DOI: 10.1002/cbic.200800547

The High Resolution NMR Structure of Parvulustat (Z-2685) from *Streptomyces parvulus* FH-1641: Comparison with Tendamistat from *Streptomyces tendae* 4158

Stephan Rehm,[a] Sigeng Han,[b] Ismail Hassani,[b] Alma Sokocevic,[b] Hendrik R. A. Jonker,[a] Joachim W. Engels,[b] and Harald Schwalbe*[a]

The protein parvulustat (Z-2685) from Streptomyces parvulus *comprises 78 amino acids and functions as a highly efficient α-amylase inhibitor. Parvulustat shares 29.6% overall amino acid sequence identity to the well-known α-amylase inhibitor tendamistat. Among the conserved residues are the two disulfide bridges (C9–C25, C43–C70) and the active-site motif (W16, R17, Y18). Here, we report the high-resolution NMR structure of parvulustat based on NOEs, J couplings, chemical shifts and hydrogen-exchange data. In addition, we studied the dynamical properties of parvulustat by heteronuclear relaxation measurements. We compare the structure of parvulustat with the structure of tendamistat in terms of secondary structure elements, charges and hydrophobicity. The overall structural composition is very similar, but there are distinct differences including the active-site region. These structural and dynamical differences indicate that for parvulustat an induced-fit mechanism for binding to α-amylase might take place, since the structure of tendamistat does not change upon binding to α-amylase.*

Introduction

In microorganisms, plants and animal secretions, α-amylases are found as a widely distributed class of enzymes. They catalyse the hydrolysis of the (α-1,4) glycosidic linkage in amylose, starch and various other polysaccharides. These enzymes and their inhibitors are known to possess a wide range of industrial and medical applications. There are two different kinds of α-amylase inhibitors: carbohydrate-based compounds, which include acarbose,[1,2] and proteinaceous α-amylase-inhibitors. Many of the proteinaceous inhibitors stem from the genus of *Streptomyces* including tendamistat from *S. tendae*,[3] Z-2685 (parvulustat) from *S. parvulus*,[4] Haim II from *S. griseosporeus*,[5] Paim I from *S. olivaceoviridis*,[6] AI-409 from *S. chartreusis*, AI-3688 from *S. aureofaciens*,[7] T-76 from *S. nitrosporeus* and MA-4680 from *S. avermitilis*.[8]

The proteinaceous α-amylase inhibitors have very similar properties. These proteins are small and exhibit a β-barrel fold that consists of two twisted sheets, each of which is formed by three antiparallel β-sheets and generally stabilized by two disulfide bridges. The proteinaceous inhibitors bind very tightly to α-amylase in a 1:1 complex. The well-conserved active site contains a loop region that includes a WRY motif.[3,7,9–11] In addition, three more regions of the inhibitor proteins are involved in the protein–protein interface; this results in a very high inhibition constant (9×10^{-12} M for tendamistat) for a protein-based inhibitor.[10]

In 1985, the structure of the α-amylase inhibitor tendamistat was independently solved by NMR spectroscopy[12,13] and X-ray crystallography.[14] These structures of tendamistat build the reference for structure-based biophysical studies on proteinaceous α-amylase inhibitors carried out since.[9,15] In 1995, the crystal structure of porcine pancreatic α-amylase in complex with tendamistat was solved by X-ray crystallography.[9] During the 1980s, tendamistat was investigated for pharmaceutical purposes in a search for α-glucosidase inhibitors against diabetes mellitus. However, because of its high allergenic potential, tendamistat could not be used for therapy purposes. Another α-amylase-inhibitor, Z-2685 (parvulustat), was identified in screening projects at that time and the culture supernatant was sequenced by Edman degradation.

Initially, parvulustat was reported to comprise 76 amino acids. By sequencing the entire gene (Figure 1), which was isolated by PCR amplification and Southern hybridization, Han et al.[16] showed that the authentic product from *S. parvulus* is two amino acids longer and has a molecular weight of 8286 Da. In analogy to tendamistat, the protein was called parvulustat. Tendamistat and parvulustat share 29.6% amino acid identity and 37.0% similarity. Their inhibition activity against animal α-amylase, however, is almost identical.

Table 1 shows the identity and similarity of parvulustat and tendamistat compared to each other and other known α-amylase inhibitors by using BLOSUM62.[17] The data show that parvulustat has a higher similarity to other α-amylase inhibitors,

[a] *S. Rehm, Dr. H. R. A. Jonker, Prof. Dr. H. Schwalbe*
Institute of Organic Chemistry and Chemical Biology
Center for Biomolecular Magnetic Resonance
Johann-Wolfgang-Goethe University, Max-von-Laue-Strasse 7
60438 Frankfurt am Main (Germany)
Fax: (+49) 69-79829515
E-mail: schwalbe@nmr.uni-frankfurt.de

[b] *S. Han, I. Hassani, A. Sokocevic, Prof. Dr. J. W. Engels*
Institute of Organic Chemistry and Chemical Biology
Johann-Wolfgang-Goethe University
Max-von-Laue-Strasse 7, 60438 Frankfurt am Main (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200800547.

A)

```
1       10        20        30        40        50        60        70        80        90
GCCACCGGATCCCCGGTTGCGGAGTGCGTGGAGTACTTCCAGAGCTGGCGTTACACAGACGTTCACAACGGGTGCGCCGATGCCGTCTCC
AlaThrGlySerProValAlaGluCysValGluTyrPheGlnSerTrpArgTyrThrAspValHisAsnGlyCysAlaAspAlaValSer

91      100       110       120       130       140       150       160       170       180
GTCACGGTCGAGTACACGCACGGCCAGTGGGCCCCCTGCCGCGTGATCGAGCCCGGCGGCTGGGCGACCTTCGCCGGCTACGGAACGGAC
ValThrValGluTyrThrHisGlyGlnTrpAlaProCysArgValIleGluProGlyGlyTrpAlaThrPheAlaGlyTyrGlyThrAsp

181     190       200       210       220       230
GGCAACTACGTCACCGGGCTGCATACCTGCGACCCTGCCACGCCCAGCGGCGTC
GlyAsnTyrValThrGlyLeuHisThrCysAspProAlaThrProSerGlyVal
```

B)

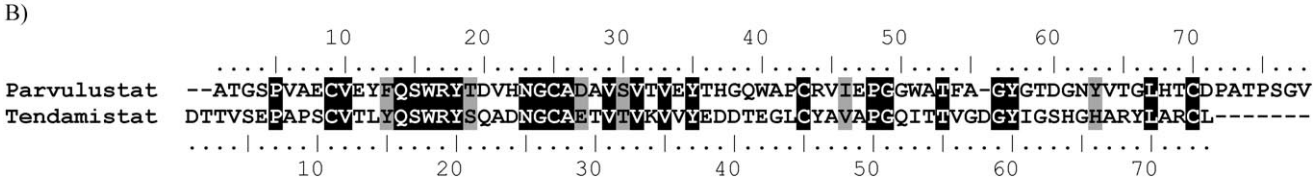

Figure 1. A) The parvulustat gene; underlined: the two amino acids found by Han et al.[16] B) Sequence alignment of tendamistat and parvulustat. The alignment was obtained by using Blosum62[17] as a matrix. Amino acid residues identical in the two proteins are shaded in black, similar amino acid residues are shaded in grey.

