Articles

Structure and Function of the Epidermal Growth Factor Domain of P-Selectin[†,‡]

Steven J. Freedman,[§,||] David G. Sanford,[§] William W. Bachovchin,[§] Barbara C. Furie,[§,||] James D. Baleja,[§] and Bruce Furie*[,§,||]

Center for Hemostasis and Thrombosis Research, Division of Hematology−Oncology, New England Medical Center, and Departments of Medicine and Biochemistry, Tufts University School of Medicine and Sackler School of Graduate Biomedical Sciences, Boston, Massachusetts 02111

Received April 30, 1996; Revised Manuscript Received August 15, 1996[⊗]

ABSTRACT: P-selectin is a multidomain adhesion protein on the surface of activated platelets and endothelial cells that functions in the recruitment of leukocytes to the site of inflammation. The amino-terminal lectin and EGF domains constitute the ligand recognition unit. We have produced a synthetic 40-residue P-selectin EGF domain (P-sel:EGF) to examine the structure and function of this domain independent of P-selectin. The peptide was folded *in vitro* and exhibited the same disulfide bonding pattern as other EGF-like domains. P-sel:EGF did not inhibit P-selectin-mediated cellular adhesion assays, indicating that the lectin domain is also required. We undertook the study of the P-selectin EGF by ^{1}H NMR to determine its structure independent of the lectin domain and to compare its structure to that of E-selectin determined crystallographically [Graves et al. (1994) *Nature 367*, 532]. Although the binding of P-selectin to its carbohydrate ligand is calcium dependent, and some EGF domains have calcium binding sites, addition of calcium had no effect on the NMR spectrum or on the pH-induced changes. Nearly complete resonance assignments were made from 2D ^{1}H NMR spectra at pH 6.0. Two sections of antiparallel β-sheet were identified on the basis of the pattern of long-range NOEs, $^3J_{HN\alpha}$ coupling constants, and slowly exchanging amides. The solution structure of the peptide backbone was determined using distance geometry and simulated annealing calculations. The backbone RMSD to the geometric average for 19 final structures is 0.64 ± 0.17 Å. The resulting fold closely resembles that of other EGF-like peptides, including the E-selectin EGF domain (RMSD ∼1.08 Å). However, compared to the E-selectin EGF structure which also contains the lectin domain, some residues from 1−11 are less ordered, and novel contacts occur between the amino terminus and the core β-sheet. Despite marked structural homology of the selectin polypeptide backbones, the selectin EGF surfaces show unique distributions of charged residues, a feature that likely correlates to the functional differences.

P-selectin is a member of the selectin family of cell adhesion proteins which mediate the interaction of platelets and endothelial cells with leukocytes [reviewed in Lasky (1995) and McEver et al. (1995)]. All selectins comprise an amino-terminal lectin domain, an EGF[1] domain, a variable number of complement regulatory protein repeats, a transmembrane domain, and a short cytoplasmic tail. Each selectin demonstrates calcium-dependent binding to its counterreceptor. Although the lectin domain is sufficient to support binding, the lectin−EGF domains comprise the optimal ligand recognition unit (Li et al., 1994; Murphy & McGregor, 1994; Gibson et al., 1995). The precise role of the EGF domain has been elusive, mostly as a result of conflicting studies regarding a direct interaction with selectin ligands (Walz et al., 1990; Bowen et al., 1990; Imai et al., 1992; Heavner et al., 1993; Kansas et al., 1994; Murphy & McGregor, 1994).

The selectin EGF domains are members of the epidermal growth factor (EGF) family. Domains in this family are related to EGF by amino acid sequence homology and structural similarity (Appella et al., 1988; Engel, 1989; Campbell & Bork, 1993). EGF domains are found in a variety of extracellular proteins of diverse biological function including coagulation and fibrinolytic proteins, components of the complement system, molecules involved in embryonic

[†] This work was supported by a grant (HL51926) from the National Institutes of Health. The NMR spectrometer was acquired with a grant (RR06282) from the National Institutes of Health.

[‡] The P-sel:EGF structure has been deposited in the Brookhaven Protein Data Bank under file name 1FSB. Distance and torsion angle restraints, including force constants, and the molecular topology file are also available as R1FSBMR.

* Corresponding author.

[§] Department of Biochemistry.

[||] Center for Hemostasis and Thrombosis Research and Department of Medicine.

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

[1] Abbreviations: EGF, epidermal growth factor; P-sel:EGF, P-selectin EGF; E-sel:EGF, E-selectin EGF; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TOCSY, total correlation spectroscopy; DQF-COSY, double-quantum-filtered correlation spectroscopy; TPPI, time-proportional phase incrementation; RMSD, root-mean-square deviation; FMOC, 9-fluorenylmethyloxycarbonyl; NMP, *N*-methylpyrrolidone; HMP, 4-(hydroxymethyl)phenoxy; Pmc, 2,2,5,7,8-pentamethylchroman-6-sulfonyl; OtBu, *tert*-butyl alcohol; tBu, *tert*-butyl; Trt, trityl; Boc, *tert*-butyloxycarbonyl; DTT, dithiothrietol; BCECF-AM, 2′,7′-bis(2-carboxyethyl)-5- (and 6-) carboxyfluorescein acetoxymethyl ester; CHO, Chinese hamster ovary; FBS, fetal bovine serum; TFA, trifluoroacetic acid.

S0006-2960(96)01025-2 CCC: $12.00

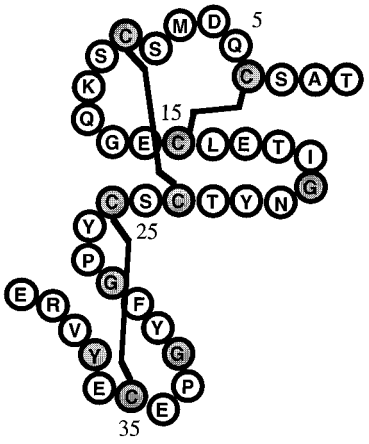

FIGURE 1: Amino acid sequence of the EGF domain of P-selectin. The 40 amino acid residue domain is located between the lectin domain and the first complement repeat of P-selectin (Johnston et al., 1989). The residue numbers indicated allow cross reference to other EGF domain sequences. EGF residue 1 is equivalent to Thr 119 in human P-selectin while EGF residue 40 is equivalent to Glu 158 (Johnston et al., 1989). Residues strictly conserved among EGF domains are shaded. The three disulfide bonds are shown as thick black bars. The schematic drawing reflects some of the structural features common to EGF domains and demonstrated in this study.

development, and cellular adhesion receptors. EGF domains vary in size between 40 and 50 residues, and their characteristic structures are partly defined by the strict conservation of six cysteines forming three disulfide bridges. In addition, nearly all EGF domains, including those of selectins, contain three conserved glycines and one conserved tyrosine at positions 20, 29, 32, and 37 (Figure 1) (Dudgeon et al., 1991). Some EGF domains bind calcium and have a characteristic Ca^{2+} binding consensus sequence composed of acidic residues (Handford et al., 1991). This sequence is absent from the selectin EGF domains despite the Ca^{2+}-dependent adhesive function of the parent proteins. There is little sequence homology between EGF domains of functionally diverse, unrelated proteins (Dudgeon et al., 1991). However, sequence identity is approximately 50% among the selectin EGF domains and about 33% compared to the first EGF domain of factor IX. Although the role of EGF domains varies depending on the effector function of their parent molecules, they appear to be primarily involved in protein−protein interactions (Campbell & Bork, 1993). The structural basis for these functions has been the subject of intense investigation (Campbell & Bork, 1993; Groenen et al., 1994).

Each of the EGF and EGF-like domain structures solved by high-resolution techniques such as NMR and X-ray crystallography provides new insights in structure and function (Harvey et al., 1991; Montelione et al., 1992; Baron et al., 1992; Hommel et al., 1992; Selander-Sunnerhagen et al., 1992; Graves et al., 1994; Rao et al., 1995; Jacobsen et al., 1996). The principal structural feature that these domains share is two regions of double-stranded antiparallel β-sheet. The major sheet is within the N-terminal subdomain, which is composed of approximately three-quarters of the sequence and contains two disulfide bonds. On the N-terminal side of this sheet are turns and loops that vary with each EGF domain. The C-terminal subdomain has a small antiparallel β-sheet and the third disulfide bond. In addition, tight turns separate each secondary structure as well as each strand of the sheets. The N-terminus can form the third strand of a triple-stranded β-sheet, as in TGF-α, murine and human EGF, and factor X:Ca^{2+}, or be disordered, as in apo factor X and apo factor IX (Harvey et al., 1991; Selander-Sunnerhagen et al., 1992; Baron et al., 1992; Cooke et al., 1990; Montellione et al., 1992; Ullner et al., 1992). On the basis of such structures certain critical residues for EGF domain function have been identified by comparing homologous structures and sequences in combination with the results of site-directed mutagenesis (Dudgeon et al., 1990; Campbell et al., 1990; Campbell & Cooke, 1990). The only structural data to date on the EGF domains of the selectin family of adhesion proteins are from the crystal structure of the lectin−EGF domains of E-selectin (Graves et al., 1994). The E-selectin EGF domain has a global fold similar to that of other EGF-like domains. The E-selectin EGF domain makes only a limited number of contacts to the amino-terminal lectin domain.

