Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structure of a proteolytically resistant analogue of (*N*Lys)^{5}SFTI-1 in complex with trypsin: evidence for the direct participation of the Ser214 carbonyl group in serine protease-mediated proteolysis

Szymon Krzywda,[a] Mariusz Jaskolski,[b,c] Krzysztof Rolka[d] and Maciej J. Stawikowski[e]*

[a]Department of Chemistry, Adam Mickiewicz University, Grunwaldzka 6, 60-780 Poznan, Poland, [b]Department of Crystallography, Faculty of Chemistry, Adam Mickiewicz University, Grunwaldzka 6, 60-780 Poznan, Poland, [c]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland, [d]Faculty of Chemistry, University of Gdansk, Stwosza 63, 80-952 Gdansk, Poland, and [e]Torrey Pines Institute For Molecular Studies, 11350 SW Village Parkway, Port St Lucie, FL 34987, USA

Correspondence e-mail: mstawikowski@tpims.org

Received 8 October 2013
Accepted 26 November 2013

PDB reference: bovine trypsin, complex with peptide–peptoid inhibitor, 4hgc

Peptide–peptoid hybrids are found to be potent inhibitors of serine proteases. These engineered peptidomimetics benefit from both types of units of the biopolymeric structure: the natural inhibitor part serves as a good binding template, while the P1-positioned peptoid component provides complete resistance towards proteolysis. In this report, the mechanism of proteolytic resistance of a P1 peptoid-containing analogue is postulated based on the crystal structure of the (*N*Lys)^5-modified sunflower trypsin inhibitor SFTI-1 in complex with bovine trypsin solved at 1.29 Å resolution. The structural differences between the (*N*Lys)^{5}SFTI-1–trypsin complex and the native SFTI-1–trypsin complex are surprisingly small and reveal the key role of the carbonyl group of the Ser214 residue of the enzyme, which is crucial for binding of the inhibitor and plays a crucial role in proteolysis mediated by serine proteases. The incorporated *N*Lys5 peptoid residue prevents Ser214 from forming a hydrogen bond to the P1 residue, and in turn Gln192 does not form a hydrogen bond to the carbonyl group of the P2 residue. It also increases the distance between the Ser214 carbonyl group and the Ser195 residue, thus preventing proteolysis. The hybrid inhibitor structure reported here provides insight into protein–protein interaction, which can be efficiently and selectively probed with the use of peptoids incorporated within endogenous peptide ligands.

1. Introduction

Serine proteases account for over one-third of all known proteolytic enzymes (Di Cera, 2009). The MEROPS classification system groups proteases into clans that typically have structural homology and/or the same linear order of catalytic triad residues (Rawlings *et al.*, 2010). Clan PA proteases, bearing the trypsin fold, are the largest family of serine proteases and are perhaps the best studied group of enzymes. Nearly all clan PA proteases utilize the canonical catalytic triad (namely His57, Asp102 and Ser195) and hydrolyze the peptide bond *via* two tetrahedral transition states (Hartley & Kilby, 1954). Aside from their various physiological roles, the unique specificities of these enzymes make them useful tools in biochemistry and molecular biology. Being cost-effective and easily available, the bovine trypsin enzyme represents an excellent research tool when it comes to studying the enzymology of serine proteases.

The development of new peptidomimetics with improved pharmacological properties is an important goal in drug discovery. Polypeptoids (poly-*N*-substituted glycine) are an important class of peptidomimetics. Since their introduction in 1992 by Bartlett and coworkers (Simon *et al.*, 1992), peptoids have been gaining momentum in drug development (Patch *et*

al., 2005). Their unique chemical properties together with their ease of synthesis make them an attractive object of research. Peptoid monomers are linked through polyimide bonds and, owing to the absence of an H atom at the imide N atom, are incapable of forming stable secondary structures that are governed by hydrogen bonds. Although the polypeptoid backbone is achiral, it is possible to introduce chirality into the side-chain functionalities and thus to obtain stable compounds with desired folding properties (foldamers; Hill *et al.*, 2001). Since peptoids are of modular structure, one could combine them into peptide–peptoid hybrid architectures. Such hybrid polymers were first introduced by Goodman and coworkers (Goodman *et al.*, 1994), and there are many examples of their application (Olsen, 2010). Unlike most peptides, polypeptoids are resistant to proteolysis. Moos and coworkers reported complete proteolytic resistance of a series of peptoid oligomers towards metallo, cysteine, aspartyl and serine proteases (Miller *et al.*, 1994). SFTI-1 is the smallest naturally occurring serine protease inhibitor in the Bowman–Birk inhibitor family (BBI; Luckett *et al.*, 1999). It is comprised of only 14 amino-acid residues, forming a cyclic peptide additionally stabilized by a disulfide bridge between Cys3 and Cys11 (Fig. 1). The rigid β-hairpin-like structure of SFTI-1 arises not only from the disulfide-bridge stabilization but also from a network of intramolecular hydrogen bonds and a *cis*-peptide bond between the Ile7 and Pro8 residues. Its small size and potent activity against many serine proteases, namely matriptase (K_i = 0.92 n*M*; Long *et al.*, 2001), trypsin (K_i = 0.1 n*M*), cathepsin G (K_i < 0.15 n*M*), chymotrypsin (K_i = 7.4 μ*M*), elastase ($K_i \simeq$ 105 μ*M*) and thrombin ($K_i \simeq$ 136 μ*M*) (Luckett *et al.*, 1999), make this inhibitor an excellent template for the design of new inhibitors with improved biological and pharmacological properties. SFTI-1 belongs to the so-called canonical serine protease inhibitors, which bind to the enzyme through an exposed convex binding loop that is complementary to the active site of the enzyme. The mechanism of inhibition in this group is always very similar and resembles that of an ideal substrate (Krowarsch *et al.*, 2003).

