An NMR-based identification of peptide fragments mimicking the interactions of the cathepsin B propeptide

Youlu Yu[a], Wim Vranken[a], Nathalie Goudreau[a], Elisa de Miguel[c], Marie-Claude Magny[c], John S. Mort[a,c], Robert Dupras[d], Andrew C. Storer[a,d], Feng Ni[a,b,*]

[a] *Protein Engineering Network of Centres of Excellence, Montreal, Que. H4P 2R2, Canada*
[b] *Biomolecular NMR, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Que. H4P 2R2, Canada*
[c] *Joint Diseases Laboratory, Shriners Hospital for Children and Department of Surgery, McGill University, Montreal, Que. H3G 1A6, Canada*
[d] *Enzyme Engineering Laboratories, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Que. H4P 2R2, Canada*

Received 18 February 1998; revised version received 23 April 1998

Abstract **Selected fragments of the 62-residue proregion (or residues 1p–62p) of the cysteine protease cathepsin B were synthesized and their interactions with cathepsin B studied by use of proton NMR spectroscopy. Peptide fragments 16p–51p and 26p–51p exhibited differential perturbations of their proton resonances in the presence of cathepsin B. These resonance perturbations were lost for the further truncated 36p–51p fragment, but remained in the 26p–43p and 28p–43p peptide fragments. Residues 23p–26p or TWQ^{25}A in the N-terminal 1p–29p fragment did not show cathepsin B-induced resonance perturbations although the same residues had strongly perturbed proton resonances within the 16p–51p peptide. Both the 1p–29p and 36p–51p fragments lack a common set of hydrophobic residues 30p–35p or F^{30}YNVDI^{35} from the proregion. The presence of residues F^{30}YNVDI^{35} appears to confer a conformational preference in peptide fragments 16p–51p, 26p–51p, 28p–43p and 26p–43p, but the same residues induce the aggregation of peptides 16p–36p and 1p–36p. The peptide fragment 26p–43p binds to the active site, as indicated by its inhibition of the catalytic activity of cathepsin B. The cathepsin B prosegment can therefore be reduced into smaller, but functional subunits 28p–43p or 26p–43p that retain specific binding interactions with cathepsin B. These results also suggest that residues F^{30}YNVDI^{35} may constitute an essential element for the selective inhibition of cathepsin B by the full-length cathepsin B proregion.**

Key words: Protein-protein interaction; Peptide design; Nuclear magnetic resonance; Cysteine protease

1. Introduction

Cathepsin B is a lysosomal cysteine protease that functions as part of the degradative apparatus within the cell [1]. Active cathepsin B is also implicated in a number of human diseases such as arthritis [2], muscular dystrophy [3], tumor metastasis [4,5] and Alzheimer's pathology [6]. The connection of cathepsin B to human diseases has spurred many investigations on the structure and function of this enzyme and on attempts to design potent and specific inhibitors [7,8]. However, it has been difficult to develop low molecular weight inhibitors highly selective for cathepsin B or other cysteine proteases.

Cathepsin B is derived from an inactive proenzyme with a 62-residue proregion and a 254-residue single-chain protease domain [9]. The active enzyme is generated following proteolytic removal of the prosegment under acidic conditions in the lysosome [10,11]. At neutral pH, the cleaved profragment remains non-covalently associated with the protease domain with a K_i of $\sim$0.4 nM [12], a phenomenon largely responsible for the inhibition and stabilization of the mature cathepsin B transported into the extracellular space [13]. The propeptide-cathepsin B interaction is very specific since the cathepsin B propeptide has a 14 000 times lower affinity for binding to papain, a protease belonging to the same cysteine protease superfamily. The crystal structures of procathepsin B determined recently indicate that the proregion folds onto the surface of the protease domain through multiple inter-domain contacts [14,15]. The cathepsin B profragment therefore utilizes a number of distinctive regions on the surface of cathepsin B to achieve the remarkable affinity and specificity for its own protease domain.

The unique mode of cathepsin B inhibition by its proregion offers possibilities to design highly selective inhibitors of cathepsin B. For example, one may construct small peptides to bind specifically to any one of the four predominant binding surfaces on cathepsin B (Table 1): the S and the S′ subsites of the active-site cleft (denoted as A and B), the occluding loop at the end of the S′ subsites (denoted as O) and the site external to the active-site cleft (denoted as E). Previously, critical segments of the proregion, namely residues 21p–46p [16] or residues 26p–50p [17], have been identified as required for high-affinity inhibition of cathepsin B by its propeptide. In this work, we used NMR spectroscopy to characterize the interactions of cathepsin B with an array of peptide fragments spanning residues 1p–51p of procathepsin B. We address the question whether the propeptide of cathepsin B can be reduced to even smaller subunits that retain some of the specific binding interactions of the full-length propeptide. Such peptide fragments, if available, may serve as starting points for the development of specific inhibitors of cathepsin B and related cysteine proteases.

2. Materials and methods

2.1. Preparation of proteins and peptides

Recombinant rat procathepsin B was expressed in the yeast *Pichia*

*Corresponding author. Fax: (1) (514) 496-5143.
E-mail: Feng.Ni@nrc.ca

PII S0014-5793(98)00548-1

Table 1
Summary of NMR results with selected fragments of the cathepsin B propeptide

HDKPS$_5$FHPLS$_{10}$DDMIN$_{15}$YINKQ$_{20}$NTTWQ$_{25}$AGRNF$_{30}$YNVDI$_{35}$SYLKK$_{40}$LCGTV$_{45}$LGGPK$_{50}$LPERVG	a
EEE E O OOO BBBB BBAAA AA	b
----- ----h hhhhh hhhhh ----- ----- ----h hhhhh hh--- ----- ------	c
(16p-51pR) *RRR*KQ$_{20}$NT***TW***Q$_{25}$***A***GRN***F***$_{30}$***YN***VD***I***$_{35}$***SYLKK***$_{40}$***LCG***T***V***$_{45}$***LGG***P***K***$_{50}$*R*	d
(26p-51p) AGRN***F***$_{30}$***YN***VDI$_{35}$S***YLKK***$_{40}$***LCG***TV$_{45}$LGGPK$_{50}$L	d
(36p-51p) SYLKK$_{40}$LCGTV$_{45}$LGGPK$_{50}$L	e
(28p-43p) RN***F***$_{30}$***YN***VDI$_{35}$SY***LKK***$_{40}$***LCG***	d
(26p-43p) AGRN***F***$_{30}$***YN***VDI$_{35}$SY***LKK***$_{40}$***LCG***	d
*RR*NKQ$_{20}$NTTWQ$_{25}$AGRNF$_{30}$YNVDI$_{35}$S (16p-36pR)	f
HDKPS$_5$FHPLS$_{10}$DDMIN$_{15}$YINKQ$_{20}$NTTWQ$_{25}$AGRNF$_{30}$YNVDI$_{35}$S (1p-36p)	f
HDKPS$_5$F***H***PLS$_{10}$DDMIN$_{15}$YINKQ$_{20}$NTTWQ$_{25}$AGRN (1p-29p)	e

