# Analysis of serine proteinase–inhibitor interaction by alanine shaving

# OLGA BUCZEK, <sup>1</sup> KATARZYNA KOSCIELSKA-KASPRZAK, <sup>1</sup> DANIEL KROWARSCH, <sup>1</sup> MICHAŁ DADLEZ, <sup>2</sup> AND JACEK OTLEWSKI <sup>1</sup>

<sup>1</sup>Laboratory of Protein Engineering, Institute of Biochemistry and Molecular Biology, University of Wroclaw, 50–137 Wroclaw, Poland

<sup>2</sup>Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02–106 Warsaw, Poland, and Department of Physics, University of Warsaw, 02–143 Warsaw, Poland

(RECEIVED August 22, 2001; ACCEPTED November 30, 2001)

# Abstract

We analyzed the energetic importance of residues surrounding the hot spot (the  $P_1$  position) of bovine pancreatic trypsin inhibitor (BPTI) in interaction with two proteinases, trypsin and chymotrypsin, by a procedure called molecular shaving. One to eight residues of the structural epitope, composed of two extended and exposed loops, were mutated to alanine(s). Although truncation of the side chains of residues surrounding the P<sub>1</sub> position to methyl groups caused a decrease in  $\Delta G_{den}$  values up to 6.4 kcal mole<sup>-1</sup>, it did not influence the overall conformation of the inhibitor. We found that the replacement of up to six residues with alanines was fully additive at the level of protein stability. To analyze the influence of the structural epitope on the association energy, we determined association constants for BPTI variants and both enzymes and applied the additivity analysis. Shaving of two binding loops led to a progressive drop in the association energy, more pronounced for trypsin (decrease up to 9.6 kcal mole<sup>-1</sup>) than chymotrypsin (decrease up to 3.5 kcal mole<sup>-1</sup>). In the case of extensively mutated variants interacting with chymotrypsin, the association energies agreed very well with the values calculated from single mutational effects. However, when  $P_1$ -neighboring residues were shaved to alanine(s), their contribution to the association energy was not fully removed because of the presence of methyl groups and main chain-main chain intermolecular hydrogen bonds. Moreover, the hot spot had a different contribution to the complex stability in the fully shaved BPTI variant compared with the wild type, which was caused by perturbations of the  $P_1-S_1$  electrostatic interaction.

Keywords: Additivity; bovine pancreatic trypsin inhibitor; protein-protein recognition; serine proteinase

Specific protein–protein recognition plays a crucial role in a great number of vital processes. Because structural information about different complexes is growing rapidly, deducing the rules that determine specificity and strength of protein–protein associations is a particularly attractive problem at this time (Jones and Thornton 1996; Otlewski and Apostoluk 1997; Lo Conte et al. 1999).

Proteins form stable complexes via numerous noncovalent interactions. Protein–protein interfaces are quite large (usually 600–1200 Å<sup>2</sup>) and typically involve 12–25 amino acid residues of each interacting protein. In several model studies of protein–protein recognition, it was shown by double or multiple additivity cycle analysis that the contribution of individual residues is often independent of other residues forming a structural epitope (Wells 1990). The additivity assumption was tested extensively for the Kazal family of protein inhibitors of serine proteinases using 92 natural and 191 recombinant variants with all coded amino acids introduced at contact positions (Lu et al. 2001). A conclusion of that study was that 90% of natural sequences

Reprint requests to: Jacek Otlewski, Laboratory of Protein Engineering, Institute of Biochemistry and Molecular Biology, University of Wroclaw, Tamka 2, 50–137 Wroclaw, Poland; e-mail: otlewski@bf.uni.wroc.pl; fax: +48 (71) 3752608.

Article and publication are at http://www.proteinscience.org/cgi/doi/ 10.1110/ps.3510102.

that contain up to six changes at contact positions show full or at least partial additivity. Clear exceptions from the additivity rule are residues on the structural epitope that are in direct contact (Wells 1990) or affect each other through a long–range electrostatic effect (Schreiber and Fersht 1995).

Alanine scanning mutagenesis studies (Wells 1991) on many model protein–protein systems showed that the functional roles of residues at a structural epitope are not equal. Mutation to alanine of less than a quarter of the contacting residues causes a significant decrease in the association energy (above 1 kcal mole<sup>-1</sup>), whereas the remainder of the interacting area plays a passive role (Bogan and Thorn 1998). The location of this hot spot is typically central and occluded from the bulk solvent by passive residues.

The two above-mentioned phenomena, additivity of residual free energies and presence of the hot spot residue(s), suggest that it should be possible to truncate side chains of passive residues without affecting global association energy. Indeed, Jin and Wells (1994), in a unique study on human growth hormone (hGH)-monoclonal antibody recognition, showed that introduction of up to 16 alanines in the passive zone of hGH (a procedure called molecular shaving) caused only marginal reduction of the affinity. The hot spot of hGH is relatively large, composed of five residues and clearly provides contact area, sufficient to outweigh the loss of surrounding side chains caused by alanine shaving. The question arises whether a more localized hot spot, for example, composed of one or two residues, could provide enough energy to stabilize the complex when the surrounding side chains are shaved.

In this study, we analyzed a single residue hot spot-the P<sub>1</sub> position (according to Schechter and Berger 1967) of bovine pancreatic trypsin inhibitor (BPTI). We chose an interaction of BPTI with two bovine proteinases: trypsin and chymotrypsin. High-resolution crystal structures are available for trypsin (Huber et al. 1974; Helland et al. 1999) and chymotrypsin (Scheidig et al. 1997) complexes with this inhibitor. A variant of BPTI, containing Lys15Arg and Met52Leu mutations, was used as the template for alanine shaving and further as the pseudo wild-type reference. The former mutation was introduced to increase the association constant for some of the most extensively mutated variants, as for both enzymes Arg is preferred about threefold over the naturally occurring Lys (Krowarsch et al. 1999). The latter mutation was made to allow CNBr cleavage of a fusion protein during the purification procedure.

Previously, we showed that truncation of the P<sub>1</sub> Arg 15, to Gly decreases the  $K_a$  for trypsin  $3 \times 10^9$ -fold. That is, it removes 68% of the total free energy of association. The effect of the Arg15Gly substitution on interaction with chymotrypsin is much smaller, a  $3.16 \times 10^3$ -fold decrease, equivalent to a 41% loss of the association energy. In addition, kinetic and thermodynamic data for binding of

15 single alanine mutants of BPTI comprising all residues that are in contact with both enzymes are available (Castro and Anderson 1996). Interaction of the pseudo wild-type BPTI with trypsin ( $K_a = 4.5 \times 10^{13} \text{ M}^{-1}$ ) is over five orders of magnitude stronger than that with chymotrypsin  $(K_a = 2.5 \times 10^8 \text{ M}^{-1})$  (Krowarsch et al. 1999). The strongest contribution to association energy for the BPTI-trypsin complex seems to come from the electrostatic interaction between the P<sub>1</sub> Lys 15, and Asp 189 in the S<sub>1</sub> pocket of the enzyme. From the comparison of crystal structures of trypsin complexes with Lys 15 and Met 15 BPTI variants and their binding energies, we could estimate an over  $4.4 \times 10^{5}$ fold increase in the association constant due to the presence of the Lys 15 side chain positive charge (Krowarsch et al. 1999). Thus, because of the presence of this extremely favorable electrostatic complementarity, the P<sub>1</sub> position serves as the true hot spot of trypsin-BPTI recognition. There is no equivalent electrostatic interaction in the chymotrypsin complex and the association energy is much weaker. In this study, we analyzed the energetics of complex formation with trypsin and chymotrypsin using 10 mutants of BPTI. In these mutants, one to eight proteinasecontacting residues were converted to alanine(s) with the aim of revealing the energetic role of the  $P_1$  residue in alanine shaved mutants.

# Results

## Alanine mutations

The analysis of the crystal structures of bovine trypsin-BPTI (Huber et al. 1974; Helland et al. 1999) and bovine chymotrypsin-BPTI (Capasso et al. 1997; Scheidig et al. 1997) complexes shows that there are 13 residues of the inhibitor forming <4 Å contacts with the enzyme. Those residues form a structural epitope in the form of two extended and exposed loops, the primary (Thr 11-Ile 19) and the secondary one (Gly 36-Arg 39; Fig. 1A). In addition, Val 34 is the only residue in both complexes that makes contacts in the range of 4.0–5.0 Å. Individual residue contacts below 4.0 Å together with a residue change in the solvent accessible surface area (ASA) accompanying complex formation are summarized in Table 1. For both complexes, contact numbers and distribution of van der Waals contacts among inhibitor residues are similar, which is a reflection of similar folds and sequences of both enzymes.

