NMR Solution Structure of the 205−316 C-Terminal Fragment of Thermolysin. An Example of Dimerization Coupled to Partial Unfolding†

Francisco Conejero-Lara,‡ Carlos González,§ M. Angeles Jiménez,§ S. Padmanabhan,§ Pedro L. Mateo,‡ and Manuel Rico*,§

Instituto Estructura de la Materia, C. S. I. C., Serrano, 119, 28006 Madrid, Spain, and Departamento de Química-Física, Facultad de Ciencias e Instituto de Biotecnología, Universidad de Granada, 18071 Granada, Spain

Received May 7, 1997; Revised Manuscript Received July 10, 1997⊗

ABSTRACT: The solution structure of the C-terminal fragment 205−316 of thermolysin has been determined by ^{1}H-NMR methods. The fragment forms a dimer in which each subunit has two different regions: the largely disordered N-terminal segment 205−260 and the structurally well-defined segment 261−316. The structured part of each subunit is composed of three helices and is largely coincident with the corresponding region in the solution structure of the dimer formed by the shorter fragment 255−316, which in turn coincides with the crystallographic structure of intact thermolysin. As with the fragment 255−316, the subunit interface is highly hydrophobic and coincides topologically with the one between the segment 255−316 and the rest of the protein in the intact enzyme. A fourth helix (residues 235−246), present in the segment 205−316 of native thermolysin, is mostly disordered in the dimer formed by the fragment 205−316. The location of the fourth helix in the native structure of intact thermolysin does not allow the formation of the dimer interface observed in the solution structure of the fragment 255−316. Under the NMR conditions, dimer formation is energetically more favorable than the dissociated monomers. The latter, based on calorimetric data, was proposed to have partial structure in the region 205−254 as in native thermolysin. Thus, it appears that the assembly of the dimer would require an initial unfolding in the region 205−254 of the monomer.

The folding and stability properties of fragments of thermolysin, a well-characterized metalloprotease isolated from *Bacillus thermoproteolyticus* (Titani et al., 1972; Holmes & Matthews, 1982), have been the subject of growing interest in recent years (Vita et al., 1979, 1984, 1989; Vita & Fontana, 1982; Dalzoppo et al., 1985; Conejero-Lara et al., 1994; Azuaga et al., 1995; Conejero-Lara & Mateo, 1996). This was due to the finding that short amino acid sequences of the protein fold into native-like conformations and undergo cooperative unfolding transitions under the appropiate conditions. Thermolysin consists of a single polypeptide chain of 316 amino acid residues with a functional zinc ion and four calcium ions and has no thiol or disulfide groups. The three-dimensional structure of thermolysin, determined by X-ray, shows that this protein is constituted of two structural domains of equal size (residues 1−157 and 158−316), with the active site located at the interface between them (Holmes & Matthews, 1982). As mentioned above, some fragments of thermolysin, namely, 121−316, 206−316, 228−316, and 255−316, all belonging to the C-terminal domain, were shown by spectroscopic and immunochemical techniques to refold in water into stable native-like structures (Fontana, 1990; Vita et al., 1989). The study of such submolecular domains (of which there are very few reports) may provide important insights on the formation of supersecondary structures, and on the interplay between and the relative contributions of short-range versus long-range forces on folding and stability.

The three-dimensional structure of the shortest fragment 255−316 in aqueous solution has been recently determined in our laboratory by ^{1}H-NMR spectroscopy (Rico et al., 1994). The fragment forms a symmetric dimer in which the structure of each subunit largely coincides with the one that this region adopts in the crystalline state of intact thermolysin (Figure 1; Holmes & Matthews, 1982). This similarity indicates that the thermolysin fragment does not require contacts with the neighboring regions in the intact protein to define its overall topology, and the study also demonstrated that autonomously folding units can be substantially smaller than the entire structural domains seen in the crystal structure of relatively large proteins. Dimerization protects from the solvent the highly hydrophobic surface created upon folding of each subunit, without involving any major conformational change, except for the rearrangement of a few side chains. Circular dichroism studies indicate that the helical structure is already present in the monomer, suggesting that dimerization is not essential for the adoption of the native-like folding of the fragment (Vita et al., 1989). The unfolding reaction of this 61 residue fragment of thermolysin has been studied by differential scanning calorimetry (DSC),[1] and the observed concentration dependence was consistent with a dimeric structure undergoing two-state unfolding with no intermediate states (Conejero-Lara et al., 1994). The thermodynamic parameters of the unfolding process were also characteristic of a compact globular structure for the dimer in agreement with the 3D structure determined by NMR.

† This work was supported by DGICYT Grants PB93-0189 and PB93-1163 from the Ministerio de Educación y Ciencia (Spain).

* Corresponding author, Instituto de Estructura de la Materia, Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain. Telephone: 34-1-5619400. Fax: 34-1-5642431. E-mail: rico@malika.iem.csic.es.

‡ Instituto Estructura de la Materia.

§ Universidad de Granada.

⊗ Abstract published in *Advance ACS Abstracts,* September 1, 1997.

S0006-2960(97)01060-X CCC: $14.00

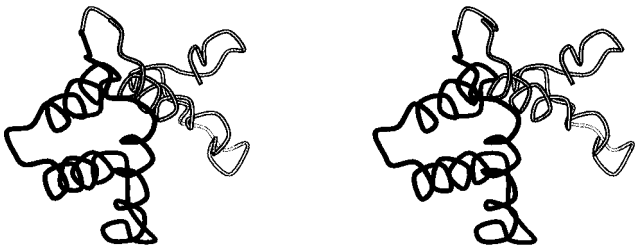

FIGURE 1: Stereo view of the backbone conformation for residues 205−316 in the crystal structure of thermolysin (Holmes & Matthews, 1982). Solid lines indicate the 255−316 segment and the grey lines the 205−254 segment.

