Role of Rigidity on the Activity of Proteinase Inhibitors and Their Peptide Mimics†

Joao R. Costa and Sophia N. Yaliraki*

Department of Chemistry, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom

Received: December 27, 2005; In Final Form: July 11, 2006

The Bowman−Birk inhibitors (BBIs) are a family of proteins that share a canonical loop structure whose presence in a conserved conformation is linked to their inhibitory activity. We study the conformational properties of the canonical loop using a graph theoretical approach as implemented in the floppy inclusions and rigid substructure topography (FIRST). We find that the canonical loop is an independent rigid cluster in the natural inhibitors. We have further used this technique to identify residues that play an important role in the structural rigidity of the protein by quantifying their contribution to the overall rigidity of the inhibitor. We find that the conserved elements among the natural and synthetic peptides are the ones that contribute the most to rigidity, even if they are located far from the active site, as rigidity effects are nonlinear and hence nonlocal. The results help to elucidate why certain mutations in the loop of the BBI produce peptides that fail to have the designed inhibitory activity.

I. Introduction

Protein flexibility is associated with a wide range of functional activities. Rather than an incidental characteristic, studies have shown that global, large scale motions are fundamental for protein function in such diverse systems as the HIV protease[1] to GroES.[2] Characterizing such motions is proving to be a valuable tool in structural biology.[3]

Conformational motions play a role in enzymatic activity as well. The subtle interplay between chemical specificity and mechanical contribution to catalysis is an active area of research.[4] Since mechanical properties are a result of the whole dynamic network of interactions in a protein including those far away from the active site, remote residues, which may not necessarily alter the binding energy directly, may nevertheless still influence activity through nonlocal interactions.

Just like flexibility, rigidity can also be a key attribute. This is of special interest to the classes of inhibitors which bind to their respective enzymes retaining their original conformation. In such cases, rigidity is an advantageous attribute as it helps to ensure that the inhibitor is locked in a conformation complementary to the enzyme. Proteins of the family of the Bowman−Birk inhibitor (BBI) (Figures 1 and 2), a class of naturally occurring serine protease inhibitors which act to regulate enzymes such as trypsin by binding stochoimetrically to them, share a conserved reactive site loop (for a review, see ref 5). Furthermore, it is striking that the most potent such inhibitor known to date found in the sunflower seed; the SFTI-1[6] consists of only this loop in a cyclic form (see Figure 1). In addition to this naturally occurring reduced representation, there is considerable experimental effort in synthesising BBI mimics[5] that are minimal in size but still functionally active, as well as designing mutated ones with different or maximal activity (Figure 3). Experimental studies[6,7] have shown that the conformation of these inhibitors remains almost unchanged after complexation with the enzyme. Its conserved elements among different peptides, such as disulfide bridges, proline residues,

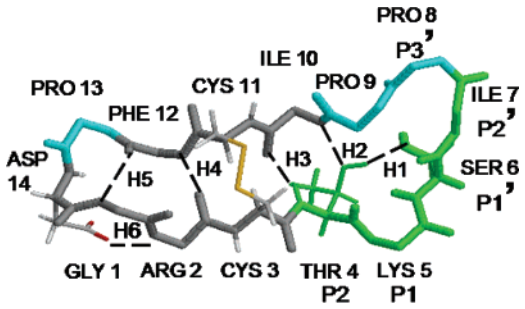

Figure 1. Typical shape of the canonical loop shared among the BBIs. The region colored in green (P2−P2′) is the active region that binds to the enzyme. The residue in position P1 is responsible for the specificity (lysine inhibits trypsin). The proline residues (blue), the disulfide bridge (yellow), and the network of hydrogen bonds (black dashed lines) are some of the elements that work together to restrict the conformational freedom of the peptide. The peptide in this figure is the sunflower trypsin inhibitor (SFTI-1), the most potent of the family known to date. All molecular images were generated with the RAMSOL visualization package.[37]

and sometimes cyclic backbones, seem to introduce clear identifiable constraints on the conformational degrees of freedom which suggests the importance of rigidity in this family of proteins. Further studies on mutated variants show that although the specific inhibitory activity is centered on position P1 (see Figure 1; nomenclature of Schechter and Berger[8]) of the loop, residues in other positions are still essential to the activity without necessarily being situated in the binding region. For example, the requirement that the cis-Pro residue be present in position P3′ for the protein to be active[9] suggests that there is a requirement for the protein to maintain an overall rigidity.

Detecting and quantifying the features that provide the appropriate rigidity to such proteins hence becomes a relevant question not only of theoretical interest but also of practical importance. Since preserving the overall structure of the original protein is crucial, it would be desirable to know a priori which residues are important to the overall structural rigidity. Questions such as the following could be answered: what is the importance

† Part of the special issue "Robert J. Silbey Festschrift".

* E-mail: s.yaliraki@imperial.ac.uk.

