doi:10.1016/j.jmb.2006.07.020 J. Mol. Biol. (2006) 362, 459–478

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Genetic Selection for Critical Residues in Ribonucleases

Bryan D. Smith[1] and Ronald T. Raines[1,2]*

[1]*Department of Biochemistry University of Wisconsin–Madison Madison, WI 53706, USA*

[2]*Department of Chemistry University of Wisconsin–Madison Madison, WI 53706, USA*

**Corresponding author*

Homologous mammalian proteins were subjected to an exhaustive search for residues that are critical to their structure/function. Error-prone polymerase chain reactions were used to generate random mutations in the genes of bovine pancreatic ribonuclease (RNase A) and human angiogenin, and a genetic selection based on the intrinsic cytotoxicity of ribonucleolytic activity was used to isolate inactive variants. Twenty-three of the 124 residues in RNase A were found to be intolerant to substitution with at least one particular amino acid. Twenty-nine of the 123 residues in angiogenin were likewise intolerant. In both RNase A and angiogenin, only six residues appeared to be wholly intolerant to substitution: two histidine residues involved in general acid/base catalysis and four cysteine residues that form two disulfide bonds. With few exceptions, the remaining critical residues were buried in the hydrophobic core of the proteins. Most of these residues were found to tolerate only conservative substitutions. The importance of a particular residue as revealed by this genetic selection correlated with its sequence conservation, though several non-conserved residues were found to be critical for protein structure/function. Despite voluminous research on RNase A, the importance of many residues identified herein was unknown, and those can now serve as targets for future work. Moreover, a comparison of the critical residues in RNase A and human angiogenin, which share only 35% amino acid sequence identity, provides a unique perspective on the molecular evolution of the RNase A superfamily, as well as an impetus for applying this methodology to other ribonucleases.

Keywords: angiogenin; error-prone PCR; genetic selection; molecular evolution; ribonuclease A

Introduction

Sequence alignments of homologous proteins suggest that only a subset of residues are critical for the formation and maintenance of their three-dimensional structures.[1,2] This supposition has been confirmed by systematic experimental analyses in which every residue in a protein is replaced individually or in combination.[3–13] Such analyses have revealed that amino acid sequences are indeed plastic, but not completely so. For example, the overall conservation of a residue is correlated with its importance, and buried residues are especially intolerant to substitution. An exhaustive search for critical residues has not, however, been done on homologous proteins.

Nearly 50 years ago, Anfinsen and co-workers employed bovine pancreatic ribonuclease (RNase A; EC 3.1.27.5) to reveal that the amino acid sequence of a protein is sufficient to encode its three-dimensional structure.[14,15] In the decades since this seminal discovery, RNase A has been used as a model to study nearly all aspects of protein structure and function, including enzymatic catalysis; protein folding, structure, and stability; and molecular evolution.[16–28] The determination of the

Abbreviations used: ALS, amyotrophic lateral sclerosis; ANG, human angiogenin; ASA, accessible surface area; B, cytosine, guanine, or thymine; CFIS, chain-folding initiation site; D, adenine, guanine, or thymine; H, adenine, cytosine, or thymine; PBS, phosphate-buffered saline; PDB, Protein Data Bank; RNase A, bovine pancreatic ribonuclease; rDNA, recombinant DNA; rRNA, ribosomal RNA; V, adenine, cytosine, or guanine.

E-mail address of the corresponding author: raines@biochem.wisc.edu

amino acid sequences of many RNase A homologs (especially by Beintema and co-workers[22,23,29]) together with related structural information[20,30] provide a unique opportunity to reveal the interplay between a sequence of amino acids, its conservation, and the structure/function of the ensuing protein.

Homologs of RNase A have been found to display intriguing biological activities.[24,31–35] For example, the amino acid sequence of angiogenin is 35% identical to that of RNase A.[36] But unlike RNase A, angiogenin promotes the growth of new blood vessels,[37–39] an activity that relies on its ribonucleolytic activity.[40,41] In addition, angiogenin acts as a transcription factor for rRNA,[42] as well as an antibacterial and antifungal protein involved in host defense.[43] Recently, mutations in the angiogenin gene have been linked to amyotrophic lateral sclerosis (ALS), a neurodegenerative disorder.[44] Such findings have led to an increasing interest in structure–function relationships within the RNase A superfamily, and have made ribonucleases both attractive targets and useful tools in disease research.[31–35]

Here, systematic mutation of the genes encoding RNase A and human angiogenin are combined with a genetic selection to reveal residues that are critical for the structure/function of these homologous ribonucleases. Even though RNase A was perhaps the most studied enzyme of the 20st century,[16–28] many "new" critical residues are identified in its amino acid sequence. Moreover, reaching across the evolutionary landscape to human angiogenin provides an unprecedented perspective on those residues that are truly important for structure/function within the RNase A superfamily.

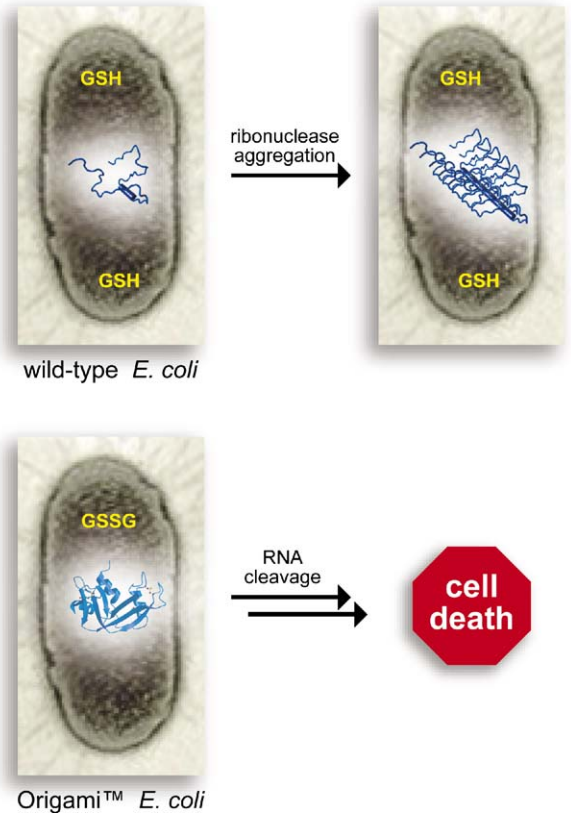

Figure 1. Genetic selection system for the RNase A superfamily. RNase A and its homologs aggregate rather than fold when produced in typical strains of *E. coli*. The Origami™ strain has a more oxidizing cytosol, allowing for disulfide-bond formation and proper folding. Accordingly, RNase A and its homologs are toxic to Origami™ cells, unless missing a residue essential for structure/function. GSH, reduced glutathione; GSSG, oxidized glutathione.

Results and Discussion

Genetic selection system for ribonucleases

Genetic selection can be used to identify a single encoded biopolymer with a particular attribute from a library as large as 10^{10}, allowing discoveries to be made that are not possible by other means.[45] To identify the critical residues in RNase A and angiogenin, it was necessary to develop a selection system for the structure/function of a pancreatic-type ribonuclease. The ability of a ribonuclease to cleave cellular RNA and thereby cause cell death provides an obvious basis for such a system. Its high transformation efficiency and other attributes make *Escherichia coli* an ideal host for a genetic selection.[46] RNase A and its homologs are not toxic, however, when produced in typical *E. coli* strains. These enzymes contain three or four disulfide bonds,[26] and do not fold in the reducing environment of the *E. coli* cytosol (Figure 1).[47]

Strains of *E. coli* have been engineered to allow for disulfide-bond formation in the cytosol.[48–50] For example, the Origami™ strain (Novagen, Madison, WI) has mutations in the thioredoxin reductase and glutathione reductase genes that combine to enable disulfide-bond formation in the cytosol. Consequently, plasmid-encoded RNase A homologs could be toxic to Origami™ cells, thus enabling the identification of inactive variants (Figure 1).

Wild-type RNase A and wild-type human angiogenin were indeed toxic to Origami™ cells (Figure 2). Leaky expression of the RNase A or angiogenin gene from an uninduced $\mathbf{P}_{tac}$ promoter was sufficient to kill those cells. To identify a threshold of activity that allows for cell growth, Origami™ cells were transformed with plasmids encoding active-site variants that were known to have a range of ribonucleolytic activities, as measured *in vitro* with defined RNA substrates. (In contrast, a cell contains a myriad of RNA sequences and structures, making the catalytic activity of a ribonuclease variant measured *in vitro* only a crude approximation of its activity within an Origami™ cell.)

The ribonucleolytic activity of K41R RNase A is ~140-fold less than that of the wild-type enzyme,[51] yet this variant was toxic to Origami™ cells. On the other hand, H12A RNase A has $>10^4$-fold

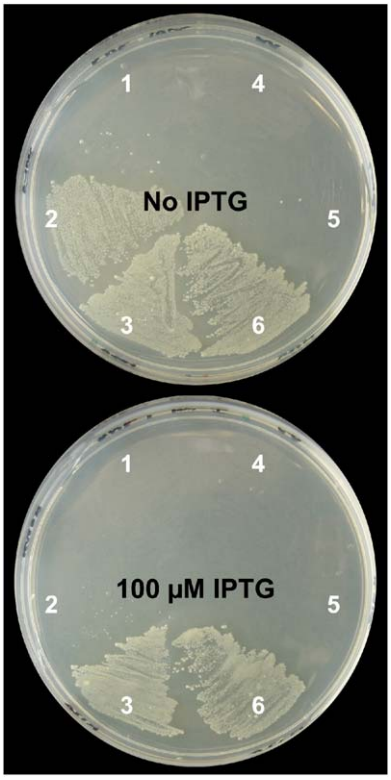

Figure 2. Threshold activity of the genetic selection system. Plasmids encoding RNase A or angiogenin were transformed into Origami™ cells and allowed to grow for two days in the absence or presence of IPTG, which induces gene expression. 1, Angiogenin; 2, K40R angiogenin (~2% catalytic activity[40]); 3, H13A angiogenin (~0.01% catalytic activity[41]); 4, RNase A; 5, K41R RNase A (~0.7% catalytic activity[51]); 6, H12A RNase A (~0.01% catalytic activity[52]).

diminished catalytic activity[52] and allowed for cell growth (Figure 2). Thus, the threshold activity of RNase A for growth in our selection system was between ~0.01% and ~0.7% of wild-type activity. This threshold activity is fateful, as the number of residues identified as "critical" depends on the stringency of a genetic selection.[11] A less stringent selection would have led to the growth of more Origami™ colonies, but revealed little about which residues are the most important to protein structure/function; whereas a more stringent selection would have led to fewer colonies, but revealed no information about residues that have a substantial but not critical impact on protein structure/function. The threshold activity of our selection system is similar to that found previously for barnase, a microbial ribonuclease that lacks disulfide bonds.[11]

Angiogenin has a ribonucleolytic activity that has been measured *in vitro* with defined substrates to be $>10^5$-fold less than that of RNase A.[53] Yet, angiogenin is toxic to Origami™ cells, even without induction of its gene. There is evidence that angiogenin could be an effective catalyst for the cleavage of specific cellular RNAs (such as rRNA[54,55]), RNA secondary structures, or protein·RNA complexes,[53,56,57] and these activities could contribute to its cytotoxicity.

The threshold activity of angiogenin, like that of RNase A, was determined empirically. A K40R substitution in the active site of angiogenin that decreases ribonucleolytic activity by ~50-fold[40] allowed for cell growth. Yet, when transcription of its gene was induced with 100 μM IPTG, K40R angiogenin was cytotoxic. H13A angiogenin, which has $>10^4$-fold less catalytic activity than the wild-type enzyme,[41] allowed for cell growth even upon induction with 100 μM IPTG (Figure 2). Thus, the threshold activity of angiogenin for growth upon induction in our selection system was between ~0.01% and ~2% of wild-type activity. To enable a comparison of critical residues in RNase A and angiogenin, all selection experiments described herein were performed in the absence of IPTG for RNase A and the presence of 100 μM IPTG for angiogenin, conditions that provide a similar threshold activity.

