ELSEVIER

Journal of Chromatography A, 777 (1997) 23–30

JOURNAL OF CHROMATOGRAPHY A

Probing protein structure using biochemical and biophysical methods
Proteolysis, matrix-assisted laser desorption/ionization mass spectrometry, high-performance liquid chromatography and size-exclusion chromatography of p21$^{Waf1/Cip1/Sdi1}$

Richard W. Kriwacki[a], Jiangyue Wu[b], Linda Tennant[a], Peter E. Wright[a,c], Gary Siuzdak[a,b,*]

[a]*Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA*
[b]*Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA*
[c]*The Skaggs Institute of Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA*

Abstract

The cyclin-dependent kinase (Cdk) inhibitor p21$^{Waf1/Cip1/Sdi1}$, important for p53 tumor suppressor-dependent cell growth control in humans and other organisms, mediates G_1/S-phase arrest through inhibition of cyclin-dependent kinases (Cdks). The enzymatic activity of these kinases is essential for progress through the cell division cycle and one level of cell cycle regulation is exerted through inhibition of Cdks by a family of small proteins, including p21. Cdk inhibition requires a sequence of approximately 60 amino acids within the p21 NH_2-terminus. Using proteolytic mapping, matrix-assisted laser desorption/ionization (MALDI) mass spectrometry, HPLC and size-exclusion chromatography, we show that p21, active as a Cdk inhibitor, exists in an extended, non-globular conformation in the absence of its biological target and that p21 lacks the hallmarks of stable secondary and tertiary structure. We have developed an efficient approach to obtain detailed proteolytic maps that takes advantage of the high accuracy and sensitivity of MALDI mass spectrometry. Our method allows a proteolytic map to be obtained from a single mass spectrum for fragments produced from a single proteolytic reaction.

Keywords: Proteins; Proteolytic mapping

1. Introduction

The cyclin-dependent kinase (Cdk) inhibitor p21$^{Waf1/Cip1/Sdi1}$ [1–3], an important component of cell growth regulation [4], is transcriptionally regulated in a p53 tumor suppressor-dependent [2] and -independent [5,6] manner and, in the former case, is directly involved in DNA damage-induced cell cycle arrest at the G_1/S checkpoint [4–7]. p21 mediates G_1/S arrest through association with and inhibition of cyclin-dependent kinases (Cdks) [1], and possibly through inhibition of DNA replication [8]. Cdk binding requires a sequence of approximately 60 amino acids within the p21 NH_2-terminus [9–11] (Fig. 1A), a region that is homologous with the NH_2-terminal domains of three other Cdk inhibitors, p27 [12,13], p28 [14,15] and p57 [16,17]. The conservation of this novel Cdk inhibitory domain

*Corresponding author.

