A Conserved *cis* Peptide Bond Is Necessary for the Activity of Bowman-Birk Inhibitor Protein†

Arnd B. E. Brauer,‡,§ Gonzalo J. Domingo,|| Robert M. Cooke,⊥ Stephen J. Matthews,§ and Robin J. Leatherbarrow*,‡

Departments of Chemistry and Biological Sciences, Imperial College of Science, Technology and Medicine, South Kensington, London SW7 2AY, U.K., Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, Washington 98109, and GlaxoSmithKline Research and Development, New Frontiers Science Park, Harlow Essex CM19 5AW, U.K.

Received April 30, 2002; Revised Manuscript Received June 24, 2002

ABSTRACT: The Bowman-Birk inhibitor (BBI) family of protease inhibitors has an inhibitory region comprising a disulfide-linked nine-residue loop that adopts the characteristic canonical motif found in many serine protease inhibitors. A unique feature of the BBI loop is the presence of a *cis* peptide bond at the edge of the inhibitory loop. BBI-related protein fragments that encapsulate this loop retain the structure and inhibitory activity of the parent protein. The most common BBI loop sequence has a proline-proline element with a *cis*−*trans* geometry at P3′−P4′. We have examined this element by analysis of the inhibitory activity and structure for a series of synthetic fragments where each of these proline residues has been systematically replaced with alanine. The results show that only when a proline is present at P3′ are potent inhibition and a *cis* peptide bond at that position in the solution structure observed, suggesting that this conformation is required for biological activity. Though a P4′ proline is not essential for activity, it effectively stabilizes the *cis* conformation at P3′ by suppressing alternative conformations. This is most evident from the Pro-Ala variant, which comprises a 1:1 mixture of slowly exchanging and structurally different *cis* and *trans* isomers. Monitoring the action of trypsin on this mixture by NMR shows that this protease interacts selectively with the *cis* P3′ structure, providing direct evidence for the link between activity and the nativelike structure of the *cis* isomer. This is, to the best of our knowledge, the first example where *cis* isomer selectivity can be demonstrated for a proteinase.

Inhibition by serine protease inhibitor proteins is often mediated by an exposed reactive site loop that is fixed in a characteristic "canonical" conformation, thought to be similar to that of a productively bound substrate (*1*, *2*). The convex-shaped canonical reactive site loop, which is complementary to the concave shape of the protease active site (*3*), is the common structural element in at least 18 different families of serine protease inhibitors from animals, plants, and microorganisms (*4*). These generally have a compact shape (*2*) but are otherwise structurally diverse (*5*). Canonical reactive site loops are constructed from a wide variety of sequences, in which, depending on the family, some positions are conserved (*3*). These conserved positions, with the exception of the hyperexposed P1 position that forms the primary specificity determinant (*4*), are generally occupied by amino acids with side chains that can act as conformational constraints. This includes disulfide-linked Cys, hydrogen-bonded amino acids such as Ser or Thr, and backbone-constrained Pro.

The Bowman-Birk inhibitor (BBI)[1] family consists of two subsets that have 9 and 11 amino acid residues in the disulfide-bridged reactive site ring, respectively (*6*). The nine-residue antitryptic reactive site loop of Bowman-Birk inhibitors (BBIs) is particularly rich in such conserved constraining residues, which constitute two-thirds of the entire sequence with a pair each of Cys, Ser/Thr, and Pro (*7*). BBIs are small, highly disulfide-linked plant proteins of typically 60−90 residues with a symmetrical structure of two tricyclic domains, each containing an independent reactive site loop (*8*−*13*).

The two reactive site loops are β hairpins with a type VI β turn, which is centered around a *cis* peptide bond for a

† We gratefully acknowledge the financial support of GlaxoSmithKline and the BBSRC.

* To whom correspondence should be addressed. E-mail: R.Leatherbarrow@ic.ac.uk. Fax: +44 1342 841 939. Telephone: +44 207 594 5752.

‡ Department of Chemistry, Imperial College of Science, Technology and Medicine.

§ Department of Biological Sciences, Imperial College of Science, Technology and Medicine.

|| Seattle Biomedical Research Institute.

⊥ GlaxoSmithKline Research and Development.

[1] Abbreviations: BBI, Bowman-Birk inhibitor; DQF-COSY, double-quantum-filtered correlated spectroscopy; Fmoc, 9-fluorenylmethyloxycarbonyl; HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HMP, 4-hydroxymethylphenoxymethyl-copolystyrene-1%-divinylbenzene; HOBt, 1-hydroxybenzotriazole; HPLC, high-pressure liquid chromatography; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; P*n*, *n*th peptide subsite of a substrate or inhibitor on the carboxyl side of the scissile bond; P*n*′, *n*th peptide subsite of a substrate or inhibitor on the amino side of the scissile bond; rmsd, root-mean-square deviation; ROE, rotating frame nuclear Overhauser effect; ROESY, rotating frame nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; TPCK, tosylphenylalanine chloromethyl ketone; TSP, 3-(trimethylsilyl)-1-propanesulfonic acid.

10.1021/bi026050t CCC: $22.00
Published on Web 07/31/2002

strictly conserved Pro at the P3′ subsite (*9*, *10*, *12*−*15*). The P4′ subsite in BBI is also frequently a Pro residue, though this has *trans* peptide bond geometry. The high content of constraining residues appears to be the basis for the retention of much of the protein's biological activity and three-dimensional structure in fragments (protein-derived or synthetic) incorporating the sequence of BBI reactive site loops (*6*, *16*). Isolated BBI-related reactive site loops are the shortest sequences shown to encompass the canonical conformation (*16*−*18*). Short peptide loops related to BBI also occur naturally: the sunflower trypsin inhibitor is a 14-residue cyclic peptide that has the same structural motif (*19*, *20*). We have already demonstrated the suitability of the BBI−proteinomimetic peptide as a scaffold for the optimization of biological activity and redirection of specificity by combinatorial methods (*21*−*24*).

In this study, we use a combination of kinetic and NMR analyses to investigate the role of the BBI P3′−P4′ *cis*-Pro-*trans*-Pro sequence. Sequential substitution of these Pro residues with Ala is used to assess the functional requirement for protease inhibition. These data are then correlated with structural information to probe directly the requirement for a *cis* peptide bond at P3′.

MATERIALS AND METHODS

Peptide Synthesis. Disulfide-cyclized peptides were synthesized as previously described (*24*). C18 reversed phase HPLC using a Gilson system was used to purify and analyze the cyclic peptides on a Waters-Millipore 25 mm × 100 mm Nova-Pak column with a 6 μm particle size and a Nova-Pak 8 mm × 100 mm column with a 4 μm silica particle size, respectively. PP, PA, AP, and AA peptides had analytical retention times of 7.4, 7.0, 7.2, and 6.6 min, respectively, using a 17 min linear gradient from 100% H_2O to a 30% H_2O/70% acetonitrile mixture (all solvents with 0.1% trifluoroacetic acid) at a flow rate of 3 mL/min. The identity of the cyclic peptides was confirmed by fast atom bombardment mass spectrometry, and for PP and PA in addition by complete assignment of the ^{1}H NMR resonances of the dominant conformers.

