Design and Solution Structure of a Well-Folded Stack of Two β-Hairpins Based on the Amino-Terminal Fragment of Human Granulin A†

Dmitri Tolkatchev, Andy Ng, Wim Vranken, and Feng Ni*

Biomolecular NMR Laboratory and the Montreal Joint Centre for Structural Biology, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Quebec, Canada H4P 2R2

Received September 13, 1999; Revised Manuscript Received December 13, 1999

ABSTRACT: Four amino acid substitutions were introduced into a peptide corresponding to the amino-terminal subdomain (30−31 residues) of human granulin A (HGA) in order to assess the contributions of a hydrophobic framework and other interactions to structure stabilization of the stack of two β-hairpins. The resulting hybrid peptide, HGA 1−31 (D1V, K3H, S9I, Q20P) with four free cysteines, spontaneously formed a uniquely disulfide-bonded isomer with an expected [1−3, 2−4] disulfide pairing pattern. This peptide was characterized in detail by use of NMR and shown to assume a highly stable structure in solution, in contrast to the amino-terminal 1−30 fragment of human granulin A. The prototype peptide, or HGA 1−30 (C17S, C27S), had lower resistance to chemical reduction and proteolysis, broad NH and H^{α} proton resonances, lower proton resonance dispersion, and no slowly exchanging amide protons. Two other peptides, HGA 1−30 (C17S, Q20P, C27S) and HGA 1−31 (D1V, K3H, S9I, C17S, C27S), with either Pro20 stabilizing a potential reverse turn or with a hydrophobic cluster consisting of Val1, His3, and Ile9, had sharper and slightly better dispersed NH and H^{α} proton resonances, but still no slowly exchanging amide protons. The solution structure of HGA 1−31 (D1V, K3H, S9I, Q20P) indicates that it adopts a well-folded conformation of a stack of two β-hairpins, as found for the amino-terminal subdomain of the prototypic carp granulin-1 with representative β-hairpin stacks. These results highlight the importance of both hydrophobic and turn-stabilizing interactions for the structural integrity of the hairpin stack scaffold. The conformational stability appears to be maintained by a combination of the well-formed second β-hairpin and two hydrophobic clusters, one located at the interface between the two β-hairpins and the other on "top" of the first β-hairpin. The implications of these findings for the design of conformationally stable hairpin stacks are discussed.

The engineering of small polypeptide ligands that bind selectively to desired biological targets has been a long-sought goal in the field of protein chemistry and is becoming an increasingly important step in the design of protein mimetics as drug candidates (for reviews see refs *1*−*3*). One of the general approaches in this direction is to optimize selected residues within a certain conformationally constrained protein scaffold (*1*, *2*). The use of a well-structured polypeptide framework is justified by the higher binding affinities that can be expected for the folded conformations due to a smaller loss of conformational entropy upon binding. In addition, once a rigid framework is constructed, it is potentially easier to establish the details of the interactions with the target protein for a folded protein ligand than for an unstructured peptide. Structure details of intermolecular contacts provide essential information for the design of a new generation of mimetics of specific protein−protein interactions.

The smallest naturally folded protein domains in the absence of disulfide bonds or metal-binding sites are 40−60 residues long, while linear peptides shorter than 40−50 residues usually assume very flexible conformations in solution. Introduction of disulfide bonds limits the conformational freedom of mimetic peptides and can significantly improve their binding affinities (*1*, *2*, *4*). For this reason, structured disulfide-rich protein domains represent an important class of folding motifs potentially suitable for the design of minimized proteins. So far, only a few examples of small (less than 40 residues) disulfide-rich scaffolds have been engineered (*5*−*10*). It is also feasible to create a new generation of minimized proteins based on a newly discovered disulfide-rich scaffold, called the hairpin stack, found in carp granulin-1, a representative member of an emerging family of disulfide-rich proteins containing the granulin/epithelin protein repeats (*11*). Indeed, it was shown recently that a peptide corresponding to the 30-residue amino-terminal subdomain of carp granulin-1, CG1[1] 1−30 (C17S, C27S), spontaneously forms an independently folded and disulfide-reinforced stack of two β-hairpins similar to that found in the native protein (*12*). However, building affinity for selected protein targets into this minimized protein scaffold must take into account molecular interactions dictating the

† This work was supported in part by the Medical Research Council of Canada (Grant MT-12556) and by the National Research Council of Canada (NRCC Publication No. 42927).

* To whom correspondence should be addressed.

[1] Abbreviations: HGA, human granulin A; PCDGF, PC cell-derived growth factor; CG1, carp granulin-1; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; rmsd, root-mean-square deviation; TCEP, tris-(2-carboxyethyl)-phosphine.

10.1021/bi992130u CCC: $19.00
Published on Web 02/23/2000

folding of the peptide into a well-defined three-dimensional framework. It is therefore very important, before attempting the design of binding activities, to understand factors that determine the unique three-dimensional structure of the stack of two β-hairpins found for a subdomain of the granulin/epithelin protein repeat.

Granulin-like protein repeats are also found in human and rat inflammatory leukocytes and bone marrow as potential autocrine and paracrine growth modulating factors for epithelial and mesenchymal cells (*13*, *14*). The amino-terminal sequence of an abundant human granulin A is almost identical to that of rat kidney epithelin 1, which stimulates the proliferation of murine keratinocytes and is known to inhibit the growth of cancerous cells, derived from a human epidermal carcinoma (*15*−*17*). Human granulin A also shares its amino-terminal sequence with the transforming growth factor type-e (TGFe) (*18*), which acts as a mitogen for epithelial and fibroblastic cells (*19*). The granulin/epithelin precursor, containing seven and a half repeats of the granulin/epithelin sequences including granulin A (*15*, *17*), has been identified as the PC cell derived growth factor, or PCDGF, secreted by a highly tumorigenic cell line, and was demonstrated to affect proliferation of both epithelial and nonepithelial cells through specific binding to cell surface receptors (*20*−*22*). Inhibition of PCDGF expression in PC cells was shown to result in a dramatic inhibition of their tumorigenicity (*23*), demonstrating the importance of PCDGF overexpression in tumor formation. Interestingly, the granulin/epithelin precursor, in particular a fragment containing granulins A and B, has been identified as a potential host protein cofactor for the human immunodeficiency virus Tat proteins (*24*). Therefore, growth-modulating proteins containing the granulin/epithelin repeats are very attractive targets for structure−function studies through protein minimization and the design of functional mimetics.

