IMPROVEMENT OF THE CHEMICAL AND PHYSICAL STABILITY OF THE EC1 DOMAIN OF E-CADHERIN BY BLOCKING ITS DISULFIDE-MEDIATED DIMERIZATION

BY

C2008
Maulik Vasant Trivedi

M.S., Pharmaceutical Chemistry, University of Kansas, 2004

Submitted to the graduate degree program in Pharmaceutical Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Chairperson

_____

_____

_____

_____

Date defended: January 21, 2008

The Dissertation Committee for Maulik Vasant Trivedi certifies

that this is the approved version of the following dissertation:

IMPROVEMENT OF THE CHEMICAL AND PHYSICAL STABILITY OF THE EC1 DOMAIN OF E-CADHERIN BY BLOCKING ITS DISULFIDE-MEDIATED DIMERIZATION

Committee:

_____
Chairperson

_____

_____

_____

_____

Date approved: January 28, 2008

ABSTRACT

The EC1 protein has an important role in the E-cadherin mediated cell-cell adhesion in epithelial and endothelial tissues. It may be used as modulator of cellular adhesion to improve paracellular delivery of macromolecules. EC1 undergoes oxidation of its Cys residue to form disulfide-linked covalent dimers. These dimers associate to form physical oligomers. The dimerization and oligomerization also lead to hydrolysis of the Asp93–Pro94 peptide bond and precipitation. To be able to study the cell adhesion-modulating activity of EC1, it is important that we block its disulfide-dimerization and subsequent oligomerization.

The strategies used to block the dimerization were addition of dithiothreitol to the EC1 solution, modification of the Cys thiol with iodoacetamide and polyethylene glycol. These derivatives were studied for their stability using HPLC, CD and fluorescence spectroscopy. All strategies applied showed improvement in the stability of EC1. The PEGylated EC1 showed the best stability.

Dedicated to my parents

ACKNOWDLEDGEMENTS

This work would not have been possible without the direct and indirect contributions of many people. I would like to acknowledge them here.

Dr. Siahaan has mentored me through this project. He taught me how to develop the appropriate attitude to conduct scientific research. Whenever I seemed to be wandering away in the wrong direction in my research, he brought me back in the right direction, reminding me what my research goals were. Without his training, I wouldn't be able to finish my project and obtain my degree. Through his own conduct, he has also taught me discipline and professionalism. I will always be grateful to him.

I would like to thank Dr. Laurence for her constant help throughout my research. She always treated me as a student in her own lab. She has also contributed to improving my thesis by reading it and suggesting the necessary changes. Dr.Middaugh has provided valuable insight into the results obtained as part of my research. I would like to give me heartfelt thanks to him, too. He read my thesis in a timely manner and helped me improve it tremendously. I would like to thank Dr. Kuczera and Dr. Berkland for being on my dissertation committee and providing me thoughtful suggestions. I would also like to thank Dr. Forrest for substituting for Dr. Berkland at my dissertation defense at very short notice.

I would like to thank Dr. Verkhivker and Ambrish Roy for working with me and performing the molecular dynamics simulations. Also, my thanks are to the entire pharmaceutical chemistry faculty for giving me such a great education.

My labmates through the last seven years have always been helpful, co-operative, friendly and dependable. I owe thanks to Dr. Irwan Makagiansar for teaching me how to conduct research when I first started working in Dr. Siahaan's lab. I have been fortunate to work closely with Kai, Phuong, Henry, Tatyana, Meagan, Anna, Naoki, Sumit, Bimo, Prakash and Barlas, Yumna and Andrew. I would also like to thank the members of Dr. Laurence's, Dr. Middaugh's, Dr.Borchardt's and Dr. Schoeneich's labs for letting me use their lab instruments and teaching me how to use them. I especially thank Dr. Tim Priddy for showing me how to use the CD and fluorescence instruments.

I would like to thank the Pharm. Chem. staff: Nancy, Karen, Ann, Nicole and Sheila, for taking care of administrative things for me and my fellow students so we can concentrate better on research.

I have been lucky to make some great friends in Lawrence, KS. Most of them have moved away, but still remain great friends. Raja, Becky, Vikram, Sandeep, Ajit, Sandipan, Pulin, Loren and all the softball teams I played on, all made life after lab much more enjoyable.

I would also like to thank Jodi for being amazingly supportive and encouraging to me. She dealt with my stress really well during my last year of research.

Lastly, I would like to thank my family for making me what I am today. My parents and sister and all my loving uncles, aunts and cousins always have been there for me. I am at this juncture today because of my parents who always stressed how important a good education is. I can never thank them enough for that.

TABLE OF CONTENTS

Chapter 1. Introduction 1

1.1. Introduction 2

1.2. Disulfide bond formation as a post-translational modification of proteins 5

1.3. Pathways of oxidative folding of proteins for disulfide formation 7

1.4. Effects of disulfide bond on protein stability 9

1.5. Analysis of free thiols and disulfide bonds 14

1.6. Degradation of disulfide bonds 19

1.7. Chemical modification of Cys thiols of proteins 25

1.8. Conclusions 29

1.9. Application of the information above to current work 30

1.10. References 32

Chapter 2. The effect of covalent dimerization on the structure and stability of EC1 domain of human E-cadherin 43

2.1. Introduction 44

2.2. Materials and methods 47

2.2.1. Expression of EC1 47

2.2.2. Purification of the EC1 protein 49

2.2.3. Chemical stability studies 52

2.2.4. Physical stability studies 53

2.2.5. Computer simulations and molecular modeling 55

2.3. Results 56

2.3.1. Chemical stability of EC1 56

2.3.2. Physical stability of EC1 65

2.3.2.1. Secondary structure evaluation using CD 65

2.3.2.2. Evaluation of tertiary structural changes by fluorescence spectroscopy 73

2.4. Molecular dynamics simulations of EC1 monomer and dimer 77

2.5. Discussion 82

2.6. References 88

Chapter 3. Improving the stability of EC1 by modification of its Cys13 thiol group by alkylation 92

3.1. Introduction 93

3.2. Materials and methods 95

3.2.1. Production of alkylated of EC1 95

3.2.2. Chemical stability studies 96

3.2.3. Physical stability studies 97

3.2.3.1 Far UV CD studies 98

3.2.3.1 Intrinsic fluorescence emission spectroscopy studies 99

3.3. Results 100

3.3.1. Alkylation of EC1 100

3.3.2. Chemical stability of EC1-IN 105

3.3.3. Physical stability of EC1-IN 108

3.3.3.1. Secondary structure studies using CD spectroscopy 108

3.3.3.2. Tertiary structure studies using fluorescence spectroscopy 118

3.4. Discussion 121

3.5. References 128

Chapter 4. PEGylation of the Cys13 thiol of EC1 improves its stability while retaining its structure 133

4.1. Introduction 134

4.2. Methods 136

4.2.1. Conjugation of EC1 with PEG 136

4.2.2. Purification of the PEGEC1 conjugate 137

4.2.3. Chemical stability studies of PEGEC1 137

4.2.4. Structural studies of PEGEC1 138

4.2.4.1. Far UV CD spectroscopy 138

4.2.4.2. Intrinsic fluorescence emission spectroscopy 139

4.3. Results 140

4.3.1. Chemical stability studies 145

4.3.2. Physical stability studies 147

4.3.2.1. Far UV CD results 147

4.3.2.2. Intrinsic fluorescence emission spectroscopy results 156

4.4. Discussion 164

4.5. References 167

Chapter 5. Summary and future work 170

5.1. Summary 171

5.2. Future work 174

5.3. Final comments 175

FIGURE LEGENDS

Figure	Page	Caption
1.1	3	Increasing trend in the use of recombinant proteins and monoclonal antibodies as drugs. The number of new biotech drugs and new indications approved for biotech drugs from 1982 to 2005.
1.2a	16	Various labeling reagents for the detection and quantification of protein thiols and disulfides: 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB or Ellman's reagent), 4,4'-dithiodipyridine (DTDP), n-octyl-5-dithio-2-nitrobenzoic acid (ODNB), and ThioGlo®1.
1.2b	17	Labeling cellular protein thiols with bimanes: monobromobimane (mBBr) and dibromobimane (bBBr) label thiols inside cells as well as on cell surfaces, whereas monobromotrimethylammoniobimane (qBBr) labels only the cell surface protein thiols.
1.3a	20	Degradation reactions of protein disulfide bonds in neutral and basic conditions that involve a direct attack on the sulfur atom by hydroxyl anions.
1.3b	21	Degradation reactions of protein disulfide bonds in neutral and basic conditions involving a β-elimination reaction. The α-proton of the Cys residue is abstracted by the hydroxyl anion.
1.3c	22	Degradation reactions of protein disulfide bonds in neutral and basic conditions involving α-elimination reaction. The β-proton of the Cys residue is abstracted by the hydroxyl anion to produce a thiolate/thioaldehyde (15).
1.4	26	Some of the reagents used for the modification of free thiols in proteins.

2.1a	46	Primary sequence of the EC1 protein obtained at the end of the purification process. The first four residues GSHM are non-native residues that result from the type of vector used for transformation into *E. coli* cells. The single Cys13 residue is equivalent to Cys9 residue in the native EC1. The potential hydrolysis site is also indicated.
2.1b	51	Mass spectrum of purified EC1 as obtained from electrospray ionization-time-of-flight mass spectrometer (ESI-TOF MS).
2.2a	57	The stability profile of EC1 incubated at 37 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 3.0. In the absence of DTT, there is some loss in the amount of EC1, presumably due to precipitation, since no new peaks were observed on the HPLC under these conditions.
2.2b	58	The stability profile of the EC1 domain incubated at 37 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 7.0. 25% EC1 was lost after 4h due to precipitation, in the absence of DTT. In the presence of 0.1 M DTT, however, only 6% of EC1 precipitated after 4 hours.
2.2c	59	The stability profile of the EC1 domain incubated at 37 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 9.0 studied by monitoring the loss in area under the peak of EC1 and appearance and increase in the area under a new peak on HPLC. In the absence of DTT, about 27% of EC1 degraded to form a new product. In the presence of 0.1 M DTT, however, only 13% of EC1 degraded to form the new product.
2.2d	60	ESI (+ve)- TOF MS of the degradation product of EC1
2.3a	62	The stability profiles of the EC1 domain incubated at 70 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 3.0 as determined by HPLC analysis. In the absence of DTT, EC1 undergoes some peptide bond hydrolysis, but

		mostly precipitation. In the presence of 0.1 M DTT, however, the extent of precipitation of EC1 after 4 h is decreased and no hydrolysis product is seen.
2.3b	63	The stability profiles of the EC1 domain incubated at 70 °C for 4 hin the absence and presence of 0.1 M DTT in buffer solutions at pH 7.0 as determined by HPLC analysis. EC1 undergoes hydrolysis of the peptide bond (12% after 4 h) and precipitation (70% after 4 h). In the presence of DTT, however, the extent of peptide bond hydrolysis is decreased (6% after 4 h) and so is the precipitation of EC1 (40% after 4 h).
2.3c	64	The stability profiles of the EC1 domain incubated at 70 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 9.0 as determined by HPLC analysis. In the absence of DTT, 53% EC1 undergoes peptide bond hydrolysis after 4 h. In the presence of 0.1 M DTT, however, the peptide bond hydrolysis is decreased to 27% after 4 h.
2.4a	66	CD spectra of EC1 in the absence and presence of DTT after incubation at 4 °C for 0, 14, or 28 days at pH 3.0. In the absence of DTT, EC1 undergoes secondary structural change presumably due to dimerization and oligomerization. In the presence of 2 mM DTT, there is no secondary structural change in EC1.
2.4b	67	CD spectra of EC1 in the absence and presence of DTT after incubation at 4 °C for 0, 14, or 28 days at pH 7.0. After 28 days, there is some loss of the β-sheet structure of EC1 in the absence of DTT as seen from the decrease in the minimum at 216 nm; addition of 2 mM DTT suppresses the secondary structural change of EC1.
2.4c	68	CD spectra of EC1 in the absence and presence of DTT after incubation at 4 °C for 0, 14, or 28 days at pH 9.0. After 28 days, the CD spectra remain unchanged in the presence and absence of DTT.

2.5a	70	Thermal unfolding profiles of EC1 in the absence and presence of DTT measured by CD at 218 nm after incubation for 14 to 28 days at 4 °C at pH 3.0. The thermal unfolding profiles in the absence of DTT are different on day 14 compared to that on day 0; in the presence of DTT, the profiles look the same on days 0 and 14.
2.5b	71	Thermal unfolding profiles of EC1 in the absence and presence of DTT measured by CD at 218 nm after incubation for 14 to 28 days at 4 °C at pH 7.0. The thermal denaturation profiles in the absence of DTT are different at different time points; however, the profiles are relatively the same in the presence of DTT.
2.5c	72	Thermal unfolding profiles of EC1 in the absence and presence of DTT measured by CD at 218 nm after incubation for 14 to 28 days at 4 °C at pH 9.0. The thermal denaturation profiles are comparable after incubation for 14 and 28 days in the presence or absence of DTT.
2.6a	74	Intrinsic fluorescence emission spectra of EC1 (after excitation at 295 nm) in the absence and presence of DTT after incubation at 4 °C for 0 or 28 days at pH 7.0. A shift in the emission maximum to a higher wavelength is observed in the absence of DTT; however, this shift is inhibited by addition of DTT.
2.6b	75	Thermal unfolding of EC1 in the absence and presence of DTT evaluated by change in the wavelength of maximum intrinsic fluorescence emission (after excitation at 295 nm) from 0 °C to 87.5 °C after incubation for 0, 14, or 28 days at 4 °C and pH 7.0.
2.7	78	RMSD values for 10 ns simulation of EC1 monomer (panel A) and EC1 dimer (panel B).
2.8	79	RMSF values for each residue in EC1 monomer (panel A) and in both A and B chains of EC1 dimer

		(panel B) at the end of 10 ns equilibrium simulations.
2.9a	80	Snapshots of EC1 monomer after 2, 5, 7 and 10 ns of molecular dynamics simulation.
2.9b	81	Snapshots of EC1 dimer after 2, 5, 7 and 10 ns of molecular dynamics simulation.
3.1a	101	The alkylation reaction of the thiol group of the Cys13 residue of EC1 with iodoacetate and iodoacetamide produces the EC1 derivatives EC1-IA and EC1-IN, respectively.
3.1b	102	MALDI-TOF spectrum of EC1 alkylated with iodoacetic acid (EC1-IA)
3.1c	103	MALDI-TOF spectrum of EC1 alkylated with iodoacetamide (EC1-IN)
3.1d	104	The effect of pH on the secondary structure of EC1-IN as determined by CD spectroscopy.
3.2a	106	The chemical stability profiles of EC1-IN after incubation for 4 h at pH 3.0, 7.0, and 9.0 at 37°C. At 37 °C, EC1-IN does not undergo any chemical degradation or precipitation at pH 3.0, 7.0, and 9.0.
3.2b	107	The chemical stability profiles of EC1-IN after incubation for 4 h at pH 3.0, 7.0 and 9.0 at 70 °C. EC1-IN shows no peptide bond hydrolysis or precipitation at pH 7.0 and 9.0. At pH 3.0, however, EC1-IN undergoes peptide bond hydrolysis as well as precipitation. The product of hydrolysis was identified as the N-terminal fragment G1-D93.
3.3a	109	Comparison of the CD spectra of EC1-IN and EC1 after their incubation at pH 3.0 for 14 days at 4 °C. The CD spectrum of EC1-IN does not change substantially after 14 days, whereas that of EC1 does.
3.3b	110	Comparison of the thermal unfolding curves of EC1-IN and EC1 proteins after incubation for 14

days. There is no change in the thermal unfolding curves of EC1-IN whereas a clear change is seen in the thermal unfolding curve of EC1. The inset shows that the thermal unfolding curves of EC1-IN can be fitted to sigmoidal functions.

3.4a 112 The CD spectra of EC1-IN after its incubation at pH 7.0 for 14 and 28 days at 4 °C. The CD spectrum of EC1-IN does not change substantially after 14 and 28 days.

3.4b 113 The CD spectra of EC1 after its incubation at pH 7.0 for 14 and 28 days at 4 °C. The CD spectrum of EC1 changes substantially after 14 and 28 days, unlike that of EC1-IN shown in Figure 3.4a.

3.4c 114 Thermal unfolding profile of EC1 and EC1-IN measured by the CD signal at 218 nm after incubation for 0, 14 and 28 days at 4 °C at pH 7.0. There are limited differences in the thermal unfolding curves of EC1-IN, but dramatic differences are seen in the thermal unfolding curves of the EC1. EC1 has a higher transition temperature than EC1-IN, indicating that EC1-IN is less stable to thermal unfolding than EC1.

3.5a 116 The CD spectra of EC1-IN after its incubation at pH 9.0 for 14 and 28 days at 4 °C. The CD spectrum of EC1-IN after 14 and 28 days is substantially different from that on day 0.

3.5b 117 Thermal unfolding profile of EC1 and EC1-IN measured by the CD signal at 218 nm after incubation for 0, 14 and 28 days at 4 °C at pH 9.0. The thermal unfolding curves of EC1-IN are substantially different from those of EC1.The inset shows that the thermal unfolding of EC1-IN can be fit to sigmoidal functions.

3.6a 119 Comparison of the intrinsic fluorescence emission spectra of EC1 and EC1-IN (after excitation at 295 nm) after incubation at 4°C for 28 days at pH 7.0. A

		red shift of 4.33 nm in the wavelength of maximum emission is observed for EC1; however, such a shift is not observed for EC1-IN.
3.6b	129	Thermal unfolding of EC1 and EC1-IN evaluated by intrinsic fluorescence emission after incubation for 0 and 28 days at 4 °C and pH 7.0. EC1-IN has a lower transition temperature than EC1. After 28 days incubation at 4 °C, EC1 shows a transition to wavelengths > 375 nm at higher temperatures, while EC1-IN does not show such a transition. The inset shows the same data, magnified by plotting a smaller wavelength range on the y-axis.
4.1a	141	Reaction of EC1 with maleimide-PEG.
4.1b	142	MALDI-TOF MS analysis of the PEGylation reaction mixture.
4.1c	143	Separation of PEGylated EC1, unreacted EC1 and maleimide-PEG using size-exclusion chromatography.
4.1d	144	HPLC profiles of EC1 (panel A, retention time = 8.4 min), M-PEG (panel B, retention time = 7.4 min), and PEGEC1 (panel C, retention time = 8.1 min)
4.2	146	Chemical stability studies of PEGEC1 at pH 3.0, 7.0, and 9.0 at 70 °C for 4 h incubation. The results for the stability of PEGEC1 are compared with those for the stability of EC1 and EC1-IN obtained from Chapters 2 and 3.
4.3a	148	Far UV CD spectrum of PEGEC1 compared to that of EC1 at pH 3.0. There are some differences between the two spectra.
4.3b	149	Far UV CD spectrum of PEGEC1 compared to that of EC1 at pH 7.0. The two spectra are very similar.
4.3c	150	Comparison of the far UV CD spectra of PEGEC1 and EC1 at pH 9.0. shows that they are very similar.

4.4a	152	Thermal unfolding profile of PEGEC1 compared to that of EC1 at pH 3.0 obtained by plotting the CD signal at 218 nm against the corresponding temperature.
4.4b	153	Thermal unfolding profile of PEGEC1 compared to that of EC1 at pH 7.0 obtained by plotting the CD signal at 218 nm against the corresponding temperature.
4.4c	154	Thermal unfolding profile of PEGEC1 compared to that of EC1 at pH 9.0 obtained by plotting the CD signal at 218 nm against the corresponding temperature.
4.4d	155	A comparison of the unfolding temperatures (T_ms) for PEGEC1 and EC1 at pH 3.0, 7.0, and 9.0 after fitting the CD thermal unfolding profiles in 4.4a, 4.4b, and 4.4c, to sigmoidal functions.
4.5a	157	Tertiary structure studies of PEGEC1 using intrinsic fluorescence emission spectroscopy. The emission spectra of PEGEC1 are compared to those of EC1 at pH 3.0. PEGEC1 and EC1 show similar tertiary structures.
4.5b	158	Tertiary structure studies of PEGEC1 using intrinsic fluorescence emission spectroscopy. The emission spectra of PEGEC1 are compared to those of EC1 at pH 7.0. PEGEC1 and EC1 show similar tertiary structures.
4.5c	159	Tertiary structure studies of PEGEC1 using intrinsic fluorescence emission spectroscopy. The emission spectra of PEGEC1 are compared to those of EC1 at pH 9.0. PEGEC1 and EC1 show similar tertiary structures.
4.6a	161	Thermal unfolding profiles of PEGEC1 compared to those of EC1 at pH 3.0 obtained by plotting the wavelength of maximum emission after excitation at 295 nm against the corresponding temperature.

4.6b	162	Thermal unfolding profiles of PEGEC1 compared to those of EC1 at pH 7.0 obtained by plotting the wavelength of maximum emission after excitation at 295 nm against the corresponding temperature.
4.6c	163	Thermal unfolding profiles of PEGEC1 compared to those of EC1 at pH 7.0 obtained by plotting the wavelength of maximum emission after excitation at 295 nm against the corresponding temperature.

Chapter 1

Introduction

1.1. Introduction

The use of proteins for therapeutic applications has increased dramatically in the last several decades. There has been a tremendous increase in the number of approved drugs derived from recombinant proteins and monoclonal antibodies since the approval of recombinant insulin in 1982 (Figure 1.1).[1] In addition, the number of approved indications is almost four times the number of approved drugs in 2005. Sales of protein drugs in 2008 are estimated to be around 71 billion USD.[2] Thus, many proteins are currently being investigated for use as therapeutic agents for different diseases (*i.e.*, cancer, autoimmune diseases). Most of the approved protein drugs that are commercially available are administered parenterally. This is due to the fact that proteins have very low bioavailability when administered orally. The low bioavailability of orally administered proteins is partly due to their chemical breakdown under the highly acidic conditions in the stomach and the mildly acidic to mildly basic conditions in the small intestine, and partly due to their very low absorption across the intestinal epithelium. The transcellular pathway is not a viable option for the absorption of proteins across the intestinal epithelia into the blood stream due to the highly charged nature of proteins. The absorption of proteins through the paracellular pathway, which utilizes the intercellular space for permeation, is also limited due to the presence of intercellular junctions. These junctions are formed through the interactions of various transmembrane proteins. One

such protein is E-cadherin. The C-terminal region of E-cadherins is located in the cytoplasm, whereas the N-terminal region is located in the extracellular matrix. This

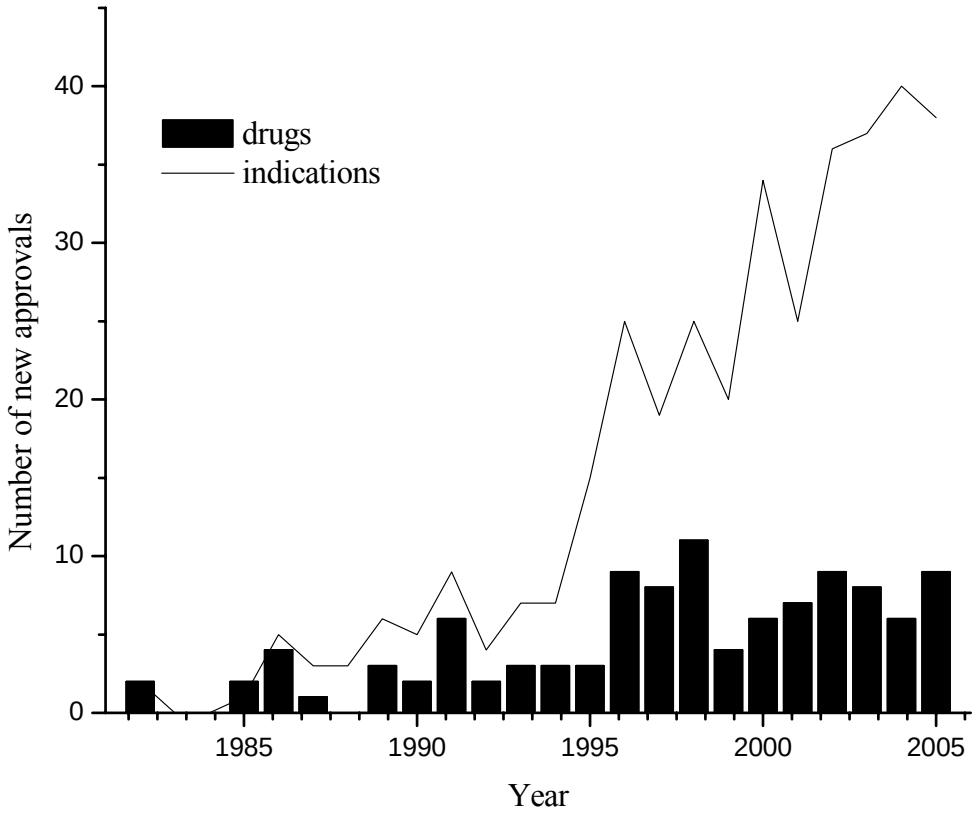

Figure 1.1: Increasing trend in the use of recombinant proteins and monoclonal antibodies as drugs. The number of new biotech drugs and new indications approved for biotech drugs from 1982 to 2005.

extracellular portion of E-cadherins is composed of five sub-domains, called EC1 through EC5. EC1 is the most N-terminal sub-domain and located farthest away from the cell surface. The extracellular portions of E-cadherins from the same cell surface interact to form *cis*-dimers; *cis*-dimers from opposite cell surfaces interact to produce *trans*-oligomers,[3-5] resulting in cell-cell adhesion. The EC1 subdomain of E-cadherins has been shown to be essential for the homophilic selectivity of E-cadherins.[6-12] Therefore, in theory, EC1 can be used to modulate the E-cadherin interactions in the intercellular junctions, thereby improving paracellular transport of protein therapeutics. Thus, our long-term goal is to study the possibility of employing EC1 as a modulator of E-cadherin interactions. In fact, peptides with sequences derived from EC1 have been shown to increase the transport of mannitol, which is a paracellular marker.[13] We will like to conduct *in vitro* studies in cultured cell monolayers to study the activity of EC1.

For this purpose, recombinant human EC1 was expressed in our laboratory.[14] The yield of the expression system was found to be low (about 2 mg EC1/L of fermentation medium used). In solution, EC1 was shown to produce covalent dimers by the formation of intermolecular disulfide bonds between the Cys9 residues of EC1 monomers.[15] Furthermore, the formation of non-covalent oligomers was also observed. Interestingly, it was found that when dithiothreitol (DTT), a thiol reducing agent, was added to the solution, the covalent dimers as well as the physical

oligomers converted to monomers of EC1. Therefore, it was concluded that formation of disulfide-linked dimers of EC1 was essential to the formation of physical oligomers of EC1. Thus, the EC1 solution, in the absence of a reducing agent, was a heterogenous mixture of monomers, covalent dimers and physical oligomers. It would be undesirable to have these different forms of EC1 in solution when performing biological activity assays of EC1, because one or more of these forms of EC1 could be inactive or worse, have a detrimental effect on the activity assay. Consequently, we decided to survey the literature for different approaches applied to overcome the problem of disulfide formation. In the following sections, we have assimilated the relevant information on disulfides and thiols in proteins. We describe how protein disulfide bonds are produced in prokaryotic and eukaryotic cells, followed by how oxidative folding of reduced denatured proteins can be achieved *in vitro*, the effect of disulfide bonds on protein stability, various qualitative and quantitative methods used to assay protein thiols and disulfides, chemical degradation reactions of protein disulfides in solution and finally, various methods that can be used to modify protein thiols so as to block their oxidation to disulfides.

1.2. Disulfide bond formation as a post-translational modification of proteins

Native disulfide bonds in proteins are assembled during the oxidative folding process. In eukaryotes, this process occurs in the endoplasmic reticulum (ER) lumen,[16] whereas in prokaryotes it takes place in the periplasm.[17] Thus, disulfide bonds are usually formed only in specific cellular compartments in healthy cells.

Proteins synthesized by ribosomes are in reduced and unfolded form, and their oxidative refolding in eukaryotes is aided by a series of enzymes such as the endoplasmic reticulum oxireductin-1 (Ero1), protein disulfide isomerases (PDI) and Erv2.[16,18-20] Many of these enzymes possess C-X-X-C (*i.e.*, PDI) and C-X-C (*i.e.*, Erv2) motifs that are involved in thiol-disulfide redox exchange reactions.[21] PDI directly oxidizes thiol groups in a protein via a series of thiol-disulfide exchange reactions by partnering with Ero1 or Erv2. The Ero1 enzyme, which is found inside the ER and is associated with the ER membrane, oxidizes PDI to form an Ero1-PDI mixed disulfide form. If the disulfide bond in the protein substrate is incorrectly formed, PDI catalyzes the reduction of these disulfide bonds and, subsequently, re-oxidizes the protein substrate to form the native disulfide bond. Thus, as a general rule, cytoplasmic proteins contain free thiols, while proteins in other compartments such as the ER can possess disulfide bonds. An exception to this rule is a family of thermophilic organisms which has been discovered to produce intracellular proteins with disulfide bonds. This is presumably due to the presence of protein disulfide oxidoreductase (PDO) in their cytoplasm, which oxidizes the cysteines to disulfides within intracellular proteins.[22]

In bacterial cells, the oxidative folding machinery parallels the machinery in mammalian cells, and the oxidation reaction is catalyzed by the enzymes DsbA, DsbB, DsbC and DsbD.[17] The disulfide bond A (DsbA) enzyme oxidizes thiol groups of the substrate protein to a disulfide bond in the periplasm with the simultaneous reduction of its own disulfide bond. Then, the reduced DsbA is reoxidized by DsbB.

The oxidized state of DsbB is maintained by transferring electrons to ubiquinones (benzoquinone derivatives) and menaquinones (naphthoquinone derivatives), components of the respiratory electron transport chain system. When the substrate protein forms a non-native disulfide bond, DsbC enzyme reduces the non-native disulfides. This is followed by the subsequent re-oxidation of the substrate to form the native disulfide by DsbA or DsbC. Finally, DsbD functions to keep DsbC in the reduced state, and the oxidized DsbD can be reduced by cytoplasmic thioredoxin.

1.3. Pathways of oxidative folding of proteins for disulfide formation

Prior to the formation of a disulfide bond, the protein may or may not be folded into a native-like structure with the Cys residues in a favorable position to form a disulfide bond.[23,24] In the case of bovine pancreatic trypsin inhibitor (BPTI), the formation of all three disulfide bonds in BPTI occurs via stable intermediates with a native conformation containing one or two native disulfide bonds.[23,24] No significant amount of non-native disulfide intermediates was observed. However, the formation of only native-like intermediates does not always predominate; for example, hirudin, with three disulfide bonds, undergoes oxidative folding via heterogeneous intermediates with two or three non-native disulfide bonds.[25] These scrambled disulfide bonds undergo reduction followed by oxidative rearrangement to form the native protein structure. The reductive unfolding of this protein follows a pathway similar to that of the folding process, but in reverse. The *in vitro* oxidative folding of human macrophage colony stimulating factor β (rhm-CSFβ) follows a pathway

similar to that of hirudin. The rhm-CSFβ protein is a dimer with three intermolecular disulfide bonds between the two monomers. Each monomer also contains three intramolecular disulfide bonds.[26] The oxidative folding of rhm-CSFβ involves the formation of monomeric isomers with native and non-native intramolecular disulfide bonds; the monomer isomerization step is the rate-determining step in this process.[26] Upon disulfide shuffling, the protein forms a stable dimeric intermediate that possesses all of the intramolecular disulfides, including a non-native intermolecular disulfide bond. Rearrangement of the non-native disulfide bond creates the native oxidized rhm-CSFβ.

Although the oxidative folding of different proteins may take different pathways, the process normally adheres to some general rules. Disulfide bond formation is usually favored at basic pH, and the presence of an oxidizing agent such as oxidized glutathione increases the rate of oxidative folding, whereas the presence of denaturants such as urea or guanidine hydrochloride (Gdn.HCl) hampers the folding process. These rules do not apply, however, to the oxidative folding of murine prion protein (mPrP 23–231), which has only one intramolecular disulfide bond at Cys179–Cys214.[27] The formation of a disulfide bond in reduced mPrP is very slow at pH 8.0 in the absence of denaturant, but mPrP refolds properly in the presence of a denaturant and glutathione at pH 8.0.[27] Contrary to intuition, the best pH condition to form a disulfide in mPrP is 4–5. This observation can be explained by the fact that, at alkaline and neutral pH, the reduced and unfolded mPrP is present in a stable intermediate conformation with its Cys residues buried and separated from each

other. The addition of glutathione, therefore, does not facilitate oxidation of Cys. The addition of denaturant and a pH of 4.0 facilitate the unfolding of the stable intermediate, bringing the two Cys residues closer and allowing them to form a disulfide bond.

1.4. Effects of disulfide bond on protein stability

Intramolecular disulfide bonds are known to contribute to protein thermal stability. Reducing all five disulfide bonds inactivates the *Aspergillus niger* phytase due to conformational change and/or unfolding.[28] Another example of the effect of disulfide bonds on stability is observed in the soybean Bowman-Birk inhibitor (BBI). The crystal structure of BBI has a bow-tie motif, with a trypsin-binding loop at one end and a chymotrypsin-binding loop at the other. In contrast to most globular proteins, it has exposed hydrophobic patches and charged residues containing bound water molecules in its interior.[29,30] Interestingly, BBI is remarkably stable against heat and chemical denaturants; this stability can be attributed to the seven intramolecular disulfide bonds within the protein's structure that "lock" it in its native conformation.[29]

Introduction of a disulfide bond by Cys mutation has been shown to improve the physical stability of some proteins. In case of Subtilisin E, mutations of Gly61Cys and Ser98Cys followed by the formation of the Cys61–Cys98 disulfide produce an active enzyme with a 4.5 °C increase in melting temperature and a half-life three times longer than that of the native protein.[31] In addition, there was no change in its

enzymatic activity due to the engineered disulfide bond. Similarly, a Cys39–Cys85 disulfide bond was engineered into dihydrofolate reductase (DHFR) without a significant alteration in its conformation. The disulfide mutant had an improved stability with its ΔG° higher than that of the native protein by 1.8 kcal/mol.[32]

At this point, it is necessary to discuss the thermodynamics behind the stabilizing effect of disulfide bonds on proteins. This stabilizing effect of disulfides is manifested by an increase in the melting point of the protein. The melting or unfolding transition of a protein is generally regarded as a reversible process between the native and the unfolded states. Differential scanning calorimetry (DSC) is usually employed to determine the reversible thermal unfolding of a protein. Superimposable DSC themorgrams resulting from repeated scans of the same protein sample imply a two-state reversible unfolding process, which is depicted as $\boldsymbol{N \rightleftharpoons U}$ **,** where ***N*** represents the native state and ***U***, the unfolded state of the protein. This is also known as the "all or none" transition in statistical thermodynamics. In such a transition, the protein is either in the ***N*** state, with all its residues in the native (**n)** conformation or it is in the ***U*** state, where all its residues are in the unfolded (**u)** conformation. The Gibbs free energy for this process is related to the equilibrium constant of this reaction by the following equation: $\Delta G = -RT \ln K_{eq}$. K_{eq} is the ratio of the folding (k_f) and unfolding (k_u) rate constants. So, $K_{eq} = k_f/k_u$. When a disulfide bond is introduced into a protein, there is likely to be a significant increase in the free energy of ***U***. The transition state (***TS***) can either (i) have a large increase in its free energy similar to the unfolded state, in which case, the disulfide bond is lost in the ***TS*** or (ii)

have a negligible change in its free energy, in this case, ***TS*** is more ***N***-like and retains the disulfide bond.[33] In (i), the folding rate does not change much, but the unfolding rate decreases significantly. In (ii), the folding rate increases while the unfolding rate remains constant. In either case, the ratio of k_f/k_u, which is equal to K_{eq}, decreases, leading to an increase in the free energy (ΔG) of the protein and making it conformationally more stable. The assumption is that there is little or no difference in the enthalpies (ΔH) of the disulfide-containing and reduced protein conformations. It has been proposed that the decrease in the entropy of ***U*** can be predicted with the help of the following equation: $\Delta S = -2.1 - (3/2)R \ln n$, where n is the number of residues in the loop formed by the intramolecular disulfide bond.[34] Similarly, reduction of disulfide bonds or mutation of disulfide-forming Cys residues in a protein can result in increased entropy of ***U***. Consequently, K_{eq} increases and ΔG of the protein decreases. This characteristic was observed in the 110 residue starch-binding domain of *Aspergillus niger* glucoamylase. The wild-type protein (SBD) possesses a disulfide bond between C3 and C98, creating a 94 residue loop between its N-terminus and the short loop between its seventh and eighth β-strands. Two different mutants C3G/C98G (GG) and C3S/C98S (SS) were constructed and their stability compared to that of SBD by determining their unfolding temperature and other thermodynamic parameters using DSC and CD techniques.[35] The unfolding temperatures were 10 °C lower for the mutants (43 °C) compared to wild-type SBD (53 °C). The unfolding processes for SBD, GG and SS were determined to be reversible with GG and SS having a free energy lower than SBD by about 10 kJ/mol. The contribution from

entropic destabilization was much larger compared to that of the enthalpic destabilization.