More recently, a similar study by DSC has been carried out on the larger 112 residue fragment 205−316 (Azuaga et al., 1995; Conejero-Lara & Mateo, 1996). The calorimetric heat capacity data and their dependences on the concentration could not be fit to the simplest $N_2 \leftrightarrow 2U$ equilibrium as was the case with the shorter 255−316 fragment, but instead required the following equilibrium model:

$$N_2 \leftrightarrow 2N \leftrightarrow 2U$$

In this model, the dimer first dissociates to native monomers and ultimately to the unfolded monomers on thermal denaturation. The analysis led to a complete thermodynamic characterization of both the association and unfolding processes of the fragment. The resulting thermodynamic functions and, in particular, the low value obtained for the specific heat capacity of unfolding as well as the negative value obtained for the dissociation molar heat capacity change, suggested some degree of disorder in the structures of both the monomer and the dimer, as well as a conformational change linked to the association process. Whether this structural disorder extended throughout the whole structure or was concentrated in specific regions of the fragment was unknown. We therefore decided to undertake a 3D determination of the solution structure of this larger fragment by NMR techniques. Here, we show that the fragment 205−316 of thermolysin forms a dimer in which each subunit has two different regions: one largely disordered and the other structured and compact. Only three of the four helices of the native structure are present in the globular structure of the dimer. Our NMR data together with those from calorimetry indicate that dimerization and partial unfolding of native structure define and stabilize the observed structure, providing an example of the coupling of association and partial unfolding. Such coupled processes are of interest because of the general role they play in folding and recognition (Spolar & Record, 1994).

MATERIALS AND METHODS

Sample Preparation. Fragment 205−316 of thermolysin was prepared by limited proteolysis of the protein purchased from Sigma as previously described (Fassina et al., 1986; Conejero-Lara & Mateo, 1996). The sample of fragment 205−316 used for NMR studies was at least 95% pure by analytical reverse-phase high performance liquid chromatography.

NMR Spectroscopy. NMR samples were prepared by dissolving the peptide in either H_2O/D_2O 9:1 or D_2O at a concentration of ca. 2 mM and a pH of 4.0 and 5.0 (direct pH meter reading without correction for isotope effects). Sodium 3-trimethylsilyl(2,2,3,3-$^2H_4$) propionate (TSP) was used as an internal reference. NMR experiments were performed on a Bruker AMX-600 spectrometer at 298 K and at 308 K. All two-dimensional spectra were acquired in the phase-sensitive mode using the time proportional phase incrementation technique (Marion & Wuthrich, 1983) with presaturation of the water signal. Scalar coupling correlated spectra (COSY; Aue et al., 1976) and nuclear Overhauser enhancement spectra (NOESY; Kumar et al., 1980), using mixing times of 50, 100, and 200 ms in the latter, were recorded using standard phase-cycling sequences. Total correlation spectra (TOCSY; Bax & Davis, 1985) were obtained using the standard MLEV17 spinlock sequence and 75 ms mixing time. The size of the acquisition data matrix was 2048 × 512 words in f_2 and f_1, respectively. Prior to Fourier transformation, the 2D data matrix was multiplied by a phase-shifted square-sine bell window function in both dimensions and zero-filled to 4096 × 1024 words. The phase shift was optimized for every spectrum. Qualitative measurements of amide hydrogen exchange at 298 K were carried out by dissolving the lyophilized fully protonated protein in D_2O at pH 5.0 and observing the intensity of peptide NH resonances after a 48 h period. Slowly exchanging protons were those still present after that period.

Structure Calculation. Distance constraints were derived from NOESY data acquired with mixing times of 50 and 100 ms. Quantitative values of the NOE cross-peak volumes were determined by using the integration routines of the program XEASY (Eccles et al., 1991; Xia & Bartels, 1994; Bartels et al., 1995). Upper distance limits were fixed from the cross-peak intensities with the program CALIBA (Güntert et al., 1991), which assumes that the NOE intensities are proportional to $1/r^n$, where r is the interproton distance. The exponent n in our case was assigned a value of 6 when neither of the protons participating in an NOE contact belonged to a methyl group and a value of 4 when at least one methyl proton is involved. The scaling factors were chosen to reproduce the correct distance limit between pairs of protons separated by known, fixed distances (i.e., Hβ2-Hβ3, Hβ-methyls in Val, etc.). Different scaling factors were used for data from NOESY spectra recorded at two different mixing times. The distances obtained with CALIBA were rounded off to 3.0, 4.0, or 5.0 Å, following conservative criteria: distances between 4.0 and 5.0 Å were set to 5.0 Å, between 3.0 and 4.0 Å were set to 4.0 Å, and lower than 3.0 were set to 3.0 Å. In cases of conflict between data from different mixing times, priority was given to the shortest mixing time spectrum.

Stereospecific assignments of Val and Leu methyl groups, Hβ2 and Hβ3 side-chain protons and Pro Hδ2 and Hδ3 protons were performed on the basis of the intensities of sequential and intraresidue NOE cross-correlations together with $^3J_{\alpha\beta}$ coupling values and were further checked for the compatibility of the assignments with the preliminary calculated structures. Only completely nonambiguous stereospecific assignments were introduced in the succeeding

[1] Abbreviations: 2D, two-dimensional; ΔCp, heat capacity change; ASA, accessible surface areas; COSY, scalar coupling correlated spectra; DSC, differential scanning calorimetry; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectra; RMD, restrained molecular dynamics; RMSD, root mean square deviations; TFE, trifluoroethanol; TOCSY, total correlation spectra; TSP, sodium 3-trimethylsilyl(2,2,3,3-$^2H_4$) propionate.

steps of the structure calculation. These were the methyl groups of Val 303, 306, 313, and 315, the methyl groups of Leu 263, 271, 275, and 284, the Hβ2 and Hβ3 of Leu 263, Tyr 268, Tyr 275, Leu 275, and Gln 301 as well as the Hδ2 and Hδ3 of Pro 277. The usual corrections for pseudoatoms (Wüthrich, 1986) were added when no stereospecific assignments could be made.

The structures were calculated by Restrained Molecular Dynamics (RMD) in the GROMOS package (van Gunsteren & Berendsen, 1987), using an annealing strategy similar to the one used for the fragment 255−316 of thermolysin (Rico et al., 1994). In each of the RMD runs, the starting structure was first energy minimized and then heated up to 1000 K during 10 ps. At this temperature, 40 ps of RMD were carried out and eight structures were extracted from the trajectory and submitted to a cooling procedure of 5 ps at 750 K, 5 ps at 500 K, and 10 ps at 300 K. The last 5 ps of each trajectory were averaged, and the resulting structures were energy minimized. Three of the eight dimeric structures previously obtained for the fragment 255−316 of thermolysin were employed as starting structures in these calculations. This is justified given the close similarity between the NOE data set of the small fragment with the one corresponding to the same residues in the large fragment. The remaining amino acids (205−254) were added to the N-terminus of the fragment 255−316 in a random conformation. A set of 24 dimer structures were calculated by using a high-temperature RMD protocol.