10.1021/jp0575299 CCC: $33.50
Published on Web 08/23/2006

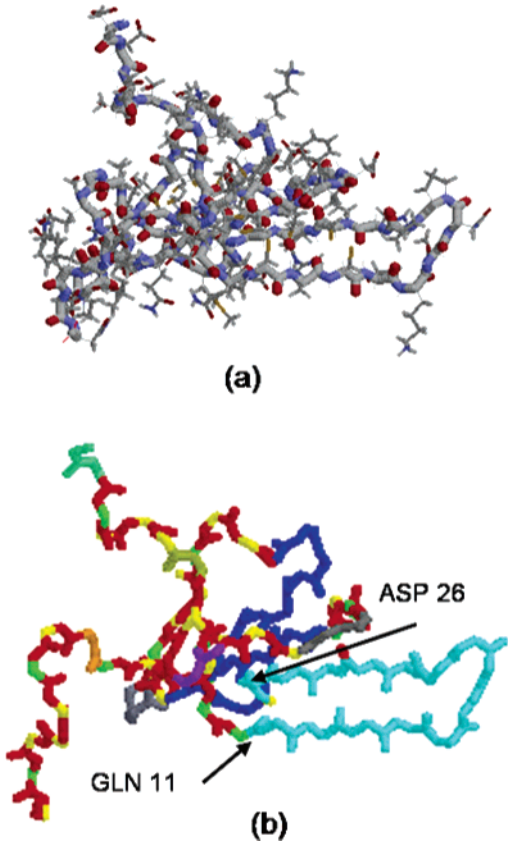

Figure 2. Rigid cluster decomposition of the soybean BBI. (a) Full-atom description as given by the pdb 1BBI (oxygen atoms are represented in red, and peptide bonds, in purple). (b) Rigid cluster decomposition obtained with FIRST. The canonical trypsin inhibitor region backbone forms a single rigid cluster independent of the rest of the protein (light blue, at the right end). Note that each rigid cluster is represented in a different color with the red and yellow parts indicating flexible regions. Removing the weakest bond of the canonical loop, the hydrogen bond between glutamine and aspartic acid with energy at the cutoff, leads to the rigid cluster becoming flexible. Note that there is another loop behind the structure, not as clearly visible, that inhibits for chymotrypsin.

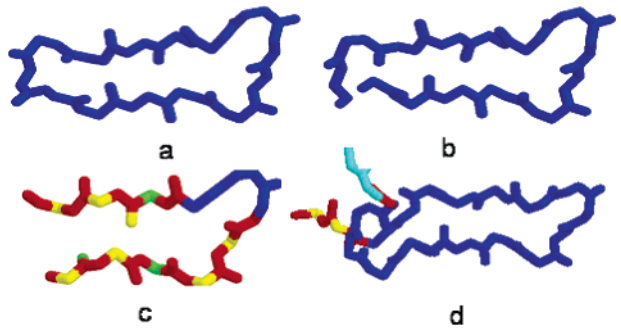

Figure 3. Rigid cluster decompositions for several small trypsin inhibitors: (a) SFTI-1, a natural inhibitor from the sunflower seed, the most potent inhibitor measured to date; (b) SFTI-1,14, the synthetic acyclic variant of SFTI-1; (c) 1GM2, a synthetic trypsin inhibitor; and (d) 1G9I, a synthetic tricyclic trypsin inhibitor. All the peptides are rigid except for 1GM2. Note that there are more clusters along the side chains, not depicted here.

of far-away residues in the flexibility of a given site or of the whole protein; which noncovalent interactions are essential in conferring the desired flexibility properties to peptides engineered in the lab; what is the shortest peptide which can hold the active region with the desired overall rigidity? Mutations of these residues could then be balanced against other requirements and help protein design by substantially narrowing the otherwise combinatorially growing number of possible candidates.

To investigate the role of rigidity of the conserved reactive loop, we perform a quantitative rigidity analysis on the various BBIs. The approach is based on obtaining the rigidity from graph representation approaches as developed in floppy inclusions and rigid substructure topography (FIRST), a computational tool for the analysis of proteins developed by Jacobs et al.[10] Starting from a full-atom description that includes covalent and noncovalent interactions, it is possible to derive a graph representation of the protein as a bond-bending network where the nodes are the atoms and the edges indicate distance constraints derived from the interatomic interactions.[10] Concepts from generic graph rigidity theory allow the identification of flexible (underconstrained) and rigid (overconstrained or isostatic) regions[11] which makes it possible to calculate for example which bonds can rotate, which segments of the protein behave as rigid bodies, and which bonds have correlated motions. Continuous regions of rigid bonds define rigid clusters which are linked by flexible bonds. This can be computed efficiently due to the pebble game algorithm that scales linearly with the number of atoms in 3D systems.[12] The output from FIRST can be used to assign a flexibility index to each amino acid[10] as well as to produce protein coarse-grained representations based on block rigidity.[13] In the past, this graph theoretical approach has been successfully applied to detect folding cores[14,15] and characterize the flexibility of several proteins[10,16] and complexes[17] and has been recently used to derive coarse-grained descriptions of supramolecular assemblies.[13]

Flexibility can be obtained by a variety of methods based on conventional normal mode analysis (NMA) on full-atom models[18,19] or on reduced models such as the elastic network model,[20−22] the Gaussian network model,[23] or other coarse-grained approaches.[24,25] It can also be deduced from principal components analysis performed on molecular dynamics (MD) simulations.[26] An advantage of the graph approach is that it uses a full-atom description to obtain the distance constraints and hence derive the set of rigid clusters instead of assuming them a priori and yet remains very efficient computationally, for example, by avoiding finding an energy minimum, the diagonalization of relatively large matrices, or long MD simulations.

An additional advantage of this approach stems from the fact that we can easily quantify the impact on rigidity of a given intramolecular interaction or group of interactions. For example, the impact of a hydrogen bond can be assessed by simply removing the corresponding edges from the graph and recalculating the rigidity of the protein. We show how this approach can be applied at the residue level of description to elucidate the nonlocal character of the mechanical properties of the proteins. This allows us to evaluate the role of residue positions within the loop and show how residue substitutions can alter the rigidity properties of the inhibitor. We can identify the impact of each residue both based on its position in the loop but also on its contribution to the global hydrogen bonding network. This feature makes it ideal for the purpose of identifying elements that are essential to achieving rigidity. It is important to point out that graph rigidity is a nonlinear property, and hence, we expect such contributions to depend heavily on the topology of the global network and the relative position of the interactions.