When multiple disulfide bonds are introduced into a protein, they can act cooperatively to stabilize the protein. The overall increase in the free energy of a protein in that case is higher than the sum of the increases in free energies due to individual disulfide pairs. This phenomenon is observed in the *Bacillus circulans* xylanase. When disulfide bonds were introduced into it by the following mutations: S110C/N148C (DS1) and A1GC/G187,C188 (cX1), the overall increase in its free energy was 5.4 kcal/mol, which is higher than the sum of the free energy increases (4.6 kcal/mol) due to each separate mutation by 0.8 kcal/mol. Consequently, the melting temperature of this double mutant was 12.4 °C higher than the native protein, compared to the 5.0 °C and 3.8 °C increase in melting points in the DS1 and cX1 mutants.[36]

It is difficult to predict whether increasing conformational rigidity upon addition of a disulfide bond will always improve the stability and/or activity of a protein. A good example that illustrates the effect of the location of a disulfide on structural stability and enzymatic activity was shown in T4 lysozyme. Four mutants of T4 lysozyme with a disulfide bond at Cys9–Cys164, Cys21–Cys142, Cys90–Cys122, and Cys127–Cys154 were designed based on molecular modeling experiments.[37] The thermal stabilities of the oxidized mutants were compared to their corresponding reduced mutants as well as to wild-type T4 lysozyme. The oxidized Cys9–Cys164 mutant has a melting temperature 6.4 °C higher than that of the wild type while

maintaining its enzymatic activity. In contrast, the Cys21–Cys142 mutant has the best thermal stability but has no enzymatic activity. Finally, the Cys90–Cys122 and Cys127–Cys154 mutants have lower thermal stabilities and enzymatic activities compared to that of the wild-type enzyme. In another case, Ser52, Ser53, Ser78, Ser79, and Ser80 in alkaline phosphatase (AP) were mutated to the respective Cys residue; a disulfide bond was formed between the new Cys residue and Cys67. All of the disulfide mutants had improved thermal stability but lower enzymatic activities than that of the native AP.[38] The decreased enzymatic activity of the mutant appeared to be due to the enhanced rigidity and lower substrate affinity of the active site imposed by the disulfide bond, which is located near the residues. Thus, the location of the non-native disulfide bond affects the physicochemical stability and biological properties of the mutant proteins in a manner difficult to predict.

To impose conformational rigidity, a disulfide bond also has been introduced in peptides to form cyclic structures (cyclic peptides), which stabilizes specific secondary structures (*i.e.*, β- and γ-turns). The formation of cyclic peptides in some cases enhances receptor-binding affinity and selectivity, as shown in opioid peptides.[39] Several cyclic peptides with a disulfide bond have been marketed as therapeutic agents, including Integrilin®[40,41] and oxytocin (Pitocin®).[42] Integrilin® is a cyclic Arg-Gly-Asp (RGD) peptide, which is used to treat thrombosis; it binds selectively to gpIIb/IIa receptors on the surface of platelets, thus blocking fibrinogen-mediated platelet aggregation.[40,41] Cyclic peptides also have been shown to be more stable to chemical and enzymatic degradation than their parent linear analogues. The

Asp-mediated chemical degradation of RGD peptides was suppressed by the formation of a cyclic RGD peptide with a disufide bond; thus, the chemical stability of the cyclic RGD peptide is higher than that of the parent linear form.[43-45]

Earlier, we discussed the role of disulfide bonds in improving protein stability. Disulfides have also been found to improve stability of antibodies. Introduction of non-native disulfide bonds, however, can also lead to a decreased protein stability. Such is the case with SBD. Previousy, we described the increase in stability of SBD by the creation its two mutants: GG and SS. When two new mutants, C3S and C3G were created, they both showed a 9 °C decrease in the melting point of wild-type SBD. The explanation for this decreased stability was that the remaining unmutated C98 of SBD was able to form disulfide dimers in the native state.[46] Thus, disulfide bonds can have a favorable or an unfavorable effect on protein stability.

1.5. Analysis of free thiols and disulfide bonds

The presence of a free thiol group in the Cys residues in a protein can be quantitatively determined by different methods, including the use of maleimide (*i.e.*, ThioGlo),[47] Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB),[48-51] and bimane reagents (Figures 1.2a and 1.2b).[52,53] The presence of a disulfide bond can be detected using the crabescein reagent.[54] Ellman's reagent (DTNB) was the earliest reagent widely used for estimating the number of thiol groups. The reaction between DTNB and thiol group(s) produces an equivalent amount of 5-thio-2-nitrobenzoic acid, the absorbance of which can be measured at 412 nm. It was found that the

DTNB method may underestimate the concentration of thiol groups in a protein due to its incomplete reaction.[47,55] Addition of cytamine as an accelerator of the DNTB-thiol reaction improves the accuracy of thiol determination (Figure 1.2a).[51] As an alternative reagent, 4,4'-dithiodipyridine (DTDP) has a performance similar to that of the DTNB/cystamine in determining thiol concentrations (Figure 1.2a).[51] Upon reaction with the thiol group, DTDP releases 4-thiopyridine (4-TP), which absorbs at 342 nm. One drawback to using DTDP is that some proteins have inherent absorbance at 342 nm (due to the ligand metal charge transfer band during the chelation of thiol groups with some metals from the *f*-block of the periodic table), which can interfere with the quantitative measurement of 4-TP. A new derivative of Ellman's reagent, n-octyl-5-dithio-2-nitrobenzoic acid (ODNB) has been designed; this reagent can react with all thiol groups of a protein, including the buried ones (as it is more hydrophobic than DTNB), in a relatively short time (Figure 1.2a).[55]

Maleimide derivatives are also used to determine the number of thiol groups in a protein. ThioGlo® reagents (ThioGlo®1, λ_{Ex}= 379 nm and λ_{Em} = 513 nm) are maleimide derivatives of naphthopyranone fluorophores that emit fluorescence upon thiol alkylation (Figure 1.2a). ThioGlo®1 is forty times more sensitive than DTNB and other maleimide reagents. The protein-ThioGlo adduct does not require separation prior to quantitative determination due to the low fluorescence of the starting reagent. One caution is that the reaction should not be carried out at a higher pH (>8) because of the possibility of amine alkylation of the maleimide and adduct degradation. Possible disadvantages include the low water solubility of the ThioGlo

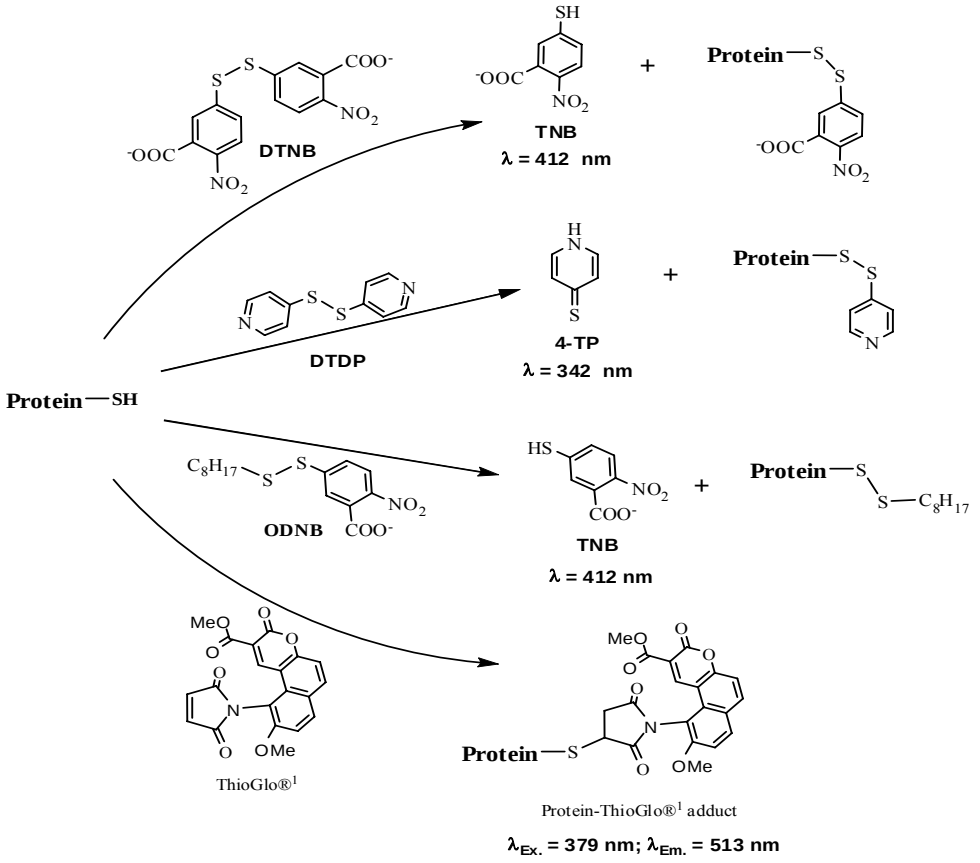

Figure 1.2a: Various labeling reagents for the detection and quantification of protein thiols and disulfides: 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB or Ellman's reagent), 4,4'-dithiodipyridine (DTDP), n-octyl-5-dithio-2-nitrobenzoic acid (ODNB), and ThioGlo®1.

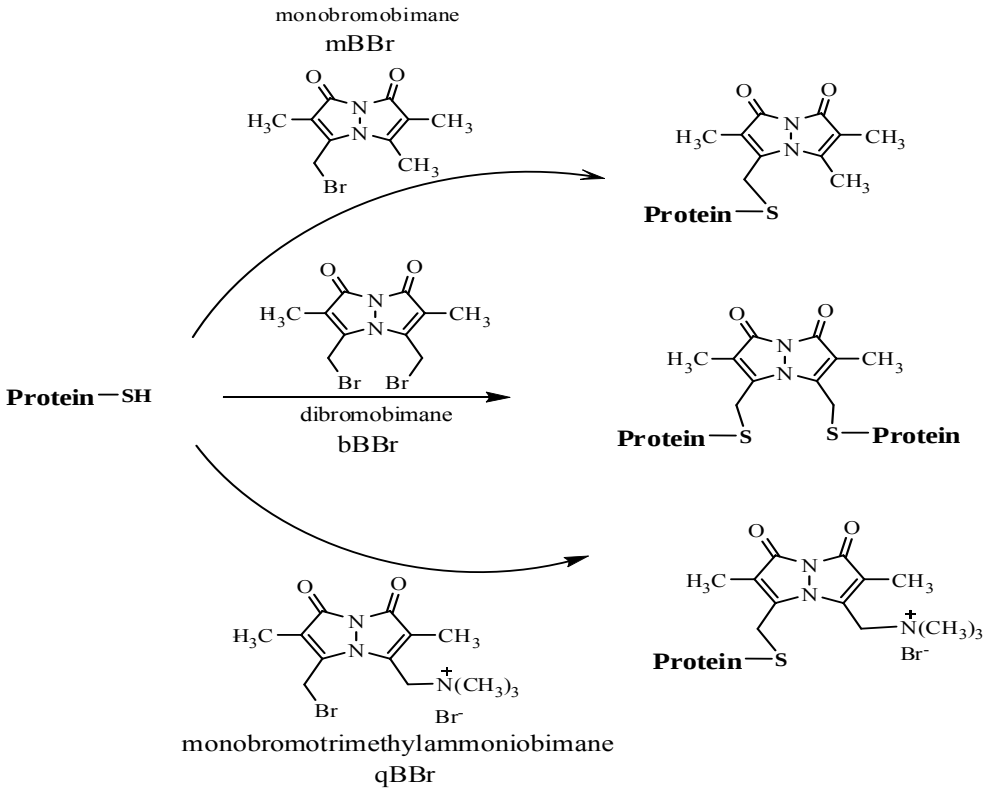

Figure 1.2b: Labeling cellular protein thiols with bimanes: monobromobimane (mBBr) and dibromobimane (bBBr) label thiols inside cells as well as on cell surfaces, whereas monobromotrimethylammoniobimane (qBBr) labels only the cell surface protein thiols.

reagent and the instability of protein-ThioGlo adducts, which may lead to inaccuracy in the quantitative determination of the thiol group.

Bimane derivatives (*i.e.*, monobromobimane (mBBr), dibromobimane (bBBr) and monobromotrimethylammoniobimane (qBBr)) also have been developed for quantitative determination of thiol groups of cellular proteins (Figure 1.2b). Bimane derivatives have no intrinsic fluorescence; upon reaction with thiol groups in proteins, they become highly fluorescent.[52] Both mBBr and bBBr can be used to measure the total cellular thiol content and qBBr can be used to measure the extracellular and membrane thiol content. Due to the presence of a charged group in qBBr, this molecule cannot permeate cell membranes, and it, therefore, may be used to measure selectively the extracellular thiols. mBBr can penetrate cell membranes and react with thiols in the cells as well as the thiols in the extracellular space. Thus, the difference in the thiol content measured by mBBr and qBBr gives the amount of thiols inside the cells. One of the advantages of using these bimane derivatives is that they are stable during exposure to air and irradiation as well as during general laboratory procedures. Bimane has also been used to determine the intracellular levels of thioredoxin after isolating the modified protein using an antibody affinity Sepharose column. After elution from the column, the amount of thioredoxin can be determined using a fluorescence spectrometer.[56]

1.6. Degradation of disulfide bonds

Chemical degradation of disulfide bonds has been studied in both peptides and proteins with most of the disulfide bond degradation observed under neutral and basic conditions (Figures 1.3a, 1.3b and 1.3c).[43,57] Degradation of disulfide bonds can be classified into three major pathways: (1) direct attack on the sulfur atom by the hydroxyl anion, opening the disulfide bond to form sulfenic acid/thiolate anion (Figure 1.3a); (2) A β-elimination reaction in which the α-proton of the Cys residue is abstracted (which produces dehydroalanine/persulfide (Figure 1.3b)); and (3) the α-elimination caused by the hydroxyl ion attacking the β-proton of the Cys residue to produce the thiolate/thioaldehyde (Figure 1.3c).[58]

The hydroxyl attack of the disulfide bond produces sufenic acid and thiolate anion (**1**, Figure 1.3a). The thiolate anion **1** can proceed to attack another disulfide bond on another protein molecule to produce a thiolate dimer intermediate **2** (pathway **a**, Figure 1.3a). This intermediate (**2**) undergoes three possible reactions as shown in pathways **b**, **c**, and **d**. In pathway **b**, intramolecular nucleophilic attack of the sulfenic acid by the thiolate anion produces dimer **3** with two intermolecular disulfide bonds. Intermediate **2** can form sulfenic acid dimer **4** by abstracting a proton from water (pathway **c**). Higher molecular weight oligomers such as trimer **5** can be produced via intermolecular disulfide bond reactions utilizing pathway **d**. The thiolate anion **1** can also undergo an intermolecular reaction between the sulfur atom of the sufinic acid group and a disulfide in another molecule to give an intermediate dimer **6** as shown in

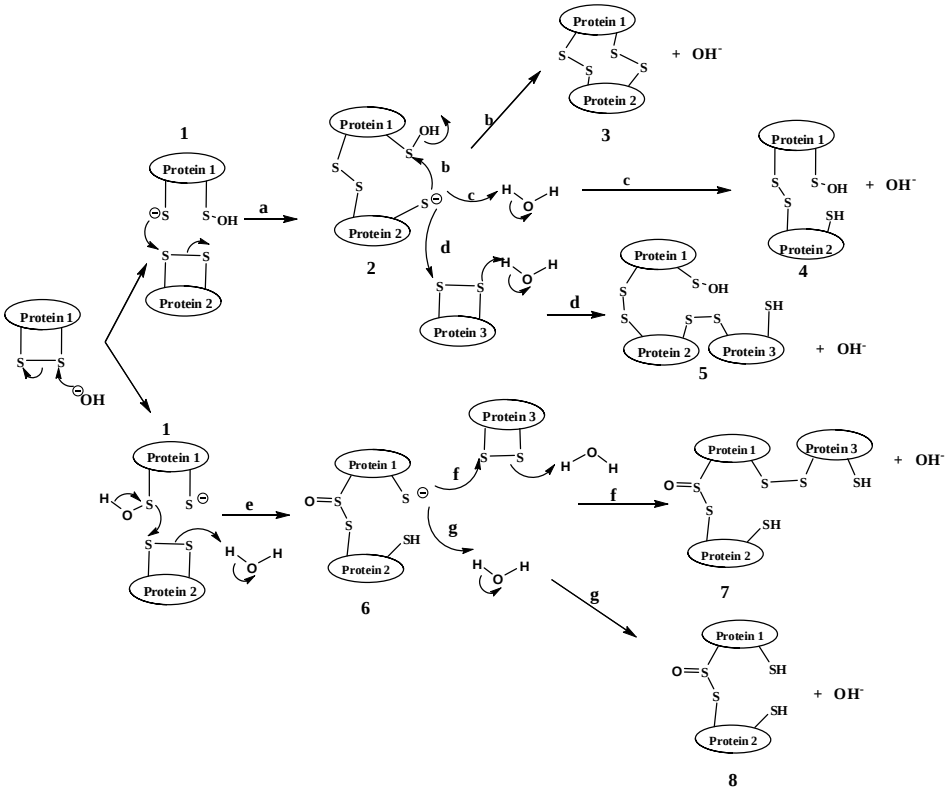

Figure 1.3a: Degradation reactions of protein disulfide bonds in neutral and basic conditions that involve a direct attack on the sulfur atom by hydroxyl anions.

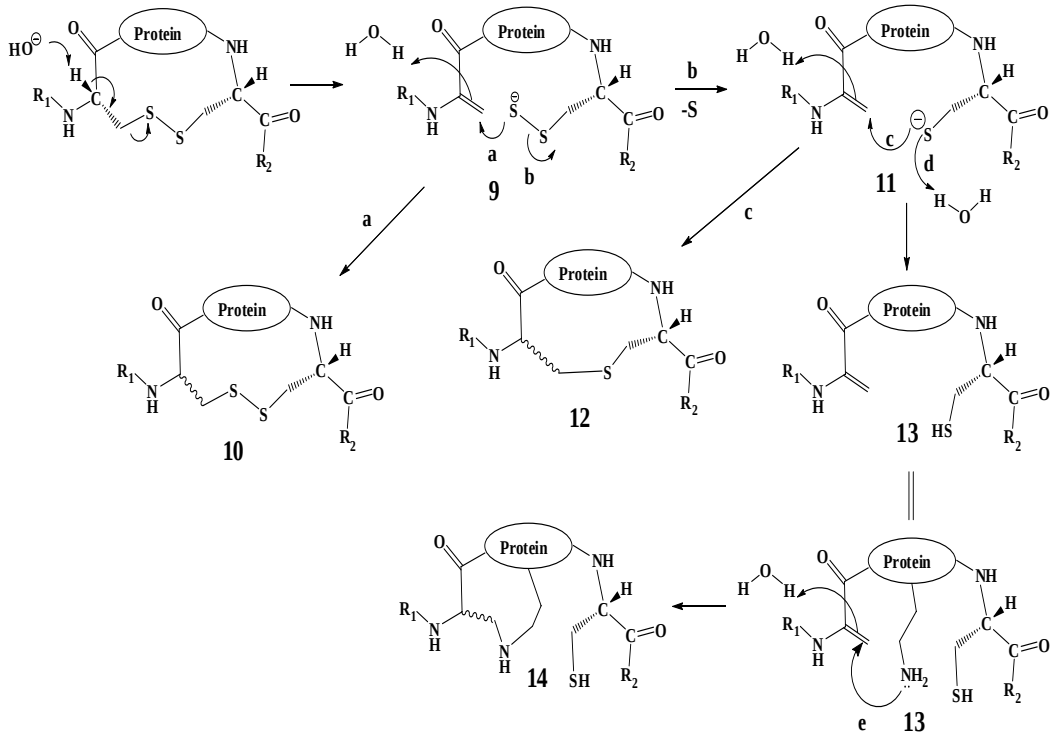

Figure 1.3b: Degradation reactions of protein disulfide bonds in neutral and basic conditions involving a β-elimination reaction. The α-proton of the Cys residue is abstracted by the hydroxyl anion.

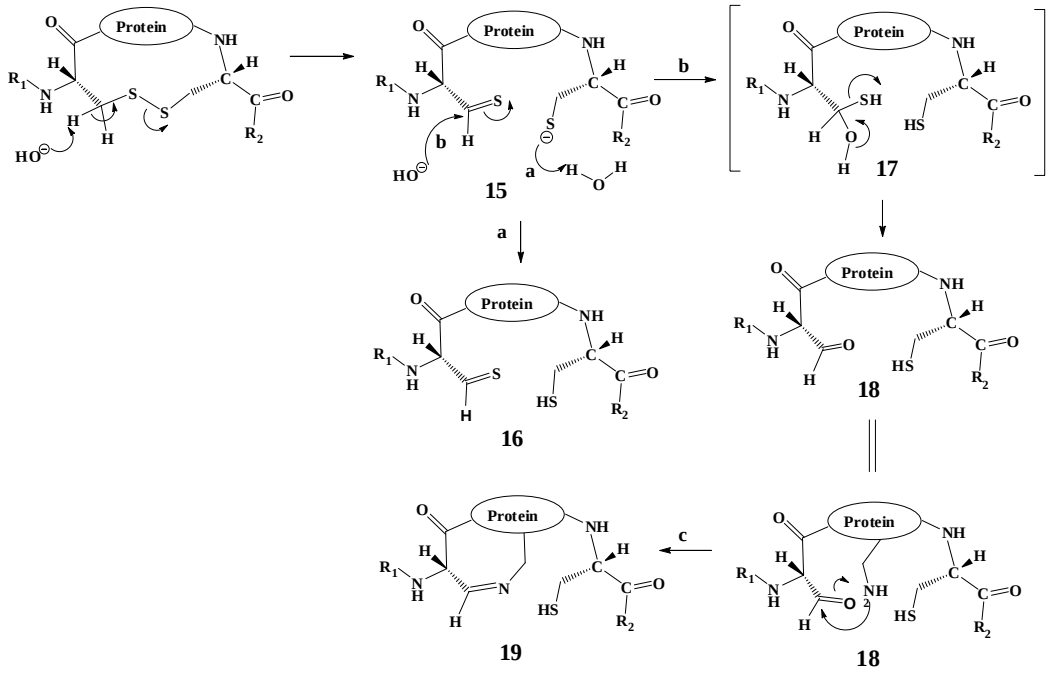

Figure 1.3c: Degradation reactions of protein disulfide bonds in neutral and basic conditions involving α-elimination reaction. The β-proton of the Cys residue is abstracted by the hydroxyl anion to produce a thiolate/thioaldehyde (15).

pathway **e**. The resulting thiolate **6** oligomerizes by reacting with another protein molecule to generate a trimer **7** (pathway **f**); further intermolecular reactions lead to higher molecular weight oligomers. Intermediate **6** also can form a dimer molecule with two free thiols (**8**).

The second degradation pathway is via β-elimination through the hydroxyl anion abstraction of the acidic Cα proton of Cys residue to yield a dehydroalanine-persulfide **9** (Figure 1.3b). The disulfide bond can be reconnected by nucleophilic attack on the dehydroalanine by the persulfide anion (pathway **a**, Figure 1.3b). Theresulting products (**10**) form a racemic mixture at the Cα of Cys residue (**10**, Figure 1.3b). The persulfide **9** can undergo a sulfur extrusion reaction via pathway **b** to yield the thiolate anion **11**. The thiolate anion in **11** reacts with the alkene functional group of the dehydroalanine to generate the thioether **12**; this reaction can also take place in an intermolecular fashion to produce oligomers. A similar reaction has been observed in an IgG1 monoclonal antibody[59] and cyclic peptides containing a disulfide bond.[43,57] In the IgG1, this thioether bond was observed at C233, which links the heavy chain to the light chain.[59] This modification was identified by LC-MS tandem mass spectrometry. Abstraction of a proton from the solvent by the thiolate anion **11** produces dehydroalanine-thiol **13**. There is a possibility that the amino group from either the Lys or the N-terminus of the protein may react with the alkene group of the dehydroalanine to produce a secondary amine **14**. This reaction has been observed as a degradation product of cyclic peptides.[43,57]

Finally, the α-elimination reaction proceeds via the abstraction of the β-proton of the Cys residue followed by disulfide bond breakage to produce thiolate/thioaldehyde **15** (Figure 1.3c). Protonation of the thiolate anion yields thiol-aldehyde **16** (pathway **a**). In the presence of the hydroxyl anion, the thioaldehyde group can be converted to an aldehyde **18** via intermediate **17** (pathway **b**). An amino group (*i.e.*, Lys or N-terminus) that is in close proximity to the aldehyde group in **18** may also produce an imine product **19**. The imine reaction may produce bimolecular reactions to create oligomers.

Intermolecular covalent bond reactions (disulfide, thioether, amine and imine in Figures 1.3a, 1.3b and 1.3c) producing oligomers can lead to the association, aggregation and precipitation of proteins. For example, intermolecular disulfide formation causes aggregation of nascent thyroglobulin (Tg) protein. Monomeric Tg is a precursor to the synthesis of thyroid hormone. It contains more than 100 Cys residues and can produce aggregates through the formation of non-native intermolecular disulfide bonds due to the oxidizing environment in the endoplasmic reticulum of thyroid cells.[60] During the transport of Tg to the golgi complex, the aggregates of Tg undergo disulfide reduction and dissociation to form monomers with assistance from molecular chaperones. This suggests that the incidence of Cys residues in the protein as well as the redox status of its environment are important factors that contribute to its disulfide-mediated aggregation. Intermolecular disulfide bond reactions in α- and β-crystallin induced by X-ray irradiation were shown to be partly responsible for cataract formation in the rabbit eye lens. It is interesting to note

that the reaction between prednisolone and the amino group(s) of crystallin alters crystallin conformation to expose the buried thiol groups.[61,62] These exposed thiol groups then undergo oxidation to form intermolecular disulfides, which cause protein aggregation and precipitation.[61] It was made clear in a subsequent study that the crystallin that underwent conformation change on reaction with predisolone was α-crystallin.[63] In most cases, the disulfide bond can be reversed by reducing agents (*i.e.*, DTT). Intermolecular disulfide bond formation has been suggested to induce the aggregation of BSA. DTT has been shown to suppress the aggregation of BSA.[64]

1.7. Chemical modification of Cys thiols of proteins

The thiol group in Cys can undergo oxidation reactions to form sulfenic (R-S-OH), sulfinic ($R\text{-}SO_2H$), and sulfonic ($R\text{-}SO_3H$) derivatives as well as the disulfide bond. Thiol groups are one of the most reactive groups in proteins and can participate in side reactions with reagents used to manipulate other functional groups in a protein and the other solution components. The presence of a free cysteine residue may cause unwanted intramolecular disulfide scrambling, and covalent oligomerization via intermolecular disulfide formation. To prevent these reactions, thiol groups can be derivatized with acetate, acetamide, 1,3-propane sultone[65], methyl methanethiosulfonate, methoxycarbonylmethyl disulfide[66], maleimide, tetrathionates, and dinitrophenyl alkyl disulfides (DNPSSR)[67] (Figure 1.4). Human serum albumin, [Lys^8] vasopressin, bovine insulin, and bovine pancreatic ribonuclease have been alkylated with 1,3-propane sultone to increase their solubility and stability to acid

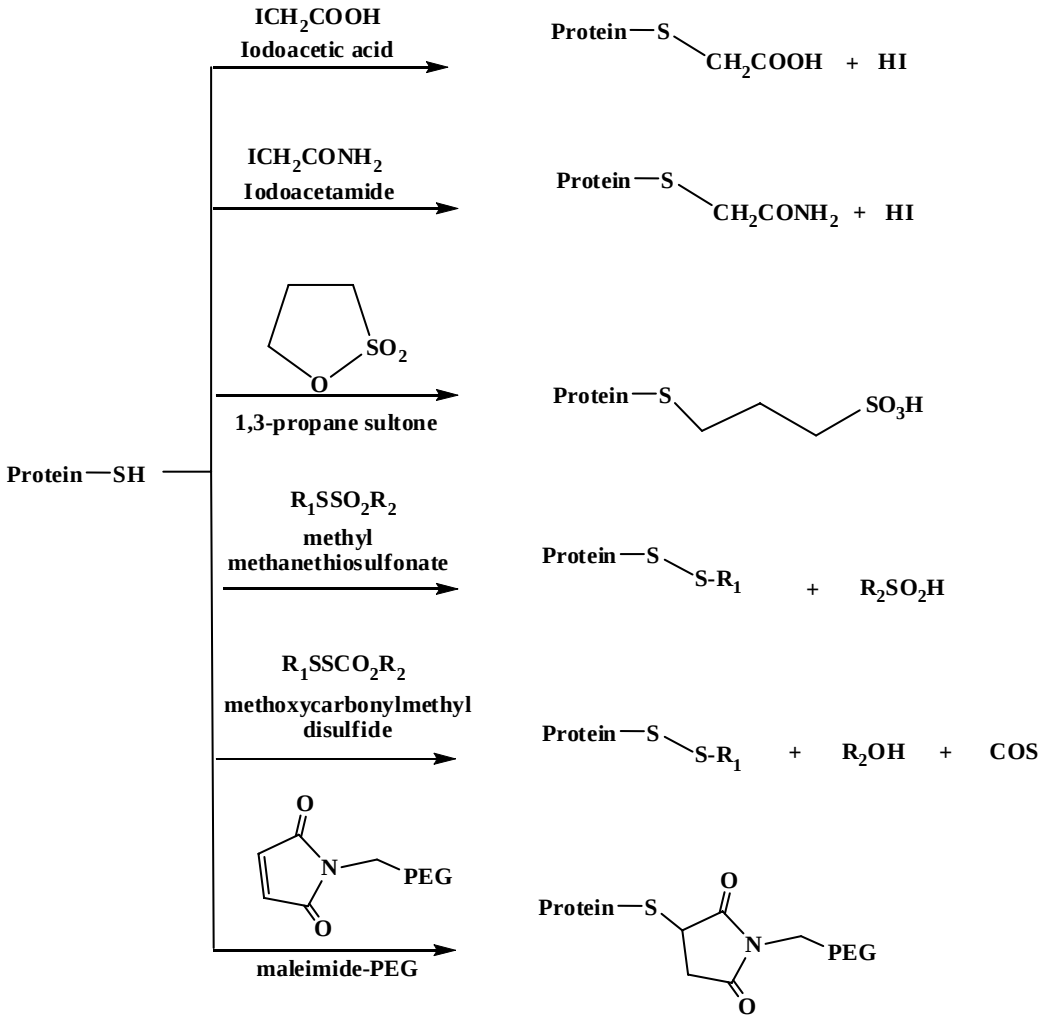

Figure 1.4: Some of the reagents used for the modification of free thiols in proteins.

hydrolysis. Alkylating the free thiol of the tetanus toxoid with N-ethylmaleimide inhibits moisture-induced aggregation of the lyophilized toxoid.[68]

To avoid side reactions of the Cys thiol group, Cys has frequently been mutated to Gly, Ala, and Ser residues in proteins. The C6A, C111S and C6S/C111A mutants of human CuZn dismutase (SOD) and the C6A mutant of bovine SOD showed greater resistance to loss of enzymatic activity than did the respective native SODs upon thermal unfolding at 70 °C.[69] This corresponds to the fact that the mutants can reversibly refold after denaturation while the unfolded native SOD refolds incorrectly. Cys mutation studies have been used to determine the role of free and disulfide-bonded Cys residues in the function of many proteins. Meprins, cell surface and secreted metalloendopeptidases, are oligomers made up of disulfide-linked multidomain subunits.

Mutations of C320A and C289A of mouse meprin-α both resulted in an increased susceptibility to proteolytic degradation and heat deactivation. Although the mutants retained their catalytic activity toward a bradykinin peptide, they displayed decreased activity toward the protein substrate azocasein. It was found that the C320A mutant only formed monomers, while the C289A mutant formed dimers and oligomers, suggesting that Cys 320 but not C289 may be involved in inter-subunit disulfide formation.[70] Mutation of either Cys residue resulted in a change in protein structure causing it to lose its enzymatic activity toward a protein substrate. Thus, Cys mutation is a technique that may not only shed light on the importance of the state of

Cys residues (paired or reduced, intra- or interprotein pairing) in a particular protein but also may help to distinguish the different roles of each Cys in the protein.

Cys residues of therapeutic proteins also have been PEGylated to improve their physicochemical stability, pharmacokinetics and pharmacodynamics properties. PEGylated proteins usually have reduced glomerular filtration and immunogenicity. The molecular weight of these PEG reagents can range from 2 to 50 kDa. Since proteins generally contain fewer unpaired Cys residues compared to Lys residues, the PEGylation of a Cys residue is more selective than PEGylation of Lys residue. PEGylation of Lys residues gives a heterogenous mixture of products with a distribution of conjugated products, which is difficult to control. The site-specific PEGylation of the Cys residue is normally carried out with PEG-maleimide, PEG-acrylamide, PEG-acrylate, PEG-vinyl sulfone, PEG-epoxide or PEG-iodoacetamide. In a protein without a free thiol, a surface residue away from the active site can be mutated to a Cys residue followed by PEGylation. This method has been successfully used to improve the half-life of granulocyte colony stimulating factor (G-CSF)[71] and the Q5C analog of interferon-α (IFN-α)[72], immunotoxin anti-Tac(Fv)-PE38 (LMB2), [73] and Eryhtropoietin (EPO).[74] When LMB2 was stabilized by PEGylation using 5– or 20– kDa PEG-maleimide, there was an increase in plasma stability (five- to eight-fold) and antitumor activity (three- to four- fold).[73] In this case, LMB2's animal toxicity and immunogenecity were significantly lower than that of the parent toxin. A native disulfide bond can also be conjugated with PEG by reaction with PEG-monosulfone; the product forms a three carbon bridge with the two sulfur atoms from

the native disulfide bond. Interferon α-2b, anti-$CD4^+$ antibody, and L-asparaginase also have been PEGylated in this manner.[75] In the case of interferon α-2b, the PEGylated product maintains its biological activity as well as its tertiary structure.[76] For an in-depth discussion of PEGylation, the readers are directed towards two excellent reviews published recently.[77,78]

1.8. Conclusions

Thiol groups and disulfides have important roles in the stability and solubility of proteins. Thiols and disulfides are two of the most reactive groups in proteins. Increased research is being devoted to studying the role of thiols and disulfides in problems associated with protein molecules. Studies have shown some disulfide bonds in proteins to be essential for either protein stability and/or activity. On the other hand, some non-native disulfide bonds in proteins are a result of altered *in vitro* or *in vivo* conditions and are detrimental to protein stability and/or activity. Various reagents have been designed to detect and quantify free thiols and disulfides in proteins. The choice of the appropriate labeling agent should be made after careful consideration because they all have advantages and disadvantages. Changes in native thiol and/or disulfide content in proteins can be blocked by either the addition of appropriate redox reagents or irreversible conjugation with various agents as described above. Mutation of the Cys residue to another amino acid (*i.e.*, Ser, Ala) is an alternative strategy when the residue is not essential to protein activity. Protein molecules are highly complex systems. Knowledge of the role of Cys residues in

these macromolecular systems is continuously expanding. It is hoped that the growing appreciation for the role of thiols and disulfides in protein structure may lead to new methods to improve the success in formulating proteins as therapeutic agents.

1.9. Application of the information above to current work

As mentioned earlier, EC1 forms disulfide-linked dimers; these dimers form physical oligomers. This heterogenous nature of EC1 in solution is undesirable during the study of its biological activity assay. In the current work, we examine the effect of intermolecular disulfide formation on the stability of EC1. We also apply various methods to block the formation of disulfide-linked dimers. In Chapter 2, we employ the simple addition of DTT to maintain EC1 in its monomeric form and compare the stability of EC1 in the absence and presence of DTT. In Chapter 3, we eliminate the need to add reducing agents such as DTT to EC1 solution to maintain EC1 in monomeric form by alkylating its Cys thiol group with iodoacetamide. We study the stability of the alkylated EC1 and compare it with that of unmodified EC1. In Chapter 4, we modify the Cys thiol group of EC1 by its reaction with maleimide-activated polyethylene glycol (PEG). PEGylation of EC1 in this manner is expected to prevent intermolecular disulfide formation; it also is expected to decrease hydrophobic interactions between EC1 molecules by attracting water molecules to the protein surface and forming a hydrophilic shell. We study the stability of the PEGylated EC1 and again, compare it with that of the unmodifed EC1. We also compare the secondary and tertiary structures of both, the alkylated derivative as well as the

PEGylated derivative with those of unmodified EC1. Retention of the secondary structure after alkylation or PEGylation is expected to be important for the retention of the biological activity of EC1.