We found that a 30 residue peptide derived from the amino-terminal domain of human granulin A has a markedly more flexible conformation in solution than its carp granulin homologue (*12*), despite that half of the residues in these peptides are identical. Analysis of the hydrophobic contacts in the three-dimensional structures of the full-length carp granulin-1 and in the corresponding amino-terminal peptide fragment of carp granulin-1 (*12*) suggests that for carp granulin-1 there exists a potentially stabilizing hydrophobic framework, or a "hydrophobic cap", formed by the three residues Val1, His3, and Ile9 in the carp granulin-1 sequence. Vranken et al. (*12*) noted that human granulin A lacks this hydrophobic cap of carp granulin-1, which could explain the reduced conformational stability of the human granulin A peptide. We therefore introduced into the amino-terminal 1−31 fragment of human granulin A selected substitutions (Asp1Val, Lys3His, and Ser9Ile), intended to restore the hydrophobic framework present in the carp analogue. Several variants of the modified peptide were prepared: in one of them Gln20 of the human granulin A sequence was left unchanged, and in another, Gln20 was replaced by Pro20, which is a part of a reverse β-turn within the central second β-hairpin present in carp granulin. We found that only the variant with the Gln20Pro substitution together with the hydrophobic cap spontaneously folded upon air oxidation into a well-defined and stable three-dimensional conformation.

MATERIALS AND METHODS

Peptide Synthesis and Purification. The amino-terminal 1−30 fragment of human granulin A, DVKCDMEVSCP-DGYTCSRLQSGAWGCSPFT, or HGA 1−30 (C17S, C27S), and five related analogues, HGA 1−30 (C17S, Q20P, C27S), HGA 1−31 (D1V, K3H, S9I), HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S), HGA 1−31 (D1V, K3H, S9I, C17S, C27S), and HGA 1−31 (D1V, K3H, S9I, Q20P), were prepared by solid-phase peptide synthesis using standard Fmoc chemistry. To avoid undesirable coupling of disulfide bonds, residues Cys17 and Cys27 of the original granulin A sequence, which are expected to connect the second and third β-hairpins in a stack of four β-hairpins, were replaced by serines in HGA 1−30 (C17S, C27S), HGA 1−30 (C17S, Q20P, C27S), HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S), and HGA 1−31 (D1V, K3H, S9I, C17S, C27S) and by *S*-acetamidomethylcysteines in HGA 1−31 (D1V, K3H, S9I) and HGA 1−31 (D1V, K3H, S9I, Q20P). Two other cysteines, Cys10 and Cys26, in HGA 1−30 (C17S, C27S), HGA 1−30 (C17S, Q20P, C27S), and HGA 1−31 (D1V, K3H, S9I, C17S, C27S) were also blocked by *S*-acetamidomethyl groups to direct the formation of disulfide bonds to those expected for a stack of two β-hairpins (*11*). The reduced forms of the peptides were prepared from the crude preparations dissolved in 0.1 M Tris-HCl, 1 mM EDTA, and 100 mM DTT, at pH 8.3 by HPLC purification using a C_{18} Vydac column and a linear 10 to 45% acetonitrile gradient in 0.1% aqueous solution of trifluroacetic acid (TFA). The regio-selective oxidation of HGA 1−30 (C17S, C27S), HGA 1−30 (C17S, Q20P, C27S), and HGA 1−31 (D1V, K3H, S9I, C17S, C27S) was performed in two steps. The first disulfide bond, or the 1−3 cystine, was formed by overnight air oxidation in 1% ammonium-acetate buffer, pH 8.5. The product was purified by HPLC, lyophilized, and redissolved in 0.5 mL of TFA. The *S*-acetamidomethyl cysteines were deprotected by silver trifluoromethanesulfonate and oxidized with DMSO in aqueous HCl as described by Hunter and Komives (*25*). To the peptide solution in 0.5 mL of TFA, anisole (10 μL) and silver trifluoromethanesulfonate (27 mg) were added and the mixture was incubated at 4 °C for 1.5 h. Ice-cold dry ether (1 mL) was added to the reaction mixture and the precipitate was collected by centrifugation. The precipitate was washed twice with the same volume of ice-cold ether and oxidized overnight in 2 mL of solution containing 0.5 M HCl and 50% dimethylsulfoxide (DMSO). The solution was then filtered, diluted 10 times with HPLC-grade water and purified by reversed-phase HPLC. HGA 1−31 (D1V, K3H, S9I), HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S), and HGA 1−31 (D1V, K3H, S9I, Q20P) were oxidized overnight by air in 1% ammonium-acetate buffer, pH 8.5, at a peptide concentration of approximately 6 μg/mL. The *S*-acetamidomethylcysteines in oxidized HGA 1−31 (D1V, K3H, S9I) and HGA 1−31 (D1V, K3H, S9I, Q20P) were left unmodified. When necessary, the peptides were repurified by HPLC using 0.1% triethylaminoacetate buffer, pH 5.2, and an acetonitrile gradient of 10−45%. The identities of the peptides were verified by the use of a SCIEX API *III* ion-spray mass spectrometer.

Disulfide Bond Mapping and Proteolytic Digestion. The mapping of the disulfide bonds was achieved by using a partial reduction method described by Gray (*26*, *27*). The

oxidized peptides were reduced for 3 min at 65 °C in an aqueous solution containing 20 mM tris-(2-carboxyethyl)-phosphine (TCEP) and 0.17 M citric acid, pH 3.0. The reaction mixture was immediately injected onto HPLC. The partially reduced species were isolated on an analytical Vydac C_{18} reversed-phase HPLC column and alkylated with a supersaturated 2.2 M solution of iodoacetamide as described by Gray (*26*). The alkylated peptides were purified on HPLC, characterized with ion-spray mass spectroscopy, and by amino-terminal sequencing.

For proteolytic digestion, an amount of 10 μg of the peptides was dissolved in 50 μL of a buffer containing the proteases and incubated at 37 °C for different periods of time. For the Glu-C protease, the digestion buffer was 50 mM sodium phosphate, pH 7.8, for the Arg-C protease, the buffer was 100 mM ammonium bicarbonate, pH 8.0, and for proteinase K, the buffer was 100 mM Tris-Cl and 1 mM $CaCl_2$, pH 7.8. For digestion with pepsin, 10 μg of the peptides were dissolved in 200 μL of 10 mM HCl, and 1 mg of pepsin linked to agarose beads was added to the peptide solution.

NMR Experiments. The peptides were dissolved in a 20 mM sodium acetate-d_3 buffer, 10% D_2O, with the pH adjusted to 5.0 with no correction for the isotope effect. The concentrations of the peptide samples were between 0.07 and 0.6 mM. For the deuterium exchange experiments, the peptide samples were lyophilized and redissolved in D_2O to the same volume as the H_2O solution. The peptide samples in D_2O were transferred to the original NMR tubes and a series of one-dimensional proton spectra were recorded immediately after the sample preparation. The first spectrum was usually recorded 12−17 min after D_2O addition. The subsequent spectra were recorded every 5−20 min.

