Structural Studies of a Neuropeptide Precursor Protein with an RGD Proteolytic Site†

Cherian Zachariah,‡ Angus Cameron,§ Iris Lindberg,§ K. J. Kao,|| Margery C. Beinfeld,⊥ and Arthur S. Edison*,‡

Department of Biochemistry and Molecular Biology, Box 100245, and Department of Pathology, University of Florida, Gainesville, Florida 32610, Department of Biochemistry and Molecular Biology, LSU Health Sciences Center, New Orleans, Louisiana 70112, and Department of Pharmacology and Experimental Therapeutics, Tufts University School of Medicine, 136 Harrison Avenue, Boston, Massachusetts 02111

Received March 6, 2001; Revised Manuscript Received May 17, 2001

ABSTRACT: The snail *Lymnaea stagnalis* produces a neuropeptide precursor protein that contains seven Arg-Gly-Asp (RGD) sites. These sites are recognized and cleaved by one or more prohormone convertases in the first processing step to yield mature neuropeptides in the secretory pathway. Conformations of two synthetic RGD-containing peptides derived from the *L. stagnalis* precursor protein were determined by NMR spectroscopy. The peptides were tested in a platelet aggregation assay for RGD activity and were processed in vitro by PC2 and furin. The native peptide with a proline following the RGD site has minimal structure around the RGD region, does not inhibit platelet aggregation, and is properly processed by the enzymes PC2 and furin. A variant of the native fragment with a serine following the RGD sequence has a significant amount of a reverse turn around the RGD region, is a potent inhibitor of platelet aggregation, and is processed with the same specificity as the native fragment. The large conformational differences between the two peptides provide a molecular mechanism for effects of proline residues following the RGD site and suggest that precursor processing is influenced more by flexibility than by the conformation of the processing site.

RGD[1] sequences bind to integrins and are involved with numerous and diverse biological functions including general cell adhesion (*1, 2*), platelet aggregation (*3*), blood pressure regulation (*4, 5*), immunity (*2*), protein sorting (*6*), multiple roles in the nervous system (*7*), and reproduction (*8*). As a result of the wide range of functions evoked by such a short amino acid sequence, a considerable amount of research has been devoted to studying the structure (*9*−*18*) and dynamics (*19*) of RGD-containing peptides and proteins. Studies from small cyclic peptides have shown that the RGD sequence can form a variety of reverse turns, most of which lead to higher integrin binding activity than flexible linear RGD peptides (*20*−*28*). RGD sequences in proteins are often associated with flexible loops (*9*−*12*), but at least one NMR study of recombinant decorsin demonstrated that although the RGD-containing loop is relatively flexible, the conformation of the RGD sequence itself is more rigid when compared to other RGD-containing proteins (*19*).

A recent study has shown that a highly conserved RGD sequence in prohormone convertase 1 (PC1) is involved in intracellular sorting, but the mechanisms underlying the sorting are unknown (*6*). Both PC1 and PC2 bind to integrin $\alpha5\beta1$, but this binding is mediated by regions of the proteins that do not contain RGD sequences (*6*). Thus, it is not yet clear that RGD-mediated sorting is related to integrin binding.

One of the primary substrates for PC enzymes are the neuropeptide precursor proteins, and proper processing of the precursors to active peptides requires that both the precursor and all associated processing enzymes are properly sorted into the regulated secretory pathway (*29*−*31*). The processing sites in many neuropeptide precursor proteins consist of basic (R or K), dibasic (RR, KK, KR, RK), or "tetrabasic" (RXXR) amino acids. These different proteolytic sites are recognized and cleaved by different members of the PC family of enzymes (*29*−*34*). In at least three neuropeptide precursor proteins [*Lymnaea stagnalis* (*35*); *Cepaea nemoralis* (*36*); *Ascaris suum* (*37*)], one or more of the processing sites is also an RGD (Arg-Gly-Asp) sequence, suggesting the possibility that the unprocessed RGD sequence could have an integrin binding function in addition to being a substrate for processing.

Previous theoretical studies on precursors of peptide hormones predicted the presence of β-turns around putative

† This study was supported by an NSF CAREER award to A.S.E. and the National High Magnetic Field Laboratory. M.C.B. was supported by NIH Grant NS31602.

* To whom correspondence should be addressed at the Department of Biochemistry and Molecular Biology, Box 100245, University of Florida, Gainesville, FL 32610-0245. Phone: 352-392-4535. FAX: 352-392-3422. Email: art@ascaris.ufbi.ufl.edu.

‡ Department of Biochemistry and Molecular Biology, University of Florida.

§ Department of Biochemistry and Molecular Biology, LSU Health Sciences Center.

|| Department of Pathology, University of Florida.

⊥ Department of Pharmacology and Experimental Therapeutics, Tufts University School of Medicine.

[1] Abbreviations: RGD, arginine-glycine-aspartate; LS-1, *Lymnaea stagnalis* peptide 1; LS-2, *Lymnaea stagnalis* peptide 2; NMR, nuclear magnetic resonance; FLP, FMRF-NH_2-like peptides; PC, prohormone convertase; PBS, phosphate-buffered saline; ADP, adenosine diphosphate; DTT, dithiothreitol; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy.

10.1021/bi010448s CCC: $20.00
Published on Web 06/30/2001

Table 1: Peptides Used in This Study

LS-1	GDPFLRF-NH_2
LS-1[S]	GDSPFLRF-NH_2
LS-2	GDPFLRFGRGDPFLRF-NH_2
LS-2[S]	GDPFLRFGRGDSPFLRF-NH_2

dibasic processing sites, providing an accessible region for peptide cleavage (*38*). Subsequent studies using prosomatostatin (*39*), synthetic peptides which mimic the pro-ocytocin−neurophysin cleavage domain (*40*), and those which mimic the cleavage site of the pro-ocytocin (*41*) indicate that processing in prohormones occurs at regions of β-turns. The relative orientations of the side chains of the dibasic amino acids at the processing site were also shown to influence processing (*42*).

In this paper, we structurally and functionally characterize two peptides derived from an FMRF-NH_2-like peptide (FLP) precursor protein from *Lymnaea stagnalis*. These peptides contain two neuropeptides separated by an intact processing site that, in the unprocessed state, includes an RGD sequence. First, we show that the two peptides have different interactions with GPIIb/IIIa integrins using platelet aggregation studies and that the native sequence does not bind to the GPIIb/IIIa integrins. Next, the conformational properties of the peptides will be presented, and we show that a single serine insertion following the RGD results in a significantly different structure at the RGD site. Finally, we present initial in vitro processing studies using two prohormone convertases, PC2 and furin.

MATERIALS AND METHODS

Peptide Synthesis. The peptides used in this study (Table 1) were synthesized by Mr. Alfred Chung (University of Florida Interdisciplinary Center for Biotechnology Research Protein Chemistry Core) at 0.25 mM scale, using an Applied Biosysytem 432A Peptide Synthesizer (Applied Biosystems, Foster City, CA) using standard 4-(fluorenylmethyloxy)-carbonyl chemistry and cleaved from the solid support with trifluoroacetic acid in the presence of appropriate scavengers. The scavengers were extracted with *tert*-butyl methyl ether, and the peptides were diluted with acetic acid and freeze-dried. The peptides were purified using HPLC to at least 95% purity, as estimated by analytical HPLC and MALDI-TOF mass spectrometry.

Platelet Aggregation Assays. Fresh platelet-rich plasma was prepared from peripheral blood of a healthy donor by centrifugation at 250*g* for 10 min at room temperature. Blood was anticoagulated with 0.38% sodium citrate. Platelet-poor plasma was prepared by centrifugation of blood at 1500*g* for 15 min. Platelet aggregation assays were conducted as reported previously (*43*) and were performed in a Chronolog Platelet Aggregometer (Chronolog Corp., Broomall, PA) at 37 °C. Briefly, the assay consisted of 400 μL of platelet-rich plasma and 100 μL of peptide or PBS. The aggregation was initiated by the addition of 5 μL of 1 mM ADP. The peptides GRGDS and GRGES were used, respectively, as positive and negative controls for inhibition of ADP-induced platelet aggregation.

