Unique Peptide Prediction of RNase Family Sequences based on Reinforced Merging Algorithms[*]

HAO-TENG CHANG, TAN-CHI FAN, and MARGARET DAH-TSYR CHANG[#]

Institute of Molecular and Cellular Biology & Department of Life Science, National Tsing Hua University, No. 101, Sec. 2, Kuang Fu Rd.
Hsinchu ,Taiwan 30013, Republic of China

TUN-WEN PAI [#], BO-HAN SU, and PEI-CHIH WU

Department of Computer Science, National Taiwan Ocean University, No. 2, Pei-Ning Rd.
Keelung ,Taiwan 20224, Republic of China

Human ribonuclease A (RNase A) superfamily consists of eight RNases with high sequence homology, in which RNase2 and RNase3 share 78% similarity. The evolutionary variation of RNases results in differential structure and function of the enzymes. To distinguish the characteristics of each RNase, we developed reinforced merging algorithms (RMA) to rapidly predict and identify the unique sequence motifs for each member of the highly conservative human RNaseA superfamily. Two unique regions in RNase3 were predicted and experimentally confirmed to contain epitopes for monoclonal antibodies (mAbs) specifically against RNase3. Our method provides a useful tool for identification of unique sequence motif for further experimental design.

1 Introduction

1.1. *Ribonuclease A superfamily*

Ribonuclease (RNase) catalyzes the degradation of ribonucleic acids (RNAs). All living organisms contain RNases which mainly function in RNA processing such as RNA maturation and turnover. Many abundant ribonucleases have been isolated and well characterized, e.g. bovine pancreatic ribonuclease (RNaseA), but their biological importance is still not clear. Recently, several proteins with known biological activities are found to have intrinsic ribonucleolytic activities and their amino acid sequences are homologous to those of RNaseA superfamily.

The ribonuclease family is composed of proteins with similar primary structures and enzymatic activities. This class of enzymes has been suggested to participate in RNA metabolism and gene-expression regulation [1]. In humans, eight ribonucleases including RNase1 (human pancreatic RNase) [2] , RNase2 (eosinophil-derived neurotoxin) [3], RNase3 (eosinophil cationic protein) [4], RNase4 [5], RNase5 (angiogenin) [6], RNase6 (RNase k6) [7], RNase7 [8], and RNase8 [9], all map to the long arm of chromosome 14[9]. The pairwise similarity comparison of protein sequences is more than 29%, especially RNase2 v.s. RNase3 (76.7%) and RNase7 v.s. RNase8 (78%). From molecular

---

[#] To whom correspondence. E-mail: dtchang@life.nthu.edu.tw; twp@mail.ntou.edu.tw

evolutionary analysis, the gene pairs of RNase2-RNase3 and RNase7-RNase8 have emerged as a result of a relatively recent gene duplication event [9]. Structurally, all RNase A ribonucleases have eight cysteines for four intra-disulfide bonds, one catalytic lysine (K41) and two catalytic histidines (H12 and H119), except RNase5 which has only six cysteines. These features are conserved in the sequence alignments by several available programs such as BLAST or SDSC Biology Workbench. In addition, three-dimensional structures of RNase1, RNase2, Rnase3, and RNase4 all show similar conformation of 3 α-helices and 5 β-sheets.

To distinguish each of the human RNases in terms of sequence specificity, quantity variation, and cellular localization, a monoclonal antibody (mAb) is often employed in biological investigation. Then monoclonal antibody is a probe to specifically recognize and distinguish an antigen. The recognition region, named as epitope, is often composed of 10 to 20 amino acids. In general, only amino acids exposed to the surface can serve as epitopes, however, it is still difficult to predict which region is the best for generation of specific mAb, especially for those proteins without any structural information. In this study, we have developed reinforced merging algorithms (RMA) to rapidly predict and identify the unique sequence motifs for each member of the highly conservative human RNaseA superfamily. Emperically we have generated two mAbs specifically against RNase3 rather than RNase2 and RNase1. Further epitope mapping experiments revealed that the two epitopes were indeed located in two separate long unique regions in RNase3 as predicted by RMA.

