The Relative Orientation of Gla and EGF Domains in Coagulation Factor X Is Altered by Ca^{2+} Binding to the First EGF Domain. A Combined NMR−Small Angle X-ray Scattering Study[†,‡]

Maria Sunnerhagen,[*,§,||] Glenn A. Olah,[⊥] Johan Stenflo,[#] Sture Forsén,[§] Torbjörn Drakenberg,[§] and Jill Trewhella[*,⊥]

Physical Chemistry 2, Chemical Center, Lund University, P.O. Box 124, S-221 00 Lund, Sweden, Chemical Science and Technology Division, Mail Stop G758, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, and Clinical Chemistry, Lund University, Wallenberg Laboratory, Malmö General Hospital, S-214 01 Malmö, Sweden

Received March 14, 1996; Revised Manuscript Received June 11, 1996[⊗]

ABSTRACT: Coagulation factor X is a serine protease containing three noncatalytic domains: an N-terminal γ-carboxyglutamic acid (Gla)[1] domain followed by two epidermal growth factor (EGF)-like domains. The isolated N-terminal EGF domain binds Ca^{2+} with a K_d of 10^{-3} M. When linked to the Gla domain, however, its Ca^{2+} affinity is increased 10-fold. In this paper, we present the NMR solution structure of the factor X Gla−EGF domain pair with Ca^{2+} bound to the EGF domain, as well as small angle X-ray scattering (SAXS) data on the Gla−EGF domain pair with and without Ca^{2+}. Our results show that Ca^{2+} binding to the EGF domain makes the Gla and EGF domains fold toward each other using the Ca^{2+} site as a hinge. Presumably, a similar mechanism may be responsible for alterations in the relative orientation of protein domains in many other extracellular proteins containing EGF domains with the consensus for Ca^{2+} binding. The results of the NMR and SAXS measurements reported in this paper confirm our previous result that the Gla domain is folded also in its apo state when linked to the EGF domain [Sunnerhagen, M., et al. (1995) *Nat. Struct. Biol. 2*, 504−509]. Finally, our study clearly demonstrates the powerful combination of NMR and SAXS in the study of modular proteins, since this enables reliable evaluation of both short-range (NMR) and long-range interactions (SAXS).

The epidermal growth factor domain is one of the most widespread domains in extracellular mosaic proteins (Campbell & Bork, 1993). With its three disulfide bridges linked in a characteristic manner (1−3, 2−4, 5−6) it is a stable platform for several functions. Numerous extracellular and membrane proteins contain arrays of tandemly arranged EGF[1]-like domains. Although the function of most of these domains is still not known, it appears that they are often involved in protein−protein interactions. Several of these proteins are involved in the regulation of cell growth as well as regulation of cell differentiation in early development (Engel, 1989; Carpenter & Wahl, 1990; Davis, 1990, Campbell & Bork, 1993).

A number of proteins involved in blood coagulation and anticoagulation contain EGF domains, whose functions are currently being investigated (Furie & Furie, 1988; Mann, 1990; Handford et al., 1991; Stenflo, 1991). In protein C, factor VII, factor IX, thrombomodulin, tissue plasminogen activator, and urokinase plasminogen activator, EGF domains have been shown to be involved in protein interactions (Appella et al., 1987; Kurosawa et al., 1988; Toomey et al., 1991; Hogg et al., 1992; Nishimura et al., 1993; Smith et al., 1994). Moreover, the EGF domains serve as spacers, placing the active site of the serine protease module at a distance above the biological membranes that is commensurate with biological activity (in factor X $\approx$ 70 Å) (Husten et al., 1987; Brandstetter et al., 1995; Banner et al., 1996). In factors VII, IX, and X, and presumably also in protein C, the C-terminal EGF domain has a large contact area with the serine protease domain, apparently necessary for the latter to fold to a native conformation (Padmanabhan et al., 1993; Brandstetter et al., 1995; Banner et al., 1996).

Ca^{2+} binding to EGF domains was first observed with the N-terminal EGF domain of the anticoagulant protein C (Öhlin et al., 1987, 1988) and has also been observed in coagulation factors VII, IX, and X as well as in the anticoagulant protein S [reviewed in Stenflo (1991)]. The Ca^{2+} binding sequence motif includes the consensus for hydroxylation of an aspartate or asparagine residue to hydroxyaspartate (Hya) or hydroxyasparagine (Hyn) Cys-Xxx-Asp/Asn-Xxx-Xxx-Xxx-Xxx-Tyr/Phe-Xxx-Cys (Stenflo et al., 1987, 1988) as well as an Asp/Asn-Ile/Val/Gly-Asp/Asn-Glu/Gln/Asp-Cys_1 consensus N-terminal to the first cysteine (Handford et al., 1991; Selander-Sunnerhagen et al., 1992). EGF domains with the consensus for Ca^{2+} binding occur in a number of extracellular

[†] This work was supported by grants from the Swedish Board of Technical Development and the Swedish Medical Research Council. The NMR spectrometer was purchased with grants from the Knut and Alice Wallenberg Foundation and the Swedish Council for Planning and Coordination of Research. Support for J.T. and G.A.O. was provided by National Institutes of Health Award GM40528 and the Department of Energy's Office of Health and Environmental Research project HA-02-02-03. M.S. acknowledges the Fredrika Bremer Society and Fysiografiska Sällskapet for generous travel grants.

[‡] The coordinates for the apo and 1 Ca forms have been deposited in the Brookhaven Protein Data Bank (access no. 1WHE and 1WHF, respectively).

* To whom correspondence should be addressed.

[§] Chemical Center, Lund University.

[||] Current address: Molecular Biophysics, Karolinska Institute, Doktorsringen 6A, S-171 77 Stockholm, Sweden.

[⊥] Los Alamos National Laboratory.

[#] Clinical Chemistry, Lund University.

[⊗] Abstract published in *Advance ACS Abstracts,* August 1, 1996.

[1] Abbreviations: EGF, epidermal growth factor; Gla, γ-carboxyglutamic acid; SAXS, small angle X-ray scattering; Hya, *erythro*-β-hydroxyaspartate.

S0006-2960(96)00633-2 CCC: $12.00 © 1996 American Chemical Society

proteins with diverse functions. These include not only coagulation factors but also receptor proteins important for the embryonal development of cells such as the *Drosophila* Notch and Delta proteins and the human Notch homolog TAN-1, fibrillin, the low-density lipoprotein receptor, transforming growth factor $\beta-1$ binding protein, and many more (Rees et al., 1988; Engel et al., 1989; Stenflo, 1991; Selander-Sunnerhagen et al., 1992; Campbell & Bork, 1993).

The solution structure of a Ca^{2+} site in the N-terminal EGF domain of factor X has been determined by NMR (Selander-Sunnerhagen et al., 1992). The isolated EGF domain binds Ca^{2+} fairly weakly, with a K_d of 10^{-3} M (Persson et al., 1989). The Ca^{2+} site only has five ligands from the protein, whereas most Ca^{2+}-binding proteins bind Ca^{2+} with six to eight ligands, seven being the favored coordination number (Strynadka & James, 1989, 1991). As the site is open to the solvent on one side, remaining coordinating positions were assumed to be occupied by water molecules. In the intact protein, however, we suggested that other protein domains might provide the lacking ligands (Selander-Sunnerhagen et al., 1992). This was also indicated by an observed 10-fold increase in Ca^{2+} affinity of the factor X EGF domain when linked to the preceding Gla domain (Persson et al., 1991; Valcarce et al., 1993). Only residues 29−44 from the Gla domain are needed for full Ca^{2+} affinity in the EGF domain (Valcarce et al., 1993). Recently, the homologous N-terminal EGF domain in coagulation factor IX was shown to have the same Ca^{2+} ligands as the factor X EGF domain (Rao et al., 1995). To complete the Ca^{2+} ligand sphere, an additional ligand is provided by a neighboring factor IX EGF domain in the crystal lattice (Rao et al., 1995), which gives further credence to our hypothesis that remaining coordinating positions might be filled by residues from other protein domains.

Not only in factor X but also in coagulation factors VII and IX and protein C are the Ca^{2+}-binding EGF domains all preceded by a Gla domain, which contains 10−12 γ-carboxyglutamic acid (Gla) residues. The Gla domains bind Ca^{2+} with low affinity and high cooperativity, thus exposing a membrane binding site which is a prerequisite for binding of Gla-containing clotting factors to biological membranes (Mann et al., 1988, 1990). Several studies have indicated that isolated Gla domains obtained by limited proteolysis and thus cleaved in the aromatic stack region do not bind Ca^{2+} with the same affinity as in the intact protein (Schwalbe et al., 1989; Persson et al., 1991). However, a synthesized Gla domain from protein C with the aromatic stack intact binds Ca^{2+} with native affinity (Zhang et al., 1992; Jacobs et al., 1994). Also, in fragments of protein C, factor VII, factor IX, and factor X containing the Gla and EGF domains, the Ca^{2+} affinity of the Gla domain is virtually the same as in the intact protein (Öhlin et al., 1990; Persson et al., 1991; Astermark et al., 1991; Persson & Petersen, 1995). Moreover, the EGF domain is less sensitive to proteolysis and reduction in the presence of the Gla domain (Valcarce et al., 1994). For coagulation factor IX, Ca^{2+}-dependent interactions between isolated Gla and EGF domains have been observed by size-exclusion chromatography (Medved et al., 1994). This suggests that there is an interaction between Gla and EGF domains, possibly stabilizing the structure of both of them.

We have determined the solution structure of the Gla−EGF domain pair of factor X in the absence of Ca^{2+} by NMR (Sunnerhagen et al., 1995). In contrast to previous assumptions, we found that the Gla domain when linked to the EGF domain is not disordered in the absence of Ca^{2+} but has similar secondary structure and global fold as in the presence of Ca^{2+} although the structure is more dynamic. By comparison with the crystal structure of the Ca^{2+} -loaded form we could identify structural changes in the Gla domain essential for membrane binding of coagulation proteins (Sunnerhagen et al., 1995). The isolated Gla domain from factor IX in the absence of Ca^{2+} is much less ordered as observed by NMR under similar conditions (Freedman et al., 1995). In the crystal structure of the apo form of intact coagulation factor IX, only the C-terminal helix of the Gla domain is observed, probably due to its flexible orientation relative to the EGF domain (Brandstetter et al., 1995).

In this work, we present the solution structure of the Gla−EGF domain pair of factor X with Ca^{2+} bound to the EGF domain only. This was possible since the Ca^{2+} affinity of the EGF domain in factor X is approximately 10 times higher than in the Gla domain (Valcarce et al., 1993). On Ca^{2+} binding to the EGF domain, an internal reorientation of the two domains is observed by both NMR and small-angle X-ray scattering (SAXS). In view of our present results, it is conceivable that Ca^{2+} binding to the N-terminal EGF domain in factor X defines a biologically active domain orientation. We suggest that this ability to confer alterations in interdomain orientations may be a general property of EGF domains with the consensus for Ca^{2+} binding.

