Same Fold with Different Mobility: Backbone Dynamics of Small Protease Inhibitors from the Desert Locust, *Schistocerca gregaria*†

Borbála Szenthe,‡ Zoltán Gáspári,§ Attila Nagy,‖ András Perczel,§ and László Gráf*,‡,⊥

Department of Biochemistry, Eötvös Loránd University, 1117 Pázmány Street 1/C, Budapest, Hungary, Department of Organic Chemistry, Eötvös Loránd University, 1117 Pázmány Street 1/A, Budapest, Hungary, Department of Biophysics, University of Pécs, 7624 Szigeti út 12, Pécs, Hungary, and Biotechnology Research Group of the Hungarian Academy of Sciences, 1117 Pázmány Street 1/C, Budapest, Hungary

Received September 18, 2003; Revised Manuscript Received November 25, 2003

ABSTRACT: SGCI (*Schistocerca gregaria* chymotrypsin inhibitor) and SGTI (*Sch. gregaria* trypsin inhibitor) are small, 35-residue serine protease inhibitors with intriguing taxon specificity: SGTI is specific for arthropod proteases while SGCI is an excellent inhibitor on both mammalian and arthropodal enzymes. Here we report the cloning, expression, and ^{15}N backbone dynamics investigations of these peptides. Successful expression could be achieved by a "dimeric" construct similar to the natural precursor of the inhibitors. An engineered methionine residue between the two modules served as a unique cyanogen bromide cleavage site to cleave the precursor and physically separate SGCI and SGTI. The overall correlation time of the precursor (5.29 ns) as well as the resulted SGCI (3.14 ns) and SGTI (2.96 ns) are as expected for proteins of this size. General order parameters (S^2) for the inhibitors are lower than those characteristic of well-folded proteins. Values in the binding loop region are even lower. Interestingly, the distribution of residues for which a chemical exchange (R_{ex}) term should be considered is strikingly different in SGCI and SGTI. Together with H−D exchange studies, this indicates that the internal dynamics of the two closely related molecules differ. We suggest that the dynamic properties of these inhibitors is one of the factors that determine their specificity.

The diverse biological roles of various serine proteases in nature are controlled by protease inhibitors of comparable, if not greater, diversity (*1*). Serpins, hirudin-like and canonical inhibitors, act at different steps of the process leading to peptide bond hydrolysis. To date, at least 18 protein families with different folds are known to belong to canonical, standard mechanism protease inhibitors (*2*). The mechanisms of canonical, standard mechanism serine protease inhibitors were studied by numerous methods, including structure determination giving rise to the widely accepted view that the protease binding loop, a common structural element of the inhibitors of otherwise unrelated fold, is the major determinant of inhibition (*2*). With the increase of the number of NMR structures of protein protease inhibitors, the dynamic properties of these molecules are also getting more and more attention. For several inhibitors, it was shown that both the conformation and dynamics of the protease binding loop are of utmost importance. Not only is the binding loop less well defined than the other parts−the structural core−of these proteins (*3*, *4*), but dynamical studies both by NMR and computational methods show that this region is indeed flexible (*5*−*7*). For CI2 and CMTI-V,[1] it was demonstrated that cleavage of the scissile P1−P1′ bond (nomenclature of ref *8*) does not alter the overall dynamic properties of regions outside the protease binding loop (*9*, *10*). The interactions stabilizing the binding loop were shown to affect its dynamics and the inhibition properties of the protein (*11*).

Members of the pacifastin inhibitor family (*12*) (also known as grasshopper family (*2*)) are small (usually about 35 residues), disulfide-rich canonical serine protease inhibitors isolated from various arthropod species. The precise biological role of these inhibitors is still elusive, despite several studies suggesting involvement in insect immunity,

† Grants form the Hungarian Scientific Research Fund (OTKA FO32772 and T032486) are acknowledged. The work was also supported by a grant of the National Széchenyi Program (NKFP 1/010/2001) and The Wellcome Trust (063822/Z/01/Z).

* Corresponding author. E-mail: graf@ludens.elte.hu. Tel: (36 1) 2090555, ext. 8781.

‡ Department of Biochemistry, Eötvös Loránd University.

§ Department of Organic Chemistry, Eötvös Loránd University.

‖ University of Pécs.

⊥ Biotechnology Research Group of the Hungarian Academy of Sciences.

[1] Abbreviations: CNBr, cyanogen bromide; CI2, chymotrypsin inhibitor 2; CMTI-V, Cucurbita maxima trypsin inhibitor V; IPTG, isopropyl-β-dithio galactopyranoside; K_I, equilibrium inhibitory constant; NMR, nuclear magnetic resonance spectroscopy; MUGB, methyl umbelliferil guanidino benzoate; MUB, methyl umbelliferone; MUTMAC, 4-methylumbelliferyl-*p*-trimethylammonium cinnamate chloride; PMP-C, pars interecerebralis major peptide C; PMP-D2, pars intercerebralis major peptide D2; R1, spin−lattice relaxation rate; R2, spin−spin relaxation rate; S^2, general order parameter; SD, standard deviation; SGCI, *Schistocerca gregaria* chymotrypsin inhibitor; SGTCI, *Schistocerca gregaria* trypsin and chymotrypsin inhibitor; SGTI, *Schistocerca gregaria* trypsin inhibitor; SGTMCI *Schistocerca gregaria* trypsin and chymotrypsin inhibitors linked together with a single Met; SGTI*, *Schistocerca gregaria* trypsin inhibitor with an extra homoserine residue at the C-terminus; T1, spin−lattice relaxation time; T2, spin−spin relaxation time; TFA, trifluoroacetic acid; τ_e, effective correlation time; τ_c; overall molecular correlation time.

10.1021/bi035689+ CCC: $27.50 © 2004 American Chemical Society
Published on Web 03/06/2004

```
SGCI     EVTCEPGTTFKDKCNTCRCGSDGKSAACTLKACPQ
PMP-C    EISCEPGKTFKDKCNTCRCGADGKSAACTLKACPNQ
SGTI     EQECTPGQTKKQDCNTCNCTPTG-VWACTRKGCPPH
PMP-D2   EEKCTPGQVKQQDCNTCTCTPTG-VWGCTRKGCQPA
         1   5    10   15   20   25   30   35
```

FIGURE 1: Sequence alignment of members of the pacifastin inhibitor family with known structure. Disulfide pairing is indicated.

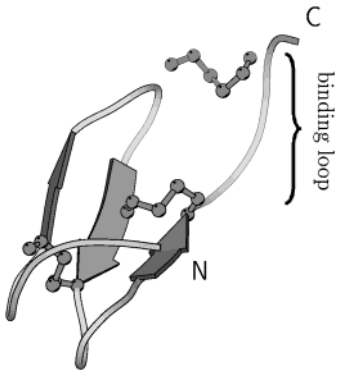

FIGURE 2: Schematic structure characteristic of the pacifastin inhibitor family. The three antiparallel β-strands are connected by three disulfide bridges (the Cα, Cβ and Sγ atoms of which are shown); the protease binding loop is located near the C terminus. Figure prepared with Molscript (*37*) from the structure of SGCI (*19*).

development, and, most recently, solitary−gregarious shift in locust species (*13*). The best characterized peptides are from locusts: PMP-C and PMP-D2 (pars intercerebralis major peptide C and D2, respectively) from *Locusta migratoria* (*14*), and SGCI (*Schistocerca gregaria* chymotrypsin inhibitor) and SGTI (*Sch. gregaria* trypsin inhibitor) from *Schistocerca gregaria* (*15*). Sequence alignment of these inhibitors is shown in Figure 1.