1.2 *Computational Methodology*

To precisely allocate the unique peptides from the RNase family employing current biological tools, it is necessary to prepare a large number of well-defined chemical derivatives and partial segments from the protein sequence derivatives for enzymatic activity tests. The complete experiments and pathway from protein sequence to monoclonal antibody and vaccine development are therefore cost-intensive and lengthy. Therefore, it is important and incentive to predict candidate peptides computationally prior to assay development. Early methods for epitope prediction focused on distinguishing possible candidates by seeking the presence of the appropriate primary/secondary anchor residues [10] [11] . These prediction methods are based on the identified and published motifs for anchor identification. Lately, the prediction strategies put emphasis on patterning matching technologies, including statistical and machinal learning methodologies [12] [13]. However, all these developed algorithms require referencing to well-known motif properties or respective structural information. Since the members of RNase family possess strong similarity and each members' structural information has yet to be unveiled completely, here we propose simple and straightforward prediction algorithms based on their associated amino acid position. With the concept of amino acid substitution, enhanced continuous antigenic properties, and statistical properties, different representative

levels of unique peptides can then be extracted manually by adjusting several parameters. More details of the proposed algorithms are introduced in the following sections.

2 Methods and Materials

2.1 *Problem Definition*

RNase family sequences are represented as strings over the 20 amino acid set, the words sequence and string are represented interchangeably. Let Z be the set of collected RNase sequences aforementioned in the RNase family. The Ith sequence in Z will be denoted by Z_I, and the total number of such a family will be indicated by $N=$**Num**(Z). More specifically, the Z set is constructed as $Z = \{Z_1, Z_2,\ldots, Z_I,\ldots, Z_N\}$. Given a subsegment u in Z_I, the natural number n is defined the length of u and is denoted by $[u]$. If u is a string and $n = [u] > 0$, $u = Z_I[i,,j]$ is the substring of Z_I by taking the ith through j^{th} symbols. Two different segments in sequence Z_I can be merged if they possess overlapping symbols in Z_I. Given any two segments u and v with length m and n respectively ($m,n > 0$), if they can be merged with respect to Z_I and denoted by $\mathcal{M}_{ZI}(u,v)$. The merged result is a segment w, hence $[w] = k = m + n - l$, where l is the number of overlapping symbols between u and v, and l is strictly greater than 0. The merged segment w is able to be constructed by

$$w(i)=\begin{cases} u(i) & \text{if } 1 \le i \le m; \\ v(i - m + l) & \text{if } m+1 \le i \le m+n-l. \end{cases}$$

If u and v do not contain any overlapping symbols between them, the merging operation will not change their original relation, i.e. $\mathcal{M}_{ZI}(u,v) = \{u,v\}$.

It is clear that if u and v are allowed to be merged with respect to Z_I, the length of the merged segment w is strictly less than the concatenated string uv, i.e. $[\mathcal{M}_{ZI}(u,v)]<[uv]$. The merging operation is proposed to enhance the discrepancy in a family set and emphasize the neighboring relationship instead of the concatenation operation mentioned in this paper. We also define the set of all unique descriptors with segment length n in Z_I by $P_n[Z_I, \cdot]$, and the jth unique descriptors with segment length n in Z_I is $P_n[Z_I, j]$. $\mathcal{M}_{ZI}(P_n[Z, \cdot])$ is a set of merged unique descriptors with segment length n over Z_I. $P_{n,k}^m[Z_I, j]$ represents the kth pattern in set $P_n^m[Z_I, j]$ which is a set of all patterns with reserving m symbols from the $P_n[Z_I, j]$ descriptor. To enhance the effectiveness of reinforced merging algorithms, we have tried to filter out the unique peptide candidates by considering the amino acid substitution relationship and their evolution features. This is completely done employing the fundamental clustering techniques and their statistical relationship. Here we define the $C(n, m)$ as the number of possible combinatorial pattern set of reserved m symbols from a unique pattern with length n, and $\mathcal{R}(P_n^m[Z_I, j])$ as the representative percentage of the jth unique descriptor $P_n[Z_I, j]$ with m symbols retained in family set Z. We simply employ the following equation for calculating the level of uniqueness of unique peptide from family sequence set Z.