0021-9673/97/$17.00
PII S0021-9673(97)00527-X

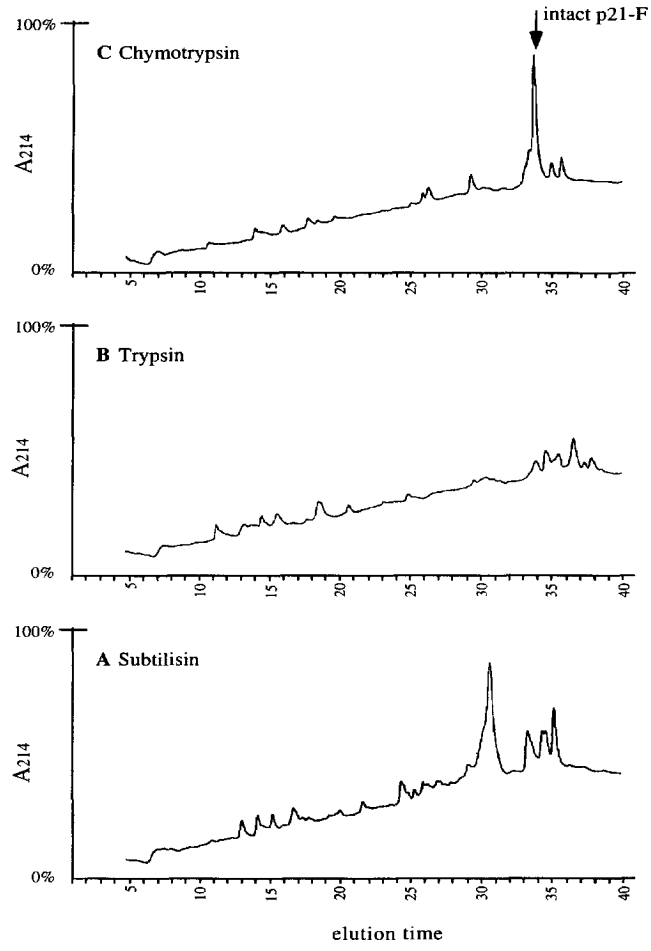

Fig. 1. HPLC chromatograms of p21-F digests using (A) subtilisin, (B) trypsin and (C) chymotrypsin.

within this protein family, coupled with its putative biological role in p53-mediated tumor suppression, identifies p21 as a potential structural prototype for novel therapeutic agents designed for the treatment of cancer through inhibition of cell proliferation.

Protease mapping is an established method for probing the primary structure of proteins [18,19] and has traditionally been performed through the use of chromatography and/or gel electrophoresis techniques in combination with Edman degradation NH_2-terminal sequencing [20]. Importantly, proteolytic cleavage can also provide indirect information on the domain structure of proteins since the existence of secondary and tertiary structure protects the polypeptide chain from proteolytic cleavage [21]. When using traditional bioanalytical methods, however, the process of separating and identifying protein fragments is time-consuming, expensive and sometimes impossible due to the low resolution of the principal analytical methods. Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry is well-suited [22] to the analysis of complex mixtures of biomolecules, offering high sensitivity, resolving power and accuracy, and thus the prospect of studying higher order protein structure when combined with protease mapping [23,24].

We have utilized MALDI mass spectrometry to analyze p21 proteolysis reactions and, moreover, have developed a streamlined approach to completely identify protease cleavage sites. First, proteolysis conditions are optimized using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–

PAGE) to monitor the extent of protein fragmentation. Secondly, under optimal proteolysis conditions samples are prepared for MALDI mass analysis by protein precipitation to reduce salt content, followed by resuspension in an organic solvent. A distinct and key advantage of this approach is that the masses of the fragments are obtained together in a single mass spectrum — without the need for individual purification using, for example, HPLC or SDS–PAGE. For a protein of known sequence and cleavage by a protease of known sequence-specificity, the mass usually identifies the exact fragment within the primary amino acid sequence, providing both N- and C-terminal cleavage sites. We report here the use of this approach to probe the secondary and tertiary structure of p21 under standard aqueous conditions and also in the presence of urea in order to monitor changes in structure due to "denaturation".

2. Experimental

Full-length p21 and two N-terminal fragments were over-expressed in *E. coli* and purified using established procedures [25]. These proteins all contain a C-terminal $(His)_6$ tag for purification purposes [it is known that the (His)6 tag does not affect either p21 structure or function] [25]. This series includes p21-F, full-length p21 [1,2] with a $(His)_6$ C-terminal tag; p21-A, p21 amino acids 1–94 plus the $(His)_6$ tag; and p21-B, amino acids 9–84 plus the $(His)_6$ tag. p21-F solutions for proteolysis were prepared by dissolving 5 mg lyophilized p21-F (after purification by HPLC) in 5 ml water. This solution was diluted into 9 volumes of buffer with a final composition of 50 m*M* Tris (pH 8.5), 500 m*M* NaCl, 1 m*M* dithiothreitol (DTT) and 1 m*M* EDTA (P buffer). This solution was concentrated to 5 ml by ultrafiltration (Amicon ultrafiltration unit, YM 10 membrane). The final protein concentration was approximately 1 mg/ml as monitored by UV absorbance using ϵ_{280} = 0.724*A* units/1 mg/ml[1]. p21-F, p21-A and p21-B solutions for size-exclusion chromatography were prepared by dissolving 1 mg lyophilized material (after purification by HPLC) in 100 μl water. This solution was diluted into 9 volumes of buffer with a final composition of 50 m*M* Tris (pH 7.5), 500 m*M* NaCl, 1 m*M* DTT and 1 m*M* EDTA (GF buffer).