Interestingly, SGTI is a taxon specific inhibitor acting orders of magnitudes more effectively on arthropodal enzymes than on mammalian ones commonly used in inhibition assays, while SGCI is an excellent inhibitor on all enzymes tested (*16*).

The solution structure of PMP-D2 and PMP-C (*17*, *18*), SGTI, and SGCI (*19*) as well as the crystal structure of the former two in complex with chymotrypsin was determined earlier (*20*). The fold characteristic of the family consists of three antiparalel β-strands with three disulfide bridges of the pattern *abcacb* (Figure 2). The protease binding loop is located near the C-terminus between two disulfides. Despite their highly similar sequence and practically identical fold, one-dimensional H−D exchange experiments suggested that the dynamical behavior of SGCI and SGTI is quite different: in contrast to SGCI, the backbone amide protons in SGTI are resistant to exchange over the surprisingly long period of a week (*16*, *19*).

In this paper, we report the cloning, expression, ^{15}N-labeling, amide nitrogen chemical shift assignment of SGCI and SGTI, as well as the examination of their backbone dynamics by ^{15}N-relaxation methods. The dynamic properties of the two inhibitors and also their common precursor molecule SGTCI are compared with each other.

MATERIALS AND METHODS

Cloning of the SGTCI mRNA. Total RNA was isolated from the fat body of the locust with RNAeasy Kit (Fermentas). cDNA was prepared using first strand cDNA Kit (Fermentas) with oligo(dT) primer. For the amplification of the gene (*13*) encoding both SGTI and SGCI (SGTCI) designed oligonucleotides (Integrated DNA Technologies INC) were used (5′ primer, 5′ TAT ATA CAT ATG GAA CAG GAA TGT ACA CCT GGC 3′, containing an NdeI site; 3′ primer, 5′ GAA TTC TCA GTG GGG GCA GGC CTT GAG 3′, with an EcoRI site) (see Figure 3). The PCR product was ligated into pET17b expression vector (Novagen) using the restriction sites and sequenced by automated dideoxy sequencing (ABI Prism) with Big Dye Terminator Kit. The natural Lys−Arg linker dipeptide connecting SGTI to SGCI in the precursor (SGTCI) was changed to a single Met (see Figure 3) by megaprimer PCR mutagenesis using megaprimer 5′ TGC CCG CCC CAT ATG GAG GTC ACC TGC GAG 3′. This primer contains a new NdeI cleavage site that facilitated screening for the mutation. The new construct is designated SGTMCI.

Protein Expression and Isolation. For cytoplasmic expression BL21 DE3 pLysS cells (Stratagene) were used. Uniform ^{15}N-labeling was achieved by growing the cells on minimal media containing 0.6% Na_2HPO_4, 0.3% KH_2PO_4, 0.05% NaCl (Sigma), 0.4% sucrose, and 0.1% $^{15}NH_4Cl$, (obtained from Cambridge Isotope Labs). Cells were grown at 37 °C to $OD_{600} = 0.8$ and were then induced with 100 μg/mL final concentration of IPTG(Sigma) for 4 h. After harvesting the cells by centrifuging with 10 000 rpm for 30 min at 4 °C (Beckman J2-MC) the pellet was resuspended in (50mMTris, 20mM EDTA, 0.1% Triton-X100 pH 8.0 (Sigma)) buffer and kept at −20 °C overnight. Cells were lysed by thawing and sonicating them (with VWR Branson Sonifier 250). After centrifuging the cell lysate with 15 000 rpm for 20 min at 4 °C, expressed SGTCI/SGTMCI appeared soluble in the supernatant (Figure 4). The supernatant was dialyzed against a buffer of pH 7.6 containing 15 mM Tris-HCl and 15 mM NaCl. Subsequently, the precursor protein was purified by affinity chromatography using α-bovine chymotrypsin (Sigma) immobilized on a CNBr activated Sepharose-4B column (Sigma). SGTCI/SGTMCI was eluted with 20 mM HCl from the column after washing with distilled water (Figure 4). SGTCI/ SGTMCI fractions were freeze-dried. For CNBr cleavage SGTMCI was redissolved in a deoxygenized solution of 70% HCOOH (Sigma) and 10% CNBr (Sigma) and left in the dark overnight at room temperature. The cleavage reaction was stopped with distilled water (90% final concentration) and the sample was freeze-dried. The resulting SGTI* (SGTI containing an extra homoserine residue on its C-terminus) and SGCI were separated by RP-HPLC (HP Aligent Series 1100 HPLC) on a Phenomenex 10.0, C-18 column with a linear gradient from 10% to 14% of solvent B was used within 20 min. The flow rate was 3 mL/min. Solvent A was Mili-Q water, containing 0.1% (v/v) TFA (Sigma), and solvent B was 80% (v/v) acetonitrile (Sigma), 0.1% (v/v) TFA (Figure 4). Fractions for NMR studies were freeze-dried and redissolved in 0.5 mL distilled water. The amount of SGTCI/ SGTMCI was determined by absorbance at 280 nm with $\epsilon_{SGTCI/SGTMCI} = 5690$.

Kinetics. Trypsin and chymotrypsin inhibition of SGTCI was measured using two different approaches. For the determination of high (from 10^{-5} to 10^{-9} M) K_I values, we made timecourse analysis using DYNAFIT software (*21*) for data processing. These measurements were performed using

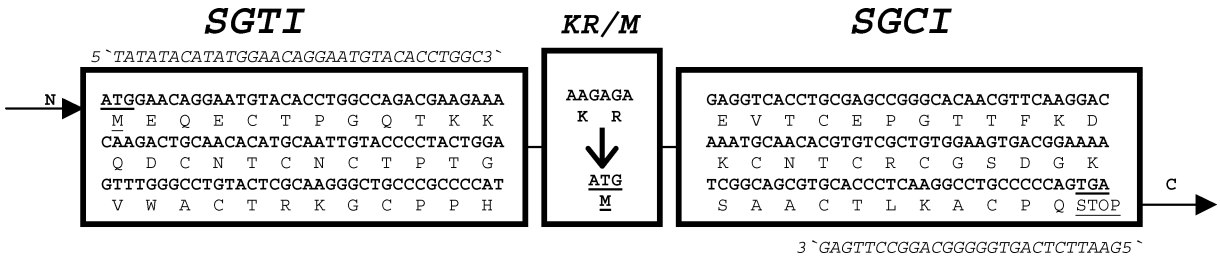

FIGURE 3: DNA and amino acid sequence of the cloned precursor for SGTI, SGCI, and the linker dipeptide and its modification (SGTCI/SGTMCI); modified regions are underlined, primers used for cloning shown in italics.