Enzyme Inhibition. Materials were supplied by Sigma Aldrich. CNBr-activated Sepharose 4B was obtained from Pharmacia Ltd. and was used as recommended by the supplier. TPCK-treated bovine trypsin was purified immediately prior to use by affinity chromatography on *p*-aminobenzamidine Sepharose 4B and subsequently analyzed for its concentration and relative content of β- and α-trypsin by differential active site titration with *p*-nitrophenyl *p*′-guanidino benzoate (*25*). Only preparations with a β-trypsin content of >75% were used for inhibition assays. Trypsin inhibition was monitored by competitive binding studies with the chromogenic substrate benzoyl-L-arginine *p*-nitroanilide (*26*) in 50 mM Tris-HCl buffer containing 50 mM NaCl and 10 mM $CaCl_2$ at pH 8.0 and 298 K. Substrate concentrations were determined by the final absorbance at 410 nm. Peptide concentrations were determined by absorbance at 280 nm (*27*).

The nonlinear regression GraFit software package (*28*) was used to analyze inhibition data. To distinguish non-tight- and tight-binding inhibition, eqs 1 and 2 (*29*) were used, respectively. The rate of enzymatic substrate hydrolysis v is expressed as a function of the uninhibited rate v_o, the apparent equilibrium dissociation constant K_d*, the total enzyme concentration $[E_o]$, and the total inhibitor concentration $[I_o]$.

$$v = v_o\left(1 - \frac{[I_o]}{[I_o] + K_d^*}\right) \quad (1)$$

$$v = v_o\{([E_o] - [K_d^* + [E_o] + x[I_o] - \sqrt{(K_d^* + [E_o] + x[I_o])^2 - 4[E_o]x[I_o]})/(2[E_o])\} \quad (2)$$

$$v = v_o - v_o x[I_o] \quad (3)$$

$$K_d = K_d^* \frac{1}{1 + [S_o]/K_M} \quad (4)$$

A factor x for the inhibitor concentration was introduced into eq 2 to determine the apparent binding stoichiometry, which is visualized in inhibition plots using eq 3 and interpreted as the inhibitory active fraction of peptide. Equation 4 was used to calculate the equilibrium dissociation constant K_d by correction for substrate competition; $[S_o]$ and K_M are the total substrate concentration and Michaelis constant, respectively.

NMR Analysis and Structure Calculation. The NMR analysis was performed in either aqueous solution (90% H_2O/10% ^2H_2O and 100% ^2H_2O) or 100 mM aqueous phosphate buffer (10% ^2H_2O), both at 298 K and pH* 3.8, at peptide concentrations ranging between 0.3 and 5 mM. Chemical shifts were referenced to TSP. DQF-COSY (*30*), TOCSY (mixing time of 80 ms) (*31*), NOESY (*32*), and ROESY (with mixing times ranging from 50 to 300 ms) (*33*) experiments were carried out on a Bruker AMX 600 spectrometer, and processed and analyzed using X-WinNMR and Aurelia software packages. Following sequential assignment (*34*), amide temperature coefficients (*35*, *36*), diastereotopic proton assignment and χ^1 conformations (*37*), $^3J_{HNH\alpha}$ coupling constants (*38*), $^3J_{H\alpha H\beta}$ coupling constants (*39*), and proline ring pucker (*40*) were analyzed as described. All coupling constants were derived from one-dimensional spectra. Interproton distances were derived from cross-peak volumes calibrated against the geminal interproton distance (*41*) or the minimal theoretical $H\alpha_i-HN_{i+1}$ interproton distance (*42*).

Peptide structures were calculated with the program X-PLOR (version 3.851) (*43*) employing the slightly modified version (*17*) of a standard simulated annealing protocol (*42*). No hydrogen bond restraints were used during the calculations. Structures were visualized and analyzed using INSIGHT II (Molecular Simulations Inc.), Swiss PDB Viewer (*44*), WebLab (Molecular Simulations Inc.), and WHAT IF (*45*) programs. All residues fall within the allowed areas of the Ramachandran plot.

RESULTS

Trypsin Inhibition Kinetics. To investigate the individual roles of Pro at P3′ and P4′ systematically, these positions were sequentially substituted with Ala as shown in Table 1. The four variants were synthesized and analyzed for their biological activity. Figure 1 shows a comparison of the inhibition curves produced by each of these BBI reactive site loop peptides immediately after mixing with trypsin. It

Table 1: Disulfide-Cyclized Peptides Synthesized for the Investigation of the Structure−Activity Relationships of the Pro Residues at the P3′ and P4′ Positions of BBI Reactive Site Loops[a]

	residue 1 subsite P4	residue 2 subsite P3	residue 3 subsite P2	residue 4 subsite P1	residue 5 subsite P1′	residue 6 subsite P2′	residue 7 subsite P3′	residue 8 subsite P4′	residue 9 subsite P5′	residue 10 subsite P6′	residue 11 subsite P7′
PP	S	C	T	K	S	I	**P**	**P**	Q	C	Y
PA	S	C	T	K	S	I	**P**	**A**	Q	C	Y
AP	S	C	T	K	S	I	**A**	**P**	Q	C	Y
AA	S	C	T	K	S	I	**A**	**A**	Q	C	Y

[a] The peptides are termed PP, PA, AP, and AA in the text to indicate their P3′ and P4′ sequences, respectively. The subsite positions are labeled according to the standard nomenclature (*67*).

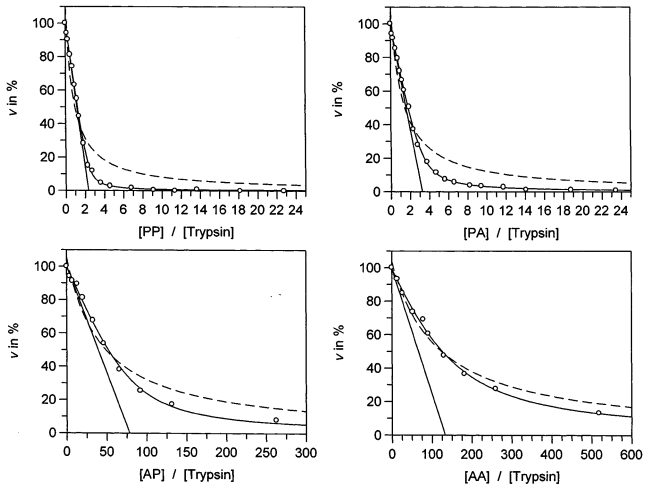

FIGURE 1: Inhibition of trypsin by the P3′/P4′ variant peptides immediately after mixing. The relative rate (v) of hydrolysis of a chromogenic substrate at varying inhibitor-to-enzyme ratios is shown. Trypsin concentrations as determined by active site titration were 1.22 μM for PP and PA and 0.61 μM for AP and AA. Solid and dashed curves represent best fits using tight-binding eq 2 and non-tight-binding eq 1, respectively. The x-intercept of the straight line (eq 3) depicts the apparent binding stoichiometry.

