Design and Chemical Synthesis of a Neoprotein Structural Model for the Cytoplasmic Domain of a Multisubunit Cell-Surface Receptor: Integrin $\alpha_{IIb}\beta_3$ (Platelet GPIIb-IIIa)†

Tom W. Muir,* Michael J. Williams,* Mark H. Ginsberg, and Stephen B. H. Kent

The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037

Received December 10, 1993; Revised Manuscript Received April 4, 1994®

ABSTRACT: Integrins are a class of heterodimeric cell adhesion receptors involved in cell migration, cell anchorage, and cell–cell interactions. The cytoplasmic domains of integrins are of key importance in these activities. We have designed and chemically synthesized a 126 amino acid model protein (MP-1) containing both cytoplasmic tails of the platelet-derived integrin $\alpha_{IIb}\beta_3$ covalently linked via a helical coiled coil. The coiled-coil tertiary structure was incorporated to mimic the membrane-spanning domain of the integrin and to act as a topological constraint fixing the two cytoplasmic tails in a parallel arrangement. This molecule, which contains two C-termini, was constructed by chemical dovetailing. The bromoacetylated and cysteinyl peptide synthons were unambiguously ligated through the formation of a thioether linkage. Ultraviolet circular dichroism (CD) spectroscopy has been performed on MP-1 and related compounds, confirming that a helical coiled coil is present within the MP-1 molecule. Significantly, the helicity apparently extends beyond the predicted amphiphilic region of MP-1. Fluorescence measurements suggest that a defined tertiary structure has formed by the association of the two cytoplasmic domains. We conclude that this is a practical design strategy for the study of the cytoplasmic domain of multisubunit cell-surface receptors.

Many integral membrane proteins transmit signals across the plasma membrane. Technical difficulties have greatly limited the application of high-resolution techniques to the structures of these proteins. Indeed molecular structures are available for only two intact transmembrane proteins, a bacterial photoreaction center (Deisenhofer et al., 1985) and a porin (Weiss et al., 1990). Structures of receptor extracellular domains have been determined using soluble truncated extracellular domains as models (De Vos et al., 1992; Milburn et al., 1991). These structures have contributed greatly to our understanding of the basis of ligand recognition, but have provided less insight into the mechanism of signal transduction. Many membrane proteins that transduce signals are members of the type I transmembrane protein family, the defining feature of which is a single membrane spanning region. These include the T cell receptor (Weiss, 1993), growth factor receptors (Patthy, 1990), and cytokine receptors (Miyajima et al., 1992). In general the cytoplasmic domain of these proteins is critical for signaling. Thus, to understand signal transduction through such receptors, it is essential to understand the structure and function of the cytoplasmic domain. This is especially difficult for multisubunit type I receptors. To address this issue we have developed a novel strategy for the study of the cytoplasmic domain of multisubunit receptors and applied this to the study of a member of the integrin family.

Integrins are heterodimeric transmembrane receptors found in almost all cell types [for review, see Hynes (1992)]. Each subunit (α or β) has a single membrane spanning domain and a cytoplasmic extension. Integrin heterodimers are centrally involved in many processes involving bidirectional transmission of signals across the plasma membrane (Hynes, 1992). One of the most intensively studied members of the integrin family is the platelet-derived integrin $\alpha_{IIb}\beta_3$. This integrin has been reported to function through both outside-in and inside-out signaling, and its cytoplasmic domain has been implicated in both processes (Ginsberg et al., 1992; Sastry & Horwitz, 1993). Structure/function studies on an *in vitro* model of the cytoplasmic domain of $\alpha_{IIb}\beta_3$ could provide a more cohesive understanding of signal transduction through this region of the protein.

The inherent hydrophobicity of membrane-spanning stretches of apolar amino acids means that structural studies on truncated receptors containing these helical regions will be impeded by insolubility in aqueous solution. In the present study, we introduce a new approach to overcome this solubility problem by replacing the hydrophobic membrane-spanning sequence with a highly soluble amphiphilic peptide designed to mimic the helical secondary structure of the native membrane-spanning domain. The membrane-spanning segments are linked together covalently in the correct orientation by means of chemical dovetailing. This novel strategy, which makes use of advances in both the *de novo* design and the chemical synthesis of proteins, has been used to construct a model protein designed to mimic the structural properties of the cytoplasmic domains of the prototypical platelet integrin $\alpha_{IIb}\beta_3$. In this paper we describe a general strategy for the design and chemical synthesis of synthetic model proteins which will allow structural and functional information to be obtained on the cytoplasmic domains of multichain receptors.

EXPERIMENTAL PROCEDURES

Materials. Boc-amino acids[1] and HBTU were obtained from Novabiochem (San Diego, CA). Preloaded Boc-amino-acyl-OCH_2-Pam-resins and diisopropylethylamine (DIEA) were purchased from Applied Biosystems (Foster City, CA). 4-Methylbenzhydrylamine-resin (Lot No. 023863) was purchased from Peninsula Labs (San Carlos, CA). *N,N*-Dimethylformamide was purchased from Mallinckrodt Chemical

† This work was supported by a grant from the Markey Foundation (S.B.H.K.) and by NIH Grants HL48728 and AR27214. T.W.M was supported by a research fellowship from Amgen.

* To whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* June 1, 1994.

0006-2960/94/0433-7701$04.50/0

Co. HPLC-grade acetonitrile was purchased from EM Science (Gibbstown, NJ). All other reagents were AR grade from Aldrich Chemical Co.

Peptide Synthesis. All peptides were synthesized manually according to our *in situ* neutralization/HBTU activation protocol for Boc solid-phase chemistry as described previously (Schnölzer et al., 1992b). Coupling yields were monitored by the quantitative ninhydrin determination of residual free amine (Sarin et al., 1981). Peptide α-carboxamides were constructed on a 4-methylbenzhydrylamine-resin, and all other peptides were synthesized on appropriate Boc-aminoacyl-OCH_2-Pam-resins. Where required, the bromoacetyl group was introduced at the N^{α}-terminal of a peptide by coupling as the preformed symmetric anhydride. In all cases, side-chain protecting groups were removed and the peptides cleaved from the resin by treatment with liquid HF containing 4% *p*-anisole, for 1 h at 0 °C. Crude peptide products were precipitated and washed with diethyl ether before being dissolved in aqueous acetic acid (10–30%) and lyophilized.

Reverse-Phase HPLC. Analytical and semipreparative gradient HPLC were performed on a Rainin dual pump high-pressure mixing system with 214-nm detection. Semi-preparative HPLC was run on a Vydac C_{18} column (10 μm, 10 × 250 mm) at a flow rate of 3 mL/min. Analytical HPLC was performed on a Whatman C_{18} column (5 μm, 4.0 × 140 mm) at a flow rate of 1 mL/min. Preparative HPLC was performed on a Waters Prep 4000 system fitted with a Waters 486 tunable absorbance detector. Preparative HPLC was run on a Vydac C_{18} column (15–20 μm, 50 × 250 mm) at a flow rate of 30 mL/min. All runs used linear gradients of 90% acetonitrile plus 0.1% TFA versus 0.1% aqueous TFA.