Our results show that the conserved loop in the natural inhibitors belongs to a single rigid cluster, which helps explain

TABLE 1: Measured Trypsin Inhibition Constants, K_i, in Nanomolar for the BBIs Considered[a]

inhibitor code	inhibition constant K_i (nM)
SFTI-1	0.5[6]
1BBI	0.6[35]
SFTI-1,14	12.1[6]
1GM2	19[7]
1G9I	120[36]

[a] SFTI-1 is the best inhibitor followed by 1BBI (we consider its trypsin inhibitor loop only). Note however that because of the experimental difficulty in measuring with accuracy the inhibition constant of potent inhibitors there is uncertainty in these values.

the success in isolating this motif. Its structural integrity is held by a tight number of successive rings of interconnected hydrogen bonds. This rigidity is marginal, which means a fluctuation that disturbs the hydrogen bond at the remote side of the active site leads to a flexible structure. Interestingly, we find that the residue which controls chemical specificity in position P1[5] plays no role in rigidity, while the positions associated with what are considered essential elements of the loop, proline residues and cysteine groups, give rise to high contributions to rigidity, in agreement with mutational experiments. The rigidity analysis helps to elucidate why certain mutations in the loop of the BBI produce peptides that fail to have the desired inhibitory activity even though they may be far from the active site or have similar binding interactions.

In section II, we describe the system in detail, and in section III, we summarize the methodology behind FIRST and define quantitative measures of rigidity that we introduce to characterize the BBIs. In section IV, we give the results and compare with experiments and end with a discussion and future directions.

II. The System

The BBIs are serine protease inhibitors (serpins) characterized by an inhibitory region that consists of a loop with a conserved shape, the so-called canonical conformation (see Figures 1 and 2) which is complementary to the protease active site.[27] They are small plant proteins studied for their anticarcinogenic properties. In many cases, the BBI consists of two homology domains (two loops) which are able to simultaneously inhibit two serine proteases, e.g., trypsin and chymotrypsin. The specificity of the inhibitor is dictated by the P1 position; for example, lysine in this position gives rise to a trypsin inhibitor whereas phenylalanine or alanine inhibit chymotrypsin instead. There are also several examples in the literature of small natural and engineered peptides that mimic the canonical loop of the BBIs.[5] Here, we focus on trypsin inhibitors only and in particular compare the full BBI (pdb code 1BBI, obtained by NMR; Figure 2) with the smallest yet most potent known to date naturally occurring one, the sunflower SFTI-1 (pdb code JBL, obtained by NMR; Figure 1). SFTI-1,14 (pdb code JBN, obtained by NMR; Figure 3b) is the synthetic acyclic version of the latter with the same sequence, but without a backbone bond between the glycine (1) and aspartic acid (14) residues. 1GM2 (pdb code 1GM2, obtained by NMR) is a yet shorter synthetic inhibitor with a mutation in position P5′ (Figure 3c), while 1G91 (pdb code 1G9I, obtained by X-ray in the complex) is a larger synthetic variant with three disulfide bridges (Figure 3d). A summary of the measured equilibrium inhibition constants, K_i, of these inhibitors is given in Table 1. Although there is agreement that all are reasonably good inhibitors in roughly that order, there is often uncertainty with the measurements of potent inhibitors.[6]

III. Methodology

Quantifying Rigidity Properties Using FIRST. From a given three-dimensional structure of a peptide, FIRST (we have used version 3.1 of the software throughout this work) identifies the salt bridges and hydrogen bonds using standard energy potentials.[10] This, together with the identification of hydrophobic contacts,[14] is a key step that allows the network of constraints to be built. The constraints are based on the bonds selected using the following criteria. There is an input parameter to the software, E_{cut}, that specifies the cutoff energy for a hydrogen bond to exist which can be loosely related to temperature.[10] The cutoff of -0.7 kcal/mol for hydrogen bonds was used throughout to model standard conditions. All peptide PDB coordinate files were downloaded from the protein databank and hydrogen atoms were added to X-ray derived structures using the software WHATIF.[28] For structures obtained from NMR experiments, where a range of possible structures was given instead, hydrogen bonds were retained only if the energy calculated by FIRST and averaged over all possible structures was above the cutoff. Since FIRST does not associate an energy value with the hydrophobic contacts, the criterion of acceptance of hydrophobic contacts was to select those that could be detected in more than 90% of the possible NMR structures.

Given the topology of the network, the analysis of the final graph representation of the protein gives the number of internal degrees of freedom (DF) of the system and identifies flexible and rigid regions of the protein which can be summed up with the number of rigid clusters. A protein can also be depicted by a rigid cluster decomposition where each color identifies a separate rigid cluster and red and yellow indicate flexible regions. Both the number of rigid clusters and DF are measures of the rigidity of the protein. The number of DF can be thought of as the number of variables necessary to specify the conformation of the protein. Removing nonredundant distance constraints from the system increases the number of degrees of freedom. However, this does not necessarily lead to an increase in the number of rigid clusters. At the same time, depending on the topology of the network, removing a constraint can substantially increase the number of clusters which increases the floppiness of the protein. If there is only one large rigid cluster, the molecule behaves as a single rigid body with no internal degrees of freedom. If there are many rigid clusters, then the different regions in the protein can move in relation to each other even though the movements may need to be correlated to ensure ring closure constraints.