It is noteworthy that colonies containing plasmids that should cause cell death appeared at a frequency of $\sim10^{-6}$ (Figure 2). In these rare colonies, insertions–deletions and other spontaneous mutations were found in the gene encoding RNase A or angiogenin, or recombination was found in the plasmid (data not shown). These mutations were likely due to errors by the *E. coli* replication machinery.[58,59]

Creation of single-base substitution libraries

Error-prone PCR was used to generate libraries of mutated RNase A and angiogenin genes. This method of systematic mutagenesis offers advantages over alternatives. For example, using primer-base methods to mutate each codon[11,13,60] or using amber-suppression strains with a library of genes containing amber codons[6,8,9] are much more arduous. Error-prone PCR has, however, an evident drawback: mutational bias.[61–63]

To overcome mutational bias, two distinct polymerases were used to generate libraries of mutated genes. The mutational bias and randomness of these libraries were analyzed by using *E. coli* DH5α cells (Table 1), which are not susceptible to RNase A and angiogenin. In agreement with previous results,[61–63] using *Taq* polymerase in the presence of Mn^{2+} led to a large excess of A/T→G/C mutations, though C/G→T/A and A/T→T/A mutations were also well represented. A/T→C/G, C/G→A/T, and C/G→G/C mutations were nearly absent from the *Taq* libraries (Table 2). To compensate for the mutational bias of *Taq* polymerase, libraries were also prepared with Mutazyme® polymerase,[61–63] which is biased towards C/G→D/H mutations.[64] Mutazyme® polymerase was found to be biased towards C/G→T/A mutations, though A/T→G/C and A/T→T/A mutations were also well represented (Table 2). Thus, for both *Taq* and Mutazyme® polymerases, the three most common mutations were A/T→G/C, A/T→T/A, and C/G→T/A, but to different extents. A/T→C/G, C/G→A/T, and

Table 1. Overview of error-prone PCR libraries

Polymerase	*Taq*		Mutazyme®	
Gene	RNase A	Angiogenin	RNase A	Angiogenin
Library size	1.1×10^6	9.6×10^5	1.5×10^5	2.6×10^5
Base-pairs sequences	33,480	33,210	30,132	35,793
Mutations	81	82	45	44
Mutation rate (%)	0.24	0.25	0.15	0.12
Insertions[a]	0 (0%)	0 (0%)	1 (2.2%)	0 (0%)
Deletions[b]	4 (4.7%)	1 (1.2%)	0 (0%)	3 (6.4%)
Sequences with 0 mutations (%)	47.8	36.7	63.0	60.8
Sequences with 1 mutation (%)	30.0	43.3	23.5	34.0
Sequences with 2 mutations (%)	13.3	13.3	8.6	4.1
Sequences with ≥3 mutations (%)	8.9	6.7	4.9	1.0

Data were obtained with *E. coli* DH5α cells, which are not susceptible to RNase A or angiogenin.
[a] Number in parentheses indicates the percentage of total mutations that resulted from insertions.
[b] Number in parentheses indicates the percentage of total mutations that resulted from deletions.

C/G→G/C mutations were underrepresented in all four libraries, but were more abundant in the Mutazyme® than the *Taq* libraries.

Mutagenesis conditions were designed to provide one mutation per gene. Under these conditions, the *Taq* libraries had a mutation frequency of ~0.24% (one mutation per ~400 bp), whereas the Mutazyme® libraries had a mutation frequency of ~0.14% (one mutation per ~700 bp) (Table 1). Of course, individual gene sequences contained a discrete number of mutations. Genes with no mutations were cytotoxic, and did not affect the results. Genes with multiple missense mutations were excluded from the analysis, as it is unclear which mutation or combination thereof was important. No genes containing only a silent mutation were isolated, indicating that rare codons did not have a discernable effect on our selection system.

RNase A and angiogenin gene sequences were obtained from Origami™ colonies that grew under selection conditions. For RNase A, 386 of the 1920 gene sequences (*Taq*: 686; Mutazyme®: 1274) were found to contain a single missense mutation. These 386 mutations led to 49 distinct changes to amino acid residues (Table 3, column 3), the other 337 being redundant. For angiogenin, 377 of the 1920 gene sequences (*Taq*: 768; Mutazyme®: 1152) contained a single missense mutation leading to 66 distinct changes to amino acid residues (Table 4, column 3), the other 311 being redundant. The large number of redundant changes found in both RNase A and angiogenin admonishes another drawback of libraries made with error-prone PCR: uncertainty that a specific mutation is represented in the library. For example, F8I RNase A should be represented in our libraries, as it arises from a common A/T→T/A mutation. An F8S substitution inactivated RNase A, allowing for growth in Origami™ cells (*vide infra*). Yet, an F8I variant was not isolated. We assume (but with confidence) that an F8I substitution is not sufficiently deleterious to enzyme structure/function to allow for growth in Origami™ cells.

Despite its drawbacks, we believe that error-prone PCR along with a suitable genetic selection (or screening) system is a superior method for identifying quickly and comprehensively those residues that are essential for protein structure/function. Reagents for error-prone PCR are inexpensive, and the protocols are neither time-consuming nor difficult. Large libraries can be created in a few days, after which experiments are limited only by the growth of colonies, and the determination and analysis of DNA sequences.

Selection for inactivating mutations

Error-prone PCR libraries were transformed into Origami™ cells. Colonies were picked after two days of growth at 37 °C. (Origami™ cells grow more slowly than do typical *E. coli* cells.) Fewer colonies were obtained at lower temperatures (data not shown), presumably because a substitution would then have had to reduce the conformational stability of the enzyme more drastically to allow for colony growth. Colonies were grown in 96 1 ml cultures for 24 h at 37 °C. Plasmid DNA was isolated, and the sequence of the RNase A or angiogenin gene was obtained in both directions.[65] These sequences were aligned, and mutations were detected computationally. Each sequence was also scanned manually for insertions and deletions, and for artifacts.

RNase A and angiogenin: a global view of inactivating mutations

The structure/function of RNase A is quite tolerant of single amino acid substitutions. Of the 124 residues in RNase A, 101 can be changed without inactivating the enzyme (Table 3; Figures 3 and 4; Movie S1 in the Supplementary Material). Moreover, only six residues in the enzyme appear to be wholly intolerant of substitution. The chance of randomly replacing an amino acid residue in RNase A and

Table 2. Bias in error-prone PCR libraries

Polymerase	*Taq*		Mutazyme®	
Gene	RNase A	Angiogenin	RNase A	Angiogenin
A/T→C/G (%)	2.5	1.2	8.9	9.1
A/T→G/C (%)	56.8	59.8	22.2	15.9
A/T→T/A (%)	12.4	26.8	13.3	18.2
C/G→A/T (%)	6.2	1.2	13.3	6.8
C/G→G/C (%)	1.2	0.0	8.9	15.9
C/G→T/A (%)	21.0	11.0	33.3	34.1
Ts/Tv[a]	3.5	2.4	1.2	1.0
A/T→B/V (%)	71.6	87.8	44.4	43.2
C/G→D/H (%)	28.4	12.2	55.6	56.8

Data were obtained with *E. coli* DH5α cells, which are not susceptible to RNase A or angiogenin.
[a] Transitions/transversions.

Table 3. Inactivating mutations in bovine pancreatic ribonuclease

Group	Residue	Mutations isolated	Mutations not isolated[a]	Conservation (%)[b]	Conservation (%)[c]	ASA (Å^2)	Notes
I	H12	D,L,Q,R,Y	–	98.1	100.0	10.8	General base catalysis
I	K41	E,M,N	R	99.3	100.0	37.8	Hydrogen bond to transition state
I	H119	L,Q,R,Y	–	98.9	100.0	82.8	General acid catalysis
II	F8	S,V	I,L,Y	98.1	100.0	3.5	Interacts with H12; hydrophobic core
II	N44	Y	D,I,S	98.1	100.0	3.2	Hydrogen bonds to H12 and K41; hydrophobic core
II	F46	C,S	I,L,Y	99.6	100.0	0.0	Hydrophobic core; CFIS
II	V54	D	A,I	72.2	50.0	0.1	Hydrophobic core
II	V57	E	A,M	63.2	50.0	0.0	Hydrophobic core
II	Y73	D	C,F,H,N	26.7	34.6	32.4	Hydrophobic core
II	S75	I,R	C,G,N	89.5	100.0	1.9	Hydrophobic core
II	Y97	C,N	F,H	99.6	100.0	10.0	Hydrogen bond to K41; hydrophobic core
II	I106	N	F,T,V	49.6	80.8	0.0	Hydrophobic core; CFIS
II	V108	E,G	A,M	77.4	92.3	0.0	Hydrophobic core; CFIS
II	A109	D	T,V	88.0	84.6	2.6	Interacts with H119; hydrophobic core; CFIS
II	F120	S	I,L,Y	45.5	61.5	25.6	Hydrogen bond to transition state; hydrophobic core
III	Q60	P	L,R	34.6	50.0	42.4	Hydrogen bonds to V57
III	T82	P	A,I,S	86.1	100.0	0.2	Hydrogen bonds to F46
IV	C26	R,S,Y	–	100.0	100.0	0.0	Forms disulfide bond with C84
IV	C58	F,G,R,S,Y	–	100.0	100.0	9.5	Forms disulfide bond with C110
IV	C65	F,Y	R,S	79.3	50.0	8.4	Forms disulfide bond with C72
IV	C72	S,W	R,Y	80.1	50.0	0.0	Forms disulfide bond with C65
IV	C84	R,S,Y	–	98.9	100.0	0.0	Forms disulfide bond with C26
IV	C110	R,S,Y	–	99.6	100.0	2.0	Forms disulfide bond with C58

[a] Mutations that were likely present in the error-prone PCR libraries, resulting from A/T→G/C, A/T→T/A, or G/C→A/T substitutions.
[b] Percentage of a set of 266 ribonucleases that contain the residue found in RNase A.
[c] Percentage of a set of pancreatic ribonucleases and angiogenins from 13 species that contain the residue found in RNase A.

inactivating the enzyme is 12.7%. This value, which has been dubbed the "*x*-factor",[12] includes the generation of stop codons. Due to the degeneracy of the genetic code, the probability of inactivating RNase A with a single nucleobase substitution (the nucleotide *x*-factor) is 9.4%.

The structure/function of angiogenin is slightly more sensitive to amino acid substitutions (Table 4; Figures 3 and 4). The *x*-factor for angiogenin is 16.2% (11.5% on the nucleotide level). This greater sensitivity could be due, at least in part, to angiogenin having lower ribonucleolytic activity than RNase A. In Origami™ cells, inducing the expression of the angiogenin gene with IPTG produced results that were qualitatively similar to those achieved with just the leaky expression of the RNase A gene (Figure 2). Of the 123 residues in angiogenin, changes at 94 residues do not inactivate the enzyme. As with RNase A, all but six residues are tolerant of some substitutions.

The *x*-factor is not an independent variable. RNase A and angiogenin are small enzymes (124 and 123 residues, respectively) that are similar in size to barnase (110 residues) and T4 lysozyme (164 residues). The low probability for protein inactivation observed here is similar to the ~5% of single amino acid replacements found to inactivate barnase[11] and the ~16% of changes found to inactivate T4 lysozyme.[6] The *x*-factor of each enzyme in this group correlates with its size. Still, the threshold activity used in the selection also contributes greatly. Mutations in the barnase gene were identified only if they reduced enzymatic activity by at least ~10^3-fold.[11] Our selection system required a ~140 and ~50-fold reduction in the ribonucleolytic activity of RNase A and angiogenin, respectively (Figure 2). T4 lysozyme activity had to be decreased by ~30-fold to be scored as defective.[6] The 33 kDa human 3-methyladenine DNA glycosylase had an *x*-factor of 34(±6)% in an assay in which positives had 5–10% of wild-type activity.[12] Thus, the *x*-factor is a variable that is largely dependent on the threshold activity.

In all selection systems, residues are identified as essential only if they are critical for a specific function. In our system, mutations were identified only if they severely decreased the ribonucleolytic activity of RNase A or angiogenin. In animals, these enzymes are known to interact with the ribonuclease inhibitor protein (RI)[66] and to have other functions, such as the promotion of angiogenesis.[67] Although we have identified the residues that are critical for ribonucleolytic activity, the number of residues critical for all biological functions of each enzyme is likely to be higher.

RNase A: essential residues

Of the 124 residues in RNase A, 23 were found to be intolerant to substitution with one or more

Table 4. Inactivating mutations in human angiogenin

Group	Residue	Mutations isolated	Mutations not isolated[a]	Conservation (%)[b]	Conservation (%)[c]	ASA (Å^2)	Notes
I	H13	D,L,Q,R,Y	–	98.1	100.0	12.2	General base catalysis
I	K40	E,I,N	R	99.3	100.0	61.3	Hydrogen bond to transition state
I	H114	L,N,P,Q,R,Y	–	98.9	100.0	92.5	General acid catalysis
II	F9	C,I,S,V	L,Y	98.1	100.0	3.4	Interacts with H13; hydrophobic core
II	M30	K	I,L,T,V	97.4	100.0	0.0	Hydrophobic core
II	R33	G,P	C,H	50.4	96.2	60.6	Shields C26 and M30 from solvent
II	N43	Y	D,I,S	98.1	100.0	0.0	Hydrogen bonds to H13 and K40; hydrophobic core
II	T44	K,P	A,I,S	98.9	100.0	0.0	Base specificity; hydrophobic core
II	F45	C,S	I,L,Y	99.6	100.0	0.0	Hydrophobic core
II	I46	N	F,T,V	22.6	46.2	0.1	Hydrophobic core
II	H47	Q,R	L,Y	84.2	100.0	10.8	Conformational change in catalysis
II	I53	N	F,T,V	17.7	50.0	3.3	Hydrophobic core
II	I56	N	F,T,V	26.7	50.0	0.0	Hydrophobic core
II	S72	N,R	C,G	89.5	100.0	5.0	Hydrophobic core
II	V78	D	A,I	16.9	42.3	1.5	Hydrophobic core
II	T79	A,N,P	I,S	86.1	100.0	0.0	Hydrophobic core
II	Y94	N	C,F,H	99.6	100.0	2.7	Hydrogen bond to K41; hydrophobic core
II	V103	D	A,I	8.7	19.2	0.1	Hydrophobic core
II	P112	H,R,S	L	98.5	100.0	0.2	Hydrophobic core
II	L115	S	M	50.0	38.5	6.9	Hydrogen bond to transition state; hydrophobic core
III	L10	P	Q	16.2	50.0	21.7	Lies in α-helix between F9 and H13
III	Q12	P	L,R	83.8	80.8	30.8	Lies in α-helix between F9 and H13
III	R95	P	Q	24.8	84.6	95.8	Interacts with Y94
IV	C26	R,S,W,Y	–	100.0	100.0	0.0	Forms disulfide bond with C81
IV	C39	F,Y	R,S	99.6	100.0	21.0	Forms disulfide bond with C92
IV	C57	F,G,R,S,Y	–	100.0	100.0	27.5	Forms disulfide bond with C107
IV	C81	G,R,S,W,Y	–	98.9	100.0	0.0	Forms disulfide bond with C26
IV	C92	R	S,Y	99.6	100.0	7.1	Forms disulfide bond with C39
IV	C107	R,S,Y	–	99.6	100.0	39.2	Forms disulfide bond with C57

[a] Mutations that were likely present in the error-prone PCR libraries, resulting from A/T→G/C, A/T→T/A, or G/C→A/T substitutions.
[b] Percentage of a set of 266 ribonucleases that contain the residue found in angiogenin.
[c] Percentage of a set of pancreatic ribonucleases and angiogenins from 13 species that contain the residue found in angiogenin.

amino acid (Table 3; Figures 3 and 4; Movie S1 in the Supplementary Material). To codify their analysis, the amino acid substitutions that inactivate RNase A were divided into four groups, which are similar to those devised in a high-throughput mutagenesis study on barnase.[11] Group I contains substitutions that alter residues directly involved in catalysis. Group II consists of changes to buried residues. Group III comprises substitutions in which proline is the only residue found at a specific position. Group IV contains changes to essential half-cystines. Barnase lacks disulfide bonds, thus this last group is new to our study. There is redundancy among the groups. For example, many buried residues (Group II) interact with a residue directly involved in catalysis (Group I).