[a]The amino acid sequence of the propeptide fragment 1p–56p of rat procathepsin B. This peptide has a K_i of 0.4 nM for the inhibition of cathepsin B at pH 6 [12].
[b]Residues of the proregion of rat procathepsin B in contact with various binding sites on cathepsin B: E, residues in contact with a site external to the catalytic active site; O, residues in contact with a surface loop (the occluding loop) at the end of the S′ sites of cathepsin B; B, residues located in the S′ subsites of cathepsin B; A, residues located in the S subsites of cathepsin B [14].
[c]Residues found in helical (h) conformations in the structure of rat procathepsin B [14].
[d]Cathepsin B binding-induced resonance perturbations were observed for peptides 16p–51pR, 26p–51p, 28p–43p and 26p–43p with significantly perturbed residues indicated in italicized bold. These residues had pronounced peak shifts, line broadening and/or intensity reductions for the NH proton and/or the side-chain proton resonances in the presence of cathepsin B (see Figs. 1–3). Peptide 28p–43p was prepared in the N-terminally acetylated form to minimize the effect of a free N-terminus on interactions with cathepsin B.
[e]Peptide 36p–51p did not show resonance perturbations in the presence of cathepsin B, but was instead cleaved by the cathepsin B mutant (see Section 2), as shown by the progressive decrease of the original proton resonances and the appearance of additional resonances. Peptide 1p–29p exhibited perturbations of only the side-chain proton resonances of the two His residues with increasing concentrations of cathepsin B.
[f]Peptides 16p–36pR and 1p–36p had very low solubilities in aqueous solution and showed broad and unresolved proton resonances, making it impossible to observe binding-induced resonance perturbations.

pastoris [18]. Residue Cys[29] in the active site was mutated to Ser to allow the production of a catalytically inactive form that is still able to bind the propeptide. The processed enzyme was purified by ion exchange chromatography using an SP-Sepharose fast flow column in 20 mM sodium acetate, pH 5.0 containing 1 mM EDTA with a 0–1.0 M NaCl gradient for protein elusion. Final purification was achieved by gel filtration chromatography through a Sepharose S-200 HR column in the same buffer containing 0.1 M NaCl. N-terminal sequencing demonstrated that the enzyme consisted of a mixture of processed forms with three- or six-residue N-terminal extensions relative to the fully processed lysosomal form as shown previously for intermolecular processing of procathepsin B [19]. The purified enzyme was concentrated to ~0.6 mM by use of a stirred-cell concentrator followed by concentration using an Amicon Centricon-10 concentrator. Prior to NMR experiments, the concentrated enzyme solution was dialysed at 4°C against the desired buffer solutions used in NMR experiments.

Selected peptide fragments of rat procathepsin B were synthesized by use of standard Fmoc chemistry on an Applied Biosystem 431A solid-phase synthesizer. The synthetic peptides were purified to homogeneity by HPLC on a Vydak C4 or C18 column using a linear 10–70% acetonitrile gradient. Weighed amounts of the purified peptides were dissolved in 450 μl of aqueous solutions containing, depending on the peptide, between 0.2–1.5 mM of the peptide, 100 mM NaCl, 20–50 mM sodium phosphate or acetate, 0.2–0.5 mM EDTA and 5 mM DTT with or without 3% DMSO. D_2O was added at 5% to each of these solutions to provide a deuterium lock signal for the NMR spectrometer. The pH values of these solutions were adjusted by titration with dilute NaOH or HCl solutions. All peptide solutions were prepared in an ice bath to avoid aggregation and precipitation of the peptides, which occurs at higher temperatures. Only freshly prepared peptide samples were used for NMR and binding experiments.

2.2. NMR experiments

All NMR experiments were carried out on a Bruker AMX-500 and a Bruker AM-500 or DRX-500 MHz spectrometers at temperatures between 10 and 25°C. Two-dimensional double quantum-filtered COSY (DQF-COSY), TOCSY, NOESY and/or ROESY spectra were acquired using time proportional phase incrementation (TPPI) and with sine modulation along the t_1 direction [20–22]. A scheme of matched acquisition was employed for DQF-COSY to improve the sensitivity of the experiment [23]. TOWNY-16 was used in the TOCSY experiments with an effective spin-lock power of 2.5–3 kHz [24]. A total of 350–400 FIDs were acquired with 2048 data points for each FID. The assignment of the proton resonances to specific residues (Tables 2–7) in the propeptide fragments 16p–51p, 26p–51p, 36p–51p, 28p–43p, 26p–43p and 1p–29p was achieved using standard procedures [20].

The interactions between peptide fragments and cathepsin B were followed by binding-induced resonance perturbations on the peptide proton resonances, including peak shifts, reduction of resonance intensities, relaxation enhancements and/or line broadening effects [25]. The proton NMR spectra of the propeptide fragments were recorded prior to the addition of the cathepsin B solution. Small volume increments (10–25 μl) of a concentrated cathepsin B solution were added into the peptide solution until the ratio of [peptide]/[cathepsin B] reached 10:1. NMR spectra were acquired ca. 30 min after each addition of cathepsin B. To ensure that the pH values remained the same after addition of cathepsin B, the pH values of both the peptide and the cathepsin B solutions were kept identical with identical salt compositions. The NMR titration experiments were carried out at either 10°C or 15°C and at pH values between 5 and 6.5. Resonance perturbations were identified through spectral overlays and, in the absence of peak shifts or line broadening effects, by the determination of the integral intensities of unique peptide resonances with each addition of cathepsin B.