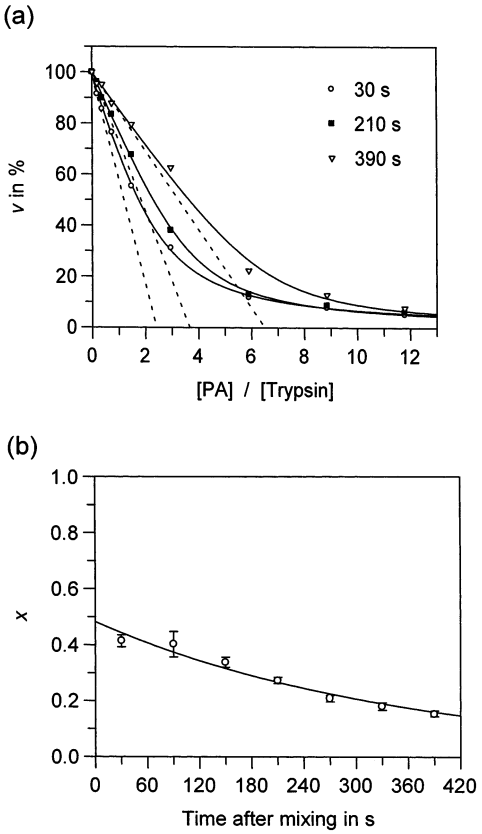

FIGURE 2: (a) Family of inhibition curves at different times after mixing PA peptide and trypsin (enzyme concentration of 0.78 μM); for clarity, only a subset of the experimental results is presented. Curves are best fits assuming tight-binding behavior (eq 2). Dotted lines depict the change in the apparent binding stoichiometry (eq 3). (b) Active fraction of the PA peptide (x) and its error as a function of time after mixing as derived from the inhibition curves in panel a. The data are described by a single-exponential curve assuming an apparent first-order process. The active fraction of peptide in the absence of enzyme and its rate of tryptic hydrolysis were estimated from the intercept at time zero and the curvature, respectively.

is apparent that peptides PP and PA are far more potent inhibitors than AP and AA. The first pair produces curves that are characteristic of tight-binding inhibition, where the extrapolated x-intercept represents the apparent binding stoichiometry. It would be anticipated that this intercept represents a 1:1 enzyme:peptide ratio, but in each case, the intercept is in the region of a 1:2 enzyme:peptide ratio. There are several reasons why this might be the case. First, the peptide concentration could be overestimated, possibly because the sample was not completely pure. Second, the peptide could be partially hydrolyzed over the measurement period, reducing the active concentration. Third, the enzyme−inhibitor interaction is unlikely to be instantaneous, and so early measurements could give misleading stoichiometries. Finally, it is possible that the peptide is heterogeneous with respect to activity as a result of having multiple conformations. The measurements of active enzyme concentration were determined by active site titration. While experimental error principally places a limit on the accuracy at which enzyme and peptide concentrations can be determined, we invariably observed an apparently substoichiometric binding behavior and believe that a combination of the latter three reasons is the most likely explanation. Figure 2a shows for the PA peptide a compilation of inhibition data collected for a range of incubation times and fitted assuming a tight-binding model. As for all four peptides that were analyzed, the apparent x-intercept grows larger with time, consistent with the amount of active peptide decreasing as a result of hydrolysis. A plot of x as a function of incubation time follows an apparent first-order process, as expected for the hydrolysis reaction (Figure 2b). This type of plot was used to estimate the active fraction of peptide in the absence of enzyme and its rate of tryptic hydrolysis from the intercept at time zero and the curvature, respectively. The extrapolated x value at time zero for PA is such that only just more than 50% of the peptide is apparently active and hydrolyzed by trypsin with a half-life of several minutes. The inhibition data acquired in the same manner for all four peptides are summarized in Table 2. These data show clear differences

Table 2: Trypsin Inhibition Parameters of the P3′/P4′ Ala Variants Derived from Data as Shown in Figure 2[a]

variant	K_d (nM)	K_i (nM)	x at time zero	$t_{1/2}$ (min)
PP	35 ± 5	19 ± 4	0.55 ± 0.04	7.2 ± 1.8
PA	167 ± 11	74 ± 8	0.44 ± 0.05	3.9 ± 1.0
AP	5500 ± 1100	116 ± 22	0.021 ± 0.002	0.6 ± 0.1
AA	18000 ± 2600	253 ± 36	0.014 ± 0.003	0.7 ± 0.2

[a] K_d is the equilibrium dissociation constant calculated using the tight-binding model. K_i is the intrinsic inhibition constant calculated for the active fraction of peptide. The inhibitory active fraction (x) of the peptide in the absence of enzyme and its estimated half-life of tryptic turnover ($t_{1/2}$) were estimated from the incubation time dependency of x in assays with an enzyme concentration of 0.31 ± 0.05 μM.

between the two peptides PP and PA and the others. For PP and PA, the data are consistent with a major fraction of the sample (roughly 50%) showing potent inhibition. In contrast, for peptides AP and AA, the weaker inhibitory activity can be fitted to either a tight-binding or a non-tight-binding model (Figure 1). The data better fit to the tight-binding curve and indicate a loss of inhibitory potency of 1−2 orders of magnitude compared with those of PP and PA. The extrapolated x values at time zero, however, would require an apparent inhibitor:enzyme stoichiometry of 48:1 (AP) or 71:1 (AA). Such a result is possible if just a small fraction of the peptide material exhibits inhibitory activity. Analysis of the inhibition at various times following addition of enzyme reveals a more rapid loss of activity in peptides with Ala at P3′ than when a Pro is present at this position. The apparent half-lives for the P3′ Ala peptides are less than a minute under the conditions of the assay, compared with several minutes for the P3′ Pro peptides. This behavior suggests that these peptides act as slowly hydrolyzed substrates.

