Effects of Disulfide Bond Formation in Production of the Recombinant Extracellular Domain of Human CD83 as a Therapeutic Protein

by

Lin Zhang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy

in

Chemical Engineering

Waterloo, Ontario, Canada, 2010

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The formation of aberrant disulfide bonds is a structural consideration for the manufacturing of the extracellular domain of human CD83 (hCD83ext), a potential therapeutic protein. In certain instances, hCD83ext protein products, even when stored frozen, tend to dimerize or even multimerize through the formation of aberrant intermolecular disulfide bonds. Herein, we discovered an analytical inconsistency and applied a modified sample preparation protocol for proper structural analysis of hCD83ext products which are heterologously expressed in *Escherichia coli* and subsequently purified. In addition, a mutant derivative with the Cys100Ser mutation was identified as an improved version which did not form dimers or multimers. The identification of this mutant variant as a more potent therapeutic protein than other hCD83ext species demonstrated that the structural variation associated with disulfide bond formation can be a critical issue for rigorous control of the quality and bioactivity of therapeutic proteins. The application of this mutant variant for protein therapeutic is currently under exploration.

As a comparative study, the hCD83ext was expressed as a glutathione-S-transferase (GST) fusion in two *E. coli* B strains, i.e. BL21 and Origami B having a reductive and oxidative cytoplasm. The final therapeutic products of hCD83ext produced by the two expression hosts exhibited significant differences in protein conformation and molecular properties, which presumably resulted from different disulfide patterns. The study highlights the importance of developing proper host/vector systems and biomanufacturing conditions for the production of recombinant therapeutic proteins with a consistent product quality.

Cys^{27} in the hCD83ext was identified as a target for molecular manipulation. Two *E. coli* strains of BL21(DE3) and Origami B(DE3) were used as the expression host to produce the Cys^{27} mutants. It was observed that Cys^{27} was involved in the *in vivo* formation of intramolecular disulfide bonds when hCD83ext was expressed in Origami B(DE3). The Origami-derived protein products had a higher tendency than the BL21-derived counterparts for multimerization via the *in vitro* formation of intermolecular disulfide bonds. Various analyses were conducted to identify the structural differences among these mutant variants. Most importantly, molecular stability was enhanced by the Cys^{27} mutations since the Cys^{27} mutants derived from either BL21 or Origami were much less susceptible to degradation compared to wild-type hCD83ext. This study highlights the implications of aberrant disulfide bond formation on the production of therapeutic proteins.

To address an inconsistent bioactivity issue that is primarily due to the aberrant formation of disulfide bonds associated with the presence of five cysteine residues, i.e. AA 27, 35, 100, 107, and 129, the molecular role that each cysteine plays upon the formation of intramolecular or intermolecular disulfide bonds was characterized, using various hCD83ext mutant variants derived by two *E. coli* expression hosts, i.e. BL21(DE3) and Origami B(DE3). Among the five cysteines, Cys^{100} and Cys^{129} can act as a bridging cysteine for *in vitro* multimerization via the formation of intermolecular disulfide bonds. The multimerization can be alleviated to some extent with less free Cys^{129} residues, associated with the possible formation of Cys^{27}-Cys^{129} intramolecular disulfide bond. As a result, introducing the Cys^{27} mutation can increase the multimerization presumably via freeing more Cys^{129} residues. In addition, protein stability can be improved in the presence of the Cys^{27}

mutation. The formation of the Cys^{27}-Cys^{129} intramolecular disulfide bond appears to be more effective in the presence of the Cys^{100} mutation, resulting in the suppression of multimerization. The two conserved cysteine residues, i.e. Cys^{35} and Cys^{107}, can be potentially linked to form an intramolecular disulfide bond, particularly when the protein is produced in Origami B(DE3).

Acknowledgements

I would like to express my sincere gratitude and appreciation to the following:

• My supervisors, Professors C. P. Chou and M. Moo-Young for their inspiration, guidance, and advice.

• Committee members, Professors M. Aucoin, L. Meiering, and J. Scharer for their valuable comments on the project and thesis.

• Professor Y-H. Lin at the University of Saskatchewan for his willingness to be my external examiner and for his valuable comments on the thesis.

• Ralph Dickhout, Bert Habicher, Ravindra Singh, Rick Hecktus and Dennis Herman for their various technical supports.

• Lorna Kelly, Liz Bevan, and Rose Guderian for their help within the duration of my study.

• My lab mates, Yali Xu, Niju Narayanan, Michael Pyne, Karan Sukhija, Luyang Zhong, Val Orr, and Kajan Srirangan for their kind collaboration in this project.

• Visiting Scholar, Dr. Abedi, for his instruction and encouragement during my study.

• My friends, Frank Bai and Wonphaka Wongrat, for their encouragement and support during my study.

• My parents for their care during my study.

Table of Contents

AUTHOR'S DECLARATION ii

Abstract iii

Acknowledgements vi

Table of Contents vii

List of Figures xii

List of Tables xxiii

Nomenclature xxiv

Chapter 1 Introduction 1

1.1 Research Background 1

1.2 Research Objectives 3

1.3 Research Approach 4

1.4 Outline of Thesis 5

Chapter 2 Literature Review 6

2.1 Expression Systems for Recombinant Therapeutic Protein Production 6

2.2 Recombinant Therapeutic Protein Production in *E. coli* 9

2.2.1 Bioprocess Development 9

2.2.2 Technical Issues and Approaches 14

2.2.3 Structural Analysis of Final Product 16

2.3 Folding and Misfolding of Recombinant Protein in *E. coli* 20

2.3.1 Folding of Recombinant Protein in Cytoplasm 20

2.3.2 Export and Folding of Recombinant Protein in Periplasm 21

2.3.3 Extracellular Secretion of Recombinant Protein 23
2.4 Disulfide Bond in Proteins 25
2.4.1 Effects on Protein Stability 25
2.4.2 Disulfide Bond Chemistry 26
2.4.3 Disulfide Bond in Protein Folding 29
2.5 Disulfide Bond Formation in Recombinant Therapeutic Protein Production 31
Chapter 3 Project Background: Bioprocess Development for Production, Purification, and Structural Characterization of Recombinant hCD83ext as a Potent Therapeutic Protein 34
3.1 Introduction 34
3.2 Materials and Methods 35
3.2.1 Bacterial Strains and Plasmids 35
3.2.2 Sample Treatment for Analyses 36
3.2.3 Downstream Processing 37
3.2.4 Analytical Methods 38
3.3 Results 40
3.4 Discussion 42
Chapter 4 Structural Identification of Recombinant Human CD83 Mutant Variant as a Potent Therapeutic Protein 47
4.1 Introduction 47
4.2 Materials and Methods 48
4.2.1 Bacterial Strains and Plasmids 48
4.2.2 Production of hCD83ext Variants 49

4.2.3 Analytical Methods 49
4.3 Results 51
4.3.1 Structural Analysis Using SDS-PAGE 51
4.3.2 Structural Analysis Using SEC 53
4.3.3 Monitoring of Conformational Changes 56
4.3.4 Bioactivity Comparison 58
4.4 Discussion 60
Chapter 5 Effect of Aberrant Disulfide Bond Formation on Protein Conformation and Molecular Property of hCD83ext 64
5.1 Introduction 64
5.2 Materials and Methods 65
5.2.1 Bacterial Strains and Plasmids 65
5.2.2 Cultivation and Downstream Processing 65
5.2.3 Sample Treatment for Analyses 65
5.2.4 Analytical Methods 66
5.3 Results 68
5.3.1 Bacterial Cultivation for hCD83ext Expression 68
5.3.2 Comparison of hCD83ext Conformation 71
5.3.3 Comparison of hCD83ext Molecular Property 73
5.4 Discussion 78
Chapter 6 Molecular Manipulation Associated with Disulfide Bond Formation to Enhance the Stability of Recombinant Therapeutic Protein 82

6.1 Introduction 82
6.2 Materials and Methods 83
6.2.1 Conserved Domain Search and Structural Prediction of hCD83ext 83
6.2.2 Bacterial Strains and Construction of hCD83ext Mutants 84
6.2.3 Protein Production and Purification 84
6.2.4 Analytical Methods 85
6.3 Results 86
6.3.1 Identification of the Key Cysteine Residue for Manipulation 86
6.3.2 Bacterial Cultivation for Mutant hCD83ext Expression 87
6.3.3 Characterization of Multimerization 94
6.3.4 Characterization of Conformational Changes 101
6.3.5 Stability Enhancement Associated with the Cys^{27} Mutations 103
6.4 Discussion 107
Chapter 7 Identification of Molecular Roles of Cysteine Residues Associated with *In Vitro* Multimerization of Therapeutic Protein Human CD83 115
7.1 Introduction 115
7.2 Materials and Methods 116
7.2.1 Bacterial Strains and Construction of hCD83ext Mutants 116
7.2.2 Protein Production and Purification 117
7.2.3 Sample Treatment for Analyses 117
7.2.4 Analytical Methods 118
7.3 Results 119

7.3.1 Rational for Mutagenesis Studies 119
7.3.2 Comparison of Molecular Properties for M3 Variants 121
7.3.3 Comparison of Molecular Properties for Extended M3 Variants 125
7.4 Discussion 131
Chapter 8 Conclusions and Recommendations 140
8.1 Conclusions 140
8.2 Recommendations 141
Appendix A DNA sequencing results and alignment 143
Appendix B Metropolis-Hastings algorithm used for parameter estimation of dimerization reactions of mutant variants hCD83ext_mut_C27E and hCD83ext_mut_C27S 148
References 156

List of Figures

Figure 2.1……10
A general bioprocess scheme for recombinant therapeutic protein production.

Figure 2.2……27
The *in vivo* disulfide bond reactions catalyzed by oxidoreductases. (A) Oxidation, (B) Reduction, and (C) Isomerization. The blue oval represents the oxidoreductase.

Figure 2.3……30
Disulfide scrambling. (A) sulfhydryl-disulfide exchange reaction occurs intramolecularly, resulting in a different intramolecular disulfide bond. (B) sulfhydryl-disulfide exchange reaction occurs intermolecularly, resulting in formation of an intermolecular disulfide bond, (C) An intermolecular disulfide bond is broken by a sulfhydryl group via sulfhydryl-disulfide exchange reaction.

Figure 3.1……43
Analysis of various hCD83ext sample lots with (A) reduced SDS-PAGE and (B) non-reduced SDS-PAGE. The protein samples were stored at -20 °C and were formulated in 50% glycerol except the lane-3 sample which was glycerol-free. The lots were made on several different dates, therefore had different "ages" (i.e. 4, 3, 3, 2, 10.5, 9.5, 6.5, 6, 6 months for lanes 2~10, respectively) upon conducting the analysis. The arrows indicated several uncommon hCD83ext species possibly mediated by unconventional intramolecular disulfide bonds.

Figure 3.2…….44
Analysis of an hCD83ext sample with (A) reduced SDS-PAGE, (B) non-reduced SDS-PAGE, and (C) Western blotting. The SDS gels were stained with silver nitrated in order to make higher multimeric species (such as trimer and tetramer) visible. All these species were verified to be associated with hCD83ext by Western blotting. Note that a small amount of GST exited in the hCD83ext product fraction as a contaminant protein.

Figure 3.3…….46
Comparison in structural change of an hCD83ext sample over 15-day storage at 0 °C using (A) reduced SDS-PAGE, (B) non-reduced SDS-PAGE, (C) CD spectroscopy, and (D) fluorescence spectroscopy.

Figure 4.1…….54
Analysis of four hCD83ext products treated with various sample preparation protocols using reduced SDS-PAGE (Panel A) and non-reduced SDS-PAGE (Panel B). For both panels: Lane 1/ molecular weight markers; lanes 2, 6, 10/ M5; lanes 3 7, 11/ M3; lanes 4, 8, 12/ wild-type-1; lane 5, 9, 13/ wild-type-2. Lanes 2~5 are samples with no NEM-treatment; lanes 6~9 are samples with NEM-treatment prior to heat denaturation; lanes 10~13 are samples with NEM-treatment after heat denaturation. The gels were stained by silver nitrate.

Figure 4.2…….55
Analysis of four hCD83ext products using SEC. The elution profiles for both reduced (with DTT treatment) and non-reduced (without DTT treatment) are shown. Panel A/ M5; panel B/ M3; panel C/ wild-type-1; panel D/ wild-type-2.

Figure 4.3…….57
Structural monitoring and comparison for M3 (Panels A and D), wild-type (Panels B and E), and M5 (Panels C and F) products stored under -20 °C using reduced SDS-PAGE (Panels A, B, and C) and non-reduced SDS-PAGE (Panels D, E, and F). Panels A and D: Lane 1/ molecular weight markers; lanes 2~5/ samples taken on day 0, 1, 7, and 15, respectively. Panel B and E: Lane 1/ molecular weight markers; lanes 2~6/ samples taken on day 0, 1, 2, 6, and 10, respectively. Panel C and F: Lane 1/molecular weight markers; lanes 2~6/ samples taken on day 0, 1, 5, 10, and 15, respectively. The gels were stained by silver nitrate.

Figure 4.4…….58
Structural comparison for M3, wild-type, and M5 products stored under -20 °C using CD (panels A and B) and fluorescent spectroscopy (panels C and D). Panels A and C: spectra of day-0 samples; panels B and D: spectra of day-15 samples.

Figure 4.5…….59
The ability of hCD83ext variants (i.e. wild-type, M5 and M3) to inhibit TNF-α production by LPS/IFN-γ stimulated primate PBMCs. The inhibition was quantified by either the percentage of monocytes that produce TNF-α (panel A) or by the mean level of TNF-α produced per cell (panel B) under various hCD83ext dosages.

Figure 5.1…….69
Comparison of bioreactor cultivation performance for the production of GST-hCD83ext in two host strains: Cell density (Panel A), specific GST activity (Panel B), and intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panels C) and Western blotting (Panels D) for the culture

samples taken prior to induction and at 6 h post-induction time are shown. Panel C: Lanes 1, 2, 5, 6, 9 and 10 represent samples produced by BL21(DE3), whereas lanes 3, 4, 7, 8, 11, and 12 represent samples produced by Origami B(DE3). Arrows represent the protein bands corresponding to disulfided and non-disulfided hCD83ext monomers. Panel D: All lanes represent culture samples taken at 6 h post-induction time. Arrows represent the protein bands corresponding to various hCD83ext structures.

Figure 5.2……....72
Analysis of protein structure for fresh hCD83ext produced by two host strains with CD spectroscopy (Panel A) and fluorescence spectroscopy (Panel B).

Figure 5.3……....74
Monitoring and comparison of molecular property for hCD83ext produced by two host strains with reducing SDS-PAGE (Panels A and B), non-reducing SDS-PAGE (Panels C and D), CD spectroscopy (Panels E and F), and fluorescence spectroscopy (Panels G and H): Samples up to 30-day storage at -20 °C and 10-day storage at room temperature were analyzed. Panels A, C, E, and G represent hCD83ext samples produced by BL21(DE3), whereas Panels B, D, F, and H represent hCD83ext samples produced by Origami B(DE3). Panels A~D: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~ 6 and 11/ -20 °C samples taken on day 1, day 2, day 6, day 10, and day 30, respectively; lanes 7 ~ 10/ room temperature samples taken on day 1, day 2, day 6, and day 10, respectively. The insets at upper right in Panel C and D represent the time profiles for dimerization of hCD83ext stored at -20 °C. Such quantification was not conducted for samples stored at room temperature since hCD83ext dimer appeared to be unstable.

Figure 5.4……..77
Time profiles for the number of titratable free sulfhydryl group in hCD83ext produced by BL21(DE3) (Panel A) and Origami B(DE3) (Panel B): Profiles for samples stored at -20 °C and room temperature are shown.

Figure 6.1……..88
Conserved domain analysis in hCD83ext sequence performed by NCBI CD-search BLAST. (A) Pairwise alignments between the hCD83ext sequence (displayed as local) and the domain-model consensus sequences from Immunoglobulin domain family (the upper panel) and Immunoglobulin V-set domain (the lower panel). The conserved cysteine residues were boxed and the cysteine at the position 27 in hCD83ext was marked with . SMART and Pfam are imported collections as the sources of conserved domain-search alignment model data [2]. Smart00409 and Pfam07686 were identified as the domain-models for hCD83ext sequence. (B) Predicted 3D structure of hCD83ext (shown in backbone style). The positions of five cysteine residues were marked as pointed by arrows.

Figure 6.2……..91
Comparison of bioreactor cultivation performance for the production of GST-hCD83ext_mut_C27E in two host strains: Cell density (Panel A), specific GST activity (Panel B), and intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panel C) and Western blotting (Panel D) for the culture samples taken prior to induction and at 6 h post-induction time were shown. Panel C: Lane 1 is molecular weight marker; Lanes 2, 3, 10, 11 and 4, 5 represent soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively,

whereas lanes 6, 7, 12, 13 and 8, 9 represent soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows represent the protein bands corresponding to monomeric fusion protein. Panel D: All lanes are corresponding to the lanes 2~13 in Panel C.

Figure 6.3…….....93
Comparison of bioreactor cultivation performance for the production of GST-hCD83ext_mut_C27S in two host strains: Cell density (Panel A), specific GST activity (Panel B), and intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panel C) and Western blotting (Panel D) for the culture samples taken prior to induction and at 6 h post-induction time were shown. Panel C: Lane 1 is molecular weight marker; Lanes 2, 3, 10, 11 and 4, 5 represent soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively, whereas lanes 6, 7, 12, 13 and 8, 9 represent soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows represent the protein bands corresponding to monomeric fusion protein. Panel D: All lanes are corresponding to the lanes 2~13 in Panel C.

Figure 6.4…….....95
Monitoring and comparison of multimerization for hCD83ext_mut_C27E produced by two host strains with reducing SDS-PAGE (Panels A and C), non-reducing SDS-PAGE (Panels B and D): Samples up to 30-day storage at -20 °C and room temperature were analyzed. Panels A and B represent hCD83ext_mut_C27S samples produced by BL21(DE3)pLysS, whereas Panels C and D represent hCD83ext_mut_C27E samples produced by Origami

B(DE3)pLysS. Panels A~B: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~ 7/ -20 °C samples taken on day 1, day 2, day 6, day 10, and day 30, respectively; lanes 7 ~ 10/ room temperature samples taken on day 1, day 2, day 6, day 10, and day30 respectively. Panels C~D: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~7/ -20 °C samples taken on day 1, day 2, day 6, day 10, and day 30, respectively; lanes 8~10/ room temperature samples taken on day 1, day 2, and day 6 respectively. The insets at upper right in Panels B and D represent the time profiles for dimerization of hCD83ext_mut_C27E stored at -20 °C and room temperature.

Figure 6.5…….....98
Monitoring and comparison of multimerization for hCD83ext_mut_C27S produced by two host strains with reducing SDS-PAGE (Panels A and C) and non-reducing SDS-PAGE (Panels B and D): Samples up to 30-day storage at -20 °C and room temperature were analyzed. Panels A and B represent hCD83ext_mut_C27S samples produced by BL21(DE3)pLysS, whereas Panels C and D represent hCD83ext_mut_C27S samples produced by Origami B(DE3)pLysS. Panels A~B: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~7/ -20 °C samples taken on day 1, day 3, day 5, day 10, and day 30, respectively; lanes 8~12/ room temperature samples taken on day 1, day 3, day 5, day 10, and day30 respectively. Panels C~D: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~6/ -20 °C samples taken on day 2, day 5, day 10, and day 30, respectively; lanes 7~10/ room temperature samples taken on day 2, day 5, day 10, and day30 respectively. The insets at upper right in Panels B and D represent the time profiles for dimerization of hCD83ext_mut_C27S stored at -20 °C and room temperature. Multimerization profiles were

further plotted according to the relative abundance of monomeric and multimeric forms quantified from their intensities on the SDS gel.

Figure 6.6……...104
Monitoring and comparison of the retention behavior of hCD83ext mutant variants in rp-HPLC: Panels A and C represent hCD83ext_mut_C27E produced by BL21(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively; Panels B and D represent hCD83ext_mut_C27E produced by Origami B(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively. Panels E and G represent hCD83ext_mut_C27S produced by BL21(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively; Panels F and H represent hCD83ext_mut_C27S produced by Origami B(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively.

Figure 6.7……...108
Monitoring and comparing the stability of hCD83ext wild-type protein with mutant proteins hCD83ext_mut_C27E and hCD83ext_mut_C27S produced by two host strains by reducing SDS-PAGE (Panel A), non-reducing SDS-PAGE (Panel B), CD spectroscopy (Panels C, D, and E), and fluorescence spectroscopy (Panels E, F, and G). Samples up to 30-day storage at -20 °C and room temperature were analyzed except that hCD83ext_mut_C27E produced in Origami B(DE3)pLysS was monitored till 6-day when stored at room temperature. Panels A and B : lane 1/ molecular weight markers; lanes 2~7/ -20 °C samples, arranged as lane 2/ wild-type from BL21(DE3)pLysS, lane 3/ wild-type from Origami B(DE3)pLysS; lane 4/ hCD83ext_mut_C27E from BL21(DE3)pLysS, lane 5/ hCD83ext_mut_C27E from Origami

B(DE3)pLysS, lane 6/ hCD83ext_mut_C27S from BL21(DE3)pLysS, lane 7/ hCD83ext_mut_C27S from Origami B(DE3)pLysS; lane 8/ wild-type from Origami B(DE3)pLysS 6-day storage at RT; lanes 9~14/ corresponding RT samples following the same arrangement. The arrows indicate the monomeric and dimeric species of different mobilities in wild-type and mutant proteins. Due to the limited amount of the sample, hCD83ext_mut_C27E produced by Origami B(DE3)pLysS was monitored for 6 days at room temperature. Accordingly, the comparison was made to the wild-type hCD83ext produced by Origami B(DE3)pLysS with the same storage conditions and period.

Figure 7.1……...120
Multiple sequence alignment of the hCD83ext (query) with the other Immunoglobulin V-set domain members. The alignment was generated via NCBI Conserved Domain Search Blast on line. Conserved cysteine residues were boxed, whereas variable cysteine residues were circled in the hCD83ext sequence.

Figure 7.2……...122
Intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panel A) and Western blotting (Panel B) for the culture samples of M3 taken prior to induction and at 6 h post-induction time were shown. Panels A & B: lane 1 was molecular weight markers; lanes 2, 3, 10, 11 and 4, 5 represented soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively, whereas lanes 6, 7, 12, 13 and 8, 9 represented soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows in two panels indicated the protein

bands corresponding to disulfide bonded and non-disulfide bonded monomeric fusion protein.

Figure 7.3……...123
Structure analysis of fresh M3 (day-0 samples) produced by two host strains. Panel A: SDS-PAGE analysis. Panel B: far-UV CD spectroscopy; Panel C: fluorescence spectroscopy.

Figure 7.4……...125
Monitoring and comparison of molecular property for M3 produced by two host strains with SDS-PAGE: Samples up to 15-day storage at -20 °C and 10-day storage at room temperature were analyzed. Lane 1/ molecular weight markers; lanes 2, 5, 8, 11/ day 0; lanes 3, 6, 9, 12/ -20 °C samples taken on day 15, lanes 4, 7, 10, 13/ room temperature samples taken on day 10. Arrows indicated protein bands corresponding to disulfide bonded multimeric forms.

Figure 7.5……...128
Intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panels A, C, and E) and Western blotting (Panels B, D, and F) for the culture samples taken prior to induction and at 6 h post-induction time were shown. Panels A & B, C & D, and E & F depicted the culture of M13, M135, and M35, respectively. Panels A ~ D: lane 1 was molecular weight marker; Lanes 2, 3, 10, 11 and 4, 5 represented soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively, whereas lanes 6, 7, 12, 13 and 8, 9 represented soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows in four panels indicated protein bands corresponding to disulfide bonded and non-disulfide bonded monomeric fusion protein.

Figure 7.6……...131
Structure analysis of fresh mutant variants produced by two host strains with SDS-PAGE. Panel A: M13; Panel B: M135; Panel C: M35.

Figure 7.7……...132
Structure analysis of fresh mutant variants produced by two host strains with CD spectroscopy (Panel A, C, and E) and fluorescence spectroscopy (Panel B, D, and F). Panels A & B, C & D, and E & F represented M13, M135 and M35, respectively.

Figure 7.8……...134
Monitoring and comparison of molecular property for mutant variants produced by two host strains with SDS-PAGE: samples up to 20-day storage at -20 °C and room temperature were analyzed. Panels A: M13, lane 1/ molecular weight markers; lanes 2, 5, 8 and 11 / day 0; lanes 3, 6, 9 and 12 / -20 °C samples taken on day 20; lanes 4, 7, 10 and 13 / room temperature samples taken on day 20. Panel B: M135, lane 1/ molecular weight markers; lanes 2, 5, 8 and 11 / day 0; lanes 3, 6, 9 and 12 / -20 °C samples taken on day 20; lanes 4, 7, 10 and 13 / room temperature samples taken on day 20. Panel C: M35, lane 1/ molecular weight markers; lanes 2, 4, 6 and 8 / -20 °C samples taken on day 20; lanes 3, 5, 7 and 9 / room temperature samples taken on day 20. Arrows indicated different multimeric forms of mutant variants.

List of Tables

Table 6.1…….....90
E-values (expectation values) corresponding to the substitution of cysteine 27 with the other 19 amino acids (AAs). The *E*-value is used to compare the similarity of a query sequence to a domain sequence [36]. A lower *E*-value implies a higher similarity.

Table 6.2……....102
Kinetic analysis of dimerization reaction

Nomenclature

Abbreviations

hCD83ext	extracellular domain of human CD83
GST	glutathione-S-transferase
CD	circular dichroism
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
rp-HPLC	reversed phase-high performance liquid chromatography

Symbols

M1	hCD83ext mutant variant with mutation of Cys27Ser
M3	hCD83ext mutant variant with mutation of Cys100Ser
M5	hCD83ext mutant variant with mutation of Cys129Ser
M13	hCD83ext mutant variant with mutation of Cys27Ser/Cys100Ser
M135	hCD83ext mutant variant with mutation of Cys27Ser/Cys100Ser/Cys129Ser
M35	hCD83ext mutant variant with mutation of Cys100Ser/Cys129Ser

Chapter 1

Introduction

1.1 Research Background

Therapeutic proteins continue to play a pivotal role in human healthcare and pharmaceutical industry. Approximately 150 therapeutic protein products with a current annual market value of US$ 90 billion have been commercialized and at least 500 candidates are under various trial stages around the world [3]. While the use of mammalian expression systems for the production of therapeutic proteins has gained practical interests over the past decade, the *Eschercihia coli* bacterium retains its popularity as the workhorse primarily due to its well-characterized genome and mature techniques for genetic manipulation. Its importance is also reflected by the fact that 9 out of the 31 therapeutic proteins approved between 2003 and 2006 are produced in *E. coli* [3]. A critical issue associated with therapeutic protein production is the batch-to-batch consistency of product quality. Since therapeutic protein products often undergo extensive storage before being clinically administered, their quality can be jeopardized by several protein deformation mechanisms, such as proteolysis, auto-degradation, misfolding, aggregation, oxidation, and other reactions, even under mild storage conditions.

Dendritic cells (DCs) are antigen presenting cells (APC) that circulate in the bloodstream and identify antigens in order to activate specific T cells for that antigen. Recently, the use of modified DC for the vaccination of cancer patients presents a promise for future development in anti-cancer vaccines. Human CD83 (hCD83) is a 45 kDa

glycoprotein expressed on the membrane of DC. Though the precise function of hCD83 is not yet known, it serves as a differential marker for mature dendritic cells and possibly mediates the display of the T cell's stimulatory capacity. Structurally, it is a member of the immunoglobulin (Ig) superfamily and composed of three domains, i.e. a cytoplasmic domain [39 amino acids (AAs)], a short transmembrane domain (22 AAs), and an extracellular Ig-like domain (125 AAs) (SwissProt Accession No. Q01151). There are three sites (i.e. AAs 79, 96, and 117) in the extracellular domain for glycosylation. The immunorelated function for hCD83 is believed to lie in the extracellular domain, since the hCD83 extracellular domain (hCD83ext) could inhibit DC-mediated T cell stimulation [4] and was clinically effective in treating autoimmune disorder using the murine experimental autoimmune encephalomyelitis (EAE) model [5]. Recently, the soluble form of hCD83ext was proposed as a potential therapeutic to treat transplantation rejection and a number of inflammatory and autoimmune disorders [4-6]. Currently, it is under preclinical trial conducted by Argos Therapeutics, USA (http://www.argostherapeutics/product_pipeline). The availability of purified hCD83ext has been crucial for work demonstrating its biological functions and applications. For this initiative, a bioprocess to produce recombinant hCD83ext using an *E. coli* expression system was developed [7], because the glycosylation is non-essential for the immunosuppressive bioactivity [8].

The cDNA encoding hCD83ext (i.e. AAs from 20 to 145 of hCD83) was cloned and overexpressed as a GST fusion protein (GST-hCD83ext) under the regulation of the *tac* promoter. The expressed GST fusion was captured using GST affinity chromatography and subsequently processed with on-column thrombin cleavage to release the hCD83ext moiety

as the target protein. Anion exchange chromatography was used to further remove trace amount of contaminant proteins and endotoxin from hCD83ext final product. Though very pure hCD83ext with low endotoxin level could be produced, inconsistent biological activity was encountered. The structural analysis revealed the presence of various disulfide-bonded monomeric and multimeric forms in several stored hCD83ext products from independent batches. Up to now, the knowledge of the structure of hCD83ext has been very limited. The three-dimensional structure or disulfide bond linkage has not yet been experimentally solved. There are five cysteine residues (AAs 27, 35, 100, 107, and 129) in the recombinant hCD83ext. Theoretically, any one of them can be involved in the intra- or intermolecular disulfide bond formation. An intramolecular disulfide bond between Cys^{35} and Cys^{107} was predicted, and Cys^{129} was previously reported to be critical for the dimerization [9]. In this research, recombinant hCD83ext was used as a target protein to address the consistent quality issue during therapeutic protein production.

1.2 Research Objectives

A comprehensive literature review on recombinant therapeutic protein production as well as understanding the structural properties of the target protein, hCD83ext, generated several critical aspects that need to be addressed for the manufacturing of hCD83ext with consistent quality.

1. What are critical structural factors leading to the inconsistency of hCD83ext product?
2. How do the factors influence the structure and molecular properties of hCD83ext?

3. What can be done to achieve consistent structure and biological activity of hCD83ext?

Therefore, the objectives of this research are:

1. Identify the critical factors for the inconsistent structure of hCD83ext.
2. Investigate the impact of the factors on the molecular properties of hCD8ext.
3. Elucidate the underlying mechanism by which these factors affect the hCD83ext.
4. Develop a strategy to modulate the structural properties of hCD83ext during bioprocessing.

1.3 Research Approach

The approach of this thesis is based on a comprehensive examination of the mechanism leading to inconsistent quality of therapeutic proteins as well as our theoretical analysis and characterization of the structure of hCD83ext.

Thus, it is hypothesized:

1. The inconsistent biological activity is associated with the unregulated formation of intramolecular and/or intermolecular disulfide bond under oxidative conditions.
2. The molecular properties are significantly affected by aberrant disulfide bond formation.
3. The presence of free cysteine residues might disrupt disulfide bond pattern via sulfhydryl/disulfide exchange reaction (known as disulfide scrambling).

1.4 Outline of Thesis

This thesis comprises eight chapters and is organized as follows:

Chapter 1 introduces the research background, research objectives as well as the hypotheses for the whole thesis.

Chapter 2 gives a comprehensive literature review on recombinant therapeutic protein production.

Chapter 3 introduces the bioprocess that was developed for production and purification of hCD83ext and the established analytical methods for the structure of hCD83ext.

Chapter 4 investigates hCD83ext mutant variants. A more potent therapeutic, mutant variant with mutation Cys100Ser (M3) was identified. A protocol for the analysis of disulfide bond formation in hCD83ext was established.

Chapter 5 examines the molecular property of hCD83ext produced by two *E. coli* host strains respectively with reductive and oxidative cytoplasm, demonstrating the impact of the *in vivo* disulfide bond formation.

Chapter 6 studies two hCD83ext mutant variants with mutation Cys27Ser and Cys27Glu, respectively, providing an approach to improve the storage stability of hCD83ext.

Chapter 7 explores a number of extended mutants of hCD83ext M3 to elucidate the roles of five cysteine residues in disulfide bond formation.

Chapter 8 summarizes the whole thesis and presents conclusions and recommendations for future work.

Chapter 2

Literature Review

Prior to the advent of modern biotechnology, medical applications involving proteins were less appreciated due to their very limited availabilities from natural sources like organs. With the development of recombinant DNA technology, therapeutic protein production has entered a new era [10]. Theoretically, therapeutic proteins from any biological source can be produced at large scale using appropriate gene expression systems. Produced by an *E. coli* expression system, Eli Lilly's recombinant DNA human insulin was the first officially approved recombinant pharmaceutical by the FDA (Food and Drug Administration, USA) in 1982. By 2006, about 150 therapeutic proteins have been approved in the USA and/or the EU, which have demonstrated important impacts on the pharmaceutical industry, diagnosis and treatment of various diseases, such as cancer, autoimmune diseases, and metabolic disorders. Currently, more than 500 candidate protein drugs are being evaluated in clinical trials, and advanced biotechnologies such as genomics, proteomics and high-throughput screening definitely will facilitate new drug development [3].

2.1 Expression Systems for Recombinant Therapeutic Protein Production

The choice of a host expression system is fundamental and critical for recombinant therapeutic protein production. Until the mid-90s, *E. coli* was the dominant host primarily

due to the economic competitiveness, a wealth of molecular knowledge, and mature genetic manipulation techniques. To date, the most successful cases are the commercialization of human insulin (Eli Lilly) and bovine growth hormone (bGH, Monsanto). Impressively, the cost of bulk production for bGH has been reduced to less than $5/g [11]. However, the high level expression in *E. coli* frequently results in the intracellular accumulation of insoluble aggregates of heterologous proteins (known as inclusion bodies), presumably due to the lack of efficient cellular machinery for protein folding. Extra refolding process is needed in industry to obtain soluble and bioactive therapeutic protein products. On the other hand, inclusion bodies facilitate downstream purification owing to a single low-speed centrifugation step after cellular homogenization to recover the inclusion bodies due to their high density. Strategies to achieve successful refolding of recombinant proteins from inclusion bodies are being continuously advanced [12, 13]. When therapeutic protein products require glycosylation for their biological activity, which is the case for 60% of the current therapeutic protein market, *E. coli* expression systems are no longer suitable.

Driven by the need to produce glycoproteins, mammalian cell production has garnered significant attention during the past decade [11, 14]. Mammalian cells possess superior capacity for proper protein folding, assembly and post-translational modifications. The development of genetic manipulation techniques now allows routine production of recombinant therapeutic proteins in such systems. They are the dominant hosts for the production of therapeutic proteins requiring specific glycosylation pattern for their biological activity. Chinese hamster ovary (CHO) cells and baby hamster kidney (BHK) cells are the most popular systems in the industry [10]. The productivity of mammalian cell cultures has

been significantly increased through medium composition, bioreactor design, and process control. And further improvement is expected along with the advances achieved in vector and host cell engineering [15].

Yeast host systems are also widely used in the pharmaceutical industry. Yeasts can be cultivated rapidly and cheaply with well-established scale-up processes. Compared to prokaryotic systems, more advanced regulation of post-translational modifications (e.g., disulfide bond formation and glycosylation) and more efficient secretion are offered. The best genetically characterized species *Saccharomyces cerevisiae* is by far the prevalent yeast host system, which is a good choice for production of therapeutic proteins that cannot be produced in functional forms by *E. coli* but do not stringently require human glycosylation patterns for the pharmacological activity. A variety of therapeutic proteins that are recombinantly produced by this system have been officially approved and commercialized, including human hirudin (Hoechst Marion Roussel, USA and Behringwerke AG, German), human glucagen (Novo Nordisk, Denmark), short-acting recombinant human insulin analog (Novo Nordisk), human platelet derived growth factor (Ortho-McNeil Pharmaceutical, USA and Janssen-Cilag, Switzerland), human growth factor (Biopartners, Switzerland), human papillomavirus recombinant vaccine (Merck, USA) and hepatitis B surface antigen (HbsAg), a component in combination vaccines Comvax (Merck), Tritanrix-HB (GSK) and Recombivax (Merck) [3].

At laboratory scale, the *Pichia pastoris* system, filamentous fungi systems (e.g., *Aspergillus nidulans* and *Aspergillus niger*), insect cell systems (e.g., *Trichoplusia*),

transgenic animals, and transgenic plants have been also under development, motivated by various advantages implicated in them [3, 16, 17].

2.2 Recombinant Therapeutic Protein Production in *E. coli*

2.2.1 Bioprocess Development

E. coli still remains an important workhorse for recombinant therapeutic protein production, given the fact that 9 out of the 31 therapeutic proteins approved during 2003 and 2006 are produced in *E. coli* [3]. A typical bioprocess for recombinant therapeutic protein production using an *E. coli* expression system consists of three technical stages, i.e. upstream for construction of the host/vector system, midstream for cultivation, and downstream for protein harvesting and purification (Figure 2.1). Protein production can be limited by gene expression, cultivation, or downstream processing. Corresponding strategies at all these levels are necessary to be established associated with molecular properties of therapeutic products.

Genetic manipulation of the host/vector system can significantly enhance recombinant protein production in terms of mediating functional expression, increasing cultivation performance, and facilitating downstream purification. Genetic modification on the chromosome of the *E. coli* host is routinely carried out based on the complete sequence of the *E. coli* chromosome and versatile genetic techniques [18]. Since proteolysis by homologous proteases often occurs to recombinant proteins, genes encoding proteases are usually deleted. In this regard, *E. coli* B strain BL21 is a popular host as it retains good

growth characteristics while lacking both *lon* and *ompT* genes that encode proteases Lon and OmpT. Also, *E. coli* can be modified to grow more vigorously by eliminating key enzymes in the metabolic pathways that generates toxic metabolites (e.g., acetate), which is of peculiar importance in the high cell-density culture.

