LOCAL AND LONG-RANGE ELECTROSTATIC INTERACTIONS IN DENATURED RIBONUCLEASE SA: COMPARING THE DENATURED PROTEIN WITH MODEL PEPTIDES

A Senior Honors Thesis

by

MICHAEL DOUGLAS DAILY

Submitted to the Office of Honors Programs
& Academic Scholarships
Texas A&M University
In partial fulfillment of the requirements of

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOWS

April 2002

Group: Life Sciences 1

LOCAL AND LONG-RANGE ELECTROSTATIC INTERACTIONS IN DENATURED RIBONUCLEASE SA: COMPARING THE DENATURED PROTEIN WITH MODEL PEPTIDES

A Senior Honors Thesis

by

MICHAEL DOUGLAS DAILY

Submitted to the Office of Honors Programs
& Academic Scholarships
Texas A&M University
In partial fulfillment for the designation of

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOW

Approved as to style and content by:

J. Martin Scholtz
(Fellows Advisor)

Edward A. Funkhouser
(Executive Director)

April 2002

Group: Life Sciences 1

ABSTRACT

Local and Long-Range Electrostatic

Interactions in Denatured Ribonuclease Sa: Comparing

the Denatured Protein with Model Peptides (April 2002)

Michael Douglas Daily
Department of Biochemistry and Biophysics
Texas A&M University

Fellows Advisor: Dr. J. Martin Scholtz
Department of Biochemistry and Biophysics

The folding of proteins into three-dimensional, biologically active conformations demands much study because protein misfolding causes many genetic diseases and because protein folding is an integral step in translating the code of DNA into the chemistry of life. While folded states of proteins are well studied, their unfolded counterparts have received little attention, partly because for many years, biochemists have treated unfolded states as unstructured. However, multiple recent studies on pH dependence of protein stability have highlighted substantial electrostatic structure in the denatured states of many proteins. To increase understanding of such electrostatic structure, unstructured pentapeptides have been used as models of the ionization behavior of a total of fourteen aspartates, glutamates, and histidines in a hypothetical random-coil conformation of the simple protein RNase Sa. The pK_as of the residues of interest in the model peptides have already been compared the intrinsic pK_as determined in control pentapeptides for the same residues to determine how local structure perturbs the ionization equilibrium of each ionizable residue from that residue's intrinsic pK_a. In

addition, it was planned to compare the model peptide pK_as to the corresponding pK_as in denatured RNase Sa to determine how any nonrandom structure in denatured RNase Sa further perturbs the ionization properties of each residue studied. Furthermore, these fourteen peptides will be used as RNase Sa-specific approximations of the denatured state of RNase Sa for pH-dependent free energy of folding calculations for RNase Sa. Previous calculations that used the intrinsic pK_as derived from generic model peptides for this approximation agreed poorly with experimentally determined $\Delta G_{folding}$ values. This study has found that out of fourteen ionizable residues researched, seven were perturbed by ± 0.1 or more from intrinsic pK_a values. However, comparison of the denatured state of RNase Sa with model compounds has not yet been achieved because of problems gathering pK_as for twelve of the fourteen residues in the denatured state. Some new approaches are presently being tried to acquire RNase Sa denatured state pK_as for the fourteen ionizable residues. No calculations of pH-dependent $\Delta G_{folding}$ values for RNase Sa using the protein-specific model compounds have yet been performed.

ACKNOWLEDGEMENTS

I thank Dr. J. Martin Scholtz and the members of his and Dr. C. Nick Pace's laboratory listed below for many helpful discussions and for providing an environment that encourages free, creative, original, and careful research. Their expert advice, scientific thoroughness, and patience has helped to transform me from the fledgling researcher I was when I began there in summer 2000 into a future graduate student in one of the best biophysics programs in this country, the Molecular Biophysics program at Johns Hopkins University.

I also thank the following members of the Pace/Scholtz group for their specific help to me in the research areas listed below:

Peptide synthesis and purification – Ron Peterson (a former lab member presently doing postdoctoral work at the University of Pennsylvania), Dr. Michael Merchant, Dr. Scholtz, Richard Thurkill, Jason Schmittschmitt

Potentiometric peptide titration – Richard Thurkill

NMR titration – Beatrice Huyghues-Despointes

RNase Sa solubility mutants – Dr. C. Nick Pace, Saul Trevino, Dr. Gerald Grimsley

TABLE OF CONTENTS

Page

LIST OF FIGURES

INTRODUCTION

The General Problem of Protein Folding

The folding of polypeptide chains into their biologically active conformations is a highly important biological process. Most proteins, especially small ones, equilibrate between two populated ensembles of conformations; the native ensemble is highly structured and held together by many short- and long-distance interactions, and the denatured ensemble, though it often retains some native contacts, is generally unstructured.

Protein folding has multiple applications. For example, genome sequencing projects in the last decade have produced many gene sequences, but the folded structures of their protein products must be known if the functions of these gene products are to be understood. Since exhaustively cloning and expressing all of these sequenced genes and determining X-ray crystal structures for all of the resultant proteins is impossible, developing algorithms for predicting protein structure directly from gene sequence will greatly help structural genomics. In addition, protein misfolding and/or unfolding causes many diseases such as Sickle cell Anemia and Alzheimer's. Furthermore, since proteins fold because of interactions between chemical groups, protein folding research may produce information useful in studying other biological phenomena like protein-protein interactions that are based on chemical group interaction.

This thesis follows the style and format of *Biochemistry*.