The extent to which removing the constraints associated with one residue affects the total number of rigid clusters and DF is dependent not only on the number of constraints formed by that residue but also, crucially, on their location in the structure. To study the importance of a particular residue position for rigidity, we remove from the graph model all the hydrogen bonds, salt bridges, and hydrophobic contacts associated with that residue. In addition, if the residue is a proline, we remove the backbone dihedral constraint on the angle imposed by its five-membered ring and, if the residue is a cysteine, we remove the constraint associated with the disulfide bridge if one was present. The number of DF contributed by the position of each residue will be an indication of the importance of topology to the overall rigidity. This will be denoted as "position contribution to rigidity" to differentiate it from a different calculation, the "residue contribution to rigidity" where the constraints associated with the backbone atoms of the residue are not removed. The idea is to isolate the contribution to the rigidity that is specific to the residue (involving the side chains) from the contribution

that in principle can be made by any substitution at the same position in the protein. Similarly, we can quantify the contribution of a particular hydrogen bond to the overall rigidity. In summary, we will quantify contributions to rigidity in three different ways defined as follows:

- Position contribution to rigidity is the impact on rigidity of removing from the graph all the constraints associated with a given residue.
- Residue contribution to rigidity is the same as the position contribution to rigidity except that constraints imposed by the backbone are not removed from the graph, since these constraints are often not residue specific.
- Hydrogen bond contribution to rigidity is the impact on rigidity when a particular hydrogen bond is removed from the graph.

IV. Results and Discussion

We first study the soybean BBI, a large inhibitor which contains the canonical loop that selectively binds to trypsin. The rigid cluster decomposition of this inhibitor is given in Figure 2. It comprises of a flexible region with a large number of small rigid clusters, represented by yellow and red, and two large rigid clusters in light and dark blue. The light blue cluster on the right is 16 residues long and includes the trypsin active region (see the atoms colored in green in Figure 1 to identify the active region in the loop). It is striking that the smaller peptides, natural and synthetic, derived from the larger inhibitor that mimic with success the BBI loop have approximately the same size as this cluster (see Figure 3). In fact, the strongest trypsin inhibitor found to date, the cyclic peptide SFTI-1, is composed of just this loop whose backbone is also a single rigid cluster (Figure 3a). All the synthetic but still good inhibitors we looked at were rigid (Figure 3), with only the 1GM2 (Figure 3d) containing any flexible parts. Furthermore, if we remove the weakest hydrogen bond in the canonical loop cluster in BBI, which is the one between glutamine (residue 11) and aspartic acid (residue 26) with energy right at the cutoff, we find that this rigid cluster breaks apart abruptly and this whole region becomes flexible. This suggests that to stabilize the active region even residues in the opposite end of the active site play an important role in maintaining the overall conformation.

The rigidity of these structures can be understood as a consequence of the topology of intramolecular interactions. The interconnected hydrogen bonding network of rings, bridged over by the disulfide backbone, provides a strong set of distance constraints. To see how a chain of rings can provide rigidity to a structure, consider a very simple example. Given a ring and assuming that the distances between adjacent nodes and the angles that the edges form at each node are fixed, any 6-fold or smaller ring is rigid. When two or more rings are interconnected, the overall structure may be rigid or flexible depending on the way they are connected. When the rings are interconnected as in Figure 4c, the 8-fold ring loses one degree of freedom and thus is still flexible. In Figure 4d, an alternative connectivity results in the 8-fold ring losing two degrees of freedom and hence leads to the whole structure becoming rigid. The way the rings interconnect is the key to achieving rigidity (for a review on rigidity theory, see ref 29). This applies to the peptides studied here, only the number of rings is larger. When we consider both the number of degrees of freedom and rigid clusters, the results indicate that the rings are tuned to achieve rigidity of the canonical loop for the strong inhibitors (Figures 5−7). This is not an obvious result since there are many ways that the rings could be connected which could lead to small

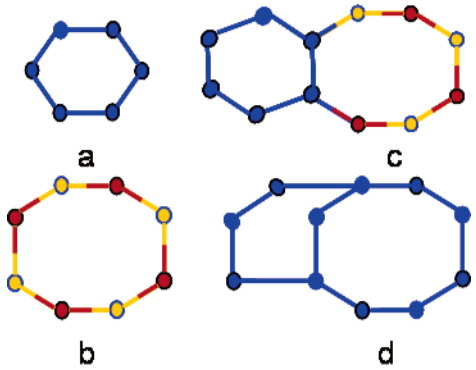

Figure 4. Effect of connectivity to overall rigidity. The rigid rings are displayed in a single color: (a) The six-fold ring is rigid; (b) The eight-fold ring is flexible with two internal degrees of freedom. Different connectivities of the six- and eight-fold rings lead to (c) a flexible or (d) a rigid eight-fold ring.

overstressed or floppy regions rather than rigid ones. Furthermore, note that this rigidity is marginal, that is, the peptides are rigid but do not have many redundant constraints. The implication of this is that any perturbation, for example, a fluctuation that removes a hydrogen bond, will immediately lead to increased flexibility which will allow for small adjustments to the conformation. Indeed, there are a number of constraints, which remove only few DF from the system but lead to a floppier structure by producing a large increase in the number of rigid clusters. This, in fact, may be a desirable compromise for the activity of the inhibitor.

We can also expect that mutations will interfere differently with the topology of the rings depending on their location and hence have different influence on the rigidity of the protein. For example, the change in the number of DF and clusters produced by adding or removing a particular amino acid is not always the same but depends on its location in the overall structure. This effect is distinct from the purely chemical contribution to binding and energetic stability. Of particular importance are substitutions that with only a small change in the number of DF produce a large effect in the number of clusters. The contribution of each residue position to the rigidity is presented in Figure 5a for the SFTI-1 peptide. The number of rigid clusters is significantly high for the positions near the termini. Similarly, the importance of the hydrogen bonds in Figure 5d seems to increase toward the termini. This indicates that the rigidity of the active region is strongly influenced by constraints on the remote side of the protein. So, it would seem that the shape of the active region is optimized to fit the inhibitor while the network of rings that extends in the opposite direction is responsible for providing rigidity to this shape.