METHODS

^{1}H NMR Assignment of Apo and the 1 Ca^{2+} Gla−EGF Domain Pair. A protein fragment comprising the Gla and EGF domains of bovine factor X, Gla−EGF (residues 1−86), was isolated and purified as described elsewhere (Persson et al., 1991). Ca^{2+} levels were below 2 mol % as determined by atomic absorption measurements (data not shown). The material used in this study was isolated from the native protein as collected from bovine blood and was therefore isotopically unlabeled but contained all native postribosomal modifications (Glu → Gla and Asp → Hya). A 1 mM sample at 17 °C and pH 7.0 was used in the assignment both at zero Ca^{2+} concentration (apo Gla−EGF) and with 1 equiv of Ca^{2+} added (1 Ca^{2+} Gla−EGF). One equivalent of $CaCl_2$ added gives 80% saturation of the EGF domain Ca^{2+} site, whereas adding more than 2 equiv of $CaCl_2$ causes aggregation (Valcarce et al., 1993). Previous titrations of Ca^{2+} into the Gla−EGF domain pair gives a continuous change of chemical shift, thus indicating that the Ca^{2+} ion is in fast exchange (Valcarce et al., 1993), and in accordance with this, only one set of resonances was observed. The experiments were run at pH 7.0 as severe aggregation occurred at pH 6.0 and below. The temperature 17 °C was chosen as a compromise between the positive effect of reduced exhange rates of the amide protons and the negative effect of line broadening at lower temperatures. At 28 °C and pH 7.0, most of the amide protons in the Gla domain exchanged so rapidly with the bulk water that assignment of ^{1}H resonances was not feasible. No additional salt or buffer was added. Spectra were obtained at 500.13 MHz on a General Electric Omega 500 spectrometer. The ^{1}H NMR assignment was made using well-established procedures and experiments (Chazin et al., 1988) as previously done for the isolated EGF domain (Selander et al., 1990; Selander-Sunnerhagen et al., 1992). Spin system assignments were obtained from COSY, R-COSY, and TOCSY (40 and 80 ms) spectra in H_2O and

D_2O. Double-quantum experiments were used to examine degeneracy of β-protons (D_2O) and glycine α-protons (H_2O). The sequential assignment was obtained from NOESY spectra (100 and 250 ms) in H_2O and D_2O. Additional COSY, R-COSY, and NOESY (100 ms) experiments were performed at 12 and 28 °C to help the assignment in crowded regions of the spectrum.

NMR Structure Determination of 1 Ca^{2+} Gla−EGF. Distance constraints for 1 Ca^{2+} Gla−EGF were obtained from NOESY spectra at 100 ms and 250 ms and 17 °C in H_2O and D_2O. In previous experiments using build-up curves for the apo form we found that the NOESY cross-peak intensities increased linearly up to a mixing time of 120 ms for all peaks except for the strongest ones, where a slightly reduced intensity was observed. Peaks with a nonlinear buildup arising from spin diffusion occurred first at 250 ms. Therefore, integration of peaks was performed in the 100 ms NOESY spectrum, and the volumes were converted to distances using the isolated spin pair approximation. The 250 ms spectrum was evaluated only for peaks not observed in the 100 ms spectrum, which were included with conservative upper distance limits of 5 Å. No peaks were included as constraints if they were sufficiently strong that they would have been observed in the 100 ms spectrum if the buildup was linear. Upper and lower bounds were set 10% above and 30% below the NOE-derived distance, and error margins were added corresponding to the scaling of equivalent and degenerate proton resonances. No pseudoatom corrections were added as distance constraints were incorporated using r^{-6} averaging for nonstereospecifically assigned protons including aromatic ring protons (Clore et al., 1986). Alternative assignments for NOE's due to spectral overlap were included using r^{-6} averaging. Backbone dihedral angle constraint intervals were [−160°, −80°] for $^3J_{HN-\alpha} > 8$ Hz and [−80°, −40°] for $^3J_{HN-\alpha} < 6$ Hz (Johansson et al., 1993). Stereospecific assignments in the EGF domain were obtained from the structure of the isolated Ca^{2+}-loaded EGF domain in regions where there were no shift differences (Selander-Sunnerhagen et al., 1992). As in the apo form, no stereospecific assignments in the Gla domain were made as NOE data combined with coupling constants did not single out any unique conformations (Sunnerhagen et al., 1995). Structure calculations were performed with X-PLOR (Brünger et al., 1992) as described for the apo Gla−EGF domain pair (Sunnerhagen et al., 1995).

SAXS Data Collection and Reduction. X-ray scattering data were collected from Gla−EGF domain pair samples without Ca^{2+} and with 1 equiv of Ca^{2+} using the X-ray scattering station described in Heidorn and Trewhella (1988). All measurements were reduced to $I(Q)$ versus Q, where $I(Q)$ is the scattered X-ray intensity and Q is the amplitude of the scattering vector. Q is equal to $4\pi \sin\theta/\lambda$, where 2θ is the scattering angle and λ is the wavelength of the incident and scattered X-rays (1.542 Å for the Cu Kα used). A sample-to-detector distance of 64 cm was used, giving a measured Q-range of 0.014−0.29 Å.

Guinier (1939) and the indirect Fourier transform [or $P(r)$] (Moore, 1980) analyses were used to calculate radius of gyration (R_g), radius of gyration of cross section (R_c), forward scatter [$I(0)$], and distance distribution function [$P(r)$], from $I(Q)$ vs Q [described in more detail in Heidorn et al. (1989), and Sosnick et al. (1991)]. $P(r)$ is the frequency of distances connecting small-volume elements weighted by the product of the scattering densities of each pair of elements within the entire volume of the scattering particle and hence goes to zero at the maximum dimension of the particle, d_{max}. R_g of a particle is the root-mean-square distance of all elemental volumes from the center of mass of the particle (weighted by their scattering densities). Similarly, R_c is the root-mean-square distance of all area elements from the center of mass of the cross-sectional area of the particle.

SAXS measurements were made at a constant temperature (21 °C) for a series of protein concentrations (14.2, 18.5, 22.7, and 27.3 mg/mL for apo Gla−EGF; 16.5, 21.6, and 26.3 mg/mL for 1 Ca^{2+} Gla−EGF). For the 1 Ca^{2+} samples, the equivalent concentration of Ca^{2+} was kept constant. Since the dilution range was narrow, the difference in saturation was small enough to be ignored. To ensure the same pH for all dilutions, a 10 mM Tris buffer (pH 7.5) was used. From these concentration series the effects of interparticle interference in the scattering profiles [$I(Q)$ vs Q], arising from repulsive forces between the negatively charged protein molecules, could be evaluated. The dominant effect of these repulsive forces is suppression of the scattering at very small angles (Chen & Bendedouch, 1986). This suppression diminishes with decreasing concentration and is usually eliminated by linear extrapolation of the scattering data to zero concentration (Pilz, 1982). For the highly charged Gla−EGF domain pair these effects are so strong that simple linear extrapolation from the concentration range that could be measured does not completely eliminate the suppression of the low-Q data. If the particles are assumed to be distributed in a disordered liquid state, then the effect of the interparticle interference arises from a broad Bragg peak resulting from the interparticle first-neighbor shell which is convoluted with the particle structure in the scattering profiles (Chen & Bendedouch, 1986). We estimated from the known particle density in our samples the mean position of the first-neighbor shell to be >100 Å for the concentrations used in these experiments. This gives rise to a diffraction peak at a Q value <0.062 Å^{-1} which could systematically affect any structural parameters derived using scattering data in this region. For larger Q values the interparticle scattering structure factor quickly converges to unity, and the total scattering becomes dominated by the particle structure factor. Thus, interparticle interference was minimized by linear extrapolation to zero protein concentration, and R_g values and $P(r)$ functions for Gla−EGF were calculated using Q values greater than those values estimated to be significantly impacted by these effects. No change in R_g or other structural parameters was measured when minimum Q values between 0.058 and 0.08 Å^{-1} were used in the analysis. Therefore, minimum Q values of 0.058−0.06 were used for all analyses. For reliable estimates of d_{max} and R_g it is important to have measured data in the low-Q Guinier region [defined to be the region such that $QR_g < 1.3$, for globular particles (Guinier, 1939)]. The Guinier region for the Gla−EGF domain pair typically ends around $Q \sim 0.09$ Å^{-1}. The Guinier region for R_c calculations was typically $0.1 < Q < 0.2$ Å^{-1}, which is unaffected by interparticle interference.

Modeling of X-ray Scattering from Gla−EGF NMR Structures. Scattering profiles and $P(r)$ functions corresponding to the NMR structures of apo and 1 Ca^{2+} Gla−EGF were computed using a Monte Carlo integration method (Heidorn & Trewhella, 1988; Olah et al., 1993; Olah & Trewhella, 1994). The simulated and measured scattering curves were scaled using $I(0)$ values (calculated from $P(r)$

analyses using a Q-range of 0.06−0.28 Å^{-1}]. A goodness of fit between simulated and measured scattering data was ascertained by calculating reduced χ^2 values for the Q-range of 0.06−0.25 Å^{-1}. A comparison was also made between the $P(r)$ functions calculated from the SAXS data and from the simulated model scattering. For comparison with experiment the simulated scattering curves were convoluted with the instrument resolution function, and equivalent Q-ranges were used for $P(r)$ analyses. A comparison was also made between the R_g, R_c, and d_{max} values derived from the measured SAXS data and the modeled NMR structures.

A previously determined crystal structure of the prothrombin fragment 1 in the absence of Ca^{2+} suggests the apo form of the Gla domain is in a disordered state (Tulinsky et al., 1988). On the other hand, NMR data show it to have a secondary structure similar to that of the Ca^{2+}-bound structure (Sunnerhagen et al., 1995). We therefore simulated a disordered Gla domain (residues 1−45) by defining a random walk in three-dimensional space between adjacent C_α carbons in the peptide backbone. Adjacent C_α carbon separation was fixed at 3.53 Å, the average value determined in the NMR structures. Spheres were built around each C_α with a volume determined by the corresponding amino acid volume (Jacrot & Zaccai, 1981; Olah et al., 1993). The random walk was constrained by not allowing spheres to overlap in the Gla domain or overlap with the intact EGF domain (EGF domain was also defined by spheres positioned at each C_α carbon in its NMR structure). A total of 1300 EGF−Gla domain structures were generated in which the Gla domain was randomly disordered. From these 1300 structures the corresponding $P(r)$ and scattering functions were calculated using the Monte Carlo modeling method and the averaged $P(r)$ and scattering functions then determined. A total of 5000 random points were used to define each structure. The averaged scattering function was convoluted with the X-ray scattering instrument resolution function and then scaled and compared with the measured scattering data for the Gla−EGF domain pair without Ca^{2+}.