$$\mathcal{R}\left(P_n^m[Z_I, j]\right) = 1 - \frac{1}{C(n,m)} \sum_{k=1}^{C(n,m)} \frac{\left\langle P_{n,k}^m \right\rangle - 1}{N - 1} \tag{1}$$

where $\left\langle P_{n,k}^m \right\rangle$ is the number of the appearing count of the kth pattern $P_{n,k}^m[Z_I, j]$ shown in the set Z and $\left\langle P_{n,k}^m \right\rangle \in [1, N]$. The number of $\left\langle P_{n,k}^m \right\rangle$ can be calculated by counting the appeared condition of the kth pattern $P_{n,k}^m[Z_I, j]$ in each Z_I, $I \in [1, N]$. The value of $\left\langle P_{n,k}^m \right\rangle$ is greater or equal to 1, since there must exist at least one sequence Z_I from set Z which contains the complete unique subsegment $P_n[Z_I, j]$. If the kth tolerant pattern $P_{n,k}^m[Z_I, j]$ can be matched in each sequence in Z, then $\dfrac{\left\langle P_{n,k}^m \right\rangle - 1}{N - 1} = 1$ and $\mathcal{R}\left(P_n^m[Z_I, j]\right) = 1 - 1 = 0\%$. This describes a special condition that all possible tolerant patterns of unique descriptor $P_n^m[Z_I, j]$ are able to be discovered in each sequence in Z. For the other extreme case, when the ratio of $\mathcal{R}\left(P_n^n[Z_I, j]\right)$ is equal to 100% it means $P_n^n[Z_I, j]$ represented as a unique descriptor from the family set Z.

2.2   *System Configuration and Description*

Figure 1 depicts the system configuration. There are three main modules designed in the prediction system, including the grouping, searching, and merging phases. In the first module, users are able to select one of the BLOSUM substitution matrices for clustering 20 amino acids into several groups with respect to threshold values. The clustering algorithms are based on a hierarchical methodology (agglomerative algorithms) which produces a sequence of groupings of a decreasing number of clusters at each step. The clustering produced at each step results from the previous one by merging two clusters into one, the algorithm stops when it meets the threshold value setting. For the second searching module, the system requires the parameters of the primitive segment and the tolerant number in a mismatching situation. Based on these settings, the module performs modified Boyer Moore algorithms to filter out all possible candidate patterns and their representative percentages $\mathcal{R}\left(P_n^m[Z_I, j]\right)$ respectively. In the last merging phase, the system reinforces merging processes to obtain predicted unique peptides. Users are able to specify the minimal length of the final merged segment with local primitive segments that satisfy the assigned representative percentage. The strict merging process provides the most rigorous properties of uniqueness. Any possible partial segment of the merged substring will satisfy the representative percentage requirement.
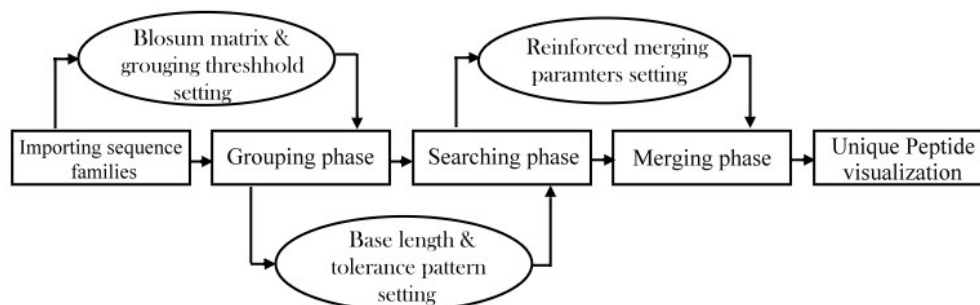
Figure 1. System configuration

The designed systems require a set of sequences first and denote the set of family sequences as symbol Z. Users are then able to decide the length of unique pattern $n=[u]$ and the number of m reserved symbols from $[u]$. $n=m$ means that no tolerant conditions are allowed, systems provide a set of unique patterns with $\mathcal{R}\left(P_n^n[Z_1,j]\right)=100\%$. $0<m<n$ represents that there will be ($n\text{-}m$) symbols in $P_n[Z_1,j]$ which are allowed to be changed from Σ, and each pattern will be calculated for its unique representative ratio $\mathcal{R}$ based on Eq. (1) respectively. If m is set as equal to 0 it means no symbol from segment u is preserved and $\mathcal{R}\left(P_n^0[Z_1,j]\right)=0\%$. After each unique representative ratio is calculated, the system offers the merging operation from the set $P_n^m[Z_1,j]$. Users are able to obtain the merging results from two ways: the first way is to provide the range of the unique representative ratio, and the second way is to provide the minimal length of output merged subsegments. For example, if a user defines the unique representative ratio between [80%, 100%], the system will select the unique patterns with $80\% \leq \mathcal{R}\left(P_n^m[Z_1,j]\right) \leq 100\%$ and merged according the rule of Eq. (1). The second way for observation is to assign the number of minimal length of merged subsegments for a specific sequence Z_1 from Z. For example, we can focus on the sequence RNase3 (Z_{RNase3}) and require at least one element from set $\mathcal{M}\left(P_n^m[Z_{RNase3},\cdot]\right)$ possessing segment length greater than or equal to 15 by decreasing the ratio $\mathcal{R}\left(P_n^m[Z_{RNase3},\cdot]\right)$. Once the merged subsegment is found, the algorithm stops. In both ways we can easily reveal the unique segments with some levels of tolerance from the specified family set. In this example, we set the length of unique pattern $n=[u]=3$ and the number of reserved symbols $m=3$ as well (i.e. without tolerance). After the searching phase, the system extracts 645 basic unique patterns from the family sequences. If users apply RMA and the set minimal merged length of unique pattern to 15, the merged results are shown in Fig. 4. Totally 15 unique patterns found and satisfied with the conditions, and more details will be discussed in the later sections.