Peptide Purification and Characterization. Crude peptides were dissolved in aqueous acetonitrile containing 0.1% TFA and purified by either semipreparative or preparative HPLC. Peptides containing cysteine residues were dissolved in HPLC buffers containing 5 mM dithiothreitol. All purified peptides were characterized by ion-spray mass spectrometry (see Table 1).

Mass Spectrometry. Ion-spray mass spectrometry of crude and purified peptide segments was performed on an API-III quadrupole ion-spray instrument (Sciex, Toronto, Ontario) as described (Schnölzer 1992a). Expected masses of peptide targets were determined using MacProMass (Terri Lee, City of Hope, Duarte, CA).

Synthesis of Model Protein MP-1. Ligation was initiated by combining Cys-helix-α_{IIb} (5.2 mg, 0.9 μmol) and BrAc-helix-β_3 (5.5 mg, 0.6 μmol) in 1.1 mL of 95% DMF/5% 0.1 M sodium phosphate, pH 7.0 containing 8 M urea at 25 °C. The ligation reaction was monitored by reverse-phase HPLC (20–50% CH_3CN over 30 min) and by ion-spray mass spectrometry. The ligation reaction was terminated after 12 h and the product purified by preparative HPLC using a gradient of 30–60% CH_3CN over 60 min. The lyophilized purified product was characterized by ion-spray mass spectrometry [calcd 14 189.3 (monoisotopic), 14 198.9 (average isotope composition), found 14 194.2 ± 2.5].

Synthesis of Model Protein MP-2. Cys-α_{IIb} (8.01 mg, 3.16 μmol) and BrAc-β_3 (11.78 mg, 2.1 μmol) were combined in 1.98 mL of 95% DMF/0.1 M sodium phosphate, pH 7.0 at 25 °C. The ligation reaction was monitored by reverse-phase HPLC using a gradient of 20–50% CH_3CN over 30 min. Peaks were collected and analyzed by ion-spray mass spectrometry. The ligation reaction was terminated after 18 h and the product purified by preparative HPLC using a gradient of 20–50% CH_3CN over 60 min. The lyophilized purified product was characterized by ion-spray mass spectrometry [calcd 8151.1 (monoisotopic), 8156.9 (average isotope composition), found 8151.4 ± 1.0].

Table 1: Mass Spectrometric Characterization of Synthetic Peptides

peptide	calcd (monoisotopic)	calcd (av isotope composition)	found
αIIb	2434.2	2435.6	2434.0 ± 0.5
Cys-αIIb	2537.2	2538.7	2537.0 ● 0.9
helix-αIIb	5452.8	5456.1	5454.3 ● 1.1
Cys-helix-αIIb	5555.8	5559.2	5558.2 ± 1.4
β_3	5572.5	5577.2	5576.1 ± 0.9
BrAc-β_3	5693.8	5698.2	5697.3 ± 1.4
BrAc-helix-β_3	8713.5	8718.8	8715.4 ± 1.6
BrAc-helix	3157.6	3160.4	3158.3 ± 0.6
Cys-helix	3140.7	3142.6	3141.2 ± 0.5

Synthesis of Helix Dimer. Ligation was carried out by combining BrAc-[G·(K·L·E·A·L·E·G)$_4$]-NH_2 (16.2 mg, 5.3 μmol) and Cys-[G·(K·L·E·A·L·E·G)$_4$]-NH_2 (13.5 mg, 4.3 μmol) in 1.5 mL of 95% DMF/0.1 M sodium phosphate, pH 7.0 at 25 °C. The ligation reaction was monitored by reverse-phase HPLC (5-μL aliquots) using a gradient of 35–50% CH_3CN over 30 min. Peaks were collected and analyzed by ion-spray mass spectrometry. The ligation reaction was thus observed to be complete after 2.5 h and the product purified by semipreparative HPLC using a gradient of 35–50% CH_3CN over 30 min. The lyophilized purified product was characterized by ion-spray mass spectrometry [calcd 6219.0 (monoisotopic), 6223.0 (average isotope composition), found 6219.4 ± 0.8].

Ultraviolet Circular Dichroism Spectroscopy. Far-UV CD spectra were recorded on an AVIV 60DS spectropolarimeter linked to an AT&T computer. CD spectra are presented as a plot of mean molar ellipticity per residue ([θ], deg cm^2 dmol^{-1}) versus wavelength in 0.5-nm increments. The digitized data was plotted using the Cricket Graph program on a Macintosh IIsi computer. All peptide and protein samples were dissolved in 50 mM boric acid at pH 7.0 and their concentrations determined by quantitative amino acid analysis.

Calculation of Protein Helicity. The percentage of helical secondary structure within a sample was estimated using eq 1. The maximal theoretical ellipticity at 222 nm $[\theta]_{max}$ was determined using relationship 2 where n is the number of residues per chain (Chen et al., 1974).

$$\% \text{ helix} = [\theta]_{222}/[\theta]_{max} \times 100 \quad (1)$$

$$[\theta]_{max} = -39500[1 - (2.57/n)] \quad ([\theta], \text{deg cm}^2 \text{ dmol}^{-1}) \quad (2)$$

Size Exclusion Chromatography. Chromatography was performed on a Pharmacia FPLC system using a Superdex 75 column [240 mm × 12 mm (i.d.)]. Peptide and protein samples (100 μL of 50 μM solution) were eluted at a flow rate of 0.5 mL/min with 0.1 M phosphate buffer at pH 7.0 containing 0.5 M NaCl. Sample elution was monitored either by absorbance at 214 nm or by fluorescence at 335 nm (excitation at 280 nm). A series of peptides of varying length derived from interleukin-8 and the fibronectin tenth type III

[1] Abbeviations: Boc, *tert*-butyloxycarbonyl; BrAc, bromoacetyl; CD, circular dichroism; DMF, *N*,*N*-dimethylformamide; HBTU, 1-hydroxybenzotriazole tetramethyluronium hexafluorophosphate; HPLC, high-performance liquid chromatography; MP-1, model protein 1; MP-2, model protein 2; NMR, nuclear magnetic resonance; standard IUPAC single-letter abbreviations for amino acids are used throughout; TFA, trifluoroacetic acid; UV, ultraviolet.