Group I residues (catalytic residues)

His12 and His119

There are three Group I replacements in RNase A. The two renowned active-site histidine residues (His12 and His119) are essential. This result is not surprising, as the mechanism of the cleavage reaction has long been known to employ these two residues,[68–70] with His12 abstracting a proton from the 2′ oxygen while His119 adds a proton to the 5″ oxygen.[25,28,71] Substitutions at either residue reduce enzymatic activity drastically.[52,72]

Lys41

Several substitutions at residue 41 inactivate RNase A. Lys41 has been shown previously to be critical for ribonucleolytic activity, as it donates a hydrogen bond to the pentavalent transition state.[51,73] K41A RNase A has a k_{cat}/K_M value that is 10^4-fold lower than that of the wild-type enzyme.[74] Yet, Lys41 can be replaced by arginine with only a 10^2-fold decrease in catalytic activity.[51] Gratifyingly, a K41R variant was not isolated, but K41M, K41N, and K41E substitutions were, all of which likely cause a marked decrease in catalytic activity. Two other possible changes at position 41 were not isolated due to the mutational bias of error-prone PCR. Either the K41Q or K41T substitution would likely inactivate RNase A. The relevant codons for glutamine and threonine are CAG and ACG, respectively, which could arise only from rare A/T→C/G mutations in the AAG codon of Lys41.

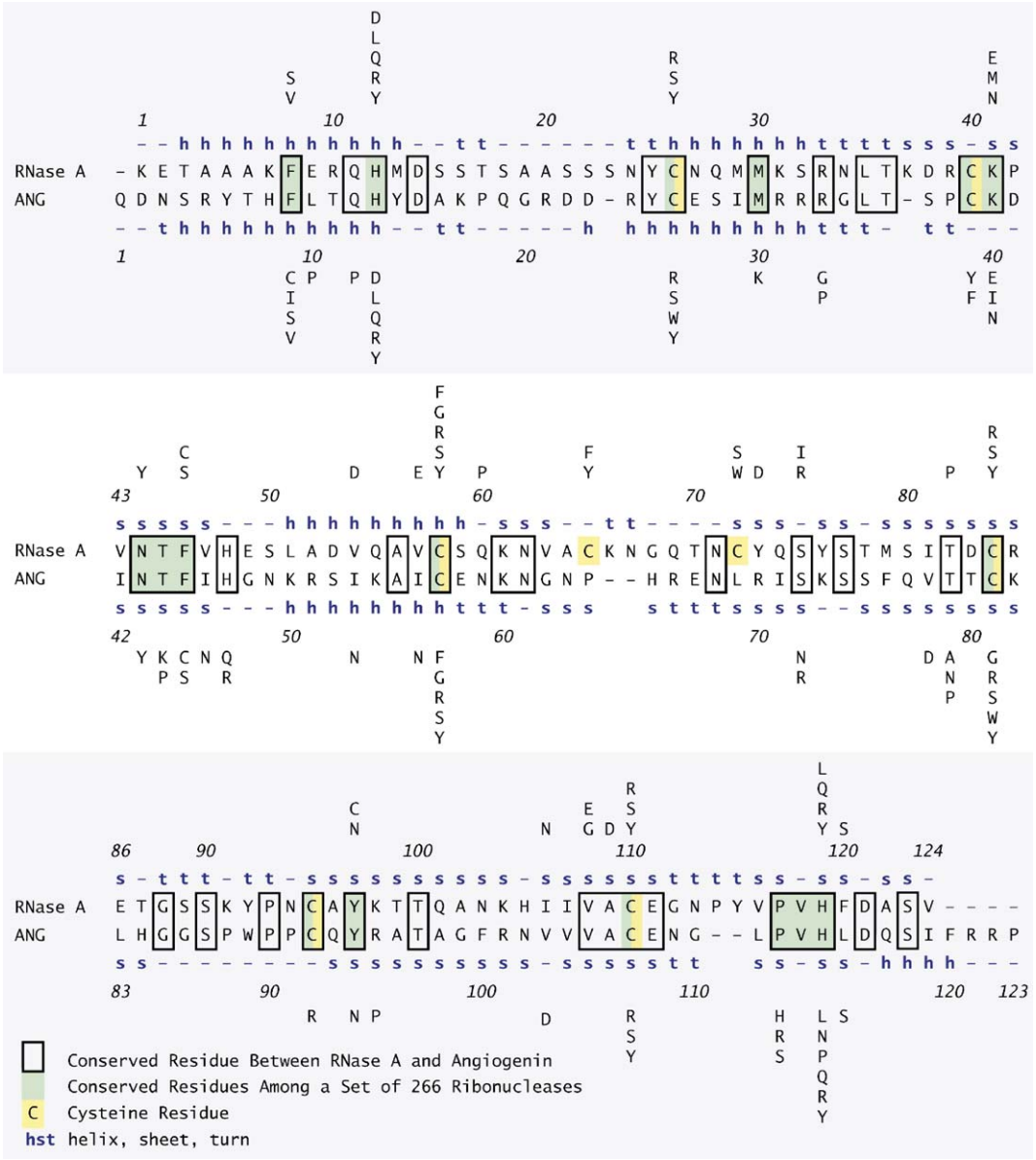

Figure 3. Amino acid sequence alignment of RNase A and angiogenin (ANG), along with substitutions found to inactivate RNase A (above) and angiogenin (below). The sequence alignment was generated with the program ClustalW.[133] Residues that are conserved between RNase A and angiogenin are boxed. Residues that are found in ≥90% of ribonuclease sequences are colored green. Cysteine residues are colored yellow. Helices, sheets, and turns are indicated by h, s, and t.

Group II residues (buried residues)

The various clusters in the hydrophobic core of RNase A have been well described,[22,23,27,75–78] though those residues that contribute most to structure/function are largely unknown. Our selection system has revealed the most important of these residues: 8, 44, 46, 54, 57, 73, 75, 97, 106, 108, 109, and 120. Many of these residues are in regions of the protein that have been shown to be important for folding and protein stability.[79–81] Typically, residues in a hydrophobic core tolerate substitution with other non-polar residues but not with polar residues.[4–6,8,10,11,82] Most substitutions in the hydrophobic core that were identified with our selection system have a hydrophobic residue replaced with a polar or charged residue. In several variants, however, large residues are replaced with much smaller residues. A hydrophobic core can contain a polar residue with a saturated hydrogen-bonding potential.[5] Substitutions to such residues were found to disrupt the structure/function of RNase A.

Phe8

The side-chain of Phe8 forms a cation–π interaction with an essential catalytic residue, His12,[83] and is in the major cluster in the hydrophobic core of the protein.[77] The F8W variant of RNase A, created for

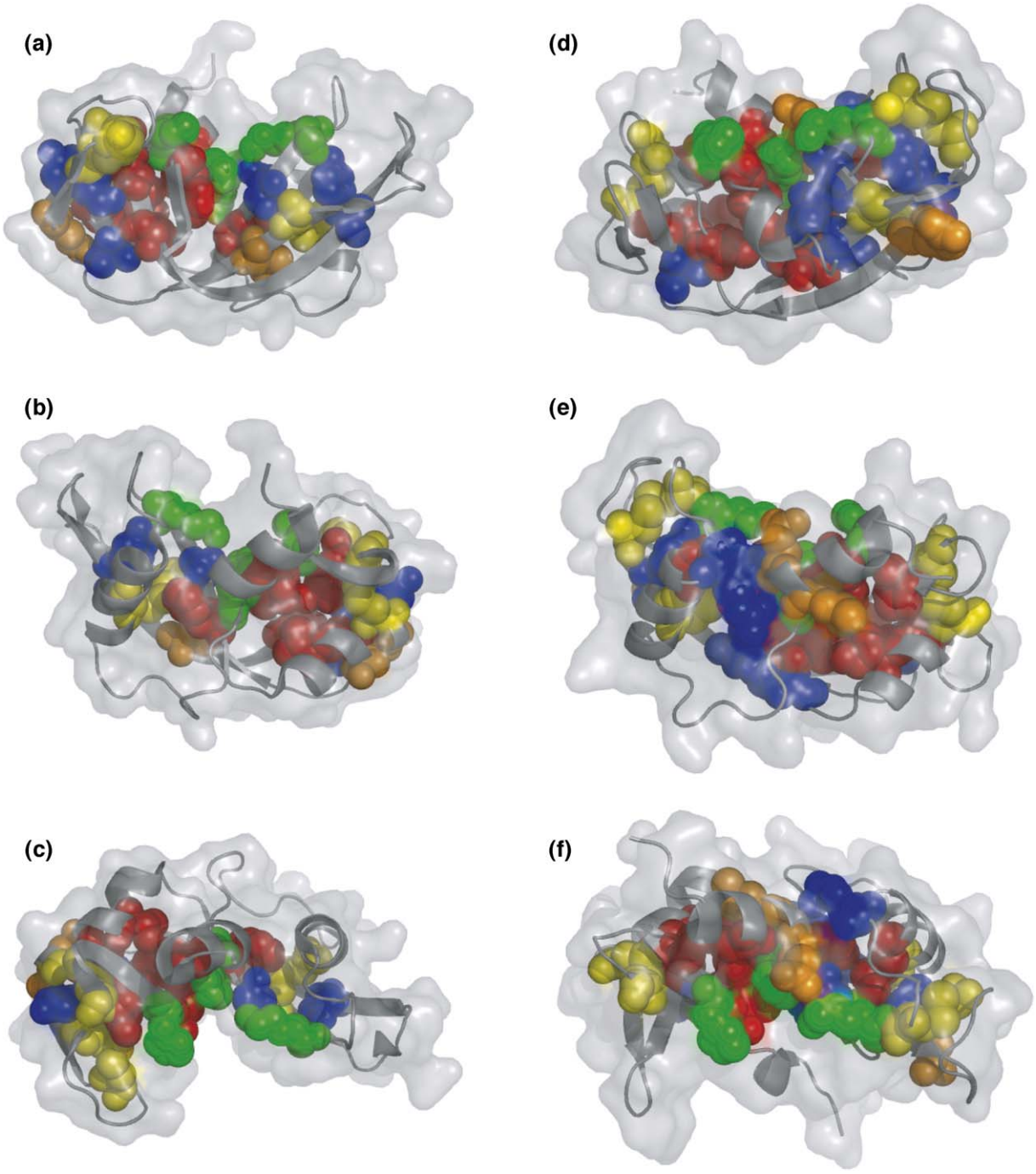

Figure 4. Essential amino acid residues in RNase A and angiogenin. Group I residues (residues involved in catalysis) are colored green. Polar residues in Group II (buried residues) are colored blue. Hydrophobic Group II residues are colored red. Group III residues (proline substitutions) are colored orange. Group IV residues (cysteine residues) are colored yellow. The surface is transparent gray, and the backbone is in cartoon representation. (Also see: Movie S1 in the Supplementary Material.) (a) Front view of RNase A (PDB: 7RSA). (b) Back view of RNase A. (c) Top view of RNase A. (d) Front view of angiogenin (PDB: 1B1I). (e) Back view of angiogenin. (f) Top view of angiogenin. Images were generated with the program PyMol (Delano Scientific, South San Francisco, CA).

the purposes of studying local and global unfolding in the enzyme, was reported to lack ribonucleolytic activity.[84] The F8S and F8V substitutions inactivate RNase A sufficiently to allow for growth in our selection system. Previous work on RNase S (which is a non-covalent complex between residues 1–20 (S-peptide) and residues 21–124 (S-protein)) showed that when Phe8 is replaced with an alanine residue in the S-peptide, the side-chain of essential residue Phe120 rotates to fill the nascent cavity.[85] Replacing Phe8 with serine or valine could also elicit this deleterious consequence.