To study the structure and function of the P-selectin EGF domain, we have prepared a synthetic peptide that exhibits the disulfide bonding pattern characteristic of EGF domains. Its three-dimensional structure has been determined and compared to the homologous E-selectin EGF domain. This study demonstrates that the P-selectin EGF domain has a fold similar to that of other EGF domains but is distinguished by a unique molecular surface which is required, but not sufficient, for cell adhesion. We propose that the linked lectin−EGF domains in the native protein generate the optimal surface for leukocyte binding to the counterreceptor PSGL-1.

MATERIALS AND METHODS

Peptide Synthesis. Synthesis was carried out on an Applied BioSystems 430A peptide synthesizer using the solid-phase method and FMOC/NMP chemistry as previously described (Jacobs et al., 1994). The crude peptide was refolded by a modification of the procedure applied to a human factor IX EGF synthetic peptide (Huang et al., 1989). Approximately 20 mg of peptide was dissolved in 20 mL of 8 M urea/0.2 M dithiothrietol (DTT) in a 0.1 M Tris-HCl buffer, pH 8.45. Residual scavengers and DTT were removed by stepwise dialysis (Spectrapor 7) against decreasing concentrations of urea in 0.1 M Tris-HCl, pH 8.45. At a final concentration of 1.5 M urea, the peptide solution was diluted 5-fold with 0.1 M Tris-HCl, pH 8.45, containing 1 mM oxidized glutathione/1 mM reduced glutathione. After 24 h the oxidized peptide was analyzed by reverse-phase HPLC using a linear gradient of acetonitrile/0.1% TFA (Beckman System GOLD; VYDAC 218TP, 5μm; 4.6 mm $\times$ 250 mm; flow rate, 1 mL/min). For large-scale preparation, urea and glutathione were removed from the folded peptide by dialysis against four changes of 50 mM ammonium bicarbonate, pH 7.6, lyophilized, and dissolved in 5 mL of H_2O. The peptide solution was injected onto a preparative reverse-phase C_{18} column and eluted with a linear gradient of acetonitrile/0.1% TFA (Bio-Rad Hi-Pore C_{18} reverse phase; 21.5 mm $\times$ 250 mm; flow rate, 8 mL/min). For NMR experiments the purified peptide was dissolved to 3.5 mM in H_2O.

Disulfide Bond Determination. The EGF peptide (10 mg) was proteolytically digested with three times recrystallized thermolysin (8% w/w) in 20 mM Tris-HCl, pH 7.8, and 150 mM NaCl at 45 °C for 24 h. Peptides in the hydrolysate

were separated on an analytical C_{18} reverse-phase HPLC column (VYDAC 218TP, 5 μm; 4.6 mm × 250 mm; flow rate, 1 mL/min) and eluted with a linear gradient of acetonitrile/0.1% TFA. Peaks were rechromatographed to ensure peptide purity and subsequently subjected to N-terminal sequencing on a Milligen/Biosearch 8800 Prosequencer.

Cell Adhesion Assays. Plates (48 well) were seeded with CHO cells (5×10^4/well) or CHO cells expressing P-selectin (CHO:P) (Larsen et al., 1992) and incubated at 37 °C in 5% CO_2 for 2−3 days until wells were nearly confluent. HL60 cells (10^7) were labeled by a 30 min incubation with the fluorescent label BCECF-AM at 37 °C in 5% CO_2 and then washed twice with RPMI. Labeled HL60 cells were resuspended in RPMI containing 1% FBS such that 100 μL had 10^5 cells. P-sel:EGF was incubated with the HL60 cells for 20 min at 25 °C. In control experiments, intact human P-selectin was used instead of P-sel:EGF. CHO cells were washed three times to remove unbound cells with cold RPMI (4 °C) containing 1% FBS. HL60 cells (100 μL; 10^5) were added to each well and incubated for 30 min at 4 °C. The wells were washed three times with RPMI containing 1% FBS to remove unbound HL60 cells, and 100 μL of media was added to each well. Adherent HL60 cells were quantitated by fluorescence emission at 530 nm using a Millipore CytoFluor 2300.

NMR Experiments. The P-sel:EGF sample for 2D NMR experiments contained 3.5 mM peptide, 20 mM NaCl, and 10% D_2O, pH 6.0. The sample was pretreated with Chelex 100 to remove trace metal ions. For a similar sample in D_2O, all reagents were lyophilized from 10% D_2O and then redissolved in 99.96% D_2O. Spectra were collected at 25 °C on a Bruker AMX-500 spectrometer with a proton frequency of 500.14 MHz. The carrier frequency was set on the water resonance, which was suppressed using presaturation or by using jump-and-return methodology (Plateau & Guéron, 1982). For preliminary sample analysis under various pH, ionic strength, and temperature conditions, 1D experiments were acquired with 4096 real data points, 8−64 summed scans, and a spectral width of 7042 Hz and then processed by applying a squared sine bell window function shifted by 60°. NOESY spectra were recorded with mixing times of 75, 100, and 250 ms. A total of 2048 (or 4096) real data points was acquired in t_2, 415−470 TPPI (or States−TPPI) increments in t_1, a spectral width of 7042 Hz in the F_2 dimension, 128 summed scans, and a relaxation delay of 1.3 s between scans. Spectra were multiplied with a sine bell window function shifted by 60° in t_2 (applied over 1024 points) and a sine bell window function shifted by 60° in t_1 (applied over all acquired points) and zero filled to a 2K by 1K (real) matrix using the Bruker NMR processing program. NOESY cross peak intensities were converted into three distance classes (strong, 1.7−3.0 Å; medium, 1.7−4.0 Å; weak, 3.0−5.0 Å) and calibrated using published methods (Detlefsen, et al., 1991; Hyberts et al., 1992). Nonstereospecifically assigned atoms were treated as pseudoatoms and given correction distances according to the guidelines presented in Wüthrich (1986). Distance restraint information was extracted from NOESY spectra with different mixing times, but comparison of the short and longer mixing time spectra was used to control for spin diffusion effects. TOCSY spectra were recorded and processed using identical parameters as in the NOESY, except with a mixing time of 35 ms, with 96 summed scans, and using an MLEV-17 mixing sequence (Bax & Davis, 1985). A DQF-COSY spectrum was recorded with 4096 real t_2 points, 64 summed scans, and 714 TPPI increments. The spectrum was multiplied by sine bell window functions shifted by 45° in t_2 and 60° in t_1 and zero filled to a 2K by 1K (real) matrix. In addition, a similar set of experiments (NOESY, TOCSY, and DQF-COSY) were performed on a 4 mM peptide sample containing 0.1 M KCl and 0.01 M potassium phosphate, pH 7.2, and recorded at 33 °C. These spectra assisted in making proton resonance assignments and measuring $^3J_{HN\alpha}$ coupling constants.

Sequence-specific resonance assignments were made in two steps: (1) identification of intraresidue spin systems using the $^1H-^1H$ through-bond connectivities found in TOCSY and DQF-COSY spectra; (2) sequential assignment of residues on the basis of sequential $d_{\alpha N}$, $d_{\beta N}$, and d_{NN} NOE connectivities (Wüthrich et al., 1982). The NOESY spectrum collected on a sample in D_2O was used to distinguish aromatic from amide protons and expose α-protons that were near the presaturated water resonance in H_2O spectra. The p11 NOESY spectrum provided full amide intensity and also exposed α protons that were near the water resonance in presaturation experiments. The NOE contacts were classified into five categories: 1, "intraresidue" for NOEs within a residue; 2, "sequential" for contacts between the backbone and side chain protons of residue *i* with the backbone amide proton of residue $i + 1$; 3, "short range" for all other contacts between residue *i* and $i+1$; 4, "medium range" for NOEs between protons on residues separated by three amino acid positions or less in the sequence; and 5, "long range" for contacts between residues that are separated by four amino acid positions or more in sequence.

The vicinal spin−spin coupling constants $^3J_{HN\alpha}$ were used to calculate ϕ torsion angles (Pardi et al., 1984). The coupling constants were measured from the splitting of amide cross peaks in a NOESY spectrum that was resolution enhanced by multiplying with a squared sine bell window function shifted by 20° and applied over 2048 (real) points in t_2. Only $^3J_{HN\alpha} > 7.5$ Hz were used, and the upper and lower limits of the constraints were extended to include all possible ϕ angles.