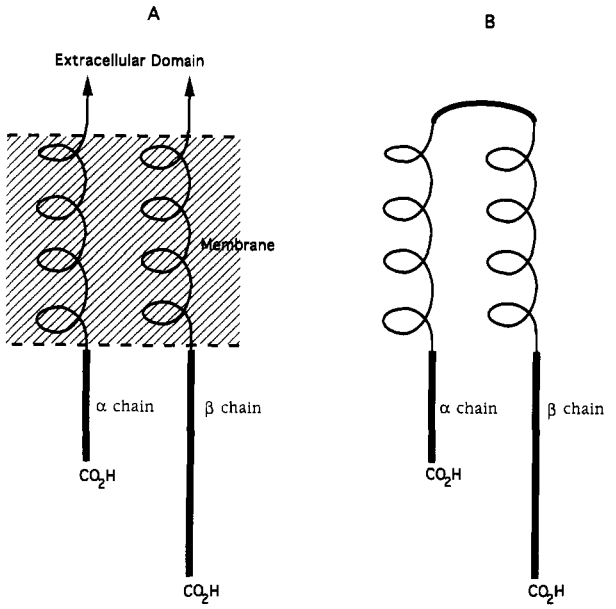

FIGURE 1: (A) Schematic representation of the presumed topology of the native integrin $\alpha_{IIb}\beta_3$ molecule. (B) Schematic representation of the designed model protein MP-1.

module, as well as the α_{IIb} and β_3 tails, were used as molecular weight standards.

Fluorescence Quenching Studies. Protein fluorescence was measured on a Jasco FP-777 spectrofluorometer. The quenching buffer comprised 50 mM NH_4OAc at pH 6.0 containing 0.16 M KCl and 0.1 mM $Na_2S_2O_3$. The KI stock solution (1.5 M) was prepared gravimetrically using quenching buffer. Protein stock solutions (typically 100 μM) were similarly prepared with quenching buffer. For each data point in the quenching experiment 200 μL of protein stock was added to the appropriate amount of KI stock and the final volume made to 1 mL with quenching buffer. All solutions were incubated for 1 h at 25 °C before measurements were taken. Fluorescence emission was monitored at 25 °C at 350 nm with excitation at 278 nm (E_m/E_x slit width = 10 nm) and expressed as $F°/F$ where $F°$ and F are the fluorescence of the protein in the absence and presence of quencher, respectively. The data are presented as direct Stern–Volmer plots of $F°/F$ *vs* quencher concentration. Stern–Volmer constants K_Q were calculated using eq 3.

$$F°/F = 1 + K_Q[I^-] \quad (3)$$

RESULTS AND DISCUSSION

Model Protein Design. In designing a model protein for the cytoplasmic and transmembrane domains of the $\alpha_{IIb}\beta_3$ receptor, various structural constraints were considered. The cytoplasmic tails of the $\alpha_{IIb}\beta_3$ receptor emerge from the plasma membrane at defined points, resulting in the shorter α_{IIb} tail being staggered relative to the longer β_3 tail (see Figure 1). Moreover, there is evidence to suggest that the α_{IIb} and β_3 cytoplasmic tails interact and that this interaction is critical for proper function (Ylänne et al., 1993). Such a quaternary interaction suggests that the cytoplasmic tails are located close to one another in the integrin heterodimer (Figure 1). Finally, there is indirect evidence to suggest that the transmembrane helices in $\alpha_{IIb}\beta_3$ interact within the membrane (Frachet et al., 1992; Bennet et al., 1993), and it is possible that the structure of the cytoplasmic domain itself is affected by the conformation of the membrane-spanning domain of $\alpha_{IIb}\beta_3$. To ensure that each of the above properties is present within our model protein,

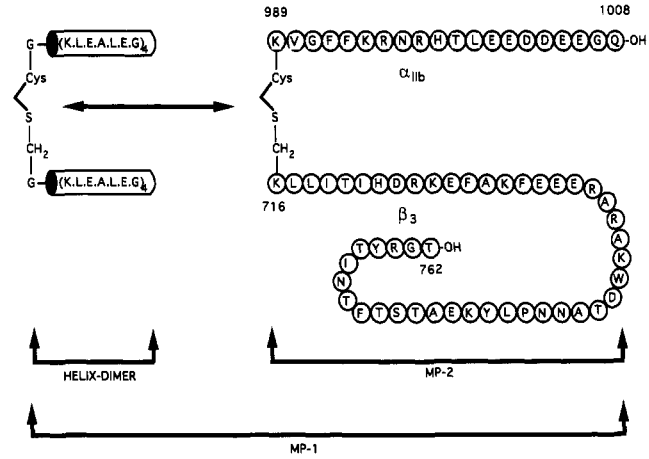

FIGURE 2: Amino acid sequences of the synthetic polypeptide targets: Helix dimer, Cys-G-(K-L-E-A-L-E-G-)$_4$CH$_2$-G-(K-L-E-A-L-E-G-)$_4$; MP-2, α_{IIb}-Cys-SCH$_2$-β_3; MP-1, α_{IIb}-helix-Cys-SCH$_2$-helix-β_3.

we have taken advantage of recent advances both in *de novo* design of peptides (Monera et al., 1993) and in the chemical synthesis of proteins (Muir & Kent, 1993). Innovations in these areas have allowed the rational design of the 126 amino acid model protein (MP-1) shown in Figure 1.

The MP-1 synthetic protein contains helical stretches designed to mimic the distinct membrane-spanning domains of the two polypeptide chains in the receptor molecule, one attached to each of the cytoplasmic tails from the α_{IIb} and β_3 subunits. The sequences of the α_{IIb} and β_3 tails correspond to residues 989–1008 (Prandini et al., 1988) and 716–762 (Fitzgerald et al., 1987), respectively (Figure 2). Note that there is evidence of post-translational side-chain amidation of the C-terminal glutamic acid in the α_{IIb} tail; therefore, a glutamine residue was incorporated at this point in the synthetic sequence (Calvete et al., 1990). The sequence designed to mimic the membrane-spanning region assumes a helical conformation in these regions of the native molecule (Deisenhofer et al., 1985). We have sought to imitate the effects of this membrane-spanning structure on the attached cytoplasmic tails by the use of a coiled-coil structure, suitably modified to provide solubility under aqueous conditions. Each chain of the coiled coil is composed of a 29 amino acid amphiphilic sequence and is itself made up of four tandem repeats of a seven-residue core peptide [G-(K-L-E-A-L-E-G)$_4$]. This particular sequence is derived from the prototypical coiled-coil protein tropomyosin (Sodek et al., 1978) and is known to form helical secondary structure in aqueous solution (Lau et al., 1984). Furthermore, synthetic peptides containing polymeric assemblies of this seven-residue core sequence are known to adopt coiled-coil tertiary and quaternary structures in aqueous systems (Zhou et al., 1992).