Asn44

The side-chain of Asn44 accepts hydrogen bonds from two critical residues involved in catalysis, His12 and Lys41. In addition, the main-chain nitrogen and oxygen of Asn44 form hydrogen bonds with the main-chain oxygen and nitrogen of the essential residue Cys84 in an antiparallel β-sheet. Asn44 is buried in RNase A and conserved universally, but has not been reported to be essential for catalysis (barring unreliable results on 63-residue RNase A analogs[86]). Only the N44Y variant of RNase A was isolated with our selection system. As Asn44 is buried completely, it is possible that a large tyrosine residue at this position is particularly disruptive.

Phe46

Phe46 is a critical residue in the hydrophobic core of RNase A, linking two separate clusters of hydrophobic residues.[78] Phe46 is also a chain-folding initiation site (CFIS) residue, important for the rapid folding of the enzyme.[78,87] Replacing Phe46 with a glutamate or lysine residue has been shown to decrease greatly the conformational stability of the protein, whereas replacement of Phe46 with valine has less of an effect.[78,87] The F46Y variant of RNase A retains nearly full enzymatic activity, but has a T_m value that is 9 deg.C less than that of the wild-type protein.[88] The F46S and F46C substitutions were isolated with our selection system. Genes that encode isoleucine, leucine, and tyrosine at position 46 should be present in our library, but were not isolated, suggesting that these substitutions are less disruptive.

Val54 and Val57

Val54, Val57, and essential residue Cys58 form an (i, $i+3$, $i+4$) triplet in the structure of RNase A.[75] These valine residues have not been studied previously, though other workers have noted that they comprise part of the hydrophobic core of RNase A.[76–78] In addition, the region of RNase A encompassing residues 54 and 57 has been shown to be important for its folding.[81] Replacing Val54 or Val57 with an anionic residue inactivates RNase A sufficiently to allow for growth in our selection system.

Tyr73 and Ile108

The side-chain of Tyr73 is largely buried, but does extend to the surface of RNase A. Interestingly, Tyr73 is not well conserved. Among the subset of ribonucleases that contain the Cys65–Cys72 disulfide bond, position 73 is always histidine, tyrosine, or phenylalanine (in that order of frequency; Table S1 in the Supplementary Material). The main chains of residues 73 and 108 form two hydrogen bonds in an antiparallel β-sheet in RNase A, and could have undergone coupled evolution.[23] Previous work revealed that Val108 is an essential residue in the hydrophobic core of RNase A and that it serves as a CFIS residue.[89,90] A V108G substitution is known to decrease the conformational stability of RNase A substantially.[89–91] Substitutions at Tyr73 or Val108 enable isolation with our selection system.

Ser75 and Ile106

In the hydrophobic core of RNase A, the side-chain hydroxyl group of Ser75 forms a hydrogen bond with the main-chain nitrogen of essential residue Ile106. In addition, the main-chain nitrogen of Ser75 forms a hydrogen bond with the main-chain oxygen of Ile106. It had been shown that Ile106 is important in RNase A, as its replacement with alanine lowers the T_m value of the enzyme by 14 deg.C.[91] The importance of Ser75 had not been demonstrated, though other workers had suggested that Ser75 could be important for the conformational stability of RNase A.[22,23] Replacing Ser75 or Ile106 allows for growth in our selection system.

Tyr97

The side-chain of Tyr97 forms a hydrogen bond with the main-chain oxygen of an active-site residue, Lys41. Replacing Tyr97 with an alanine or glycine residue lowers the value of k_{cat}/K_M by ~500-fold, and (possibly more significantly) lowers the value of T_m by 34 deg.C.[92] Replacement of Tyr97 with phenylalanine lowers the value of T_m by 10 deg.C, but only lowers k_{cat}/K_M by threefold.[92] Y97N and Y97C variants of RNase A were isolated with our selection system. These smaller side-chains leave a void and cannot form a hydrogen bond with Lys41. Y97F RNase A was likely in our libraries but was not isolated with selection system, probably due to its minimal effects on enzymatic activity and conformational stability.

Ala109

Ala109 is in van der Waals contact with catalytic residue His119 and could be involved in imparting substrate specificity in the B_2 subsite of RNase A[93] (B. D. S., J. E. Lee & R. T. R., unpublished results). An A109R substitution in human pancreatic ribonuclease decreases catalytic activity by ~100-fold (R. J. Johnson and R. T. R., unpublished results). Ala109 is also a part of a CFIS, and an A109G variant of RNase A has slightly lower conformational stability.[91] Ala109 is well conserved among diverse ribonucleases, though frog ribonucleases have a threonine or lysine residue in this position (and have different base specificities in the B_2 subsite).[94] An A109D substitution in RNase A inactivates the enzyme sufficiently to allow for growth in Origami™ cells.

Phe120

The main-chain nitrogen of Phe120 forms a hydrogen bond with a non-bridging oxygen of the

pentavalent transition state. Previous work showed that replacement of Phe120 with a glutamate or alanine residue decreases catalytic activity by <50-fold and lowers the T_m value by 12–15 deg.C.[95] An F120S variant was isolated *via* genetic selection. The enzyme likely tolerates other changes to this residue, as genes encoding isoleucine, leucine, and tyrosine at position 120 should be present in the libraries but were not isolated with our selection system. Residue 120 is quite conserved in the RNase A superfamily. Angiogenin and other human homologs have leucine in this position; pronghorn and giraffe ribonuclease have tyrosine.[22] As with RNase A, the only substitution in angiogenin that was isolated at this position (Leu115) was to serine (*vide infra*).

Hydrophobic nuclei

Several hydrophobic nuclei (γ1, γ2, and γ3) and microclusters in RNase A have been identified.[76] In gratifying agreement, the residues that we identified as being essential in our selection system coincide with the most important hydrophobic nuclei. Essential residues Phe8, His12, Val54, Val57, Cys58, Tyr73, Ile106, Val108, Cys110, and Phe120 constitute a subset of residues located in the γ1 nucleus, which is well conserved. The γ2 nucleus is smaller and not well conserved. The only residues that were identified as being essential in that nucleus are Cys65 and Cys72, which form an important disulfide bond.[96] The γ3 nucleus is similar to the γ1 nucleus in that its residues are well conserved and have a high number of contacts (on average), and its side-chain hydrocarbon atoms have low temperature factors (on average).[76] Variants were isolated with our selection system at a subset of the residues involved in the γ3 nucleus: Cys26, Asn44, Phe46, Cys84, and Tyr97. Thus, our results not only agree well with previous work on the hydrophobic nuclei of RNase A, but also extend that work by identifying those residues that are likely to be the most important in each nucleus.

Group III residues (proline substitutions)

Replacing a residue with proline can be particularly deleterious to protein structure.[6,8,11] In RNase A, proline can be substituted at 47 residues by changing a single nucleotide (37 of which would require a rare error-prone PCR mutation). Group III substitutions were identified in RNase A at two of those residues: Gln60 and Thr82. Gln60 lies between an α-helix and a β-strand; its main-chain nitrogen forms a hydrogen bond with the main-chain oxygen of critical residue Val57. Thr82 lies in the middle of a β-strand; its main chain forms two hydrogen bonds to the main-chain atoms of essential residue Phe46.

Group IV residues (half-cystines)

Four cysteine residues that contribute to the two most important disulfide bonds in RNase A (Cys26–Cys84 and Cys58–Cys110)[26,96] appear to be wholly intolerant of substitution. *In vitro*, C26A/C84A RNase A (T_m = 27 °C) and C58A/C110A RNase A (T_m = 24 °C) have T_m values that are ~35 deg.C lower than that of the wild-type enzyme.[96] It is not surprising that substitutions at these residues were identified in this study, as genetic selection experiments were performed at 37 °C, a temperature at which both these enzymes would be substantially unfolded.

RNase A has two other disulfide bonds (Cys65–Cys72 and Cys40–Cys95). Previous studies have shown that replacing these cysteine residues with alanine reduces the T_m values of these proteins to ~40 deg.C, and lower catalytic activity by 40-fold.[96] Other work has shown that RNase A variants lacking either disulfide bond are stable at room temperature and remain catalytically active.[97] With our selection system, we isolated substitutions at the cysteine residues in the Cys65–Cys72 disulfide bond, but not the Cys40–Cys95 disulfide bond. This result is surprising from an evolutionary perspective, as the Cys40–Cys95 bond is absolutely conserved in the RNase A superfamily, whereas the Cys65–Cys72 bond is not. Yet, given that an RNase A variant lacking the Cys40–Cys95 disulfide bond has a T_m value greater than 37 deg.C, is catalytically active, and has only local disorder around the missing disulfide bond,[97,98] it is not surprising that substitutions were not isolated at residue 40 or 95. An RNase A variant lacking the Cys65–Cys72 disulfide bond also has a T_m value greater than 37 deg.C and is catalytically active,[97] thus it seems to be surprising that substitutions at these residues were isolated with our selection system. The Cys65–Cys72 disulfide is, however, particularly important in the oxidative folding pathway of RNase A,[99] which could account for its decreased tolerance to substitution. Still, the Cys65–Cys72 disulfide bond is somewhat intolerant to change, as only two substitutions per half-cystine residue were identified with our selection system. Substitutions that should be present in our libraries (such as Cys65 to serine or arginine or Cys72 to tyrosine or arginine) were not isolated with our selection system.

Implications for human health

Several homologs of RNase A have been found to be toxic to cancer cells.[24,31–35] The frog ribonuclease, Onconase®, is in Phase III clinical trials for the treatment of malignant mesothelioma.[100] Variants of RNase A and its human homolog can likewise be toxic to cancer cells upon their engineering to evade RI.[101,102] Such cytotoxic mammalian ribonucleases could have desirable attributes compared to an amphibian ribonuclease. In engineering new toxins based on RNase A, changes to the essential residues identified herein should be avoided, as the structure/function of a ribonuclease is essential for its cytotoxic activity.[103]

Angiogenin: essential residues

Whereas 23 of the 124 residues in RNase A were found to be intolerant to substitution with one or more amino acid, 29 of the 123 residues in angiogenin were likewise intolerant (Tables 3 and 4; Figures 3 and 4; Movie S1 in the Supplementary Material). A lower conformational stability could lead to a less stringent selection for angiogenin. To our knowledge, the conformational stability of angiogenin had not been measured previously. By monitoring thermal denaturation using circular dichroism spectroscopy, we find that the T_m value of angiogenin in PBS is 63.6(±0.6) °C. The T_m value for RNase A in PBS is 64(±2) °C.[104] Thus, the lower stringency with angiogenin does not result from its having less conformational stability.

Unlike its conformational stability, the ribonucleolytic activity of angiogenin is much lower than that of RNase A. Although the angiogenin gene was expressed at a higher level in our selection system to compensate for this difference (Figure 2), the enzymatic activity of angiogenin in Origami™ cells could still be lower than that of RNase A. This lower activity could give rise to a less stringent selection for angiogenin.

The set of substitutions that inactivate angiogenin largely intersect with that for RNase A. The discussion that follows focuses on the differences found in angiogenin.

Group I residues (catalytic residues)

The three residues critical for catalysis by RNase A (His12, His119, and Lys41) are likewise essential for the ribonucleolytic activity of angiogenin (His13, His114, and Lys40).

Group II residues (buried residues)

γ-1 nucleus

The hydrophobic core of angiogenin is similar to that of RNase A, though angiogenin appears to rely more on residues in the γ-1 and γ-3 hydrophobic nuclei for conformational stability, and less on those in the γ-2 nucleus. This shift could arise because angiogenin lacks the loop containing the Cys65–Cys72 disulfide bond and other important hydrophobic residues in the γ-2 nucleus, and instead contains a loop necessary for its promotion of angiogenesis.[105,106] Substitutions in the γ-1 nucleus that can inactivate angiogenin occur at residues 9, 13, 44, 46, 53, 56, 57, 72, 78, 103, 112, and 115. Replacing Thr44, Ile46, Val78, or Pro112 (corresponding to Thr45, Val47, Ile81, and Pro117 in RNase A) inactivates angiogenin, but not RNase A.

Thr44

Substitution of Thr44 in angiogenin has been shown to decrease catalytic activity by up to 50-fold and alter substrate specificity.[107] This threonine residue is buried in both angiogenin and RNase A. The equivalent residue in RNase A, Thr45, imparts substrate specificity by forming hydrogen bonds with a nucleobase in the B_1 subsite of the enzyme.[108,109] Thr45 could also be involved in ground-state destabilization during catalysis.[110] No substitutions at residue 45 were isolated in RNase A with our selection system. In contrast, replacing Thr44 with lysine or proline inactivates angiogenin sufficiently to allow for growth in Origami™ cells.

Ile46 and Val78

Ile46 and Val78 are within 5 Å of each other. Val78 is in van der Waals contact with essential residue Phe45. Ile46 interacts with essential hydrophobic core residue Ile53. In addition, the main-chain nitrogen of Ile46 forms a hydrogen bond with the main-chain oxygen of active-site residue His13. Replacing Ile46 or Val78 inactivates angiogenin sufficiently to allow for growth in Origami™ cells.