NMR experiments were carried out at 500 MHz on Bruker Avance-500 or DRX-500 NMR spectrometers. TOCSY (τ_{SL} = 35−65 ms) spectra (*28*) and water-flipback NOESY (τ_m = 150−250 ms) spectra (*29*) were obtained at 288 K and 298 K. The water signal was suppressed using WATERGATE (*30*) for the samples in H_2O and using low-power presaturation for those in D_2O. The spin-lock used for TOCSY experiments was an MLEV-17 sequence (*31*) with a field strength of 8.3 kHz preceded and followed by a 2.0 ms trim pulse. The collected FID data were processed using the Bruker XWIN-NMR program. Spectral visualization and analysis were carried out manually on plotted spectra and/or on Silicon Graphic workstations with the program PRONTO (*32*). Residue-specific assignment of the proton resonances was achieved by spin system identification using TOCSY, followed by sequential assignment through NOE connectivities. All chemical shift values are referenced to that of an internal DSS standard which was set to 0 ppm.

The $^3J_{HNH\alpha}$ coupling constants were calculated from the separation of the cross-peak extrema measured with two forms of DQF-COSY spectra, one phased in pure absorption and the other in pure dispersion (*33*). All values between 6.0 and 8.0 Hz were discarded. An error of ±1 Hz was assumed in converting the remaining coupling constants to dihedral ϕ angle constraints.

Structure Calculation. The experimental NOE distances for HGA 1−31 (D1V, K3H, S9I, Q20P) were divided into three categories with upper bounds of 2.5, 3.0, and 5.0 Å, based on the intensity of the corresponding cross-peaks. Individual distances were categorized by an approximate calibration against the intensities of the NOE peaks between geminal protons. The number of unambiguous distance constraints totaled 174, including those involving prochiral centers treated using the summation procedure (*34*, *35*). In addition, two disulfide bonds and eight ϕ angle constraints were used in the structure calculations. A total of 100 initial structures were calculated using the XPLOR program (*36*) using a standard protocol of distance geometry followed by simulated annealing (DG/SA). The NOE penalty function is a square-well potential with a weight of 50 kcal mol^{-1} Å^{-2} throughout the calculations. Ten lowest energy structures were used to define five hydrogen bonds that were included in the final list of constraints. A final set of 200 structures was generated by DG/SA and refined with XPLOR using a simulated annealing protocol. The stereochemical quality of the final set of computed structures was verified using PROCHECK (*37*). The Sybyl and InsightII molecular graphics programs were used for structure manipulation and viewing.

RESULTS

Folding Properties of the N-Terminal Subdomain of Human Granulin A and Its Analogues. Upon oxidation, a peptide with four cysteines can theoretically form three different disulfide-bonded isomers. To eliminate two of the possibilities, we prepared the reduced form of HGA 1−30 (C17S, C27S) with Cys10 and Cys26 blocked by S-acetamidomethyl groups. The selection of these blocked cysteines was intended to assess the folding properties of HGA 1−30 (C17S, C27S) within the [1−3, 2−4] disulfide framework of the stack of two β-hairpins. After two-step oxidation of the peptide (see Materials and Methods), the [1−3, 2−4] disulfide pairing of HGA 1−30 (C17S, C27S) was reconfirmed by a partial reduction of the peptide with the subsequent alkylation of the produced semireduced species (*26*, *27*). After 3 min of incubation of HGA 1−30 (C17S, C27S) with an excess amount of TCEP, as described in the Materials and Methods, 42% of the total peptide was completely reduced. In addition to the fully oxidized form of HGA 1−30 (C17S, C27S), two more species are observed (Figure 1): the fully reduced peptide, with the HPLC retention time of 21.8 min, and the semireduced species, with the HPLC retention time of 17.2 min. At reduction times below 2 min, it is possible to observe another semireduced isomer, with the HPLC retention time of 18 min (data not shown), but the yield of the semireduced species is markedly lower. The semireduced species was collected, alkylated with iodoacetamide, and purified. The mass of the alkylated peptide was 3357, corresponding to the expected mass of HGA 1−30 (C17S, C27S) with two labeled cysteines. The amino-terminal analysis of the first 18 residues demonstrated that the first and the third cysteines were labeled, while the second cysteine was not labeled (data not shown). These data prove conclusively that upon regio-selective oxidation the HGA 1−30 (C17S, C27S) disulfide pairing pattern is, as expected, [1−3, 2−4]. The regio-selective oxidations of HGA 1−30 (C17S, Q20P, C27S) and HGA 1−31 (D1V, K3H, S9I, C17S, C27S) were performed in a similar fashion (data not shown).

To establish if the other three peptide analogues, HGA 1−31 (D1V, K3H, S9I), HGA 1−31 (D1V, K3H, S9I,

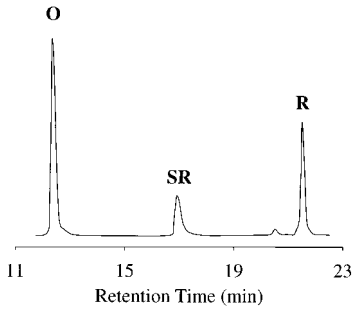

FIGURE 1: HPLC separation of the partially reduced products of HGA 1−30 (C17S, C27S) by TCEP. The peptides were reduced as described in Materials and Methods and immediately applied to a 25 × 0.46 cm Vydac C18 reversed-phase HPLC column equilibrated with 90% solution A (0.1% TFA) and 10% solution B (acetonitrile, 0.1% TFA). The separation was performed at a flow rate of 1.5 mL/min using an extended gradient of 10−50% solution B over 40 min. The start of the gradient was marked as zero time. The detector wavelength was set at 278 nm. Peaks O and R correspond to the fully oxidized and fully reduced species of HGA 1−30 (C17S, C27S), respectively. The SR peak corresponds to the semireduced species.

Q20P), and HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S), can spontaneously form the similar [1−3, 2−4] disulfide pairing pattern upon air oxidation, the peptides were oxidized as described in the Materials and Methods, and the products were analyzed using reversed-phase HPLC. The two peptides HGA 1−31 (D1V, K3H, S9I) and HGA 1−31 (D1V, K3H, S9I, Q20P) contain Acm-blocked Cys17 and Cys27, which may better mimic the disulfide bonds that connect the second and the third β-hairpins in the consensus three-dimensional structure for granulin repeats (*11*). Air oxidation of HGA 1−31 (D1V, K3H, S9I) produced two intense and well-separated peaks on HPLC with a retention time of 21.5 and 22.5 min. Both peptide fractions had a molecular mass of 3562, corresponding to the fully oxidized peptide. The oxidation of HGA 1−31 (D1V, K3H, S9I, Q20P) produced a unique species with a molecular mass of 3530 and a retention time of 15 min. Subsequent repurification of HGA 1−31 (D1V, K3H, S9I, Q20P) using the triethylaminoacetate-acetonitrile solvent system confirmed the homogeneity of the peak. Therefore, HGA 1−31 (D1V, K3H, S9I, Q20P) upon air oxidation folds into a single molecular species with a unique disulfide pairing. The peptide HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S), with two serine residues replacing cysteines 17 and 27, also spontaneously formed a uniquely folded species upon free-air oxidation (data not shown).