Interestingly, the incorporation of *N*-substituted glycine into the structure of naturally occurring peptide inhibitors also results in complete resistance to proteolysis of such peptide–peptoid hybrids, as reported by Stawikowski *et al.* (2005). The proteolytic resistance of peptoids has inspired us to incorporate their building blocks into the sunflower trypsin inhibitor (SFTI-1) sequence at the P1 position. The substitution retained most of the SFTI-1 activity and made the P1–P1′ bond resistant to proteolysis (Stawikowski *et al.*, 2005). This observation was very interesting, since upon the introduction of the peptoid residue at the P1 position the (scissile) P1–P1′ bond remains a regular peptide bond while the P1–P2 bond becomes the imide bond. In other words, cleavage should occur at the C-terminus of the peptoid moiety, but does not. This intriguing fact motivated us to determine the crystal structure of the (*N*Lys)^{5}SFTI-1 inhibitor in complex with bovine trypsin in order to elucidate the structural basis of this phenomenon. Such a structure is presented here.

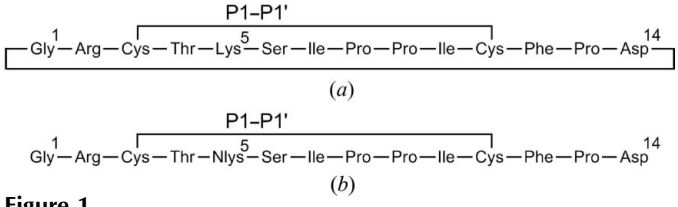


Figure 1
Chemical formulae of (*a*) native SFTI-1 and (*b*) [*N*Lys]^{5}SFTI-1.

2. Materials and methods

2.1. Peptide–peptoid hybrid synthesis

The (*N*Lys)^{5}SFTI-1 inhibitor was synthesized manually using Fmoc/tBu chemistry and the DIC/HOBt coupling procedure as described previously (Stawikowski *et al.*, 2005). The C-terminal amino acid (Asp) was attached to chlorotrityl polystyrene resin (Rapp Polymere, Germany) by using 1.5 equiv molar excess of Fmoc-Asp(OtBu) in the presence of *N*,*N*-diisopropylethylamine. Incorporation of the peptoid residue *N*Lys was achieved by the submonomeric method. Briefly, bromoacetic acid (10 equiv) and *N*,*N*-diisopropylcarbodiimide (10 equiv) were used to acetylate the peptidyl-resin. The side chain of *N*Lys was incorporated using 1,4-diaminobutane (10 equiv) and trityl chloride (1.1 equiv) was used to protect the ε-amino group of the *N*Lys residue. The subsequent peptide residue was coupled using 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) as the coupling reagent. The progress of incorporation of the peptoid residue and incorporation of the Trt protecting group was monitored by MALDI-TOF analysis along with the Kaiser and chloranil tests. After completion of the synthesis, the peptide–peptoid hybrid was cleaved from the resin and protecting groups were removed in one step using TFA/phenol/triisopropylsilane/H_2O (88:5:2:5, *v*:*v*:*v*:*v*). The crude linear analogue was oxidized in a mixture of 5% AcOH and 20% DMSO, pH 6 (adjusted with ammonium carbonate) with gentle stirring for 24 h, (Tam *et al.*, 1991) and subsequently desalted by solid phase extraction on C18 LiChrolut SPE columns (Merck, Germany). The final compound was purified by RP-HPLC on a Beckman Gold System chromatograph (Beckman, USA) using Kromasil-100 and a C8 column (8 × 250 mm, Knauer, Germany). The solvent system was (*A*) 0.1% TFA in water, and (*B*) 80% acetonitrile in *A*. A linear gradient from 20 to 80% *B* at 1.5 ml min^{-1} for 30 min and detection at 220 nm were used.

2.2. Crystallization of the (*N*Lys)^{5}SFTI-1–trypsin complex

Commercially available bovine trypsin (Sigma) was dissolved in a buffer consisting of 60 m*M* benzamidine, 1 m*M* calcium chloride, 50 m*M* MES pH 6.0 to a final concentration of 30 mg ml^{-1} without prior purification. The unsolubilized fraction was removed by centrifugation. Single crystals of the enzyme were grown within one week using the hanging-drop vapour-diffusion technique at 292 K by mixing 2 µl trypsin with 2 µl reservoir solution consisting of 2.3 *M* ammonium sulfate, 1 m*M* benzamidine, 0.1 *M* MES pH 6.0. As previously

Table 1
X-ray data-collection and model-refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Crystal dimensions (mm)	0.2 × 0.2 × 0.7
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 60.12, b = 64.14, c = 70.04
X-ray source	X11, EMBL/DESY
Wavelength (Å)	0.8148
Temperature (K)	100
Mosaicity (°)	0.69
Oscillation angle (°)	0.5
No. of images	360
Crystal-to-detector distance (mm)	100
Resolution (Å)	37.17–1.29 (1.36–1.29)
$R_{\rm int}$† (%)	5.7 (80.4)
$\langle I/\sigma(I)\rangle$	16.5 (2.0)
Reflections	
Measured	488927
Unique	67689
Multiplicity	7.2 (6.8)
Completeness (%)	99.4 (98.6)
Wilson B factor (Å^2)	19.74
Refinement	
Resolution (Å)	30.74–1.29
No. of reflections in working set	65592
No. of reflections in test set	2036
$R_{\rm work}/R_{\rm free}$‡ (%)	12.62/15.67
No. of atoms	
Protein	1751
Solvent	287
B factors (Å^2)	
Protein	18.64
Solvent	30.03
R.m.s.d. from ideal	
Bond lengths (Å)	0.018
Bond angles (°)	1.98
Ramachandran statistics (%)	
Favoured	97.8
Allowed	2.2
PDB code	4hgc

† $R_{\rm int} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith measurement of the intensity of reflection hkl and $\langle I(hkl)\rangle$ is the mean intensity of reflection hkl. ‡ $R = \sum_{hkl}||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are the observed and calculated structure-factor amplitudes, respectively.