RESULTS

Assignment of Apo and the 1 Ca^{2+} Gla−EGF Domain Pair. NMR assignments were made at pH 7.0 and 17 °C with 0 (apo Gla−EGF) and 1 equiv of Ca^{2+} added (1 Ca^{2+} Gla−EGF). The assignment of the EGF domain was straightforward whereas the sequential assignment of the Gla domain posed a challenge due to the presence of 12 surface-exposed Gla residues often occurring pairwize and with very similar side chain shifts. The assignment lists are given as Supporting Information. To obtain the postribosomally modified Gla and Hya residues, we used native, isotopically unlabeled protein in the study.

The chemical shift effects in the Gla−EGF domain pair on Ca^{2+} binding to the EGF domain are displayed in Figure 1. It is clear that shift changes occur not only in the EGF domain but also in the Gla domain. Especially, residues 33−44 are affected by Ca^{2+} binding to the EGF domain. These residues are located in the C-terminal helix of the Gla domain, the so-called hydrophobic stack helix, and include residues 29−44 previously shown to be essential for high-affinity Ca^{2+} binding to the EGF domain (Valcarce et al., 1993). Other significant shift changes in the Gla domain are apparent for residues 12−14, 19, and 23−25. Several of these residues are in contact with residues in the hydrophobic stack helix, such as Leu 13 which interacts with Phe 31 located on top of the hydrophobic stack (Sunnerhagen et al., 1995). Since, at 1 equiv of Ca^{2+} added, the EGF domain is only 80% saturated, it cannot be excluded that some shift changes in the Gla domain are due to Ca^{2+} binding to this domain. However, the dominant effect must be due to Ca^{2+} binding to the EGF domain in the domain pair.

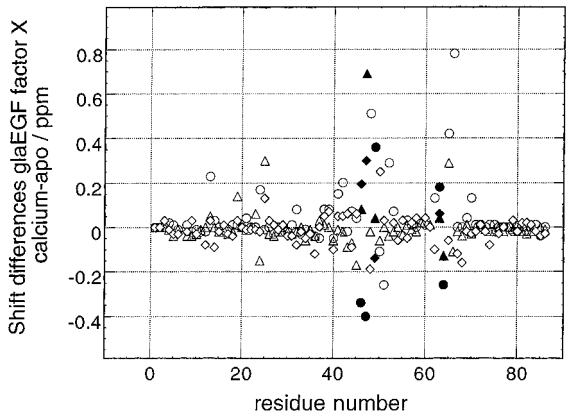

FIGURE 1: Shift changes on Ca^{2+} binding to the Gla−EGF module pair (1 Ca^{2+} form vs apo). The Gla domain includes residues 1−44 and the EGF domain residues 45−86. For the side chains, only the largest shift changes are displayed. Residues involved in Ca^{2+} binding to the isolated N-terminal EGF module (Sunnerhagen et al., 1992) are labeled with filed symbols. Symbols: (○) NH; (△) $C^\alpha H$; (◇) side chain hydrogens.

Table 1: Distribution of Number of NOE Constraints in Factor X Gla−EGF

	total	EGF	Gla	interdomain
1 Ca^{2+} Gla−EGF				
intraresidue	333	169	164	
sequential	210	142	68	
medium range	95	33	58	4
long range	112	84	19	9
apo Gla−EGF				
intraresidue	320	164	156	
sequential	180	112	68	
medium range	72	28	44	2
long range	113	86	27	6

NMR Structure of the 1 Ca^{2+} Gla−EGF Domain Pair. The structure determination of the 1 Ca^{2+} Gla−EGF domain pair was based on a constraint set of 750 NOE's (Table 1) and 55 dihedral angle constraints (44 ϕ and 11 χ^1) in the final structure calculations. The number of constraints per residue is less for the Gla domain than for the EGF domain in both cases. This is primarily due to a much better shift dispersion in the EGF domain than in the Gla domain but could also be due to the higher level of dynamics in the Gla domain as judged by the high amide proton exchange rates (Sunnerhagen et al., 1995). Alternative assignments for NOE's caused by spectral overlap were included using r^{-6} averaging in the first rounds of calculation. For many of these, only one of the alternative assignments was found structurally possible, and the NOE was thereby assigned. In the final constraint set, 28 alternative assignments remained. Four of these involve more than two residues, where separate NOE's were observed at 12 and/or 28 °C (18 HN→17 HN or 16 HN,27 HN→26 HN or 25 HN,37 HN→34 HA or 36 HA, and 49 HG*→46 HA or 67 HA). The other alternative assignments involve either of residues 5, 13, 18, 24, 58, or 65 which have severe side chain shift overlaps (see Supporting Information). From 100 initial structures, 37 converged to structures with distance violations <0.5 Å, sensible

Table 2

	ensemble	minimized av structure
energetic parameters (kcal mol^{-1})		
total energy	356 ± 34	316
bonds	33 ± 14	31
angles	135 ± 14	125
impropers	30 ± 6	24
van der Waals	58 ± 9	44
dihedral restraints	2.5 ± 1.5	1
NOE restraints	98 ± 10	92
deviations from exptl constraints		
rms NOE (Å)	0.036 ± 0.002	0.035
rms dih (deg)	0.69 ± 0.21	0.46
deviations from ideal geometry		
rms bond (Å)	0.005 ± 0.003	0.005
rms angles (deg)	0.60 ± 0.04	0.60
rms impropers (deg)	0.53 ± 0.05	0.47

folds, and comparably low energy parameters. A subset of 15 structures had no distance, violations greater than 0.3 Å, no dihedral angle violations above 5°, no obvious misfolds, and total energies below 400 kcal/mol. These structures were chosen to represent the 1 Ca^{2+} Gla−EGF domain pair. Energetic parameters and rms deviations to experimental constraints are given in Table 2.

The NMR structure family of factor X Gla−EGF with 1 equiv of Ca^{2+} added is shown in Figure 2, and corresponding rmsd values are in Table 3. It is clear from both the structures and rmsd values that the Gla and EGF domains are individually more well defined than the entire structure, as was also found for Ca^{2+}-free Gla−EGF. Within the domains, the α-helical and β-sheet regions are even more well defined. Especially for the EGF domain, the structure allows determination of the exact location of many of the side chains. As a whole, the 1 Ca^{2+} structure family has a significantly lower rmsd compared to the apo structure family. However, the rmsd's of the Gla and EGF domains taken separately are similar in the apo and 1 Ca^{2+} forms, which indicates a similar level of definition. This implies that the relative orientation of the Gla and EGF domains is better defined in the 1 Ca^{2+} structure than in the apo structure (Sunnerhagen et al., 1995).

Reorientation of Gla and EGF Domains As Observed by NMR. A comparison of the average structures of apo and 1 Ca^{2+} structures is given in Figure 3. In the 1 Ca^{2+} form, the Gla domain is more bent toward the EGF domain than in the apo form. The structural reorientation mainly involves a change in orientation of the hydrophobic stack helix (residues 34−42) in the Gla domain relative to the EGF domain. This is consistent with the fact that the shift changes observed on Ca^{2+}-binding mainly occur in the hydrophobic stack helix of the Gla domain. Although the orientation between the two domains is poorly defined due to the limited number of interdomain constraints observed (see below), the difference in relative orientation is present in the whole family of NMR structures. The rmsd for the average Gla−EGF domain structure in its apo form compared to the 1 Ca^{2+} form is significantly higher than the rmsd's of the structure families (Table 3). In contrast, the average structures of the individual domains are the same within experimental error in the apo and 1 Ca^{2+} forms as judged by the rmsd's between the average apo and 1 Ca^{2+} structures, which are similar to those of the separate structure families (Table 3).

Residues that show interdomain NOE's are primarily located in the linker region between the Gla and EGF domains, in the hydrophobic stack helix in the Gla domain, and in the top of the major β-sheet in the EGF domain (an overview of the interdomain constraints for the apo and 1 Ca^{2+} forms is given in Table 1 in Supporting Information). Only two constraints are present in both apo and 1 Ca^{2+} forms. As a rule, NOE interdomain distances observed in one ligand state are violated in the other, unless the NOE is observed in both ligand states. The location of interdomain constraints and participating residues in the three-dimensional structure is given in Figure 4. Special care was taken to confirm that the NOE constraint Phe 40 HN to Gly 66 HA* that seems to be responsible for pulling the Gla domain down in the apo form is *not* present in the 1 Ca^{2+} form. The NOE contacts between the two domains mainly involve hydrophobic residues. At the current resolution, no hydrogen bonds between the two domains have been identified.

Interestingly, the Gla domain folds over onto the EGF domain Ca^{2+} site region (Figure 5). Although the structure of the EGF domain in general is well determined, the side chains in the Ca^{2+} site show multiple conformations (Figures 2 and 5) . This may be due to the nonoptimal sample conditions for NMR as well as to the fact that the EGF domain Ca^{2+} site is only 80% occupied at 1 equiv of Ca^{2+} added (Valcarce et al., 1994). Therefore, at the current resolution details on the coordination of Ca^{2+} in the domain pair remain unclear. Interestingly, Asp 48 and Ile 65 show NOE's to each other, which is not the case for the isolated Ca^{2+}-loaded EGF domain where Asp 48 only shows tertiary NOE's to Glu 51 (Selander-Sunnerhagen et al., 1992). This could indicate that, for Gla−EGF, Asp 48 is included in the ligand sphere or indirectly in the Ca^{2+}-mediated interaction between Gla and EGF domains. At the current resolution, it is also not possible to identify additional ligands from the Gla domain that may contribute to Ca^{2+} binding in the EGF domain. Actually, all Gla residues but one (Gla 14) point away from the Ca^{2+} site as defined in the isolated EGF domain (Figure 5). This may be a true observation but could also be due to the fact that NOE's to the Gla side chains are notoriously hard to define due to severe side chain overlaps among the Gla residues. Furthermore, no Ca^{2+} ion has been introduced in the calculations to help fold the Ca^{2+} site. This may result in carboxyl groups pointing away from a plausible ligating position due to the absence of NOE's to carbons and oxygens, as for Asp 46 in the isolated EGF domain (Selander-Sunnerhagen et al., 1992) which was recently observed to be a Ca^{2+} ligand in the X-ray structure of the homologous factor IX EGF domain (Rao et al., 1995). The proximity of Gla 14 to the presumed Ca^{2+} site is primarily due to the presence of an NOE between residues Arg 15 and Gly 64.