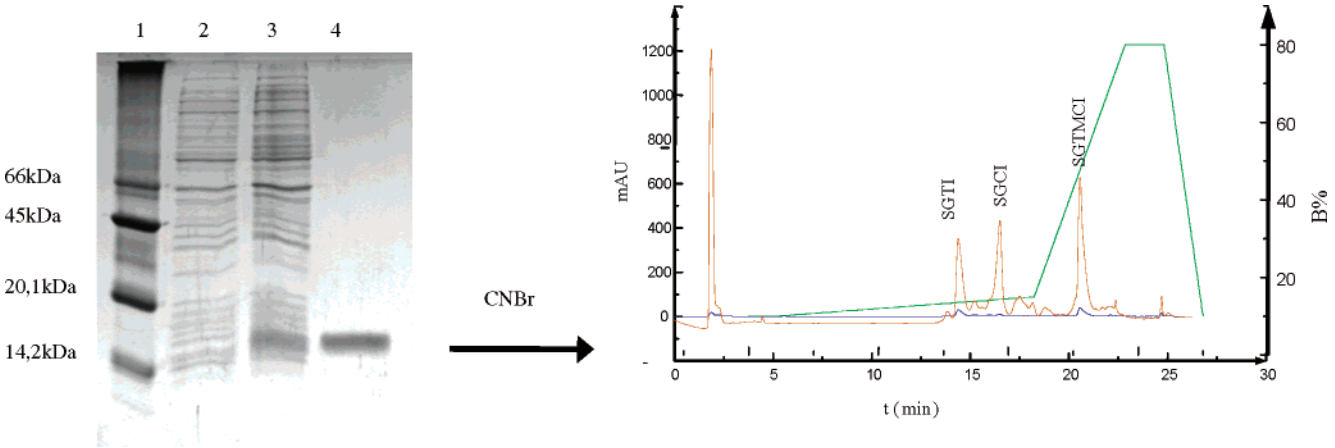

FIGURE 4: Expression, isolation, cleavage, and purification of the two inhibitors. On the SDS−PAGE lane 1 is MWL-VL; lane 2, *E. coli* cell lysate before induction; lane 3, after induction; lane 4, eluted fraction of SGTMCI from the inmobilized α-bovine-chymotrypsin with 20mM HCl. On the RP-HPLC chromatogram, the cleaved inhibitors and the remained precursor are separated from each other, and detection at 280 nm is shown in red, at 220 nm in blue, and the gradient in green.

a Shimadzu spectrophotometer at 405 nm on CBZ-Gly-Pro-Arg-pNA (Sigma) substrate (50 μM final concentration) for bovine trypsin with a final concentration of 10 nM and 0.9 nM final concentration of SGTCI at 37 °C. Low K_I values (10^{-10}−10^{-12} M) were determined by preincubating 10^{-10} M enzymes (α-bovine chymotrypsin (Sigma) and crayfish trypsin (isolated from the animal)) with SGTCI from 10^{-8} to 10^{-11} M for 6 h. After incubation we measured the remaining activity on a fluorometer (Fluoromax Spex 320) using Succ-Ala-Ala-Pro-Lys/Tyr-AMC substrate (Sigma). The liberation of AMC was followed by emission at 460 nm with excitation at 380 nm. K_I values were determined by fitting $(K_I/E_f + 1)(E_t - E_f) = I_t$ equation (where E_t = total enzyme amount, E_f = free enzyme amount = remaining activity in percentages, I_t = total inhibitor amount) on the measured data with Origin 5.0. The enzyme concentration was determined by active site titration by MUGB and MUTMAC, and MUB for calibration (Sigma). The SGTCI/SGTMCI concentration was determined by titration with chymotrypsin.

NMR Spectroscopy. NMR spectra were recorded on a Bruker DRX 500 spectrometer at 300 K and pH 3. Data processing was done using Bruker's Xwin-NMR and also Felix97 (Molecular Simulations, Inc.) Sample concentration was ∼1.5 mM. To assign the ^{15}N resonances, 3D TOCSY−HSQC and NOESY−HSQC experiments were carried out (*22*, *23*). For *T*1 and *T*2 relaxation rate measurements a series of spectra were collected using parametric relaxation delays of 11.86, 59.3, 213.48 (2×), 355.8, 593, 806.48 (2×), 1186, 2134.8 (2×), and 2965 ms for *T*1 and 8 (2×), 24, 40, 64 (2×), 80, 120, 160, 240, 480, and 640 ms for *T*2, respectively (*24*). The experiments indicated with (2×) were recorded twice in order to estimate peak volume uncertainties. NOE enhancement was detected using the interleaved pulse sequence described in ref *24*.

Peak picking and all subsequent analysis of the spectra were done with SYBYL (Tripos, Inc.) and XEASY. All software were run on an SGI Octane R10000 workstation.

Data Analysis. Relaxation data analysis was carried out by means of the Lipari-Szabo model-free approach (*25*, *26*). For obtaining R1, R2 and NOE values, peak picking from the two-dimensional spectra and integration was performed using SYBYL (TRIAD). Relaxation curves were fitted with the Levenberg−Marquardt algorithm using the program Curvefit (Version 1.23 for SGI, A.G Palmer, http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/curvefit.html). Peak volume uncertainties were estimated from duplicated measurements (for T1 and T2) and from the noise level of the spectra (NOE measurement). Overall correlation time and internal mobility values were calculated with the program Tensor V2.0 (*27*).

RESULTS

Cloning, Expression, and Isolation. Since our preliminary attempts to separately express, in monomeric forms, SGTI and SGCI failed, we decided to clone the cDNA encoding both SGTI and SGCI as a precursor (SGTCI). Primers for 5′ and 3′ end were designed on the basis of the known cDNA sequence of the precursor protein (*13*). The 5′ primer used contains an NdeI restriction site for insertion and also a codon for an extra Met as a translational start. The 3′ primer contains a TGA stop codon and an EcoRI restriction site (Figure 3). For the sake of convenient separation of the two

Table 1: Comparison of the K_I Values of Synthetic SGTI and SGCI[a] and the Newly Expressed Precursor SGTCI[b]

enzyme	K_I (M)		
	SGTI	SGCI	SGTCI
bovine trypsin	2×10^{-7}	2×10^{-5}	2×10^{-7}
bovine chymotrypsin	2×10^{-6}	6×10^{-12}	5.5×10^{-12}
crayfish trypsin	1.3×10^{-12}	7.4×10^{-7}	4.3×10^{-12}

[a] Data from ref *15*. [b] SGTMCI inhibited α-bovine chymotrypsin with about the same K_I value as SGTCI did.

inhibitors, the Lys−Arg dipeptide linker via site-directed mutagenesis was replaced by Met, providing the only CNBr cleavage site of the precursor. The SGTCI and SGTMCI precursor proteins were successfully expressed in soluble and active forms in the cytoplasm of the *Escherichia coli* cells. This is an unexpected result as the used expression system, due to overexpression, usually provides the recombinant protein in insoluble and denatured form (inclusion bodies). We believe that the relatively small size (73 residues, ∼8 kDa) of the SGTCI/SGTMCI precursor together with its "easy to fold" (6 disulfide bridges) structure could explain this result. As expressed in soluble from, isolation and purification of SGTCI/SGTMCI could be readily achieved with affinity chromatography using a Sepharose 4B column containing immobilized α-bovine-chymotrypsin (Figure 4). Cleavage of the SGTMCI form with CNBr at the sole Met and separating with RP-HPLC made isolation of the two inhibitors, SGTI* and SGCI possible (Figure 4). Thus, we achieved a yield of 25 mg pure protein and 5−5 mg separated SGCI and SGTI*/l bacterial culture suitable for NMR studies.