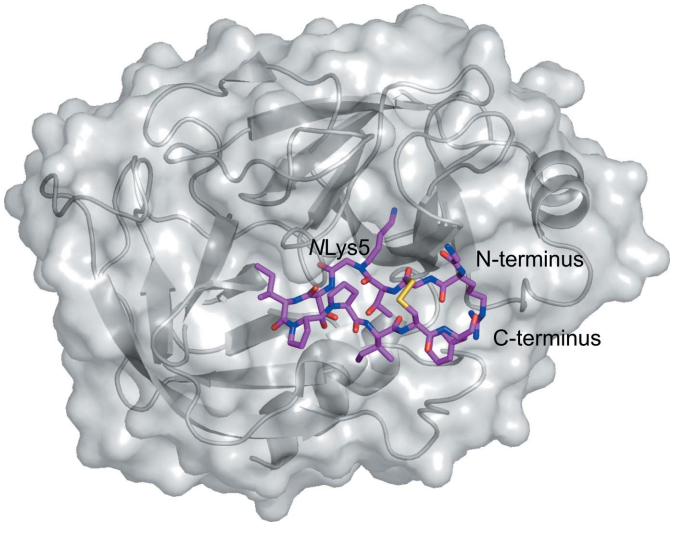


Figure 2
Overall structure of (*N*Lys)^{5}SFTI-1 in complex with bovine trypsin. The enzyme surface is coloured grey; the inhibitor molecule bound to the active site is shown in magenta.

observed (Bartunik *et al.*, 1989), trypsin crystals of low and high density appear in the same drops. Although both forms are morphologically similar, the low-density crystals grew to much larger dimensions, reaching 0.4 × 0.4 × 1.4 mm. To remove benzamidine, the low-density crystals were transferred to a well containing 2.3 *M* ammonium sulfate, 1 m*M* calcium chloride, 0.1 *M* MES buffer pH 6.0. The washing procedure was repeated three times over a period of 24 h. Finally, lyophilized (*N*Lys)^{5}SFTI-1 inhibitor was added until saturation (as marked by an undissolved residue). The soaking process took several hours. For cryoprotection, the inhibitor-soaked crystal was transferred to a solution consisting of the final soaking solution supplemented with 25%(*v*/*v*) glycerol.

2.3. X-ray diffraction data collection

Diffraction data were collected at 100 K on the EMBL beamline X11 at the Deutsches Elektronen-Synchrotron (EMBL c/o DESY, Hamburg). Diffraction data were processed using *MOSFLM* (Leslie & Powell, 2007) and scaled with *SCALA* (Evans, 2006). Crystal parameters and data-collection and processing statistics are summarized in Table 1.

2.4. Structure solution and refinement

The structure was solved by molecular replacement using *MOLREP* (Vagin & Teplyakov, 2010) from the *CCP*4 suite (Winn *et al.*, 2011) with the structure of the SFTI-1–trypsin complex (PDB entry 1sfi; Luckett *et al.*, 1999) as the search model. The amino-acid sequence of the present structure is identical (except for the *N*Lys5 residue) to that of the 1sfi model. Maximum-likelihood restrained structure refinement was carried out in *REFMAC* (Murshudov *et al.*, 2011) using all intensity data, with the exception of 2036 reflections (3.0%) flagged for $R_{\rm free}$ testing. No σ cutoff was applied. To account for diffuse solvent effects, a correction according to the Babinet principle was applied (Moews & Kretsinger, 1975). The stereochemical restraints for the nonstandard *N*Lys residue were generated using the *CCP*4 software. A two-dimensional diagram of the *N*Lys ligand was drawn and converted into a structure and a restraint set. Manual rebuilding of the model was performed in *Coot* (Emsley & Cowtan, 2004). Water and ligand molecules which were present either in the crystallization buffers (calcium and sulfate ions) or in the cryoprotectant solution (glycerol) were added manually in *Coot*. In the final stages the model was refined with anisotropic atomic displacement parameters and with H atoms added at riding positions. The final model contains 287 water molecules (nine in alternative locations), five glycerol molecules, three sulfate anions and one calcium cation. The progress of the refinement was monitored and the model was validated using the $R_{\rm free}$ parameter (Brünger, 1992). The quality of the final structure was assessed with *MolProbity* (Chen *et al.*, 2010). The refinement converged with a final R factor of 12.7% ($R_{\rm free}$ = 15.7%) for all data. The final model is characterized by a root-mean-square deviation (r.m.s.d.) from ideal bond lengths of 0.018 Å, with 97.8% of all residues in the most favoured areas of the Ramachandran plot

and with no residues in disallowed regions. The refinement statistics are reported in Table 1.

3. Results and discussion

3.1. Overview of the structure

Since our initial report of the successful incorporation of a peptoid residue into a native peptide inhibitor (Stawikowski *et al.*, 2005), there have been several additional publications describing various SFTI-1 analogues containing peptoid residuse at the P1 and P1′ sites (reviewed by Lesner *et al.*, 2011). Although a number of analogues have been reported, there are no structural studies addressing the proteolytic stability of peptoid-containing SFTI-1 analogues. In this report, we present the X-ray crystal structure of the acyclic analogue (*N*Lys)^{5}SFTI-1 in complex with trypsin determined at 1.29 Å resolution (Figs. 2 and 3, Table 1). The structure presented reveals high structural similarity to the native SFTI-1–trypsin complex. The structural differences between the two complexes measured as the r.m.s.d. between aligned C$^\alpha$ atoms is only 0.155 Å. The terminal ε-amino groups of the P1 residue are only 0.2 Å apart and the two conserved water molecules at the bottom of the catalytic pocket show perfect overlap (Fig. 4).

In our structure, (*N*Lys)^{5}SFTI-1 binds to the active site of trypsin in the same way as native SFTI-1 and other BBI inhibitors (Luckett *et al.*, 1999). The fully exposed binding loop of the inhibitor presents the P1 site with the peptoid residue (*N*Lys), which mimics the native Lys5 residue of the wild-type inhibitor (Figs. 3 and 4). The *N*Lys side chain is buried within the substrate pocket of the enzyme and makes contacts with Ser190, Asp189 and the two water molecules at the bottom of this pocket. One of the two antiparallel β-strands of (*N*Lys)^{5}SFTI-1 (residues Gly1–Cys3) is aligned with the enzyme, and the backbone atoms of Cys3 form canonical hydrogen bonds to Gly216 of the enzyme (Krowarsch *et al.*, 2003). Table 2 and Supplementary Fig. S1[1] summarize the hydrogen bonds that are formed between (*N*Lys)^{5}SFTI-1 and trypsin and compare them with the native SFTI-1–enzyme contacts. The last two residues of the (*N*Lys)^{5}SFTI-1 structure, namely Pro13 and Asp14, could not be reliably modelled in the electron-density maps and therefore are not included in the model. In the present structure, the electron-density map clearly shows that the P1–P1′ bond is intact, with no indication of hydrolysis (Fig. 3). Interestingly, the geometry as well as the distance between O$^\gamma$ of Ser195 and the carbonyl group of the P1 residue where cleavage normally takes place is the same (2.8 Å) in the native structure (PDB entry 1sfi) and in the present peptoid-containing structure.