NMR Spectroscopy. NMR data were collected at 600 and 750 MHz using Bruker Avance 600 or 750 at the University of Florida Center for Structural Biology or using a Varian Unity 600 at the National High Field Magnet Laboratory (Tallahassee, FL). Approximately 3−5 mM samples of the peptides were prepared in 90% H_2O, 10% D_2O, and 0.3 mM 3-trimethylsilyl[2,2,3,3-$^2H_4$]propionic acid for an internal chemical shift standard (0.0 ppm). All NMR data were collected at 4 °C.

One-dimensional ^{1}H pH titrations were carried out from pH 1.9 to 5.5 (uncorrected for D_2O) at pH increments of about 0.16. The pH was adjusted by the addition of small volumes of concentrated KOH or HCl. The spectral widths were between 10 and 13 ppm. One-dimensional data were collected with 16 384 points; two-dimensional data were collected with 2048 or 4096 complex points in the acquisition dimension and 512 complex points in the indirect dimension. ^{1}H resonances were assigned using total correlation spectroscopy (TOCSY) (*44*) using the MLEV-17 mixing sequence (*45*) with a 60 ms mixing time and nuclear Overhauser effect spectroscopy (NOESY) with a mixing time of 150−300 ms (*46*). Water in all spectra was eliminated using presaturation, excitation sculpting (*47*), or 3-9-19 WATERGATE sequences (*48*).

NMR data were processed using the software NMRPipe (*49*) by first eliminating the residual water resonance by spectral deconvolution, multiplying the data by a squared cosine function, zero filling once, Fourier transformation, and baseline correction. Data were sequentially assigned using standard methods (*50*) using NMRview (*51*). Peak positions and intensities were obtained using the automatic routines included in NMRview.

Molecular Dynamics Simulations. Models of LS-2 and LS-2[S] were made in Insight II (Molecular Simulations Inc., San Diego, CA). Energy minimization and dynamics simulations were done with Discover or Discover 3 on an Onyx 2 computer (Silicon Graphics) at the McKnight Brain Institute of the University of Florida. The peptides were constructed in an extended conformation and briefly minimized, and NMR distance restraints were added as pseudo-potentials to the cvff force field. NOE distances were calibrated by the relationship: $r_{ab}^6 = r_{cal}^6(I_{cal}/I_{ab})$, where r_{ab} is the distance between atoms *a* and *b*, r_{cal} is the known distance, and $I_{ab(cal)}$ is the intensity of the corresponding NOESY cross-peak. Calibrations used were the following: Arg-9 H$^\beta$s (1.75 Å) for LS-2 and Asp-2 H$^\beta$s (1.76 Å) for LS-2[S]. One (1.0) angstrom was added to each experimental distance to define flat-bottom potential energy wells and allow for conformational averaging. Hydrogen bond restraints were added between one of the carboxylate groups of Asp and titrating amide residues. The O to H restraints were 2.0 or 2.2 Å, depending on the change in chemical shift in the pH titration. Following initial minimization in vacuum, the peptides were soaked with a 7 Å layer of water. The assembly was subjected to around 200 ps of dynamics simulations with 1 fs time steps at 300 K, and history files were stored every 100 fs. Charges, cross terms, and Morse potentials were used with water, and charges were turned off for calculations in vacuum. The dynamics simulations were analyzed by plotting (ϕ,ψ) angles for each amino acid and all long-range NMR distance restraints vs time. At steady state (e.g., when the total energy equilibrated), clusters of 30 equally spaced structures from the dynamics trajectory were energy mini-

mized and superimposed using various indicated backbone atoms.

Enzyme Assays: Furin. Sixty microliters of 1 mM LS-2 or LS-2[S] was incubated with 1.44 μg of furin, equivalent to an activity against pERTKR-AMC of 0.0304 μmol of MCA/h (for assay and purification details, see ref *52*) in 50 mM HEPES, pH 7.0, 2.5 mM $CaCl_2$, 0.1% Brij 35 at 37 °C, in a total volume of 120 μL. At the indicated times, 10 μL aliquots were removed, and the reaction was stopped by reducing the pH to 2.0 with 90 μL of 0.1% TFA. The aliquots were immediately frozen at −80 °C prior to analysis.

Enzyme Assays: PC2. Sixty microliters of 1 mM LS-2 or LS-2[S] was incubated with 1.71 μg of PC2 (equivalent to an activity against pERTKR-AMC of 0.0497 μmol of MCA/h: for assay and purification details, see ref *33*) in 50 mM sodium acetate, pH 5.0, 2.0 mM $CaCl_2$, 0.1% octyl glucoside at 37 °C, in a total volume of 120 μL. At the indicated times, 10 μL aliquots were removed, and the reaction was stopped by reducing the pH to 2.0 with 90 μL of 0.1% TFA. The aliquots were immediately frozen at −80 °C prior to analysis.

HPLC was performed on the furin and PC2 digests using a Thermo Separation Products P400 quaternary pump, UV1000 detector, and AS3000 autosampler under the control of a PC running PC1000 software. The column was a Vydac 218TP54 with a Phenomenex C18 SecurityGuard guard column. The sample (50 μL total) was eluted with 100% A for 10 min, followed by a gradient of 0−100% B over 40 min, where A was 0.1% (v/v) TFA in water and B was 0.1% TFA in 80% (v/v) ACN.

RESULTS

The Native RGD-Containing Peptide Does Not Inhibit Platelet Aggregation. Lymnaea stagnalis produces a FLP precursor protein with two alternatively spliced transcripts, one that produces tetrapeptides (primarily FMRF-NH_2) and another that produces heptapeptides (*35*); the two transcripts are expressed in different sets of neurons (*53*). The heptapeptide-producing precursor protein contains seven RGD sequences that separate a series of adjacent neuropeptides. To test for RGD activity, we synthesized a fragment of the native precursor containing one RGD site separating two duplicate neuropeptides: GDPFLRFGRGDPFLRF-NH_2 (LS-2).

LS-2 had no effect on platelet aggregation at concentrations as high as 0.8 mM (Figure 1). GRGDS is part of the native fibronectin sequence and has become a frequently used positive control for inhibition of ADP-induced platelet aggregation assays (*43*); therefore, we synthesized a variant of LS-2 with a serine substitution following the RGD sequence: GDPFLRFGRGDSPFLRF-NH_2 (LS-2[S]). LS-2[S] inhibited platelet aggregation at concentrations as low as 0.2 mM (Figure 1), and the positive control GRGDS produced similar inhibition. As negative controls, the mature peptides produced by LS-2 and LS-2[S] (GDPFLRF-NH_2 and GDSPFLRF-NH_2, respectively) (data not shown) and GRGES were assayed and produced no effect on platelet aggregation.

Thus, the native LS-2 sequence has no interaction with GPIIb/IIIa integrins, but the serine substitution in LS-2[S] produced a molecule that does bind to GPIIb/IIIa integrins. Clearly, the proline following the RGD site in LS-2 plays a role in the different integrin binding activities (*1*), and in the next section we describe structural changes between the two peptides that can explain the different functions.

