The boundaries of a DNA-binding domain as defined by proteolytic digestion

Marina Voulova

under the direction of
Dr. Ky Lowenhaupt and Thomas Schwartz
Massachusetts Institute of Technology

Research Science Institute
August 6, 1997

Abstract

Zα, a Z-DNA-binding domain, is assumed to have a dense, folded tertiary structure which makes it resistant to proteolytic cleavage. In order to determine its structual boundaries, proteolytic digestions were performed on a 131 amino acid peptide that included the studied Zα domain. From analysis of both single and double protease digests, it was possible to define a 77 amino acid core region, which is sufficient to bind Z-DNA. Further studies will be necessary to understand the role of both N-terminal and C-terminal sequences that may play a role in increasing the stability of the Zα domain.

1 Introduction

Human double-stranded RNA adenosine deaminase (hADAR) is a protein involved in post-transcriptional RNA editing [1]. Since its discovery in 1987, scientists have been trying to define its structure in order to better understand its function. Recent data have shown that hADAR contains a catalytic domain, a dsRNA-binding domain, and a Z-DNA-binding domain [2]. Although it is known that the dsRNA-binding domain and the catalytic domain cofunction to perform RNA editing, how the Z-DNA-binding domain interacts with them remains unclear.

Previous studies have shown that this Z-DNA-binding domain consists of two structurally similar subdomains: Zα and Zβ, both of which are able to bind Z-DNA. The Zα domain is interesting because it has shown a high binding affinity for Z-DNA. Analysis of its structure and function has been limited by the vaguely defined boundaries of the domain. The current study attempts to determine the exact boundaries of Zα.

The smallest segment of hADAR that binds Z-DNA had been previously identified by deleting amino acids from the N and the C termini of the protein and checking the binding affinity of the remaining residues [2]. This 78 amino acid fragment(Z78) was assumed to be the core of Zα. In the current study Z78 is extended by 20 amino acids in each direction in order to construct a segment (Z131) which spans beyond the binding domain. To facilitate the cleavage analysis and the protein purification, a histidine tag was added to the protein's C-terminus(Z131HC) and N-terminus(Z131HN).

The method used in this study for determining the location of Zα within the Z131 constructs was based on the DNA-binding property of the domain. In order to be able to

interact with DNA, the Zα domain must be properly folded in a compact and dense tertiary structure. Such a folded structure would make Zα resistant to proteolytic digestion because it would prevent enzymes from reaching the residues situated in the inner part of its coils. Therefore, although the proteases used in these experiments had potential cleavage sites all over the Z131 protein, those positioned on the Zα domain were expected to remain intact. In other words, when analyzing the results of the digestions, it was assumed that most of the actual digestion sites were outside the Zα domain while the relatively stable regions were probably within it. PACE gels and Southwestern blotting proved this presupposition to be true.

2 Materials and Methods

2.1 Preparation of proteins

Two expression plasmids containing the 131 amino acid fragment of hADAR had previously been constructed [3]. Z131HN has a His-tag expressed at the N-terminus while Z131HC includes a C-terminal His-tag. The plasmids were then cloned into *E.coli*. The bacteria were grown until their OD_{600} in the medium reached 0.07 and the cells were then induced by addition of 1mM IPTG. After 4 hours they were harvested by centrifugation. The supernatant was discarded and the pellets were resuspended in lysis buffer (20mM Tris pH 8.0, 300mM NaCl, 5mMβ-mercaptoethanol, 1mM PMSF). The cells were lysed using French Press and the obtained lysate was centrifuged to remove the unbroken cells and cellular frag-

ments. Small amounts of the resuspended pellets were analyzed by SDS-PAGE to ensure that Z131HN and Z131HC were soluble in the buffer used.

2.1.1 Affinity Chromatography

The proteins of interest were purified by binding the His-tags to Ni^{++}. Each supernatent was incubated with 4mL of 50% Ni-NTA resin and was shaken at 4°C for one hour. Three washing steps with 20mL of lysis buffer were performed to remove the unbound protein. The suspensions were then centrifuged for 3 min. The pelleted resins were resuspended with 15mL of lysis buffer and poured onto columns. The columns were washed with 40mL of wash buffer#1(20mM Tris pH 8.0, 1M NaCl, 5mM β-ME) followed by 20mL of wash buffer#2(20mM Tris pH 8.0, 2M NaCl, 5mM β-ME). The bound proteins were eluted from the Ni-resin by 3 washing steps(five times 4mL of lysis buffer containing 30mM Imidazole; five times 4mL of lysis buffer containing 50mM Imidazole and five times with 4mL of lysis buffer containing 250mM Imidazole). The collected fractions were then analyzed on a SDS-PAGE gel.

2.1.2 Cation Exchange Chromatography

The fractions which contained Z131HN were combined, concentrated and dialysed into a low salt buffer. The same was done with the Z131HC fractions. These partially purified fractions were further purified by FPLC chromatography on a Mono S column. The fractions showing the highest A_{280} peaks were analyzed on an SDS-PAGE gel and those containing the pure Z131HN and Z131HC were collected.