Table 1. Comparison of parvulustat (PL) and tendamistat (TM) to other known α-amylase inhibitors from various *Streptomyces* species with respect to identity and similarity of their amino acids (aa) by using Bioedit[67] with BLOSUM62.[17]

Inhibitor	From	Size [aa]	Identity to PL [%]	Similarity to PL [%]	Identity to TM [%]	Similarity to TM [%]
parvulustat	*S. parvulus*	78	100	100	29.6	37.0
tendamistat	*S. tendae*	74	29.6	37.0	100	100
haim II	*S. griseosporeus*	78	38.3	42.0	29.1	36.7
paim I	*S. corchorusii*	73	34.2	43.0	51.3	67.1
T-76	*S. nitrosporeus*	77	68.4	78.5	31.2	43.8
AI-409	*S. chartreusis*	78	71.3	77.5	35.0	43.8
MA-4680	*S. avermitilis*	92	34.8	38.0	22.6	31.2
AI-3688	*S. aureofaciens*	36	33.3	38.5	25.7	31.1

and is therefore a better representative of the class of proteinaceous α-amylase inhibitors, than tendamistat.

To provide a basis for future studies regarding, for example, protein stability or the mechanism of oxidative folding of parvulustat, we report the NMR structure determination of parvulustat. Based on a large number of NOEs and *J*-coupling restraints obtained by using heteronuclear multidimensional NMR spectroscopy, a high resolution structure of parvulustat could be calculated. The quality of the previous structures of tendamistat and the structure of parvulustat solved here allow for a detailed comparison of the structural properties of both α-amylase inhibitors.

The overall structure of the two proteins is very similar with a backbone RMSD of 0.51 Å; this includes the region involved in the protein–protein interface. The identical high-inhibition potency of parvulustat and tendamistat is likely to have evolved through compensating amino acid mutations. The dynamics of parvulustat are more pronounced compared to tendamistat. In particular, slow millisecond dynamics are observed for the three amino-acid motif, WRY, during binding to the α-amylase active site. In detail, the charge distribution of the parvulustat surface shows distinct differences. Given the increased subnanosecond and millisecond dynamics in parvulustat, binding of parvulustat to α-amylase most probably involves an additional induced-fit binding step compared to tendamistat, the structure of which, free in solution, is very similar to the α-amylase bound form.

Results and Discussion

NMR spectroscopy measurements

For the 78 residue protein parvulustat, the backbone resonance assignment is essentially complete with the exception of A1, G3, S4, Q14, W16 and R17 and all the nitrogens of the five prolines. The line widths of the two residues in the active site (W16, R17) are broadened beyond detectability.

Figure 2 shows a $^{1}H,^{15}N$ HSQC spectrum of parvulustat recorded at 600 MHz with annotated $^{1}H,^{15}N$-resonance assignments obtained from HNCA, HNCACB and CBCACONH experiments. The NMR resonances are well-dispersed with backbone amide signals resonating from 6.9 to 9.8 ppm. Such chemical-shift dispersion is typical for a protein rich in β-sheet secondary structure. The amide resonances for amino acids 25, 26 and 39 overlap in the 2D $^{1}H,^{15}N$ HSQC spectrum and could only be resolved in 3D experiments. Furthermore, four side chain and

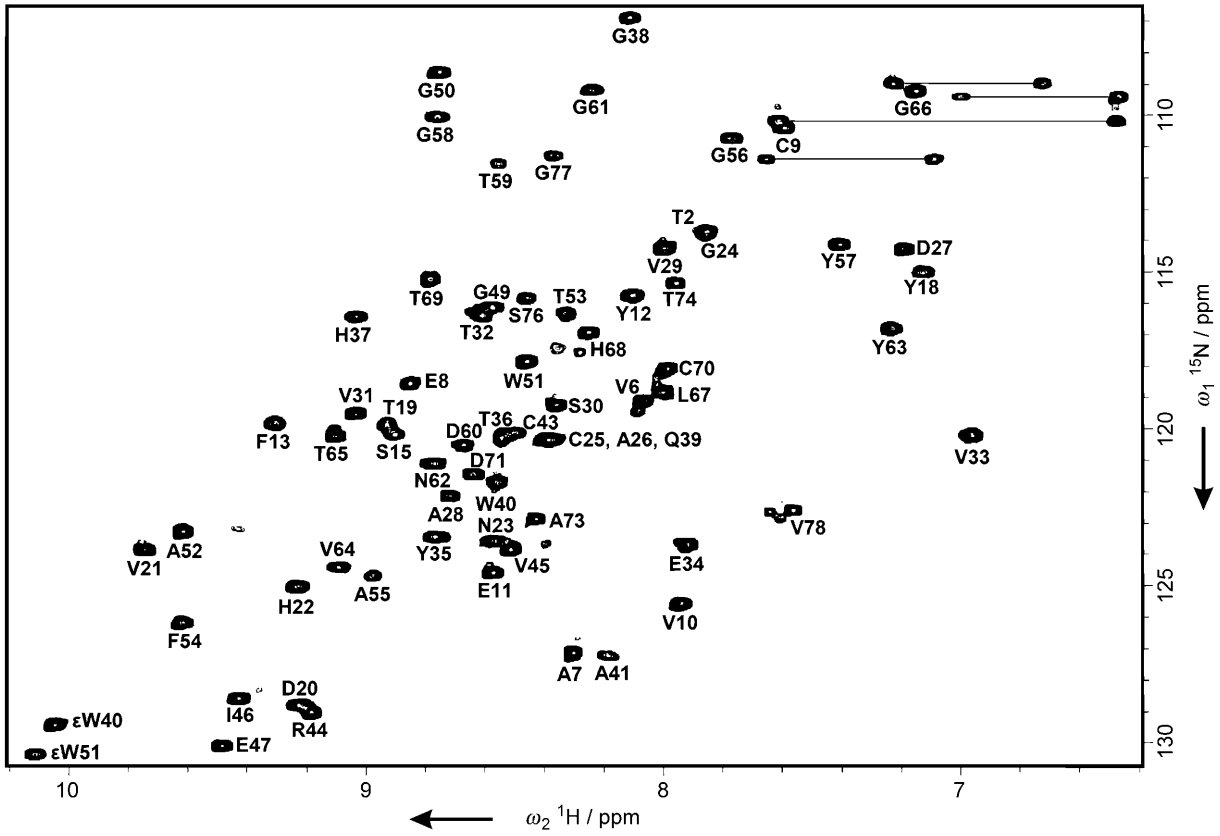

Figure 2. The 600 MHz ^{1}H,^{15}N HSQC of parvulustat (0.5 mM) in H_2O (92.5%)/D_2O (7.5%), sodium phosphate buffer (30 mM, pH 6.6), NaCl (170 mM), NaN_3 (0.01%) at 45 °C. All the assigned backbone ^{1}H,^{15}N peaks are indicated as well as the assigned Hε1 protons of W40 and W51.

two tryptophan indole H_ε signals are visible. These H_ε resonances could be assigned to W40 and W51, while the indole side chain resonance of the tryptophan in the WRY motif is not observable; this is in agreement with the lack of the backbone resonance. By using triple resonance experiments, backbone assignment and nonaromatic side-chain assignments were made to the following extents: H_N (93%), C_α (95%), H_α (94%), C_β (96%), H_β (91%), C_γ (44%), H_γ (69%), C_δ (24%), H_δ (72%) and C′ (85%).