Table 2
^{1}H chemical shifts of peptide 16p–51pR in water (pH 5.5, 50 mM sodium phosphate, 1 mM EDTA, 100 mM NaCl and 5 mM DTT-d_{10}) at 15°C

Residue	NH	C_αH	C_βH	Others
Arg^{16}	–	3.96	1.80, 1.80	γ1.52; δ3.09; ε7.09
Arg^{17}	8.79	4.26	1.53, 1.53	γ1.68, 1.68; δ3.08; ε7.15
Arg^{18}	8.56	4.20	1.51, 1.51	γ1.66, 1.66; δ3.05; ε7.11
Lys^{19}	8.45	4.14	1.65, 1.65	γ1.30; δ1.55; ε2.86
Gln^{20}	8.48	4.20	1.84, 1.94	γ2.24, 2.24; ε6.80, 7.46
Asn^{21}	8.55	4.64	2.67, 2.70	δ6.85, 7.54
Thr^{22}	8.18	4.36	4.09	γ1.01
Thr^{23}	8.11	4.21	4.08	γ1.02
Trp^{24}	8.17	4.47	3.11, 3.11	δ1, 7.10; ε3, 7.32; ζ3, 7.06; η2, 6.96; ζ2, 7.40; ε1, 10.02
Gln^{25}	7.93	3.95	1.61, 1.75	γ1.89, 1.96; ε6.72, 7.29
Ala^{26}	8.02	3.94	1.21	–
Gly^{27}	8.23	3.76, 3.82	–	
Arg^{28}	7.95	4.12	1.58, 1.71	γ1.45, 1.45; δ 3.06; ε6.98
Asn^{29}	8.30	4.50	2.50, 2.50	δ6.81, 7.42
Phe^{30}	8.17	4.34	2.82, 2.82	δ7.15; ζ7.15; ε 6.93
Tyr^{31}	8.01	4.37	2.56, 2.75	δ6.94; ε6.68
Asn^{32}	8.18	4.52	2.52, 2.65	δ6.82, 7.48
Val^{33}	7.90	3.96	1.93	γ0.77, 0.77
Asp^{34}	8.29	4.53	2.67, 2.67	–
Ile^{35}	8.15	4.00	1.82	γ11.26, 1.09; γ2 0.74; δ 0.74
Ser^{36}	8.36	4.13	3.75, 3.75	–
Tyr^{37}	7.87	4.30	2.95, 2.95	δ6.98; ε6.69
Leu^{38}	7.68	3.97	1.51, 1.51	γ1.38; δ0.71, 0.76
Lys^{39}	7.98	4.05	1.53, 1.53	γ1.22; δ1.48; ε2.94
Lys^{40}	8.04	4.16	1.57, 1.69	γ1.35; δ 1.45; ε 2.93
Leu^{41}	8.36	4.25	1.54, 1.54	γ1.46; δ 0.73, 0.73
Cys^{42}	8.23	4.37	2.83, 2.83	–
Gly^{43}	8.30	3.89, 3.89	–	
Thr^{44}	7.98	4.24	4.05	γ1.05
Val^{45}	8.18	3.99	1.93	γ0.78, 0.78
Leu^{46}	7.88	4.09	1.69, 1.69	γ1.32, 1.32; δ0.77, 0.77
Gly^{47}	8.35	3.85	–	
Gly^{48}	8.06	3.93, 4.01	–	
Pro^{49}	–	4.30	1.88, 2.15	γ1.80, 1.80; δ3.48
Lys^{50}	8.40	4.17	1.71, 1.63	γ1.34, 1.35; δ1.55; ε2.88
Arg^{51}	7.95	4.04	1.56, 1.71	γ1.46, 1.46; δ3.04; ε7.10

3. Results

3.1. Interactions of peptide 16p–51pR with cathepsin B

The 16p–51p fragment contains all the sequence elements of the proregion in contact with various binding sites on the cathepsin B domain in procathepsin B (Table 1). To increase solubility and facilitate NMR experiments, the 16p–51p peptide was modified to incorporate four arginine residues replacing 16p, 17p, 18p and 51p of the native sequence (Table 1). For this 16p–51pR peptide, there were differential perturbations of many of the proton resonances after the addition of cathepsin B (Fig. 1), indicating binding. Fig. 2 shows the fingerprint region of a TOCSY spectrum, correlating the amide proton with the αCH proton resonance of every residue in peptide 16p–51pR before (in blue) and after (in red) the addition of cathepsin B (10:1 peptide/enzyme ratio). Many of the correlation peaks were not perturbed upon binding (except for slight broadening resulting from lowered contouring, Fig. 2), such as those of residues Arg-18p to Thr-23p and Arg-51p at both ends of the peptide. In contrast, peaks with significant shifts corresponded to resonances from within the I^{35}SYLKK^{40}LCG and the V^{45}LGGPK^{50} sequences. In addition, residues Asp-34p and Thr-44p had minor peak shifts and some line broadening effects. Interestingly, residues Leu-38p, Lys-39p and Lys-40p had the most dramatically perturbed cross peaks along both the amide and the αCH proton chemical shift dimensions. There were almost no perturbations on the NH-αCH correlations of residues in the TWQ^{25}AGRNF^{30}YNV sequence. These same residues did exhibit binding-induced changes in their side-chain proton resonances (Fig. 1). The most obvious are perturbations on the side-chain resonances of Thr-23p, Trp-24p, Ala-26p, Phe-30p, Tyr-31p and on one of the two side-chain δNH protons of Asn-32p in contrast to the absence of perturbations on the same protons of the other two Asn residues, Asn-21p and Asn-29p. The residues in the 16p–51pR peptide perturbed by cathepsin B binding thus closely match the contact residues between the proregion and the protease domain observed in the X-ray structure of procathepsin B (Table 1).

3.2. Binding of the 26p–51p and 36p–51p peptides to cathepsin B

Peptide 26p–51p contains the F^{30}YNV and the YLKK^{40}LCGTV^{45}LG regions of the prosegment that may still bind to the occluding loop and the active-site regions of cathepsin B, respectively (Table 1). Peptide 36p–51p only has the propeptide residues in contact with the active-site (S′- and S-) regions of cathepsin B. For many of the residues of the 26p–51p peptide, line broadening was again observed as a result of cathepsin B binding. The observed effects on the 26p–51p peptide were similar to the 16p–51pR peptide (Table 1), except for ambiguities with residues V^{45}LGGPK^{50}L due to resonance overlaps.