Kinetic Measurements. The determined K_I values of the precursor protein SGTCI/SGTMCI are comparable with those previously published (*15*) (Table 1). This confirmed that both inhibitory active parts of the precursor molecule are expressed in a native form.

Resonance Assignment. The proton chemical shift values of SGTI and SGCI have already been described (*19*, BMRB entries 5274 and 5272, respectively). The assignment of ^{15}N resonances using both the previous data and the information from the 3D-TOCSY and 3D-NOESY spectra was straightforward. Assignments for all residues could be obtained except Glu1 and His35 in SGTI*, Gln35 in SGCI, Glu1 and Gln72 in SGTCI as well as Glu1, Met36, and Cys40 in SGTMCI.

The ^{1}H chemical shift values of the dimeric precursor molecules do not differ significantly from the previously determined values of the individual SGTI and SGCI. Notably, the amide nitrogen and proton chemical shifts belonging to the majority of conserved residues in structurally identical positions in SGTI and SGCI are remarkably similar. Most typical examples are those in the structural vicinity of the binding loop, Cys14, Cys17, Cys28 (Cys27), Thr29 (Thr28), and Cys33 (Cys32) (throughout the text, the residue number for SGTI is given in parentheses when it differs from the corresponding one in SGCI). In contrast, some conserved residues are in a different chemical environment reflected by their divergent chemical shift values, e.g., Cys4, Gly7, Cys19, and Gly23. This can possibly be explained by the structural vicinity of the aromatic Trp side chain in SGTI.

Relaxation Data Analysis. R1, R2 and NOE enhancement values for SGCI and SGTI are shown in Figure 5. It is apparent that the $R2$ value distributions of the two molecules differ, as there are much higher $R2$ values for some residues in SGCI than in SGTI. NOE enhancements are generally small (|NOE| < 0.5), and a negative value for residue 2 (Val2 in SGCI, Gln2 in SGTI) could be observed.

For SGCI, preliminary calculations showed that exclusion of residues 4, 19, 20, 26, 28, 29, and 30 from calculation of the overall correlation time was necessary, since all the above residues require inclusion of a significant (>1 Hz) R_{ex} term. However, only four of these (19, 20, 26, and 28) have $T1/T2$ values greater than the mean value + 1SD, (see ref *28*). The calculated τ_c is 3.14 ns.

In SGTI, residues 13, 14, and 32 have $T1/T2$ values above the threshold value of mean + 1 SD. However, final parameter fitting showed that inclusion of R_{ex} term was only required for Cys14 and its disulfide-bonded pair Cys32. The calculated overall correlation time for SGTI is 2.96 ns. Fitted S^2, τ_ϵ, and R_{ex} values are shown in Figure 6.

In the precursor molecule SGTCI (Figure 7), order parameters are somewhat higher than in the individual inhibitors ($\tau_c = 5.93$ ns). Differences regarding the residues with motions on the microsecond−millisecond time scale are also observed; however, the similarity to the dynamics of the inhibitors is still apparent. Order parameters for residues in the linker region are low ($S^2 < 0.5$), comparable to that of the N-terminal residues. Results for SGTMCI are similar (data not shown).

DISCUSSION

Dynamics on the Picosecond−Nanosecond Time Scale. Despite of their well-defined solution structure, S^2 values for both SGCI and SGTI are generally low ($S^2 < 0.8$), resembling those of flexible peptides rather than well-folded globular proteins (see, e.g., refs *29* and *30*) and in contrast to the values reported for other canonical inhibitors investigated (e.g., refs *5* and *31*). Most similar but still slightly higher order parameters were observed for the 29-residue CMTI−III of the squash family (*6*). The order parameters of SGTI and SGCI are comparable, in SGCI the average being lower (Table 2, Figure 6). It should be noted here that none of the observed intra- or intermolecular differences can be regarded as statistically significant. Residues in the protease binding loop of SGTI (0.57 ± 0.07) have lower order parameters than the corresponding ones in SGCI even when compared to the rest of the molecule. Interestingly, on the basis of S^2 values, the loop regions in both molecules are slightly more constrained than the β-strands. This is in contrast with the calculated RMSD values of the 10 best conformers of the inhibitors, as these values are lowest for the β-strands (*19*). The subtle difference between the overall S^2 values of SGTI* and SGCI does not reflect the difference in their calculated RMSD values either (*19*); thus, in the case of SGTI*, the lower RMSD cannot be regarded indicative for a less constrained structure.

The order parameter for Trp25 Nϵ1 in SGTI*, the only side chain nitrogen atom for which relaxation parameters could be reliably derived, is 0.54 ± 0.04. This is consistent with the NOE-derived distance restraints between Trp25 and Lys10 indicating conformational heterogeneity of these side chains (*19*).

Motions on the Microsecond−Millisecond Time Scale. The number and distribution of residues with considerable