The specific contribution of each residue to the rigidity of the SFTI-1 inhibitor (Figure 5b) reveals that the other high contributions are made by the cysteines (3 and 11) and the proline in position P3′ (8). Also having an important contribution is the residue 4 threonine (P2) and residues 6, 7, and 9. These results correlate well with the fact that most synthetic and natural sequences have threonine conserved at position P2, a proline in the position P3′, and the two cysteine residues linked by a disulfide bridge. Threonine is present in the P2 locus in more than 80% of the BBI loop sequences,[30] and McBride et al.[31] have found it to be the optimum residue at the P2 position when screening a library of 8000 peptides against a different but closely related enzyme, chymotrypsin. Perhaps more striking is the experimental finding of Brauer et al.[9] who have revealed that the proline at position P3′ in the cis-configuration is necessary for activity even though this residue is not in the active

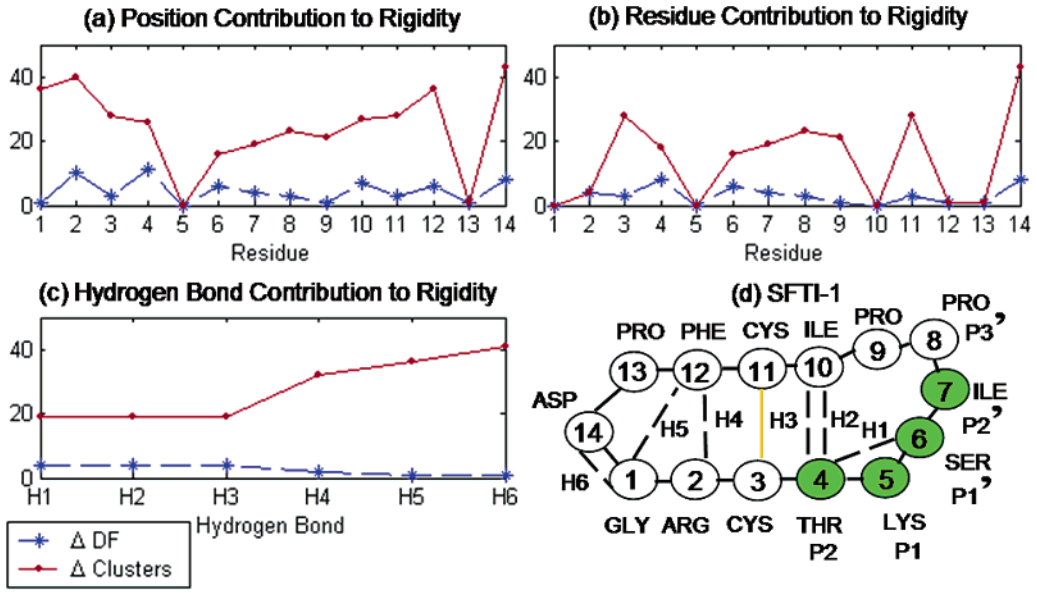

Figure 5. Rigidity study of the peptide SFTI-1. The *y* axis in a, b, and c represents the increase in the number of rigid clusters (Δ clusters, red circles) and in the number of degrees of freedom (Δ DF, blue asterisks) in comparison with the case where all the constraints are present as in d, where the number of DF is DF = 13 and the number of clusters is clusters = 14. The backbone forms a single rigid cluster (Figure 3a) while the remaining clusters occur in the side chains. (a) Position contribution to rigidity (all constraints associated with the residue are removed). (b) Residue contribution to rigidity (backbone constraints are preserved). (c) Hydrogen bond contribution to rigidity. (d) Topology of the network of hydrogen bonds. Here, we depict only those just above the cutoff, $E_{H1} = -3.9$, $E_{H2} = -4.9$, $E_{H3} = -3.0$, $E_{H4} = -2.6$, $E_{H5} = -1.7$, and $E_{H6} = -4.5$ kcal/mol. Note that the position P1 responsible for specificity plays no role in rigidity. On the contrary, the highest contributions occur when removal of a few DF leads to high Δ clusters, as is the case of the proline (8) in position P3′, the cysteins, and the hydrogen bonds at the termini. It has been shown experimentally that the proline in P3′ is essential for activity.