We prepared 10 mutants of BPTI in which alanines were progressively introduced into the primary and/or secondary binding loop (Table 2) except for two disulfide-bonded cysteine and two glycine positions. We did not substitute Gly 12 as its torsion angles are not compatible with amino acids containing  $\beta$ -carbons, and we showed previously that substitution of Gly 12 always caused a strong decrease in  $K_a$ 



Fig. 1. Space-filling model of the pseudo wild-type BPTI (*A*) and  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  mutant (*B*) surface in contact with trypsin. Residues replaced in this work are highlighted and shaded and other contact residues (Gly 12, Cys 14, Ala 16, Gly 36, and Cys 38) are highlighted but not shaded.

and stability (Grzesiak et al. 2000b). Also, we did not introduce the Gly36Ala mutation because it is known to produce a very large, 4000-fold, decrease in  $K_a$  with trypsin (Castro and Anderson 1996). Hence, Gly36Ala, together with several other mutations, would lead to variants with unmeasurably low  $K_a$  values.

In a free inhibitor, side chains of mutated residues are at least 40% solvent exposed (Arg 15, Arg 17, and Arg 39 in >75%) and do not form many interactions with the remainder of the inhibitor. The substituted residues form only one hydrogen bond, Thr 11  $O^{\gamma 1}$ –Val 34 O (2.64 Å), and seven van der Waals contacts, Ile 18  $C^{\delta 1}$ –Gly 37  $C^{\alpha}$  (3.59 Å), Val 34  $C^{\gamma 1}$ –Arg 17  $C^{\alpha}$  (3.74 Å), Thr 11  $C^{\beta}$ –Tyr 10 C (3.67 Å), Val 34  $C^{\gamma 2}$ –Phe 33 C (3.74 Å), Val 34  $C^{\beta}$ –Phe 33 C (3.63 Å), Val 34  $C^{\gamma 2}$ –Ile 19  $C^{\delta 1}$  (3.48 Å) and Ile 19  $C^{\gamma 2}$ –Thr 32  $C^{\beta}$  (3.72 Å). Thus, alanine mutation(s) are not expected to break many stabilizing interactions.

The sequences of the mutants are shown in Table 2. All the variants were refolded from inclusion bodies to their native structures under conditions enabling correct pairing of disulfide bonds. For those that contain more than three alanine substitutions and are much less stable (see below), we had to apply more gentle refolding conditions (Table 2).

# Association of alanine mutants with trypsin and chymotrypsin

The association constant values for interaction of all mutants with trypsin or chymotrypsin are shown in Table 3. We decided to determine  $K_a$  values at pH 8.2 to enable direct comparison with the single alanine mutants of Castro and Anderson (1996). However, the  $K_a$  values could be directly measured at pH 8.2 only for three trypsin and seven chymotrypsin mutant complexes. Partial digestion of the inhibitor observed upon incubation precluded direct  $K_a$  measurements for the remaining ones. To obtain  $K_a$  values at pH 8.2 for all mutant trypsin complexes, we measured  $K_a$  values at pH 5.5, at which the complex is still strong but no digestion of the inhibitor was observed due to decreased activity of the enzyme. As the ratio of  $K_a$  values at pH 8.2 and pH 5.5 was similar for all three mutants for which direct measurements were possible (70, 62.5, and 70, see Table 3), the remaining  $K_a$  values at pH 8.2 were calculated using the average ratio factor of 68. We also found similar factors, 19.7, 19.7, and 19.1, respectively, for the three chosen mutants in interaction with chymotrypsin. Therefore, we calculated the  $K_a$  value for  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  at pH 8.2 using an average factor of 19.5. The analysis of alanine mutation effects presented below will be based on  $K_a$  values at pH 8.2 because they allow direct and extensive comparison with the alanine scan data of Castro and Anderson (1996).

Generally, introduction of alanine(s) into the contact region led to a progressive decrease in the association energy with trypsin and chymotrypsin (Table 3). For the most extensively mutated variant compared with the pseudo wild type, the effect was much more pronounced for trypsin  $(1.3 \times 10^7$ -fold decrease). In the case of chymotrypsin binding, it was only  $3.8 \times 10^2$ -fold weaker. For single alanine mutants,  $R^{15}A^{17}$  and  $R^{15}A^{34}$ , no decrease in binding constants was observed for both enzymes. However, a single substitution, Pro13Ala, produced very different effects—it was neutral for chymotrypsin and highly deleterious for trypsin.

Multiple mutants, consisting of more than two alanine substitutions, behaved differently with the two proteinases. Generally, these six mutants were worse inhibitors, however,  $K_a$  decreases were always more pronounced for trypsin than chymotrypsin. Additionally, the ratio of decreases changed from mutant to mutant. For example, the ratio is the smallest for  $\mathbf{A^{13}R^{15}A^{17}A^{18}A^{19}}$  ( $1.1 \times 10^3 \div 1.7 \times 10^2 = 6.7$ ) and the highest for  $\mathbf{A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{37}A^{39}}$  ( $1.3 \times 10^7 \div 3.8 \times 10^2 = 3.5 \times 10^4$ ). In other words, the ratio of trypsin to chymotrypsin association constants is  $1.8 \times 10^5$  for the pseudo wild-type BPTI, but, for the most extensive  $\mathbf{A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{37}A^{39}}$  mutant, it is only 5.2.

			Trypsin		Chymotrypsin				
	vdW <sup>a</sup>	HH <sup>b</sup>	ASA <sup>c</sup> (Å <sup>2</sup> )	ASA change <sup>d</sup>	vdW <sup>a</sup>	$\mathrm{HH}^{\mathrm{b}}$	ASA <sup>c</sup> (Å <sup>2</sup> )	ASA change <sup>d</sup>	
Thr 11	1	0	64	25%	1	0	64	56%	
Gly 12	1	0	20	21%	2	0	18	88%	
Pro 13	9	1	72	76%	14	1	78	75%	
Cys 14	16	1	55	100%	9	0	48	100%	
Arg 15	62	7	239	100%	60	4	224	98%	
Ala 16	15	0	46	100%	17	0	55	100%	
Arg 17	21	3	182	78%	28	0	201	88%	
Ile 18	2	0	65	89%	1	0	67	52%	
Ile 19	8	1	105	36%	0	0	118	33%	
Gly 36	2	0	1	100%	2	0	1	100%	
Gly 37	2	0	37	60%	7	0	34	68%	
Cys 38	3	0	32	88%	2	0	36	91%	
Arg 39	5	2	179	48%	2	1	168	41%	
Total	147	15	1097	790 $Å^2$	145	6	1112	$808 \text{ Å}^2$	

**Table 1.** Number of noncovalent interactions formed by 13 residues comprising the binding epitope of BPTI, their ASA in free inhibitor, and ASA change accompanying complex formation with trypsin or chymotrypsin

All values were calculated using NACCESS (Hubbard and Thornton 1993) or HBPLUS program (McDonald and Thornton 1994) and structural data for respective trypsin–BPTI (Helland et al. 1999) and chymotrypsin–BPTI complexes (Scheidig et al. 1997); in case of Arg 15, values were calculated from data for APPI complexed with trypsin or chymotrypsin (Scheidig et al., 1997). Both ASA value sets for free inhibitor slightly differ, since they were calculated from inhibitor portion obtained from respective complexes (Helland et al. 1999; Scheidig et al. 1997).

<sup>a</sup> Number of intermolecular van der Waals contacts below 4 Å. <sup>b</sup> Number of intermolecular hydrogen bonds defined according to standard parameters.

<sup>c</sup> ASA in free inhibitor calculated using 1.4 Å water radius.

<sup>d</sup> ASA change accompanying complex formation.

ASA change accompanying complex formation.