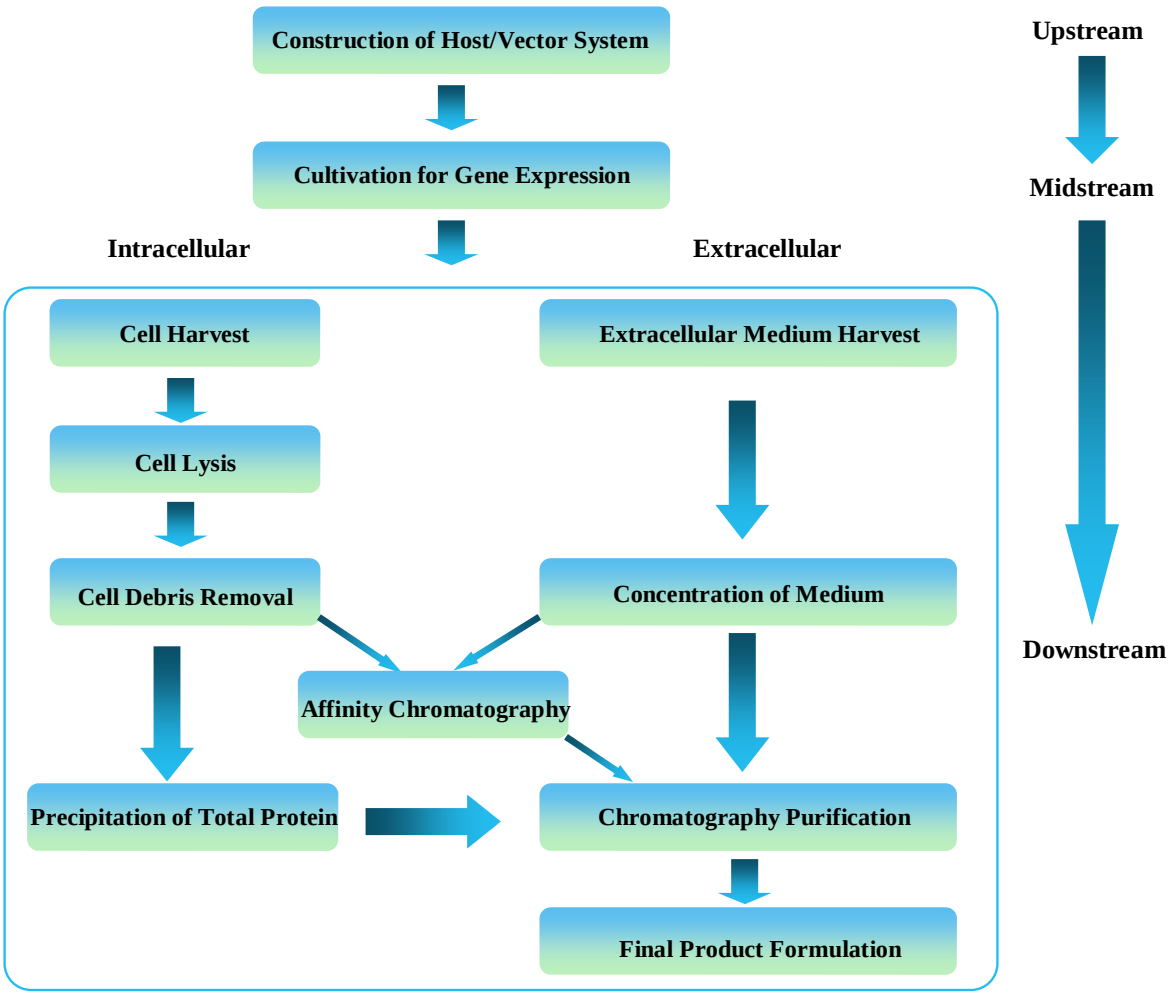


Figure 2.1 A general bioprocess scheme for recombinant therapeutic protein production.

During the construction of expression vector, the use of a strong promoter system can increase the protein yield up to 10~30% or more of the total cellular protein [19]. Also, the use of a protein fusion tag not only facilitates protein purification but also enhances gene expression [20]. The targeting of recombinant protein needs to be determined and manipulated at this stage as well. Three compartments in *E. coli* are available (i. e. cytoplasm, periplasm and extracellular secretion) with well studied mechanisms and pathways [21, 22]. If targeting the protein to the cytoplasm, the cDNA sequence encoding the therapeutic protein lies downstream of the promoter. If targeting the protein to the periplasm, a proper signal sequence mediating the translocation across the inner membrane needs to be fused with the protein of interest. If targeting the protein to the extracellular secretion, in addition to the signal sequence, a carrier protein is frequently fused with target recombinant protein, for example, outer membrane protein F (OmpF) [23]. Otherwise, coexpression of genes encoding permeabilizing proteins (e.g., *tolAIII*, the gene encoding the third topological domain of the transmembrane protein TolA [24]) or physicochemical treatment to increase the permeability of the outer membrane needs to be performed. Alternatively, fusing the target protein directly to the C-terminal of hemlysin secretion signal sequence can straightforwardly secrete the product from the cytoplasm to the extracellular medium through the hemlysin transport system [25, 26]. The advantages and disadvantages of protein folding in these compartments will be discussed later.

During cultivation in a bioreactor, the gene of interest is heterologously expressed in the exponential phase of cell growth upon induction. It is desirable to achieve both high-level gene expression and high cell-density culture. Corresponding strategies have been developed

on the basis of boosting cell growth, alleviating physiological deterioration, and enhancing functional gene expression. The medium recipe needs to be chosen by considering both the cost and the culture performance. In the case of extracellular production, less complex medium composition is preferred to simplify subsequent purification. The design of bioreactor and process control for *E. coli* cultivation has been well established with a long history [27]. Fedbatch cultivation is widely applied in industry to reach high cell-density culture [>100 g (dry cell weight)/L] but often limited by oxygen transfer and accumulation of growth-inhibiting metabolites. Cultivation parameters, such as pH, temperature, dissolved oxygen, and feeding profile, need to be monitored and adjusted to maximize culture performance. In addition, tradeoff between cell growth and functional gene expression needs to be carefully decided. For example, low temperature (e.g., 28 °C) is often chosen rather than high temperature (e.g., 37 °C) because insoluble expression often results at high temperature.

Following cultivation, a series of separation steps are undertaken during downstream processing. Most of these steps require high volume solid-liquid separation and are usually performed by centrifugation or filtration. In the case of intracellular proteins, cells are harvested at the end of cultivation followed by several pretreatment steps, including cell lysis, removal of cell debris, and recovery of total protein, to obtain a concentrated crude protein extract. Cell lysis is typically achieved by homogenization in large-scale production. At laboratory-scale, chemical and enzymatic treatment is commonly employed. In the case of extracellular product, cell-free media are directly harvested and concentrated following cultivation. After concentration, a number of high-resolution chromatographic steps are

usually carried out under low, medium or high pressure, to achieve higher purity of therapeutic product. A range of chromatography systems are available to fulfill the purification task relying on the physiochemical interaction between the protein product and the chromatographic media, including gel-filtration, affinity, metal-chelate, ion-exchange, ceramic hydroxyapatite and hydrophobic interaction. Affinity chromatography is welcomed anywhere, particularly the first chromatographic step, as its high biospecificity facilitates to achieve a very high degree of purification in one-step operation. A number of protein affinity fusion tags with their corresponding chromatographic media have been developed [20]. Fusion tag cleavage can be conducted simultaneously with fusion protein binding to the chromatography media in order to facilitate separation of the target protein and fusion tag, an approach known as on-column cleavage [28]. A final polishing step is usually implemented to remove trace contaminant proteins. Therapeutic proteins produced in *E. coli* and other Gram-negative bacterial systems are easily contaminated by endotoxin (i.e. lipopolysaccharides on the outer membrane of bacteria), which is harmful for humans at dosage rates as low as 0.5 ng/kg body weight [10]. Colorimetric-based *Limulus* amoebocyte lysate (LAL) assay is commonly used to quantify the endotoxin in the final product in the unit of EU/mg (specific activity) or EU/mL (volumetric activity). Specific membrane products have been commercialized to remove endotoxin on the basis of charge reaction, for example, Mustang® E membrane from Pall Life Science. One more membrane chromatography step to remove endotoxin may be performed, if necessary. While solid bead-packed columns are normally used as the media for chromatographic operation (i.e. column chromatography), the use of polymeric membranes (i.e. membrane chromatography) has

become popular, particularly in the polishing step (e.g., ion-exchange chromatography), due to much higher volumetric binding capacity and disposability [29].

Finally, the purified therapeutic product is formulated for storage. The final formulation has significant impacts on product quality of therapeutic proteins. Huge efforts have been made in this process, especially in the pharmaceutical industry. The basic goal of formulation is to maintain or enhance the stability and efficacy of the protein product during storage, transportation and administration. Its success depends on the understanding of the physicochemical factors of protein instability and the thermodynamic mechanism for protein-solvent interaction [30]. There are certain general principles to follow. Extremely high or low temperature will accelerate protein unfolding or other structure destructions. Proteins ubiquitously exhibit extraordinary pH dependence because charged groups play an important role in determining the structure and function of the protein in solution. Agitation and shear can disrupt protein structure to some extent and become more severe upon exposure of the protein to hydrophobic interfaces [31, 32]. A solid formulation of a biopharmaceutical is preferred sometimes to liquid formulation in order to avoid chemical degradations in solutions. Lyophilization (freeze-drying) is the most common method to achieve this purpose. Also, a wide variety of excipients have been often added during the formulation, either liquid or solid form, to stabilize therapeutic proteins [33].

2.2.2 Technical Issues and Approaches

While progress is continuously being made in recombinant therapeutic protein production, certain technical issues pertinent to assuring consistent quality of therapeutic products are still challenging. Protein molecules are marginally stable, and thus their structure is easily

partially or completely altered due to various physicochemical stresses during downstream processing and subsequent storage and transportation. Destabilizing mechanisms have been reviewed in detail and employed as the general guidelines for the formulation of therapeutic protein products [31, 34]. A summary overview is provided below.

The side-chain amide groups of asparagine and glutamine amino acids are peculiarly susceptible to hydrolysis which yields aspartic acid or glutamic acid, respectively. This reaction is known as deamidation, a major mechanism by which insulin products degrade. Imide formation frequently occurs to asparagine, aspartic acid, glutamine or glutamic acid reacting in aspartimides or glutarimides. Imide formation is promoted by X-Gly, X-Ser and X-Ala sequences (X represents asparagine, aspartic acid, glutamine or glutamic acid), resulting in increased vulnerability to hydrolysis. The studies conducted on a case-to-case basis reveal that deamidated proteins usually exhibit shortened half-life, because they become more susceptible to proteolysis. The protein conformation can be partially changed owing to increased net negative charge on the surface. Consequently, disassociation or unfolding might occur. Refolding can be inhibited by deamidation due to the increased repulsion, which might lead to aggregation of unfolded proteins.

Several amino acids are very sensitive to oxidative stress. Methionine can be oxidized to give methionine sulfoxide and even methionine sulfone. The most prominent reaction products of tryptophan oxidation are *N'*-formylkynurenine and 3-hydroxykynurenine. Tyrosine may be oxidized to dityrosine which can lead to covalently linked dimer. Histidine can be changed to 2-oxo-imidazoline through metal–catalyzed oxidation. Oxidation of cysteine can result in disulfide bond as well as sulfenic acid, sulfinic acid and sulfonic acid.

Although conformational changes do not always occur except when there are modifications of the disulfide bond, altered pharmaceutical behaviors are observed due to oxidation, including loss in bioactivity, reduced half-lives, as well as changed pharmacokinetics and immunogenicity.

In the end, inconsistent biological activity may result. While formulation stabilizes final products during extensive storage, corresponding bioprocessing strategies are necessary to improve the performance of therapeutic protein production. Modifications on protein primary structure are frequently conducted to help attain structure consistency. A number of molecular manipulation techniques can be adopted, including site-directed mutagenesis, directed evolution mutagenesis, and error-prone PCR. Since proteins undergo various environmental stresses during downstream processing, shortened operation time and mild operation conditions (e.g., avoiding high temperature and extreme pH) should be managed as much as possible. Due to the diversity and speciality of structure and function of protein molecules, establishment of effective approaches entails identification of critical factors leading to inconsistent conformation and biological activity on a case-by-case basis.

2.2.3 Structural Analysis of Final Product

1. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is the most commonly used technique in the assessment of final product purity. It separates proteins based on their different mobility in an SDS gel, which is primarily determined by the molecular weight of protein. Bands containing as little as 100 ng proteins can be visualized by staining the gel with dyes such as Coomassie blue. The detection

sensitivity can be even increased up to 100-fold with the use of silver nitrate staining. The bands additional to those associated with the protein product generally correspond to protein impurities. Quantification of purity can be conducted on the basis of densitometry.

SDS-PAGE is usually run under reducing conditions. Supplement of a reducing agent, such as β-mercaptoethanol or dithiothreitol, is conducted during SDS sample preparation to break down the intra- and intermolecular disulfide bonds. It can be run under non-reducing conditions where disulfide bonds are retained so that more information on protein structure can be obtained. Multimerization formed via intermolecular disulfide bonds can be characterized. Also, the presence of intramolecular disulfide bond(s) can be detected since the mobility of protein with intramolecular disulfide bond is greater than that with free sulfhydryl groups [35].

With the development of electrophoresis apparatus, two-dimensional (2-D) electrophoresis can be conveniently conducted to analyze the purity and structure homogeneity of final product. The mostly utilized method entails separation of proteins by isoelectric focusing (IEF) in the first dimension, with separation in the second dimension being undertaken in the presence of SDS.

2. High Performance Liquid Chromatography (HPLC)

Chromatography systems are widely applied in analyzing the purity and structure of protein product and usually operated under high pressure, such as size-exclusion (SEC) and reversed-phase (rp-HPLC). Sophisticated models of HPLC systems with automatic sampling and band analyzing are commercially available, equipped with on-line UV detectors set at ~220 or ~280 nm wavelengths. The stationary phase in the SEC is cross-linked agarose of

defined pore size supported by silica. It separates proteins primarily based on the size and is frequently utilized to analyze multimers or aggregates, as well as proteolyzed or hydrolyzed degradation products. rp-HPLC is a very sensitive analytical technique. It is capable of separating proteins of nearly identical structure but with minor differences in their surface hydrophobicity. The stationary phase in the rp-HPLC column normally consists of silica or polymeric support material to which hydrophobic alkyl chains are chemically attached. Three types of rp-HPLC columns, C4 (butyl group), C8 (octyl group), and C18 (octadecyl group), are correspondingly suitable in separation of proteins or peptides from high to low molecular weights. Mobile phase are commonly organic solvents, such as acetonitrile, trifluoroacetic acid and methanol. Proteins entering the column adsorb to the hydrophobic surface and remain adsorbed until the organic solvent concentration reaches a specific concentration (known as elution gradient), when the protein desorbs from the surface. Hence, the elution gradient represents the surface hydrophobicity of proteins.

3. Mass spectrometry

With the advent of desorption/ionization techniques, it is possible to analyze large biomolecules using mass spectrometry (MS), such as fast atom bombardment (FAB), electrospray (ESI), and matrix-assisted laser desorption ionization (MALDI) MS analysis, which provides a rapid and sensitive approach to characterize protein structural alterations as well as plays a pivotal role in peptide mapping.

4. Circular Dichroism (CD)

The chiral nature of peptide bond gives rise to CD signal when circularly polarized light beams pass through proteins. The protein secondary structure can be accurately characterized by far-UV CD measurement (180~250 nm) because peptide bonds are the dominant chromophores in this region. The tertiary structure can be reflected by the spectra obtained in the near-UV region (250~300 nm), in which aromatic amino acids and disulfide bond contribute to the CD signal to some extent. In addition to identification of banding features of each secondary structure (i.e. α-helix, β-sheet, and unordered structure) in spectra, fractional contents of each type can be deconvoluted by a number of programs which are accessible on line, e.g., DichroWeb, (http://dichroweb.cryst.bbk.ac.uk/html/process.shtml). CD measurement is conducted on a spectropolarimeter equipped with CD detector and widely applied in thermal stability characterization and ligand binding of protein.

5. Fluorescence Spectroscopy

Amino acids, like tryptophan, tyrosine and phenylalanine, give rise to fluorescence signals at the excitation wavelengths ranging from 280 to 300 nm. Among them, the emission fluorescence from the tryptophan residue is most sensitive to structural changes and thereby most often employed to monitor the local tertiary structural changes. Accordingly, longer excitation wavelengths (> 295 nm) that specifically energize tryptophan residue are chosen. Fluorescence is measured using a spectrofluorometer at a designated excitation wavelength with emission spectra being recorded. The emission maximum wavelength and corresponding intensity are subsequently used to analyze the tertiary structure.

2.3 Folding and Misfolding of Recombinant Protein in *E. coli*

E. coli has naturally developed mechanisms to help fold homologous proteins that suffer from misfolding. These mechanisms can also be used to assist the correct folding of heterologous proteins expressed in this system. *E. coli* has two sets of folding modulators, including molecular chaperones that prevent aggregation and misfolding by shielding the proteins in their cavity, and folding catalysts that accelerate rate-limiting steps during protein folding, such as the isomerization of peptidyl-prolyl bonds from an abnormal *cis* to a *trans* conformation.

2.3.1 Folding of Recombinant Protein in Cytoplasm

Upon being released from the ribosome, the newly synthesized polypeptides instantaneously enter the *de novo* folding pathways in the cytoplasm mediated by a number of chaperones. First, the short chain proteins primarily interact with trigger factor (TF), which moderately binds the ribosome at a location close to the peptide exit site. TF displays peptidyl-prolyl cis/trans isomerase activity (PPIase), which is the only PPIase identified in the cytoplasm. The longer chain proteins preferably bind to chaperone DnaK and subsequently are processed by the DnaK-DnaJ-GrpE system. Following undocking from two systems, a nascent protein may reach a native conformation, cycle back to the two systems until it folds, or flow to the downstream GroEL-GroES system which constructs a chamber in the center for proteins to undergo conformational changes at infinite dilution. *E. coli* has also evolved mechanisms to refold proteins which fold improper because of environmental stress. Partially folded or misfolded proteins are stabilized by holding chaperones, such as IbpA, IbpB, Hsp31 and

Hsp33, while awaiting DnaK and GroEL to refold them to correct conformation. In the case of formation of large and insoluble aggregates, *E. coli* has developed a mechanism to solubilize them which is primarily mediated by a disaggregating chaperone ClpB. Renaturation of solubilized aggregates is specifically performed by DnaK-DnaJ system.

Misfolding and inclusion body formation remain serious issues for cytoplasmic expression. This situation can be attributed to the overexpression of recombinant proteins which is mediated by strong promoters and high inducer concentrations. The folding modulators are rapidly used up, thereby leaving certain amount of protein misfolded. In addition, *E. coli* lacks the cytoplasmic machinery to perform all post-translational modifications that may be critically required for correct protein folding.

2.3.2 Export and Folding of Recombinant Protein in Periplasm

While some proteins remain in the cytoplasm, proteins with amino-terminal signal sequences are destined for export to the periplasm. There are three mechanisms for translocation across the inner membrane. The Sec-dependent pathway is responsible for secreting most proteins, which is composed of SecB, SecA, SecYEG, and SecYajC complex. SecB, a cytoplasmic secretory chaperone, possessing two 70-Å-long hydrophobic channels along its sides, is generally adopted to maintain preproteins in an extended form. Subsequently via its acidic top region, SecB docks and transfers the preprotein to the peripheral membrane protein SecA. Driven by ATP-hydrolysis, SecA and the cargo preprotein are transported to the pores formed by the integral membrane protein SecYEG.

Assisted by the SecDFYajC complex, the translocation to the periplasm is attained. During this process, the signal peptide is cleaved by the membrane-associated Lep or Lsp signal

peptidases. The second export pathway is the signal recognition particle (SRP)-dependent pathway which specifically translocates preproteins with highly hydrophobic signal sequences and transmembrane segments of inner membrane proteins. While both Sec- and SRP-dependent pathways are usually engaged in secreting unfolded preproteins, the twin-arginine (Tat)-dependent pathway exclusively exports folded or partially folded preproteins. Proteins translocated via the Tat-dependent pathway are frequently fused with a signal sequence containing a twin arginine motif.

After being translocated, mature proteins (after signal sequences are cleaved) are subject to folding in the periplasm assisted by a variety of periplasmic folding modulators. Skp primarily captures outer membrane proteins and misfolded periplasmic proteins. A group of modulators function as PPIases with rather broad substrate spectra, such as SurA, PpiD, FkpA, and PpiA. There are a number of modulators interacting with specific substrates to carry out chaperone activity, including SurA, LolA, PapD, and FimC. The coexpression of these modulators has been frequently employed to improve the extracytoplasmic expression. In the meantime, cell physiology was observed to be significantly enhanced by overexpression of Skp and FkpA during surface display, presumably due to alleviated stress response [36, 37].

One of the most distinguishable features of the periplasm is its oxidizing environment which enables disulfide bond formation. A set of sulfhydryl-disulfide oxidoreductase known as Dsb proteins have evolved to catalyze disulfide bond formation and isomerization in newly translocated proteins. Most periplasmic Dsb proteins contain a catalytic CXXC motif. This motif is kept in either reduced form (Dsb-$(SH)_2$) or oxidized form (Dsb-S_2), depending

on their primary function. While the DsbA and DsbB are engaged in the oxidation of cysteine residues, DsbC and DsbD mediate the isomerization of disulfide bond. The oxidative form of DsbA is a powerful oxidant and quickly reacts with unfolded proteins as they enter the periplasm, membrane-bound DsbB functions to maintain the oxidative form of DsbA. DsbA is not a specific oxidant and can introduce nonnative disulfide bonds into protein. These incorrect disulfide bonds need to be corrected to avoid misfolding and aggregation. This task is fulfilled by isomerization process. The DsbC is maintained in a reduced form by the inner membrane DsbD at the expense of NADPH oxidation in the cytoplasm. Recently, a new family member DsbG was identified that functions similar to DsbC. The overexpression of Dsb proteins, particularly DsbC has been used to improve the correct disulfide bond formation during the production of recombinant therapeutic proteins [11].

In the meantime, the periplasm also develops proteolysis mechanism to discard misfolded proteins, although the level of proteases is much lower than that of the cytoplasm. DegP and Prc are the most well studied periplasmic proteases together with several other proteases, including DegS, DegQ, and Protease III.

2.3.3 Extracellular Secretion of Recombinant Protein

Secretion to the extracellular medium may be the most ideal way to produce recombinant proteins. It does not only simplify the downstream processing, but also promotes soluble and bioactive expression of the target therapeutic proteins, in particular their disulfide bond formation in the periplasm as well as outside the cell. *E. coli* have evolved mechanisms for secretion to discard detrimental proteins [22]. The Sec-dependent system is the most widely used secretion pathway which relies on the Sec system for translocation across the inner

membrane and both the characteristics of signal sequence and proteins for extracellular secretion. There is no universal principle in selecting a proper signal sequence for a given recombinant protein to guarantee its successful secretion. Several commonly used signal peptides, either homologous or heterologous, need to be examined by trial and error. A detailed list of these signal peptides and their applications are summarized by Choi [21]. To improve the secretion efficiency of this system, it is essential to manipulate the folding of recombinant protein in the periplasm, such as co-expression of periplasmic folding modulators and minimizing periplasmic proteolysis. In the meantime, its extracellular secretion efficiency can be enhanced by fusing the target protein to a carrier protein, either normally secreted or integrated into the outer membrane. In addition, increasing the permeability of the outer membrane can help efficient excretion. Both chemical and genetic strategies have been developed for this purpose. Supplement of glycine and Triton X-100 to the medium and physicochemical treatment of cell membrane (e.g., osmotic shock, freezing/thawing, chloroform) are frequently used to create "leaks" in the outer membrane. Also, co-expression of *kil* [38], *TolAIII* [24], *out* [39], or bacteriocin release protein (BRP) genes can augment the yield of extracellular production [40].

The hemolysin transporter system (Hly) is another popular secretion system [21]. The complex formed by inner membrane protein HlyD trimer and HlyB interacts with the outer membrane protein TolC to form an export conduit to transport the HlyA toxin/HlyA-fusion protein from the cytoplasm to the extracellular medium. By fusing to the C-terminal of Hly secretion signals, a number of therapeutic proteins have been successfully secreted in an active form, including single chain Fv antibody [25] and human interleukin-6 [26].

2.4 Disulfide Bond in Proteins

Proteins are linear polypeptide chains that fold into a 3-D structure to yield a functioning molecule. The principles governing protein folding have been postulated by Christian Anfinsen, who received the Nobel prize for this work. He postulated that all the information required for the proper folding of a protein to its native 3-D structure is contained in its amino acid sequence [41]. The 20 constituent amino acids mainly interact via electrostatic, Van der Waal, hydrogen bonding, and hydrophobic forces. Cysteine is the only amino acid that allows the formation of covalent bonds. This bond is known as a disulfide bridge and is formed via the cross-linking of two sulfhydryl groups in two cysteine residues through oxidization. Disulfide bridges are present in numerous proteins across a wide variety of organisms ranging from simple unicellular prokaryotes to complex multi-cellular organisms.

2.4.1 Effects on Protein Stability

Owing to its covalent nature, disulfide bonds are generally believed to stabilize proteins, especially under extremely harsh conditions. Bovine pancreatic ribonuclease A [42] is a classical example of protein stabilization through disulfide bonds. Efforts have been made to improve protein stability via *de novo* engineering of disulfide bonds. However, successful cases are very limited mainly due to the fact that the mechanism by which the disulfide bond stabilizes protein remains imprecise [43]. By far, the stabilizing effects of disulfide bonds are proposed to be a result of reduction in conformational entropy in the unfolded state [44], combinatory entropic and enthalpic effects that involve not only interactions within protein chains but also interactions with surrounding solvents [45], or primarily enthalpic effects

[46]. In any case, it is a well established fact that the protein with native (or correct) disulfide bonds are more stable than variants without disulfide bonds [45]. Furthermore, the pattern (or position) of disulfide bonds is believed to significantly affect the stability of the protein. While correct disulfide bonds stabilize proteins, incorrect disulfide bonds may result in a destabilizing effect. In many cases, the formation of correct disulfide bond in proteins is essential for them to reach or stabilize a biologically active conformation. For example, human insulin requires correct formation of intra- and interchain disulfide bonds to function properly. The inability to form disulfide bond and/or aberrant disulfide bond formation is the key factor leading to the protein misfolding, e.g., inclusion body formation in *E. coli*.

2.4.2 Disulfide Bond Chemistry

The *in vivo* formation of disulfide bonds is carried out during post-translational processing and is assisted by a series of oxidoreductases in prokaryotes [47] or protein disulfide bond isomerases (PDI) in eukaryotes [48]. There are three kinds of reactions involved, including oxidation, reduction, and isomerization (Figure 2.2).

By reacting with the oxidase, a disulfide bond is introduced to the substrate protein. Equivalently, a disulfide bond formed in the substrate protein can be reduced by the reductase. The rearrangement of disulfide bonds within a substrate protein can occur through two different mechanisms: a two-step process with a reduction reaction followed by an oxidation reaction or a single-step isomerization process. The two-step mechanism involves the presence of two oxidoreductases, an isomerase that reduces the disulfide bond and an oxidase that forms the disulfide bond. On the other hand, the single step mechanism involves only the isomerase enzyme. In this case, the assisting oxidoreductase forms a mixed disulfide

by attacking an existing disulfide bond. Depending on the conformation, such mixed disulfide bond may be resolved by an adjacent cysteine residue resulting in the formation of a different disulfide bond (Figure 2.2C).

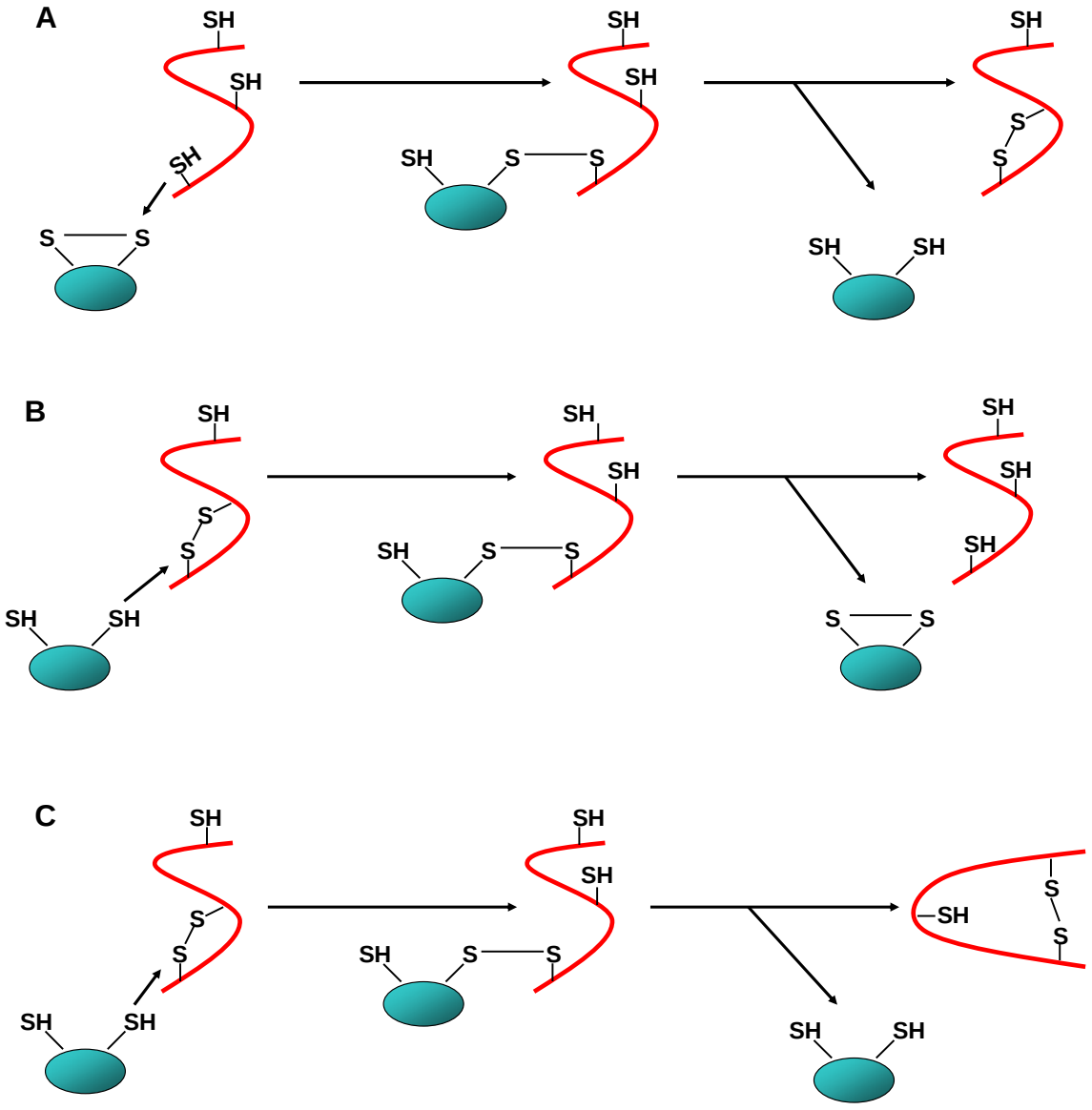


Figure 2.2 The *in vivo* disulfide bond reactions catalyzed by oxidoreductases. (A) Oxidation, (B) Reduction, and (C) Isomerization. The blue oval represents the oxidoreductase.

All the above reactions can take place *in vitro* as well, for example, during downstream processing and/or storage. Atmospheric oxygen and small sulfhydryl-disulfide chemical reagents such as glutathione and DTT are usually employed to function in a similar way as the oxidoreductases. The rearrangement of disulfide bonds within the protein may even occur in the absence of an exogenous sulfhydryl-disulfide reagent. The protein sulfhydryl groups may attack the existing disulfide bond, resulting in sulfhydryl-disulfide exchange. This is known as disulfide scrambling that can take place intra- and intermolecularly, potentially leading to protein conformational changes (Figure 2.3). It is particularly aggravated under protein destabilizing conditions, e.g., in the presence of denaturant or high temperature. Disulfide bonds may also undergo irreversible β-elimination reactions in alkaline solutions as shown in Eq. a and b:

$$R_1CH_2SSCH_2R_2 + OH^- \rightleftharpoons R_1CH_2SS^- + CH_2 == CR_2 + H_2O \quad (a)$$

$$R_1CH_2SS^- + OH^- \rightleftharpoons R_1CH_2SO^- + HS^- \quad (b)$$

Hydroxides attack protein disulfide bond, and as a result, disulfide bonds are cleaved concomitant with the formation of dehydroalanine and persulfide. Following an electrophilic addition reaction, ε-amino group of lysine or the sulfhydryl of cysteine can be inserted to the double bond of dehydroalanine, yielding lysinoalanyl or lanthionyl cross-links, respectively. Persulfides may further react with hydroxide to release hydrosulfides that open chances to introduce additional modifications to the protein, such as disulfide scrambling [49]. The β-elimination develops very rapidly at 90~ 100 °C in the pH range of 6 to 8 or at room

temperature in extremely basic solutions. Hence, β-elimination is recognized as a major protein degradation mechanism.

2.4.3 Disulfide Bond in Protein Folding

In many cases, proteins do not fold to their native conformation until nearly all of its native disulfide bonds have been formed, suggesting that disulfide bond can be a probe to investigate protein folding pathways [50, 51]. Bovine pancreatic trypsin inhibitor (BPTI) has long been a model protein to study the mutual effects of disulfide bond formation and conformational folding on each other [50, 52, 53]. Oxidative folding processes of fully reduced and denatured proteins are usually initiated by introducing sulfhydryl-disulfide reagents. Mixed disulfide bonds are then formed between protein molecules and reagents. Followed by numerous sulfhydryl-disulfide exchange reactions, substantial disulfide bonded folding intermediates are captured and characterized to elucidate disulfide-coupled folding pathways. Similar methodology has been successfully applied to decipher the folding pathways of a number of proteins, including porcine insulin [54], human proinsulin [55], and RNase A [56]. Great detail on folding pathways only can be obtained if the disulfide bonded intermediates are stable and can accumulate to significant level. Disulfide bond patterns in the intermediates are usually explored by peptide mapping relying on chemical/enzymatic digestion followed by rp-HPLC separation and MS analysis.

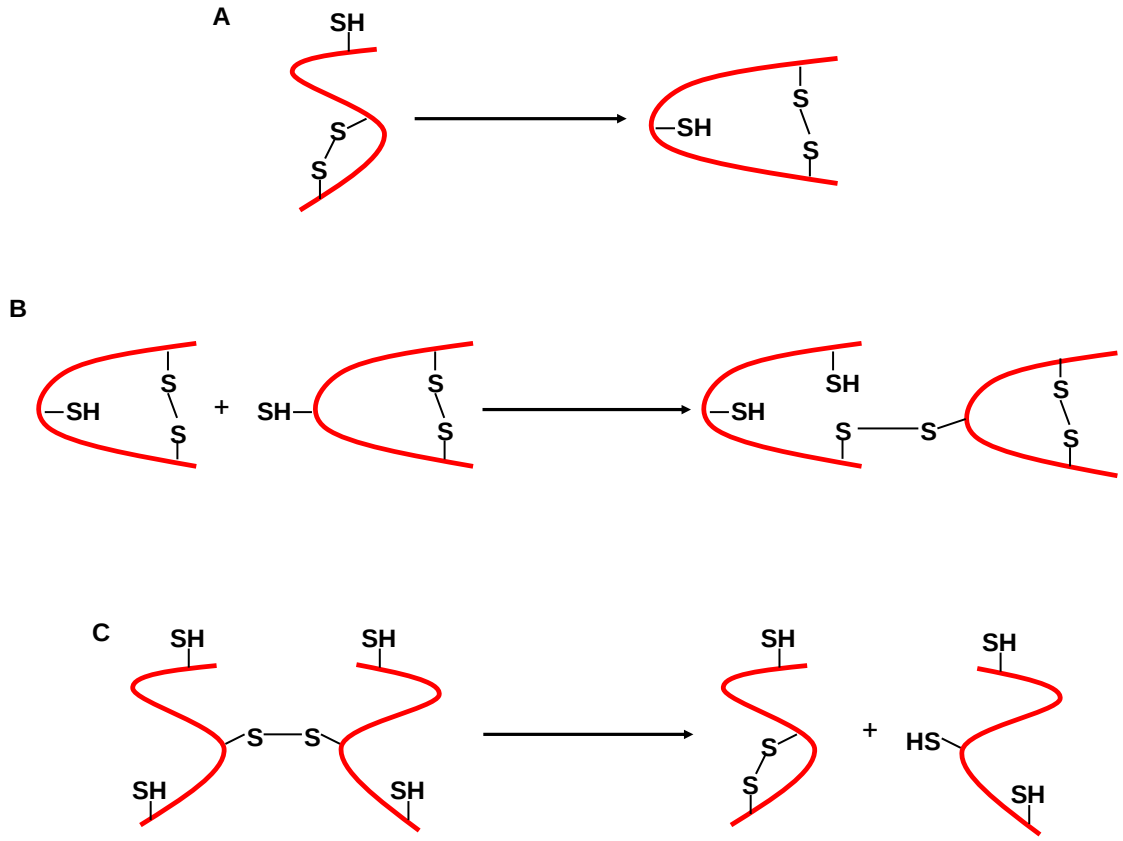


Figure 2.3 Disulfide scrambling. (A) sulfhydryl-disulfide exchange reaction occurs intramolecularly, resulting in a different intramolecular disulfide bond. (B) sulfhydryl-disulfide exchange reaction occurs intermolecularly, resulting in formation of an intermolecular disulfide bond, (C) An intermolecular disulfide bond is broken by a sulfhydryl group via sulfhydryl-disulfide exchange reaction.

2.5 Disulfide Bond Formation in Recombinant Therapeutic Protein Production

Many commercialized therapeutic proteins contain disulfide bonds. The correct formation of disulfide bonds is often a prerequisite for these proteins to achieve their biologically active three-dimensional (3D) conformation [56-58]. On the other hand, aberrant formation of disulfide bonds has been recognized as a critical factor leading to protein misfolding, aggregation, and malfunction [59-62]. Therefore, the production of bioactive therapeutic proteins can be limited by the correct formation of disulfide bonds. This technical issue is particularly more challenging when *E. coli* is used as the expression system since it lacks the cellular machinery as efficient as that of the mammalian expression system to regulate this posttranslational event. Subsequent exposure of the protein intermediate or final product to an oxidative environment during downstream processing and storage stages can also result in the aberrant formation of disulfide bonds and/or the interchange between free sulfhydryl groups and disulfide bonds, known as disulfide scrambling [51, 56].

Rearrangement of disulfide bond, known as disulfide shuffling, can be performed *in vitro* to achieve a desirable disulfide bond pattern. It starts with denaturing and reducing protein polypeptides followed by exposing polypeptides to redox reagents, such as glutathione, dithiothreitol and cysteine. Theoretically, disulfide bonds are first formed between polypeptides and redox reagents. Subsequently, free cysteine residues in the polypeptides will be involved in sulfhydryl/disulfide exchange reactions to form desirable intramolecular disulfide bonds. The redox state of the protein solution can be precisely controlled by the concentrations of the reduced and oxidized forms of the redox reagent as

well as their ratio to the polypeptide-cysteine concentration. Using this approach, several disulfide-bond-containing therapeutic proteins, including human interleukin-4 (with three disulfide bonds) [63] and tissue plasminogen activator (t-PA; with 17 disulfide bonds) [64], have been successfully renatured to achieve their correct disulfide bond patterns. Likewise, the inclusion of redox reagents in downstream processing steps is anticipated to be helpful [65]. Protein disulfide isomerases (PDIs) from different sources have been exploited, either in a free or an immobilized form [66], to promote the correct formation of disulfide bonds *in vitro* [67, 68], particularly in cell-free systems for protein synthesis [69].

From a technical point of view, it is still desirable to express proteins with a correct disulfide bond pattern in *E. coli* since no extra processing will be required for disulfide shuffling and protein refolding. In addition, the intracellular environment is supposed to be more consistent than the *in vitro* one for disulfide bond formation. The periplasm of *E. coli* is an ideal intracellular compartment to express proteins with disulfide bonds because it is oxidative and possesses the Dsb-family proteins for mediating the *in vivo* formation disulfide bonds [47]. The overexpression of the Dsb-family proteins appears to significantly facilitate the formation of disulfide bonds for the functional expression of several therapeutic proteins, such as leptin, single-chain Fv antibody fragment, and t-PA [70-72]. Nevertheless, heterologous protein expression targeting this compartment is often limited by ineffective translocation. As a result, *E. coli* expression systems enabling the formation of disulfide bonds in the cytoplasm have been developed [73-75] and several strains, e.g. Origami (Novagen; with two deletions of the *trxB* gene encoding thioredoxin reductase and the *gor* gene encoding glutathione reductase) and SHuffle (New England Biolab; similar to Origami

but with an overexpression of DsbC), have been made commercially available. The approach of cytoplasmic expression has been explored for the functional expression of several proteins, including Fab antibody fragment [76], t-PA [73], lipase B [77], and human 11-β-hydrooxysteroid dehydrogenase [78].