Pro112

Replacement of Pro117 with alanine in RNase A destabilizes the enzyme,[111,112] though substitutions at this position in RNase A are not deleterious enough to be isolated with our selection system. Yet, replacement of the equivalent proline in angiogenin, Pro112, with histidine, serine, or arginine allows for growth in Origami™ cells. We did not isolate a P112L variant with our selection system, possibly indicating that this substitution (which arises from a common C/G→T/A mutation) is not sufficiently deleterious.

Substitutions that inactivate RNase A but not angiogenin

Substitution at Arg70, Val105, or Ala106 does not inactivate angiogenin, though substitution at a corresponding residue (Tyr73, Val108, or Ala109) inactivates RNase A. Tyr73 is found only in ribonucleases with the Cys65–Cys72 disulfide bond. Further, residues 73 and 108 in RNase A form two hydrogen bonds between their main-chain atoms and are coupled by evolution.[23] It is surprising that we did not isolate substitutions at the corresponding positions in angiogenin. Ala109 in RNase A could be part of the B_2 nucleobase-binding subsite, but angiogenin lacks some of the residues in that subsite.[113] Thus, Ala106 could be less important to angiogenin than Ala109 is to RNase A.

γ-3 nucleus. Substitutions in the γ-3 nucleus that inactivate angiogenin occur at residues 26, 30, 33, 40, 43, 45, 47, 79, 81, and 94. Replacing Met30, Arg33, or His47 inactivates angiogenin, though substitution at a corresponding residue (Met30, Arg33, or His48) does not inactivate RNase A.

Met30. Met30 is conserved in almost all ribonucleases in the RNase A superfamily. An M30L angiogenin variant is fully active, though the conformational stability of the enzyme is unknown.[114] Replacing Met23 (corresponding to Met30 in angiogenin and RNase A) with leucine in the frog ribonuclease, Onconase®, decreases conformational stability and increases sensitivity to proteolysis.[115] We isolated an M30K substitution in angiogenin. As residue 30 is in the hydrophobic core of angiogenin, a lysine there would likely be disruptive. Variants with other hydrophobic residues at position 30 were not isolated with our selection system.

Arg33. Replacing Arg33 in angiogenin with alanine has been shown to decrease catalytic activity by sevenfold.[116] Arg33 interacts with essential residue Phe45, which could account for this decrease in activity. Arg33 also shields Met30 and Cys26 from solvent.[75] Substituting a glycine or proline residue at residue 33 in angiogenin inactivates the enzyme sufficiently to allow for growth in Origami™ cells. Interestingly, glycine places the least limitation on main-chain torsion angles and proline places the most, suggesting that the guanidinopropyl group of Arg33 is critical.

His47. His47 is located at the bottom of a hinge in the enzyme. Studies on the equivalent residue in RNase A, His48, have found that this residue is involved in a conformational change of the enzyme during catalysis.[19,22,117] Substitutions at His47 of angiogenin (but not His48 of RNase A) enable isolation with our selection system.

Group III residues (proline substitutions)

In angiogenin, proline can be substituted at 44 residues by changing a single nucleotide (34 of which would require a rare error-prone PCR mutation). There are three examples in which proline is the only substitution at a particular residue that inactivates angiogenin: residues 10, 12, and 95. Residues 10 and 12 are in the first α-helix in angiogenin and lie between essential residues Phe9 and His13. Disruption of this α-helix with a proline residue could disrupt the important cation–π interaction between residues Phe9 and His13. Arg95 lies in a β-sheet. Replacing Arg95 with proline could disrupt the orientation of an adjacent essential residue, Tyr94.

Group IV residues (half-cystines)

Unlike RNase A, angiogenin has only three disulfide bonds, lacking Cys65–Cys72. All three of these disulfide bonds are essential, though the Cys39–Cys92 disulfide bond tolerates some amino acid changes. This result indicates that the Cys26–Cys81 and Cys57–Cys107 disulfide bonds are the two most important in angiogenin, as are the corresponding disulfide bonds in RNase A.

Although angiogenin lacks the Cys65–Cys72 disulfide that is found in most other ribonucleases, its conformational stability is equivalent to that of RNase A. Other ribonucleases, such as Onconase®, lack the Cys65–Cys72 disulfide bond, but have compensated for this loss, as Onconase® has a T_m value of ~90 deg.C.[115] This and other frog ribonucleases contain the three disulfide bonds of angiogenin, but have an additional disulfide bond that is distinct from any in RNase A and that contributes greatly to conformational stability.[118]

Implications for human health

The data listed in Tables 3 and 4 could be useful for the interpretation of genotypic information. For example, substitutions at Cys39 or Ile46 of human angiogenin have been linked with familial ALS.[44] Substitutions at these residues were likewise identified with our selection system, indicating that these residues are essential to the structure/function of angiogenin. Similarly, substitutions at Gln12, Lys17, Arg31, Cys39, and Ile46 have been linked with sporadic ALS,[44] and substitutions at Gln12, Cys39, and Ile46 were identified herein.

Surface accessibility *versus* essentiality

The accessible surface area (ASA) of a residue in RNase A or angiogenin is not sufficient to

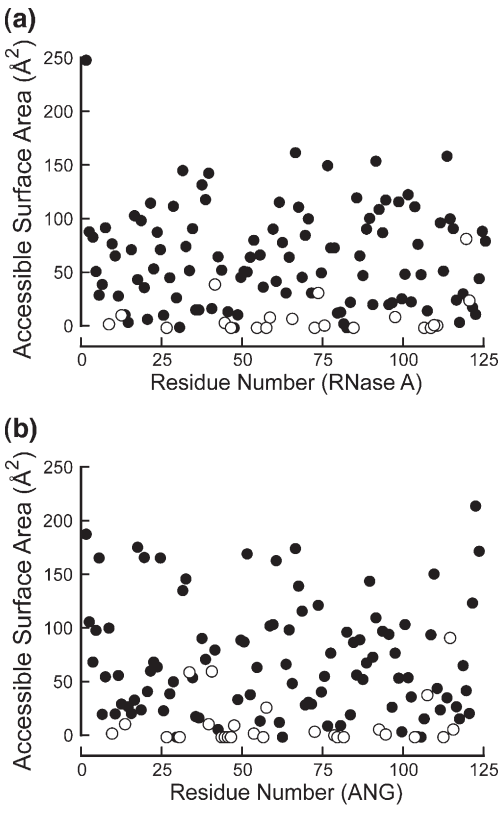

Figure 5. Accessible surface area of amino acid residues in RNase A (a) and angiogenin (b). Essential residues are indicated by open circles.

establish its importance. Most essential residues in RNase A and angiogenin are buried (Tables 3 and 4; Figure 5). A few residues that are completely buried and several residues that are substantially buried are not essential, and tolerate non-conservative changes.

Among the essential residues that are not completely buried are Lys41 and His119 (RNase A numbering), which are involved directly in catalysis. The side-chain of essential residue Tyr73 in RNase A is mostly buried but does extend to the surface. Phe120 is partially surface exposed in RNase A, as its main-chain nitrogen forms a hydrogen bond with the pentavalent phosphoryl group during catalysis. In angiogenin, Arg33 is accessible and serves to shield Cys26 and Met30 from solvent. The Cys57–Cys107 disulfide bond in angiogenin is somewhat exposed (more so than is the equivalent disulfide bond in RNase A). Whether this disulfide bond is more susceptible to reduction in angiogenin is unknown.

Sequence conservation *versus* essentiality

An alignment of 338 ribonuclease amino acid sequences was downloaded from the HOVERGEN database (gene family HBG008396).[119] Duplicate entries and fragments were removed, leaving a set of 266 sequences (Table S1 in the Supplementary Material). This set of diverse sequences includes pancreatic ribonucleases, angiogenins, eosinophil-associated ribonucleases, and other ribonucleases in the RNase A superfamily from many mammals, several amphibians, and a chicken. The molecular evolution of pancreatic ribonucleases has been well studied, and pancreatic ribonuclease amino acid sequences have been determined from many species,[22,23,77] thus sequence conservation numbers are biased towards pancreatic ribonucleases. The percentage of this set of sequences that contains the residue found in RNase A or angiogenin at each position is shown in Figure 6(a) and (b).

To alleviate bias problems in the alignment for the purposes of this study, a second alignment of angiogenin and pancreatic ribonuclease amino acid sequences from 13 selected species was generated (Table S2 in the Supplementary Material). The percentage of this set of sequences that contains the residue found in RNase A or angiogenin at each position is shown in Figure 6(c) and (d). Many residues are 50% conserved (100% conserved among pancreatic ribonucleases and/or 100% conserved among angiogenins). These residues could be important for the biological functions of each enzyme.

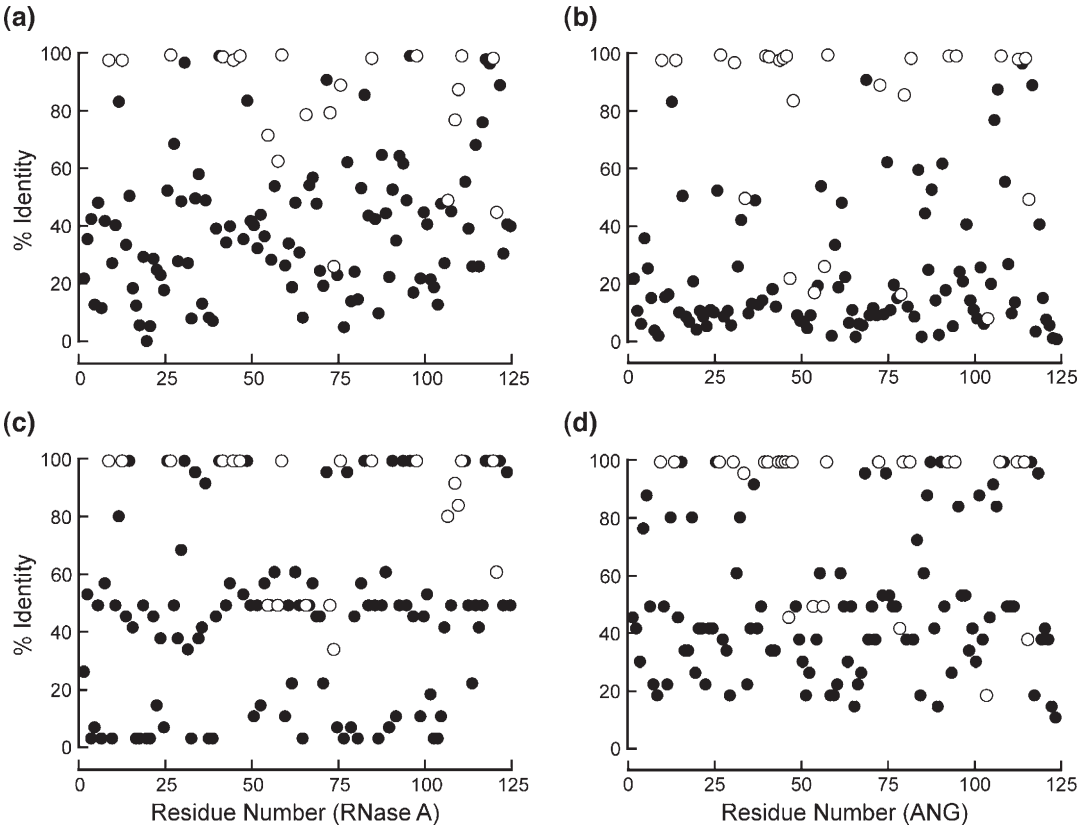

Figure 6. Sequence conservation of amino acid residues in RNase A and angiogenin. Percent identity per residue of RNase A (a) and angiogenin (b) to a set of 266 ribonuclease amino acid sequences. Percent identity per residue of RNase A (c) and angiogenin (d) to a set of pancreatic ribonuclease and angiogenin amino acid sequences from 13 species. Essential residues are indicated by open circles.

Most of the residues that are >95% conserved among the set of 266 ribonucleases are essential for ribonucleolytic activity (Figure 6). For RNase A, the exceptions are Met30, Cys40, Thr45, Cys95, and Val118. Several of these residues are important for RNase A structure/function, but substitutions to these residues might not be sufficiently deleterious to reach the threshold of our selection system. For angiogenin, the only residue with >95% sequence conservation that is not essential is Val113 (which corresponds to Val118 in RNase A). Several substitutions at this position should be present in our error-prone PCR libraries, including a non-conservative replacement with aspartate. Val118 has been replaced with a half-cystine in both RNase A and human pancreatic ribonuclease variants containing five disulfide bonds. The resulting enzymes retain >10% of wild-type catalytic activity and have increased conformational stability.[120,121] Val118 and Val113 are substantially buried in RNase A and angiogenin, respectively, making unlikely a role in a protein–protein interaction. Hence, the impetus for the nearly complete conservation of Val118 is unclear.