# Additivity of association energies

The  $K_a$  values for alanine mutants were converted into free energy of association values ( $\Delta G_a$ ) and grouped in seven thermodynamic additivity cycles, denoted A–G (Fig. 2). In the cycle, the effects of the same alanine substitution(s) ( $\Delta\Delta G_a$ ) in two different sequence contexts were compared. The key value of the cycle is  $|\Delta\Delta\Delta G_a|$ , an absolute value of a difference of two  $\Delta\Delta G_a$ s, called an interaction or coupling energy (Wells 1990; LiCata and Ackers 1995).  $|\Delta\Delta\Delta G_a|$ determines whether the effect of substitution is additive. For an ideal additivity cycle,  $|\Delta\Delta\Delta G_a|$  should equal 0. In our cycles  $|\Delta\Delta\Delta G_a|$  ranges from 0 to 2.9 kcal mole<sup>-1</sup> for trypsin and 0.1 to 1.78 kcal mole<sup>-1</sup> for chymotrypsin, suggesting that nonadditivity occurs in most of the considered cycles. However, the cycle comprises four independent  $K_a$ s, each with an estimated error of ±20%, equivalent to ±0.1 kcal mole<sup>-1</sup> error in  $\Delta G_a$  value at the 1 $\sigma$  level and, consequently,

Table 2.	Alanine	mutants	of	BPTI	used	in	the	study
----------	---------	---------	----	------	------	----	-----	-------

Lys15Arg	R <sup>15</sup>	80% <sup>a</sup>
Pro13Ala, Lys15Arg	A <sup>13</sup> R <sup>15</sup>	70% <sup>a</sup>
Thr11Ala, Pro13Ala, Lys15Arg	$A^{11}A^{13}R^{15}$	$60\%^{\mathrm{a}}$
Lys15Arg, Arg17Ala	R <sup>15</sup> A <sup>17</sup>	70% <sup>a</sup>
Lys15Arg, Val34Ala	$R^{15}A^{34}$	70% <sup>a</sup>
Lys15Arg, Val34Ala, Gly37Ala, Arg39Ala	$R^{15}A^{34}A^{37}A^{39}$	40% <sup>a</sup>
Pro13Ala, Lys15Arg, Arg17Ala, Ile18Ala, Ile19Ala	A <sup>13</sup> R <sup>15</sup> A <sup>17</sup> A <sup>18</sup> A <sup>19</sup>	40% <sup>a</sup>
Thr11Ala, Pro13Ala, Lys15Arg, Arg17Ala, Ile18Ala, Ile19Ala	A <sup>11</sup> A <sup>13</sup> R <sup>15</sup> A <sup>17</sup> A <sup>18</sup> A <sup>19</sup>	30% <sup>a</sup>
Pro13Ala, Lys15Arg, Arg17Ala, Ile18Ala, Ile19Ala, Val34Ala	A <sup>13</sup> R <sup>15</sup> A <sup>17</sup> A <sup>18</sup> A <sup>19</sup> A <sup>34</sup>	30% <sup>b</sup>
Thr11Ala, Pro13Ala, Lys15Arg, Arg17Ala, Ile18Ala, Ile19Ala, Val34Ala	A <sup>11</sup> A <sup>13</sup> R <sup>15</sup> A <sup>17</sup> A <sup>18</sup> A <sup>19</sup> A <sup>34</sup>	30% <sup>b</sup>
Thr11Ala, Pro13Ala, Lys15Arg, Arg17Ala, Ile18Ala, Ile19Ala, Val34Ala, Gly37Ala, Arg39Ala	$A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$	30% <sup>b</sup>

Alanine mutations introduced into the primary binding loop are shown in bold and those in the secondary binding loop are in italic. All mutants contain additional Met52Leu substitution. Refolding yield is shown in the third column.

<sup>a</sup> Refolding conditions:  $1 \times 10^{-4}$  M BTPI, 4 M urea, 20 mM Tris, pH 8.7,  $6 \times 10^{-4}$  M GSH,  $3 \times 10^{-4}$  M

GSSG, 3 mM EDTA, 1 h, 47°C.

<sup>b</sup> Refolding conditions:  $1 \times 10^{-5}$  M BPTI, 0.4 M urea, 20 mM Tris, pH 8.7,  $6 \times 10^{-4}$  M GSH,  $3 \times 10^{-4}$  M GSSG, 3 mM EDTA, 1 h, 37°C.

		Trypsin			Chymotrypsin	
BPTI mutant	$\frac{K_a (M^{-1})}{pH 5.5}$	$K_a (M^{-1})$ pH 8.2	effect	$\frac{K_{\rm a}  ({\rm M}^{-1})}{{ m pH}  5.5}$	$K_{\rm a} ({\rm M}^{-1})$ pH 8.2	effect
R <sup>15</sup>		$4.5 \times 10^{13a}$	_	$1.27 \times 10^{7}$	$2.5 \times 10^{8}$	
$A^{13}R^{15}$	$7.2 \times 10^{9}$	$4.5 \times 10^{11}$	$1.0 \times 10^2 \downarrow$	$1.22 \times 10^{7}$	$2.4 \times 10^{8}$	1
$A^{11}A^{13}R^{15}$	$6.0 \times 10^{9}$	$4.2 \times 10^{11}$	$1.1 \times 10^2 \downarrow$	$1.15 \times 10^{7}$	$2.2 \times 10^{8}$	1.1 ↓
R <sup>15</sup> A <sup>17</sup>	$4.1 \times 10^{11}$	$2.8 \times 10^{13b}$	1.6↓	_	$1.8 \times 10^{8}$	1.4 ↓
$R^{15}A^{34}$	$6.5 \times 10^{11}$	$4.4 \times 10^{13b}$	1.0	_	$2.2 \times 10^{8}$	1.1 ↓
$R^{15} A^{34} A^{37} A^{30}$	$7.4 \times 10^{8}$	$5.2 \times 10^{10}$	$8.7 \times 10^2 \downarrow$	_	$6.0 \times 10^{6}$	41.7 ↓
A <sup>13</sup> R <sup>15</sup> A <sup>17</sup> A <sup>18</sup> A <sup>19</sup>	$5.9 \times 10^{8}$	$4.0 \times 10^{10b}$	$1.1 \times 10^3 \downarrow$	_	$1.5 \times 10^{6}$	$1.7 \times 10^2 \downarrow$
$A^{11}A^{13}R^{15} A^{17}A^{18}A^{19}$	$6.1 \times 10^{8}$	$4.1 \times 10^{10b}$	$1.1 \times 10^3 \downarrow$	_	$3.7 \times 10^{6}$	67.6↓
$A^{13} R^{15} A^{17} A^{18} A^{19} A^{34}$	$3.5 \times 10^{6}$	$2.4 \times 10^{8b}$	$1.9 \times 10^5 \downarrow$	_	$2.0 \times 10^{5}$	$1.3 \times 10^3 \downarrow$
$A^{11}A^{13}R^{15} A^{17}A^{18}A^{19}A^{34}$	$3.6 \times 10^{6}$	$2.4 \times 10^{8b}$	$1.9 \times 10^5 \downarrow$	_	$3.7 \times 10^{6}$	67.6↓
$\mathbf{A^{11}A^{13}R^{15}} \mathbf{A^{17}A^{18}A^{19}} A^{34} A^{37} A^{39}$	$5.1 \times 10^{4}$	$3.4 \times 10^{6b}$	$1.3 \times 10^7 \downarrow$	$3.4 \times 10^{4}$	$6.6 \times 10^{5a}$	$3.8 \times 10^2 \downarrow$

**Table 3.** Association constants  $K_{\alpha}$  determined for the interaction of alanine mutants of BPTI with bovine  $\beta$ -trypsin and  $\alpha$ -chymotrypsin at pH 8.2 and 5.5, 22°C

Inhibition effects were compared with the pseudo wild-type BPTI. Downward arrow indicates a decrease in the association constant value upon mutations. Estimated error of  $K_a$  values is  $\pm 20\%$ .

<sup>a</sup> From Krowarsch et al. (1999).

<sup>b</sup>  $K_a$  values calculated from the equation  $K_{a,pH \ 8.2} = K_{a,pH \ 5.5} \times 68$  (see text). <sup>c</sup>  $K_a$  values calculated from the equation  $K_{a,pH \ 8.2} = K_{a,pH \ 5.5} \times 19.5$  (see text).

0.8 kcal mole<sup>-1</sup> as an acceptable additivity limit of  $|\Delta\Delta\Delta G_a|$ values at the  $2\sigma$  level (Qasim et al. 1997; Krowarsch et al. 1999). Cases in which  $|\Delta\Delta\Delta G_a|$  is <0.8 kcal mole<sup>-1</sup> will be called additive and >0.8 kcal mole<sup>-1</sup> – nonadditive. Three cycles: A, B, and C are additive, and four cycles: D, E, F, and G are nonadditive for trypsin. In the case of chymotrypsin, two cycles (B and E) are additive and five cycles (A, C, D, F, and G) are nonadditive.