The Unfolded State is not Random Coil

The denatured state is the subject of much protein folding research since this state is the beginning of protein folding pathways. Unfolded state structure is especially important to thermodynamic state functions like enthalpy (H), entropy (S), and free energy (G), which depend only on the initial (unfolded) and final (folded or native) states. For many years, researchers considered the unfolded state to be completely unstructured or "random coil" and thus insensitive to mutations that alter the native state of the protein. However, experiments on barley chymotrypsin inhibitor 2 (*1*), L9 (*2*) and Staphylococcal Nuclease (SNase)(*3*) have shown that the denatured conformation of a protein is not always a random coil. The dependence of the free energy of folding, $\Delta G_{folding}$, on pH, is determined mostly by the differences between the pK_as of ionizable residues in the folded and unfolded states. The pK_as of ionizable residues can be easily determined by NMR in the folded state for small proteins, but the corresponding pK_as in the unfolded states cannot be easily determined because chemical shifts for individual ionizable residues often overlap. Thus, denatured state pK_as are often approximated using short glycine or alanine-based model peptides that have no structure. pH-dependent free energies of folding ($\Delta G_{folding}$) calculated for the proteins in the above three studies using the assumption that denatured state ionizable residues possess the pK_as of these unstructured model peptides do not agree with experimentally determined pH-dependent $\Delta G_{folding}$ values. Thus, these proteins probably have perturbed pK_as in the denatured state that result from secondary interactions beyond the random-coil structures of the model peptides. In addition, Shortle *et al.* (*4*) have used NMR techniques to directly detect the

presence of long-range interactions of SNase in 8M urea. Baldwin and Rose (*5*) offer strong theoretical arguments that nonlocal interactions are required to determine the structure of secondary elements as proteins fold, and thus, nonlocal interactions probably exist in proteins' denatured states. Furthermore, in an extraordinary example of how the structure of a protein secondary element can depend on nonlocal interactions, a secondary element of human α-lactalbumin studied by Raleigh *et al.* (*6*) adopted nonnative structure in water when it was isolated from the rest of the protein in a model peptide. Since nonlocal interactions are significant in denatured proteins, they must be identified when studying the denatured state of a protein and separated from local interactions that do not depend on higher order protein structure. This project takes three approaches to understanding local and nonlocal electrostatic interactions in the denatured state structure of Ribonuclease Sa.

Ribonuclease Sa the Model Protein

We use Ribonuclease (RNase) Sa as a model protein in this experiment. As a 96-residue monomer, RNase Sa has no quaternary interactions and fewer secondary and tertiary interactions than the average microbial protein of about 300 amino acids. It has one helix and one antiparallel β-sheet, and its structure is shown in figure 1. RNase Sa has 7 aspartates, 5 glutamates, 2 histidines, 4 arginines, and no lysines for a total charge of about –7 at pH 7 and a pI of 3.5. The pI or isoelectric pH is the pH at which a protein has a net charge of zero. RNase Sa is easy to express and purify in *E. coli* (*7*), and the thermodynamics of its folding are well understood (*8*).

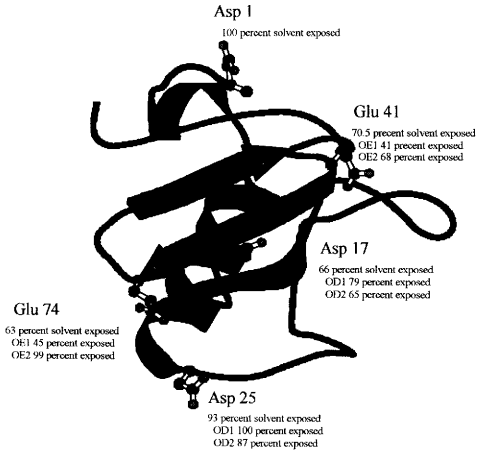

Figure 1: Structure of RNase Sa. RNase Sa is 96 residues long and contains one helix and one antiparallel β-sheet. Figure taken from (*9*).

First Objective

This study has two major objectives, both of which seek to contribute to the knowledge of how ribonuclease Sa folds by improving understanding of this protein's unfolded state. The most important objective is to identify the short-range electrostatic interactions of 14 ionizable residues in RNase Sa, the sequence of which is given in figure 2 along with the sequences of the peptides Ac-abXde-NH_2 representing each of the 14 residues. The arginines of RNase Sa are not examined in this experiment because they have full positive charges up to pH 10, which is well above the physiological pH range. In addition, arginine titration procedures are difficult, and peptide hydrolysis can become a problem in the range of pH 12-13, where arginine titrates. A short-range electrostatic interaction for these pentapeptides is defined as an electrostatic interaction with an i±1 or i±2 residue. In Ac-abXde-NH_2, X, which is either aspartate, glutamate, or histidine, is the ionizable residue, and a, b, d, and e are X's flanking residues in RNase Sa. To determine the intrinsic pK_as of aspartate, glutamate, and histidine residues, that is, their respective pK_as in the absence of ionic influences from neighboring residues, the control peptides Ac-AADAA-NH_2, Ac-AAEAA-NH_2, and Ac-AAHAA-NH_2 will also be made. The N- and C-termini of all control peptides except DVS have been blocked with acetyl and carboxamide groups, respectively, so that these groups will be in amide bonds as in RNase Sa and will not ionize and interfere with pK_a determinations. DVS was not capped on the N-terminus because this peptide sequence begins with aspartate 1, which has an ionizable α-amino group in RNase Sa. From this point on, a peptide will be written only as its sequence