NOE cross-peak intensities were obtained by manual peak picking each strip of the 800 MHz 3D ^{15}N NOESY-HSQC spectrum taken at the assigned ^{1}H,^{15}N amide resonances as well as in each strip of the 900 MHz 3D ^{13}C NOESY-HSQC spectrum taken at the assigned ^{1}H,^{13}C frequencies. In addition, four 2D ^{1}H,^{1}H NOESY (900 MHz) spectra were manually peak picked (Table 2).

The backbone assignments from parvulustat (^1H_N, ^{15}N, ^{13}C′, $^{13}C_\alpha$, $^{13}C_\beta$, $^1H_\alpha$) were used to assess the secondary structure of the protein. A protocol that is based on the joint probability of each of the three secondary-structure motifs (β-strand, α helix and random coil) derived from chemical-shift data (PSSI[18]) was used to identify the secondary structure and revealed the presence of six β-sheets (see the Supporting Information).

In the H/D-exchange experiments measured at 41.5 °C, the amide resonances disappeared rapidly: after 29 min, 53 of 67 resonances were not detectable anymore; after 150 min, only seven amide resonances (11, 25, 29, 31, 32, 46 and 52) remained and after 18 h only four signals (11, 25, 29 and 46) were left.

The ^{15}N NMR relaxation data (^{15}N-T_1, ^{15}N-T_2, {^{1}H}-^{15}N heteronuclear NOE, all at 600 MHz) were measured to determine the correlation time for the overall tumbling and to detect local subnanosecond backbone dynamics as well as slow chemical exchange of parvulustat. Figure 3A–C shows the mean values of the hetNOE, R_2/R_1 and S^2. Due to broadening beyond detection, no values were obtained for residues 14, 16 and 17 and the N terminus (residues 1–5). For residues 24, 25, 26 and 39, the relaxation data could not be reliably resolved due to strong chemical-shift overlap. Residues 15 and 59 were not used in the analysis of hetNOE data, because of an insufficient signal-to-noise ratio.

The heteronuclear NOE (hetNOE; Figure 3A) shows that the C terminus (starting from amino acid 73) is flexible, since these values are very small and most are even negative. For the core, most of the R_2/R_1 ratios (Figure 3B) are around a value of 5 (with some outliers to higher values), and for the flexible C terminus around a value of 2. The order parameter (S^2) is mapped on the structure (as shown in Figure 3C); this indicates more flexibility for the loop region around residue 37 and 60 as well as for the C terminus, for which very low values are observed. Since resonances from N-terminal residues are not visible, one can speculate that the N terminus is also flexible, based on the slightly lower S^2 values of residues 6–9.

Table 2. Structural statistics for the ensemble of 20 NMR structures of parvulustat.

Experimental restraints	
Distance restraints	
intraresidue ($i-j=0$)	526
sequential ($\|i-j\|=1$)	361
medium range ($2\leq\|i-j\|\leq4$)	111
long range ($\|i-j\|\geq5$)	723
hydrogen bond	19
total	1740
Dihedral angle restraints	
torsion angle (ϕ/ψ)	2×32
J-coupling constants	
$^3J_{HNH\alpha}$ coupling restraints	51
Diffusion anisotropy restraints	
T_1/T_2 values for HN	54
CNS energies (kcal mol^{-1})	
E_{total}	-1704 ± 70
E_{vdw}	-123 ± 11
E_{elec}	-2498 ± 65
E_{NOE}	129 ± 8
$E_{\phi/\psi}$	0 ± 0
E_{coup}	46 ± 6
$E_{dani(T1/T2)}$	51 ± 5
RMSD from average for residue [Å][a]	
backbone N, CA, C′	0.514
heavy atoms	0.962
Ramachandran plot[a]	
most favoured regions [%]	75.5
additional allowed regions [%]	20.1
generously allowed regions [%]	3.1
disallowed regions [%]	1.3

[a] The values were calculated for residues 8–71 by using MOLMOL[26] and PROCHECK.[22,23]

The experimental value for global rotational tumbling ($\tau_c^{exp}=6.2$ ns) of parvulustat is consistent with the predictions ($\tau_c^{exp}=6.48$ ns, derived from HYDRONMR 5a[19]), and show that under the conditions applied in solution NMR studies, parvulustat is monomeric and its tendency to aggregate at higher concentrations and different buffer conditions could successfully be suppressed. In addition, the overall diffusion tensor is well described by the determined structural model.

The 3D NMR solution structure of parvulustat shows six β-strands comprising two twisted sheets of three antiparallel strands each. The overall structure of parvulustat is very similar to the known structure of the α-amylase inhibitor tendamistat[14,20] and is a member of the Greek-key β-barrel protein structures[12,21] (Figure 4). In the following, we compare in detail the structures of parvulustat and tendamistat.

Structural analysis: β-sheets

In the NMR structure of tendamistat, the β-strands were assigned to involve amino acids 12 to 17, 20 to 25 and 52 to 58 in the first sheet, and amino acids 30 to 37, 41 to 49 and 67 to 73 in the second sheet.[21] The predominantly hydrophobic sides of the two β-sheets face each other to form a hydrophobic core.

The following regions are assigned to be β-strands in the water-refined parvulustat structural bundle (20 structures) by