Besides the analysis presented above, there is another way to estimate additivity effects in our system. Castro and Anderson (1996) reported  $K_a$  values for chymotrypsin association with 15 single mutants of BPTI in which all contact residues were mutated to alanine. Their data for trypsin are less useful here, because, for many of the single mutants, only a lower limit of  $K_a$  could be estimated. To combine both data sets, we applied identical experimental conditions of buffer composition, pH, and temperature. Table 4 compares experimental  $\Delta G_a$  values for chymotrypsin determined in this work and  $\Delta G_{a,calc}$  values calculated by Castro and Anderson (1996) assuming full additivity. The nonadditivity term is equal to a difference between the calculated and measured value:  $\Delta\Delta G_{a,calc} = \Delta G_{a,calc} - \Delta G_{a}$ .

Estimation of additive behavior requires a careful error analysis. The error in  $\Delta G_a$  values for 15 single alanine

**Table 4.** Experimental  $(\Delta G_a)$  and calculated  $(\Delta G_{a,calc})$  values of free energy of binding for 10 alanine mutants and chymotrypsin at 22°C and pH 8.2

			AAG b	$\Delta\Delta G_{a,calc}$ error			
BPTI mutant	$\Delta G_a$ (kcal mole <sup>-1</sup> )	$\Delta G_{a,calc}$ (kcal mole <sup>-1</sup> )	$\Delta\Delta G_{a,calc}$ (kcal mole <sup>-1</sup> )	kc	(kcal mole <sup>-1</sup> )	additivity <sup>d</sup>	
A <sup>13</sup> R <sup>15</sup>	-11.25	-11.35	-0.1	1	±0.40	+	
$A^{11}A^{13}R^{15}$	-11.20	-11.15	0.05	2	±0.40	+	
R <sup>15</sup> A <sup>17</sup>	-11.08	-10.75	0.33	1	±0.40	+	
$R^{15}A^{34}$	-11.20	-11.24	-0.04	1	±0.40	+	
$R^{15}A^{34}A^{37}A^{39}$	-9.10	-10.26	-1.16	3	±0.57	_	
$A^{13}R^{15}A^{17}A^{18}A^{19}$	-8.29	-9.34	-1.05	4	±0.75	_	
$A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}$	-8.82	-9.15	-0.33	5	±0.94	+	
$A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$	-7.12	-9.31	-2.19	5	±0.94	_	
$A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$	-8.82	-9.15	-0.33	6	±1.13	+	
$A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$	-7.81	-8.13	-0.32	8	±1.52	+	

 $^{a}\Delta G_{a,calc}$  was calculated from data for single alanine mutants (Castro and Anderson 1996).

 ${}^{b}\Delta\Delta G_{a,calc} = \Delta G_{a,calc} - \Delta G_{a}$ .  ${}^{c}$  k is a number of introduced alanine mutations.

<sup>d</sup> (+) Additivity; (–) non-additivity.



Fig. 2. Thermodynamic additivity cycles between alanine mutants of BPTI used in this work. The values for trypsin interaction are shown above (bold) and those for chymotrypsin are below (grey). The influence of the same alanine substitution(s) is described by two  $\Delta\Delta G_a$  values equal to  $\Delta G_{a, \text{ mutant with less Ala}} - \Delta G_{a, \text{ mutant with more Ala}} \Delta\Delta\Delta G_a$  is a difference between  $\Delta\Delta G_a$  values calculated for the same alanine mutation(s) in the two compared sequence contexts.

mutants was similar to that in our work:  $\pm 0.2$  kcal mole<sup>-1</sup> at the  $2\sigma$  level (Castro and Anderson, 1996). For  $\Delta\Delta G_{a,calc}$ , the global error depends on the number of introduced alanines, according to the equation (Lu et al. 2001):

$$\Delta\Delta G_{a \ calc} \ (\text{error}) = \pm 0.2 \ \text{kcal} \cdot \text{mol}^{-1} \cdot \sqrt{k^2 - k + 2}$$
(1)

where *k* is the number of introduced alanine mutations. This error analysis is valid for k > 2. If *k* equals 1 or 2, the error is calculated as a sum of individual errors (Lu et al. 2001).

Inspection of Table 4 shows that seven out of 10 alanine mutants are fully additive and three are nonadditive. Interestingly, these three mutants  $R^{15}A^{34}A^{37}A^{39}$ ,  $A^{13}R^{15}A^{17}A^{18}A^{19}$ , and  $A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$  contain three, four, and five alanine substitutions—less than the two most extensive variants (six and eight alanines), which are additive. Upon addition of the Thr11Ala mutation to the latter two variants, additivity is recovered, as observed in the  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}$  and  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$  variants (Table 4). Respective  $\Delta G_a$  values are decreased by 0.53 and 1.70 kcal mole<sup>-1</sup>. On the contrary, introduction of the single Thr11Ala substitution increases  $\Delta G_a$  by 0.2 kcal mole<sup>-1</sup> (Castro and Anderson 1996). It should be stressed that the most extensive mutant containing eight alanines binds chymotrypsin exactly as predicted from single mutant effects.

#### Conformation and stability of alanine mutants

Although all our variants inhibited both proteinases, we were aware that some of them might be structurally altered, as we introduced multiple alanine substitutions into a small protein. We tested the structural properties of all the variants using CD spectroscopy and thermodynamic stability analysis.

CD spectra recorded in a near and far UV did not differ significantly from the pseudo wild-type protein spectra, indicating a properly folded structure of the mutants (Fig. 3). Only the spectrum of the most extensive mutant,  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  BPTI, showed significant changes in the far UV. Further, transitions of thermal denaturation curves of all mutants, monitored both with CD and DSC, were highly cooperative. We analyzed denaturation of mutants at three pH values: 2.0, 2.5, and 2.7. Denaturation temperatures for all mutants were 6.4°C to 34.8°C lower compared with the pseudo wild type. Generally, introduction of a larger number of alanines led to a more severe destabilization of the molecule. In this respect, mutant  $\mathbf{A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}}$  behaves unusually, as it is more stable than several less extensive mutants, which contain four to six alanines (Table 5). Additionally, we determined stability parameters for all the mutants using the DSC method.  $T_{den}$  values were on average 1.5°C higher than those determined from CD-monitored curves, and the

812



**Fig. 3.** Circular dichroism spectra of alanine mutants of BPTI in far (*top*) and near (*bottom*) UV range. Spectra were recorded in 10 mM Gly/HCl, (pH 2.0 at  $25^{\circ}$ C). Details are described in the Materials and Methods.

ratio of calorimetric to van't Hoff enthalpy was close to unity, suggesting the two-state transition.

 $\Delta H_{den}$  versus  $T_{den}$  data points for individual mutants do not fall on a single line, rather they provide a set of (roughly) parallel lines (Fig. 4). Therefore, to compare the thermodynamic stability of the mutants, we calculated  $\Delta G_{den}$  values from the Gibbs-Helmholtz equation at a common temperature (25 °C). The mutants showed decreases in  $\Delta G_{den}$  values up to >6.4 kcal mole<sup>-1</sup>, strongly manifested in the case of those containing three or more alanines. Thus, up to 77% of the stabilization energy was lost upon introduction of alanine mutations. Two single mutants: R<sup>15</sup>A<sup>17</sup> and R<sup>15</sup>A<sup>34</sup>, despite lower  $T_{den}$  values, were marginally more stable than the pseudo wild type because of larger  $\Delta H_{den}$ values (Table 5 and Fig. 4).