Since disulfide bonds are formed by crosslinking cysteine pairs transiently close to each other, it is plausible to assume that the formation probability can be explicitly determined by protein structure or even intrinsically encoded by protein sequence. However, up to now, there is limited knowledge and experimental demonstration as to how protein conformation and/or protein sequence can affect the formation mechanism associated with aberrant disulfide bonds. Theoretically, to ensure consistent quality of therapeutic products, protein molecules should be designed and produced in such a way to both mediate a fixed disulfide pattern and to prevent the formation of aberrant disulfide bonds. In this regard, understanding various structural factors associated with disulfide bond formation becomes critical for quality production of therapeutic proteins at both the stages of molecular design and biomanufacturing.

Chapter 3

Project Background: Bioprocess Development for Production, Purification, and Structural Characterization of Recombinant hCD83ext as a Potent Therapeutic Protein

The work in this chapter was published in *Protein Expr. Purif.* 65 (2009) 92-99.

Authors: Yali Xu, **Lin Zhang**, Weifang Yao, Shreyas S. Yedahalli, Stephen Brand, Murray Moo-Young, C. Perry Chou

Declaration: my contribution to this paper is the structural characterization of final protein product. Yali Xu chiefly contributes to the bioprocess development for production and purification. The inclusion of this work in the thesis is to describe the production protocol of recombinant hCD8ext and illustrate the inconsistent structure associated with inconsistent disulfide bond formation.

3.1 Introduction

In this work, we developed a bioprocess for effective production of hCD83ext with a high quality suitable for various *in vivo* and *in vitro* trials for bioactivity evaluation. Culture conditions were determined based on maximization of the volumetric recombinant protein titer by manipulating various operating parameters, including medium recipe, host/vector system, induction condition (e.g. inducer concentration and induction timing), temperature, and dissolved oxygen. Downstream processing steps were established based on several

purification criteria, such as product throughput, yield, purity, and endotoxin level. In addition, extensive monitoring of the structural change, multimerization, and stability associated with the final hCD83ext product was conducted.

3.2 Materials and Methods

3.2.1 Bacterial Strains and Plasmids

The plasmid of pGEX2ThCD83ext containing the hCD83ext cDNA fused with the *gst* gene was used as the expression vector for the production of recombinant GST-hCD83ext under the regulation of an isopropyl β-D-thiogalactopyranoside (IPTG)-inducible *tac* promoter [8]. Due to the design of a thrombin cleavage site at the junction between GST and hCD83ext, the final hCD83ext product has four extra AAs, i.e. Gly-Ser-Pro-Gly, in the amino terminus and this modification was presumed to have no effect on the property of hCD83ext. BL21 (F^- *ompT gal [dcm] [lon] hsdS*$_B$ (r_B^- m_B^-)) [79] was used as the expression host. Purification of plasmid DNA was performed using a spin-column kit (BD Biosciences Clontech, Mississauga, Ontario, Canada). Plasmid transformation was carried out using a chemical method. Cultivation cells were revived by streaking the stock culture, stored at -80 °C, on LB agar plates (5 g/L NaCl, 5 g/L Bacto yeast extract, 10 g/L Bacto tryptone, and 15 g/L Bacto agar). The plate was incubated at 37 °C for approximately 15 h. An isolated single colony was picked to inoculate 100 mL of LB medium, which was then incubated at 37 °C on a rotary shaker at 200 rpm for approximately 12 h. The medium was supplemented with 50 μg/mL ampicillin (Ap) when necessary. The seed culture (80 mL) was used to inoculate a

bench-top bioreactor (Omni-Culture, VirTis, Gardiner, NY, USA) containing 1-L working volume of culture medium (5 g/L NaCl, 20 g/L Bacto yeast extract, 20 g/L Bacto tryptone, 5 g/L glucose, and 10 μL/L Antifoam 289 (Sigma, St. Louis, MO, USA)). When the culture cell density reached ~2 OD_{600}, IPTG at 0.5 mM was added for induction. The bioreactor was purged with filter-sterilized air at 2 L/min for aeration. The culture pH was regulated at 7.0 ± 0.1 by adding 3 N NH_4OH or 3 N HCl using a combination of pH electrode (Mettler-Toledo, Switzerland), a pH controller (PC310, Suntex, Taipei, Taiwan), and two peristaltic pumps (101 U/R, Watson Marlow, Falmouth, UK). The bioreactor was operated at 28 °C and 650 rpm for approximately 6 h after induction.

3.2.2 Sample Treatment for Analyses

The culture sample was appropriately diluted with saline solution for measuring cell density in OD_{600} with a spectrophotometer (DU® 520, Beckman Coulter, Fullerton, CA, USA). For the preparation of cell extract, cells in the amount of 30 OD_{600}-units (defined as 'OD_{600} × mL') were centrifuged at 6000 × *g* and 2 °C for 10 min. The cell pellet was resuspended in 0.75 mL phosphate-buffered saline (pH 7.3), sonicated intermittently (i.e. 0.5 s/0.5 s on/off) for 4 min using an ultrasonic processor (Misonix, Farmingdale, NY, USA) with a microtip, and then centrifuged at 15,000 × *g* and 2 °C for 15 min to remove cell debris. The supernatant containing soluble proteins was used for various analyses, including GST assay, SDS–PAGE, and Western blotting. The pellet containing insoluble proteins and cell debris was washed with phosphate buffer, resuspended in TE/SDS buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, 1% SDS), and heated to 100 °C for 5 min for dissolution. The protein content of the pellet was analyzed as the insoluble fraction.

3.2.3 Downstream Processing

After cultivation, the cells were harvested as described above, weighed, and stored at -80 °C for later processing. Typically, cell paste at approximately 20 g wet cell weight (wcw) could be obtained from 1-L culture. Cell suspension at 0.05 g wcw/mL was prepared by suspending the cell paste in PBS. A batch of 20-mL cell suspension at approximately 20 OD_{600} was sonicated intermittently (i.e. 0.5 s/0.5 s on/off) for 4 min using an ultrasonic processor with a regular tip (XL-2020; Misonix, Farmingdale, NY, USA). The processed cell lysate was then centrifuged at 15,000 × *g* and 2 °C for 15 min to remove cell debris. The supernatant containing total soluble proteins was filtered with a 0.45 μm filter before subsequent chromatographic processing for protein purification.

To capture the GST-hCD83ext fusion, the above-prepared lysate with PBS as a binding buffer was loaded into a GST affinity chromatographic column (GE Healthcare, Baie d'Urfé, Québec, Canada). The bound GST-hCD83ext fusion was in-situ cleaved with thrombin (Sigma). To conduct on-column cleavage, the binding buffer in the GST affinity chromatographic column bound with GST-hCD83ext fusion was replaced by the same volume of the binding buffer containing thrombin at a designated concentration through manual injection. The cleavage was conducted at room temperature for approximately 2.5 h and the bulk liquid in the GST affinity chromatographic column was ejected using 1 column volume of the binding buffer. The GST moiety along with undigested GST-hCD83ext was then eluted using the elution buffer (50 mM Tris, 10 mM glutathione, pH 8.0).

The polishing step was conducted to remove thrombin and other contaminant proteins from hCD83ext. A low-pressure chromatographic system (BioLogic LP, BioRad, Hercules,

CA, USA) equipped with a strong anion exchange column (Q, GE Healthcare) was used for the processing. Tris buffer (20 mM, pH 7.5) with 50 mM NaCl and Tris buffer with 1 M NaCl were used as the loading and elution buffers, respectively.

The fractions containing the pure hCD83ext were pooled and concentrated with ultrafiltration using a high-pressurized stirred cell (Amicon, Model 8010 with YM10 disk, Millipore Canada, Cambridge, Ontario, Canada). The hCD83ext final product was filter-sterilized before storage. Using this final product fraction, the extinction coefficient of hCD83ext was determined to be approximately 1.16 OD_{280}-mL/mg/cm.

3.2.4 Analytical Methods

GST was assayed at ambient temperature using 1-chloro-2,4 dinitrobenzene (CDNB) and glutathione as the substrates. GST (or its fusion) catalyzes the reaction of the two substrates to form a conjugate product that can be monitored at 340 nm. One unit (U_{GST}) is defined as the amount of enzyme that causes an increase in absorbance at 1 OD_{340}/min. The volumetric activity (in U_{GST}/mL) is the product of the specific activity (in U_{GST}/OD_{600}-unit) and cell density (in OD_{600}). Total proteins were assayed using a bicinchoninic acid kit (Sigma-Aldrich, St. Louis, MO, USA). LAL assay was conducted to measure the endotoxin level using a commercial kit from Cambrex (East Rutherford, NJ, USA).

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) was performed in a Mini-PROTEAN® II electrophoresis cell (Bio-Rad, Hercules, CA, USA) using a 12.5% polyacrylamide separating gel stacked by a 4% polyacrylamide stacking gel. Electrophoresis was conducted at a constant voltage of 120 V for approximately 2 h. The gel was stained with Coomassie blue or silver nitrate and dried. The dried gel was scanned.

To conduct Western blotting, the proteins on the polyacrylamide gel were electroblotted to a poly(vinylidene difluoride) (PVDF) membrane (Pall Life Sciences, Pensacola, FL, USA) after SDS–PAGE using a Mini Trans-Blot® Cell (Bio-Rad) according to a standard protocol [80]. The electrophoretic transfer was conducted at a constant voltage of 100 V for 1 h. Protein–antibody binding was performed as described by Sambrook et al. [81]. Mouse monoclonal anti-CD83 antibody (CD83-1G11, Cedarlane laboratories limited, Hornby, Ontario, Canada) was used as the primary antibody. On the other hand, the purified hCD83ext was used as an antigen for raising polyclonal anti-CD83 antibodies in rabbits. It appears that the two types of antibodies have the same binding affinity to various hCD83 species (data not shown). The secondary antibody was goat anti-mouse IgG conjugated with horseradish peroxidase (Sigma-Aldrich). hCD83-related polypeptides (e.g. GST-hCD83ext and hCD83ext) were detected by a colorimetric method using 3,3'-diaminobenzidine tetrahydrochloride (DAB) as the substrate. The processed membrane was scanned.

Far-UV circular dichroism (CD) spectra of protein samples were obtained using a spectropolarimeter (Jasco-815; Jasco, Tokyo, Japan). Fluorescence emission spectra (over the range of 300~420 nm) of protein samples were obtained using a spectrofluorometer (FP-6500; Jasco) with an excitation wavelength at 295 nm. All the spectroscopic measurements were made at a room temperature of 23 °C.

3.3 Results

Using a bench-top bioreactor for cultivation for the production of recombinant GST-hCD83ext, several culture parameters, including medium recipe, induction condition (i.e. induction timing and IPTG concentration), agitation speed, and temperature, were found to critically affect the culture performance. The cultivation was terminated at approximately 6 h after induction since the protein yield reached a plateau and would decrease upon extended cultivation. The downstream processing primarily consists of three stages for hCD83ext purification. Initially, the expressed GST-hCD83ext was captured in a GST-affinity column. Subsequently, instead of being eluted, GST-hCD83ext was *in situ* cleaved with thrombin while it bound to the GST-affinity column medium and the hCD83ext moiety was released into the bulk liquid phase containing thrombin. Finally, the liquid mixture was processed with anion exchange chromatography so that thrombin, other contaminant proteins, and endotoxin could be removed for hCD83ext purification. Ultrafiltration was conducted to concentrate this fraction and the final hCD83ext product with a high purity was obtained. To optimize the *in situ* cleavage of GST-hCD83ext, the optimum thrombin concentration and the cleavage time were determined. More than 95% of GST-hCD83ext was cleaved when the thrombin concentration was above 10 $U_{thrombin}$/mL column medium, which was eventually designated as the optimum thrombin concentration for conducting in-situ cleavage. With such a relatively low thrombin concentration, the incubation time was prolonged to 2.5 h to ensure complete cleavage. It appeared that hCD83ext was the only protein minimally interacting with the anion-exchange medium under this operating condition and a pool of the

flowthrough containing extremely pure and low-endotoxin hCD83ext was collected and concentrated by ultrafiltration [7].

The purified hCD83ext, formulated either in the presence or absence of 50% glycerol, tended to degrade when being stored at 4 °C perhaps due to the intrinsic instability of this protein molecule. Therefore, the protein had to be stored at -20 °C to ensure its stability. Nevertheless, dimerization could still occur under this frozen condition during long-term storage. The SDS–PAGE analytical results for hCD83ext samples from different batches, hence with different storage ages, are summarized in Figure 3.1. Though these samples showed an identical pattern when analyzed by reduced SDS–PAGE (Figure 3.1A), they showed different patterns in term of the dimer composition when analyzed by non-reduced SDS–PAGE (Figure 3.1B). Generally, the dimer composition increased monotonically with the sample age, indicating the progressive conversion of monomer to dimer under this storage condition. Note that there were at least two different monomer forms (lane 6, Figure 3.1B) and three different dimer forms (lanes 3 and 7, Figure 3.1B) derived from various sample preps. Cys^{129} has been identified as the key AA residue mediating an intermolecular disulfide bond to form hCD83ext dimers [9]. In addition to dimers, protein bands corresponding to trimers, tetramers, and even multimers were observed in certain hCD83ext samples and these bands were justified to behCD83ext-related when probed with anti-CD83 antibodies (Figure 3.2), implying other cysteine residues could also mediate unconventional disulfide bonds for multimerization.

The structural change of hCD83ext was further investigated using a sample stored at 4 °C for 15 days and the analytical results are summarized in Figure 3.3. There was slight degradation

and significant dimerization over the storage period (Figure 3.3A and B). The far-UV CD spectra of 0-day and 15-day samples (Figure 3.3C) displayed a negative band at 220 nm and a positive band of comparable magnitude near 203 nm, which is a typical pattern of β-sheet mixed with type II β-turn [82]. This is rather consistent with our prediction of hCD83ext 3-D structure, which is comprised by nine antiparallel β-strands in a typical b-sandwich V-type immunoglobulin fold (data not shown).

Though the CD spectra for the two samples appeared to be distinctive, the spectroscopic analysis using CDSSTR method [83] only showed a slight difference in β-sheet content (i.e. 4% α-helix, 23% β-turn, 39% β-sheet and 33% unordered for the 0-day sample, 4% α-helix, 23% β-turn, 35% β-sheet and 37% unordered for the 15-day sample). Fluorescence emission spectra of the two samples using 295 nm as the excitation wavelength were compared in Figure 3.3D. The emission maximum had a slight blue-shift from 352 nm (0-day) to 342 nm (15-day), indicating that the tryptophan residues (i.e. W39 and W49) became more buried inside the hydrophobic core of the protein molecule when dimerization occurred. These results suggest that the structural variation of hCD83ext, associated with disulfide bond formation, would be a critical issue for rigorous control of the quality or even bioactivity of hCD83ext as a therapeutic protein.

3.4 Discussion

Due to the intrinsic instability of hCD83ext, the protein needs to be stored under a low temperature. Though degradation could be prevented as a result, the protein underwent

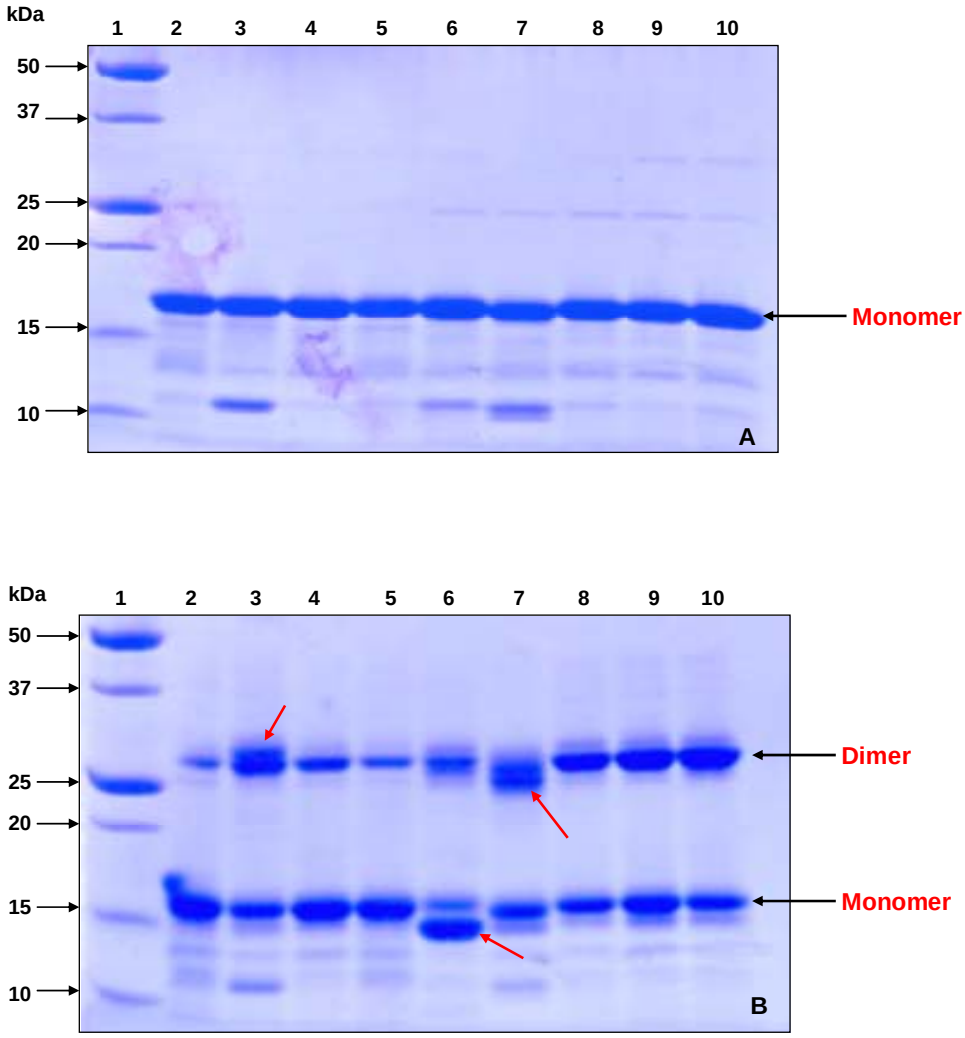


Figure 3.1 Analysis of various hCD83ext sample lots with (A) reduced SDS-PAGE and (B) non-reduced SDS-PAGE. The protein samples were stored at -20 °C and were formulated in 50% glycerol except the lane-3 sample which was glycerol-free. The lots were made on several different dates, therefore had different "ages" (i.e. 4, 3, 3, 2, 10.5, 9.5, 6.5, 6, 6 months for lanes 2~10, respectively) upon conducting the analysis. The arrows indicated several uncommon hCD83ext species possibly mediated by unconventional intramolecular disulfide bonds.

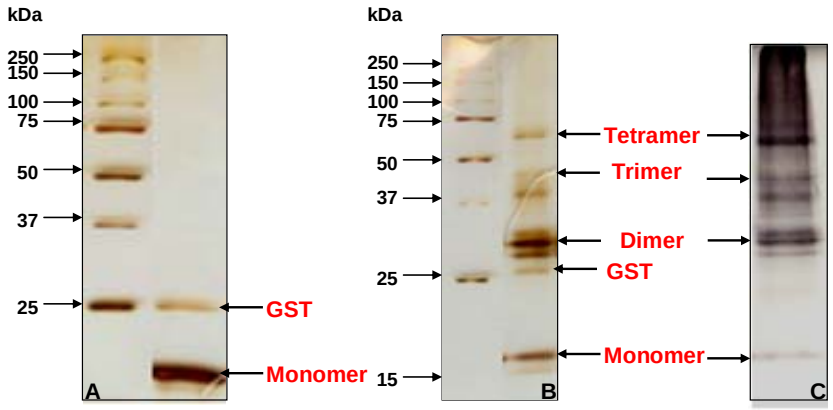


Figure 3.2 Analysis of an hCD83ext sample with (A) reduced SDS-PAGE, (B) non-reduced SDS-PAGE, and (C) Western blotting. The SDS gels were stained using silver nitrate in order to make higher multimeric species (such as trimer and tetramer) visible. All these species were verified to be associated with hCD83ext by Western blotting. Note that a small amount of GST exited in the hCD83ext product fraction as a contaminant protein.

multimerization via the formation of intermolecular disulfide bonds under this relatively oxidative storage condition. There are five cysteine residues (AAs 27, 35, 100, 107, and 129) in hCD83ext and Cys^{129} was identified as the major one mediating the intermolecular disulfide bond to form dimers [9]. However, the other four cysteine residues can be involved in the formation of intramolecular [9] or even intermolecular disulfide bonds. The multiple forms of monomers and dimers in Figure 3.1 could possibly result from different pairing for intramolecular disulfide bond formation among these five cysteine residues. Recently, we have demonstrated that hCD83ext could form dimers via unconventional intermolecular

disulfide bonds using cysteine residues other than Cys^{129}. Interestingly, in addition to dimer, higher multimeric forms (i.e. trimer, tetramer, etc.) were also observed in a few hCD83ext samples (Figure 3.2). The results suggest that, though Cys^{129} is the key residue driving the formation of the intermolecular disulfide bond for dimerization, other cysteine residues could also be involved in the extensive formation of intermolecular disulfide bonds for multimerization.

Theoretically, the GST-hCD83ext fusion originally harvested from the cytoplasm (which is a reductive compartment) of producing cells should not contain any disulfide bonds. However, disulfide bond formation could occur in any downstream processing steps which were normally conducted in an oxidative environment. In fact, we observed the existence of hCD83ext dimers in the post-GST fraction (i.e. the eluent fraction from GST affinity chromatographic column after thrombin cleavage) though at a small amount (data not shown). With the potential application of hCD83ext as a therapeutic protein, this posttranslational issue of disulfide bond formation during downstream processing could have an impact on the product quality and needs to be properly addressed. In addition, developing a proper formulation recipe leading to a specific pattern of disulfide bond formation for hCD83ext product might be required to obtain consistent therapeutic activity.

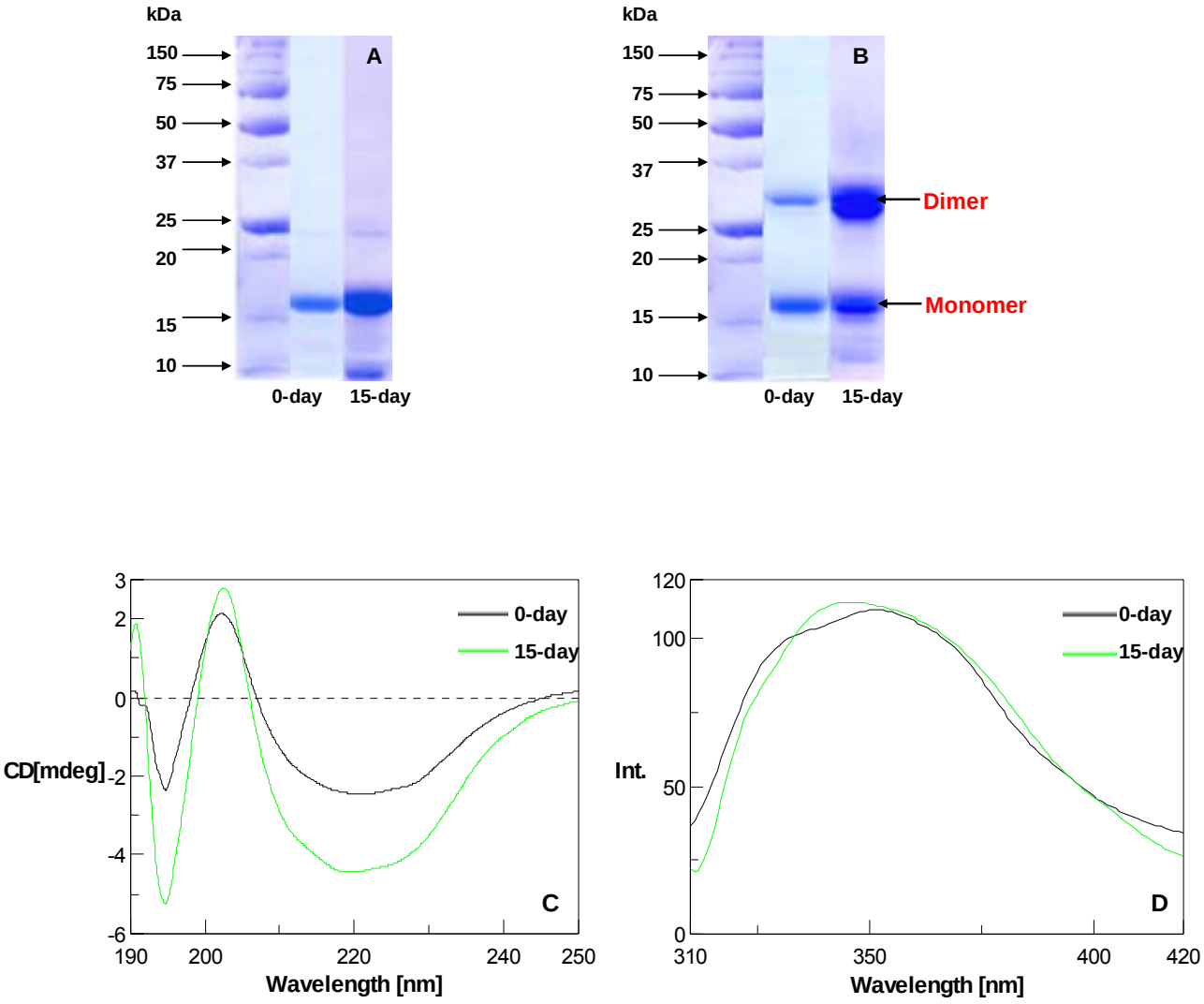


Figure 3.3 Comparison in structural change of an hCD83ext sample over 15-day storage at 0 °C using (A) reduced SDS-PAGE, (B) non-reduced SDS-PAGE, (C) CD spectroscopy, and (D) fluorescence spectroscopy.

Chapter 4

Structural Identification of Recombinant Human CD83 Mutant Variant as a Potent Therapeutic Protein

The work in this chapter was published in *Protein Expr. Purif.* 73 (2010) 140-146.

Authors: **Lin Zhang**, Niju Narayanan, Stephen R. Brand, Charles A. Nicolette, Miren Baroja, Jacqueline Arp, Hao wang, Murray Moo-Young, C. Perry Chou

Declaration: I initiate and conduct the structural study of wild-type and mutant variants of hCD8ext, which is the main content of this chapter. Niju Narayanan kindly derives the mutant variant with mutation of Cys100Ser. Miren Baroja, Jacqueline Arp and Hao wang in the Department of Surgery, University of Western Ontario, contribute to the bioactivity assay.

4.1 Introduction

There are five cysteine residues (AAs 27, 35, 100, 107, and 129) in the recombinant hCD83ext and, theoretically, any one of them can be involved in the aberrant formation of disulfide bonds, either intermolecularly or intramolecularly. Previously, Cys^{129} was identified as the major cysteine mediating the intermolecular disulfide bond to form hCD83ext dimers and the Cys129Ser mutant derivative still had comparable bioactivity [84] without forming dimers [9]. Nevertheless, the protein product of this mutant derivative produced in our lab had an inconsistent bioactivity (unpublished data), similar to the case of wild-type hCD83ext.

The presence of higher multimeric forms, such as trimers and tetramers, for wild-type hCD83ext [7] suggests that cysteine residues other than Cys^{129} could also be involved in disulfide bond formation.

In this study, we re-examined and modified the sample preparation protocol to analyze hCD83ext products. In contrast to a previous report [9], our results based on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and size exclusion chromatography (SEC) demonstrate that the Cys129Ser mutant derivative (i.e. M5) can still form dimers, presumably via the formation of intermolecular disulfide bond(s) between cysteine residues other than Cys^{129}. Possible underlying reasons for such analytical discrepancy were proposed. Most importantly, we identified another mutant derivative, Cys100Ser mutation (i.e. M3), which remains monomeric even after prolonged storage.

4.2 Materials and Methods

4.2.1 Bacterial Strains and Plasmids

Plasmid pGEX2ThCD83ext, containing the cDNA encoding hCD83ext (i.e. AA 20-145 of hCD83ext) fused with the *gst* gene, was used for the expression of GST-hCD83ext regulated by the IPTG-inducible *tac* promoter [8]. Due to the design of a thrombin cleavage site at the junction between GST and hCD83ext, the final hCD83ext product has four extra AAs, i.e. Gly-Ser-Pro-Gly, in the amino terminus and this modification was presumed to have no effect on hCD83ext bioactivity [8]. Plasmid pGEX2ThCD83ext_mut129_CtoS, a derivative from pGEX2ThCD83ext, was used for the expression of GST-hCD83ext_mut129_CtoS (i.e.

GST-M5). Both pGEX2ThCD83ext and pGEX2ThCD83ext_mut129_CtoS were gifts from A. Steinkasserer. Plasmid pGEX2ThCD83ext_mut100_CtoS, another pGEX2ThCD83ext derivative constructed in this study, was used for the expression of GST-hCD83ext_mut100_CtoS (i.e. GST-M3). To construct the plasmid, site-directed mutagenesis was performed using a QuikChange® Multi Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA, USA) with the PCR template of pGEX2ThCD83ext and the primer of 5'-ATCCGAAACACTAC*G*AGC**TC**CAACTCGGGG-3' (nucleotidic mutations are denoted in Italic). The primer contains the Cys100Ser mutation (in bold) and a silent mutation generating an extra *SacI* restriction site (underlined). Purification of plasmid DNA was performed using a spin-column kit (BD Biosciences Clontech, Mississauga, ON, Canada). Plasmid transformation was carried out using a chemical method [85]. *Escherichia coli* BL21(DE3)(pLysS) (Novagen, New Canaan, CT, USA) was used as the expression host.

4.2.2 Production of hCD83ext Variants

The production of hCD83ext variants, i.e. wild-type hCD83ext, M5, and M3, was performed using the fermentation and downstream processing protocols developed previously [7]. The protein products were formulated in 50% glycerol and stored at -20 °C.

4.2.3 Analytical Methods

Protein concentration for hCD83ext variants was quantified by the absorbance at 280 nm using the extinction coefficient of 1.1 OD280-mL-mg-1cm-1. Limulus Amoebocyte Lysate (LAL) assay was carried out to measure the endotoxin level using a commercial kit from Cambrex (East Rutherford, NJ, USA).

SEC was performed using ÄKTApurifer (GE Healthcare, Baie d'Urfé, QC, Canada) equipped with a TSK-GEL® G2000SWXL column (Tosoh Bioscience LLC, Montgomeryville, PA, USA) and a UV detector. The mobile phase was 100 mM phosphate buffer (pH 7.0) with 50 mM NaCl at a flow rate of 0.7 ml/min. Protein samples, either reduced with dithiothreitol (DTT) or non-reduced, were loaded and subsequently detected at 280 nm.

SDS-PAGE and Western blotting were performed using the protocols described previously [7]. Where indicated, protein samples (0.1 mg/mL, pH 7.3) were pretreated with 5 mM N-ethylmaleimide (NEM) at room temperature for 30 min for alkylation of free sulfhydryl groups. Samples for reducing and non-reducing SDS-PAGE were prepared by mixing NEM-pretreated protein samples with 2X sample buffer (65 mM Tris-HCl, 25% v/v glycerol, 2% w/v SDS, 0.01% w/v bromophenol blue) respectively with and without 5% v/v 2-mercaptoethanol. All SDS-PAGE samples were boiled for 4 min after mixing with 2X sample buffer.

The number of free sulfhydryl groups was determined according to the Ellman's assay [35, 86] by mixing 100 µL protein sample and 20 µL 5, 5'-Dithio-bis-(2-nitrobenzoic acid) (DTNB) (4 mg/mL in 0.1 M phosphate buffer, pH 8.0) with 1 mL reaction buffer (0.1 M phosphate buffer, pH 8.5). The reaction was carried out for 15 min at room temperature. The absorbance at 412 nm was measured and the concentration of sulfhydryl groups was quantified using the molar extinction coefficient of 5-thio-2-nitrobenzoic acid (TNB2-) (14,150 OD_{412}-M-1cm-1).

Far-UV circular dichroism (CD) spectra of protein samples were obtained using a spectropolarimeter (Jasco-815; Jasco, Japan). Fluorescence emission spectra (over the range of 300 ~ 420 nm) of protein samples were obtained using a spectrofluorometer (FP-6500; Jasco, Japan) with an excitation wavelength at 295 nm. All the spectroscopic measurements were made at a room temperature of 23 °C.

Biological activity of hCD38ext variants was evaluated using an in vitro assay based on the determination of the protein's ability to inhibit tumor necrosis factor alpha (TNF-α) production by lipopolysaccharide/interferon-gamma (LPS/IFN-γ) stimulated peripheral blood mononuclear cells (PBMCs). Specifically, PBMCs from cynomolgus macaques (Macaca fascicularis) were isolated and pretreated with hCD83ext variants under different concentrations (i.e. 0.5, 5, 25, and 100 μg/mL) for 12 h. The cells were then activated with LPS (1 μg/mL) and IFN-γ (100 U/mL) in the presence of Brefeldin A (4 μg/mL) for 6 h and intracellularly stained for TNF-α using a commercial Fix/Perm kit. Flow cytometry was performed to quantify the amount of TNF-α production using monocytes.

4.3 Results

4.3.1 Structural Analysis Using SDS-PAGE

Due to the presence of free sulfhydryl groups and disulfide bonds in the product samples of hCD83ext variants, we hypothesized that thiol-disulfide exchange could possibly occur during denaturing heat treatment via a mechanism known as disulfide scrambling [51, 56]. As a result, the intermolecular disulfide bonds could be broken by the free sulfhydryl group

on another cysteine residue to result in the formation of an intramolecular disulfide bonds, and thereby release hCD83ext monomers. If this were to occur, the observed ratio of monomer to dimer upon the SDS PAGE analysis would not be representative of the actual ratio of monomer to dimer that was present prior to denaturation. To prevent such disulfide scrambling, NEM alkylation of free sulfhydryl groups was conducted prior to protein denaturation. Using different sample preparation protocols to process four hCD83ext samples (two independent wild-type samples, M5, and M3), the results of SDS-PAGE analysis under both reducing and non-reducing conditions are summarized in Figure 4.1. While there was only a single band for all the samples using reducing SDS-PAGE (Figure 1A), non-reducing SDS-PAGE enabled the analysis of various monomeric and multimeric forms (Figure 4.1B). By comparing the results of non-NEM-treated samples (lanes 2-5, Figure 4.1B) and the corresponding NEM-treated samples (lanes 6-9), it is clear that the dimer/monomer ratio increased significantly as a result of the NEM treatment, particularly for the two wild-type samples (lane 5 vs. 9) and M5 (lane 2 vs. 6). These analytical discrepancies can be associated with disulfide scrambling which broke intermolecular disulfide bonds within dimers and higher multimers to form intramolecular disulfide bonds. In other words, NEM-treated samples analyzed with non-reducing SDS-PAGE reflected the actual species distribution in the protein sample. Interestingly, multimerization was not observed with M3 (lane 7), unlike the wild-type and M5 species. There were at least two monomeric bands observed for M5 (lane 2) and M3 (lane 3) without NEM treatment, but the two bands merged into one when samples were treated with NEM (lanes 6 and 7). This result may be due to disulfide scrambling between free sulfhydryl groups and intramolecular disulfide bonds within

monomers, which can be also prevented by NEM treatment. We also note that the mobility of M3 monomer (lane 7) was approximately the same as that of wild-type monomer (lanes 8 and 9) but different from that of M5 monomer (lane 6). Such mobility differences presumably resulted from different intramolecular disulfide patterns since all reduced monomers had the same mobility (lanes 6-9 in Figure 4.1A).

Interestingly, the above-mentioned effects associated with the NEM treatment were no longer observed when samples were treated with NEM after protein denaturation [i.e. no analytical discrepancies were observed between samples without NEM treatment (lanes 2-5) and samples with NEM treatment after protein denaturation (lanes 10-13)], suggesting that disulfide scrambling occurred primarily during heat treatment and, therefore, the NEM treatment has to be performed prior to protein denaturation.

4.3.2 Structural Analysis Using SEC

To further verify the above results obtained based on non-reducing SDS-PAGE (particularly in monomer/dimer distribution), both reduced and non-reduced protein samples were subjected to SEC analysis under their native condition and the results are summarized in Figure 4.2. The monomeric and dimeric forms are represented by the two peaks eluted at 12.6 and 11.5 min, respectively. Apparently, M3 is the only variant that did not form dimers whereas dimerization was observed for both wild-type and M5. The monomer/dimer distributions obtained by SEC analysis were quantitatively more consistent with those obtained by non-reducing SDS-PAGE using samples with proper NEM treatment (i.e. lanes 6-9 in Figure 4.1B). While SEC was adopted here for qualitative or even quantitative analysis of dimer/monomer distribution, it could not be used for differentiating different monomeric

species, which were observed as the bands with different mobilities using non-reducing SDS-PAGE.

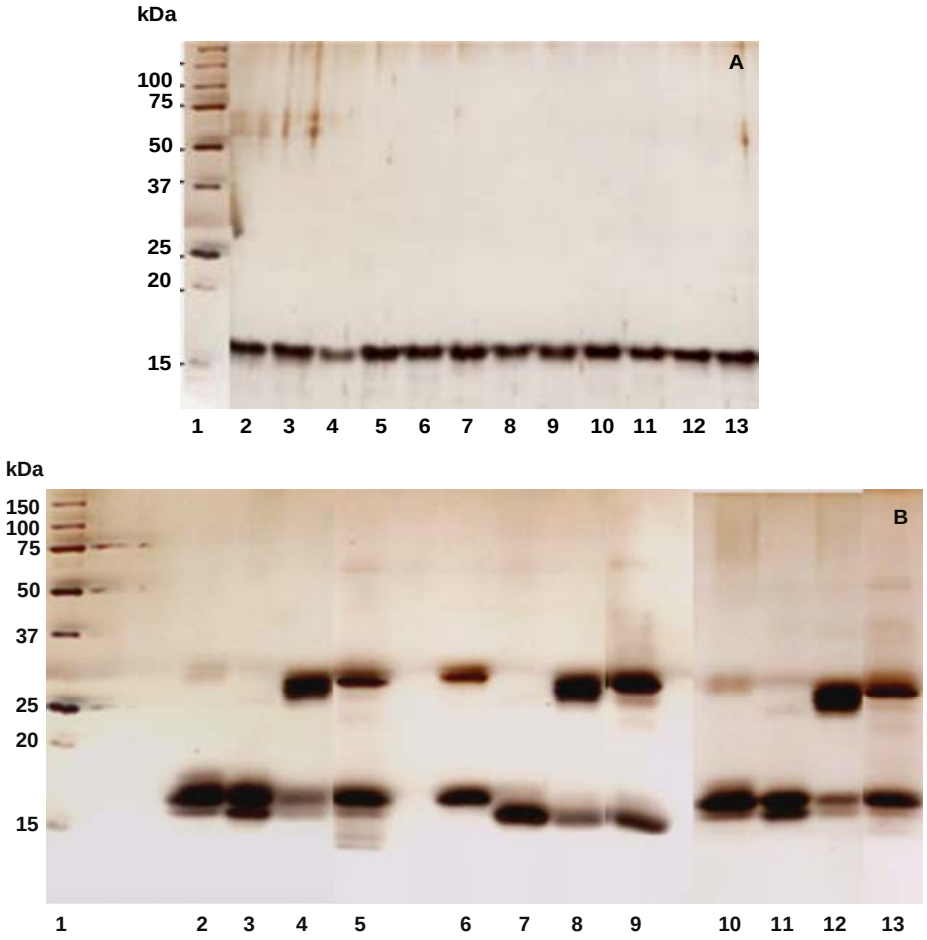


Figure 4.1 Analysis of four hCD83ext products treated with various sample preparation protocols using reduced SDS-PAGE (Panel A) and non-reduced SDS-PAGE (Panel B). For both panels: Lane 1/ molecular weight markers; lanes 2, 6, 10/ M5; lanes 3 7, 11/ M3; lanes 4, 8, 12/ wild-type-1; lane 5, 9, 13/ wild-type-2. Lanes 2~5 are samples with no NEM-treatment; lanes 6~9 are samples with NEM-treatment prior to heat denaturation; lanes 10~13 are samples with NEM-treatment after heat denaturation. The gels were stained by silver nitrate.