The 36p–51p peptide produced well-resolved and sharp proton resonances, but no resonance perturbations were detected in the presence of cathepsin B, in contrast to the 16p–51pR and 26p–51p peptides. It is possible that binding of cathepsin B to 36p–51p does not induce resonance perturbations as a result of higher affinity and slower exchange [25], but this is unlikely considering that 36p–51p is a subfragment of the 26p–51p peptide. Therefore, the mode of interactions for residues 36p–51p within the longer 26p–51p peptide is almost certainly not present in the truncated 36p–51p peptide. Instead, peptide 36p–51p was slowly cleaved by the inactivated cathepsin B with the mutation of the active site residue Cys^{29} to a Ser residue. Interestingly, peptides 16p–51pR, 26p–51p and 36p–51p were all cleaved by the cathepsin B mutant, although the cleavage rates became significantly reduced with longer peptides: 16p–51pR (no cleavage after days of incubation) $\ll$26p–51p (cleaved within hours) $\ll$36p–51p (cleaved within minutes). In other words, the rate of cleavage drops as the F^{30}YNVDI^{35} and TWQ^{25}A sequences are progressively included in the peptide fragments. The lack of binding-induced resonance perturbations for peptide 36p–51p points to the importance of the F^{30}YNVDI^{35} sequence for the interactions of the longer peptides 16p–51pR and 26p–51p with cathepsin B.

3.3. The 26p–43p and 28p–43p peptides retain specific binding to cathepsin B

Since the 26p–51p fragment still retained resonance perturbations by cathepsin B similar to those observed with peptide 16p–51pR (Table 1), the next step was to assess whether C-terminal truncated fragments of this peptide were still able to interact with cathepsin B. The truncated peptide 26p–43p still contains residues for binding to the occluding loop (the O site) and residues mostly located in the S′ subsites of procathepsin

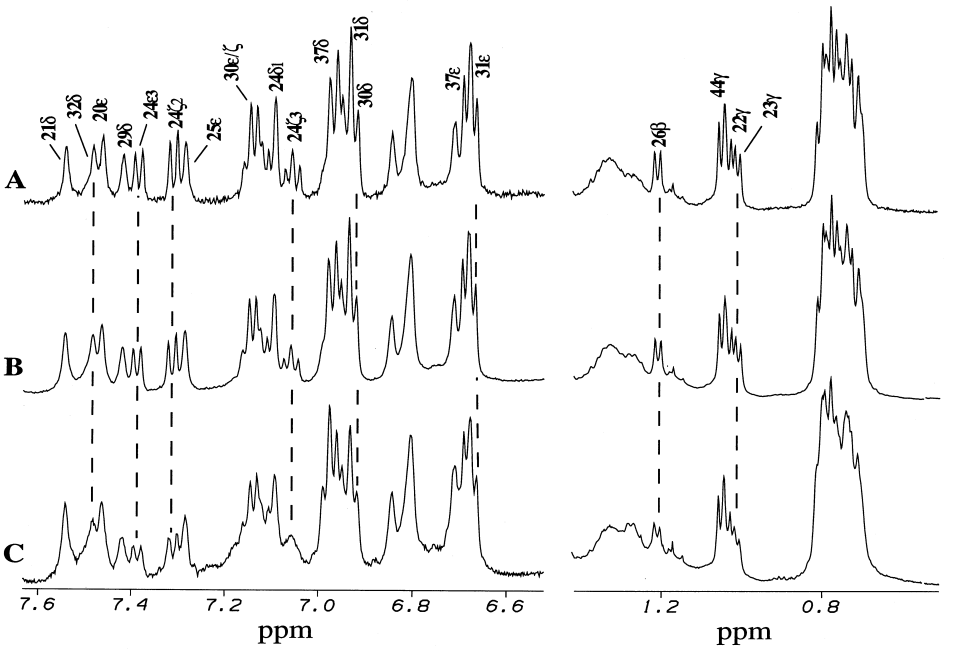

Fig. 1. The aromatic (left panel) and aliphatic (right panel) regions of the proton NMR spectra of 16p–51pR in the absence (A) and in the presence of cathepsin B (B and C) at pH 6. The ratio of [16p–51pR]/[cathepsin B] was 20:1 in B and 10:1 in C with a peptide concentration of 0.5 mM. Proton resonances are labeled by the residue numbers (Table 1) followed by the positions of the attached carbon atoms of each residue. Indicated by dashed lines are those resonances (and residues) with significant perturbations in their linewidths and/or peak intensities. Similar, but less pronounced resonance perturbations were observed at pH 5.5.

B. The 28p–43p peptide was synthesized and its interaction with cathepsin B examined since the Ala-26p and Gly-27p residues in 26p–51p did not show resonance perturbations (Table 1).

Fig. 3 shows that cathepsin B binding still induced significant perturbations on the NH resonances of Lys-39p and Lys-40p of the 28p–43p peptide. The same two residues exhibit the largest binding-induced resonance shifts when they are part of the 16p–51pR peptide (Fig. 2). Line broadening was also observed for the NH resonances of residues Leu-38p, Leu-41p and Cys-42p in the 28p–43p peptide. In addition, the side-chain ring protons of residues Phe-30p, Tyr-31p and Tyr-37p were broadened in the presence of cathepsin B (right panel of Fig. 3). Of the two Asn residues in the sequence, only the side chain NH resonances of Asn-32p were significantly broadened, while Asn-29p was not much affected. The similar patterns of differential resonance perturbations for peptides 16p–51pR, 26p–51p and 28p–43p (Table 1) suggest that a similar binding mode is adopted by these peptide fragments in their complexes with cathepsin B. The distinctive resonance broadening of Asn-32p reflects its unique environ-

Table 3
^{1}H chemical shifts of peptide 26p–51p in water (pH 5.5, 50 mM sodium phosphate, 1 mM EDTA, 10 mM DTT-d_{10}) at 15°C