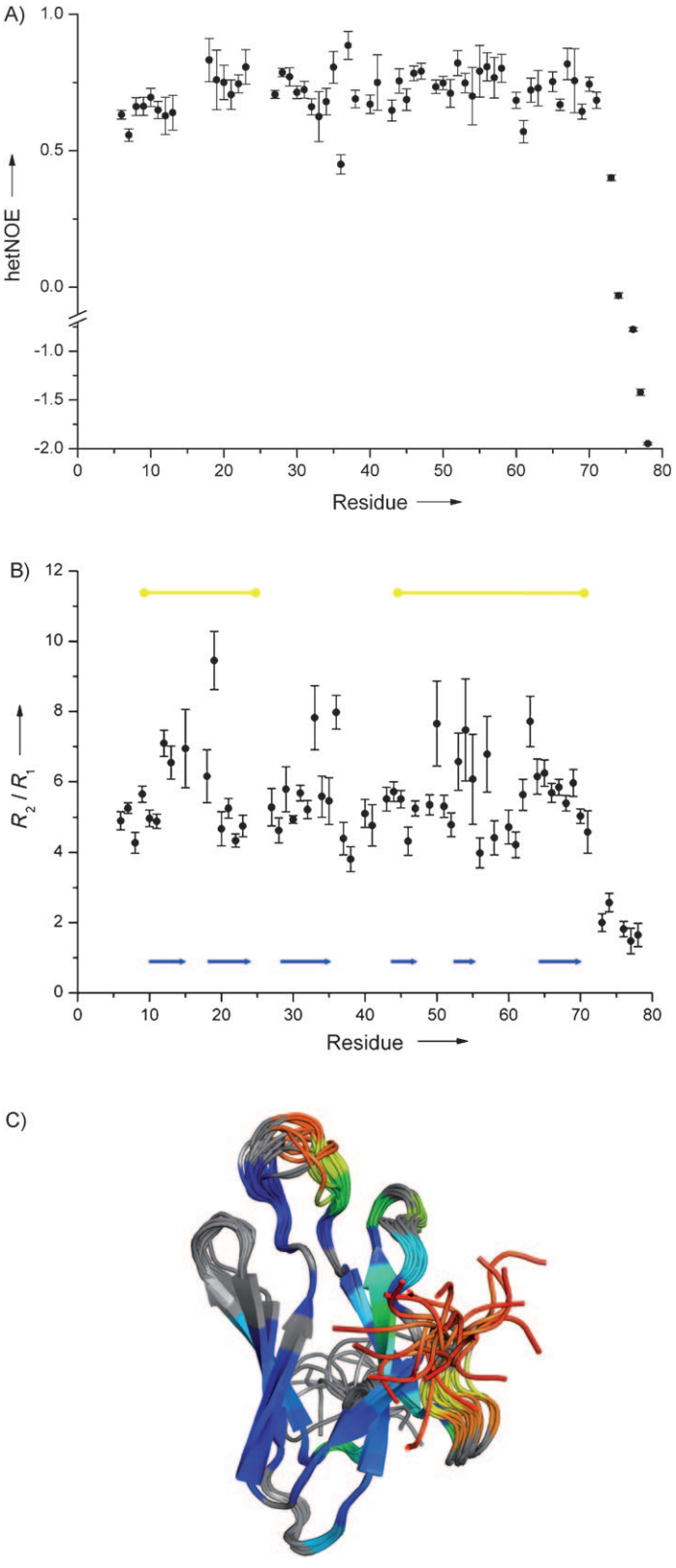

Figure 3. Backbone dynamics of parvulustat measured at 600 MHz and 45 °C: A) {^{1}H},^{15}N hetNOE. B) ^{15}N R_2/R_1 ratios; blue arrows: β-sheets, yellow bars: disulfide bridges. C) Order parameters (S^2) obtained from the model-free analysis[24,25] were mapped on the structure and show flexible parts in the loop regions around amino acids 37 and 60, probably also for the loop region at amino acid 16. Colours range from blue ($S^2=1$, very rigid) to red ($S^2=0$, very flexible); grey: no data/no reliable fit.

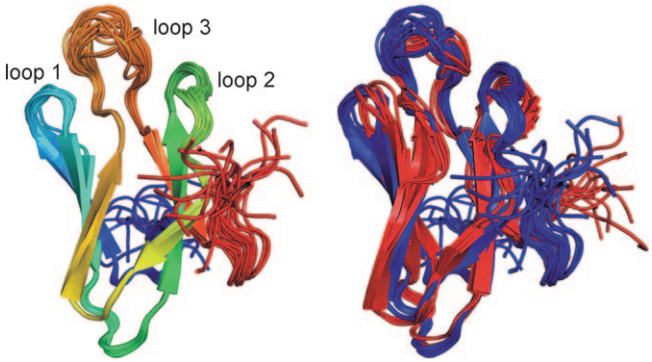

Figure 4. Left: backbone bundle of parvulustat (20 structures); right: overlay of parvulustat (blue, 20 structures) with the backbone bundle of tendamistat (PDB ID code: 2AIT; red, nine structures). The numbers indicate the successive loops in the sequence.

using PROCHECK 3.5.4[22,23] analysis: 10 to 15, 18 to 23 and 53 to 54 for the first β-sheet, and from 29 to 35, 44 to 46 and 64 to 69 in the second β-sheet. Taking the individual numbering of the two proteins into account, the first two β-strands in the first β-sheet are one amino acid longer each, while the third one (52 to 58 and 53 to 54) is five amino acids shorter in parvulustat. The third β-strand in tendamistat is distorted, both in the X-ray and in the NMR structure, and induces a twist in the three-stranded antiparallel β-sheet. In parvulustat, these strands are straight, most probably because they are a bit shorter.

Parvulustat forms a β-barrel structure for the same reasons that tendamistat does. In Figure 5 it is clearly visible that in the interior of the protein no charges (red: negatively charged, blue: positively charged, grey: no charge) appear and all these noncharged side chains constitute a hydrophobic core in parvulustat.

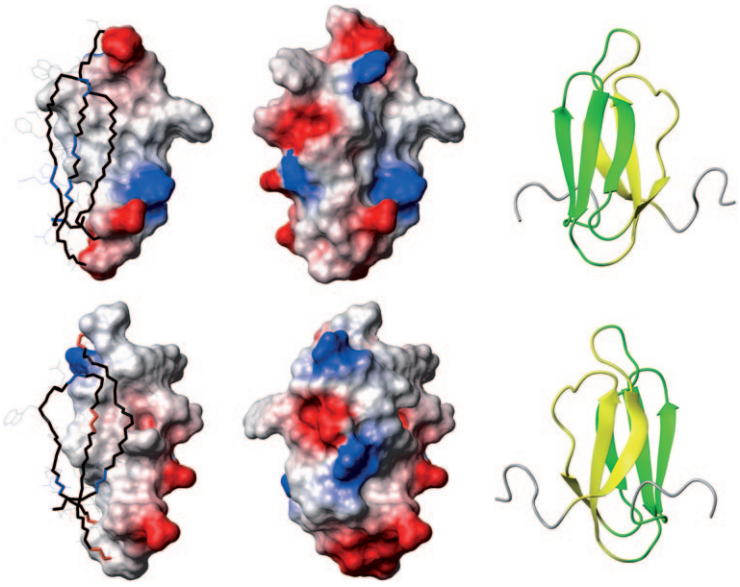

Figure 5. Left: the interior of parvulustat from both sides of the protein with respect to charges (red: negatively charged, blue: positively charged, grey: no charge); middle: charge distribution of the full protein shown from the same angle as on the left; right: ribbon illustration of the protein in the same angle.

For tendamistat, we compared the secondary structure analysis given in the literature with the PROCHECK[22,23] analysis of that structure and found that the first three β-strands are mostly the same, while for β-strand 4 the PROCHECK[22,23] analysis gives a significantly shorter β-strand than originally described. The NMR structure of parvulustat determined by ARIA was also analysed by PROCHECK[22,23] and compared with the results we gained from H/D-exchange experiments. Here, we find that due to the rather high temperature (41.5 °C) the exchange of the amide protons is fast. After 29 min, only 14 of 67 signals were still detectable. These 14 signals stem from residues that are located in β-sheets of parvulustat found by PROCHECK[22,23] (Figure 6), with just three notable exceptions: amino acids 24, 25 and 50, which are located in the lower very defined part of the protein, are not present in the β-sheet regions that were found by PROCHECK,[22,23] but are nevertheless more stable than two thirds of the other amide protons. Further support for differences in the length of β-sheet regions between tendamistat and parvulustat can also be derived, because none of the amide protons of the last β-strand remains after 8 min. Over time, nearly all of the signals vanish, except for seven signals that are still visible after 150 min (11, 25, 29, 31, 32, 46 and 52). Residues 29, 31 and 46, which are located close to each other, seem to build a very strong hydrogen bonded part of the β-sheet, while 25 is the only remaining amino acid not in a β-sheet.

In the second β-sheet in parvulustat, the second β-strand (41 to 49 and 44 to 46) is six amino acids shorter than the one in tendamistat, while the other two β-strands are only one amino acid shorter. Close to this β-strand, residue 42 is a proline in parvulustat, which corresponds to a leucine at this position in tendamistat. Because of the unique proline conformation, this changes the secondary structural element of parvulustat at the end of the β-sheet and the following loop in comparison to tendamistat (Figure 7). In general, the structures of tendamistat and parvulustat are very similar, but for the longer β-sheets found in tendamistat.

Structural analysis: the loops

The loops that connect the distinct β-strands of both β-sheets with each other play an important role in the structure of parvulustat and tendamistat. The active site containing the WRY motif is found within a flexible loop. In parvulustat, the loops comprise amino acids 15 to 18, 23 to 28, 36 to 40, 47 to 50 and 55 to 63.