Protease reactions in P buffer totaling 240 μl contained ~330 μg/ml p21-F; 1 m*M* $CaCl_2$; and 8.3, 33, or 133 ng/ml subtilisin (Boehringer Mannheim); or 16, 48 or 256 ng/ml trypsin (Sigma); or 16, 48 or 256 ng/ml chymotrypsin (Sigma). Reactions at room temperature were initiated by addition of protease and were terminated after 30 min by the addition of 2 m*M* PMSF and 10% (w/v) trichloroacetic acid (TCA), followed by incubation on ice for 30 min. The addition of TCA quantitatively precipitates polypeptides from the relatively high ionic strength solution. These samples, contained in 0.6 ml plastic tubes, were centrifuged at 13 000 rpm for 20 min followed by quantitative removal of supernatant in a procedure that effectively desalts the polypeptide fraction. Samples for SDS–PAGE were resuspended in SDS-containing buffer [26] and samples for MALDI mass analysis or HPLC analysis were resuspended in water–CH_3CN–trifluoroacetic acid (TFA) (50:50:0.1). Protease reactions were also carried out in the presence of 2.0 and 4.0 *M* urea, with other conditions unchanged.

Reversed-phase HPLC analysis was performed using a Vydac C_4 column (25 cm×4.6 mm) with a solvent gradient from water–CH_3CN–TFA (82:18:0.1) to water–CH_3CN–TFA (52:48:0.1) over 50 min at a flow rate of 2 ml/min. Detection was at 214 nm.

Size-exclusion chromatography was performed using a Superose-12 column (30 cm×1.0 cm, Pharmacia) operating at 0.5 ml/min using GF buffer. Loaded sample volumes ranged from 60–100 μl and detection was at 280 nm. The column performance was calibrated using a standard set of low molecular mass standards (Bio-Rad), including γ-globulin (M_r 158 000), ovalbumin (44 000), myoglobin (17 000) and vitamin B_{12} (1400). The elution volumes for these proteins were determined in triplicate and, after subtracting the void volume (4.49 ml), fit to Eq. (1):

$$M_r(\text{act})(\text{stds.}) = 1.53 \cdot 10^6 \cdot e^{[-1.33 \cdot (V_e - V_v)]} \quad (1)$$

where: V_e is elution volume and V_v is the void volume.

[1] R.W. Kriwacki, P.E. Wright, The ϵ_{280} value was determined by measuring the absorbance for a p21-F solution at 280 nm followed by determination of the exact protein concentration through amino acid analysis.

Using this equation, the correlation coefficient (R) for the fit of the $V_e - V_v$ data to the actual molecular masses is 0.999. The elution times for p21-F, p21-A and p21-B were similarly determined in triplicate. The apparent molecular weights [M_r(app)] for p21 proteins were calculated using the experimental $V_e - V_v$ values and the above equation. The actual molecular masses [M_r(act)] were determined from the molecular formulae.

MALDI mass measurements were performed with a PerSeptive Voyager Elite MALDI mass spectrometer. Digest samples (40 μM) were combined with the matrix 2,5-dihydroxybenzoic acid [~0.2 M in water–CH_3CN–TFA (50:50:0.5)] at a 1:1 volume ratio. Two micoliters of the sample-matrix mixture solution were deposited onto the MALDI sample plate and inserted into the ionization source. Samples were irradiated with a nitrogen laser (Laser Science Inc.) operated at 337 nm. Spectra shown were typically an average of spectra from 128 laser pulses.