Residue	NH	$C_\alpha H$	$C_\beta H$	Others
Ala^{26}	–	3.57	1.07	–
Gly^{27}	8.60	3.92, 3.92	–	
Arg^{28}	8.33	4.16	1.60, 1.60	γ1.45, 1.49; δ3.04; ε7.08
Asn^{29}	8.41	4.52	2.57, 2.57	δ7.49, 6.85
Phe^{30}	8.16	4.40	2.79, 2.92	δ7.02; ε7.23; ζ7.20
Tyr^{31}	8.04	4.40	2.79, 2.93	δ6.93; ε6.68
Asn^{32}	8.21	4.52	2.56, 2.68	δ7.51, 6.86
Val^{33}	7.99	3.97	1.99	γ0.82, 0.82
Asp^{34}	8.34	4.55	2.57, 2.70	–
Ile^{35}	8.17	4.02	1.86, 1.86	γ1 1.13, 1.33; γ2 0.84; δ0.79
Ser^{36}	8.36	4.16	3.77, 3.77	–
Tyr^{37}	7.89	4.32	2.99, 2.99	δ7.01; ε 6.73
Leu^{38}	7.68	4.01	1.56, 1.55	γ1.45, δ 0.74, 0.79
Lys^{39}	7.96	4.07	1.72, 1.70	γ1.39, 1.31; δ1.54; ε2.87
Lys^{40}	7.91	4.10	1.73, 1.69	γ1.37, 1.31; δ1.56; ε2.90
Leu^{41}	8.05	4.25	1.57, 1.52	γ1.52; δ0.76, 0.81
Cys^{42}	8.20	4.40	2.86, 2.86	–
Gly^{43}	8.33	3.91, 3.91	–	
Thr^{44}	8.00	4.25	4.08	γ1.08
Val^{45}	8.19	4.01	1.98	γ0.83, 0.83
Leu^{46}	8.37	4.18	1.58, 1.57	γ1.36; δ0.83, 0.77
Gly^{47}	8.37	3.87, 3.88	–	
Gly^{48}	8.06	3.94, 4.07	–	
Pro^{49}	–	4.34	2.17, 1.90	γ1.78, 1.81; δ3.52
Lys^{50}	8.37	4.26	1.76, 1.67	γ1.37, 1.36; δ1.57; ε2.92
Leu^{51}	7.92	4.11	1.56, 1.49	γ1.38, 1.30; δ0.80, 0.75

Table 4
^{1}H chemical shift of peptide 36p–51p in water (pH 5.5, 50 mM sodium phosphate, 1 mM EDTA and 10 mM DTT-d_{10}) at 15°C

Residue	NH	$C_\alpha H$	$C_\beta H$	Others
Ser^{36}	–	4.31	3.52	–
Tyr^{37}	8.6	4.52	2.91, 2.92	δ7.03, ε6.74
Leu^{38}	8.11	4.17	1.43, 1.40	γ1.43, δ0.74, 0.76
Lys^{39}	8.13	4.11	1.68, 1.63	γ1.33; δ1.57; ε2.90
Lys^{40}	8.26	4.18	1.68, 1.65	γ1.34; δ1.58; ε2.90
Leu^{41}	8.34	4.28	1.54, 1.50	γ1.49; δ0.78, 0.83
Cys^{42}	8.36	4.41	2.84, 2.86	–
Gly^{43}	8.46	3.92, 3.92	–	
Thr^{44}	8.04	4.26	4.06	1.09
Val^{45}	8.23	4.02	1.97	0.83
Leu^{46}	8.40	4.26	1.58, 1.51	?, δ 0.77, 0.83
Gly^{47}	8.41	3.87, 3.87	–	
Gly^{48}	8.06	3.94, 4.06	–	
Pro^{49}	–	4.32	1.92, 2.18	γ1.79; δ3.52
Lys^{50}	8.37	4.21	1.75, 1.64	γ1.36; δ1.56; ε2.90
Leu^{51}	7.90	4.10	1.50, 1.50	γ1.49; δ0.76, 0.80

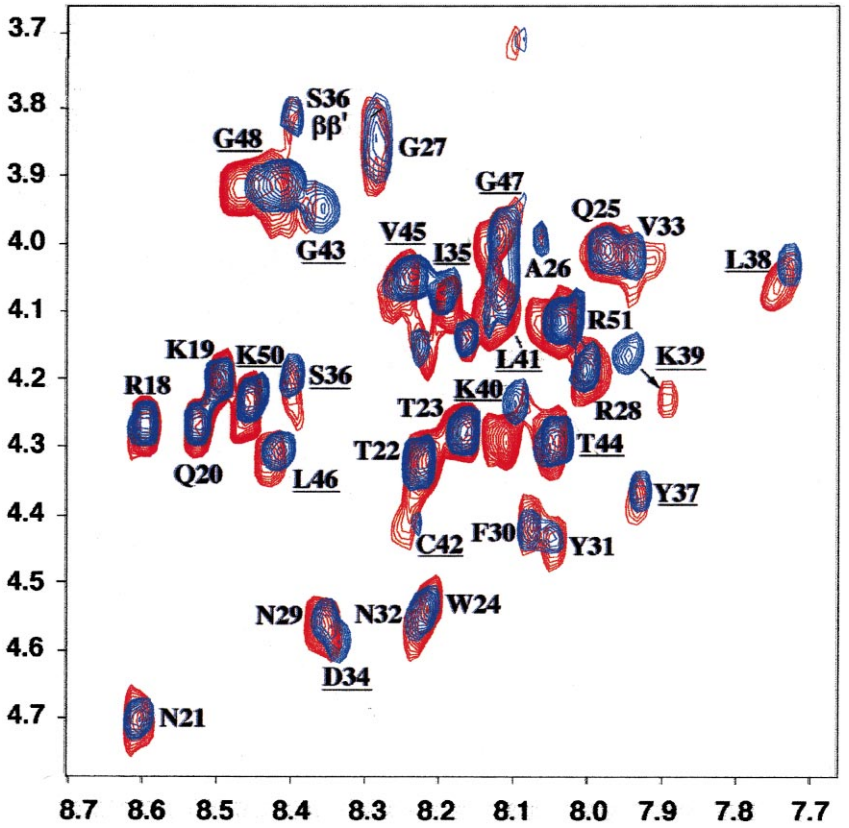

Fig. 2. The fingerprint regions of the two-dimensional TOCSY spectra of 16p–51pR in the absence (blue) and in the presence (red) of cathepsin B (at a 10:1 ratio of [16p–51pR]/[cathepsin B]). The spectra were recorded at pH 5.5 and at 15°C with a peptide concentration of 0.5 mM. Lower contour levels (red) were chosen in the presence of cathepsin B to better visualize broadened peaks with reduced intensity. The amide to the αCH proton correlations are labelled by the corresponding residues with their one-letter codes followed by the residue numbers (Table 1). Underlined residues are those with significant perturbations (peak shifts and/or line broadening effects) of their NH-αCH cross peaks.

ment in the complexes of cathepsin B with all propeptides containing the sequence $F^{30}YNVDI^{35}$.