BPTI mutant	pН	$T_{den}$ (°C)	$\Delta T_{den}$ (°C)	$\frac{\Delta H_{\rm den}}{\rm (kcal\ mole^{-1})}$	$\Delta C_{p,den}^{a}$ (kcal mole <sup>-1</sup> deg <sup>-1</sup> )	$\Delta G_{den}^{(25^{\circ}\mathrm{C}) \mathrm{b}}$ (kcal mol <sup>-1</sup> )	$\Delta\Delta G_{den}^{(25^{\circ}\mathrm{C})}$ (kcal mol <sup>-1</sup> )
R <sup>15</sup>	2.0	81.9	_	63.0		8.27	_
	2.5	83.4	_	64.5			
	2.7	89.3		66.1	0.38 (±0.12)		
$A^{13}R^{15}$	2.0	69.9	12.0	53.7		5.95	2.32
	2.5	71.9	11.5	53.9			
	2.7	74.7	14.6	55.3	0.35 (±0.1)		
$A^{11}A^{13}R^{15}$	2.0	69.7	12.2	54.4		6.06	2.21
	2.5	72.1	11.3	55.9			
	2.7	75.1	14.2	56.3	0.34 (±0.14)		
R <sup>15</sup> A <sup>17</sup>	2.0	75.5	6.4	69.4		8.67	-0.4
	2.5	77.3	6.1	70.1			
	2.7	83.3	6.0	72.2	0.36 (±0.004)		
$R^{15}A^{34}$	2.0	74.8	7.1	68.1		8.44	-0.17
	2.5	76.6	6.8	68.9			
	2.7	82.1	7.2	70.7	0.35 (±0.01)		
$R^{15}A^{34}A^{37}A^{39}$	2.0	64.5	17.4	35.3		3.37	4.9
	2.5	68.9	14.5	37.7			
	2.7	75.4	13.9	39.0	0.33 (±0.087)		
$A^{13}R^{15}A^{17}A^{18}A^{19}$	2.0	47.1	34.8	30.6		1.87	6.4
	2.5	49.1	34.3	31.7			
	2.7	53.0	36.3	32.5	0.31 (±0.09)		
A <sup>11</sup> A <sup>13</sup> R <sup>15</sup> A <sup>17</sup> A <sup>18</sup> A <sup>19</sup>	2.0	47.5	34.4	30.6		1.90	6.37
	2.5	48.1	35.3	31.5			
	2.7	53.1	36.2	32.6	0.31 (±0.11)		
$A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$	2.0	47.3	34.6	34.1		2.18	6.09
	2.5	50.9	32.5	34.9			
	2.7	51.9	37.4	35.5	0.25 (±0.003)		
$A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$	2.0	48.6	33.3	35.1		2.35	5.92
	2.5	49.2	34.2	35.9			
	2.7	52.0	37.3	36.5	0.25 (±0.067)		
<b>A<sup>11</sup>A<sup>13</sup></b> R <sup>15</sup> <b>A<sup>17</sup>A<sup>18</sup>A<sup>19</sup></b> A <sup>34</sup> A <sup>37</sup> A <sup>39</sup>	2.0	60.9	21.0	41.7		_	_
	2.5	N.D.	_	N.D.			
	2.7	N.D.	_	N.D.	_		

 Table 5. CD spectroscopy determined thermal denaturation parameters for BPTI variants

The unfolding experiments were performed in 10 mM Gly/HCl, pH 2.0, 2.5, and 2.7 with final protein concentration of 20  $\mu$ g ml<sup>-1</sup>; ellipticity was followed at 223 nm, with temperature scan rate of 1 deg min<sup>-1</sup> (Jasco model J-715 spectropolarimeter).

<sup>a</sup>  $\Delta C_{p,den}$  calculated from  $\Delta H_{den}$  dependence on  $T_{den}$  (Fig. 4).

<sup>b</sup> Denaturation free energy changes were calculated at 25°C at pH 2.0 using Gibbs-Helmholz equation; (N.D.) Not determined.

We checked whether the stability effects were additive through thermodynamic cycle analysis (Fig. 5). To analyze additivity, we first estimated errors in  $T_{den}$  (±0.1°C),  $\Delta H_{den}$ (±1 kcal mole<sup>-1</sup>) and  $\Delta C_{p,den}$  (±20% of the experimental value). Hence, the error in  $\Delta G_{den}$  from the Gibbs-Helmholtz equation equals ±0.34 kcal mole<sup>-1</sup>, and the limit on additivity is  $|\Delta\Delta\Delta G_{den}| = 4 \times 0.34$  kcal mole<sup>-1</sup> = 1.36 kcal mole<sup>-1</sup>.

 $|\Delta\Delta\Delta G_{den}|$  values reported for five thermodynamic cycles in Figure 5 show that introduction of alanine(s) is fully additive, despite very large effects on the stability of individual mutants. This observation suggests that alanine substitutions are independent of each other in terms of free energy changes. However, we noticed unusual behavior of  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  BPTI. Compared to the  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$  variant ( $T_{den} = 48.6^{\circ}$ C at pH 2.0), the previous one contains two additional mutations, Gly37Ala and Arg39Ala. Yu et al. (1995) showed that the former substitution reduced  $T_{den}$  by 17.2°C, and the latter had no effect on BPTI stability. Although these two effects were observed in the context of the single Cys5–Cys55 disulfide BPTI variant and at pH 7.0, we feel that severe nonadditivity occurs for  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  BPTI. The two mutations should decrease its  $T_{den}$  to about 31.4°C (48.6°C–17.2°C), and the observed  $T_{den}$  at pH 2.0 is 60.9°C. The determination of the NMR solution structure for this variant is currently underway in our laboratory.

# Discussion

The main aim of this study was to analyze the influence of surrounding residues on the energy of interaction of a functional epitope formed by the single  $P_1$  residue of a protein inhibitor, BPTI, with trypsin and chymotrypsin. To accom-



**Fig. 4.** Denaturation temperature  $(T_{den})$  dependence of van't Hoff enthalpy  $(\Delta H_{\nu H})$  for 10 alanine mutants of BPTI determined in 10 mM Gly/HCl buffer at pHs 2.0, 2.5, and 2.7.

plish that, we applied additivity cycle analysis using 10 single or multiple BPTI mutants, in which different sets of contact residues were converted to alanine(s). Serine proteinase-protein inhibitor interaction seems to be well suited for such analysis. In all known complexes of this type, the shape of the contact area and the geometry of the interaction, including the network of hydrogen bonds, are well preserved, despite differences in the sequences of proteinases and inhibitors (Bode and Huber 1992). For example, geometry of trypsin recognition by 10 P<sub>1</sub> mutants of BPTI that differ in strength of association with trypsin by nine orders of magnitude is identical, even at the level of the length of individual intermolecular hydrogen bonds (Helland et al. 1999). It has also been shown that the same mutations at the P<sub>1</sub> position of inhibitors from families of totally different scaffoldings led to similar energetic effects (Qasim et al. 1997; Krowarsch et al. 1999). An analogous intrascaffolding additivity rule works when contact residues of one interacting component are mutated and analyzed through a thermodynamic cycle (Wells, 1990; Lu et al. 2001). Finally, Kuroda and Kim (2000) showed that BPTI scaffolding is tolerant of extensive mutagenesis as it is able to accept up to 50% of alanines without major conformational changes thus justifying extensive alanine mutagenesis of the inhibitor.

The following analysis of protein–protein association data is based on the assumption that, for all mutants, no conformational adjustments occur upon alanine(s) introduction when compared with the pseudo wild-type BPTI, which was verified by CD spectra measurements. Additionally, the atomic resolution structure for  $A^{11}A^{13}R^{15}$  BPTI showed a similar conformation of its primary binding loop to that of the wild-type BPTI with an rmsd of 0.4 Å for backbone atoms (Czapinska et al. 2000; Addlagatta et al. 2001). We are aware that the conformation of  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$ BPTI may differ from all others, as its CD spectrum was changed in the far UV range. Also, the thermodynamic stability of this mutant showed significant deviation from additivity.

We feel that the additivity tests performed here are rigorous as the mutations compared lead to severe changes in the complex contact area. Figure 6 shows estimated changes in the size of the contact area accompanying complex formation calculated for all alanine mutants interacting with trypsin and chymotrypsin. To illustrate the extent of the changes, a comparison of structural epitopes of the pseudo wild type and of the most shaved  $\mathbf{A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}}$  mutant is shown (Fig. 1). In this mutant, the total contact area with both trypsin and chymotrypsin is reduced by about 22% compared with the respective pseudo wild-type complexes (Fig. 6). In all modeled trypsin complexes Arg 15 remains 100% buried, as it is in the pseudo wild-type complex. Arg 15 in chymotrypsin complexes, including the pseudo wild type, is 92% buried, and exposes about 15  $Å^2$  to the solvent. Thus, in all trypsin and chymotrypsin complexes, interactions involving Arg 15 should not be, in principle, affected by alanine shaving. Also, the relative contribution of Arg 15 is larger in more extensively mutated variants. For example, it provides 16% of the total contact area in the wild-type complex with trypsin and 20% of the total area in the  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$ trypsin complex.