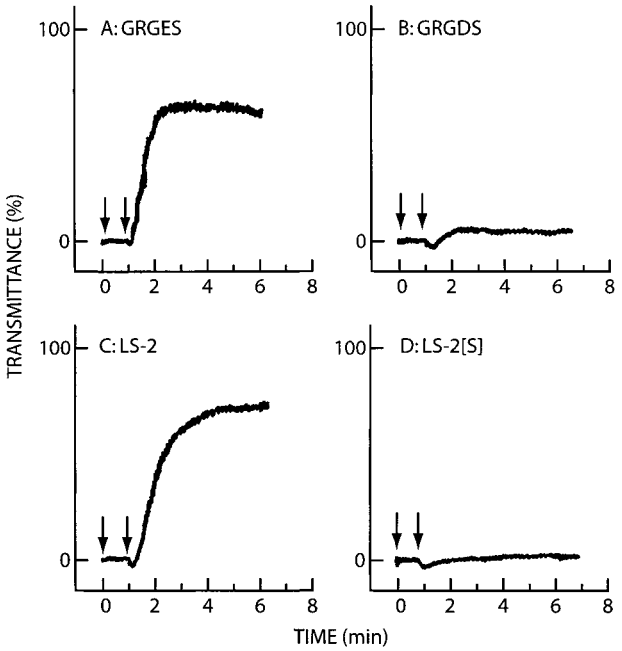

FIGURE 1: Plots of percent transmission vs time indicating platelet aggregation in the presence of the peptides used in this study. (A) 0.4 mM GRGES (negative control); (B) 0.4 mM GRGDS (positive control); (C) 0.2 mM LS-2; (D) 0.2 mM LS-2[S]. The first arrow indicates the addition of peptide, and the second arrow indicates the addition of ADP.

LS-2 and LS-2[S] Have Significantly Different RGD Conformations. Both LS-2 and LS-2[S] are relatively short linear peptides that have significant flexibility in solution. However, NMR can be used to identify highly populated regions of secondary structure (*54*, *55*), as recently demonstrated by an analysis of reverse turn conformations in a series of mature neuropeptides (*56*, *57*).

Figure 2 shows the amide to amide and amide to side-chain regions of NOESY spectra for LS-2 and LS-2[S]. NOESY spectra provide information on interatomic distances and, in conjunction with a TOCSY spectrum, allow for the determination of sequential resonance assignments. Chemical shifts are sensitive indicators of molecular environment, and differences from random-coil values (*58*) can provide a qualitative assessment of peptide structure (Figure 3). Several features of Figure 3 are worth noting. First, LS-2 contains two adjacent and identical neuropeptides (GDPFLRF) separated by a processing site. Disregarding the N-terminus, the two halves of LS-2 have nearly identical patterns of chemical shifts that are very similar to those measured in the mature peptide, GDPFLRF-NH_2 (*56*). Second, the N-terminal half of LS-2[S] is almost identical to the corresponding region in LS-2. The C-terminal half that contains the serine insertion differs somewhat from LS-2, especially the H^N of F14 in LS-2[S] compared to the H^N of F13 in LS-2.

The medium- and long-range NOEs in Figure 2 are summarized in Figure 4. Both peptides have significant numbers of long-range NOEs around the prolines, consistent with highly populated turns around those regions. The major difference between the NOE patterns of the two peptides is in the region spanning the central RGD site. LS-2 has no

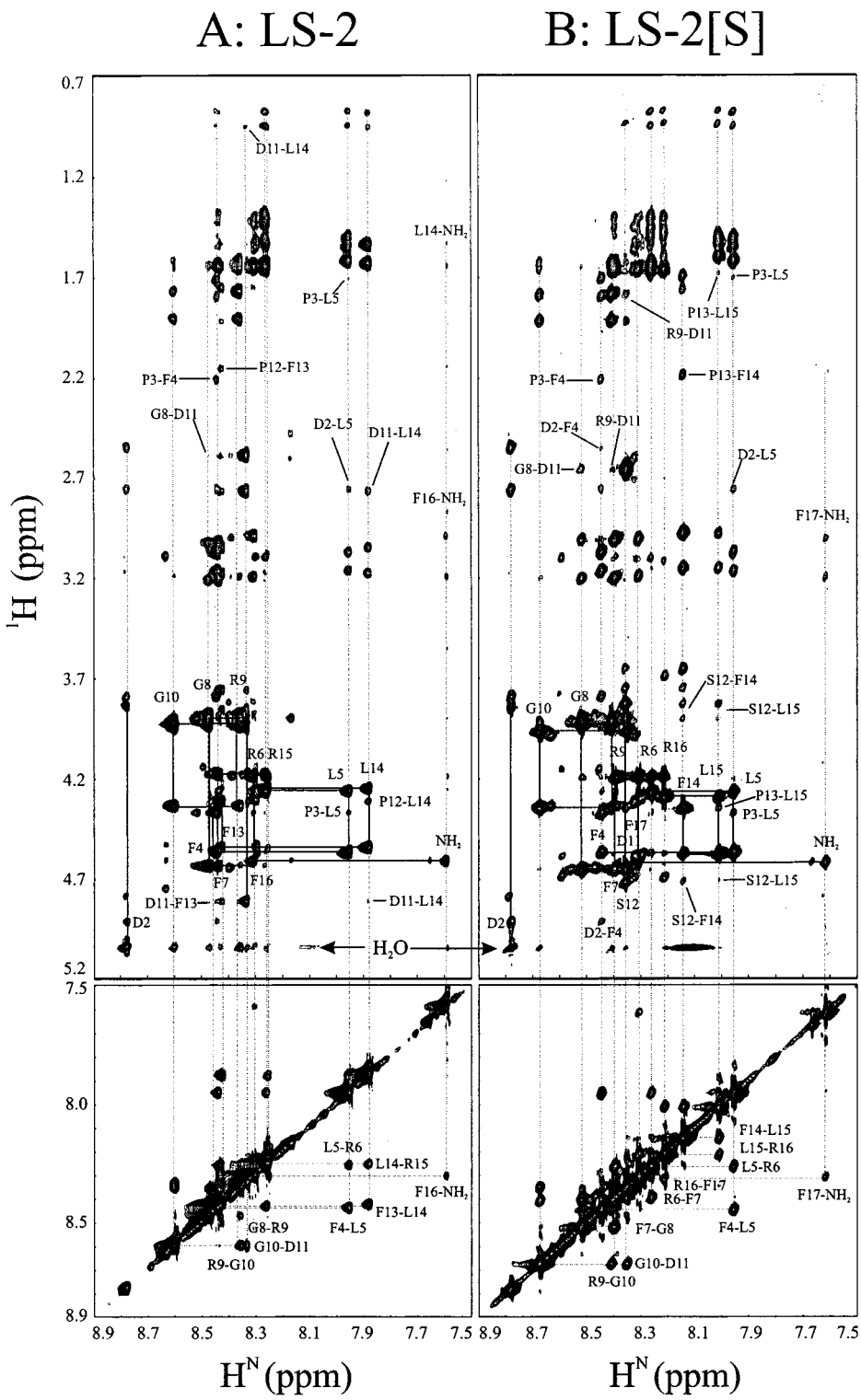

FIGURE 2: Side chain to amide (top) and amide to amide (bottom) regions of NOESY spectra of LS-2 (A) and LS2-2[S] (B). The sequential assignments are indicated in the fingerprint region of the top panels, with the amino acid labels referring to positions of the amide protons. The data were collected at a field strength of 750 MHz (17.6 T) and a temperature of 4 °C at pH 5.5 with a NOESY mixing time of 250 ms.

long-range NOEs across the RGD residues, but LS-2[S] has several, including two *i*,*i*+3 interactions (Figure 4).