Reorientation of Gla and EGF Domains As Observed by SAXS. We used SAXS as an independent technique to compare the global folds and the relative dispositions of the domains in the Gla−EGF pair with and without Ca^{2+}. The SAXS technique is particularly powerful when one is interested in the relationships between domains or modules in a protein, as demonstrated by our previous comparisons of the solution and crystal forms of calmodulin and troponin C (Heidorn & Trewhella, 1988). R_g and d_{max} values obtained from the $P(r)$ analysis of the SAXS data are given in Table 4, along with the corresponding values determined for the NMR structures and for the model in which the Gla domain

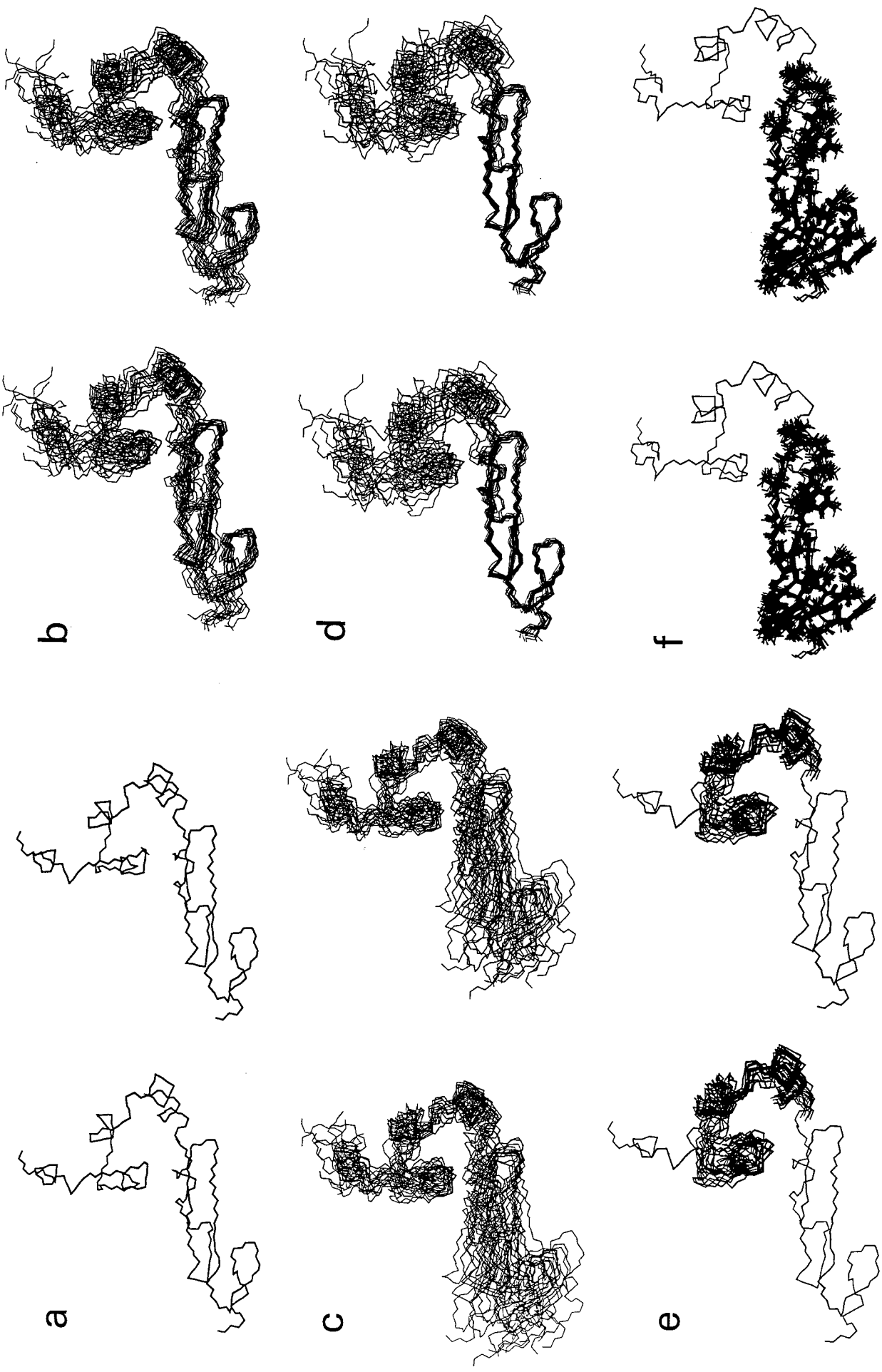

FIGURE 2: Structure of the Gla−EGF module pair with 1 equiv of Ca^{2+} added. (a) Stereo representation of the backbone average structure. (b−f) Stereo representations of the family of NMR structures superimposed onto the average structure by minimizing the rmsd for the heavy atoms of (b) residues 4−86, the Gla−EGF domain pair omitting residues 1−3 due to lack of NOE's (backbone atoms only), (c) residues 4−44, the Gla domain (backbone), (d) residues 45-86, the EGF domain (backbone), (e) residues 13−42, i.e., the α-helical part of the Gla domain (backbone), and (f) residues 55−84, i.e., the β-sheet part of the EGF domain (backbone and side chain atoms) not displaying side chains His 60, Lys 62, Glu 77, and Lys 79 since they were disordered.

Table 3

RMS Deviations (Å) to the Average Structure for Apo and 1 Ca^{2+} Gla−EGF NMR Structure Families[a]

section of protein	1 Ca^{2+}, backbone	1 Ca^{2+}, all	apo, backbone	apo, all
Gla−EGF (4−86)	1.69	2.37	2.54	3.14
Gla (4−44)	1.55	2.39	1.47	2.27
EGF (45−86)	0.67	1.28	0.92	1.47

RMS Deviations (Å) between Apo and 1 Ca^{2+} Gla−EGF Average Structures

section of protein	1 Ca^{2+} versus apo, backbone
Gla−EGF (4−86)	5.95
Gla (4−44)	1.72
EGF (45−86)	1.59
EGF without N-terminal residues involved in interdomain interactions (50−86)	0.98

[a] The three N-terminal residues (1−3) are not taken into account as they lack NOE's. Backbone includes N, C_α, and C′ whereas all includes all heavy atoms.

was disordered. A systematic error is observed in the R_g and d_{max} values due to the finite Q-range used in our analyses. The impact of omitting the lower Q-values affected by interparticle interference is seen by comparing the NMR model analyses using two different Q_{min} values. There is a slight apparent decrease in the R_g values when data between $Q = 0.001$ and 0.06 $Å^{-1}$ are omitted. This decrease is approximately 0.3−0.4 Å for the two NMR models and 1 Å for the model with a disordered Gla domain. The finite Q_{max} also impacts the apparent R_g and d_{max} values, calculated to be an ∼5% decrease from those calculated using an infinite Q-range. Table 5 shows a similar comparison of R_g and R_c values obtained from a Guinier analysis. Using the same Q-ranges for analysis, good agreement is found between the R_g and R_c values calculated from the NMR and SAXS data for both apo and 1 Ca^{2+} conditions, and the relative changes measured for these parameters are well predicted by the NMR structures. In contrast, the R_g, d_{max}, and R_c values for the Gla−EGF domain pair assuming a disordered Gla domain (in the absence of Ca^{2+}) are significantly different than the measured values—the disordered Gla domain predictably giving rise to larger R_g and d_{max} values and a smaller R_c.

Figure 6 shows the $I(Q)$ vs Q data for the apo and 1 Ca^{2+} concentration series, while Figure 7 shows the scattering profiles extrapolated to zero protein concentration compared with the scattering profiles calculated from the apo and 1 Ca^{2+} NMR structures. The down turn in the measured scattering profiles at Q-values below 0.058 $Å^{-1}$ is due to the interparticle interference effects and is clearly present even for the profiles extrapolated to zero concentration. Good fits were obtained between the scattering profiles calculated from each NMR structure and its corresponding SAXS measurement as evaluated by reduced χ^2values ≤ 1.5 (Table 6). In contrast, the comparison between the apo (or 1 Ca^{2+}) NMR structure and the 1 Ca^{2+} (or apo) SAXS measurement shows relatively poor agreement, as does the comparison between the model structure with the disordered Gla domain and the apo SAXS measurement (reduced χ^2 value of 6.0).

$P(r)$ functions for apo and 1 Ca^{2+} Gla−EGF are shown in Figure 8. The $P(r)$ functions for apo and 1 Ca^{2+} NMR structures are again in good agreement with the SAXS measurements. Both the apo and 1 Ca^{2+} forms have peaks at ∼14 Å with the apo form being a little more asymmetric and showing an additional shoulder at ∼30 Å. This shoulder is slightly lower for the profile calculated from the apo NMR structure compared to the measurement; this is reasonable, however, since the error bars shown for the $P(r)$ functions are based only on counting statistics and are calculated as one standard deviation. The dominant effect in omitting the scattering below Q of 0.06 $Å^{-1}$ is larger uncertainty in longer pair−distance vectors. This effect is apparent in the $P(r)$ functions and is manifested by large error bars for lengths approaching the d_{max} values.

The structural parameters determined from the SAXS data using both the Guinier and $P(r)$ analyses show that a contraction of the Gla−EGF domain pair occurs upon addition of Ca^{2+}. This result is consistent with the NMR data in which the two domains move toward each other about some sort of hinge between the two domains upon addition of Ca^{2+}. Such a movement would result in the observed reductions in both R_g and d_{max} values and the observed changes in the $P(r)$ functions. This last point is also supported by a comparison of the measured scattering profiles and the scattering profiles calculated from the apo and 1 Ca^{2+} NMR structures.

Structure in the Gla Domain—SAXS and NMR Evaluation. In order to see if a model with a disordered Gla domain could fit the experimental scattering data, a model structure with a disordered Gla domain but a folded EGF domain was simulated. For the apo case in Figure 8a, it is evident that the simulated structure with the disordered Gla domain does not fit the data. Therefore, the SAXS data strongly suggest that the Gla domain, when part of the Gla−EGF domain pair, is not disordered.

In the NMR structure of the 1 Ca^{2+} Gla−EGF domain pair, the Gla domain has a fold similar to that of the Ca^{2+}-loaded Gla domain as determined by X-ray crystallography. This was also previously observed for factor X apo Gla−EGF (Sunnerhagen et al., 1995). Although a complete set of constraints could not be obtained for the Gla domain due to shift overlap and rapid amide proton exchange, the number of medium- and long-range constraints in combination with the dihedral angle constraints is still sufficient to give the global fold of the Gla domain (Table 1). A survey of medium- and long-range distance constraints projected onto the tertiary structure of the Gla domain from the 1 Ca^{2+} Gla−EGF NMR structure is given in Figure 9. The distribution of constraints is very similar in the apo Gla−EGF NMR structure. All constraints used in the structure determination are being deposited together with the structures in the Brookhaven Protein Data Bank.