χ_1 torsion angles and β-methylene proton stereospecific assignments were determined from a measurement of $^3J_{\alpha\beta}$ and a comparison of the βH−NH and αH−βH cross peak intensities in the NOESY spectra for each stereospecific β-proton of AMX residues. χ_1 torsion angle data for valines, threonines, and isoleucines were obtained from a measurement of the coupling constant, $^3J_{\alpha\beta}$, and an analysis of the αH−βH NOE cross peak shapes for remaining residues (Hyberts et al., 1987; Driscoll et al., 1989). The coupling constants, $^3J_{\alpha\beta}$, were measured from the splitting of β-proton cross peaks in a D_2O NOESY spectrum that was resolution enhanced by multiplying with a squared sine bell window function shifted by 30° and applied over 2048 (real) points in t_2. In this manner, the χ_1 conformations were designated as gauche/trans, trans/gauche, and gauche/gauche for AMX residues, gauche/gauche or non-gauche/gauche for non-AMX residues, and trans or gauche for threonines, valines, and isoleucines. Stereospecific assignments of glycine α protons (residues 13, 20, and 29) and Val 38 methyl groups were determined by comparing the quality and final energies of the structures having the two possible chiralities.

Hydrogen bond restraints were incorporated on the basis of the following two criteria. First, structures in the absence of such restraints were analyzed using InsightII which defines hydrogen bonds as a distance of less than 2.5 Å between the donor proton and heavy atom acceptor and by an angle between 120° and 180° of the heavy atom donor, proton, and heavy atom acceptor. Second, the labile hydrogen involved in the hydrogen bond is resistant to exchange in the presence of D_2O (from 20 min to 24 h). If these criteria were fulfilled, then hydrogen bond restraints were given upper and lower limit distances of 2.5 and 1.8 Å, respectively, between the donor proton and heavy atom acceptor.

Structure determination used a set of 859 distance restraints (intraresidue and sequential, 324; short, medium, and long range, 535), 14 ϕ angle and 29 χ_1 angle restraints, 10 stereospecific assignments, and 11 hydrogen bond restraints. A combination of distance geometry and simulated annealing methods (Havel, 1991) was used to generate 19 converging structures using the DGII program of InsightII (Biosym Technologies, San Diego, CA). The simulated annealing protocol has been described elsewhere (Freedman et al., 1995). The final total error function value was 0.33 ± 0.14 kcal/mol. Nonconverged structures had energies of >1 kcal/mol. The average structure for the ensemble was calculated using the Analysis program of InsightII. Average root-mean-square deviation (RMSD) values following superimposition of the backbone atoms of each structure with the geometric average reflected the quality of the structures determined. In addition, the coherence of torsion angles among different structures was evaluated. The average torsion angle was measured by a vector addition method (Hyberts et al., 1992). An order parameter, *S*, equals 1.0 when the torsion angle is the same in all structures, whereas an order parameter near 0 indicates disorder.

Selectin EGF Models. A molecular model of the EGF domain of L-selectin was prepared from the P-selectin EGF domain coordinates. Given the marked structural homology of the P-selectin EGF domain and the E-selectin EGF domain (*vide infra*) and the sequence homology between the EGF domains of P-selectin and L-selectin, these models were prepared using QUANTA (Molecular Simulations). The backbone structure and side chain orientations of the P-selectin EGF domain were preserved in the L-selectin EGF domain model.

RESULTS

The P-selectin EGF domain (P-sel:EGF) was prepared by solid-phase peptide synthesis. This 40-residue peptide corresponds to the intervening sequence between the lectin domain and the first complement repeat (Johnston et al., 1989) (Figure 1). As in all EGF domains, there are six conserved cysteines that form three disulfide bonds in the native molecules. The peptide was folded by a redox denaturation−renaturation protocol as applied to other EGF-like peptides (Huang et al., 1989). The peptide isomers were purified by HPLC, and the isomer in largest yield was subjected to disulfide bond analysis. Peptides from the thermolysin hydrolysate were isolated and subjected to N-terminal sequencing. Peptides connected by one or more disulfide bonds yield more than one peptide sequence, thus allowing positive identification of the disulfide bonding pattern (Figure 2). The disulfide pattern determined, Cys 4−Cys 15, Cys 9−Cys 24, and Cys 26−Cys 35, matches that of EGF and other EGF-like peptides (Cooke et al., 1987).

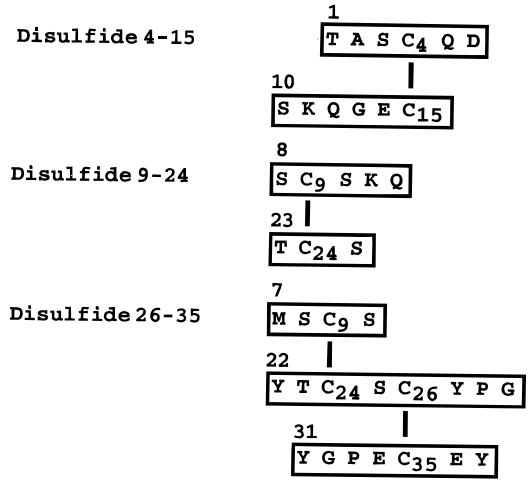

FIGURE 2: Disulfide bonding pattern of peptides derived from the thermolysin hydrolysate of the P-selectin EGF domain. The peptides were isolated by reverse-phase HPLC.

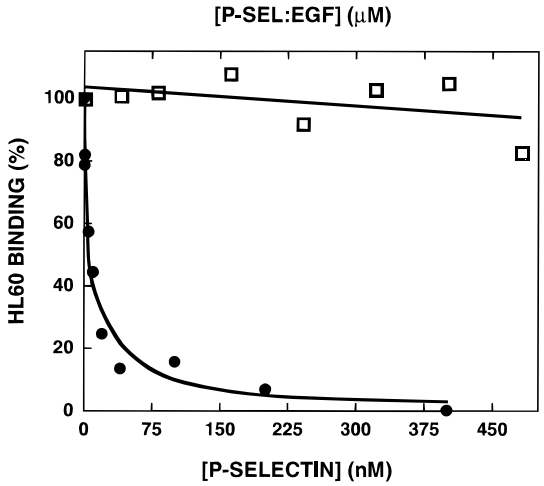

FIGURE 3: Inhibition of P-selectin-mediated cell adhesion by P-selectin and the P-selectin EGF domain. Adherent CHO or CHO−P-selectin (CHO:P) were grown to confluence in 48 well plates. Nonspecific sites were blocked with RPMI/1% BSA. HL60 cells (100 μL; 10^5), labeled with BCECF-AM, were incubated for 20 min at 25 °C with either P-selectin or P-sel:EGF at the indicated concentration. The HL60 cell suspension was added to each well containing adherent CHO cells and incubated for 30 min at 4 °C. After washing, bound HL60 cells were detected by fluorescence using an excitation wavelength at 485 nm and the emission wavelength at 530 nm. Results are the mean of two wells. Symbols: P-selectin EGF domain, □; P-selectin, ●.

A cell adhesion assay was employed to evaluate the ability of P-sel:EGF to compete with P-selectin expressed on the surface of CHO cells for the P-selectin ligand on HL60 cells. CHO or CHO:P-sel were grown to near confluence in 48 well plates, and fluorescently labeled HL60 cells, either alone or preincubated with purified human P-selectin or P-sel:EGF, were allowed to adhere. After the wells were washed to remove nonadherent HL60 cells, the number of HL60 cells remaining was quantitated by fluorescence measurement (Figure 3). P-selectin was inhibitory in the nanomolar range (half-maximal inhibition ~10 nM). However, P-sel:EGF was not inhibitory of P-selectin-mediated cell adhesion even at concentrations of peptide to 4 mM. Thus, the EGF domain alone is not sufficient to inhibit P-selectin binding to PSGL-1. ^{1}H NMR spectroscopy was used to study the structure

of P-sel:EGF. As a preliminary step to determining its three-dimensional structure, we analyzed the peptide over a range of pH and temperature conditions. The degree of dispersion in the amide region and the aliphatic region over a wide range of pH (2.5−10.4) and temperature (2.7−50.8 °C) indicates a well-ordered, stable structure. The overall pattern of the spectrum remains largely unchanged under all conditions, except for a few undefined small resonance shifts at various pH values. However, a notable change was the acid-induced upfield shift and broadening (and ultimate disappearance from the spectrum) of the Glu 34 amide proton resonance on lowering the pH from 6 to 5 ($\Delta\delta = \sim 0.2$ ppm) and of the Glu 36 amide proton resonance between pH 6 and pH 4 ($\Delta\delta = \sim 1.0$ ppm). On the basis of these observations the peptide structure was studied at pH 6.0 or above where spectral dispersion was greatest and line width broadening was minimized. The Glu 36 amide was downfield of 10 ppm at this pH, characteristic of other EGF domains studied (Baron et al., 1992; Tappin et al., 1989).

Since P-selectin-mediated adhesion to its ligand is a calcium-dependent interaction, and since some EGF domains such as those from the vitamin K-dependent blood coagulation proteins bind calcium ions, we determined whether calcium ions influence the P-sel:EGF ^{1}H NMR spectrum. The 1D spectrum was unchanged in the presence of 8 mM $CaCl_2$ compared to a sample pretreated with chelating agents to remove trace metal ions. In addition, $CaCl_2$ did not alter the pH-dependent resonance shifts, particularly that of the amide proton of Glu 36. Thus, we have no evidence that the P-sel:EGF peptide binds calcium ions.