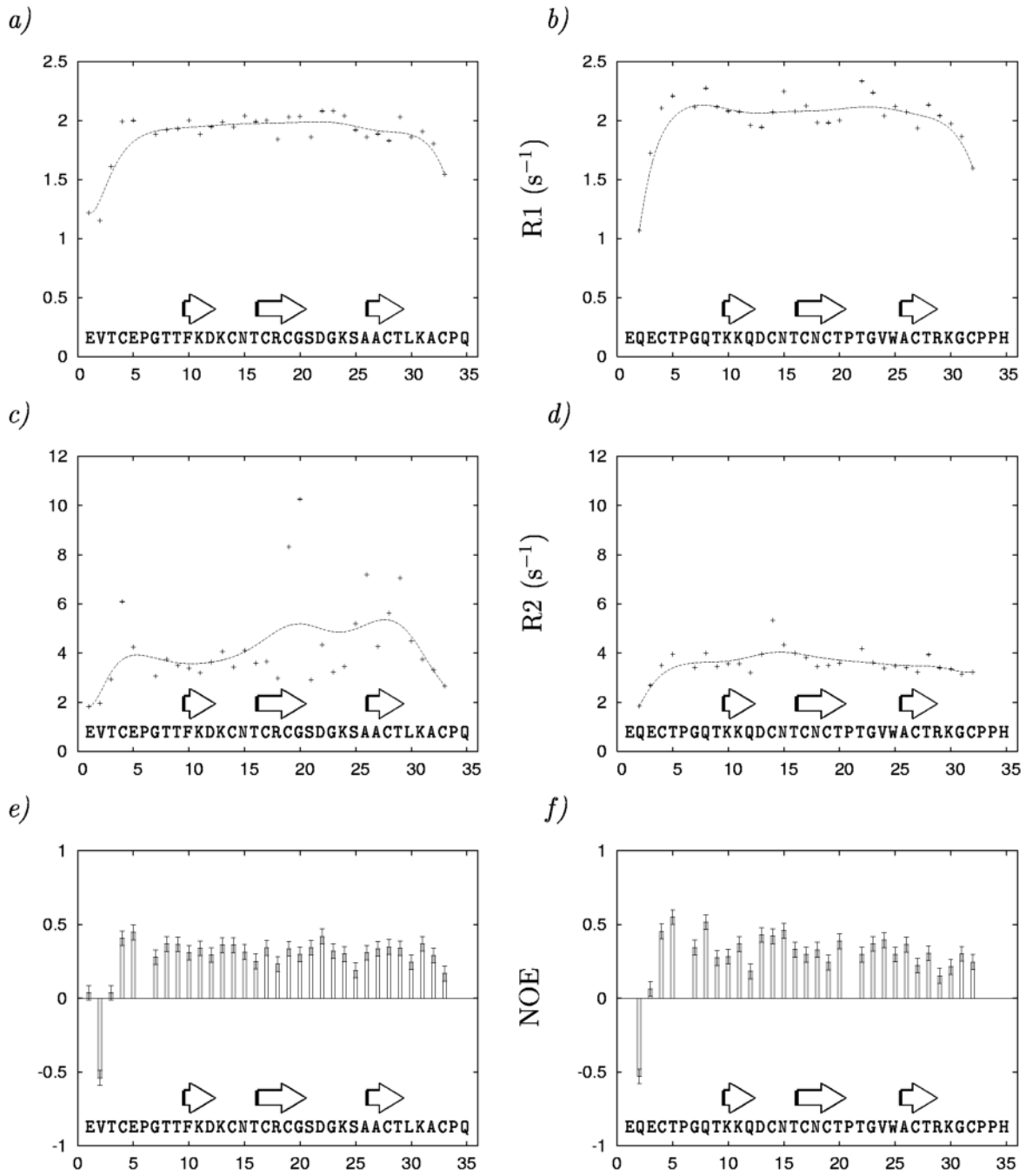

FIGURE 5: NMR relaxation parameters of SGCI (a, c, e) and SGTI (b, d, f) as a function of the amino acid sequence: R1 (a, b), R2 (c, d), and NOE (e, f) values. The dotted line in a−d is a bezier curve smoothed to the data values only to show the tendencies in the variation of the parameters along the sequences. The position of the β-strands is indicated by arrows.

chemical exchange ($T1/T2$ > mean + 1SD) term is strikingly different in the two inhibitors. In SGCI, these residues form two groups: Cys19, Gly20, and Ala26 form a small "cluster" in the loop connecting the second and third β-strands, whereas Thr29 is located at the C-terminus of the third β-strand and is part of the protease binding loop (the P2 residue). For the fitting of motional parameters, inclusion of an R_{ex} term for residues Cys4, Ala27 and Leu30 was also required, extending both groups. The presence of the former group of residues 4, 19, 20, 26, and 27 indicates that there are motions on the μs-ms time scale in the loop connecting β-strands 2 and 3 and also linked to the N-terminal part by the Cys4−Cys19 disulfide bridge. The other group consisting of the P2 and P1 residues Thr29 and Leu30, respectively, is located C-terminally to Cys28.

In SGTI, there are only three residues with possible motions on the microsecond−millisecond time scale, Asp13, Cys14, and its disulfide pair, Cys32. However, final parameter fitting required an explicit R_{ex} term for Cys14 and Cys32 only (for Asp13, $T1/T2$ is only slightly above the average). In contrast to Thr29 in SGCI, Cys32 is located at the C-terminus of the binding loop (position P3′). Moreover, the R_{ex} contribution to $T2$ is remarkably smaller for all three residues in SGTI than for the others in SGCI.

In the heteronuclear 3D spectra an extra group of peaks is observed that could possibly be assigned to an alternative