We then performed a comparative analysis of the susceptibility of two representative peptides, i.e., the oxidized HGA 1−30 (C17S, C27S) and HGA 1−31 (D1V, K3H, S9I, Q20P), to partial reduction by TCEP and proteolysis with proteases. After the peptides were treated with TCEP and different proteases, the products were immediately injected onto and analyzed with reversed-phase HPLC. The TCEP reduction of these two peptides showed that HGA 1−30 (C17S, C27S) is reduced faster than HGA 1−31 (D1V, K3H, S9I, Q20P). As mentioned above, after 3 min of HGA 1−30 (C17S, C27S) incubation with excess TCEP, 42% of the total peptide was completely reduced. HGA 1−31 (D1V, K3H, S9I, Q20P) was more difficult to reduce under the same conditions, after 5 min of reduction, only 6% of total peptide

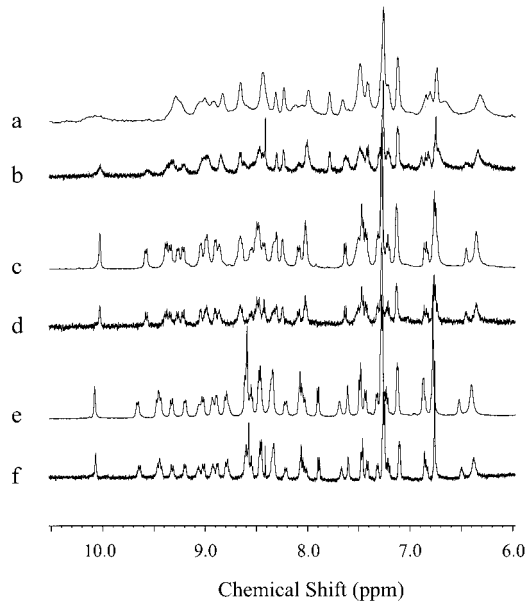

FIGURE 2: The amide proton resonances of (a) 0.4 mM HGA 1−30 (C17S, C27S); (b) 0.068 mM HGA 1−30 (C17S, C27S); (c) 0.6 mM HGA 1−30 (C17S, Q20P, C27S); (d) 0.07 mM HGA 1−30 (C17S, Q20P, C27S); (e) 0.6 mM HGA 1−31 (D1V, K3H, S9I, Q20P); (f) 0.068 mM HGA 1−31 (D1V, K3H, S9I, Q20P). The one-dimensional NMR spectra of these peptides were recorded in 20 mM sodium-acetate-d_3 buffer, 10% D_2O, pH 5.0, at 288 K.

was reduced, and only 28% was reduced after 10 min of reduction.

Both peptides were not cleaved after 6 h of incubation with Glu-C and Arg-C at 37 °C, while they were equally well degraded by proteinase K for 2 h at 37 °C. However, the peptides showed different susceptibility to pepsin digestion. HGA 1−30 (C17S, C27S) was completely digested with pepsin after 1.5 h of incubation at 37 °C, while HGA 1−31 (D1V, K3H, S9I, and Q20P) did not show an observable change in HPLC profile after 4 h of incubation at 37 °C.

NMR Characterization of the Uniquely Disulfide-Bonded Peptides. The oxidized peptide HGA 1−30 (C17S, C27S), derived from human granulin A, shares a 50% sequence identity with its carp analogue CG1 1−30 (C17S, C27S). The human granulin A peptide, however, has a tendency to aggregate in solution at concentrations higher than 100 μM, producing broadened NMR spectra (Figure 2, spectra a and b). Therefore, we studied the behavior of this peptide at concentrations below 70 μM. One-dimensional NMR spectrum and two-dimensional TOCSY NMR spectrum of HGA 1−30 (C17S, C27S) showed that at the concentration of 67 μM the amide proton and side-chain proton resonance dispersion is lower than that of CG1 1−30 (C17S, C27S), with an NH resonance dispersion of 2.26 ppm as compared to 2.69 ppm for the carp analogue at pH 5.0. The NH resonance peaks are comparatively broad, suggesting a conformational exchange (Figure 2, panels a and b).

HGA 1−30 (C17S, Q20P, C27S) and HGA 1−31 (D1V, K3H, S9I, Q20P) appear to be better folded in solution than its prototype, HGA 1−30 (C17S, C27S). The dispersion of the resonance peaks is higher for both HGA 1−30 (C17S, Q20P, C27S) and HGA 1−31 (D1V, K3H, S9I, Q20P), with an NH peak dispersion of 2.27 (Figure 2, spectra c and d) and 2.36 ppm (Figure 2, panels e and f), respectively. In addition, the NH resonance signals were sharper for both

peptides, with no changes observed in the range of the peptide concentrations between 70 and 600 μM (Figure 2, spectra c to f). Interestingly, both HGA 1−31 (D1V, K3H, S9I, C17S, C27S) and HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S) are also better folded than HGA 1−30 (C17S, C27S) on the basis of higher resonance dispersion and sharper NH proton resonance peaks (spectra not shown).

The conformational properties of the disulfide-bonded peptides were further characterized by following the course of exchange of the amide protons in the peptides with D_2O. No evidence was found for slowly exchanging amide protons in HGA 1−30 (C17S, C27S), HGA 1−30 (C17S, Q20P, C27S), and HGA 1−31 (D1V, K3H, S9I, C17S, C27S). However, both HGA 1−31 (D1V, K3H, S9I, Q20P) and HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S) contained slowly exchanging amide protons, some of which were still present after hours of incubation with D_2O. For example, the residues with slowly exchanging amide protons were identified as Tyr14, Thr15, Cys17, Leu19, Gly25, and Cys27 in HGA 1−31 (D1V, K3H, S9I, Q20P) by recording a TOCSY spectrum in D_2O. Slowly exchanging amide protons were also observed within the N-terminal subdomain of the full-length carp granulin-1 (*11*) and for a peptide corresponding to its amino-terminal 1−30 fragment, CG1 1−30 (C17S, C27S) (unpublished observations). The corresponding residues in the full-length carp granulin-1 have been shown to participate in the formation of interstrand backbone−backbone hydrogen bonds within the second β-hairpin (*11*). The absence of slowly exchanging amide protons in HGA 1−30 (C17S, C27S), HGA 1−30 (C17S, Q20P, C27S), and HGA 1−31 (D1V, K3H, S9I, C17S, C27S) provides conclusive evidence for a reduced conformational stability of these three peptides as compared to both HGA 1−31 (D1V, K3H, S9I, Q20P) and HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S). These results, along with the spontaneous folding of and very similar proton NMR spectra for HGA 1−31 (D1V, K3H, S9I, Q20P) and HGA 1−31 (D1V, K3H, S9I, C17S, Q20P, C27S), suggest that the conformational stability of the human granulin A peptide conferred by the Asp1Val, Lys3His, Ser9Ile, and Gln20Pro substitutions is independent of whether Cys17 and Cys27 are replaced by serines or *S*-acetamidomethylcysteines. We therefore focused further structural characterization on the HGA 1−31 (D1V, K3H, S9I, Q20P) variant with Acm-blocked cysteines which can be used for the formation of disulfide bonds in future studies.