The stability of these coiled coils is largely a result of strong interchain hydrophobic interactions between leucine residues in the seven-residue repeat (Lau et al., 1984). By incorporating two of these amphiphilic segments into MP-1 we meet the absolute requirement for helicity, but not at the price of insolubility as might be the case if the natural hydrophobic membrane-spanning sequence were used. Moreover, the tendency of these helical elements to associate to form coiled coils may better mimic the proximity of transmembrane helices in the natural system and also ensures that a defined topology is maintained between the α and β cytoplasmic tails. In other words, the coiled coil acts as a structural template onto which the cytoplasmic domain of the integrin can be attached. This

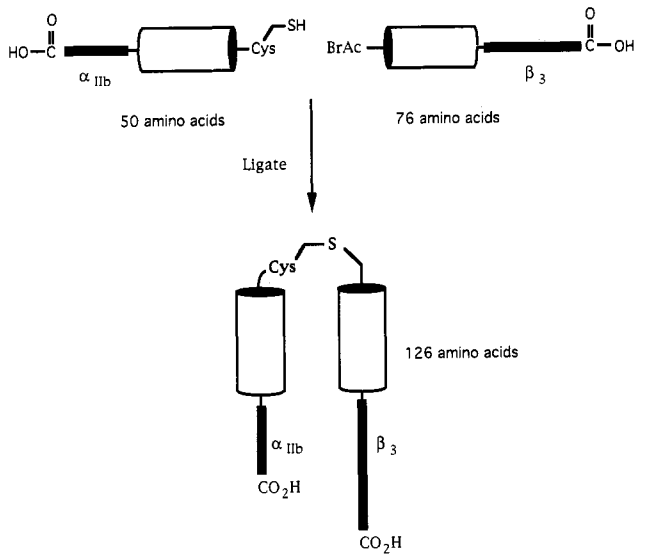

FIGURE 3: Synthesis of the model protein, MP-1, by chemical dovetailing. Schematic equation illustrating the intended reaction between Cys-helix-α_{IIb} and BrAc-helix-β_3.

ensures that the two cytoplasmic tails are staggered with respect to one another in a manner which approximates the intact protein (see Figure 1). Helical coiled-coil structures have previously been used as templates for the presentation of a small peptide motif (Engel et al., 1991). To better gauge the structural impact of the coiled-coil template we also designed a series of reference compounds. These included a second model protein, designated MP-2, containing the cytoplasmic tails of the $\alpha_{IIb}\beta_3$ receptor linked together in a head-to-head manner. This molecule contains no coiled-coil structural template. The individual cytoplasmic tails of the two subunits, on their own and with the additional amphiphilic sequence on their N-termini, were also deemed useful control molecules, as were the head-to-head linked coiled-coil segments themselves. The primary structures of all the component peptides and the assemblies of them are shown in Figure 2.

Synthetic Strategy. The model proteins illustrated in Figures 1 and 2 have a somewhat unusual architecture, namely, two C-termini. This feature renders them not directly accessible via either recombinant technology or standard chemical polypeptide synthesis. Clearly, a novel strategy must be employed. This synthetic challenge can be met by using the chemoselective ligation principle developed in this laboratory (Schnölzer & Kent, 1992). This strategy simply involves the chemical dovetailing of two fully unprotected peptide segments by a chemoselective ligation reaction to form the mature target compound. The selectivity of this ligation reaction is imposed by incorporating unique, mutually reactive groups, one within each peptide segment to be joined.

In the case of MP-1, each half of the target molecule was individually constructed, and then these two intermediates were joined together through the N-termini of the helices. In this case, the ligation chemistry takes advantage of the fact that neither cytoplasmic tail of the $\alpha_{IIb}\beta_3$ receptor contains a cysteine residue. Thus by including a unique cysteine residue with its nucleophilic sulfhydryl at the N-terminus of one half of MP-1 and an electrophilic bromoacetyl unit at the N-terminus of the other half we can chemically dovetail the two pieces in the desired manner (see Figure 3). The same principle applies to the synthesis of the control protein, MP-2, in which the two cytoplasmic tails are directly linked together.

Optimum Ligation Conditions. The feasibility of this approach was tested by studying the chemical ligation of two

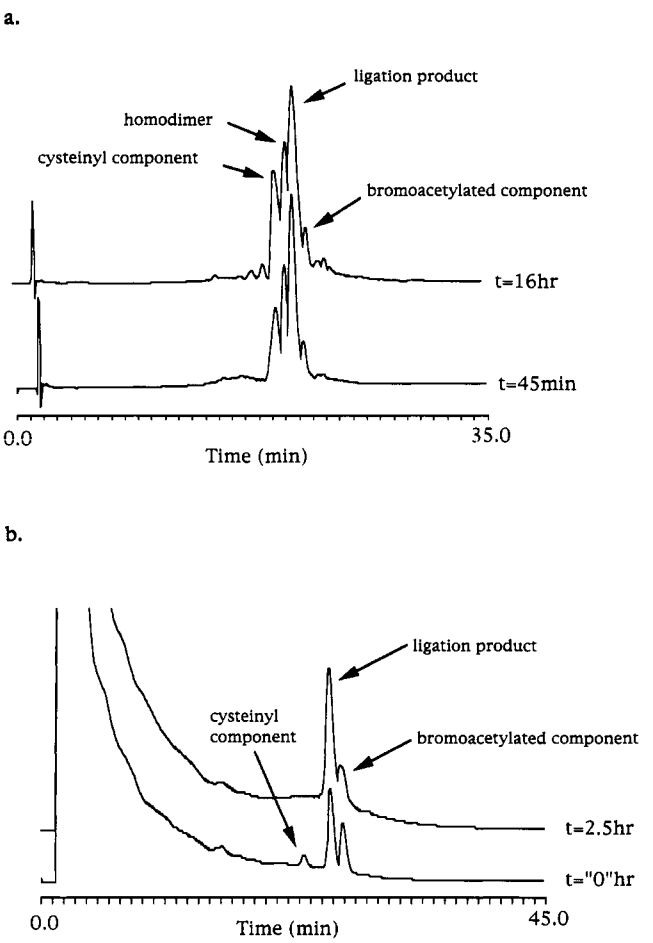

FIGURE 4: Chemical dovetailing of two peptide segments by means of a thioether bond. Study of reaction conditions using the pro-(coiled coil) peptides as model segments: (a) aqueous solution (0.1 M phosphate at pH 7.0); note high level of disulfide-linked peptide homodimer and relatively low yield of ligation product; (b) DMF–0.1 M phosphate at pH 7.0 (95:5 v/v). The absence of the cysteinyl component after 2.5 h (only residual bromoacetylated component remains) indicates essentially quantitative formation of the desired thioether-linked ligation product. Note the absence of disulfide-linked homodimer.

derivatives of the amphiphilic helical peptide. Thus both the nucleophilic peptide segment Cys-G·(K·L·E·A·L·E·G)$_4$ and the electrophilic component BrAc-G·(K·L·E·A·L·E·G)$_4$ were synthesized as peptide amides, and their mutual reactivity was investigated. The choice of pH for the reaction is critical since it must be sufficiently high to ionize the side-chain SH of Cys and render it nucleophilic. Preliminary studies had revealed that ligations could not be brought about between cysteine-containing peptides and bromoacetylated peptides at low pH (T. W. Muir, M. Schnölzer, and S. B. H. Kent, unpublished data). However, too high a pH will also deprotonate the ϵ-amino groups in lysine side chains (pK_a $\sim$10.5), causing them to react with electrophiles such as the bromoacetyl group, and leading to undesired reaction products. Obviously a compromise must be struck between these two effects. Similar observations have been made by Wetzel and co-workers (Wetzel et al., 1990).