3.2. The peptoid residue at position P1

There are two major consequences of the incorporation of N-substituted glycine into the peptide sequence of the native inhibitor: (i) a side-chain shift and (ii) increased flexibility of the modified moiety (Butterfoss *et al.*, 2009). A C$^\alpha$-to-N side-chain relocation should, in principle, disrupt the network of contacts between the inhibitor and the cognate enzyme. Surprisingly, however, the *N*Lys5 side chain ends at almost exactly the same position as the native residue and the ε-amino group makes identical hydrogen-bond contacts at the bottom of the binding pocket (Fig. 4). Clearly, the increased flexibility of the peptoid residue together with the length of the *N*Lys5 side chain can successfully compensate for the C$^\alpha$-to-N shift. When comparing the χ angles of the P1 residue, it is clear that the χ_3 and χ_4 torsion angles in the SFTI-1 and (*N*Lys)^{5}SFTI-1 structures are the same within ∼1° (Supplementary Table S1).

In general, peptoids are much more flexible than peptides (Butterfoss *et al.*, 2009). The potential backbone flexibility of the P1 peptoid residue, on the other hand, is restricted by the

Table 2
Comparison of inhibitor–enzyme contacts in the structures of the native SFTI-1 (PDB entry 1sfi) and (*N*Lys)^{5}SFTI-1 (this work) inhibitors in complex with trypsin.

Inhibitor residue/atom		Trypsin residue/atom	SFTI-1 Distance (Å)	(*N*Lys)^{5}SFTI-1 Distance (Å)
P3	Cys3 N	Gly216 O	3.08	2.93
	Cys3 O	Gly216 N	3.15	3.07
P2	Thr4 O	Gln192 N$^{\varepsilon2}$	3.00	Not present
P1	Lys5 O	Gly193 N	2.60	2.65
	Lys5 O	Ser195 N	3.06	2.99
	Lys5 N	Ser214 O	3.31	Not present
	Lys5 N	Ser195 O$^\gamma$	2.90	Not present
	Lys5 N$^\zeta$	Asp189 O$^{\delta1}$	3.19	3.22
	Lys5 N$^\zeta$	Ser190 O	3.11	2.94
	Lys5 N$^\zeta$	Ser190 O$^\gamma$	2.99	3.01
P2′	Ile7 N	Phe41 O	3.01	3.36

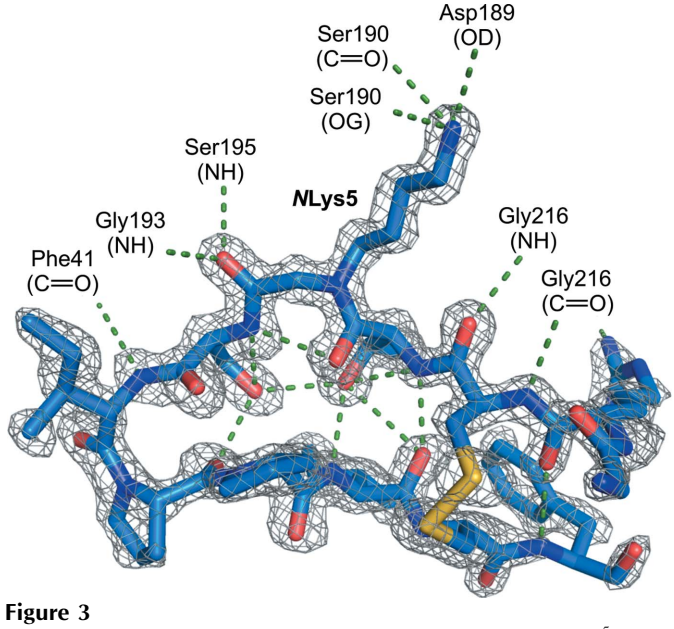


Figure 3
$2F_o - F_c$ electron density contoured at 1σ around the (*N*Lys)^{5}SFTI-1 inhibitor. Intramolecular and intermolecular hydrogen bonds are shown as dashed lines.

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: DZ5312).

rigidity of the active-site loop located between residues Cys3 and Cys11. In the investigated complex, the *N*Lys5 peptoid residue has a similar geometry to the native lysine residue. A comparison of the main-chain φ/ψ dihedral angles of the residues around the P1 position is given in Supplementary Table S2. In the BBI family, an inhibitory region is comprised of a disulfide-bridged nine-residue loop that adopts a characteristic canonical conformation. A unique feature of the BBI inhibitory loop is the presence of a *cis*-peptide bond between residues Ile7 and Pro8 at the N-terminal end of this loop. The following Pro9 residue effectively stabilizes the *cis* configuration (Brauer *et al.*, 2002). Interestingly, in the (*N*Lys)^{5}SFTI-1 structure the Pro9 residue possesses a different ring pucker to that in the native SFTI-1 structure. Such a compact and constrained architecture of the BBI loop does not leave much room for backbone flexibility and thus limits the movement of the incorporated peptoid residue, or the different ring pucker of Pro9 is a result of the presence of a flexible peptoid moiety within the inhibitory loop.

3.3. Enzyme flexibility: double conformation of the Cys191–Cys220 disulfide bridge

An additional interesting feature of the (*N*Lys)^{5}SFTI-1–trypsin complex structure is a dual right-handed conformation of the Cys191–Cys220 disulfide bridge of the enzyme molecule. The disulfide-bond conformation is determined by the analysis of five torsion angles of the two linked Cys residues, namely χ_1, χ_2, χ_3, $\chi_{2'}$, $\chi_{1'}$ (Schmidt *et al.*, 2006). Right- and left-handed structures are determined by the sign of the χ_3 angle. In our structure, the dual disulfide conformation is classified as −RHSpiral and +/−RHHook (according to the accepted nomenclature; Schmidt *et al.*, 2006), with S-atom occupancies of 0.53 (−RHSpiral) and 0.47 (+/−RHHook), respectively (Fig. 5). This phenomenon could indicate an increased flexibility of the enzyme framework upon the accommodation of a peptoid residue at the P1 position of the inhibitor. Among the 351 bovine trypsin structures deposited in the PDB, only 20 have the Cys191–Cys220 disulfide bond in a double conformation, including that presented here.