Analysis of the Structures. All the molecular graphics manipulations were carried out with the program INSIGHT II version 2.2.0 (San Diego: Biosym Technologies, 1993). This program was also used for the analysis of the resulting structures. Solvent-accessible polar and nonpolar surfaces (Richards, 1977) were calculated using the VADAR package which implements the program ANAREA (Richmond, 1984).

RESULTS

The assignment of the ^{1}H-NMR spectra of the fragment 205−316 of thermolysin was performed by application of the standard sequential assignment strategy (Wüthrich, 1986). A complete assignment was possible without any requirement for ^{15}N- or ^{13}C-labeled samples. Figure 2 shows a summary of the observed sequential NN, αN, and/or βN NOE connectivities used in the spectral assignment. A list of the ^{1}H δ values is available as Supporting Information. Due to signal overlap, the sequential NN or αN NOE was not observed for G212, which was assigned on the basis of the NOE between its α protons and the ortho protons of Y211, a NOE which is commonly observed. The large upfield shifts (relative to random coil values) of the C_αH protons within regions 260−272, 283−296, and 301−310 (Figure 2) indicate that fragment 205−316 of thermolysin contains three helices. The two long stretches of $NN_{i,i+2}$, $\alpha N_{i,i+3}$, $\alpha N_{i,i+4}$, and $\alpha\beta_{i,i+3}$ NOE connectivities characteristic of helices (Figure 2) also confirm the presence of the last two helices. For the remaining helix, the observation of these characteristic NOE was hindered by signal overlap, but the large number of *i*,*i*+3 and *i*,*i*+4 NOE connectivities involving at least one side-chain proton is consistent with the presence of a helix spanning residues 260−274. A large number of such *i*,*i*+3 and *i*,*i*+4 NOE connectivities are also observed for the helices spanning residues 281−296 and 301−315. The slow hydrogen-exchange rates for most of the NH amide protons within the three segments provide additional support for the location and extension of the three helices, which match very closely the three C-terminal helical segments of crystalline thermolysin. In contrast, the segment 235−246, which exists as an α-helix in native thermolysin, shows very low helical population in aqueous solution, similar to that observed for the short fragment 233−248 (Jiménez et al., 1993), and, in agreement with that, the δ values of segment 235−246 in both fragments 205−316 and in 233−248 are almost identical. Both the pattern of NOE connectivities, with only a few medium-range NOEs (Figure 2), and the profile of $\Delta\delta$ C_αH shifts, very small in absolute value (Figure 2), indicate that the peptide chain is largely disordered within segment 205−259. The only indication of some nonrandom conformation in this region (205−259) is the slow exchange of some NH amide protons (Figure 2). This observation is, however, very puzzling since the absence of any medium- or long-range NOEs and random-coil-like δ for the residues in this region indicate absence of defined structure. We are unable to come up with a satisfactory explanation for this. The presence of low populations of turn conformations around Pro residues, as found in many short peptides (Dyson et al., 1988; Stroup & Gierasch, 1990; Chandrasekhar et al., 1991; Blanco et al., 1993), could be one possible origin of the protection of those protons against exchange. The simultaneous presence of structured and unstructured regions explains the striking appearance of the one-dimensional ^{1}H-NMR spectrum of thermolysin 205−316 fragment, which showed characteristics of denatured proteins, i.e., narrow and intense signals for the ortho and meta protons of Tyr residues, together with indications of a well-defined structure, such as a broad δ range for the NH amide protons and some very high-field methyl groups.

The existence of a monomer−dimer equilibrium has been discussed elsewhere for 205−316 thermolysin fragment at low concentrations (Azuaga et al., 1995; Conejero-Lara & Mateo, 1996). Nevertheless, we expected this fragment to be mostly dimeric under the high protein concentration used in our NMR study. In fact, using the dimerization constants reported by Azuaga et al. (1995), the dimer population turns out to be 92% under our NMR conditions. The presence of a series of NOE connectivities which are not compatible with the structure of a single subunit confirmed the dimeric state of the fragment 205−316 of thermolysin (Figure 3). Structural calculation of symmetric dimers from NMR data is complicated by the ambiguity between intra- and intersubunit cross-peaks. In this study, we used a similar strategy to the one employed in the calculation of the structure of fragment 255−316. The analysis of the NOESY data in the structure calculations was simultaneously checked for the two possibilities (intra or intersubunit) for those constraints which were systematically violated. The close similarity between the spin systems of the residues 255−316 in the two fragments facilitated the process. The final set of constraints consisted of 1113 intrasubunit and 53 intersubunit restraints, making an overall set of 2332 constraints for the whole dimeric molecule (Table 1). The sequential and medium-range NOE cross-peaks are displayed in Figure 2, and the complete list of distance constraints are available upon request from the authors. Most of the medium- and long-range NOEs detected belong to residues in the region 256−