Thermal stability curves recorded at three pH values provided thermodynamic parameters,  $\Delta H_{den}$ ,  $\Delta C_{p,den}$ , and  $T_{den}$ for all mutants. The  $\Delta H_{den}$  versus  $T_{den}$  plots for individual mutants did not provide a single but a series of (roughly) parallel lines, in a marked difference to previously analyzed P<sub>1</sub> mutants of BPTI (Krowarsch and Otlewski 2001). Thus, to compare their stability, we turned to  $\Delta G_{den}$  values. Arrangement of  $\Delta G_{den}$  values in five thermodynamic additivity cycles showed that all the effects of substitutions are fully independent of each other and, therefore, additive. In the free inhibitor, there is a hydrogen bond Thr 11  $O^{\gamma}$ -Val 34 O that involves a side chain of a mutated residue and may cause nonadditive stability behavior (Wells 1990). However, elimination of this bond as a result of the Thr11Ala substitution had a marginal effect on stability of the respective variants and was fully additive (cycles B and C).

We also applied a simpler method of additivity estimation, suggested by Kuroda and Kim (2000), through comparison of respective  $\Delta T_{den}$  values at a common pH, rather



Fig. 5. Thermodynamic stability additivity cycles for alanine substitutions in BPTI. All  $\Delta G_{den}$  values were calculated at common a temperature of 25°C using the Gibbs-Helmholtz equation.  $\Delta\Delta G_{den}$  is equal to  $\Delta G_{den, \text{ mutant with less Ala}} - \Delta G_{den, \text{ mutant with more Ala}}$  and describes the stability effect caused by introduction of alanine(s). Similarly as defined for protein–protein additivity,  $\Delta\Delta\Delta G_{den}$  is a difference of two  $\Delta\Delta G_{den}$  values and determines whether the effect of the substitution is additive.

than through  $\Delta G_{den}$  (data not shown). This method did not allow us to close cycles D and E.

Data obtained on interaction of 10 alanine mutants with chymotrypsin enable more extensive calculations of additivity effects, also including data for the single alanine mutants described by Castro and Anderson (1996). It should be stressed that our data for all three single mutants ( $A^{13}R^{15}$ ,  $R^{15}A^{17}$ , and  $R^{15}A^{34}$ ) agree very well with alanine scanning data justifying joint analysis of two data sets. Table 4 shows that  $\Delta G_{a,calc}$  values estimated from single mutation effects agree very well with experimental  $\Delta G_a$ s for seven mutants. However, for three mutants:  $A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$ ,

 $R^{15}A^{34}A^{37}A^{39}$ , and  $A^{13}R^{15}A^{17}A^{18}A^{19}$ , the discrepancies from additivity excess allowed error by -1.25, -0.62, and -0.3 kcal mole<sup>-1</sup>, respectively. It is intriguing that these deviations occur just for these three mutants and not for the more extensive ones. Lack of respective structures for inhibitor mutants and their complexes with chymotrypsin does not allow more detailed interpretation of this phenomenon.

The association constant for the  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}$  $A^{34}A^{37}A^{39}$  BPTI–chymotrypsin interaction is  $6.6 \times 10^5$  M<sup>-1</sup> and agrees extremely well with the value calculated from single mutation effects. The same agreement is observed for



**Fig. 6.** Contact area values for the complexes between individual alanine mutants of BPTI and trypsin (*A*) or chymotrypsin (*B*). Total contact area (black bar) and contact area of enzyme (gray) and inhibitor (light gray) components are shown.

mutant containing six alanines  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}$ . We conclude that these two extensively mutated variants fulfill a very strong additivity test.

It could be assumed that the  $K_a$  values for  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  BPTI interacting with both enzymes should be equal to the contribution of Arg 15 (the  $P_1$  residue) to the total energy of association. In the case of chymotrypsin, this contribution was calculated to be 4.7 kcal mole<sup>-1</sup> from the effect of the Arg15Gly substitution on the binding energy (Krowarsch et al. 1999). Thus, we observed a large gap between the two values, equivalent to 3 kcal mole<sup>-1</sup>. This means that the association energy for the pseudo wild-type BPTI cannot be reproduced from a sum of single residue mutational effects. There are two possible reasons for this discrepancy. First, we suppose that the alanine background surrounding the  $P_1$  site might not be fully

neutral and might still provide some binding energy as a result of the presence of numerous methyl groups of alanines. For example, we observe that the Ala15Gly mutation removes an additional1.9 kcal mole<sup>-1</sup> of binding energy with chymotrypsin (Krowarsch et al. 1999), and the analogous mutation at  $P_1'$  in BPTI removes 1.5 kcal mole<sup>-1</sup> (Grzesiak et al. 2000a). This residual energy could easily be removed if we performed glycine shaving analysis. However, it appears to be a formidable task because of the much higher susceptibility to proteolysis of glycine mutants. Also, introduction of glycines would significantly increase the conformational entropy of the contact loops complicating the interpretation of the binding data.

The second possible reason for the observed discrepancy is related to the presence of the main chain-main chain intermolecular hydrogen bond network, which is well conserved among different proteinase-inhibitor complexes (Jackson 1999). In our opinion, preservation of the binding mode is the main reason for the high nonadditivity described above. The total energetic contribution of these hydrogen bonds remains unknown, but should be substantial. A study on the chemical variant of the turkey ovomucoid third domain in which the  $P_2-P_1$  amide bond has been replaced with an ester shows that a single intermolecular hydrogen bond provides 0.8-2.0 kcal mole<sup>-1</sup> of free energy, depending on a proteinase (Lu et al. 1997). Thus, the retained 3 kcal mole<sup>-1</sup> can come from binding energy provided by the methyls of alanines and/or by intermolecular backbone hydrogen bonds.

Because of the extremely strong electrostatic interaction, analysis of additivity and hot spot energetics is more complicated in the case of the trypsin interaction.  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  BPTI binds trypsin with a  $K_a$  of  $3.4 \times 10^6$  M<sup>-1</sup>, and the contribution of Arg 15 estimated from the effect of the Arg15Gly substitution on binding energy with this proteinase approaches -12.7 kcal mole<sup>-1</sup> (Krowarsch et al. 1999; Helland et al. 1999). Thus, this fully shaved variant binds to trypsin 3.95 kcal mole<sup>-1</sup> weaker than expected from the contribution of Arg 15 to the association energy. Interestingly, as the effect is opposite to the discrepancy described above for chymotrypsin, we suppose that, when corrected for the presence of alanine methyls rather than glycines and main chain hydrogen bonds, the total disagreement can be about 3 kcal mole<sup>-1</sup> larger, that is, up to about 7 kcal mole<sup>-1</sup>. This indicates that in  $A^{11}A^{13}R^{15}A^{17}A^{18}A^{19}A^{34}A^{37}A^{39}$  BPTI even 45% of the Arg 15 contribution can be destroyed. Clearly, this points to large nonadditivity comprising the  $P_1$  side chain.

This nonadditivity can also be shown through the following reasoning. Castro and Anderson (1996) reported one single alanine substitution effect: The IIe18Ala mutation decreases  $\Delta G_a$  by 4.94 kcal mole<sup>-1</sup>. We measured the effects of four single alanine substitutions: Thr11Ala (0 kcal mole<sup>-1</sup>), Pro13Ala (2.7 kcal mole<sup>-1</sup>), Arg17Ala (0.2 kcal mole<sup>-1</sup>) and Val34Ala (0 kcal mole<sup>-1</sup>). The former effect is calculated on the basis of the additivity assumption, which works in all three cycles (A, B, and C) including the single Thr11Ala substitution. The sum of all these effects (-7.84 kcal mole<sup>-1</sup>) subtracted from  $\Delta G_a$  for the pseudo wild type (-18.3 kcal mole<sup>-1</sup>) equals -10.86 kcal mole<sup>-1</sup>. This value is already 2 kcal mole<sup>-1</sup> higher than expected from Arg15Gly (-12.7 kcal mole<sup>-1</sup>) and still does not cover effects from all contact residues.

We suppose that the nonadditivity observed in the F and G cycles and particularly large nonadditivity observed in the D and E cycles results from weakening of the electrostatic  $P_1-S_1$  interaction. It can be speculated that the nonadditivity observed in these four cycles may result from different perturbations of the electrostatic interaction caused by various numbers of alanine substitutions introduced in the individual mutants. On the other hand, the Thr11Ala substitution does not seem to influence the strength of this interaction, as the A, B, and C cycles are fully additive.

