzide, and 1 mM DTT. The same buffer without Triton-X100 was used for the final sonication and wash of the pellet. The insoluble inclusion bodies containing edin were solubilized in 8 M urea, 10 mM EDTA, and 0.1 mM DTT. The solubilized protein was mixed 1:1 with 3 M Guanidine HCl (pH 4.2), 10 mM NaAcetate, and 10 mM EDTA. The mixture was forcefully injected through a 26-gauge needle into a stirring folding buffer containing 400 mM L-Arginine, 100 mM TrisHCl (pH 8.0), and 2 mM EDTA, with (10 μM) or without biotin. After incubation overnight at 4°C with stirring, the folding solution was centrifuged to remove insoluble material and concentrated with Millipore's Centriprep YM-10. Edin was purified by gel filtration chromatography with AKTA FPLC (Pharmacia).

Mass spectrometry

A sample of purified edin was analyzed by matrix-assisted laser desorption ionization-time of flight-mass spectrometry (MALDI-TOF-MS) using a Shimadzu/Kratos Axima CFR in linear mode. Best results were obtained with a matrix of α -cyano-4-hydroxycinnamic acid and a protein concentration of roughly 10 pmol/ μL . Lysozyme was used as an external calibrant. All masses were averages determined for smoothed data (20 point) (Savitsky-Golay) for duplicate acquisitions.

Circular dichroic spectral measurements

Spectra of edin folded with or without the presence of biotin were recorded (Aviv 62DS spectrometer, 1-mm path length for far-UV spectra) at 4°C, using 15 μM of edin in 10 mM PBS (pH 7.4). Data were averaged over 20 sec and collected using 1-nm intervals. CD spectra of edin were corrected by subtracting the buffer spectrum obtained under identical conditions.

Biotin binding ability of edin

Approximately 10 μg of purified edin (about 680 pmol of subunits) was mixed in PBS (pH 7.4), with 740 pmol and 74 pmol of D-[8,9- ^3H]biotin (27 Ci/mmol) (Amersham) in a total volume of 100 μL . After incubation at 4°C for overnight, the mixtures were loaded into Ultrafree-MC centrifugal filtration units (10 kDa cutoff) (Millipore), and centrifuged at 4°C at 5000g for 15 min. The amount of biotin in the starting mixture and the filtrate fraction were measured by liquid scintillation counting. The difference of radioactivity between the total biotin and the free biotin in the filtrate fraction was used for determining the biotin binding ability of edin. A titration assay was performed to determine the biotin

binding ability of edin, in which a fixed amounts of D-[8,9- ^3H]biotin (27 Ci/mmol) (Amersham) with concentrations of 370 nM and 740 nM were mixed with different amounts of purified edin with concentrations ranging from 68 nM to 68 μM of subunits in PBS (pH 7.4). After incubation at 4°C for overnight, the mixtures were loaded into Ultrafree-MC centrifugal filtration units (10 kDa cutoff) (Millipore) and centrifuged at 4°C at 5000g for 15 min. The amount of biotin in the starting mixture and the filtrate fraction were measured by liquid scintillation counting. The difference of radioactivity between the total biotin and the free biotin in the filtrate fraction was used for determining the biotin binding ability of edin.

Dispo-Equilibrium Biodialyzers (10 kDa cutoff) from Harvard Bioscience were also used for measuring biotin binding ability of edin. Approximately 5 μg of edin (about 340 pmol of subunits) in 50 μL of PBS (pH 7.4) was placed to one chamber of the dialyzer according to the manufacturer's instruction. About 370 pmol or 37 pmol of D-[8,9- ^3H]biotin (27 Ci/mmol) (Amersham) in 50 μL of PBS was applied to another chamber. The dialyzers were incubated for 24 h at 4°C to reach equilibrium. The samples from both chambers were measured by liquid scintillation counting. The difference of radioactivity between two chambers was used for determining the biotin binding ability of edin.

Molecular weight of edin

The molecular weight of edin was determined by gel filtration using Superdex 75 HR 10/13 with Pharmacia's AKTA FPLC. The column was calibrated with a gel filtration LMW calibration kit (Pharmacia).

SDS-PAGE analysis

SDS-PAGE analysis was performed with 12% polyacrylamide gels. The samples were boiled for 5 min in the presence of 10 mM β -mercaptoethanol. Proteins were stained with Coomassie brilliant blue.

Acknowledgments

We thank Dr. George Tarr for doing the MALDI-TOF-MS work, Li Li for helping with energy calculations, and Kevin Weihe for helping with SCWRL calculations. We thank Rudolf Raff for kindly supplying us with the SpEGF III cDNA clone. We thank Dr. Jonathan Lee for the use of circular dichroism spectrometer. We also thank Gonzalo Hernandez for his assistance with the circular dichroism spectrometer. We also thank Dr. Meir Wilchek for helpful discussions and encouragement. I.Y. is a Koshland Scholar. This work was partially funded by NSF Grants DBI-0078194, DBI-0133834, and DBI-0116574.

References

- Appella, E., Weber, I.T., and Blasi, F. 1988. Structure and function of epidermal growth factor-like regions in proteins. *FEBS Lett.* **231**: 1–4.
- Bisgrove, B.W. and Raff, R.A. 1993. The SpEGF III gene encodes a member of the fibropellins: EGF repeat-containing proteins that form the apical lamina of the sea urchin embryo. *Dev. Biol.* **157**: 526–538.
- Bisgrove, B.W., Andrews, M.E., and Raff, R.A. 1991. Fibropellins, products of an EGF repeat-containing gene, form a unique extracellular matrix structure that surrounds the sea urchin embryo. *Dev. Biol.* **146**: 89–99.
- Bork, P. and Beckmann, G. 1993. The CUB domain. A widespread module in developmentally regulated proteins. *J. Mol. Biol.* **231**: 539–545.

- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**: 187–217.
- Burke, R.D., Lail, M., and Nakajima, Y. 1998. The apical lamina and its role in cell adhesion in sea urchin embryos. *Cell Adhes. Commun.* **5**: 97–108.
- Busby, T.F. and Ingham, K.C. 1990. NH₂-terminal calcium-binding domain of human complement C1s- mediates the interaction of C1r- with C1q. *Biochemistry* **29**: 4613–4618.
- Cameron, R.A. and Holland, N.D. 1985. Demonstration of the granular layer and the fate of the hyaline layer during the development of a sea urchin (*Lytechinus variegatus*). *Cell Tissue Res.* **239**: 455–458.
- Chen, R. and Weng, Z. 2002. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins* **47**: 281–294.
- . 2003. A novel shape complementarity scoring function for protein-protein docking. *Proteins* **51**: 397–408.
- Chen, R., Li, L., and Weng, Z. 2003a. ZDOCK: An initial-stage protein-docking algorithm. *Proteins* **52**: 80–87.
- Chen, R., Mintseris, J., Janin, J., and Weng, Z. 2003b. A protein-protein docking benchmark. *Proteins* **52**: 88–91.
- Chen, R., Tong, W., Mintseris, J., Li, L., and Weng, Z. 2003c. ZDOCK predictions for the CAPRI challenge. *Proteins* **52**: 68–73.
- Davis, C.G. 1990. The many faces of epidermal growth factor repeats. *New Biol.* **2**: 410–419.
- Dunbrack Jr., R.L. and Cohen, F.E. 1997. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci.* **6**: 1661–1681.
- Engel, J. 1989. EGF-like domains in extracellular matrix proteins: Localized signals for growth and differentiation? *FEBS Lett.* **251**: 1–7.
- . 1991. Domains in proteins and proteoglycans of the extracellular matrix with functions in assembly and cellular activities. *Int. J. Biol. Macromol.* **13**: 147–151.
- Engel, J. and Kammerer, R.A. 2000. What are oligomerization domains good for? *Matrix Biol.* **19**: 283–288.
- Gould, S.J. and Vrba, E.S. 1982. Exaptation: A missing term in the science of form. *Paleobiology* **8**: 4–15.
- Green, N.M. 1990. Avidin and streptavidin. *Methods Enzymol.* **184**: 51–67.
- Guex, N. and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723.
- Hall, H.G. and Vacquier, V.D. 1982. The apical lamina of the sea urchin embryo: Major glycoproteins associated with the hyaline layer. *Dev. Biol.* **89**: 168–178.
- Hunt, L.T. and Barker, W.C. 1989. Avidin-like domain in an epidermal growth factor homolog from a sea urchin. *FASEB J.* **3**: 1760–1764.
- Hursh, D.A., Andrews, M.E., and Raff, R.A. 1987. A sea urchin gene encodes a polypeptide homologous to epidermal growth factor. *Science* **237**: 1487–1490.
- Kraulis, P.J. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**: 946–950.
- Laitinen, O.H., Airene, K.J., Marttila, A.T., Kulik, T., Porkka, E., Bayer, E.A., Wilchek, M., and Kulomaa, M.S. 1999. Mutation of a critical tryptophan to lysine in avidin or streptavidin may explain why sea urchin fibropellin adopts an avidin-like domain. *FEBS Lett.* **461**: 52–58.
- Nardone, E., Rosano, C., Santambrogio, P., Curnis, F., Corti, A., Magni, F., Siccardi, A.G., Paganelli, G., Losso, R., Aprea, B., et al. 1998. Biochemical characterization and crystal structure of a recombinant hen avidin and its acidic mutant expressed in *Escherichia coli*. *Eur. J. Biochem.* **256**: 453–460.
- Peitsch, M.C., Wells, T.N., Stampf, D.R., and Sussman, J.L. 1995. The Swiss-3DImage collection and PDB-Browser on the World-Wide Web. *Trends Biochem. Sci.* **20**: 82–84.
- Romero, A., Romao, M.J., Varela, P.F., Kolln, I., Dias, J.M., Carvalho, A.L., Sanz, L., Topfer-Petersen, E., and Calvete, J.J. 1997. The crystal structures of two spermadhesins reveal the CUB domain fold. *Nat. Struct. Biol.* **4**: 783–788.
- Spiegel, E., Howard, L., and Spiegel, M. 1989. Extracellular matrix of sea urchin and other marine invertebrate embryos. *J. Morphol.* **199**: 71–92.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**: 60–89.
- Szuskowski, J.D. and Weng, Z. 2000. Protein structure alignment using a genetic algorithm. *Proteins* **38**: 428–440.
- Wang, Y., Geer, L.Y., Chappey, C., Kans, J.A., and Bryant, S.H. 2000. Cn3D: Sequence and structure views for Entrez. *Trends Biochem. Sci.* **25**: 300–302.
- Zhang, C., Vasmatzis, G., Cornette, J.L., and DeLisi, C. 1997. Determination of atomic desolvation energies from the structures of crystallized proteins. *J. Mol. Biol.* **267**: 707–726.