Nuclear Overhauser Effects for the Well-Folded HGA 1−31 (*D1V, K3H, S9I, Q20P*). A summary of the NOE connectivities observed for the HGA 1−31 (D1V, K3H, S9I, Q20P) is shown in Figure 4. Similar to the full-length carp granulin-1, HGA 1−31 (D1V, K3H, S9I, Q20P) exhibits a pattern of interstrand NOEs, characteristic of a regular antiparallel β-sheet from residues Thr14 to Pro28 (Figure 4, panels b and c). Residues 1−11 of HGA 1−31 (D1V, K3H, S9I, Q20P) form a β-sheetlike subdomain with some interstrand NOE connectivities consistent with the formation of another β-hairpin. There were, however, no slowly exchanging amide protons in this region, suggesting the absence of a regular β-hairpin in solution. Also, the H^{α} chemical shift deviations from the random coil values are lower in this part compared to the second, well-defined β-hairpin. Similarly, the first β-hairpin appears to be

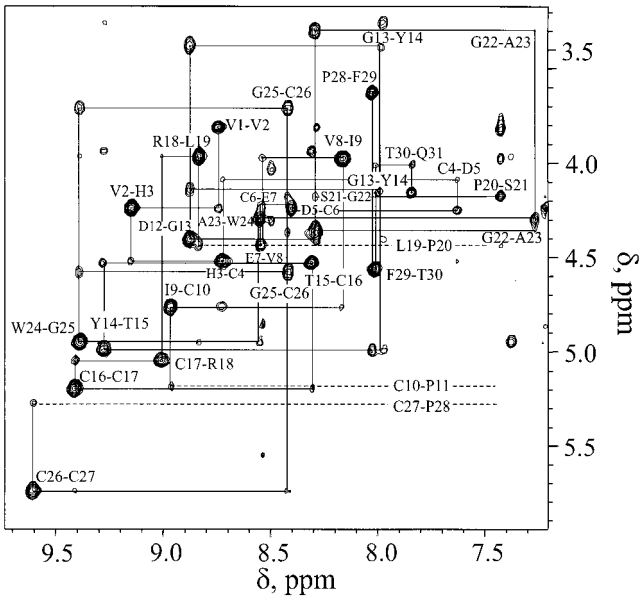

FIGURE 3: Sequential NOE walk for the HGA 1−31 (D1V, K3H, S9I, Q20P). The experimental conditions were as in Figure 2e. The NOESY spectrum was acquired with a mixing time of 250 ms at 288 K.

structurally less defined than the second β-hairpin in both carp granulin-1 and its amino-terminal fragment (*11*, *12*).

The structural similarities between HGA 1−31 (D1V, K3H, S9I, Q20P), the full-length carp granulin-1, and its N-terminal fragment were further confirmed by the observation of long-range NOEs typical of the hydrophobic interactions in the stack of two β-hairpins. The presence of the hydrophobic cap in HGA 1−31 (D1V, K3H, S9I, Q20P) was suggested by the NOE connectivities between the side-chain protons of Ile9 and the $H^{\alpha,\beta}$ protons of His3 and between the side-chain protons of Ile9 and the H^{β} protons of Val1. The $H^{\beta,\gamma}$ protons of Val2 make close contacts with the H^{β} protons of Cys10. Numerous NOE connectivities are observed between the side chain of Trp24 and the side chains of Arg18, Cys4, and Cys16, suggesting the presence of the Trp24−Arg18−(Cys4Cys16)−(Cys10Cys26)−Val2 hydrophobic cluster resembling that in carp granulin-1. In addition, the NOE connectivities between the heterocycle protons of Trp24 and the H^{γ} protons of Val8 are consistent with the relative orientation of the N-terminal and C-terminal β-hairpins observed in carp granulin-1.

On the basis of the unambiguous NOEs between the $H^{\beta}{}_2$ protons of Cys26 and the downfield H^{β} proton of Cys10, Vranken et al. (*12*) established that the disulfide pairing in the amino-terminal fragment of carp granulin-1 is [1−3, 2−4], a pattern identical to that of the native full-length carp granulin-1. Similarly, we determined the disulfide pairing in HGA 1−31 (D1V, K3H, S9I, Q20P) by searching for H^{β} to H^{β} and H^{α} to H^{β} NOEs between different cysteines. The H^{α} of Cys4 and H^{α} of Cys26 were unambiguously connected to the $H^{\beta}{}_2$ of Cys16 and to the upfield H^{β} of Cys10, respectively. No other H^{α} to H^{β} contacts between cysteines were found. Therefore, we concluded that the disulfide pairing in HGA 1−31 (D1V, K3H, S9I, Q20P) is [1−3, 2−4], the same as in the native carp granulin. The NOEs consistent with the H^{β} to H^{β} connectivities between Cys4 and Cys16 were also observed, but they could not be unambiguously assigned because of the spectral overlaps.

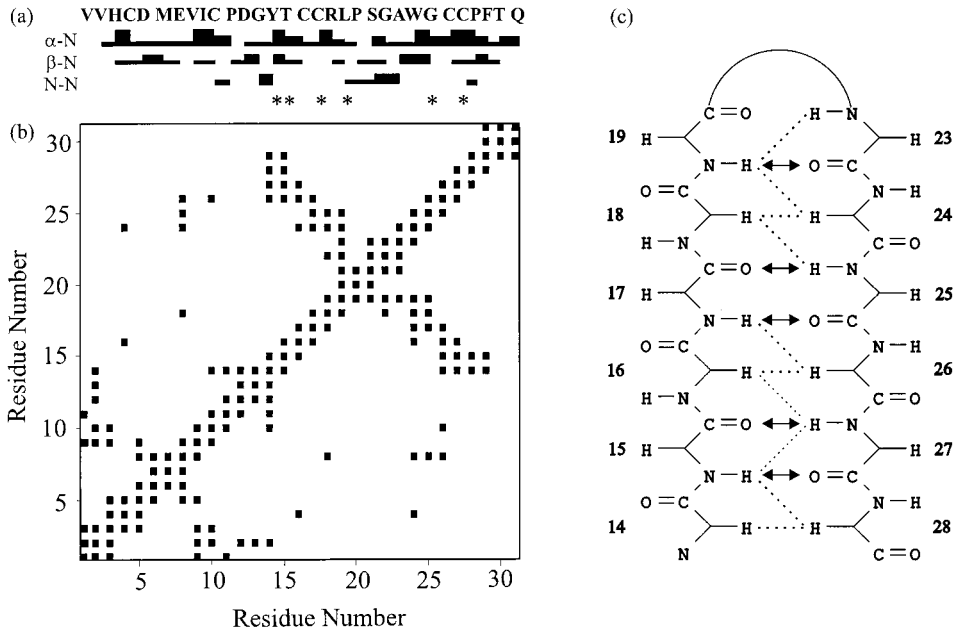

FIGURE 4: A summary of the NOE connectivities and other NMR data for HGA 1−31 (D1V, K3H, S9I, Q20P). (a) Sequential NOE contacts (indicated by α-N, *β*-N, and N−N) between the H^{α}, H^{β}, and NH protons of the current residue and the following residue. The thickness of the lines reflects the relative intensity of the NOE cross-peaks (strong, medium, and weak). Residues with slowly exchanging protons are marked with asterisks. (b) Diagonal plot of the NOE contacts observed for HGA 1−31 (D1V, K3H, S9I, Q20P). A filled square in position (*i*,*j*) indicates that an NOE was observed between residues *i* and *j*. (c) Identification of an antiparallel *β*-sheet involving residues 14−28 of HGA 1−31 (D1V, K3H, S9I, Q20P). Dotted lines indicate the observed interstrand NOEs. Arrows show the assigned hydrogen bonds with slowly exchanging amide protons (indicated by asterisks in Figure 4a).