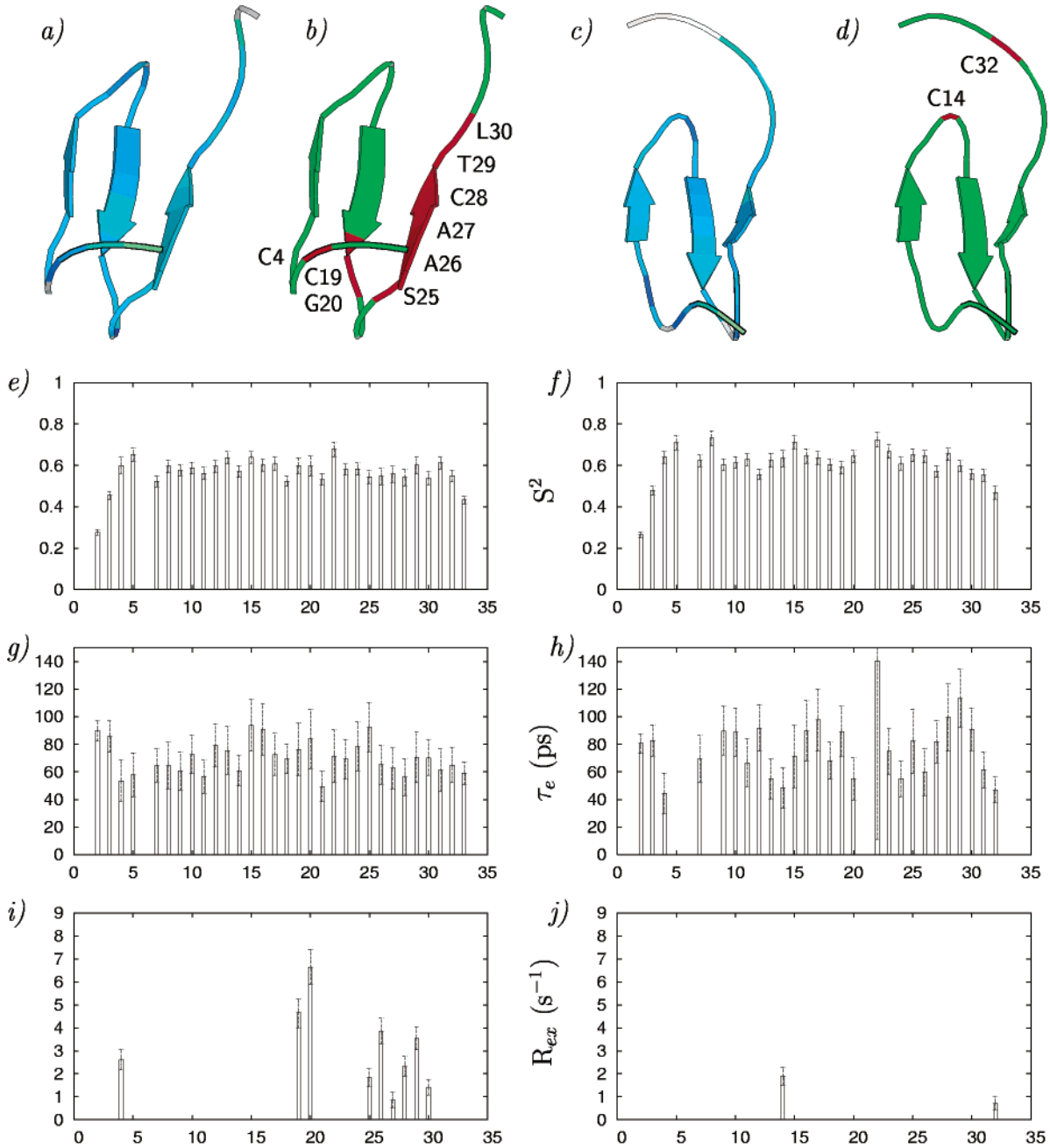

FIGURE 6: Dynamical parameters of SGCI (a, b, e, g, i) and SGTI (c, d, f, h, j): cartoon representation of the backbone folds colored according to the S^2 values (a and c, darker blue means greater S^2) and R_{ex} contribution (b and d, residues with considerable R_{ex} are shown in red and are labeled; figures prepared with Molscript (*37*)), S^2 (e, f), τ_e (g, h), and R_{ex} values (i, j).

conformer of His35. This could arise from cis−trans isomerization of the preceding Pro33−Pro34 amide bond. This could be further supported by slow conformational motions of Cys32.

Dynamics from Microseconds to Minutes. Proton−deuterium exchange experiments revealed striking differences between SGCI and SGTI (*16*, *19*). In the latter peptide, the exchange of amide protons is orders of magnitudes slower than in SGCI, especially in the loop connecting the second and third β-strand (residues 19, 20, and 23−27/23−26; Figure 8). In SGTI, there are amide protons almost unaffected by exchange even after 180 h of incubation in D_2O, whereas in SGCI, although the corresponding residues also show relatively slow exchange, all protons exchange within this time. This could indicate that the dynamics of the two peptides differ in the time scale from seconds to minutes.

H−D exchange and ^{15}N-dynamics give information about the dynamical properties of the molecules on different time scales. However, care should be taken when interpreting H−D exchange data, as the H−D exchange rate at a particular site depends not only on dynamical but also structural properties of the protein (*32*).

In our previous paper we argued that the H−D exchange observations indicate a more rigid structure for SGTI than SGCI (*19*). In light of the present study, this view needs to be refined in the sense that SGTI may be somewhat less flexible on the picosecond−nanosecond time scale but displays no extensive movements on the μs-ms scale characteristic of SGCI. Assuming that motions of the latter type may facilitate H−D exchange, the two results can be considered to be in agreement in that sense. It should be noted, however, that the outstanding resistance of SGTI to

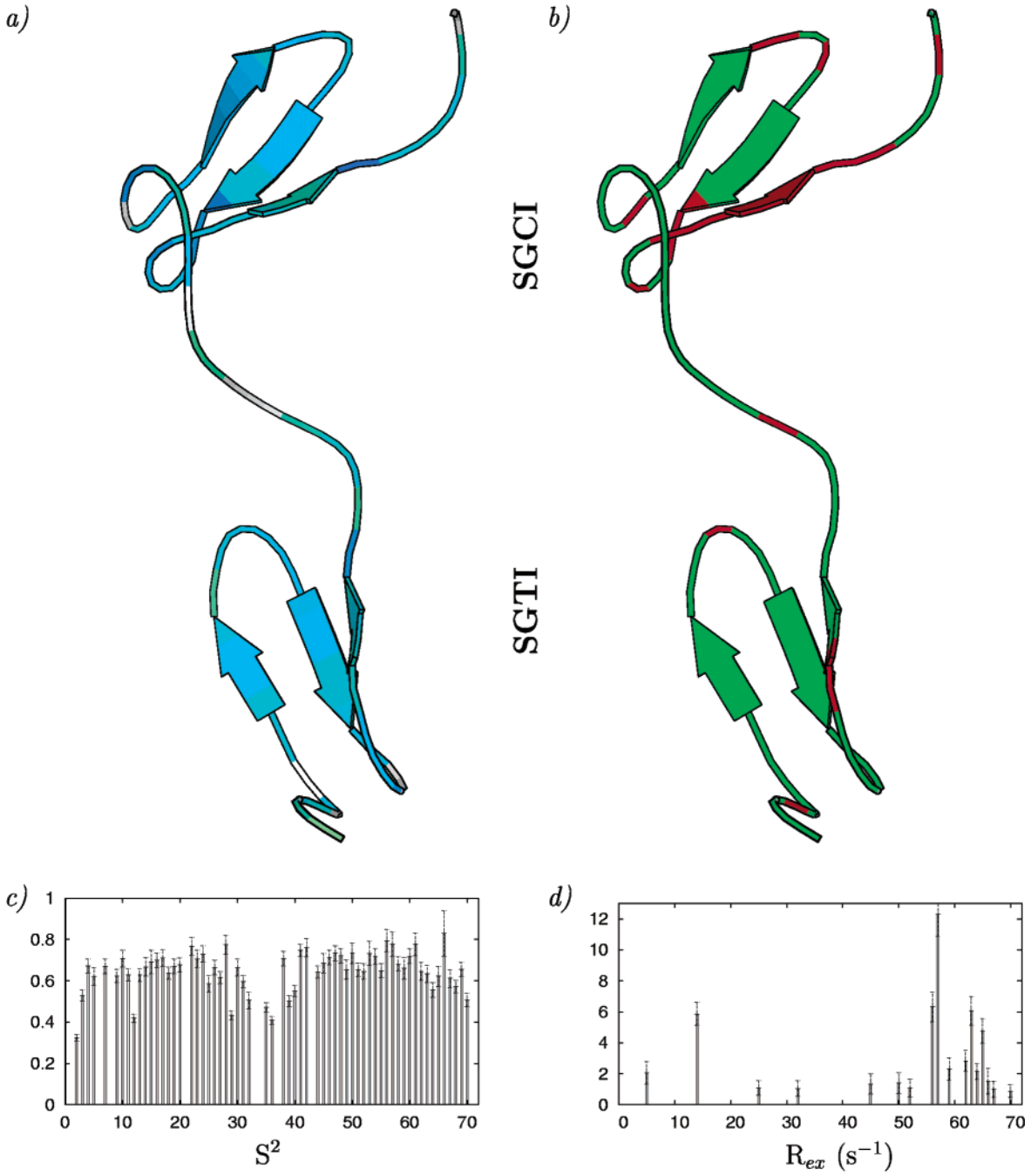

FIGURE 7: Dynamical parameters of the dimeric precursor SGTCI: cartoon representation of the backbone fold colored according to the S^2 values (a, darker blue means greater S^2) and R_{ex} contribution (b, residues with considerable R_{ex} are shown in red; figures prepared with Molscript (*37*)), S^2 (c), and R_{ex} values (d). The presented structure of SGTCI is a model based on the experimentally determined atomic coordinates of SGCI and SGTI.