The inhibition analysis of PP suggests that the biological activity of this peptide, even after accounting for the effect of tryptic hydrolysis by back-extrapolation to zero exposure time, is not entirely uniform. While one fraction of the peptide is a tight-binding inhibitor with an estimated intrinsic inhibition constant of 19 nM, there appears to be another less or completely inactive fraction. Affinity chromatography was used to separate these fractions. An excess of immobilized trypsin was used to deplete a solution of the PP peptide of the biological active fraction. Aliquots of the peptide remaining in solution were then repeatedly tested for trypsin inhibition. The results are depicted in Figure 3 and show that this fraction of peptide regains inhibitory potency with time. Such behavior would be expected for interchanging conformers that are distinguished by different biological activity.

NMR Analysis. The amide region of the ^{1}H NMR spectrum given by each of the four peptides is compared in Figure 4, which also shows the spectrum following incubation of the peptides with trypsin. The NMR spectrum of the intact PP peptide is characteristic of a well-structured dominant conformer with a wide dispersion of amide resonances. In marked contrast, the PA peptide exhibits roughly twice the expected number of amide peaks, which are also well-dispersed. This is due to the presence of two dominant conformers, each present in roughly equal proportions. Full assignment of the spectrum (Supporting Information) reveals that these represent *cis* and *trans* isomers of the P3′ Pro, which are in slow exchange (no exchange cross-peaks correlating the two conformers were observed in ROESY spectra). The high populations and good spectral resolution of the two sets of amide resonances originating from the two *cis*−*trans* isomers made it possible to monitor the time course of tryptic hydrolysis of the PA variant peptide directly by NMR. Figure 5 shows examples at two different inhibitor: enzyme ratios. At a low enzyme concentration (a), both PA

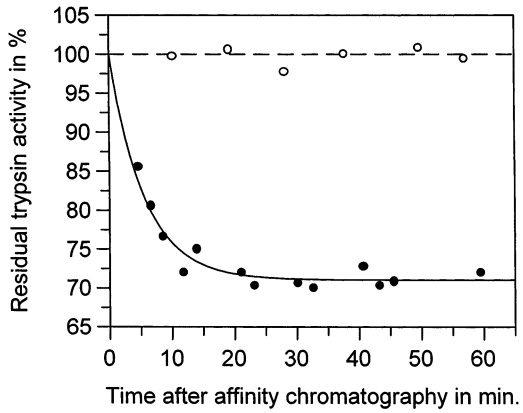

FIGURE 3: Time course of interconversion of the inhibitory inactive form of PP peptide and its inhibitory active form monitored by the residual trypsin activity. The inactive fraction was isolated by exposing the PP peptide solution for 1 min to a severalfold excess of trypsin immobilized on Sepharose 4B. The solution containing the nonbinding fraction of PP peptide was rapidly separated from the immobilized trypsin, and aliquots were repeatedly assayed for trypsin inhibition. The filled circles show the increasing level of inhibition of trypsin with time, and are described by a single-exponential decay curve with an estimated relaxation time of 5.6 min (solid line). Empty circles are results of control experiments in the absence of peptide. Peptide and enzyme concentrations were 25 and 5 μM, respectively.

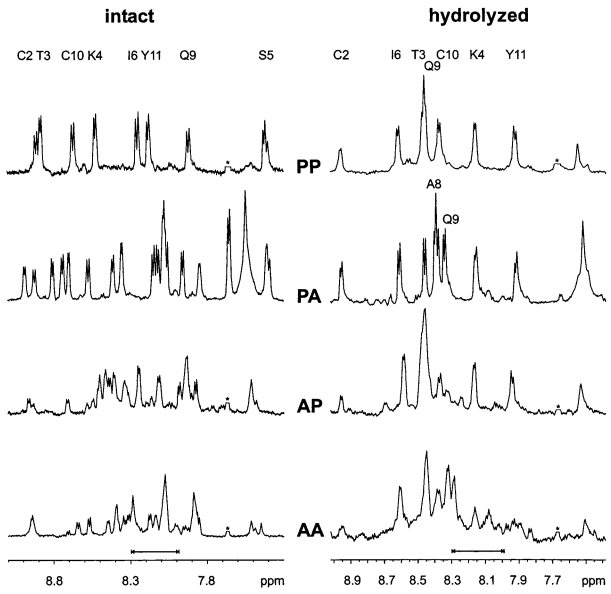

FIGURE 4: Amide region of one-dimensional NMR spectra of the four P3′/P4′ variant peptides before and after tryptic hydrolysis in 100 mM phosphate buffer at pH* 3.8 and 298 K. Backbone amide resonances are annotated for PP. The resonances in hydrolyzed PP and PA are virtually identical except for the annotated additional A8 and the shifted Q9 resonances. The absence of an S5 amide resonance is consistent with hydrolysis at the scissile peptide bond. An asterisk marks the cropped resonance of an impurity in the phosphate buffer. The arrow at the bottom indicates the random coil chemical shift range for the residues constituting the peptides (*47*).

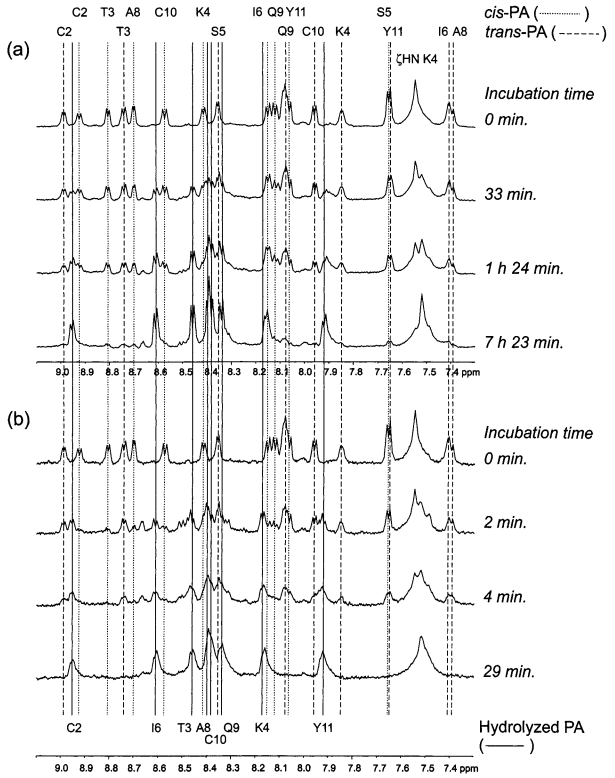

FIGURE 5: Time course of tryptic hydrolysis of the PA variant monitored by NMR. The experiment was performed at inhibitor:enzyme ratios of 735:1 (a) and 60:1 (b) in 100 mM phosphate buffer at pH* 3.8 and 298 K. The HN resonances of *cis*-PA, *trans*-PA, and hydrolyzed PA are labeled.

conformers appear to be converted into the same form of hydrolyzed material at the same rate. This is in sharp contrast to the behavior observed if the enzyme concentration is closer to stoichiometry (b). Here, the *cis*-PA conformer is processed at a significantly higher rate, with the *trans*-PA conformer lagging behind. This can best be seen in the spectrum at the 4 min incubation time. While the *cis*-PA conformer resonances are at the detection limit, intense resonances of the *trans*-PA conformer still remain. At 29 min, no resonances of intact PA can be detected. This demonstrates that trypsin preferably, possibly even selectively, binds and processes the *cis*-PA isomer.