To further understand the important contribution of the EGF domain to P-selectin function, and to compare its structure to related EGF molecules in the database, including the homologous E-selectin lectin−EGF crystal structure, we determined the structure of P-sel:EGF by 2D homonuclear ^{1}H NMR spectroscopy. Standard NOESY, TOCSY, and DQF-COSY experiments were used to assign the resonances (Wüthrich et al., 1982; Bax & Davis, 1985) (Figure 4). The starting point for making sequential assignments in the NOESY spectra was the conserved Glu 36 residue whose backbone amide proton is positioned downfield of 10 ppm in nearly all EGF peptide spectra. For the majority of residues, TOCSY spectra showed backbone amide to α and β connectivities, the only exceptions being those α-protons buried underneath the water resonance and connectivities to the amide proton of Cys 9 (see below), and the intraresidue connectivities to Thr 1, Pro 28, and Pro 33. In some cases, the NH−γH connectivities were also observed. α-Protons bleached out by presaturation were identified using a jump-and-return NOESY spectrum of the H_2O sample and NOESY and DQF-COSY spectra of the D_2O sample. Some amide resonances were very broad (width of peak at half-height, 75−140 Hz), including residues 7, 9, 10, and 11, whereas a few others were moderately broad (width of peak at half-height, 40−50 Hz), including residues 5, 34, and 36. The α-proton and/or side chain proton resonances were moderately broad for residues 4, 8, and 9. The amide proton resonance of Cys 9 was identified by its sequential connectivities to Ser 8 since none of the intraresidue connectivities were observed; the β-protons were identified in the aliphatic region of the NOESY and TOCSY spectra from the characteristic chemical shifts of the cysteine spin system, as well as being the only unassigned intraresidue cross peaks

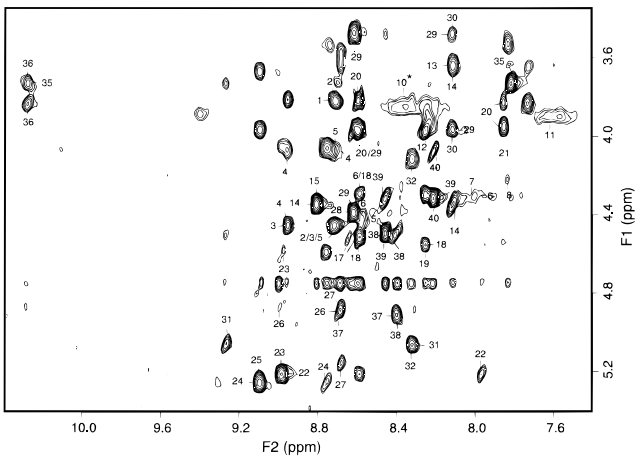

FIGURE 4: Two-dimensional ^{1}H NMR spectrum of the P-sel:EGF peptide in the fingerprint (αH−NH) region. The cross peaks involving intraresidue and sequential αH and backbone NH protons are labeled by residue number. The experimental conditions for the presaturation NOESY experiment are described in Materials and Methods. The mixing time was 75 ms. The asterisk (*) next to residue 10 is the βH−NH cross peak and is used to indicate the broad width of the amide proton resonance (the αH−NH cross peak does not appear at short mixing times). Other amide proton resonances shown here are particularly broad, including residues 7 and 11. The αH−NH cross peaks not observed in this spectrum were observed in other spectra using different pulse programs and mixing times except for Cys 9.

in the spectra. The α-proton of Cys 9 was the only unidentified proton resonance, possibly as a result of extreme resonance broadening. The proton resonance assignments are presented in Table S-1 (Supporting Information). From this information all remaining non-intraresidue, nonsequential NOE cross peaks could be identified.

Torsion angle data, ϕ and χ_1, on the P-sel:EGF peptide were derived from a measurement of both $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants as described in the Materials and Methods section. The spectral resolution precluded determination of $^3J_{HN\alpha}$ for amide protons that had extremely broad line widths, and for other amides measurements were limited to $^3J_{HN\alpha} > 7.5$ Hz. In sum, 14 ϕ angles were determined from the data. For spectroscopically distinguishable β-methylene protons of AMX residues, $^3J_{\alpha\beta}$ in conjunction with $d_{\alpha\beta}$ and $d_{HN\beta}$ provided restraints for one of the three possible χ_1 rotameric conformations. For the valine, isoleucine, and three threonines, the $^3J_{\alpha\beta}$ size limited the χ_1 angle to gauche or trans unless an averaged conformation existed. The χ_1 angle for all other residues (except the four glycines, two prolines, single alanine, and Cys 9) could be limited to either the gauche/gauche or non-gauche/gauche conformation on the basis of cross peak shapes. For the AMX residues, three are gauche/gauche (Tyr 22, Ser 25, and Tyr 31), and five are trans/gauche (Cys 24, Cys 26, Tyr 27, Phe 30, and Tyr 37). For the non-AMX residues, Thr 18 and Thr 23 are gauche, Val 38 is trans, and Glu 36 is gauche/gauche. The prolyl rings of residues 28 and 33 exist in the trans conformation on the basis of the NOE intensities: $d(\alpha H_i - \delta H_{(i+1)}) > d(\alpha H_i, \alpha H_{(i+1)})$ (Wüthrich, 1986). The major secondary structure in the P-sel:EGF peptide was assessed as antiparallel β-sheet by analysis of sequential $\alpha H - NH_{(i,i+1)}$ distances, $NH_{(i,j)}$ and $\alpha H_{(i,j)}$ NOEs, $^3J_{HN\alpha}$, and amide exchange data in D_2O (Figure 5). First, an alternating pattern of interstrand NH−NH and αH−αH contacts was observed between residues 12 and 18 with residues 21−27 and

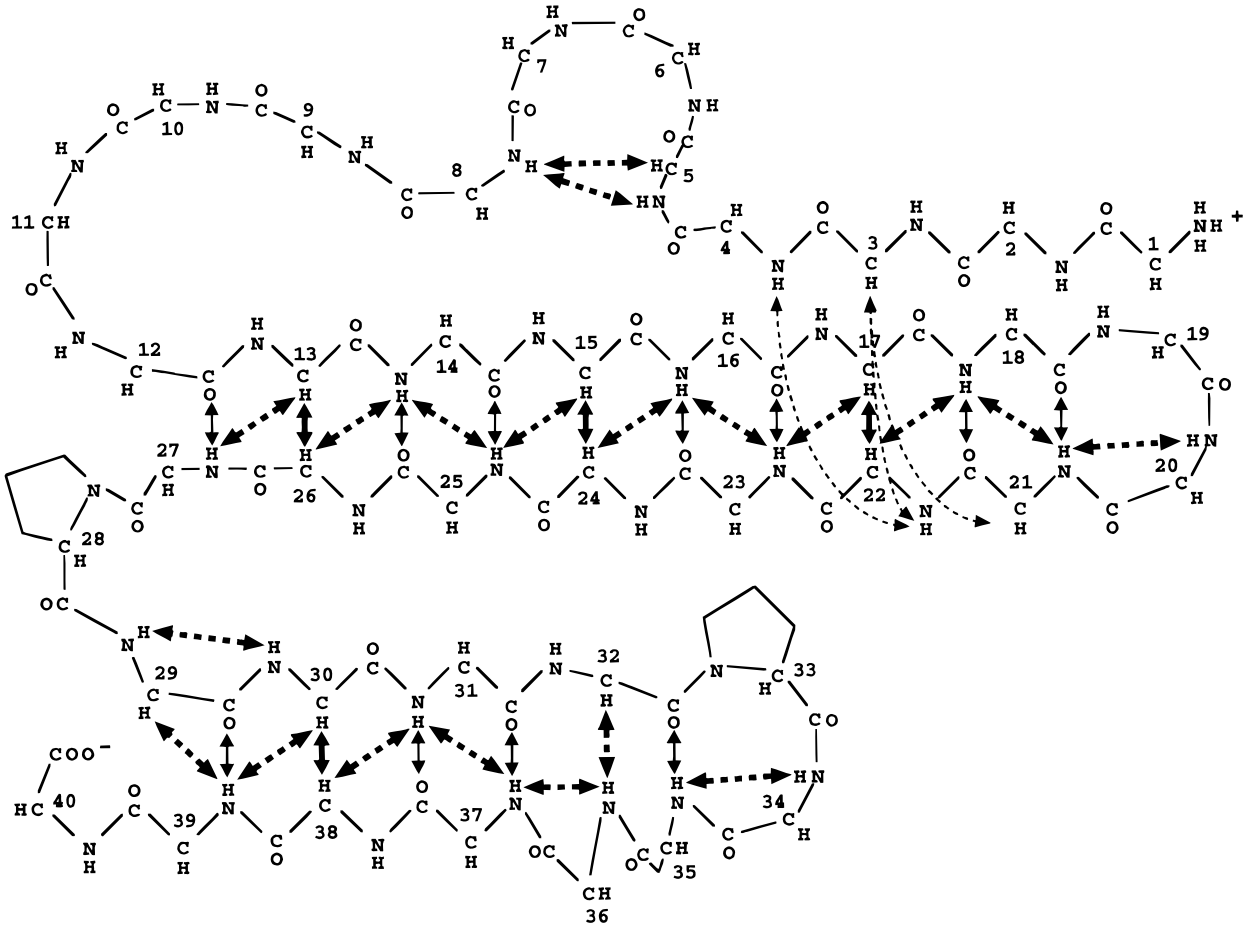

FIGURE 5: Antiparallel β-sheets in P-sel:EGF. The backbone atoms for the entire 40-residue sequence are shown. The arrows between atoms illustrate contacts derived from the experimental data that support the presence of antiparallel β-sheet structure from residues 12 to 27 and residues 29 to 39. Symbols: thick continuous arrows, $d_{\alpha\alpha(i,j)}$; thin continuous arrows, hydrogen bond contacts based on slowly exchanging amide protons in D_2O and structure analysis; thick broken arrows, $d_{\alpha N(i,j)}$ and $d_{NN(i,j)}$; thin broken arrows, $d_{\alpha N(i,j)}$ and $d_{NN(i,j)}$ between the amino terminus and the 12−27 β-sheet that suggests a triple-stranded sheet. Note that some arrows come from residues not in the sheets per se but are nearest neighbors. Similar types of contacts shown between residues 5 and 8 partly define their participation in a loop.