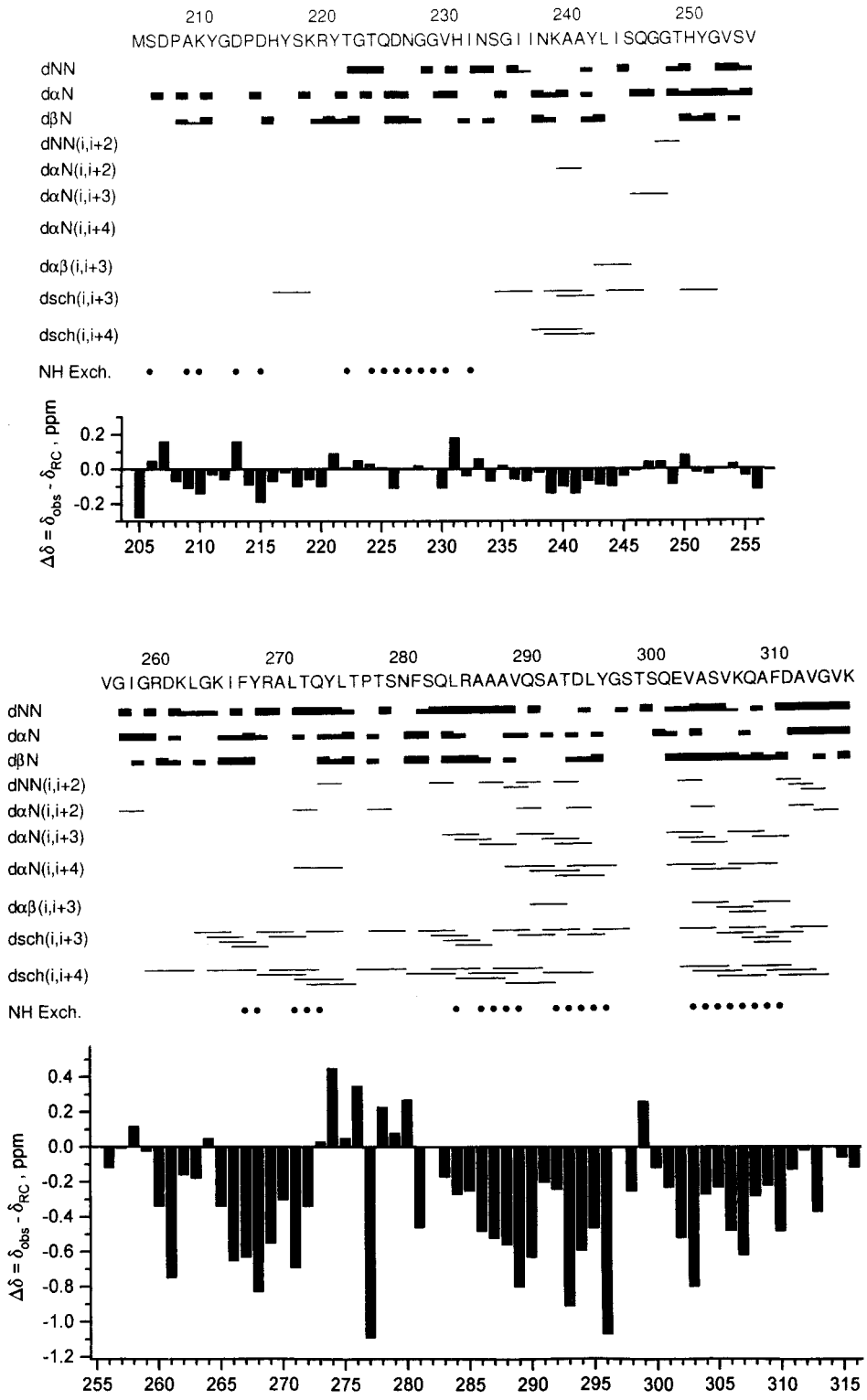

FIGURE 2: Sequence of 205−316 fragment of thermolysin together with a summary of observed sequential and medium-range NOEs, amide exchange data and CαH conformational shifts (Δδ). The NOEs classified into strong, medium, and weak are indicated by the thickness of the line. dsch indicates NOE connectivities involving at least one side chain proton. Amide NH protons which exchange slowly with solvent are indicated by filled circles. The CαH conformational shifts were evaluated as the difference between the observed *d* and that accepted as random coil value for each amino acid residue (Bundi & Wüthrich, 1979).

316, where the three helices are present. In contrast, there are only a few medium-range NOE cross-correlations, mostly weak, among residues within the segment 205−255. Most of them involve side-chain and backbone protons of residues within segment 234−249 (Figure 2) and are compatible with the presence of a low population of helical conformations within that segment similar to that found in the short fragment 233−248, as deduced from the profile of $\Delta\delta$ $C_\alpha H$ shifts (see above; Jiménez et al., 1993). The number of intersubunit constraints obtained for fragment 205−316 is equal to that obtained for the fragment 255−316, thus indicating that the spatial disposition of the intersubunit interface is very similar in the large and the short fragments. As expected from chemical shift data, the final structures

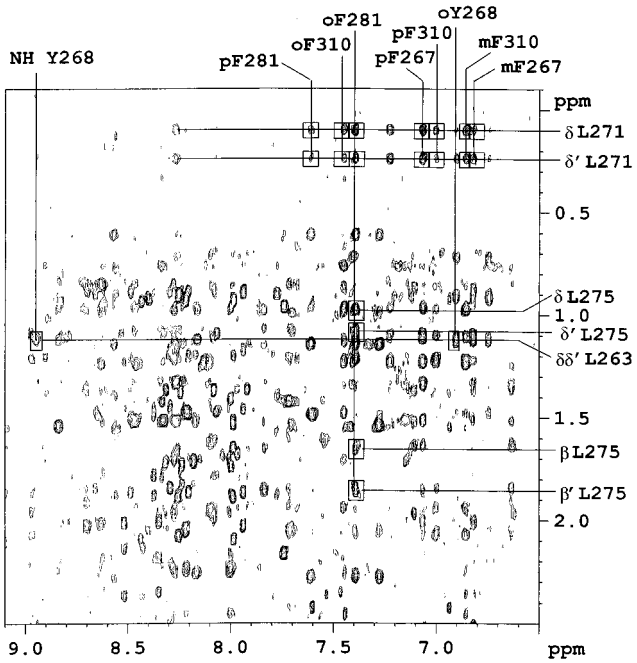

FIGURE 3: Region of the NOESY spectrum of the 205−316 fragment of thermolysin (1 mM, H_2O/D_2O 9:1, pH 4.0, 25 °C, mixing time 150 ms). Intersubunit NOE cross-correlations are boxed and labeled outside the vertical and horizontal scales.