2.3 *Molecular Cloning*

The DNA fragments, mature RNase3, mature RNase2, RNase3$_{1\text{-}73}$, RNase3$_{24\text{-}133}$, and RNase3$_{51\text{-}73}$, were prepared by PCR and subsequently cloned into plasmids. The mature RNase3 and RNase2 inserts were ligated into the pET3a and pRSETA vectors, respectively. The other gene fragments were fused to eGFP in pET23a (+) vectors.

Between RNase3 fragments (RNase3$_{1-73}$ and RNase3$_{51-73}$) and eGFP, a linker (222 bps) was inserted. The recombinants, pET3a-mRNase3-H$_6$, pRSETA-mRNase2, pET23a-RNase3-eGFP-H$_6$, pET23a-RNase3$_{1-73}$-eGFP-H$_6$, pET23a-RNase3$_{24-133}$-eGFP-H$_6$, and pET23a-RNase3$_{51-73}$-eGFP-H$_6$, were transformed into *E. coli* BL21 (DE3) for protein expression.

2.4   *Protein Purification, Monoclonal Antibody Generation and Western Blotting*

The inclusion bodies of recombinant proteins were isolated by Novagen user instrument and resuspended in binding buffer under denaturing condition. After centrifugation, the supernatant was loaded into His-Bind® metal chelating chromatography (Novagen). To remove other proteins, the column was washed with binding buffer and wash buffer. The recombinant RNase3 was eluted with elution buffer and the purified proteins were analyzed by SDS-PAGE stained with Coomassie Blue. BALB/c mice were immunized with 10 μg recombinant mature RNase3 suspended in Freund's complete adjuvant. These animals were boosted twice with the same amount of the antigen in Freund's incomplete adjuvant after a 2-week interval. The final boost was done by injection with 10 μg of the antigen without adjuvant. Three days after the last immunization, splenocytes were collected and fused with NS-1 myeloma cells. Hybridoma producing anti- RNase3 mAb were obtained by indirect ELISA screening with purified recombinant RNase3 as the antigen while recombinant RNase2 was used as the negative control. Positive clones were subjected to single-cell cloning twice by the limiting dilution method. mAbs were purified from hybridoma supernatant using the Montage® Antibody Purification Kit (Millipore). For detection of protein, the proteins were transferred to a nitrocellulose membrane. The membrane was incubated in 3% non-fat milk at room temperature for blocking. One hour later, the Ab was diluted with 1.5% non-fat milk and reacted with the target protein at room temperature for 2 hr. The membrane was washed with TTBS (20 mM Tris-HCl, pH 7.6, 0.8% NaCl, 1% Tween-20) three times for 15 min each. Then, the secondary Ab, goat anti-mouse Ab conjugated with horseradish peroxidase, was diluted in the 1.5% non-fat milk and reacted with the first Ab for 2 hr. The membrane washing procedure was repeated three times and the target proteins were visualized using SuperSignal West Pico.