2.2 Proteolytic digestions

Trypsin, chymotripsin, thermolysin and *Staphylococcus aureus (V8) protease* were used for the digestions of the two z131 constructs [4,5]. Prior to digestion, the proteins were dialysed into NH_4HCO_3 buffer(50mM NH_4HCO_3, 150mM NaCl, 1mM DTT) and were concentrated to 1mg/mL .

2.3 Gel Electrophoresis

Polyacrylamide-Gel Electrophoresis(PAGE) was used to determine the sizes of the proteins found in both the purified fractions and the digestion reactions. Depending on the samples, gels with different separating gradient were used. For purified fractions, the gel contained a separating solution of 6.5mL Protogel, 2.5mL separating gel buffer, 800μL H_2O, 100μL 10% SDS, 100μL 10% APS, and 100μL TEMED and a stacking solution of 650μL Protogel, 1.25mL stacking gel buffer, 3mL H_2O, 50μL 10% SDS, 50μL 10% APS,and 5μL TEMED. These gels were run in Laemmli buffer at 200V for 45min [6]. For the products of digestion reactions, the gel contained a separating solution of 5.5mL Protogel, 3.3mL 3x Gel buffer, 1.1mL H_2O, 100μL 10% APS, and 10μL TEMED and a stacking gel solution of 1.3mL Protogel, 1.65mL 3x Gel buffer, 2mL H_2O, 50μL 10% APS,and 5μL TEMED. These gels were run in Tris-tricine buffer at 90V [7]. Gels were stained in Coomassie blue.

Polyacrylamide-Gel Coelectrophoresis(PACE) was performed with some of the digested proteins to determine whether they could still bind Z-DNA [8]. The gel solution contained 6mL Protogel, 1mL 10x TB and 3mL H_2O. A buffer (10mM Tris pH 8.0, 100mM

NaCl) was used to dilute the stock protein solution. Ten different horizontal lanes were loaded. Each contained 2μL 10% APS, 1μL TEMED, 225μL gel solution and 225μL protein solution of known concentration. ^{32}P labeled Z-DNA was then run through the gel and, depending on the degree of binding, was retarded in a different way. The results of the gels were seen by covering it with a film and developing a picture.

2.4 Southwestern blotting

Southwestern blots were used to determine whether the fragments remaining after the proteolytic digestion of z131 still bound to Z-DNA. The cleaved fragments were first run on a PAGE gel. They were then transferred overnight from the gel to a nitrocellulose membrane in a buffer without SDS to allow them to return to their original folded shape. The membrane with the blotted protein was blocked with BSA to prevent DNA from binding to its surface. The membrane was then rinsed with buffer to remove any unbound BSA. A mixture of 100μL of H_2O, 100μL of 40mM MgCl, 20μL of radoactively labelled DNA, and 5μL of salmon sperm DNA competitor was preincubated to allow the labelled DNA to assume the Z-confirmation and added to the probe in 5mL of buffer. After one hour it was washed with TBST until the radiation level of the membrane part containing no protein was bellow background. The membrane was left to dry and then was covered with a film and an intensifying screen. The film was developed after 90 min.

2.5 Mass Spectroscopy

Mass spectroscopy was used to determine the molecular weights of the V8/Chymotrypsin and V8/Thermolysin double digests as well as the V8 and trypsin single digests of z131hn. In two different tubes, 50μL of protein was mixed with V8 or trypsin respectively. The reactions were stopped after two hours. Chymotrypsin and thermolysin were added separately to samples of the V8-digest and were allowed again to react for two hours. The prepared solutions were then serially diluted with cyanamic acid which induced crystallization. One μL of each dilution was put on the sample plate and examined by MALDI spectroscopy [9]. Cytochrome C (MW 12360.5) was used for calibrating the apparatus.

3 Results and Data

SDS-PAGE gels were used to analyze the digestion products for each protease and for several double digestions. The number, size, and stability of the bands produced in each digestion are shown in Table 1:

(Note: TY=trypsin; CH=chymotrypsin; TL=thermolysin)

Proteases used	proteins digested	Bands observed on SDS-PAGE gel
single digestions with TY, V8, CH, TL	Z78	Digested by every protease except TY The most stable cleavage product was produced by V8
TY	Z131HN Z131HC	In both cases one, very stable band was observed
V8	Z131HN Z131HC	In both cases two major, very stable bands that add up to the full size of the digested protein were observed. The distance between the bands of Z131HC was larger than the distance between the bands of Z131hn All bands were smaller that the TY product
TY/V8	Z131HN Z131HC	The TY digests were further cleaved by V8 The typical upper band of the V8 digests was present but the lower one was missing
TL	Z131HN Z131HC	Z131HN was completely digested while Z131HC was intact
V8/TL	Z131HN Z131HC	TL digested the upper band of the V8 cleavage of Z131HN The upper band of the V8 digest of Z131HC was left intact In both constructs the lower V8 band was digested A longer time was required for TL to digest the predigested Z131HN than to digest the full sized protein
CH	Z131HN Z131HC	In both cases three major, not very stable bands were created
V8/CH	Z131HN Z131HC	The lower V8 band of both constructs was further digested A band of intermediete lenght between the two V8 bands was produced It took CH a longer time to completely digest the predigested proteins than to digest the whole sized Z131s
V8/CH (+/- Ni)	Z131HC	The presence or absence of Ni made no noticeable difference