Many other well-conserved residues are not critical. These residues could be involved in a function other than ribonucleolytic activity. Alternatively, substitutions that decrease catalytic activity but not below the threshold of our selection system could fall prey to stricter selection *in vivo*. This explanation is especially plausible for Gln11, Asn71, and Asp121, each of which is highly conserved and known to be important for the productive binding of substrate, substrate specificity, and catalytic activity, respectively.[47,122–124]

Interestingly, many residues that are not well conserved were found to be essential (Tables 3 and 4; Figure 6). Most, but not all, of these are hydrophobic core residues that do tolerate replacement with other hydrophobic residues. Examples include residues Val54, Val57, Ile106, and Phe120 in RNase A (corresponding to Ile53, Ile56, Val103, and Leu115 in angiogenin). Exceptions include Tyr73 in RNase A and Arg33 in angiogenin.

Conclusions

We have identified the residues in RNase A and angiogenin that are essential for ribonucleolytic activity. After more than 180 million years of evolution,[125] the two enzymes still share most of the residues important for their structure/function. Yet, many significant changes have occurred. Although research on RNase A has been extensive,[16–28] we have identified several "new" residues that are critical for its structure/function. The biological activities of members of the RNase A superfamily, such as the role of angiogenin in the growth of new blood vessels, are now being elucidated. Moreover, new ribonucleases are being discovered, even in humans.[126] The work described herein could aid in deciphering the myriad functions of these fascinating enzymes.

Materials and Methods

Materials

Taq polymerase was from Promega (Cat. no. M166B; Madison, WI). Mutazyme® DNA polymerase was from Stratagene (La Jolla, CA). *Escherichia coli* Origami™ B (DE3) cells and DH5α cells were from Novagen (Madison, WI). Phosphate-buffered saline (PBS) contained (in 1.00 l) NaCl (8.0 g), KCl (0.2 g), $Na_2HPO_4 \cdot 7H_2O$ (2.16 g), and KH_2PO_4 (0.2 g), and had pH 7.4. All other commercial chemicals and biochemicals were of reagent grade or better, and were used without further purification.

Plasmids

Genes encoding RNase A and human angiogenin were inserted into pSH12,[127] which is a pGEX-4T3 based plasmid, at its NdeI and SalI restriction enzyme sites. The resulting plasmids were named pGEX-RNase A and pGEX-Ang. Both plasmids code for a methionine residue at the start of the gene, designated Met(–1).

Random mutation libraries

Due to the inherent codon bias of the techniques used for generating random mutations during the PCR, libraries were produced in two complementary ways, using Mn^{2+} with *Taq* DNA polymerase (A/T→B/V mutation bias) and Mutazyme® DNA polymerase (C/G→D/H mutation bias).

Error-prone PCR with *Taq* DNA polymerase

Error-prone PCR was performed by the methods described,[128] with the following modifications. Error-prone PCR reaction mixtures (50 µl) contained 1× Cloned *Pfu* Buffer (which was 20 mM Tris–HCl buffer (pH 8.8), containing $MgSO_4$ (2 mM), KCl (10 mM), $(NH_4)_2SO_4$ (10 mM), Triton® X-100 (0.1% w/v), and nuclease-free BSA (0.1 mg/ml); Stratagene, La Jolla, CA), dNTPs (200 µM), forward and reverse primers (200 nM each), template DNA (20 fmol), *Taq* DNA Polymerase (five units), and varying amounts of 1× Mutagenic PCR buffer (which contained dTTP (0.8 mM), dCTP (0.8 mM), $MgCl_2$ (4.8 mM), and $MnCl_2$ (0.5 mM)). The amount of Mutagenic PCR buffer added to the reaction mixture affected the rate of mutation. (For example, adding 12.5 µl of 1× Mutagenic PCR buffer to a 50 µl PCR mixture was found to generate approximately one mutation per 500 bp of DNA, for a mutation rate of 0.2%.) After the PCR was performed (25 cycles at 95 °C for 30 s, 54 °C for 60 s, and 72 °C for 60 s), *Taq* DNA polymerase was eliminated by adding proteinase K (50 µg/ml), EDTA (5 mM), and SDS (0.5% w/v) at pH 8.5, and incubating for 15 min at 65 °C. The PCR product was then purified with the Wizard SV Gel and PCR Clean-Up kit (Promega; Madison, WI).

Error-prone PCR with Mutazyme®

Mutazyme® DNA Polymerase was used to perform error-prone PCR according to the protocol of the

manufacturer (Stratagene; La Jolla, CA). Briefly, 50 μl PCR mixtures contained 1× Mutazyme® Reaction Buffer, dNTPs (800 μM), forward and reverse primers (200 nM each), target DNA (10 or 100 ng), and Mutazyme® DNA Polymerase (2.5 units). After the PCR was performed (30 cycles; 95 °C for 30 s, 54 °C for 30 s, 72 °C for 60 s), DpnI (ten units; Promega, Madison, WI) was added to digest template DNA, and the samples were incubated at 37 °C for 1–2 h. DNA Polymerase was eliminated by adding proteinase K (50 μg/ml), EDTA (5 mM), and SDS (0.5% w/v) at pH 8.5, and incubating for 15 min at 65 °C. The PCR product was then purified with the Wizard SV Gel and PCR Clean-Up kit (Promega; Madison, WI).

Ligation of mutagenized PCR product

Purified mutagenized PCR products were digested with NdeI and SalI (20 units each) at 37 °C for 2 h. Plasmid pGEX-RNase A (20 μg) was digested with NdeI and SalI (30 units each). (Unlike an empty expression vector, any uncut pGEX-RNase A will be toxic to Origami™ cells at later steps, and hence will not generate any colonies.) The restriction enzymes were inactivated by heating at 65 °C for 15 min. Digests were purified with the Wizard SV Gel and PCR Clean-Up kit (Promega; Madison, WI), and eluted with ddH_2O (25 μl). Additionally, cut pGEX-RNase A was purified using S-500 Sephacryl HR (Pharmacia, Piscataway, NJ) to remove the RNase A gene. S-500 Sephacryl (800 μl) was added to a microspin column. The column was subjected to centrifugation at 2000***g*** for 1 min. The tip of the column was dried to remove any excess buffer, and the column was then placed into a new microfuge tube. Cut plasmid DNA (20 μl) was added, and the column was subjected to centrifugation again at 2000***g*** for 1 min. The eluate was then run over another S-500 Sephacryl microspin tube as above to ensure that all of the excised fragment had been separated from the plasmid. The linear plasmid (4 μg) and mutated PCR product (10 μl) were added to a ligation reaction mixture (30 μl) containing 1×ligase buffer, DNA ligase (six units; Promega, Madison, WI), and additional ATP (1 mM). The ligation reaction mixtures were placed at 14 °C overnight. Sodium acetate (0.3 M) and three volumes of ice-cold ethanol were added to the reaction mixture, and the tube was placed at –20 °C for 1 h to precipitate the DNA. The ligated DNA was subjected to centrifugation at 10,000***g*** for 10 min. The ethanol was removed by aspiration, and 250 μl of an ice-cold aqueous solution of ethanol (70% v/v) was added. The sample was subjected to centrifugation again. The ethanol was removed by aspiration, and the pellet was allowed to dry. The dry pellet was dissolved in ddH_2O (10 μl) and desalted over an AutoSeq G50 column (GE Healthcare, Piscataway, NJ). Ligated DNA was transformed by electroporation into competent *E. coli* DH5α cells to analyze the quality and randomness of the mutagenic PCR libraries. Cells were electroporated (1.80 kV, 200 Ω, and 25 μF) with 1 μl of the desalted and purified ligated DNA. SOC (1.0 ml) was added immediately, and the cells were allowed to recover at 37 °C for 1 h before being plated on LB agar containing ampicillin (100 μg/ml). Ligated DNA was transformed into competent *E. coli* Origami™ cells for selection experiments. Origami™ cells were electroporated as above, but allowed to recover for 1.5 h before being plated on LB agar containing ampicillin (100 μg/ml), kanamycin (15 μg/ml), and tetracycline (12.5 μg/ml). Plates were placed at 37 °C for 2 d.

Library analysis

Colonies were grown in deep-well 96-well plates, and DNA was isolated with the Wizard SV 96 Plasmid DNA Purification System (Promega, Madison, WI). Plasmid DNA was eluted in ddH_2O (100 μl). DNA sequencing reactions were performed in 96-well plates. A DNA sequencing reaction mixture (5 μl) contained Big Dye 3.1 (0.5 μl), Big Buffer (0.75 μl), ddH_2O (0.75 μl), primer (0.5 μl from a 10 μM stock), and plasmid DNA (2.5 μl). Reaction mixtures were subjected to thermocycling (36 cycles; 96 °C for 20 s, 48 °C for 30 s, and 58 °C for 5 min). The sequencing reaction mixtures were purified with the CleanSEQ Dye-terminator Removal kit (Agencourt Bioscience, Beverly, MA). DNA sequences were obtained in the forward and reverse directions.

As an alternative to the expensive and time-consuming step of isolating plasmids, some DNA sequencing reactions were performed directly on crude lysates by using a heat-lysis technique[65] with the following modifications. Briefly, 1.0-ml cultures grown in LB medium for 24 h were subjected to centrifugation at 1500***g*** for 15 min. The supernatant was decanted, and cells were resuspended in ddH_2O (500 μl). The cells were spun again, and then washed again with ddH_2O (500 μl). After another centrifugation step, the supernatant was carefully decanted, as the pellets were loose. Cells were suspended in ddH_2O (125 μl) and transferred to a 96-well PCR plate. The plate was sealed, and cells were placed in boiling water for 5 min. The resulting crude lysate was subjected to centrifugation at 1000***g*** for 10 min. Sequencing reactions were performed as above, substituting 2.5 μl of crude lysate for purified plasmid DNA.

DNA sequences were determined with automated capillary-based instruments at the University of Wisconsin Biotechnology Center. Sequence data were analyzed with the Staden software package†, and its Mutation Scanner module was used to detect mutations.[129] Each sequence was then scanned manually for insertions–deletions or any artifacts that would lead to a false positive (e.g. a cloning artifact, star activity, etc.).

x-Factor

Due to the mutational bias of error-prone PCR, the protein and nucleotide *x*-factors were calculated for A/T→G/C, A/T→T/A, and C/G→T/A mutations (which occur at high frequencies in our libraries).

Accessible surface area calculations

The ASA of each residue was calculated for Protein Data Bank entries 7RSA[130] and 1B1I[131] with the module AREAIMOL in the program CCP4.[132]

Assay of conformational stability

The conformational stability of angiogenin was assessed by circular dichroism spectroscopy using an Aviv 202SF spectrometer (Aviv Associates, Lakewood, NJ) at the University of Wisconsin–Madison Biophysics Instrumentation Facility. A solution of angiogenin (20 μM) was heated from 25 to 80 °C in 3 deg.C increments, and the change in molar ellipticity at 215 nm was monitored after a

† http://Staden.sourceforge.net

5-min equilibration at each temperature. Spectra were fitted to a two-state model for denaturation to determine the value of T_m.

Acknowledgements

We are grateful to J. G. McCoy for assistance with ASA calculations. We thank D. R. McCaslin for assistance with circular dichroism spectroscopy, and for advice. We thank T.-Y. Chao for a sample of purified human angiogenin. We thank B. G. Miller, R. J. Johnson, J. E. Lee, and T. J. Rutkoski for contributive discussions. This work was supported by Grant CA73808 (NIH). B.D.S. was supported by Biotechnology Training Grant 08349 (NIH). The University of Wisconsin–Madison Biophysics Instrumentation Facility was established with grants BIR-9512577 (NSF) and RR13790 (NIH).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.07.020