3.   Results

3.1. *Specificity of anti- RNase3 Monoclonal Antibodies*

The recombinant constructs were transformed into *E. coli* BL21 (DE3) and expressed by induction with 0.5 mM IPTG (Fig. 2). The recombinant RNase3 was purified with metal chelating chromatography under denatured condition (Fig. 2A). The purity of recombinant RNase3 was more than 95% and sufficient for monoclonal antibody production. The monoclonal antibodies against RNase3 were successfully produced. To analyze the specificity of anti-RNase3 mAbs, the RNase1, BSA and bacterial extract which contained recombinant RNase3-H$_6$, RNase3-eGFP-H$_6$, eGFP-H$_6$, and RNase2 were separated by 15%

SDS-PAGE (Fig. 2B, upper panel). The proteins were transferred onto nitrocellulose membrane and probed by α3 clone-1 and α3 clone-2 mAbs (Fig. 2B, middle and lower panels). Figure 1B indicated that these two mAbs could specifically recognize RNase3-H_6 and RNase3-eGFP-H_6, but not cross-reacted to RNase1, RNase2 and eGFP tag. Both α3 clone-1 and α3 clone-2 could specifically distinguish RNase3 from RNase2, the two most conservative human RNases.
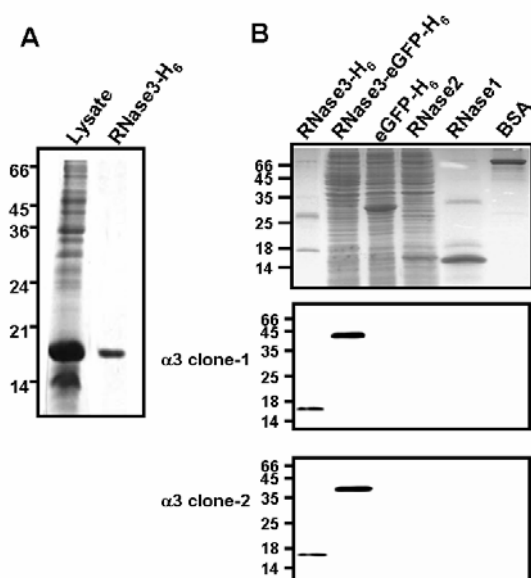

Figure 2. The purification of mature RNase3 and specificities of anti-RNase3 monoclonal antibodies. The mature RNase3 was purified with metal chelating chromatography and located at 18 kDa (A). The RNase3-H_6, RNase3-eGFP-H_6, eGFP-H_6, RNase2, RNaseA and BSA were analyzed by 15% SDS-PAGE (B, upper panel). Both the α3 clone-1 and α3 clone-2 could probe RNase3-H_6 and RNase3-eGFP-H_6, but not eGFP-H_6, RNase2, RNase1 and BSA (negative control) (B, middle and lower panel)

3.2. *Epitopes screening of anti-RNase3 monoclonal antibodies*

For efficient expression, the mature or truncated RNase3 were fused to eGFP-H_6 tag and named as RNase3-eGFP-H_6, RNase3$_{1-73}$-eGFP-H_6, RNase3$_{24-133}$-eGFP-H_6 and RNase3$_{51-73}$-eGFP-H_6 (Fig 3A). These proteins could be expressed in *E. coli* and recognized by anti-His antibody as a positive control (Fig. 3B, upper panel). After the membrane was re-probed by α3 clone-1, it could recognize all four proteins, indicating that the epitope of α3 clone-1 was located within RNase3$_{51-73}$ (Fig. 3B, middle panel). By the same method, the epitope of α3 clone-2 was determined to be located within RNase3$_{24-50}$ (Fig 3B, lower panel). According to the previously mentioned experiments, both mAbs could specifically distinguish RNase3 from RNase2, and the two mAbs recognized different epitopes. These two epitopes may be the most distinguishable regions between RNase2 and RNase3. The