Table 1: Structural Statistics[a]

	NOE restraints per subunit at a given distance (Å)			
type[b]	total	<3.5	<4.5	≤4.5
intra	414	75	162	177
SEQ	188	70	64	54
MBB	91	7	38	46
LBB	20	1	2	17
LGN	400	11	30	359
all intrasubunit	1113			
intersubunit	53			

intrasubunit constraint violations[c]	
range (Å)	
0.10−0.25	106.0
0.25−0.50	78.5
0.50−0.75	18.5
0.75−1.00	4.4
>1.0	0.9
maximal violation (Å)	1.0
sum of violations (Å)	60.0

intersubunit constraint violations	
maximal violation (Å)	0.8
sum of violations (Å)	2.4
averaged total energy[d]	−9513
total energy range	−7914 to −11359
averaged Lennard-Jones energy	−7044
Lennard-Jones energy range	−7471 to −6138
averaged NOE term	1057
NOE term range	933−1187

[a] Average values for the 24 final structures. [b] Classified according to Wüthrich (1986): intra, intraresidue; SEQ, sequential backbone or backbone-β; MBB, medium-range backbone−backbone or backbone-β; LBB, long-range backbone−backbone or backbone-β; LNG, all other interresidue NOEs. [c] Average values for the two subunits. [d] Energy units are kJ/mol.

are well-defined in the region 261−314 (Figure 4), but very poorly defined in the rest of the protein. The atomic RMSD for the backbone atoms in the 261−314 region is 0.64 and 1.35 Å for all heavy atoms in that region. The RMSD for the whole molecule is obviously very large, since the segment 205−255 is largely disordered (Figure 4). The backbone torsion angle deviations along the sequence show two separated regions (Figure 5a): a flexible part, comprising residues 205−261 (perhaps with the exception of the small segment 244−247 which presents a relatively ordered structure), and a region with a well-defined structure, comprising residues 261−312. In this region, the torsional angles in helical segments are extremely well-defined, whereas those in the loops are comparatively more relaxed. In the well-defined region of the molecule, the backbone conformation is very similar to the one found in the solution structure of fragment 255−316, which in turn is largely coincident with the crystal structure of intact thermolysin (Figure 1 and Figure 5, panels c and d). Although no symmetry conditions were explicitly included in the calculations, the average torsion angles in the backbone for the segment 261−314 are almost identical in both subunits, as can be seen in Figure 5b.

Both the backbone and the side chains in the region 261−314 of the fragment 205−316 are better defined here that in the smaller fragment 255−316. This is due to the larger number of constraints obtained in this study, as a consequence of analyzing the NOESY spectra using the semiautomatic assignment program XEASY. The latter program yields all possible assignments for each cross-peak in different stages of the calculation enabling a more extensive list of distance restraints than that previously obtained for the smaller 255−316 fragment by manual assignment. Accordingly, the lower atomic RMSDs obtained here do not mean that the segment 261−316 in the fragment 205−316 is more rigid than in the fragment 255−316, but rather reflect a larger constraint list.

DISCUSSION

Description of the Structure of Fragment 205−316 and Comparison with the X-ray and the Solution Structure of Thermolysin 255−316. The fragment 205−316 of thermolysin is a dimer in aqueous solution at the concentrations used in this NMR study. Each dimeric subunit consists of a disordered 56 residue long N-terminus, extending from residue 205 to 261, followed by a well-defined region between residues 261 and 314 (Figure 4). The last two residues (315, 316) in the chain are disordered. The well-defined region adopts a globular and compact structure, containing three helices spanning residues 261−274, 281−295, and 303−313, almost coincident with those observed in the native crystal structure (Holmes & Matthews, 1982; see Figure 1). The residues connecting the first two helices form a type I β-turn, which is better defined structurally than those between helices II and III. It may be noted that the peptides comprising each of the four helices in the segment 205−316 of native thermolysin were all shown to adopt, in aqueous solution, low but significant populations of native-like helical conformations, which increase upon addition of TFE (Jiménez et al., 1993). The conformational behavior of segment 236−247 in the short 233−248 peptide appears to be the same as that in the large 205−316 protein fragment. In contrast to the other three helices, the presence of the rest of the sequence does not promote the stabilization of the helix in segment 236−247 nor its further inclusion in a superstructure.

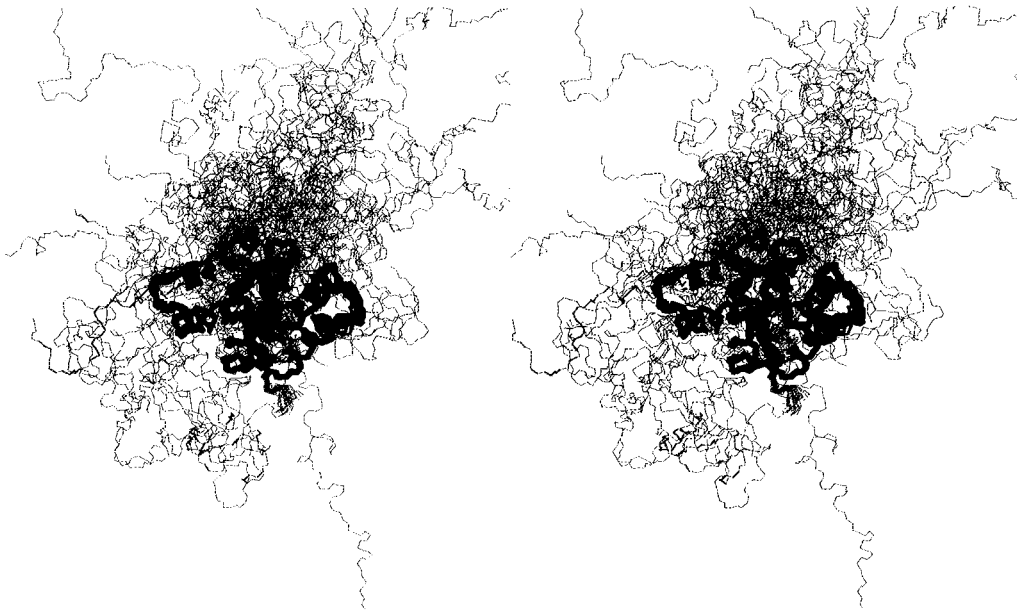

FIGURE 4: Stereo view of the the backbone conformations in the 24 final structures of the 205−316 fragment of thermolysin superposed for minimal RMSD over the structurally well-defined 255−316 region. Solid lines indicate the 255−316 segment and the grey lines the 205−254 segment.