Table 3
Comparison of hydrogen-bond distances (Å) between the side chains of Gln192 and Ser214 and residues P1 and P2 of different inhibitor families, together with the distance between the Ser214 carbonyl group and Ser195 O$^{\gamma}$.

Inhibitor family	PDB code	Gln192 N$^{\varepsilon}$–P2 CO	Ser214 CO–P1 NH	Ser214 CO–Ser195 O$^{\gamma}$
BBI	1smf	2.7	3.3	3.6
Squash	2btc	3.2	3.1	3.5
Kunitz (STI)	1avw	3.2	3.0	3.5
Antistasin	1eja	3.4	3.3	3.6
Kazal	1ldt	3.1	3.1	3.7

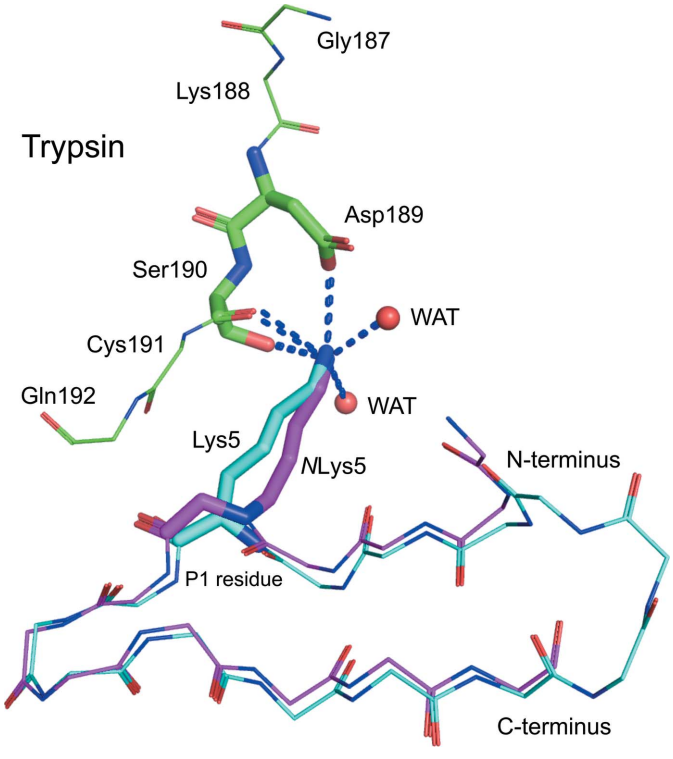
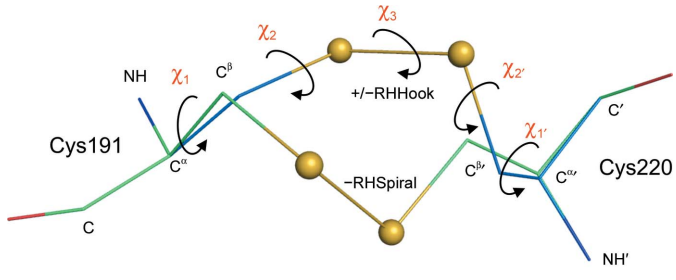


Figure 4
Comparison of the side chains of the *N*Lys5 residue (thick, magenta; this work) and the Lys5 residue of the native inhibitor (thick, cyan; PDB entry 1sfi) after C$^{\alpha}$ superposition of their trypsin complexes. The terminal ε-NH$_3^+$ group of *N*Lys5 makes the same contacts (dashed lines) with the bottom of the catalytic pocket of the enzyme, including hydrogen bonds to two water molecules (WAT; red balls). An analysis of the side-chain χ angles of P1 residues is presented in Supplementary Table S1.

3.4. Role of the trypsin Gln192 residue

The geometry of the catalytic site of trypsin in the (*N*Lys)^{5}SFTI-1 and native SFTI-1 (PDB entry 1sfi) complex structures is nearly identical, with one striking difference. Specifically, the side chain of Gln192 in the (*N*Lys)^{5}SFTI-1 structure has a very different conformation, with the terminal four atoms having only partial occupancy refined at 0.65. In this conformation the Gln192 side chain does not form any hydrogen bonds (Fig. 6). In the 1sfi structure, the Gln192 side chain forms a hydrogen bond to the carbonyl group of Thr4 at



Disulfide-bond classification	Dihedral angle (°) χ_1	χ_2	χ_3	$\chi_{2'}$	$\chi_{1'}$
−RHSpiral	−160.33	56.80	77.74	170.25	−43.49
+/−RHHook	−57.62	−143.35	87.77	78.03	160.16

Figure 5
Double Cys191–Cys220 disulfide conformation. The measured χ angles along with the S—S bond classification are provided below the figure.

the P2 position. Since the structure factors for entry 1sfi are not available in the PDB, it is not possible to examine the electron-density map for this side chain. The anisotropic atomic displacement parameters (ADPs) suggest that the side chain of Gln192 in the 1sfi structure is well ordered. The hydrogen bond between the Gln192 side chain and the carbonyl group of the P2 residue stabilizes (locks) the P1 residue in place. It is intriguing that the Gln192 side chain in the present structure does not form a hydrogen bond to the Thr4 carbonyl group of (*N*Lys)^{5}SFTI-1 (Table 2). To investigate this, we searched the PDB to analyze the position of the side chain of Gln192 of trypsin when in complex with an inhibitor. We found that in many inhibitor complexes the Gln192 side chain forms a hydrogen bond to the carbonyl group of the amino-acid residue at the P2 position. This type of interaction is found not only within the BBI family but also in other families of canonical protein inhibitors, such as those represented by squash, Kunitz, antistasin and Kazal family inhibitors. Representative crystal structures of such complexes and the respective hydrogen-bond distances are provided in Table 3. Another invariant feature of the analyzed complexes is the formation of a short antiparallel β-sheet interaction between residues P1–P3 of the canonical inhibitors and residues 214–216 of the protease. This intermolecular β-sheet between the protease and the inhibitor includes the following three highly conserved hydrogen bonds between the two main chains: (i) the carbonyl of Ser214 and the amide NH group of the P1 residue, (ii) the amide NH group of Gly216 and the carbonyl of the P3 residue and (iii) the carbonyl of Gly216 and the amide NH group of the P3 residue (Coombs *et al.*, 1999). Analyzing the trypsin–inhibitor complexes available in the PDB, one can observe the simultaneous interaction of residues 214 and 192 that locks the P1 residue of the inhibitor in the required position and allows the O$^{\gamma}$ atom of Ser195 to launch a nucleophilic attack on the carbonyl group of P1 (Fig. 6).