between residues 29 and 32 with residues 36−39. Some of the α-protons in these regions (residues 22, 24, 27, 30, and 31) are shifted downfield of 5 ppm, which is characteristic of antiparallel β-sheet (Wishart et al., 1991). Second, all the sequential αH−NH distances in these regions are short relative to the intraresidue αH−NH distances. Similarly, all measurable $^3J_{HN\alpha}$ in these regions are large (>7.5 Hz) including residues 14−16, 18, 23−25, 30, and 36−38. Third, several of the backbone amides are protected from exchange in the presence of D_2O, suggesting that they are stabilized by interstrand hydrogen bonds, including residues 14, 16, 18, 21, 23, 25, 27, 31, 37, and 39. Long-range backbone contacts also occurred between residues 3 and 4 with residues 21 and 22.

Structure calculations by distance geometry and simulated annealing methods employed 859 distance restraints of which 324 were derived from intraresidue and sequential backbone NOEs and 535 were derived from short-, medium-, and long-range NOEs (Figure 6). In addition, 14 ϕ and 29 χ_1 angle restraints, 10 stereospecific assignments, and 11 hydrogen bond restraints were used. The resulting structures are displayed in Figure 7. The pairwise RMSD to the average is 0.64 ± 0.17 Å for the well-defined backbone residues of 19 final structures and 1.10 ± 0.18 Å for all heavy atoms (non-hydrogen atoms) of those residues. The correlation and

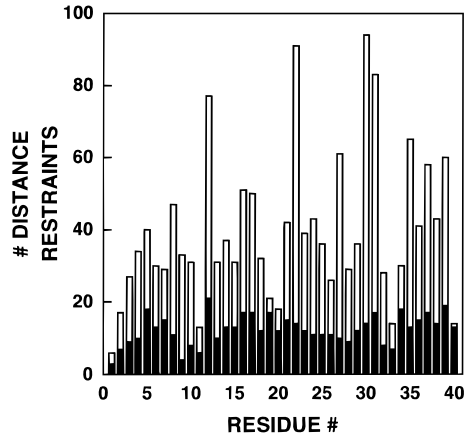

FIGURE 6: Number of distance restraints versus residue number. The top of each open bar represents the total number of restraints derived from the experimental data for each residue. The top of each closed bar represents the number of intraresidue and sequential backbone distances for each residue. The difference between the closed and open bars represents the number of short-, medium-, and long-range distances (as defined in Materials and Methods) for each residue.

mean ϕ and ψ angles for each residue were measured for the family of structures (Figure 8 and Table 1). The majority of backbone torsion angles had correlations near 100% for residues 12−38 with the exception of the ψ angle of Ile 19

FIGURE 7: Stereoview overlay of the 19 calculated backbone structures for the P-sel:EGF peptide structure. All 19 final structures are shown superimposed with the geometric average using the backbone atoms of the well-defined residues. The backbone root-mean-square deviation compared to the average is 0.64 ± 0.17 Å. The N- and C-termini are labeled.

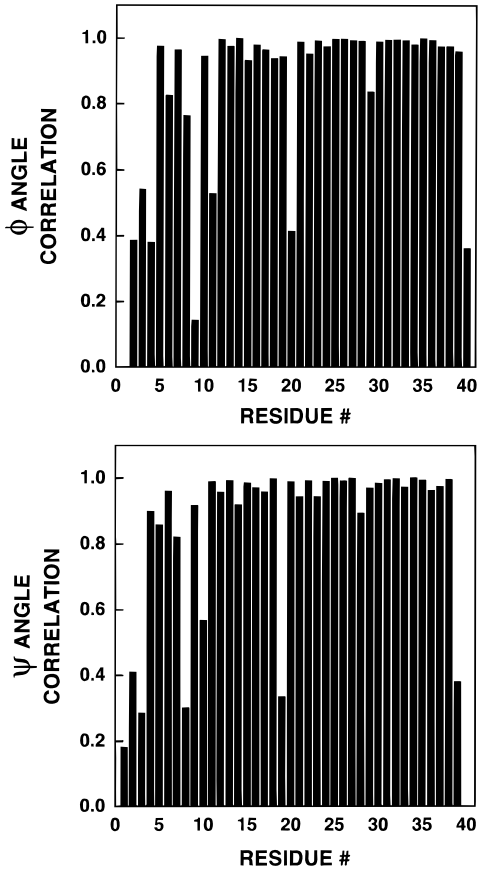

FIGURE 8: Torsion angle correlations for the P-sel:EGF peptide structure. The ϕ and ψ angle correlations for the family of 19 P-sel:EGF calculated structures are shown as a function of residue number. The correlation is a measure of the degree of agreement among structures for the ϕ or ψ angles of a particular residue. A value of 1 indicates a perfect correlation, whereas a value of 0 indicates disorder.

and the ϕ angle of Gly 20. Otherwise, the remaining residues have poorly defined backbones with the exception of well-defined torsion angles for residues 5−7, the ψ angles of residues 4, 9, and 11, and the ϕ angles of residues 8 and 39. A Ramachandran plot in Figure 9 shows the distribution of

Table 1: Average Backbone ϕ and ψ Angles[a]

residue	mean ϕ	mean ψ	residue	mean ϕ	mean ψ
Thr 1		−159	Asn 21	178	−178
Ala 2	−155	−105	Tyr 22	−136	173
Ser 3	−151	−152	Thr 23	−131	180
Cys 4	−60	112	Cys 24	−121	135
Gln 5	−123	−161	Ser 25	−105	82
Asp 6	−57	−54	Cys 26	−55	134
Met 7	−100	41	Tyr 27	−55	160
Ser 8	−32	168	Pro 28	−65	98
Cys 9	−2	54	Gly 29	135	−4
Ser 10	62	15	Phe 30	−131	161
Lys 11	60	39	Tyr 31	−161	174
Gln 12	−133	15	Gly 32	99	155
Gly 13	−176	164	Pro 33	−46	−30
Glu 14	−102	159	Glu 34	−120	20
Cys 15	−92	128	Cys 35	47	44
Leu 16	−114	81	Glu 36	−93	−6
Glu 17	−68	139	Tyr 37	−116	146
Thr 18	−125	−171	Val 38	−110	143
Ile 19	−58	8	Arg 39	−69	−167
Gly 20	179	33	Glu 40	26	

[a] The average ϕ and ψ angles for the 19 final structures are displayed. The information presented here should be considered together with the correlations presented in Figure 6.

ϕ and ψ angles for the family of structures, which indicates that the majority of the defined angles fall within the energetically favorable regions. Many of the ϕ/ψ angles are clustered in the negative ϕ/positive ψ angle region consistent with a β-sheet structure. Notably, Gly 29 and Gly 32 lie outside these regions since glycines can assume a larger number of energetically favorable conformations. Also, the ϕ/ψ angles of Cys 35 are found within the region typical of the unusual left-handed helix. Finally, correlation values measured for the χ_1 angles indicate that approximately 50% of the residues are ordered ($S > 0.9$) at least as far as the β-carbon.