Three-Dimensional Structure of HGA 1−31 (D1V, K3H, S9I, Q20P). Three-dimensional structure of HGA 1−31 (D1V, K3H, S9I, Q20P) was calculated with eight ϕ angle constraints, 72 sequential NOEs, 102 medium plus long-range NOEs, five hydrogen bonds, and two disulfide bond constraints. The eight ϕ angle constraints for residues Val2, Glu7, Cys10, Thr15, Cys17, Cys26, Cys27, and Thr30 were determined from the $^3J_{HNH^{\alpha}}$ coupling constants as described in the Materials and Methods. Six slowly exchanging amide protons were examined for possible candidates of hydrogen bond formation on the basis of the first rounds of structure calculation. Five hydrogen bonds were unambiguously assigned and incorporated in the final list of distance constraints. For every assigned hydrogen bond, two constraints of 1.8−2.0 and 2.7−3.0 Å were imposed for the amide proton-carbonyl oxygen and for the amide nitrogen−carbonyl oxygen distances, respectively. Of the final 200 calculated structures, 10 of the lowest NOE energy were chosen for further analysis (see Materials and Methods for more details). The calculation statistics demonstrates the high quality of these 10 structures (Table 1).

The superposition of the 10 best structures is shown in Figure 5. There is a well-defined second (central) *β*-hairpin from Thr14 to Pro28. The two strands of the central *β*-sheet in HGA 1−31 (D1V, K3H, S9I, Q20P) are linked by a reverse turn, with residues Leu19, Pro20, Ser21, Gly22, and Ala23 adopting a conformation of the type I + G1 *β*-bulge turn. These residues are located in the $\beta\alpha_R\gamma_R\gamma_L\beta$ region of the Ramachandran plot defining the backbone conformations (*38*). This turn is stabilized by a hydrogen bond connecting the NH proton of Leu19 and the carbonyl oxygen of Ala23. The amide proton of Leu19 is shielded from solvent exchange (see above); there were, however, no other slowly exchanging protons within the turn. The whole central

Table 1: Experimental Constraints and Refinement Statistics[a] for HGA 1−31 (D1V, K3H, S9I, Q20P)

distance constraints	intraresidue		none used
	sequential		72
	medium and long range		102
	hydrogen bonds		5
	disulfide bonds		2
angle constraints	dihedral angles		8
Cartesian coordinate pairwise rmsd[b] (Å)	residues	2−28	1.08 ± 0.24
		1−11	0.77 ± 0.23
		14−28	0.58 ± 0.19
		19−23	0.20 ± 0.09
deviations from covalent geometry	bonds		0.002 ± 0.0001Å
	valence angles		0.53 ± 0.01°
XPLOR energies (kcal mol^{-1})	total		44.03 ± 1.99
	NOE		1.37 ± 0.26
PROCHECK statistics	residues in the		
	most favored regions		66.1%
	additional allowed regions		28.7%
	generously allowed regions		2.2%
	disallowed regions		3.0%

[a] Values were calculated over the entire sequence and expressed as the averages of the 10 lowest energy structures. [b] Pairwise rmsd was calculated by superposition of the 10 lowest energy over the sequence specified.

β-sheet can be classified as a 3:5 type *β*-hairpin, which is widespread in protein structures (*38*). The hairpin also exhibits the right-hand twist characteristic of this class of hairpins (*39*). The corresponding 3:5 type hairpin was also found in the full-length carp granulin and its independently folded amino-terminal fragment (*11*, *12*). The average pairwise rmsd for the backbone atoms of residues 14−28 is 0.58 ± 0.19 Å, and of residues 19−23, defining the turn, is 0.20 ± 0.09 Å. The first hairpin (residues 1−11) is not as

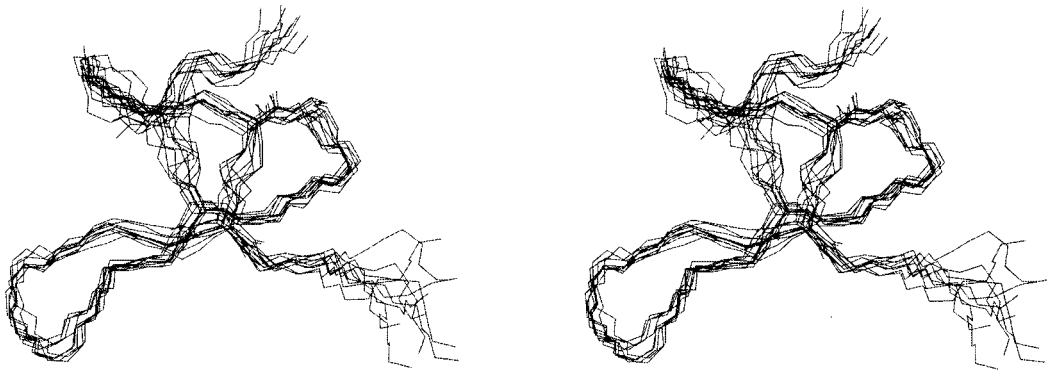

FIGURE 5: A cluster of the 10 best structures of HGA 1−31 (D1V, K3H, S9I, Q20P) shown in stereoview. Only backbone atoms and two disulfide bonds are displayed. Superposition of all the structures was carried out using the C^{α} atoms of residues 2−28.