The 26p–43p peptide, as expected, binds to cathepsin B similarly to peptide 28p–43p (Table 1). The inclusion of the two N-terminal residues Ala-26p and Gly-27p in the 26p–43p peptide increased the solubility and decreased the peptide precipitation observed with 28p–43p. The binding properties of the 26p–43p peptide to cathepsin B were then examined further by inhibition assays against the catalytic properties of active cathepsin B measured using the substrate CBZ-Phe-Arg-MCA [16]. Analysis of the inhibition data yielded an estimated inhibition constant (K_i) of ~60 μM for 26p–43p. The binding of 26p–43p to cathepsin B is therefore rather weak, but still significant since the C-terminally extended 26p–56p fragment only had a K_i value of ~10 μM [16]. The short 26p–43p fragment, which includes mostly propeptide residues in contact with the S′ subsites of cathepsin B (Table 1), therefore preserves the binding mode of the cathepsin B propeptide. The resonance perturbations and inhibitory activity observed for peptide 26p–43p are in sharp contrast to the lack of resonance perturbations with peptide 36p–51p, suggesting again that residues $F^{30}YNVDI^{35}$ may be responsible for the unique mode of binding of peptide 26p–43p (or 28p–43p) and other peptides containing the 30p–35p propeptide residues.

Apart from restoring the binding interactions, the presence of the $F^{30}YNVDI^{35}$ sequence in propeptide fragments may also induce the formation of secondary structures required for proper binding to the cathepsin B active site. An analysis of the proton resonance assignments (Tables 2–7) shows that the NH and αCH chemical shifts of the $F^{30}YNVDI^{35}SYLKK^{40}LCG$ sequence are comparable for the 16p–51pR, 26p–51p, 28p–43p and 26p–43p peptides, but differ systematically for the $YLKK^{40}LCG$ sequence in the

Table 5
^{1}H chemical shifts of peptide 28p–43p in water (pH 5.5, 50 mM sodium phosphate, 0.5 mM EDTA and 5 mM DTT-d_{10}) at 15°C

Residue	NH	$C_\alpha H$	$C_\beta H$	Others
Ac		?		
Arg^{28}	8.26	4.15	1.61, 1.63	γ1.50, 1.48; δ3.10; ε6.8, 7.14
Asn^{29}	8.23	4.58	2.62, 2.74	δ6.88, 7.55
Phe^{30}	8.21	4.48	2.86, 2.98	δ7.28; ε6.08; ζ7.25
Tyr^{31}	8.04	4.47	2.85, 3.00	δ7.04; ε6.76
Asn^{32}	8.41	4.59	2.59, 2.59	δ6.85, 7.51
Val^{33}	7.93	4.04	2.05	γ0.87, 0.87
Asp^{34}	8.34	4.61	2.60, 2.73	–
Ile^{35}	8.15	4.12	1.91	γ1 1.16, 1.55; γ2 0.84; δ0.84
Ser^{36}	8.39	4.24	3.81, 3.81	–
Tyr^{37}	7.93	4.41	3.02, 3.02	ε6.79; δ7.09
Leu^{38}	7.72	4.11	1.60, 1.58	γ1.50, 1.48; δ0.80, 0.83
Lys^{39}	8.03	4.22	1.73, 1.80	γ1.63; δ1.38, 1.42; ε2.94
Lys^{40}	7.98	4.17	1.73, 1.78	γ1.63; δ.37, 1.40; ε2.94
Leu^{41}	8.16	4.30	1.65, 1.65	γ1.59, 1.59; δ0.82, 0.85
Cys^{42}	8.21	4.50	2.91, 2.93	–
Gly^{43}	8.18	4.12, 4.12	–	

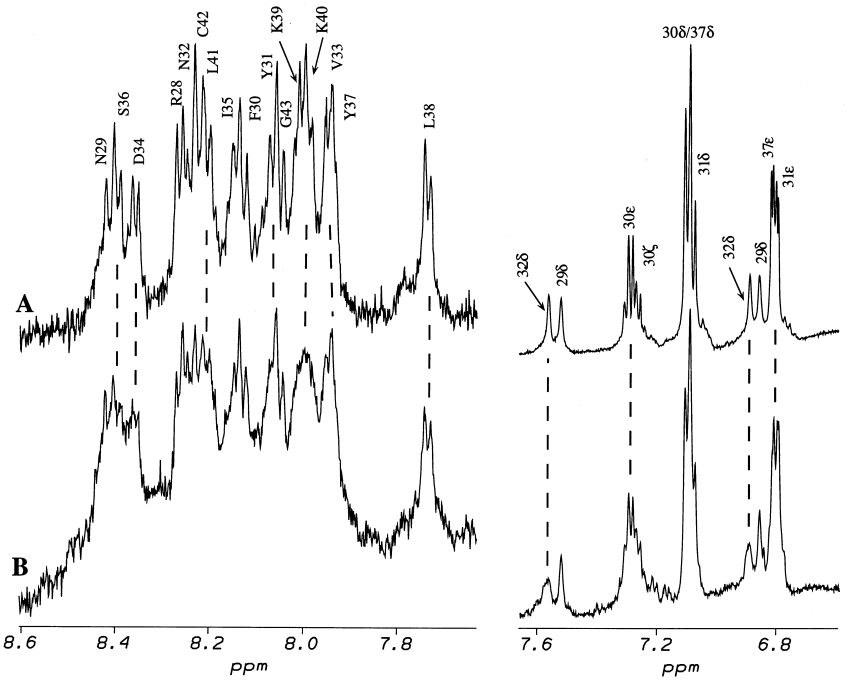

Fig. 3. The NH (left panel) and aromatic (right panel) regions of the one-dimensional proton NMR spectra of the propeptide fragment 28p–43p in the absence (A) and in the presence (B) of cathepsin B. The ratio of [28p–43p]/[cathepsin B] was kept at 10:1 with a peptide concentration of 0.5 mM at pH 5.5. Indicated by dashed lines are those resonances (and residues) with the most significant line broadening effects induced by cathepsin B binding.

36p–51p peptide. This suggests that truncation after Ile-35p, as in peptide 36p–51p, may interrupt the local conformations of the F^{30}YNVDI^{35}SYLKK^{40}LCG fragment. In the 26p–43p peptide, NOESY spectra showed that helix-like conformations, indicated by medium-range backbone-to-backbone NOEs, are present in the I^{35}SYLKK^{40} fragment. Therefore, the I^{35}SYLKK^{40} sequence has a strong tendency to maintain helical conformations in peptides 16p–51pR, 26p–51p, 26p–43p and 28p–43p as well as in the intact procathepsin B (Table 1).