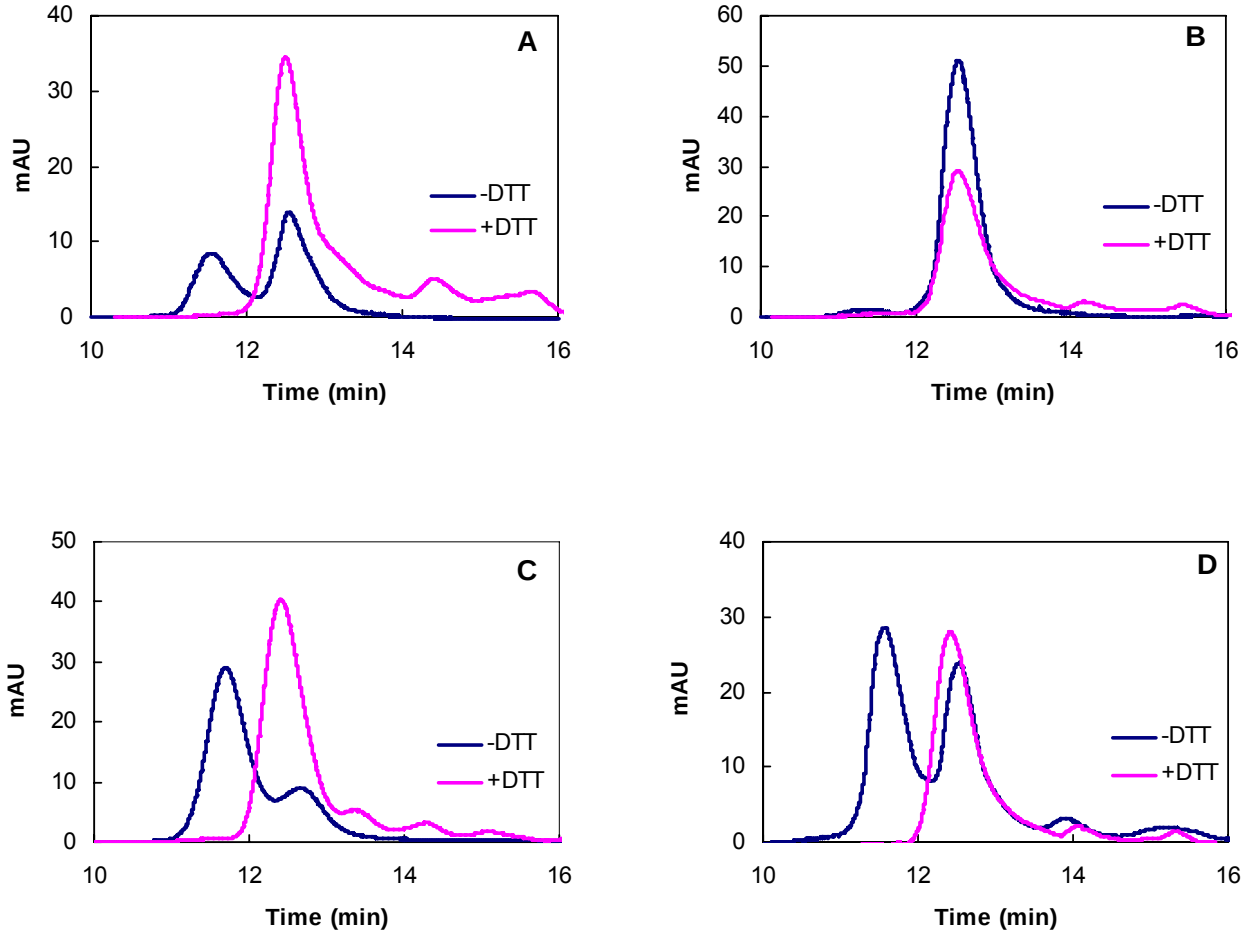


Figure 4.2 Analysis of four hCD83ext products using SEC. The elution profiles for both reduced (with DTT treatment) and non-reduced (without DTT treatment) are shown. Panel A/ M5; panel B/ M3; panel C/ wild-type-1; panel D/ wild-type-2.

4.3.3 Monitoring of Conformational Changes

Molecular behavior of the three hCD83ext variants was compared by monitoring the conformational changes under a frozen storage condition and the results are summarized in Figures 4.3 and 4.4. While the three protein samples shared an identical band pattern using reducing SDS-PAGE (Figure 4.3A~4.3C), they had rather distinctive dimerization profiles over the storage period using non-reducing SDS-PAGE with proper NEM treatment (Figure 4.3D~4.3F). No dimerization was observed for M3, whereas progressive dimerization was observed for both wild-type and M5. The rate of dimerization for M5 was comparable to that of wild-type. Notice that there are at least two distinctive dimer bands and even several higher multimers for M5, implying a more complex scenario for M5 multimerization. Again, the mobility of M5 monomer band was significantly lower than that of wild-type or M3 possibly due to their different disulfide patterns.

Structural monitoring and comparison of hCD83ext variants were also performed using far-UV CD and fluorescence spectra and the results are summarized in Figure 4.4. Interestingly, M3 displayed certain major spectroscopic changes (i.e. CD spectra at 195, 203 and 220 nm and fluorescence spectra at λ_{max} of 343 nm) during the 15-day frozen storage whereas only minor changes were observed for both wild-type and M5. The results suggest that dimerization could not be properly monitored using these two spectroscopic analyses. Notice that the significant increase in fluorescence intensity and the shift of λ_{max} for M3 as compared to wild-type imply that the protein conformational changes associated with the Cys100Ser mutation might contribute to the prevention of dimer formation.

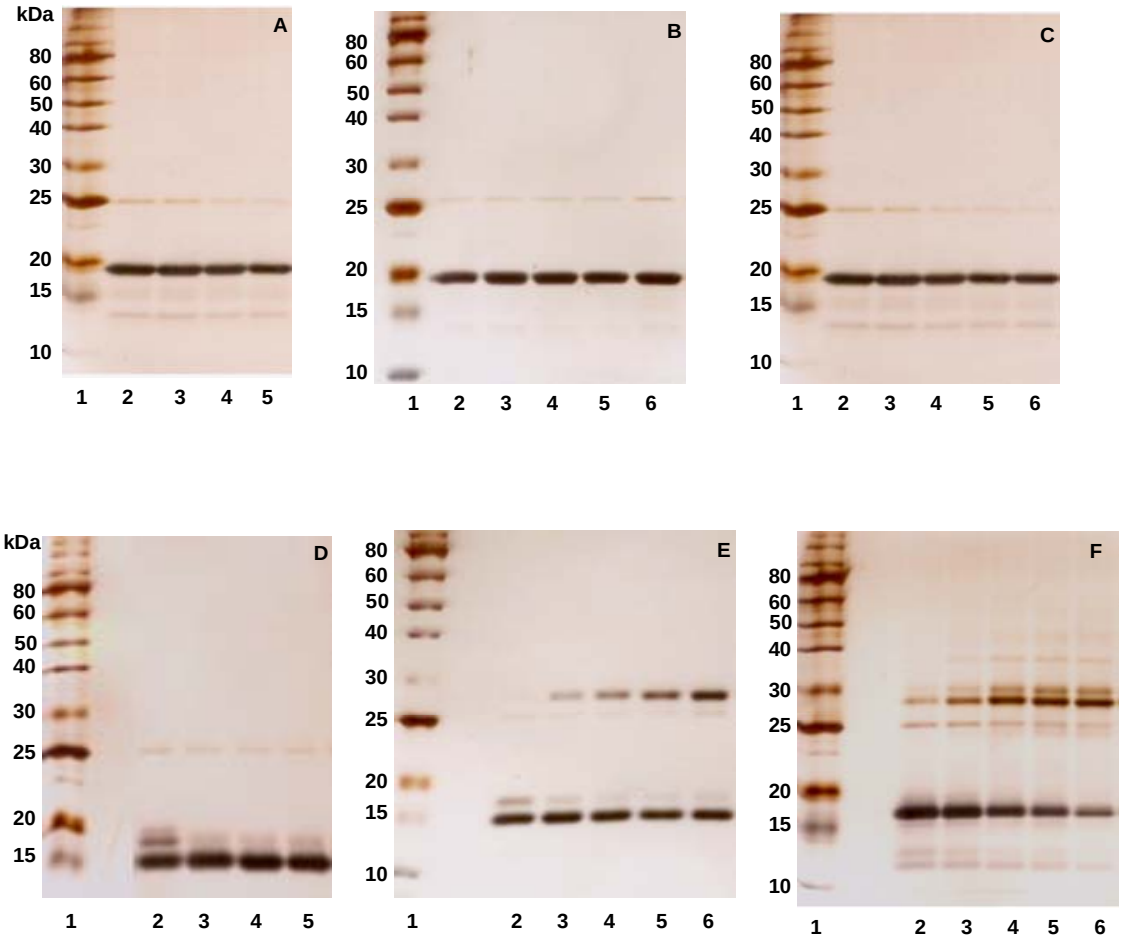


Figure 4.3 Structural monitoring and comparison for M3 (Panels A and D), wild-type (Panels B and E), and M5 (Panels C and F) products stored under -20 °C using reduced SDS-PAGE (Panels A, B, and C) and non-reduced SDS-PAGE (Panels D, E, and F). Panels A and D: Lane 1/ molecular weight markers; lanes 2~5/ samples taken on day 0, 1, 7, and 15, respectively. Panel B and E: Lane 1/ molecular weight markers; lanes 2~6/ samples taken on day 0, 1, 2, 6, and 10, respectively. Panel C and F: Lane 1/molecular weight markers; lanes 2~6/ samples taken on day 0, 1, 5, 10, and 15, respectively. The gels were stained by silver nitrate.

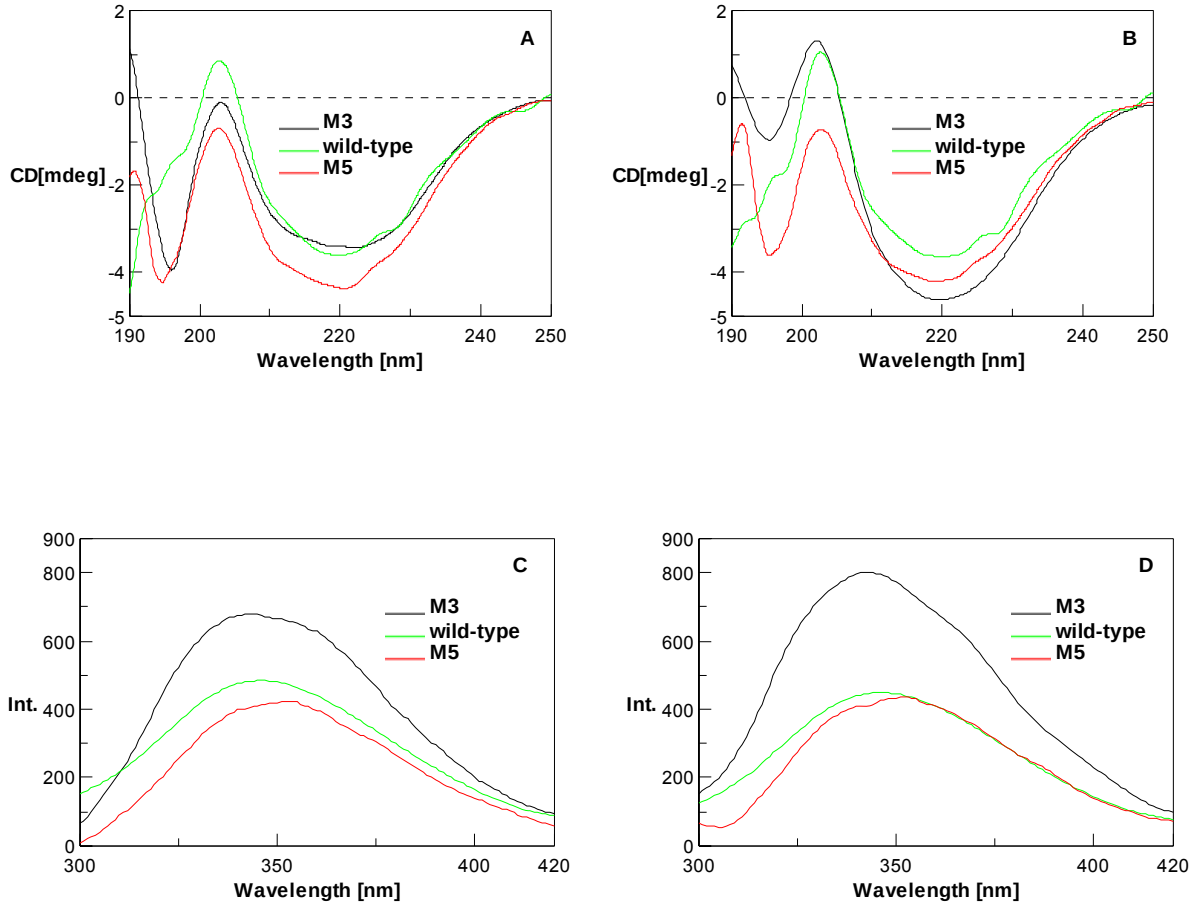


Figure 4.4 Structural comparison for M3, wild-type, and M5 products stored under -20 °C using CD (panels A and B) and fluorescent spectroscopy (panels C and D). Panels A and C: spectra of day-0 samples; panels B and D: spectra of day-15 samples.

4.3.4 Bioactivity Comparison

Based on the above analytical results, M3 appeared to be the most structurally stable among the three hCD83ext variants and was not subjected to multimerization during the frozen storage. The bioactivity of the three hCD83ext variants was compared using an *in vitro* immunological assay and the results are summarized in Figure 4.5.

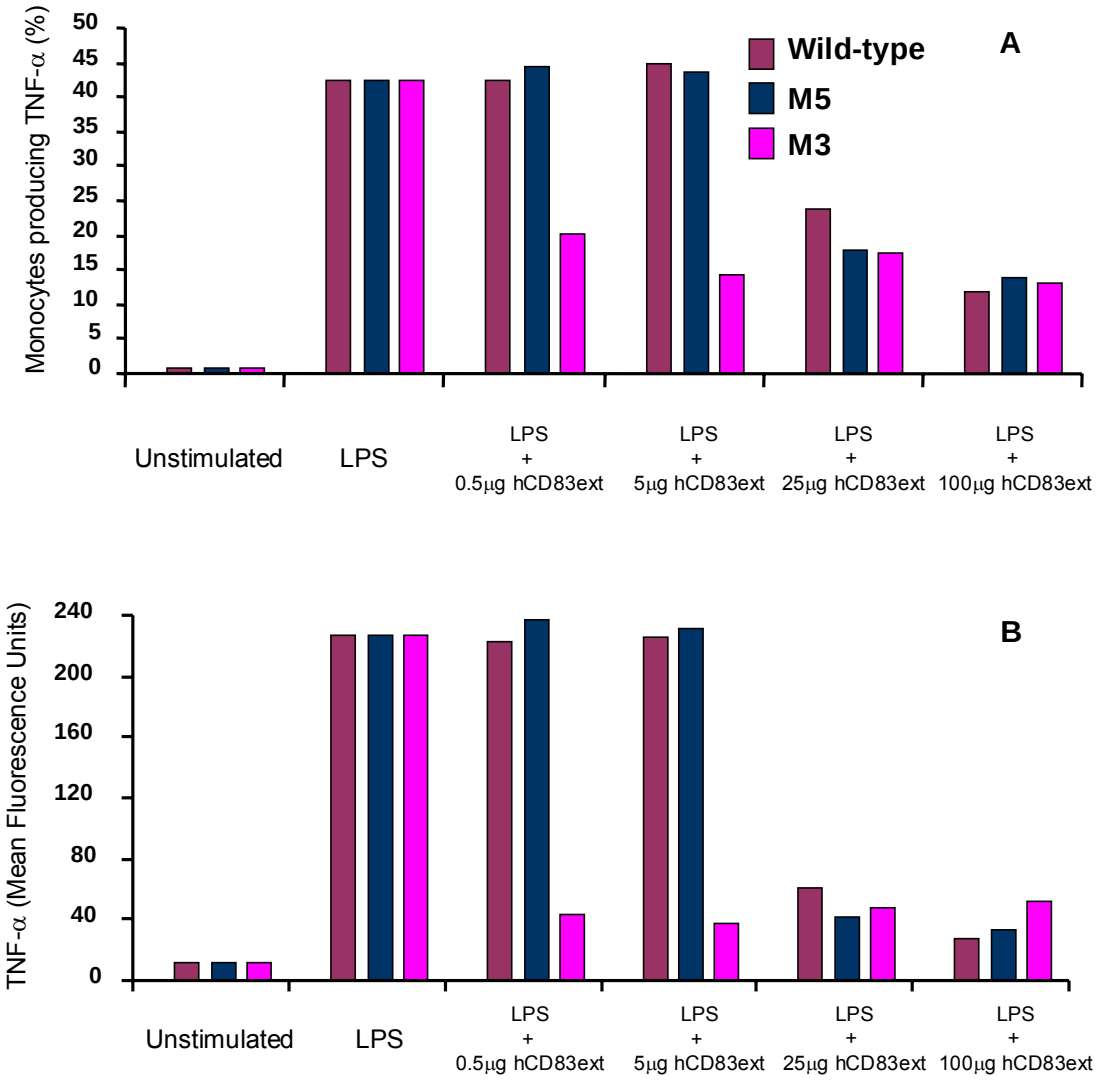


Figure 4.5 The ability of hCD83ext variants (i.e. wild-type, M5 and M3) to inhibit TNF-α production by LPS/IFN-γ stimulated primate PBMCs. The inhibition was quantified by either the percentage of monocytes that produce TNF-α (panel A) or by the mean level of TNF-α produced per cell (panel B) under various hCD83ext dosages.

While both wild-type and M5 could inhibit TNF-α production by LPS/IFN-γ stimulated PBMCs when present at high concentrations (e.g. 25 or 100 μg), M3 displayed the inhibitory effect at all concentrations.

4.4 Discussion

Since therapeutic protein products are often subject to an extensive storage period before clinical administration, their quality can be jeopardized by possible conformational changes and other factors during the storage period. Theoretically, the GST-hCD83ext fusion protein originally harvested from *E. coli* cytoplasm (which is a reductive compartment) should contain no disulfide bonds. It is clear that aberrant disulfide bonds, either intramolecular or intermolecular, can be formed *in vitro* during downstream processing and/or storage. Such unregulated disulfide bond formation could potentially limit consistent bioactivity. Hence, identifying hCD83ext variants that can avoid the formation of aberrant disulfide bonds is warranted.

There are five cysteine residues in wild-type hCD83ext and, theoretically, any cysteine pair spatially close to one another can form a disulfide bond. We hypothesized that the formation probability can be explicitly determined by the transient protein structure under storage conditions and/or intrinsically encoded by protein sequence. Up to now, there is rather limited knowledge and experimental demonstration as to how protein conformation and/or protein sequence can affect the formation of aberrant disulfide bonds. Since disulfide bond formation is considered an uncontrolled variable, all cysteine residues become a

legitimate target for protein manipulations aimed at improving the robustness and consistency of the manufacturing process. The effects of two mutations, i.e. Cys100Ser and Cys129Ser, on hCD83ext conformation, multimerization, and bioactivity were specifically investigated in this study.

Our results suggest that, for proper analysis by non-reducing SDS-PAGE, hCD83ext protein samples should be treated with NEM prior to protein denaturation to avoid disulfide scrambling by which the intermolecular disulfide bonds within hCD83ext multimers (particularly dimers) could be broken with concomitant formation of intramolecular disulfide bonds. The modified protocol led to an analytical discrepancy with the previous report [9]. Notably, there was only a faint dimeric band for M5 without NEM treatment (lane 2, Figure 4.1B), but the signal for the dimeric band became much stronger for M5 with NEM treatment (lane 6, Figure 4.1B). M5 dimerization was observed consistently using both non-reducing SDS-PAGE and SEC in this study. The results suggest that disulfide scrambling occurred and consequently all intermolecular disulfide bonds connecting M5 monomers were broken, leading to the analytical misinterpretation. Structurally, the intermolecular disulfide bond for M5 dimers could be rather unique and close to free sulfhydryl group(s) to cause disulfide scrambling. On the other hand, there could be more than one type of intermolecular disulfide bond linking monomers of wild-type hCD83ext, consistent with what was suggested in our previous report [7]. The intermolecular disulfide bond similar to the one formed within M5 dimers was scrambled easily, whereas others (which might be distant from free sulfhydryl groups) remained intact. The current results suggest that Cys^{129} in wild-type hCD83ext could

be involved in forming intermolecular disulfide bonds that are distant from free sulfhydryl groups.

Consistent with the results for structural monitoring of the final product samples of hCD83ext variants under frozen storage conditions (Figure 4.3), dimerization of M5 was further confirmed whereas no dimerization was observed with M3, even after long-term storage. These data confirm that the molecular behavior for therapeutic proteins can be significantly altered by minor structural changes such as single amino acid mutations. Apparently, Cys^{100} is a key amino acid involved in the formation of intermolecular disulfide bonds for hCD83ext multimerization. Hence, Cys^{100} in M5 could be involved in forming intermolecular disulfide bonds close to free sulfhydryl groups, leading to effective disulfide scrambling. However, cysteine residues other than Cys^{100} could still be involved in the formation of aberrant disulfide bonds. The structural effects of Cys→Ser mutation for the other three cysteine residues (i.e. Cys^{27}, Cys^{35}, Cys^{107}) are currently under investigation. Preliminary results indicate that dimerization still occurs to the hCD83ext variant with any of these single mutations.

Theoretically, to ensure the consistent quality of therapeutic products, protein molecules should be optimally designed and produced in such a way as to incorporate a fixed disulfide pattern and to prevent the formation of aberrant disulfide bonds. The consistent molecular behavior without dimerization over an extensive storage period made M3 a feasible protein therapeutic candidate. As shown by the current results, excellent *in vitro* biological activity was observed for M3. Recently, therapeutic bioactivity for M3 was successfully demonstrated in several *in vivo* studies using rodent models (unpublished data).

Though no dimerization was detected by SDS-PAGE and SEC for M3, the conformational changes associated with spectroscopic analyses and the decreasing number of free sulfhydryl groups (data not shown) suggests the possible formation of intramolecular disulfide bond(s) within M3 monomers. Notice that M3 monomers with different intramolecular disulfide patterns cannot be differentiated using non-reducing SDS-PAGE. Fortunately, it appears that such formation of intramolecular disulfide bonds had a minimal impact on bioactivity.

Chapter 5

Effect of Aberrant Disulfide Bond Formation on Protein Conformation and Molecular Property of hCD83ext

The work in this chapter was published in *Pure Appl. Chem.* 82 (2010): 149-159.

5.1 Introduction

Theoretically, *in vivo* disulfide bond formation should be more consistent than *in vitro* formation since biological cells are able to maintain their intracellular environment rather steadily. In addition, intracellular disulfide bonds can be consistently processed by the Dsb-family members with rigorous formation mechanisms [47]. Though *E. coli* contains an oxidative periplasm in which disulfide bonds can be formed, heterologous protein expression targeting this compartment can be possibly limited by ineffective translocation. It is known that disulfide bond formation hardly occurs in the reductive cytoplasm of wild-type *E. coli*. However, the double mutation on the genes encoding thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) [or glutathione synthetase (*gshB*)] can convert the reductive cytoplasm into an oxidative one, which offers an alternative expression compartment for proteins requiring disulfide bonds [87].

In this study, we used two *E. coli* B strains, i.e. BL21 and Origami B respectively with reductive and oxidative cytoplasm, to produce hCD83ext. With these two product preparations for parallel comparison, we demonstrated the significant effect of aberrant

disulfide formation on protein conformation which subsequently affected the molecular property of this therapeutic product.

5.2 Materials and Methods

5.2.1 Bacterial Strains and Plasmids

The plasmid pGEX2ThCD83ext containing the hCD83ext cDNA fused with the *gst* gene was used as the vector for expression of GST-hCD83ext under the regulation of an IPTG-inducible *tac* promoter [8]. Due to the design of a thrombin cleavage site at the junction between GST and hCD83ext, the final hCD83ext product has four extra AAs, i.e. Gly-Ser-Pro-Gly, in the amino terminus and this modification was presumed to have no effect on the biochemical property of hCD83ext. BL21(DE3) and its *trxB* and *gor* double mutant derivative of Origami B(DE3) (EMD Biosciences/Novagen, La Jolla, CA, USA) were used as the expression hosts.

5.2.2 Cultivation and Downstream Processing

The production and purification of hCD83ext were conducted using the protocols developed previously [7].

5.2.3 Sample Treatment for Analyses

The culture sample was appropriately diluted with saline solution for measuring cell density at OD_{600} with a spectrophotometer (DU®520, Beckman Coulter, Fullerton, CA, USA). For the preparation of cell extract, cells in the amount of 30 OD_{600}-units (defined as 'OD_{600}×mL')

were centrifuged at 6000×*g* and 2 °C for 10 min. The cell pellet was resuspended in 0.75 mL phosphate-buffered saline (pH 7.3), sonicated intermittently (i.e. 0.5 s/0.5 s on/off) for 4 min using an ultrasonic processor (Misonix, Farmingdale, NY, USA) with a microtip, and then centrifuged at 15,000×*g* and 2 °C for 15 min to remove cell debris. The supernatant containing soluble proteins was used for various analyses, including GST assay, SDS-PAGE, and Western blotting. The pellet containing insoluble proteins and cell debris was washed with phosphate buffer, resuspended in TE/SDS buffer (10 mM Tris HCl, pH 8.0, 1 mM EDTA, 1% SDS), and heated to 100 °C for 5 min for dissolution. The protein content of the pellet was analyzed as the insoluble fraction.

5.2.4 Analytical Methods

GST was assayed at ambient temperature using 1-chloro-2,4 dinitrobenzene (CDNB) and glutathione as the substrates. GST (or its fusion) catalyzes the reaction of the two substrates to form a conjugate product that can be monitored at 340 nm. One unit (U_{GST}) is defined as the amount of enzyme that causes an increase in absorbance at 1 OD_{340}/min. The protein concentration of purified hCD83ext was quantified with its absorbance at 280 nm (OD_{280}) and the extinction coefficient of 1.2 OD_{280}-ml/mg/cm. The number of titratable free sulfhydryl group was determined using the Ellman assay [51].

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed in a Mini-PROTEAN®II electrophoresis cell (Bio-Rad) under both reducing and non-reducing conditions using a 12.5% polyacrylamide separating gel stacked by a 4% polyacrylamide stacking gel. Reduced and non-reduced protein samples were prepared using

the 2X sample buffer respectively supplemented with 5% 2-mercaptoethanol and 20 mM *N*-ethylmaleimide (NEM). Electrophoresis was conducted at a constant voltage of 200 V for approximately 45 min. The gel was stained with Coomassie blue or silver nitrate and dried. The dried gel was scanned. Compositions of monomer and dimer were quantified by measuring the pixels of their representing bands in the scanned gel images using UN-SCAN-IT software (Silk Scientific, Orem, UT, USA).

To conduct Western blotting, the proteins on the polyacrylamide gel were electroblotted to a PVDF membrane after SDS-PAGE using a Mini Trans-Blot® Cell (Bio-Rad) according to a standard protocol [80]. The electrophoretic transfer was conducted at a constant voltage of 100 V for 1 h. Protein-antibody hybridization was performed as described by Sambrook et al. [81]. The purified hCD83ext was used as an antigen for raising polyclonal anti-CD83 antibodies in rabbits. The secondary antibody was goat anti-rabbit IgG conjugated with horseradish peroxidase (Sigma-Aldrich). All hCD83-related polypeptides (e.g. GST-hCD83ext and hCD83ext) were detected by a colorimetric method using 3,3'-diaminobenzidine tetrahydrochloride (DAB) as the substrate. The processed membrane was scanned.

Far-UV circular dichroism (CD) spectra of protein samples were obtained using a spectropolarimeter (Jasco-815; Jasco, Japan). Fluorescence emission spectra (over the range of 300 ~ 420 nm) of protein samples were obtained using a spectrofluorometer (FP-6500; Jasco, Japan) with an excitation wavelength at 295 nm. All the spectroscopic measurements were conducted at ambient temperature of approximately 23 °C.

5.3 Results

5.3.1 Bacterial Cultivation for hCD83ext Expression

Using BL21(DE3) and Origami B(DE3) as the host, cultivation for the expression of GST-hCD83ext was conducted in a bench-top bioreactor under identical operating conditions and the comparative results are summarized in Figure 5.1. The growth rate of recombinant Origami B(DE3) was significantly lower than that of recombinant BL21(DE3), resulting in approximately 20% reduction in both final cell density and specific GST activity (Panels A and B of Figures 5.1). The reduction in growth rate for Origami strains was previously reported and attributed to possible malfunction of many cytoplasmic proteins under an oxidative condition [73]. However, the GST activity might not reflect the actual expression level of GST-hCD83ext since SDS-PAGE results indicated otherwise (Figure 5.1C, Lanes 2 vs. 4 and Lanes 6 vs. 8) and this discrepancy appeared to be reproducible, suggesting that the oxidative environment in Origami B(DE3) could result in a certain level of reduction in GST activity.

Notice that, upon SDS-PAGE analysis, GST-hCD83ext expressed in Origami B(DE3) had a higher electrophoretic mobility under a non-reducing condition compared to that of its corresponding monomeric form under a reducing condition (Figure 5.1C, Lanes 4 vs. 8), indicating that certain intramolecular disulfide bond(s) formed in the oxidative cytoplasm. Unsurprisingly, such observation was not made when BL21(DE3) was used as the expression host since disulfide bonds cannot be formed in the reductive cytoplasm of wild-type *E. coli*. In addition, GST-hCD83ext expressed in both strains was not completely soluble though the

expression in Origami B(DE3) was slightly more soluble than that in BL21(DE3) (Figure 5.1C). Finally, it appeared that multimeric forms of GST-hCD83ext were produced via

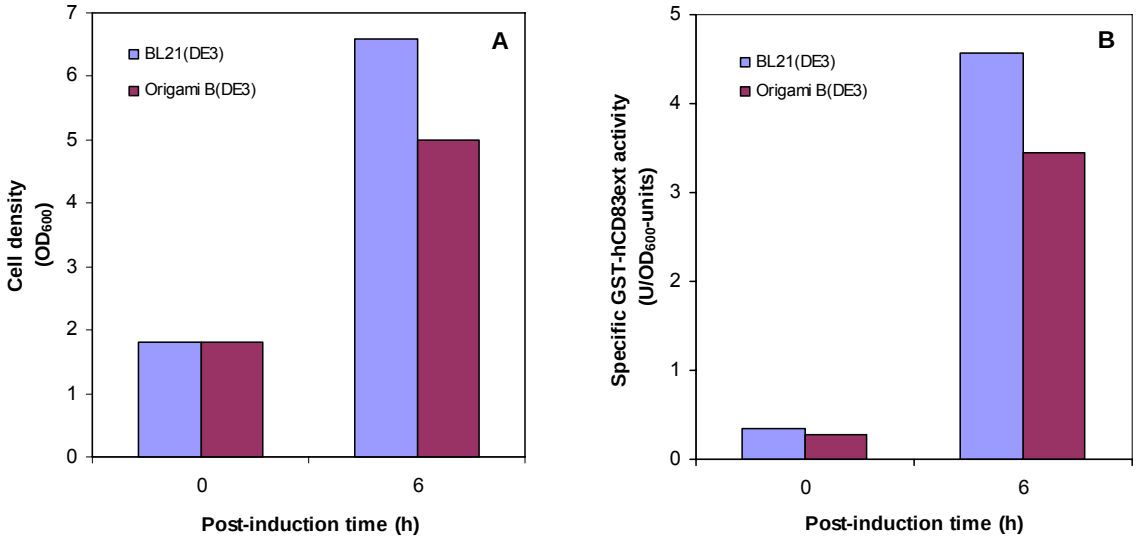


Figure 5.1 Comparison of bioreactor cultivation performance for the production of GST-hCD83ext in two host strains: Cell density (Panel A), specific GST activity (Panel B), and intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panels C) and Western blotting (Panels D) for the culture samples taken prior to induction and at 6 h post-induction time are shown. Panel C: Lanes 1, 2, 5, 6, 9 and 10 represent samples produced by BL21(DE3), whereas lanes 3, 4, 7, 8, 11, and 12 represent samples produced by Origami B(DE3). Arrows represent the protein bands corresponding to disulfided and non-disulfided hCD83ext monomers. Panel D: All lanes represent culture samples taken at 6 h post-induction time. Arrows represent the protein bands corresponding to various hCD83ext structures.

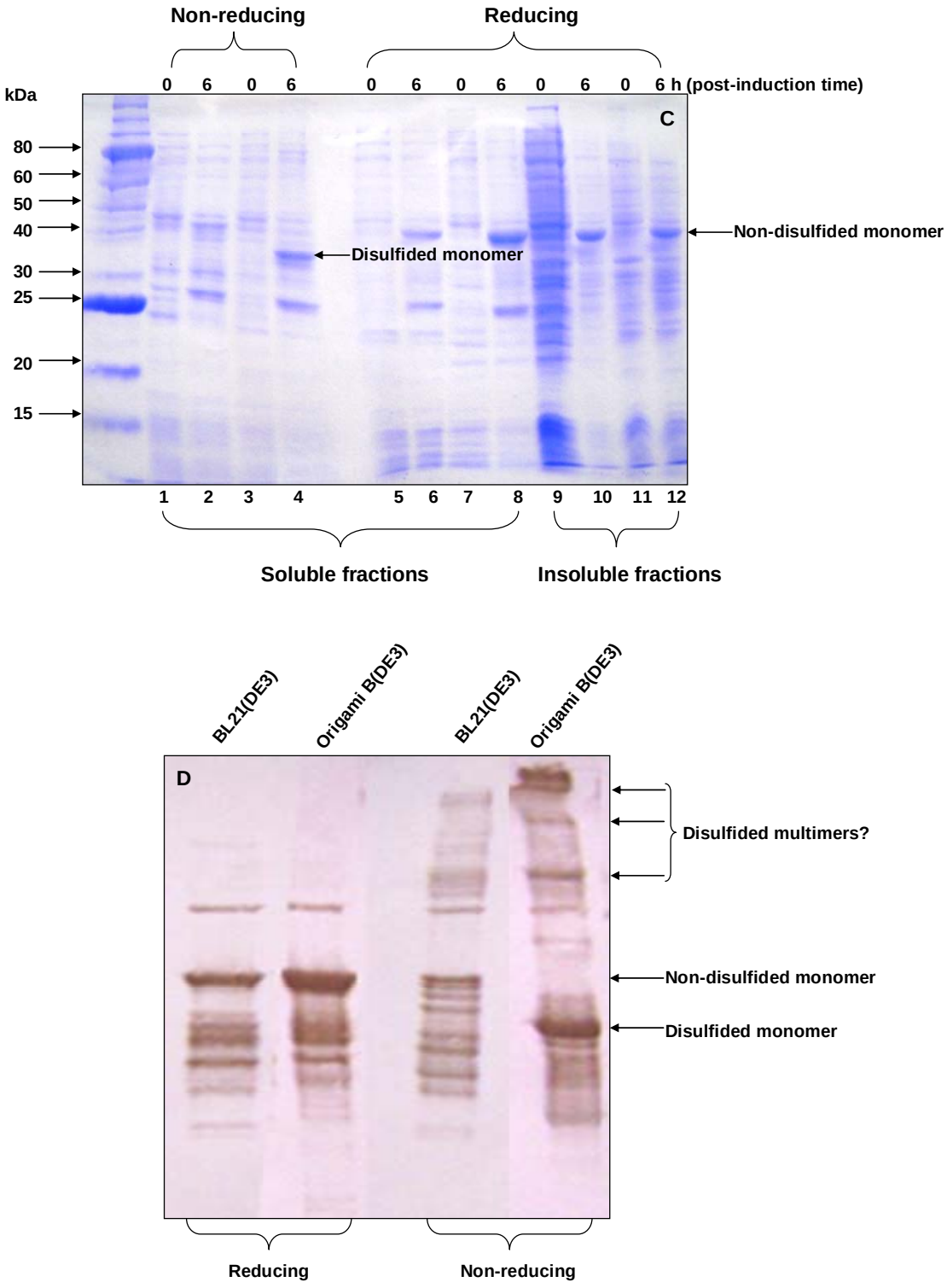


Figure 5.1 (Continued)

intermolecular disulfide bond(s) in Origami B(DE3) and these bands were verified to be associated with hCD83ext using Western blotting (Lane 4, Figure 5.1D).

5.3.2 Comparison of hCD83ext Conformation

The expressed GST-hCD83ext from both production strains was trapped with GST-affinity chromatography and was subject to on-column cleavage with thrombin to release the hCD83ext moiety, which was then polished with anion exchange chromatography, concentrated, and formulated in 50% glycerol for subsequent structural characterization. It appeared that the yield for the polishing step was moderately increased by *in vivo* disulfide bond formation (data not shown). Far-UV CD and fluorescent spectra of hCD83ext produced by the two strains were compared and the results are summarized in Figure 5.2. Both hCD83ext samples displayed a CD spectrum with a negative band at 220 nm and a positive band near 203 nm (Figure 5.2A), which is a typical pattern for β-sheets mixed with type II β-turns [82]. The spectrum of hCD83ext produced by BL21(DE3) had a lower band at 203 nm and a deeper trough at 220 nm, compared to that of hCD83ext produced by Origami B(DE3). Based on the spectroscopic simulation using the K2D method [88], the two hCD83ext preparations had slightly different compositions in the secondary structure, i.e. the ratio of α-helix: β-sheet: unordered structure at 9%: 43%: 48% for hCD83ext produced by BL21(DE3) and 4%: 48%: 48% for hCD83ext produced by Origami B(DE3). Since hCD83ext has two tryptophan residues (i.e. AAs 39 and 49), the emission spectrum reflecting the local environment near the tryptophan residues can shed light on the tertiary structure. The maximum emission intensity (at 345 nm) of hCD83ext produced by BL21(DE3) was approximately 3-fold that (at 350 nm) of hCD83ext produced by Origami B(DE3) with a

slight red-shift (Figure 5.2B), implying significantly different 3-D conformation for the two protein preparations.

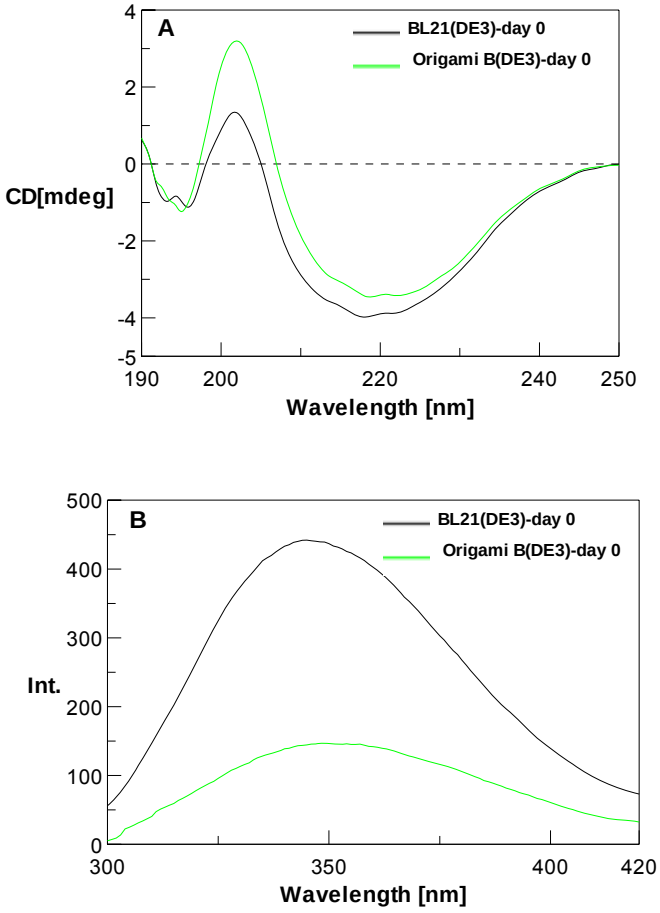


Figure 5.2 Analysis of protein structure for fresh hCD83ext produced by two host strains with CD spectroscopy (Panel A) and fluorescence spectroscopy (Panel B).