DISCUSSION

Over the past 10 years EGF domains have received considerable attention due to their wide distribution in a variety of proteins (Campbell & Bork, 1993). In contrast

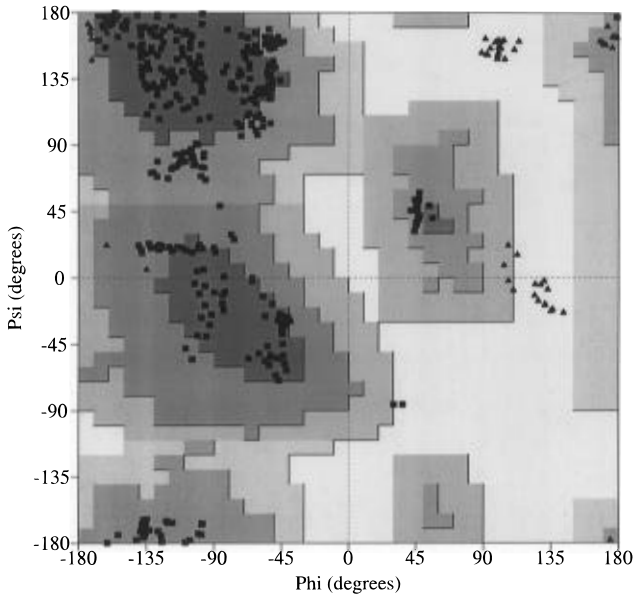

FIGURE 9: Ramachandran plot of the ϕ and ψ angles for all calculated structures. Only the well-defined ϕ/ψ angles are shown. The shaded regions of the plot are the energetically favorable ones, with the darkest regions being the most favored. All ϕ/ψ angles are shown as closed squares, except for the glycines which are shown as closed triangles. The majority of ϕ/ψ angles are found within the most favorable regions. Many are clustered in the negative ϕ/positive ψ angle region representing β-sheet-containing amino acids. Note that Gly 29, Gly 32, and Cys 35 have positive angles (see text for details). The plot was generated using the program Procheck (Laskowski et al., 1993).

to the growth factor subset of domains, other EGF domains have diverged in function. Some have become components of larger multidomain proteins with varied roles but are clearly crucial, however, to the function of the parent molecules. They appear to either serve a structural role and form interdomain interactions or contact other proteins directly (Campbell & Bork, 1993). Several groups have identified critical amino acids and the structure for EGF domain function (Dudgeon et al., 1990; Campbell et al., 1990; Campbell & Cooke, 1990), but few such studies have been applied to selectin EGF domains. The structure of the E-selectin lectin and EGF domains determined by crystallography has been reported (Graves et al., 1994), but little is known about the contribution of the E-selectin EGF domain to cell adhesion. In this report we have examined the P-selectin EGF domain in order to study its structure and function independent of the lectin domain.

In the presence of calcium, the selectins bind a counter-receptor, mucin-like glycoproteins containing sialyl Lewis x (Lasky, 1995; McEver et al., 1995). The selectins require both the lectin and EGF domains for optimal adhesion (Li et al., 1994; Gibson et al., 1995). Indeed, the EGF domain may have a direct role in ligand binding; when the EGF domain of P-selectin is substituted into L-selectin, the chimera has P-selectin ligand binding specificity (Kansas et al., 1994). Also, peptide 127−139 derived from the P-selectin EGF domain blocked adhesion (Murphy & McGregor, 1994). Monoclonal antibodies directed against the L-selectin EGF domain block adhesion, thereby also implicating the EGF in a direct role in cell recognition (Spertini et al., 1991; Kansas et al., 1991). The E-selectin lectin−EGF domain structure shows few interdomain contacts (Graves et al., 1995). On the other hand, the EGF domain may play an indirect role since epitopes in the lectin domain of L-selectin and E-selectin are only expressed when the EGF domain is present (Walz et al., 1990; Bowen et al., 1990). In addition, antibodies that map to the lectin domain or EGF domain prevented the binding of small glycosylated L-selectin ligand peptides to L-selectin, thereby suggesting that the EGF domain may alter the structure of the lectin domain since both epitopes would not likely interact with the small ligands simultaneously (Imai et al., 1992). In this report, we could not demonstrate any binding of the P-selectin EGF domain to PSGL-1 on HL60 cells. This result is in contrast to the report of the inhibitory activity of fragment 127−139 derived from the P-selectin EGF domain (Murphy & McGregor, 1994). Our results can be interpreted in two ways. Either the EGF domain requires the lectin domain for its functional conformation or the EGF domain influences the structure of the lectin domain to support ligand binding.

EGF domain structures have been extensively studied using ^{1}H NMR (Harvey et al., 1991; Montelione et al., 1992; Baron et al., 1992; Hommel et al., 1992; Selander-Sunnerhagen et al., 1992; Jacobsen et al., 1996). We observed that some amide and side chain proton resonances were broad over a range of pH and temperature conditions. The structure was determined at ambient temperature as a compromise between slower molecular rotation at lower temperatures and the loss of amide proton intensity due to more rapid exchange at higher temperatures. Furthermore, the structure was determined at pH 6.0 because this was the minimum pH where the amide proton resonances of Glu 34 and Glu 36 were shifted furthest downfield and had the narrowest line widths. This was also a compromise between loss of amide proton intensity at higher pH values, and the broadening of the glutamate amide proton resonance line widths at lower pH values. The resonance line widths and positions of other amide and side chain protons in P-sel:EGF do not appear to be influenced by pH, but this conclusion would require a detailed study using 2D spectra recorded at different pH values. In contrast, work on other EGF domains has shown that resonances may broaden as the pH is increased *or* decreased and may reflect a conformational change (Tappin et al., 1989; Handford et al., 1990). On the other hand, the structures of murine EGF at pH 2.0 and pH 6.8 were identical for the region conserved with P-sel:EGF (Kohda & Inagaki, 1992). We may be observing an intermediate type of chemical exchange on the NMR time scale, but we have no independent evidence for this interpretation. Line broadening leads to weaker or absent NOE intensity which, in turn, results in fewer NOEs per residue. Furthermore, the line broadening makes torsion angle measurement impossible for the broadest residues. Specific secondary structures in this region are defined more on the basis of NOEs than measured torsion angles. Thus the amino-teminal 11 residues are less well determined than the rest of the molecule.

The conserved Glu 36 amide proton resonates downfield of 10 ppm in other EGF domain NMR spectra, including, for example, factor IX, factor X, TGF-α, and human and murine EGF (Baron et al., 1992; Tappin et al., 1989; Montelione et al., 1988; Selander-Sunnerhagen et al., 1992). In addition, a similarly sized pH-dependent shift of the Arg 42 and Glu 44 amide proton resonances of TGF-α (Glu 34 and Glu 36 in P-sel:EGF) has been reported (Tappin et al., 1989). Also, the amide proton resonances of Glu 34 and Glu 36 tend to be moderately broad in factor IX at pH 4.5

(Baron et al., 1992). In E-selectin, the Lys 34 and Glu 36 amides form hydrogen bonds with the Glu 36 terminal carboxyl group (Graves et al., 1994). Although our structures are not defined sufficiently to interpret hydrogen bonding for the side chain of Glu 36, we do have NMR evidence to suggest that similar hydrogen bonding occurs. First, the amide proton of Glu 36 is slowly exchanging and present at 24 h in D_2O. Second, the amide protons of Glu 34 and Glu 36 are in close proximity to the Glu 36 carboxyl group since NOEs were found between the two residues, and the Glu 36 γH−NH NOEs were unusually strong, more so than the βH−NH NOEs. We propose that the Glu 34 and Glu 36 amide protons are shifted downfield at high pH by the Glu 36 carboxyl group with which they form hydrogen bonds and that the pH-dependent shifts as the hydrogen ion concentration is increased are due to titration of the carboxyl group at its pK_a. Since the amide proton of Glu 34 disappears rapidly in D_2O, and the pH-dependent transition occurs over a smaller change in acidity than that of Glu 36, the Glu 34−Glu 36 hydrogen bond may be less stable than the intraresidue one. The structural result of the intraresidue hydrogen bond is to produce a β-bulge at Glu 36 (see below) since the amide proton is not available as a hydrogen-bonding partner in the regular β-sheet pattern.

The solution structure of P-sel:EGF has the major structural features of all EGF domains. The global fold is a compact globular structure with an unusual amount of hydrophilic character (McCaldon & Argos, 1988). The disulfide structure consists of three disulfide bonds that project upward from one surface of the molecule and form an inverted tripod-like structure. The well-ordered region of the molecule consists of two segments of antiparallel β-sheet from residues 12 to 27 and from residues 29 to 39 as predicted from the spectral analysis. The β-sheets are analogous to those in other EGF molecules and are referred to using the same nomenclature. The major β-sheet undergoes an appreciable right-handed twist in contrast to the minor one. The major and minor sheets are separated by a β-turn comprising residues 27−30, which most resembles a type II turn. The consensus angles for a type II turn (ϕ_2 ~−60°, ψ_2 ~120°, ϕ_3 ~80°, ψ_3 ~0°) are met for ϕ_2 and ψ_3, but deviate by approximately 20° and 55° for the ψ_2 and ϕ_3 angles, respectively (Richardson, 1981); however, they are not excluded based on the lower correlations for these angles. Glu 36 interrupts the minor β-sheet producing a β-bulge. β-Bulges have also been described for other EGF molecules at this residue including the first EGF domain of factor IX (Baron et al., 1992). The β-strands of the major and minor sheets are separated by turns from residues 18 to 21 and from residues 32 to 35, respectively. Within the turns the amide protons of Asn 21 and Cys 35 are slowly exchanging and form hydrogen bonds with the carbonyls of Thr 18 and Gly 32, respectively. The 18−21 turn does not fit any classical type, but the ψ angle of Ile 19 and the ϕ angle of Gly 20 are less defined. The 32−35 turn most closely approximates a type VIII turn (consensus angles, ϕ_2 ~−60°, ψ_2 ~−30°, ϕ_3 ~−120°, ψ_3 ~120°) for the ϕ_2, ψ_2, and ϕ_3 angles but not for ψ_3 angle, which is a 100° difference from the canonical (20°). The ϕ angles of the conserved Gly 32 and Cys 35 residues in this turn are positive. The major and minor sheets are oriented relative to each other by intersubdomain contacts between Gln 12, Cys 24, Cys 26, and Tyr 27 with Phe 30, Tyr 31, and Cys 35. Residues 1−13 are less ordered but consist of two turns from residues 5 to 8 and from residues 8 to 13. The 5−8 turn best ressembles a type I turn (consensus angles, ϕ_2 ~−60°, ψ_2 ~−30°, ϕ_3 ~−90°, ψ_3 ~0°), noting that the ψ_2 and ψ_3 angles deviate by ~20° and ~40°, respectively. The 8−13 turn is a large, less-defined turn. The regions from residues 1 to 4 and from residues 8 to 10 project above the major β-sheet, making contacts with Cys 15, Glu 17, Asn 21, Tyr 22 and Gly 13, Cys 15, Cys 35, respectively. Residues 1−4 may be partially forming the third strand of a triple-stranded β-sheet based on some of these contacts with the twisted major β-sheet. In particular, interstrand-like contacts occur between the α-protons of Ser 3 and Asn 21, between the α-proton of Ser 3 and the amide proton of Tyr 22, and between the amide protons of Cys 4 and Tyr 22. Other long-range NOEs from the N-terminus occurred between Cys 4 and Glu 17 and between Thr 1 and Ala 2 with Glu 17 and Tyr 22.