3. Results

The results of typical digests of p21-F with subtilisin, trypsin and chymotrypsin analyzed using reversed-phase HPLC and MALDI mass spectrometry are shown in Fig. 1 and Fig. 2, respectively. A large number of p21-F fragments with a wide range of elution times are seen in Fig. 1A–C,

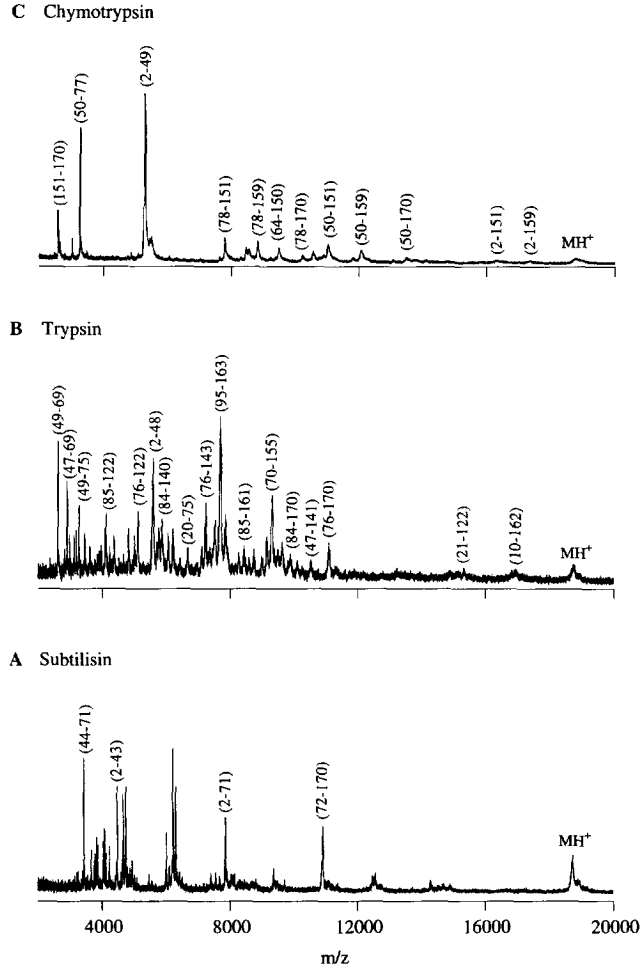

Fig. 2. MALDI mass spectral data of p21-F digests using (A) subtilisin, (B) trypsin and (C) chymotrypsin.

indicating that proteolytic cleavage is extensive. This finding is supported by the results of SDS–PAGE analysis (data not shown). The elution-time range for fragments produced by the three different proteases is similar, extending from approximately 10 to 38 min, while the exact peak profiles produced by the different proteases differ due to the different sequence preferences of these enzymes. Cleavage is more extensive with subtilisin and trypsin versus chymotrypsin, as judged by the intensity of the intact p21-F peak that elutes at ~34 min.

The corresponding MALDI mass spectra show peaks over a wide mass range, from ~2000 to 20 000, consistent with the extensive digestion seen by HPLC. Based on the sequence specificity of trypsin (cleavage after Lys or Arg residues) and chymotrypsin (cleavage after Phe, Tyr and Trp residues), we identified greater than 95% of the peaks in Fig. 2B and Fig. 2C. The process of assigning a particular peak at a specific mass to a p21-F fragment involved first generating a table of all possible fragments and their masses based on the p21-F sequence and protease specificity, and then comparing the observed masses with those in the table. Due to the high accuracy of MALDI mass analysis (better than 0.1% at M_r 10 000), virtually all peaks in the trypsin and chymotrypsin spectra could be uniquely assigned. Interestingly, cleavage by trypsin is observed at all but one predicted site (cleavage predicted after amino acids 9, 16, 19, 20, 32, 46, 48, 67, 69, 75, 83, 84, 86, 93, 122, 140, 141, 142, 143, 154, 155, 156, 161 and 162; cleavage observed at all sites except 32). The same is true for cleavage by chymotrypsin (cleavage predicted after amino acids 22, 49, 51, 53, 65, 77, 150, 151 and 159; cleavage observed at all sites except 65). The appearance of relatively fewer fragments with chymotrypsin (Fig. 1C and Fig. 2C) versus trypsin (Fig. 1B and Fig. 2B) reflects the relative abundance of cleavage sites, 8 for chymotrypsin and 23 for trypsin. The trypsin and chymotrypsin data are summarized in the form of a linear protease map in Fig. 3. Subtilisin fragments could not be assigned using MALDI analysis alone due to the lack of sequence specificity for this enzyme, however we identified a few sites using traditional methods, as illustrated in Fig. 2A. Despite this, the HPLC chromatogram and MALDI mass spectrum of subtilisin fragments (Fig. 1A and Fig. 2A) can be qualitatively interpreted and clearly show that many sites throughout p21-F are accessible to this protease.