1 ***D***VS GTVCLSAL **PP*EA*T** 16

15 **AT*D*TL** NLI **AS*D*GP** FPY **SQ*D*GV** VGQ **NR*E*SV** LP

46 TQSYG **YY*H*EY** 55

52 **YH*E*YT** TVITPGARTRGTRRII **TG*E*AT** 76

76 **TQ*E*DY** 80

77 **QE*D*YY TG*D*HY** 86

82 **GD*H*YA** 87

88 TFS **LI*D*QT** C

Figure 2: Native sequence of RNase Sa with locations of model peptides shown.
The native sequence of RNase Sa is given above. The sequences represented by model peptides are bolded with residue X italicized. The sequence is displayed in this manner since several peptides overlap. The sequence is taken from (*8*).

abXde, and the presence of the N- and C-terminal blocking groups will be treated as understood. While literature peptide pK_a values for aspartate, glutamate, and histidine in similar control peptides could be used, we have used our own control peptides since these peptides match the backbone structure of our model peptides. The pK_a of each of the 14 residues in abXde will be compared with its pK_a in AAXAA, and the residues flanking X can be examined to determine the most likely source(s) of the perturbation. The pK_a of an ionizable group is the pH at which that group is 50 percent ionized.

The sequence of RNase Sa as given in figure 2 can be used to predict peptides in which the pK_a of X in abXde might be perturbed in from X's pK_a in AAXAA because of Coulombic interactions. The pK_as of glutamate 78 and aspartate 79 in TQEDY and QEDYY, respectively, should be raised above the pK_as of AAEAA and AADAA, respectively, by interactions with adjacent anionic residues aspartate 79 and glutamate 78, respectively; that is, the neutral ($-CO_2H$) form of these carboxylate residues should be

favored. The pK_as of glutamates 41 and 54 and aspartate 84 in NRESV, YHEYT, and TGDHY, respectively, should be depressed below the control pK_as in AAEAA and AADAA, respectively, by the influences of adjacent cationic residues arg 40, his 53, and his 85, respectively; that is, the anionic ($-CO_2^-$) forms of these carboxylates should be favored. In addition, the pK_as of histidines 53 and 85 in YYHEY and GDHYA, respectively, should be raised above the pK_a of AAHAA by adjacent anionic residues glu 54 and glu 84, respectively, that is, the cationic (his^+) form of these histidines should be favored. The pK_as of some of the ionizable residues may be perturbed by nearby hydrophobic and/or hydrogen bonding residues.

Second Objective

The second objective of this study is to determine which of the 14 ionizable residues participate in long-range electrostatic interactions in unfolded RNase Sa in addition to any local interactions found in the first objective of this project. A long-range interaction is an interaction with a residue that is close in structure but not sequence in the protein or an interaction with the overall charge environment of the protein. Such interactions are only possible in the presence of secondary or tertiary protein structure that brings distant amino acids close to each other. The 14 ionizable residues of interest cannot participate in long-range interactions in the unstructured model peptides, but some of them may participate in such interactions in denatured RNase Sa if this denatured state is compact. If a given X residue is involved in a nonrandom structure element, X's pK_a in RNase Sa in 8M urea as measured by NMR should differ from the corresponding pK_a in abXde in 8M urea. Essentially, objective 2 tests the hypothesis that RNase Sa in 8M urea

has nonrandom structure, which in this study would be reflected by nonlocal electrostatic interactions. The null hypothesis in this experiment is that RNase Sa is completely random-coil in 8M urea, that is, that the only the local electrostatic interactions observed in the abXde model peptides. will be observed in denatured RNase Sa.

Third Objective

A third goal of this project is the development of a set of protein-specific model compounds for RNase Sa that can be used to improve the accuracy of $\Delta G_{folding}$ determinations above the accuracy of $\Delta G_{folding}$ calculations that use nonspecific alanine- or glycine-based model peptides that do not take into account the local sequence context of each ionizable residue. As will be discussed later, both local and nonlocal interactions of ionizable residues in the denatured state can perturb the pK_as of these residues from nonspecific model compound values. Our protein-specific model compounds, while not taking into account nonlocal interactions in the denatured state of RNase Sa, should lead to a more accurate $\Delta G_{folding}$ vs. pH determination for RNase Sa because they take into account local interactions.

When combined with other studies being done by my supervisor J. Martin Scholtz and his colleague C. Nick Pace, this study on unfolded RNase Sa will improve the understanding of the folding of RNase Sa and possibly help lead to a folding pathway for this protein and a better understanding of protein folding in general.

MATERIALS AND METHODS

Peptide Synthesis

The fourteen RNase Sa peptides abXde labeled in figure 2 and the three control peptides AADAA, AAEAA, and AAHAA have been synthesized and purified. Each peptide was synthesized at 0.1 millimole scale using rink amide resin solid phase support, fmoc-amino acids, NMP as a solvent, piperidine as a deprotecting reagent, and HOBt/HBTU/DIPEA coupling chemistry. To facilitate purification, acetic anhydride was added between coupling and deprotection steps to acetylate the N-termini of any chains which did not add the appropriate amino acid in the coupling step. Such blocking would stop such a chain from growing any larger and would minimize the size of a side product produced by such a chain. The N-termini of all completed peptides but DVS were also acetylated for reasons discussed in the introduction. While most peptides were made manually, a few were synthesized using an Applied Biosystems automated peptide synthesizer. After the final acetylation, each peptide was cleaved from rink resin with a trifluoroacetic acid (TFA) solution containing the appropriate scavengers to prevent the oxidation of the peptide, and the cleavage reaction released each peptide from rink resin with the C-terminal carboxamide cap. The TFA extract of each peptide was added to –80°C tert-butyl methyl ether, and each peptide was extracted from ether with double distilled water (ddH_2O).