In previous studies involving the fully processed peptides of LS-2, we showed that pH titrations provided detailed information about specific interactions involving titrating groups, and thus helped to better define the type of turn present in solution (*56*). Figure 5 shows the amide regions of 1D ^{1}H pH titrations of LS-2 and LS-2[S], and Figure 6 is a summary of the changes in chemical shifts of the titrating peaks between high and low pH. The summary was made by completely assigning each peptide at low (∼2.3), medium (∼4.0), and high (∼5.5) values of pH (data not shown). Like the chemical shifts at high pH (Figure 3), the N- and C-terminal halves of LS-2 and the N-terminal half of LS-2[S] are nearly identical in their pH dependence. Most prominent is the large downfield shift of the amide protons

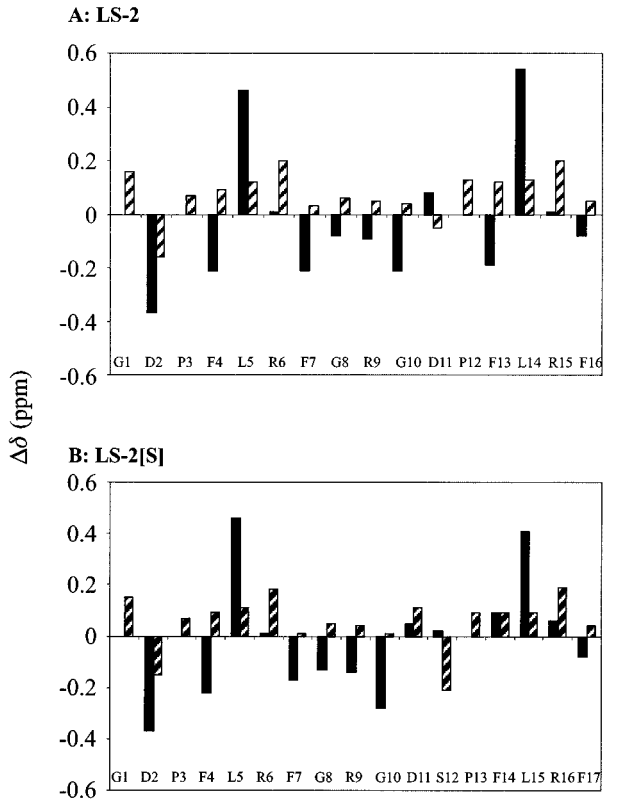

FIGURE 3: Deviations from random-coil values (*58*) of H^N (solid) and H^α (hatched) chemical shifts. Vertical axes are in ppm, and horizontal axes represent the amino acid sequence of LS-2 (A) and LS-2[S] (B). The data were collected at 750 MHz, pH 5.5, and 4 °C.

FIGURE 4: Summary of the NOEs from NMR data: Short-, medium-, and long-range NOE cross-peaks are for LS-2 (top) and LS-2[S] (bottom). The heights of the bars indicate the magnitude of the interaction.

of F4 and F13. As with the mature processed peptide, GDPFLRF-NH_2, the shift in the Phe (F) following the Pro (P) represents a high population of hydrogen bonded species of the Asp (D) side chain with the amide of Phe (F) (*56*). The insertion of Ser (S) between the Asp (D) and Pro (P) in LS-2[S] eliminates the pH effect on F14, but in LS-2[S] the amide proton of G8 is significantly shifted with pH. We interpret this shift as arising from interactions between the side chain of D11 and the amide proton of G8, and this interpretation is consistent with the patterns of NOEs shown in Figure 4.

Distances from the NOE data and pH titrations (Table 2) were estimated as described under Materials and Methods and used as extra restraining potential functions in molecular dynamics simulations. The dynamics simulations were done for about 200 ps at a temperature of 300 K and with a 7 Å layer of water solvating the peptides. Families of 30 structures spaced evenly over the dynamics trajectory for both LS-2 and LS-2[S] were energy minimized in the presence of water and superimposed, as shown in Figure 7.

In the left panels (A and C) of Figure 7, the N-terminal heavy atoms from residues G1 through G8 for both peptides are superimposed, and in the right panels (B and D) of Figure 7, the C-terminal heavy atoms from residues P12 through F16 for LS-2 and from S12 through F17 for LS-2[S] are superimposed. The superpositions show that each neuropeptide, corresponding to the two halves of each molecule, is relatively well-structured. The overall structure of each half is consistent with a type I reverse turn, as described in previous studies on the fully processed peptides (*56*, *57*). However, the two halves of both LS-2 and LS-2[S] are disordered with respect to each other, as shown by the complete lack of overlap in the half that was not used to define each superposition.

The major structural difference between LS-2 and LS-2[S] is the orientation of the Asp in the RGD. These differences are a result of and consistent with the different NOE and pH constraints used in the molecular dynamics simulations. Figure 8 shows the central region of single energy minimized structures from LS-2 and LS-2[S]. In LS-2, the Asp of the RGD is interacting strongly with C-terminal residues, especially the amide proton of the Phe following P12. In LS-2[S], the Asp of the RGD is interacting strongly with N-terminal residues, especially G8. The different interactions of the Asp of the RGD site lead to large differences in the overall conformation of the RGD. In LS-2, the RGD is extended, but in LS-2[S] the RGD forms a reverse turn because of the interactions of the Asp side chain with residues preceding the RGD. These differences in conformation correlate well with differences in integrin binding described above: the extended RGD in LS-2 is unable to inhibit ADP-induced platelet aggregation, but the reverse turn RGD in LS-2[S] is a potent inhibitor.

LS-2 and LS-2[S] Are Selectively Processed by Furin and PC2. The processing enzyme(s) for the *L. stagnalis* heptapeptide precursor protein has (have) not yet been identified, but based on their widespread neuronal distribution and presence in *L. stagnalis*, possible candidate enzymes include PC2 (*59*) and furin (*60*, *61*). Therefore, we decided to test both LS-2 and LS-2[S] as substrates for these enzymes. Given that LS-2 is only a small portion of the full-length precursor, our aim in these experiments was to qualitatively answer the following questions. First, do the enzymes selectively produce the expected products? Second, are there any major differences in processing between LS-2 and LS-2[S] for a given enzyme?

Figure 9 shows results of the in vitro processing of LS-2 (left) and LS-2[S] (right) using PC2 (top) or furin (bottom). Substrates were incubated with each enzyme for various lengths of time, and the resulting products were then separated by HPLC and identified by mass spectrometry. Both substrate sequences contain, in addition to the predicted monobasic processing site, two additional Arg residues that form the C-terminus of the mature peptides. In fact FLPs are often defined as peptides containing a C-terminal RF-

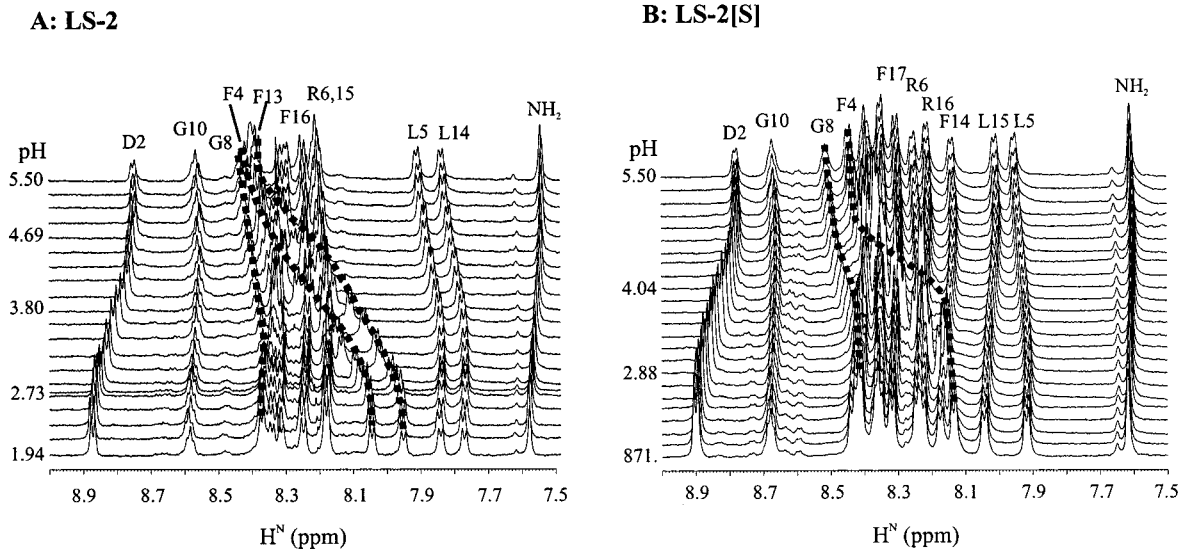

FIGURE 5: Amide regions of 1D pH titrations for LS-2 (A) and LS-2[S] (B). Series of 1D ^{1}H NMR spectra are drawn from the lowest to the highest pH (bottom to top), with some pH values indicated. All data were collected at 600 MHz and 4 °C.