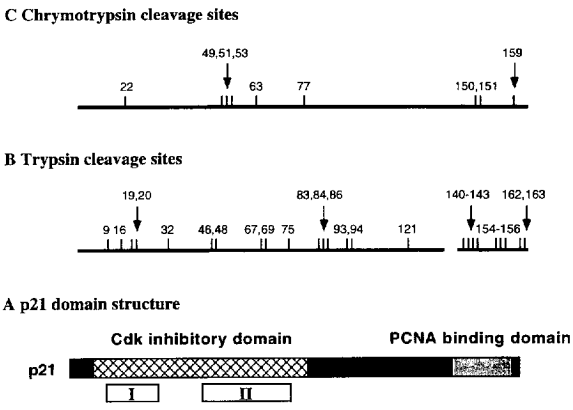

Fig. 3. p21 domain structure (A), and trypsin (B) and chymotrypsin cleavage sites (C). (A) The boxed regions I and II represent p21 residues 17–33 and 49–77, respectively, which show 76 and 48% identity, respectively, with p27 [12,13].

The effect of urea, a strong protein denaturant, on p21-F cleavage by subtilisin was determined by monitoring the products of reactions carried out at two different concentrations of subtilisin in the presence of 0, 2.0 and 4.0 *M* urea. Subtilisin was chosen as the probe in these experiments because it was less susceptible to inactivation by 2.0 or 4.0 *M* urea than either trypsin or chymotrypsin. However, qualitative analysis of p21-F digested with trypsin or chymotrypsin in the presence of urea are consistent with the results using subtilisin (data not shown). Subtilisin fragments were analyzed using SDS–PAGE (Fig. 4) and HPLC (Fig. 5). Fig. 4 shows that the same principal bands are produced with or without urea (compare lanes 4, 5 and 6, and lanes 7, 8 and 9), strongly suggesting that cleavage site accessibility is not changed by urea. Most likely due to gradual protease denaturation, subtilisin specific activity is reduced in the presence of 2.0 *M* urea and is further reduced with 4.0 *M* urea, as revealed by the relatively darker band for p21-F in lanes 5 and 6 versus 4, and 8 and 9 versus 7 in Fig. 4. However, the similarity of the fragment band positions with or without urea is readily apparent. The HPLC chromatograms obtained for the same samples (Fig. 5) show similar behavior. The peak profiles in Fig. 5A–C are very similar and differ only slightly in

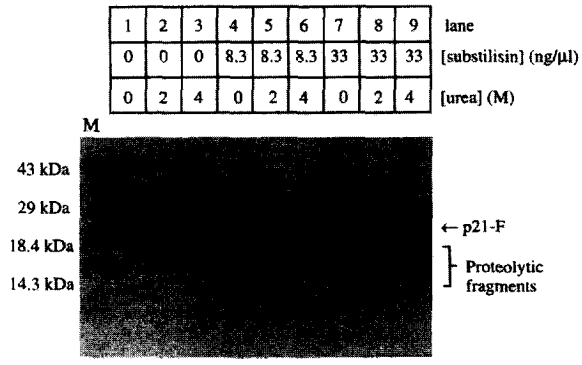

Fig. 4. SDS–PAGE of the p21 subtilisin digest. Subtilisin p21 fragments produced in the presence of the denaturant urea (lanes 4, 5 and 6) and in the absence of urea (lanes 7, 8 and 9). kDa = kilodaltons.

some peak intensities but not in peak positions. The relatively larger parent peak for p21-F in Fig. 5B and Fig. 5C, again, reflects the reduced specific activity of the enzyme in the presence of urea. It can be clearly seen, however, that the same major and minor fragments are produced with or without urea.