The P-sel:EGF structure is very similar to the analogous region of the crystal structure of the E-selectin lectin−EGF domains (average RMSD of the P-sel:EGF-E-sel:EGF structures is 1.08 ± 0.18 Å for the well-defined backbone residues) (Graves et al., 1994) (Figure 10). Moreover, the regions of defined secondary structure from residues 5 to 38 are nearly identical, including the regions of β-sheet and the approximate β-turns (the residue 5−8 type I turn, the residue 27−30 type II turn, and the residue 32−35 type VIII turn). The large deviations from consensus turn angles occur in both structures, including the 100° difference in the ψ_3 angle of the type VIII turn. In contrast to the less-defined turn from residues 8 to 13 in P-sel:EGF (consisting of the residues in which resonance broadening was greatest), the analogous structure in E-sel:EGF is a left-handed helical twist. This surface turn is hydrophilic in both structures. In sum, it is apparent that the strict disulfide bonding pattern imposed on these domains allows some sequence divergence but little structural change for the conserved β-sheet-containing portions of the protein. The P-sel:EGF structure is also similar to the factor IX and factor X EGF structures (RMSD between 1 and 2 Å) (Baron et al., 1992; Selander-Sunnerhagen et al., 1992). Differences in structure for TGF-α are mostly the result of the insertion of a residue between cysteines (Hommel et al., 1992; Montellione et al., 1992; Harvey et al., 1991). Disparity in secondary structure among related EGF molecules occurs primarily in the amino-terminal 11 residues in which small turns, large turns (Ω loops), and helices have all been described (Baron et al., 1992; Selander-Sunnerhagen et al., 1992; Hommel et al., 1992; Montellione et al., 1992; Harvey et al., 1991; Graves et al., 1994). The residues preceding the first cysteine are the most variable in that they form the third strand of a triple-stranded β-sheet in TGF-α and murine and human EGF (Harvey et al., 1991; Cooke et al., 1987; Montellione et al., 1992). However, these groups acknowledge that the interstrand NOEs are weak for such interactions and amide protons are not protected from exchange in D_2O. Thus, the amino terminus *transiently* forms the third strand of a β-sheet structure in these molecules. In contrast, the human EGF structure of Hommel et al. (1992) had an unstructured amino terminus—no non-intraresidue, nonsequential contacts were observed in NOESY spectra prior to the first cysteine. Similarly, the Ca^{2+}-free forms of the first EGF domains of factor IX and factor X are unstructured prior to the first

FIGURE 10: Overlay of the P-sel:EGF NMR structure with the analogous region of the E-sel:lectin−EGF crystal structure. P-sel:EGF is shown in black and E-sel:EGF in gray. The lowest energy P-sel:EGF structure was used for superimposition. The backbone structures are similar for the well-defined residues (RMSD is approximately 1.1 Å). The primary differences illustrated are at the amino terminus and the turn from residues 18−21, which are not well-defined in the P-sel:EGF structure.

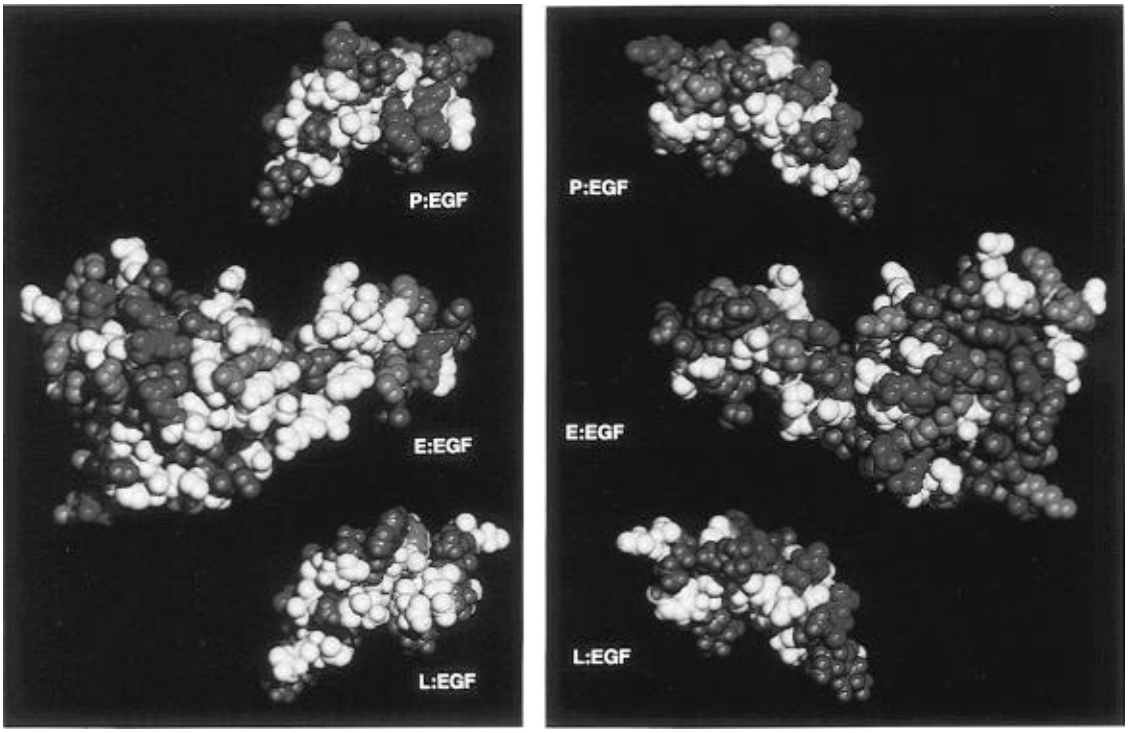

FIGURE 11: Comparison of the EGF domains of the selectins. The E-selectin lectin−EGF domains from the X-ray coordinates (Graves et al., 1994) are displayed in the center. The P-selectin EGF domain structure, based on the current work, and the L-selectin EGF domain, based upon construction of a homology model, are shown above and below the E-selectin EGF domain. Coloring denotes functional properties of the residues: red, positively charged; blue, negatively charged; white, hydrophilic; dark gray, hydrophobic. Code: P:EGF, P-selectin EGF domain; E:EGF, E-selectin EGF domain; L:EGF, L-selectin EGF domain. Panels: A (left), front view; B (right), back view.

cysteine (Baron et al., 1992; Ullner et al., 1992). The Ca^{2+}-bound forms do exhibit long-range contacts to the major β-sheet (Selander-Sunnerhagen et al., 1992; Rao et al., 1995); these amino-terminal structures form part of a Ca^{2+}-binding site and are stabilized in this manner. The P-sel:EGF structure appears to be a variant of these other EGF domains

with regard to the amino terminus. In P-sel:EGF peptide spectra, the amino terminus, including Thr 1, Ala 2, Ser 3, *and* the first cysteine, makes intermediate intensity NOEs with Cys 15, Glu 17, Asn 21, and Tyr 22 of the major β-sheet. Amide hydrogens in these contacts are not protected from exchange, however, in the presence of D_2O. In E-selectin, the amino terminus forms a random coil junction between the lectin and EGF domains. None of the long-range contacts observed between the amino terminus of P-sel:EGF and the major β-sheet are short distances in the E-sel:EGF domain (>5 Å apart). Since most of the residues involved are conserved between P- and E-selectin, the structural difference is attributable to the presence of the lectin domain in E-sel:lectin-EGF. It remains possible that the disorder of the N-terminus of the P-selectin EGF domain is a consequence of the analysis of a free EGF peptide. In the intact P-selectin molecule, the N-terminal region of the EGF domain may be structured because, when covalently attached to the lectin domain, the EGF domain is constrained.