The loops from 36 to 40 and from 55 to 63 are flexible. The flexibility of these regions was further supported by analysis of heteronuclear relaxation data by using the model-free analysis.[24,25] The general order parameter (S^2) given in Figure 3C, reveals two flexible loops in the protein. For the loop region at the active site, NMR resonances are broadened beyond detection at all available NMR frequencies; this suggests conformational exchange in the milli-

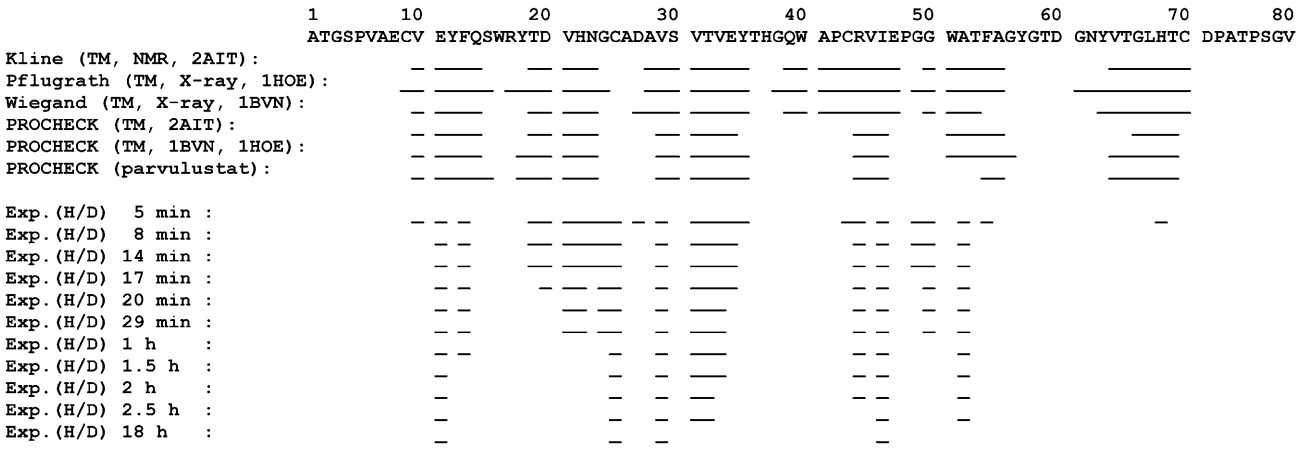

Figure 6. Comparison of β-sheets in tendamistat and parvulustat from the literature and PROCHECK[22,23] with H/D-exchange experiments, respectively, measured at 600 MHz and 41.5 °C. After 150 min, only the amide resonances of amino acids 11, 25, 29, 31, 32, 46 and 52 were visible, after 18 h only 11, 25, 29 and 46 were visible.

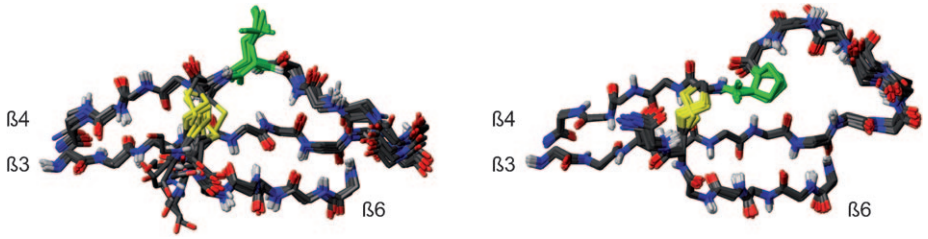

Figure 7. View of the backbone of the second β-sheet in tendamistat (left) and parvulustat (right), including the disulfide bridge (yellow); highlighted in green is the difference for residue P42 (parvulustat numbering) and L44 (tendamistat numbering).

second time regime. The notion of increased flexibility—yet in a different time regime—in this loop region is further supported by the very low order parameter (S^2) for amino acid 18. The lack of amide signals for the two other residues in the WRY motif (W16, R17) within the active-site loop, however, is quite remarkable as it points to millisecond dynamics and therefore enables an induced-fit mechanism of binding to α-amylase.

The other two loops of the protein (23 to 28 and 47 to 50) located on the opposite side of the protein, however, are very well defined in the backbone as well as in the side chains, which is in very good agreement with the structure determination of tendamistat, in which these parts are also highly defined.

Structural analysis: implications for the complex of parvulustat with α-amylase

The reported inhibition constant of tendamistat of 9×10^{-12} M is in the same range as the one determined for parvulustat, which is 2.8×10^{-11} M (Han et al., unpublished results). The fact that almost identical inhibition constants are found for the proteinaceous α-amylase inhibitors is interesting, since the amino acid sequence, as well as the parts of the charge distribution and to some extent the structure, especially the orientation of the side chains, are different. The free form structure of tendamistat fits very well to the structure when bound to α-amylase, which implies that parvulustat and α-amylase have to undergo some changes for binding.

Tendamistat binds to α-amylase in an extended groove with interactions in four segments: segment I: Y15, W18–Y20; segment II: L44, Y46; segment III: Q52–T55; segment IV: D58–G62. In Figure 8, these four segments are mapped on the structures of tendamistat and parvulustat (yellow). Also, in these α-amylase binding sites, there are substantial differences in the sequences of parvulustat and tendamistat. The largest differences are observed in segment II. With P42 in parvulustat and L44 in tendamistat, as well as R44 in parvulustat and Y46 in tendamistat (Figure 7), substantially different amino acids are present at the same locations in the proteins and this is the reason for the structural differences between the two inhibitors. Interestingly, in the WRY motif differences in the orientation of the side chains are clearly notable. However, it has to be noted that the structure of the WRY motif is mostly based on the force field during the structure calculation as no NMR signals were observed for this region in parvulustat.

Structural analysis: charge and hydrophobictiy

Charge and hydrophobicity distribution were calculated by using MOLMOL[26] based on the solution structure with the lowest energy. Figure 8 shows the charge distributions at the surface of tendamistat (free and bound form) and parvulustat (free form). Most of the residues that are involved in binding (yellow) are not substantially different in the two proteins; only one part (red arrow) is significantly different. At this position, we find R44 in parvulustat and Y46 in tendamistat, which

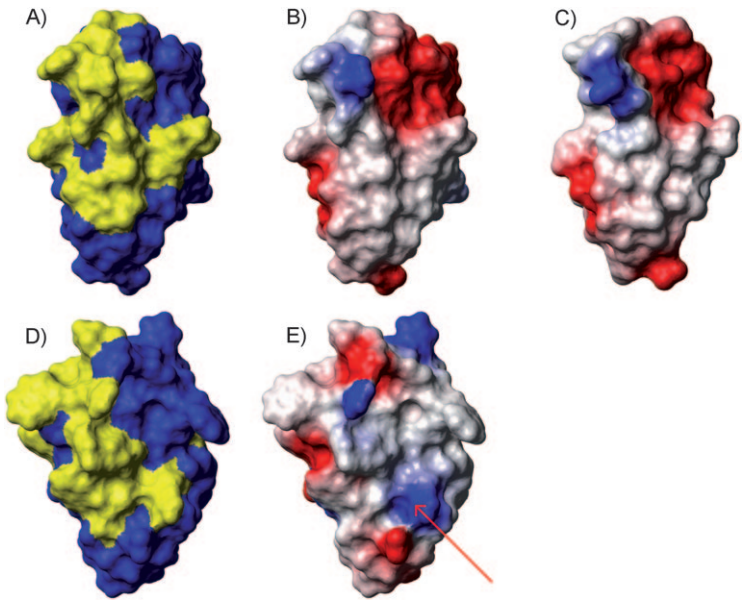

Figure 8. A) Surface representation of the tendamistat structure (yellow: binding segments to α-amylase). B) Surface potential for tendamistat free form (PDB ID code: 2AIT), and C) bound form (PDB ID code: 1BVN); red: negatively charged, blue: positively charged. D) Surface representation of the parvulustat structure in which the potential binding segments are mapped from tendamistat (yellow). E) Surface potential for parvulustat; the red arrow indicates R44, which corresponds to Y46 in tendamistat. The surface potentials were calculated with MOLMOL.[26]

explains the changes at the surface. The remaining binding-relevant residues of parvulustat and tendamistat fit reasonably well with respect to their charge distribution. On the back-side surface, which is not involved in the binding to α-amylase, substantial differences occur due to the difference in the side chains of both proteins. Tendamistat possesses a hydrophobic core, formed by the predominantly hydrophobic sides of the two β-sheets that face each other.[13] In parvulustat this is very similar especially for F54, for which the aromatic ring resides in the middle of the protein as defined by 37 NOEs (see the Supporting Information).