Though significant conformational differences between the two hCD83ext variants, the SDS-PAGE analysis revealed a rather similar pattern (Lane 2 in Panels A~D of Figure 5.3) except the observation that hCD83ext produced by BL21(DE3) had two monomeric forms (Lane 2, Figure 5.3C) whereas hCD83ext produced by Origami B(DE3) only had a single monomeric form (Lane 2, Figure 5.3D). On the other hand, the number of titratable free sulfhydryl group was 0.96 and 0.63 under a native condition for hCD83ext produced by BL21(DE3) and Origami B(DE3), respectively (Panels A and B of Figure 5.4).

5.3.3 Comparison of hCD83ext Molecular Property

The molecular properties of the two hCD83ext variants were compared by monitoring the conformational changes under storage conditions (either -20 °C or room temperature) and the results are summarized in Figures 5.3 and 5.4. While the two protein samples had an identical band pattern upon reducing SDS-PAGE (Panels A and B of Figure 5.3), two rather different dimerization profiles over the incubation period were observed upon non-reducing SDS-PAGE, particularly for the samples stored at -20 °C (Panels C and D of Figure 5.3). Compared to hCD83ext produced by BL21(DE3), hCD83ext produced by Origami B(DE3) appeared to have a higher tendency for initial dimerization. The monomeric doublet corresponding to hCD83ext produced by BL21(DE3) disappeared over the storage period. The rate of dimerization for hCD83ext appeared to increase noticeably when incubated at room temperature, particularly for hCD83ext produced by BL21(DE3). Notice that, while the hCD83ext monomers corresponding to the two protein samples were rather stable at room temperature, the dimers started to degrade after 2 days of incubation (Lanes 8, 9 and 10 in Panels C and D of Figure 5.3).

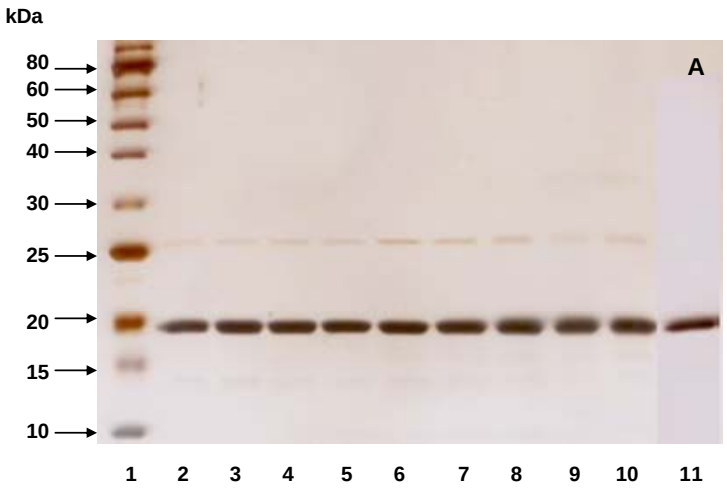


Figure 5.3 Monitoring and comparison of molecular property for hCD83ext produced by two host strains with reducing SDS-PAGE (Panels A and B), non-reducing SDS-PAGE (Panels C and D), CD spectroscopy (Panels E and F), and fluorescence spectroscopy (Panels G and H): Samples up to 30-day storage at -20 °C and 10-day storage at room temperature were analyzed. Panels A, C, E, and G represent hCD83ext samples produced by BL21(DE3), whereas Panels B, D, F, and H represent hCD83ext samples produced by Origami B(DE3). Panels A~D: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~ 6 and 11/ -20 °C samples taken on day 1, day 2, day 6, day 10, and day 30, respectively; lanes 7 ~ 10/ room temperature samples taken on day 1, day 2, day 6, and day 10, respectively. The insets at upper right in Panel C and D represent the time profiles for dimerization of hCD83ext stored at -20 °C. Such quantification was not conducted for samples stored at room temperature since hCD83ext dimer appeared to be unstable.

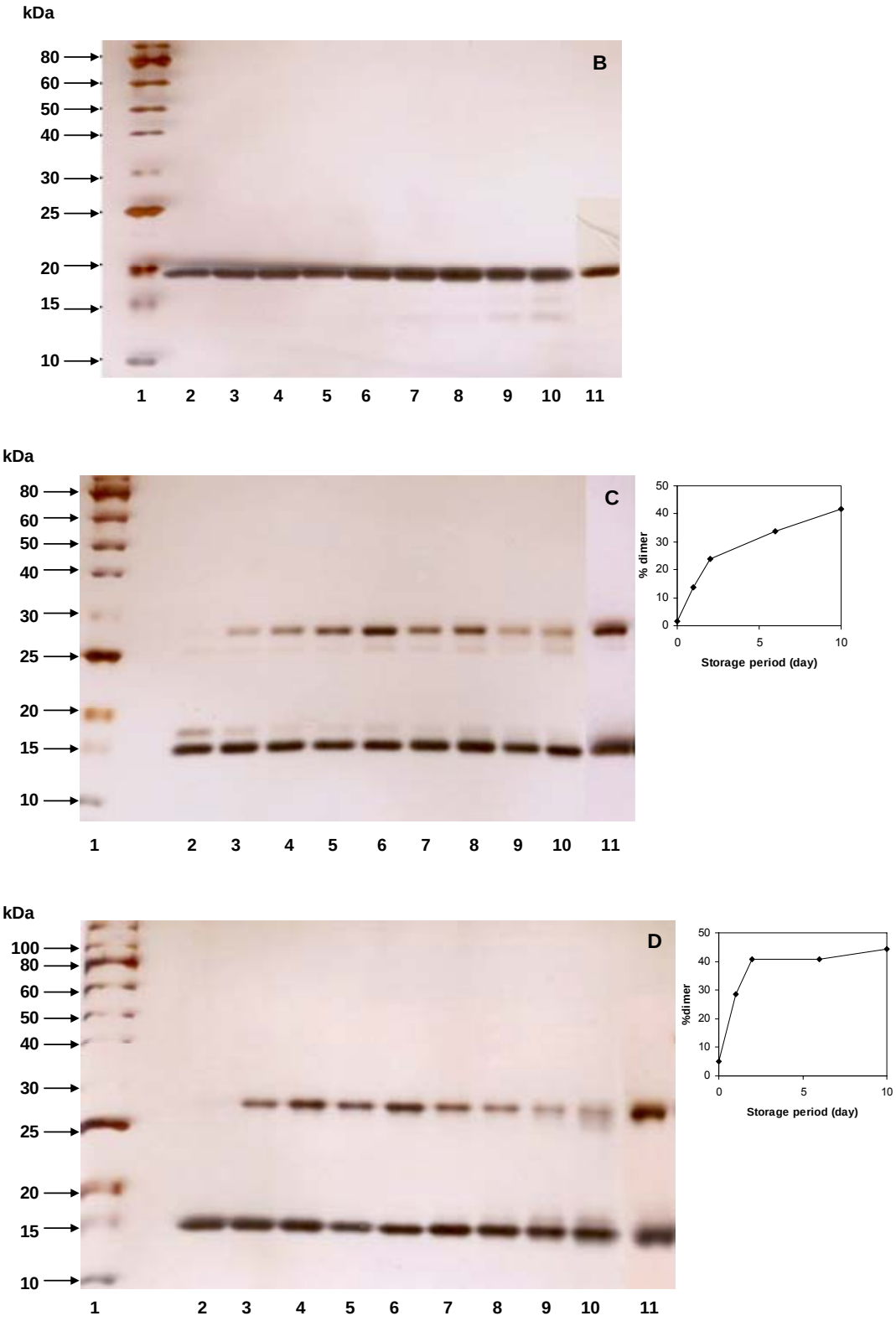


Figure 5.3 (Continued)

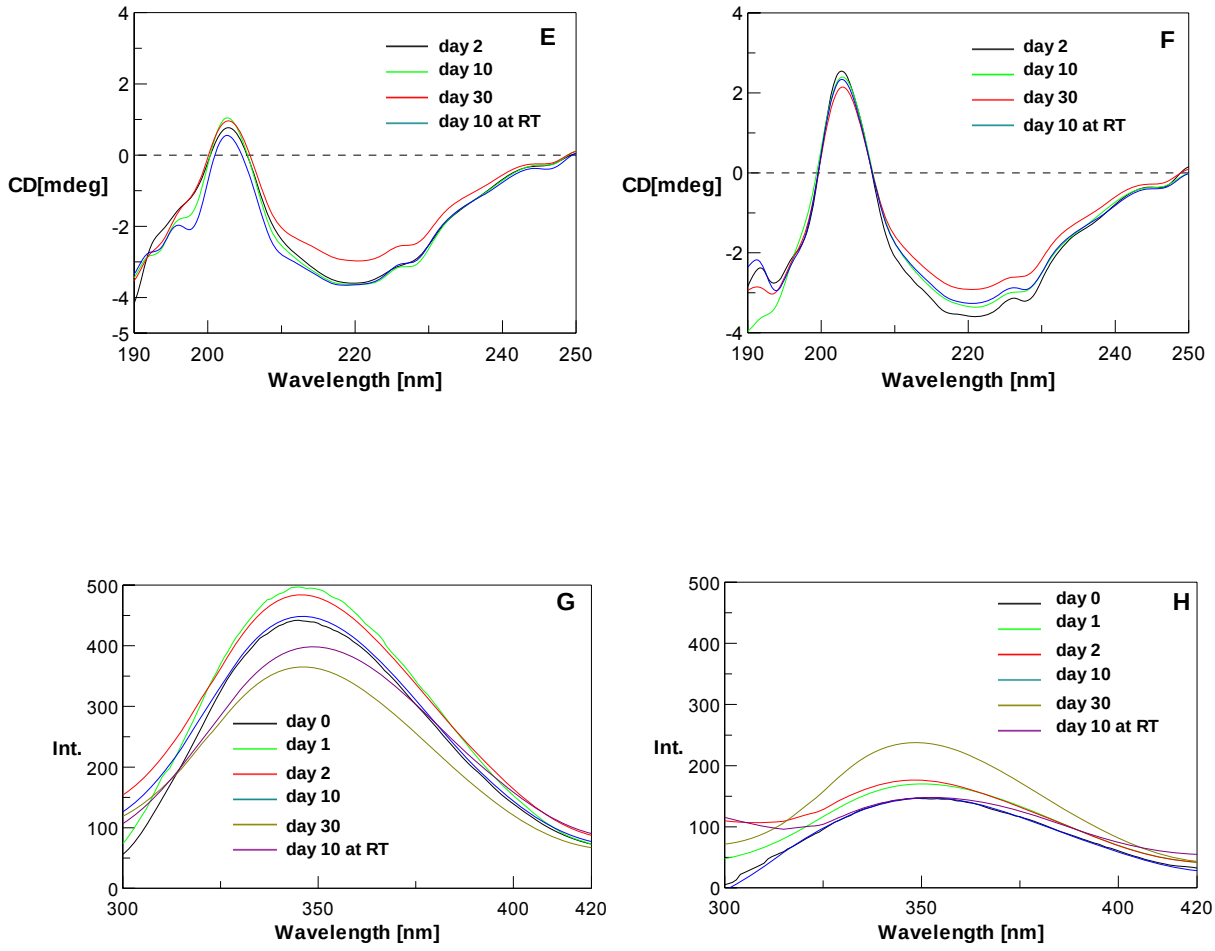


Figure 5.3 (Continued)

The far-UV CD spectra with respect to storage time suggest that both hCD83ext variants did not experience significant changes in the secondary structure (Panels E and F of Figure 5.3) during the storage, even in the presence of dimerization and dimer degradation. Results of fluorescence spectra revealed a slight red-shift upon 10 days of incubation at room temperature, presumably associated with dimer degradation (Panels G and H of Figure 5.3).

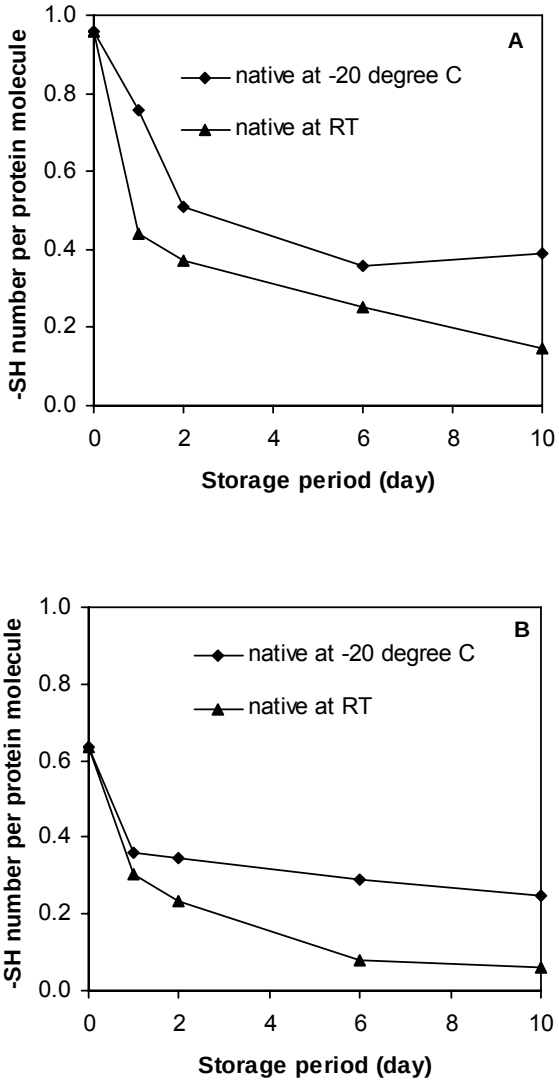


Figure 5.4 Time profiles for the number of titratable free sulfhydryl group in hCD83ext produced by BL21(DE3) (Panel A) and Origami B(DE3) (Panel B): Profiles for samples stored at -20 °C and room temperature are shown.

The maximum intensity of emission spectrum for hCD83ext produced by BL21(DE3) monotonically decreased with incubation time; whereas that for hCD83ext produced by Origami B(DE3) appeared to fluctuate with a significant increase for the 30-day sample stored at -20 °C. The decreasing trend in the number of titratable free sulfhydryl group suggests that disulfide bonds were continuously being formed (Figure 5.4). Overall, the above results suggest the conformational changes in the tertiary structure over the storage period.

5.4 Discussion

There are five cysteine residues in hCD83ext and it was reported that hCD83ext can form dimers *in vitro* via a major intermolecular disulfide bond at Cys^{129} [9]. However, *in vivo* disulfide bond formation within GST-hCD83ext at the protein expression level was not previously explored. The present results show that GST-hCD83ext fusion contained no disulfide bonds when expressed in the reductive cytoplasm of BL21(DE3), but contained intramolecular or even intermolecular disulfide bond(s) when expressed in the oxidative cytoplasm of Origami B(DE3). The Cys^{27} residue in the hCD83ext moiety could play a key role for such *in vivo* disulfide bond formation since no monomer with a higher electrophoretic mobility was observed for the GST fused with the Cys^{27} mutant derivative of hCD83ext expressed in Origami B(DE3) [89]. Though the precise location(s) of such *in vivo* disulfide bond formation was not characterized, the protein conformation and molecular

properties of the final hCD83ext product were significantly affected by this post-translational processing at the protein expression level.

With respect to the protein conformation, it appeared that the two hCD83ext variants had rather distinct secondary structures (based on CD spectra) and tertiary structures (based on fluorescent spectra) though SDS-PAGE revealed a very similar pattern. Since the expression host was the only difference in Materials & Methods upon producing these two variants, it is plausible to assume that the structural differences could result from different disulfide patterns, namely hCD83ext produced by Origami B(DE3) might contain an extra unconventional intramolecular disulfide bond(s) bridged *in vivo*. Analysis of hCD83ext sequence reveals an immunoglobulin (Ig) domain with a conserved disulfide bond between Cys^{35} and Cys^{107}. In addition, our prediction of 3-D structure (data not shown) also suggests Cys^{27} and Cys^{100} are spatially close enough to form a possible disulfide bond. However, the argument cannot be verified by reducing SDS-PAGE since the major monomeric bands for both variants had approximately the same electrophoretic mobility (Panels C and D of Figure 5.3). Interestingly, an extra monomeric band with a lower electrophoretic mobility was observed in day-0 sample for hCD83ext produced by BL21(DE3) and we speculated that it could represent un-disulfided hCD83ext. If this argument were true, the major monomeric band for hCD83ext produced by BL21(DE3) would also contain a disulfide bond(s), which were presumably formed *in vitro* and rather effectively during downstream processing normally conducted in an oxidative environment. To address these technical issues, characterization of disulfide pattern for hCD83ext variants using mass spectroscopy is currently under extensive exploration in our lab.

The presence of the extra intramolecular disulfide bond(s) for hCD83ext produced by Origami B(DE3) mediated a slight shift in the composition of the secondary structure of hCD83ext from α-helix to β-sheet with a minimum influence on the composition of the unordered structure. While the local environment surrounding the two tryptophan residues in the two hCD83ext product variants appeared to be similar in terms of tryptophan exposure (λ_{max} > 340 nm), the emission fluorescence was slightly red-shifted ($\Delta\lambda_{max}$ = 5 nm; implying a more polar local environment surrounding the two tryptophan residues) and significantly quenched for hCD83ext produced by Origami B(DE3). Emission quenching via disulfide bonds which acted as electron scavengers absorbing electrons from excited indole ring of tryptophan was previously reported [90]. Alternatively, the hCD83ext variant with an intramolecular disulfide bond(s) could possibly develop a conformation that facilitates the transfer of emission energy, resulting in the above quenching. Apparently, these conformational changes potentially associated with intramolecular disulfide bond formation enhanced the subsequent formation of intermolecular disulfide bond for dimerization, possibly by shifting the local environment and orientation of cysteine residues to facilitate their spatial contact. The results indirectly suggest that transient protein structure or even protein sequence likely correlate with the probability of having cysteine residues to appear within a contact range to form aberrant disulfide bonds.

It was previously reported that hCD83ext should be stored under frozen conditions due to its intrinsic instability [7]. Though protein degradation was prevented at -20 °C, hCD83ext tended to form dimers or even multimers via intermolecular disulfide bonds [7]. The present results show that the rate of dimerization increased with temperature. In addition,

both monomers and dimers were rather stable at -20 °C, whereas dimers were less stable than monomers at room temperature. Apparently, prolonged incubation of hCD83ext at a relatively high temperature resulted in faster dimerization and subsequent degradation, which could be a major cause for the intrinsic instability. Nevertheless, the stability issue appeared to be unaffected by the presence of unconventional disulfide bonds.

Chapter 6

Molecular Manipulation Associated with Disulfide Bond Formation to Enhance the Stability of Recombinant Therapeutic Protein

The work in this chapter was accepted by *Protein Expr. Purif.*, in press

Authors: **Lin Zhang**, Murray Moo-Young, and C. Perry Chou

6.1 Introduction

A technical issue closely related to protein quality is disulfide bond formation. Because a disulfide bond is formed by crosslinking two cysteine residues transiently adjacent to each other, it is plausible to assume that the formation probability can be explicitly determined by protein three-dimensional (3D) structure or even intrinsically encoded by protein sequence. To ensure a consistent performance, the therapeutic protein molecule should be designed and produced with a constant disulfide pattern. Manipulating cysteine residues within a given protein sequence becomes a rational strategy to either introduce desired disulfide bonds or prevent aberrant disulfide bonds which are often undesired. Because disulfide bond formation can potentially affect protein stability [45, 46, 91], the above cysteine manipulation strategy can have implications on protein stabilization [43, 92-98].

The therapeutic product hCD83ext produced using the developed bioprocess was subject to proteolysis, particularly during a prolonged storage (unpublished data). Structural monitoring of the therapeutic product suggested these cysteine residues were involved in forming aberrant (intramolecular and/or intermolecular) disulfide bonds under oxidative

storage conditions [7]. Formation of aberrant disulfide bonds was thus assumed to be associated with protein instability. To address this issue, we explored the derivation of mutant variants with enhanced stability by manipulating cysteine residues.

Through conserved domain analysis in this study, Cys^{27} was identified as a key residue potentially involved in aberrant disulfide bond formation and, therefore, was designated as the target for mutagenesis. Two amino acids were selected to replace Cys^{27}, i.e. glutamic acid which results in the highest statistic significance of homology to conserved domain families and serine which is a common substitution for cysteine. Two *E. coli* B strains, i.e. BL21 and Origami B with a reductive and oxidative cytoplasm, respectively, were used to express the mutant proteins since significantly different molecular properties were observed for the wild-type hCD83ext expressed in these two strains [99]. The mutant proteins were produced using the previously developed bioprocess [7] for structural monitoring and characterization.

6.2 Materials and Methods

6.2.1 Conserved Domain Search and Structural Prediction of hCD83ext

Conserved domain search was performed against the Conserved Domain Database (CDD; http://www.ncbi.nlm.nih.gov/cdd) using the default parameters. The 3D structure of hCD83ext was predicted using RAPTOR (Bioinformatics Solutions, Waterloo, Ontario, Canada), a bioinformatics software for predicting protein structures through protein threading [100]. The A chain of rat myelin oligodendrocyte glycoprotein (PDB code: 1pkoA) was identified as the best template for modeling.

6.2.2 Bacterial Strains and Construction of hCD83ext Mutants

Plasmid pGEX2ThCD83ext contains the wild-type hCD83ext cDNA fused with the *gst* gene [8]. The two plasmid derivatives constructed in this study, i.e. pGEX2ThCD83ext_mut_C27E and pGEX2ThCD83ext_mut_C27S, were used for the production of hCD83ext_mut_C27E and hCD83ext_mut_C27S. To construct these two plasmids, site-directed mutagenesis was performed using the QuikChange® Multi Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) with pGEX2ThCD83ext as the template. Two primers of 5'-GAGGTGAAGGTGGC***GAGC***TCCGAAGATGTG-3' and 5'-CCGGAGGTGAAGGTGGCT***GAG***TCCGAAGATGTGGAC-3' were used to replace Cysteine27 with serine and glutamic acid, respectively (Note: Mutations are in italics and the codons corresponding to serine and glutamic acid are in bold). The primer containing the Cys27Ser mutation has a silent mutation generating an extra *Sac*I restriction site and the primer containing the Cys27E mutation has a silent mutation generating an extra *Ple*I restriction site (underlined). Purification of plasmid DNA was performed using a spin-column kit (QIAGEN, Mississauga, Ontario, Canada). Mutations were verified by DNA sequencing. BL21(DE3)pLysS and its *trxB* and *gor* double mutant derivative Origami B(DE3)pLysS (EMD Biosciences/Novagen, La Jolla, CA, USA) were used as the expression hosts. Plasmid transformation was carried out by electroporation.

6.2.3 Protein Production and Purification

The production and purification of hCD83ext variants, i.e. wild-type hCD83ext, hCD83ext_mut_C27S and hCD83ext_mut_C27E, in the two host strains was performed

using the bioprocess developed previously [7]. The protein products were formulated in 50% glycerol.

6.2.4 Analytical Methods

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed in a Mini-PROTEAN®II electrophoresis cell (Bio-Rad, Hercules, CA, USA) under both reducing and non-reducing conditions using a 12.5% polyacrylamide separating gel stacked by a 4% polyacrylamide stacking gel. Reduced and non-reduced protein samples were prepared using the 2X sample buffer [65 mM Tris-HCl, pH6.8, 25% glycerol, 2% (w/v) SDS, 0.01% (w/v) bromophenol blue] respectively supplemented with 5% 2-mercaptoethanol and 10 mM *N*-ethylmaleimide (NEM). Electrophoresis was conducted at a constant voltage of 200 V for approximately 50 min. The gel was stained with silver nitrate, dried, and then scanned for documentation.

Protein samples at 0.1 mg/mL were prepared in 20 mM Tris-HCl buffer (pH 7.0) for spectroscopic analysis. Circular dichroism (CD) spectra were obtained using a spectropolarizer (J-810; Jasco, Tokyo, Japan). A cuvette with a path length of 0.1 cm was used for measurement in the far ultraviolet (UV) region from 250 to 190 nm. Fluorescence spectra were obtained using a spectrofluorometer (FP-6500; Jasco). A cuvette with a path length of 1 cm was used for fluorescent measurement in the range of 300~420 nm at an excitation wavelength of 295 nm. All spectra were recorded at ambient temperature. The average of three scans were taken and corrected by subtracting the buffer spectrum obtained under the same conditions.

To conduct the analysis using reversed-phase high performance liquid chromatography (rp-HPLC), 20 μL of the purified protein product was acidified by adding 1 μL 4% TFA and then injected to ÄKTApurifier™ 10 (GE Healthcare, Baie d'Urfé, Québec, Canada) equipped with a C8 column (ACE 222-2546, 5 μm particle size, 300 Å pore, 4.6 mm × 250 mm; Mac-Mod, Chadds Ford, PA, USA) for separation and a zinc lamp for UV detection at 214 nm. This separation was performed through a linear gradient elution of 30 to 40% solvent B over the course of 25min at a flow rate of 0.7 mL/min [Solvent A: 0.1% trifluoroacetic acid (TFA) in water; solvent B: 0.1% TFA in acetonitrile (ACN)].

6.3 Results

6.3.1 Identification of the Key Cysteine Residue for Manipulation

The results of the conserved domain search, including the length of the conserved domains, the significance of similarity (*E*-value), and the sequence alignment, are documented in Figure 6.1A. The local sequence, i.e. hCD83ext, was found to have a high structural similarity to the immunoglobulin (Ig) domain superfamily, especially the Ig variable-set (V-set) domain. Note that there is a 15-AA difference between the cloned hCD83ext sequence number in Figure 6.1A and the actual hCD83ext sequence number (adopted elsewhere in this paper) because of the lack of the 19-AA signal peptide and the additional 4 AAs resulting from the design of a thrombin cleavage site joining GST and hCD83ext. Two cysteine residues in hCD83ext, i.e. Cys^{35} and Cys^{107} (boxed and in red), are conserved, whereas $Cys^{27}$and Cys^{100} are non-conserved (in blue). Cys^{129} is not located in the conserved domain.

The predicted 3D protein structure of hCD83ext with labeled cysteine residues was shown in Figure 6.1B. The two cysteine pairs, i.e. Cys^{35}/Cys^{107} and Cys^{27}/Cys^{100}, were spatially close within a β-sheet, whereas Cys^{129} was relatively distant from the other cysteines. An intramolecular disulfide bond between Cys^{35} and Cys^{107} was previously predicted [9]. Recently, we observed that several free cysteines in hCD83ext can be involved in thiol-disulfide exchange known as disulfide scrambling [101], implying that the presence of free cysteines can possibly jeopardize the structural consistency. In that regard, the presence of Cys^{27} might be unfavorable and, therefore, was selected as the target for mutagenesis. The substituted AA was selected on the basis of the sequence homology to the Ig domain family and V-set domain. The calculated *E*-values for all Cys^{27} substitutions are summarized in Table 6.1. Noticeably, C27E had the highest homology to both conserved domains. In addition, C27S was selected since serine is a common AA substitution for cysteine.

6.3.2 Bacterial Cultivation for Mutant hCD83ext Expression

Using BL21(DE3)pLysS and Origami B(DE3)pLysS as the host, cultivation for the expression of GST-hCD83ext_mut_C27E and GST-hCD83ext_mut_C27S was conducted in a bench-top bioreactor and the results are summarized in Figures 6.2 and 6.3. While there is no significant host effect on the expression level, the growth rate and biomass yield of recombinant Origami B(DE3) was approximately 30% lower than that of recombinant BL21(DE3). In addition, approximately 50% of the GST-hCD83ext fusion was expressed insolubly in both hosts. Interestingly, the electrophoretic mobility of the GST-hCD83ext fusion expressed in BL21(DE3) was approximately the same as that expressed in Origami B(DE3) (Lanes 11 vs. 13 in Panels C and D of Figures 6.2 and 6.3), suggesting that the *in vi-*

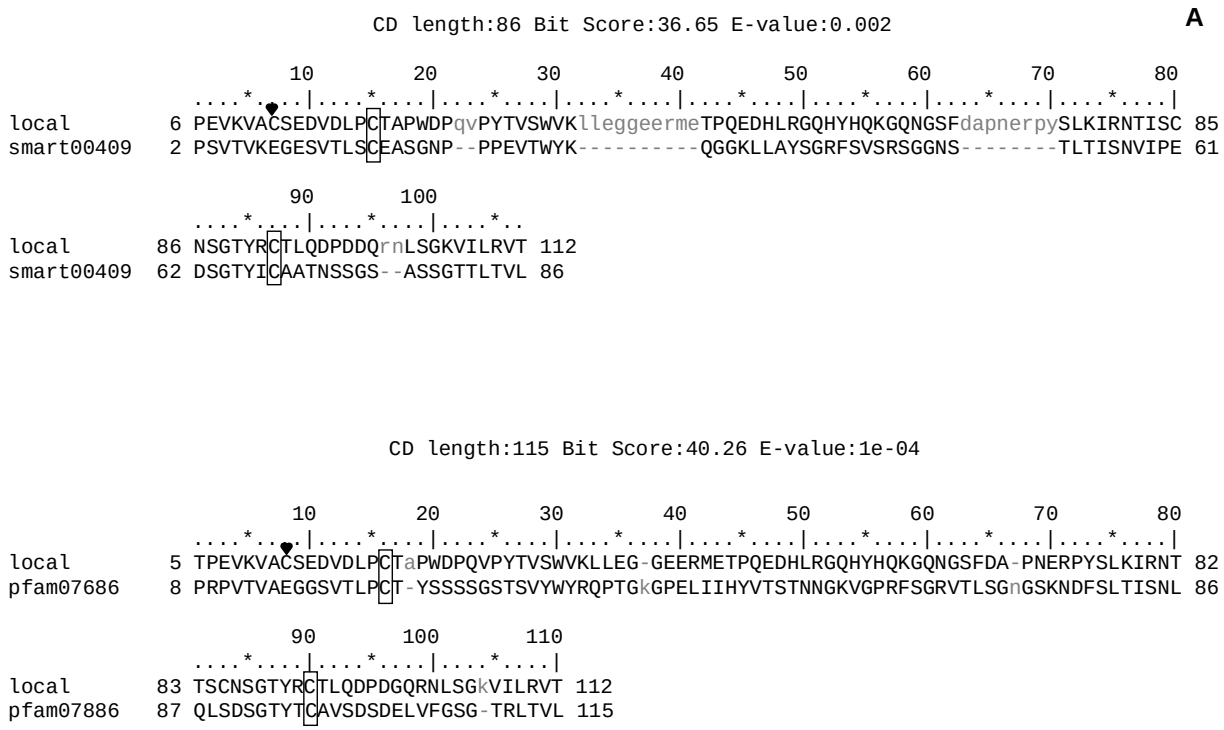


Figure 6.1 Conserved domain analysis in hCD83ext sequence performed by NCBI CD-search BLAST. (A) Pairwise alignments between the hCD83ext sequence (displayed as local) and the domain-model consensus sequences from Immunoglobulin domain family (the upper panel) and Immunoglobulin V-set domain (the lower panel). The conserved cysteine residues were boxed and the cysteine at the position 27 in hCD83ext was marked with black heart. SMART and Pfam are imported collections as the sources of conserved domain-search alignment model data [2]. Smart00409 and Pfam07686 were identified as the domain-models for hCD83ext sequence. (B) Predicted 3D structure of hCD83ext (shown in backbone style). The positions of five cysteine residues were marked as pointed by arrows.

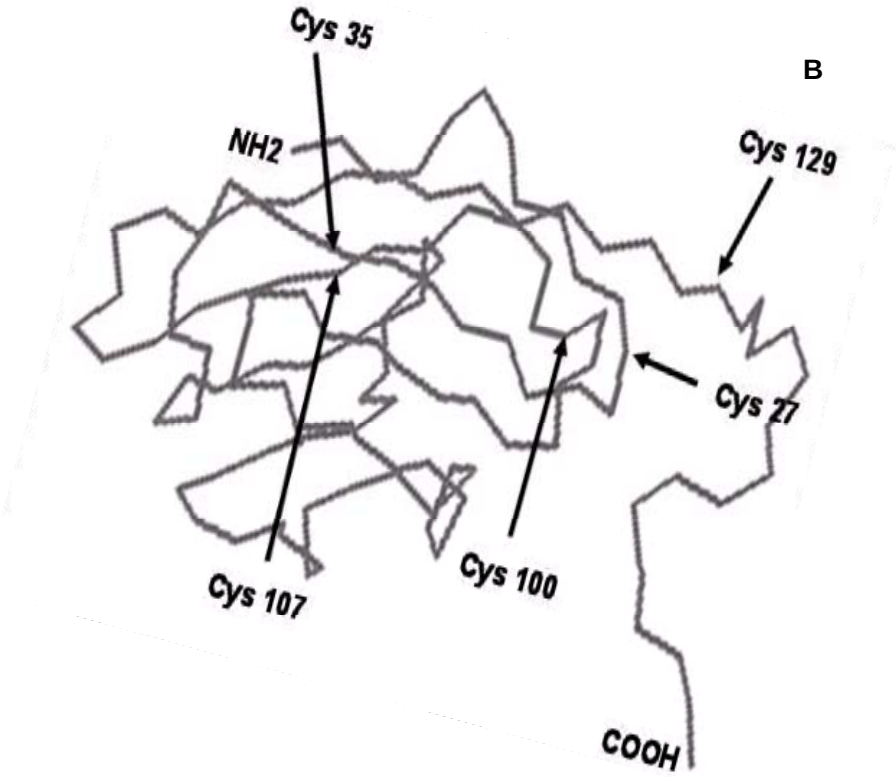


Figure 6.1 (Continued)

vo formation of intramolecular disulfide bond in Origami B(DE3) was rather ineffective. The observation is distinct from that when the fusion of GST and wild-type hCD83ext was expressed in Origami B(DE3) [99].

Table 6.1 *E*-values (expectation values) corresponding to the substitution of cysteine 27 with the other 19 amino acids (AAs). The *E*-value is used to compare the similarity of a query sequence to a domain sequence [36]. A lower *E*-value implies a higher similarity.

AA	Ig family	V-set domain
C	0.002	1e-04
G	0.001	7e-05
A	4e-04	3e-05
V	4e-04	1e-05
L	4e-04	1e-05
I	5e-04	2e-05
M	6e-04	2e-05
F	0.002	4e-05
W	0.003	8e-05
P	2e-04	1e-05
S	**7e-04**	**3e-05**
T	0.001	2e-05
Y	0.002	7e-05
N	0.001	2e-05
Q	5e-04	2e-05
D	7e-04	3e-05
E	**2e-04**	**6e-06**
K	7e-04	3e-05
R	5e-04	1e-05
H	0.001	2e-05

The selected substitutions and corresponding *E*-values are denoted in bold.

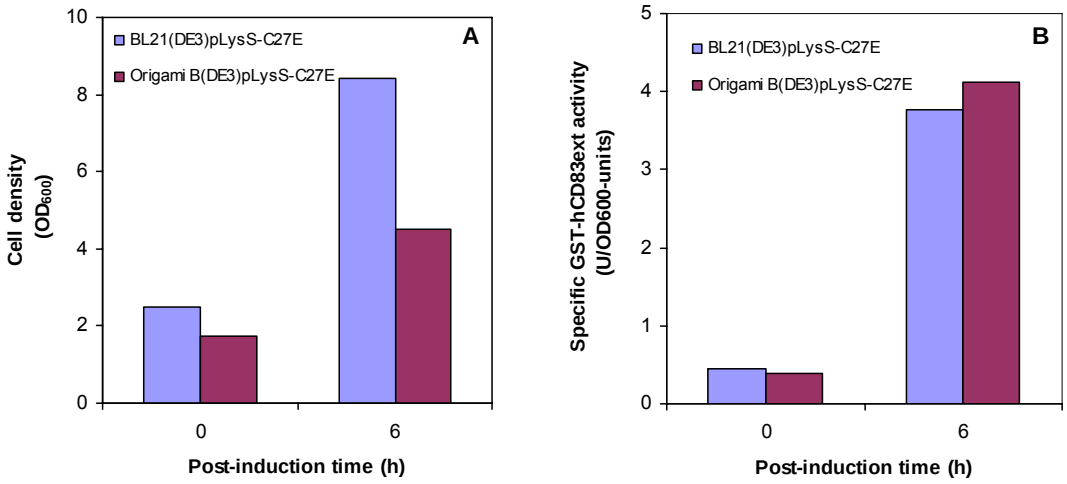


Figure 6.2 Comparison of bioreactor cultivation performance for the production of GST-hCD83ext_mut_C27E in two host strains: Cell density (Panel A), specific GST activity (Panel B), and intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panel C) and Western blotting (Panel D) for the culture samples taken prior to induction and at 6 h post-induction time were shown. Panel C: Lane 1 is molecular weight marker; Lanes 2, 3, 10, 11 and 4, 5 represent soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively, whereas lanes 6, 7, 12, 13 and 8, 9 represent soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows represent the protein bands corresponding to monomeric fusion protein. Panel D: All lanes are corresponding to the lanes 2~13 in Panel C.