The spectrum of the PA hydrolysis product is virtually identical to that of PP, except for the substituted position (Figure 4). The retention of well-dispersed amide resonances implies that hydrolysis of these peptides is not accompanied by a complete loss of structural integrity, although increased conformational averaging is evident from mostly intermediate $^3J_{HNH\alpha}$ coupling constants ranging between 6.6 and 8.3 Hz.

The AP and AA peptides generate spectra that also indicate conformational heterogeneity, but appear to comprise a more complex mixture of structures. Hydrolysis of AP and AA increases the similarity with the spectra of hydrolyzed PP and PA while still retaining a larger element of complexity, compared with the spectra of intact peptides.

Structural Characterization of cis-PA and trans-PA Isomers. Figure 6 summarizes the NMR parameters characterizing the dominant conformers of PP, *cis*-PA, and *trans*-PA. The NMR parameters of PP and *cis*-PA, including chemical shifts, NOE connectivities, coupling constants, and temperature coefficients, are almost identical. Minor differences can be traced to the substituted position. The backbone amide of the neighboring Gln9 has lost its hydrogen bond donor characteristic as expressed in the changes in the temperature coefficient and chemical shift. The downfield shift that is typically induced by an adjacent Pro can account for the chemical shift change of the preceding Pro7 Hα (*46*). We conclude that *cis*-PA adopts the proteinlike β hairpin structure of PP with a type VI β turn centered at the Ile6-Pro7 pair (Figure 7a). The backbone geometry appears to be slightly altered at the substituted position, which disrupts one of the cross-strand hydrogen bonds.

The NMR-derived parameters of *trans*-PA are very different from those of both PP and *cis*-PA. The increased diversity of its patterns of chemical shift deviations (less pronounced downfield deviation from random coil values), of sequential NOE correlations (increased number of HN_i-HN_{i+1} NOEs at the expense of attenuated $H\alpha_i-HN_{i+1}$ NOEs), and of attenuated $^3J_{HNH\alpha}$ coupling constants all suggest a decreased element of regular β-strand conformation. However, the retention of significant deviations from random coil values with regard to chemical shifts (*47*), $^3J_{HNH\alpha}$ coupling constants (*48*), ratios of sequential backbone NOE intensities (*49*), and backbone amide temperature coefficients (*36*) in combination with the observation of some nonsequential NOEs suggest elements of significant structural integrity within the *trans*-PA isomer. To identify such elements, putative simulated structures were calculated for *trans*-PA based on NMR-derived parameters. Figure 7 presents a view of a family of lowest-energy structures for *trans*-PA in an orientation chosen to highlight both similarities and differences with the family of lowest-energy structures calculated for PP, which were both calculated with the same simulated annealing protocol. The family of putative *trans*-PA structures is superimposed over the backbone of the turn region (Ile6−Gln9), which is shifted by one residue toward the C-terminus compared with PP. The turn in *trans*-PA is experimentally determined by a pair of nonsequential NOEs between the side chain of Ile6 and HN of Gln9 (Figure 6c). As the latter exhibits a temperature coefficient characteristic of a hydrogen bond donor, the family of structures, which was calculated without hydrogen bond restraints, allows the identification of the carbonyl of Ile6 as the acceptor. This confirms the location of the turn. Unexpectedly, the Lys4−Ile6 backbone arrangement in the family of *trans*-PA structures conforms with the requirements for the P2−P1′ stretch of the canonical conformation (*2*). This would place Ser5 in a P1-like hyperexposed position, while Lys4, which occupies the P1 position in PP but a P2-like position in *trans*-PA, is intramolecularly oriented toward the C-terminus located on the opposite strand. The family of *trans*-PA structures shows increased disorder toward the termini, which hampered the unambiguous identification of hydrogen bond acceptors for experimentally determined donors other than HN of Gln9.

DISCUSSION

While it is difficult to relate exactly the proportions of different conformational isomers with the apparent inhibitory stoichiometry described above (the pH and concentration conditions are necessarily different), it is likely that conformational heterogeneity is a major factor in the lower-than-expected inhibition stoichiometries extrapolated for time zero