Interestingly, lack of additivity connected with the P<sub>1</sub> electrostatic interaction has been also observed for the ecotin–trypsin system (Yang et al. 1998). In this case, the interaction is extremely weak, as the effect of the substitutions at this position on  $K_a$  is only several fold, compared with 10<sup>6</sup>-fold in BPTI–trypsin system (Krowarsch et al. 1999). In ecotin, the P<sub>1</sub>–S<sub>1</sub> interaction appears to be highly coupled and nonadditive with interactions formed by two remote secondary binding loops that take over the majority of the association energy and relax the loop comprising the P<sub>1</sub> site.

# Materials and methods

#### Materials

Guanidinium chloride (GdmCl), urea, dimethylsulfoxide (DMSO), *N*,*N*-dimethylformamide (DMF), methanol, and acetonitrile were purchased from Merck. Trifluoroacetic acid (TFA), formic acid (88%) and cyanogen bromide (CNBr) were obtained from Fluka. Tris, sodium acetate, DTT, chloramphenicol, ampicillin, reduced and oxidized glutathione (GSH and GSSG), and 2–[[N-morpho-lino]]ethanesulfonic acid (MES) were supplied by Sigma. Sephadex G-25 was obtained from Amersham Pharmacia Biotech.

The components of the culture media were purchased from Merck. DNA-modifying enzymes (T4 DNA polymerase, T4 DNA ligase, and T4 polynucleotide kinase) and the DNA sequencing kit was from Amersham Pharmacia Biotech, and the DNA purification kit was from Qiagen. Oligonucleotides were synthesized by Ransom Hill. IPTG was obtained from Bachem.

Bovine trypsin and bovine chymotrypsin were purchased from Worthington. 4-Nitrophenyl 4'-guanidinobenzoate (NPGB) was from Merck and N-Bz-L-Arg-4-nitroanilide (Bz-Arg-pNA) was from Sigma. N-suc-Gly-Gly-L-Phe-4-nitroanilide and N-suc-Ala-Ala-Pro-L-Phe-4-nitroanilide were purchased from Bachem. N-suc-L-Phe-4-nitroanilide, N-p-tosyl-Gly-Pro-L-Arg-4-nitroanilide and N-Bz-Pro-Phe-L-Arg-4-nitroanilide were from Sigma.

## Plasmid construction

Single or multiple alanine substitutions were introduced individually into the gene coding for Lys15Arg, Met52Leu BPTI by oligonucleotide-directed mutagenesis (Kunkel 1985), and the introduced mutations were confirmed by DNA sequencing.

#### Expression and purification of BPTI mutants

All 10 alanine mutants of BPTI and the pseudo wild type (Lys15Arg, Met52Leu) were overexpressed in *Escherichia coli* strain BL 21 (DE3) pLysS by use of the T7 polymerase promoter system (Studier et al. 1990). Purification of BPTI mutants was performed as described (Krokoszynska et al. 1998). Protein samples were >96% pure, as judged from C<sub>18</sub> reverse-phase HPLC runs. A typical yield of in vitro oxidative refolding was 30%–80% (1–3 mg of pure protein per liter of culture medium). Electrospray mass spectrometry of all mutants used in this study was performed with a Finningan MAT TSQ-700 equipped with an ESI source. Samples were dissolved in methanol/water (1:1, vol/vol) containing 1% acetic acid. Molecular masses of all mutants were within 1.0 atomic mass unit from those expected from amino acid sequences.

#### Association constant determination

The association constant  $(K_a)$  values were determined as described by Empie and Laskowski (1982) and Otlewski and Zbyryt (1994). All measurements were performed in 100 mM Tris-HCl, 20 mM CaCl<sub>2</sub>, 0.005% Triton X-100 (pH 8.2) or in 100 mM MES, 20 mM CaCl<sub>2</sub>, 0.005% Triton X-100 (pH 5.5) at 22°C. To determine an association constant increasing amounts of the inhibitor were added to a constant concentration of a given proteinase. The initial enzyme concentration  $(E_0)$  was chosen so that the condition 1 < 1 $[E_o] \cdot K_a < 50$  was fulfilled. For the determination of low  $K_a$  values  $(10^3 - 10^7 \text{ M}^{-1})$  inhibitors were used at much higher concentrations than enzymes to force complex formation. After  $8-10t_{1/2}$  incubation time, the residual enzyme activity was measured by monitoring linear release of *p*-nitroaniline for the period of 300 sec using a diode array spectrophotometer HP 8452A. The average value of absorbance in the 380-410 nm range was recorded and corrected for the background average absorbance in the 480-510 nm range. The value of  $K_a$  was calculated by a three-parameter algorithm  $(E) = f(E_0, K_a, F)$  using the nonlinear regression analysis program GraFit, according to the equation

$$[E] = \frac{1}{2} \cdot ([E_0] - F \cdot [I_0] - K_a^{-1} + \sqrt{([E_0] + F \cdot [I_0] + K_a^{-1})^2 - 4 \cdot [E_0] \cdot F \cdot [I_0]})$$
(2)

where  $[E_o]$  and  $[I_o]$  are the total enzyme and inhibitor concentrations, respectively, [E] is the residual enzyme concentration, and F is the enzyme:inhibitor equimolarity factor. In the case of weak associations ( $K_a < 10^7 \text{ M}^{-1}$ ), the default value of F = 1 was applied.

#### Calorimetry

Differential scanning calorimetry (DSC) experiments were performed with a Nano II calorimeter (CSC Corp.). The experiments were carried out at protein concentrations of  $100-200 \ \mu g \ mL^{-1}$  (total cell volume 323 µL), at a scan rate of 1.0 deg min<sup>-1</sup>, under a pressure excess of 2.5 atm. Before a DSC run, the protein solution was extensively dialyzed against 10 mM Gly/HCl (pH 2.0). Repeated measurements of the same sample demonstrated ~70% reproducibility of  $\Delta H_{cal}$  values. The baseline was determined by running the calorimeter with both cells filled with the dialysis buffer. Experiments were performed in triplicate. The partial specific volume of BPTI was assumed to be 0.71 cm<sup>3</sup> g<sup>-1</sup> (Makhatadze et al. 1993).  $T_{den}$ ,  $\Delta H_{cal}$  and  $\Delta H_{vH}$  values were calculated from thermogram analysis using Cpcalc software provided by CSC Corp.

#### CD measurements

#### Spectra

CD spectra of BPTI variants were recorded in 10 mM Gly/HCl buffer (pH 2.0), at 25°C on a Jasco J-715 spectropolarimeter. Measurements were performed at a protein concentration of  $2 \times 10^{-5}$  M using a 10-mm cuvette (240–360 nm range) or at  $2 \times 10^{-4}$  M protein in a 1-mm pathlength cuvette (190–260 nm range). Spectra were averaged from seven separate scans.

#### Thermal transitions

Thermal denaturation was monitored following the ellipticity at 223 nm using a band slit of 2 nm and a 4-sec response time. An automatic Peltier accessory PFD 350S allowed continuous monitoring of the thermal transition at a constant rate of 1 deg min<sup>-1</sup>. The temperature of the sample was monitored directly using a probe immersed in the cuvette and controlled with a PFD-350S/350L Peltier type FDCD attachment. Protein was dissolved in 10 mM Gly/HCl buffer, (pH 2.0, 2.5, and 2.7). Before measurement, a protein solution was passed through a 0.22– $\mu$ m Millipore filter. Each denaturation experiment was performed in duplicate. The data were analyzed assuming a two-state reversible equilibrium transition:

$$\theta(T) = \frac{\left\{ \exp\left[\frac{\Delta H_{\nu H}}{R} \cdot \left(\frac{1}{T_{den} + 273.15} - \frac{1}{T + 273.15}\right)\right] \right\}}{\left\{ 1 + \exp\left[\frac{\Delta H_{\nu H}}{R} \cdot \left(\frac{1}{T_{den} + 273.15} - \frac{1}{T + 273.15}\right)\right] \right\}}$$
(3)

where *T* is the temperature, *R* is the gas constant,  $\theta(T)$  is the ellipticity signal at 223 nm;  $\Delta H_{\nu H}$  is the van't Hoff enthalpy,  $\theta_F$  is the value of the folded signal extrapolated to -273.15 °C,  $m_F$  is the slope of the temperature dependence of the CD signal for the folded protein,  $\theta_U$  is the value of the unfolded signal extrapolated to -273.15 °C,  $m_U$  is the slope of the temperature dependence of the CD signal for the folded protein,  $\theta_U$  is the slope of the temperature dependence of the CD signal for the unfolded protein, and  $T_{den}$  is denaturation temperature.