The segment 261−314 in the large 205−316 fragment adopts a three-dimensional structure almost identical to that found in the shorter 255−316 fragment (Rico et al., 1994) and in the native structure of intact thermolysin determined by X-ray diffraction (Figure 1; Holmes & Matthews, 1982). The comparison of segment 261−314 in the solution structure of thermolysin 205−316 with the same region in thermolysin 255−316 and in intact thermolysin gives RMSDs of 0.8 and 1.2 Å for the backbone and heavy atoms, respectively. The differences in average backbone torsion angles are not significant because they do not exceed the RMSD between the torsion angles in the 24 final structures (Figure 5a). The same similarity is observed for the side chains of the residues buried in the core of the fragment 261−314 in both subunits.

The partial disorder of the 205−316 fragment as well as the close similarity in the structured regions between this and the 255−316 fragment is fully consistent with the whole set of thermodynamic functions characterizing the stability and unfolding reported elsewhere for both dimeric fragments (Conejero-Lara et al., 1994; Azuaga et al., 1995). For example, both the molar heat capacity change of unfolding together with molar unfolding enthalpy at the convergence temperature (140 °C) are the same within experimental uncertainty for the two fragments. A similar conclusion would also be valid for the molar heat capacity of the structured regions in the two folded dimers.

Dimer Interface. Most of the intersubunit NOEs identified in the fragment 205−316 of thermolysin are coincident with those observed in the small one, with no intersubunit contacts observed for the region 205−254. As a consequence, the resulting structures present a dimer interface very similar to the one obtained in the fragment 255−316 (Rico et al., 1994). The dimer interface is formed by an extensive hydrophobic surface comprising residues Leu 263, Phe 267, Tyr 268, Leu 271, Leu 275, Phe 281, Phe 310, and Val 313. A detailed description of such a dimer interface is given in our previous paper (Rico et al., 1994). A few side chains change their conformation with respect to the native structure in order to achieve a perfect matching of the two complementary surfaces in the dimer. The changes observed are conserved in the shorter and larger fragments. Only two intersubunit hydrogen bonds were detected (Phe 281 NH−Ser 279 CO and its symmetry related one). Hydrophobic interactions appear to be the primary driving force responsible for dimer formation.

On the basis of our results and the previous DSC studies (Azuaga et al., 1995; Conejero-Lara et al., 1996), the dimer of the 205−316 fragment is energetically more stable than the monomer at the high concentrations used for NMR. The dimer appears to find a favorable Gibbs energy balance by establishing extensive intersubunit interactions (mainly hydrophobic) without sacrificing a large amount of configurational entropy corresponding to the region 205−254, which remains disordered. It is interesting to note that the conformational state of side chains at the interface is extremely well-defined. The interdigitation of side chains is comparable to that found in native globular structures. This is most remarkable because of the nonnative character of these interactions, for which more relaxed long-range interactions are normally expected.

Monomer State. Using the solution structure reported here for the dimer of the fragment 205−316, we have calculated the changes in accessible surface areas (ASA) produced by dissociation. If it is assumed that the dissociation of the dimer leads to the three-helix monomeric subunits, then 2679 Å^2 of nonpolar surface and 1163 Å^2 of polar surface become exposed to the solvent. Conversely, these would be the areas that are buried when the two subunits interact to form the dimer (Figure 6). However, we do not know whether the dissociation follows such a simple process, since we could not determine the structure in solution of the monomeric form of the 205−316 fragment, because it is stable only at very low sample concentrations.

In a previous paper (Conejero-Lara & Mateo, 1996), we have reported on the far- and near-UV−CD spectra of the monomeric and dimeric states of the 205−316 fragment. Far UV−CD spectra for monomeric and dimeric fragments are