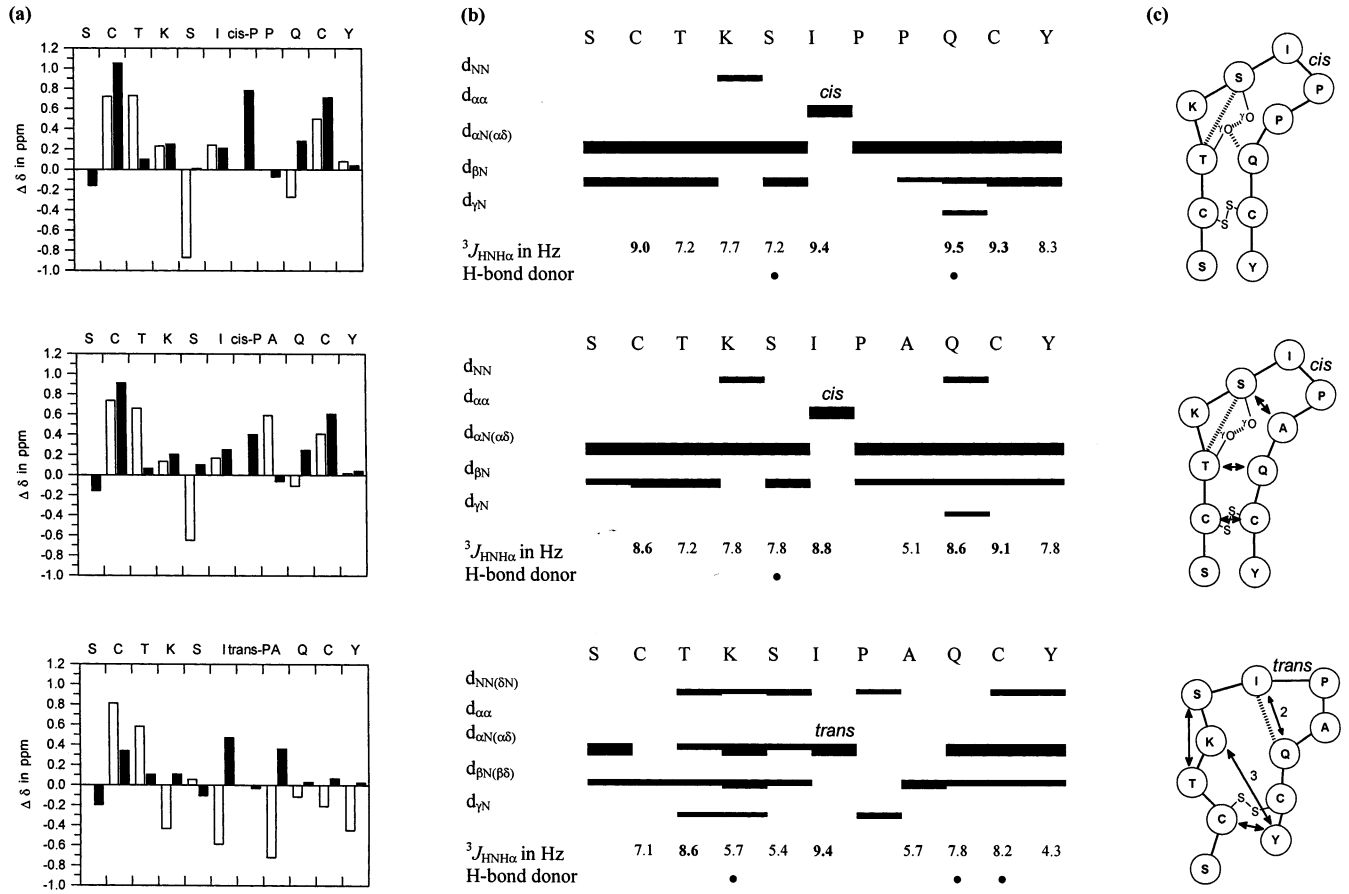

FIGURE 6: NMR parameters characterizing the backbones of the dominant conformations of the PP and PA peptides in a 90% H_2O/10% ^2H_2O mixture at pH* 3.8 and 298 K. (a) Deviation $\Delta\delta$ of HN (□) and Hα (■) chemical shifts from random coil values (*47*). (b) Sequential interproton NOE connectivities (relative intensities are represented by the thicknesses of the lines), $^3J_{HNH\alpha}$ coupling constants [values of >8.5 Hz characteristic of extended or β-strand conformation are highlighted in bold (*48*)], and backbone amide temperature coefficient indicative of hydrogen bonding [● (*36*)]. (c) Schematic representation of the structures. Dashed lines represent potential hydrogen bonds. Arrows indicate nonsequential NOE connectivities. All nonsequential NOEs observed for *cis*-PA and none of the nonsequential NOEs observed for *trans*-PA were observed for PP.

(Table 2). There appears to be an overall correlation between the presence of conformers that have a *cis* P3′ and inhibitory activity. For the peptides that were studied that contain a P3′ Pro, the dominant conformer adopts the *cis* conformation found in native BBI proteins. Replacement of the P3′ Pro with Ala results in peptides that lack well-defined structure and inhibitory potency, which explains the absolute requirement for a Pro in P3′ of BBI proteins. It is interesting to speculate about whether the small residual activity of 1−2% in P3′ Ala peptides originates from a small population of a nativelike conformer with a *cis*-Ala. This would be consistent with an experimental study that found populations of *cis*-peptide bonds of 0.1−1% at aromatic residues in linear di- and oligopeptides in aqueous solution at 295 K (*50*).

The role of a Pro at the P4′ position becomes particularly clear from a comparison of the NMR spectra of the intact PP and PA variants. The P4′ Pro → Ala substitution gives rise to an additional highly populated conformer, which is characterized by a *trans*-Pro at P3′. The rigidifying effect of the P4′ Pro can be considered to be a direct consequence of Pro's cyclic nature that drastically restricts the allowed backbone conformations in Pro residues (*51*). These additional restraints on the backbone of the peptide also appear to reduce the rate of tryptic hydrolysis. The PP peptide is approximately twice as resistant to tryptic hydrolysis as the *cis*-PA variant (Table 1). We further attribute this attenuation of resistance to the associated observed corruption of the cross-strand hydrogen bond which originates from the P5′ backbone amide in PP. This hydrogen bond has previously been shown to confer resistance against proteolytic turnover (*7*). In a very recent publication, Descours et al. have performed an alanine scan of the sunflower trypsin inhibitor sequence, which is a short cyclic peptide having an inhibitory loop that is highly related to the BBI family (*18*). They found that activity was abolished following replacement of the first but not the second proline, in line with the results presented in this paper.

Trypsin preferably binds the *cis*-PA rather than the *trans*-PA isomer. This isomer selectivity can be explained in structural terms. *cis*-PA adopts a nativelike structure with a hyperexposed P1 Lys, which is rather buried in the *trans*-PA structure. Instead, Ser5 appears to occupy an equivalent hyperexposed position in the *all-trans* structure. This is not favored for an interaction with trypsin's S1 binding pocket and would in addition result in a further unfavorable interaction between Pro7 and the S2′ subsite (*52*).

trans-Pro selective hydrolysis has been described for several proteases. Pepsin (*53*), proline-specific endopeptidase (*54*), prolidase, aminopeptidase P, carboxypeptidase P, and HIV-1 protease (*55*) all hydrolyze substrates with a P1 Pro