Table 1: Digestion products observed on SDS-PAGE gels

	trypsin	V8	V8/CH	V8/TL
mass of	12224	10287	7585	5223
digestion	6110	6376	8064	8582
products	12068	5142	7949	8381
(m/z)	11640	6226	7684	5022
	12426	8326	3791	4289
	6214	16631	4030	4188
	16630		2941	10283
			3973	10081
			4601	7799
			1508	7592
				7199
				6895
				3375

Table 2: Mass spectroscopy results

Mass Spectroscopy was used to determine the molecular weights of the digestion products. Table 2 gives the masses of the digestion products, listed in the order of decreasing abundance for each protease.

The digestion products of Z131HN cleaved with V8 or V8/CH were also subjected to N-terminal sequencing to confirm the location of the cleavage sites. V8 was shown to cut before the LSIYQ amino acid sequence. The results of the sequencing of the V8/CH digests, were uninterpretable.

To ensure that the protease cleavage sites on Z131 were outside the Zα domain, the binding affinity of the digested products for Z-DNA was determined. A PACE gel was used to study the single V8 as well the V8/CH double-digestions. In both digestions there were products which bound Z-DNA.

A second method, Southwestern blotting, was applied to most of the digestions as well as to the undigested Z131hn, Z131HC and Z78 proteins. All of the cleavage products larger

than or equal to the length of Z78 could still bind to Z-DNA. The cleaved N- and C-termini seen on the SDS-PAGE gels did not bind Z-DNA. When digested, Z78 also lost its binding affinity.

4 Discussion

A list of all the possible cleavage sites for each enzyme was created(figure 2). Then the SDS-PAGE results of each digestion were analyzed to find which were the most probable sites. The interpretation of the gels was of crucial importance in order to correctly define the various digestions products with respect to their primary amino acid sequence. Only after the fragments were located, could the specific cleavage sites be identified from the mass spectroscopy results.

Because V8 digestion produced two stable, equally intense bands that added up to the full size protein, it was concluded that V8 has one major cleavage site. The difference between the distance separating the two bands of Z131HN and Z131HC was explained by the hypothesis that this cleavage site is N-terminal of the center(i.e. the inconsistency is caused by the additional His-tags). The fact that the smaller fragment of Z131HN showed great binding affinity to the Ni resin confirms the assumption that the cleavage site is closer to the N-terminus. The digestion of Z78 suggested that the cleavage is actually within the presumed core, which further decreased the region of the possible locations of the V8 site.

The same methods were used to anlyse the trypsin digestions. The fact that it created just one stable band, suggests that the protease has only one major cleavage site near one of the termini, cutting Z131 into two parts-a long and a short one (not visible on the gel).

Because the trypsin/V8 double digest has a higher V8 band but has lost its lower one, we conclude that the trypsin digestion site is located on the smaller V8 cleavage product. Since the V8 digestion is N-terminal, the trypsin site is therefore also N-terminal. Based on the facts that the trypsin product was larger than the upper band of V8 and that trypsin did not digest Z78, it was also concluded that the trypsin site is outside the Z78 core and is closer to the N-terminus than the V8 site.

Very interesting results were obtained from the thermolysin digestions. The fact that it completely digested Z131HN, suggests that the protease has multiple cleavage sites. Because the enzyme digested the lower band of V8, we conclude that one minor site is N-terminal. Most interesting is the fact that Z131HC is resistant to thermolysin. Since the V8 in double V8/TL digestions created two peptides with the same N-terminus, the difference in the digestion resistance could have been caused only by C-terminal protection. The fact that Z131HC's C terminus was protected against thermolysin but Z78's C-terminus was not suggests that the Zα domain extends further in the C direction. It is possibile that its C-terminal boundary may not even be within the Z131 construct. This seems unlikely, however, since other C-terminal cleavage sites were mapped within Z131. It is more probable that the C-terminal protection is caused by some property of the C-His-tag. However, only future studies using a polypeptide that extends further than Z131 C-terminus can completely resolve this issue.

The digestions with chymotripsin produced three major cleavage products. The addition of Ni-resin to the Z131HC digestion did not make any noticeable difference on the SDS-PAGE gel which showed that the C-terminal fragment, containing the His-tag, is very small,

i.e., the cleavage is C-terminal.