Peptide Purification

Each crude peptide extract was then lyophilized and analyzed by a Perseptive Biosystems matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy

(MALDI/TOF MS) system to check for the presence of the peptide and to detect side products. Peptides found by MALDI to be present in acceptable yield and purity were then redissolved in about 5 mL ddH_2O and purified by FPLC. The purification apparatus used a Resource RPC 3 mL column, 0.1% TFA/ddH_2O as the polar eluent (A), and 95% acetonitrile/4.9% ddH_2O/0.1% TFA as the nonpolar eluent (B). Most peptides have eluted from the column between 90% A/10% B and 80% A/ 20% B, though some of the more hydrophobic peptides have required a gradient rising as high as 35% B. Fractions from each peptide are screened for the desired peptide by MALDI/TOF MS.

Potentiometric Titrations

Purified peptides have been potentiometrically titrated using an automated system designed by a Texas A&M University engineering class. Figure 3 gives a schematic of this system. This system uses a water bath to maintain the peptide at 25°C, a pH meter, a 5.00 mL titration vessel, and custom-designed controlling software based on National Instruments Labview 6i. For histidine (pK_a~6.4) peptide titrations, about 2 μmol of peptide were added to 5.00 mL of double-distilled water with no salt and titrated with 0.01 M HCl from about pH 8.5 down to about pH 4.5 with 1-2μL aliquots of acid. For glu (pK_a~4.3) and asp (pK_a~3.9) titrations, about 20 μmol of peptide were added to 2.50 mL double-distilled water with no salt and were titrated with 0.2M HCl from about pH 6.5 to about pH 2.7 with 1-2μL aliquots of acid. According to the Henderson-Hasselbach equation, an acid titration like those used in this experiment, which covered the pH range of about pK_a±2, titrated the peptide from

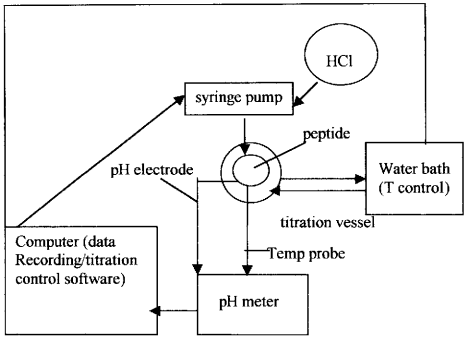

Figure 3: Apparatus for titration of model peptides. The syringe pumps HCl into the titration vessel, using volume and time increments programmed into the computer. From the pH meter and temperature probe, the computer records the pH and temperature of the titration vessel, respectively, at each point. The water bath keeps the titration vessel at 25±1°C.

about 99.0% deprotonated to about 99.0% protonated. All acid solutions in the titrations were degassed to remove CO_2, which can interfere with titration of the analyte. Toward the end of my work, I began running a continuous flow of nitrogen (not shown on figure) through the titration vessel to keep CO_2 out during the titration. For each titration, a blank of the same volume, acid concentration, and pH range was subtracted from the sample titration curve.

The resultant volume vs. pH data is fitted using Microcal Origin 6 to the

$$V = (V_{A^-} + V_{HA} * 10^{n(pKa-pH)})/(1+10^{n(pKa-pH)})$$

equation where V is the blank-corrected titrant volume, V_{A^-} is the blank-corrected volume of titrant at the basic baseline, V_{HA} is the corrected volume of titrant at the acidic baseline, n is the Hill coefficient or number of protons transferred during the titration, and pK_a is the pH at which the analyte residue X is half-ionized. Each titration was run in duplicate, and reproducibility of ± 0.03 pK_a units could generally be obtained between the two runs.

The potentiometric titration method described above does not work for the peptides TQEDY and QEDYY because the two carboxylates in each of these peptides titrate too closely to be resolved. However, a proton NMR-based method was successfully used to determine these two pK_as. About 5 μmol of the peptide QEDYY was dissolved in 100% D_2O and titrated from pH 10 to pH 2.5 with 2.5% DCl, and the chemical shifts of the C_γ protons on glu 78 and the C_β protons on asp 79 were recorded at each pH. These chemical shifts were recorded in reference to sodium 3-trimethylsilyl propionate (TSP). The slightly different environments of these two sets of protons allowed their chemical shifts to be sufficiently separated that they could be resolved. Plots of chemical shift vs. pH were prepared for glu 78 and asp 79 and analyzed in the same way as the potentiometric titration curves for the other residues. Both pK_a values were derived from the QEDYY peptide since the TQEDY peptide could not be sufficiently purified for NMR.

Data Interpretation

The pK_a values found for the fourteen ionizable residues in the model peptides and for the AADAA, AAEAA, and AAHAA control peptides will be sufficient to accomplish the first major objective of this project. If, for a given ionizable residue, the pK_a of X in abXde is within experimental error (±0.05 - ±0.10) of the corresponding value for AAXAA, then a, b, d, and e either exert no significant electrostatic influence on the pK_a of X or exert balancing electrostatic influences. If the pK_as in abXde and AAXAA differ by more than ±0.10, then a, b, d, and e can be analyzed to determine how these flanking residues perturb the ionization equilibria of X.