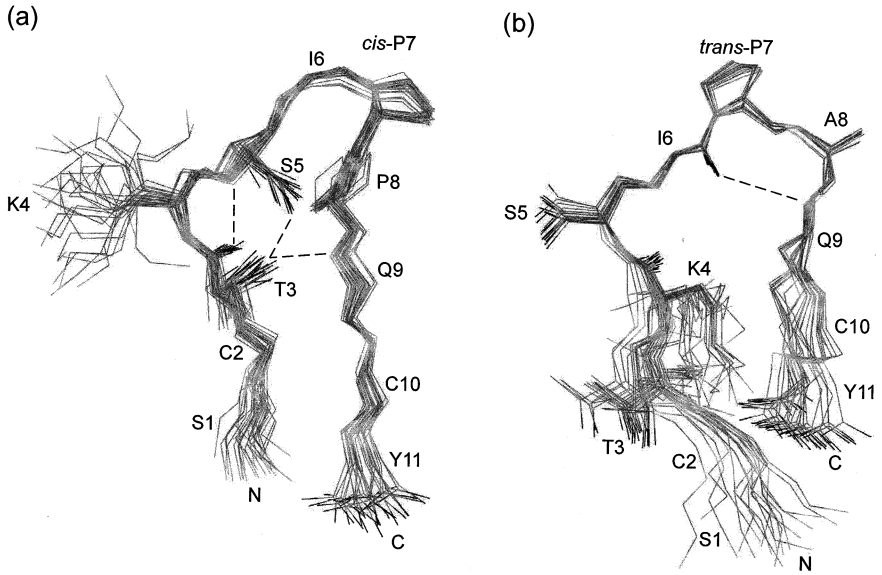

FIGURE 7: Families of the 30 lowest-energy structures calculated from NMR data for PP (a) (PDB entry 1GM2) and *trans*-PA. The structures are superimposed over the backbone of residues 2−10 for PP (pairwise rmsd of 0.43 ± 0.14 Å) and over the backbone of turn residues 6−9 for *trans*-PA (pairwise rmsd of 0.59 ± 0.37 Å). Complete backbones and selected side chains are shown (O, C, and N atoms in increasingly lighter shades of gray; the disulfide bridge is omitted for clarity). The termini and residues are labeled. Dashed lines represent selected potential hydrogen bonds. No hydrogen bond restraints were used in the structure calculations; potential hydrogen bonds were identified with the program HBPLUS (*68*). The orientation of the structures is to highlight the retention of a P2−P1′ stretch of canonical-like conformation in the putative *trans*-PA structure (residues K4−I6; pairwise rmsd over the backbone of 0.19 ± 0.26 Å). While K4 is hyperexposed in PP at the P1 position for the primary interaction with trypsin, this residue is rather shielded in *trans*-PA, in which S5 appears to adopt a P1-like prominence.

selectively in the *trans* conformation. Trypsin hydrolyzes substrates with P2′ Pro residues only in the *trans* conformation (*56*, *57*). Substrates with P2 Pro residues are cleaved more efficiently in the *trans* than in the *cis* conformation by chymotrypsin and trypsin (*57*, *58*). However, selective interaction and hydrolysis of a ligand defined by a *cis*-Pro conformation in the P3′ position is, to the best knowledge of the authors, described here for the first time. The isomer selectivity is particularly remarkable because the P3′ *cis*-Pro is distant from the scissile bond and does not make any direct contact with the enzyme surface in trypsin-complexed crystal structures (*10*, *12*, *20*, *59*, *60*). It can therefore be deduced that its role is purely to direct and restrain the backbone conformation of the reactive site loop. The isomer-selective interaction with trypsin implies that the biological activity of the proteinomimetic is essentially determined by the proteinlike isomer. Thus, for example, the 4-fold weaker inhibition of PA compared with that of PP may in parts be attributed to the loss of the hydrophobic interaction between the P4′ Pro side chain and the γ-methylene of trypsin's Gln192, which has been observed in trypsin-complexed crystal structures (*10*, *12*, *20*, *59*, *60*).

The isomer-selective hydrolysis of the PA peptide and the relaxation time of approximately 340 s of the PP conformers (Figure 3), which is consistent with relaxation times of tens to hundreds of seconds observed for Pro *cis*−*trans* isomerization in peptides and proteins under comparable experimental conditions (*61*), suggest a hydrolytic pathway as follows. Inhibitors with *cis* and *trans* P3′ are in slow equilibrium; trypsin binds and hydrolyzes selectively the *cis* P3′ form. At high enzyme concentrations, the *trans* to *cis* conversion will be rate-limiting and result in the observed lag phase for the turnover of *trans*-PA. At low enzyme concentrations, hydrolysis would be rate-limiting, resulting in the loss of *cis* and *trans* signals at apparently equal rates. This behavior is observed in Figure 5. Once hydrolyzed, the peptide is released, possibly still in the *cis* conformation, and is free to re-equilibrate.

In conclusion, a *cis* Pro at the P3′ position of the reactive site loop is essential for biological activity within BBI-derived protease inhibitors. It constitutes the center of a type VI β turn, which allows for a register of the antiparallel β strands that support the bioactive canonical conformation. The P3′ *cis*-Pro is supported by a second Pro at P4′, the substitution of which results in the emergence of a highly populated *all-trans* isomer. This isomer is characterized by a non-native turn, which is shifted by one residue toward the C-terminus. This finding supports the notion that the turn sequence is a critical determinant of the register of β hairpins (*62*−*66*). The P3′ *cis*-Pro residue, however, does not interact directly with the enzyme surface. Its role is therefore purely that of a powerful conformational determinant.

ACKNOWLEDGMENT

We thank Dr. H. Toms and P. Haycock of the University of London 600 MHz NMR facility at Queen Mary and Westfield College for their expert advice and assistance. We thank John Barton from Imperial College, Department of Chemistry, for performing the mass spectrometry.