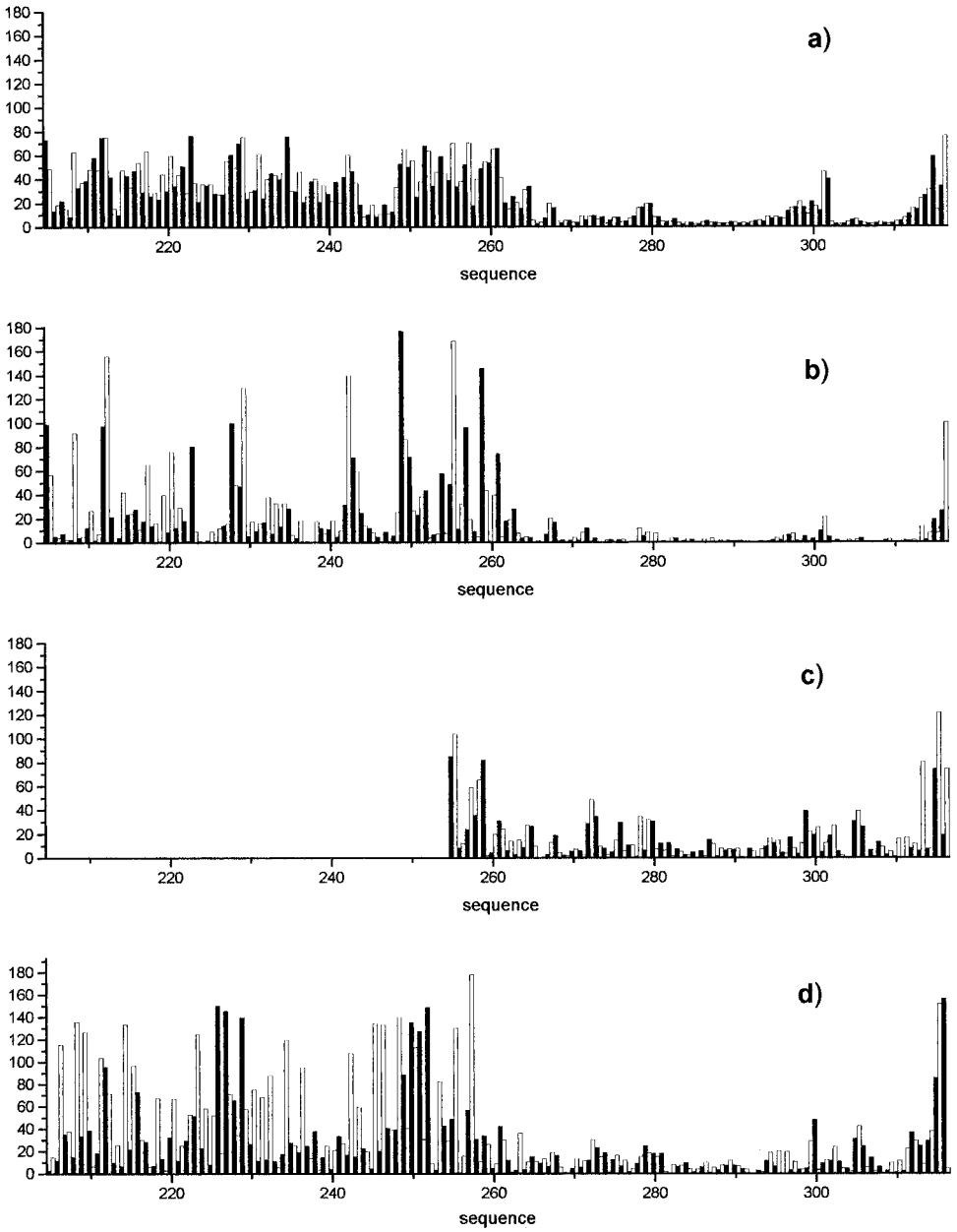

FIGURE 5: Root mean square deviations of the backbone torsional angles ϕ (black) and ψ (white) of the 24 final structures calculated for the 205−316 fragment of thermolysin. Average values for the two subunits (a). Absolute value of the difference between the average torsion angles, ϕ (black) and ψ (white), of the two subunits in the fragment 205−316 (b), of the solution structures of 205−316 and 255−316 fragments (c), and between the solution and crystal structures (d).

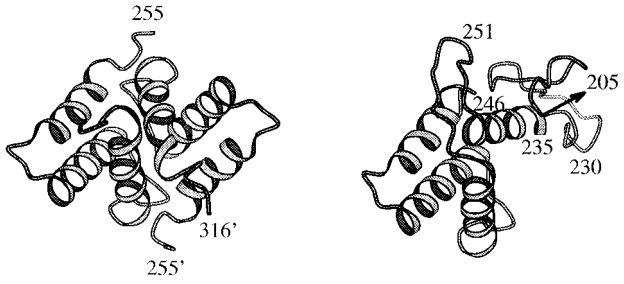

FIGURE 6: Dimeric structure of the 255−316 region of the 205−316 fragment of thermolysin (left) and structure of the 205−316 region in crystalline thermolysin (right) in a ribbon representation constructed with MOLSCRIPT (Kraulis, 1991).

very similar suggesting that they do not differ in helical structure. However, the 205−316 thermolysin fragment contains 14 aromatic residues (three Phe, eight Tyr, and three His) with five of the eight Tyr residues being present in the 205−254 region. The aromatic Tyr, Trp, and (to a smaller extent) Phe contribute to the far UV−CD signal at 222 nm, indicative of helix content, but the dependences of this contribution on helix content and length, on the sequence position, and on the orientation of the aromatic chromophores remain to be determined (Chakrabartty et al., 1993). It is likely that this aromatic contribution in the monomer is different from that in the dimer since near-UV−CD spectra (characteristic of ordered tertiary structure around aromatic residues) suggest that the dimer has a higher content of fixed tertiary structure than the monomer (Conejero-Lara & Mateo, 1996). Therefore, aromatic contributions complicate the interpretation of the CD data so that no clear-cut evidence can be drawn about actual differences in helical structure between the monomer and dimer.

In the native structure of thermolysin, some residues in the region 205−254, in particular those in the segment 235−

246, are in contact with the region 255−316 and bury part of its hydrophobic surface (Figure 6). This burial can contribute to the stability of helix 235−246 in native thermolysin and might also stabilize the same helix in the monomeric 205−316 fragment. In fact, in the calorimetric studies of the dissociation and unfolding of this fragment discussed above (Azuaga et al., 1995; Conejero-Lara et al., 1996), it was proposed that in the monomer the region 205−254 would have some structure and that this structure would probably involve the 235−246 helix. If we assume that the dissociated subunits undergo a rearrangement in their conformations to yield the four-helical structure found in native thermolysin, then the changes in ASA upon dissociation are −307 Å^2 of nonpolar surface (a burial) and +1079 Å^2 of polar surface (a solvent exposure) per dimer. It has been shown that the energetics of folding/unfolding can be estimated from structural information by using the changes in both polar and nonpolar ASA values (Murphy & Freire, 1992). Using that approach based on the dissolution of solid model compounds, the calculated heat capacity change, ΔCp, for the dissociation of the dimeric 205−316 fragment is −0.88 kJ/(K mol of monomer), whereas the dissociation of the dimer as a simple undocking into two three-helix monomers would lead to a ΔCp value of +1.89 kJ/(K mol of monomer). The values are somewhat smaller if expression based on the transfer of model compounds from the liquid state to water are used (Spolar & Record, 1994). The experimental ΔCp value obtained by DSC is -1.1 ± 0.3 kJ/(K mol of monomer) at acid pH (Azuaga et al., 1995). Both the theoretical value (−0.88) and the calorimetric one agree within experimental error, which is consistent with the view that, to all ends and purposes, dissociation of the dimer must be coupled to an increase in the structure of the segment 205−254, as was proposed earlier (Azuaga et al., 1995). The presence of native-like structure in the region 235−246 would sterically hinder the formation of the dimer, by burying the dimerization interface. Therefore, we may conclude that the eventual assembly of the dimer requires some previous unfolding in the overall region 205−254 of the monomer.