1.10. References

1. Bio. 2007. Approved biotechnology drugs. Guide to Biotechnology, ed.

2. Cyran R. 2004. The market for bioengineered protein drugs BCC Research Report and Reviews, ed.

3. Takeichi M 1993. Cadherins in cancer: implications for invasion and metastasis. Curr Opin Cell Biol l5:806–811.

4. Yap AS, Niessen CM, Gumbiner BM 1998. The juxtamembrane region of the cadherin cytoplasmic tail supports lateral clustering, adhesive strengthening, and interaction with p120ctn. J Cell Biol l141:779–789.

5. Corada M, Mariotti M, Thurston G, Smith K, Kunkel R, Brockhaus M, Lampugnani MG, Martin-Padura I, Stoppacciaro A, Ruco L, McDonald DM, Ward PA, Dejana E 1999. Vascular endothelial-cadherin is an important determinant of microvascular integrity in vivo. Proceedings of the National Academy of Sciences l96:9815–9820.

6. Nose A, Tsuji K, Takeichi M 1990. Localization of specificity determining sites in cadherin cell adhesion molecules. Cell l61:147–155.

7. Klingelhofer J, Laur OY, Troyanovsky RB, Troyanovsky SM 2002. Dynamic interplay between adhesive and lateral E-cadherin dimers. Mol Cell Biol l22:7449–7458.

8. Nagar B, Overduin M, Ikura M, Rini JM 1996. Structural basis of calcium-induced E-cadherin rigidification and dimerization. Nature l380:360–364.

9. Overduin M, Harvey TS, Bagby S, Tong KI, Yau P, Takeichi M, Ikura M 1995. Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. Science lv267:p386(384).

10. Shapiro L, Fannon AM, Kwong PD, Thompson A, Lehmann MS, Grubel G, Legrand J-F, Als-Nielsen J, Colman DR, Hendrickson WA 1995. Structural basis of cell-cell adhesion by cadherins. Nature l374:327–337.

11. Troyanovsky RB, Sokolov E, Troyanovsky SM 2003. Adhesive and lateral E-cadherin dimers are mediated by the same interface. Mol Cell Biol l23:7965–7972.

12. Zheng K, Trivedi M, Siahaan TJ 2006. Structure and function of the intercellular junctions: barrier of paracellular drug delivery. Current Pharmaceutical Design l12:2813–2824.

13. Makagiansar I, Avery M, Hu Y, Audus KL, Siahaan TJ 2001. Improving the selectivity of HAV-peptides in modulating E-cadherin-E-cadherin interactions in the intercellular junction of MDCK cell monolayers. Pharm Res l18:446–553.

14. Makagiansar IT, Ikesue A, Duc Nguyen P, Urbauer JL, Bieber Urbauer RJ, Siahaan TJ 2002. Localized production of human E-cadherin-derived first repeat in Escherichia coli. Protein Expression and Purification l26:449–454.

15. Makagiansar IT, Nguyen PD, Ikesue A, Kuczera K, Dentler W, Urbauer JL, Galeva N, Alterman M, Siahaan TJ 2002. Disulfide bond formation promotes the cis- and trans-dimerization of the E-cadherin-derived first repeat. Journal of Biological Chemistry l277:16002–16010.

16. Gruber CW, Cemazar M, Heras B, Martin JL, Craik DJ 2006. Protein disulfide isomerase: the structure of oxidative folding. Trends Biochem Sci l31:455–464.

17. Kadokura H, Katzen F, Beckwith J 2003. Protein disulfide bond formation in prokaryotes. Annual Review of Biochemistry l72:111–135.

18. Frand AR, Cuozzo JW, Kaiser CA 2000. Pathways for protein disulphide bond formation. Trends in Cell Biology l10:203–210.

19. Sevier CS, Kaiser CA 2006. Conservation and diversity of the cellular disulfide bond formation pathways. Antioxid Redox Signal l8:797–811.

20. Kulp MS, Frickel EM, Ellgaard L, Weissman JS 2006. Domain architecture of protein-disulfide isomerase facilitates its dual role as an oxidase and an isomerase in Ero1p-mediated disulfide formation. J Biol Chem l281:876–884.

21. Fomenko DE, Gladyshev VN 2003. Genomics perspective on disulfide bond formation. Antioxid Redox Signal l5:397–402.

22. Beeby M, O'Connor BD, Ryttersgaard C, Boutz DR, Perry LJ, Yeates TO 2005. The genomics of disulfide bonding and protein stabilization in thermophiles. PLoS biology l3:e309.

23. Weissman JS, Kim PS 1995. A kinetic explanation for the rearrangement pathway of BPTI folding. Nat Struct Biol l2:1123–1130.

24. Dadlez M, Kim PS 1995. A third native one-disulphide intermediate in the folding of bovine pancreatic trypsin inhibitor. Nat Struct Biol l2:674–679.

25. Arolas JL, Aviles FX, Chang JY, Ventura S 2006. Folding of small disulfide-rich proteins: clarifying the puzzle. Trends Biochem Sci l31:292–301.

26. Zhang YH, Yan X, Maier CS, Schimerlik MI, Deinzer ML 2002. Conformational analysis of intermediates involved in the in vitro folding pathways of recombinant human macrophage colony stimulating factor beta by sulfhydryl group trapping and hydrogen/deuterium pulsed labeling. Biochemistry l41:15495–15504.

27. Lu B-Y, Beck PJ, Chang J-Y 2001. Oxidative folding of murine prion mPrP(23-231). European Journal of Biochemistry l268:3767–3773.

28. Wang X-Y, Meng F-G, Zhou H-M 2004. The role of disulfide bonds in the conformational stability and catalytic activity of phytase. Biochemistry and cell biology l82:329–334.

29. Singh RR, Appu Rao AG 2002. Reductive unfolding and oxidative refolding of a Bowman-Birk inhibitor from horsegram seeds (*Dolichos biflorus*): evidence for 'hyperreactive' disulfide bonds and rate-limiting nature of disulfide isomerization in folding. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology l1597:280–291.

30. Voss RH, Ermler U, Essen L-O, Wenzl G, Kim Y-M, Flecker P 1996. Crystal structure of the bifunctional soybean Bowman-Birk inhibitor at 0.28-nm resolution. Structural peculiarities in a folded protein conformation. European Journal of Biochemistry l242:122–131.

31. Takagi H, Takahashi T, Momose H, Inouye M, Maeda Y, Matsuzawa H, Ohta T 1990. Enhancement of the thermostability of subtilisin E by introduction of a

disulfide bond engineered on the basis of structural comparison with a thermophilic serine protease. J Biol Chem l265:6874–6878.

32. Villafranca JE, Howell EE, Oatley SJ, Nguyen Huu X, Kraut J 1987. An engineered disulfide bond in dihydrofolate reductase. Biochemistry l26:2182–2189.

33. Wedemeyer WJ, Welker E, Narayan M, Scheraga HA 2000. Disulfide bonds and protein folding. Biochemistry l39:7032–7032.

34. Pace CN, Grimsley GR, Thomson JA, Barnett BJ 1988. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. J Biol Chem l263:11820–11825.

35. Sugimoto H, Nakaura M, Kosuge Y, Imai K, Miyake H, Karita S, Tanaka A 2007. Thermodynamic effects of disulfide bond on thermal unfolding of the starch-binding domain of Aspergillus niger glucoamylase. Bioscience, Biotechnology, and Biochemistry l71:1535–1541.

36. Davoodi J, Wakarchuk WW, Carey PR, Surewicz WK 2007. Mechanism of stabilization of Bacillus circulans xylanase upon the introduction of disulfide bonds. Biophysical Chemistry l125:453–461.

37. Matsumura M, Becktel WJ, Levitt M, Matthews BW 1989. Stabilization of phage T4 lysozyme by engineered disulfide bonds. PNAS l86:6562–6566.

38. Asgeirsson B, Adalbjornsson BV, Gylfason GA 2007. Engineered disulfide bonds increase active-site local stability and reduce catalytic activity of a cold-adapted alkaline phosphatase. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics l1774:679–687.

39. Berezowska I, Chung NN, Lemieux C, Wilkes BC, Schiller PW 2006. Cyclic dermorphin tetrapeptide analogues obtained via ring-closing metathesis. Acta Biochim Pol l53:73–76.

40. Thibault G, Tardif P, Lapalme G 2001. Comparative specificity of platelet alpha(IIb)beta(3) integrin antagonists. J Pharmacol Exp Ther l296:690–696.

41. Rossi ML, Zavalloni D 2004. Inhibitors of platelets glycoprotein IIb/IIIa (GP IIb/IIIa) receptor: rationale for their use in clinical cardiology. Mini Rev Med Chem l4:703–709.

42. Gimpl G, Burger K, Politowska E, Ciarkowski J, Fahrenholz F 2000. Oxytocin receptors and cholesterol: interaction and regulation. Exp Physiol l85 Spec No:41S–49S.

43. Bogdanowich-Knipp SJ, Chakrabarti S, Williams TD, Dillman RK, Siahaan TJ 1999. Solution stability of linear vs. cyclic RGD peptides. J Pept Res l53:530–541.

44. Bogdanowich-Knipp SJ, Jois DS, Siahaan TJ 1999. The effect of conformation on the solution stability of linear vs. cyclic RGD peptides. J Pept Res l53:523–529.

45. Bogdanowich-Knipp SJ, Jois SD, Siahaan TJ 1999. Effect of conformation on the conversion of cyclo-(1,7)-Gly-Arg-Gly-Asp-Ser-Pro-Asp-Gly-OH to its cyclic imide degradation product. J Pept Res l54:43–53.

46. Tanaka A, Karita S, Kosuge Y, Senoo K, Obata H 1999. Thermal unfolding of mutant forms C509G and C509S of starch binding domain-fragment of Aspergillus niger glucoamylase. Netsusokutei l26:136.

47. Wright SK, Viola RE 1998. Evaluation of methods for the quantitation of cysteines in proteins. Analytical Biochemistry l265:8–14.

48. Ellman GL 1959. Tissue sulfhydryl groups. Archives of Biochemistry and Biophysics l82:70–77.

49. Robyt JF, Ackerman RJ, Chittenden CG 1971. Reaction of protein disulfide groups with Ellman's reagent: a case study of the number of sulfhydryl and disulfide groups in Aspergillus oryzae -amylase, papain, and lysozyme. Arch Biochem Biophys l147:262–269.

50. Sedlak J, Lindsay RH 1968. Estimation of total, protein-bound, and nonprotein sulfhydryl groups in tissue with Ellman's reagent. Anal Biochem l25:192–205.

51. Riener CK, Kada G, Gruber HJ 2002. Quick measurement of protein sulfhydryls with Ellman's reagent and with 4,4'-dithiodipyridine. Anal Bioanal Chem l373:266–276.

52. Kosower NS, Kosower EM, Newton GL, Ranney HM 1979. Bimane fluorescent labels: labeling of normal human red cells under physiological conditions. Proceedings of the National Academy of Sciences of the United States of America l76:3382–3386.

53. Kosower NS, Newton GL, Kosower EM, Ranney HM 1980. Bimane fluorescent labels. Characterization of the bimane labeling of human hemoglobin. Biochimica et Biophysica Acta l622:201–209.

54. Packard B, Edidin M, Komoriya A 1986. Site-directed labeling of a monoclonal antibody: targeting to a disulfide bond. Biochemistry l25:3548–3552.

55. Faulstich H, Tews P, Heintz D 1993. Determination and derivatization of protein thiols by n-octyldithionitrobenzoic acid. Analytical Biochemistry l208:357–362.

56. Chinn PC, Pigiet V, Fahey RC 1986. Determination of thiol proteins using monobromobimane labeling and high-performance liquid chromatographic analysis: application to Escherichia coli thioredoxin. Analytical Biochemistry l159:143–149.

57. He HT, Gürsoy RN, Kupczyk-Subotkowska L, Tian J, Williams T, Siahaan TJ 2006. Synthesis and chemical stability of a disulfide bond in a model cyclic pentapeptide: Cyclo(1,4)-Cys-Gly-Phe-Cys-Gly-OH. Journal of Pharmaceutical Sciences l95:2222–2234.

58. Florence MT 1980. Degradation of protein disulphide bonds in dilute alkali. Biochem J l189:507–520.

59. Tous GI, Wei Z, Feng J, Bilbulian S, Bowen S, Smith J, Strouse R, McGeehan P, Casas-Finet J, Schenerman MA 2005. Characterization of a novel modification to monoclonal antibodies: Thioether cross-link of heavy and light Chains. Anal Chem l77:2675–2682.

60. Kim PS, Kim KR, Arvan P 1993. Disulfide-linked aggregation of thyroglobulin normally occurs during nascent protein folding. Am J Physiol l265:C704–711.

61. Bucala R, Manabe S, Urban RC, Cerami A 1985. Nonenzymatic modification of lens crystallins by prednisolone induces sulfhydryl oxidation and aggregate formation: In vitro and in vivo studies. Experimental Eye Research l41:353–363.

62. Manabe S, Bucala R, Cerami A 1984. Nonenzymatic addition of glucocorticoids to lens proteins in steroid-induced cataracts. The Journal of Clinical Investigation l74:1803–1810.

63. Hook DWA, Harding JJ 2002. The effect of modification of alpha-crystallin by prednisolone-21-hemisuccinate and fructose 6-phosphate on chaperone activity. Developments in Ophthalmology l35:150–160.

64. Kelly ST, Zydney AL 1994. Effects of intermolecular thiol-disulfide interchange reactions on BSA fouling during microfiltration. Biotechnology and Bioengineering l44:972–982.

65. Ruegg UT, Rudinger J editors. Alkylation of cysteine thiols with 1,3-propane sultone. ed. p 116–122.

66. Smith DJ, Miggio ET, Kenyon GL 1975. Simple alkanethiol groups for temporary blocking of sulfhydryl groups of enzymes. Biochemistry l14:766–771.

67. Koizumi S, Suzuki T, Takahashi S, Satake K, Takeuchi T, Umezawa H, Nagatsu T 1987. Sulfhydryl modification and activation of phenylalanine hydroxylase by dinitrophenyl alkyl disulfide. Biochemistry l26:6461–6465.

68. Schwendeman SP, Lee JH, Gupta RK, Costantino HR, Siber GR, Langer R 1994. Inhibition of moisture-induced aggregation of tetanus toxoid by protecting thiol

groups. Proceedings of the International Symposium on Controlled Release of Bioactive Materials l21ST:54.

69. Hallewell RA, Imlay KC, Lee P, Fong NM, Gallegos C, Getzoff ED, Tainer JA, Cabelli DE, Tekamp-Olson P, Mullenbach GT, Cousens LS 1991. Thermostabilization of recombinant human and bovine CuZn superoxide dismutases by replacement of free cysteines. Biochemical and Biophysical Research Communications l181:474–480.

70. Marchand P, Volkmann M, Bond JS 1996. Cysteine mutations in the MAM domain result in monomeric meprin and alter stability and activity of the proteinase. J Biol Chem l271:24236–24241.

71. Rosendahl MS, Doherty DH, Smith DJ, Bendale AM, Cox GN 2005. Site specific protein PEGylation: Application to cysteine analogs of recombinant human granulocyte colony-stimulating factor. Bioprocess international l3:52–62.

72. Rosendahl MS, Doherty DH, Smith DJ, Carlson SJ, Chlipala EA, Cox GN 2005. A long acting, highly potent interferon α-2 conjugate created using site-specific PEGylation. Bioconjugate Chemistry l16:200–207.

73. Tsutsumi Y, Onda M, Nagata S, Lee B, Kreitman RJ, Pastan I 2000. Site-specific chemical modification with polyethylene glycol of recombinant immuntoxin anti-Tac(Fv)-PE38 (LMB-2) improves antitumor activity and reduces animal toxicity and immunogenicity. l97:8548–8553.

74. Long DL, Doherty DH, Eisenberg SP, Smith DJ, Rosendahl MS, Christensen KR, Edwards DP, Chlipala EA, Cox GN 2006. Design of homogeneous,

monopegylated erythropoietin analogs with preserved in vitro bioactivity. Exp Hematol l34:697–704.

75. Brocchini S, Balan S, Godwin A, Choi JW, Zloh M, Shaunak S 2006. PEGylation of native disulfide bonds in proteins. Nat Protoc l1:2241–2252.

76. Balan S, Choi Jw, Godwin A, Teo I, Laborde CM, Heidelberger S, Zloh M, Shaunak S, Brocchini S 2007. Site-specific PEGylation of protein disulfide bonds using a three-carbon bridge. Bioconjugate Chem l18:61–76.

77. Morar AS, Schrimsher JL, Chavez MD. 2006. PEGylation of proteins: A structural approach. Biopharm International, ed. p 34–.

78. Parveen S, Sahoo SK 2006. Nanomedicine: Clinical applications of polyethylene glycol conjugated proteins and drugs. Clinical Pharmacokinetics l45:965–.

Chapter 2

The effect of covalent dimerization on the structure and stability of EC1 domain of human E-cadherin

2.1. Introduction

The EC1 domain of cadherins has been shown to be important for the homophilic selectivity of cadherins.[1-7] E-cadherin, a member of this cadherin family, is a transcellular protein located in the adherens junctions of epithelial and endothelial tissues. It creates a paracellular barrier to transepithelial transport of large molecules through homophilic interactions on the lateral surface of same cell surface as well as opposing cell surfaces. Two E-cadherin molecules from the same cell surface form a *cis*-dimer through interaction of their extracellular segments; interaction of *cis*-dimers from opposing cell surfaces forms *trans*-oligomers. It has been proposed that the formation of *cis*-dimers of E-cadherin is necessary for the formation of *trans*-oligomers.[3,4] The extracellular portion of E-cadherin is composed of five sub-domains EC1 through EC5. Each of these sub-domains has an immunoglobulin-like β-sheet-rich structure. All five sub-domains as well as the linker sequences between the sub-domains possess Ca^{2+} binding residues. Ca^{2+} binding has been shown to be essential for E-cadherin mediated cell adhesion. The EC1 domain is the N-terminal domain and is located farthest from the cell surface. The EC1 domain has been shown to interact with an EC1 domain of a neighboring E-cadherin molecule to form cell-cell adhesion sub-structures in the intercellular junctions. E-cadherin-mediated homophilic cell-cell adhesion can be inhibited by EC1-derived peptides,[8,9] These peptides temporarily open the intercellular junctions of MDCK monolayers to enhance the transport of ^{14}C-mannitol.[10,11] Although the EC1 domain can modulate the E-cadherin-mediated Caco-2 single cell adhesion to modified Caco-2 cell

monolayers, it is difficult to store this molecule in solution due to its instability as described below. Thus, it is of interest to study the stability of the EC1 domain so that its behavior can be better understood and it can be stabilized for future studies.

The recombinantly expressed EC1 domain in this study is from human E-cadherin. It contains 105 amino acid residues (MW = 11,628 Da), including four additional non-native amino acids (GSHM) at the N-terminus. The Cys9 residue in the native EC1 is located at Cys13 in the EC1 domain studied here (Figure 1a).[12] The monomeric form of the EC1 domain can readily form a covalent dimer via intermolecular disulfide bond formation of the Cys13 residues followed by the formation of physical oligomers as well as precipitates.[13] The dimeric and oligomeric forms cannot be directly used to study the biological activity of the EC1. We propose that the formation of physical oligomers is induced by the presence of a covalent dimer; thus, blocking the formation of the covalent dimer may inhibit the oligomerization and precipitation of the EC1 domain. Therefore, the goal of this study was to investigate the effect of covalent dimerization as it pertains to oligomerization, precipitation and the chemical stability of EC1. Stabilization of the monomer should help our effort in (a) elucidating the mechanism of interaction between the EC1 domain and other recombinantly expressed domains (EC2, EC3, EC4, and EC5) of E-cadherin, (b) evaluating the activity of the EC1 domain in blocking homotypic and heterotypic cell-cell adhesion, and (c) determining the binding site(s) of E-cadherin peptides on the EC1 domain. Furthermore, EC1 can also provide a general model for studying the oligomerization process of a protein in solution.

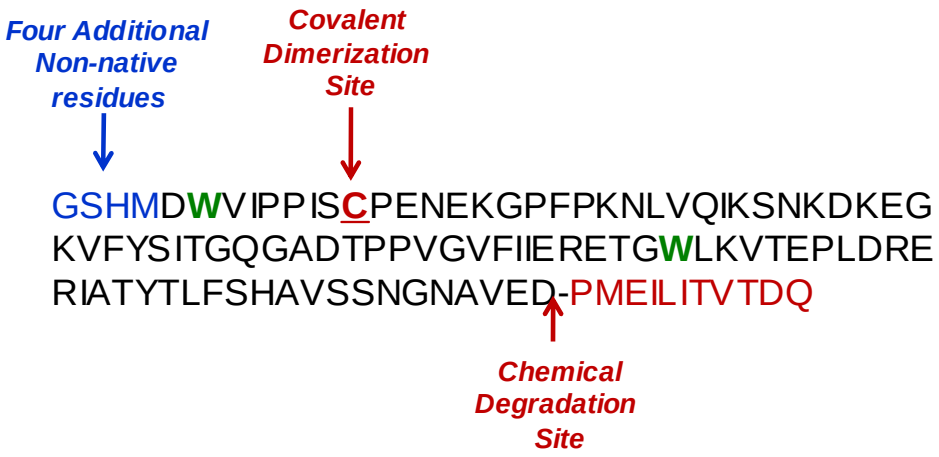

Figure 2.1a: Primary sequence of the EC1 protein obtained at the end of the purification process. The first four residues GSHM are non-native residues that result from the type of vector used for transformation into *E. coli* cells. The single Cys13 residue is equivalent to Cys9 residue in the native EC1. The potential hydrolysis site is also indicated.

In this work, we studied the accelerated solution physical and chemical stability of the EC1 domain at different temperatures and pH values in the absence and presence of dithiothreitol (DTT). The chemical degradation process was evaluated by high pressure liquid chromatography (HPLC) and mass spectrometry (MS). The physical stability and structural properties of the EC1 domain were evaluated using spectroscopic methods (i.e., CD and fluorescence). The effect of pH and incubation time on the secondary structure was studied by performing CD spectroscopy. Intrinsic fluorescence emission studies also were conducted to follow changes in the tertiary structure of EC1 upon incubation under different conditions. Molecular dynamics simulations were used to investigate the dynamic properties of the EC1 monomer and dimer; these simulations suggest that the EC1 dimer may undergo conformational changes more readily than the monomer. The results also indicate that the formation of the intermolecular disulfide bond by EC1 is unfavorable to its physical and chemical stability.

2.2. Materials and methods

2.2.1. Expression of EC1:

The DNA of EC1 was produced by performing a polymerase chain reaction (PCR) of the cDNA of human E-cadherin (obtained from David Rimm, M.D., Ph.D., Yale University School of Medicine). The nucleotide sequence 5'-GTAGCATATGGACTGGGTTATTCCTC-3' was used as the forward primer and 5'-GAATCATATGTCACTGATCGGTTACCGTG-3' was used as the reverse primer.

The EC1 DNA thus amplified was then ligated into pET-15b vector (obtained from EMD Biosciences) using the restriction site Nde I. The pET-15b vector has a sequence that codes for an N-terminal His-tag (6 Histdine residues in a row) and allows us to use the metal binding property of this His-tag for eventual purification of the recombinantly expresed EC1. DNA sequencing was performed to confirm the ligation product contained EC1 sequence. The ligation product was then transformed into competent BL21 (DE3) *E. coli* cells (obtained from Novagen, Inc). These *E. coli* cells were grown in LB medium at 37 °C and 250 rpm shaking. When the optical density of the *E. coli* reached 0.5 units at 600 nm, the cells were induced with isopropyl-β-D-thiogalactopyranoside (IPTG). The cell growth was monitored until their UV absorbance at 600 nm reached a plateau (usually about 1.5 units after 3 h). The *E. coli* cells were then harvested by centrifugation at a speed of 10,000 x g for 6 min. The pellets obtained were resuspended in a buffer containing 100 mM tris-(hydroxymethyl)aminomethane (Tris), 100 mM NaCl, and 2.0 mM β-mercaptoethanol (BME) at pH 8.0. The suspension of cells was passed through a French pressure cell at a pressure of more than 1000 psi to lyse the cells and release the expressed EC1 into the buffer.[14] The lysed cell suspension was then centrifuged for 45 min at a speed of 48,000 x g at 4 °C to separate the insoluble cell debris from the dissolved EC1. The supernatant obtained in this manner was further cleared of any particulate matter by passage through a 0.45 micron filter. The EC1 obtained in this manner has a His-tag linked to its N-terminus with a sequence (LVPRGS) that is a substrate for Thrombin. The cleavage of the His-tag with Thrombin results in

retention of a four-amino acid extension (GSHM) preceding the EC1 protein. As such, the numbering is adjusted by +4 compared to human EC1.

2.2.2. Purification of the EC1 Protein:

The lysed cell suspension above containing His-tagged EC1 was loaded onto a HisTrap HP nickel affinity column (GE Healthcare).[14] This was followed by elution with washing buffer containing 100 mM Tris and 100 mM NaCl at pH 8.0 with increasing concentration of imidazole (from 10 mM to 50 mM) to wash off other non-specifically bound proteins. The EC1-Histag protein bound to the nickel affinity column was then eluted by increasing the concentration of imidazole to 150 mM. The eluted EC1-Histag was dialyzed against a buffer containing 20 mM Tris, 100 mM NaCl, and 10 mM EDTA at pH 8.0 to remove any residual nickel. The EC1-Histag solution was then dialyzed against buffer containing 20 mM Tris and 100 mM NaCl at pH 8.0 to remove the EDTA. The pure EC1-Histag was concentrated to 1.0 mL followed by dilution with 8 mL of deionized water. To 8.0 mL of diluted protein, thrombin (0.5 units/mg of EC1-Histag) was added in cleavage buffer (200 mM Tris, 1.5 M NaCl, 25 mM $CaCl_2$, 2 mM BME) and incubated with gentle shaking for 2 h at room temperature to cleave the Histag from the EC1 domain. At the end of 2 h incubation, the mixture was run through a benzamidine column to trap the thrombin; a mixture of EC1 and cleaved Histag was eluted from the column using buffer (20 mM Tris, 100 mM NaCl, 2 mM BME) at pH 8.0. The eluted mixture was then passed through the nickel affinity column to separate the pure EC1 from the Histag. The pure

EC1 obtained was then dialyzed in a buffer containing 20 mM Tris, 100 mM NaCl, 2 mM BME, and 10 mM EDTA at pH 8.0 to remove any residual nickel. Finally, the EC1 solution was dialyzed against a buffer containing 20 mM Tris, 100 mM NaCl, and 2 mM BME at pH 8.0 to remove any remaining EDTA. SDS-PAGE of the protein was performed to confirm its purity, and the identity of EC1 was confirmed by determining its molecular weight with electrospray ionozation (ESI) positive mode time of flight (TOF) mass spectrometry (MS). The molecular weight of the recombinant EC1 determined was 11,628 amu which is the same as its theoretical molecular weight (Figure 1b).

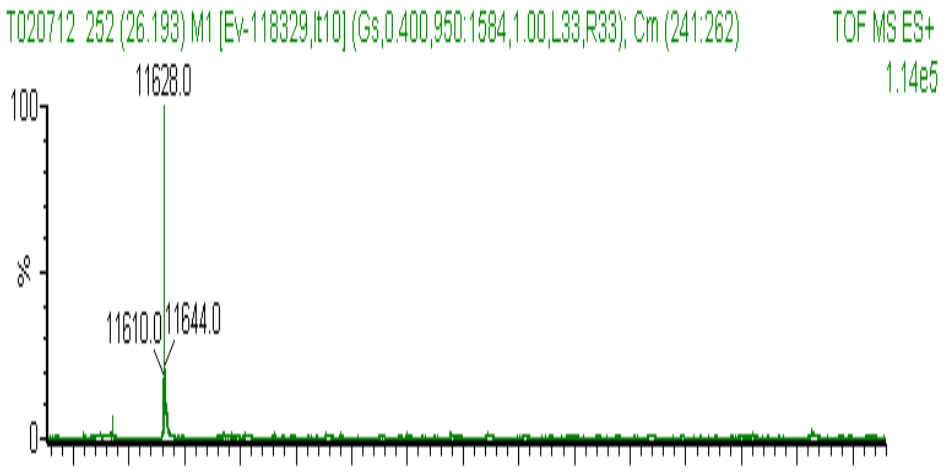

Figure 2.1b: Mass spectrum of purified EC1 as obtained from electrospray ionization-time-of-flight mass spectrometer (ESI-TOF MS).

2.2.3. Chemical Stability Studies:

Chemical stability studies of 86 μM EC1 (1 mg/mL) were performed at pH 3.0, 7.0, and 9.0 and temperatures of 4 °C, 37 °C, and 70 °C. The buffers used were 100 mM phosphate for pH 3.0, 50 mM phosphate for pH 7.0, and 100 mM borate containing 0.08 M NaCl for pH 9.0. EC1 was dialyzed into each buffer using Amicon centrifugal filter devices containing a membrane with molecular weight cut-off of 10,000 Da. Samples containing the EC1 protein at different pH values were incubated at 4 °C, 37 °C or 70 °C in the absence and presence of 0.1 M DTT. The molar ratio of DTT: EC1 in these studies is approximately 10,000:1. Hence, a 99.9% loss in the amount of DTT at 70 °C would be expected to still leave enough DTT to maintain EC1 in its reduced monomeric form. Three samples were drawn out at each time interval and immediately frozen at –70°C prior to analysis. Each of the frozen samples was thawed and centrifuged at 12,000 x g for 10 mins at 4 °C before HPLC analysis described below.

The degradation profiles of the EC1 protein were evaluated using SDS PAGE and an HPLC system (Dynamax SD-200). Sample (25 μL) were injected onto an analytical HPLC column (C18, Varian Microsorb, pore-size 300Å, dimensions 250 x 4.6 mm) followed by gradient elution at a rate of 1.0 mL/min. Solvent A (94.9% double distilled water, 5% acetonitrile (ACN) and 0.1% TFA) and solvent B (100% ACN) were used in various proportions for the separation and resolution of eluting peaks. The sequence of the gradient elution was as follows: 0% B to 45% B in 2.0 min, 45% B to 51% B in 10.0 min, at 51% B for 5.0 min, 51% B to 100% B in 2.0

min, at 100% B for 2.0 min, and 100% B to 0% B in 1.0 min. The eluted protein and its degradation products were detected by a Varian Prostar UV detector at a wavelength of 220 nm. The degradation products were identified by LCMS employing a C4 micro-column with an ESI positive mode MS with TOF detection.

2.2.4. Physical Stability Studies:

Long-term physical stability studies of EC1 were performed at a concentration of 86 μM (1.0 mg/mL) and pH 3.0, 7.0, and 9.0 under the same conditions as the chemical stability studies (see above). The sealed EC1 samples (300 μL in each vial) were stored at 4 °C for 14 or 28 days in the absence and presence of 2.0 mM DTT. The CD spectra of the samples prior to incubation (day 0) and after incubation were collected in the far-UV spectrum (200–250 nm) at 1 nm bandwidth and 0.1 cm pathlength using a Jasco spectropolarimeter (J–720) equipped with a Peltier temperature controller. The thermal unfolding profile of the protein samples was monitored in two ways: 1) by collecting CD spectra every 5 °C with increase in temperature from 10 °C to 65 °C at the rate of 15 °C per hour after equilibrating the sample for 300 s at each temperature and 2) by monitoring the CD signal at 218 nm (for change in β-sheet structure) at every 0.5 °C rise in temperature.[15] Each sample was run in triplicate and a blank sample spectrum containing the buffer alone was subtracted from each sample spectrum. The ellipticity value at 218 nm was plotted for each temperature point to examine the thermal unfolding profile of the sample. This

plot was fitted to a sigmoidal function using Origin 7.0 software, and the midpoint of each plot was taken as the thermal unfolding (denaturation) temperature.

For fluorescence spectroscopy studies, 1.0 ml 86 μM EC1 (1 mg/mL) was incubated at pH values of 3.0, 7.0 and 9.0 each for a period of 28 days at 4 °C. At days 0, 14, and 28, EC1 (0.3 mL) was withdrawn from the incubated samples and diluted to 0.1 mg/mL with buffer of the corresponding pH. This solution (0.9 mL) was loaded into a quartz fluorescence cuvette of 1 cm pathlength and intrinsic fluorescence spectra were collected on a PTI QuantaMaster spectrofluorometer at 2.5 °C intervals from 10 °C to 87.5 °C. The analyte was equilibrated for 5 min at every temperature prior to the spectrum collection. An emission spectrum was obtained from 305 nm to 405 nm after excitation of the analyte at 295 nm (>95% Trp emission).[16] The excitation and emission slits were adjusted so that the emission at its maximum was between 500,000 and 1,000,000 counts/sec. Each experiment was performed in triplicate, and the appropriate buffer was run as a control.

The entire analysis described above was repeated in the presence of 2.0 mM DTT. To analyze the data, the signals from the blanks were subtracted from emission spectra of the samples. Each fluorescence emission spectrum obtained was fitted to an extreme assymetric peak function using the non-linear curve-fitting wizard of Origin 7.0 software to obtain the wavelength of maximum emission. The wavelength of maximum emission at each temperature was plotted against temperature to calculate the transition temperature.

2.2.5. Computer Simulations and Molecular Modeling:

Molecular dynamics simulations of the E-cadherin EC1 monomer (PDB code: 1EDH)[1,17] and the EC1 domain dimer containing a disulfide bond were performed for 10 ns using Nanoscale Molecular Dynamics 2.6 (NAMD 2.6).[18] Both systems were solvated by cubic boxes of TIP3P water with a margin of ~15 Å between the protein and the boundaries of the periodic box. The Na^+ and Cl^- counterions were added to neutralize the system. Protein, water, and ions were modeled with the CHARMM 22 force field.[18] The particle mesh Ewald method was used to treat long-range electrostatics interactions. Periodic boundary conditions, 10 Å cutoff, and a switching cutoff distance of 8 Å for nonbonded van der Waals (vdw) interactions were applied. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm and a time-step of 2 fs was employed to integrate the equations of motion. The equilibration stage includes energy minimization of the protein for 5,000 steps followed by minimization of the complete system for 10,000 steps to remove close contacts. Finally, the entire system was subjected to a gradual temperature increase from 0 K to 300 K in intervals over 30 ps by increasing the temperature of Langevin damping and Langevin piston by 30 K in each step. The complete system was then equilibrated for 300 ps. Pressure and temperature were maintained at 1 atm and 300 K using a Langevin barostat and Langevin thermostat.

2.3. Results

2.3.1. Chemical Stability of EC1:

The EC1 protein was incubated in the presence and absence of 0.1 M DTT at pH 3.0, 7.0, and 9.0 at 4 °C, 37 °C, and 70 °C. Samples were drawn out at 0.5, 1, 2 and 4 h time-points and the supernatant after centrifugation of each sample at 12,000 X g was injected into HPLC system. The decrease in the area under the EC1 peak as well as the increase in the area under any new peaks, if present, was monitored. Degradation products, represented by new peaks, were identified using LCMS (ESI-TOF). Potential products include deamidation and peptide bond hydrolysis products. The results of this study are described below.

Stability at 4 °C: When incubated at 4 °C for 4 h at pH values of 3.0, 7.0, and 9.0, there was no substantial decrease in the amount of EC1 protein as detected by the HPLC analysis described before (data not shown). There was no observable difference in the stability of EC1 in the absence and presence of DTT, suggesting that the protein was stable at low temperature.

Stability at 37 °C: The EC1 protein was stable at pH 3.0 for 4 h at 37 °C, and it was difficult to distinguish the effect of DTT on the stability of EC1 (Figure 2.2a).

At pH 7.0, the amount of EC1 decreased to about 76 % after 4 h in the absence of DTT (Figure 2.2b). No new peak was observed, suggesting that the loss of EC1 was presumably due to precipitation. In contrast, the loss of EC1 to precipitation after 4 h in the presence of DTT at pH 7.0 was only 6% (Figure 2.2b).