SUPPORTING INFORMATION AVAILABLE

Assignment tables of the two isomers of the PA peptide (including coupling constants and amide temperature coefficients), statistics for the calculated family of *trans*-PA structures, and a Ramachandran plot of the family of *trans*-PA structures. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Laskowski, M., Jr., and Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593−626.
2. Bode, W., and Huber, R. (1992) *Eur. J. Biochem. 204*, 433−451.
3. Apostoluk, W., and Otlewski, J. (1998) *Proteins: Struct., Funct., Genet. 32*, 459−474.
4. Laskowski, M., and Qasim, M. A. (2000) *Biochim. Biophys. Acta 1477*, 324−337.
5. Jackson, R. M., and Russell, R. B. (2000) *J. Mol. Biol. 296*, 325−334.
6. McBride, J. D., and Leatherbarrow, R. J. (2001) *Curr. Med. Chem. 8*, 909−917.
7. McBride, J. D., Brauer, A. B. E., Nievo, M., and Leatherbarrow, R. J. (1998) *J. Mol. Biol. 282*, 447−457.
8. Prakash, B., Selvaraj, S., Murthy, M. R., Sreerama, Y. N., Rao, D. R., and Gowda, L. R. (1996) *J. Mol. Evol. 42*, 560−569.
9. Werner, M. H., and Wemmer, D. E. (1992) *Biochemistry 31*, 999−1010.
10. Lin, G. D., Bode, W., Huber, R., Chi, C. W., and Engh, R. A. (1993) *Eur. J. Biochem. 212*, 549−555.
11. Ikenaka, T., and Norioka, S. (1986) in *Proteinase Inhibitors* (Barrat, A. J., and Salvensen, G., Eds.) pp 361−374, Elsevier, Amsterdam.
12. Koepke, J., Ermler, U., Warkentin, E., Wenzl, G., and Flecker, P. (2000) *J. Mol. Biol. 298*, 477−491.
13. Voss, R. H., Ermler, U., Essen, L. O., Wenzl, G., Kim, Y. M., and Flecker, P. (1996) *Eur. J. Biochem. 242*, 122−131.
14. Tsunogae, Y., Tanaka, I., Yamane, T., Kikkawa, J., Ashida, T., Ishikawa, C., Watanabe, K., Nakamura, S., and Takahashi, K. (1986) *J. Biochem. 100*, 1637−1646.
15. Song, H. K., Kim, Y. S., Yang, J. K., Moon, J., Lee, J. Y., and Suh, S. W. (1999) *J. Mol. Biol. 293*, 1133−1144.
16. Brauer, A. B. E., Kelly, G., Matthews, S. J., and Leatherbarrow, R. J. (2002) *J. Biomol. Struct. Dyn.* (in press).
17. Brauer, A. B. E., Kelly, G., McBride, J. D., Cooke, R. M., Matthews, S. J., and Leatherbarrow, R. J. (2001) *J. Mol. Biol. 306*, 799−807.
18. Descours, A., Moehle, K., Renard, A., and Robinson, J. A. (2002) *ChemBioChem 3*, 318−323.
19. Korsinczky, M. L., Schirra, H. J., Rosengren, K. J., West, J., Condie, B. A., Otvos, L., Anderson, M. A., and Craik, D. J. (2001) *J. Mol. Biol. 311*, 579−591.
20. Luckett, S., Garcia, R. S., Barker, J. J., Konarev, A. V., Shewry, P. R., Clarke, A. R., and Brady, R. L. (1999) *J. Mol. Biol. 290*, 525−533.
21. McBride, J. D., Freeman, N., Domingo, G. J., and Leatherbarrow, R. J. (1996) *J. Mol. Biol. 259*, 819−827.
22. McBride, J. D., Freeman, H. N. M., and Leatherbarrow, R. J. (1999) *Eur. J. Biochem. 266*, 403−412.
23. McBride, J. D., Freeman, H. N. M., and Leatherbarrow, R. J. (2000) *J. Pept. Sci. 6*, 446−452.
24. Domingo, G. J., Leatherbarrow, R. J., Freeman, N., Patel, S., and Weir, M. (1995) *Int. J. Pept. Protein Res. 46*, 79−87.
25. Focault, G., Seydoux, F., and Yon, J. (1974) *Eur. J. Biochem. 47*, 295−302.
26. Nakata, H., and Ishii, S. (1972) *J. Biochem. 72*, 281−290.
27. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
28. Leatherbarrow, R. J. (1992) *GraFit*, Erithacus Software Ltd., Staines, U.K.
29. Tonomura, B., Senda, M., Tsura, D., Komiyama, T., Miwa, M., Akaska, K., Kainosho, M., Tsuji, T., Takahashi, K., Hiromi, K., Ikenaka, T., and Murao, S. (1985) in *Protein Protease Inhibitor: The Case of Streptomyces Subtilisin Inhibitor* (Hiromi, K., Akaska, K., Mitsui, Y., Tonomura, B., and Murao, S., Eds.) pp 291−362, Elsevier, Amsterdam.
30. Derome, A. E., and Williamson, M. P. (1990) *J. Magn. Reson. 88*, 177−185.
31. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
32. Jeener, J., Meier, B. H., Bachman, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
33. Griesinger, C., and Ernst, R. R. (1987) *J. Magn. Reson. 75*, 261−271.
34. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
35. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., and Wright, P. E. (1988) *J. Mol. Biol. 201*, 161−200.
36. Andersen, N. H., Neidigh, J. W., Harris, S. M., Lee, G. M., Liu, Z. H., and Tong, H. (1997) *J. Am. Chem. Soc. 119*, 8547−8561.
37. Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., and Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611−639.
38. Pardi, A., Billeter, M., and Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741−751.
39. DeMarco, A., Llinás, M., and Wüthrich, K. (1978) *Biopolymers 17*, 617−636.
40. Cai, M. G., Huang, Y., Liu, J. H., and Krishnamoorthi, R. (1995) *J. Biomol. NMR 6*, 123−128.
41. Kessler, H., and Seip, S. (1994) in *Two-Dimensional NMR Spectroscopy* (Croasmun, W. R., and Carlson, R. M. K., Eds.) pp 619−654, VCH Publishers, Weinheim, Germany.
42. Hofmann, M., Gondol, D., Bovermann, G., and Nilges, M. (1989) *Eur. J. Biochem. 186*, 95−103.
43. Brünger, A. T. (1993) *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
44. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis 18*, 2714−2723.
45. Vriend, G. (1990) *J. Mol. Graphics 8*, 52−56.
46. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1996) *J. Biomol. NMR 5*, 67−81.
47. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
48. Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., and Dobson, C. M. (1996) *J. Mol. Biol. 225*, 494−506.
49. Maynard, A. J., Sharman, G. J., and Searle, M. S. (1998) *J. Am. Chem. Soc. 120*, 1996−2007.
50. Scherer, G., Kramer, M. L., Schutkowski, M., Reimer, U., and Fischer, G. (1998) *J. Am. Chem. Soc. 120*, 5568−5574.
51. Milner-White, E. J., Bell, L. H., and Maccallum, P. H. (1992) *J. Mol. Biol. 228*, 725−734.
52. Laskowski, M., Jr. (1980) *Biochem. Pharmacol. 29*, 2089−2094.
53. Vance, J. E., LeBlanc, D. A., and London, R. E. (1997) *Biochemistry 36*, 13232−13240.
54. Lin, L. N., and Brandts, J. F. (1983) *Biochemistry 22*, 4480−4485.
55. Vance, J. E., LeBlanc, D. A., Wingfield, P., and London, R. E. (1997) *J. Biol. Chem. 272*, 15603−15606.
56. Lin, L. N., and Brandts, J. F. (1983) *Biochemistry 22*, 553−559.
57. Lin, L. N., and Brandts, J. F. (1985) *Biochemistry 24*, 6533−6538.
58. Fischer, G., Bang, H., and Mech, C. (1984) *Biomed. Biochim. Acta 43*, 1101−1111.
59. Li, Y. L., Huang, Q. C., Zhang, S. W., Liu, S. P., Chi, C. W., and Tang, Y. Q. (1994) *J. Biochem. 116*, 18−25.
60. Zhu, Y. S., Huang, Q. C., and Chi, C. W. (1999) *J. Biomol. Struct. Dyn. 16*, 1219−1224.
61. Fischer, G. (2000) *Chem. Soc. Rev. 29*, 119−127.
62. de Alba, E., Rico, M., and Jimenez, M. A. (1999) *Protein Sci. 8*, 2234−2244.
63. de Alba, E., Rico, M., and Jimenez, M. A. (1997) *Protein Sci. 6*, 2548−2560.
64. de Alba, E., Jimenez, M. A., and Rico, M. (1997) *J. Am. Chem. Soc. 119*, 175−183.
65. Jourdan, M., Griffiths-Jones, S. R., and Searle, M. S. (2000) *Eur. J. Biochem. 267*, 3539−3548.
66. Griffiths-Jones, S. R., Maynard, A. J., and Searle, M. S. (1999) *J. Mol. Biol. 292*, 1051−1069.
67. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
68. McDonald, I. K., and Thornton, J. M. (1994) *J. Mol. Biol. 238*, 777−793.

BI026050T