Initially, the ligation reaction was performed under aqueous conditions (pH 7.0, 0.1 M phosphate, 10 mg/mL in each reactant) and the progress of the ligation reaction was monitored by analytical HPLC (see Figure 4a). Individual HPLC peaks were collected and examined by mass spectrometry. Two chemical reactions were observed under these conditions. The desired ligation product (thioether-linked helix

dimer) was formed in significant amounts after only 45 min. However, substantial amounts of disulfide-linked homodimer of the cysteinyl component also formed under these conditions. Formation of this oxidized species effectively protects the reactive sulfhydryl group, thereby substantially reducing the ligation yield.

Next we sought to enhance the nucleophilic ligation reaction relative to the unwanted oxidation reaction. Air oxidation of a mercaptan to a disulfide follows a multistep mechanism involving sulfur radical formation (Wallace et al., 1963), whereas the nucleophilic attack of a thiolate ion on a primary alkyl halide, such as the bromoacetyl unit, occurs by an S_N2 mechanism. Thus, the reaction rate for *ligation* should be greatly increased relative to oxidation by the use of a dipolar aprotic solvent. This is because of two effects: enhanced reactivity of the charged nucleophile due to reduced solvation, and better solvation of the charge-separated transition state (Parker, 1969).

To test this hypothesis, the ligation reaction was repeated using a solvent system composed of 95% dimethylformamide (DMF) and 5% 0.1 M phosphate at pH 7.0, the aqueous phosphate being added to ensure thiolate formation (Baca & Kent, 1993). To further disfavor dimerization of the cysteinyl component, the concentration of each reactant was reduced from 10 mg/mL to 5 mg/mL. Under these conditions the nucleophilic ligation reaction proceeded extremely quickly. Considerable amounts of product were present at the earliest time points examined, after less than 1 min of reaction. After 2.5 h the reaction had already gone to completion, as indicated by the complete disappearance of the cysteinyl component (see Figure 4b). Importantly, absolutely no disulfide-linked dimer was observed. Thus, by use of the dipolar aprotic solvent DMF, disulfide formation can be eliminated in favor of the desired ligation reaction.

Synthesis of the Model Proteins MP-1 and MP-2. The synthesis of MP-2, by reaction of H-Cys-[α_{IIb}]-OH with BrAc-[β_3]-OH, was similarly performed in 95% DMF/5% phosphate (pH 7.0) using a concentration of 5 mg/mL in each reactant. As expected from the model ligation studies, the reaction proceeded virtually to completion under these conditions. The ligation product, MP-2, was subsequently purified by preparative HPLC and its covalent structure confirmed by ion-spray mass spectrometry (see Figure 5a).

The neoprotein MP-1 was prepared by chemical dovetailing of the peptides H-Cys-G·(K·L·E·A·L·E·G)$_4$-[α_{IIb}]-OH and BrAc-G·(K·L·E·A·L·E·G)$_4$-[β_3]-OH. Chemoselective ligation was carried out in 95% DMF/5% 0.1 M phosphate (pH 7.0) containing 8 M urea, using a concentration of 5 mg/mL in each reactant. The inclusion of urea in the mixed solvent system was found to greatly aid solubility of the two polypeptide fragments. The ligation product, MP-1, was purified to homogeneity by preparative HPLC and characterized by ion-spray mass spectrometry (see Figure 5b). Significantly, this ligation reaction was observed to proceed much slower than for the model ligation studies and for the synthesis of MP-2, both described above. After 12 h, the yield of product was approximately 60%. At this point the reaction was terminated since the residual bromoacetylated peptide had almost all been converted to the unreactive chloroacetyl adduct (as indicated by mass spectrometry) by chloride ions of unknown origin. In the model ligation studies and the synthesis of MP-2, this formation of chloroacetyl peptide was not a problem since the ligation reaction itself was extremely rapid. We have no explanation as to the reasons behind this unexpected sluggish ligation.

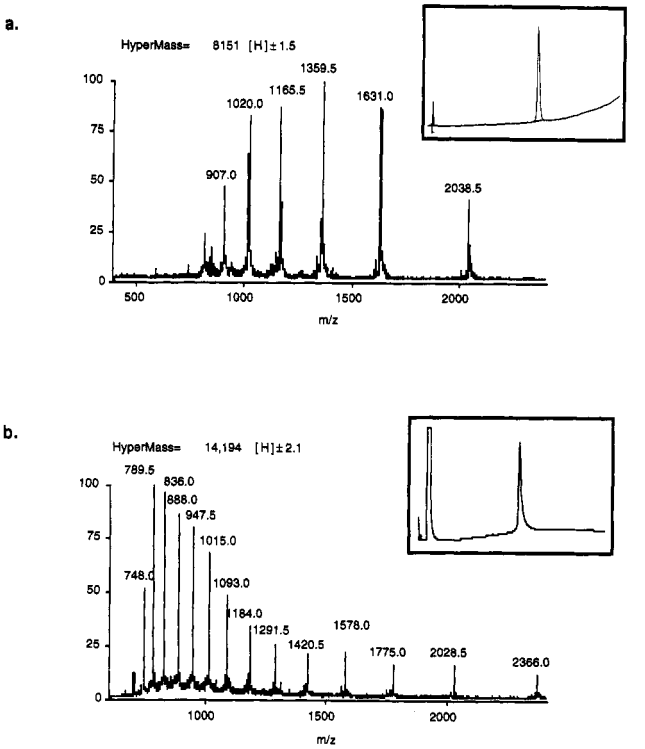

FIGURE 5: Characterization of the purified synthetic neoproteins: (a) MP-2; (b) MP-1. Main panels: ion-spray mass spectra showing charge state distribution. Inserts: analytical reverse-phase HPLC spectra. Data clearly reveal that MP-1 and MP-2 both possess the desired covalent structure and are highly homogeneous.

Circular Dichroism Studies. Preliminary structural studies on the synthetic molecules described were performed using ultraviolet circular dichroism (UV CD) spectroscopy.