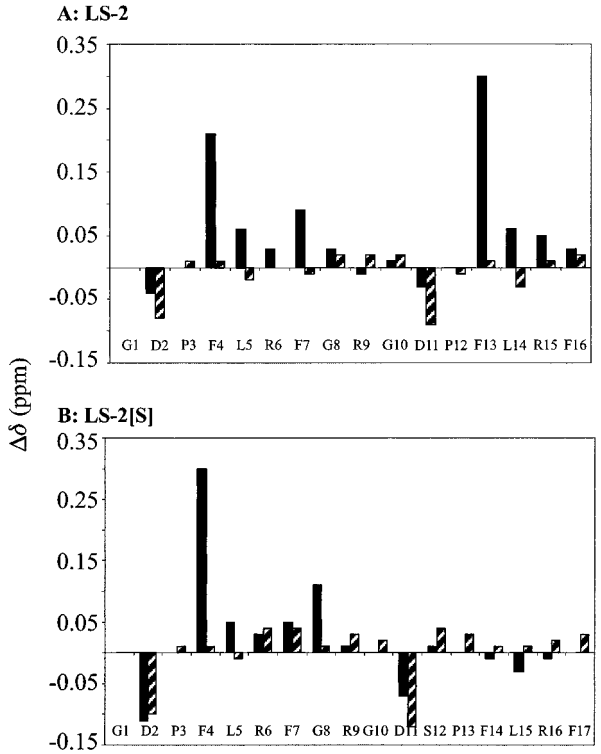

FIGURE 6: Summary of changes in H^N (solid) and H^α (hatched) chemical shifts with change in pH. Vertical axes are in ppm, and horizontal axes represent the amino acid sequence of LS-2 (A) and LS-2[S] (B). Histograms represent differences between high- and low-pH NMR chemical shifts.

Table 2: NMR-Based Restraints Used in Molecular Dynamics Simulations

LS-2		LS-2[S]	
residue	distance (Å)	residue	distance (Å)
D2βL5 H^N	3.4,4.1	D2αF4 H^N	3.4
D2βL5δ	3.1,3.1	D2βL5δ	3.4,3.4
P3αL5 H^N	3.4	D2βF4 H^N	3.5,3.9
P3αR6 H^N	3.3,3.5	P3αR6 H^N	3.8
P3βL5β	3.6,3.8	P3γL5 H^N	4.0
F4αR6 H^N	3.6	F4αR6 H^N	3.5
L5δG8 H^N	3.5,3.6	L5αF7 H^N	3.3
R6βG8 H^N	3.4	R6αG8 H^N	3.5
F7βR9 H^N	3.2,3.8	F7βG10 H^N	3.8
D11αF13 H^N	3.1	G8αG10 H^N	3.5
D11βF13 H^N	3.2,3.3	G10αL15δ	3.5,3.8
D11αL14 H^N	2.6,3.0	D11βR9 H^N	3.6
D11βL14 H^N	3.2,3.9	D11βG8 H^N	3.5,3.6
D11βL14β	3.0,3.4	D11βL15δ	3.5,3.8
D11βL14δ	3.3,3.4	S12αF14 H^N	3.4
P12αL14 H^N	3.1	S12βF14 H^N	3.4,3.6
P12αR15 H^N	3.7,3.8	S12αL15 H^N	3.8
P12βL14β	3.9,4.1	S12βL15 H^N	3.4,3.8
F13αR15 H^N	3.3	S12βL15β	3.5,3.7
L14δD11H^N	3.5,4.2	S12βL15δ	3.6,3.9
D2-COO^-F4 H^N	2.0	P13αL15 H^N	3.4
D11-COO^-F13 H^N	2.0	P13γL15 H^N	3.9
		F14αR16 H^N	3.6
		D2-COO^-F4 H^N	2.0
		D11-COO^-G8 H^N	2.2

NH_2 sequence (*62*). In the unprocessed FLP precursor protein, therefore, all peptides contain the consensus sequence "RFG(R/K)".

PC2 is known for its ability to cleave both dibasic and monobasic sites (*63*), and furin has a consensus processing site of RXXR (*64*). Figure 9 demonstrates that both LS-2 and LS-2[S] are properly cleaved at the central Arg processing site by both PC2 and furin. Despite the presence of extra Arg residues that, in theory, might be cleaved by PC2, both PC2 and furin produced only the expected products. Both enzymes appear to cleave with similar rates, with furin cleaving both substrates slightly more slowly than PC2 in these assays. A quantitative determination of specific rates of processing of each site is beyond the scope of this study.

DISCUSSION

Because of their ability to bind to integrins, RGD peptides have been the subject of intense academic and commercial investigation for several years. Recent research has shown that RGD peptides are also involved in protein sorting in the secretory pathway, but the molecular mechanism of the sorting remains unknown (*6*). The observation that several neuropeptide precursor proteins contain RGD sequences at their proteolytic processing sites suggests the possibility that precursor processing and sorting could be coupled. Our