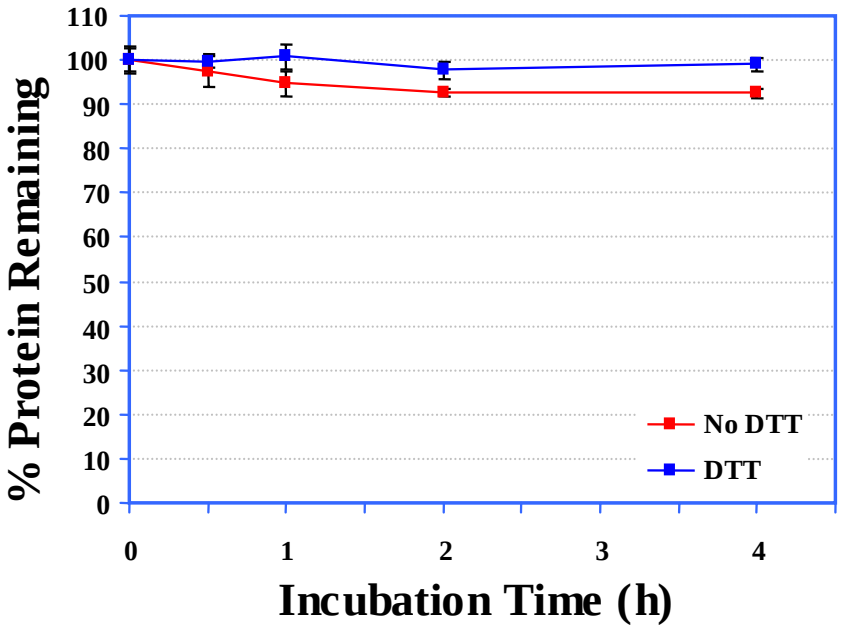

Figure 2.2a: The stability profile of EC1 incubated at 37 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 3.0. In the absence of DTT, there is some loss in the amount of EC1, presumably due to precipitation, since no new peaks were observed on the HPLC under these conditions.

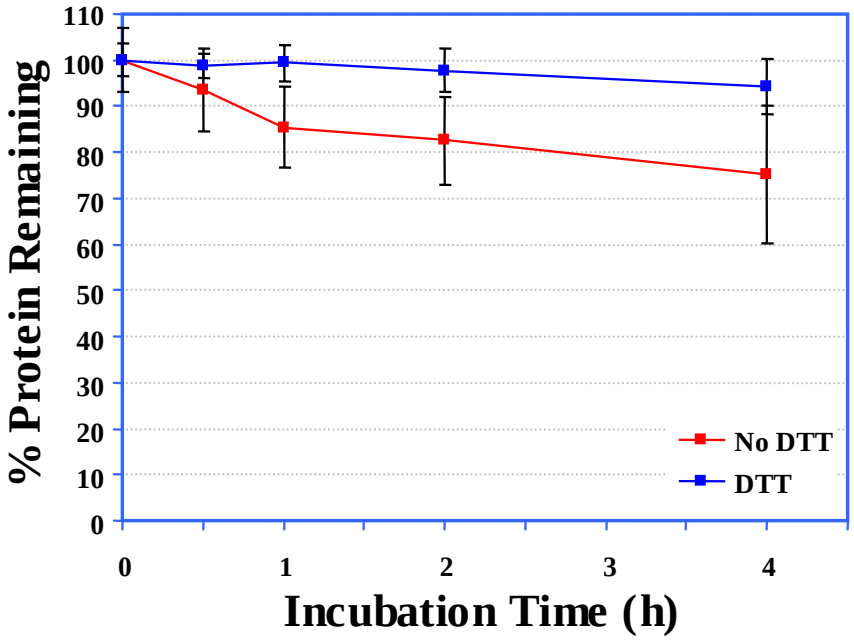

Figure 2.2b: The stability profile of the EC1 domain incubated at 37 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 7.0. 25% EC1 was lost after 4h due to precipitation, in the absence of DTT. In the presence of 0.1 M DTT, however, only 6% of EC1 precipitated after 4 hours.

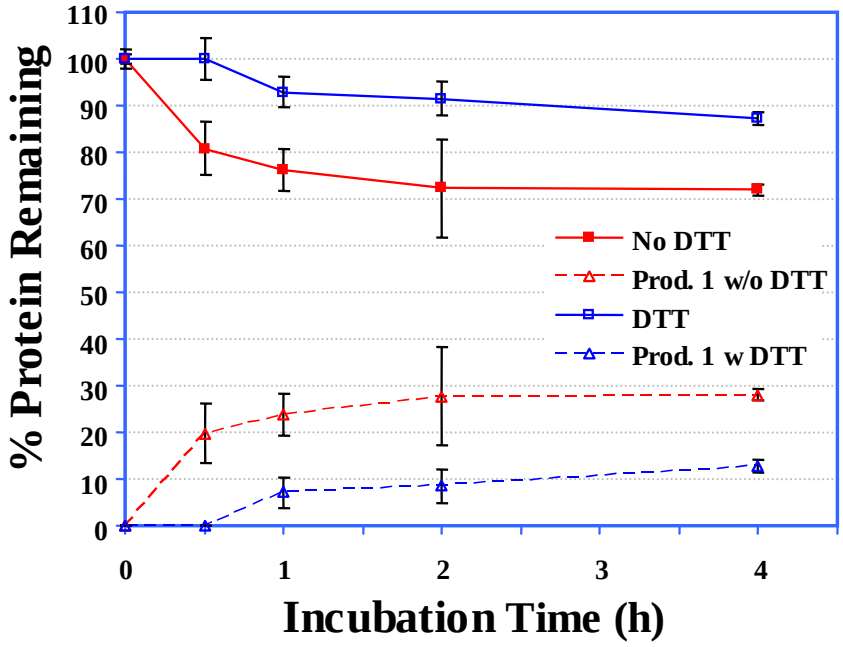

Figure 2.2c: The stability profile of the EC1 domain incubated at 37 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 9.0 studied by monitoring the loss in area under the peak of EC1 and appearance and increase in the area under a new peak on HPLC. In the absence of DTT, about 27% of EC1 degraded to form a new product. In the presence of 0.1 M DTT, however, only 13% of EC1 degraded to form the new product.

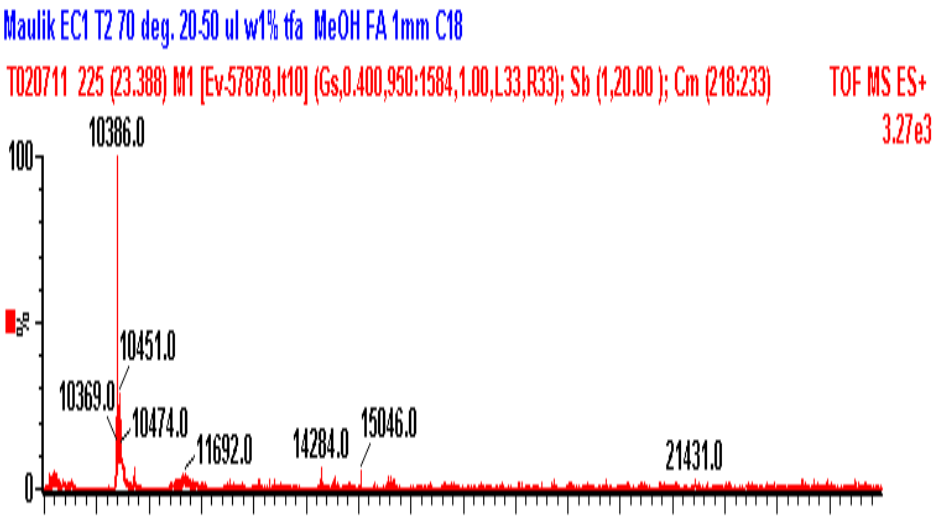

Figure 2.2d: ESI (+ve)- TOF MS of the degradation product of EC1

At pH 9.0, however, a chemical degradation product appeared over time (Figure 2.2c) with a molecular weight of 10,386 Da as determined by ESI positive TOF-MS (Figure 2.2d). In the absence of DTT, the amount of the degradation product was almost 27% and the amount of EC1 decreased to 73% (Figure 2.2c). In the presence of DTT, however, the amount of the degradation product was only about 12%, and the amount of DTT decreased to 87% (Figure 2.2c). Thus, no precipitation is observed at pH 9.0. The degradation product corresponds to a protein fragment resulting from the cleavage of the peptide bond between the Asp93 and Pro94 residues (Figure 2.1a).

Thus, DTT supresses the precipitation as well as the peptide bond hydrolysis degradation of EC1.

Stability at 70 °C: The degradation of EC1 at pH 3.0 and 70 °C was rapid compared to that at 4 °C and 37 °C with about 25% and 5% of EC1 remaining in solution after 2 h and 4 h incubation, respectively (Figure 2.3a). After 30 min, the degradation product with a molecular weight of 10,386 Da was observed as previously found at pH 9.0 and 37 °C. This product was identified previously as the product of hydrolysis of the Asp94-Pro94 peptide bond. An increase in the degradation product, from 20% to 40%, was seen during 0.5 to 2.0 h incubation followed by a drop in product amount to 12% at the 4 h incubation time. The decrease in the degradation product in solution and the lack of mass balance suggest that besides peptide bond hydrolysis, the degradation of EC1 also involves a precipitation

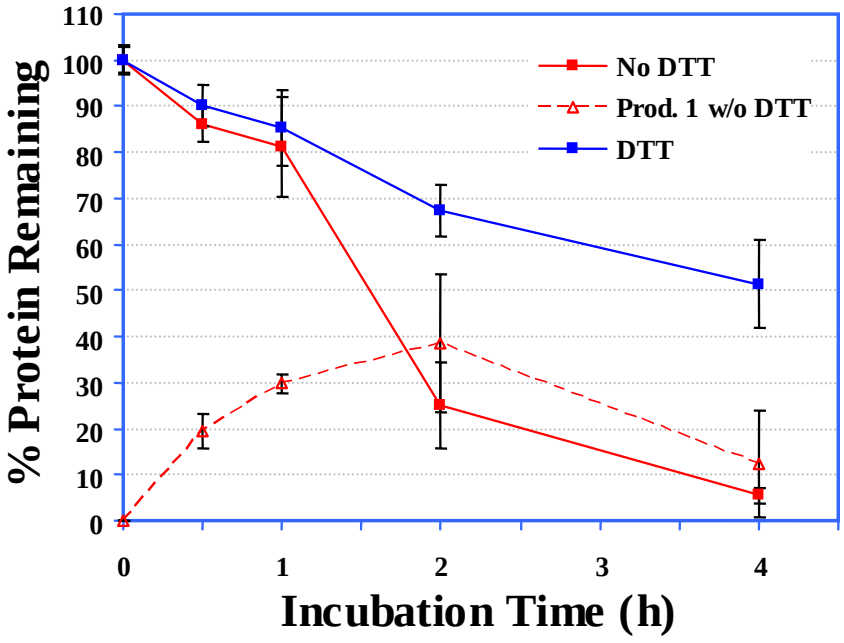

Figure 2.3a: The stability profiles of the EC1 domain incubated at 70 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 3.0 as determined by HPLC analysis. In the absence of DTT, EC1 undergoes some peptide bond hydrolysis, but mostly precipitation. In the presence of 0.1 M DTT, however, the extent of precipitation of EC1 after 4 h is decreased and no hydrolysis product is seen.

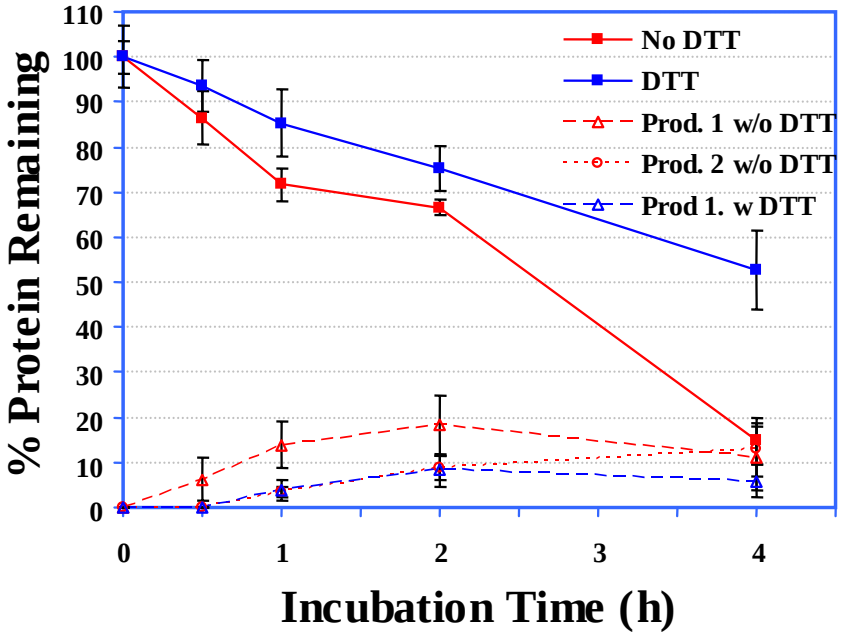

Figure 2.3b: The stability profiles of the EC1 domain incubated at 70 °C for 4 hin the absence and presence of 0.1 M DTT in buffer solutions at pH 7.0 as determined by HPLC analysis. EC1 undergoes hydrolysis of the peptide bond (12% after 4 h) and precipitation (70% after 4 h). In the presence of DTT, however, the extent of peptide bond hydrolysis is decreased (6% after 4 h) and so is the precipitation of EC1 (40% after 4 h).

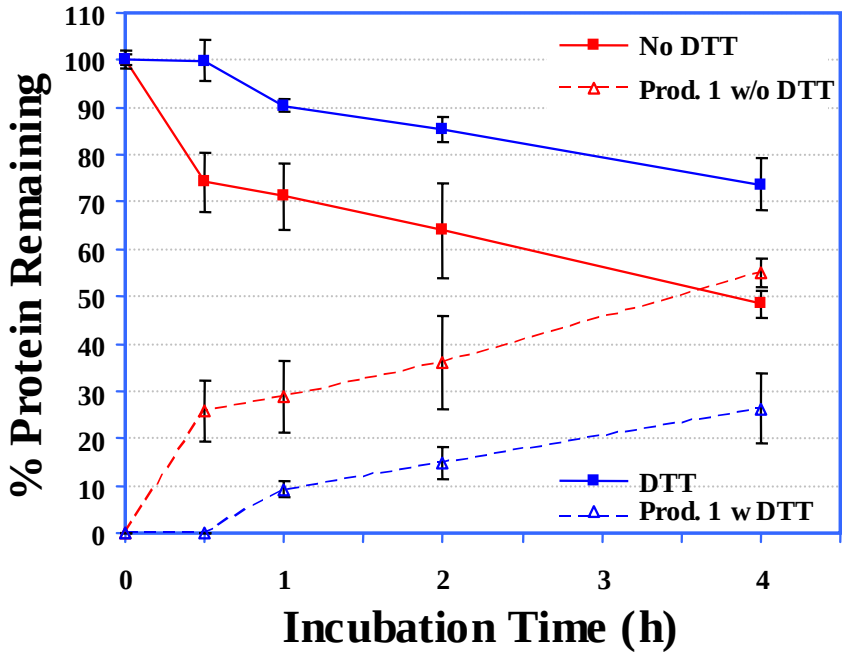

Figure 2.3c: The stability profiles of the EC1 domain incubated at 70 °C for 4 h in the absence and presence of 0.1 M DTT in buffer solutions at pH 9.0 as determined by HPLC analysis. In the absence of DTT, 53% EC1 undergoes peptide bond hydrolysis after 4 h. In the presence of 0.1 M DTT, however, the peptide bond hydrolysis is decreased to 27% after 4 h.

reaction. Again, it is found that DTT decreases the precipitation and completely suppress the peptide bond hydrolysis degradation. The degradation of EC1 at pH 7.0 (Figure 2.3b) also showed a two-stage profile of loss of EC1 in which an initial stage is seen from 100% to 67% between 0 and 2 h and a second from 67% to 15% between 2 and 4 h. A peptide bond hydrolysis product, similar to that seen at pH 3.0 but in a lower amount (20% after 2 hours), was found at pH 7.0, which implies that the degradation rate was slower at pH 7.0 than pH 3.0. Also, under these conditions, 70% of EC1 was presumably lost to precipitation, due to the lack of mass balance. The addition of DTT inhibited the rate of peptide bond hydrolysis as well as precipitation of EC1.

EC1 had the greatest amount of chemical degradation product at pH 9.0 (Figure 2.3c) compared to that at pH 3.0 and 7.0. Again, the production of the peptide bond hydrolysis product was suppressed by the presence of DTT. In addition, the amount of starting material remaining in solution is higher at pH 9.0 than at pH 3.0 or 7.0 at the 4 h time point, which suggests less precipitation of EC1 at pH 9.0 than at pH 3.0 and pH 7.0.

2.3.2. Physical Stability of EC1:

2.3.2.1. Secondary structure evaluation using CD: At pH 3.0, the CD spectrum of EC1 has a negative peak around 216 nm before incubation (Figure 2.4a, day 0), suggesting that it has a predominantly β-sheet structure. This is consistent with the known structures of EC domains. Addition of DTT caused some change in the

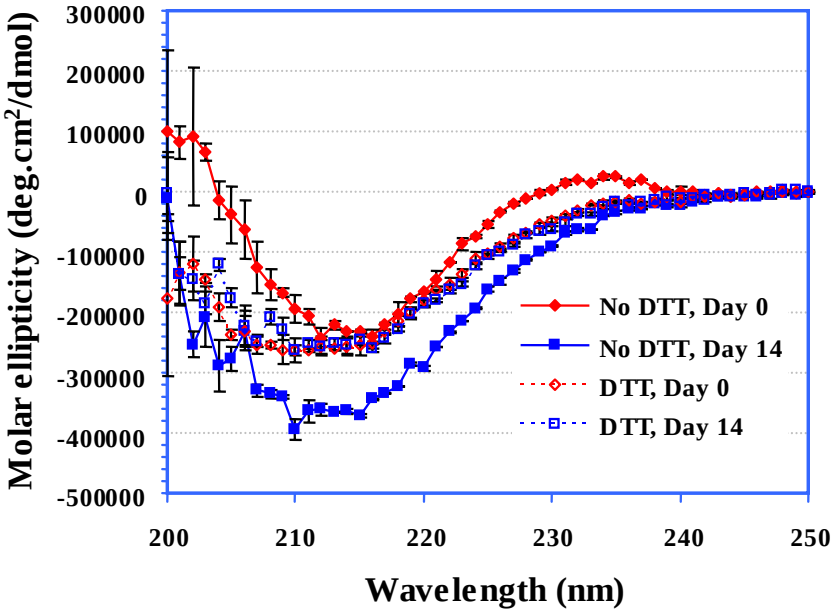

Figure 2.4a: CD spectra of EC1 in the absence and presence of DTT after incubation at 4 °C for 0, 14, or 28 days at pH 3.0. In the absence of DTT, EC1 undergoes secondary structural change presumably due to dimerization and oligomerization. In the presence of 2 mM DTT, there is no secondary structural change in EC1.

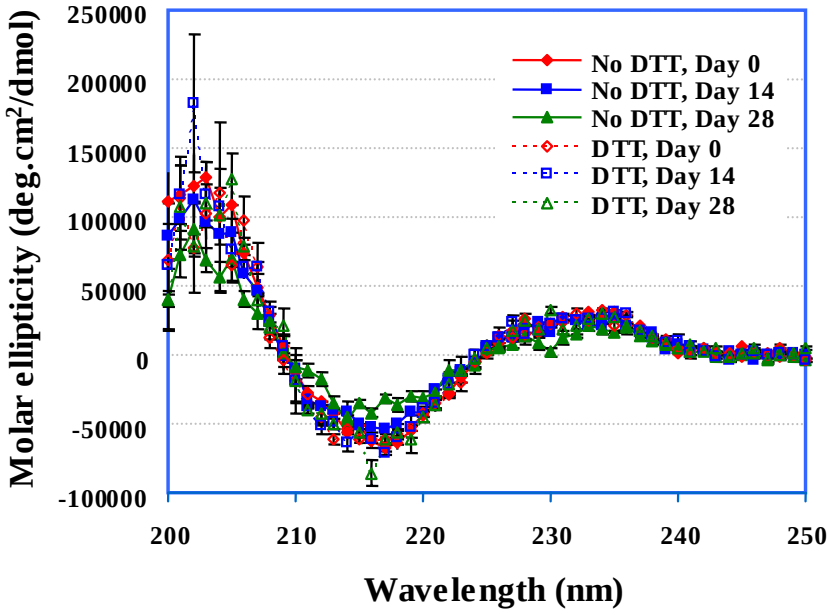

Figure 2.4b: CD spectra of EC1 in the absence and presence of DTT after incubation at 4 °C for 0, 14, or 28 days at pH 7.0. After 28 days, there is some loss of the β-sheet structure of EC1 in the absence of DTT as seen from the decrease in the minimum at 216 nm; addition of 2 mM DTT suppresses the secondary structural change of EC1.

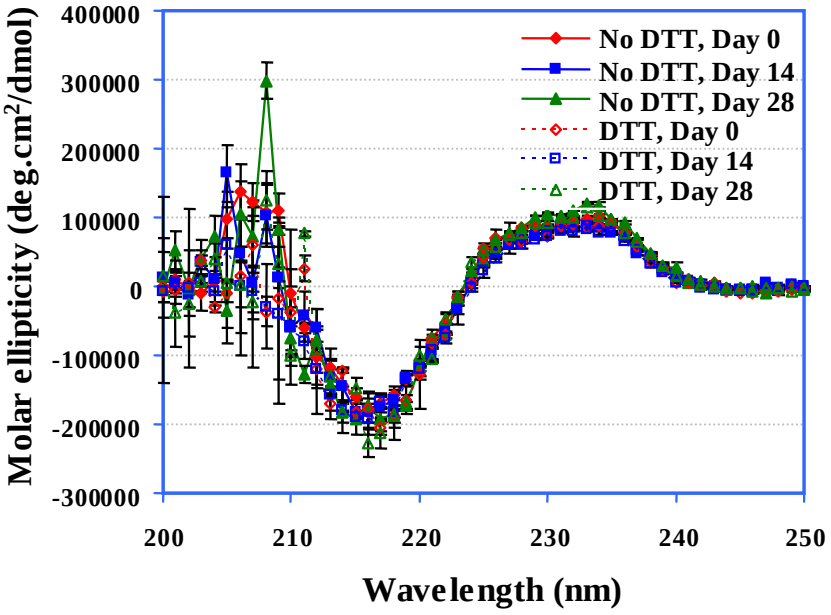

Figure 2.4c: CD spectra of EC1 in the absence and presence of DTT after incubation at 4 °C for 0, 14, or 28 days at pH 9.0. After 28 days, the CD spectra remain unchanged in the presence and absence of DTT.

secondary structure of EC1. Upon storage of EC1 for 14 days at 4 °C without DTT, its CD spectrum was dramatically changed compared to the original spectrum and showed increasing β-sheet character. This could be due to oligomerization via β-sheet intermolecular interactions. Addition of DTT suppressed the change in the CD spectra of EC1 between day 0 and day 14 (Figure 2.4a), suggesting that DTT inhibits the aggregation of the EC1 protein. At pH 7.0 (Figure 2.4b), the CD spectral changes were less dramatic compared to that at pH 3.0 (Figure 2.4a). The major changes in the CD spectra of EC1 at pH 7.0 after incubation for 14 and 28 days without DTT were in the minimum at 216 and the maxima at 205nm (Figure 2.4b). Addition of 2.0 mM DTT into the buffer suppressed the spectral change upon incubation for 28 days. There were no observable changes in the CD spectra of EC1 at pH 9.0 after 28 days in the presence and absence of DTT (Figure 2.4c).

The melting curves of EC1 were monitored at 218 nm and compared in the absence and presence of DTT (Figure 2.5a). At pH 3.0 without DTT, the melting curves of EC1 were dramatically different between day 0 and day 14, with a more negative ellipticity minimum observed at day 14 than that at day 0 (Figure 2.5a). The melting temperatures (T_m) were 28.8 °C on day 0 and 33.1 °C on day 14, indicating that the protein could form oligomers with a higher T_m. In contrast, the thermal unfolding curves of EC1 on days 0 and 14 were very similar in the presence DTT (Figure 2.5a). At the final melting temperature, the protein solution containing DTT has a lower minimum than that of protein without DTT, indicating that the β structure

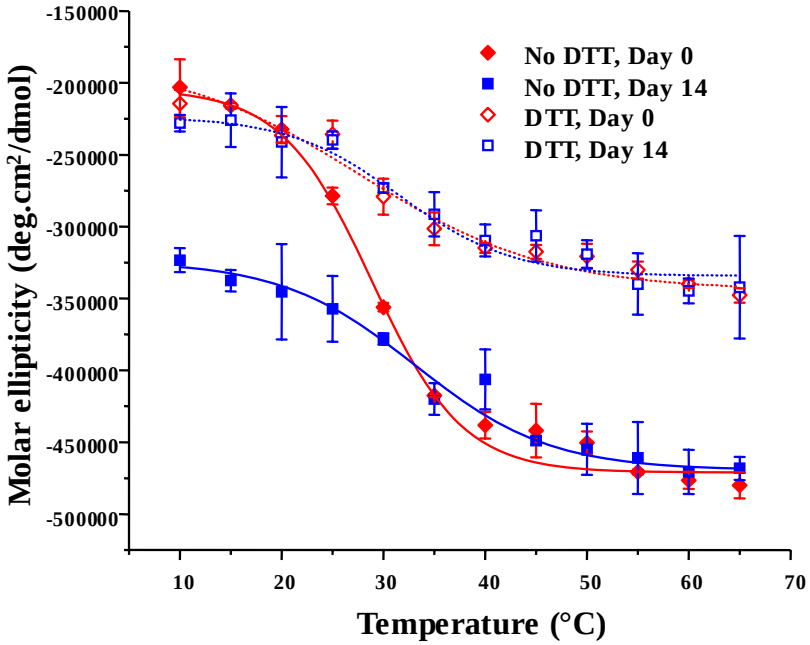

Figure 2.5a: Thermal unfolding profiles of EC1 in the absence and presence of DTT measured by CD at 218 nm after incubation for 14 to 28 days at 4 °C at pH 3.0. The thermal unfolding profiles in the absence of DTT are different on day 14 compared to that on day 0; in the presence of DTT, the profiles look the same on days 0 and 14.

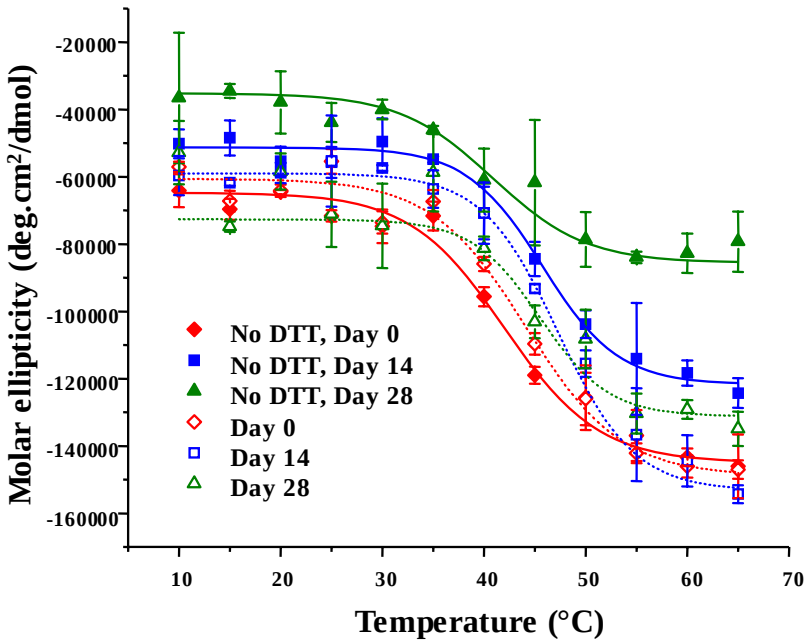

Figure 2.5b: Thermal unfolding profiles of EC1 in the absence and presence of DTT measured by CD at 218 nm after incubation for 14 to 28 days at 4 °C at pH 7.0. The thermal denaturation profiles in the absence of DTT are different at different time points; however, the profiles are relatively the same in the presence of DTT.

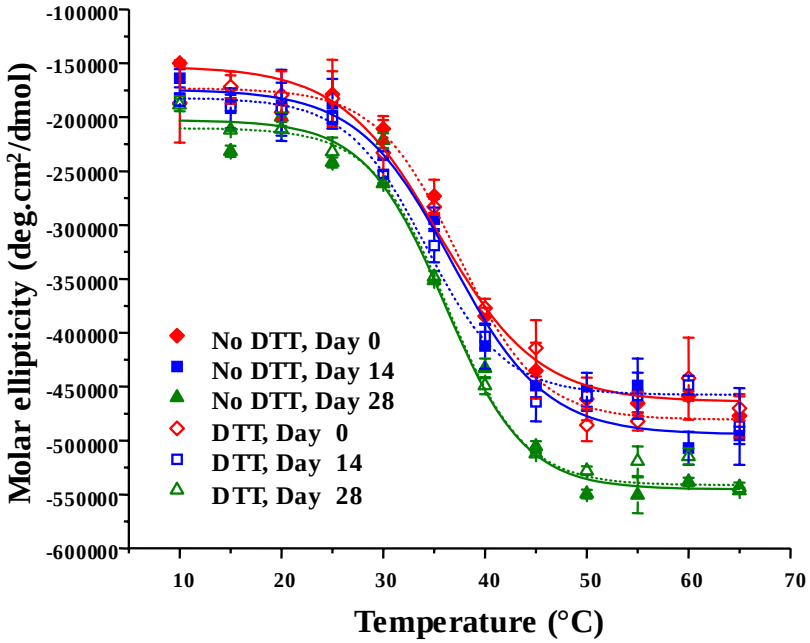

Figure 2.5c: Thermal unfolding profiles of EC1 in the absence and presence of DTT measured by CD at 218 nm after incubation for 14 to 28 days at 4 °C at pH 9.0. The thermal denaturation profiles are comparable after incubation for 14 and 28 days in the presence or absence of DTT.

of the protein at high temperature in the absence of DTT was increased compared to that in the presence of DTT.

At pH 7.0, thermal unfolding of EC1 at days 0, 14, and 28 in the presence of DTT did not show a difference in melting temperature (T_m = 44.6 °C, Figure 2.5b). Without DTT, thermal denaturation curves at pH 7.0 gradually shifted from day 0 to day 28 to the increased initial and final ellipticities; this result is different than the one observed at pH 3.0, suggesting that the structure of the intermediates at pH 7.0 is different than those at pH 3.0. The presence of DTT prevented the change in thermal unfolding curves over time, again suggesting that DTT prevented the physical degradation of EC1. Thermal denaturation of EC1 at pH 9.0 showed a T_m at 37.8 °C (Figure 2.5c) with a slightly higher T_m when incubated up to day 28. At pH 9.0, DTT did not have a dramatic effect on the melting curve profiles of the EC1 domain at the different time points, which remained around 38 °C.

2.3.2.2. Evaluation of tertiary structural changes by fluorescence spectroscopy: The intrinsic fluorescence emission spectrum of EC1 at pH 3.0 on day 0 without DTT showed a maximum emission at 339.93 ± 0.04 nm. There was no observable change upon incubation for 14 days (data not shown). After 28 days, a slight blue shift of 3.0 nm (337.09 ± 0.04 nm) in the maximum emission was observed, implying that the Trp residue environment became more apolar. In the presence of DTT on day 0, EC1 showed an emission profile (339.85 ± 0.04 nm) similar to that in the absence of DTT. A blue shift to 337.35 ± 0.04 nm was also observed after incubation for 28 days. At pH 7.0 without DTT, the emission maximum was 339.27 ± 0.03 nm; after incubation

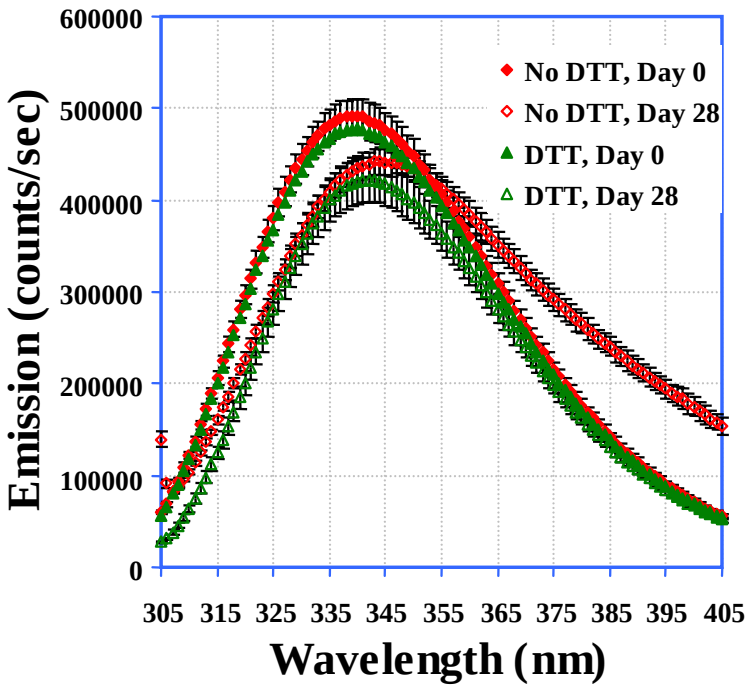

Figure 2.6a: Intrinsic fluorescence emission spectra of EC1 (after excitation at 295 nm) in the absence and presence of DTT after incubation at 4 °C for 0 or 28 days at pH 7.0. A shift in the emission maximum to a higher wavelength is observed in the absence of DTT; however, this shift is inhibited by addition of DTT.

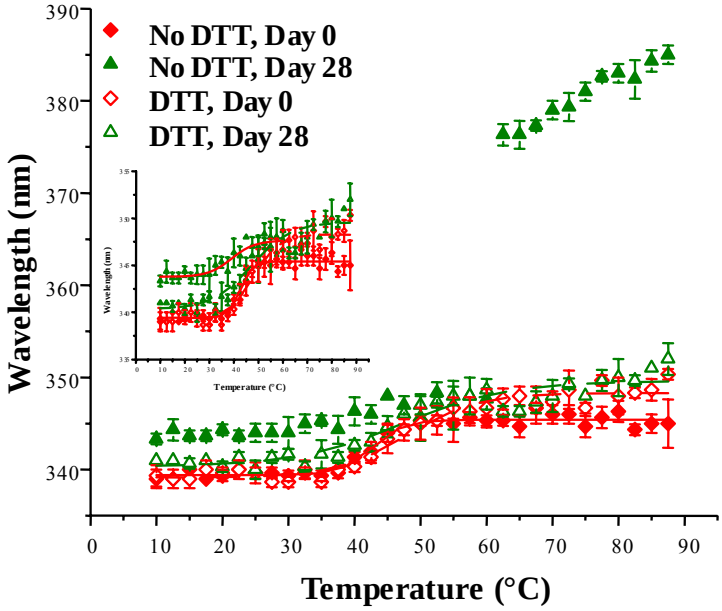

Figure 2.6b: Thermal unfolding of EC1 in the absence and presence of DTT evaluated by change in the wavelength of maximum intrinsic fluorescence emission (after excitation at 295 nm) from 0 °C to 87.5 °C after incubation for 0, 14, or 28 days at 4 °C and pH 7.0.

for 28 days, there was a considerable red shift in the emission maximum to 345.81 ± 0.13 nm, presumably due to increased solvent exposure of the Trp residues (Figure 2.6a). The presence of DTT prevented the red shift in the emission maximum from 339.26 ± 0.03 nm on day 0 to 340.66 ± 0.04 nm on day 28. Without DDT at pH 9.0, the intrinsic fluorescence spectra showed a slight blue shift from 341.32 ± 0.03 nm on day 0 to 340.32 ± 0.03 nm and 340.08 ± 0.03 nm on days 14 and 28, respectively (data not shown). DTT also did not prevent the shift in maximum emission from 341.06 ± 0.03 nm on day 0 to 341.19 ± 0.02 nm and 341.42 ± 0.03 nm on days 14 and 28, respectively. Thus, DTT could not prevent the change in EC1 tertiary structure upon incubation for 28 days at different pH values.

The thermal transitions of EC1 at pH 3.0, 7.0, and 9.0 were also monitored by fluorescence spectroscopy at the emission wavelength of 340 nm upon heating from 10 °C to 87.5 °C. At pH 3.0, there was no clear and observable transition seen in the absence or presence of DTT at all time points (data not shown). Distinct thermal transitions, however, were observed at pH 7.0 in both the absence and presence of DTT (Figure 2.6b). In the absence of DTT, the transition temperature decreased from 42.63 ± 0.56 °C at day 0 to 39.10 ± 2.31 °C at day 28. On the other hand, relatively constant transition temperatures were observed in the presence of DTT between day 0 (45.36 ± 0.79 °C) and day 28 (45.14 ± 0.65 °C). At pH 9.0, transitions were observed on day 0 in the absence and presence of DTT; these transitions did not change significantly after incubation for 28 days with and without DTT (data not shown).