(*A*) *Cytoplasmic Tails.* The far UV CD spectra of the cytoplasmic tails of both α_{IIb} and β_3 were taken in boric acid, pH 7.0 at 25 °C (see Figure 6a). Significantly, in each case the absence of distinct minima at 208 and 222 nm revealed that neither cytoplasmic tail contained appreciable helical secondary structure in aqueous solution. However, in other respects the two CD spectra were strikingly different. The spectrum of the 47 amino acid β_3 cytoplasmic tail contained a strong minimum at 203–205 nm as opposed to that of the 20 amino acid α_{IIb} cytoplasmic tail which was characterized by a small maximum at 219–220 nm and a weak negative ellipticity at 235 nm (Figure 6a).

The far-UV CD of the α_{IIb} cytoplasmic tail clearly indicates a nonhelical structure. The spectrum has many features in common with classical random-coil poly(amino acids) (Townend et al., 1966). The primary sequence of the α_{IIb} tail contains a series of six acidic amino acids in a row (Glu·Glu·Asp·Asp·Glu·Glu), and at pH 7.0 there is potential for electrostatic repulsion between neighboring carboxylates, perhaps accounting for this disordered structure. The β_3 cytoplasmic tail exhibits a far-UV CD spectrum which differs significantly from the classical random-coil spectrum exhibited by the α_{IIb} cytoplasmic tail. The spectrum indicates that the β_3 cytoplasmic tail also contains little if any α-helix. Thus, while both isolated cytoplasmic tails are clearly nonhelical, these spectral differences point to the cytoplasmic tails having dissimilar structural properties.

(*B*) *Helix-Cytoplasmic Tail Molecules.* The far-UV CD spectra of both H-G·(K·L·E·A·L·E·G)$_4$-[α_{IIb}]-OH (helix-α_{IIb}) and H-G·(K·L·E·A·L·E·G)$_4$-[β_3]-OH (helix-β_3) were markedly different from the corresponding cytoplasmic domains

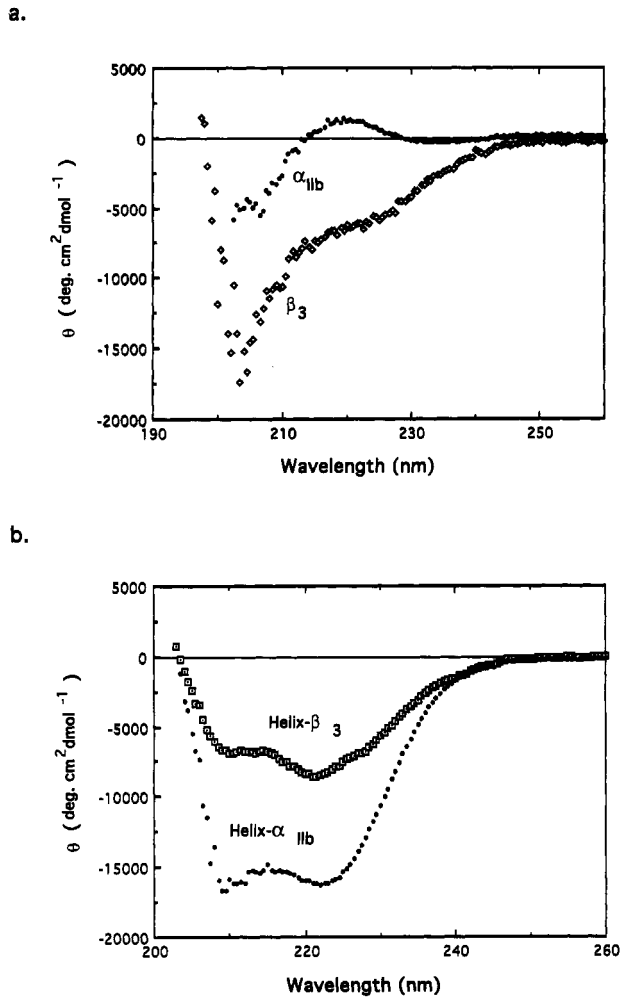

FIGURE 6: Circular dichroism studies of synthetic polypeptides containing the isolated cytoplasmic domains of the integrin $\alpha_{IIb}\beta_3$. (a) Isolated α_{IIb} and β_3 cytoplasmic domains. (b) H-G-(K-L-E-A-L-E-G)$_4$-[α_{IIb}]-OH (helix-α_{IIb}) and H-G-(K-L-E-A-L-E-G)$_4$-[β_3]-OH (helix-β_3).

Table 2: Percent Helicity of Synthetic Proteins, Calculated from Far-UV CD Spectra Using Eqs 1 and 2

sample	protein helicity (%)	$\theta_{222}/\theta_{208}$
helix-αIIb	44	1.02
helix-β_3	23	1.25
helix dimer	85	1.00
MP-1	74	1.08

alone (see Figure 6b). Both helix-α_{IIb} and helix-β_3 exhibited bimodal UV CD spectra with minima at 208 and 222 nm, indicative of helical secondary structure (Holzwarth & Doty, 1965). The protein helicity estimated from $[\theta]_{222}$ using eq 1 and 2 (see Experimental Procedures) was 44% for helix-α_{IIb} and 23% for helix-β_3 (Table 2). The absence of detectable helicity in the individual α_{IIb} and β_3 cytoplasmic tails suggests that the α-helix is restricted to the N-terminal pro-(coiled coil) sequence in helix-α_{IIb} and helix-β_3. The $[\theta]_{222}/[\theta]_{208}$ ratio has previously been used to assess the number of helical strands within a molecule (Lau et al., 1984). A value for $[\theta]_{222}/[\theta]_{208}$ of around 0.8 is associated with single-stranded α-helix, whereas a value of about unity is suggestive of a two-stranded coiled coil. By this measure helix-α_{IIb} with a $[\theta]_{222}/[\theta]_{208}$ value of 1.02 would appear to contain a coiled-coil structural unit. Similarly, helix-β_3 ($[\theta]_{222}/[\theta]_{208} = 1.20$) would also contain a coiled coil by this criterion.

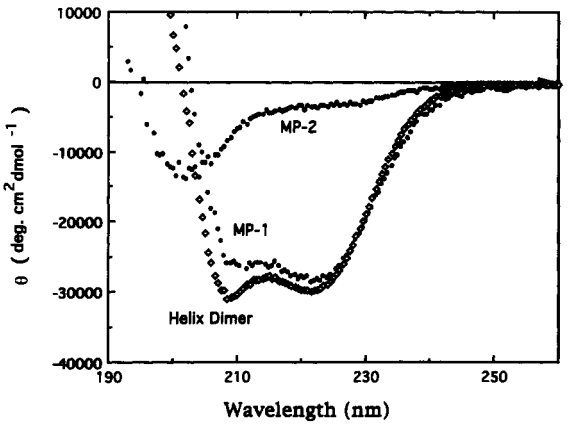

FIGURE 7: Circular dichroism studies of the synthetic neoprotein structural models MP-1, MP-2, and the covalent helix dimer. From these spectra it is clear that MP-1 and helix dimer contain high levels of helicity, whereas MP-2 is essentially nonhelical.