Table 2: Order Parameters (S^2) for SGTI and SGCI

region	SGCI	SGTI
all residues	0.56 ± 0.07	0.61 ± 0.09
β-strands (9−11,16−19, 26−28/25−27)	0.57 ± 0.03	0.62 ± 0.03
binding loop (28−33/27−32)	0.55 ± 0.03	0.57 ± 0.07
12−15 loop (λ1)	0.61 ± 0.03	0.63 ± 0.06
20−25 loop (λ2)	0.59 ± 0.05	0.66 ± 0.05
N-terminal part (3−8)	0.57 ± 0.08	0.64 ± 0.10

amide proton exchange cannot be explained by our NMR dynamical data and thus remains elusive.

Dynamics of the Protease Binding Loop. Residues in the protease binding loop exhibit generally low S^2 values ($S^2 <$ 0.6). This is not uncommon in canonical serine protease inhibitors (*5*, *6*, *31*). The difference with other inhibitors is that these values are not exceptionally low compared to the rest of the backbone. However, the relative lowering of the order parameters in the binding loop (corresponding to about 90% of the values of residues in constrained regions) is comparable to that observed in other inhibitors (*5*, *31*), especially in SGTI. The fact that chemical exchange terms should be considered for residues in the binding loop is also not unprecedented; e.g., in CMTI-V, the P1 residue Lys44 was found to exhibit motions on the millisecond−second time scale (*5*).

Comparison of the Precursor Molecule with the Mature Peptides. Chemical shifts obtained for the dimeric precursor

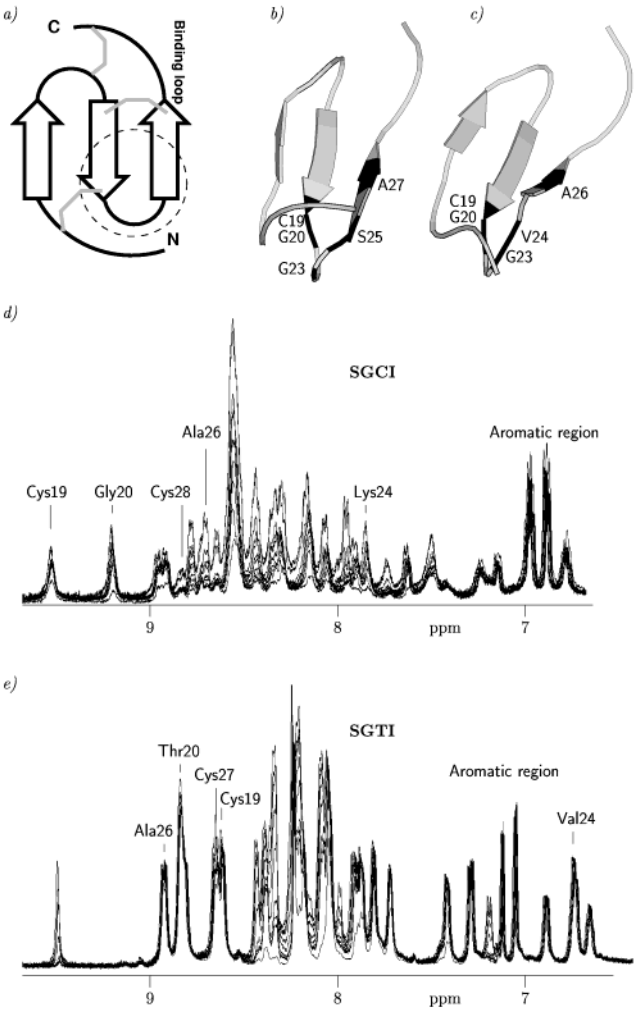

FIGURE 8: Proton−deuterium exchange experiments of SGCI and SGTI [19]. Schematic structure of the inhibitors, the region where the two molecules exhibit considerable differences in H−D exchange is circled (a); cartoon representation of SGCI (b) and SGTI (c); residues showing differences in H−D exchange are labeled; series of one-dimensional spectra recorded to study the H−D exchange properties of the inhibitors SGCI (d) and SGTI (e), the first spectrum recorded 3 min and the last 180 h after dissolving the samples in D_2O. Some of the clearly distinguishable peaks are labeled. For the assignment shown in panels b and c, results of preliminary 2D exchange experiments were also used. Panels b and c were prepared with Molscript (*37*).

molecules (SGTCI and SGTMCI) were highly similar to those of the individual molecules. This suggests that there are no significant interactions between the two modules (SGTI and SGCI in the precursor). This is further supported by the enzyme inhibition properties of SGTCI, as the determined K_I values are close to those of the individual inhibitors. The backbone dynamics of the dimer are also comparable to the observed motional properties of the monomers, except that the order parameters are generally higher than in the individual inhibitors. Order parameters of residues between the modules indicate that the linker is flexible and allows independent motion of the modules. Functional independence of the modules is further supported by the presence of multiple biochemically active homologous inhibitor domains in the pacifastin light chain (*33*).

Biological Implications. One of our long-term aims is to find the structural basis, if any, for the intriguing difference in enzyme specificity of SGCI and SGTI. Taxon-specificity of PMP-D2, the *Locusta* orthologue of SGTI, was also noted (*20*). On the basis of the X-ray structures of PMP-C and PMP-D2 in complex with chymotrypsin, it was suggested that in PMP-D2 the loop comprising residues 20−25 (the $\lambda 2$ loop) makes contacts with the enzyme but not in PMP-C (*20*). This is a result of a slightly different orientation of the two inhibitors with respect to the enzymes in the crystals, which is unexpected given the close structural similarity of the two molecules.

In comparison of the NMR and X-ray structures of PMP-C, a conformational difference in the binding loop region was observed. This was explained on the basis of binding loop flexibility, supported by the low number (four) of medium and long-range NOEs for this region. The deduced cascade of events involving intramolecular hydrogen bond formation upon protease binding is challenged by the finding that the solution structure of SGCI (*19*) compares better with the PMP-C structure in the complex than the PMP-C NMR structure (ref *19* and Figure S2 in Supporting Information). Moreover, one of the three hydrogen bonds between Asn15 and the binding loop observed in the PMP-C chymotrypsin complex is present in the solution structure of SGCI. Thus, a subtle conformational rearrangement of the binding loop and its structural neighborhood is likely to occur but clarifying it exactly will need further investigations.

We speculate that the different dynamical behavior of the molecules is a consequence of the different residues involved in the core interaction characteristic of the inhibitor family (*34*) and may be one of the structural factors affecting protease specificity. In a recent paper Kellenberger et al. (*35*) showed that PMP-D2 (the *Locusta migratoria* homologue of SGTI, exhibiting similar taxon specificity) could be converted into a potent inhibitor of mammalian trypsins by three mutations involving a Pro→Ala replacement in the 20−24 loop (designated as P6−P10 loop in the paper) that also results in redesigning the core of the molecule (PMP-D2-[K10W, P21A, W25A) designated C-like PMP-D2). The authors conclude that possible steric clashes with mammalian proteases could be the reason for the taxon selectivity of PMP-D2 and these mutations act by reshaping the 20−24 loop. Moreover, deleting residues 1−12 of PMP-D2 yields a peptide still showing considerable activity toward *Locusta* trypsins. These observations are consistent with our findings and suggestions about the importance of the different type of stabilization of SGCI and SGTI. It should also be mentioned that wild-type SGCI is fucosylated on Thr9. The role of this modification is not yet clear, as the nonfucosylated form retains full inhibitory activity (*15*), but fucosylation was shown to stabilize the peptide by decreasing its dynamic fluctuations (*36*). Deciphering the exact role of molecular flexibility in the enzyme−inhibitor interactions will need further investigations with different techniques.