Pre-digestion with V8 made the polypeptides more stable to cleavage by other proteases. This very interesting observation is explained by the hypothesis that the V8 cleavage site is situated just outside the Zα core domain. If that is true, pre-digestions with V8 will remove the extra N-terminal part and make the remaining peptide chain more stable. Since the V8 site is within Z78, it must also be concluded that the whole Z78 is not necessary for Z-DNA binding.

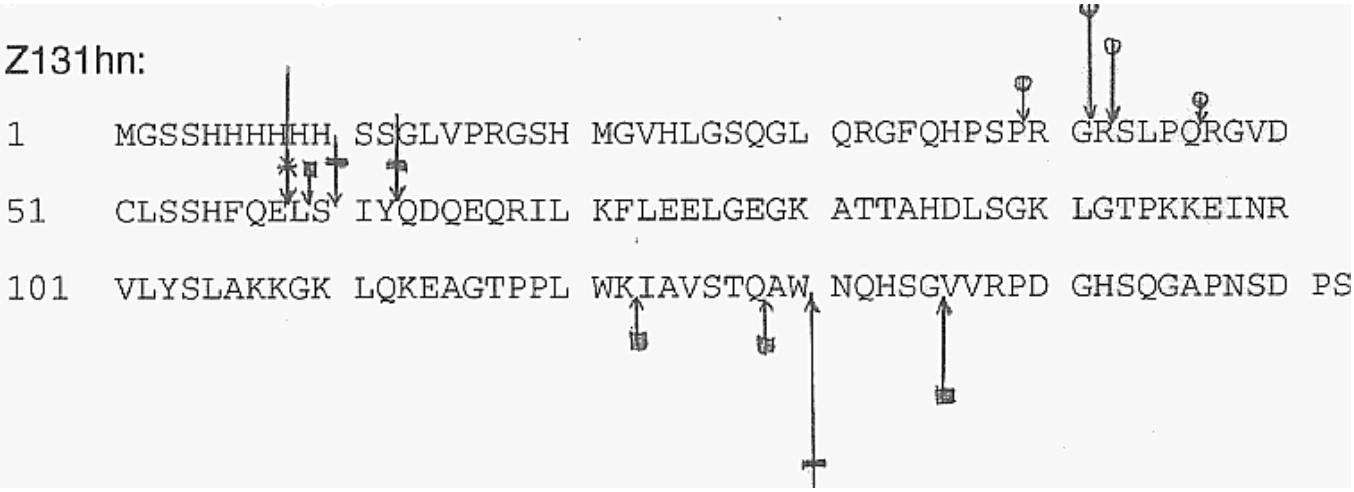

Figure 1: The figure shows the cleavage sites on Z131HN. The height of arrows indicates the extent of cutting at that site. The highest arrows show the most prefered cleavage sites for each protease(see legend of fig.2).

By analyzing the SDS-PAGE results of the digestions, we were able to identify the most probable regions in which the cleavage sites of each protease are located. All cleavage sites in the regions were considered to be possible. Using the information from the gel analysis revealing whether each product was C- or N-terminal, the expected weights of the most probable cleavage products were calculated(see Appendix). Then, by comparing these values with the actual values determined by the mass spectroscopy, we were able to pinpoint the

actual cleavage sites for each protease as well as the relative contribution of each site. A map of the cleavage sites is shown in Figure 1.