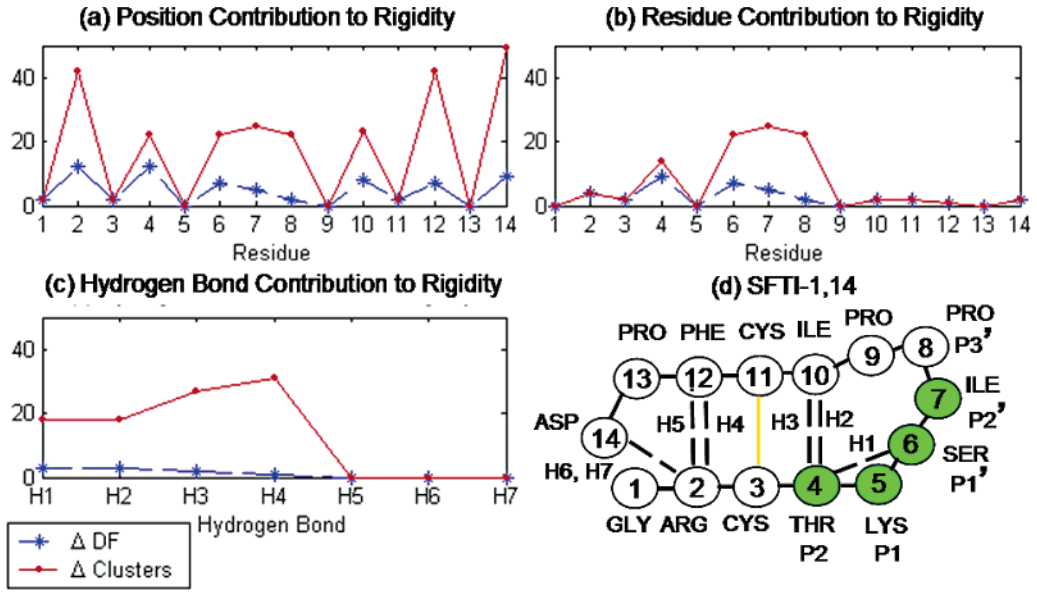

Figure 6. Rigidity study of the peptide SFTI-1,14, the acyclic variant of SFTI-1. The *y* axis in a, b, and c represents the increase in number of rigid clusters (Δ clusters, red circles) and in the number of degrees of freedom (Δ DF, blue asterisks) in comparison with the case where all the constraints are present as in d, where the number of DF is DF = 12 and the number of clusters is clusters = 13. (a) Position contribution to the rigidity. (b) Residue contribution to rigidity (backbone constraints are preserved). (c) Hydrogen bond contribution to rigidity. (d) Topology of the network of hydrogen bonds. Only those just above the cutoff are depicted here, $E_{H1} = -3.2$, $E_{H2} = -5.8$, $E_{H3} = -2.3$, $E_{H4} = -1.3$, $E_{H5} = -0.7$, $E_{H6} = -1.2$, and $E_{H7} = -1.1$ kcal/mol. The backbone is still rigid with the proline in P3′ still playing an important role in rigidity. Removing the cyclic constraint introduces new distance constraints among residues 4−10 which lead to a higher number of hydrophobic contacts and help ensure the rigidity of the structure.

site. Furthermore, they also found that the second proline in position P4′, although not essential, contributes to biological activity. On the other hand, note that the residue responsible for specificity of the inhibitor activity in the P1 position (residue 5) does not contribute to rigidity. This is consistent among all the inhibitors we considered. Additionally, isoleucine (residue 10) occupies a position that contributes to rigidity but in itself does not contribute. This implies that it should be possible to replace it without affecting the biological activity of the inhibitor and has indeed often been done experimentally without a loss of the desired inhibitory properties.[9]

Although backbone cyclization should be an important constraint to achieving rigidity, it is known that SFTI-1,14, an acyclic version of SFTI-1 with a break in the backbone chain

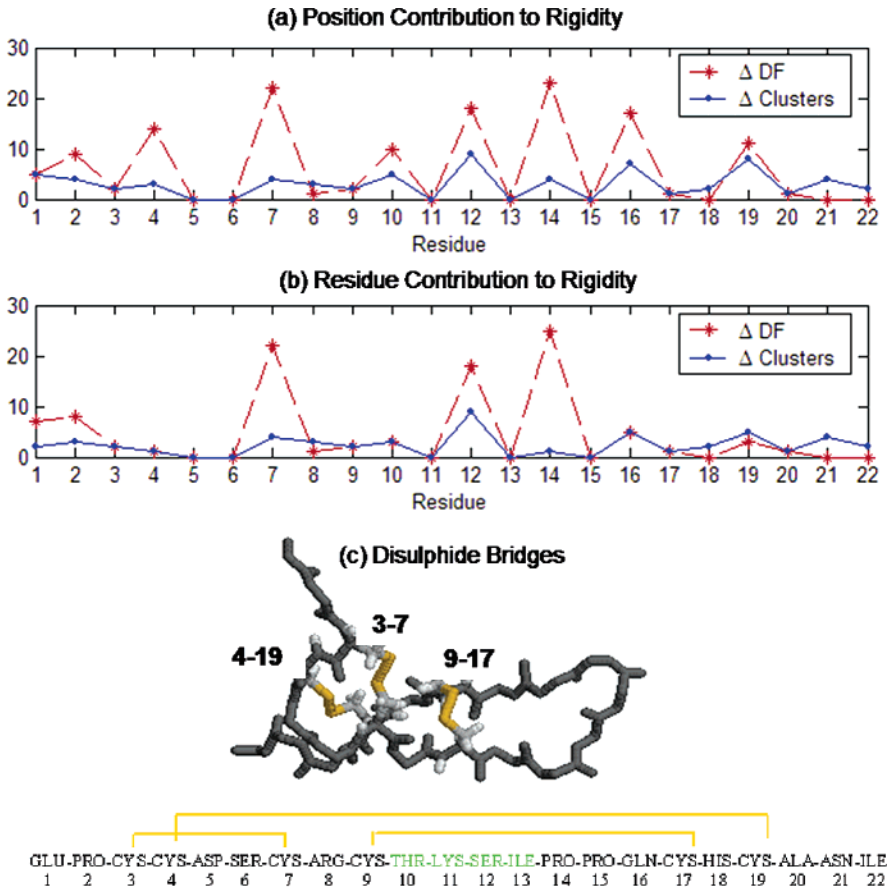

Figure 7. Rigidity study of the synthetic peptide 1G9I. The *y* axis in a and b represents the increase in the number of rigid clusters (Δ clusters, red circles) and in the number of degrees of freedom (Δ DF, blue asterisks) in comparison with the case where all the constraints are present as in c, where DF = 26 and clusters = 42. (c) Sequence and structure of the peptide including the additional disulfide bridges. The cysteines and the proline (14) in position P3′ play an important role in rigidity. There are four hydrogen bonds with energy at the cutoff whose removal would lead to a flexible structure.

between residues 1 and 14, also has a high affinity for trypsin, even though its K_i value is weaker than the cyclic form (see Table 1). We have analyzed this peptide, and the results in Figure 6 show that the peptide is rigid even though the backbone is not closed. In agreement with the results obtained for the SFTI-1, we can see that the residues far from the active region also play an important role (Figure 6a) although the effect of the distant hydrogen bonds is not so critical (Figure 6d). Just as for the case of SFTI-1, the threonine (residue 4) at position P2 and the proline (residues 8) at position P3′ are important for rigidity. However, removing the requirement of the cyclic structure produces changes in the structure that create new distance constraints. In particular, more hydrophobic contacts are formed between residues 4−10, which help explain the retention of rigidity which we would expect to be lost based on the results of the soybean BBI. These contributions also mask the effect of the cysteines. If the cyclic backbone is not vital to rigidity, it may perhaps play a different role in acitvity. Korsinczky et al.[6] have suggested that the cyclic backbone could extend the in vivo lifetime of the peptide by protecting it against degradation by proteases.