DISCUSSION

In this paper, we have presented the structure of the Gla−EGF fragment from factor X in two ion ligation states: without Ca^{2+} and with 1 equiv of Ca^{2+} bound to the EGF domain (80% site saturation). Using two independent techniques for structure determination in solution, SAXS and NMR, we show that on Ca^{2+} binding to the EGF domain in solution, the Gla and EGF domains fold toward each other. The relative orientation of the Gla and EGF domains thus depends on metal ion binding to the EGF domain. The Ca^{2+} site, as identified in the isolated domain structure (Selander-Sunnerhagen et al., 1992), seems to be part of the interdomain interaction surface. Finally, SAXS data and NMR data consistently show that the Gla domain is not disordered when

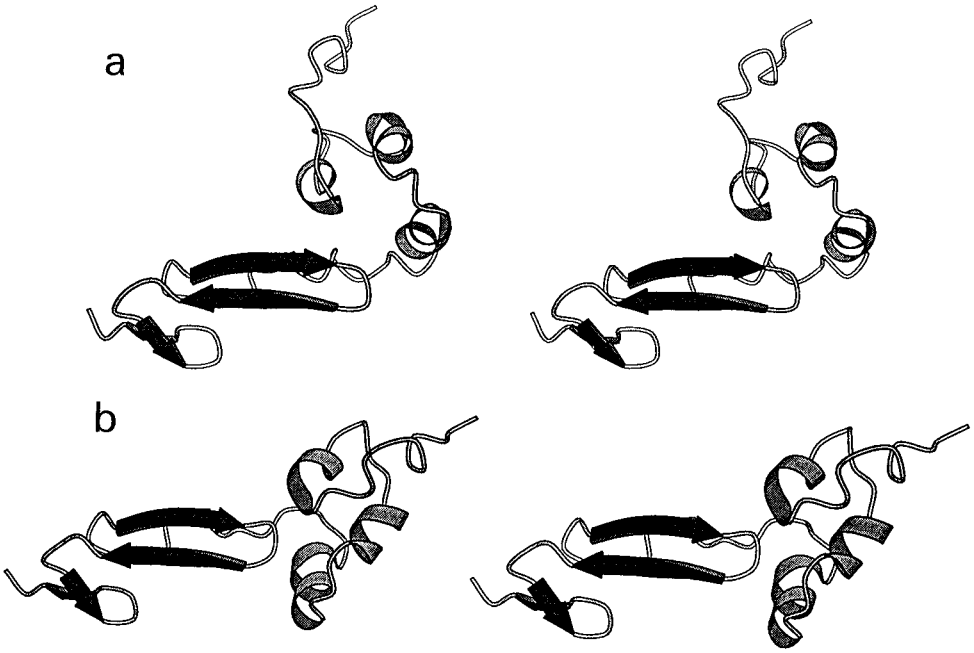

FIGURE 3: Stereo representation of the secondary structure elements in the Gla−EGF domain pair (a) with 1 equiv of Ca^{2+} added and (b) in the absence of Ca^{2+} (Sunnerhagen et al., 1995). The figure was drawn with the program Molscript (Kraulis, 1991).

part of the Gla−EGF domain pair. Since NMR uses only short-range distance constraints for structure determination, an accurate relative orientation of two domains is hard to achieve if the interaction area is small, as for the Gla−EGF domain pair. SAXS, on the other hand, is relatively insensitive to differences in short-range atomic distances but is sensitive to long-range structural alterations such as rearrangements of protein domains. Thus these highly complementary methods are a powerful combination—especially for studying modular proteins like coagulation factor X. The agreement found in our studies between the NMR and SAXS data strengthens our conclusions concerning the Gla−EGF interactions.

The present structure is the first one solved for a Ca^{2+}-bound EGF domain covalently linked by its N-terminus to another domain. Several structures have, however, been solved for domain pairs containing EGF domains without the consensus for Asp/Asn hydroxylation and Ca^{2+} binding and with other domains linked N-terminal to the EGF domain: the lectin−EGF domain pair (Graves et al., 1994) and prostaglandin H2 synthase-1 (Picot et al., 1994), both determined by X-ray crystallography, and the fibronectin type 1−EGF domain pair in tissue type plasminogen activator determined by NMR (Smith et al., 1994, 1995). Interestingly, all these domain pairs have well-defined relative orientations, often stabilized by interdomain hydrogen bonds. Furthermore, the EGF domain contact area is similar to that of the N-terminal EGF domain in the factor X Gla−EGF domain pair.

The introduction of a Ca^{2+} site in the N-terminal part of an EGF domain seemingly allows the relative orientation of domains to be altered in a Ca^{2+}-dependent manner instead of having a fixed relative orientation as observed for non-Ca^{2+}-binding EGF domain pairs. The biological importance of this structural effect is crucial, as shown by several naturally occurring mutations in the Ca^{2+} site regions. For example, mutations that destroy the Ca^{2+}-binding site in the N-terminal EGF domain in coagulation factor IX cause hemophilia B (Rees et al., 1988; Handford et al., 1991). In *Drosophila*, mutations in the Ca^{2+} site region alter the direction of embryonal development (Kelley et al., 1989; Rebay et al., 1991), whereas a mutation of the conserved hydroxyaspartic acid in fibrillin causes Marfan's syndrome (Hewett et al., 1993; Glanville et al., 1994; Knott et al., 1995; Wu et al., 1995).

In the apo structure, the contact area between the EGF and Gla domains is restricted to the linker region, the top of the major β-sheet in the EGF domain, and one face of the hydrophobic stack helix in the Gla domain. In the 1 Ca^{2+} structure, the contact area is somewhat increased, but still involves residues in the same region. For both apo and Ca^{2+} structures, the observed connectivities are mainly due to hydrophobic interactions, and no interdomain hydrogen bonds were identified. In a recent crystal structure of intact factor IX in the absence of Ca^{2+}, the same regions are involved in Gla−EGF domain interactions (Brandstetter et al., 1995). We have previously noted that EGF domains with the Ca^{2+}-binding consensus have a strong preference for hydrophobic residues in the top of the major β-sheet in the EGF domain, which would indicate that this part of the EGF domain might be involved in protein−domain interactions (Sunnerhagen et al., 1992). This now seems to be confirmed for factors IX and X and thus also for homologous coagulation factors with the Gla−EGF domain pair.

The hydrophobic stack helix involved in the Gla−EGF domain interaction is the major structural element of the Gla domain (Soriano-Garcia et al., 1992; Sunnerhagen et al., 1995). Whether the interaction between the Gla domain and a neighboring domain is necessary for folding of the Gla domain, particularly in the absence of Ca^{2+}, is still unclear. In the absence of Ca^{2+}, the NMR structure of the isolated Gla domain from factor IX is much less ordered than that of the Gla domain of factor X when linked to the EGF domain (Freedman et al., 1995a; Sunnerhagen et al., 1995). Therefore, it is possible that the hydrophobic helix of the Gla domain is stabilized by the interaction with the EGF domain which may thus facilitate folding of the entire Gla domain. In the presence of Ca^{2+}, however, the NMR structure of the isolated Gla domain from factor IX is very similar to the crystal structure of the Ca^{2+}-bound Gla domain

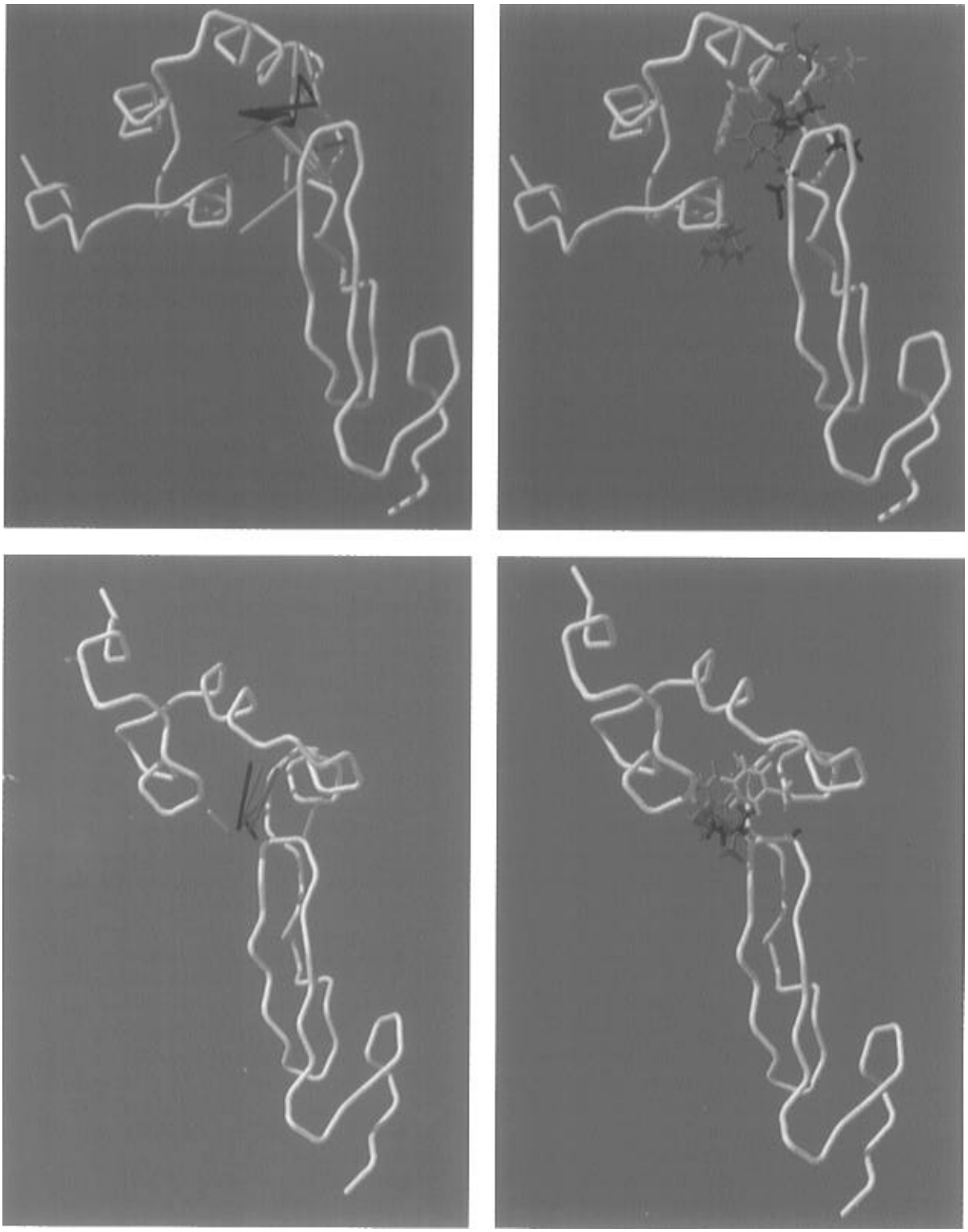

FIGURE 4: (a, top left) Display of interdomain NOE constraints between the Gla domain (residues 1−44) and the EGF domain (residues 45−86) for the 1 Ca^{2+} form. Constraints are subgrouped into unique constraints for the 1 Ca^{2+} form (yellow), constraints also present in the apo form (red), and constraints unique for the 1 Ca^{2+} form, but where constraints between other atoms in the same residue pair are present in the apo form (green). (b, top right) Residues involved in interdomain NOEs in the 1 Ca^{2+} form: Arg 15 (green), Gla 39 (light green), Phe 40 (yellow), Lys 43 (light orange), Tyr 44 (dark orange), Asp 46 (red), Asp 48 (red), Gly 64 (magenta), and Ile 65 (magenta). (c, bottom left) Display of interdomain NOE constraints between the Gla domain (residues 1−44) and the EGF domain (residues 45−86) for the apo form (Sunnerhagen et al., 1995). Constraints are subgrouped into unique constraints for the apo form (yellow), constraints also present in the 1 Ca^{2+} form (red), and constraints unique for the apo form, but where constraints between other atoms in the same residue pair are present in the 1 Ca^{2+} form (green). (d, bottom right) Residues involved in interdomain NOE's in the apo form: Gla 14 (green), Phe 40 (yellow), Tyr 44 (orange), Asp 46 (dark orange), Ile 65 (magenta), and Gly 66 (magenta).