#### Acknowledgments

This work was supported by grant 6 PO4B 014 17 from the Polish Committee for Scientific Research. We thank Professor Zbigniew Szewczuk for mass spectra analysis and Marta Gladysz for help with mutant refolding. Daniel Krowarsch is a recipient of the Young Scholar Award from the Foundation for Polish Science. The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

#### References

Addlagatta, A., Krzywda, S., Czapinska, H., Otlewski, J., and Jaskolski, M. 2001. Ultra high resolution structure of BPTI. Acta Cryst. D57: 649–663.

- Bode, W. and Huber, R. 1992. Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.* 204: 433–451.
- Bogan, A.A. and Thorn, K.S. 1998. Anatomy of hot spots in protein interfaces. J. Mol. Biol. 280: 1–9.
- Capasso, C., Rizzi, M., Menegatti, E., Ascenzi, P., and Bolognesi, M. 1997. Multiple protein–protein recognition sites in the crystal structure of α-chymotrypsin in complex with basic pancreatic trypsin inhibitor. J. Mol. Recogn. 10: 26–35.
- Castro, M.J. and Anderson, S. 1996. Alanine point-mutations in the reactive region of bovine pancreatic trypsin inhibitor: Effects on kinetics and thermodynamics of binding to β-trypsin and α-chymotrypsin. *Biochemistry* 35: 11435–11446.
- Czapinska, H., Otlewski, J., Krzywda, S., Sheldrick, G.M., and Jaskólski, M. 2000. High resolution structure of bovine pancreatic trypsin inhibitor with altered binding loop sequence. J. Mol. Biol. 295: 1237–1249.
- Empie, M.W. and Laskowski, Jr., M. 1982. Thermodynamics and kinetics of single residue replacements in avian ovomucoid third domains: Effect on inhibitor interactions with serine proteinases. *Biochemistry* 21: 2274–2284.
- Grzesiak, A., Helland, R., Smalås, A.O., Krowarsch, D., Dadlez, M., and Otlewski, J. 2000a. The P<sub>1</sub>' position in BPTI is a major determinant of the association energy with serine proteinases. J. Mol. Biol. 301: 207–218.
- Grzesiak, A., Krokoszynska, I., Krowarsch, D., Buczek, O., Dadlez, M., and Otlewski, J. 2000b. Inhibition of six serine proteinases of human coagulation system by mutants of bovine pancreatic trypsin inhibitor. *J. Biol. Chem.* 275: 33346–33352.
- Helland, R., Otlewski, J., Sundheim, O., Dadlez, M., and Smalas, A.O. 1999. The crystal structures of the complexes between bovine β-trypsin and ten P1 variants of BPTI. J. Mol. Biol. 287: 923–942.
- Hubbard, S.J. and Thornton, J.M. 1993. A computer algorithm to calculate surface accessibility. Department of Biochemistry and Molecular Biology, University College, London.
- Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., and Steigemann, W. 1974. Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 89: 73–101.
- Jackson, R.M. 1999. Comparison of protein–protein interaction in serine proteinase–inhibitor and antibody–antigen complexes: Implications for the protein docking problem. *Protein Sci.* 8: 603–613.
- Jin, L. and Wells, J.A. 1994. Dissecting the energetics of an antibody-antigen interface by alanine shaving and molecular grafting. *Protein Sci.* 3: 2351– 2357.
- Jones, S. and Thorton, J.M. 1996. Principles of protein-protein interactions. Proc. Natl. Acad. Sci.93: 13–20.
- Krokoszynska, I., Dadlez, M., and Otlewski, J. 1998. Structure of single-disulphide variants of bovine pancreatic trypsin inhibitor (BPTI) as probed by their binding to bovine β-trypsin. J. Mol. Biol. 275: 503–513.
- Krowarsch, D. and Otlewski, J. 2001. Amino-acid substitutions at the fully exposed P<sub>1</sub> site of bovine pancreatic trypsin inhibitor affect its stability. *Protein Sci.* 10: 715–724.
- Krowarsch, D., Dadlez, M., Buczek, O., Krokoszynska, I., Smalas, A.O., and Otlewski, J. 1999. Interscaffolding additivity: Binding of P<sub>1</sub> variants of bovine pancreatic trypsin inhibitor to four serine proteases. *J. Mol. Biol.* 289: 175–186.
- Kunkel, T.A. 1985. Rapid and efficient site-specific mutagenesis without phenotopic selection. Proc. Natl. Acad. Sci. 82: 488–492.
- Kuroda, Y. and Kim, P.S. 2000. Folding of bovine pancreatic trypsin inhibitor (BPTI) variants in which almost half the residues are alanine. J. Mol. Biol. 298: 493–501.
- LiCata, V.J. and Ackers, G.K. 1995. Long-range, small magnitude non-additivity of mutational effects in proteins. *Biochemistry* 34: 3133–3139.
- Lo Conte, L., Chothia, C., and Janin, J. 1999. The atomic structure of protein– protein recognition sites. J. Mol. Biol. 285: 2177–2198.
- Lu, W., Qasim, M.A., Laskowski, Jr., M., and Kent, S.B.H. 1997. Probing intermolecular main chain hydrogen bonding in serine proteinases-protein inhibitor complexes: chemical synthesis of backbone-engineered turkey ovomucoid third domain. *Biochemistry* 36: 673–679.

- Lu, S.M., Lu, W., Qasim, M.A., Anderson, S., Apostol, I., Ardelt, W., Bigler, T., Chiang, Y.W., Cook, J., James, M.N.G., et al. 2001. Predicting the reactivity of proteins from their sequence alone. Kazal family of protein inhibitors of serine proteinases. *Proc. Natl. Acad. Sci.* 98: 1410–1415.
- Makhatadze, G.I., Kim, K., Woodward, C. and Privalov, P.L. 1993. Thermodynamics of BPTI folding. *Protein Sci.* 2: 2028–2036.
- McDonald, I. K. and Thornton, J. M. 1994. Satisfying hydrogen bonding potential in proteins. J. Mol. Biol. 238: 777–793.
- Otlewski, J. and Apostoluk, W. 1997. Structural and energetic aspects of protein-protein recognition. Acta Biochim. Polon. 44: 367-388.
- Otlewski, J. and Zbyryt, T. 1994. Single peptide bond hydrolysis/resynthesis in squash inhibitors of serine proteinases. I. Kinetics and thermodynamics of the interaction between squash inhibitors and bovine β-trypsin. *Biochemistry* 33: 200–207.
- Qasim, M.A., Ganz, P.J., Saunders, C.W., Bateman, K.S., James, M.N.G., and Laskowski, M.Jr 1997. Interscaffolding additivity. Association of P<sub>1</sub> variants of eglin c and ovomucoid third domain with serine proteinases. *Biochemistry* 36: 1598–1607.
- Schechter, P. and Berger, A. 1967. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. 27: 157–162.

Scheidig, A.J., Hynes, T.R., Pelletier, L.A., Wells, J.A., and Kossiakoff, A.A.

1997. Crystal structures of bovine chymotrypsin and trypsin complexed to the inhibitor domain of Alzhaimer's amyloid  $\beta$ -precursor (APPI) and basic pancreatic trypsin inhibitor (BPTI): Engineering of inhibitors with altered specificities. *Protein Sci.* **6:** 1806–1824.

- Schreiber, G. and Fersht, A.R. 1995. Energetics of protein–protein interactions: Analysis of the barnase-barstar interface by single mutations and double mutant cycles. J. Mol. Biol. 248: 478–486.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 polymerase to direct expression of cloned genes. *Methods Enzymol.* 185: 60–89.
- Wells, J.A. 1990. Additivity of mutational effects in proteins. *Biochemistry* 29: 8509–8517.
- ——. 1991. Systematic mutational analyses of protein-protein interfaces. Methods Enzymol. 202: 390–411.
- Yang, S.Q., Wang, C.-I., Gillmor, S.A., Fletterick, R.J., and Craik, C.S. 1998. Ecotin: A serine protease inhibitor with two distinct and interesting binding sites. J. Mol. Biol. 279: 945–957.
- Yu, M.-H., Weissman, J.S., and Kim, P.S. 1995. Contribution of individual side-chains to the stability of BPTI examined by alanine-scanning mutagenesis. J. Mol. Biol. 249: 388–397.