Conclusions

This study reveals the similarities and differences between two α-amylase inhibitors: parvulustat and tendamistat. Tendamistat was already well described by both NMR spectroscopy and X-ray data. We were able to use advanced 3D heteronuclear NMR spectroscopy methods to solve the solution structure of ^{13}C,^{15}N-labelled parvulustat in high quality as well as elucidate its dynamical properties from relaxation data. For tendamistat only 2D homonuclear spectra were available, therefore, differences appearing in the structure should be interpreted with care, since they could partially be due to the difference in the used method. Both proteins, which share 29.6% overall amino acid sequence identity are very similar with respect to their overall structure as well as their inhibition function against α-amylase. However, we found some distinct differences in the β-sheet structure. Parvulustat generally has shorter β-sheets and one of the sheets exhibits a different bending because of the presence of a proline instead of a leucine. With respect to the charges, both protein folds are stabilized by a hydrophobic core and mainly the same charges are solvent exposed. At one position in parvulustat R44 puts a positive charge on the surface, while in tendamistat there is Y46.

The structures of tendamistat in its free form as well as in complex with α-amylase reveal very little differences.[9,21] Since the structure of parvulustat differs from tendamistat, we propose either an induced-fit mechanism for binding to α-amylase or a slightly different binding mode. The conformational changes of the α-amylase upon binding of tendamistat are small,[27] but more substantial changes might accompany binding to parvulustat.

Experimental Section

Expression of parvulustat in *Streptomyces*: The parvulustat gene was subcloned in *Streptomyces lividans* TK24.[28] The expression cassette—as optimized by Schmitt-John and Engels[29] and Hoffmann[30]—was particularly developed for the expression of tendamistat in *S. lividans* TK24 and the episomal vector system pAX5a were found to be optimal for the expression of parvulustat in *S. lividans* TK24. In shaking-flask cultures, we obtained an initial protein yield of 50 mg L^{-1}. The detailed cloning procedure and optimisation of expression will be published elsewhere (Han et al., unpublished results; Figure 9).

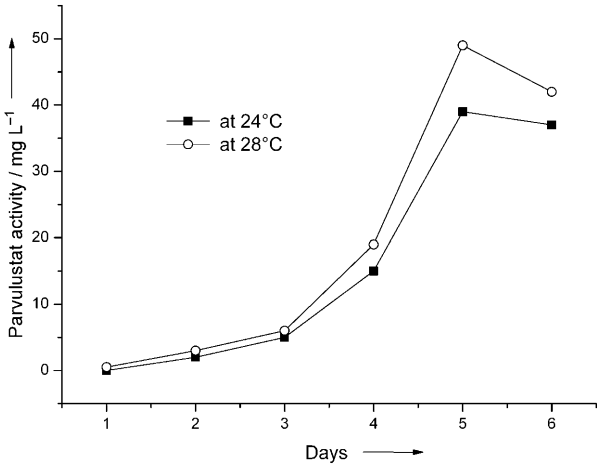

Figure 9. Expression level at 24 and 28 °C of parvulustat by using inhibition of cleavage activity of α-amylase as assay.

Ammonium sulfate precipitation (55% saturation) was used as the first step for protein purification, followed by the purification procedure reported earlier by Haas-Lauterbach et al.[31] and O'Connell et al.[32] Protein purity was analysed by RP-HPLC chromatography and SDS/tricine-PAGE; a protein purity of ≥95% was found (see the Supporting Information). Parvulustat was also characterized by MALDI-MS. For NMR spectroscopy measurements, parvulustat was both ^{15}N- and ^{15}N,^{13}C-labelled (see the Supporting Information).

The ^{15}N-labelled parvulustat was produced by using an isotope ^{15}N-enriched *Streptomyces* minimal medium (MM; for 50 mL: 22.95 mL sterile distilled water, 12.5 mL 20% (*w/v*) glucose, 10 mL 10% (*w/v*) ^{15}N-labelled amino acids (Spectra Stable Isotopes, Andover, MA, USA), 5 mL 300 mM sodium potassium phosphate buffer, 1.5 mL 1 M $MgSO_4$, 0.5 mL 20% (*w/v*) $(^{15}NH_4)_2SO_4$ (Spectra Stable Isotopes), 50 μL trace element solution).

The expression of the ^{15}N-isotope labelled parvulustat was also accomplished with the bifunctional shuttle vector pAX5a and with the supplemented *S. lividans* MM (see above). Isolation of the protein was started by filtration of the culture (Sartorius, 0.45 μm membrane filter) and concentration by ultrafiltration (Amicon YM3 membrane) followed by ammonium sulfate precipitation (45% saturation). Purification was carried out by anion-exchange chromatography (MonoQ HR5/5, GE Healthcare) and RP-HPLC (RP8 7 μm, 10×250 mm, Macherey–Nagel). The theoretical molecular mass of ^{15}N-labelled parvulustat is 8379.04 Da; by using MALDI-MS the mass was determined to be 8376.61 Da. The comparison of the two masses confirmed for parvulustat a ^{15}N-isotope labelling of 99.97%. For further NMR spectroscopy investigations, ^{15}N,^{13}C-labelled parvulustat was also expressed in *S. lividans* TK24. For the production of ^{15}N,^{13}C-labelled parvulustat the supplemented MM was identical to that used for the ^{15}N labelling (described above), except that the glucose used in the media was replaced by ^{13}C-glucose (Spectra Stable Isotopes). The purification of the doubly labelled protein, which was carried out according to the method described above, was successfully accomplished. The agreement of the molecular mass of the doubly labelled parvulustat with its specified mass was verified and confirmed by MALDI-MS with a discrepancy of 9.7 Da.