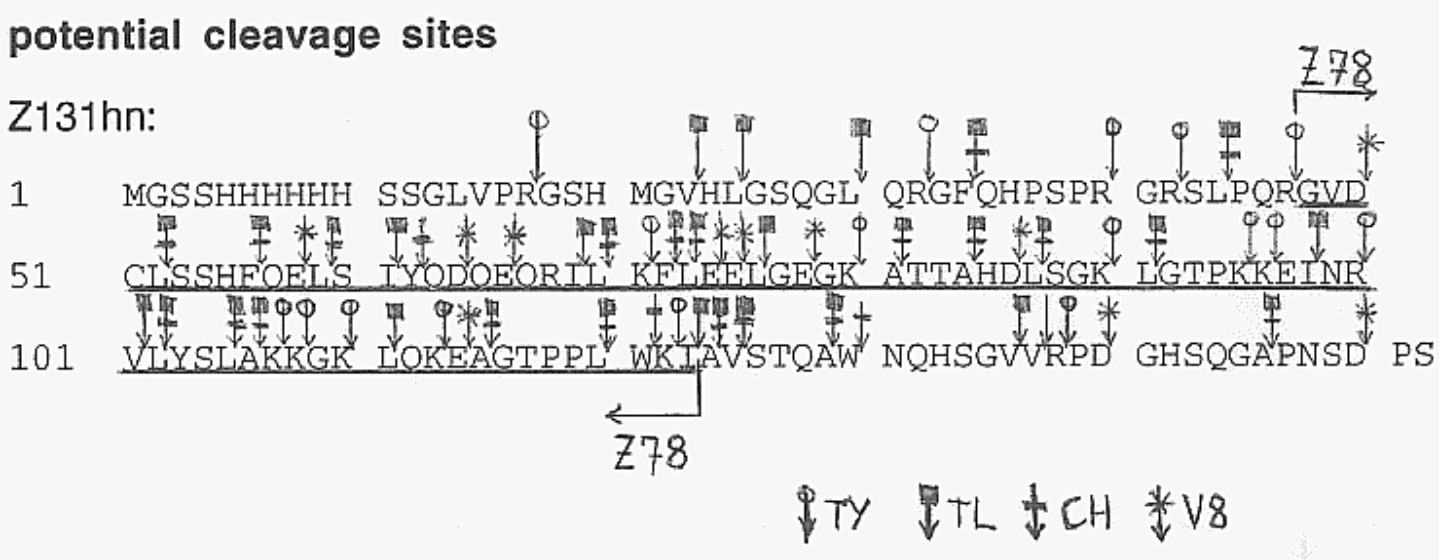

Figure 2: The figure shows all potential cleavage sites on Z131HN. The potential sites on Z131HC are identical. The sequence of Z78 is underlined.

If the actual cleavage sites are compared with all possible sites (Figure 2), a stable core region is clearly identified. It is located between the major V8 and major chymotrypsin sites which mark its N-terminal(LSIYQ) and C-terminal(STQAW) boundaries. Since the described region does not comprise any major or moderate actual cleavage sites, we conclude that it has an extremely dense tertiary structure. Although this core does not include the whole Z78 domain, PACE as well as Southwestern blotting showed that it still binds Z-DNA. From the facts that it is very stable to proteolytic cleavage and at the same time has great binding affinity for Z-DNA, we conclude that this region is the core of the Zα domain.

On both of the terminals of this newly defined core, there are regions that may affect the stability of the whole Zα domain. In the N-direction, the core is adjacent to a relatively stable region with a N-terminal boundary at the major trypsin site. Since there is only one

major site that separates these two regions, the case that they may belong to the same folded region has to be considered. This is possible because although the Zα domain has a dense tertiary structure, digestion occurs if a cleavage site is situated on its surface. Further studies with the RSLPQ..STQAW sequence should be performed to determine the exact function of the N-terminal elongation of the core.

A stable region is also adjacent to the C-terminus of the discovered core. Although there are originally fewer possible cleavage sites on the C-terminus, its proteolytic resistance is of particular interest because of the thermolysin protection of Z131HC. Although we assume that the thermolysin cleavage is restricted by the His-tag, studies of a construct extending further in C-direction should be performed in order for the significance of the adjacent to the core region to be defined.

5 Conclusion

By digesting the extended polypeptide which contained the previously defined Zα sequence, we were able to identify a stable region within the studied peptide. We suggest that this region corresponds to the full sized Zα domain.

6 Acknowledgments

I would like to thank my mentor Dr. Ky Lowenhaupt and Thomas Schwartz for the ir patience and support in quiding me through this project. I wish to thank also Megan Moore and Steve Fein for revising the previous drafts of this pa per. I am also grateful to the "Cyril

and Metodius foundation" and the Center for Excellence in Education for providing me the opportunity to paricipate in this science program.