of prothrombin fragment 1 (Freedman et al., 1995b). Additional work is required to resolve whether the different folds of Ca^{2+}-free Gla domains in factors IX and X reflect inherent differences between the Gla domains of the two proteins or a requirement of an adjacent EGF domain to attain a native conformation.

The Gla domain folds over onto the EGF domain on Ca^{2+} binding in such a way that Ca^{2+} should be sandwiched

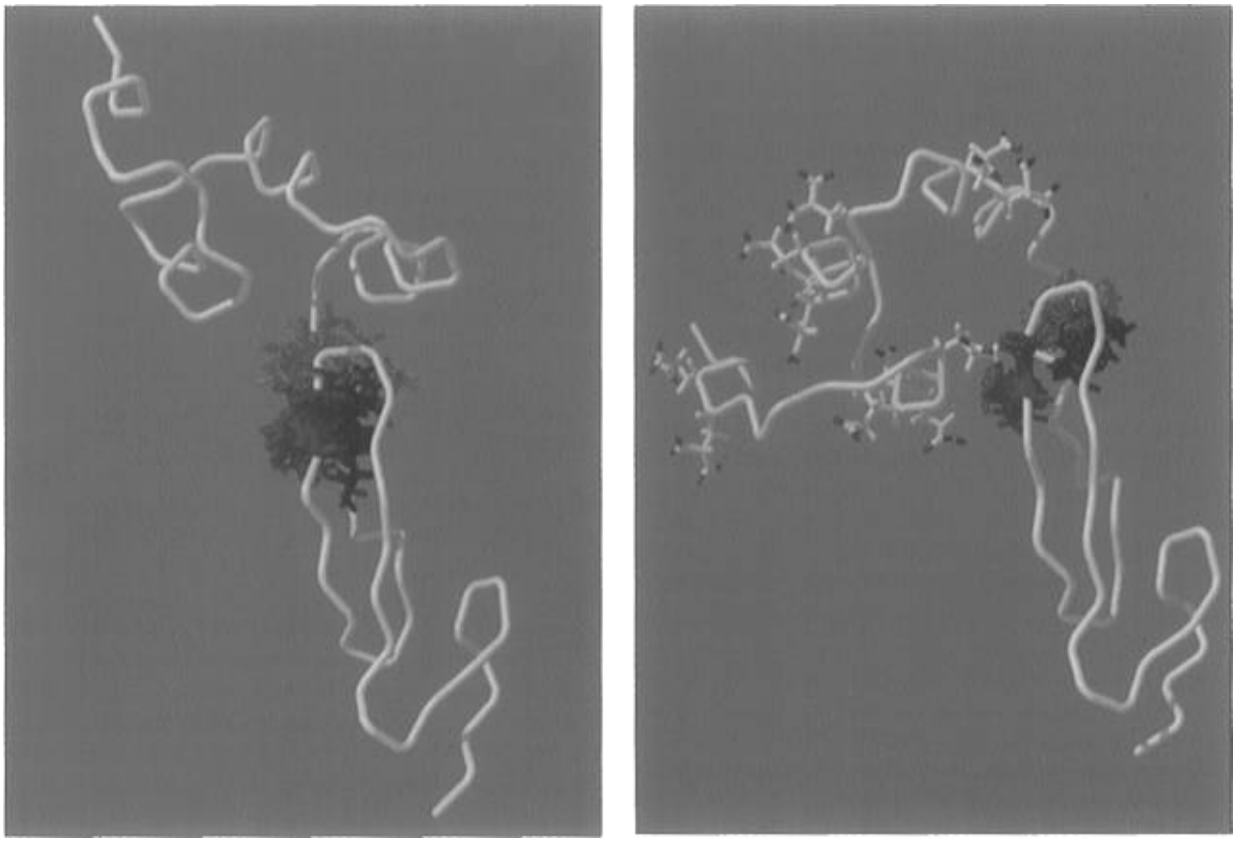

FIGURE 5: Residues Asp 46 (green), Gly 47 (light blue), Gln 49 (dark blue), Hya 63 (magenta), and Gly 64 (pink), which are Ca^{2+} ligands in the isolated EGF domain pair of factor X, are displayed on the average structure of (a, left) apo and (b, right) 1 Ca^{2+} forms of the Gla−EGF domain pair. Backbone heavy atoms of residues 46−66 were used for the local superposition of the NMR structure families onto the average structure. For the 1 Ca^{2+} form in (b), the Gla residues in the energy-minimized average structure are also displayed.

Table 4: R_g and d_{max} Values from $P(r)$ Analysis

	ΔQ (Å^{-1})	R_g (Å)	d_{max} (Å)
SAXS data			
apo	0.058−0.28	16.0 ± 0.4	52 ± 4
1 Ca^{2+}	0.058−0.28	14.2 ± 0.2	45 ± 3
NMR structures			
apo	0.058−0.28	15.6	51
apo	0.001−0.28	16.2	51
1 Ca^{2+}	0.058−0.28	14.6	48
1 Ca^{2+}	0.001−0.28	14.9	48
disordered Gla domain model			
apo	0.058−0.28	18.6	64
1 Ca^{2+}	0.001−0.28	19.6	64

between the two domains. It is highly suggestive that an extra Ca^{2+} ligand is provided by the Gla domain, especially since the Ca^{2+} affinity is 10-fold increased in the Gla−EGF domain pair (Valcarce et al., 1994) and the Ca^{2+} coordination sphere is not complete for the isolated EGF domain (Selander-Sunnerhagen et al., 1992). Unfortunately, due to the low resolution of the present structure in the Ca^{2+} site region, no details on the ligation of Ca^{2+} in the Gla−EGF domain pair can be given. The only Gla residue in our structures that comes close to the suggested Ca^{2+}-binding region is Gla 14, due to the presence of an NOE between Arg 15 and Gly 64. In the apo form, an NOE between Gla 14 and Ile 65 was identified. Thus, in both the apo and Ca^{2+} state there seems to be an interdomain interaction involving the Gla 14−Arg 15 residues. Interestingly, a mutation of Gla 14 in factor X which gave rise to bleeding disorders was recently identified (Kim et al., 1995). Arg 15 is strictly conserved in Gla domains, and an Arg 15→Leu mutation in protein C severely reduces the anticoagulant activity (Zhang et al., 1992). It is therefore possible that this interdomain interaction is biologically important. It is less likely, however, that either of these residues participate in Ca^{2+} binding, since a factor X peptide including residues 29−86 has the same Ca^{2+} affinity as intact factor X (Valcarce et al., 1993). A previously suggested ligand from the Gla domain, Gla 39 (Rao et al., 1995), clearly points away from the interaction region as defined by interdomain constraints and interacting residues (Figure 5) and would not be able to participate in the ligation of Ca^{2+} according to the hypothetical model suggested by Rao et al. Due to the presence of new NOE's between residues Asp 48 and Ile 65, it is possible that Asp 48 may be involved either in Ca^{2+} binding or indirectly in the Gla−EGF domain interaction. In this context it is noteworthy that an Asp-to-Glu mutation at position 48 in factor IX is destructive for the biological activity of the intact protein, i.e., where the EGF and Gla domains are covalently linked (Rees et al., 1988).

The change in domain orientations on Ca^{2+} binding to the EGF domain, as well as the stabilizing effect on the Gla domain provided by the EGF domain, may be important for protein−protein interactions. Several examples of protein interactions involving Gla−EGF domain pairs where the EGF domain has a Ca^{2+} site are known. Coagulation factor VII binds to tissue factor with high specificity in the presence of Ca^{2+}. Experiments using chimeric proteins and analysis of mutations have identified determinants for tissue factor binding in both Gla and EGF domains of factor VII (Ruf et al., 1991; Toomey et al., 1991; Chaing et al., 1994; Chang et al., 1995), and a recent X-ray structure of the factor VII−tissue factor complex confirms that both Gla and EGF

Table 5: R_g and R_c Values from Guinier Analysis[a]

	ΔQR_g ($Å^{-1}$)	R_g (Å)	ΔQR_c ($Å^{-1}$)	R_c (Å)
SAXS data				
apo	0.060−0.082	16.2 ± 0.9	0.10−0.19	7.0 ± 0.3
1 Ca^{2+}	0.061−0.099	13.9 ± 0.7	0.10−0.19	7.3 ± 0.3
NMR structures				
apo	0.060−0.082	15.7	0.10−0.19	7.0
apo	0.004−0.082	16.2		
1 Ca^{2+}	0.060−0.099	14.6	0.10−0.19	7.4
1 Ca^{2+}	0.004−0.099	14.8		
disordered Gla domain model				
apo	0.060−0.078	17.2	0.10−0.25	5.6
apo	0.004−0.071	18.8		

[a] ΔQR_g and ΔQR_c are the Q-ranges used for the R_g and R_c determinations, respectively.

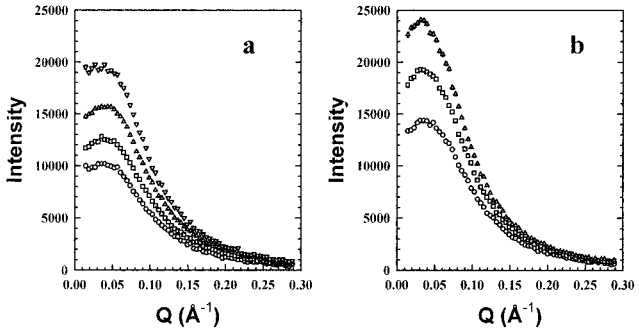

FIGURE 6: SAXS data measured for the apo form (a) and the 1 Ca^{2+} form (b) of the Gla−EGF domain pair. For the apo form concentrations are (O) 14.2, (□) 18.5, (△) 22.7, and (▽) 27.3 mg/mL and for the 1 Ca^{2+} form concentrations are (○) 16.5, (□) 21.6, and (▽) 26.3 mg/mL.