sequence comparisons made by RMA are shown in Fig. 4. The epitopes of α3 clone-1, α3 clone-2, and Boix's mAb [14] were underlined in red, green, and yellow, respectively. In this comparison, the orange and brown characters represent the long-sequential unique sequences (≥ 15 residues), the dark and light blue characters are the short-sequential sequences (< 15 residues), and the black ones are the conserved region in the eight RNases. The sequence comparison by RMA also reveals that the key residues in the enzymatic active sites are located in the conserved regions among most of the RNases.
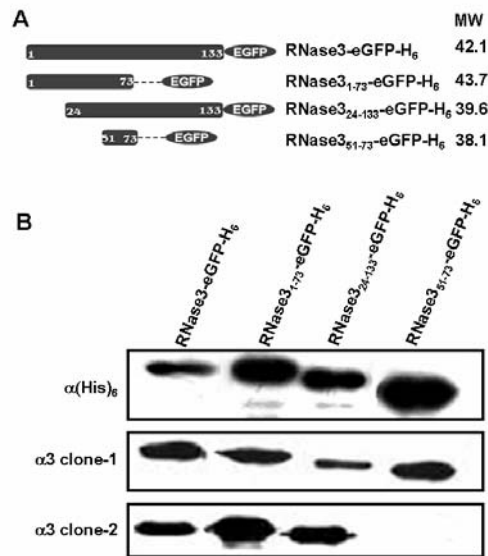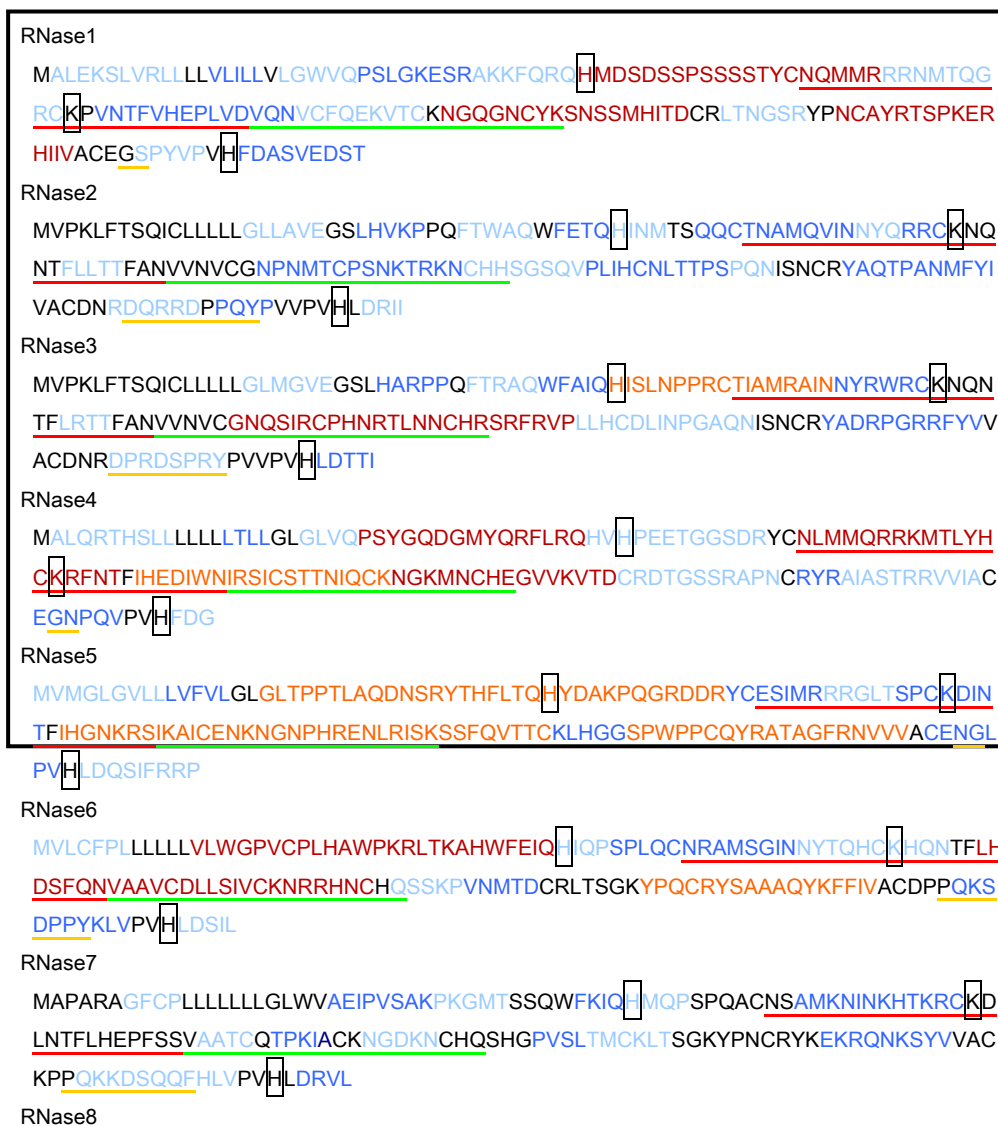

Figure 3. Epitope screening of anti-RNase3 monoclonal antibodies. The cartoon of mature or truncated RNase3 fused to eGFP-H$_6$ was shown (A). Anti-His antibody could probe all the recombinant proteins to prove the protein expression and loading control (B, upper panel). α3 clone-1 could recognize all four recombinant proteins (B, middle panel). α3 clone-2 could probe the first 3 proteins, but not RNase3$_{51-73}$-eGFP-H$_6$ (B, lower panel).