A SDS-PAGE gels

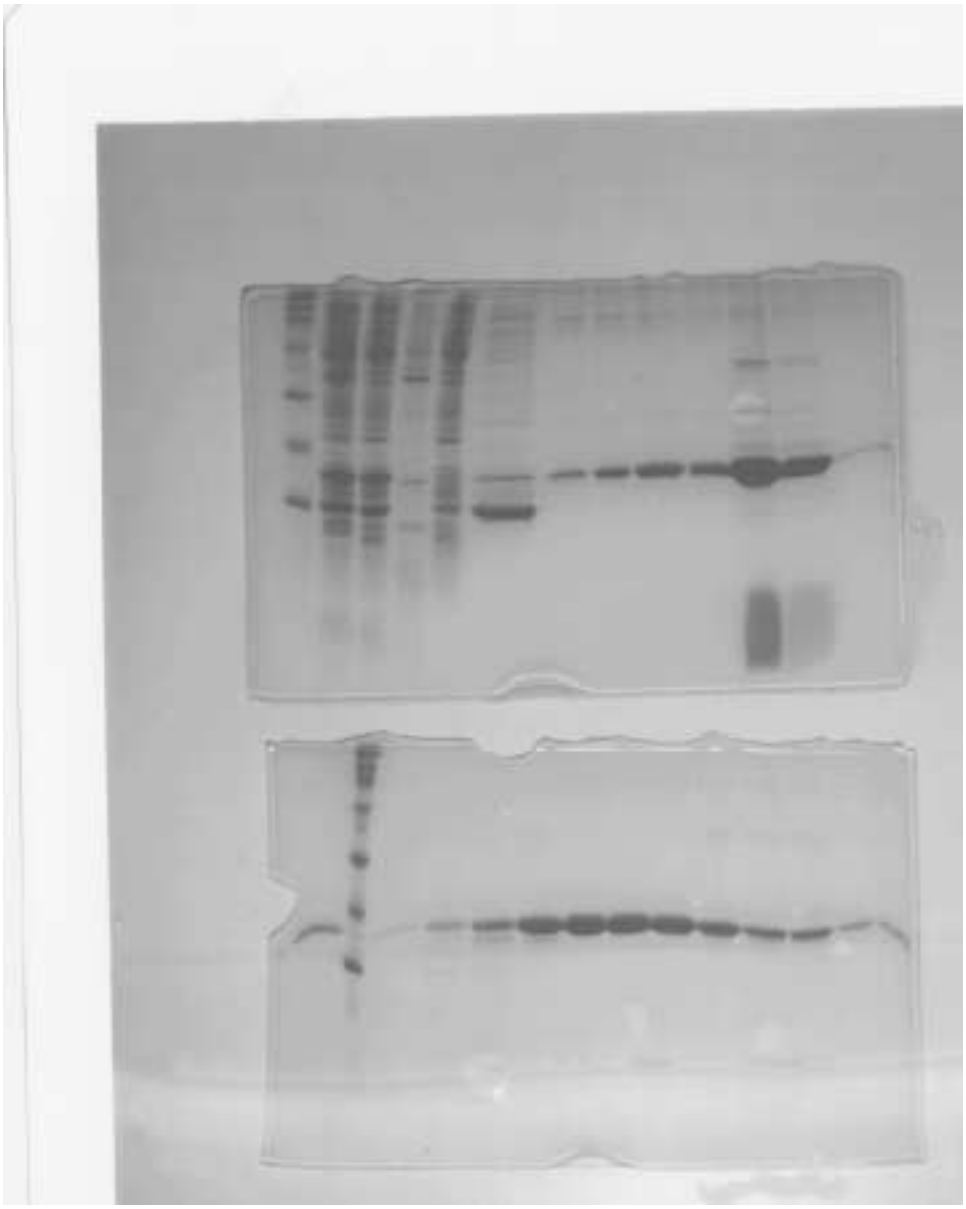

Figure 3: Purification of Z131HN: Affinity chromatography, Cation chromatography

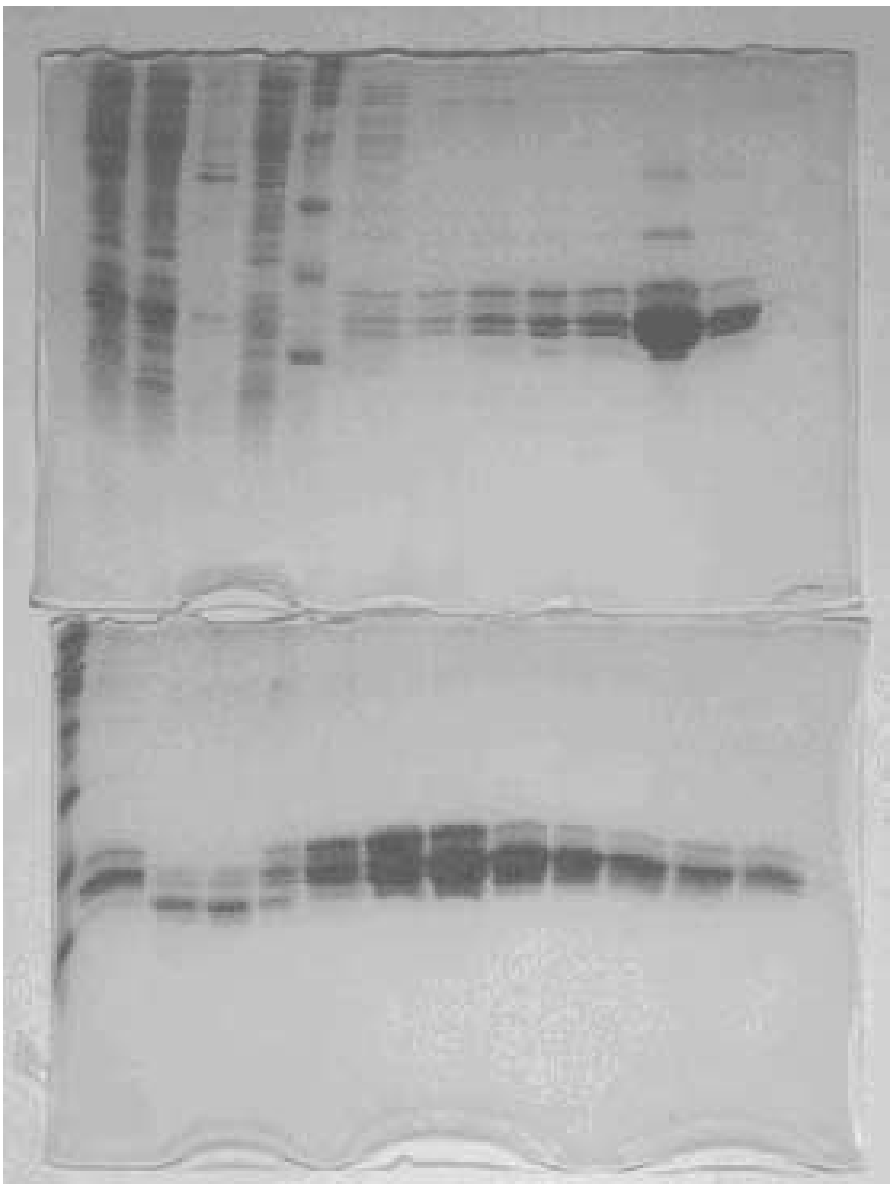

Figure 4: Purification of Z131HC: Affinity chromatography, Cation chromatography

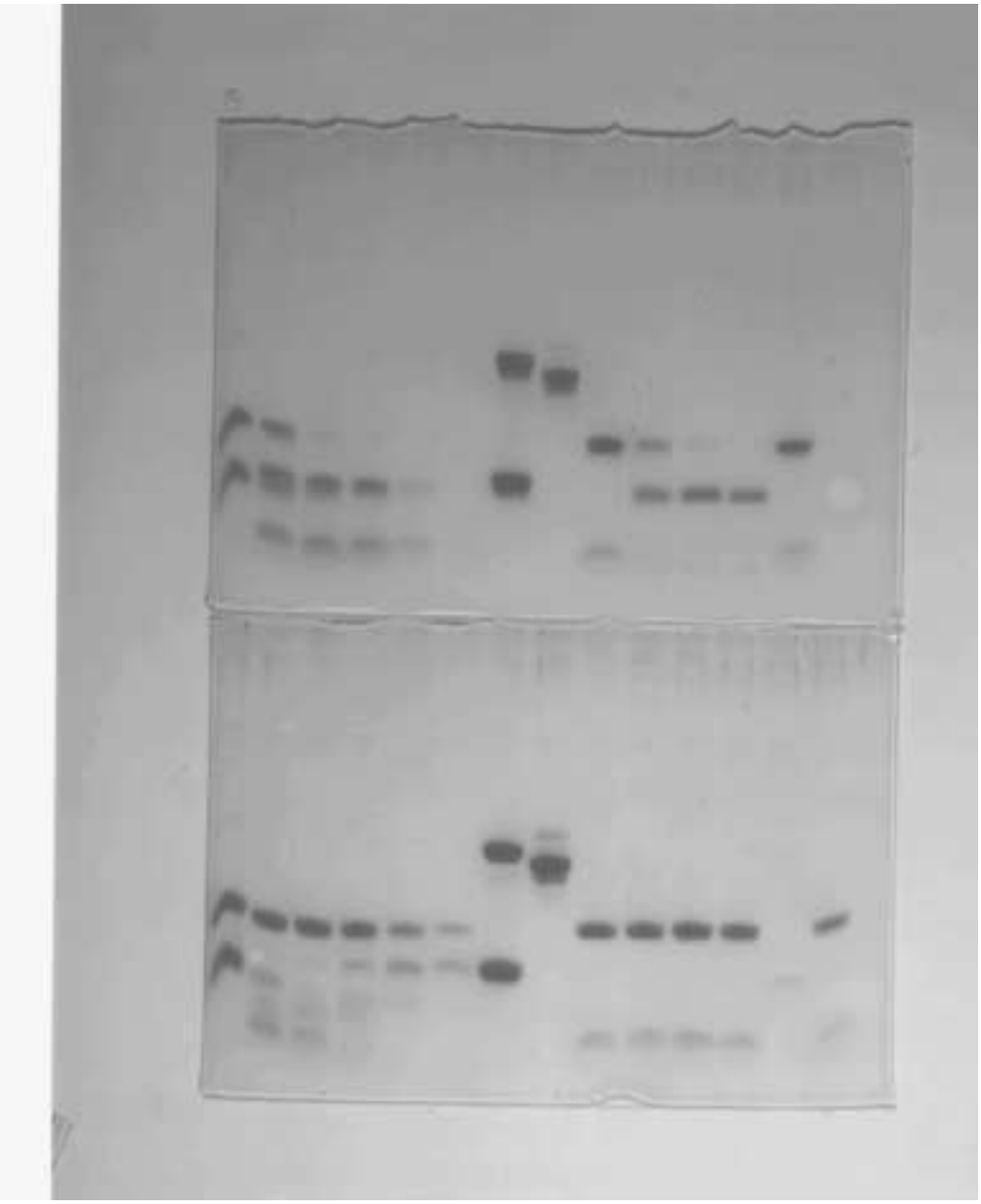

Figure 5: V8/CH double digests: Z131HN, Z131HC

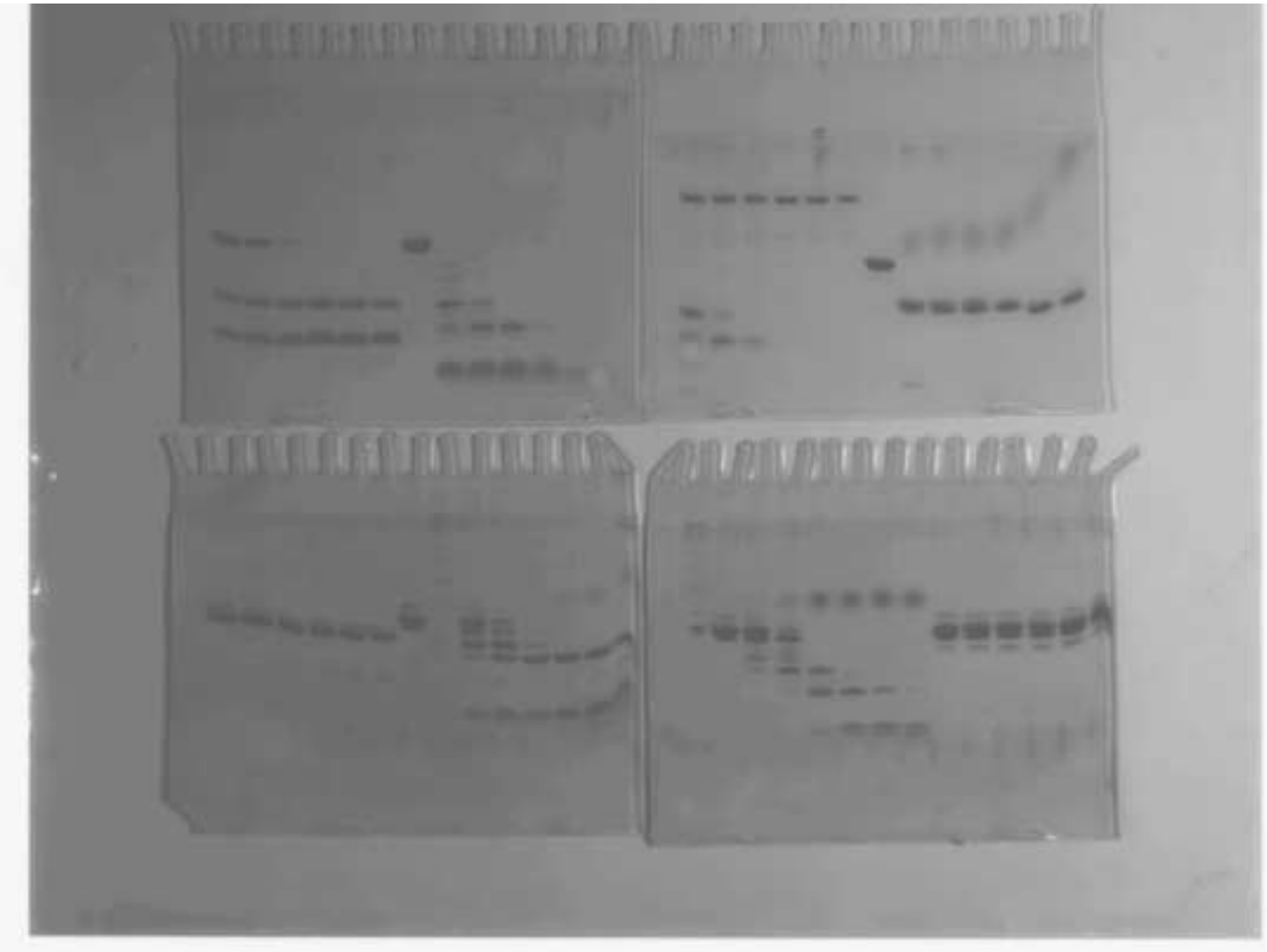

Figure 6: Single digests: Z131HN(V8,CH,TL,TR)-upper gels Z131HC(V8,CH,TL,TR)-lower gels

References