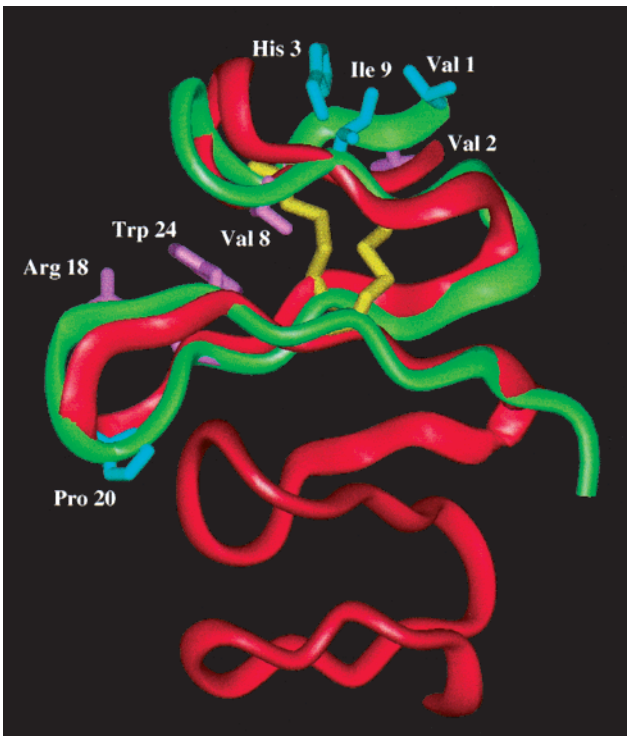

FIGURE 6: Overlay of the backbone ribbons of the full carp granulin-1 protein (in red) and the amino terminal fragment HGA 1−31 (D1V, K3H, S9I, Q20P) (in green). Also shown are the side-chain heavy atoms of the residues forming the first (Val1, His3, and Ile9) and the second (Val2, Val8, Arg18, Trp24, Cys4−Cys16, and Cys10−Cys26) cluster of hydrophobic residues. Cystines are shown in yellow, and substituted residues are shown in blue.

well-defined as the second one, and the average pairwise rmsd for this region is 0.77 ± 0.23 Å.

The overall fold of the peptide is very similar to that of full carp granulin-1 and represents another example of a well-structured stack of two β-hairpins linked together by one reverse β-turn and two disulfide bonds. There is a network of hydrophobic interactions composed of two hydrophobic clusters in HGA 1−31 (D1V, K3H, S9I, Q20P) that is also present in the native carp granulin-1 (Figure 6). One cluster is located between the two β-hairpins and consists of the side chains of Val2, Trp24, and the two disulfide bridges. The second cluster is located on top of the first β-hairpin and is formed by the side chains of Val1, His3, and Ile9, or the so-called hydrophobic cap (*11*, *12*). The aromatic side chain of Trp24 also makes characteristic contacts with Asp5 and Val8, linking the two β-hairpins. The same interactions were observed in the amino-terminal fragment of carp granulin-1 (*12*). The side chain of Pro20 and the bulky Acm groups do not have obvious NOE contacts and do not seem to participate in hydrophobic interactions. Also, there do not seem to be any side chain−side chain hydrophobic interactions on the lower side of the second β-hairpin.

DISCUSSION

To study the role of hydrophobic interactions for the conformational stability of a stack of two β-hairpins, we introduced a hydrophobic cap found in an independently folded amino-terminal fragment of carp granulin-1 into a related but not well-folded peptide, derived from human granulin A. In addition, residue Gln20 was replaced with proline present in a strategic location within the central β-hairpin of carp granulin-1. Proline residues have a defined ϕ angle because of the pyrrolidine ring and are, therefore, expected to reduce unfavorable entropy contributions in structure formation (*40*). It also may, due to restriction of the ϕ angle to about −60°, direct the folding of the peptide by favoring the tertiary interactions stabilizing the hairpin and the overall fold. As a result of proline substitution, HGA 1−31 (D1V, K3H, S9I, Q20P) spontaneously forms a unique disulfide species with the expected "native" [1−3, 2−4] pairing. Narrower proton resonances, larger resonance dispersion and slow deuterium exchange in the central β-hairpin region together with a decreased susceptibility to chemical reduction and proteolytic degradation, provided conclusive evidence that the resulting hybrid peptide is markedly more structured in solution than its native human sequence with the same disulfide pairing. Introduction of either proline or the hydrophobic cap into the human granulin A peptide alone improved its folding characteristics, but was not enough to create a network of slowly exchanging hydrogens in the second β-hairpin. Only incorporation of the hydrophobic cap together with Pro20 recreated a peptide comparable in its conformational stability and similar in its three-dimensional structure to the corresponding carp granulin-1 peptide. These findings highlight the importance of both hydrophobic and turn-stabilizing interactions for the structural integrity of the stack of two β-hairpins.

Apparently, spontaneous oxidative folding of the peptide into the correct crisscross disulfide pattern is related to the conformational stability of the stack of two β-hairpins. It is possible that during the oxidation tertiary interactions defining the fold are strong enough to bring the cysteine side chains together in a fashion that favors formation of the

correct disulfide pairing. The ability of the small cysteine-rich peptides to spontaneously fold into the correct disulfide pairing is particularly valuable if the structure scaffold is to be used with molecular evolution techniques, such as phage display, to identify binding proteins and to optimize binding affinity. The cysteine-rich scaffolds are usually avoided in functional optimization through phage display due to the uncertainties in cystine formation (*2*). However, the disulfide-rich knottin architecture was shown recently to be suitable for the creation of small proteins with novel binding activities (*9*). Although the authors demonstrated that reduction of the selected ligands abolished their binding activities, indicating a dependence on the formation of disulfide bonds, there was no information concerning the actual disulfide pairing in the created molecules. Examples of the EGF-like repeats are known which differ severalfolds in binding affinities toward their targets depending on their disulfide pairing pattern (*41*−*43*). Hence, if the search for novel binding activities is desired to be performed within well-folded representatives of a certain scaffold, phage display optimization is best to be combined with rational design, ensuring proper topology of the ligand molecules (*4*, *44*).

The stack of two β-hairpins satisfies the criteria and may be a subclass of the so-called disulfide β-cross (*45*, *46*), a highly abundant scaffold shared by members of the growth factor family, small disulfide-rich toxins, and growth factor-like protein repeats. Although the disulfide β-cross was proposed to be a folding nucleus for small disulfide-rich proteins (*45*), no stable and independently folded disulfide β-crosses were reported until recently, when the amino-terminal 30 residue fragment of carp granulin-1 was shown to have an identical fold to that found in the full-length protein (*12*). In contrast, another possible candidate for the formation of a β-cross, the amino-terminal fragment of the sixth epidermal growth factor-like domain from thrombomodulin, was demonstrated to have no stable fold free in solution (*43*). The peculiarity of the carp granulin sequence is that it provides enough information for the correct folding of its amino-terminal subdomain in the absence of its interaction with the carboxy-terminal fragment or with other binding proteins. Two hydrophobic clusters, one between two β-hairpins and another on top of the first hairpin (the hydrophobic cap), seem to be responsible for the exceptional conformational stability of the amino-terminal carp granulin-1 fragment (*11*, *12*). The conformational stability of the carp granulin N-terminal subdomain may also contribute to the resistance of the full-length protein to thermal denaturation: indeed, heating the native protein up to 358 K followed by cooling to 298 K did not result in any changes to the proton resonances (unpublished observations).