4. Discussion

The determination of protein epitopes is important for generation of peptide antibodies or antibody drugs. The epitopes should possess some properties such as hydrophobicity, antigenicity, and specificity. Several available web programs all predict the epitopes according to these properties. In our RMA program, we enhance the accumulation of unique features to distinguish different regions from eight RNases. The unique sequential sequences labeled by orange, brown (long-sequential sequence) or blue (short- sequential sequences) distinguished RNase3 from all other RNases without considering the biological properties. Two monoclonal antibodies against RNase3 were generated and characterized to specifically recognize RNase3 but not the highly homologous RNase2. Furthermore, the three dimensional structure of RNase3 revealed that three epitopes were all exposed to the surface, and the correspondent regions in RNase1, RNase2, and RNase4 labeled in red,

green, and yellow were evidently different in their structures (Fig. 5). The epitope of α3 clone-1 should be noticed especially. Comparison of RNase2 and RNase3 reveals 77.7% identity. Within this region, 21 amino acids are exactly the same in the stretch of 27 residues, and the 6 different amino acids are separate in such a way that they are usually ignored by biologists in epitope design. However, RMA can rapidly pick up the unique signature from the RNases with high sequence identity. Hence RMA can be used as a powerful tool to identify unique region(s) among several highly homologous proteins.

RNase1

MALEKSLVRLLLLVLILLVLGWVVQPSLGKESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQG
RQKPVNTFVHEPLVDVQNVCFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKER
HIIVACEGSPYVPVHFDASVEDST

RNase2

MVPKLFTSQICLLLLLLGLLAVEGSLHVKPPQFTWAQWFETQHINMTSQQCTNAMQVINNYQRRCKNQ
NTFLLTTFANVVNVCGNPNMTCPSNKTRKNCHHSGSQVPLIHCNLTTPSPQNISNCRYAQTPANMFYI
VACDNRDQRRDPPQYPVVPVHLDRII

RNase3

MVPKLFTSQICLLLLLLGLMGVEGSLHARPPQFTRAQWFAIQHISLNPPRCTIAMRAINNYRWRCKNQN
TFLRTTFANVVNVCGNQSIRCPHNRTLNNCHRSRFRVPLLHCDLINPGAQNISNCRYADRPGRRFYVV
ACDNRDPRDSPRYPVVPVHLDTTI

RNase4

MALQRTHSLLLLLLLTLLGLGLVQPSYGQDGMYQRFLRQHVHPEETGGSDRYCNLMMQRRKMTLYH
CKRFNTFIHEDIWNIRSICSTTNIQCKNGKMNCHEGVVKVTDCRDTGSSRAPNCRYRAIASTRRVVIAC
EGNPQVPVHFDG

RNase5

MVMGLGVLLLVFVLGLGLTPPTLAQDNSRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLTSPCKDIN
TFIHGNKRSIKAICENKNGNPHRENLRISKSSFQVTTCKLHGGSPWPPCQYRATAGFRNVVVACENGL
PVHLDQSIFRRP

RNase6

MVLCFPLLLLLLVLWGPVCPLHAWPKRLTKAHWFEIQHIQPSPLQCNRAMSGINNYTQHCKHQNTFLH
DSFQNVAAVCDLLSIVCKNRRHNCHQSSKPVNMTDCRLTSGKYPQCRYSAAAQYKFFIVACDPPQKS
DPPYKLVPVHLDSIL

RNase7

MAPARAGFCPLLLLLLLGLWVAEIPVSAKPKGMTSSQWFKIQHMQPSPQACNSAMKNINKHTKRCKD
LNTFLHEPFSSVAATCQTPKIACKNGDKNCHQSHGPVSLTMCKLTSGKYPNCRYKEKRQNKSYVVAC
KPPQKKDSQQFHLVPVHLDRVL

RNase8

MAPARAGCCPLLLLLLGLWVAEVLVRAKPKDMTSSQWFKTQ⬚VQPSPQACNSAMSIINKYTERC⬚KDL
NTFLHEPFSSVAITCQTPNIACKNSCKNCHQSHGPMSLTMGELTSGKYPNCRYKEKHLNTPYIVACDP
PQQGDPGYPLVPV⬚HLDKVV

Figure 4. The sequence comparison of 8 ribonucleases by RMA. The eight RNases have three conserved active site residues, His, Lys, and His (boxes). The red, green and yellow underlines show the 3 epitopes of mAb against RNase3. The orange and brown characters are long-sequential unique sequences, and the dark and light blue characters are short-sequential sequences.