2.3.3. Molecular Dynamics Simulations of EC1 Monomer and Dimer:

To explain the difference between the stability of EC1 monomer and dimer, 10 ns equilibrium simulations were conducted on both in solution. EC1 from the structure PDB code: 1EDH was used as the starting point. To explore their dynamic stability, root-mean-square displacement (RMSD) values for the protein C_α atoms were calculated and plotted during the production phase relative to the starting structures for the EC1 monomer (Figure 2.7, panel A) and dimer (Figure 2.7, panel B). The RMSD plots indicate that the structures of the EC1 monomer (Figure 2.7, panel A) achieve equilibrium much faster than those of the dimer (Figure 2.7, panel B); the dimer achieves equilibrium after approximately 3 ns of simulations. The RMSD fluctuations of the monomer are within 1.5 Å of the initial structure while the RMSD deviations of the dimer are in the range of 3.5–4.0 Å. Root-mean-square fluctuations (RMSF) were also plotted against the residue number for the monomer (Figure 2.8, panel A) and both chains of the dimer (Figure 2.8, panel B). The average RMSF value per residue of the dimer is larger than that of the monomer, demonstrating that the dimer has larger overall flexibility than the EC1 monomer. Structural snapshots of the monomer and dimer after 2, 5, 7, and 10 ns simulations are shown in Figures 2.9a and 2.9b, respectively. All four structures are very stable without noticeable differences. In contrast, there are significant differences among the snapshot structures of the dimer, with some of the regions showing increased dynamic properties. The increased dynamic properties of the disulfide-mediated EC1 dimer may explain the decreased chemical and physical stability of the dimer.

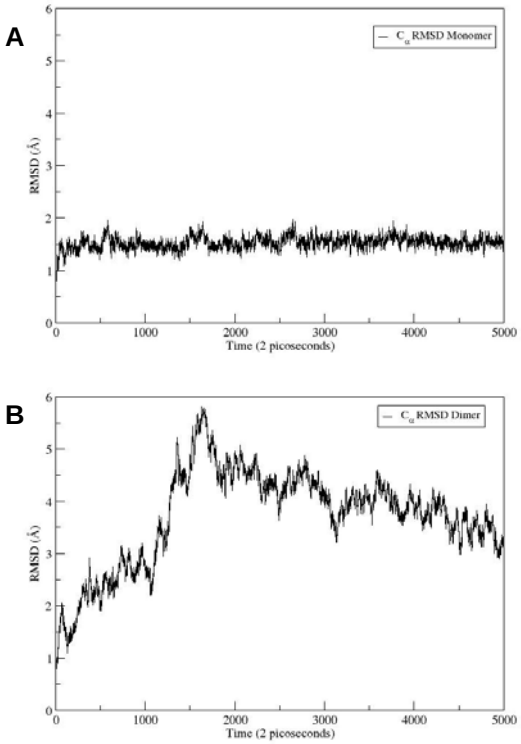

Figure 2.7: RMSD values for 10 ns simulation of EC1 monomer (panel A) and EC1 dimer (panel B).

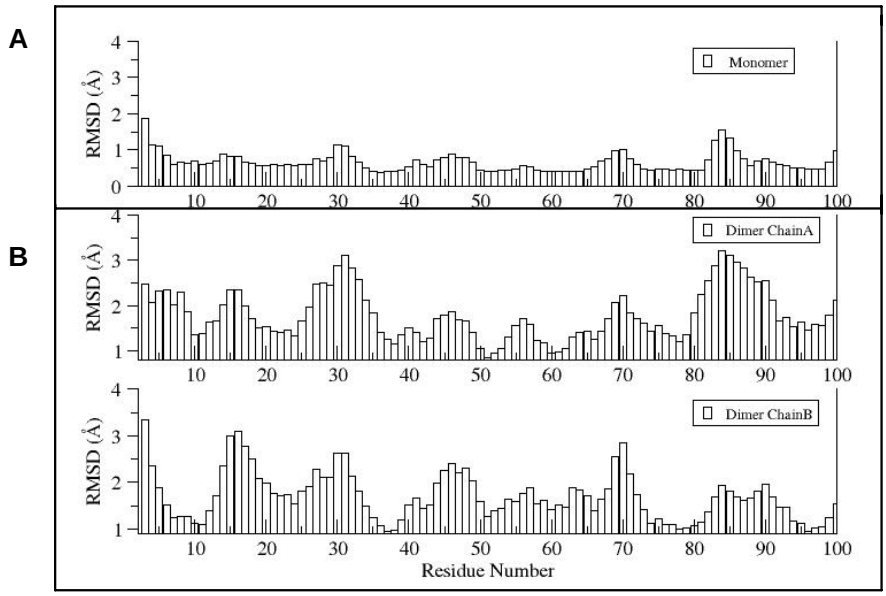

Figure 2.8: RMSF values for each residue in EC1 monomer (panel A) and in both A and B chains of EC1 dimer (panel B) at the end of 10 ns equilibrium simulations.

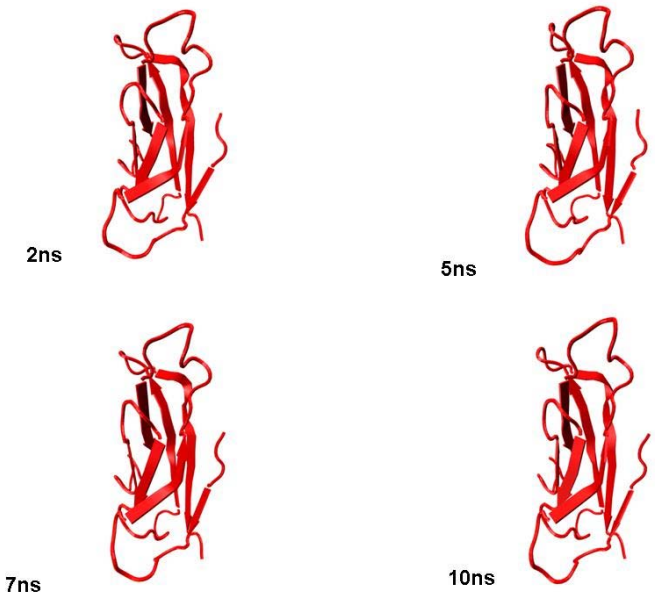

Figure 2.9a: Snapshots of EC1 monomer after 2, 5, 7 and 10 ns of molecular dynamics simulation.

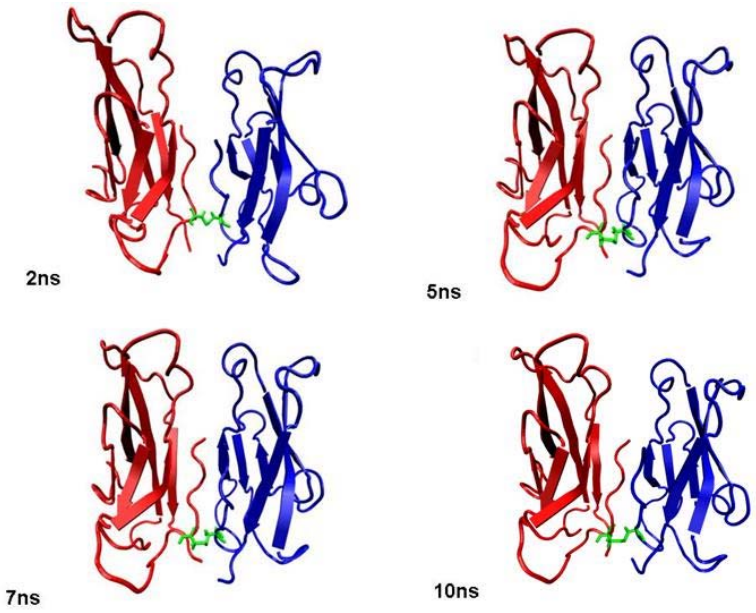

Figure 2.9b: Snapshots of EC1 dimer after 2, 5, 7 and 10 ns of molecular dynamics simulation.

2.4. Discussion

During the expression of the EC1 domain, the protein is found as a mixture of monomers, dimers, and oligomers, which can be separated by size exclusion chromatography (SEC).[13] The dimer is formed by an intermolecular disulfide bond between two Cys13 residues from each monomeric unit. It is proposed that the oligomers are formed by physical interaction of the dimer molecules. Under reducing conditions, the EC1 domain shows only one monomeric band upon SDS-PAGE, confirming that the dimer is the result of intermolecular disulfide formation at Cys13. Previously, we have also shown that an Ala mutation of the Cys residue prevents the formation of both dimers and oligomers.[13] The fact that EC1 monomers do not oligomerize before forming covalent dimers leads us to hypothesize that the structures of EC1 monomer and covalent dimer are different. In this work, we evaluated the chemical stability of EC1 and the changes in its structure in the absence and presence of the reducing agent, DTT.

In general, the accelerated thermal degradation mechanism of EC1 in solution involves chemical degradation and aggregation that leads to protein precipitation. The mass balance between the degradation products and the remaining starting material in solution did not add up to 100%; this difference in mass balance is presumably due to protein precipitation. The EC1 domain also showed hydrolysis products at all pH values at 70 °C while at 37 °C, this hydrolysis was only observed at pH 9.0. The

degradation products were the result of a peptide bond hydrolysis between Asp93 and Pro94 residues as identified by ESI MS. Base-catalyzed hydrolysis of the Asp-Pro peptide bond has been reported previously in peptides and proteins, including peptide bond cleavage of Asp133–Pro134 in recombinant human interleukin 11[19] and Asp169-Pro170 in recombinant human macrophage colony stimulating factor.[20] It has been previously reported that the peptide and protein degradation mechanisms at Asp residues were different at different pH values (acidic, neutral, and basic).[21,22] At pH < 5.0, the major degradation products are derived from peptide bond hydrolysis at the N- and C-termini of target Asp residue via the anchimeric assistance of the carboxylic acid side chain of the Asp residue. At neutral and basic conditions (pH > 6.0), the degradation mechanism is through a cyclic-imide intermediate formation that can produce iso-Asp as well as the related hydrolysis products.[21,22] Due to the absence of the amide hydrogen in the Asp-Pro peptide bond of EC1, the degradation of EC1 in neutral and basic conditions can not proceed through the cyclic-imide intermediate; thus, it was expected that Iso-Asp product would not be observed as a degradation product of EC1. An important finding here is that addition of DTT suppresses or completely eliminates the Asp93-Pro94 peptide bond hydrolysis. Because in the presence of DTT only the monomeric form exists in solution, the results suggest that the hydrolysis of the EC1 monomer may be catalyzed by dimeric and/or oligomeric forms of EC1. In other words, the monomeric form is chemically more stable than the dimer and oligomer forms; thus, the EC1 domain can be stabilized against peptide bond hydrolysis by trapping it in its monomeric form.

The effect of 4 °C storage on the secondary structure and physical stability of EC1 in the absence and presence of DTT was evaluated by far UV CD. EC1 shows a negative CD peak at 216 nm at all pH values, which is characteristic for β-sheet proteins; the X-ray[1] and NMR[4] structures of EC1 show that it has 7 β-strands and two small α-helices. After storage at 4 °C for 14 days at pH 3.0 in the absence of DTT, the CD spectrum changed dramatically with the peak at 216 nm becoming more negative. This could imply the formation of aggregates via the formation of intermolecular β-sheet interactions. In contrast, there was no change in the CD spectrum after 14 days incubation in the presence of DTT, indicating that there was no secondary structural change. It should be noted that addition of DTT at pH 3.0 changes the spectrum of EC1 on day 0, suggesting that DTT induces a secondary structural change in EC1 that prevents the physical degradation of EC1 at pH 3.0 upon 4 °C storage (Figure 2.5a). At pH 7.0 and 9.0, there were no observable changes in the CD spectra of EC1 with DTT on day 0 and day 28. Upon incubation of EC1 in the absence of DTT, there were observable changes in the CD spectrum on day 28 compared to that on day 0.

To further differentiate the effect of DTT on the thermal stability and possible oligomerization of EC1, thermal unfolding profiles of EC1 after incubation at 4 °C at 0, 14, and 28 days were determined by monitoring the CD spectra at 218 nm. At pH 3.0, thermal denaturation profiles at days 0 and 14 were very similar in the presence of DTT but were very different in its absence. This suggests that there are protein structural changes (i.e., oligomerization) upon incubation in the absence of DTT. In the absence of DTT at pH 7.0, CD spectra of EC1 changed after storage for 14 and 28

days; in contrast, there was no observable change in the presence of DTT. At pH 9.0, thermal unfolding properties of EC1 did not change significantly upon storage at 4 °C in the absence and presence of DTT. These results suggest that the secondary structures of monomer and covalent dimer are very similar at pH 9.0; however, they are different at lower pH.

Changes in the tertiary structure of EC1 upon incubation under different conditions were also monitored by observing the emission spectra of Trp6 and Trp63 in EC1 using fluorescence spectroscopy. The Trp6 residue is located close to Cys13. We hypothesize that the formation of intermolecular disulfide between two Cys13 residues can cause a change in the conformation of EC1 that alters the environment of one or both Trp residues. There was considerable change in the environment of at least one of the Trp residues over a period of 28 days incubation at 4 °C at pH 3.0 and 7.0. In the case of pH 7.0, there was a red shift indicating a more apolar environment for one or both Trp residue; in contrast, a blue shift was observed at pH 3.0, indicating a more apolar environment of the Trp residue. Thus, the resulting tertiary structures or association states of EC1 were different at these two pH values (i.e., 3.0 and 9.0). The addition of DTT prevented the fluorescence change at pH 7.0 but not at pH 3.0. This is in contrast to the CD data in which addition of DTT prevented a change in the secondary structure of EC1 at both pH 3.0 and 7.0. The Trp6 residue has been shown previously to be involved in domain swapping to form EC1 physical dimers, suggesting that the fluorescence changes seen could be due to the contribution of Trp6 domain swapping upon covalent dimer formation. At pH 3.0, the

formation of the intermolecular disulfide may be slowed, resulting in a predominance of domain swapped physical dimers, which subsequently oligomerize. In such structures, the Trp6 would be expected to be buried, and hence the blue shift at pH 3.0

Our findings show that the formation of a covalent dimer induces chemical and physical instability; in an attempt to explain this observation, the dynamic properties of the EC1 monomer and dimer were examined using computer simulations. Previously, molecular dynamics simulations were done to determine the dynamic properties of the EC1–2 portion of E-cadherin in the presence and absence of Ca^{2+} ions. These simulations confirmed that apo-cadherin shows much higher conformational flexibility on a nanosecond timescale than does the calcium-bound form.[23] The results from the current molecular dynamics simulations showed that the EC1 dimer had a greater mobility than did the EC1 monomer. In the EC1 dimer, some hydrophobic residues that would be expected to be folded into the core of the protein were found to be solvent-exposed at the end of the molecular simulations. Intermolecular interactions of these exposed hydrophobic residues on several EC1 dimers could be the driving force for the formation of oligomers in solution. These dramatic conformational changes were not observed in the simulation of the EC1 monomer; thus, keeping the EC1 domain in the monomeric form with DTT prevents its oligomerization. The formation of a covalent dimer could make the Asp93 and Pro94 residues more exposed to the solvent, leaving the Asp93–Pro94 peptide bond susceptible to hydrolysis. This supports our experimental observation that an increase

in dimeric forms enhances peptide bond hydrolysis products between Asp93 and Pro94. These molecular simulations data support the hypothesis that the formation of oligomers may be via the covalent EC1 dimer. The formation of a covalent dimer of EC1 induces a change in conformation that results in the Asp-Pro peptide bond becoming more solvent-exposed and thus susceptible to hydrolysis.

In conclusion, this study shows that EC1 degradation is mediated by the formation of an EC1 covalent dimer via its Cys13 residue. This covalent dimer induces a conformational change in EC1 that is prone to produce oligomers, which in turn lead to precipitation, and to facilitate chemical hydrolysis. The results of the molecular dynamics simulations support the experimental findings that the monomeric form is chemically more stable than the dimeric and oligomeric forms. Thus, prevention of covalent dimer formation by alkylating the thiol group could be an alternative method to stabilize the EC1 domain in solution without addition of a reducing agent. The alkylation of the thiol group of Cys13 will provide EC1 monomer that can potetntially be used to block E-cadherin-mediated homotypic and heterotypic cell-cell adhesion for pharmaceutical applications.

2.5. References

1. Nagar B, Overduin M, Ikura M, Rini JM 1996. Structural basis of calcium-induced E-cadherin rigidification and dimerization. Nature l380:360-364.

2. Nose A, Tsuji K, Takeichi M 1990. Localization of specificity determining sites in cadherin cell adhesion molecules. Cell l61:147-155.

3. Shapiro L, Fannon AM, Kwong PD, Thompson A, Lehmann MS, Grubel G, Legrand J-F, Als-Nielsen J, Colman DR, Hendrickson WA 1995. Structural basis of cell-cell adhesion by cadherins. Nature l374:327-337.

4. Overduin M, Harvey TS, Bagby S, Tong KI, Yau P, Takeichi M, Ikura M 1995. Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. Science lv267:p386(384).

5. Troyanovsky RB, Sokolov E, Troyanovsky SM 2003. Adhesive and lateral E-cadherin dimers are mediated by the same interface. Mol Cell Biol l23:7965-7972.

6. Klingelhofer J, Laur OY, Troyanovsky RB, Troyanovsky SM 2002. Dynamic interplay between adhesive and lateral E-cadherin dimers. Mol Cell Biol l22:7449-7458.

7. Zheng K, Trivedi M, Siahaan TJ 2006. Structure and function of the intercellular junctions: barrier of paracellular drug delivery. Current Pharmaceutical Design l12:2813-2824.

8. Lutz KL, Siahaan TJ 1997. Modulation of the cellular junctions protein E-cadherin in bovine brain microvessel endothelial cells by cadherin peptides. Drug Del l10:187–193.

9. Kobayashi N, Ikesue A, Majumdar S, Siahaan TJ 2006. Inhibition of E-cadherin-mediated homotypic adhesion of Caco-2 cells: a novel evaluation assay for peptide activities in modulating cell-cell adhesion. J Pharmacol Exp Ther l317:309–316.

10. Sinaga E, Jois SD, Avery M, Makagiansar IT, Tambunan US, Audus KL, Siahaan TJ 2002. Increasing paracellular porosity by E-cadherin peptides: Discovery of bulge and groove regions in the EC1-domain of E-cadherin. Pharm Res l19:1170–1179.

11. Makagiansar I, Avery M, Hu Y, Audus KL, Siahaan TJ 2001. Improving the selectivity of HAV-peptides in modulating E-cadherin-E-cadherin interactions in the intercellular junction of MDCK cell monolayers. Pharm Res l18:446–553.

12. Makagiansar IT, Ikesue A, Duc Nguyen P, Urbauer JL, Bieber Urbauer RJ, Siahaan TJ 2002. Localized production of human E-cadherin-derived first repeat in Escherichia coli. Protein Expression and Purification l26:449-454.

13. Makagiansar IT, Nguyen PD, Ikesue A, Kuczera K, Dentler W, Urbauer JL, Galeva N, Alterman M, Siahaan TJ 2002. Disulfide bond formation promotes the cis- and trans-dimerization of the E-cadherin-derived first repeat. Journal of Biological Chemistry l277:16002-16010.

14. Zabell KM, Laurence JS, Kinch MS, Knapp DW, Stauffacher CV 2006. Expression and purification of the intact cytoplasmic domain of the human ephrin receptor A2 tyrosine kinase in Escherichia coli. Protein Expression and Purification l47:210-216.

15. Derrick TS, Kashi RS, Durrani M, Jhingan A, Middaugh CR 2004. Effect of metal cations on the conformation and inactivation of recombinant human factor VIII. Journal of Pharmaceutical Sciences l93:2549-2557.

16. Rexroad J, Wiethoff CM, Green AP, Kierstead TD, Scott MO, Middaugh CR 2003. Structural stability of adenovirus type 5. Journal of Pharmaceutical Sciences l92:665-678.

17. Pertz O, Bozic D, Koch AW, Fauser C, Brancaccio A, Engel J 1999. A new crystal structure, Ca2+ dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. Embo J l18:1738-1747.

18. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K 2005. Scalable molecular dynamics with NAMD. J Comput Chem l26:1781-1802.

19. Kenley RA, Warne NW 1994. Acid-catalyzed peptide bond hydrolysis of recombinant human interleukin 11. Pharmaceutical Research l11:72-76.

20. Schrier JA, Kenley RA, Williams R, Corcoran RJ, Kim Y, Jr RPN, D'Augusta D, Huberty M 1993. Degradation pathways for recombinant human macrophage colony-stimulating factor in aqueous solution. Pharmaceutical Research l10:933-944.

21. Oliyai C, Borchardt RT 1993. Chemical pathways of peptide degradation. IV. Pathways, kinetics, and mechanism of degradation of an aspartyl residue in a model hexapeptide. Pharmaceutical Research l10:95-102.

22. Bogdanowich-Knipp SJ, Chakrabarti S, Williams TD, Dillman RK, Siahaan TJ 1999. Solution stability of linear vs. cyclic RGD peptides. J Pept Res l53:530–541.

23. Cailliez F, Lavery R 2005. Cadherin mechanics and complexation: The importance of calcium binding. The Biophysical Journal l89:3895-3903.

Chapter 3

Improving the stability of EC1 by modification of its Cys13 thiol group by alkylation

3.1. Introduction

E-cadherins are intercellular proteins found in the adherens junctions of the intestinal mucosa and the blood-brain barrier. They create cell-cell adhesion by *cis*- and *trans*-homophilic interactions. The *cis*-interactions are generated between E-cadherins on the same cell and the *trans*-interactions are produced by E-cadherins from opposing cells. It has been proposed that the *cis*-interactions are necessary for the formation of the *trans*-interactions that generate intercellular adhesion.[1,2] The extracellular portion of E-cadherin involved in the *cis*- and *trans*- interactions is composed of five domains, numbered EC1-EC5. EC1 is the N-terminal domain of E-cadherin and is located farthest from the cell surface. The EC1 domain has been shown to have an important role in the formation of *cis*- and *trans*-interactions. EC1 has also been shown to be important for the selectivity of E-cadherin in creating the homophilic interactions.[1-7] There are several proposed mechanisms of E-cadherin interactions that mediate cell-cell adhesion, and some of these mechanisms involve domain swapping between the Trp2 residue of EC1 from one E-cadherin molecule and a hydrophobic pocket of EC1 from another E-cadherin molecule. Due to the important role of the EC1 domain, we expressed and evaluated the recombinant human EC1 protein for its use as an inhibitor of E-cadherin-mediated cell-cell adhesion.

EC1 and peptides derived from EC1 have been shown to inhibit cell-cell adhesion mediated by homophilic interactions of E-cadherins.[8-11] Unfortunately, the EC1 protein is unstable during storage (*i.e.*, 4 weeks at 4 °C) after its purification. It

produces a covalent dimer via an intermolecular disulfide bond between the Cys13 residues of two monomers, followed by the generation of oligomers and precipitation. The presence of oligomers and precipitates may interfere with studies designed for the evaluation of the biological activity of EC1. The covalent dimerization and physical oligomerization of EC1 have been determined by SDS-PAGE and size exclusion chromatography (SEC),[12,13] as well as by monitoring changes in its far UV circular dichroism (CD) spectra and intrinsic fluorescence emission spectra after storage at 4 °C. Accelerated stability studies at different pH values (3.0, 7.0, and 9.0) and temperatures (37 °C and 70 °C) have shown that the EC1 protein undergoes hydrolysis at the D93–P94 peptide bond. Molecular dynamics simulation studies have indicated that the formation of the intermolecular disulfide at Cys13 of EC1 is a highly feasible reaction and that the covalent dimer has an altered structure, which is more dynamic than the structure of monomeric EC1. The increased flexibility of the EC1 dimer could explain the susceptibility of the D93–P94 peptide bond to hydrolysis. The oligomerization, precipitation and hydrolysis of the D93–P94 peptide bond of EC1 have been shown to be significantly inhibited or completely eliminated by the addition of dithiothreitol (DTT), a thiol reducing agent. The presence of thiol reducing agents (*i.e.*, DTT or glutathione) would be undesirable during biological activity studies (*i.e.*, inhibition of cell-cell adhesion) of EC1 since they may affect the results of the cell adhesion assays. Thus, an alternative method is needed to stabilize the monomeric form of the EC1 protein so it can be evaluated in biological assays and stored for a long period.

The formation of covalent dimers by Cys residues can be prevented by modifying the Cys thiol group irreversibly[14] or by mutating the Cys residue with an Ala, Ser, or Thr residue. Previously, we have shown that mutation of the Cys residue to an Ala in EC1 prevents the formation of the EC1 oligomers.[13] In this work, the thiol group of the Cys residue was modified to a thioether by reacting it with iodoacetic acid or iodoacetamide to produce the EC1-IA and EC1-IN derivatives, respectively. The reactions of Cys thiol with iodoacetic acid and iodoacetamide have been known for a long time[15,16] and have been extensively applied to modify protein structure and stability.[17-24] The iodoacetamide derivative of EC1 (EC1-IN) has the same total charge as the parent EC1 protein, but EC1-IA has an additional negative charge introduced in EC1. The secondary structure of EC1-IN was evaluated by far UV CD spectroscopy under different pH conditions. The chemical and physical stability of the EC1-IN derivative were compared to that of EC1. The EC1-IN derivative showed a better chemical and physical stability profile than did the parent EC1 only at pH 7.0.

3.2. Materials and methods

3.2.1. Production of Alkylated EC1:

The EC1 protein (M.W. = 11,628 Da) was expressed in *E. coli* and purified in the presence of β-mercaptoethanol (BME) as previously described in Chapter 2. The EC1 protein (0.2 mg/mL) was reacted in the dark with iodoacetamide (1.0 mM) for 4 h at pH 8.0 in a buffer containing 0.02 mM BME, 20 mM Tris and 4.0 M guanidine

hydrochloride (Gdn.HCl). The completion of the reaction was monitored and confirmed by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry with time-of-flight (TOF) detection. The alkylated EC1 (EC1-IN) was refolded by employing a step-wise dilution method in which the concentration of Gdn.HCl was reduced from 4.0 M to 0.5 M with a 0.5 M decrease in each dilution step. The dilution step consists of the addition of an appropriate volume of a buffer containing 20 mM Tris at pH 8.0, followed by incubation at 4 °C for 8 hours. When the Gdn.HCl concentration reached 0.5 M, dialysis procedure was performed with an 8-kD membrane to remove the Gdn.HCl, BME, and the unreacted iodoacetamide. The entire procedure described above was repeated by replacing iodoacetamide with iodoacetic acid to produce the EC1-IA derivative. Far UV CD spectroscopy was used to determine the success of protein refolding. The CD spectra of EC1-IA and EC1-IN were compared with the CD spectrum of EC1. The EC1-IN protein has a spectrum identical to that of EC1 at pH 7.0. The EC1-IA protein, however, has a spectrum substantially different form that of the native EC1; thus, it was not investigated further in the stability studies. The EC1-IN protein was evaluated further for its stability.

3.2.2. Chemical Stability Studies:

The EC1-IN protein was dialyzed into pH 3.0, pH 7.0, and pH 9.0 buffers and incubated at a concentration of 86 μM in sealed vials at 4 °C, 37 °C, and 70 °C for 4 hours in solutions buffered at pH 3.0, 7.0, and 9.0. The buffers used were 100 mM

phosphate for pH 3.0, 50 mM phosphate for pH 7.0, and 100 mM borate containing 0.08 M NaCl for pH 9.0. Samples were drawn at 0, 0.5, 1.0, 2.0, and 4.0 h and immediately stored at –70 °C until subsequent High Performance Liquid Chromatography (HPLC) analysis was performed. The HPLC was equipped with a C18 column (Varian Microsorb; pore-size: 300 Å; dimensions 250 X 4.6 mm) for seaparation of the analytes. The protein solution (25 μL) was injected into the column and eluted at a rate of 1.0 mL/min. A gradient elution method incorporating solvent A (94.9% double distilled water, 5% acetonitrile (ACN), and 0.1% TFA) and solvent B (100% ACN) was used. The gradient elution followed the sequence: 0% to 45% B from 0.0 to 2.0 min, 45% to 51% B from 2.0 to 12.0 min, at 51% B from 12.0 to 17.0 min, 51% to 100% B from 17.0 to 19.0 min, at 100% B from 19.0 to 21.0 min, and 100% B back down to 0% B from 21.0 to 22.0 min. A Varian Prostar UV detector at a wavelength of 220 nm was used to detect the proteins eluted from the column. MALDI-TOF MS analysis was used to determine the molecular weight of any new peaks observed on the HPLC chromatogram.

3.2.3. Physical Stability Studies:

Physical stability studies of EC1-IN were performed by monitoring the far UV CD spectra and the intrinsic fluorescence emission spectra on day 0 and after 14 and 28 days after incubation at 4 °C at pH 3.0, 7.0 and 9.0. The buffering agents used

were the same as described in the chemical stability studies section. Concentration of EC1-IN used in these studies was 86 μM.

3.2.3.1. Far UV CD studies: The secondary structural change of the EC1-IN protein was analyzed by far UV CD spectroscopy with a Jasco spectropolarimeter (J-720) equipped with a Peltier temperature controller by loading 300 μL of EC1-IN into a sealed CD cuvette with 0.1 cm pathlength.[25] The protein CD spectrum was measured at 10 °C between 200 and 250 nm. Triplicate spectra were obtained for EC1-IN, and the contribution of the buffer was eliminated by subtracting the the blank spectrum (containing the buffer alone) from the spectrum of protein in the same buffer. The same samples were also used for the thermal unfolding studies by recording their CD spectra at every 5 °C increase in temperature. The temperature of the cuvette holder was increased from 10 °C to 65 °C at a rate of 15 °C/h. The samples were equilibrated at the target temperature for 300 s before obtaining their spectra. The entire procedure was repeated after 14 and 28 days for EC1-IN samples stored at pH values of 3.0, 7.0, and 9.0 at 4 °C. The ellipticity at 218 nm (used to monitor changes in the β-sheet content) was plotted against the corresponding temperature and a sigmoidal function was used to fit the thermal unfolding curves using Origin 7.0 software. The midpoint of each transition in the sigmoidal fit was defined as the thermal unfolding temperature (T_m).

3.2.3.2. Intrinsic fluorescence emission studies: Changes in the tertiary structure of EC1-IN after 14 and 28 day incubation at 4 °C were monitored using fluorescence spectroscopy and compared to the emission profile on day 0. For each experiment, 90 μL of the EC1-IN protein (86 μM) was diluted to 0.9 mL with the appropriate buffer and transferred to a quartz fluorescence cuvette. The protein was excited at 295 nm (>95% Trp emission) and the fluorescence emission was observed between 305 and 405 nm using a PTI QuantaMaster spectrofluorometer.[26] Emission was set between 500,000 and 1,000,000 counts/sec by adjusting the excitation and emission slits. The emission spectra were obtained in triplicate and the blank spectrum was subtracted from each of them. The thermal unfolding of the EC1-IN protein was followed by detecting the change in the wavelength of maximum emission for every 2.5 °C increase in temperature. The temperature was increased at a rate of 15 °C/h and the samples were equilibrated at each analysis temperature for 300 s before collecting their spectra. Again, three samples containing EC1-IN and one containing the buffer alone were used to perform each thermal unfolding experiment. Each fluorescence emission spectrum obtained was fitted to an extreme assymetric peak function using the non-linear curve-fitting wizard of Origin 7.0 software to obtain the wavelength of maximum emission. The entire procedure mentioned above was repeated once for each pH studied.

3.3. Results

3.3.1. Alkylation of EC1:

Alkylation of the Cys thiol group of EC1 with iodoacetamide or iodoacetic acid is represented in figure 3a. The alkylation produced the desired product in a quantitative yield with one band in an SDS PAGE gel and a single peak on a reversed-phase HPLC C18 column. The products show the respective thioether of acetate (EC1-IA) and acetamide (EC1-IN) in the EC1 protein as determined by MALDI mass spectrometry (Figures 3.1b and 3.1c). The secondary structure of EC1-IN has characteristics similar to those of the parent EC1 at pH 7.0, showing a minimum at 216 nm and maxima at 235 and 205 nm (Figure 3.1d). The secondary structure of EC1-IN is sensitive to pH changes; the CD spectra of EC1-IN at pH 3.0 and 9.0 are different than its spectrum at pH 7.0 (Figure 3.1d). The minimum at 216 nm at pH 7.0 shifted to 210 nm at pH 3.0 and 9.0, and this minimum is more pronounced at pH 3.0 than at pH 9.0. Furthermore, the spectra at pH 3.0 and 9.0 lack the maxima at 205 and 235 nm. Thus, these results suggest that EC1-IN has a secondary structure similar to that of the parent EC1 at pH 7.0, but at pH 3.0 and 9.0, it has a different secondary structure than does EC1.

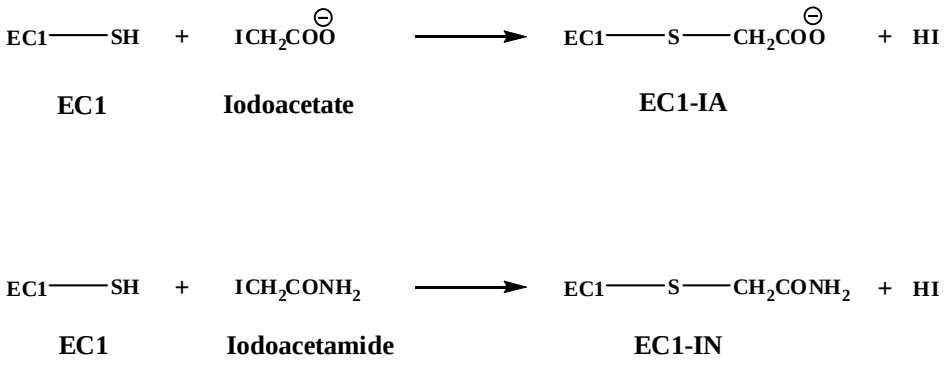

Figure 3.1a: The alkylation reaction of the thiol group of the Cys13 residue of EC1 with iodoacetate and iodoacetamide produces the EC1 derivatives EC1-IA and EC1-IN, respectively.