3.4. Behavior of the 16p–36p, 1p–36p and 1p–29p peptides

The apparent importance of residues 26p–35p, in particular the F^{30}YNVDI^{35} segment, prompted an examination of the behavior of peptides 16p–36pR and 1p–36p (Table 1) containing both the TWQ^{25}A and the F^{30}YNVDI^{35} sequences. Unexpectedly, the removal of residues 37p–51p in 16p–36pR

Table 6
^{1}H chemical shifts of peptide 26p–43p in water (pH 5.5, 50 mM sodium phosphate, 0.5 mM EDTA and 5 mM DTT-d_{10}) at 5°C

Residue	NH	C_αH	C_βH	Others
Ala^{26}		4.02	1.42	–
Gly^{27}	8.68	3.91	–	
Arg^{28}	8.44	4.15	1.58	γ1.44, 1.38; δ3.04; ε7.08
Asn^{29}	8.51	4.55	2.57	δ6.91, 7.57
Phe^{30}	8.27	4.39	2.81, 2.91	δ7.02; ε7.21; ζ7.19
Tyr^{31}	8.12	4.40	2.79, 2.92	δ7.01; ε6.71
Asn^{32}	8.27	4.53	2.68, 2.55	δ6.90, 7.60
Val^{33}	8.04	3.97	1.99	γ0.82
Asp^{34}	8.41	4.54	2.67, 2.53	–
Ile^{35}	8.22	4.08	1.86	γ1 1.10, 1.30; γ2 0.76; δ0.76
Ser^{36}	8.44	4.20	3.75	–
Tyr^{37}	8.00	4.37	2.98, 2.91	ε6.72; δ7.02
Leu^{38}	7.76	4.05	1.40, 1.31	γ1.50; δ0.78, 0.74
Lys^{39}	8.04	4.11	1.72, 1.67	γ1.37, 1.30; δ1.57; ε2.87
Lys^{40}	8.16	4.15	1.74, 1.67	γ1.33, 1.29; δ 1.57; ε2.87
Leu^{41}	8.30	4.24	1.58, 1.54	γ1.58; δ0.82, 0.76
Cys^{42}	8.29	4.44	2.85	–
Gly^{43}	8.07	3.66	–	

Table 7
^{1}H chemical shifts of peptide 1p–29p in water (pH 5.5, 50 mM sodium phosphate, 0.5 mM EDTA) at 15°C

Residue	NH	C_αH	C_βH	Others
His^{1}	–	?	?	?
Asp^{2}	?	4.61	2.64, 2.50	–
Lys^{3}	8.61	4.56	1.79, 1.68	γ1.44; δ1.65; ε2.95
Pro^{4}	–	4.37	2.23, 1.78	γ1.97; δ1 3.79; δ2 3.61
Ser^{5}	8.40	4.34	3.75	–
Phe^{6}	8.27	4.54	2.97, 2.93	δ7.14; ε7.25
His^{7}	8.30	4.81	3.08, 2.96	δ1 7.15; ε2 8.49
Pro^{8}	–	4.27	2.23, 1.89	δ1 3.48; δ2 3.36; γ1.89
Leu^{9}	8.52	4.30	1.64	γ1.53; δ1 0.89; δ2 0.84
Ser^{10}	8.20	4.36	3.93, 3.80	–
Asp^{11}	8.45	4.49	2.64, 2.59	–
Asp^{12}	8.26	4.49	2.66, 2.60	–
Met^{13}	8.13	4.38	2.05	γ1 2.57; γ2 2.47
Ile^{14}	8.01	3.91	1.79	γ11.44; γ2 1.11; γ2, δ0.74
Asn^{15}	8.26	4.58	2.69	δ1 7.68; δ2 6.92
Tyr^{16}	8.01	4.37	3.03, 2.97	δ7.04; ε6.75
Ile^{17}	8.02	3.87	1.77	γ1 1.45, 1.11; γ2 0.80; δ0.80
Asn^{18}	8.30	4.56	2.79, 2.71	δ1 7.61; δ2 6.93
Lys^{19}	8.14	4.15	1.78, 1.70	γ11.36; γ2 1.29; δ1.57; ε2.88
Gln^{20}	8.24	4.18	2.03, 1.93	γ2.24; δ1 7.37; δ2 6.82
Asn^{21}	8.32	4.69	2.81, 2.73	δ1 7.61; δ2 6.91
Thr^{22}	8.16	4.30	4.19	γ1.115
Thr^{23}	8.13	4.26	4.16	γ1.113
Trp^{24}	8.17	4.51	3.21	δ1 7.20; ε1 10.12; ε3 7.48; ζ2 7.41; ζ3 7.07; η2 7.17
Gln^{25}	7.93	4.01	1.84, 1.70	γ1 2.08; γ2 2.02
Ala^{26}	8.04	4.01	1.31	–
Gly^{27}	8.33	3.86	–	
Arg^{28}	8.03	4.30	1.80, 1.64	γ1.51; δ3.04
Asn^{29}	8.09	4.42	2.70, 2.58	–

drastically changed its characteristics and made the peptide highly insoluble in aqueous solution. The 1p–36p peptide also aggregated and precipitated in solution, as evidenced by strongly increased line widths for its proton resonances (spectra not shown). The lack of proton resonance dispersion in the spectra of both 16p–36pR and 1p–36p made it impossible to assess their binding to cathepsin B by proton resonance perturbations. The increased aggregation with both 16p–36pR and 1p–36p, on the other hand, implicates that residues $F^{30}YNVDI^{35}$ may have a negative impact on the solution properties of all propeptide fragments containing this sequence.

An N-terminal fragment 1p–29p was then synthesized to further assess the impact of residues $F^{30}YNVDI^{35}$ on the binding properties of propeptide fragments. The removal of residues $F^{30}YNVDI^{35}S$ from the 1p–36p peptide resulted in a dramatic improvement of spectral resolution and peptide behaviour for the 1p–29p fragment. The resonance perturbations observed for 1p–29p were limited to small upfield shifts and some line broadening effects on the side chain NH protons of His-1p and His-7p (spectra not shown). Most unexpectedly, residues $TWQ^{25}A$ did not show any binding-induced resonance perturbations when part of the 1p–29p peptide with increasing concentrations of cathepsin B (Table 1). The loss of resonance perturbations for both the 1p–29p and 36p–51p peptides indicates that residues $F^{30}YNVDI^{35}$ must be responsible for the mode of binding observed for the complex of cathepsin B with the 16p–51pR peptide. The observation that both the 1p–29p and 36p–51p fragments produced well-resolved proton NMR spectra suggests that residues $F^{30}YNVDI^{35}$ may also contribute negatively to the solution behavior of the propeptide fragments containing this important sequence.