References

1. Perutz, M. F., Kendrew, J. C. & Watson, H. C. (1965). Structure and function of haemoglobin II. Some relations between polypeptide chain configuration and amino acid sequence. *J. Mol. Biol.* **13**, 669–678.
2. Chothia, C. & Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5**, 823–826.
3. Loeb, D. D., Swanstrom, R., Everitt, L., Manchester, M., Stamper, S. E. & Hutchison, C. A., III (1989). Complete mutagenesis of the HIV-1 protease. *Nature*, **340**, 397–400.
4. Bowie, J. U. & Sauer, R. T. (1989). Identifying determinants of folding and activity for a protein of unknown structure. *Proc. Natl Acad. Sci. USA*, **86**, 2152–2156.
5. Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A. & Sauer, R. T. (1990). Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science*, **247**, 1306–1310.
6. Rennell, D., Bouvier, S. E., Hardy, L. W. & Poteete, A. R. (1991). Systematic mutation of bacteriophage T4 lysozyme. *J. Mol. Biol.* **222**, 67–88.
7. Poteete, A. R., Rennell, D. & Bouvier, S. E. (1992). Functional significance of conserved amino acid residues. *Proteins: Struct. Funct. Genet.* **13**, 38–40.
8. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S. & Miller, J. H. (1994). Genetic studies of the lac repressor. XIV. Analysis of 4000 altered *Escherichia coli lac* repressors reveals essential and non-essential residues, as well as "spacers" which do not require a specific sequence. *J. Mol. Biol.* **240**, 421–433.
9. Suckow, J., Markiewicz, P., Kleina, L. G., Miller, J., Kisters-Woike, B. & Muller-Hill, B. (1996). Genetic studies of the *lac* repressor. XV. 4000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure. *J. Mol. Biol.* **261**, 509–523.
10. Axe, D. D., Foster, N. W. & Fersht, A. R. (1996). Active barnase variants with completely random hydrophobic cores. *Proc. Natl Acad. Sci. USA*, **93**, 5590–5594.
11. Axe, D. D., Foster, N. W. & Fersht, A. R. (1998). A search for single substitutions that eliminate enzymatic function in a bacterial ribonuclease. *Biochemistry*, **37**, 7157–7166.
12. Guo, H. H., Choe, J. & Loeb, L. A. (2004). Protein tolerance to random amino acid change. *Proc. Natl Acad. Sci. USA*, **101**, 9205–9210.
13. Campbell-Valois, F. X., Tarassov, K. & Michnick, S. W. (2005). Massive sequence perturbation of a small protein. *Proc. Natl Acad. Sci. USA*, **102**, 14988–14993.
14. Sela, M., White, F. H., Jr & Anfinsen, C. B. (1957). Reductive cleavage of disulfide bridges in ribonuclease. *Science*, **125**, 691–692.
15. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, **181**, 223–230.
16. Barnard, E. A. (1969). Ribonucleases. *Annu. Rev. Biochem.* **38**, 677–732.
17. Richards, F. M. & Wyckoff, H. W. (1971). Bovine pancreatic ribonuclease. In *The Enzymes*, vol. 4, pp. 647–806.
18. Karpeisky, M. Y. & Yakovlev, G. I. (1981). Topochemical principles of the substrate specificity of nucleases. *Sov. Sci. Rev. sect. D*, **2**, 145–257.
19. Blackburn, P. & Moore, S. (1982). Pancreatic ribonuclease. *The Enzymes*, vol. 15, pp. 317–433.
20. Wlodawer, A. (1985). Structure of bovine pancreatic ribonuclease by X-ray and neutron diffraction. In *Biological Macromolecules and Assemblies, vol. II, Nucleic Acids and Interactive Proteins* (Jurnak, F. A. & McPherson, A., eds), pp. 333–375, Wiley, New York.
21. Eftink, M. R. & Biltonen, R. L. (1987). Pancreatic ribonuclease A: the most studied endoribonuclease. In *Hydrolytic Enzymes* (Neuberger, A. & Brocklehurst, K., eds), pp. 333–375, Elsevier, New York.
22. Beintema, J. J. (1987). Structure, properties and molecular evolution of pancreatic-type ribonucleases. *Life Chem. Rep.* **4**, 333–389.
23. Beintema, J. J., Schüller, C., Irie, M. & Carsana, A. (1988). Molecular evolution of the ribonuclease superfamily. *Prog. Biophys. Mol. Biol.* **51**, 165–192.
24. D'Alessio, G. & Riordan, J. F. (eds) (1997). *Ribonucleases: Structures and Functions*. Academic Press, New York.
25. Raines, R. T. (1998). Ribonuclease A. *Chem. Rev.* **98**, 1045–1065.
26. Wedemeyer, W. J., Welker, E., Narayan, M. & Scheraga, H. A. (2000). Disulfide bonds and protein folding. *Biochemistry*, **39**, 4207–4216.
27. Chatani, E. & Hayashi, R. (2001). Functional and structural roles of constituent amino acid residues of bovine pancreatic ribonuclease A. *J. Biosci. Bioeng.* **92**, 98–107.
28. Raines, R. T. (2004). Active site of ribonuclease A. In *Artificial Nucleases.* (Zenkova, M. A., ed.) pp. 19–32, Springer–Verlag, Heidelberg, Germany.
29. Beintema, J. J. & Kleineidam, R. G. (1998). The ribonuclease A superfamily: general discussion. *Cell. Mol. Life Sci.* **54**, 825–832.
30. Leonidas, D. D., Chavali, G. B., Oikonomakos, N. G., Chrysina, E. D., Kosmopoulou, M. N., Vlassi, M. *et al.* (2003). High-resolution crystal structures of

ribonuclease A complexed with adenylic and uridylic nucleotide inhbitors. Implications for structure-based design of ribonucleolytic inhibitors. *Protein Sci.* **12**, 2559–2574.
31. Schein, C. H. (1997). From housekeeper to microsurgeon: the diagnostic and therapeutic potential of ribonucleases. *Nature Biotechnol.* **15**, 529–536.
32. Leland, P. A. & Raines, R. T. (2001). Cancer chemotherapy—Ribonucleases to the rescue. *Chem. Biol.* **8**, 405–425.
33. Matoušek, J. (2001). Ribonucleases and their antitumor activity. *Comp. Biochem. Physiol.* **129C**, 175–191.
34. Loverix, S. & Steyaert, J. (2003). Ribonucleases: from prototypes to therapeutic targets? *Curr. Med. Chem.* **10**, 779–785.
35. Makarov, A. A. & Ilinskaya, O. N. (2003). Cytotoxic ribonucleases: molecular weapons and their targets. *FEBS Letters*, **540**, 15–20.
36. Strydom, D. J., Fett, J. W., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F. & Vallee, B. L. (1985). Amino acid sequence of human tumor derived angiogenin. *Biochemistry*, **24**, 5486–5494.
37. Riordan, J. F. (1997). Structure and function of angiogenin. In *Ribonucleases: Structures and Functions* (D'Alessio, G. & Riordan, J. F., eds), pp. 445–489, Academic Press, New York.
38. Strydom, D. J. (1998). The angiogenins. *Cell. Mol. Life Sci.* **54**, 811–824.
39. Badet, J. (1999). Angiogenin, a potent mediator of angiogenesis. Biological, biochemical and structural properties. *Pathol. Biol.* **47**, 345–351.
40. Shapiro, R., Fox, E. A. & Riordan, J. F. (1989). Role of lysines in human angiogenin: chemical modification and site-directed mutagenesis. *Biochemistry*, **28**, 1726–1732.
41. Shapiro, R. & Vallee, B. L. (1989). Site-directed mutagenesis of histidine-13 and histidine-114 of human angiogenin. Alanine derivatives inhibit angiogenin-induced angiogenesis. *Biochemistry*, **28**, 7401–7408.
42. Xu, Z. P., Tsuji, T., Riordan, J. F. & Hu, G. F. (2003). Identification and characterization of an angiogenin-binding DNA sequence that stimulates luciferase reporter gene expression. *Biochemistry*, **42**, 121–128.
43. Hooper, L. V., Stappenbeck, T. S., Hong, C. V. & Gordon, J. I. (2003). Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nature Immunol.* **4**, 269–273.
44. Greenway, M. J., Andersen, P. M., Russ, C., Ennis, S., Cashman, S., Donaghy, C. *et al.* (2006). ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nature Genet.* **38**, 411–413.
45. Woycechowsky, K. J. & Hilvert, D. (2004). Deciphering enzymes. Genetic selection as a probe of structure and mechanism. *Eur. J. Biochem.* **271**, 1630–1637.
46. Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucl. Acids Res.* **16**, 6127–6145.
47. delCardayré, S. B., Ribó, M., Yokel, E. M., Quirk, D. J., Rutter, W. J. & Raines, R. T. (1995). Engineering ribonuclease A: production, purification, and characterization of wild-type enzyme and mutants at Gln11. *Protein Eng.* **8**, 261–273.
48. Derman, A. I., Prinz, W. A., Belin, D. & Beckwith, J. (1993). Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*. *Science*, **262**, 1744–1747.
49. Prinz, W. A., Åslund, F., Holmgren, A. & Beckwith, J. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J. Biol. Chem.* **272**, 15661–15667.
50. Bessette, P. H., Aslund, F., Beckwith, J. & Georgiou, G. (1999). Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc. Natl Acad. Sci. USA*, **96**, 13703–13708.
51. Messmore, J. M., Fuchs, D. N. & Raines, R. T. (1995). Ribonuclease A: revealing structure–function relationships with semisynthesis. *J. Am. Chem. Soc.* **117**, 8057–8060.
52. Thompson, J. E. & Raines, R. T. (1994). Value of general acid-base catalysis to ribonuclease A. *J. Am. Chem. Soc.* **116**, 5467–5468.
53. Leland, P. A., Staniszewski, K. E., Park, C., Kelemen, B. R. & Raines, R. T. (2002). The ribonucleolytic activity of angiogenin. *Biochemistry*, **41**, 1343–1350.
54. St. Clair, D. K., Rybak, S. M., Riordan, J. F. & Vallee, B. L. (1987). Angiogenin abolishes cell-free protein synthesis by specific ribonucleolytic inactivation of ribosomes. *Proc. Natl Acad. Sci. USA*, **84**, 8330–8334.
55. St. Clair, D. K., Rybak, S. M., Riordan, J. F. & Vallee, B. L. (1988). Angiogenin abolishes cell-free protein synthesis by specific ribonucleolytic inactivation of 40S ribosomes. *Biochemistry*, **27**, 7263–7268.
56. Shapiro, R., Riordan, J. F. & Vallee, B. L. (1986). Characteristic ribonucleolytic activity of human angiogenin. *Biochemistry*, **25**, 3527–3532.
57. St Clair, D. K., Rybak, S. M., Riordan, J. F. & Vallee, B. L. (1987). Angiogenin abolishes cell-free protein synthesis by specific ribonucleolytic inactivation of ribosomes. *Proc. Natl Acad. Sci. USA*, **84**, 8330–8334.
58. Schimmel, P. (1990). Hazards and their exploitation in the applications of molecular biology to structure–function relationships. *Biochemistry*, **29**, 9495–9502.
59. Fujii, S., Akiyama, M., Aoki, K., Sugaya, Y., Higuchi, K., Hiraoka, M. *et al.* (1999). DNA replication errors produced by the replicative apparatus of *Escherichia coli*. *J. Mol. Biol.* **289**, 835–850.
60. Hermes, J. D., Blacklow, S. C. & Knowles, J. R. (1990). Searching sequence space by definably random mutagenesis: improving the catalytic potency of an enzyme. *Proc. Natl Acad. Sci. USA*, **87**, 696–700.
61. Rowe, L. A., Geddie, M. L., Alexander, O. B. & Matsumura, I. (2003). A comparison of directed evolution approaches using the β-glucuronidase model system. *J. Mol. Biol.* **332**, 851–860.
62. Neylon, C. (2004). Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucl. Acids Res.* **32**, 1448–1459.
63. Vanhercke, T., Ampe, C., Tirry, L. & Denolf, P. (2005). Reducing mutational bias in random protein libraries. *Anal. Biochem.* **339**, 9–14.
64. Cline, J. & Hogrefe, H. (2000). Randomize gene sequences with new PCR mutagenesis kit. *Strategies Newsletter*, **13**, 157–162.
65. Ganguly, T., Chen, P., Teetsel, R., Zhang, L. P., Papaioannou, E. & Cianciarulo, J. (2005). High-throughput sequencing of high copy number plasmids from bacterial cultures by heat lysis. *Biotechniques*, **39**, 304–308.
66. Dickson, K. A., Haigis, M. C. & Raines, R. T. (2005). Ribonuclease inhibitor: structure and function. *Prog. Nucl. Acid Res. Mol. Biol.* **80**, 349–374.
67. Acharya, K. R., Shapiro, R., Allen, S. C., Riordan, J. F. & Vallee, B. L. (1994). Crystal structure of human