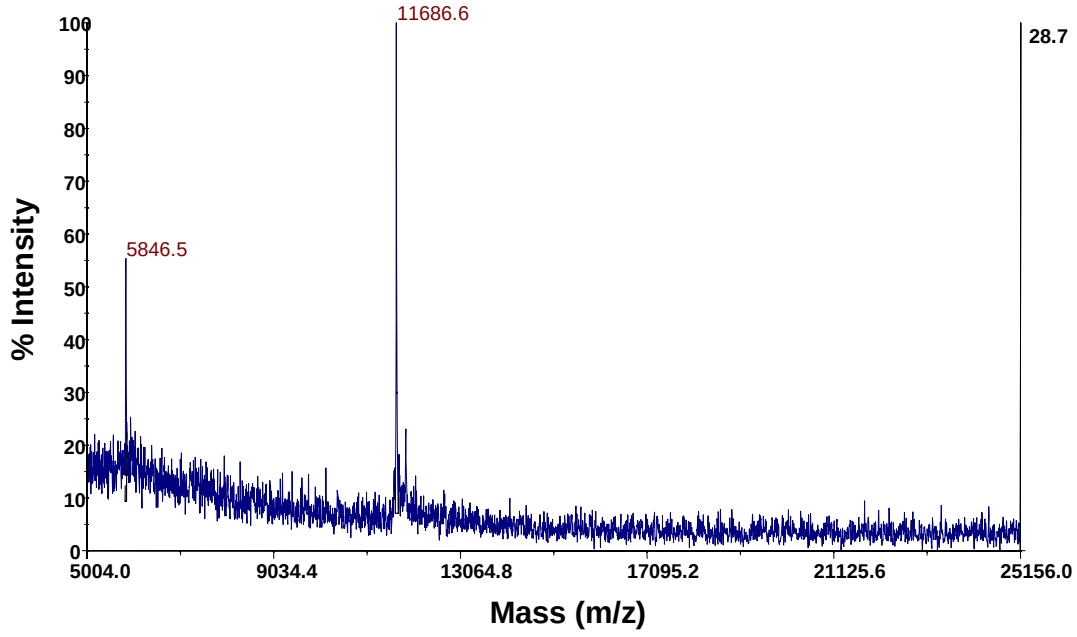

Figure 3.1b: MALDI-TOF spectrum of EC1 alkylated with iodoacetic acid (EC1-IA)

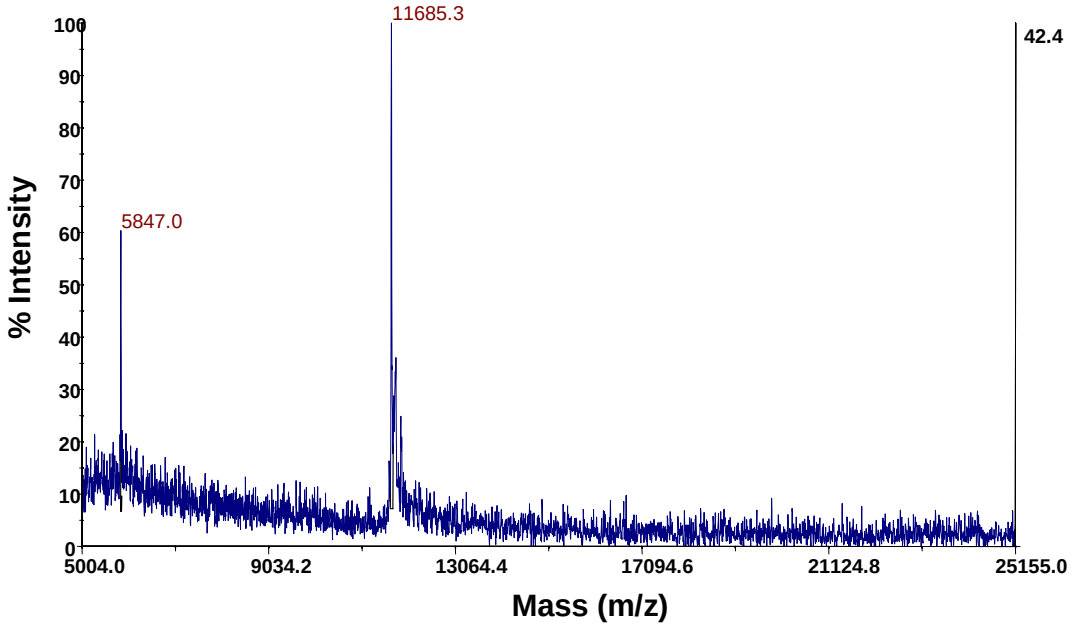

Figure 3.1c: MALDI-TOF spectrum of EC1 alkylated with iodoacetamide (EC1-IN)

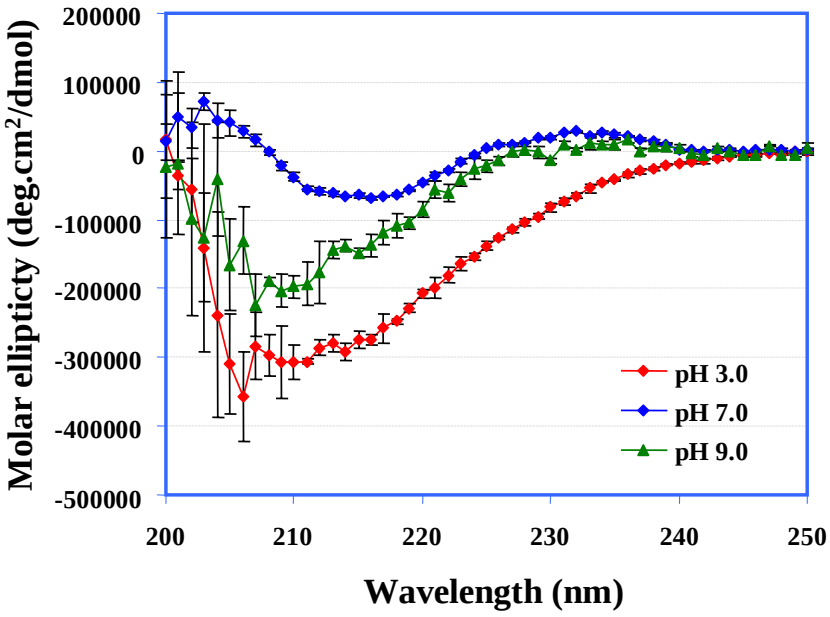

Figure 3.1d: The effect of pH on the secondary structure of EC1-IN as determined by CD spectroscopy.

3.3.2. Chemical stability of EC1-IN:

The chemical stability of the EC1-IN protein was monitored by HPLC after incubation for 4 h at pH 3.0, 7.0, and 9.0 at 4 °C. Under these conditions, there was no observable decrease in the chromatographic peak of the EC1-IN protein. The area under the peak of EC1-IN did not show substantial changes. There was no appearance of a new peak for any degradation product over the 4 h incubation period (data not shown). These results indicate that the EC1-IN protein is stable upon incubation for 4 h at 4 °C at all three pH values studied.

The EC1-IN protein was also chemically stable at 37 °C at pH 3.0, 7.0, and 9.0 (Figure 3.2a). Upon HPLC, there was no change in the area under the peak of EC1-IN and no new peaks representing degradation products were detected. In contrast, our previous studies showed that the parent EC1 was unstable when incubated for 4 h at pH 7.0 and 9.0 at 37 °C. Native EC1 degraded via hydrolysis of its D93-P94 peptide bond. These results imply that the EC1-IN protein is more stable to hydrolytic cleavage than the parent EC1 protein at 37 °C.

The EC1-IN protein was also stable for 4 h at pH 7.0 and 9.0 at 70 °C (Figure 3.2b), which is again in contrast to the results obtained for the stability study of the parent EC1 protein. Previously, the EC1 protein was shown to undergo hydrolysis of its D93-P94 peptide bond under similar conditions. At pH 3.0, however, a significant decrease of EC1-IN to 22% of the original amount was observed after 4 h incubation at 70 °C (Figure 3.2b). The disappearance of the HPLC peak of EC1-IN was accompanied by the appearance of a new peak with a molecular weight of 10,443.5

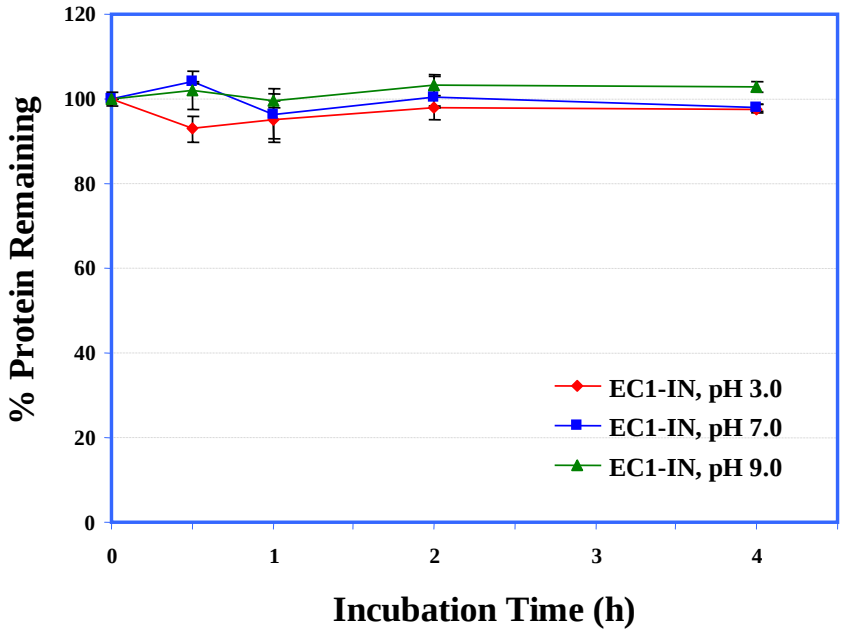

Figure 3.2a: The chemical stability profiles of EC1-IN after incubation for 4 h at pH 3.0, 7.0, and 9.0 at 37°C. At 37 °C, EC1-IN does not undergo any chemical degradation or precipitation at pH 3.0, 7.0, and 9.0.

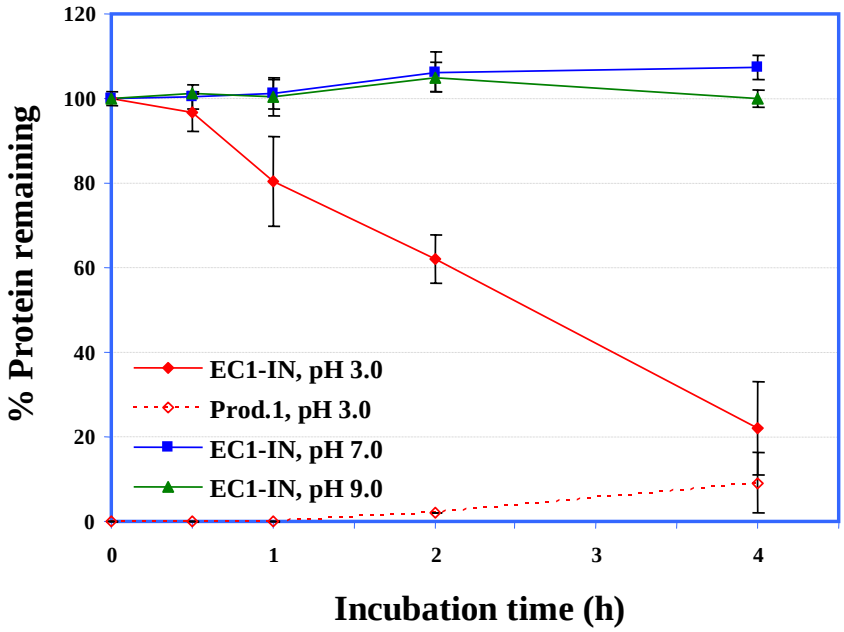

Figure 3.2b: The chemical stability profiles of EC1-IN after incubation for 4 h at pH 3.0, 7.0 and 9.0 at 70 °C. EC1-IN shows no peptide bond hydrolysis or precipitation at pH 7.0 and 9.0. At pH 3.0, however, EC1-IN undergoes peptide bond hydrolysis as well as precipitation. The product of hydrolysis was identified as the N-terminal fragment G1-D93.

Da (the M.W. was determined using MALDI-TOF MS). This fragment is the product of hydrolysis of EC1-IN at its Asp93–Pro94 peptide bond. This degradation product is similar to the degradation product observed for the parent EC1 protein. Moreover, at pH 3.0, the increase in the peak area of the hydrolysis product did not correspond quantitatively to the decrease in the peak area of the EC1-IN protein. This suggests that at pH 3.0, the degradation of EC1-IN also involves aggregation that leads to precipitation.

3.3.3. Physical stability of EC1-IN:

3.3.3.1. Secondary structure studies using CD spectroscopy: The effect of incubating EC1-IN at pH 3.0, 7.0 and 9.0 at 4 °C for up to 28 days on its physical stability was studied by observing changes in its CD spectra and comparing the data to the CD spectra of the parent EC1 protein. At pH 3.0, the CD spectra of the EC1-IN protein did not change after incubation for 14 days. The CD spectrum of EC1, however, changed from day 0 to day 14 with a shift in the minimum at 216 nm and disappearance of the maxima at 235 and 205 nm (Figure 3.3a), suggesting an increase in β-sheet character that could be due to the formation of oligomer. The melting curve of EC1-IN was compared to that of EC1 by monitoring the change in the CD signal at 218 nm (Figure 3.3b). At pH 3.0, there is no observable difference in the melting curve of EC1-IN after 14 day incubation, whereas a dramatic change is detected in

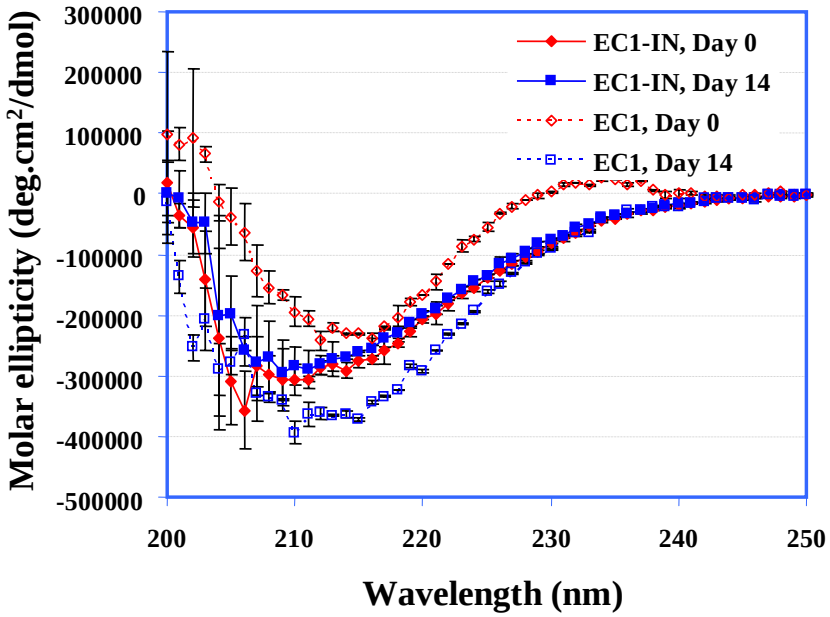

Figure 3.3a: Comparison of the CD spectra of EC1-IN and EC1 after their incubation at pH 3.0 for 14 days at 4 °C. The CD spectrum of EC1-IN does not change substantially after 14 days, whereas that of EC1 does.

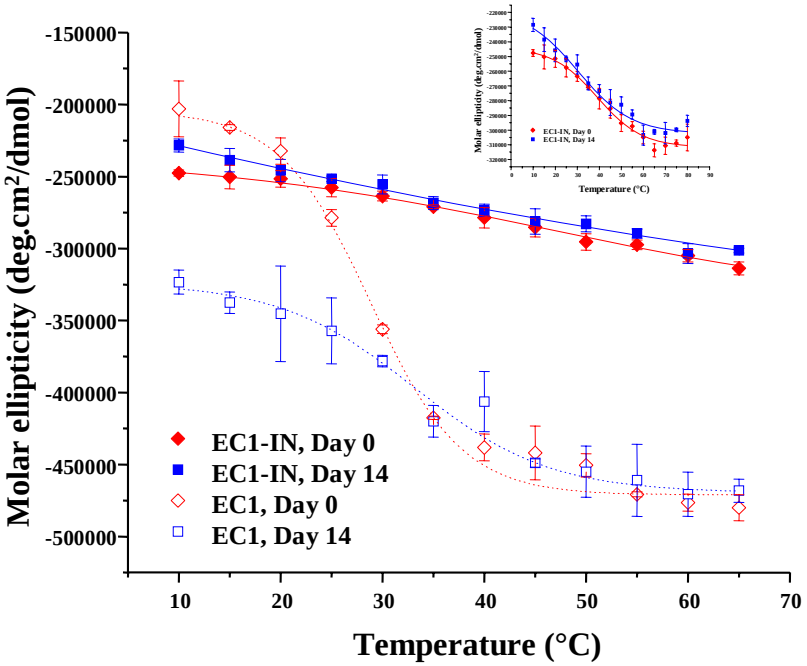

Figure 3.3b: Comparison of the thermal unfolding curves of EC1-IN and EC1 proteins after incubation for 14 days. There is no change in the thermal unfolding curves of EC1-IN whereas a clear change is seen in the thermal unfolding curve of EC1. The inset shows that the thermal unfolding curves of EC1-IN can be fitted to sigmoidal functions.

the melting curve of EC1 after 14 day incubation. These data imply that EC1-IN does not undergo changes, such as covalent dimerization or physical oligomerization when incubated at 4 °C for 14 days as does EC1.

At pH 7.0, the CD spectra of EC1-IN protein did not change upon incubation at 0, 14, and 28 days at 4 °C (Figure 3.4a). On the other hand, incubation of the EC1 protein under the same conditions produced changes in the CD spectra after 28 days (Figure 3.4b). The melting curves of EC1-IN after incubation for 0, 14, and 28 days did not change, but the melting curves of the EC1 protein showed dramatic changes (Figure 3.4c), suggesting that EC1-IN did not undergo physical degradation upon incubation.

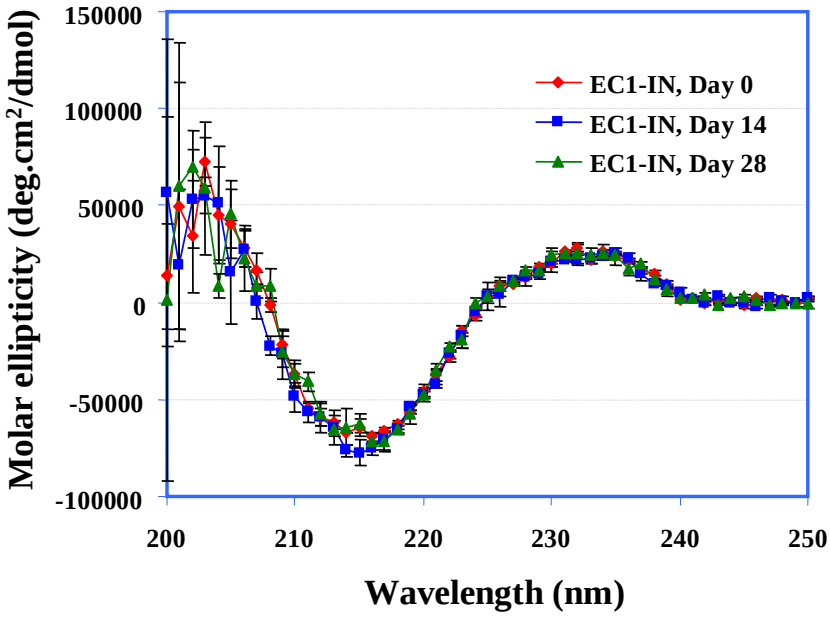

Figure 3.4a: The CD spectra of EC1-IN after its incubation at pH 7.0 for 14 and 28 days at 4 °C. The CD spectrum of EC1-IN does not change substantially after 14 and 28 days.

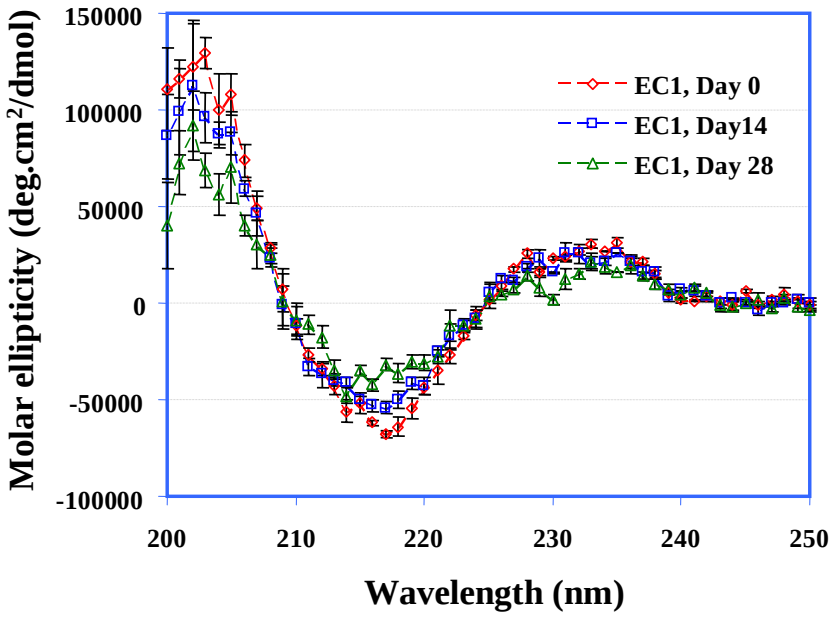

Figure 3.4b: The CD spectra of EC1 after its incubation at pH 7.0 for 14 and 28 days at 4 °C. The CD spectrum of EC1 changes substantially after 14 and 28 days, unlike that of EC1-IN shown in Figure 3.4a.

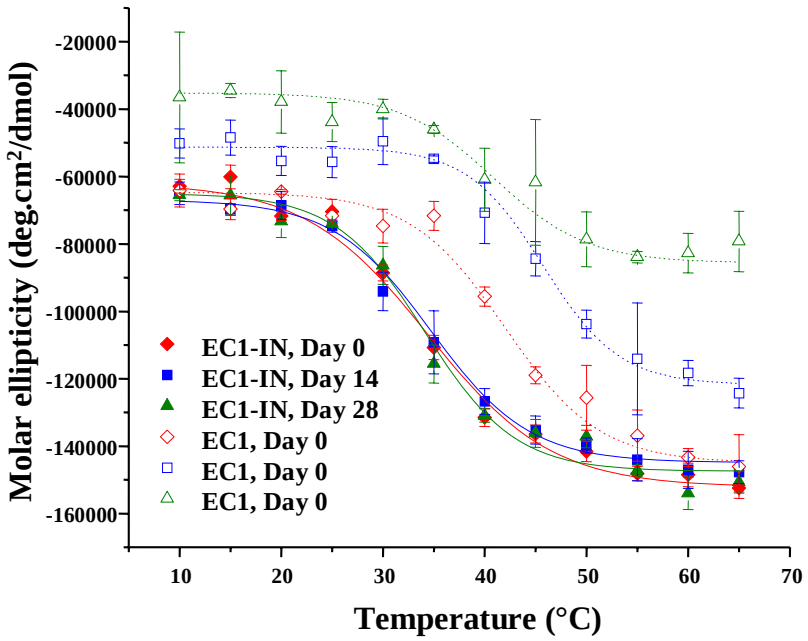

Figure 3.4c: Thermal unfolding profile of EC1 and EC1-IN measured by the CD signal at 218 nm after incubation for 0, 14 and 28 days at 4 °C at pH 7.0. There are limited differences in the thermal unfolding curves of EC1-IN, but dramatic differences are seen in the thermal unfolding curves of the EC1. EC1 has a higher transition temperature than EC1-IN, indicating that EC1-IN is less stable to thermal unfolding than EC1.

At pH 9.0, the CD spectra of the EC1-IN protein after 14 and 28 day incubation showed a dramatic shift of the minimum at 210 nm to 202 nm, indicating a change in the secondary structure upon incubation (Figure 3.5a). Conversely, the CD spectrum of EC1 did not change after incubation for 28 days (data not shown). The melting curves of both EC1-IN and EC1 did not show any dramatic change after different incubation periods. Overall, alkylation of the Cys13 residue of EC1 with iodoacetamide improves its physical stability at pH 3.0 and 7.0. This is probably due to the prevention of covalent dimer formation of EC1-IN. EC1-IN, however, is physically less stable than EC1 at pH 9.0.

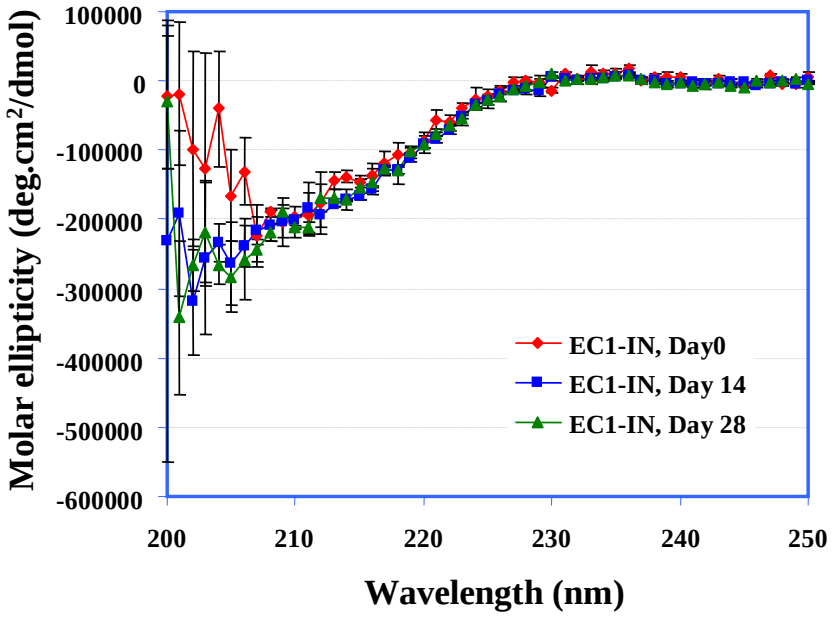

Figure 3.5a: The CD spectra of EC1-IN after its incubation at pH 9.0 for 14 and 28 days at 4 °C. The CD spectrum of EC1-IN after 14 and 28 days is substantially different from that on day 0.

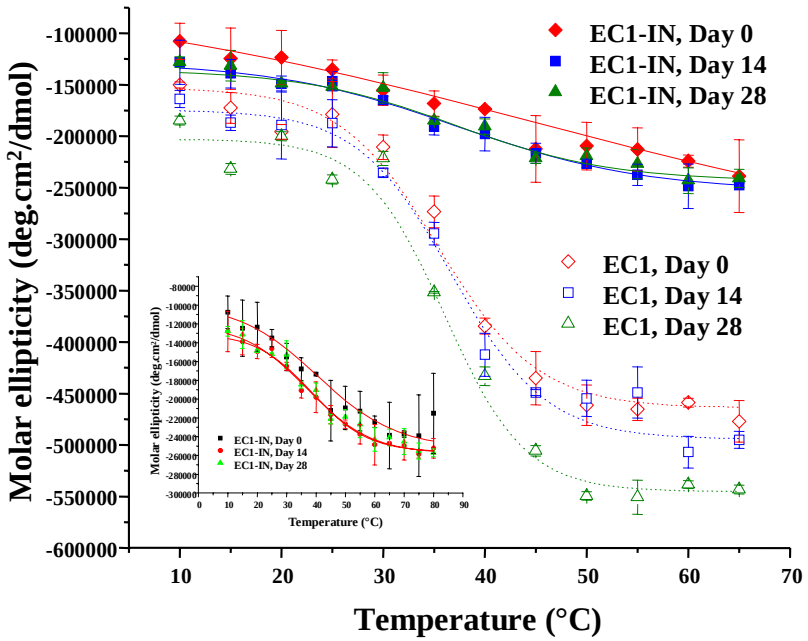

Figure 3.5b: Thermal unfolding profile of EC1 and EC1-IN measured by the CD signal at 218 nm after incubation for 0, 14 and 28 days at 4 °C at pH 9.0. The thermal unfolding curves of EC1-IN are substantially different from those of EC1.The inset shows that the thermal unfolding of EC1-IN can be fit to sigmoidal functions.

3.3.3.2. Tertiary structure studies using fluorescence spectroscopy: Intrinsic fluorescence emission spectroscopy was employed to monitor changes in the microenvironment of the Trp residues in EC1-IN and EC1. The Trp emission profile indicates how buried or exposed the Trp residues of the protein are. This, in turn, can be an indicator of the state of the tertiary structure of the protein being studied. The primary structure of EC1 contains two Trp residues located at positions 6 and 66 from the N-terminus Trp6 and Trp 66. At pH 3.0, EC1-IN exhibited fluorescence emission spectra with maximum intensity at wavelengths 340.61 ± 0.08 nm on day 0 and 338.91 ± 0.06 nm on day 28. This amounts to a blue shift of about 0.7 nm. The fluorescence emission spectra of EC1 showed a more pronounced blue shift of almost 3 nm in the wavelength of maximum emission from 339.93 ± 0.04 nm to 337.09 ± 0.04 nm after 28 day incubation at 4 °C. A blue shift indicates a possible burial of Trp residues, which could be a result of protein aggregation. Hence, our data suggest that at pH 3.0, EC1 undergoes more aggregation than EC1-IN.

A similar comparison of the wavelength of maximum intensity between the fluorescence emission spectra of EC1-IN and EC1 at pH 7.0 showed a small blue shift of approximately 1 nm from 342.13 ± 0.05 nm on day 0 to 341.21 ± 0.05 nm on day 28 for EC1-IN, but a large red shift of more than 6 nm from 339.27 ± 0.03 nm to 345.81 ± 0.13 nm for EC1 (Figure 3.6a). Thermal unfolding curves of EC1-IN and EC1 were obtained by plotting the wavelength of maximum intensity against temperature. The thermal unfolding profiles for EC1-IN at pH 7.0 on day 0 and day 28 show one observable transition around 35 °C (Figure 3.6b). In contrast, EC1 has

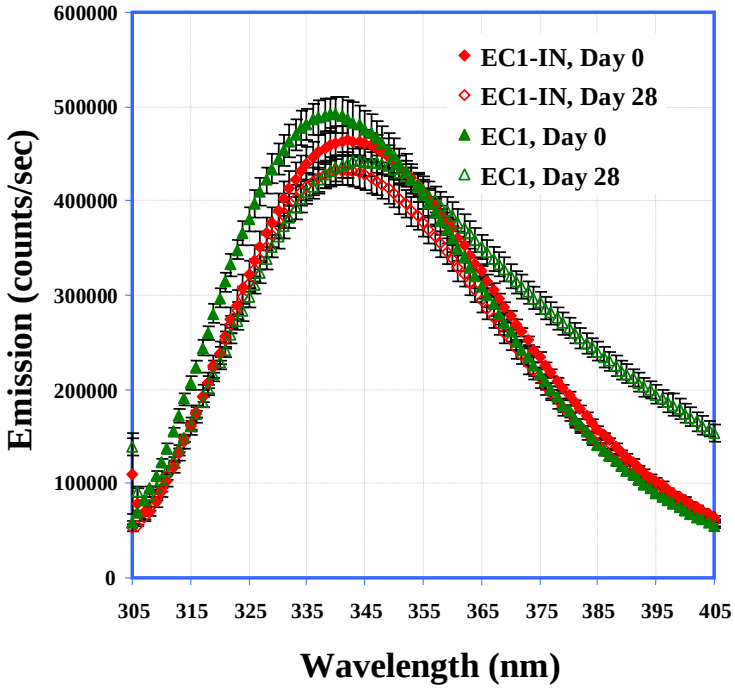

Figure 3.6a: Comparison of the intrinsic fluorescence emission spectra of EC1 and EC1-IN (after excitation at 295 nm) after incubation at 4°C for 28 days at pH 7.0. A red shift of 4.33 nm in the wavelength of maximum emission is observed for EC1; however, such a shift is not observed for EC1-IN.

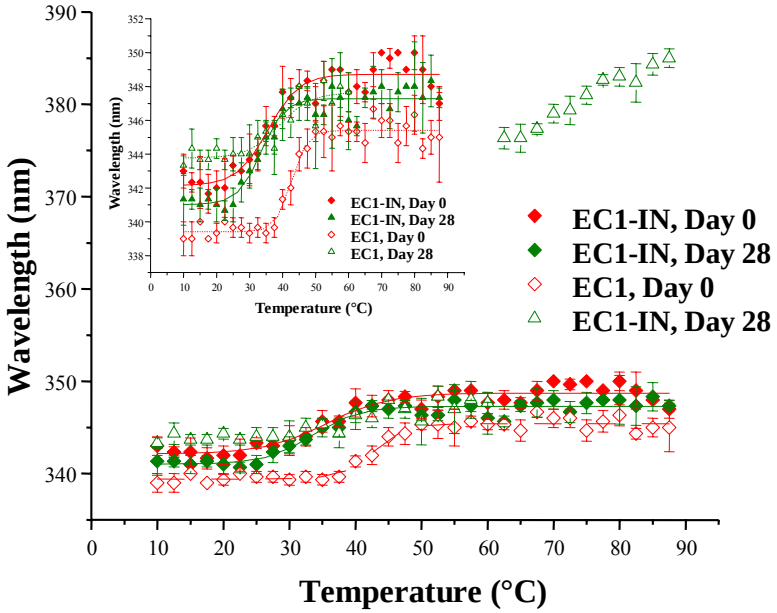

Figure 3.6b: Thermal unfolding of EC1 and EC1-IN evaluated by intrinsic fluorescence emission after incubation for 0 and 28 days at 4 °C and pH 7.0. EC1-IN has a lower transition temperature than EC1. After 28 days incubation at 4 °C, EC1 shows a transition to wavelengths > 375 nm at higher temperatures, while EC1-IN does not show such a transition. The inset shows the same data, magnified by plotting a smaller wavelength range on the y-axis.

one transition around 35 °C on day 0, but displays two transitions after incubation for 28 days at 4 °C with the second transition to very high wavelengths (>375 nm) above 60 °C.

At pH 9.0, the fluorescence emission spectra of EC1-IN and EC1 do not show significant differences. The wavelength of maximum EC1 emission shifts from 342.0 ± 0.0 nm on day 0 to 339.67 ± 1.15 nm on day 28. For EC1-IN emission spectra, the corresponding change is from 343.0 ± 1.0 nm on day 0 to 341.67 ± 0.58 nm. Overall, the fluorescence data obtained suggest that at pH 3.0, 7.0 and 9.0, the microclimate of the Trp residues in EC1-IN does not change significantly during an incubation period of 28 days at 4 °C. In contrast, emission spectra of EC1 at pH 3.0 and 7.0 and its thermal unfolding at pH 7.0 showed significant changes in the same incubation conditions.

3.4. Discussion

EC1 is a protein of interest because it has an important role in intercellular adhesion in epithelial and endothelial junctions. Improving the physicochemical stability of the EC1 protein is necessary for studying its binding properties to E-cadherin peptides and other EC domains of E-cadherin. The EC1 domain of endothelial E-cadherin also has a role in the inflammatory process; it binds to $\alpha_E\beta_7$ integrins on T-lymphocytes during their extravasation through the vascular endothelium.[27] Thus, the EC1 protein has potential uses as a therapeutic agent to block T-cell-endothelium adhesion during inflammatory responses. It is, therefore,

necessary to stabilize the EC1 protein for its long-term use in *in vitro* and *in vivo* biological assays.

The EC1 domain of E-cadherin has previously been shown to form covalent dimers *in vitro* through intermolecular disulfide bonds between the Cys13 residues of EC1 monomers.[13] The covalent dimers subsequently oligomerize via physical interactions that lead to precipitation.[13] It has also been shown conclusively that covalent dimer formation is necessary for the oligomerization process. Blocking the covalent dimerization with reducing agents (i.e., DTT) suppresses the oligomerization of EC1.[13] Stabilization of EC1 with DTT, however, would not be a preferred strategy because it may interfere with the biological evaluation of EC1. To block intermolecular disulfide bond formation, we decided to modify the thiol group of the Cys13 residue with iodoacetate or iodoacetamide to form the alkylated products, EC1-IA and EC1-IN, respectively. Although the X-ray structure of EC1 shows that the Cys13 residue is exposed on the surface of EC1, the alkylation reaction of the thiol group of the folded form of EC1 was slow. The alkylation reaction was accelerated in the presence of 4 M Gdn.HCl. The subsequent slow removal of Gdn.HCl refolded the iodoacetamide derivative (EC1-IN) to a structure similar to the parent EC1 protein as determined by CD. Unfortunately, the iodoacetate derivative (EC1-IA) did not refold correctly since it failed to exhibit the expected β-sheet character similar to that of the parent protein (EC1). This may be due to the additional negative charge on EC1-IA, which could prevent its proper refolding by destabilizing a refolding intermediate through charge repulsion with other negatively charged

residues (such as Asp or Glu) or by stabilizing an undesirable intermediate through charge-charge attractive interactions (*i.e.* salt bridge). A survey of the literature shows that additional negative charges introduced by modification of proteins with iodoacetate can result in significant differences in protein structure and activity between the iodoacetate- and iodoacetamide- modified proteins. In a study on the S8C mutant of sorghum phosphoenolpyruvate carboxylase (PEPC), modification of the C8 residue with iodoacetate decreased the sensitivity of the mutant to L-malate by 50%.[22] This decrease in the activity of the mutant PEPC was attributed to the additional negative charge introduced by modification with iodoacetate. In contrast, modification of the C8 residue with iodoacetamide, which did not introduce any additional charges in the protein, did not change the sensitivity of the mutant PEPC to L-malate. Due to the difficulty in folding EC1-IA, only the stability of EC1-IN was evaluated in this work.