4. Discussion

Residues $TWQ^{25}A$, which interact with the exo-site region of cathepsin B in the structure of procathepsin B (Table 1), have a very different behaviour when they are part of either the N-terminal (1p–29p) or the C-terminal (16p–51pR) fragment of the 1p–51p propeptide. No binding was observed within the 1p–29p peptide, while this region had binding-induced line broadening in the 16p–51pR peptide. The $TWQ^{25}A$ residues are not essential for binding, however, as the 26p–51p peptide still had a binding pattern similar to 16p–51pR. Inhibition studies, on the other hand, revealed that there is a dramatic decrease in the binding affinity of 26p–56p or upon removal of residues $NTTWQ^{25}$ [16]. The crystal structure of rat procathepsin B shows that the Trp-24p residue forms an aromatic cluster and has a large contact area ($\sim$100 Å^2) with the enzyme surface [14]. Furthermore, over 150 fold of binding affinity was lost when Trp-24p was replaced by an alanine residue within the 21p–56p propeptide [16]. Taken together, it appears that Trp-24p or the $TWQ^{25}A$ sequence contributes to the binding affinity of the cathepsin B propeptide in a highly context-dependent manner. Interactions with cathepsin B involving Trp-24p of the propeptide may be important only in propeptide fragments containing residues $F^{30}YNVDI^{35}$ and $YLKK^{40}LCGTV^{45}LG$ occupying both the occluding loop and the active site of cathepsin B.

The patterns of resonance perturbations for the $YLKK^{40}LCGTV^{45}LG$ region, or part of it, in peptides 16p–51pR, 26p–51p, 28p–43p and 26p–43p (Table 1) indicate that these peptides preserve the same unique mode of binding interactions as in procathepsin B. The 36p–51p peptide, on the other hand, lacks binding-induced resonance perturbations, even though it contains the propeptide residues located in the substrate-binding grove (S and S′ subsites) of cathepsin B (Table 1). Thus, the Cys-42p residue or the $CGTV^{45}L$ sequence, which is required for inhibitory binding [16], and the $LKK^{40}L$ sequence, which showed the most pronounced resonance perturbations in 16p–51pR (Fig. 2), are not able to retain the same mode of binding for 36p–51p as in 16p–51pR. However, the 36p–50p peptide was reported to inhibit the catalytic activity of cathepsin B [17], suggesting that a different binding mode may be adopted by the 36p–50p fragment in comparison to peptides containing the $F^{30}YNVDI^{35}$ sequence.

All these results indicate the importance of residues $F^{30}YNVDI^{35}$ for maintaining specific binding interactions between the propeptide fragments and cathepsin B. The active-site binding region can be partly removed while specific binding is still retained if the $F^{30}YNVDI^{35}$ fragment is included in the peptide. The 28p–43p fragment, as a result, is currently the smallest propeptide fragment that still retains predominant binding contacts with cathepsin B. Since cathepsin B is the only member of the papain superfamily of cysteine proteases possessing an occluding loop, the presence of the $F^{30}YNVDI^{35}$ segment in peptide analogs may result in highly selective binding to this enzyme.

Acknowledgements: We thank Zhigang Chen for useful discussions, Yamin Chen for assistance in inhibition assays and Mirek Cygler for access to the X-ray coordinates of rat procathepsin B. Betty Zhu is acknowledged for help with NMR data processing and with molecular graphics. This work was supported in part by the Protein Engineering Network Centres of Excellence, sponsored by the Government of Canada, and by the National Research Council of Canada.

References

[1] Barrett, A.J. and Kirschke, H. (1981) Methods Enzymol. 80, 535.
[2] Mort, J.S., Recklies, A.D. and Poole, A.R. (1984) Arthritis Rheum. 27, 509–515.
[3] Katunuma, N. and Kominami, E. (1987) Rev. Physiol. Biochem. Pharmacol. 108, 1–20.
[4] Elliott, E. and Sloane, B.F. (1996) Perspect. Drug Discovery Design 6, 12–32.
[5] Sloane, B.F. (1990) Semin. Cancer Biol. 1, 137–152.
[6] Nakamura, Y., Takeda, M., Suzuki, H., Hattori, H., Tada, K., Hariguchi, S., Hashimoto, S. and Nishimura, T. (1991) Neurosci. Lett. 130, 195–198.
[7] Rasnick, D. (1996) Perspect. Drug Discovery Design 6, 47–63.
[8] Storer, A.C. and Menard, R. (1996) Perspect. Drug Discovery Design 6, 33–46.
[9] Chan, S.J., Segundo, B.S., McCormick, M.B. and Steiner, D.F. (1986) Proc. Natl. Acad. Sci. USA 83, 7721–7725.
[10] Hara, K., Kominami, E. and Katunuma, N. (1988) FEBS Lett. 231, 229–231.
[11] Kominami, E., Tsukahara, T., Hara, K. and Kutumuma, N. (1988) FEBS Lett. 231, 225–228.
[12] Fox, T., de Miguel, E., Mort, J.S. and Storer, A.C. (1992) Biochemistry 31, 12571–12576.
[13] Mach, L., Mort, J.S. and Glossl, J. (1994) J. Biol. Chem. 269, 13030–13035.
[14] Cygler, M., Sivaraman, J., Grochulski, P., Coulombe, R., Storer, A.C. and Mort, J.S. (1996) Structure 4, 405–416.
[15] Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M. and Turk, V. (1996) FEBS Lett. 384, 211–214.

[16] Chen, Y., Plouffe, C., Ménard, R. and Storer, A.C. (1996) FEBS Lett. 393, 24–26.
[17] Chagas, J.R., Ferrer-Di Martino, M., Gauthier, F. and Lalmanach, G. (1996) FEBS Lett. 392, 233–236.
[18] Sivaraman, J., Coulombe, R., Magny, M.C., Mason, P., Mort, J.S. and Cygler, M. (1996) Acta Crystallogr. D52, 874–875.
[19] Rowan, A.D., Mason, P., Mach, L. and Mort, J.S. (1992) J. Biol. Chem. 267, 15993–15999.
[20] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley-Interscience, New York.
[21] Ni, F. (1992) J. Magn. Reson. 99, 391–397.
[22] Otting, G., Wider, H., Wagner, G. and Wüthrich, K. (1986) J. Magn. Reson. 66, 187–193.
[23] Kumar, A., Brown, S.C., Donlan, M.E., Meier, B.U. and Jeffs, P.W. (1991) J. Magn. Reson. 95, 1–9.
[24] Kadhodaei, M., Hwang, T., Tang, J. and Shaka, A.J. (1993) J. Magn. Reson. 105, 104–107.
[25] Ni, F. (1994) Progr. NMR Spectrom. 26, 517–606.