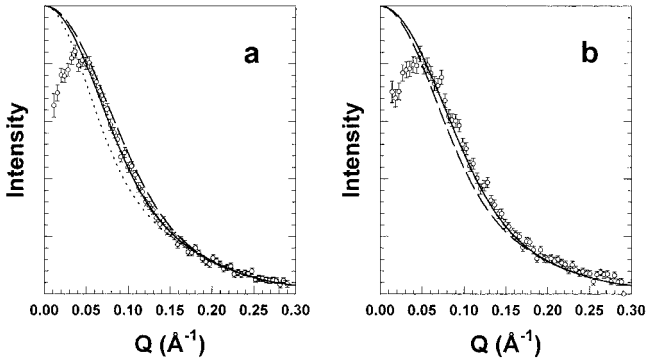

FIGURE 7: SAXS data extrapolated to zero concentration for the apo (a) and 1 Ca^{2+} form (b). Scattering curves simulated from the NMR data were corrected for instrument resolution and then scaled by $I(0)$ obtained from the Guinier analysis. The solid line in (a) is that from the apo NMR structure and the dashed line that from the 1 Ca^{2+} NMR structure. Inversely, in (b), the solid line is that from the 1 Ca^{2+} NMR structure and the dashed line that from the apo NMR structure. The dotted line in (a) is from the simulated average structure with disordered Gla domain.

Table 6: Reduced χ^2 Values for $I(Q)$ versus Q Fits between Simulated NMR Structures and Measured SAXS Data[a]

	reduced χ^2
apo SAXS data compared to	
apo NMR structure	0.6
1 Ca^{2+} NMR structure	1.9
simulated apo structure with disordered Gla	6.0
1 Ca^{2+} SAXS data compared to	
apo NMR structure	4.7
1 Ca^{2+} NMR structure	1.5

[a] A Q-range of 0.06−0.28 $Å^{-1}$ was used for reduced χ^2 calculations.

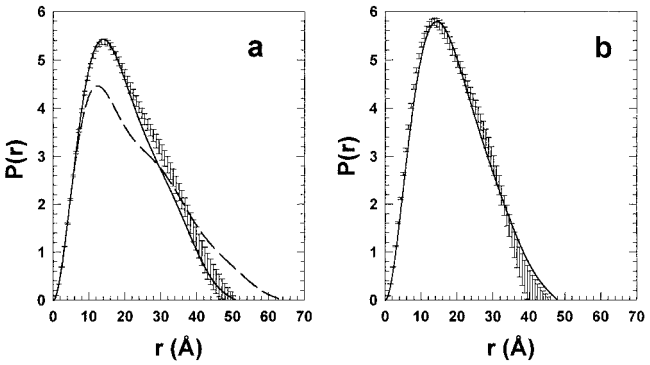

FIGURE 8: $P(r)$ functions determined from the extrapolated X-ray scattering data from the apo form (a) and 1 Ca^{2+} form (b) are plotted with standard error bars representing one standard deviation. The solid lines are calculated from the respective NMR structures. The dashed line in (a) is from the simulated average structure with disordered Gla domain. The Q-range used in this analysis was 0.058−0.28 $Å^{-1}$.

domains participate in the interaction (Banner et al., 1996). In protein C, which interacts with the thrombin−thrombomodulin complex, only the intact Gla−EGF region and not the isolated domains inhibit protein C activation in the presence of Ca^{2+} (Hogg et al., 1992). Due to the high extracellular free Ca^{2+} concentrations in blood (≈1 mM) the EGF domain Ca^{2+} site should always be nearly saturated, and thereby, Ca^{2+} binding to the EGF domain would not have any regulatory effect on the coagulation process for proteins VII, IX, and X. Ca^{2+} binding rather seems to be needed to stabilize the biologically active interdomain orientation in a way not achievable by the peptide chain only. In another environment, however, a protein with a Ca^{2+}-binding EGF domain could possess a passive as well as an active conformation, using the metal ion to stabilize intermodular orientations in order to expose or hide a recognition surface for protein interactions.

Although this is the first direct structural observation of a Ca^{2+}-mediated interaction between two different extracellular domains in the same protein, it has been observed by X-ray crystallography how Ca^{2+} ion sites can form part of a protein interaction surface. In phospholipase A_2, Ca^{2+} ligates both to the protein and to the bound substrate (Scott et al., 1990). In the assembly of the tobacco mosaic virus, the viral capsid is rigidified by Ca^{2+} ions binding to sites in the interface between capsid proteins (Ilag et al., 1994). The surface-exposed Ca^{2+} sites of the annexins have been demonstrated to mediate interactions between the annexin and a phospholipid surface (Swairjo et al., 1995). In cadherin, surface Ca^{2+} site-mediated interactions between identical domains in the crystal lattice result in the buildup of a zipper-like supermolecular ribbon believed to be the fundamental unit of cadherin adhesive function (Shapiro et al., 1995). Ca^{2+} sites

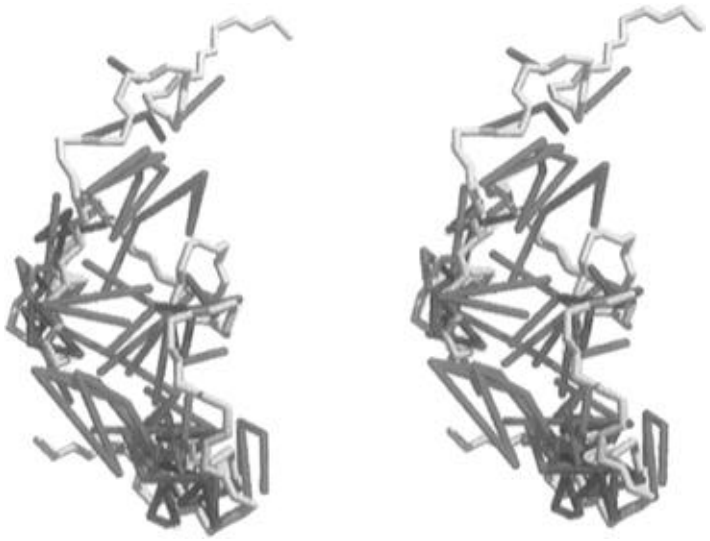

FIGURE 9: Stereo representation of the secondary and tertiary NOE constraints in the Gla domain in the 1 Ca^{2+} Gla−EGF domain pair displayed on the average NMR structure. For simplification, only the backbone atoms are shown although the constraints are drawn between their respective atoms. Specific α-helix constraints [$d_{NN}(i, i+2)$, $d_{\alpha N}(i, i+2)$, $d_{\alpha N}(i, i+3)$, $d_{\alpha N}(i, i+4)$, $d_{\alpha\beta}(i, i+3)$] are drawn in red, whereas other constraints are in green.

identified by NMR are suggested to act in a similar way for a cadherin homologue, bacterial protein S (Bagby et al., 1994). For the isolated factor IX EGF domain, a neighboring identical domain in the crystal lattice provides a ligand to the EGF domains Ca^{2+} site. This interaction between two isolated EGF domains has formed the basis for hypothetical models of Ca^{2+}-dependent interactions between Ca^{2+}-binding domains in fibrillin (Handford et al., 1995; Wu et al., 1995).

We have shown that Ca^{2+} binding to the EGF site induces domain interaction between the Gla and EGF domains. It is probable that the role of the Ca^{2+} binding is to position these domains in proper orientations relative to one another. The severely reduced biological function observed when Ca^{2+} site residues are mutated in factor IX and protein C as well as *Drosophila* Notch and fibrillin stresses the structural and biological importance of the EGF domain Ca^{2+} site. In recognition mechanisms for protein interactions that involve Gla and EGF domains, we suggest that Ca^{2+} binding can induce a biologically active conformation either by exposing an interaction surface or by enabling the domain pair to fit into an active cleft in the receptor protein. In view of our results, it seems highly likely that Ca^{2+} binding to EGF domains in proteins with many sequentially linked EGF domains may be involved in positioning these EGF domains relative to each other so that a biologically active conformation is obtained.

ACKNOWLEDGMENT

The technical assistance of Ann-Marie Thämlitz is gratefully acknowledged. The Swedish NMR Center is acknowledged for generously making available a Silicon Graphics workstation for the structure calculations. Dr. Bryan Finn is acknowledged for expert assistance in the making of Figures 4, 5, and 9.

SUPPORTING INFORMATION AVAILABLE

Table 1 containing the ^1H NMR chemical shifts of the Ca^{2+}-free Gla−EGF domain pair, Table 2 containing the chemical shifts observed in the presence of 1 equiv of Ca^{2+}, Table 3 giving the interdomain NOE's in the Gla−EGF without Ca^{2+}, and Table 4 giving NOE's in the presence of 1 equiv of Ca^{2+} (8 pages). Ordering information is given on any current masthead page.

REFERENCES

Appella, E., Robinson, L. A., Ullrich, S. J., Stoppelli, M. P., Corti, A., Cassani, G., & Blasi, F. (1987) *J. Biol. Chem. 262*, 4437−4440.

Astermark, J., Björk, I., Öhlin, A.-K., & Stenflo, J. (1991) *J. Biol. Chem. 266*, 2430−2437.

Bagby, S., Harvey, T. S., Eagle, S. G., Inouye, S., & Ikura, M. (1994) *Structure 2*, 107−122.

Banner, D. W., D'Arcy, A., Chène, C., Winkler, F. K., Guha, A., Konigsberg, W. H., Nemerson, Y., & Kirchhofer, D. (1996) *Nature 380*, 41−46.

Brandstetter, H., Bauer, M., Huber, R., Lollar, P., & Bode, W. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9796−9800.

Brünger, A. T. (1992) *X-PLOR version 3.1*, Yale University, New Haven, CT.

Barlow, P. N., Steinkasserer, A., Norman, D. G., Kieffer, B., Wiles, A. P., Sim, R. B., & Campbell, I. D. (1993) *J. Mol. Biol. 232*, 268−284.

Campbell, I. D., & Bork, P. (1993) *Curr. Opin. Struct. Biol. 3*, 385−392.

Carpenter, G., & Wahl, M. I. (1990) *Handbook Exp. Pharmacol. 95*, 69−171.

Chaing, B. S., Clarke, B., Sridhara, S., Chu, K., Friedman, P., VanDusen, W., Roberts, H. R., Blajchman, M., Monroe, D. M., & High, K. A. (1994) *Blood 83*, 3524−3535.