The results of size-exclusion chromatography offer another indication of the non-compact nature of p21. After calibrating the performance of the size-exclusion column using molecular mass standards ranging from 158 000–1400, the elution time and apparent molecular mass for p21-F, p21-A and p21-B were determined (Table 1). The apparent molecular mass for p21-F (24 100) is greater than the actual mass (18 900) and indicates that p21-F's hydrodynamic behavior is not well-modelled by the globular standards. This finding is consistent with findings from proteolysis, both without and with urea, and strengthens the view that p21 is disordered and non-globular. The size-exclusion result for p21-A (M_r(app), 18 000; M_r(act), 11 300) is consistent with that for p21-F while that for p21-B (M_r(app), 9400; M_r(act), 9600) suggests a more compact structure for this smaller N-terminal fragment of p21.

4. Discussion

Protease mapping is often used to identify stable domains within proteins and is often used to optimize protein constructs for detailed structural analysis using X-ray crystallography and/or NMR spectroscopy. With p21, however, stable domains are not observed. Rather, we observe extensive fragmentation using trypsin, chymotrypsin and subtilisin. With the former two enzymes, cleavage sites are evenly distributed throughout the primary sequence indicating that within the limits of these sites, stable elements of secondary and/or tertiary structure do not exist so as to protect some sites from cleavage. These results are inconsistent with the existence of stable, structured domains within p21-F but provide evidence for extended, non-globular polypeptide conformations. A realistic model would represent p21-F as a dynamic ensemble of extended conformations that interconvert rapidly on the timescale of the protease reactions (minutes). This picture is consistent with results for p21-F from circular dichroism (CD) spectropolorimetry and for N-terminal p21 fragments from CD spectropolorimetry and NMR spectroscopy [25]. The hydrodynamic properties of p21-F and p21-A, based on size-exclusion chromatography, strongly suggest that these proteins are non-globular, and that they may be best modelled by a prolate ellipsoid rather than a sphere [27]. This conclusion was based on the departure from unity for the ratio of the apparent and actual mass for p21-F (M_r(app):M_r(act) = 1.28) (Table 1). The physical picture reveals that, due to their non-spherical shape, the p21 molecules migrate more quickly than expected through the gel. Assuming this ellipsoidal model [27], the M_r(app):M_r(act) values can be used to estimate values of ellipsoidal Perrin shape factors (Table 1), which in turn are related to the axial ratio (length:width) for a prolate ellipsoid. These estimates range between 5:1 (p21-F) and 10:1 (p21- A), and suggest that p21-F and p21-A are highly elongated, this idea being consistent with the generally extended character of disordered polypeptide chains. The fact that the hydrodynamic behavior of p21-B (the smallest p21 fragment studied here) falls in line with that of the globular standards may indicate that it possesses a more compact, but still disordered structure than p21-F or p21-A. The gel filtration results for p21-F and p21-A, and the proteolysis results for p21-F are especially striking in the context of the biochemical activity of p21-F as a potent (<1 nM) cyclin-dependent kinase inhibitor.