angiogenin reveals the structural basis for its functional divergence from ribonuclease. *Proc. Natl Acad. Sci. USA*, **91**, 2915–2919.
68. Findlay, D., Herries, D. G., Mathias, A. P., Rabin, B. R. & Ross, C. A. (1961). The active site and mechanism of action of bovine pancreatic ribonuclease. *Nature*, **190**, 781–784.
69. Findlay, D., Mathias, A. P. & Rabin, B. R. (1962). The active site and mechanism of action of bovine pancreatic ribonuclease. 7. The catalytic mechanism. *Biochem. J.* **85**, 139–144.
70. Crestfield, A. M., Stein, W. H. & Moore, S. (1963). Alkylation and identification of the histidine residues at the active site of ribonuclease. *J. Biol. Chem.* **238**, 2413–2419.
71. Cuchillo, C. M., Vilanova, M. & Nogués, M. V. (1997). Pancreatic ribonucleases. In *Ribonucleases: Structures and Functions* (D'Alessio, G. & Riordan, J. F., eds), pp. 271–304, Academic Press, New York.
72. Park, C., Schultz, L. W. & Raines, R. T. (2001). Contribution of the active site histidine residues of ribonuclease A to nucleic acid binding. *Biochemistry*, **40**, 4949–4956.
73. Messmore, J. M. & Raines, R. T. (2000). Pentavalent organo-vanadates as transition state analogues for phosphoryl transfer reactions. *J. Am. Chem. Soc.* **122**, 9911–9916.
74. Thompson, J. E., Kutateladze, T. G., Schuster, M. C., Venegas, F. D., Messmore, J. M. & Raines, R. T. (1995). Limits to catalysis by ribonuclease A. *Bioorg. Chem.* **23**, 471–481.
75. Lenstra, J. A., Hofsteenge, J. & Beintema, J. J. (1977). Invariant features of the structure of pancreatic ribonuclease. A test of different predictive models. *J. Mol. Biol.* **109**, 185–193.
76. Kolbanovskaya, E. Y., Sathyanarayana, B. K., Wlodawer, A. & Karpeisky, M. Y. (1992). Intramolecular interactions in pancreatic ribonucleases. *Protein Sci.* **1**, 1050–1060.
77. Kolbanovskaya, E. Y., Morozov, N. Y., Gavryushov, S. A., Ilin, V. A., Beintema, J. J., Wlodawer, A. & Karpeisky, M. Y. (1993). Structural–functional analysis of ribonuclease A and related proteins. *Mol. Biol.* **27**, 821–836.
78. Chatani, E., Nonomura, K., Hayashi, R., Balny, C. & Lange, R. (2002). Comparison of heat- and pressure-induced unfolding of ribonuclease A: the critical role of Phe46 which appears to belong to a new hydrophobic chain-folding initiation site. *Biochemistry*, **41**, 4567–4574.
79. Matheson, R. R. & Scheraga, H. A. (1978). Method for predicting nucleation sites for protein folding based on hydrophobic contacts. *Macromolecules*, **11**, 819–829.
80. Matheson, R. R. & Scheraga, H. A. (1979). Steps in the pathway of the thermal unfolding of ribonuclease A. A nonspecific photo-chemical surface-labeling study. *Biochemistry*, **18**, 2437–2445.
81. Montelione, G. T. & Scheraga, H. A. (1989). Formation of local structures in protein folding. *Acc. Chem. Res.* **22**, 70–76.
82. Balaji, S., Aruna, S. & Srinivasan, N. (2003). Tolerance to the substitution of buried apolar residues by charged residues in the homologous protein structures. *Proteins: Struct. Funct. Genet.* **53**, 783–791.
83. Shoemaker, K. R., Fairman, R., Schultz, D. A., Robertson, A. D., York, E. J., Stewart, J. M. & Baldwin, R. L. (1990). Side-chain interactions in the C-peptide helix: Phe 8···His 12^{+}. *Biopolymers*, **29**, 1–11.
84. Navon, A., Ittah, V., Laity, J. H., Scheraga, H. A., Haas, E. & Gussakovsky, E. E. (2001). Local and long-range interactions in the thermal unfolding transition of bovine pancreatic ribonuclease A. *Biochemistry*, **40**, 93–104.
85. Ratnaparkhi, G. S. & Varadarajan, R. (2000). Thermodynamic and structural studies of cavity formation in proteins suggest that loss of packing interactions rather than the hydrophobic effect dominates the observed energetics. *Biochemistry*, **39**, 12365–12374.
86. Gutte, B. (1978). Synthetic 63-residue RNase A analogs. Simultaneous exchange of asparagine 44 by leucine and of threonine 45 by valine. *J. Biol. Chem.* **253**, 3837–3842.
87. Kadonosono, T., Chatani, E., Hayashi, R., Moriyama, H. & Ueki, T. (2003). Minimization of cavity size ensures protein stability and folding: structures of Phe46-replaced bovine pancreatic RNase A. *Biochemistry*, **42**, 10651–10658.
88. Markert, Y., Köditz, J., Vriend, G., Mansfeld, J., Arnold, U. & Ulbrich-Hofmann, R. (2001). Increased proteolytic resistance of ribonuclease A by protein engineering. *Protein Eng.* **14**, 791–796.
89. Torrent, J., Connelly, J. P., Coll, M. G., Ribo, M., Lange, R. & Vilanova, M. (1999). Pressure versus heat-induced unfolding of ribonuclease A: the case of hydrophobic interactions within a chain-folding initiation site. *Biochemistry*, **38**, 15952–15961.
90. Torrent, J., Rubens, P., Ribo, M., Heremans, K. & Vilanova, M. (2001). Pressure versus temperature unfolding of ribonuclease A: an FTIR spectroscopic characterization of 10 variants at the carboxy-terminal site. *Protein Sci.* **10**, 725–734.
91. Coll, M. G., Protasevich, I. I., Torrent, J., Ribo, M., Lobachov, V. M., Makarov, A. A. & Vilanova, M. (1999). Valine 108, a chain-folding initiation site-belonging residue, crucial for the ribonuclease A stability. *Biochem. Biophys. Res. Commun.* **265**, 356–360.
92. Eberhardt, E. S., Wittmayer, P. K., Templer, B. M. & Raines, R. T. (1996). Contribution of a tyrosine side chain to ribonuclease A catalysis and stability. *Protein Sci.* **5**, 1697–1703.
93. Holloway, D. E., Chavali, G. B., Hares, M. C., Subramanian, V. & Acharya, K. R. (2005). Structure of murine angiogenin: features of the substrate- and cell-binding regions and prospects for inhibitor-binding studies. *Acta Crystallog. sect. D*, **61**, 1568–1578.
94. Leu, Y. J., Chern, S. S., Wang, S. C., Hsiao, Y. Y., Amiraslanov, I., Liaw, Y. C. & Liao, Y. D. (2003). Residues involved in the catalysis, base specificity, and cytotoxicity of ribonuclease from Rana catesbeiana based upon mutagenesis and X-ray crystallography. *J. Biol. Chem.* **278**, 7300–7309.
95. Tanimizu, N., Ueno, H. & Hayashi, R. (1998). Role of Phe120 in the activity and structure of bovine pancreatic ribonuclease A. *J. Biochem. (Tokyo)*, **124**, 410–416.
96. Klink, T. A., Woycechowsky, K. J., Taylor, K. M. & Raines, R. T. (2000). Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *Eur. J. Biochem.* **267**, 566–572.
97. Laity, J. H., Shimotakahara, S. & Scheraga, H. A. (1993). Expression of wild-type and mutant bovine pancreatic ribonuclease A in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **90**, 615–619.

98. Pearson, M. A., Karplus, P. A., Dodge, R. W., Laity, J. H. & Scheraga, H. A. (1998). Crystal structures of two mutants that have implications for the folding of bovine pancreatic ribonuclease A. *Protein Sci.* **7**, 1255–1258.
99. Shin, H. C., Narayan, M., Song, M. C. & Scheraga, H. A. (2003). Role of the [65-72] disulfide bond in oxidative folding of bovine pancreatic ribonuclease A. *Biochemistry*, **42**, 11514–11519.
100. Mikulski, S. M., Costanzi, J. J., Vogelzang, N. J., McCachren, S., Taub, R. N., Chun, H. *et al.* (2002). Phase II trial of a single weekly intravenous dose of ranpirnase in patients with unresectable malignant mesothelioma. *J. Clin. Oncol.* **20**, 274–281.
101. Leland, P. A., Schultz, L. W., Kim, B.-M. & Raines, R. T. (1998). Ribonuclease A variants with potent cytotoxic activity. *Proc. Natl Acad. Sci. USA*, **98**, 10407–10412.
102. Leland, P. A., Staniszewski, K. E., Kim, B.-M. & Raines, R. T. (2001). Endowing human pancreatic ribonuclease with toxicity for cancer cells. *J. Biol. Chem.* **276**, 43095–43102.
103. Kim, J.-S., Souèek, J., Matoušek, J. & Raines, R. T. (1995). Catalytic activity of bovine seminal ribonuclease is essential for its immunosuppressive and other biological activities. *Biochem. J.* **308**, 547–550.
104. Rutkoski, T. J., Kurten, E. L., Mitchell, J. C. & Raines, R. T. (2005). Disruption of shape-complementarity markers to create cytotoxic variants of ribonuclease A. *J. Mol. Biol.* **354**, 41–54.
105. Harper, J. W. & Vallee, B. L. (1989). A covalent angiogenin/ribonuclease hybrid with a fourth disulfide bond generated by regional mutagenesis. *Biochemistry*, **28**, 1875–1884.
106. Raines, R. T., Toscano, M. P., Nierengarten, D. M., Ha, J. H. & Auerbach, R. (1995). Replacing a surface loop endows ribonuclease A with angiogenic activity. *J. Biol. Chem.* **270**, 17180–17184.
107. Curran, T. P., Shapiro, R. & Riordan, J. F. (1993). Alteration of the enzymatic specificity of human angiogenin by site-directed mutagenesis. *Biochemistry*, **32**, 2307–2313.
108. delCardayré, S. B. & Raines, R. T. (1994). Structural determinants of enzymatic processivity. *Biochemistry*, **33**, 6031–6037.
109. delCardayré, S. B. & Raines, R. T. (1995). A residue to residue hydrogen bond mediates the nucleotide specificity of ribonuclease A. *J. Mol. Biol.* **252**, 328–336.
110. Kelemen, B. R., Schultz, L. W., Sweeney, R. Y. & Raines, R. T. (2000). Excavating an active site: the nucleobase specificity of ribonuclease A. *Biochemistry*, **39**, 14487–14494.
111. Dodge, R. W. & Scheraga, H. A. (1996). Folding and unfolding kinetics of the proline-to-alanine mutants of bovine pancreatic ribonuclease A. *Biochemistry*, **35**, 1548–1559.
112. Juminaga, D., Wedemeyer, W. J. & Scheraga, H. A. (1998). Proline isomerization in bovine pancreatic ribonuclease A. 1. Unfolding conditions. *Biochemistry*, **37**, 11614–11620.
113. Russo, N., Acharya, K. R., Vallee, B. L. & Shapiro, R. (1996). A combined kinetic and modeling study of the catalytic center subsites of human angiogenin. *Proc. Natl Acad. Sci. USA*, **93**, 804–808.
114. Shapiro, R., Harper, J. W., Fox, E. A., Jansen, H. W., Hein, F. & Uhlmann, E. (1988). Expression of Met-(–1) angiogenin in *Escherichia coli*: conversion to the authentic <Glu-1 protein. *Anal. Biochem.* **175**, 450–461.
115. Notomista, E., Catanzano, F., Graziano, G., Dal Piaz, F., Barone, G., D'Alessio, G. & Di Donato, A. (2000). Onconase: an unusually stable protein. *Biochemistry*, **39**, 8711–8718.
116. Shapiro, R. & Vallee, B. L. (1992). Identification of functional arginines in human angiogenin by site-directed mutagenesis. *Biochemistry*, **31**, 12477–12485.
117. Hammes, G. G. (2002). Multiple conformational changes in enzyme catalysis. *Biochemistry*, **41**, 8221–8228.
118. Leland, P. A., Staniszewski, K. E., Kim, B. & Raines, R. T. (2000). A synapomorphic disulfide bond is critical for the conformational stability and cytotoxicity of an amphibian ribonuclease. *FEBS Letters*, **477**, 203–207.
119. Duret, L., Mouchiroud, D. & Gouy, M. (1994). HOVERGEN: a database of homologous vertebrate genes. *Nucl. Acids Res.* **22**, 2360–2365.
120. Futami, J., Tada, H., Seno, M., Ishikami, S. & Yamada, H. (2000). Stabilization of human RNase 1 by introduction of a disulfide bond between residues 4 and 118. *J. Biochem. (Tokyo)*, **128**, 245–250.
121. Klink, T. A. & Raines, R. T. (2000). Conformational stability is a determinant of ribonuclease A cytotoxicity. *J. Biol. Chem.* **275**, 17463–17467.
122. Tarragona-Fiol, A., Eggelte, H. J., Harbron, S., Sanchez, E., Taylorson, C. J., Ward, J. M. & Rabin, B. R. (1993). Identification by site-directed mutagenesis of amino acids in the B2 subsite of bovine pancreatic ribonuclease A. *Protein Eng.* **6**, 901–906.
123. Fontecilla-Camps, J. C., de Llorens, R., le Du, M. H. & Cuchillo, C. M. (1994). Crystal structure of ribonuclease A•d(ApTpApApG) complex. *J. Biol. Chem.* **269**, 21526–21531.
124. Schultz, L. W., Quirk, D. J. & Raines, R. T. (1998). His···Asp catalytic dyad of ribonuclease A: structure and function of the wild-type, D121N, and D121A enzymes. *Biochemistry*, **37**, 8886–8898.
125. Cho, S. & Zhang, J. (2006). Ancient expansion of the ribonuclease A superfamily revealed by genomic analysis of placental and marsupial mammals. *Gene*, **373**, 116–125.
126. Cho, S., Beintema, J. J. & Zhang, J. (2005). The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. *Genomics*, **85**, 208–220.
127. Park, S. H. & Raines, R. T. (2000). Genetic selection for dissociative inhibitors of designated protein-protein interactions. *Nature Biotechnol.* **18**, 847–851.
128. Matsumura, I. & Ellington, A. D. (2002). Mutagenic polymerase chain reaction of protein-coding genes for *in vitro* evolution. *Methods Mol. Biol.* **182**, 259–267.
129. Staden, R., Beal, K. F. & Bonfield, J. K. (2000). The Staden package, 1998. *Methods Mol. Biol.* **132**, 115–130.
130. Wlodawer, A., Anders, L. A., Sjölin, L. & Gilliland, G. L. (1988). Structure of phosphate-free ribonuclease A refined at 1.26 Å. *Biochemistry*, **27**, 2705–2717.
131. Leonidas, D. D., Shapiro, R., Allen, S. C., Subbarao, G. V., Veluraja, K. & Acharya, K. R. (1999). Refined crystal structures of native human angiogenin and two active site variants: implications for the unique functional properties of an enzyme involved in

neovascularisation during tumour growth. *J. Mol. Biol.* **285**, 1209–1233.
132. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
133. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalities and weight matrix choice. *Nucl. Acids Res.* **22**, 4673–4680.

Edited by F. Schmid

(Received 9 May 2006; received in revised form 3 July 2006; accepted 11 July 2006)
Available online 15 July 2006