In general, in serine proteases residue 192, three positions before the catalytic Ser195, plays an important role in determining the substrate specificity. In trypsin and most trypsin-like enzymes with relatively broad specificity, Gln occupies this position. In thrombin, activated protein C and factor Xa, residue 192 contributes to the specificity of these enzymes towards their substrates and inhibitors (Rezaie & Esmon, 1996). In the plasma serine protease factor VIIa, residue 192 (Lys) is responsible for governing the substrate and inhibitor specificities (Neuenschwander & Morrissey, 1995). There are several reports suggesting that the residue at position 192 of coagulation proteases plays an important role in mediating the inhibitor- and substrate-specificity of these enzymes (Le Bonniec & Esmon, 1991).

3.5. Role of the trypsin Ser214 residue in catalysis

Of the approximately 200 mammalian and bacterial trypsin-like serine proteases, all but three have a serine residue at position 214. Ser214 is solvent-inaccessible, and its side-chain O$^{\gamma}$ atom forms a hydrogen bond to the essential Asp102 of the catalytic triad of trypsin. The influence of this residue on catalysis was examined in a mutagenesis study of trypsin by McGrath *et al.* (1992). When Ser214 was replaced with Lys, Glu or Ala to alter the polar environment, and consequently the electrostatic potential of Asp102, the mutations had a substantial effect on the catalytic activity. The S214K mutant had only 1% of the catalytic activity towards a tripeptide substrate, while the S214E mutant retained 44% of the activity. Remarkably, S214A trypsin was more active than the native enzyme. Crystallographic studies confirmed that the catalytic power of the other variants is compromised by a combination of structural and electrostatic perturbations (McGrath *et al.*, 1992). Krem and coworkers also used mutagenesis studies to elucidate the role of Ser214 in thrombin (Krem *et al.*, 2002). They reported the mutation of Ser214 to Ala, Thr, Cys, Asp, Glu and Lys. None of the mutants seriously compromised the catalytic function as measured by the kinetic parameter $k_{\rm cat}$. In contrast, the S214C mutant was 95% inactive. The Cys residue presumably engages in an improper disulfide bond with another Cys residue, although there is no further experimental evidence to confirm this hypothesis.

Apart from establishing the role of the side chain of residue 214 in catalysis, the role of the carbonyl group of Ser214 has also been described in the literature. For example, Derewenda and coworkers suggested the participation of the C$^{\varepsilon1}$—H group of His57 in a C—H···O=C hydrogen bond to Ser214 and showed how this interaction might contribute to what is now known as the ring-flip hypothesis, which proposes that a 180° rotation of the His57 imidazole ring facilitates catalysis (Derewenda *et al.*, 1994). These authors compiled a list of (His57)C$^{\varepsilon1}$—H···O=C(Ser214) distances in 22 trypsin-related native enzymes and found that they were in the range 2.04–2.59 Å, with an average of 2.3 Å, and proposed how the resonance structure of the imidazole ring is stabilized by this hydrogen-bond interaction. NMR measurements reported by Bachovchin and coworkers also confirmed that a 180° rotation of the His57 ring might be an important event in the catalytic process (Ash *et al.*, 2000). Recently, Schneider analyzed the

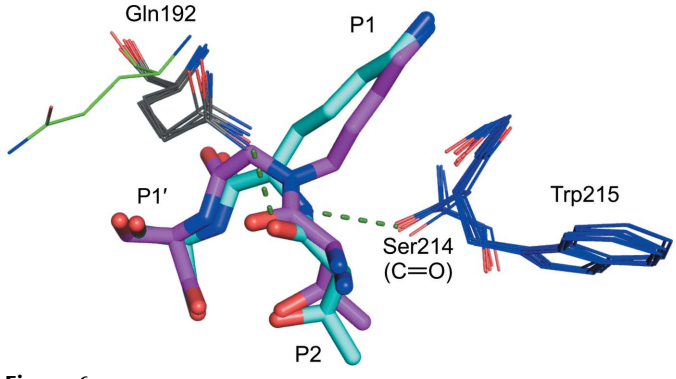


Figure 6
Cooperation of Ser214 (blue) and Gln192 (grey) of the enzyme with the P1–P2 residues of the inhibitor upon binding. This characteristic 'clip' is found within many families of canonical protein inhibitors, including SFTI-1 (cyan). Representative structures from Table 3 were superposed using the C$^{\alpha}$ atoms of the enzyme. In the present (*N*Lys)^{5}SFTI-1 structure (magenta) there is no hydrogen bond between Ser214 and NH of the P1 residue; as a consequence, Gln192 (green) does not interact with the carbonyl group of the P2 residue.

viability of the ring-flip hypothesis using quantum-chemical calculations (Scheiner, 2008). The (His57)$C^{\varepsilon 1}$—H···O=C(Ser214) distance computed for the active configuration is 2.30 Å, in agreement with the experimental observations.

Since it is not possible to probe the role of the Ser214 carbonyl group on ligand binding by site-directed mutagenesis, other approaches have to be used, in particular selective modification of the interacting ligand partner. To map the contacts that are crucial for protein–ligand interactions, a peptidomimetic modification can be introduced in a systematic manner, analogous to the commonly used alanine scanning (Cunningham & Wells, 1989). From the synthetic point of view, two methods that are feasible and highly compatible with peptide synthesis can be utilized: (i) the introduction of *N*-methylated amino acids (Türker *et al.*, 1972; Slon-Usakiewicz *et al.*, 1997) or (ii) amino-acid replacement with peptoid residues. The latter approach is known as a 'peptoid scan', in which peptoid-containing analogues of peptidic ligands can provide valuable information (Park *et al.*, 2013).