The moderate protein helicity exhibited by the helix-[cytoplasmic tail] molecules would seem to preclude the intramolecular formation of a coiled-coil architecture, implying that the molecules exist as homodimeric coiled coils. To test this hypothesis, size exclusion chromatography was performed on both molecules. When comapred with a number of peptide standards of varying length, the retention volumes of helix-α_{IIb} (monomer = 5.4 kDa) and helix-β_3 (monomer = 8.7 kDa) corresponded to apparent masses of approximately 9.9 and 16 kDa, respectively. These data support the existence of homodimeric species. The putative intermolecular coiled coils within these homodimers presumably involve the noncovalent association of the two N-terminal amphiphilic helices stabilized by hydrophobic interactions. It is important to note that the amphiphilic sequences within these coiled coils can potentially associate in either a parallel or antiparallel arrangement, and that a low-resolution technique such as circular dichroism cannot distinguish between these two possible orientations.

(C) Helix Dimer, MP-1, and MP-2. The CD spectrum of the 60 amino acid Cys-G-(K-L-E-A-L-E-G)$_4$CH$_2$-G-(K-L-E-A-L-E-G)$_4$ (helix dimer) was indicative of an α-helical coiled-coil protein (Figure 7). Both the estimated protein helicity of 85% (Table 2) and the $[\theta]_{222}/[\theta]_{208}$ value of exactly 1.00 confirm the presence of this tertiary structure in the covalently linked molecule. This observation is of key importance since the presence of a coiled coil in such a parallel architecture is a fundamental design feature of the model protein MP-1. The far-UV CD spectrum of MP-2 reveals that this 69-residue molecule does not contain substantial helical secondary structure (Figure 7), an observation which is entirely consistent with the CD studies on the individual cytoplasmic tails.

The far-UV CD spectrum of the 126-residue neoprotein MP-1 contained minima at 208 and 222 nm, again indicating the presence of helical secondary structure (Figure 7). The calculated $[\theta]_{222}/[\theta]_{208}$ value of 1.08 suggests that, as expected, the molecule does contain an area of α-helical coiled coil. The retention volume of the MP-1 (monomer = 14 kDa) molecule on gel permeation chromatography corresponded to an apparent mass of 18 kDa. This suggests that the molecule exists as a monomer in aqueous solution, implying that this coiled coil is intramolecular. In stark contrast to the nonhelical MP-2, the protein helicity of MP-1 was estimated to be some 74%. On the basis of the covalent structure of the MP-1 molecule, the expected coiled-coil region makes up only 47% of the amino acid content of the molecule. Thus the 74% observed helicity (Table 2) indicates that helicity extends

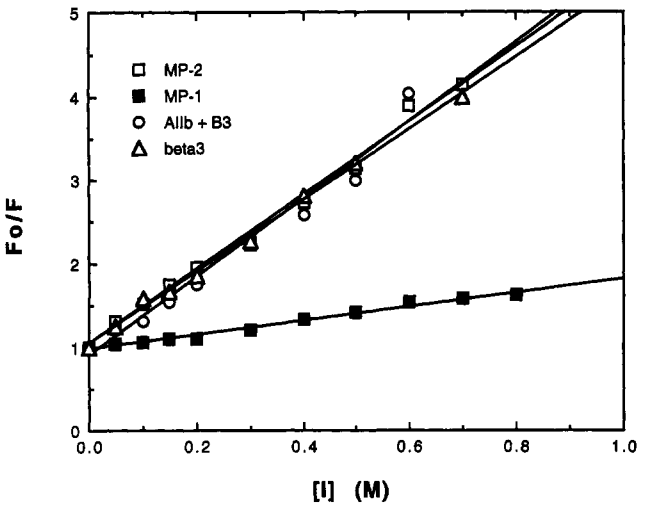

FIGURE 8: Fluorescence quenching studies as a conformational probe of the synthetic neoprotein structural models. Stern–Volmer plots of MP-1, MP-2, isolated β_3, and isolated β_3 (1 equiv) + isolated α_{IIb} (1 equiv). F° and F refer to the tryptophan fluorescence in the absence and presence of quencher, respectively. $[I^-]$ is the concentration of iodide quencher. These data reveal a significant difference in the solvent exposure of the single tryptophan in MP-1 compared to the other molecules studied.

beyond the expected amphiphilic coiled-coil region of the molecule into the cytoplasmic domains. It is significant that this additional helicity was not observed for either helix-α_{IIb} or helix-β_3, both of which appear to exist as homodimers containing coiled coils. This argues that the coiled coil unit is not itself sufficient to induce helicity into the cytoplasmic tails. Therefore, the additional helicity which is unique to the MP-1 molecule must be a result of the α_{IIb} and β_3 tails being present in a staggered and parallel orientation within this molecule. For this reason, we sought to further define the structures of the "cytoplasmic domains" of the neoprotein MP-1.

Fluorescence Quenching Studies. Quenching of protein fluorescence by addition of external heavy atoms such as the iodide ion can provide valuable information regarding the local structural environment of the tryptophan or tyrosine side chain fluorophores (Lehrer & Leavis, 1978). Quenching occurs as a result of diffusion-controlled encounters of the heavy atom with the fluorophore and so can be related to the solvent exposure of the residue. The MP-1 neoprotein construct contains a single tryptophan fluorophore located 24 residues from the C-terminal of the β_3 tail. This can thus be used as a convenient structural probe into this area of the protein.

Iodide quenching data was obtained for MP-1, MP-2, the β_3 cytoplasmic tail, and the β_3 tail mixed with a stoichiometric amount of the α_{IIb} cytoplasmic tail. The data sets are presented as plots of F°/F versus iodide concentration, where F° and F are fluorescence in the absence and presence of quencher, respectively (Figure 8). In each case the fluorescence quenching appears to follow the simple Stern–Volmer relationship (eq 3), suggesting that the quenching of the tryptophan fluorescence is not complicated by energy transfer to tyrosine residues (the nearest of which is eight residues downstream). The resulting Stern–Volmer constants K_Q, calculated from the graphs using eq 3, reveal a significant difference between MP-1 and the other molecules studied (Table 3).

There appears to be little difference in the solvent exposure of the lone tryptophan of MP-2 and the β_3 cytoplasmic tail alone as indicated by the similar values of K_Q calculated for each. Furthermore, addition of the α_{IIb} cytoplasmic tail to a solution of the β_3 cytoplasmic tail has no apparent effect on the tryptophan quenching. The MP-1 fluorescence, on the other hand, is quenched to a considerably lesser extent, indicating that the single tryptophan in this neoprotein molecule is substantially protected from solvent. In light of the measurements on control compounds, we think that this solvent shielding can be attributed to tertiary interactions between the α_{IIb} and β_3 cytoplasmic tails with the formation of a hydrophic core involving the tryptophan. This suggests that the neoprotein, MP-1, has a defined tertiary structure.