The SDS–PAGE and HPLC results for subtilisin

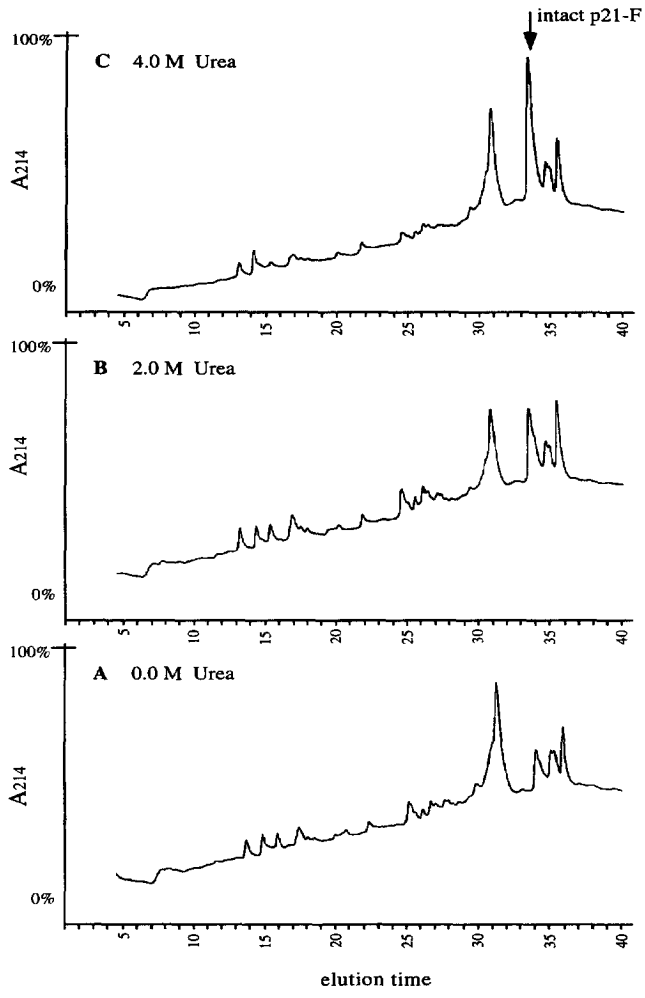

Fig. 5. HPLC of the p21 subtilisin digest with and without urea.

cleavage of p21-F in the presence of urea show that protease accessibility, and therefore protein structure, does not change due to the denaturant. This result is consistent with former data from CD spectropolorimetry and NMR spectroscopy results and further strengthens the view that p21-F is highly disordered, even in the absence of urea.

Previous proteolysis data for p21-B (an 81 amino

Table 1
Results from Superose 12 size-exclusion chromatography for standard globular proteins and p21-F, p21-A and p21-B

Protein	M_r(act) ($\cdot 10^3$)	Avg. $V_e - V_v$ (ml)	M_r(app) ($\cdot 10^3$)	M_r(app):M_r(act)
γ-Glogulin	158	1.76±0.02		
Ovalbumin	44	2.67±0.02		
Myoblogin	17	3.31±0.02		
p21-F	18.9	3.12±0.02	24.1±0.7	1.28
p21-A	11.4	3.34±0.02	18.0±0.5	1.58
p21-B	9.6	3.83±0.02	9.4±0.3	0.98

acid N-terminal fragment of p21 showing a high degree of protease accessibility) in the free solution state [24] is entirely consistent with the proteolysis data for p21-F reported here. In contrast, trypsin proteolysis studies of p21-B in association with Cdk2 reveal protection of Arg residues at positions 46, 48, 67 and 69, indicating that these amino acids lie at the protein–protein interface in the complex, and/or are indirectly involved in intermolecular contacts through the formation of secondary structure. Based on X-ray crystallography data for p27 [28], a related kinase inhibitor [12,13] bound to cyclin A/Cdk2, the protected amino acids of p21 in complex with Cdk2 are probably involved in β-strand secondary structures. That these sites in p21-F are accessible to trypsin indicates that, in the absence of Cdk2, these amino acids are not involved in β-strand or other secondary structures.

The data we report herein are based on detailed analyses of complex mixtures of polypeptides that are produced by proteolysis. Despite the complexity of these mixtures, we have identified 95% or more of the polypeptides that compose these mixtures. This has been done through the application of MALDI mass analysis coupled with new sample handling procedures. For proteins of known and potentially unknown [29] sequence, these techniques offer access to detailed protease maps with only modest investments of time and material. The accuracy and speed of our approach surpasses traditional methods based on HPLC/SDS–PAGE analysis and Edman degradation. Furthermore, due to the high resolution and sensitivity of MALDI mass analysis, a much greater number of protein fragments can be identified than was previously possible, offering more detailed "maps" of protein structure. The application reported here involves the study of a kinase inhibitor, p21, that exhibits unusual structural properties. However, the proteolysis/MALDI methods are equally well-suited to the study of highly structured proteins.