The second objective requires additional data. Other members of the laboratories of my supervisor and a colleague have tried to determine by NMR the pK_as of the 14 ionizable residues in RNase Sa in 8M urea. While pK_as for all 14 ionizable residues in denatured RNase Sa have not been successfully gathered by NMR, two other approaches that will be discussed in the results section may be tried to gather the denatured state pK_as for these residues. The pK_as of each of the 14 ionizable residues in unfolded RNase Sa need to be taken both with and without an agent to block the native disulfide linkage between cysteines 7 and 96. If a given region abXde of unfolded RNase Sa is in the random coil state represented by its model peptide, then X should have the same pK_a (within ± 0.2) in both the model peptide in 8M urea/ddH_2O and in RNase Sa in 8M urea/ddH_2O. If the pK_a of X in RNase Sa in 8M urea/ddH_2O is more than ±0.03 different from the pK_a in abXde in 8M urea/ddH_2O, then X is probably involved in long-range interactions in denatured RNase Sa. If X's pKa in denatured RNase Sa is closer to its

native conformation pK_a than to its pK_a in abXde, then X in denatured RNase Sa may be involved in nativelike secondary structure.

For the third objective, other members of the Pace/Scholtz group will calculate $\Delta G_{folding}$ vs. pH using the RNase Sa model peptides for the denatured state pK_as of the ionizable residues.

RESULTS AND DISCUSSION

First Objective

Out of fourteen pK_a measurements desired, thirteen could be measured. Table 1 shows the pK_as measured for these thirteen peptides and for the three AAXAA control peptides. Figure 4 shows a blank-corrected potentiometric peptide titration curve for AAEAA, which is a typical curve for a carboxylate titration. Curves for the histidine peptides have the same shape but use 0.01M HCl instead of 0.2M HCl.

For histidine pK_as, it was expected that the charged form (his^+) would be favored for histidine 53 in YYHEY and histidine 85 in GDHYA because both of these histidines are adjacent to anionic carboxylate residues. The pK_a of histidine 85 was elevated by 0.26 relative to AAHAA, in accordance with the prediction, but the pK_a of histidine 53 was not elevated relative to AAHAA; only an elevation of +0.1 or larger would be considered significant. The effect of glu 54 on histidine 53 may be canceled by other interactions in YYHEY, such as glu 54 hydrogen bonding with tyr 55 so that glu 54 does not interact with his 53. A repetition of the titration of YYHEY would be necessary to clarify that his 53 does not have a perturbed pK_a relative to AAHAA.

Both expected and unpredicted pK_a effects were observed for the glutamate peptides. Glu 41 in NRESV and glu 54 in YHEYT were expected to be favored in the charged (glu^-) forms relative to AAEAA because glu 41 is adjacent to cationic arg 40, and because glu 54 is adjacent to cationic his 53. The pK_as of glu 41 in NRESV and glu 54 in YHEYT were depressed by 0.15 and 0.10, respectively, relative to AAEAA, in accordance with these predictions. However, an unexpected elevation of the pK_a of

Table 1: pK_a values for selected ionizable residues in RNase Sa in control and model peptides and for native and denatured RNase Sa

		peptide[a]	peptide[a]	ΔpKa (peptide)[b]	RNase Sa	
Sa residue	model pep.	0M salt, 25°C	8M urea, 25°C	(from AAXAA)	native[c]	denatured[d]
H control	AA*H*AA	6.39	-		-	-
H53	YY*H*EY	6.42	-	0.03	8.27	7.2(6.6)
H85	GD*H*YA	6.65	6.78	0.26	6.34	7.03(6.4)
E control	AA*E*AA	4.31	4.92		-	-
E14	PP*E*AT	4.40	-	0.09	5	-
E41	NR*E*SV	4.16	-	-0.15	3.41	-
E54	YH*E*YT	4.21	-	-0.10	4.14	-
E74	TG*E*AT	4.31	-	0	3.53	-
E78	TQ*E*DY[e]	4.27	-	-0.04	3.17	-
D control	AA*D*AA	3.95	-		-	-
D1	*D*VS	3.08	-	-0.87	3.45	-
D17	AT*D*TL	3.82	-	-0.13	3.71	-
D25	AS*D*GP	3.98	-	0.03	4.87	-
D33	SQ*D*GV	3.85	-	-0.10	2.3	-
D79	QE*D*YY[e]	3.90	-	-0.05	7.2	-
D84	TG*D*HY	3.33	-	-0.62	-	-
D93	LI*D*QT	-	-	-	3.2	-

[a]These individual values have an error of approximately ±0.03. All values but AADAA, TGDHY, and DVS are the average of two measurements. All peptides but TQEDY and QEDYY were titrated potentiometrically. For a sample titration curve, see figure 4 below.
[b]=pK_a(peptide abXde in 0M salt, 25.0°C) - pK_a(AAXAA in same conditions). A ΔpK_a of ±0.10 or more is considered significant.
[c]determined by NMR in 100% D_2O, 20.0°C, 0.1M salt (*10*)
[d]determined by NMR in 8M urea, 100% D_2O, 20.0°C, 0.1M salt (*10*)
[e]Determined by NMR, 100% D_2O, 25.0°C, 0M salt. Both E78 and D79 pK_as were determined from a single experiment on the QEDYY peptide. See figure 5 for a sample titration.

glu 14 in PPEAT by 0.09 relative to AAEAA was observed. This elevation for glu 14 is just below the 0.10 significance threshold, but it is unclear if this elevation is a consequence of experimental error or of an unpredicted electrostatic effect. This elevation means the uncharged (glu^0) form of glu 14 is favored. This effect could not result from the nonpolar effect of a nearby bulky hydrophobic residue or from a repulsive electrostatic interaction with another anionic residue since glu 14 is not adjacent to either of these. It is possible that this pK_a perturbation could be a