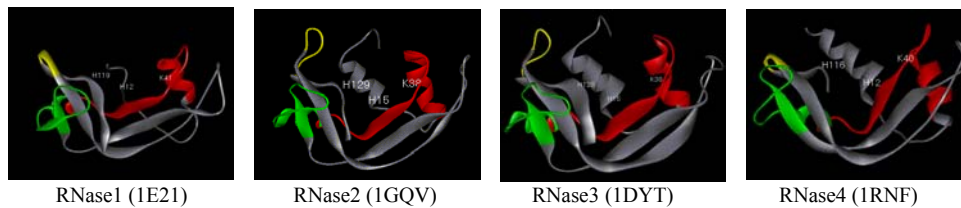RNase1 (1E21) RNase2 (1GQV) RNase3 (1DYT) RNase4 (1RNF)

Figure 5. The three-dimensional structures of human RNase1 (N7-truncated), RNase2, RNase3 and RNase4. The structures of ribonuclease A superfamily contains 3 α-helices and 5 β-sheets. The catalytic residues of RNase are His, Lys and His (white). The epitope region of α3 clone-1, α3 clone-2, and the antibody generated by Boix [14] is labeled in red, green and yellow, respectively. These structures are derived from Protein Data Bank.

5. References

1. D'Alessio, G., *New and cryptic biological messages from RNases.* Trends Cell Biol, 1993. **3**(4): p. 106-9.
2. Weickmann, J.L., M. Elson, and D.G. Glitz, *Purification and characterization of human pancreatic ribonuclease.* Biochemistry, 1981. **20**(5): p. 1272-8.
3. Slifman, N.R., et al., *Ribonuclase activity associated with human eosinophil-derived neurotoxin and eosinophil cationic protein.* J. Immunol., 1986. **137**: p. 2913-2917.
4. Olssom, I. and P. Venge, *Cationic proteins of human granulocytes. I. Isolation of the cationic proteins from the granules of leukaemic myeloid cells.* Scand. J. Haematol., 1972. **9**: p. 204-214.
5. Rosenberg, H.F. and K.D. Dyer, *Human ribonuclease 4 (RNase 4): coding sequence, chromosomal localization and identification of two distinct transcripts in human somatic tissues.* Nucleic Acids Res, 1995. **23**(21): p. 4290-5.
6. Strydom, D.J., et al., *Amino acid sequence of human tumor derived angiogenin.* Biochemistry, 1985. **24**(20): p. 5486-94.
7. Rosenberg, H.F. and D.K. D., *Molecular cloning and characterization of a noval human ribonuclease (RNase k6): increasing diversity in the enlarging ribonuclease gene family.* Nucleic Acids Res., 1996. **24**: p. 3507-3513.
8. Zhang, J., K.D. Dyer, and H.F. Rosenberg, *Human RNase 7: a new cationic ribonuclease of the RNase A superfamily.* Nucleic Acids Res, 2003. **31**(2): p. 602-7.
9. Zhang, J., K.D. Dyer, and H.F. Rosenberg, *RNase 8, a novel RNase A superfamily ribonuclease expressed uniquely in placenta.* Nucleic Acids Res, 2002. **30**(5): p. 1169-75.
10. Collins, E.J., D.N. Garboczi, and D.C. Wiley, *Three-dimensional structure of a peptide extending from one end of a class 1 MHC binding site.* Nature, 1994. **371**(6498): p. 626-9.

11. Ruppert, J., et al., *Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules.* Cell, 1993. **74**(5): p. 929-37.
12. Donnes, P. and A. Elofsson, *Prediction of MHC class I binding peptides, using SVMHC.* BMC Bioinformatics, 2002. **3**(1): p. 25.
13. Gulukota, K., et al., *Two complementary methods for predicting peptides binding major histocompatibility complex molecules.* J Mol Biol, 1997. **267**(5): p. 1258-67.
14. Boix, E., et al., *Identification and characterization of human eosinophil cationic protein by an epitope-specific antibody.* J Leukoc Biol, 2001. **69**(6): p. 1027-35.