The secondary structure of EC1-IN was similar to EC1 at pH 7.0 as shown in the similarity of the CD spectra of both proteins. The estimated secondary structure content of EC1-IN was 47% β-sheet, 5% α-helix, and 48% random coil and that of EC1 was 51% β-sheet, 2% α-helix, and 47% random coil. These values are well within the experimental error of the estimation method. The CD spectra of EC1-IN, were, however, dramatically different from those of the EC1 protein at pH 3.0 and 9.0. Altered secondary structures of proteins due to carboxymethylation or carboxyamidomethylation have been previously reported. Alkylation of the thiol groups of ovalbumin with iodoacetate and iodoacetamide resulted in a structural

transition from α-helix to β-sheet.[20] In another study, alkylation of completely reduced bovine serum albumin with iodoacetate and iodoacetamide led to a significant decrease in the α-helical content and a corresponding increase in β-sheet and random coil.[21] The CD spectra of EC1-IN and EC1 imply that EC1-IN is more stable than EC1 after their incubation at pH 3.0 and 4 °C for 14 days (Figure 3.2a). This finding is corroborated by the changes in the fluorescence emission spectra of EC1-IN and EC1. There is no change in the emission spectra of EC1-IN, but there is a blue shift in the spectra of EC1 after incubation at 4 °C for 28 days. The melting temperature (T_m) of EC1-IN estimated from the change in its β-sheet character with increase in temperature is higher than the T_m of EC1 at both pH 3.0 and 9.0. At pH 3.0, the T_m of EC1-IN is 39.9 ± 1.0 °C, whereas the T_m of EC1 is 28.853 ± 0.27 °C. At pH 7.0, however, the T_m of EC1-IN (33.9 ± 0.7 °C) is lower than the T_m of EC1 (42.0 ± 0.6 °C). Decreased thermal stability of a protein due to its alkylation has been reported previously. For example, thermal unfolding curves for alpha alpha tropomyosin (AAT) obtained by monitoring the α-helical change in protein structure with increasing temperature revealed that AAT modified by iodoacetate and iodoacetamide showed decreased thermal stability compared to that of unmodified AAT.[24] Although EC1-IN has a lower T_m at pH 7.0, there is no change in its CD spectra after incubation for 28 days. In contrast, these incubation conditions cause a change in the secondary structure of the parent EC1. The advantage of alkylating EC1 is that the alkylation maintains the conformation of EC1 during long-term, low

temperature incubation, possibly by preventing formation of covalent dimers, which have an altered structure.

The EC1-IN protein also has a higher chemical stability than that of native EC1 with the peptide bond hydrolysis reaction only observed during incubation of the protein at pH 3.0 and 70 °C. In this case, the amount of EC1-IN was decreased to 22% after 4 h incubation. Because the total amount of EC1-IN (22%) and the degradation product (10%) does not add up to 100%, the remaining 78% of EC1-IN could be in the form of insoluble material. Compared to EC1-IN, the EC1 is very unstable at 70 °C at all pH values studied. The precipitation of EC1 or EC1-IN may occur due to formation of an insoluble intermediate during their thermal unfolding process at 70 °C. At 37 °C, EC1-IN was stable to chemical degradation after 4 h at pH values of 3.0, 7.0 and 9.0. In contrast, EC1 is chemically stable only at pH 3.0 but not at pH 7.0 and 9.0, indicating that unmodified EC1 is more prone to form a covalent dimer at neutral and basic pH. Because alkylation of the thiol group enhanced the chemical and physical stability of EC1-IN compared to EC1 and addition of DTT to EC1 improved its stability, these results support the proposed hypothesis that the degradation of EC1 is mediated by dimer formation. Alkylating the thiol group thus eliminates the formation of the unstable covalent dimer. With the help of computational modeling, we have shown previously that the dimerization of EC1 is a facile reaction and it causes a different conformation of the two EC1 chains in the dimer compared to that of the monomeric form of EC1. The EC1 dimer is also more flexible than the monomer as shown by molecular dynamics simulations, thus

making the D93–P94 peptide bond susceptible to hydrolysis. Addition of DTT maintains EC1 in its monomeric form and decreases the hydrolysis of the D93–P94 bond as well as maintains its structure during long-term (28 days) incubation at 4 °C.

The intrinsic fluorescence emission spectra of EC1 at pH 7.0 previously had revealed a red shift of 4.33 nm from day 0 to day 28 in its wavelength of maximum emission (Figure 3.5a). The thermal unfolding of EC1 monitored by observing the change in wavelength of maximum emission with temperature also showed a significant difference between unfolding profiles on day 0 and day 28 (Figure 3.5b). This suggested a change in the microenvironment of at least one of the two Trp residues of EC1 and, therefore, a change in the tertiary structure of EC1. In contrast, the intrinsic fluorescence emission spectra of EC1-IN at pH 7.0 did not show a red shift from day 0 to day 28 (Figure 3.5a). The thermal unfolding profiles of EC1-IN at pH 7.0 on day 0 and day 28 were also very similar (Figure 3.5b), suggesting that there was no observable change in the tertiary structure of EC1.

In conclusion, modification of EC1 by alkylating the thiol group improves the chemical stability of EC1-IN as well as its primary, secondary and tertiary structure during long-term storage at low temperatures at pH 7.0. At pH 3.0 and 9.0, however, the secondary structure of EC1-IN was significantly different than that of EC1. EC1-IN also precipitated at pH 3.0 at high temperature. We hypothesize that modification of the Cys thiol of EC1 with a hydrophilic polymer such as PEG may decrease the precipitation of EC1 and EC1-IN observed at pH 3.0. It has been shown that PEGylation of a protein does not disrupt its structure, thereby maintaining its

activity.[28,29] We therefore, performed site-specific PEGylation of the Cys13 thiol of EC1 and studied its stability profile. These results are described in Chapter 4.

3.5. References

1. Shapiro L, Fannon AM, Kwong PD, Thompson A, Lehmann MS, Grubel G, Legrand J-F, Als-Nielsen J, Colman DR, Hendrickson WA 1995. Structural basis of cell-cell adhesion by cadherins. Nature l374:327–337.

2. Overduin M, Harvey TS, Bagby S, Tong KI, Yau P, Takeichi M, Ikura M 1995. Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. Science lv267:p386(384).

3. Nagar B, Overduin M, Ikura M, Rini JM 1996. Structural basis of calcium-induced E-cadherin rigidification and dimerization. Nature l380:360–364.

4. Klingelhofer J, Laur OY, Troyanovsky RB, Troyanovsky SM 2002. Dynamic interplay between adhesive and lateral E-cadherin dimers. Mol Cell Biol l22:7449–7458.

5. Nose A, Tsuji K, Takeichi M 1990. Localization of specificity determining sites in cadherin cell adhesion molecules. Cell l61:147–155.

6. Troyanovsky RB, Sokolov E, Troyanovsky SM 2003. Adhesive and lateral E-cadherin dimers are mediated by the same interface. Mol Cell Biol l23:7965–7972.

7. Zheng K, Trivedi M, Siahaan TJ 2006. Structure and function of the intercellular junctions: barrier of paracellular drug delivery. Current Pharmaceutical Design l12:2813–2824.

8. Kobayashi N, Ikesue A, Majumdar S, Siahaan TJ 2006. Inhibition of E-cadherin-mediated homotypic adhesion of Caco-2 cells: a novel evaluation assay for

peptide activities in modulating cell-cell adhesion. J Pharmacol Exp Ther l317:309–316.

9. Lutz KL, Siahaan TJ 1997. Modulation of the cellular junctions protein E-cadherin in bovine brain microvessel endothelial cells by cadherin peptides. Drug Del l10:187–193.

10. Makagiansar I, Avery M, Hu Y, Audus KL, Siahaan TJ 2001. Improving the selectivity of HAV-peptides in modulating E-cadherin-E-cadherin interactions in the intercellular junction of MDCK cell monolayers. Pharm Res l18:446–553.

11. Sinaga E, Jois SD, Avery M, Makagiansar IT, Tambunan US, Audus KL, Siahaan TJ 2002. Increasing paracellular porosity by E-cadherin peptides: Discovery of bulge and groove regions in the EC1-domain of E-cadherin. Pharmaceutical research l19:1170–1179.

12. Makagiansar IT, Ikesue A, Duc Nguyen P, Urbauer JL, Bieber Urbauer RJ, Siahaan TJ 2002. Localized production of human E-cadherin-derived first repeat in Escherichia coli. Protein Expression and Purification l26:449–454.

13. Makagiansar IT, Nguyen PD, Ikesue A, Kuczera K, Dentler W, Urbauer JL, Galeva N, Alterman M, Siahaan TJ 2002. Disulfide bond formation promotes the cis- and trans-dimerization of the E-cadherin-derived first repeat. Journal of Biological Chemistry l277:16002–16010.

14. Pace CN, Grimsley GR, Thomson JA, Barnett BJ 1988. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. J Biol Chem l263:11820–11825.

15. Dickens F 1933. Interaction of halogenacetates and SH compounds

The reaction of halogenacetic acids with glutathione and cysteine. The mechanism of iodoacetate poisoning of glyoxalase. Biochem Journal l27:1141–1151.

16. Smythe CV 1936. The reaction of iodoacetate and iodoacetamide on various sulfydyl groups, with urease, and with yest preparations. Journal of Biological Chemistry l114:601–612.

17. Caccia P, Nitti G, Cletini O, Pucci P, Ruoppolo M, Bertolero F, Valsasina B, Roletto F, Cristiani C, Cauet G, Sarmientos P, Malorni A, Marino G 1992. Stabilization of recombinant human basic fibroblast growth factor by chemical modifications of cysteine residues. European Journal of Biochemistry l204:649–655.

18. Jörnvall H, Fowler AV, Zabin I 1978. Probe of beta-galactosidase structure with iodoacetate. Differential reactivity of thiol groups in wild-type and mutant forms of beta-galactosidase. Biochemistry l17:5160–5164.

19. Schrooyen PMM, Dijkstra PJ, Oberthur RC, Bantjes A, Feijen J 2000. Partially carboxymethylated feather keratins. 1. Properties in aqueous systems. J Agric Food Chem l48:4326–4334.

20. Batra PP, Sasa K, Ueki T, Takeda K 1989. Circular dichroic study of conformational changes in ovalbumin induced by modification of sulfhydryl groups and disulfide reduction. J Protein Chem l8:609–6117.

21. Batra PP, Sasa K, Ueki T, Takeda K 1989. Circular dichroic study of the conformational stability of sulfhydryl-blocked bovine serum albumin. Int J Biochem l21:857–862.

22. Duff SM, Lepiniec L, Crétin C, Andreo CS, Condon SA, Sarath G, Vidal J, Gadal P, Chollet R 1993. An engineered change in the L-malate sensitivity of a site-directed mutant of sorghum phosphoenolpyruvate carboxylase: the effect of sequential mutagenesis and S-carboxymethylation at position 8. Archives of biochemistry and biophysics l306.

23. Erneux DCaC 1996. Identification of an active site cysteine residue in human type I Ins(1,4,5)P3 5-phosphatase by chemical modification and site-directed mutagenesis. Biochem J l320:181–186.

24. Holtzer ME, Holtzer A, Crimmins DL 1990. The effect of sulfhydryl blocking groups on the thermal unfolding of alpha alpha tropomyosin coiled coils. Biochem Biophys Res Commun l166:1279–1283.

25. Derrick TS, Kashi RS, Durrani M, Jhingan A, Middaugh CR 2004. Effect of metal cations on the conformation and inactivation of recombinant human factor VIII. Journal of Pharmaceutical Sciences l93:2549–2557.

26. Rexroad J, Wiethoff CM, Green AP, Kierstead TD, Scott MO, Middaugh CR 2003. Structural stability of adenovirus type 5. Journal of Pharmaceutical Sciences l92:665–678.

27. Karecla PI, Green SJ, Bowden SJ, Coadwell J, Kilshaw PJ 1996. Identification of a binding site for integrin alphaEbeta7 in the N-terminal domain of E-cadherin. Journal of Biological Chemistry l271:30909–30915.

28. Digilio G, Barbero L, Bracco C, Corpillo D, Esposito P, Piquet G, Traversa S, Aime S 2003. NMR Structure of two novel polyethylene glycol conjugates of the

human growth hormone-releasing factor, hGRF(1-29)-NH_2. J Am Chem Soc l125:3458–3470.

29. Hinds KD, Kim SW 2002. Effects of PEG conjugation on insulin properties. Advanced Drug Delivery Reviews l54:505–530.

Chapter 4

PEGylation of the Cys13 thiol of EC1 improves its stability while retaining its structure

4.1. Introduction

E-cadherins are cell adhesion glycoproteins located in the adherens region of intercellular junctions of the epithelial or endothelial tissues of biological barriers. The extracellular portion of E-cadherin consists of five extracellular (EC) domains (EC1-to-EC5) that are involved in mediating cell-cell adhesion by *cis*- and *trans*-interactions. The *cis*-interactions (*cis*-dimers) are formed between two extracellular domains protruding from the same cell, and the *trans*-interactions are produced by protein-protein interactions from the opposing cells.[1,2] The EC1 domain at the N-terminal of E-cadherin has been shown to be important for the homophilic specificity of E-cadherin interactions in the *adherens* junctions.[1-7] The EC1 protein and peptides derived from EC1 can inhibit E-cadherin-mediated cell-cell adhesion. The EC1 protein, however, is very unstable upon long-term storage, contributing to difficulties in studying its *in vitro* and *in vivo* biological activities.

The instability of the EC1 protein is due to covalent dimer formation followed by oligomerization and precipitation; thus, any method that prevents the dimerization of the EC1 protein might be beneficial for future studies of the EC1 domain. The covalent dimers of EC1 are produced by the formation of disulfide bonds between Cys13 residues of EC1 monomers; this is followed by oligomerization of the covalent dimers through non-covalent interactions. In addition, this protein also undergoes hydrolysis of its D93–P94 peptide bond. The presence of covalent dimers, physical oligomers, and hydrolysis products of EC1 in solution was characterized with the use of SDS-PAGE, size exclusion chromatography (SEC) and mass spectrometry.[8,9]

Molecular dynamics studies suggest that the formation of a covalent dimer makes the protein more dynamic, thereby contributing to the instability of the protein. The stability of the EC1 protein was shown to improve with the addition of dithiothreitol (DTT) to the protein solution as well as by modification of its Cys13 thiol group to thioacetamide (EC1-IN).

In Chapter 3, we showed that the EC1-IN protein has the same conformation as the EC1 protein at pH 7.0. Its secondary structure, however, is sensitive to pH changes; at pH 3.0, and 9.0, its secondary structure is significantly different from that at pH 7.0. As an alternative approach to alkylation of the EC1 thiol, modification of the thiol group with polyethylene glycol (PEG) was attempted to improve the stability of EC1.[10] PEGylation of therapeutic proteins has been shown to improve their physical and biological properties and lower their immunogenicity. PEGylation of a protein generally leads to an increase in its hydrodynamic radius, resulting in lower glomerular filtration and also increases its enzymatic stability. Thus, a higher systemic circulation of the parent protein is achieved, leading to a reduction in the dosing amount and dosing frequency. It has been proposed that protein PEGylation enhances the hydrophilic hydration shell around the protein thereby increasing the physicochemical stability of the protein.

In this work, the EC1 protein was modified by alkylating the thiol group of Cys13 with a PEG group activated with a maleimide functional group. Conjugation via the Cys residue is more selective than using the amino group of the Lys residue or the carboxylic acid groups of the Asp and Glu residues. Because EC1 has 7 Lys, 6 Asp,

and 9 Glu residues, the conjugation of a PEG group to these residues can produce a mixture of products with different numbers of attached PEG molecules. Due to the presence of only one Cys in the EC1 sequence, the PEGylation of EC1 produces a single product (PEGEC1), which has one molecule of PEG covalently linked to a molecule of EC1. PEGEC1 can no longer form disulfide-mediated covalent dimers. The structural characteristics of PEGEC1 were analyzed by circular dichroism (CD) and fluorescence spectroscopy, and the chemical stability of PEGEC1 was compared with that of the parent EC1 protein. The PEGEC1 shows a better physicochemical stability than the parent EC1.

4.2. Methods

4.2.1. Conjugation of EC1 with PEG:

Maleimide-PEG (M.W. = 5,000 Da) was obtained from Sigma Chemicals Co. 250 mg of maleimide-PEG was added to 50 mL of 20 mM Tris buffer at pH 8.0 containing 0.2 mg/mL (0.17 mM) of recombinant human EC1 and 0.2 mM DTT. The reaction between the maleimide-PEG and EC1 was allowed to proceed by incubating this solution at 22 °C. After 2 hours, 50 μL of this solution was subjected to MALDI-TOF MS analysis. The remaining solution was concentrated to 10 mL by passing it through a 10,000 Da cutoff membrane in an Amicon centrifuge tube and centrifuging at 4000 X g. The concentrated solution was stored at 4 °C until it was purified by (SEC).

4.2.2. Purification of the PEGEC1 conjugate:

Three mL of the above concentrated solution was injected onto a Superdex 200 size exclusion column. The mobile phase was 20 mM Tris buffered at pH 8.0. The column volume was 150 mL and the flow rate was 0.5 mL/min. The volume of the fractions collected was 3 mL. The separated PEG-EC1 conjugate (PEGEC1), EC1, and the maleimide-PEG were detected as separate peaks by their UV absorbance at 280 nm. To estimate the yield of PEGylation, the unreacted EC1 was quantified by the area under its SEC peak and subtracted from the area under the peak of EC1 on the SEC before the reaction. From these calculations, the yield of the PEGylated product was determined to be approximately 75%. The PEGEC1 fractions were pooled and concentrated to 7.5 mL. The concentration of PEGEC1 in this final solution was equal to 1 mg/mL of EC1 (86 μM). Fifty μL of this PEGEC1 solution was analyzed by MALDI-TOF analysis. The remaining PEGEC1 was used for chemical stability studies using high performance liquid chromatography (HPLC) analysis and structural studies using far UV CD spectroscopy and intrinsic fluorescence emission spectroscopy.

4.2.3. Chemical stability studies of PEGEC1:

Eighty six μM of PEGEC1 (equivalent to 1 mg/mL EC1) was dialyzed into solutions buffered at pH 3.0, 7.0 and 9.0 and incubated at 70 °C for 4 h. The composition of the buffers was 100 mM phosphate for pH 3.0, 50 mM phosphate for pH 7.0, and 100 mM borate containing 0.08 M NaCl for pH 9.0. After incubation at

70 °C for 4 h, the samples were transferred to a –70 °C freezer and stored until further analysis. HPLC was employed for the analysis of these samples. The HPLC system used was a Dynamax SD-200. The column used was a Varian Microsorb C18 (pore size 300Å, dimensions: 250 X 4.6 mm). Gradient elution at a rate of 1 mL/min was used for the separation. The two mobile phases used were A: 94.9% water, 5.0% acetonitrile, 0.1% TFA and B: 100% acetonitrile. A Varian Prostar UV detector was used to detect the eluted peaks at a wavelength of 220 nm.

4.2.4. Structural studies of PEGEC1:

Far UV CD spectroscopy and intrinsic fluorescence emission spectroscopy were used to characterize the secondary and tertiary structure, respectively, of PEGEC1.

4.2.4.1. Far UV CD spectroscopy: A Jasco spetropolarimeter (J-720) equipped with a Peltier temperature controller was used to perform the CD spectroscopy studies. The CD spectra of PEGEC1 were obtained at 10 °C between 200 and 250 nm at pH 3.0, 7.0, and 9.0. For evaluation at each pH, 300 mL of 86 μM PEGEC1 was transferred to a CD cuvette of pathlength 0.1 cm. Three such cuvettes were used to obtain the CD spectra and the spectra were averaged after subtracting the blank signal from each spectrum. The blank signal was the CD spectrum obtained for a cuvette containing the corresponding buffer alone. Thermal unfolding studies were also performed for PEGEC1 at pH 3.0, 7.0, and 9.0.[11] The temperature of the cuvette holder was increased from 10 °C to 65 °C at the rate of 15 °C/h, and the CD spectra of three cuvettes containing PEGEC1 in a buffer and one cuvette containing the

buffer alone were recorded. Thermal unfolding profiles were obtained for PEGEC1 at each pH studied by plotting the CD signal at 218 nm (to monitor changes in the β-sheet content) against the corresponding temperature. The thermal unfolding plots were fitted to sigmoidal functions using Origin 7.0. The midpoint of each plot was used as the thermal unfolding temperature (T_m) of PEGEC1.

4.2.4.2. Intrinsic fluorescence emission spectroscopy: A PTI Quantamaster spectrofluorometer was used to study the intrinsic fluorescence emission of PEGEC1. 900 μL aliquots of 8.6 μM PEGEC1 at pH 3.0, 7.0, or 9.0 were transferred to a sealed quartz fluorescence cuvette. Three cuvettes containing protein and an additional cuvette containing the corresponding buffer were prepared and loaded into four position cuvette holder in the spectrofluorometer and equilibrated at 10 °C. These samples were then excited with UV light of wavelength 295 nm (>95% Trp emission), and their emission spectra were recorded between 305 and 405 nm.[12] Thermal unfolding of PEGEC1 also was studied at pH 3.0, 7.0, and 9.0 by increasing the temperature of the cuvette chamber from 10 °C to 87.5 °C at the rate of 15 °C/h, and the emission spectra were recorded at every 2.5 °C increase in temperature. After subtracting the blank emission signal from each PEGEC1 spectrum, the wavelength of maximum emission at each temperature point was plotted against the corresponding temperature. This plot, which provides the thermal unfolding profile of PEGEC1, was then fitted to a sigmoidal function using Origin 7.0 software. The wavelength of maximum emission for each fluorescence emission spectrum was

obtained by fitting the spectrum to an extreme asymmetric peak function using the non-linear curve-fitting wizard of Origin 7.0 software. Thermal unfolding profiles for PEGEC1 were plotted for each pH studied.

4.3. Results

The reaction between Cys thiol of EC1 and maleimide-activated PEG is shown in Figure 4.1a. The MALDI spectrum of the reaction mixture is shown in Figure 1b. The PEGEC1 was separated from EC1 and PEG by SEC, as mentioned previously (Figure 4.1c). The HPLC peak of purified PEGEC1 (8.1 min) is shown in comparison to that of EC1 (8.4 min) and maleimide-activated PEG (7.9 min) (Figure 1d).

Reaction:

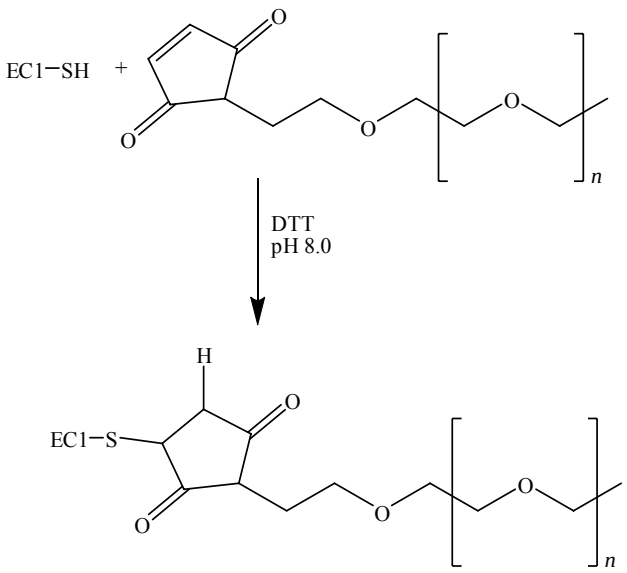

Figure 4.1a: Reaction of EC1 with maleimide-PEG.

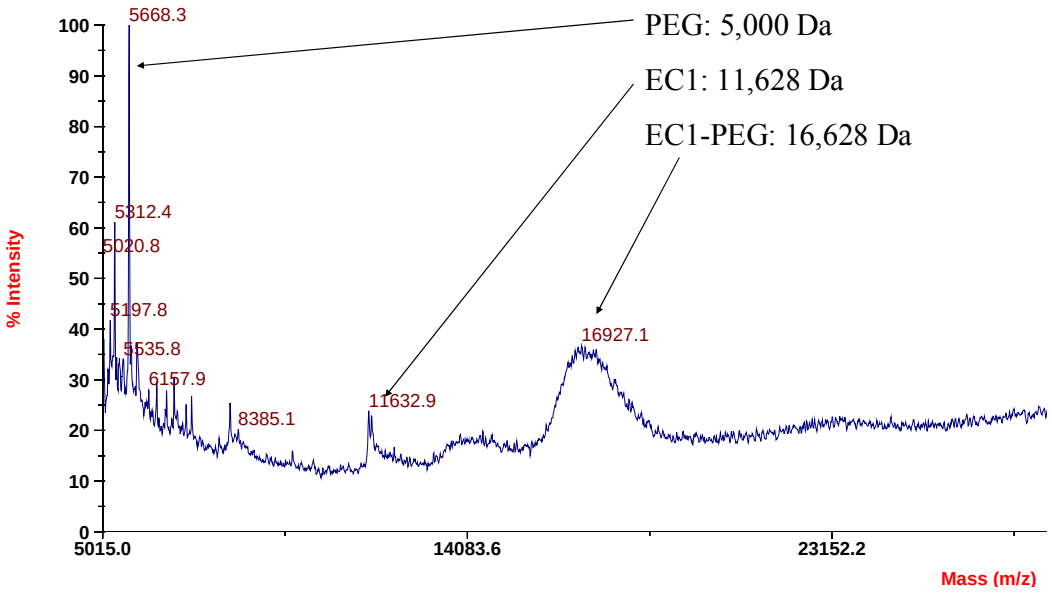

Figure 4.1b: MALDI-TOF MS analysis of the PEGylation reaction mixture.

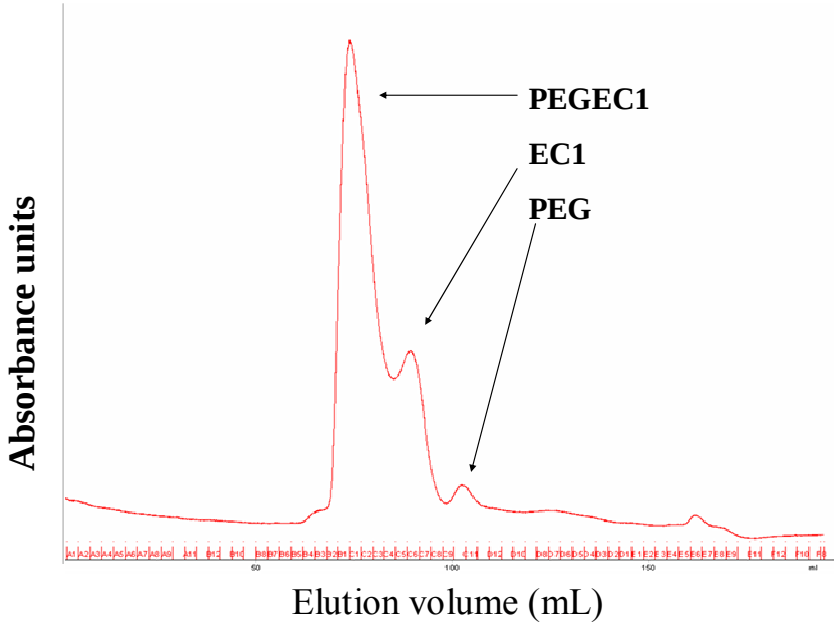

Figure 4.1c: Separation of PEGylated EC1, unreacted EC1 and maleimide-PEG using size-exclusion chromatography.

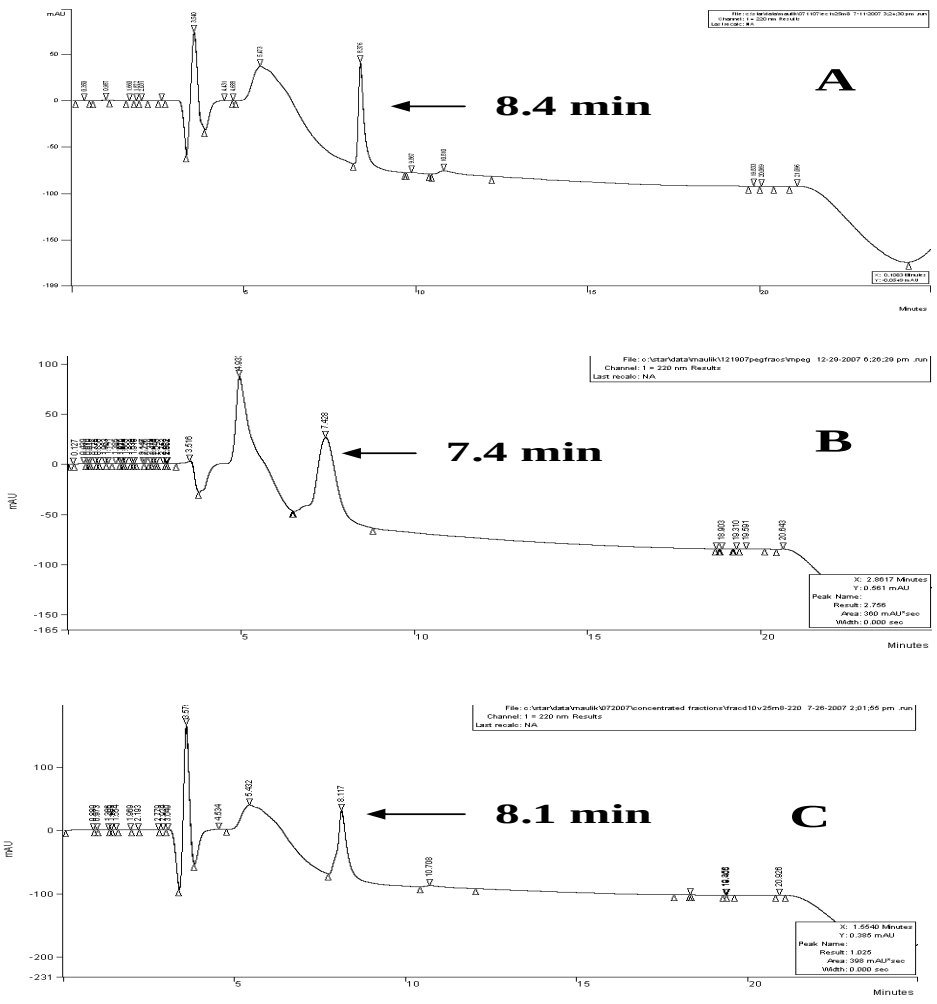

Figure 4.1d: HPLC profiles of EC1 (panel A, retention time = 8.4 min), M-PEG (panel B, retention time = 7.4 min), and PEGEC1 (panel C, retention time = 8.1 min)

4.3.1. Chemical stability studies:

The chemical stabilities of PEGEC1 at 70 °C incubation for 4 h at pH 3.0, 7.0, and 9.0 were analyzed by HPLC. At all three pH values studied, we found that the amount of PEGEC1, as given by the area under the HPLC peak, did not change substantially after incubation for 4 h at 70 °C (Figure 4.2). Moreover, no new peaks appeared in the chromatograms. This suggests that, at high-temperature, PEGEC1 is stable to precipitation and to chemical degradation reactions such as peptide bond hydrolysis. This result is in contrast to our previous studies that showed EC1 and its alkylated derivative (EC1-IN) underwent a peptide bond hydrolysis and/or precipitate formation under similar conditions.

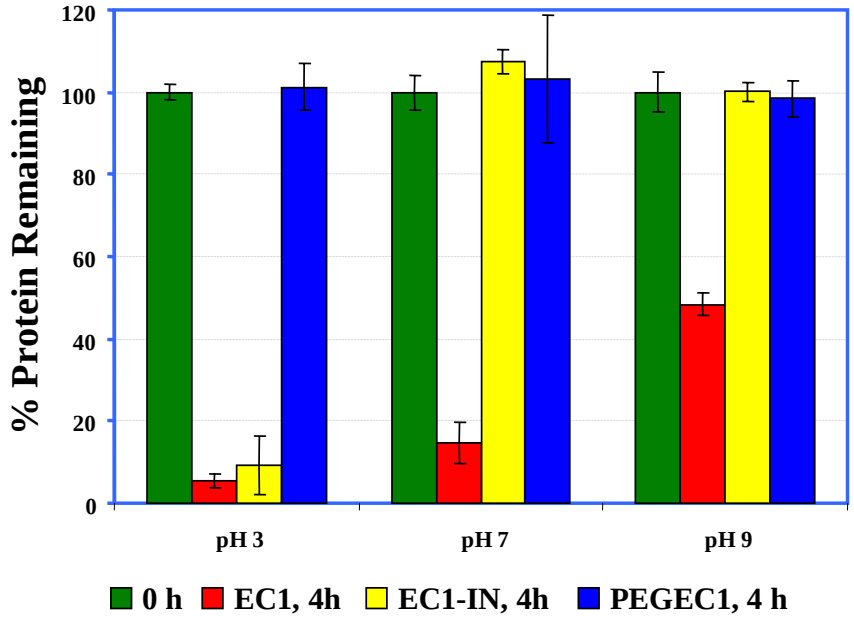

Figure 4.2: Chemical stability studies of PEGEC1 at pH 3.0, 7.0, and 9.0 at 70 °C for 4 h incubation. The results for the stability of PEGEC1 are compared with those for the stability of EC1 and EC1-IN obtained from Chapters 2 and 3.

4.3.2. Physical stability studies:

4.3.2.1. Far UV CD results: Secondary structural changes in PEGEC1 were evaluated by far UV CD spectroscopy between 200 and 250 nm. Comparison between the CD spectra of PEGEC1 and EC1 at pH 3.0 showed minor differences between the two (Figure 4.3a). The PEGEC1 spectrum had more pronounced positive and negative peaks at 235 nm and 216 nm, respectively. At pH 7.0 and 9.0, however, there was no significant difference between the CD spectra of PEGEC1 and EC1 (Figures 4.3b and 4.3c).

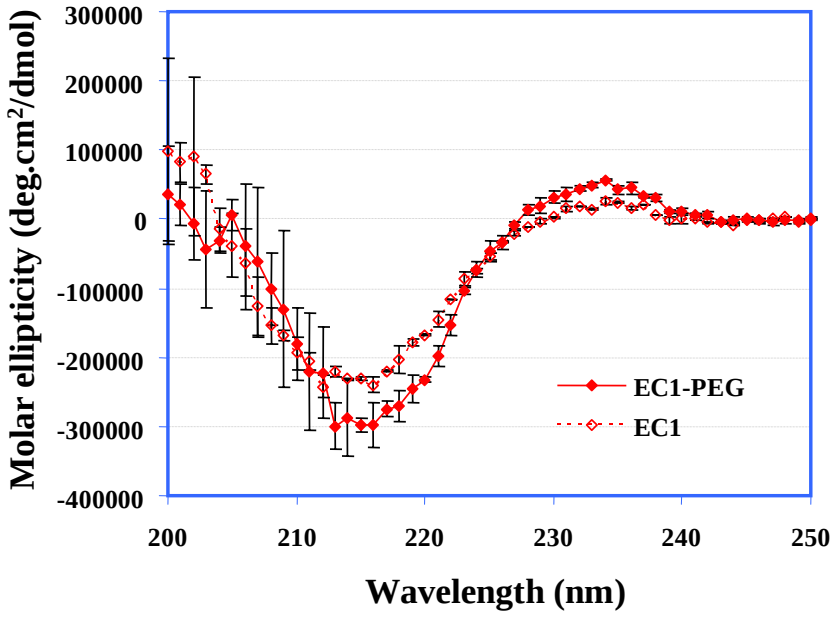

Figure 4.3a: Far UV CD spectrum of PEGEC1 compared to that of EC1 at pH 3.0. There are some differences between the two spectra.

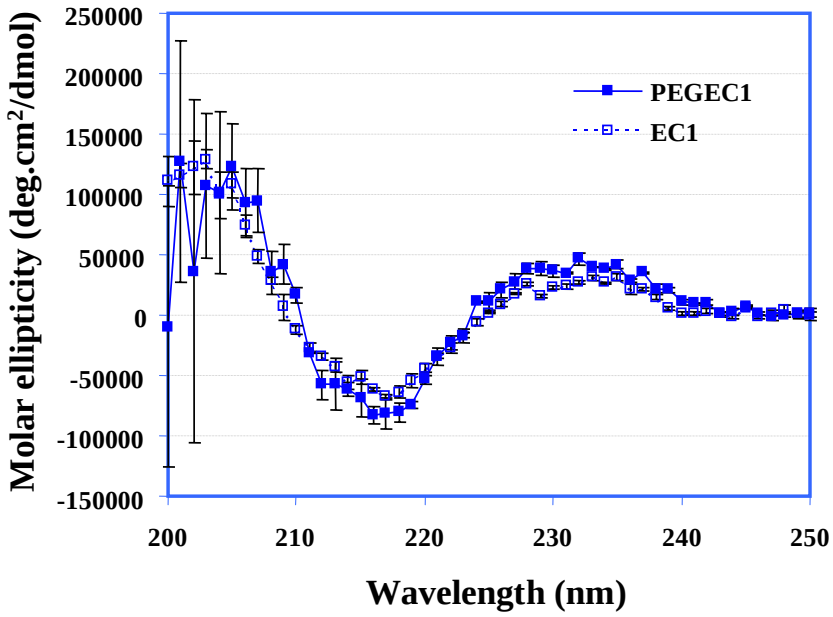

Figure 4.3b: Far UV CD spectrum of PEGEC1 compared to that of EC1 at pH 7.0. The two spectra are very similar.