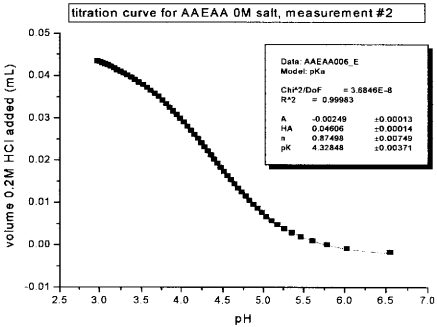

Figure 4: Blank-corrected titration curve for the peptide AAEAA with 0.200 M HCl in water at 0M salt. This particular curve, which is a good representation of a titration curve for a carboxylate side chain, yielded a pK_a value of 4.33±0.01 for the glu side chain in AAEAA. Similar-shaped curves are observed for histidine side chain titrations, but an HCl concentration of 0.01 M was used for the histidine peptides. Microcal Origin 6.0 was used to plot and analyze this curve.

consequence of experimental error. The measurements for both AAEAA and glu 14 and PPEAT could be reproduced within ±0.02, but instrument drift between the times of pK_a measurements for AAEAA and glu 14 in PPEAT might have introduced enough experimental error to cause this apparent pK_a shift. No significant perturbation was observed for glu 74 in TGEAT, but the local sequence context of glu 74 does not suggest that such an effect should exist. The pK_a of glu 78 in QEDYY (TQEDY was not used for reasons discussed in the methods section) was depressed by 0.04 from the pK_a of glu in AAEAA, contrary to the expectation that it would be elevated by the adjacent asp 79. This could result from the fact that QEDYY was only titrated once or from the fact that

other interactions in QEDYY may suppress or cancel the expected interaction of glu 78 and asp 79.

For the aspartate pK_as, both expected and unexpected electrostatic effects were observed. It was expected that the pK_a of asp 1 (side chain) in DVS would be depressed relative to AADAA by the cationic α-amino group of asp 1 and that the pK_a of asp 84 in TGDHY would be depressed relative to AADAA by the adjacent histidine 85. As expected, the pK_a of asp 1 (side chain) in DVS was depressed by 0.87, and the pK_a of asp 84 in TGDHY was depressed by 0.62, relative to AADAA. Unexpected pK_a depressions relative to AADAA were observed with asp 17 in ATDTL (0.13) and asp 33 in SQDGV (0.10). The titrations for asp 17 in ATDTL, asp 25 in ASDGP, and asp 33 in SQDGV were all done with good quantities of peptide and with reproducibilities of at worst 0.05 between pK_a measurements. However, the titrations of AADAA and TGDHY were all done with small quantities of peptides using a hand-titration procedure that is probably not as reliable as the automated procedure discussed in the methods section. In addition, the DVS was titrated only once because of a lack of time. The ΔpK_as measured for asp 1 (side chain) in DVS and asp 84 in TGDHY are so large that they probably represent real electrostatic effects, though it is unlikely that these effects are as large as the ΔpK_as represent. For example, the ΔpK_a for glu 54 in YHEYT, which is adjacent to his 53 as asp 84 is adjacent to his 85, is only –0.10 compared to –0.62 for asp 84. Thus, there may be significant error in the pK_as for AADAA, TGDHY, and DVS. It may be possible that asp 1 has a larger ΔpK_a than other carboxylates with nearby cationic groups because the α-amino group of asp 1 in DVS is on the same residue as asp 1 and thus closer to the

perturbed side chain than any of the other cationic groups are to the carboxylates whose pK_as they perturb. If the pK_a of AADAA is too high because of the above reasons, then the depressions in the pK_as of asp 17 in ATDTL and asp 33 in SQDGV may not be real effects. However, possible hydrogen bonding of asp 17 with thr 16 and thr 18 and of asp 33 with gln 32 may favor the asp^- forms of these two residues, which would depress of the pK_as of these two carboxylates relative to AADAA. Since for aspartates in proteins there is an inverse correlation between number of hydrogen bonds and pK_a (*11*), it is reasonable to suspect that in peptides, the pK_as of hydrogen-bonding aspartates like asp 17 and asp 33 might be depressed relative to AADAA. No significant ΔpK_a was observed for asp 25 in ASDGP, but none was expected from the local sequence context of asp 25. The pK_a of asp 79 in QEDYY was depressed by 0.05 from the pK_a of asp in AADAA, contrary to the expectation that it would be elevated by the adjacent glu 78. Since the titration curve for asp 79 was determined by NMR rather than by potentiometry, this titration curve is given in figure 5. The titration curve for glu 78 closely resembles the titration curve for asp 79. Other electrostatic or hydrogen bonding effects in QEDYY may cancel or prevent the interaction of glu 78 and asp 79. Because the I/D and L/I couplings in the peptide LIDQT involved coupling with the β-branched side chain of isoleucine, not enough LIDQT could be recovered to allow for titration of LIDQT to determine the pK_a of asp 93.