Table 3: Fluorescence Quenching Studies of the Synthetic Neoprotein MP-1 and Related Control Compounds: Stern–Volmer Constants (K_Q) Calculated Using Eq 3

sample	K_Q (M^{-1})
β_3	4.3 ± 0.2
β_3 + αIIb	4.7 ± 0.3
MP-2	4.4 ± 0.2
MP-1	0.84 ± 0.06

It is important to mention that two slightly different sequences for the cytoplasmic domain of α_{IIb} are present in the literature. All the studies reported above refer to the α_{IIb} sequence shown in Figure 2. However, model proteins containing the other α_{IIb} cytoplasmic tail sequence were also constructed using the same procedure as before. Note that this variant differs at only two positions, His^{998} and Thr^{999} (Prandini et al., 1988) to Pro^{998} and Pro^{999} (Poncz et al., 1987). These molecules were found to possess essentially identical structural properties to the above when studied by both circular dichroism and fluorescence spectroscopies.

SIGNIFICANCE

In summary, these preliminary structural results suggest that the synthetic protein MP-1 is a promising soluble model for the cytoplasmic domains of the heterodimeric receptor, integrin $\alpha_{IIb}\beta_3$. MP-1 has a defined tertiary structure, with unique features present only in the model protein, namely, the coiled-coil template mimics the transmembrane helices of the two subunits, at the same time constraining the cytoplasmic tails in a staggered, parallel and proximal topology. Further studies are now underway to determine the structure and biological activity of the MP-1 construct through high-resolution nuclear magnetic resonance (NMR) spectroscopy techniques and a range of functional assays.

The ability to study a putative functional model of the cytoplasmic domains of integrin $\alpha_{IIb}\beta_3$ in non membrane bound conditions will allow systematic exploration of the structural origins of biological effects. The use of chemically synthesized neoprotein structural models of nonlinear architecture to mimic complex membrane receptors may be a useful approach to study signaling processes in multisubunit transmembrane receptors.

REFERENCES

Baca, M., & Kent, S. B. H. (1993) *Proc. Natl. Acad. Sci. U.S.A.*, in press.

Bennet, J. S., Kolodziej, M. A., Vilaire, G., & Poncz, M. (1993) *J. Biol. Chem. 268*, 3580–3585.

Calvete, J. J., Schafer, W., Henschen, A., & Gonzalez-Rodriguez, J. (1990) *FEBS Lett. 263*, 43–46.

Chen, Y. H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350–3359.

Deisenhofer, J., Epp, O., Miki, K., Huber, R., & Michel, H. (1985) *Nature 318*, 618–624.

De Vos, A. M., Ultsch, M., & Kossiakoff, A. A. (1992) *Science 255*, 306–312.
Engel, M., Williams, R. W., & Erickson, B. W. (1991) *Biochemistry 30*, 3161–3168.
Fitzgerald, L. A., Steiner, B., Rall, S. C., Jr., Lo, S.-S., & Philips, D. R. (1987) *J. Biol. Chem. 262*, 3936–3939.
Frachet, P., Duperray, A., Delachanal, E., & Marguerie, G. (1992) *Biochemistry 31*, 2408–2415.
Ginsberg, M. H., Du, X., & Plow, E. F. (1992) *Curr. Opin. Cell Biol. 4*, 766–771.
Holzwarth, G., & Doty, P. (1965) *J. Am. Chem. Soc. 87*, 218–228.
Hynes, R. O. (1992) *Cell 69*, 11–25.
Lau, S. Y. M., Taneja, A. K., & Hodges, R. S. (1984) *J. Biol. Chem. 259*, 13253–13261.
Lehrer, S. S., & Leavis, P. C. (1978) *Methods Enzymol. 49*, 222–236.
Milburn, M. V., Prive, G. G., Milligan, D. L., Scott, W. G., Yeh, J., Jancarik, J., Koshland, D. E., & Kim, S.-H. (1991) *Science 254*, 1342–1347.
Miyajima, A., Hara, T., & Kitamura, T. (1992) *TIBS 17*, 378–382.
Monera, O. D., Zhou, N. E., Kay, C. M., & Hodges, R. S. (1993) *J. Biol. Chem. 268*, 19218–19227.
Muir, T. W., & Kent, S. B. H. (1993) *Curr. Opin. Biotechnol. 4*, 420–427.
Muir, T. W., Schnolzer, M., & Kent, S. B. H. (1993) unpublished data.
Parker, P. (1969) *Chem. Rev. 69*, 1–32.
Patthy, L. (1990) *Cell 61*, 13–14.
Phillips, D. R., Charo, I. F., & Scarborough, R. M. (1991) *Cell 65*, 359–362.
Poncz, M., Eisman, R., Heidenreich, R., Silver, S. M., Vilaire, G., Surrey, S., Schwartz, E., & Bennet, J. S. (1987) *J. Biol. Chem. 262*, 8476–8482.
Prandini, M. H., Denarier, E., Frachet, P., Uzan, G., & Marguerie, G. (1988) *Biochem. Biophys. Res. Commun. 156*, 595–601.
Sarin, V. K., Kent, S. B. H., Tam, J. P., & Merrifield, R. B. (1981) *Anal. Biochem. 117*, 147–157.
Sastry, S. K., & Horwitz, A. F. (1993) *Curr. Opin. Cell. Biol. 5*, 819–831.
Schnölzer, M., & Kent, S. B. H. (1992) *Science 256*, 221–225.
Schnölzer, M., Jones, A., Alewood, P. F., & Kent, S. B. H. (1992a) *Anal. Biochem. 204*, 335–343.
Schnölzer, M., Alewood, P., Jones, A., Alewood, D., & Kent, S. B. H. (1992b) *Int. J. Pept. Protein Res. 40*, 180–193.
Sodek, J., Hodges, R. S., & Smillie, L. B. (1978) *J. Biol. Chem. 253*, 1129–1136.
Townend, R., Kumosinski, T. F., Timasheff, S. N., Fasman, G. D., & Davidson, B. (1966) *Biochem. Biophys. Res. Commun. 23*, 163–169.
Wallace, T. J., Schriesheim, A., & Bartok, W. (1963) *J. Org. Chem. 28*, 1311–1314.
Weiss, A. (1993) *Cell 73*, 209–212.
Weiss, M. S., Wacker, T., Weckesser, J., Welte, W., & Schulz, G. E. (1990) *FEBS Lett. 267*, 268–272.
Wetzel, R., Halualani, R., Stults, J. T., & Quan, C. (1990) *Bioconjugate Chem. 1*, 114–122.
Ylänne, J., Chen, Y., O'Toole, T. E., Loftus, J. C., Takada, Y., & Ginsberg, M. H. (1993) *J. Cell Biol. 122*, 223–233.
Zhou, N. E., Kay, C. M., & Hodges, R. S. (1992) *J. Biol. Chem. 267*, 2664–2670.