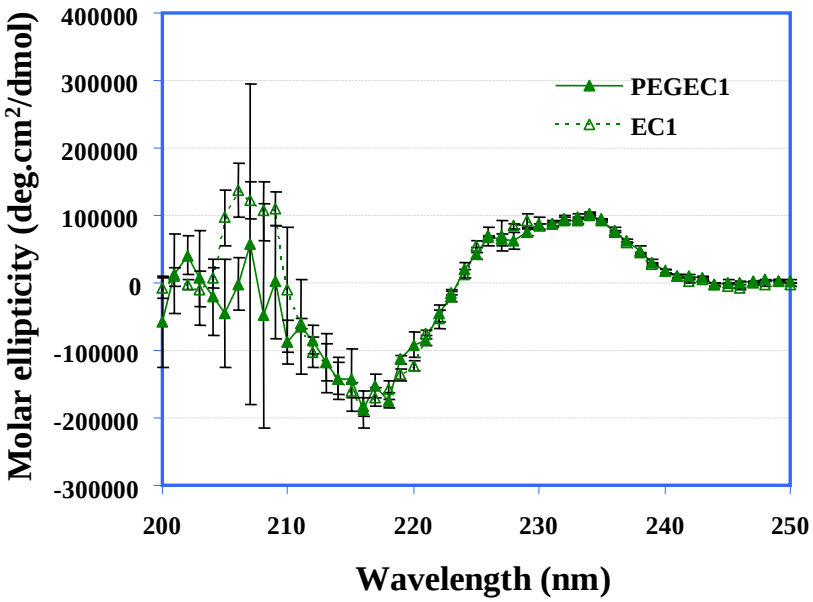

Figure 4.3c: Comparison of the far UV CD spectra of PEGEC1 and EC1 at pH 9.0. shows that they are very similar.

Thermal unfolding of PEGEC1 also was studied from 10 °C to 65 °C at pH 3.0, 7.0, and 9.0. The data obtained for the thermal unfolding transitions were fitted to a sigmoidal function (Figures 4.4a, 4.4b, and 4.4c), and the mid-point of each fit was defined as the thermal unfolding temperature (T_m). T_ms obtained for PEGEC1 were compared to those obtained for EC1. PEGEC1 showed a higher T_m at each pH studied (Figure 4.4d). Thus, PEGEC1 has better thermal stability than EC1.

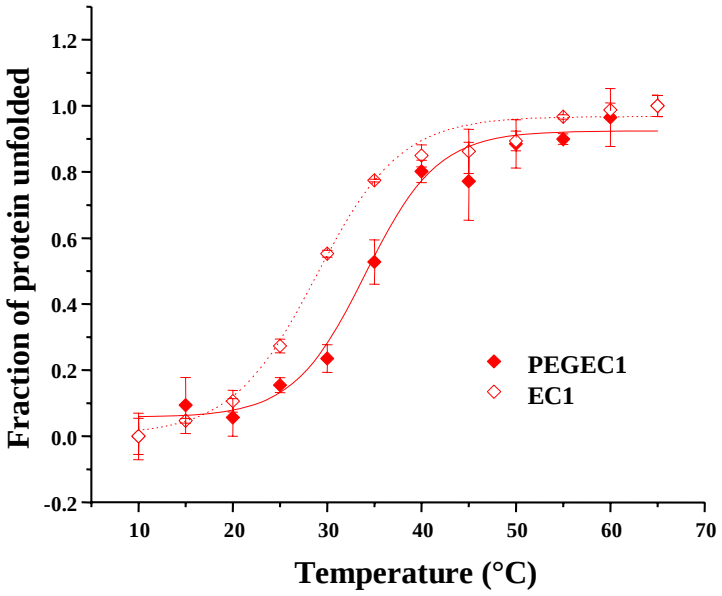

Figure 4.4a: Thermal unfolding profile of PEGEC1 compared to that of EC1 at pH 3.0 obtained by plotting the CD signal at 218 nm against the corresponding temperature.

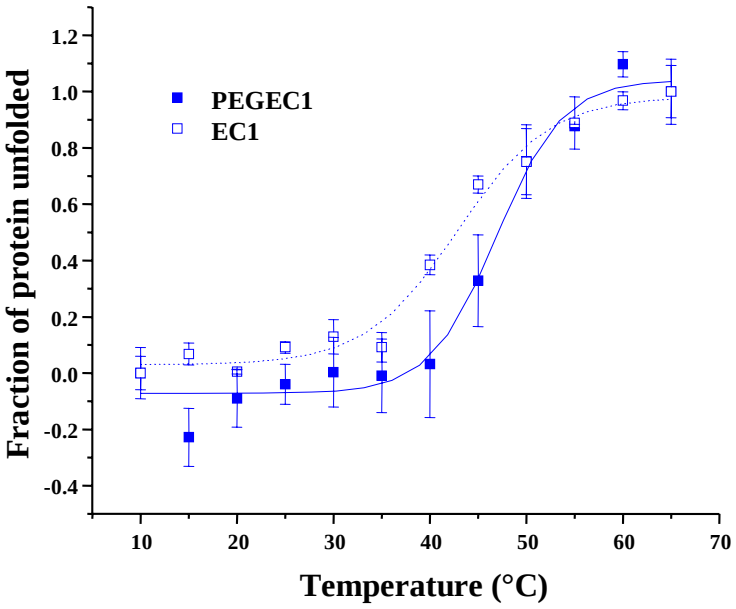

Figure 4.4b: Thermal unfolding profile of PEGEC1 compared to that of EC1 at pH 7.0 obtained by plotting the CD signal at 218 nm against the corresponding temperature.

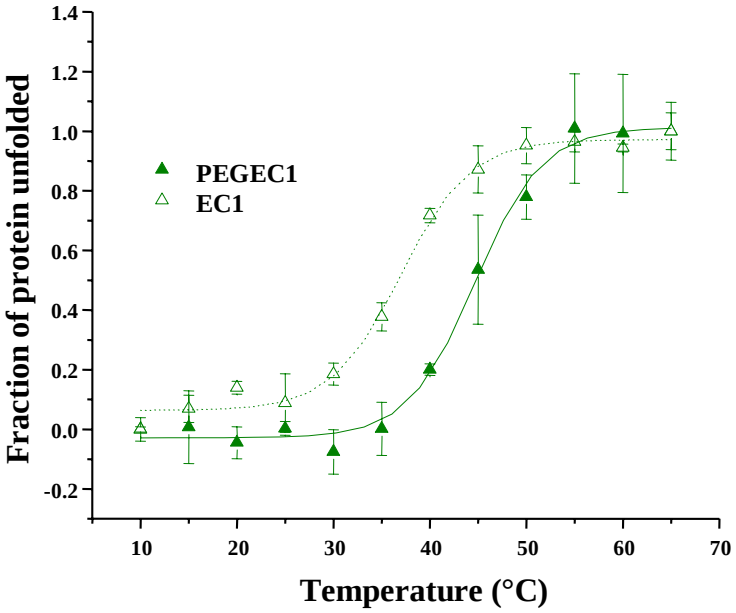

Figure 4.4c: Thermal unfolding profile of PEGEC1 compared to that of EC1 at pH 9.0 obtained by plotting the CD signal at 218 nm against the corresponding temperature.

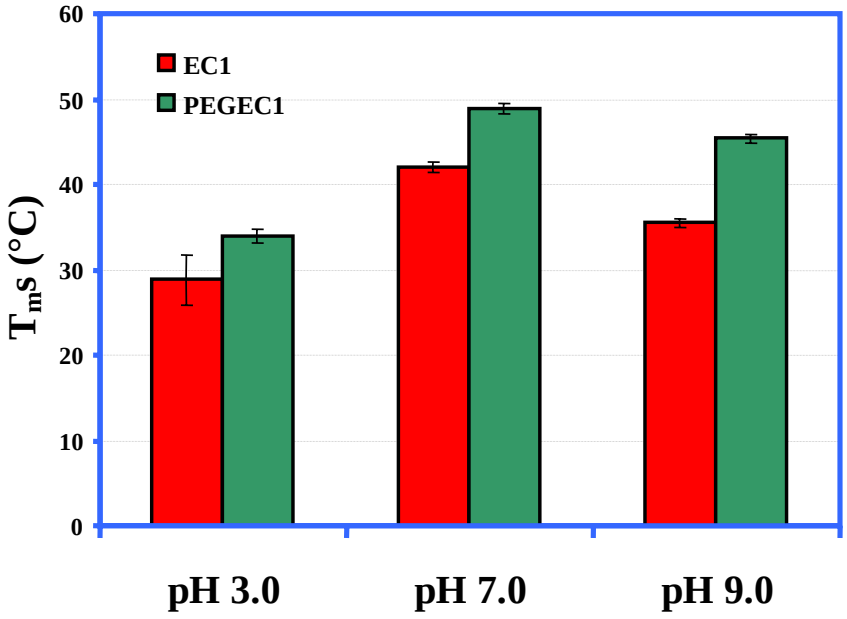

Figure 4.4d: A comparison of the unfolding temperatures (T_ms) for PEGEC1 and EC1 at pH 3.0, 7.0, and 9.0 after fitting the CD thermal unfolding profiles in 4.4a, 4.4b, and 4.4c, to sigmoidal functions.

4.3.2.2. Intrinsic fluorescence emission spectroscopy results: Intrinsic fluorescence emission spectra of PEGEC1 were obtained at pH 3.0, 7.0, and 9.0 between 305 nm and 405 nm after excitation at 295 nm. These spectra were compared to the corresponding spectra of EC1. The wavelength of maximum emission at pH 3.0 was 339.87 ± 0.03 nm for PEGEC1, similar to that for EC1 (339.93 ± 0.04 nm) at pH 3.0 (Figure 4.5a). At pH 7.0, it was 339.83 ± 0.05 nm, similar to that of EC1 (339.27 ± 0.03 nm) at pH 7.0 (Figure 4.5b). At pH 9.0, however, there was a small but significant difference in the wavelengths of maximum emission of PEGEC1 (339.97 ± 0.02 nm) and EC1 (341.31 ± 0.03 nm) (Figure 4.5c). A possible explanation for this difference at pH 9.0 is that the PEG interacts with at least one of the two Trp residues, probably the Trp6 residue that is near the site of PEGylation at the Cys13 residue. The presence of the bulky PEG group near the Cys13 residue makes Trp6 less exposed to the solvent in PEGEC1 than in EC1.

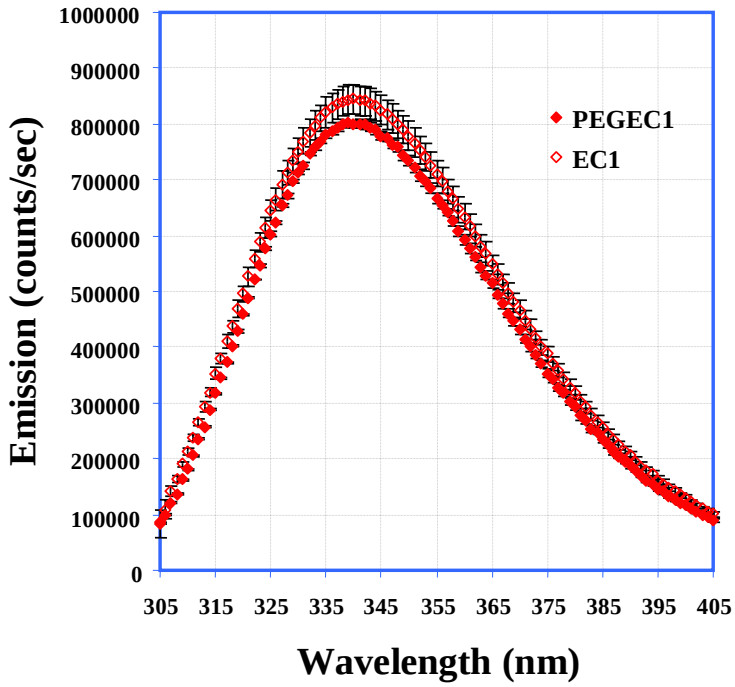

Figure 4.5a: Tertiary structure studies of PEGEC1 using intrinsic fluorescence emission spectroscopy. The emission spectra of PEGEC1 are compared to those of EC1 at pH 3.0. PEGEC1 and EC1 show similar tertiary structures.

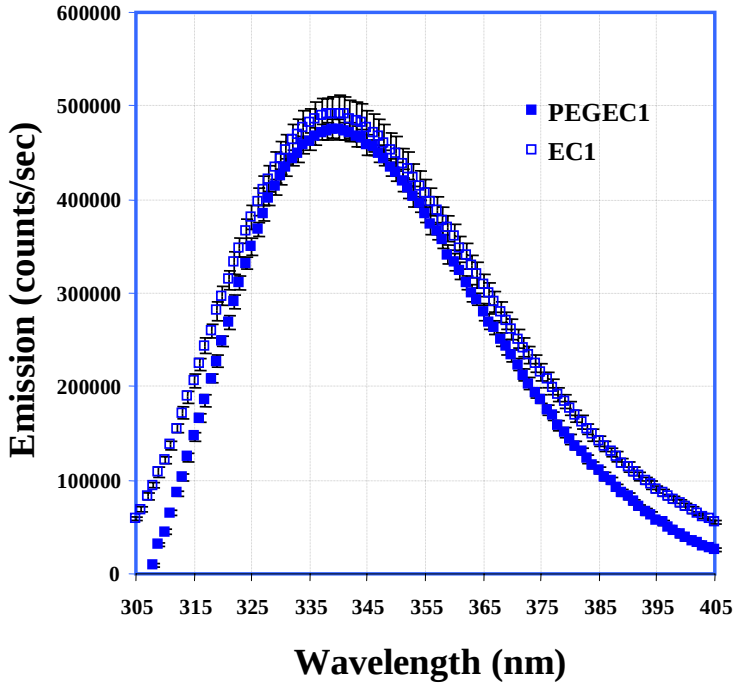

Figure 4.5b: Tertiary structure studies of PEGEC1 using intrinsic fluorescence emission spectroscopy. The emission spectra of PEGEC1 are compared to those of EC1 at pH 7.0. PEGEC1 and EC1 show similar tertiary structures.

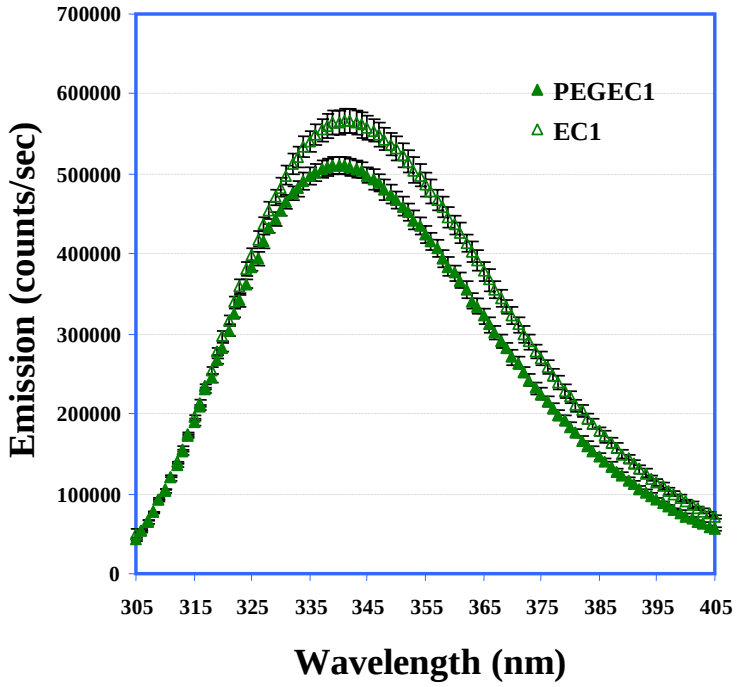

Figure 4.5c: Tertiary structure studies of PEGEC1 using intrinsic fluorescence emission spectroscopy. The emission spectra of PEGEC1 are compared to those of EC1 at pH 9.0. PEGEC1 and EC1 show similar tertiary structures.

The thermal unfolding transition of PEGEC1 was followed also by monitoring the shift in wavelength of maximum emission with increasing temperature. At pH 3.0, the thermal unfolding profiles of PEGEC1 and EC1 appear similar (Figure 4.6a), with no transition observed until 75 °C. At pH 7.0, the thermal unfolding profiles of PEGEC1 and EC1 were different, with PEGEC1 showing a transition to higher wavelengths (~ 350 nm) than EC1 (~346 nm) (Figure 4.6b). The transition temperature of PEGEC1 was estimated as 47.38 ± 1.40 °C, about 5 °C higher than that of EC1 (42.73 ± 0.60 °C). At pH 9.0, the thermal unfolding profiles for PEGEC1 and EC1 were similar (Figure 4.6c). The transition temperature of PEGEC1 was estimated to be 40.58 ± 0.95 °C, which is not significantly different than the T_m obtained for EC1 (38.34 ± 1.02 °C).

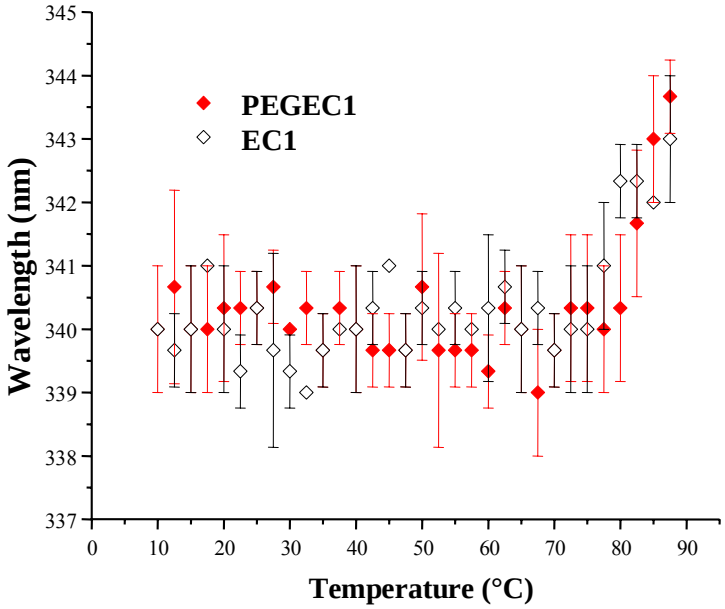

Figure 4.6a: Thermal unfolding profiles of PEGEC1 compared to those of EC1 at pH 3.0 obtained by plotting the wavelength of maximum emission after excitation at 295 nm against the corresponding temperature.

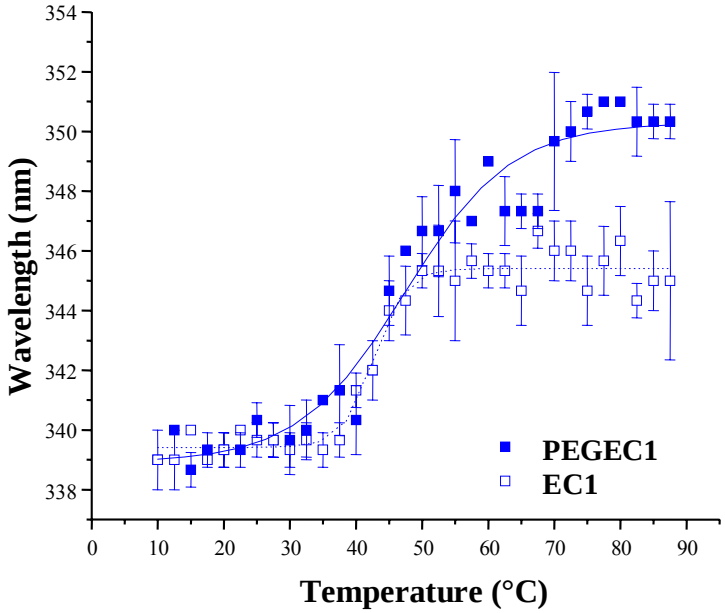

Figure 4.6b: Thermal unfolding profiles of PEGEC1 compared to those of EC1 at pH 7.0 obtained by plotting the wavelength of maximum emission after excitation at 295 nm against the corresponding temperature.

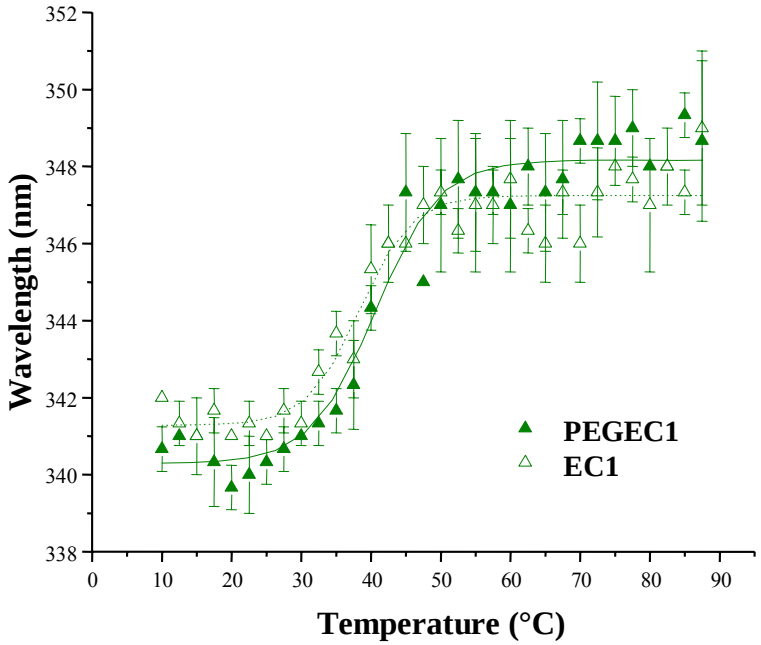

Figure 4.6c: Thermal unfolding profiles of PEGEC1 compared to those of EC1 at pH 7.0 obtained by plotting the wavelength of maximum emission after excitation at 295 nm against the corresponding temperature.

4.4. Discussion

We successfully synthesized and purified the PEGylated product of EC1 (PEGEC1). The PEGylation was accomplished by nucleophilic attack on the maleimide group of PEG by the thiol group of the Cys13 residue of the EC1 protein. This reaction produced a homologous mono-PEGylated product as determined by HPLC and mass spectrometry. PEGylation of the Cys residues of proteins has been employed previously to improve their *in vitro* and *in vivo* stabilities.[13-17] It has been proposed that protein stabilization with linear PEG is generally a result of the formation of a polymer shell around the protein that recruits water to form an outer water shell that minimizes intermolecular protein interactions.[10] PEGylation at several sites on a protein with higher molecular weight branched PEGs (as opposed to linear PEG) can shield the entire protein molecule. The hydrophobic and hydrophilic portions of PEG can interact with the hydrophobic and hydrophilic regions, respectively of the protein conjugated to the PEG. This is how a PEG-shell is created around the protein surface. This shell can mask the surface interactions between protein molecules. Because a 5,000 Da linear PEG is used in PEGEC1, there is a possibility that the PEG group may not be large enough to form a shell around the entire EC1 molecule. This could be significant because the peptide sequences (HAV and ADT) of EC1 which have been shown to modulate E-cadherin interactions are located on the surface of EC1 opposite the surface of the PEGylation site. This aspect is discussed in more detail later in this chapter. The X-ray structure of EC1 indicates that the surface of EC1 where the Cys13 residue is located is also the surface that is

involved in dimerization of EC1. The Trp6 residue on this surface is involved in the formation of domain swapped physical dimers; in the domain swapping process, the Trp6 residue of one EC1 domain is buried in a hydrophobic pocket of another EC1 domain. Thus, we propose that the presence of the PEG near the domain swapping region may prevent the formation of physical and chemical dimers of EC1.

Chemical stability studies at 70 °C showed that PEGEC1 was stable after 4 h at pH 3.0, 7.0, and 9.0. In contrast, the unmodified EC1 protein can undergo precipitation and peptide bond hydrolysis. The alkylated EC1 protein (EC1-IN) has a better stability at pH 7.0 and 9.0 than does the parent EC1; at pH 3.0, however, EC1-IN undergoes precipitation. Thus, these results suggest that, of all the methods examined so far, PEGylation is the best strategy for stabilizing the EC1 protein. It is also possible that the PEGylation of EC1 decreases the dynamic properties of PEGEC1, which contributes to the decrease in the hydrolysis reaction; this proposal needs further investigation.

Although PEGylation improves the physical and chemical stability of EC1, this modification may alter the ability of EC1 to inhibit E-cadherin-mediated cell-cell adhesion. Our previous studies have indicated that peptides containing the His47–Ala48–Val49 (HAV) and Ala83–Asp84–Thr85 (ADT) sequences derived from EC1 inhibit E-cadherin-mediated cell-cell adhesion. The HAV and ADT sequences are located on a surface of EC1 different from that of the PEGylation site. Therefore, PEGylation is not expected to alter the activity of the HAV and ADT sequences. It has been suggested that EC1 could have two homophilic binding surfaces. One is the

domain-swapping N-terminal region and the other is the region where the HAV and ADT sequences are located. In the future, the activity of PEGEC1 can be studied to determine the importance of either the domain swapping or the HAV/ADT region of EC1 in inhibiting the E-cadherin-mediated cell-cell adhesion.

PEGylation of EC1 has some impact on the secondary and tertiary structure of EC1. Far UV CD studies showed some minor differences between CD spectra of PEGEC1 and EC1 at pH 3.0. At pH 7.0 and 9.0, however, no substantial differences were observed between the CD spectra of PEGEC1 and EC1. These results suggest that compared to EC1-IN, PEGEC1 is less sensitive to pH changes. Studies of other PEGylated proteins also show that PEGylation does not change the secondary structure of the protein.[18,19] The thermal unfolding profiles determined by CD spectroscopy suggest that PEGEC1 was significantly more stable than EC1 with an increase in its T_m by 5.1 °C at pH 3.0, 4.0 °C at pH 7.0, and 7.7 °C at pH 9.0 after PEGylation. It was also found that the tertiary structure of PEGEC1 studied at pH 3.0, 7.0, and 9.0 was similar to that of EC1.

In conclusion, PEGylation of EC1 improves its physical and chemical stability. Unlike the modified EC1-IN protein, PEGEC1 has a structure similar to that of the parent EC1 at different pH values. In the future, the biological activity of PEGEC1 will be compared to that of EC1-IN and EC1 in inhibiting E-cadherin-mediated cell-cell adhesion in *in vitro* cell culture systems developed in our laboratory.

4.5. References

1. Overduin M, Harvey TS, Bagby S, Tong KI, Yau P, Takeichi M, Ikura M 1995. Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. Science lv267:p386(384).

2. Shapiro L, Fannon AM, Kwong PD, Thompson A, Lehmann MS, Grubel G, Legrand J-F, Als-Nielsen J, Colman DR, Hendrickson WA 1995. Structural basis of cell-cell adhesion by cadherins. Nature l374:327–337.

3. Klingelhofer J, Laur OY, Troyanovsky RB, Troyanovsky SM 2002. Dynamic interplay between adhesive and lateral E-cadherin dimers. Molecular Cell Biology l22:7449–7458.

4. Nagar B, Overduin M, Ikura M, Rini JM 1996. Structural basis of calcium-induced E-cadherin rigidification and dimerization. Nature l380:360–364.

5. Nose A, Tsuji K, Takeichi M 1990. Localization of specificity determining sites in cadherin cell adhesion molecules. Cell l61:147–155.

6. Troyanovsky RB, Sokolov E, Troyanovsky SM 2003. Adhesive and lateral E-cadherin dimers are mediated by the same interface. Mol Cell Biol l23:7965–7972.

7. Zheng K, Trivedi M, Siahaan TJ 2006. Structure and function of the intercellular junctions: barrier of paracellular drug delivery. Current Pharmaceutical Design l12:2813–2824.

8. Makagiansar IT, Ikesue A, Duc Nguyen P, Urbauer JL, Bieber Urbauer RJ, Siahaan TJ 2002. Localized production of human E-cadherin-derived first repeat in Escherichia coli. Protein Expression and Purification l26:449–454.

9. Makagiansar IT, Nguyen PD, Ikesue A, Kuczera K, Dentler W, Urbauer JL, Galeva N, Alterman M, Siahaan TJ 2002. Disulfide bond formation promotes the cis- and trans-dimerization of the E-cadherin-derived first repeat. Journal of Biological Chemistry l277:16002–16010.

10. Morar AS, Schrimsher JL, Chavez MD. 2006. PEGylation of proteins: A structural approach. Biopharm International, ed.

11. Derrick TS, Kashi RS, Durrani M, Jhingan A, Middaugh CR 2004. Effect of metal cations on the conformation and inactivation of recombinant human factor VIII. Journal of Pharmaceutical Sciences l93:2549–2557.

12. Rexroad J, Wiethoff CM, Green AP, Kierstead TD, Scott MO, Middaugh CR 2003. Structural stability of adenovirus type 5. Journal of Pharmaceutical Sciences l92:665–678.

13. Long DL, Doherty DH, Eisenberg SP, Smith DJ, Rosendahl MS, Christensen KR, Edwards DP, Chlipala EA, Cox GN 2006. Design of homogeneous, monopegylated erythropoietin analogs with preserved in vitro bioactivity. Experimental Hematology l34:697–704.

14. Rosendahl MS, Doherty DH, Smith DJ, Bendale AM, Cox GN 2005. Site specific protein PEGylation: Application to cysteine analogs of recombinant human granulocyte colony-stimulating factor. Bioprocess international l3:52–62.

15. Rosendahl MS, Doherty DH, Smith DJ, Carlson SJ, Chlipala EA, Cox GN 2005. A long acting, highly potent interferon a-2 conjugate created using site-specific PEGylation. Bioconjugate Chemistry l16:200–207.

16. Tsutsumi Y, Onda M, Nagata S, Lee B, Kreitman RJ, Pastan I 2000. Site-specific chemical modification with polyethylene glycol of rcombinant immuntoxin anti-Tac(Fv)-PE38 (LMB-2) improves antitumor activity and reduces animal toxicity and immunogenicity. l97:8548–8553.

17. Vanwetswinkel S, Plaisance S, Zhi-yong Z, Vanlinthout I, Brepoels K, Ladters I, Collen D, Jespers L 2000. Pharmacokinetic and thrombolytic properties of cysteine-linked polyethylene glycol derivatives of staphylokinase. Blood l95:936–942.

18. Digilio G, Barbero L, Bracco C, Corpillo D, Esposito P, Piquet G, Traversa S, Aime S 2003. NMR structure of two novel polyethylene glycol conjugates of the human growth hormone-releasing factor, hGRF(1-29)-NH_2. J Am Chem Soc l125:3458–3470.

19. Hinds KD, Kim SW 2002. Effects of PEG conjugation on insulin properties. Advanced Drug Delivery Reviews l54:505–530.

Chapter 5

Summary and future work

5.1. Summary

A large number of peptides and proteins currently are being investigated for their potential as therapeutic agents. One of the major obstacles to their development as future drugs is their inability to permeate biological barriers such as the intestinal mucosa and the blood brain barrier. The possibility of improving the paracellular transport of these drugs across such barriers currently is being studied. The intercellular junctions of these biological barriers contain a variety of protein-protein interactions, which function as cell adhesion mechanisms and act as 'filters' to the transport of large molecules through the junctions. E-cadherin is one such transmembrane protein located in the adherens region of the intercellular junctions of epithelial tissue in the intestinal mucosa as well as the blood-brain barrier. Understanding the mechanism of E-cadherin-mediated homophilic interactions is essential for developing strategies to modulate them. Studies have shown that the EC1 domain of E-cadherin plays an important role in the selectivity of these interactions. Thus, it is desirable to study the interaction of the EC1 domain with other EC domains in *in vitro* cell culture systems such as Caco-2 cell monolayers, which are mediated by E-cadherin. We therefore decided to express the recombinant human EC1 domain and study its stability.

In Chapter 2, we developed a system to express recombinant human EC1 in *E. coli* cells. This system provided a yield approximately 10 times higher than that of the system previously in place. It had also been shown previously that EC1 formed disulfide-mediated dimers that, in turn, formed physical oligomers. We conducted

studies to evaluate the effect of this covalent dimerization and physical oligomerization on the structure and stability of EC1. Short-term chemical stability studies, also known as "accelerated stability testing", were performed by incubating the protein under different pH conditions at high temperatures and quantifying the amount of protein remaining using HPLC. We found that at 37 °C and 70 °C, a significant amount of EC1 was lost to precipitation and hydrolysis of the peptide bond between the D93 and P94 residues. Addition of dithiothreitol (DTT), a disulfide reducing agent, to EC1 before performing the accelerated stability tests significantly decreased the level of precipitation as well as the peptide bond hydrolysis. We concluded that disulfide-mediated dimerization and physical oligomerization of EC1 reduced the stability of EC1. This result hinted at a difference in the structures of EC1 monomer, dimer, and oligomers. We sought to clarify this possibility by running molecular dynamics simulations of the EC1 monomer and dimer. These simulation studies revealed that disulfide bond formation between the Cys13 residues of two EC1 monomers is a facile reaction. The resulting covalent dimer is structurally different than the monomer and also has higher flexibility. The increased dynamics of the dimer possibly expose the D93–P94 peptide bond to the solvent, thereby making it more susceptible to hydrolysis. To confirm the molecular dynamics simulation results, we investigated the change in the structure of EC1 due to covalent dimerization and physical oligomerization by incubating EC1 for a relatively long period (28 days) at 4 °C. After incubation for 14 and 28 days, the changes in its secondary and tertiary structures were monitored by far UV circular dichroism (CD)

spectroscopy and intrinsic fluorescence emission spectroscopy, respectively. The results revealed that the secondary structure of EC1 changed significantly after 14 to 28 days. When DTT was added to the EC1 solution, however, before repeating the experiments under the same conditions, no changes were detected in the CD spectra. The tertiary structure studies revealed essentially the same results. The intrinsic fluorescence emission spectra of EC1 changed substantially after 28 days incubation at 4 °C, but those of the DTT-supplemented EC1 solution did not. We therefore concluded that the disulfide-mediated dimerization and physical oligomerization of EC1 are the main routes of its degradation.

In Chapter 3, we modified the Cys13 thiol of EC1 irreversibly by alkylating it with iodoacetamide. This modification can serve the purpose of maintaining EC1 in its monomeric form without DTT. This is particularly useful because the addition of DTT may interfere with cell culture systems. An eventual goal of stabilizing EC1 is to study the biological activity of the stabilized EC1 in cell culture systems. The alkylated EC1 (EC1-IN) was resistant to hydrolysis of its D93–P94 peptide bond. The secondary and tertiary structures of EC1-IN are similar to those of EC1 at pH 7.0. An important finding was that the secondary and tertiary structures of EC1-IN did not change significantly after long-term incubation at 4 °C, unlike the structure of EC1. This result would be expected because the irreversible modification of the thiol group in EC1-IN does not permit intermolecular disulfide formation between EC1 monomers. At pH 3.0 and 9.0, however, EC1-IN showed secondary structures different from those of EC1. Although not unexpected, it provided us with the

challenge of producing a thiol derivative of EC1 that retains a secondary structure similar to EC1 over a wide range of pH. Maintaining the secondary structure of EC1 after modification of its thiol could be important for retaining its biological activity.

Chapter 4 describes our results in attempting to make another thiol derivative of EC1 that is even more stable. In this case, the thiol group of Cys13 in EC1 was reacted with a maleimide-activated polyethylene glycol (M.W. 5,000 Da). The PEGylated EC1 (PEGEC1) showed a secondary structure very similar to that of EC1 at pH 3.0, 7.0, and 9.0. Moreover, PEGEC1 has higher thermal stability than the parent EC1 and also is resistant to peptide bond hydrolysis at 70 °C. PEGEC1 has a structure similar to that of the parent EC1 at different pH values; this is in contrast to EC1-IN.

In summary, EC1 provided us with a model for studying disulfide-mediated protein degradation. Using various strategies, we were able to dramatically inhibit this degradation by the production of a PEGylated derivative.

5.2. Future work

As mentioned previously, the long-term goal behind our attempts at stabilizing EC1 is to study its *in vitro* biological activity as well as its binding properties to other EC domains of E-cadherin. Our studies on the stability of EC1 provide useful information for the handling and storage of EC1. Because EC1-IN retained its secondary structure at pH 7.0, this molecule is expected to retain its biological activity. EC1-IN can also be studied for its interaction with peptides derived from E-

cadherin. Different EC domains can be immobilized on solid surfaces, and their binding to EC1-IN in solution can be studied using techniques such as surface plasmon resonance. Alternatively, multi-dimensional NMR spectroscopy can be used to study the interaction of ^{15}N-labeled EC1-IN with peptides derived from the EC domains. Furthermore, the interaction of EC1-IN with the entire extracellular segment of E-cadherin in living cells also can be studied in cell culture systems such as Caco-2 cell monolayers.

5.3. Final comments

It has been difficult to handle and study EC1 due to the heterogeneity produced by its spontaneous dimerization mediated by intermolecular disulfide bonds, followed by its physical oligomerization. By blocking its covalent dimerization and physical oligomerization through modification of its Cys13 thiol, we have provided EC1 derivatives that can replace EC1 in future studies.