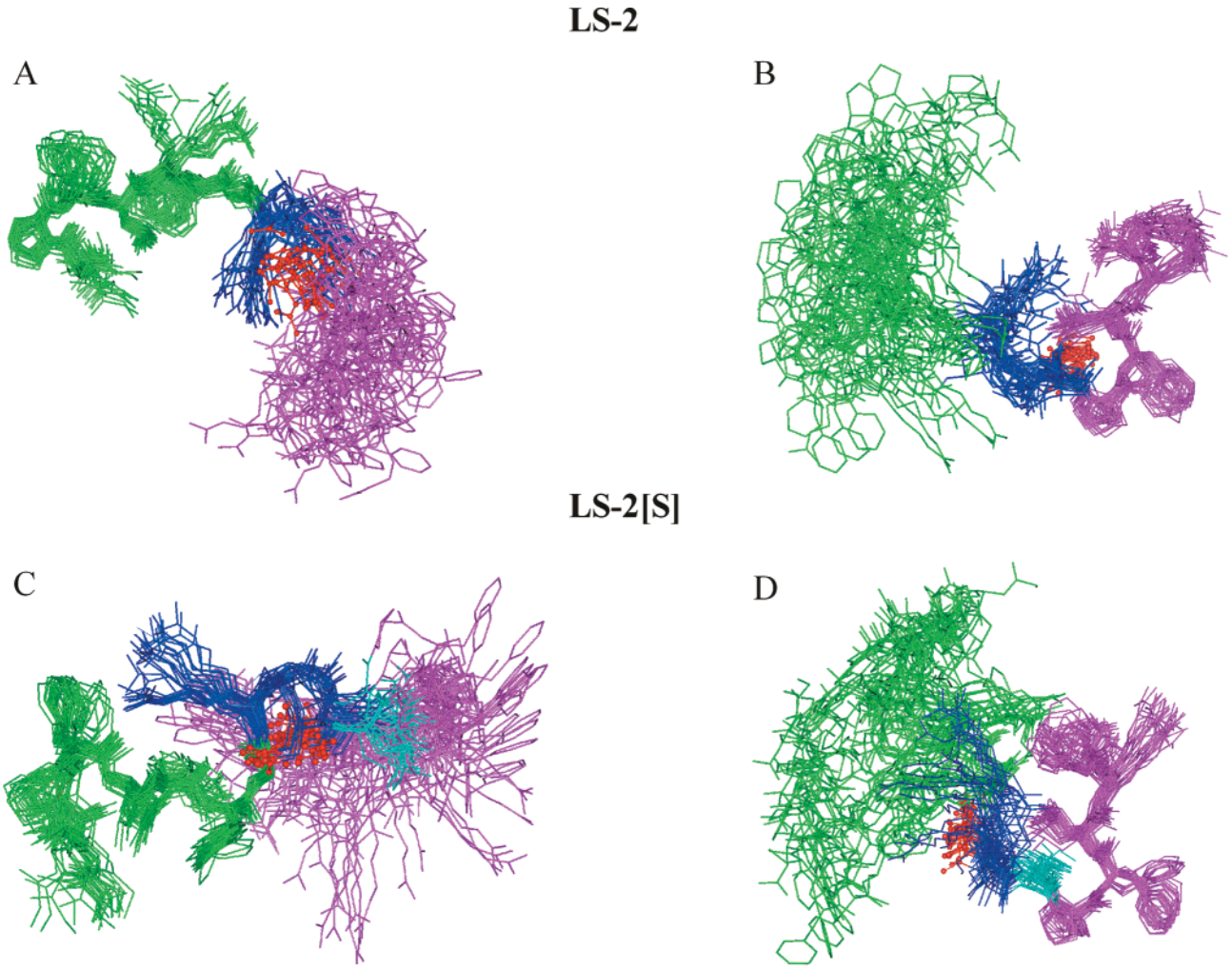

FIGURE 7: Families of structures from molecular dynamics simulations. Structures from 30 equally spaced time points in the molecular dynamics trajectory were energy minimized and superimposed as follows: In panels A and C, the heavy atoms of G1 through G8 were superimposed for both LS-2 (top) and LS-2[S] (bottom). In panels B and D, the heavy atoms of P12−F16 (LS-2, top) or S12−F17 (LS-2[S], bottom) were superimposed. The first neuropeptide (GDPFLRF) is colored green, the RGD sequences are in blue, and the second neuropeptide is shown in purple. The carboxylate groups of the Asp in the RGD sequences are shown as balls and are in red. The Ser of LS-2[S] is light blue. All calculations were done in the presence of a 7 Å layer of water which was removed in the figure for clarity.

results show that LS-2, a fragment of a longer precursor protein, does not bind to GPIIb/IIIa integrins. However, we cannot rule out the possibility that LS-2 could interact with other types of integrins.

The structural differences between LS-2 and LS-2[S] provide a mechanism for the lack of GPIIb/IIIa integrin binding activity of RGDP sequences. An earlier study showed that insertion of Pro (P) after RGD eliminated inhibition of cell adhesion (*1*), and we can explain this phenomenon by the following observations. The LS-2 sequence, GDPFLRFGRGDPFLRF-NH_2, is essentially a duplicate of two GDPFLRFG sequences separated by R. We have shown in previous studies that the peptide GDPFLRF-NH_2 forms a type I reverse turn involving residues DPFL (*56*, *57*). The NMR spectra of LS-2 and GDPFLRF-NH_2 are very similar, including the large pH dependence of the Phe (F) residues in the DPFL sequence (Figure 5 and ref *56*). The large pH effects demonstrate quite clearly that in the DPFL sequence, the Asp (D) side chain is hydrogen bonding to the amide proton of Phe (F). This pH effect is duplicated for both halves of LS-2. Thus, the structure of LS-2 is identical to two linked GDPFLRF-NH_2 halves, and the linkage is flexible or disordered (Figure 7). In contrast, insertion of a Ser between the Asp and Pro in LS-2[S] (GDPFLRFGRGDSPFLRF-NH_2) blocks the ability for the Asp side chain to interact in a reverse turn across Pro. In LS-2[S], the Asp (D) of the RGD sequence interacts with the previous amino acids as far back as G8 to help stabilize a reverse turn around the RGD site (Figure 8). The activity of RGD sequences is enhanced by a more fully populated reverse turn around the RGD site (*24*, *25*). A Pro residue following an RGD sequence creates the perfect geometry for the Asp to H-bond across the Pro into a type I reverse turn and thus *prevents* the Asp from interacting with previous residues to define a turn around the RGD sequence.

In addition to demonstrating that the neuropeptide GDP-FLRF-NH_2 has a high percentage of type I reverse turn, our previous work also found that in a series of related similar neuropeptides, the population of reverse turn was strongly anti-correlated to receptor binding affinities (*56*). The overall conformations of the LS-2 and LS-2[S] peptides are similar in the regions containing the mature peptides, which appear to adopt type I reverse turns around the DPFL and SPFL sequences. These conformations are similar to those found

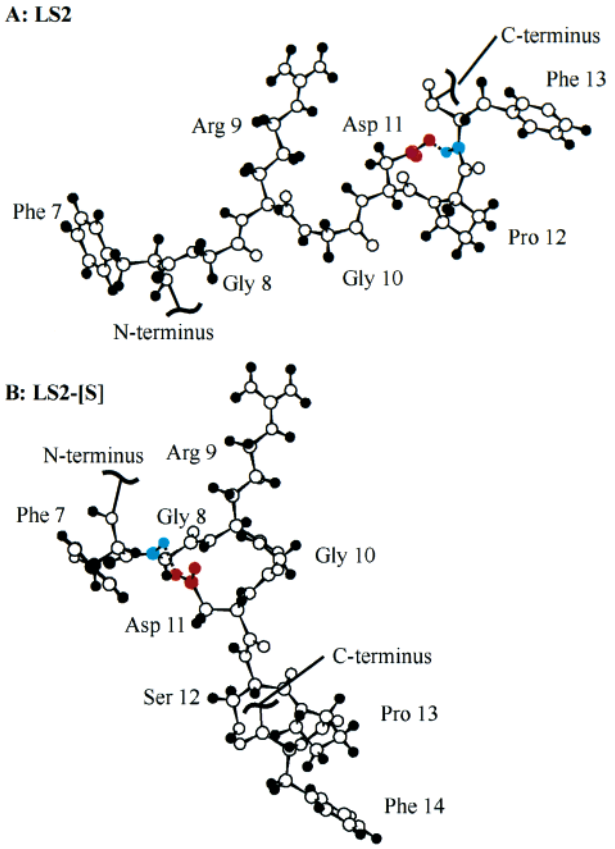

FIGURE 8: Schematic representation of differences in structure between LS-2 (top) and LS-2[S] (bottom). The two structures are representative minimized structures from the superposition in Figure 7. The Asp carboxylate groups are colored in red, and their nearest H-bonded neighbors are shown in blue. The Asp in the LS-2 peptide interacts with groups C-terminal to the RGD sequence, leaving the RGD sequence disordered or flexible. In contrast, the Asp in the LS-2[S] peptide interacts with residues N-terminal to the RGD sequence, and this interaction stabilizes a turn around the RGD region.