Implications for Protein Folding. As seen above, the final native structure of fragment 205−316 of thermolysin is a dimer, in which the subunits show a peculiar feature: the C-terminal half of the molecule has a well-defined globular three-dimensional structure, whereas the N-terminal half is mainly disordered (Figure 4). On the other hand, the starting point in the folding reaction would be the one in which there are small helical populations in the regions corresponding to the native helices. The picture emerging for the folding of the fragment would then be the following: mutual interactions among the three C-terminal helices would stabilize the 255−316 subdomain, which would constitute an initial nucleus of folding; the large hydrophobic surface created by the folding of the 255−316 subdomain would promote the interaction with the rest of the chain and particularly with the native helix 235−246 in order to become stabilized. On increasing the concentration, the subdomain 255−316 would, however, find a higher stabilization by dimerization, burying its exposed hydrophobic surface and simultaneously hampering the interactions with the 205−254 chain, which in turn remains unfolded. This is an example in which dimerization and formation of the native structure compete in the definition and stabilization of the final structure of the fragment. This sort of coupling between partial unfolding and association of two protein fragments may play a general role in folding and molecular recognition processes. Spolar and Record (1994) have reported a number of processes involving protein oligomerization and protein−DNA complexation, where conformational changes coupled to binding could be inferred from a thermodynamic analysis of ΔCp and entropy change associated with the binding. Almost all of their examples involved conformational changes where structure was induced in disordered regions on binding. Spolar and Record (1994) proposed an approach to estimate the number of residues that fold on binding by the dissection of the binding entropy into individual contributions. If we apply this approach to the 205−316 fragment of thermolysin, the result is that overall 14 residues unfold concomitantly upon the dimerization of the fragment although the specific residues involved cannot be identified from the available data. Recently, it has been shown that DNA binding could drive unfolding of an α-helix present in the native protein in studies on the structure and DNA binding of *Bam*HI-endonuclease (Newman et al., 1995) and the transcription factor Ets-1 (Petersen et al., 1995). Therefore, the dimerization-induced partial unfolding we propose to occur in the 205−316 thermolysin fragment may now be added to the above examples.

ACKNOWLEDGMENT

The excellent technical assistance of Apolo Gómez, Cristina López, and Luis de la Vega is gratefully acknowledged. We are indebted to Vincenzo de Filippis and Angelo Fontana (University of Padua) for their introduction to the problem and help in obtaining and purifying the thermolysin fragment.

SUPPORTING INFORMATION AVAILABLE

Table S1, ^{1}H chemical shifts of thermolysin 205−316 fragment (5 pages). Ordering information is given in any current masthead page.

REFERENCES

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229−2246.

Azuaga, A. I., Conejero-Lara, F., Rivas, G., De Filippis, V., Fontana, A., & Mateo, P. L. (1995) *Biochim. Biophys. Acta 1252*, 95−102.

Bartels, C., Xia, T. H., Billeter, M., Güntert, P., & Wüthrich, K. (1995) *J. Biomol. NMR 6*, 1−10.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson.* 65, 355−360

Blanco, F. J., Jiménez, M. A., Herranz, J., Rico, M., Santoro, J., & Nieto, J. L. (1993) *J. Am. Chem. Soc. 115*, 5887−5888.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285−297.

Chakrabartty, A., Kortemme, T., Padmanabhan, S., & Baldwin, R. L. (1993) *Biochemistry 32*, 5560−5565.

Charasekhar, K., Profy, A. T., & Dyson, H. J. (1991) *Biochemistry 30*, 9187−9194.

Conejero-Lara, F., & Mateo, P. L. (1996) *Biochemistry 35*, 3477−3486.

Conejero-Lara, F., De Filippis, V., Fontana, A., & Mateo, P. L. (1994) *FEBS Lett. 344*, 154−156.

Creighton, T. E., Ed. (1992) *Protein Folding*, Freeman WH & Co., New York.

Dalzoppo, D., Vita, C., & Fontana, A. (1985) *J. Mol. Biol. 182*, 331−340.

Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988). *J. Mol. Biol. 201*, 161−200.

Eccles, C., Güntert, P., Billeter, M., & Wüthrich, K. (1991) *J. Biomol. NMR 1*, 111−130.
Fassina, G., Vita, C., Dalzoppo, D., Zamai, M., Zambonin, M., & Fontana, A. (1986) *Eur. J. Biochem. 156*, 221−228.
Fontana, A. (1990) in *Peptides: Chemistry, Structure and Biology* (Rivier J., & Marshall, G. E., Eds.) pp 557−564, Escom, Leiden.
Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
Holmes, M. A., & Matthews, B. W. (1982) *J. Mol. Biol. 160*, 623−639.
Jiménez, M. A., Bruix, M., González, C., Blanco, F. J., Nieto, J. L., Herranz, J., & Rico, M. (1993) *Eur. J. Biochem. 211*, 569−581.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
Marion, D., & Wüthrich, K. (1983). *J. Magn. Reson. 79*, 352−356.
Murphy, K. P., & Freire, E. (1992) *Adv. Prot. Chem. 43*, 313−361.
Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I,. & Aggarwal, A. (1995) *Science 269*, 656−663.
Petersen, J. M., Skalicky, J. J., Donaldson, L. W., McIntosh, L. P., Alber, T., & Graves, B. J. (1995) *Science 269*, 1866−1869.
Richards, F. M. (1977) *Ann. Rev. Biophys. Bioeng. 6*, 151−176.
Richmond, T. J. (1984) *J. Mol. Biol. 178*, 63−89.
Rico, M., Jiménez, M. A., González, C., De Filippis, V., & Fontana, A. (1994) *Biochemistry 33*, 14834−14847.
Spolar, R. S., & Record, M. T. (1994) *Science 263*, 777−784.
Stroup, A. N., & Gierasch, L. M. (1990) *Biochemistry 29*, 9765−9771.
Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1972). *Nature 238*, 35−37.
Van Gunsteren, W. F., & Berendsen, H. J. C. (1987) *Groningen Molecular Simulation (GROMOS) Library Manual*. Biomos, Groningen, The Netherlands.
Vita, C., Fontana, A., Seeman, J. R., & Chaiken, I. M. (1979) *Biochemistry 18*, 3023−3031.
Vita, C., & Fontana, A. (1982) *Biochemistry 21*, 5196−5202.
Vita, C., Dalzoppo, D., & Fontana, A. (1984) *Biochemistry 23*, 5512−5519.
Vita, C., Fontana, A., & Jaenicke, R. (1989) *Eur. J. Biochem. 183*, 513−518.
Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, J. and Sons, New York.
Xia, T. H., & Bartels, C. (1994) *XEASY. ETH automated spectroscopy for X windows systems.*

BI971060T