The reported crystal structure of an enzyme in complex with (*N*Lys)^{5}SFTI-1 illustrates how the *N*Lys peptoid residue selectively disables the Ser214 interaction (owing to its inability to form a hydrogen bond to the backbone of the P1 residue) which is normally present in the native SFTI-1–trypsin complex. Our crystal structure reported here also provides evidence for the importance of the main-chain carbonyl group of the Ser214 residue for inhibitor/substrate binding and catalysis. The structural observations are supported by previously reported enzyme kinetic data for (*N*Lys)^{5}SFTI-1, showing that this particular analogue has a 1000-fold lower potency against trypsin than the native acyclic inhibitor (Stawikowski *et al.*, 2005). There are several interesting examples in the literature supporting the major role of the carbonyl group of Ser214 in ligand binding and catalysis. For example, selective blocking of the Ser214 carbonyl group was reported in the case of thrombin, where a series of bivalent thrombin inhibitors consisting of a DPhe-Pro-N^{α}(Me)-Arg active-site-blocking segment were used (Steinmetzer *et al.*, 2000). Although methylation of the P1 residue led to a moderate decrease in affinity, it also prevented thrombin-catalyzed proteolysis (as measured by HPLC analysis), regardless of the P1′ residue used (Steinmetzer *et al.*, 2000). The crystal structure of the N^{α}(Me)-Arg-containing inhibitors in complex with thrombin revealed that the N^{α}(Me) group is directed toward the carbonyl O atom of Ser214, also pushing the Ser195 O^{γ} atom out of its normal position (PDB entry 1eb1; Friedrich *et al.*, 2002). On the other hand, the electron-density maps showed that the peptide bond of the inhibitor had been cleaved in the crystal, presumably owing to the long incubation time of 14 d that was needed for crystallization and data collection (Friedrich *et al.*, 2002). Bode and coworkers speculated that even though the scissile peptide bond of the inhibitor and the O^{γ} atom of Ser195 were separated by a distance (5.0 Å) that theoretically makes nucleophilic attack unlikely, on the timescale of two weeks this crystal-favoured geometry could allow catalysis to occur at a slow rate. In our case, the (*N*Lys)^{5}SFTI-1 inhibitor was stable in solution for more than a week (Stawikowski *et al.*, 2005) and the distance between the O^{γ} atom of Ser195 and the carbonyl O atom of Ser214 is 4.3 Å. Interestingly, the average distance between these two atoms in the set of analyzed complexes (Table 3), including native SFTI, is ∼3.6 Å. In our reported structure, there is no methyl moiety to interfere with the carbonyl group of Ser214 and this permits a slightly shorter distance between these two atoms.

4. Conclusions

We have determined the crystal structure of the *N*Lys5 peptoid-containing analogue of the SFTI-1 inhibitor in complex with bovine trypsin at 1.29 Å resolution. To the best of our knowledge, this is the first structure demonstrating how peptoid residues might confer resistance to proteases on peptidomimetic inhibitors by misguiding the enzyme-recognition/interaction patterns. In the case of monocyclic SFTI-1 analogues, a flexible side chain of the peptoid P1 residue is capable of interacting with the catalytic pocket of trypsin, and by analogy with other serine proteases as well, in a native-like fashion. The absence of a main-chain amide proton at the P1 residue results in structural changes within the catalytic site. The carbonyl group of Ser214 is no longer able to form a hydrogen bond to the P1 residue, which in turn makes the side chain of Gln192 unable to form a hydrogen bond to the carbonyl group of residue P2. Also, the fact that the interaction between P1 amide group and the carbonyl group of Ser214 is crucial for enzyme–inhibitor interaction is supported by kinetic data (association equilibrium constants of K_a = $1.0 \times 10^8\ M$ for acyclic SFTI-1 and $K_a = 9.9 \times 10^{10}\ M$ for the (*N*Lys)^{5}SFTI-1 analogue; Stawikowski *et al.*, 2005). The electron-density maps also show that a tight fit of the P1 side chain within the catalytic pocket separates the C=O group of Ser214 from the nucleophilic O^{γ} atom of Ser195 by an extra distance of ∼1 Å compared with the native SFTI-1–trypsin structure. Thus, all of the above structural changes within the catalytic site resulting from blocking the Ser214 carbonyl group from interactions render P1 peptoid-containing SFTI-1 analogues resistant to proteolysis. In summary, the role of the carbonyl group of Ser214 in enzyme–substrate complex formation is unquestionable; its interactions can be efficiently blocked by the presence of a peptoid residue at the P1 position of the inhibitor, rendering hybrid inhibitors with a P1 peptoid completely resistant to hydrolysis. Proteolytic stability of the P1–P1′ bond was also observed in numerous peptoid-containing monocyclic SFTI-1 analogues tested against bovine trypsin, α-chymotrypsin and human leukocyte elastase (Stawikowski *et al.*, 2004; Łukajtis *et al.*, 2011). By engineering a peptoid P1 residue into inhibitors with a peptidic scaffold, it is possible to modulate the specificity and/or affinity of such analogues (Lesner *et al.*, 2011). For example, while the (*N*Lys)^{5}SFTI-1 analogue is a 1000-fold less potent trypsin inhibitor in comparison with native SFTI-1, the (*N*Phe)^{5}SFTI-1 analogue turned out to be two orders of magnitude more

potent than its Phe5-containing peptidic counterpart (Stawikowski *et al.*, 2005).

The side-chain shift which takes place when amino acids are replaced with *N*-substituted glycine residues is the main factor responsible for the activity modulation of peptoid-containing peptidomimetics. In many cases such peptidomimetics are simply not recognized by the respective enzymes and are not processed; hence, they exhibit stability toward proteolytic degradation (Miller *et al.*, 1994). Serine proteases are special in this respect because, in contrast to other proteases such as matrix metalloproteases (Stawikowski & Fields, 2014), they tolerate the accommodation of peptoid residues at the P1 position of native peptidic ligands.

The authors are grateful to Agilent Technologies for excellent support during preliminary X-ray data collection on a SuperNova diffractometer.