[1] Herbert, A.: RNA editing, introns and evolution. *TIG* **12**: 1

[2] Herbert, A., Alfken, J., Kim, Y., Mian, I.S., Nishikura,K. and Rich, A.: A Z-DNA-binding domainpresent in the human editing enzyme, duoble stranded RNA adenosine deaminase. *Proc. Natl. Acad. Sci. USA* . (1997): in press

[3] Schwartz, T., unpublished data

[4] Liu, T.Y. and Chang, Y. H.:*J. Biol. Chem.*. **216** (1971): 2842

[5] Penke, B., Ferenci, R. and Kovacs, K.: *Anal. Biochem.*. **60** (1974): 45

[6] Laemmli,U.K.: Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage. *Nature*. **227** (1997): 680-685

[7] Schagger, H. and v.Jagow, G.: Tricine-SDS-PAGE for Separation of Proteins in the Range from 1 to 100kDa.*Anal. Biochem.*. **166** (1987): 368-378

[8] Cilley, C.D. and Williamson, J.R.: Analysis of bacteriophage N protein and peptide binding to boxB RNA using polyacrylamide gel electrophoresis(PACE). *RNA*. (1997), 3:57-67

[9] Stults, J.T.: Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). *Current Opinion in Structural Biology*. **5** (1995): 691-698