IG92 is a larger synthetic trypsin inhibitor with three disulfide bridges added with the goal to keep the conformation of the protein rigid (Figure 7c). The peptide is in fact still rigid (Figure 3d). In this case, the residues that occupy the positions closer to the termini do not have a significant impact on the number of clusters (Figure 7b) which suggests that they are not critical to the rigidity. There does not seem to be an increase in rigidity over the shorter peptides, such as SFTI-1, which have been proven to be better inhibitors. In Figure 7b, one can observe again the importance of two of the cysteines and the proline (residue 14) in position P3′. It is also interesting to note that there is a large number of hydrogen bonds with energies at the threshold of the cutoff, which, if removed, would lead to a more flexible cluster. These observations are consistent with the fact that this is a weaker inhibitor than the previous ones.

The shortest peptide 1GM2 is the most flexible of all the peptides studied (Figure 3d). This is a result of the lack of the extra ring at the terminus that is present in all others and essential to overall rigidity (Figure 8). The importance of the lack of the extra ring is compatible with experimental findings that a minimum of nine residues, larger than the active site, is reported in the literature as a minimum requirement for active inhibition.[5] In principle though, this peptide should be a weak inhibitor due to its flexibility. There are widely varying experimental values of K_i reported,[7] from very weak to reasonably good although not as strong as SFTI-1 and SFTI-1,14. This may be attributed to the fact that its structure, less rigid now, can explore a wider range of conformations. For example, the absence of the hydrogen bond between threonine (3) and serine (4), present in the rest of the inhibitors, leads to a substantial reduction in rigidity. In fact, the existence of inactive conformers of the 1GM2 has been experimentally confirmed with only one of them showing activity[7] where it appears that, in the active conformer, the hydrogen bond between threonine and serine plays a role.

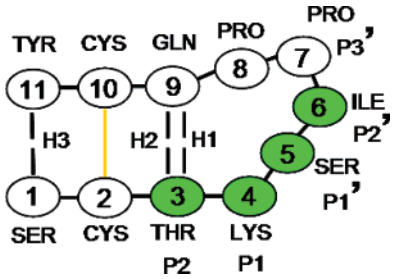

Figure 8. Synthetic peptide 1GM2. The topology of the network of hydrogen bonds ($E_{H1} = -2.3$, $E_{H2} = -4.4$, and $E_{H3} = -1.4$ kcal/mol) is shown as averaged over the NMR structure. The peptide in this configuration appears flexible (Figure 3c) due to the absence of the residues at the termini that provide the necessary ring of hydrogen bonds. Additionally, the hydrogen bond between threonine and serine is below the cutoff. This additional constraint would increase the rigidity of the peptide. In fact, the existence of inactive conformers have been experimentally confirmed with only one of them showing activity. It appears that this bond is present in the active conformer.

IV. Conclusions and Outlook

The presence of the conserved canonical loop in intact conformation in a variety of natural and synthetic potent trypsin inhibitors of the BBI family suggests that maintaining the rigidity of this structure is essential to activity. To test this hypothesis, we have quantified for the first time, to our knowledge, the rigidity properties of representative inhibitors by using FIRST. The results have shown that the canonical loop in the large soybean BBI, which includes the trypsin active region, represents an independent rigid cluster. It correlates well with the size of smaller potent inhibitors and may help explain its success in acting as an independent motif. The rings of constraints in this loop seem to be tuned to provide rigidity. Moreover, we have studied the importance of each residue in providing rigidity to the overall structure and found that the most important residues are also those that are well conserved in this family of peptides and are independent of the active site. These residues which show the maximal contribution to rigidity correlate well with those found experimentally to be essential for good inhibitory activity. These effects are nonlocal (since they are nonlinear) and the global topology of the network must be considered in order to assess the effect of a mutation in disrupting the rings of constraints that maintain rigidity. In contrast, the residue that is mostly responsible for specificity in position P1 plays no role in rigidity. Our conclusions seem to agree with a recent work that shows that enzymatic activity is often a result of an interplay of chemical and mechanical properties of the proteins.[4]

The relative rigidity of the proteins seems to correlate well with their activity as captured by K_i measurements. It would be desirable to convert these rigidity differences directly to relative K_i rates when comparing inhibitors. Since these inhibitors tend to preserve their conformation on binding and most changes occur away from the active site, at first approximation, we can assume that the ratio of the rates would come mainly from the difference in entropy of the inhibitors,[32] which is directly related to their rigidity attributes. Quantifying these entropy changes, which, in general, is a nonadditive property[33] is a future direction of this work. More quantitative free energy differences could then be further investigated.[34]

We must also point out that the selection of the intramolecular interactions that make up the graph whose rigidity is analyzed depends both on the potentials used and the cutoff. More detailed potentials could be used to evaluate energies. Increasing the cutoff energy for hydrogen bonds decreases the number of hydrogen bonds in the system which produces the same overall effect as increasing the temperature. We have tried to minimize against any artifacts by choosing constraints in a conservative way so as to bias against rigidity, for example, by demanding that they be present in a majority of possible structures. However, a better approach may be to obtain an average rigidity that comes from an ensemble of graphs, rather than just the average graph. This would remove some of the drawbacks of the static approach used here.

In conclusion, we have shown through a rigidity analysis that maintaining overall rigidity plays an important role in the activity of BBI protease inhibitors. Through this efficient computational approach, rigidity analysis may be used as an additional criterion in optimizing protein design.