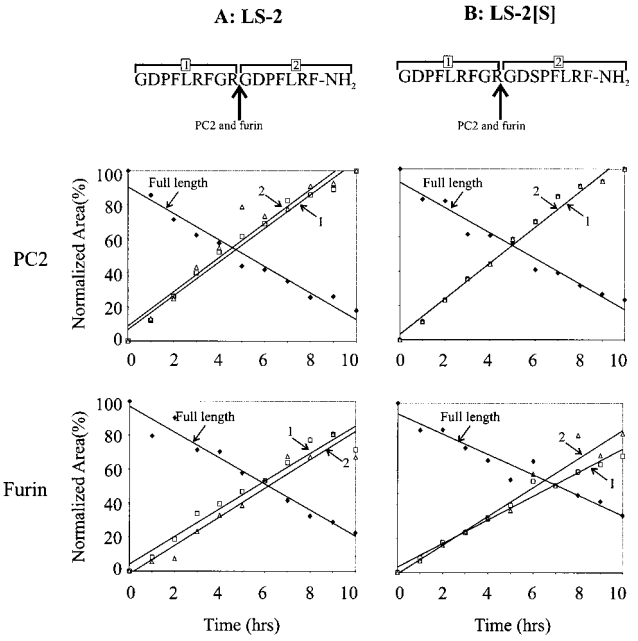

FIGURE 9: Processing experiments of LS-2 (A: left panels) and LS-2[S] (B: right panels) with PC2 (top panels) and furin (bottom panels). The digests are presented as normalized to the starting absorbance of the full-length precursor. The products were analyzed by MALDI-TOF mass spectrometry. The cleavage site and products of the incubations are indicated above.

in the fully processed neuropeptides (*56*, *57*). LS-2 and LS-2[S] have significant differences around the central RGD region, which is also the site of proteolytic processing. Our results provide convincing evidence that the native RGDP sequence of LS-2 does not interact with platelet GPIIb/IIIa integrin receptors. LS-2 is representative of all of the RGD sequences on the full-length *L. stagnalis* heptapeptide precursor, so we consider it unlikely that the precursor interacts with GPIIb/IIIa integrins in vivo. However, there are many other types of integrin receptors that participate in a wide variety of biological functions (*2*, *65*). Our results do not rule out the possibility that the RGD sequences of the *L. stagnalis* precursor interact with another integrin. The alternatively spliced *L. stagnalis* tetrapeptide precursor protein (*66*) also contains an RGDE sequence, and we suggest that structurally this region might behave more like the RGDS in LS-2[S] and thus might be capable of binding to GPIIb/IIIa integrins.

Much of the previous work on structural requirements for prohormone processing was based on dibasic processing sites, and these studies suggest that reverse turns are required at the site of processing (*38*−*42*). Both LS-2 and LS-2[S] possess monobasic processing sites which form part of the RGD sequence, and are processed correctly by PC2 and furin. As noted before, LS-2 and LS-2[S] have extended and reverse turn structures around the RGD sequence, respectively. However, given these conformational differences, we see very little difference in processing by PC2 or furin. Perhaps the structural requirements for monobasic processing sites are less important than flexibility in the region (Figure 7).

The native processing enzyme of the *L. stagnalis* FMRF-amide-like protein (FLP) from which LS-2 is derived has not been identified, but based on the general neuronal distribution of convertases (reviewed in *29*, *30*, and *32*) and known *L. stagnalis* enzymes, PC2 and furin are good in vivo candidates. Previous studies have shown that PC2 or LPC2 (*59*), furin or Lfurin2 (*60*), and Lfurin 1/Lfurin 1-X (*61*) are present in *Lymnaea*. FLPs have been identified with both mono- and dibasic sites between multiple neuropeptides. Examples of the former are flp-1A and flp-1B from *C. elegans* (*67*), precursors from *D. melanogaster* (*68*), and the *L. stagnalis* precursor from which LS-2 is derived. Examples of the latter are flp-1A and flp-1B from *C. elegans* (*67*) or afp-1 from *A. suum* (*37*). Common to all FLP precursor proteins is another Arg residue positioned four amino acids N-terminal to the processing site (e.g., RFG[K/R]), and a significant outstanding question in FLP processing has been how the upstream Arg escapes proteolysis. This question is especially intriguing in monobasic FLP precursors such as LS-2 that contain, for each neuropeptide, two closely spaced single Arg residues, only one of which is processed. Our results with furin suggest that the specificity of FLP processing might include both Arg residues and that it is incorrect to consider FLP precursors as strictly “monobasic” or “dibasic”.

PC2 cleaves the LS-2 and LS-2[S] substrates at the central Arg of the RGD sequence but not at the first Arg of the RFGR sequence. This is the same specificity generally predicted—and observed in this study—for furin. Many FLP

precursors, including the full-length *L. stagnalis* heptapeptide precursor, contain a single tetrabasic site that is thought to be cleaved by furin. These unique furin sites have perhaps diverted attention from the fact that all FLP processing sites are consensus furin sites (RXXR). In the case of *Aplysia* pro egg-laying hormone (ELH), the first cleavage is by furin and occurs at the RRKR tetrabasic site, and subsequent cleavages are at dibasic sites by PC1 enzymes (*69*, *70*). The furin and PC2 processing results (Figure 9) for both LS-2 and LS-2[S] produce the expected neuropeptide products, i.e., those observed in vivo. Thus, our results clearly demonstrate that, in addition to PC2, furin or a furin-like protein needs to be considered as a candidate for any FLP protease.

In conclusion, we have demonstrated significant structural differences between two RGD peptides that differ by insertion of a single serine following the RGD sequence; these structural changes appear to explain differences in integrin binding. In contrast, these structural changes appear to have little effect on processing. The RGDP sequences in the *L. stagnalis* heptapeptide precursor protein do not interact with GPIIb/IIIa integrins, but we cannot rule out interactions with other integrins. We also demonstrated that LS-2 and LS-2[S] are both processed properly by PC2 and furin. Thus, we suggest that both enzymes are possible candidates for FMRF-amide-like peptide processing.

ACKNOWLEDGMENT

We thank Dr. Adrian Roitenberg for helpful discussions about molecular modeling and Dr. Nabil Seidah for discussions about furin specificity.