NMR spectroscopic experimental data: The following conditions were identified after considerable optimization to yield high quality NMR spectra of parvulustat: 0.5 mM in H_2O (92.5%)/D_2O (7.5%), sodium phosphate buffer (30 mM, pH 6.6), NaCl (170 mM) and NaN_3 (0.01%). At higher protein concentrations, aggregation was observed. All NMR spectra (except the H/D exchange) were recorded at 45 °C by using Bruker spectrometers ranging from 600 to 900 MHz and equipped with *z*-axis gradient ^{1}H{^{13}C,^{15}N} triple resonance cryogenic probes. The spectrometers were locked on D_2O. Sequential backbone resonance assignments were obtained by using HNCO,[33,34] HNCA,[34,35] HNCACB[36,37] and CBCACONH[35] triple resonance experiments. Side-chain assignments were obtained from HBHACONH[36,38] and HCCH-TOCSY[39] experiments. NOE-based distance restraints were obtained from the following experiments: four 2D NOESY[40–43] spectra (900 MHz, mixing times: 50, 100, 150 ms (90% H_2O/10% D_2O), 325 ms (100% D_2O)), 3D ^{15}N NOESY-HSQC[44] (mixing time: 150 ms) and 3D ^{13}C NOESY-HSQC[45–47] (mixing time: 130 ms). H/D-exchange experiments were measured at 600 MHz and 41.5 °C with SOFAST-HMQCs.[48] The *J* couplings were measured by using the HNHA experiment (Supporting Information).[49,50] Heteronuclear ^{15}N relaxation experiments ({^{1}H}-^{15}N hetNOE, T_1 and T_2)[51–53] were performed with a ^{15}N-labelled parvulustat sample (0.5 mM) at 600 MHz and a temperature of 318 K. The longitudinal ^{15}N relaxation rates were determined from a series of spectra with delays of 10, 20, 80, 120, 200, 400, 600, 800 and 1000 ms. Relaxation delays used for determining the transverse relaxation rates were 15.7, 31.4, 47.0, 62.7, 78.4, 109.8, 141.1, 188.2, 235.2, 313.6 and 1000 ms. The {^{1}H}-^{15}N hetNOEs were determined from the ratio of peak intensities (I^{on}/I^{off}) with and without the saturation of amide protons with a recovery delay and a length of saturation of 5.0 s. The dynamics calculation by using hetNOE values and T_1/T_2 ratios was performed with TENSOR2[54] and model-free,[24,25] for which the models were selected as proposed by d'Auvergne and Gooley for the model selection.[55] The ^{1}H chemical shifts were referenced to TMS at 0.00 ppm and ^{13}C and ^{15}N chemical shifts were calculated from the ^{1}H frequency.[56]

All NMR spectroscopy data were processed by using XWIN-NMR 3.5 and TOPSPIN 1.3-2.0 (Bruker Biospin, Karlsruhe, Germany) and analyzed by XEASY[57] and SPARKY 3.113.[58]

Structure calculation: All calculations were performed with CNS 1.1[59] by using the ARIA 1.2[60] setup and protocols with inclusion of diffusion anisotropy (DANI) restraints.[61] The protein allhdg 5.3 force field[62] was used with PROLSQ parameters. ARIA protocols were adapted to allow input of more than five spectra and to include half-assigned (ambiguous) peaks. After each of the first eight iterations (0–7) in which 50 structures were calculated, the NOE distance restraints were recalibrated by ARIA based on the ten lowest energy structures. The violation tolerance was progressively reduced to 0.1 Å in the last iteration (8) in which 200 structures were calculated. For the structure calculations, the standard simulated annealing (SA) protocol in ARIA was used with torsion angle dynamics (TAD). The initial high temperature stage consisted of 10000 steps at 10000 K and a high-temperature conformational search of 8000 steps to 2000 K. These high temperature stages were followed by two slow cooling stages in which the temperature was linearly decreased in 5000 steps to 1000 K and then in 10000 steps to 50 K. During the SA protocol the force constants for the NOE distance restraints and hydrogen bond restraints was set to 10 kcal mol^{-1} Å^{-2} during the high-temperature stages and 50 kcal mol^{-1} Å^{-2} during the cooling stages.

The ^1H_N, $^1H_\alpha$, ^{15}N, $^{13}C'$, $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts served as input for TALOS[63] to extract ϕ and ψ angles. Based on the secure TALOS[63] predictions, chemical-shift values (PSSI[18]) and NOE contacts, a total of 32 amply defined ϕ and ψ angle restraints were included for the β-sheet regions. The force constants for dihedral angle restraints were set to 5, 25 and 200 kcal mol^{-1} rad^{-2} in the high-temperature SA stage and the two cooling stages. In addition, 51 $^3J_{HNH\alpha}$ coupling constants were included for direct refinement by using the Karplus equation[64] with a force constant of 0, 0.2 and 1.0 kcal mol^{-1} Hz^{-2} during the high-temperature and the two cooling SA stages.

For refinement based on diffusion anisotropy,[61] 54 T_1/T_2 values revealed by the analysis of heteronuclear relaxation data were included with a force constant of 1 kcal mol^{-1}. The T_1/T_2 ratios were excluded for residues with a hetNOE value lower than 0.5 and for residues in flexible regions. The initial diffusion tensor anisotropy and rhombicity components and rotational correlation time were estimated by using TENSOR2.[54] A grid search procedure[61] was used to optimize the diffusion tensor based on the calculated structures after each iteration. The final values at the water-refinement stage of the calculation for the anisotropy and rhombicity were 2.0 and 1.0, respectively. The final 20 lowest-energy structures were further analyzed with PROCHECK 3.5.4[22,23] and refined in explicit water[65] by using OPLS parameters.[66]

Abbreviations: DANI: diffusion anisotropy; HSQC: heteronuclear single quantum coherence; MALDI-MS: matrix assisted laser desorption ionization mass spectroscopy; NMR: nuclear magnetic resonance; NOESY: nuclear Overhauser enhancement and exchange spectroscopy; RMSD: root mean square deviation; RP-HPLC: reversed-phase high-performance liquid chromatography; SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis; TMS: tetramethylsilane; TOCSY: total correlation spectroscopy

Acknowledgements

We are grateful to Sridhar Sreeramulu for helpful discussions. This work was supported by the Center for Biomolecular Magnetic Resonance (BMRZ). Support by the Center of Excellence: Macromolecular Complexes, is acknowledged. A.S. gratefully acknowledges financial support by the Graduate Study Program "Biologicals" DFG, GK 1172.

Keywords: inhibitors · NMR spectroscopy · parvulustat · protein structure · tendamistat