Direct proof of the importance of the P-selectin EGF domain in leukocyte recognition is based on two independent observations. Kansas et al. (1994) demonstrated that COS cells overexpressing a selectin chimera composed of L-selectin with the P-selectin EGF domain replacing the L-selectin EGF domain bind to HL60 cells. COS cells expressing L-selectin do not bind to HL60 cells, in contrast to those cells expressing P-selectin. These results suggest a role for the P-selectin EGF domain in leukocyte recognition, at least at the supraphysiologic selectin cell density realized in these experiments. Furthermore, Gibson et al. (1995) demonstrated that the lectin and EGF domains together represent the optimal recognition unit for the P-selectin ligand, PSGL-1, when P-selectin chimeras are expressed on cell surfaces at physiologic density. On the basis of the important role of the EGF domain in cell recognition, we have analyzed models of the selectin EGF domains to identify structural correlates to these functional properties. The X-ray structure of the E-selectin lectin−EGF domains (Graves et al., 1994) and the 2D NMR structure of the P-selectin EGF domain provide a starting point for this analysis. Given the identical size and the marked sequence homology, including 23 identical homologous amino acids in P-selectin and L-selectin, we prepared an homology model of the L-selectin EGF domain based upon the P-selectin EGF domain three-dimensional structure. The EGF domains of P-selectin, E-selectin, and L-selectin demonstrate marked differences in electrostatic charge distribution despite the common polypeptide backbone structure (Figure 11). It is particularly striking that the EGF domain of P-selectin has two prominent glutamic acid residues on the surface, Glu 34 and Glu 36, while L-selectin has no charged groups on this surface (Figure 11A). Differences such as these are likely responsible, at least in part, for the specificity of binding to the selectin counterreceptors.

ACKNOWLEDGMENT

We thank Margaret Jacobs for synthesizing the P-sel:EGF peptide and B. J. Graves for providing the coordinates of the E-selectin lectin−EGF domain shortly after publication.

SUPPORTING INFORMATION AVAILABLE

Table S-1 containing the proton resonance assignments for the P-sel:EGF peptide (1 page). Ordering information is given on any current masthead page.

REFERENCES

Appella, E., Weber, I. T., & Blasi, F. (1988) *FEBS Lett. 231*, 1−4.
Baron, M., Norman, D. G., Harvey, T. S., Handford, P. A., Mayhew, M., Tse, A. G. D., Brownlee, G. G., & Campbell, I. D. (1992) *Protein Sci. 1*, 81−90.
Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
Campbell, I. D., & Cooke, R. M. (1990) *J. Cell Sci., Suppl. 13*, 5−10.
Campbell, I. D., & Bork, P. (1993) *Curr. Opin. Struct. Biol. 3*, 385−392.
Campbell, I. D., Baron, M., Cooke, R. M., Dudgeon, T. J., Fallon, A., Harvey, T. S., & Tappin, M. J. (1990) *Biochem. Pharmacol. 40*, 35−40.
Cooke, R. M., Wilkinson, J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H., & Sheard, B. (1987) *Nature 327*, 339−341.
Cooke, R. M., Tappin, M. J., Campbell, I. D., Kohda, D., Miyake, T., Fuwa, T., Miyazawa, T., & Inagaki, F. (1990) *Eur. J. Biochem. 193*, 807−815.
Detlefsen, D. J., Thanabal, V., Pecoraro, V. L., & Wagner, G. (1991) *Biochemistry 30*, 9040−9046.
Driscoll, P. C., Clore, G. M., Beress, L., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2178−2187.
Dudgeon, T. J., Cooke, R. M., Baron, M., Campbell, I. D., Edwards, R. M., & Fallon, A. (1990) *FEBS Lett. 261*, 392−396.
Dudgeon, T. J., Hommel, U., Cooke, R. M., Fallon, A., & Campbell, I. D. (1991) *Techniques in Protein Chemistry II*, pp 251−261, Academic Press, Inc., New York.
Freedman, S. J., Furie, B. C., Furie, B., & Baleja, J. D. (1995) *Biochemistry 34*, 12126−12137.
Gibson, R. M., Kansas, G. S., Tedder, T. F., Furie, B., & Furie, B. C. (1995) *Blood 85*, 151−158.
Graves, B. J., Crowther, R. L., Chandran, C., Rumberger, J. M., Li, S., Huang, K.-S., Presky, D. H., Familletti, P. C., Wolitzky, B. A., & Burns, D. K. (1994) *Nature 367*, 532−538.
Groenen, L. C., Nice, E. C., & Burgess, A. W. (1994) *Growth Factors 11*, 235−257.
Handford, P. A., Baron, M., Mayhew, M., Willis, Beesley, T., Brownlee, G. G., & Campbell, I. D. (1990) *EMBO J. 9*, 475−480.
Handford, P. A., Mayhew, M., Baron, M., Winship, P. R., Campbell, I. D., & Brownlee, G. G. (1991) *Nature 351*, 164−167.
Harvey, T. S., Wilkinson, A. J., Tappin, A. J., Cooke, R. M., & Campbell, I. D. (1991) *Eur. J. Biochem. 198*, 555−562.
Havel, T. F. (1991) *Prog. Biophys. Mol. Biol. 56*, 43−78.
Heavner, G. A., Falcone, M., Kruszynski, M., Epps, L., Mervic M., Riexinger, D., & McEver, R. P. (1993) *Int. J. Pept. Protein Res. 42*, 484−489.
Hommel, U., Harvey, T. S., Driscoll, P. C., & Campbell, I. D. (1992) *J. Mol. Biol. 227*, 271−282.
Huang, L. H., Ke, X.-H., Sweeney, W., & Tam, J. P. (1989) *Biochem. Biophys. Res. Commun. 160*, 133−139.
Hyberts, G., Märki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625−635.
Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.
Imai, Y., Lasky, L. A., & Rosen, S. D. (1992) *Glycobiology 2*, 373−381.
Jacobs, M., Freedman, S. J., Furie, B. C., & Furie, B. (1994) *J. Biol. Chem. 269*, 25494−25501.
Jacobsen, N. E., Abadi, N., Sliwkowski, M. X., Reilly, D., Skelton, N. J., & Fairbrother, W. J. (1996) *Biochemistry 35*, 3402−3417.
Johnston, G. I., Cook, R. G., & McEver, R. P. (1989) *Cell 56*, 1033−1044.
Kansas, G. S., Spertini, O., Stoolman, L. M., & Tedder, T. F. (1991) *J. Cell Biol. 114*, 351−358.

Kansas, G. S., Saunders, K. B., Ley, K., Zakrzewicz, A., Gibson, R. M., Furie, B. C., Furie, B., & Tedder, T. F. (1994) *J. Cell Biol. 124*, 609−618.
Kohda, D., & Inagaki, F. (1992) *Biochemistry 31*, 11928−11939.
Larsen, G., Sako, D., Ahern, T., Shaffer, M., Erban, J., Sajer, S. A., Gibson, R. M., Wagner, D. D., Furie, B. C., & Furie, B. (1992) *J. Biol. Chem. 266*, 11104−11110.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Lasky, L. A. (1995) *Annu. Rev. Biochem. 64*, 113−139.
Li, S. H., Burns, D. K., Rumberger, J. M., Presky, D. H., Wilkinson, V. L., Anostario, M., Jr., Wolitzky, B. A., Norton, C. R., Familletti, P. C., Kim, K. J., et al. (1994) *J. Biol. Chem. 269*, 4431−4437.
McCaldon, P., & Argos, P. (1988) *Proteins 4*, 99−122.
McEver, R. P., Moore, K. L., & Cummmings, R. D. (1995) *J. Biol. Chem. 270*, 11025−11028.
Montelione, G. T., Wüthrich, K., & Scheraga, H. A. (1988) *Biochemistry 27*, 2235−2243.
Montelione, G. T., Wüthrich, K., Burgess, A. W., Nice, E. C., Wagner, G., Gibson, K. D., & Scheraga, H. A. (1992) *Biochemistry 31*, 236−249.
Murphy, J. F., & McGregor, J. L. (1994) *Biochem. J. 303*, 619−624.
Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741−751.
Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.
Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G. G., & Stuart, D. (1995) *Cell 82*, 131−141.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
Selander-Sunnerhagen, M., Ullner, M., Persson, M., Teleman, O., Stenflo, J., & Drakenberg, T. (1992) *J. Biol. Chem. 267*, 19642−19649.
Spertini, O., Kansas, G. S., Reimann, K. A., Mackay, C. R., & Tedder, T. F. (1991) *J. Immunol. 147*, 942−949.
Tappin, M. J., Cooke, R. M., Fitton, J. E., & Campbell, I. D. (1989) *Eur. J. Biochem. 179*, 629−637.
Ullner, M., Selander, M., Persson, E., Stenflo, J., Drakenberg, T., & Teleman, O. (1992) *Biochemistry 31*, 5974−5983.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York.
Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311−319.

BI9610257