Chang, J.-Y., Stafford, D. W., & Straight, D. L. (1995) *Biochemistry 34*, 12227−12232.

Chazin, W. J., Rance, M., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 603−622.

Chen, S.-H., & Bendedouch, D. (1986) *Methods Enzymol. 130*, 79−116.

Cheung, W.-F., Straight, D. L., Smith, K. J., Lin, S.-W., Roberts, H. R., & Stafford, D. W. (1991) *J. Biol. Chem. 266*, 8797−8800.

Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. (1986) *J. Mol. Biol. 191*, 523−551.

Davis, C. G. (1990) *New Biol. 2*, 410−419.

Engel, J. (1989) *FEBS Lett. 251*, 1−7.

Familletti, P. C., Wolitzky, B. A., & Burns, D. K. (1994) *Nature 367*, 532−538.

Freedman, S. J., Furie, B. C., Furie, B., & Baleja, J. D. (1995a) *J. Biol. Chem. 270*, 7980−7987.

Freedman, S. J., Furie, B. C., Furie, B., & Baleja, J. D. (1995b) *Biochemistry 34*, 12126−12137.

Furie, B., & Furie, B. C. (1988) *Cell 53*, 505−518.

Glanville, R. W., Qian, R.-Q., McClure, D. W., & Maslen, C. L. (1994) *J. Biol. Chem. 269*, 26630−26634.
Graves, B. J., Crowter, R. L., Chandran, C., Rumberger, J. M., Li, S., Huang, K.-S., Presky, D. H., & Guinier, A. (1939) *Ann. Physiol. (Paris) 12*, 161−237.
Handford, P. A., Mayhew, M., Baron, W., Winship, P. R., Campbell, I. D., & Brownlee, G. G. (1991) *Nature 351*, 164−167.
Hansen, A. P., Petros, A. M., Meadows, R. P., Nettesheim, D. G., Mazar, A. P., Olejinczak, E. T., Xu, R. X., Pederson, T. M., Henkin, J., & Fesik, S. W. (1994) *Biochemistry 33*, 4847−4864.
Heidorn, D. B., & Trewhella, J. (1988) *Biochemistry 27*, 909−915.
Heidorn, D. B., Seeger, P. A., Rokop, S. E., Blumenthal, D. K., Means, A. R., Crespi, H., & Trewhella, J. (1989) *Biochemistry 28*, 6757−6764.
Hewett, D. R., Lynch, J. R., Smith, R., & Sykes, B. C. (1993) *Hum. Mol. Genet. 2*, 475−477.
Hogg, P. J., Öhlin, A. C., & Stenflo, J. (1992) *J. Biol. Chem. 267*, 703−706.
Husten, E. J., Esmon, C. T., & Johnsson, A. E. (1987) *J. Biol. Chem. 262*, 12953−12961.
Jacobs, M., Freedman, S. J., Furie, B. C., & Furie, B. (1994) *J. Biol. Chem. 269*, 25494−25501.
Jacrot, B., & Zaccai, G (1981) *Biopolymers 20*, 2413−2426.
Johansson, C., Ullner, M., & Drakenberg, T. (1993) *Biochemistry 32*, 8429−8438.
Kim, D. J., Thompson, A. R., & James, H. L. (1995) *Hum. Genet. 95*, 212−214.
Knott, V., Downing, A. K., Cardy, C. M., & Handford, P. (1996) *J. Mol. Biol. 255*, 22−27.
Kotkow, K. J., Furie, B., & Furie, B. C. (1993) *J. Biol. Chem. 268*, 15633−15639.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kurosawa, S., Stearns, D. J., Jackson, K. W., & Esmon, C. T. (1988) *J. Biol. Chem. 263*, 5993−5996.
Mann, K. G., Jenny, R. G., & Krishnaswamy, S. (1988) *Annu. Rev. Biochem. 57*, 915−956.
Mann, K. G., Nesheim, M. E., Church, W. R., Haley, P., & Krishnaswamy, S. (1990) *Blood 76*, 1−16.
Medved, L. V., Vysotchin, A., & Ingham, K. C. (1994) *Biochemistry 33*, 478−485.
Moore, P. B. (1980) *J. Appl. Crystallogr. 13*, 168−175.
Nishimura, H., Takeya, H., Miyata, T., Suehiro, K., Okamura, T., Niho, Y., & Iwanaga, S. (1993) *J. Biol. Chem. 268*, 24041−24046.
Öhlin, A.-K., & Stenflo, J. (1987) *J. Biol. Chem. 262*, 13798−13804.
Öhlin, A.-K., Linse, S., & Stenflo, J. (1988) *J. Biol. Chem. 263*, 7411−7417.
Öhlin, A.-K., Björk, I., & Stenflo, J. (1990) *Biochemistry 29*, 644−651.
Olah, G. A., & Trewhella, J. (1994) *Biochemistry 33*, 12800−12806.
Olah, G. A., Trakhanov, S., Trewhella, J., & Quiocho, F. A. (1993) *J. Biol. Chem. 268*, 16241−16247.
Olah, G. A., Rokop, S. E., Wang, C.-L., Blechner, S. L., & Trewhella, J. (1994) *Biochemistry 33*, 8233−8239.
Padmanabhan, K., Padmanabhan, K. B., Tulinsky, A., Park, C. H., Bode, W., Huber, R., Blankenship, D. T., Cardin, A. D., & Kisiel, W. (1993) *J. Mol. Biol. 232*, 947−966.
Persson, E., & Petersen, L. C. (1995) *Eur. J. Biochem. 234*, 293−300.
Persson, E., Selander, M., Linse, S, Drakenberg, T., Öhlin, A.-K., & Stenflo, J. (1989) *J. Biol. Chem. 264*, 16897−16904.
Persson, E., Björk, I., & Stenflo, J. (1991) *J. Biol. Chem. 266*, 2444−2452.
Picot, D., Loll, P. J., & Garavito, R. M. (1994) *Nature 367*, 243−249.
Pilz, I. (1982) in *Small-Angle X-ray Scattering* (Glatter, O., & Kratky, O., Eds.) pp 244−247, Academic Press, New York.
Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G. G., & Stuart, D. (1995) *Cell 82*, 131−141.
Rebay, I., Fleming, R. J., Fehon, R. G., Cherbas, L., Cherbas, P., & Artawanis-Tsakonas, S. (1991) *Cell 67*, 687−699.
Rees, D. J. G., Jones, I. M., Handford, P. A., Walter, S. J., Esnouf, M. P., Smith, K. J., & Brownlee, G. G. (1988) *EMBO 7*, 2053−2061.
Ruf, W., Kalnik, M. W., Lund-Hansen, T., & Edgington, T. S. (1991) *J. Biol. Chem. 266*, 15719−15725.
Ryan, J., Wolitzky, B., Heimer, E., Lambrose, T., Felix, A., Tam, J. P., Huang, L. H., Nawroth, P., Wilner, G., Kisiel, W., Nelsestuen, G. L., & Stern, D. M. (1989) *J. Biol. Chem. 264*, 20283−20287.
Schwalbe, R. A., Ryan, J., Stern, D. M., Kisiel, W., Dahlbäck, B., & Nelsestuen, G. L. (1989) *J. Biol. Chem. 264*, 20288−20296.
Scott, D. L., White, S. P., Otwinowski, Z., Yuan, W., Gelb, M. H., & Sigler, P. B. (1990) *Science 250*, 1541−1546.
Selander, M., Persson, E., Stenflo, J., & Drakenberg, T. (1990) *Biochemistry 29*, 8111−8118.
Selander-Sunnerhagen, M., Ullner, M., Persson, E., Teleman, O., Stenflo, J., & Drakenberg, T. (1992) *J. Biol. Chem. 267*, 19642−19649.
Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, L.-F., Als-Nielsen, J., Colman, D. R., & Hendrickson, W. A. (1995) *Nature 374*, 327−337.
Smith, B. O., Downing, A. K., Dudgeon, T. J., Cunningham, M., Driscoll, P. C., & Campbell, I. D. (1994) *Biochemistry 33*, 2422−2429.
Smith, B. O., Downing, A. K., Driscoll, P. C., Dudgeon, T. J., & Campbell, I. D. (1995) *Structure 3*, 823−833.
Soriano-Garcia, M., Padmanabhan, K., deVos, A. M., & Tulinsky, A. (1992) *Biochemistry 31*, 2554−2566.
Sosnick, T., Charles, S., Stubbs, G., Yau, P., Bradbury, E. M., Timmins, P., & Trewhella, J. (1991) *Biophys. J. 60*, 1178−1189.
Stenflo, J. (1991) *Blood 78*, 1637−1651.
Stenflo, J., Lundwall, Å., & Dahlbäck, B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 368−372.
Stenflo, J., Öhlin, A.-K., Owen, W., & Schneider, W. J. (1988) *J. Biol. Chem. 263*, 21−24.
Strynadka, N. C. J., & James, M. N. G. (1989) *Annu. Rev. Biochem. 58*, 951−998.
Strynadka, N. C. J., & James, M. N. G. (1991) *Curr. Opin. Struct. Biol. 1*, 905−914.
Sunnerhagen, M., Forsèn, S., Hoffren, A.-M., Drakenberg, T., Teleman, O., & Stenflo, J. (1995) *Nat. Struct. Biol. 2*, 504−509.
Swairjo, M. A., Concha, N. O., Kaetzel, M. A., Dedman, J. R., & Seaton, B. A. (1995) *Nat. Struct. Biol. 2*, 968−974.
Toomey, J. R., Smith, K. J., & Stafford, D. W. (1991) *J. Biol. Chem. 266*, 19198−19202.
Toomey, J. R., Smith, K. J., Roberts, H. R., & Stafford, D. W. (1992) *Biochemistry 31*, 1806−1808.
Tulinsky, A., Park, C. H., & Skrzypczak-Jakun, E. (1988) *J. Mol. Biol. 202*, 885−901.
Ullner, M., Selander, M., Persson, E., Stenflo, J., Drakenberg, T., & Teleman, O. (1992) *Biochemistry 31*, 5974−5983.
Valcarce, C., Selander-Sunnerhagen, M., Taemlitz, A.-M., Drakenberg, T., Björk, I., & Stenflo, J. (1993) *J. Biol. Chem. 268*, 26673−26678.
Valcarce, C., Holmgren, A., & Stenflo, J. (1994) *J. Biol. Chem. 269*, 26011−26016.
Wu, Y.-S., Bevilaqua, V. L. H., & Berg, J. M. (1995) *Chem. Biol. 2*, 91−97.
Zhang, L., & Castellino, F. J. (1992) *J. Biol. Chem. 267*, 26078−26084.
Zhang, L., Jhingan, A., & Castellino, F. J. (1992) *Blood 80*, 942−952.

BI960633J