[1] W. Puls, U. Keup, H. Krause, G. Thomas, F. Hoffmeister, *Naturwissenschaften* **1977**, *64*, 536–537.
[2] D. Schmidt, W. Frommer, B. Junge, L. Müller, W. Wingender, E. Truscheit, D. Schäfer, *Naturwissenschaften* **1977**, *64*, 535–536.
[3] H. Aschauer, L. Vértesy, G. Nesemann, G. Braunitzer, *Hoppe-Seyler's Z. Physiol. Chem.* **1983**, *364*, 1347–1356.
[4] O. Hofmann, L. Vertesy, G. Braunitzer, *Biol. Chem. Hoppe-Seyler* **1985**, *366*, 1161–1168.
[5] S. Murao, A. Goto, Y. Matsui, K. Ohyama, *Agric. Biol. Chem.* **1980**, *44*, 1679–1681.
[6] S. Murao, N. Oouchi, A. Goto, M. Arai, *Agric. Biol. Chem.* **1983**, *47*, 453–454.
[7] L. Vértesy, D. Tripier, *FEBS Lett.* **1985**, *185*, 187–190.
[8] H. Ikeda, J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, S. Omura, *Nat. Biotechnol.* **2003**, *21*, 526–531.
[9] G. Wiegand, O. Epp, R. Huber, *J. Mol. Biol.* **1995**, *247*, 99–110.
[10] L. Vertesy, V. Oeding, R. Bender, K. Zepf, G. Nesemann, *Eur. J. Biochem.* **1984**, *141*, 505–512.
[11] H. Murai, S. Hara, T. Ikenaka, A. Goto, M. Arai, S. Murao, *J. Biochem.* **1985**, *97*, 1129–1133.
[12] A. D. Kline, K. Wüthrich, *J. Mol. Biol.* **1985**, *183*, 503–507.
[13] A. D. Kline, W. Braun, K. Wüthrich, *J. Mol. Biol.* **1986**, *189*, 377–382.
[14] J. W. Pflugrath, G. Wiegand, R. Huber, L. Vertesy, *J. Mol. Biol.* **1986**, *189*, 383–386.
[15] H. Kessler, P. Schmieder, W. Bermel, *Biopolymers* **1990**, *30*, 465–475.
[16] S. Han, Diploma thesis, Goethe University Frankfurt (Germany), **1998**.
[17] S. Henikoff, J. G. Henikoff, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10915–10919.
[18] Y. Wang, O. Jardetzky, *Protein Sci.* **2002**, *11*, 852–861.
[19] J. Garcia de la Torre, M. L. Huertas, B. Carrasco, *Biophys. J.* **2000**, *78*, 719–730.
[20] V. Konig, L. Vertesy, T. R. Schneider, *Acta Crystallogr. Sect. A* **2003**, *59*, 1737–1743.
[21] A. D. Kline, W. Braun, K. Wüthrich, *J. Mol. Biol.* **1988**, *204*, 675–724.
[22] R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
[23] R. A. Laskowski, J. A. C. Rullmannn, M. W. MacArthur, R. Kaptein, J. M. Thornton, *J. Biomol. NMR* **1996**, *8*, 477–486.
[24] A. M. Mandel, M. Akke, A. G. Palmer, *J. Mol. Biol.* **1995**, *246*, 144–163.
[25] A. G. Palmer, M. Rance, P. E. Wright, *J. Am. Chem. Soc.* **1991**, *113*, 4371–4380.
[26] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graph.* **1996**, *14*, 51–55.
[27] M. Machius, L. Vertesy, R. Huber, G. Wiegand, *J. Mol. Biol.* **1996**, *260*, 409–421.
[28] K. T. Hopwood, D. A. Bibb, M. J. Buttner, K. F. Chater, *The John Innes Foundation Norwich* **2000**.
[29] T. Schmitt-John, J. W. Engels, *Appl. Microbiol. Biotechnol.* **1992**, *36*, 493–498.
[30] H. Hofmann, PhD thesis. **1998**, Goethe University Frankfurt.
[31] S. Haas-Lauterbach, M. Scharf, B. Sprunkel, M. Neeb, K. P. Koller, J. W. Engels, *Appl. Microbiol. Biotechnol.* **1993**, *38*, 719–727.
[32] J. F. O'Connell, R. Bender, J. W. Engels, K. P. Koller, M. Scharf, K. Wüthrich, *Eur. J. Biochem.* **1994**, *220*, 763–770.
[33] R. Clubb, V. Thanabal, G. Wagner, *J. Magn. Reson.* **1992**, *97*, 213–217.
[34] L. E. Kay, G. Y. Xu, T. Yamazaki, *J. Magn. Reson. Ser. A* **1994**, *109*, 129–133.
[35] S. Grzesiek, A. Bax, *J. Magn. Reson.* **1992**, *96*, 432–440.
[36] D. R. Muhandiram, L. E. Kay, *J. Magn. Reson. Ser. B* **1994**, *103*, 203–216.
[37] M. Wittekind, L. Mueller, *J. Magn. Reson. Ser. B* **1993**, *101*, 201–205.
[38] S. Grzesiek, A. Bax, *J. Biomol. NMR* **1993**, *3*, 185–204.
[39] L. E. Kay, G. Y. Xu, A. U. Singer, D. R. Muhandiram, J. D. Forman-Kay, *J. Magn. Reson. Ser. B* **1993**, *101*, 333–337.
[40] M. Piotto, V. Saudek, V. Sklenar, *J. Biomol. NMR* **1992**, *2*, 661–665.
[41] V. Sklenar, *J. Magn. Reson. Ser. A* **1995**, *114*, 132–135.
[42] V. Sklenar, M. Piotto, R. Leppik, V. Saudek, *J. Magn. Reson. Ser. A* **1993**, *102*, 241–245.
[43] G. Lippens, C. Dhalluin, J. Wieruszeski, *J. Biomol. NMR* **1995**, *5*, 327–331.
[44] A. L. Davis, J. Keeler, E. D. Laue, D. Moskau, *J. Magn. Reson.* **1992**, *98*, 207–216.
[45] A. G. Palmer, J. Cavanagh, P. E. Wright, M. Rance, *J. Magn. Reson.* **1991**, *93*, 151–170.
[46] L. E. Kay, P. Keifer, T. Saarinen, *J. Am. Chem. Soc.* **1992**, *114*, 10663–10665.
[47] J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. Glaser, O. Sørensen, C. Griesinger, *J. Biomol. NMR* **1994**, *4*, 301–306.
[48] P. Schanda, B. Brutscher, *J. Am. Chem. Soc.* **2005**, *127*, 8014–8015.
[49] G. W. Vuister, A. Bax, *J. Am. Chem. Soc.* **1993**, *115*, 7772–7777.
[50] G. W. Vuister, A. Bax, *J. Biomol. NMR* **1994**, *4*, 193–200.
[51] K. T. Dayie, G. Wagner, *J. Magn. Reson. Ser. A* **1994**, *111*, 121–126.
[52] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, L. E. Kay, *Biochemistry* **1994**, *33*, 5984–6003.
[53] L. E. Kay, L. K. Nicholson, F. Delaglio, A. Bax, D. A. Torchia, *J. Magn. Reson.* **1992**, *97*, 359–375.
[54] P. Dosset, J. C. Hus, M. Blackledge, D. Marion, *J. Biomol. NMR* **2000**, *16*, 23–28.
[55] E. J. d'Auvergne, P. R. Gooley, *J. Biomol. NMR* **2003**, *25*, 25–39.
[56] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, B. D. Sykes, *J. Biomol. NMR* **1995**, *6*, 135–140.
[57] C. Bartels, T. Xia, M. Billeter, P. Güntert, K. Wüthrich, *J. Biomol. NMR* **1995**, *6*, 1–10.
[58] T. D. Goddard, D. G. Kneller Sparky 3, University of California, San Francisco.
[59] A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Crystallogr. D Biol. Crystallogr.* **1998**, *54*, 905–921.
[60] J. P. Linge, S. I. O'Donoghue, M. Nilges, *Methods Enzymol.* **2001**, *339*, 71–90.
[61] K. Houben, C. Dominguez, F. M. van Schaik, H. T. M. Timmers, A. M. Bonvin, R. Boelens, *J. Mol. Biol.* **2004**, *344*, 513–526.
[62] J. P. Linge, M. Nilges, *J. Biomol. NMR* **1999**, *13*, 51–59.
[63] G. Cornilescu, F. Delaglio, A. Bax, *J. Biomol. NMR* **1999**, *13*, 289–302.
[64] D. S. Garrett, J. Kuszewski, T. J. Hancock, P. J. Lodi, G. W. Vuister, A. M. Gronenborn, G. M. Clore, *J. Magn. Reson. Ser. B* **1994**, *104*, 99–103.
[65] J. P. Linge, M. A. Williams, C. A. Spronk, A. M. Bonvin, M. Nilges, *Proteins* **2003**, *50*, 496–506.
[66] W. L. Jørgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
[67] T. A. Hall, *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.

Received: August 14, 2008
Published online on December 9, 2008