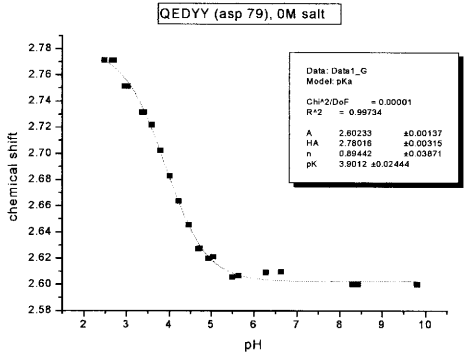

Figure 5: Titration curve for asp 79 in QEDYY. This curve was determined by NMR in 100% D_2O, 0M salt at 25°C using 2.5% DCl as a titrant, and the chemical shift is relative to sodium 3-trimethylsilyl propionate (TSP). The pK_a for asp 79 was found from this measurement to be 3.90 ± 0.02. The titration curve for glu 78 closely resembles this titration curve. Microcal Origin 6.0 was used to plot and analyze this curve.

Second Objective

Work on this objective was not successful because pK_as could not be acquired in the urea-denatured, disulfide-reduced state of RNase Sa by NMR. In addition, there was only time to determine the pK_as in 8M urea of AAEAA and his 85 in GDHYA. The model peptide pK_as taken in distilled water differ substantially from the corresponding pK_as in 8M urea for these peptides; the pK_as are 0.13 higher and 0.61 higher for his 85 in GDHYA and glu in AAEAA, respectively, in 8M urea than in water. However, it is not

worth expending the effort to acquire the other 12 peptide pK_as in 8M urea until the denatured state pK_as for each ionizable residue can be found.

pK_as for histidines 53 and 85 have been found in urea-denatured RNase Sa by NMR, but similar attempts to determine the pK_as for the 12 carboxylate residues in this study have been unsuccessful because the NMR experiments to determine the pK_a of the carboxylates take substantially longer than the NMR experiments to determine the histidine pK_as. With the carboxylate NMR experiments, some of the RNase Sa begins to precipitate before the experiments can be completed. Two mutants of RNase Sa have been constructed with the objective of increasing the solubility of RNase Sa in 8M urea so that it will not precipitate before the experiments can be completed. The two mutants, V35Q and I58N, have replaced surface hydrophobic residues so that the stability of the wild type protein and thus its expression would not be severely altered, and have mutated the two hydrophobic residues to uncharged polar residues since mutation to an ionizable residue would introduce additional electrostatic effects into the mutant proteins. Asn and gln were chosen since they can participate in the most hydrogen bonds of any polar uncharged residue. These two mutants have been successfully expressed and purified according to the protocol in (7), and both mutations increase the solubility of RNase Sa at pH 3.5 (the pI and solubility minimum of wild type RNase Sa), 0.1M KCl, by five times or so. A third mutant was made combining these two mutants, but this mutant was less soluble than wild type, and it was less soluble than either mutant was alone. The solubility results for these mutants are highly reproducible for two mutants and reproducible within a factor of two for the other mutant, within one batch of each mutant.

I later plan to test the solubility of each of these three mutants in 8M urea. If these mutants are substantially more soluble in 8M urea than wild type RNase Sa, it might be possible to retain the mutants in 8M urea long enough to gather the carboxylate pK_as by NMR.

Another possible approach to gather the twelve pK_as for the carboxylates is to prepare aspartate to asparagine mutants for all the asp residues and glutamate to glutamine mutants for each of the glu residues in the protein. For an aspartate, the potentiometric titration curve in 8M urea of the D->N mutant could be subtracted from the corresponding curve for wild type; in theory, the net curve should be the titration curve of the aspartate of interest. A similar approach could be applied for the glutamates. This approach has been used in native RNase Sa to determine the pK_a of asp 79 by subtracting the titration curve for a D79F mutant from the titration curve of wild type to acquire the titration curve of D79 by itself. The principal problem with this mutation approach is that twelve RNase Sa mutants would take a long time to make. It would be interesting, however, to try this approach for one or two of the carboxylates.

Third Objective

The third objective of this project was to use the fourteen protein-specific model compound pK_as for all of the ionizable residues in RNase Sa to more accurately calculate the free energy of folding of RNase Sa at various pH values. Since seven ionizable residues in RNase Sa have ΔpK_as of ± 0.1 or greater from their control peptides AAXAA, failure of nonspecific model compounds to account for local pK_a perturbations may partially cause the inaccuracy of pH-dependent $\Delta G_{folding}$ calculations for RNase Sa. The

inaccuracy caused by the use of nonspecific model compounds may be less significant in larger proteins with more ionizable residues, but since most protein folding model proteins are small monomers, the use of nonspecific model compounds in protein folding studies will probably frequently cause inaccurate pH-dependent $\Delta G_{folding}$ determinations. Even protein-specific model compounds, however, will not produce fully accurate pH-dependent $\Delta G_{folding}$ determinations because these model compounds do not take into account long-range denatured state electrostatic structure. These protein-specific model compounds have not yet been used to calculate $\Delta G_{folding}$ for RNase Sa, but their use will probably improve the accuracy of these calculations.

CONCLUSIONS

Seven residues in the RNase Sa model peptides abXde have pK_as for X perturbed by ±0.1 or more from the corresponding pK_as in AAXAA. Five of these can be easily attributed to attractive charge/charge interactions, and two to possible hydrogen bonding interactions. Three expected pK_a perturbations, the expected elevation of the pK_a of histidine 53 by the presence of glutamate 54 and the expected elevation of the pK_as of glu 78 and asp 79 because of their mutually repulsive interaction, were not found. Thus, local electrostatic interactions in RNase Sa exert significant influences on the pK_as of ionizable residues in this protein apart from the presence of any higher order structure.