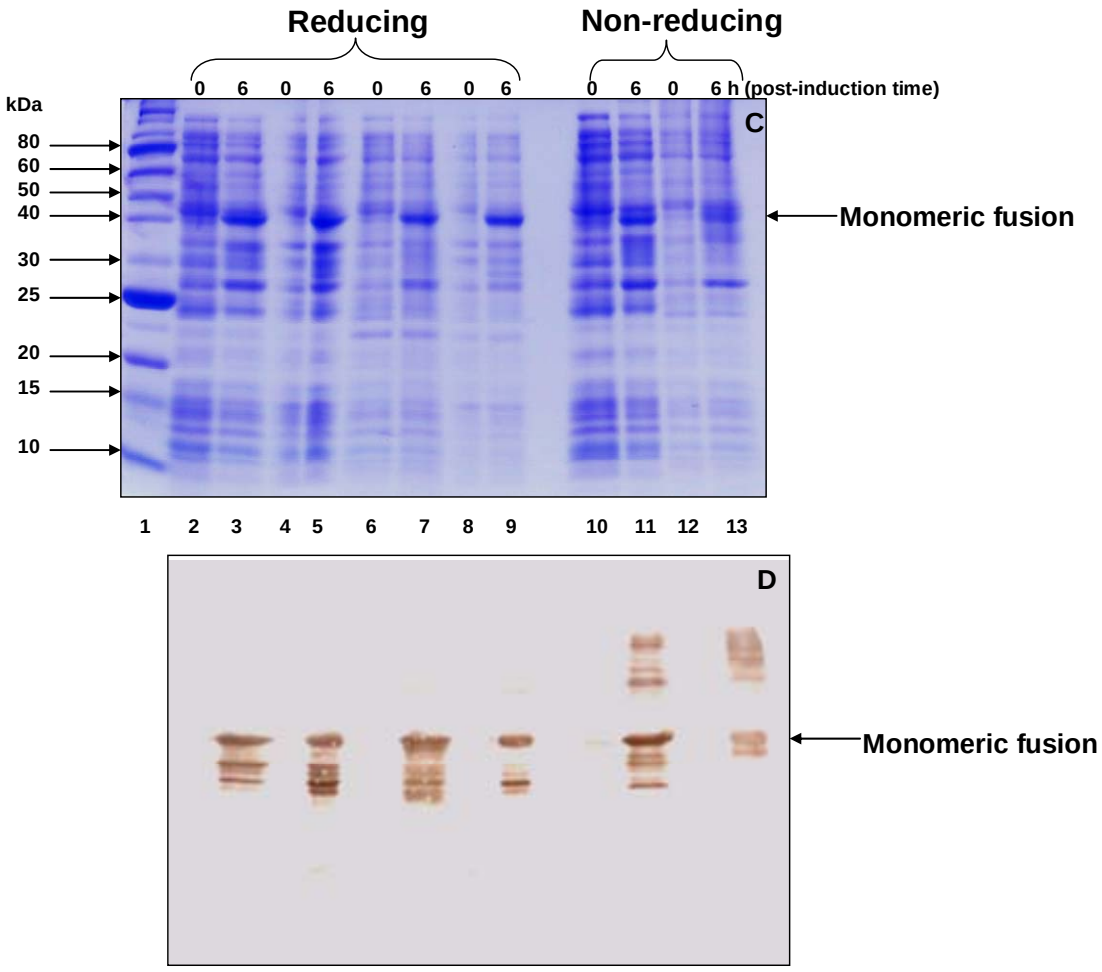


Figure 6.2 (Continued)

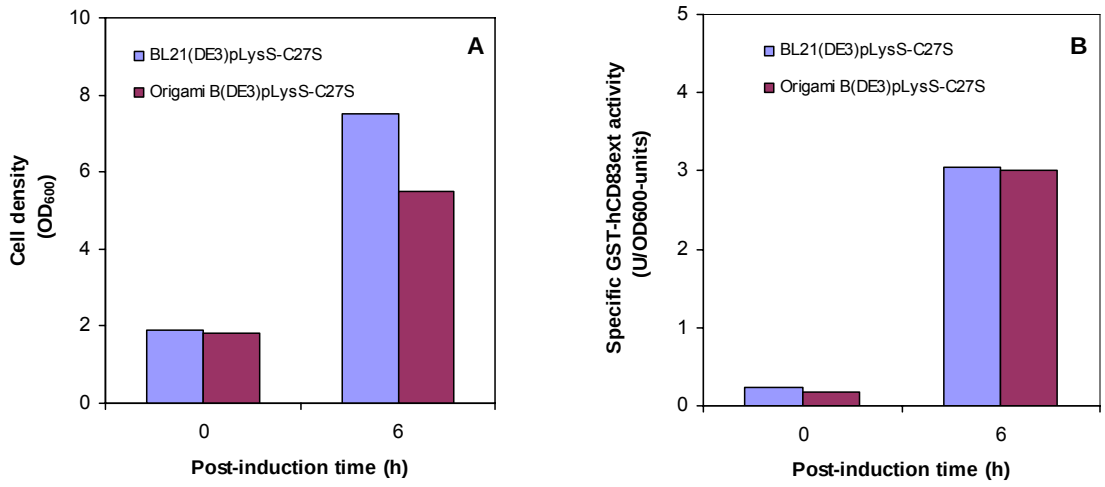


Figure 6.3 Comparison of bioreactor cultivation performance for the production of GST-hCD83ext_mut_C27S in two host strains: Cell density (Panel A), specific GST activity (Panel B), and intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panel C) and Western blotting (Panel D) for the culture samples taken prior to induction and at 6 h post-induction time were shown. Panel C: Lane 1 is molecular weight marker; Lanes 2, 3, 10, 11 and 4, 5 represent soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively, whereas lanes 6, 7, 12, 13 and 8, 9 represent soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows represent the protein bands corresponding to monomeric fusion protein. Panel D: All lanes are corresponding to the lanes 2~13 in Panel C.

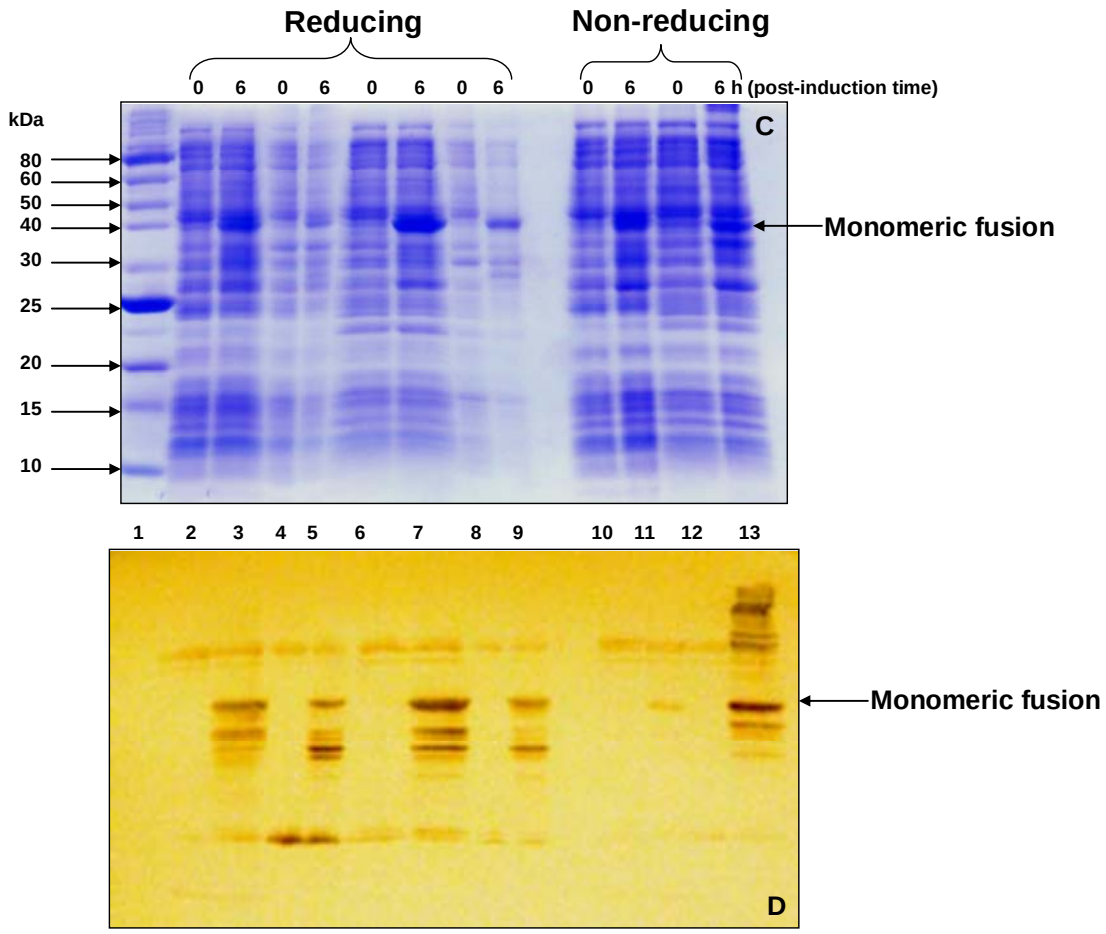


Figure 6.3 (Continued)

6.3.3 Characterization of Multimerization

Structural monitoring of the four hCD83ext mutant variants was conducted under -20 °C and a room temperature and the results are summarized in Figures 6.4 and 6.5.

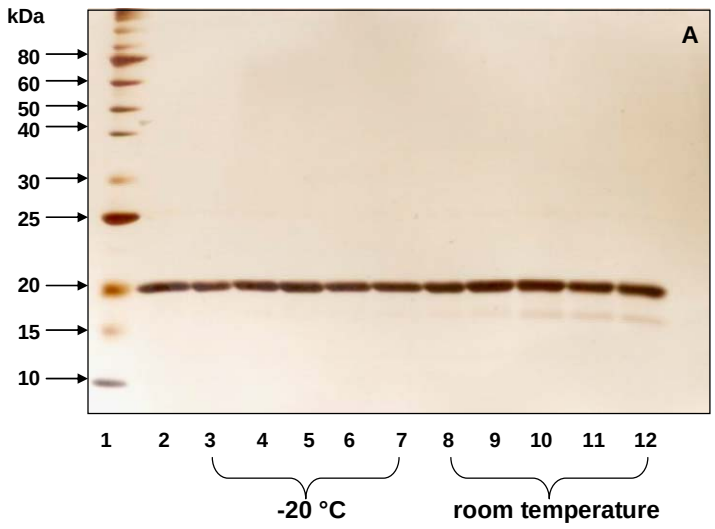


Figure 6.4 Monitoring and comparison of multimerization for hCD83ext_mut_C27E produced by two host strains with reducing SDS-PAGE (Panels A and C), non-reducing SDS-PAGE (Panels B and D): Samples up to 30-day storage at -20 °C and room temperature were analyzed. Panels A and B represent hCD83ext_mut_C27S samples produced by BL21(DE3)pLysS, whereas Panels C and D represent hCD83ext_mut_C27E samples produced by Origami B(DE3)pLysS. Panels A~B: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~ 7/ -20 °C samples taken on day 1, day 2, day 6, day 10, and day 30, respectively; lanes 7 ~ 10/ room temperature samples taken on day 1, day 2, day 6, day 10, and day30 respectively. Panels C~D: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~7/ -20 °C samples taken on day 1, day 2, day 6, day 10, and day 30, respectively; lanes 8~10/ room temperature samples taken on day 1, day 2, and day 6 respectively. The insets at upper right in Panels B and D represent the time profiles for dimerization of hCD83ext_mut_C27E stored at -20 °C and room temperature.

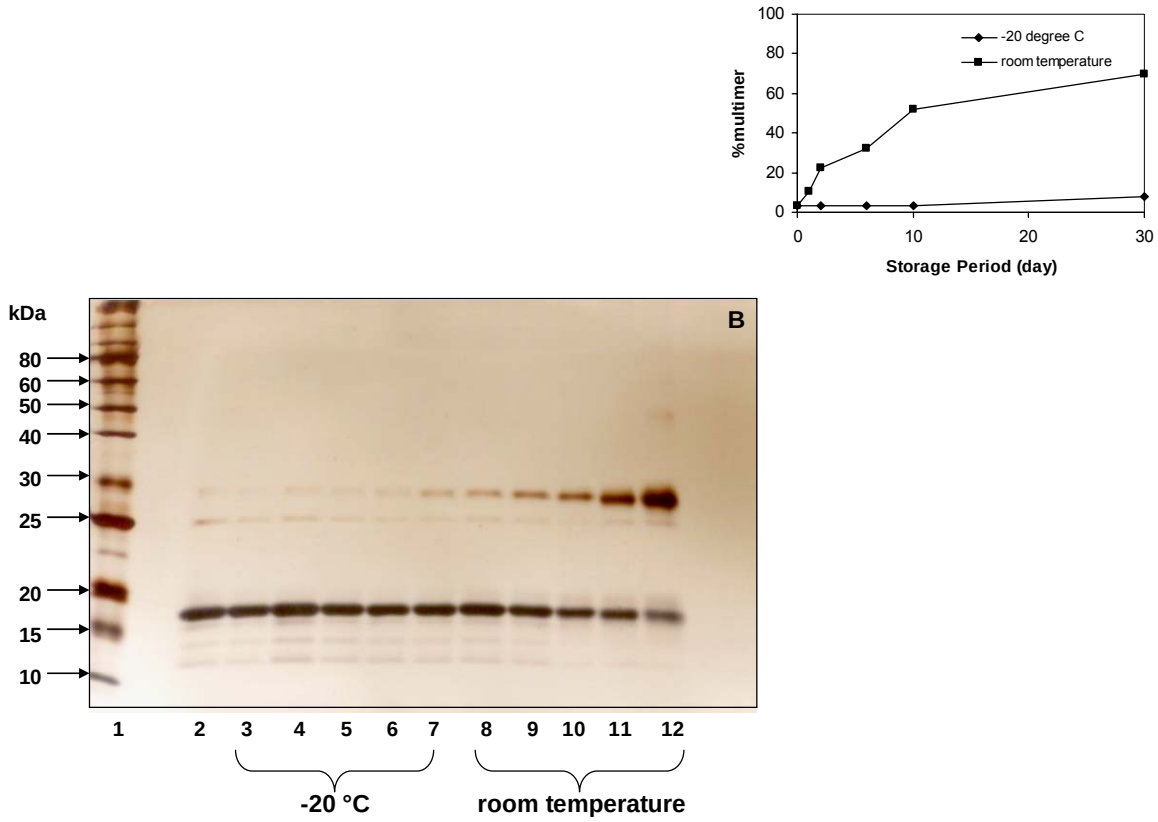
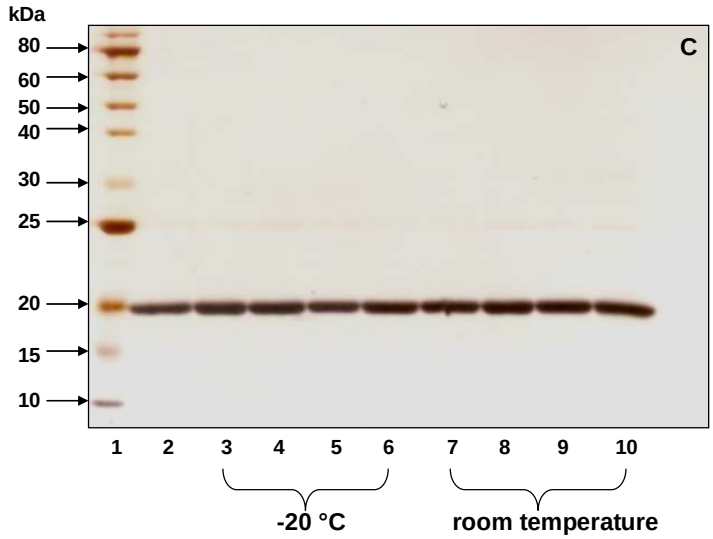
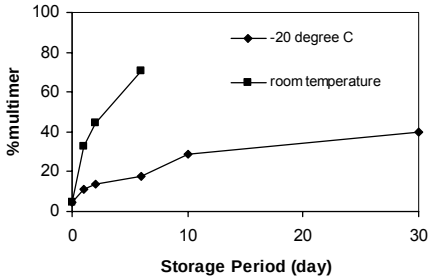


Figure 6.4 (Continued)





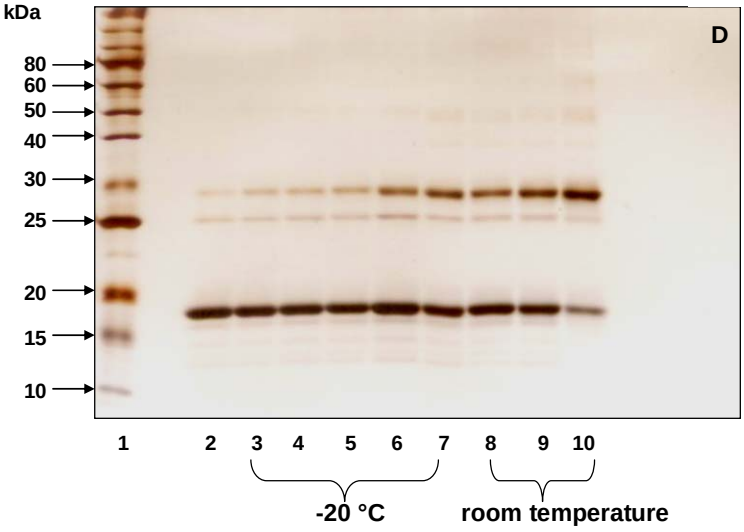


Figure 6.4 (Continued)

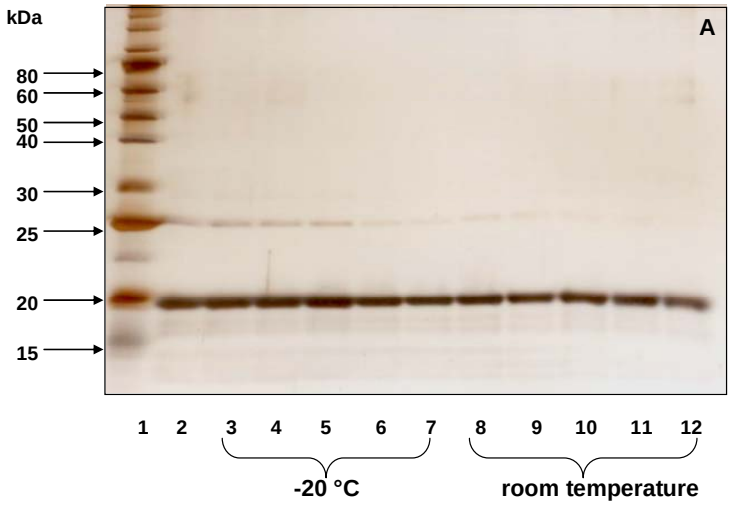


Figure 6.5 Monitoring and comparison of multimerization for hCD83ext_mut_C27S produced by two host strains with reducing SDS-PAGE (Panels A and C) and non-reducing SDS-PAGE (Panels B and D): Samples up to 30-day storage at -20 °C and room temperature were analyzed. Panels A and B represent hCD83ext_mut_C27S samples produced by BL21(DE3)pLysS, whereas Panels C and D represent hCD83ext_mut_C27S samples produced by Origami B(DE3)pLysS. Panels A~B: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~7/ -20 °C samples taken on day 1, day 3, day 5, day 10, and day 30, respectively; lanes 8~12/ room temperature samples taken on day 1, day 3, day 5, day 10, and day30 respectively. Panels C~D: lane 1/ molecular weight markers; lane2/ day 0; lanes 3 ~6/ -20 °C samples taken on day 2, day 5, day 10, and day 30, respectively; lanes 7~10/ room temperature samples taken on day 2, day 5, day 10, and day30 respectively. The insets at upper right in Panels B and D represent the time profiles for dimerization of

hCD83ext_mut_C27S stored at -20 °C and room temperature. Multimerization profiles were further plotted according to the relative abundance of monomeric and multimeric forms quantified from their intensities on the SDS gel.

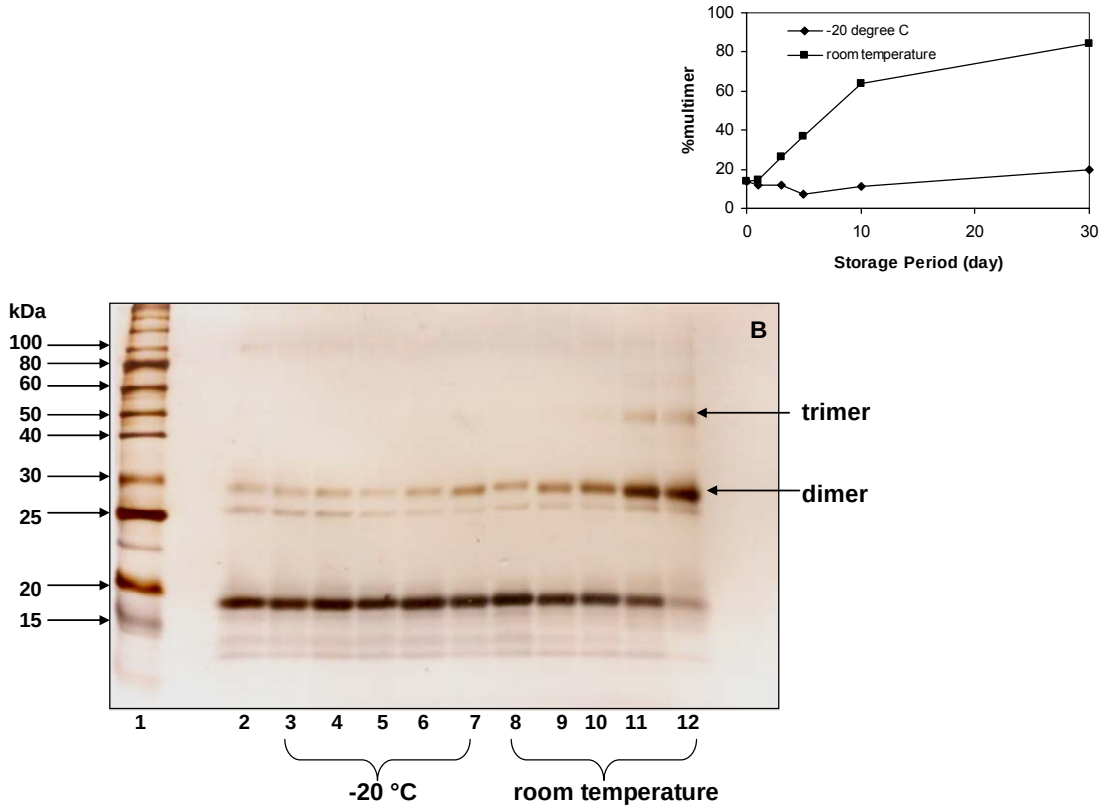


Figure 6.5 (Continued)

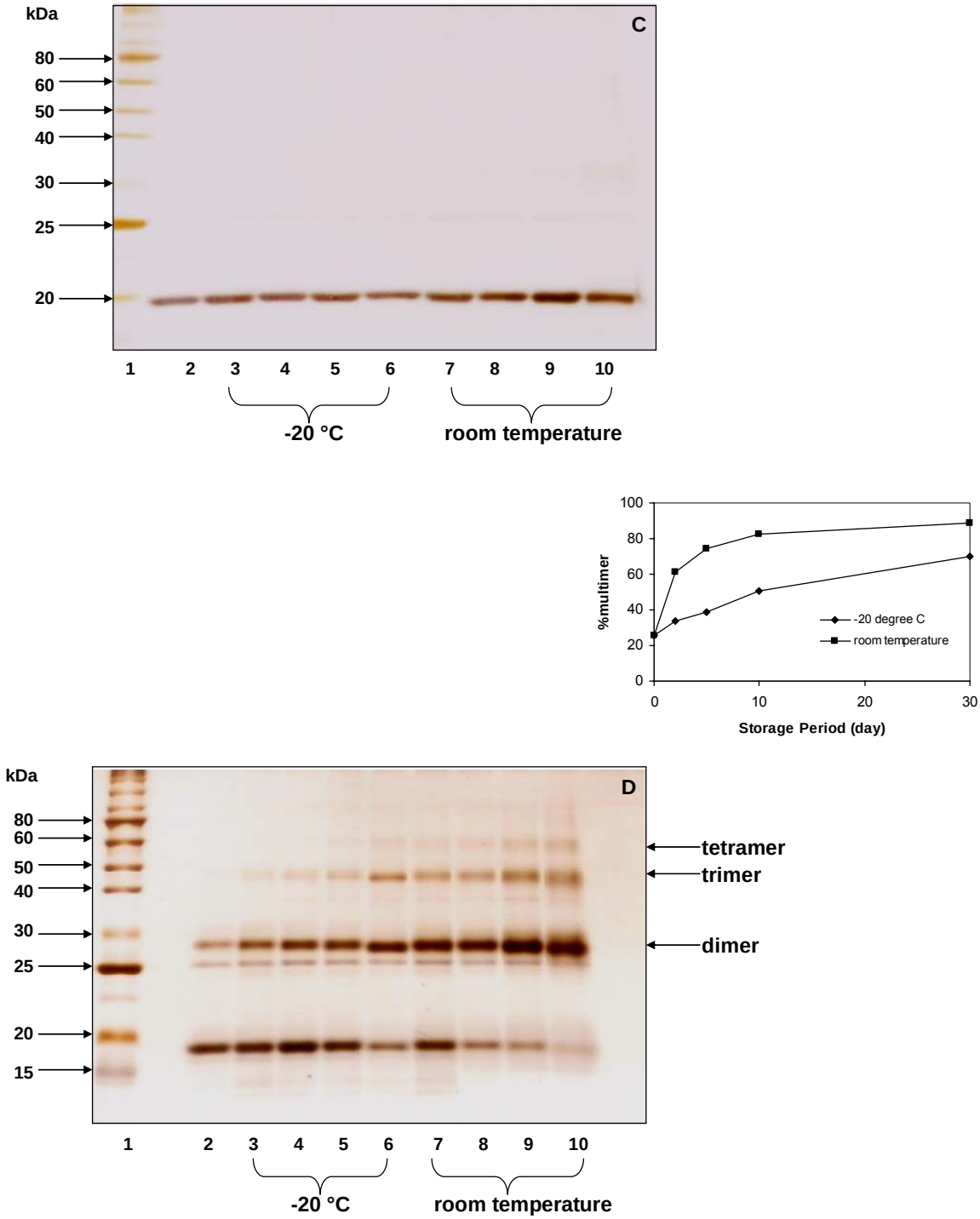


Figure 6.5 (Continued)

Monomeric and multimeric forms of hCD83ext were quantified by scanning the band intensity and various time profiles of multimerization were obtained. Kinetic simulation was conducted with an assumption that dimerization occurred exclusively since dimer represents the majority of multimeric forms. The two kinetic parameters, i.e. reaction order and rate constant, were estimated using the Metropolis-Hastings algorithm [102]. The reaction orders based on the eight scenarios in Figures 6.4 and 6.5 were determined to be 1.9~2.6. Therefore, a second-order reaction kinetics was used for data fitting to obtain rate constants (Table 6.2). Apparently, the multimerization was enhanced by an elevated temperature for all mutant variants. The hCD83ext mutant variants with a serine substitution or produced by Origami B(DE3) had a higher tendency for multimerization as compared to those counterparts with a glutamic acid substitution or produced by BL21(DE3).

6.3.4 Characterization of Conformational Changes

We also used rp-HPLC to characterize the change in protein surface hydrophobicity and the chromatograms of protein products stored at -20 °C are documented in Figure 6.6. Several peaks eluted at approximately 32.8%, 33.6%, 34.1% and 34.5% ACN were identified as structurally distinctive forms, implying that all mutant variants experienced structural changes. They all tended to come into the conformation with more surface hydrophobicity associated with the storage. The structural changes in the Origami-derived proteins appear to be more significant as compared to the BL21(DE3)-derived counterparts, presumably due to their enhanced tendency for multimerization. We also monitored the number of titratable free sulfhydryl group using Ellman's assay [35] (data not shown). In general, the number for all protein samples decreased monotonically, implying that disulfide bonds formed in vitro duri-

Table 6.2 Kinetic analysis of dimerization reaction

Species Name	Rate Constant (k)	R^2	R^2 –critical
hCD83ext_mut_C27S from BL21(DE3)pLysS at -20 °C	3.04×10^{-3}	0.9011	0.6583 (N=6)
hCD83ext_mut_C27S from BL21(DE3)pLysS at RT	1.33×10^{-1}	0.9675	0.6583 (N=6)
hCD83ext_mut_C27S from Origami B(DE3)pLysS at -20 °C	7.25×10^{-2}	0.9933	0.7714 (N=5)
hCD83ext_mut_C27S from Origami B(DE3)pLysS at RT	5.45×10^{-1}	0.9954	0.7714 (N=5)
hCD83ext_mut_C27E from BL21(DE3)pLysS at -20 °C	2.62×10^{-3}	0.799	0.6583 (N=6)
hCD83ext_mut_C27E from BL21(DE3)pLysS at RT	8.80×10^{-2}	0.9811	0.6583 (N=6)
hCD83ext_mut_C27E from Origami B(DE3)pLysS at -20 °C	2.21×10^{-2}	0.9234	0.6583 (N=6)
hCD83ext_mut_C27E from Origami B(DE3)pLysS at RT	3.39×10^{-1}	0.9969	0.9025 (N=4)

N is the number of independent sample measurements. The goodness of fit was checked by coefficient R^2 using R^2-critical calculated through critical values of the Student's t-distribution with probability =0.05 following the formula: t-critical^2 / [t-critical^2 + (N-2)].

ng the storage. The disulfide bond formation was aggravated by an elevated temperature, particularly in the initial incubation. Spectroscopic analysis of the conformational change was conducted using far-UV CD and fluorescence spectroscopy. Based on the deconvolution of far-UV CD spectra using the CDSSTR method on DichroWeb website [103, 104], it appears that no hCD83ext mutant variant had a major change in the secondary structure other than a slight increase in β-sheets (1~4%). However, overlaying far-UV CD spectra with different ages suggests otherwise, particularly in light of the signal differences at 195, 203 and 220 nm wavelengths which are typically associated with β-sheets. The Origami-derived proteins had more variations at 220 nm, whereas the BL21-derived counterparts displayed blue-shifts at shorter wavelengths. In the meantime, Origami-derived proteins had a significantly lower fluorescence intensity as compared to the BL21(DE3)-derived counterparts, suggesting the formation of certain intramolecular disulfide bond(s) could result in a major structural change in the local environment surrounding tryptophan residues. However, no specific pattern of fluorescence change associated with storage was observed.

6.3.5 Stability Enhancement Associated with the Cys^{27} Mutations

The changes in molecular structure for the wild-type hCD83ext and two Cys^{27} mutants upon a long-term storage at -20 °C and a room temperature was monitored and the results are summarized in Figure 6.7. Wild-type hCD83ext were stable at -20 °C, but subject to degradation at a room temperature. The degradation was significantly alleviated for the two Cys^{27} mutants. Interestingly, the trend of multimerization for hCD83ext variants was affected by several factors, including the Cys^{27} mutation, the expression host, and even the storage temperature.

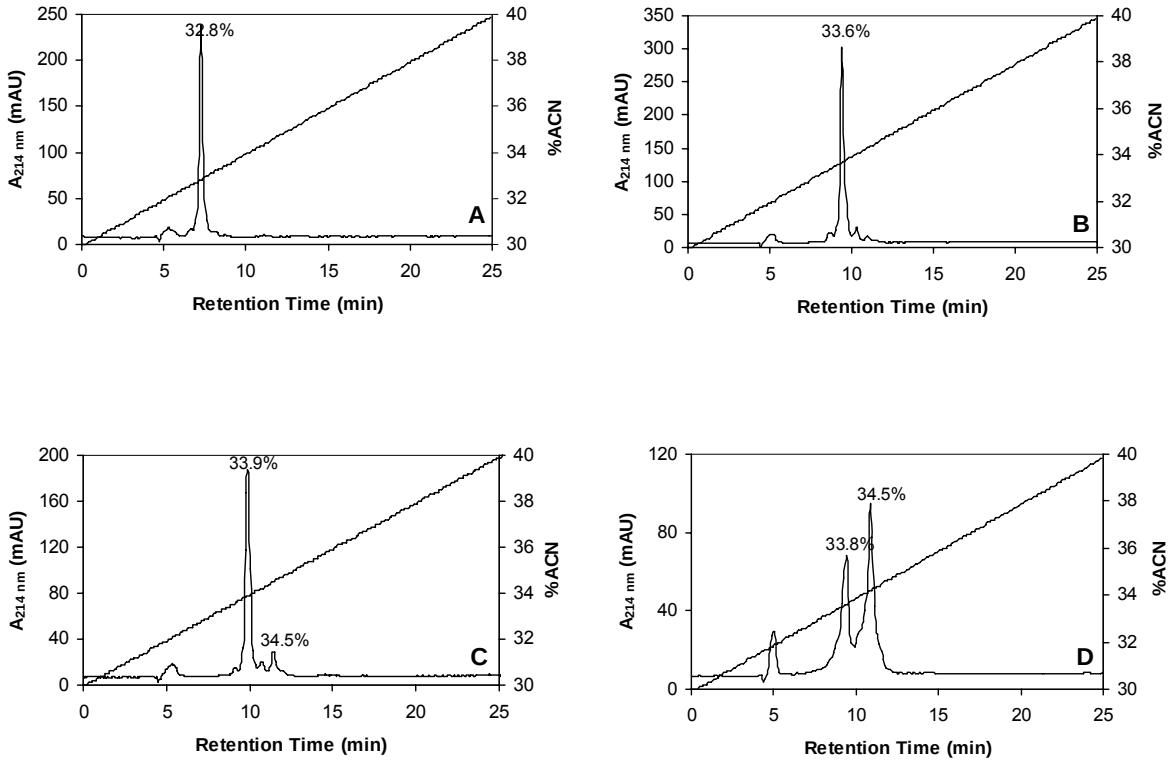


Figure 6.6 Monitoring and comparison of the retention behavior of hCD83ext mutant variants in rp-HPLC: Panels A and C represent hCD83ext_mut_C27E produced by BL21(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively; Panels B and D represent hCD83ext_mut_C27E produced by Origami B(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively. Panels E and G represent hCD83ext_mut_C27S produced by BL21(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively; Panels F and H represent hCD83ext_mut_C27S produced by Origami B(DE3)pLysS and samples were taken at day 0 and day 30 under -20 °C storage, respectively.

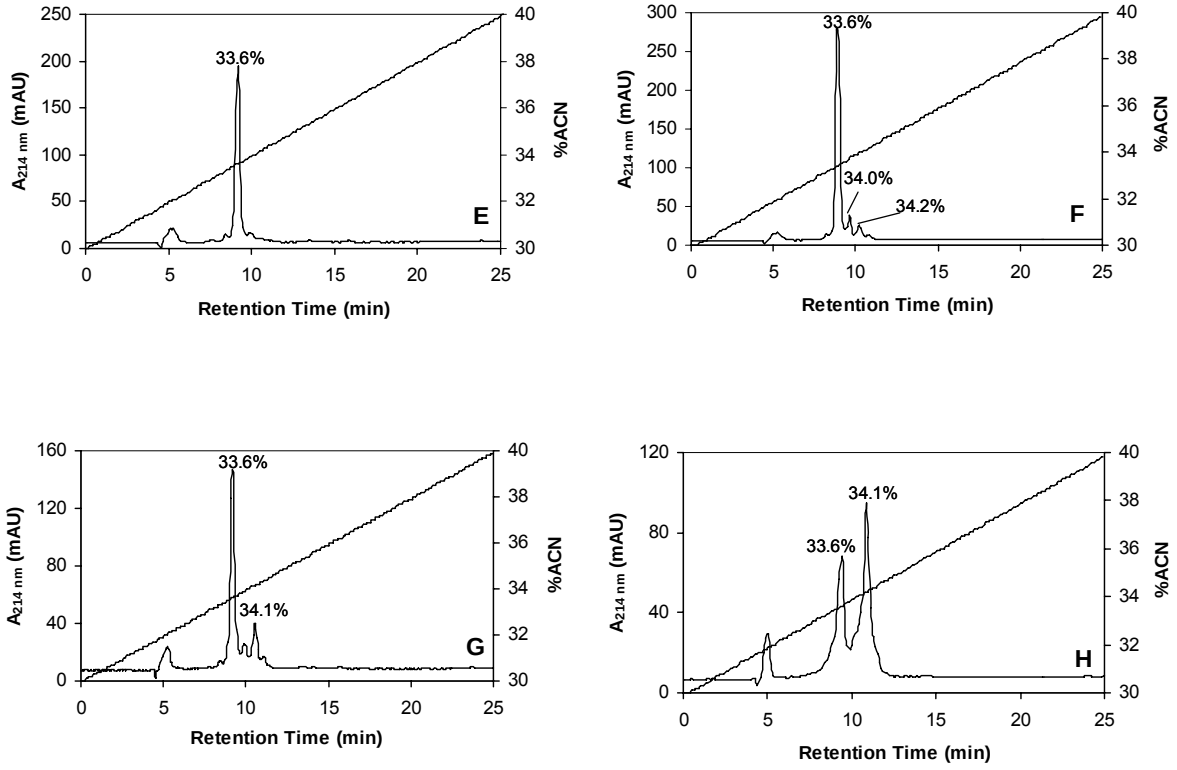


Figure 6.6 (Continued)

Compared to wild-type hCD83ext stored at -20 °C, the two Cys^{27} mutants derived from BL21(DE3) showed a lower tendency for dimerization (lanes 4 and 6 vs. lane 2). However, using Origami B(DE3) as the expression host increased the dimerization (Lane 5 vs. lane 4 and lane 7 vs. lane 6). The dimerization enhancement associated with the Cys^{27} mutations and the use of Origami B(DE3) as the expression host can be clearly observed for hCD83ext variants stored at a room temperature (lanes 9-14). It appears that the Cys^{27} mutations

increased dimerization (lanes 11 and 12 for hCD83ext_mut_C27E) or even multimerization (lanes 13 and 14 for hCD83ext_mut_C27S), which could prevent the degradation and subsequently enhanced the structural stability. The electrophoretic mobility of the non-reduced mutant monomers (18 kDa) or dimers (30 kDa) was higher than that of the corresponding non-reduced wild-type monomers (15 kDa) or dimers (27 kDa) (Figure 7B). The difference in the electro-mobility is presumably associated with the intramolecular disulfide bonding patterns and/or the number of intramolecular disulfide bonds. The stability enhancement was supported by spectroscopic analyses with far-UV CD and fluorescence. The CD spectra associated with the two storage temperatures for the wild-type hCD83ext had major differences, such as different signals at 203 nm and different band patterns at 195 nm, whereas these differences were insignificant for the two Cys^{27} mutants. In particular, the red-shift at 220 nm wavelength, which was previously suggested to be associated with the degradation for the Origami-derived wild-type sample at a room temperature [105], was not observed for the two Cys^{27} mutants. The signal variations within the 203~220 nm range of the far-UV CD spectra suggest that the β-sheet structure of hCD83ext might be twisted rather than planar [82], which is very common in globular proteins [106, 107]. The twist was suggested to affect the stability of the β-sheet structure through mediating the burial of hydrophobic surfaces [108] and, therefore, might have implications in the structural stability of hCD83ext which has more than 60% β-sheet structure (i.e. β-strands and β-turn). Similarly, the fluorescence spectra associated with the two storage temperatures for the wild-type hCD83ext were rather distinguishable, such as the increased signal at 400~420 nm and major red-shift of the emission maximum for room-temperature samples. The fluorescence

emission maximum at ~350 nm suggests that tryptophan residues are located on the protein surface in contact with free water molecules [109], whereas the oxidized forms of tryptophan emit fluorescence at distinctively longer wavelengths (400~420 nm) [110]. This features associated with the structural instability [105] were not observed for the two Cys^{27} mutants.

6.4 Discussion

The presence of free cysteines within a protein molecule can possibly lead to the aberrant formation of disulfide bonds in an oxidative storage condition. These cysteines become a rational target for molecular manipulation. On the other hand, a fixed disulfide pattern is presumably critical for developing consistent bioactivity for protein therapeutics. As a result, bioprocessing strategies leading to the desired disulfide pattern can play a pivotal role to ensure the quality of biopharmaceuticals. For example, controlling the oxidative state of a protein solution has been adopted to regulate the in vitro formation of disulfide bonds [111]. However, it is generally believed disulfide bond formation can be more consistently mediated in vivo (i.e. in biological cells) even though the desired disulfide pattern might not necessarily be developed. Since Origami B(DE3) is genetically manipulated to transform its cytoplasm into an oxidative compartment allowing disulfide bond formation [11, 73], the strain was applied to express the two Cys^{27} mutants of hCD83ext in this study. As opposed to the expression of wild-type hCD83ext in Origami B(DE3) where intramolecular disulfide bonds formed in vivo [99], the efficiency of disulfide bond formation was significantly reduced when the two Cys^{27} mutants were expressed in Origami B(DE3) (Figures 2 and 3).

The results suggest that Cys^{27} might be involved in the formation of intramolecular disulfide bonds in Origami B(DE3).