An elegant example demonstrating that the formation of an appropriate scaffold may be a "sufficient" requirement for binding activities, was recently provided by Blanco-Aparicio et al. (*47*). It was shown that the potato carboxypeptidase inhibitor (PCI) is able to suppress the growth of several human pancreatic adenocarcinoma cell lines, both in vitro and in nude mice. The PCI contains a disulfide-rich T-knot structural motif that is shared by several mammalian growth factors, including the epidermal growth factor and the transforming growth factor α (*45*, *46*). The authors suggested that structural similarities of PCI and the epidermal growth factor might be the reason for the antitumor activity of PCI. This latest finding unquestionably emphasizes the importance of creating of a family of well-folded scaffolds present in growth-modulating proteins. The stack of two β-hairpins studied here represents, therefore, an important alternative and expands our knowledge in the building of novel and conformationally stable protein mimetics as selective ligands for biologically important targets.

ACKNOWLEDGMENT

We are grateful to Dr. Ping Xu for help with acquiring the NMR data.

SUPPORTING INFORMATION AVAILABLE

^{1}H Chemical shifts (from DSS) of HGA 1−31 (D1V, K3H, S9I, Q20P) in 20 mM sodium acetate-d_3, pH 5.0, at 288 K. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Cunningham, B. C., and Wells, J. A. (1997) *Curr. Opin. Struct. Biol. 7*, 457−462.
2. Nygren, P.-A., and Uhlen, M. (1997) *Curr. Opin. Struct. Biol. 7*, 463−469.
3. Zuckerman, R. N. (1993) *Curr. Opin. Struct. Biol. 3*, 580−584.
4. Starovasnik, M. A., Braisted, A. C., and Wells, J. A. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 10080−10085.
5. Drakopoulou, E., Zinn-Justin, S., Guenneugues, M., Gilquin, B., Menez, A., and Vita, C. (1996) *J. Biol. Chem. 271*, 11979−11987.
6. Vita, C., Roumestand, C., Toma, F., and Menez, A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6404−6408.
7. Zinn-Justin, S., Guenneugues, M., Drakopoulou, E., Gilquin, B., Vita, C., and Menez, A. (1996) *Biochemistry 35*, 8535−8543.
8. Chiche, L., Heitz, A., Padilla, A., Le-Nguyen, D., and Castro, B. (1993) *Protein Eng. 6*, 675−682.
9. Smith, G. P., Patel, S. U., Windass, J. D., Thornton, J. M., Winter, G., and Griffiths, A. D. (1998) *J. Mol. Biol. 277*, 317−332.
10. Heitz, A., Le-Nguen, D., and Chiche, L. (1999) *Biochemistry 38*, 10615−10625.
11. Hrabal, R., Chen, Z., James, S., Bennett, H. P. J., and Ni, F. (1996) *Nat. Struct. Biol. 3*, 747−752.
12. Vranken, W., Chen, Z., Xu, P., James, S., Bennett, H. P. J., and Ni, F. (1999) *J. Pept. Res. 53*, 590−597.
13. Bateman, A., Belcourt, D., Bennett, H., Lazure, C., and Solomon, S. (1990) *Biochem. Biophys. Res. Commun. 173*, 1161−1168.
14. Bateman, A., and Bennett, H. P. (1998) *J. Endocrinol. 158*, 145−151.
15. Bhandari, V., Palfree, R. G. E., and Bateman, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1715−1719.
16. Shoyab, M., McDonald, V. L., Byles, C., Todaro, G. J., and Plowman, G. D. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7912−7916.
17. Plowman, G. D., Green, J. M., Neubauer, M. G., Buckley, S. D., McDonald, V. L., Todaro, G. J., and Shoyab, M. (1992) *J. Biol. Chem. 267*, 13073−13078.
18. Parnell, P. G., Wunderlich, J., Carter, B., and Halper, J. (1992) *Growth Factors 7*, 65−72.
19. Brown, C. A., and Halper, J. (1990) *Exp. Cell Res. 190*, 233−242.
20. Zhou, J., Gao, G., Crabb, J. W., and Serrero, G. (1993) *J. Biol. Chem. 268*, 10863−10869.
21. Xia, X., and Serrero, G. (1998) *Biochem. Biophys. Res. Commun. 245*, 539−543.

22. Xu, S.-Q., Tang, D., Chamberlain, S., Pronk, G., Masiarz, F. R., Kaur, S., Prisco, M., Zanocco-Marani, T., and Baserga, R. (1998) *J. Biol. Chem. 273*, 20078−20083.
23. Zhang, H., and Serrero, G. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 14202−14207.
24. Trinh, D. P., Brown, K. M., and Jeang, K. T. (1999) *Biochem. Biophys. Res. Commun. 256*, 299−306.
25. Hunter, M. J., and Komives, E. A. (1995) *Anal. Biochem. 228*, 173−177.
26. Gray, W. (1993) *Protein Sci. 2*, 1732−1748.
27. Gray, W. (1993) *Protein Sci. 2*, 1749−1755.
28. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
29. Lippens, G., Dhalluin, C., and Wieruszeski, J.-M. (1995) *J. Biomol. NMR 5*, 327−331.
30. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
31. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 335−360.
32. Kjaer, M., Andersen, K. V., and Poulsen, F. M. (1994) *Methods Enzymol. 29*, 288−318.
33. Kim, Y., and Prestegard, J. H. (1989) *J. Magn. Reson. 84*, 9−13.
34. Ni, F., Meinwald, Y. C., Vasquez, M., and Scheraga, H. A. (1989) *Biochemistry 28*, 3094−3105.
35. Ni, F. (1994) *Progr. NMR Spectrosc. 26*, 517−606.
36. Brunger, A. T. (1992) X-PLOR (Version 3.1) A System for X-ray crystallography and NMR, Yale University Press, New Haven.
37. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
38. Sibanda, B. L., and Thornton, J. M. (1991) *Methods Enzymol. 202*, 59−82.
39. De Alba, E., Rico, M., and Jimenez, M. A. (1997) *Protein Sci. 6*, 2548−2560.
40. Matthews, B. W., Nicholson, H., and Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663−6667.
41. Hunter, M. J., and Komives, E. A. (1995) *Protein Sci. 4*, 2129−2137.
42. Benitez, B. A. S., Hunter, M. J., Meininger, D. P., and Komives, E. A. (1997) *J. Mol. Biol. 273*, 913−926.
43. Tolkatchev, D., and Ni, F. (1998) *Biochemistry 37*, 9091−9100.
44. Braisted, A. C., and Wells, J. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5688−5692.
45. Harrison, P. M., and Sternberg, M. J. E. (1996) *J. Mol. Biol. 264*, 603−623.
46. Lin, L., and Nussinov, R. (1995) *Nat. Struct. Biol. 2*, 835−837.
47. Blanco-Aparicio, C., Molina, M. A., Fernandez-Salas, E., Frazier, M. L., Mas, J. M., Querol, E., Aviles, F. X., and Llorens, R. (1998) *J. Biol. Chem. 273*, 12370−12377.

BI992130U