REFERENCES

1. Pierschbacher, M. D., and Rouslahti, E. (1987) *J. Biol. Chem. 262* (36), 17294−17298.
2. Hynes, R. O. (1992) *Cell 69*, 11−25.
3. Ginsberg, M. H., Pierschbacher, M. D., Ruoslahti, E., Marguerie, G., and Plow, E. (1985) *J. Biol. Chem. 260*, 3931−3936.
4. Mogford, J. E., Davis, G. E., Platts, S. H., and Meininger, G. A. (1996) *Circ. Res. 79*, 821−826.
5. Platts, S. H., Mogford, J. E., Davis, M. J., and Meininger, G. A. (1998) *Am. J. Physiol. 275*, H1449−H1454.
6. Rovère, C., Luis, J., Lissitzky, J. C., Basak, A., Marvaldi, J., Chretien, M., and Seidah, N. G. (1999) *J. Biol. Chem. 274*, 12461−12467.
7. Jones, L. S. (1996) *Trends Neurosci. 19*, 68−72.
8. Bowen, J. A., and Hunt, J. S. (2000) *Proc. Soc. Exp. Biol Med. 223*, 331−343.
9. Saudek, V. R., Atkinson, A., and Pelton, J. T. (1991) *Biochemistry 30*, 7369−7372.
10. Chen, Y., Pitzenberger, S. M., Garsky, V. M., Lumma, P. K., Sanyal, G., and Baum, J. (1991) *Biochemistry 30*, 11625−11636.
11. Adler, M., Lazarus, R. A., Dennis, M. S., and Wagner, G. (1991) *Science 253*, 445−448.
12. Adler, M., and Wagner, G. (1992) *Biochemistry 31*, 1031−1039.
13. McDowell, R. S., Dennis, M. S., Louie, A., Shuster, M., Mulkerrin, M. G., and Lazarus, R. A. (1992) *Biochemistry 31*, 4766−4772.
14. Senn, H., and Klaus, W. (1993) *J. Mol. Biol. 232*, 907−925.
15. Jaseja, M., Lu, X., Williams, J. A., Sutcliffe, M. J., Kakkar, V. V., Parslow, R. A., and Hyde, E. I. (1994) *Eur. J. Biochem. 226*, 861−868.
16. Krezel, A. M., Wagner, G., Ulmer, J. S., and Lazarus, R. A. (1994) *Science 264*, 1944−1947.
17. Atkinson, R. A., Saudek, V., and Pelton, J. T. (1994) *Int. J. Pept. Protein Res. 43*, 563−572.
18. Smith, K. J., Jaseja, M., Lu, X., Williams, J. A., Hyde, E. I., and Trayer, I. P. (1996) *Int. J. Pept. Protein Res. 48*, 220−228.
19. Krezel, A. M., Ulmer, J. S., Wagner, G., and Lazarus, R. A. (2000) *Protein Sci. 9*, 1428−1438.
20. Mickos, H., Bahr, J., and Lüing, B. (1990) *Acta Chem. Scand. 44*, 161−164.
21. Aumailley, M., Gurrath, M., Müller, G., Calvete, J., Timpl, R., and Kessler, H. (1991) *FEBS Lett. 291*, 50−54.
22. Bogusky, M. J., Naylor, A. M., Pitzrnberger, S. M., Nutt, R. F., Brady, S. F., Colton, C. D., Sisko, J. T., Anderson, P. S., and Veber, D. F. (1992) *Int. J. Pept. Protein Res. 39*, 63−76.
23. Gurrath, M., Müller, G., Kessler, H., Aumailley, M., and Timpl, R. (1992) *Eur. J. Biochem. 210*, 911−921.
24. Kopple, K. D., Baures, P. W., Bean, J. W., D'Ambrosio, C. A., Hughes, J. L., Peishoff, C. E., and Eggleston, D. S. (1992) *J. Am. Chem. Soc. 114*, 9615−9623.
25. Peishoff, C. E., Ali, F. E., Bean, J. W., Calvo, R., D'Ambrosio, C. A., Eggleston, D. S., et al. (1992) *J. Med. Chem. 35*, 3962−3969.
26. Johnson, C. W., Jr., Pagano, T. G., Basson, C. T., Madri, J. A., Gooley, P., and Armitage, I. M. (1993) *Biochemistry 32*, 268−273.
27. Mayo, K. H., Fan, F., Beavers, M. P., Eckardt, A., Keane, P., Hoekstra, W. J., and Andrade-Gorden, P. (1996) *Biochim. Biophys. Acta 1296*, 95−102.
28. Miyashita, M., Akamatsu, M., Hayashi, Y., and Ueno, T. (2000) *Bioorg. Med. Chem. Lett. 10*, 859−863.
29. Beinfeld, M. C. (1998) *Endocrine 8* (1), 1−5.
30. Seidah, N. G., and Chrétien, M. (1999) *Brain Res. 848*, 45−62.
31. Zhau, A., Webb, G., Zhu, X., and Steiner, D. F. (1999) *J. Biol. Chem. 274* (30), 20745−20748.
32. Steiner, D. F. (1998) *Curr. Opin. Chem. Biol. 2*, 31−39.
33. Johanning, K., Juliano, M. A., Juliano, L., Lazure, C., Lamango, N. S., Steiner, D. F., and Lindberg, I. (1998) *J. Biol. Chem. 273*, 22672−22680.
34. Day, R., Lazure, C., Basak, A., Boudreault, A., Limperis, P., Dong, W. J., and Lindberg, I. (1998) *J. Biol. Chem. 273*, 829−836.
35. Saunders, S. E., Bright, K., Kellet, E., Benjamin, P. R., and Burke, J. F. (1991) *J. Neurosci. 11*, 740−745.
36. Price, D. A. (1993) GenBank Accession No. AAA03426.
37. Edison, A. S., Messinger, L. A., and Stretton, A. O. (1997) *Peptides 18*, 929−935.
38. Rholam, M., Nicolas, P., and Cohen, P. (1986) *FEBS Lett. 207*, 1−6.
39. Brakch, N., Boileau, G., Simonetti, M., Nault, C., Joseph-Bravo, P., Rholam, M., and Cohen, P. (1993a) *Eur. J. Biochem. 216*, 39−47.
40. Brakch, N., Rholam, M., Boussetta, H., and Cohen, P. (1993b) *Biochemistry 32*, 4925−4930.
41. Di Bello, C., Simonetti, M., Dettin, M., Paolillo, L., D'Aurla, G., Falcigno, L., Saviano, M., Scatturin, A., Vertuani, G., and Cohen, P. (1995) *J. Pept. Sci. 1*, 251−265.
42. Brakch, N., Rholam, M., Simonetti, M., and Cohen, P. (2000) *Eur. J. Biochem. 267*, 1626−1632.
43. Ugen, K. E., Mahalingam, M., Klien, P. A., and Kao, K. J. (1988) *J. Nat. Cancer Inst. 80* (18), 1461−1466.
44. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
45. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
46. Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
47. Callihan, D., West, J., Kumar, S., Scheitzer, B. I., and Logan, T. M. (1996) *J. Magn. Reson. 112*, 82−85.
48. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
49. Delagalio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.

50. Wüthrich, K. (1986) in *NMR of proteins and nucleic acids*, Wiley, New York.
51. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−614.
52. Cameron, A., Appel, J., Houghten, R. A., and Lindberg, I. (2000) *J. Biol. Chem. 47*, 36741−36749.
53. Saunders, S. E., Kellet, E., Bright, K., Benjamin, P. R., and Burke, J. F. (1992) *J. Neurosci. 12*, 1033−1039.
54. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., and Wright, P. E. (1988a) *J. Mol. Biol. 201*, 161−200.
55. Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., and Lerner, R. A. (1988b) *J. Mol. Biol. 201*, 201−217.
56. Edison, A. S., Espinoza, E., and Zachariah, C. (1999) *J. Neurosci. 19*, 6318−6326.
57. Carlacci, L., and Edison, A. S. (2000) *Proteins: Struct., Funct., Genet. 40*, 367−377.
58. Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.
59. Smit, A. B., Spijker, S., and Geraerts, W. P. M. (1992) *FEBS Lett. 312*, 213−218.
60. Smit, A. B., Spijker, S., Nagle, G. T., Knock, S. L., Kurosky, A., and Geraerts, W. P. M. (1994) *FEBS Lett. 343*, 27−31.
61. Spijker, S., Smit, A. B., S. L., Sharp-Baker, H. E., Van Elk, R., Van Kesteren, E. R., Van Minnen, J., Kurosky, A., and Geraerts, W. P. M. (1999) *J. Neurobiol. 41*, 399−413.
62. Espinoza, E., Carrigan, M., Thomas, S. G., Shaw, G., and Edison, A. S. (2000) *Mol. Neurobiol. 21*, 35−56.
63. Dupuy, A., Lindberg, I., Zhou, Y., Akil, H., Lazure, C., Chrétien, M., Seidah, N. G., and Day, R. (1994) *FEBS Lett. 337*, 60−65.
64. Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R., and Thomas, G. (1992) *J. Biol. Chem. 267*, 16396−16402.
65. Clarke, E. A., and Brugge, J. S. (1995) *Science 268*, 233−239.
66. Linacre, A., Kellett, E., Saunders, S., Bright, K., Benjamin, P. R., and Burke, J. F. (1990) *J. Neurosci. 10*, 412−419.
67. Rosoff, M. L., Bürglin, T. R., and Li, C. (1992) *J. Neurosci. 12*, 2356−2361.
68. Schneider, L. E., and Taghert, P. H. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1993−1997.
69. Sossin, W. S., Fisher, J. M., and Scheller, R. H. (1989) *Neuron 2*, 1407−1417.
70. Chun, J. Y., Korner, J., Kreiner, T., Scheller, R. H., and Axel, R. (1994) *Neuron 12*, 831−844.

BI010448S