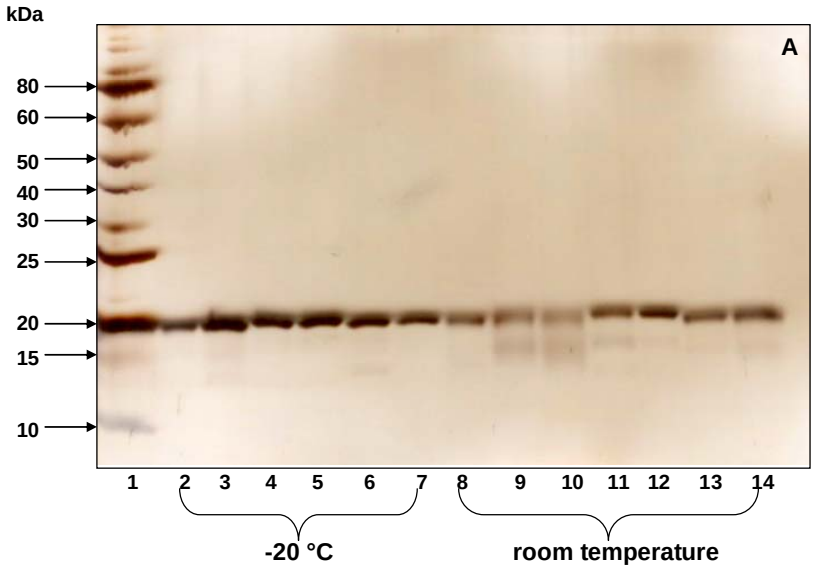


Figure 6.7 Monitoring and comparing the stability of hCD83ext wild-type protein with mutant proteins hCD83ext_mut_C27E and hCD83ext_mut_C27S produced by two host strains by reducing SDS-PAGE (Panel A), non-reducing SDS-PAGE (Panel B), CD spectroscopy (Panels C, D, and E), and fluorescence spectroscopy (Panels E, F, and G). Samples up to 30-day storage at -20 °C and room temperature were analyzed except that hCD83ext_mut_C27E produced in Origami B(DE3)pLysS was monitored till 6-day when stored at room temperature. Panels A and B : lane 1/ molecular weight markers; lanes 2~7/ -20 °C samples, arranged as lane 2/ wild-type from BL21(DE3)pLysS, lane 3/ wild-type from Origami B(DE3)pLysS; lane 4/ hCD83ext_mut_C27E from BL21(DE3)pLysS, lane 5/

hCD83ext_mut_C27E from Origami B(DE3)pLysS, lane 6/ hCD83ext_mut_C27S from BL21(DE3)pLysS, lane 7/ hCD83ext_mut_C27S from Origami B(DE3)pLysS; lane 8/ wild-type from Origami B(DE3)pLysS 6-day storage at RT; lanes 9~14/ corresponding RT samples following the same arrangement. The arrows indicate the monomeric and dimeric species of different mobilities in wild-type and mutant proteins. Due to the limited amount of the sample, hCD83ext_mut_C27E produced by Origami B(DE3)pLysS was monitored for 6 days at room temperature. Accordingly, the comparison was made to the wild-type hCD83ext produced by Origami B(DE3)pLysS with the same storage conditions and period.

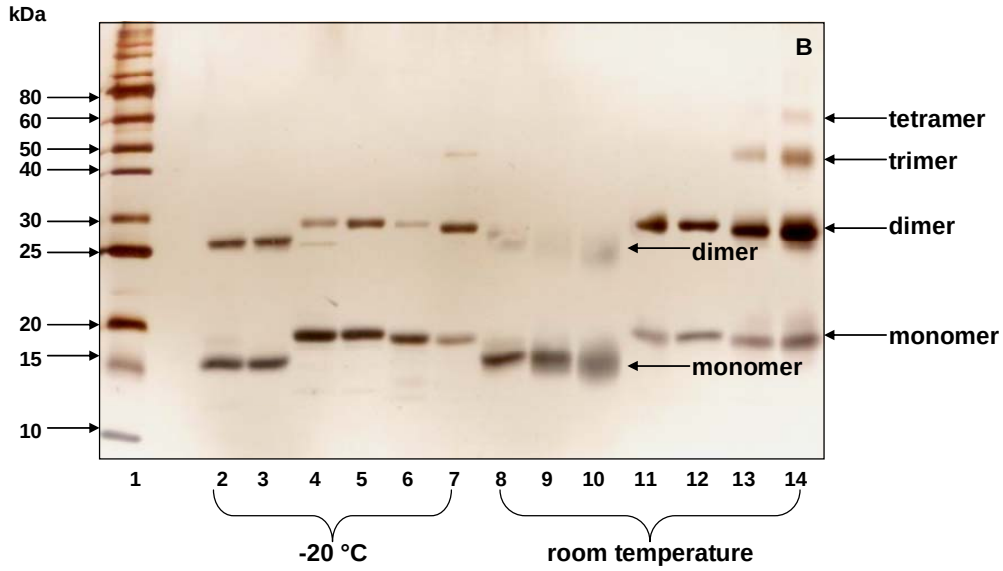


Figure 6.7 (Continued)

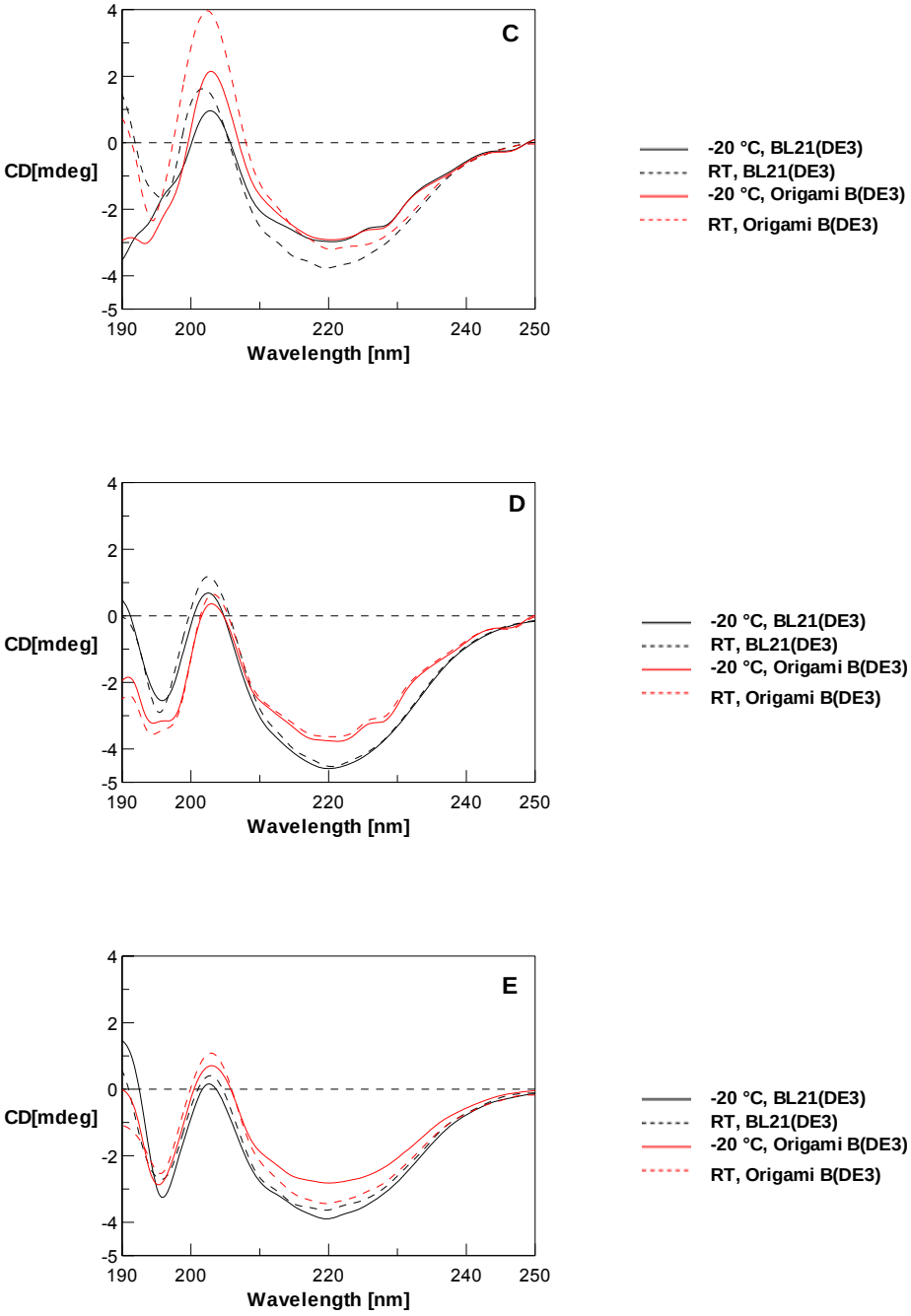


Figure 6.7 (Continued)

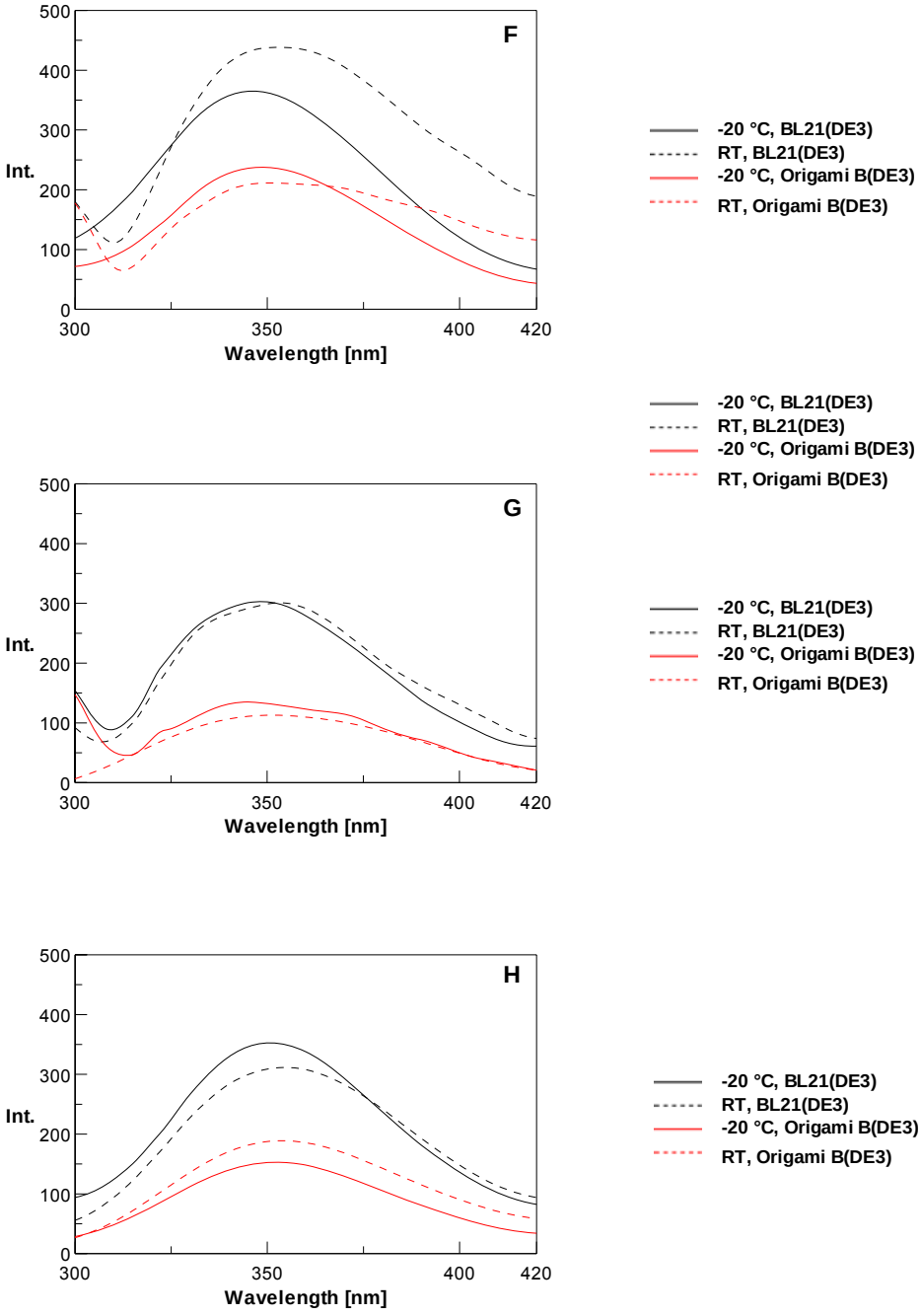


Figure 6.7 (Continued)

In addition to the *in vivo* effect, the Cys^{27} mutations significantly affected the *in vitro* molecular behavior of the protein product during the storage. Similar to wild-type hCD83ext [99], multimerization via the *in vitro* formation of intermolecular disulfide bonds still occurred to the two Cys^{27} mutants derived from either BL21(DE3) or Origami B(DE3). The Origami-derived Cys^{27} mutant variants had a higher tendency for multimerization than the BL21-derived counterparts and certain structural variations in the fresh protein products could be associated with this different multimerization behavior. However, such structural variations are expected to be less significant than those associated with the fresh protein products of wild-type hCD83ext derived from BL21(DE3) and Origami B(DE3) [99] based on at least two observations. First, the efficiency for the *in vivo* formation of intramolecular disulfide bonds in Origami B(DE3) was seriously impaired by the Cys^{27} mutations (Figures 6.2 and 6.3). Second, the analysis based on the fluorescence spectroscopy (i.e. quenching of the tryptophan fluorescence reflected by the different intensity at the emission maximum) revealed fewer conformational differences between the Cys^{27} mutants derived from BL21(DE3) and Origami B(DE3), as compared to the case of wild-type hCD83ext [99]. Note that the electrophoretic mobility of the monomer of the Origami-derived products was approximately the same as that of the BL21-drived counterparts. Because the intramolecular disulfide bonds cannot be formed in the reducing cytoplasm of BL21(DE3), the results imply they are effectively formed *in vitro* during the downstream processing (unpublished data). The *in vitro* multimerization of the Cys^{27} mutant also appears to be aggravated more by Ser substitution than Glu substitution (for both BL21-derived and Origami-derived products) and this could be due to increased steric hindrance and impulsion associated with the bulky and

negatively charged side chain of glutamic acid. The argument is supported by the 3D structural prediction that Cys^{27} is lies in a turn connecting two β-strands near the surface.

Molecular stability of protein therapeutics has been recognized as a key issue in biopharmaceutical research and development since denatured and instable proteins can potentially jeopardize the safety and efficacy of therapeutic products [30]. Recombinant hCD83ext has been found to be susceptible to degradation during a prolonged storage particularly at elevated temperatures [105]. In this study, molecular stability of hCD83ext was found to be enhanced by the Cys^{27} mutations with a concomitantly altered disulfide pattern. Based on the reduced electrophoretic mobility, the Cys^{27} mutant variants could have less intramolecular disulfide bonds in average, compared to that of the wild-type hCD83ext [35]. This argument is supported by the Ellman assay results that more titratable free sulfhydryl groups are presented in the Cys^{27} mutant variants. Therefore, mutating Cys^{27} might somehow prevent the formation of intramolecular disulfide bond, both *in vivo* and in vitro, and this preventing effect was independent of the expression host. The stability improvement associated with the Cys^{27} mutations appears to be mediated through multimerization since all multimeric forms of the Cys^{27} mutants were much less susceptible to degradation. On the other hand, the multimerization did not improve the molecular stability for wild-type hCD83ext, suggesting the presence of structural variations between the Cys^{27} mutant and wild-type hCD83ext. The results also suggest that Cys^{27} might not be critically involved in the formation of intermolecular disulfide bonds for multimerization. Recently, we have identified Cys^{100} as a key cysteine responsible for multimerization [101], implying that the sulfhydryl group of Cys^{100} should be retained in monomeric hCD83ext

prior to multimerization. Based on the predicted 3D structure of hCD83ext (Figure 6.1B), Cys^{27} is spatially close to both Cys^{100} and Cys^{129}, suggesting possible formation of intramolecular disulfide bonds among these cysteine residues. In particular, the formation of an intramolecular disulfide bond between Cys^{27} and Cys^{100} could inactivate multimerization. This argument is supported by the observation that multimerization was enhanced upon introducing the Cys^{27} mutation, which can potentially free more Cys^{100} sulfhydryl groups for multimerization. Though molecular stability was improved upon introducing the Cys^{27} mutations, consistent bioactivity tends to be affected by the aggravated multimerization which demolishes the structural homogeneity of the protein therapeutics.

Chapter 7

Identification of Molecular Roles of Cysteine Residues Associated with *In Vitro* Multimerization of Therapeutic Protein Human CD83

7.1 Introduction

An hCD83ext mutant (i.e. M3) with a Cys100Ser mutation was derived to exhibit a more consistent bioactivity without disulfide-driven multimerization [101]. The underlying mechanisms associated with the improvement in molecular property and therapeutic bioactivity for M3 remained uncharacterized. In this study, more extended mutants of M3, i.e. M13 (Cys27Ser/Cys100Ser), M135 (Cys27Ser/Cys100Ser/Cys129Ser) and M35 (Cys100Ser/Cys129Ser), were derived and two *E. coli* B strains, i.e. BL21 and Origami B, were used as the expression host to produce these mutant derivatives. The molecular structures of these hCD83ext mutant variants were characterized and compared. The characterization data enable us to gain a partial insight regarding the structural role that these key cysteine residues play.

7.2 Materials and Methods

7.2.1 Bacterial Strains and Construction of hCD83ext Mutants

Plasmid pGEX2ThCD83ext contains the wild-type hCD83ext cDNA fused with the *gst* gene [8]. The two plasmid derivatives of pGEX2ThCD83ext_mut_C100S and pGEX2ThCD83ext_mut_C27S/C100S were used to produce hCD83ext_mut_C100S (i.e. M3) and hCD83ext_mut_C27S/C100S (i.e. M13). To construct these plasmids, site-directed mutagenesis was performed using QuikChange® Multi Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) with pGEX2ThCD83ext as the template. Two primers of 5'-ATCCGAAACACTAC*G*AGC**TC**CAACTCGGGG-3' for mutating Cys residue at the position AA100 to Ser and 5'-GAGGTGAAGGTGGC*G***AGC**TCCGAAGATGTG-3' for mutating Cys residue at the position AA27 to Ser (nucleotidic mutations are denoted in italics and the codons corresponding to serine are in bold). Both primers have a silent mutation generating an extra *SacI* restriction site (underlined). By using the same primers with the PCR template of pGEX2ThCD83ext_mut_C129S (kindly presented by Prof. Lechmann), two plasmid derivatives pGEX2ThCD83ext_mut_C100S/C129S and pGEX2ThCD83ext_mut_C27S/C100S/C129S were constructed and used for the production of hCD83ext_mut_C100S/C129S (i. e. M35) and hCD83ext_mut_C27S/C100S/C129S (i.e. M135). Purification of plasmid DNA was performed using a spin-column kit (QIAGEN, Mississauga, Ontario, Canada). The correct nucleotide sequences were verified by sequencing. BL21(DE3)pLysS and its *trxB* and *gor* double mutant derivative Origami

B(DE3)pLysS (EMD Biosciences/Novagen, La Jolla, CA, USA) were used as the expression hosts. Plasmid transformation was carried out by electroporation.

7.2.2 Protein Production and Purification

The production and purification of hCD83ext mutant variants, i.e. M3, M13, M35 and M135, in the two host strains was performed using the bioprocess developed previously [7]. The protein products were formulated in 50% glycerol.

7.2.3 Sample Treatment for Analyses

The culture sample was appropriately diluted with saline solution for measuring cell density at OD_{600} with a spectrophotometer (DU®520, Beckman Coulter, Fullerton, CA, USA). For the preparation of cell extract, cells in the amount of 20 OD_{600}-units (defined as 'OD_{600}×mL') were centrifuged at 6000 × *g* and 4 °C for 10 min. The cell pellet was resuspended in 0.75 mL phosphate-buffered saline (pH 7.3), sonicated intermittently (i.e. 0.5 s/0.5 s on/off) for 4 min using an ultrasonic processor (Misonix, Farmingdale, NY, USA) with a microtip, and then centrifuged at 15,000 × *g* and 4 °C for 15 min to remove cell debris. The supernatant containing soluble proteins was used for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting analyses. The pellet containing insoluble proteins and cell debris was washed with phosphate buffer, resuspended in TE/SDS buffer (10 mM Tris HCl, pH 8.0, 1 mM EDTA, 1% SDS), and heated to 100 °C for 5 min for dissolution. The protein content of the pellet was analyzed as the insoluble fraction.

7.2.4 Analytical Methods

SDS-PAGE was performed in a Mini-PROTEAN®II electrophoresis cell (Bio-Rad, Hercules, CA, USA) under both reducing and non-reducing conditions using a 12.5% polyacrylamide separating gel stacked by a 4% polyacrylamide stacking gel. Reduced and non-reduced protein samples were prepared using the 2X SDS sample buffer [65 mM Tris-HCl, pH6.8, 25% glycerol, 2% (w/v) SDS, 0.01% (w/v) bromophenol blue] respectively supplemented with 5% 2-mercaptoethanol and 10 mM *N*-ethylmaleimide (NEM). Electrophoresis was conducted at a constant voltage of 200 V for approximately 50 min. For the analysis of samples from cultivation, 0.0175 OD_{600}-units of soluble and insoluble fractions were loaded and the gel was stained with Coomassie blue. For the monitoring of molecular properties during storage, 0.25 µg of protein was loaded and silver nitrate staining was applied for higher sensitivity. All gels were dried and scanned.

To conduct Western blotting, the proteins on the polyacrylamide gel were electroblotted to a poly(vinylidene difluoride) (PVDF) membrane (Pall Life Sciences, Pensacola, FL, USA) after SDS-PAGE using a Mini Trans-Blot® Cell (Bio-Rad) according to a standard protocol [80]. The electrophoretic transfer was conducted at a constant voltage of 100 V for 1 h. Protein-antibody binding was performed as described by Sambrook et al. [81]. The purified hCD83ext was used as an antigen for raising polyclonal anti-CD83 antibodies in rabbits. The secondary antibody was goat anti-rabbit IgG conjugated with horseradish peroxidase (Sigma-Aldrich, St. Louis, MO, USA). All hCD83-related polypeptides (GST-hCD83ext mutant variants) were detected by a colorimetric method using 3,3'-

diaminobenzidine tetrahydrochloride (DAB) as the substrate. The processed membrane was scanned.

Protein samples at 0.1 mg/mL were prepared in 20 mM Tris-HCl buffer (pH 7.0) for spectroscopic analysis. Circular dichroism (CD) spectra were obtained using a spectropolarizer (J-810; Jasco, Tokyo, Japan). A cuvette with a path length of 0.1 cm was used for measurement in the far ultraviolet (UV) region from 250 to 190 nm. Fluorescence spectra were obtained using a spectrofluorometer (FP-6500; Jasco). A cuvette with a path length of 1 cm was used for fluorescence measurement in the range of 300~420 nm at an excitation wavelength of 295 nm. All spectra were recorded at an ambient temperature. The average of three scans was taken by subtracting the buffer spectrum obtained under the same conditions.

7.3 Results

7.3.1 Rational for Mutagenesis Studies

Because hCD83ext showed a high structural similarity to the Immunoglobulin V-set (IgV) domain [89], multiple sequence alignment between hCD83ext and several IgV members was conducted (Figure 7.1). Of the five cysteine residues in hCD83ext, Cys^{35} and Cys^{107} (boxed and in red) are highly conserved, whereas Cys^{27} and Cys^{100} (circled and in blue) are not. Note that Cys^{129} is not even included in the alignment. An intramolecular disulfide bond between Cys^{35} and Cys^{107} was previously predicted [9]. Because the presence of free cysteines can potentially result in the formation of aberrant disulfide bonds leading to inconsistent

molecular properties, Cys[27], Cys[100], and Cys[129] were selected as the mutagenesis targets to eliminate possible free cysteines in hCD83ext in this study. In fact, the single mutation effects associated with each of these cysteines have been investigated [89, 101]. The combinatorial effects of these individual cysteine mutations and the expression host presented in this study enabled us to partially explore the underlying mechanism associated with our previous observation of the improved molecular properties of M3 [101].

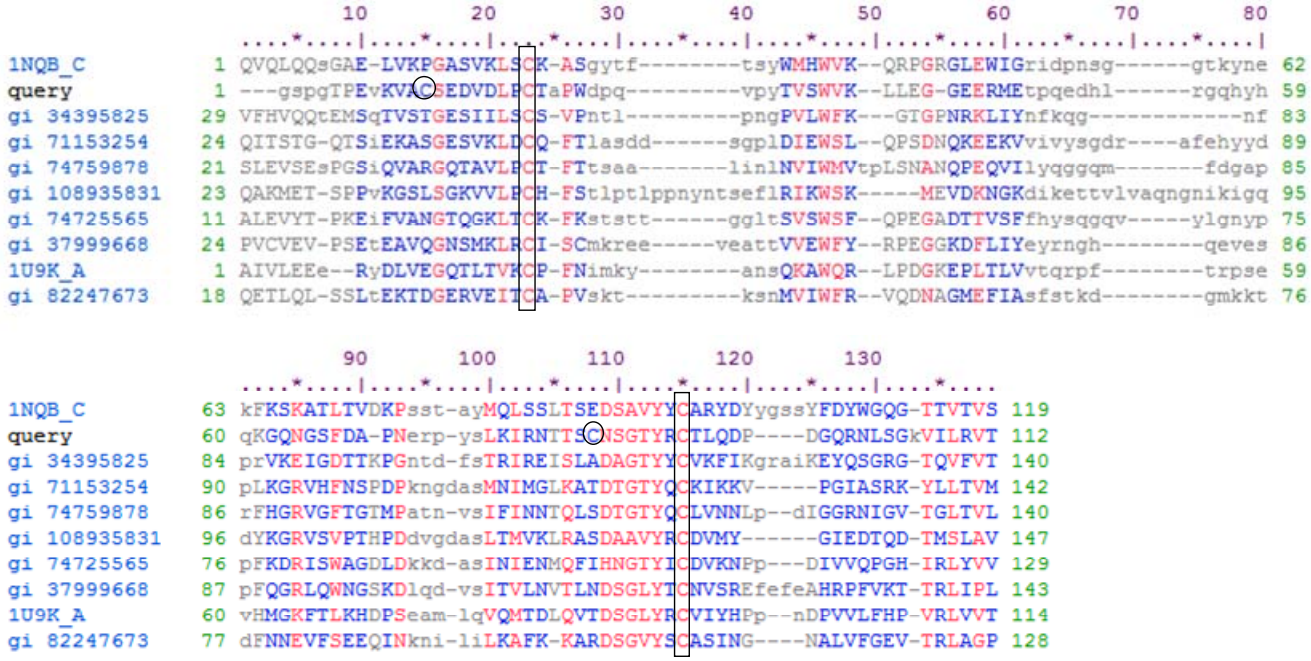


Figure 7.1 Multiple sequence alignment of the hCD83ext (query) with the other Immunoglobulin V-set domain members. The alignment was generated via NCBI Conserved Domain Search Blast on line. Conserved cysteine residues were boxed, whereas variable cysteine residues were circled in the hCD83ext sequence.

7.3.2 Comparison of Molecular Properties for M3 Variants

Using BL21(DE3)pLysS and Origami B(DE3)pLysS as the hosts, cultivation for the expression of GST-hCD83ext_mut_C100S (M3) was conducted in a bench-top bioreactor and the results are summarized in Figure 7.2. The M3-GST fusion expressed in Origami B(DE3)pLysS had a higher electrophoretic mobility under a non-reducing condition compared to that of the corresponding monomeric form under a reducing condition (lane 7 vs. 13), implying that certain intramolecular disulfide bond(s) was formed in the oxidative cytoplasm of Origami B(DE3). On the other hand, the same electrophoretic mobility for the M3-GST fusion expressed in BL21(DE3) under both reducing and non-reducing conditions (lane 3 vs. 11) suggested that no intramolecular disulfide bond was formed in BL21(DE3) host as expected.

Structural analysis of two fresh M3 product samples was conducted using SDS-PAGE, far-UV CD, and fluorescence spectroscopy, and the results are summarized in Figure 7.3. Under a non-reducing condition, multimers, i.e. dimers and trimers formed presumably via intermolecular disulfide bonds, were distinctively present for the Origami B(DE3)-derived M3 (Figure 7.3A). While BL21(DE3)-derived M3 had two disulfide bonded monomeric forms (i.e. two bands at 18 kDa and 15 kDa, Figure 7.3A), Origami B(DE3)-derived M3 had one monomeric form (i.e. a band at 15 kDa, Figure 7.3A). Overlaying the far-UV CD and fluorescence spectra of the two M3 variants revealed significant structural differences at both secondary and tertiary levels. By comparing the far-UV CD spectra of the two variants (Figure 7.3B), the Origami B(DE3)-derived M3 showed more positive signals at 195 nm and 203 nm wavelengths but a more negative signal at 220 nm wavelength, although

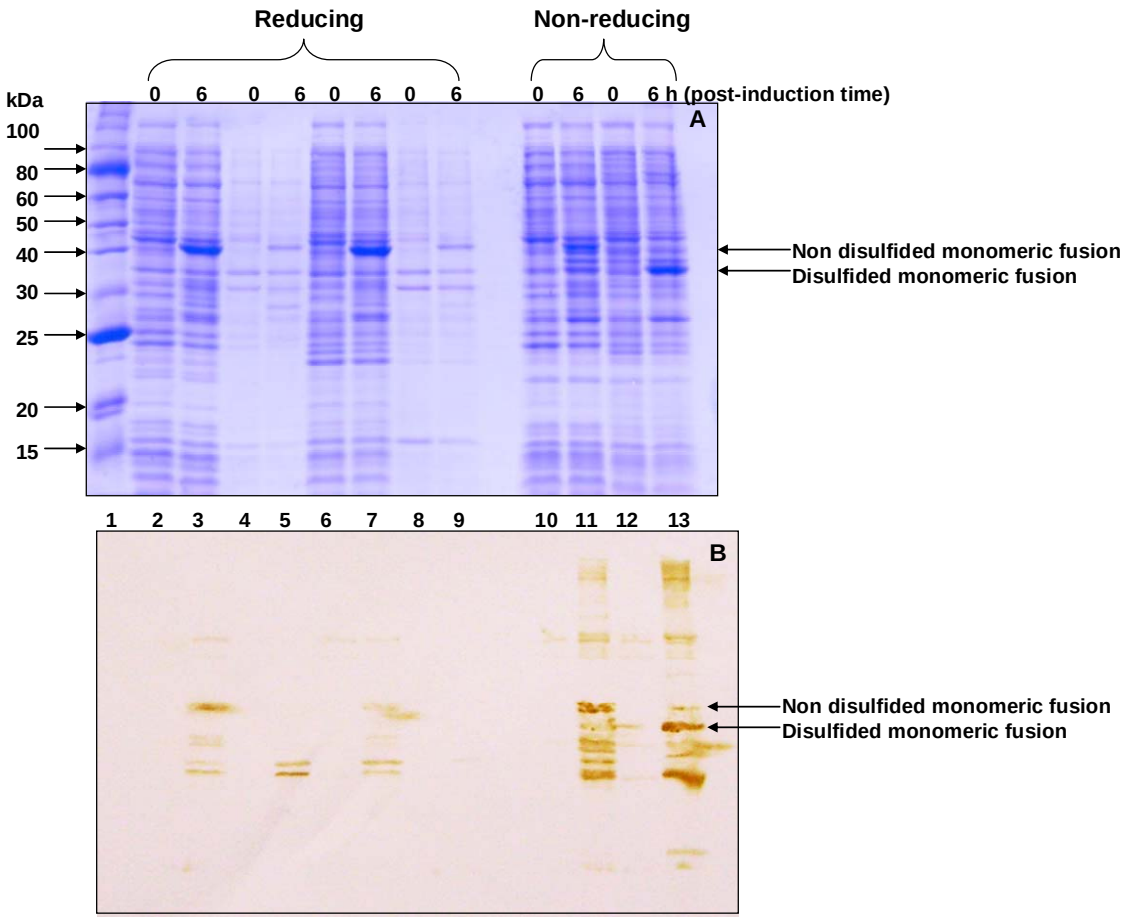


Figure 7.2 Intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panel A) and Western blotting (Panel B) for the culture samples of M3 taken prior to induction and at 6 h post-induction time were shown. Panels A & B: lane 1 was molecular weight markers; lanes 2, 3, 10, 11 and 4, 5 represented soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively, whereas lanes 6, 7, 12, 13 and 8, 9 represented soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows in two panels indicated the protein bands corresponding to disulfide bonded and non-disulfide bonded monomeric fusion protein.

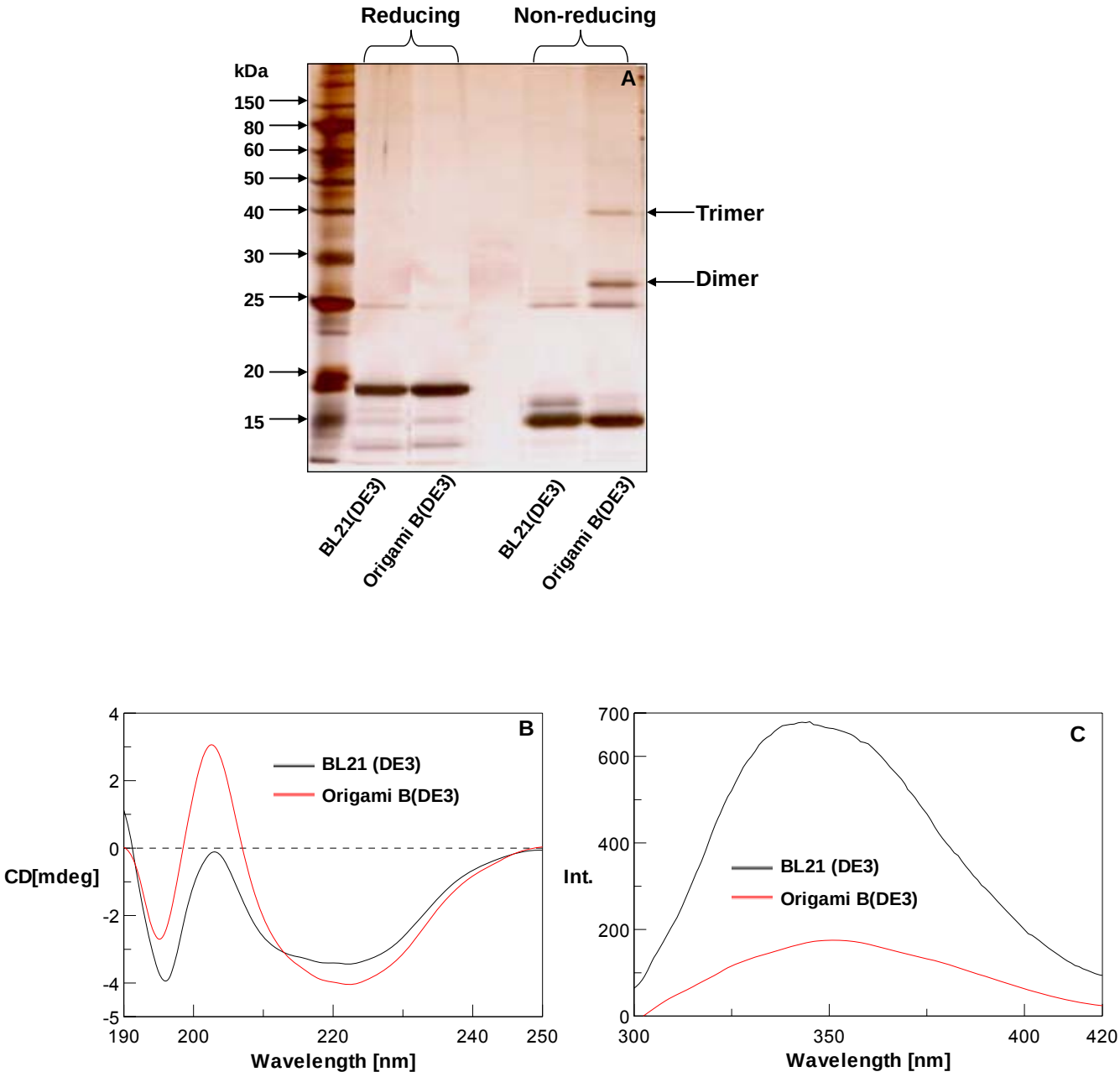


Figure 7.3 Structure analysis of fresh M3 (day-0 samples) produced by two host strains. Panel A: SDS-PAGE analysis. Panel B: far-UV CD spectroscopy; Panel C: fluorescence spectroscopy.

they both displayed a β-sheet structure mixed with type II β-turn based on ~5 nm red-shift of all the featuring bands as compared to β-sheet CD spectra [82]. By comparing the tryptophan fluorescence spectra of the two variants (Figure 7.3C), the maximum emission intensity (at 345 nm) of BL21(DE3)-derived M3 was approximately 4-fold that (at 350 nm) of Origami B(DE3)-derived M3 with a slight red-shift, suggesting that the formation of certain intramolecular and/or intermolecular disulfide bond(s) could result in a major conformational change in the local environment surrounding the tryptophan residues within M3.

The molecular properties of the two M3 variants were compared by monitoring the conformational changes under storage conditions (either -20 °C or a room temperature) and the results are documented in Figure 7.4. While the two protein samples had an identical band pattern upon the analysis with reducing SDS-PAGE, they behaved rather distinctively in terms of disulfide bond formation upon the analysis with non-reducing SDS-PAGE. The monomeric doublet corresponding to BL21(DE3)-derived M3 merged into a single band with a higher electrophoretic mobility during the storage with no multimerization being observed. On the other hand, while the electrophoretic mobility of the monomeric form of Origami B(DE3)-derived M3 remained unchanged, significant multimerization was observed even at a frozen temperature of -20 °C presumably through the formation of intermolecular disulfide bonds. Both M3 variants were subject to degradation when stored at a room temperature.

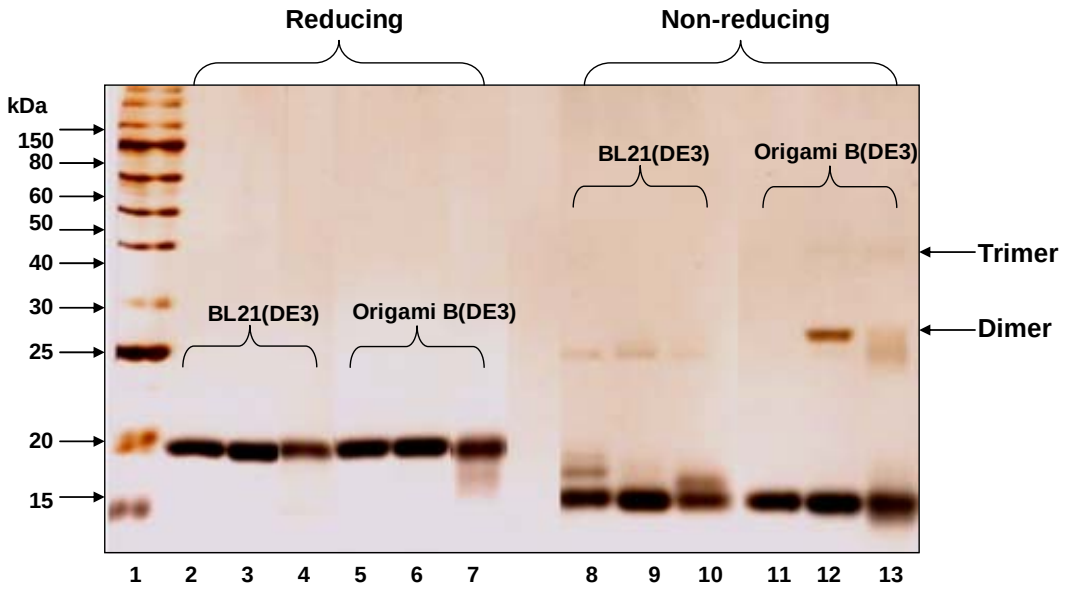


Figure 7.4 Monitoring and comparison of molecular property for M3 produced by two host strains with SDS-PAGE: Samples up to 15-day storage at -20 °C and 10-day storage at room temperature were analyzed. Lane 1/ molecular weight markers; lanes 2, 5, 8, 11/ day 0; lanes 3, 6, 9, 12/ -20 °C samples taken on day 15, lanes 4, 7, 10, 13/ room temperature samples taken on day 10. Arrows indicated protein bands corresponding to disulfide bonded multimeric forms.