CONCLUSION

In our study, we have shown that the internal dynamics of SGCI and SGTI differ significantly. This is intriguing in the light of the close sequence and structural similarity of the two small proteins. Even conserved residues (Cys4, Cys19) and the protease binding loop exhibit different motions on multiple time scales. This may be explained by

the different stabilizing interactions within the two molecules: in SGCI the aromatic side-chain of Phe10, located in the first β-strand, organizes a structural core reminiscent of the hydrophobic core of larger proteins, whereas in SGTI the interaction between Lys10 and Trp25 of the third β-strand should be considered (*19*, *34*). The picture may be more complex with fucosylation, yet another type of stabilizing interaction in SGCI. We suggest that the different internal mobility of the inhibitors may be one of the factors determining their different taxon specificity.

ACKNOWLEDGMENT

We thank András Patthy for his valuable discussions and his help in protein purification on RP-HPLC and William M. Westler for his help in recording the NMR experiments. The generous help of Tripos, Inc., and Antal Lopata (Chemicro Ltd) in obtaining and using SYBYL is also acknowledged.

SUPPORTING INFORMATION AVAILABLE

Figure S1 showing a $^{15}N-^{1}H$ HSQC spectrum of SGTCI, with several peaks exemplifying the similarities and differences in the chemical shifts labeled, and Figure S2 showing comparisons of structures of *Scistocerca* and *Locusta* inhibitors. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bode, W., and Huber, R. (1991) *Eur. J. Biochem. 204*, 433−451.
2. Laskowski, M., Jr., and Qasim, M. A. (2000) *Biochim. Biophys. Acta 1477*, 324−337.
3. Cai, M., Gong, Y., Kao, J. L., and Krishnamoorthi, R. (1995) *Biochemistry 34*, 5201−5211.
4. Francart, C., Duachez, M., Alix, A. J., and Lippens G. (1997) *J. Mol. Biol. 268*, 666−677.
5. Liu, J., Prakash, O., Cai, M., Gong, Y. X., Huang, Y., Wen, L., Wen, J. J., Huang, J.-K., and Krishnamoorthi, R. (1996) *Biochemistry 35*, 1516−1524.
6. Liu, J., Gong, Y., Prakash, O., Wen, L., Lee, I., Huang, J.-K., and Krishnamoorthi (1998) *Protein. Sci. 7*, 132−141.
7. Li, A., and Daggett, V. (1995) *Protein Eng. 8*, 1117−1128.
8. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
9. Shaw, G. L., Davis, B., Keeler, J., and Fersht, A. R. (1995) *Biochemistry 34*, 2225−2233.
10. Liu, J., Prakash, O., Huang, Y., Wen, L., Wen, J. J., Huang, J.-K., and Krishnamoorthi, R. (1996) *Biochemistry 35*, 12503−12510.
11. Cai, M., Gong, Y. X., Wen, L., and Krishnamoorthi, R. (2002) *Biochemistry 41*, 9572−9579.
12. Simonet, G., Claeys, I., and Vanden Broeck, J. (2002) *Comp. Biochem. Biophys. B 132*, 247−255.
13. Vanden Broeck, J., Chiou, S.-J., Schoofs, L., Hamadaoui, A., Vandenbussche, F., Simonet, G., Wataleb, S., and De Loof, A. (1998) *Eur. J. Biochem. 254*, 90−95.
14. Nakakura, N., Hietter, H., van Dorsselaer, A., and Luu, B. (1992) *Eur. J. Biochem. 204*, 147−153.
15. Malik, Z., Amir, S., Pál, G., Buzás, Zs., Várallyay, É., Antal, J., Szilágyi, Z., Vékey, K., Asbóth, B., Patthy, A., and Gráf, L. (1999) *Biochim. Biophys. Acta. 1434*, 143−150.
16. Patthy, A., Amir, S., Malik, Z., Bódi, Á., Kardos, J., Asbóth, B., and Gráf, L. (2002) *Arch. Biochem. Biophys. 398*, 179−187.
17. Mer, G., Kellenberger, C., Koehl, P., Stote, R., Sorokine, O., Van Dorsellaer, A., Luu, B., Hietter, H., and Lefèvre, J.-F. (1994) *Biochemistry 33*, 154397−154407.
18. Mer, G., Hietter, H., Kellenberger, C., Renatus, M., Luu, B., and Lefèvre. J.-F. (1996) *J. Mol. Biol. 258*, 158−171.
19. Gáspári, Z., Patthy, A., Gráf, L., and Perczel, A. (2002) *Eur. J. Biochem. 269*, 527−537.
20. Roussel, A., Mathieu, M., Dobbs, A., Luu, B. Cambillau, C., and Kellenberger, C. (2001) *J. Biol. Chem. 276*, 38893−38898.
21. Kuzmic, P. (1996) *Anal. Biochem. 237*, 260−273.
22. Talluri, S., and Wagner, G. (1996) *J. Magn. Reson., Ser. B 112*, 200−205.
23. Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zilj, P. C. (1995) *J. Magn. Reson. B. 108*, 94−98.
24. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
25. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.
26. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4559−4570.
27. Dosset, P., Huss, J. C., Marion, D., and Blackledge, M. (2001) *J. Biomol. NMR*, *20*, 223−231.
28. Clore, G. M., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) *Biochemistry 29*, 7387−7401.
29. Orekhov, V. Y., Korzhnev, D. M., Pervushin, K. V., Hoffmann, E., and Arseniev, A. S. (1999) *J. Biomol. Struct. Dynam. 17*, 157−174.
30. Campbell, A. P., Spyracopoulos, L., Irvin, R. T., and Sykes, B. D. (2000) *J. Biomol. NMR*, *17*, 239−255.
31. Song, J., and Markley, J. L. (2003) *Biochemistry 42*, 5186−5197.
32. Rachke, T. M., and Marqusee, S. (1998) *Curr. Opin. Biotechnol. 9*, 80−86.
33. Liang, Z., Sottrup-Jensen, L., Aspán, A., Hall, M., and Söderhäll, K. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 6682−6687.
34. Gáspári, Z., Ortutay, C., and Perczel, A. (2004) *Bioinformatics*, in press.
35. Kellenberger, C., Ferrat, G., Leone, P., Darbon, H., and Roussel, A. (2003) *Biochemistry 42*, 13605−13612.
36. Mer, G., Hietter, H., and Lefèvre, J.-F. (1996) *Nat. Struct. Biol. 3*, 45−53.
37. Kraulis, P. J. (1991) *J. Appl. Cryst. 24*, 946−950.

BI035689+