Acknowledgments

R.W.K. would like to thank Dr. Tom Steitz of Yale University for suggesting the proteolysis work on p21, Steve Reed and Ludger Hengst for stimulating discussions, Dr. Lisa Bibbs for protein analysis services and Ms. Martine Reymond for technical assistance with HPLC. R.W.K. is a Leukemia Society Fellow. This work was supported by grant GM38794 (P.E.W.) from the National Institutes of Health.

References

[1] J.W. Harper, G.R. Adami, N. Wei, K. Keyomarsi, S.J. Elledge, Cell 75 (1993) 805.
[2] W.S. El-Deiry et al., Cell 75 (1993) 817.
[3] A. Noda, Y. Ning, S.F. Venable, O.M. Pereira-Smith, J.R. Smith, Exp. Cell Res. 211 (1994) 90.
[4] C. Deng, P. Zhang, J.W. Harper, S.J. Elledge, P. Leder, Cell 82 (1995) 675.
[5] S.B. Parker et al., Science 267 (1995) 1024.
[6] K.F. Macleod et al., Genes Dev. 9 (1995) 935.
[7] V. Dulic et al., Cell 76 (1994) 1013.
[8] S. Waga, G.J. Hannon, D. Beach, B. Stillman, Nature 369 (1994) 574.
[9] M. Nakanishi, R.S. Robetorye, G.R. Adami, O.M. Pereira-Smith, J.R. Smith, EMBO J. 14 (1995) 555.
[10] J. Chen, P.K. Jackson, M.W. Kirschner, A. Dutta, Nature 374 (1995) 386.
[11] Y. Luo, J. Hurwitz, J. Massagué, Nature 375 (1995) 159.
[12] K. Polyak, M.H. Lee, H. Erdjument-Bromage, A. Koff, J.M. Roberts, P. Tempst, J. Massagué, Cell 78 (1994) 59.
[13] H. Toyoshima, T. Hunter, Cell 78 (1994) 67.
[14] J.-Y. Su, R.E. Rempel, E. Erikson, J.L. Maller, Proc. Natl. Acad. Sci. USA 92 (1995) 10187.
[15] W. Shou, W.G. Dunphy, Mol. Biol. Cell 7 (1995) 457.
[16] M.-H. Lee, I. Reynisdottir, J. Massagué, Genes Dev. 9 (1995) 639.
[17] S. Matsuoka et al., Genes Dev. 9 (1995) 650.
[18] A. Fontana, G. Fassina, C. Vita, D. Dalzoppo, M. Zamai, M. Zambonin, Biochemistry 25 (1986) 1847.
[19] S.J. Hubbard, F. Eisenmenger, J.M. Thornton, Protein Sci. 3 (1994) 757.
[20] P. Edman, G. Begg, Eur. J. Biochem. 1 (1967) 80.
[21] T.E. Creighton (Ed.), Protein Structure, A Practical Approach, IRL Press, New York, 1989.
[22] G. Siuzdak, Proc. Natl. Acad. Sci. USA 91 (1994) 11290.
[23] S.L. Cohen, A.R. Ferre-D'Amare, S.K. Burley, B.T. Chait, Protein Sci. 4 (1995) 1088.
[24] R.W. Kriwacki, J. Wu, G. Siuzdak, P.E. Wright, J. Amer. Chem. Soc. 118 (1996) 5320.
[25] R.W. Kriwacki, L. Hengst, L. Tennant, S.I. Reed, P.E. Wright, Proc. Natl. Acad. Sci. USA 93 (1996) 00.
[26] T. Maniatis, E.F. Frisch, J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.
[27] C.R. Cantor, P.R. Schimmel, Biophysical Chemistry, W.H. Freeman and Co., New York, NY, 1980.
[28] A.A. Russo, P.D. Jeffrey, A.K. Patten, J. Massague, N.P. Pavletich, Nature (London) 382 (1996) 325.
[29] B.T. Chait, R. Wang, R.C. Beavis, S.B. Kent, Science 262 (1993) 89.