Accomplishing the second objective of this project would have required pK_as for the 14 ionizable residues of interest in the urea-unfolded state of RNase Sa. Though solubility problems of the wild type protein in 8M urea have prevented the gathering of all but two of these pK_as, a mutational approach has been tried with promising results that may allow RNase Sa to be kept in solution long enough to obtain by NMR the pK_as of the other twelve residues in the denatured state. If this solubility approach fails, it may be possible to gather some of the unfolded state pK_as by a combination mutation/potentiometric titration approach.

The prediction that recalculating a pH dependence of $\Delta G_{folding}$ for RNase Sa using these protein-specific model compounds instead of using nonspecific model compounds would improve the accuracy of this calculation has not yet been tested. However, the significant number of perturbed pK_as identified raises strong suspicion that the use of nonspecific model compounds may contribute strongly to the error of the $\Delta G_{folding}$

calculation. Some additional error in the $\Delta G_{folding}$ calculation may result from failure of both nonspecific model compounds or protein-specific model compounds to take into account the possible presence of nonlocal structure in the denatured state of RNase Sa.

If and when it becomes possible to gather pK_as for all fourteen ionizable residues in the unfolded state of RNase Sa, we will have gathered very specific information about the level of nonlocal electrostatic structure in the denatured state of proteins. This will improve substantially on the knowledge of protein denatured states, which though currently known to often contain nonrandom structure, are not well-known in terms of specific interactions. Specific information about the denatured states of proteins will be required if science is to fully understand and predict this process and its effects on biochemistry and medicine.

REFERENCES

(*1*) Fersht, A.R., Tan, Y.J., Oliveberg, M., & Davis, B. (1995). Perturbed pK_a Values in the Denatured States of Proteins. *Journal of Molecular Biology* **254**, 980-992.

(*2*) Raleigh, D.P., Kuhlman, B., Luisi, D.L., & Young, P. (1999). pK_a Values and the pH Dependent Stability of the N-Terminal Domain of L9 as Probes of Electrostatic Interactions in the Denatured State. Differentiation between Local and Nonlocal Interactions. *Biochemistry* **38**, 4896-4903.

(*3*) Garcia-Moreno, E. & Whitten, S.T. (2000). pH Dependence of Stability of Staphylococcal Nuclease: Evidence of Substantial Electrostatic Interactions in the Denatured State. *Biochemistry* **39**, 14292-14304.

(*4*) Shortle, D. & Ackerman, M.S. (2001). Persistence of Native-Like Topology in a Denatured Protein in 8 M Urea. *Science* **293**, 487-489.

(*5*) Baldwin, R.L. & Rose, G.D. (1999). Is protein folding hierarchic? I. Local structure and peptide folding. *Trends in Biochemical Sciences* **24**, 26-33.

(*6*) Raleigh, D.P., Demarest, S.J., & Hua, Y. (1999). Local Interactions Drive the Formation of Nonnative Structure in the Denatured State of Human α-Lactalbumin: A High Resolution Structural Characterization of a Peptide Model in Aqueous Solution. *Biochemistry* **38**, 7380-7387.

(*7*) Pace, C.N., Hebert, E.J., Grimsley, G.R., Hartley, R.W., Horn, G., Schell, D., Garcia, S., Both, V., & Sevcik, J. (1997). Purification of Ribonucleases Sa, Sa2, and Sa3 after Expression in *Escherichia coli*. *Protein Expression and Purification* **11**, 162-168.

(*8*) Pace, C.N., Hebert, E.J., Shaw, K.L., Schell, D., Both, V., Krajcikova, D., Sevcik, J., Wilson, K.S., Dauter, Z., Hartley, R.W., & Grimsley, G.R. (1998). Conformational Stability and Thermodynamics of Folding of Ribonucleases Sa, Sa2, and Sa3. *Journal of Molecular Biology* **279**, 271-286.

(*9*) Pace, C.N., Shaw, K.L., Grimsley, G.R., Yakovlev, G.I., & Makarov, A.A. (2001). The effect of net charge on the solubility, activity, and stability of ribonuclease Sa. *Protein Science* **10**, 1206-1215.

(*10*) Pace, C.N. and Huyghues-Despointes, B.M. (2000). Unpublished data.

(*11*) Robertson, A.D., Forsyth, W.R, and Antosiewicz, J.M. (2002). Empirical Relationships Between Protein Structure and Carboxyl pK_a Values in Proteins. *Protein Structure, Function, and Genetics*, publication pending.

VITA

Michael Douglas Daily

15930 Aberdeen Trails Drive

Houston, TX 77095

I will graduate from Texas A&M University in May 2002 with a B.S. degree in Biochemistry and Genetics. In summer 2002, I will begin work in the Ph. D. program in Molecular Biophysics at Johns Hopkins University. Thereafter, I will pursue a postdoctoral position and eventually a faculty position in biophysics at a university.

During my career at Texas A&M, I have been awarded a 2001 Barry M. Goldwater scholarship. I have held a President's Endowed Scholarship, a Director's Excellence Award, a Merit Plus Award, and a Texas A&M National Merit Sponsorship during my time here. I have also received the Honor Society of Gamma Sigma Delta Texas A&M University Chapter's Outstanding College of Agriculture and Life Sciences Senior Award for 2002. I have also been on the College of Agriculture and Life Sciences Dean's Honor Roll throughout my time here.

I have been a member of the Biochemistry and Genetics Societies, the Executive Council of Health Organizations, and the Honor Society of Phi Eta Sigma at Texas A&M. I have also advised freshman scholarship students under the ASPIRE peer mentor program.