References

Ash, E. L., Sudmeier, J. L., Day, R. M., Vincent, M., Torchilin, E. V., Haddad, K. C., Bradshaw, E. M., Sanford, D. G. & Bachovchin, W. W. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 10371–10376.

Bartunik, H. D., Summers, L. J. & Bartsch, H. H. (1989). *J. Mol. Biol.* **210**, 813–828.

Brauer, A. B., Domingo, G. J., Cooke, R. M., Matthews, S. J. & Leatherbarrow, R. J. (2002). *Biochemistry*, **41**, 10608–10615.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Butterfoss, G. L., Renfrew, P. D., Kuhlman, B., Kirshenbaum, K. & Bonneau, R. (2009). *J. Am. Chem. Soc.* **131**, 16798–16807.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Coombs, G. S., Rao, M. S., Olson, A. J., Dawson, P. E. & Madison, E. L. (1999). *J. Biol. Chem.* **274**, 24074–24079.

Cunningham, B. C. & Wells, J. A. (1989). *Science*, **244**, 1081–1085.

Derewenda, Z. S., Derewenda, U. & Kobos, P. M. (1994). *J. Mol. Biol.* **241**, 83–93.

Di Cera, E. (2009). *IUBMB Life*, **61**, 510–515.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Friedrich, R., Steinmetzer, T., Huber, R., Stürzebecher, J. & Bode, W. (2002). *J. Mol. Biol.* **316**, 869–874.

Goodman, M., Wang, L. L. & Feng, Y. (1994). *Polymer Prepr.* **35**, 767–768.

Hartley, B. S. & Kilby, B. A. (1954). *Biochem. J.* **56**, 288–297.

Hill, D. J., Mio, M. J., Prince, R. B., Hughes, T. S. & Moore, J. S. (2001). *Chem. Rev.* **101**, 3893–4012.

Krem, M. M., Prasad, S. & Di Cera, E. (2002). *J. Biol. Chem.* **277**, 40260–40264.

Krowarsch, D., Cierpicki, T., Jelen, F. & Otlewski, J. (2003). *Cell. Mol. Life Sci.* **60**, 2427–2444.

Laskowski, R. A. & Swindells, M. B. (2011). *J. Chem. Inf. Model.* **51**, 2778–2786.

Le Bonniec, B. F. & Esmon, C. T. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 7371–7375.

Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.

Lesner, A., Łęgowska, A., Wysocka, M. & Rolka, K. (2011). *Curr. Pharm. Des.* **17**, 4308–4317.

Long, Y.-Q., Lee, S.-L., Lin, C.-Y., Enyedy, I. J., Wang, S., Li, P., Dickson, R. B. & Roller, P. P. (2001). *Bioorg. Med. Chem. Lett.* **11**, 2515–2519.

Luckett, S., Garcia, R. S., Barker, J. J., Konarev, A. V., Shewry, P. R., Clarke, A. R. & Brady, R. L. (1999). *J. Mol. Biol.* **290**, 525–533.

Łukajtis, R., Łęgowska, A., Wysocka, M., Dębowski, D., Lesner, A. & Rolka, K. (2011). *J. Pept. Sci.* **17**, 281–287.

McGrath, M. E., Vásquez, J. R., Craik, C. S., Yang, A. S., Honig, B. & Fletterick, R. J. (1992). *Biochemistry*, **31**, 3059–3064.

Miller, S. M., Simon, R. J., Ng, S., Zuckermann, R. N., Kerr, J. M. & Moos, W. H. (1994). *Bioorg. Med. Chem. Lett.* **4**, 2657–2662.

Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–225.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Neuenschwander, P. F. & Morrissey, J. H. (1995). *Biochemistry*, **34**, 8701–8707.

Olsen, C. A. (2010). *Chembiochem*, **11**, 152–160.

Park, M., Wetzler, M., Jardetzky, T. S. & Barron, A. E. (2013). *PLoS One*, **8**, e58874.

Patch, J. A., Kirshenbaum, K., Seurynck, S. L., Zuckermann, R. N. & Barron, A. E. (2005). *Pseudo Peptides in Drug Development*, edited by P. E. Neilsen, pp. 1–31. Weinheim: Wiley-VCH.

Rawlings, N. D., Barrett, A. J. & Bateman, A. (2010). *Nucleic Acids Res.* **38**, D227–D233.

Rezaie, A. R. & Esmon, C. T. (1996). *Eur. J. Biochem.* **242**, 477–484.

Scheiner, S. (2008). *J. Phys. Chem. B*, **112**, 6837–6846.

Schmidt, B., Ho, L. & Hogg, P. J. (2006). *Biochemistry*, **45**, 7429–7433.

Simon, R. J. *et al.* (1992). *Proc. Natl Acad. Sci. USA*, **89**, 9367–9371.

Slon-Usakiewicz, J. J., Purisima, E., Tsuda, Y., Sulea, T., Pedyczak, A., Féthière, J., Cygler, M. & Konishi, Y. (1997). *Biochemistry*, **36**, 13494–13502.

Stawikowski, M. & Fields, G. B. (2014). In preparation.

Stawikowski, M., Stawikowska, R., Brzozowski, K., Zablotna, E., Jaskiewicz, A. & Rolka, K. (2004). *Peptides 2004: Bridges Between Disciplines. Proceedings of the Third International and Twenty-eighth European Peptide Symposium, September 5-10, 2004, Prague, Czech Republic*, edited by M. Flegel, M. Fridkin, C. Gilon & J. Slaninova, pp. 121–297. Geneva: Kenes International.

Stawikowski, M., Stawikowska, R., Jaskiewicz, A., Zablotna, E. & Rolka, K. (2005). *Chembiochem*, **6**, 1057–1061.

Steinmetzer, T., Batdordshjin, M., Pineda, F., Seyfarth, L., Vogel, A., Reissmann, S., Hauptmann, J. & Stürzebecher, J. (2000). *Biol. Chem.* **381**, 603–610.

Tam, J. P., Wu, C. R., Liu, W. & Zhang, J. W. (1991). *J. Am. Chem. Soc.* **113**, 6657–6662.

Türker, R. K., Hall, M. M., Yamamoto, M., Sweet, C. S. & Bumpus, F. M. (1972). *Science*, **177**, 1203–1205.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.