Acknowledgment. We would like to thank D.J. Jacobs, B.M. Hespenheide, and M.F. Thorpe for permission to use the FIRST software. We acknowledge helpful discussions with Martin Hemberg and Mauricio Barahona. J.R.C. is supported by a Marie Curie Fellowship from the European Union. It is a great pleasure to dedicate this work to Bob Silbey, thesis advisor to S.N.Y., on his 65th birthday.

References and Notes

(1) Nicholson, L. K.; Yamazaki, T.; Torchia, D. A.; Grzesiek, S.; Bax, A.; Stahl, S. J.; Kaufman, J. D.; Wingfield, P. T.; Lam, P. Y.; Jadhav, P. K. *Nat. Struct. Biol.* **1995**, *2*, 274−280.
(2) Shewmaker, F.; Maskos, K.; Simmerling, C.; Landry, S. J. *J. Biol. Chem.* **2001**, *276*, 31257−31264.
(3) Bahar, I.; Rader, A. J. *Curr. Opin. Struct. Biol.* **2005**, *15*, 586−592.
(4) Yang, L. W.; Bahar, I. *Structure (Cambridge, MA, U.S.)* **2005**, *13*, 893−904.
(5) McBride, J. D.; Watson, E. M.; Brauer, A. B.; Jaulent, A. M.; Leatherbarrow, R. J. *Biopolymers* **2002**, *66*, 79−92.
(6) Korsinczky, M. L.; Schirra, H. J.; Rosengren, K. J.; West, J.; Condie, B. A.; Otvos, L.; Anderson, M. A.; Craik, D. J. *J. Mol. Biol.* **2001**, *311*, 579−591.
(7) Brauer, A. B.; Kelly, G.; Matthews, S. J.; Leatherbarrow, R. J. *J. Biomol. Struct. Dyn.* **2002**, *20*, 59−70.
(8) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157−162.
(9) Brauer, A. B.; Domingo, G. J.; Cooke, R. M.; Matthews, S. J.; Leatherbarrow, R. J. *Biochemistry* **2002**, *41*, 10608−10615.
(10) Jacobs, D. J.; Rader, A. J.; Kuhn, L. A.; Thorpe, M. F. *Proteins* **2001**, *44*, 150−165.
(11) Jacobs, D. J. *J. Phys. A: Math. Gen.* **1998**, *31*, 6653−6668.
(12) Jacobs, D. J.; Thorpe, M. F. *Phys. Rev. Lett.* **1995**, *75*, 4051−4054.
(13) Hemberg, M.; Yaliraki, S. N.; Barahona, M. *Biophys. J.* **2006**, *90*, 3029−3042.
(14) Hespenheide, B. M.; Rader, A. J.; Thorpe, M. F.; Kuhn, L. A. *J. Mol. Graph. Model.* **2002**, *21*, 195−207.
(15) Rader, A. J.; Hespenheide, B. M.; Kuhn, L. A.; Thorpe, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3540−3545.
(16) Mamonova, T.; Hespenheide, B.; Straub, R.; Thorpe, M. F.; Kurnikova, M. *Phys. Biol.* **2005**, *2*, S137−47.
(17) Gohlke, H.; Kuhn, L. A.; Case, D. A. *Proteins* **2004**, *56*, 322−337.
(18) Levitt, M.; Sander, C.; Stern, P. S. *J. Mol. Biol.* **1985**, *181*, 423−447.
(19) Brooks, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6571−6575.
(20) Tirion, M. M. *Phys. Rev. Lett.* **1996**, *77*, 1905−1908.
(21) Tama, F.; Wriggers, W.; Brooks, C. L., 3rd *J. Mol. Biol.* **2002**, *321*, 297−305.
(22) Ming, D.; Kong, Y.; Wakil, S. J.; Brink, J.; Ma, J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7895−7899.
(23) Haliloglu, T.; Bahar, I.; Erman, B. *Phys. Rev. Lett.* **1997**, *79*, 3090.
(24) Tama, F.; Gadea, F. X.; Marques, O.; Sanejouand, Y. H. *Proteins* **2000**, *41*, 1−7.
(25) Suhre, K.; Sanejouand, Y. H. *Nucleic Acids Res.* **2004**, *32*, W610−4.
(26) Teodoro, M. L.; Phillips, G. N., Jr; Kavraki, L. E. *J. Comput. Biol.* **2003**, *10*, 617−634.

(27) Hedstrom, L. *Chem. Rev.* **2002**, *102*, 4501−4524.

(28) Vriend, G. *J. Mol. Graph.* **1990**, *8*, 52−6, 29.

(29) Thorpe, M. F., Duxbury, P. M., Eds. In *Rigidity Theory and Applications*; Springer: New York, 1999; p 428.

(30) McBride, J. D.; Brauer, A. B.; Nievo, M.; Leatherbarrow, R. J. *J. Mol. Biol.* **1998**, *282*, 447−458.

(31) McBride, J. D.; Freeman, N.; Domingo, G. J.; Leatherbarrow, R. J. *J. Mol. Biol.* **1996**, *259*, 819−827.

(32) Janin, J. *Proteins* **1997**, *28*, 153−161.

(33) Jacobs, D. J.; Dallakyan, S.; Wood, G. G.; Heckathorne, A. *Phys. Rev. E* **2003**, *68*, 061109.

(34) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. *Science* **1989**, *244*, 1069−1072.

(35) Voss, R. H.; Ermler, U.; Essen, L. O.; Wenzl, G.; Kim, Y. M.; Flecker, P. *Eur. J. Biochem.* **1996**, *242*, 122−131.

(36) Zhu, G.; Huang, Q.; Zhu, Y.; Li, Y.; Chi, C.; Tang, Y. *Biochim. Biophys. Acta* **2001**, *1546*, 98−106.

(37) Sayle, R. A.; Milner-White, E. J. *Trends Biochem. Sci.* **1995**, *20*, 374.