7.3.3 Comparison of Molecular Properties for Extended M3 Variants

Molecular properties of the three extended M3 mutant variants, i.e. M13, M135 and M35, were characterized in a similar way to the above and the results are summarized in Figures 7.5~7.8, respectively. While the M13-GST fusion expressed in Origami B(DE3)pLysS had an identical electrophoretic mobility under both reducing and non-reducing conditions (lane

7 vs. 13 in Panels A & B of Figure 7.5), the M135-GST and M35-GST fusions had a higher mobility under a non-reducing condition than those under a reducing condition (lane 7 vs. 13 in Panels C & D and E & F of Figure 7.5), suggesting that certain intramolecular disulfide bond(s) was formed within the M135-GST and M35-GST fusions, but not the M13-GST fusion, upon their expression in Origami B(DE3). Note that the mobility increase for the Origami B(DE3)-derived M135-GST was less than that of Origami B(DE3)-derived M3-GST or M35-GST fusions, implying different intramolecular disulfide bond patterns existed in these fusions.

There were two monomeric forms for M135 derived from BL21(DE3) (Figure 7.6B). Hence, the upper band with a mobility of 20 kDa represents the form with no disulfide bond and the lower band with a mobility of 18 kDa represents the form with a Cys^{35}-Cys^{107} bridge which is the only possible disulfide bond that can be formed for M135. Since this disulfide bond cannot be formed in BL21(DE3), it must be formed *in vitro* during the downstream processing stage. On the other hand, there was only a single monomeric form with a mobility of 18 kDa, representing the form with the Cys^{35}-Cys^{107} bridge, for M135 derived from Origami B(DE3). The enhanced efficiency in disulfide bond formation can be attributed to the *in vivo* oxidizing capacity of Origami B(DE3), which possibly mediated the formation of the Cys^{35}-Cys^{107} bridge when M135-GST was expressed. Interestingly, no dimer was formed in BL21(DE3)-derived M135, whereas dimerization was somewhat observed for Origami B(DE3)-derived M135 (Figure 7.6B) and no progressive in vitro multimerization was observed during a prolonged storage (Figure 7.8B). The results suggest that neither Cys^{35} nor Cys^{107} was directly involved in the *in vitro* formation of intermolecular disulfide bonds for

multimerization. The M135 derived from the two hosts exhibited rather different spectroscopic profiles (Panels C and D of Figure 7.7). However, deconvolution of the CD spectra using the Contin method (DichroWeb: http://dichroweb.cryst.bbk.ac.uk/) only revealed a slight difference in the secondary structure composition of α-helix and β-sheet. The significant signal difference at short wavelengths of 195 nm and 203 nm (Figure 7.7C) can be associated with different levels of the splitting of the peptide ππ* transition and the twist of β-sheet [82]. Compared to BL21(DE3)-derived M135, the fluorescence intensity was significantly quenched for Origami B(DE3)-derived M135, presumably due to the presence of the Cys^{35}-Cys^{107} disulfide bond which could act as an electron scavenger to absorb the electrons released from an excited indole ring of tryptophan [75, 90]. The results suggest that the significant changes in the spectroscopic profiles are primarily associated with the presence of the Cys^{35}-Cys^{107} bridge.

Dimers were observed for both M13 variants, but more abundant for the Origami B(DE3)-derived one (Figure 7.6A). Multimerization was significantly enhanced for M13 compared to M135, implying the critical role that Cys^{129} played during the *in vitro* formation of intermolecular disulfide bonds for multimerization. Compared to M3 (Figure 7.4), M13 appears to be less subject to degradation during the prolonged storage (Figure 7.8A) and the improved structural stability was associated with the Cys^{27} mutation and the aggravated trend for multimerization, particularly for Origami B(DE3)-derived M13. On the other hand, neither multimerization nor improved stability was observed for M35 (Figure 7.6C & 7.8C), supporting both the above arguments that the Cys^{129} was critically involved in the formation

of intermolecular disulfide bonds for multimerization and the Cys^{27} mutation could improve the structural stability.

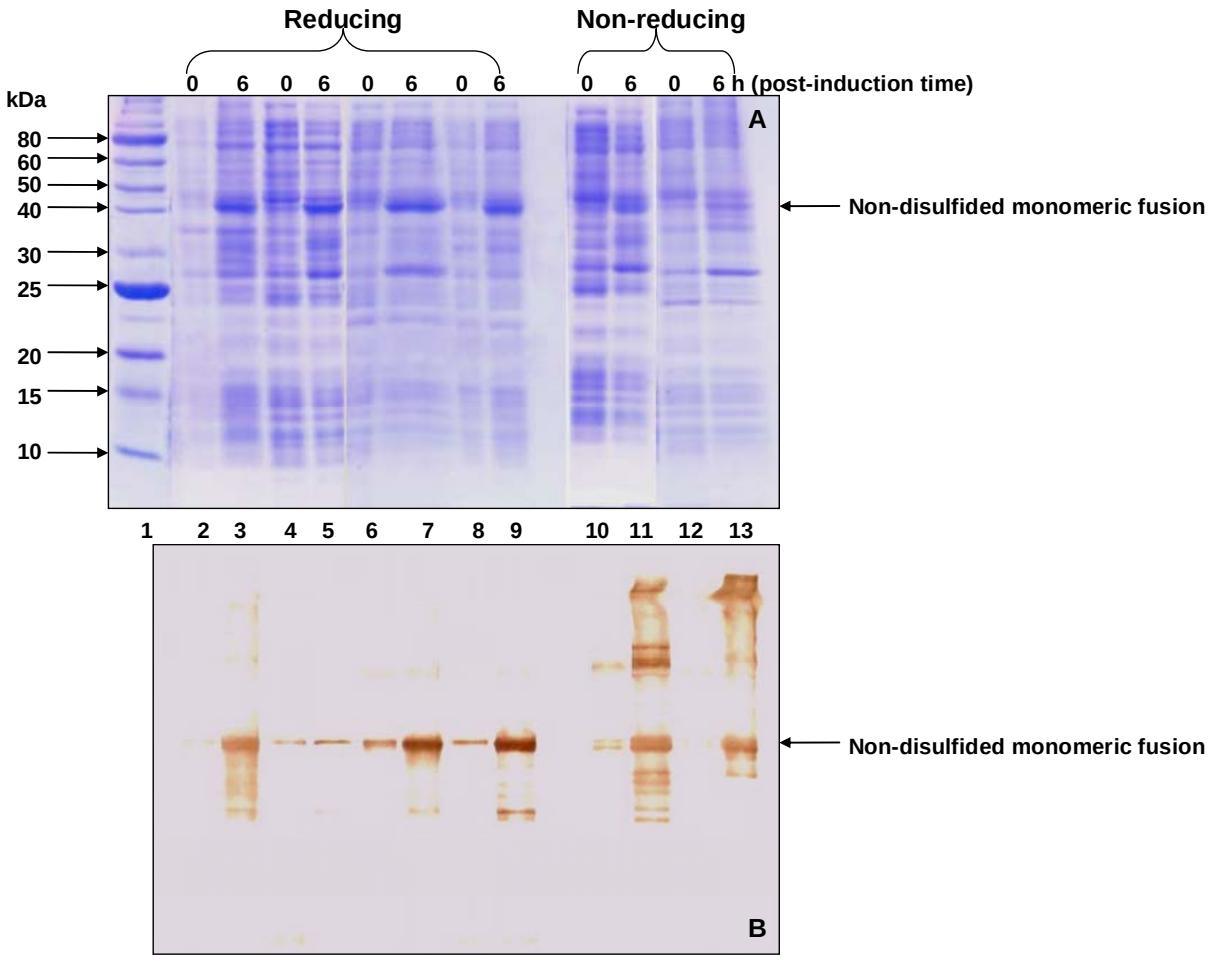


Figure 7.5 Intracellular protein analysis of both soluble and insoluble fractions under reducing and non-reducing conditions with SDS-PAGE (Panels A, C, and E) and Western blotting (Panels B, D, and F) for the culture samples taken prior to induction and at 6 h post-induction time were shown. Panels A & B, C & D, and E & F depicted the culture of M13,

M135, and M35, respectively. Panels A ~ D: lane 1 was molecular weight marker; Lanes 2, 3, 10, 11 and 4, 5 represented soluble and insoluble fractions of samples produced by BL21(DE3)pLysS, respectively, whereas lanes 6, 7, 12, 13 and 8, 9 represented soluble and insoluble fractions of samples produced by Origami B(DE3)pLysS, respectively. Arrows in four panels indicated protein bands corresponding to disulfide bonded and non-disulfide bonded monomeric fusion protein.

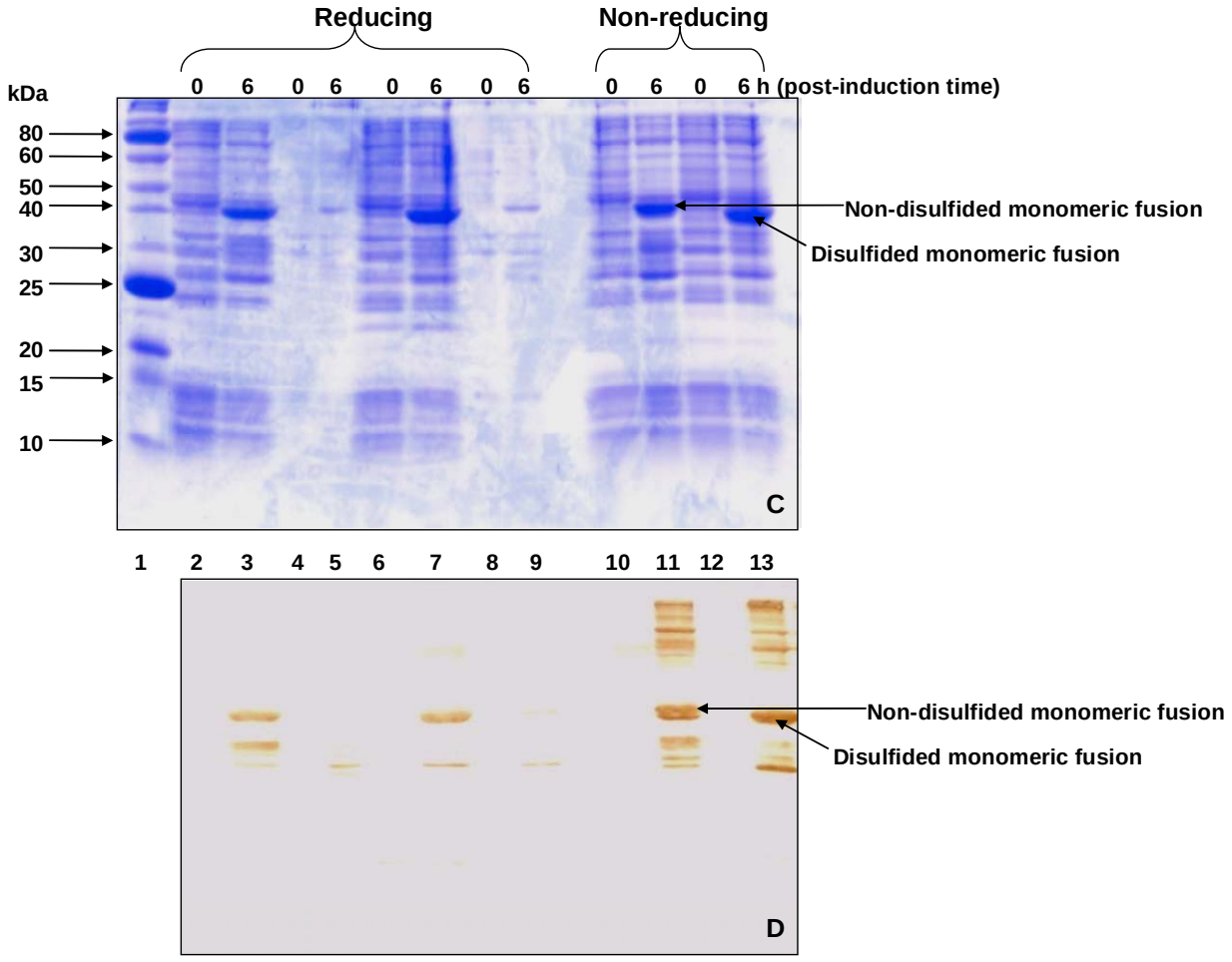


Figure 7.5 (Continued)

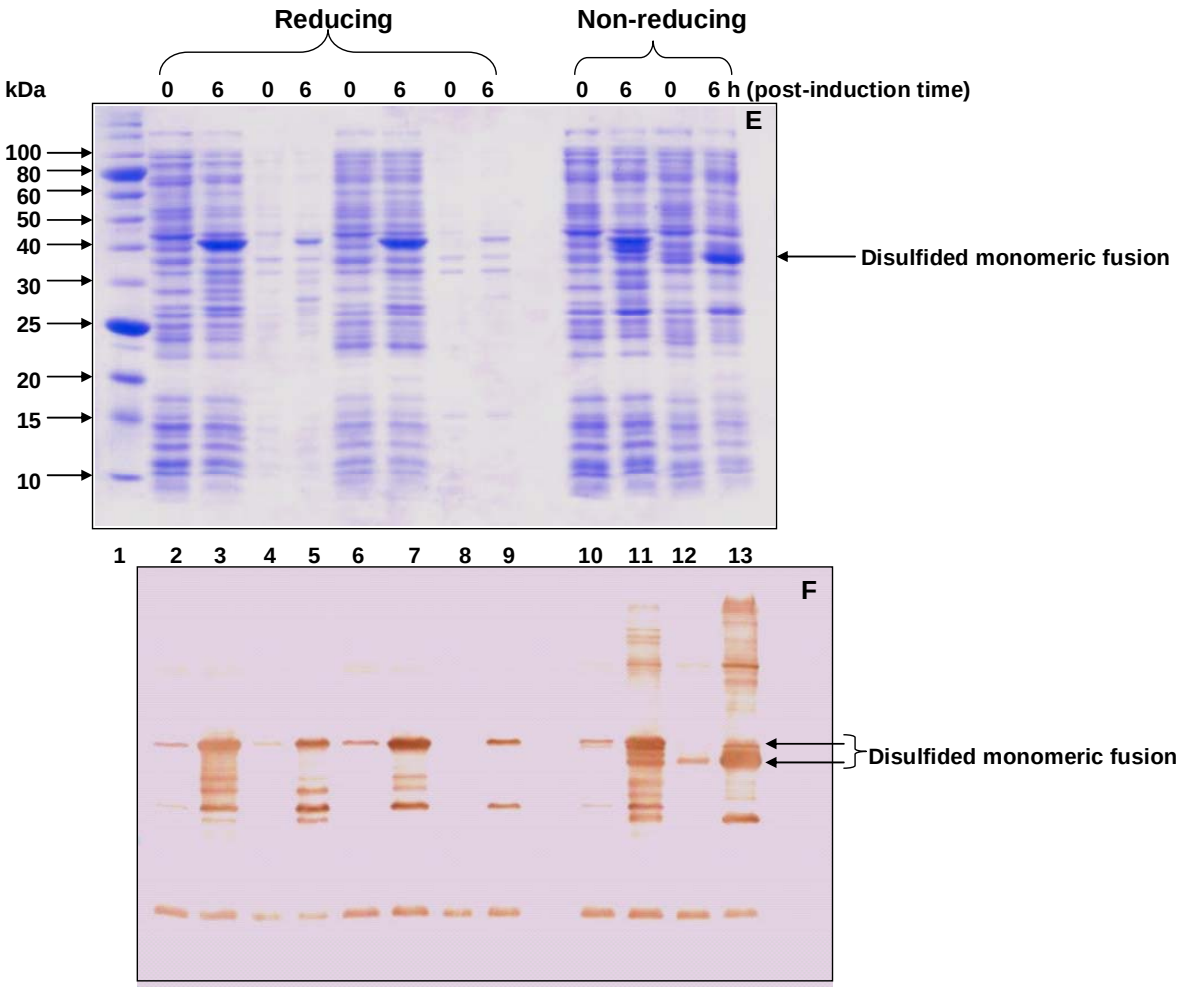


Figure 7.5 (Continued)

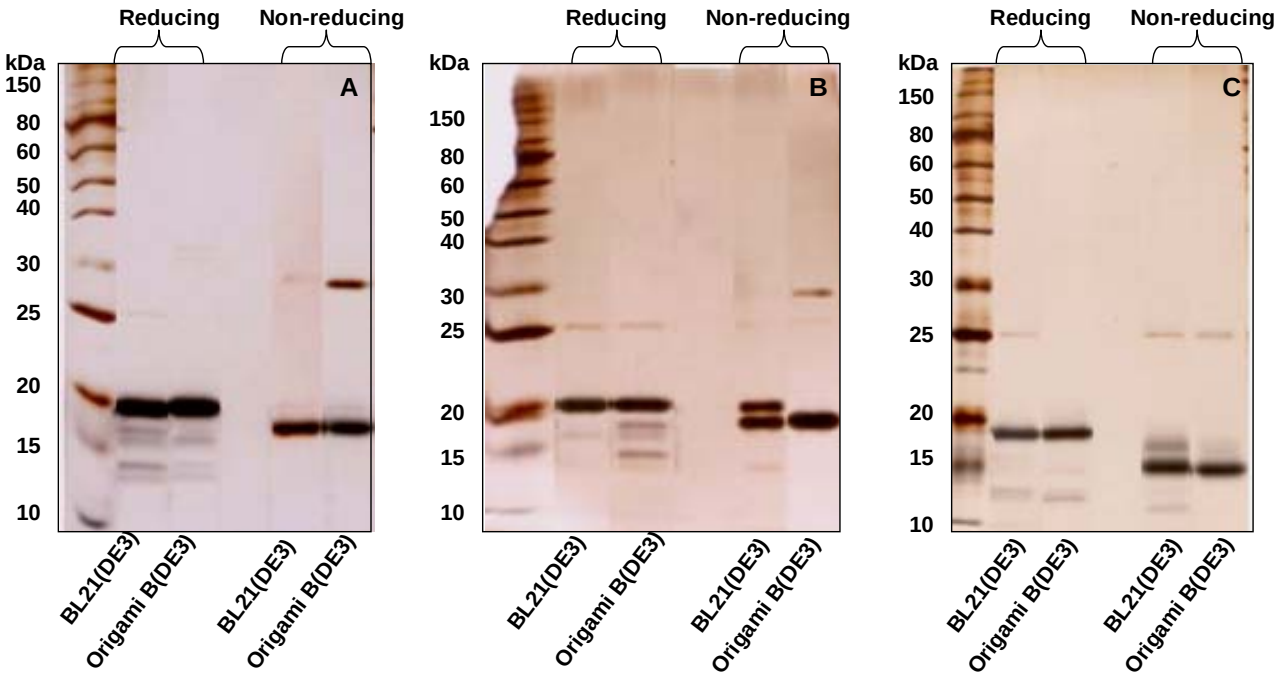


Figure 7.6 Structure analysis of fresh mutant variants produced by two host strains with SDS-PAGE. Panel A: M13; Panel B: M135; Panel C: M35.

7.4 Discussion

Recently, an hCD83ext mutant variant of M3 with the Cys100Ser mutation was identified to outperform the wild-type hCD83ext by exhibiting a consistent bioactivity. Structurally, no multimerization via the aberrant formation of intermolecular disulfide bonds was observed for the BL21(DE3)-derived M3 mutant [101].

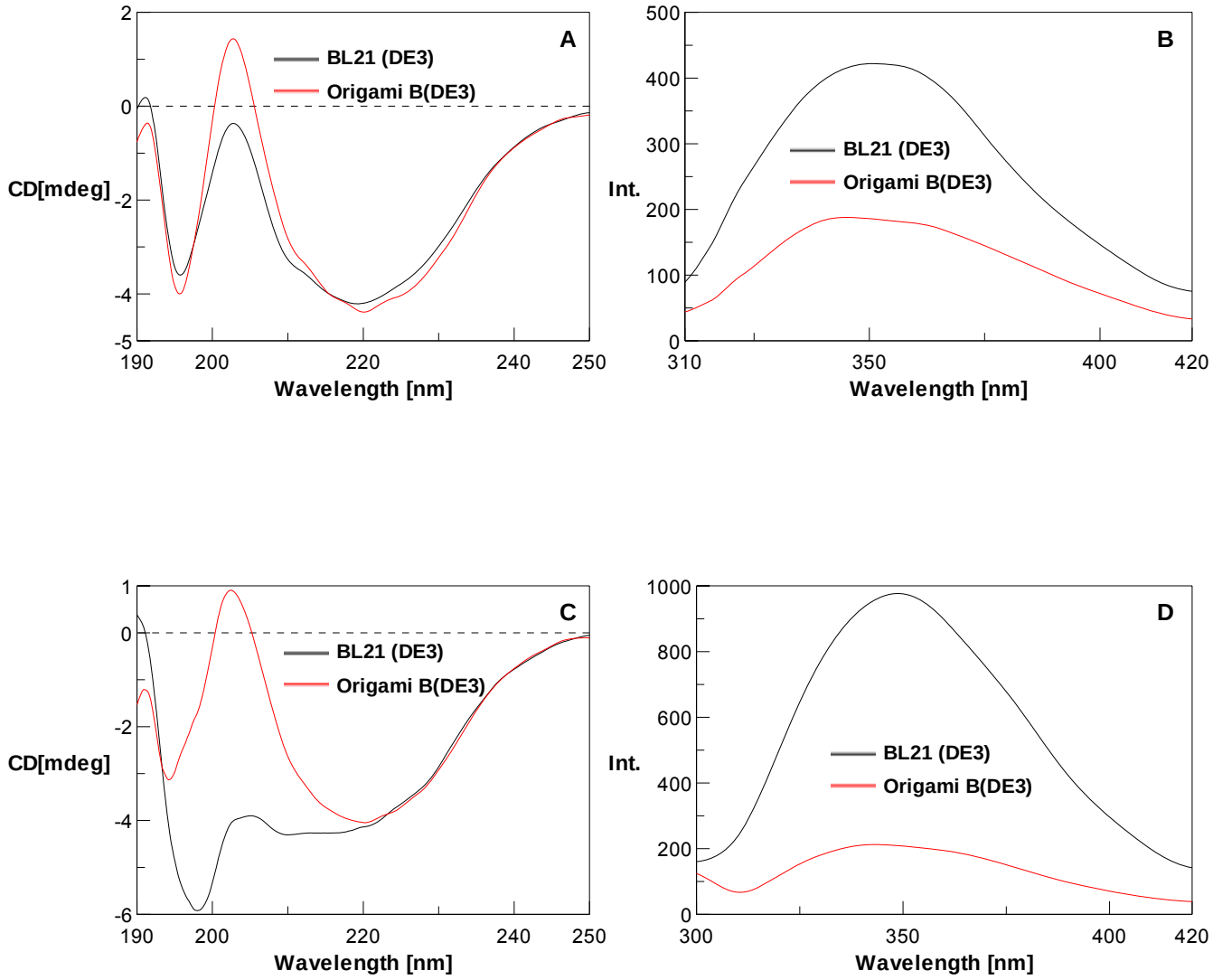


Figure 7.7 Structure analysis of fresh mutant variants produced by two host strains with CD spectroscopy (Panel A, C, and E) and fluorescence spectroscopy (Panel B, D, and F). Panels A & B, C & D, and E & F represented M13, M135 and M35, respectively.

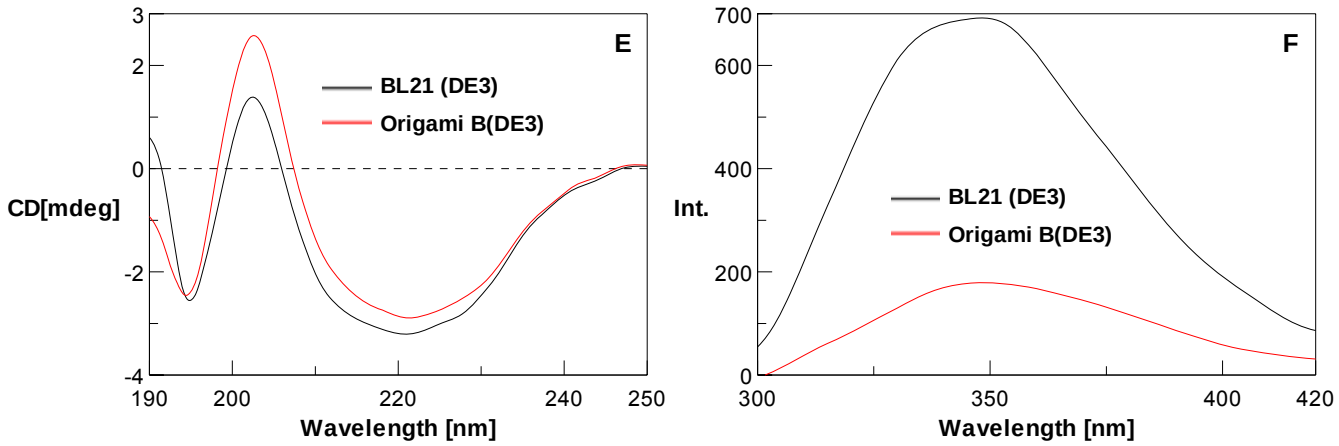


Figure 7.7 (Continued)

In this study, multimerization was observed for the Origami B(DE3)-derived M3 mutant (Figure 3) as well as the M13 mutant derived from both BL21(DE3) and Origami B(DE3) (Figure 8A). The results suggest that Cys^{100} was not directly involved in the aberrant formation of intermolecular disulfide bonds for these mutant variants. We then propose that the Cys^{100} mutation might induce a major conformational change through which the *in vitro* multimerization was prevented. The preventing effect of such a conformational change on the multimerization could be suppressed by the presence of the Cys^{27} mutation or the adoption of Origami B(DE3) as the production host.

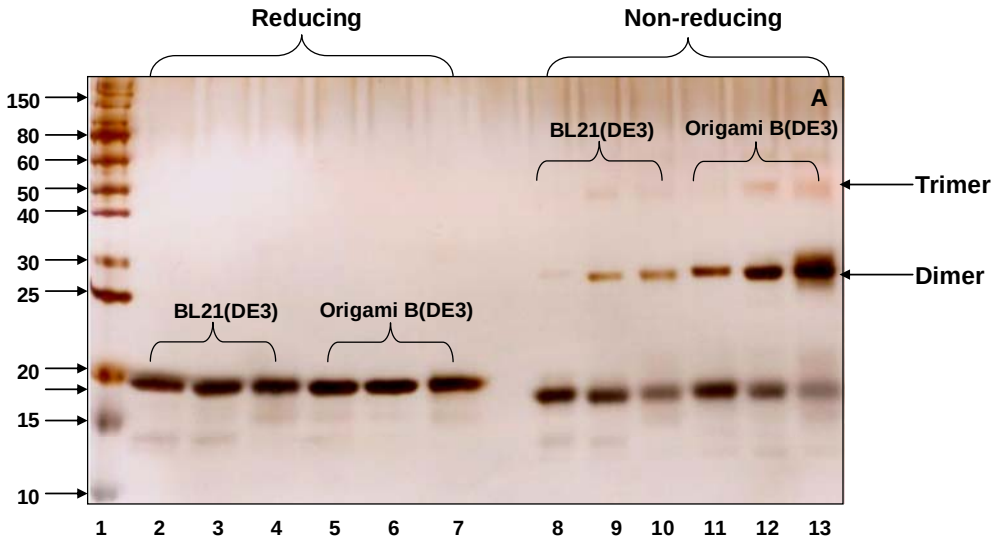


Figure 7.8 Monitoring and comparison of molecular property for mutant variants produced by two host strains with SDS-PAGE: samples up to 20-day storage at -20 °C and room temperature were analyzed. Panels A: M13, lane 1/ molecular weight markers; lanes 2, 5, 8 and 11 / day 0; lanes 3, 6, 9 and 12 / -20 °C samples taken on day 20; lanes 4, 7, 10 and 13 / room temperature samples taken on day 20. Panel B: M135, lane 1/ molecular weight markers; lanes 2, 5, 8 and 11 / day 0; lanes 3, 6, 9 and 12 / -20 °C samples taken on day 20; lanes 4, 7, 10 and 13 / room temperature samples taken on day 20. Panel C: M35, lane 1/ molecular weight markers; lanes 2, 4, 6 and 8 / -20 °C samples taken on day 20; lanes 3, 5, 7 and 9 / room temperature samples taken on day 20. Arrows indicated different multimeric forms of mutant variants.

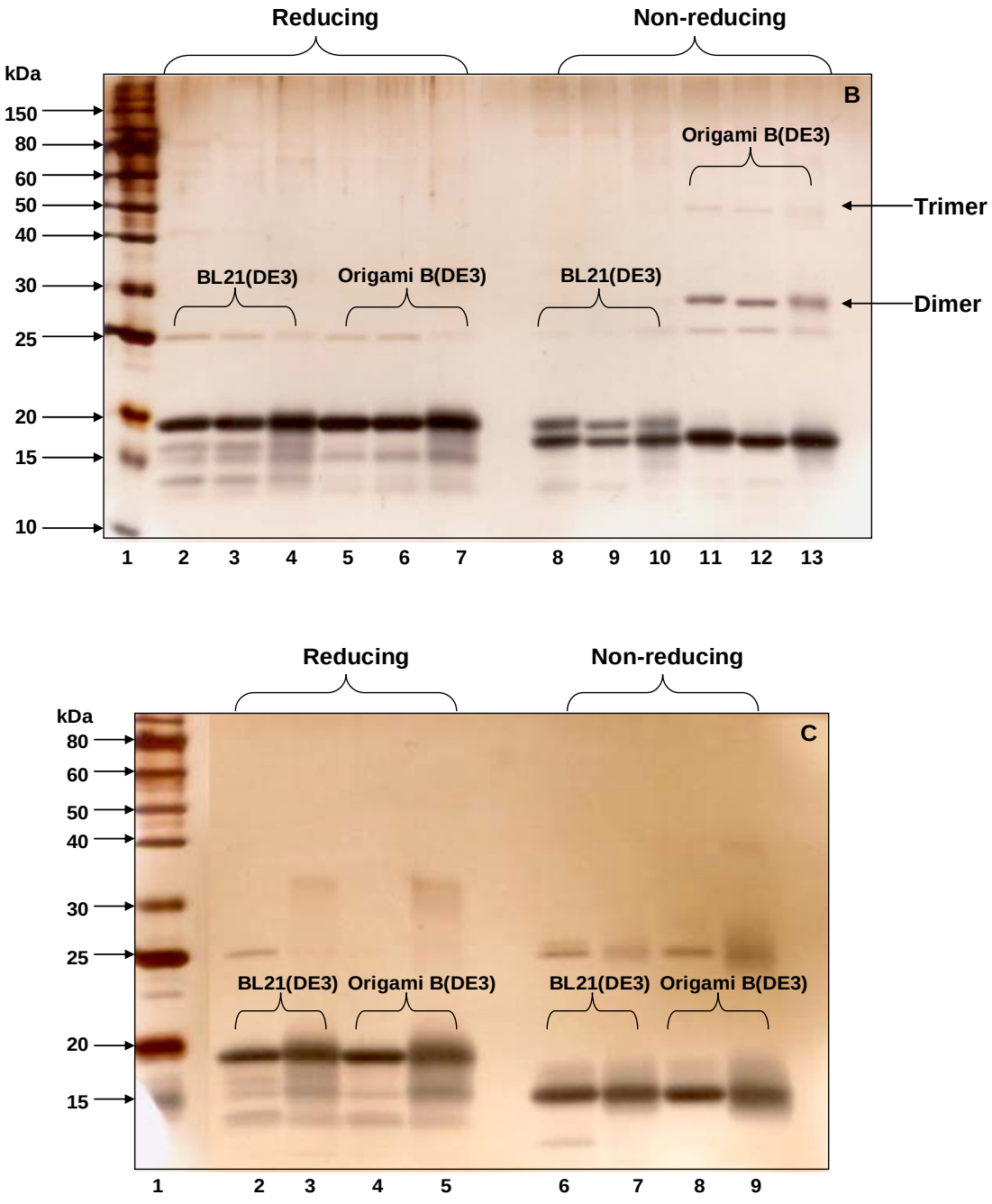


Figure 7.8 (Continued)

We also propose that Cys^{129} could be a primary cysteine residue directly involved in the formation of intermolecular disulfide bonds based on the following observations. First, the presence of the Cys^{27} mutation enhanced the multimerization (i.e. M13 vs. M3), precluding Cys^{27} from being the major bridging cysteine. This observation is also consistent to our previous results (i.e. M1 vs. wild-type hCD83ext) [89]. Second, no multimerization was observed for BL21(DE3)-derived M135 with the presence of two free cysteine residues of Cys^{35} and Cys^{107} (Figure 7.8B), precluding these two as the major bridging cysteine as well. Third, multimerization associated with M13, derived from either BL21(DE3) or Origami B(DE3), disappeared upon the introduction of the Cys^{129} mutation (i.e. M13 vs. M135). In fact, our current proposal that Cys^{129} could act as a primary bridging cysteine was previously reported based on an observation that no multimerization was observed for BL21(DE3)-derived M5 [9]. However, we recently reported a conflicting result, i.e. multimerization can still occur to BL21(DE3)-derived M5, based on a modified analytical protocol to prevent disulfide scrambling through which the intermolecular disulfide bonds could be broken [101]; suggesting that a cysteine residue(s) other than Cys^{129} can also act as the bridging cysteine for multimerization. With the above arguments precluding Cys^{27}, Cys^{35}, and Cys^{107} as the bridging cysteine, Cys^{100} appears to be another possible one. The conclusion is supported by the observation that multimerization associated with M5 was prevented by introducing the Cys^{100} mutation (i.e. M5 vs. M35). The overall results suggest that both Cys^{100} and Cys^{129} can possibly act as a bridging cysteine for *in vitro* multimerization and this conclusion is further supported by another two observations. First, no *in vitro* multimerization was observed for BL21(DE3)-derived or Origami B(DE3)-

derived M35 (Figure 8C). Second, the presence of the Cys^{27} mutation did not suppress the prevention of the multimerization for M35 (i.e. M135 vs. M35), as it did for M3 (i.e. M13 vs. M3).

As discussed above, the presence of the Cys^{27} mutation significantly enhances the multimerization (i.e. M13 vs. M3 and M1 vs. wild-type) presumably using Cys^{129} as the major bridging cysteine. Based on this observation, we hypothesize that the formation of the Cys^{27}-Cys^{129} intramolecular disulfide bond *in vitro* might be thermodynamically or kinetically favorable [particularly in the presence of the Cys^{100} mutation (i.e. BL21(DE3)-derived M3)], thus limiting the availability of free Cys^{129} for multimerization. Introducing the Cys^{27} mutation would free Cys^{129} and, as a result, multimerization enhanced. Also, notice that hCD83ext variants with the Cys^{27} mutation were less subject to degradation [89], implying that the formation of the Cys^{27}-Cys^{129} disulfide bridge might destabilize the protein molecule.

For all the hCD83ext mutant variants investigated in this study, there was a major conformational difference between the BL21(DE3)-derived and Origami B(DE3)-derived products based on the analyses with CD and fluorescence spectroscopy. By comparing the monomeric bands using non-reducing SDS-PAGE for M135 derived from the two hosts (Figure 8B), the conformational change should be associated with the formation of the Cys^{35}-Cys^{107} intramolecular disulfide bond. Though, as proposed above, the formation of the Cys^{35}-Cys^{107} intramolecular disulfide bond might not be thermodynamically or kinetically favorable *in vitro*, it could be rather effective in the oxidative cytoplasm of Origami B(DE3). In addition, the *in vivo* formed Cys^{35}-Cys^{107} intramolecular disulfide bond significantly

enhanced *in vitro* multimerization since Origami B(DE3)-derived products (e.g. M3 and M13) tended to form multimers more effectively than their BL21(DE3)-derived counterparts. The multimerization associated with Origami B(DE3)-derived M3 and M13 was mediated through Cys^{129} since the multimerization was prevented by the presence of the Cys^{129} mutation [i.e. Origami (DE3)-derived M3 vs. Origami B(DE3)-derived M35 and Origami B(DE3)-derived M13 vs. Origami B(DE3)-derived M135].

In summary, the mechanism of disulfide bond formation in hCD8ext was here elucidated. (1) Cys^{100} is critically but indirectly involved in the formation of intermolecular disulfide bonds for multimerization [Origami B(DE3)-derived M3 and two M13 variants develop multimerization]. (2) The presence of the Cys^{100} mutation prevents multimerization. Such prevention can be, however, suppressed by either the *in vivo* formation of the disulfide bond Cys^{35}-Cys^{107} [Origami B(DE3)-derived M3 forms multimers] or by the introduction of the Cys^{27} mutation (two M13 variants form multimers). (3) Cys^{129} is primarily, possibly directly but not exclusively, engaged in the formation of intermolecular disulfide bonds for multimerization (I. M13 variants form dimers whereas M135 variants eliminate dimerization; II M5 forms dimers possibly linked via Cys^{100}). (4) Cys^{27} might possibly form an intramolecular disulfide bond with Cys^{129} to prevent the multimerization which is enhanced by the presence of Cys^{100} mutation [BL21(DE3)-derived M3 eliminates multimerization vs. two M13 variants significantly develop multimerization]. (5) The presence of the Cys^{27} mutation enhances multimerization, possibly via freeing Cys^{129} [I. M1 variants demonstrate increased multimerization compared to the wild-type variants; II. M13 variants display augmented multimerization compared to M3; III, M135 variants eliminate multimerization].

(6) Cys^{35} and Cys^{107} are unlikely to participate in intermolecular disulfide bond formation; particularly in vitro (two M135 variants eradicate dimerization *in vitro*, even in the presence of free Cys^{35} and Cys^{107}). (7) The formation of intramolecular disulfide bond between Cys^{35} and Cys^{107} is peculiarly favored *in vivo* [Origami B(DE3)-derived M135 has 100% Cys^{35}-Cys^{107} bridged monomer vs. BL21(DE3)-derived M135 has an equilibrium between two monomeric isoforms (i.e. monomer with Cys^{35}-Cys^{107} and monomer without Cys^{35}-Cys^{107})]. (8) The *in vivo* formed Cys^{35}-Cys^{107} results in conformational changes that enhance subsequent multimerization *in vitro* [the Origami B(DE3)-derived variants, wild-type, M1, M3, and M13, exhibit higher tendency for multimerization than their BL21(DE3)-derived counterparts].

Chapter 8

Conclusions and Recommendations

8.1 Conclusions

In this thesis, structural properties of hCD83ext produced by the developed bioprocess were studied. A series of cysteine mutant variants of hCD83ext were constructed to investigate the disulfide bond formation in this recombinant therapeutics. The following can be concluded:

1. The disulfide bond formation significantly affects the structural property and biological activity of hCD83ext. While the established bioprocess produces hCD83ext product with high purity and low endotoxin level, inconsistent biological activity remains an issue, which primarily resulted from the lack of control on consistent disulfide bond formation. Therefore, modulation of disulfide bond formation is absolutely necessary in the production and extensive storage of the protein.
2. Aberrant intermolecular disulfide bond formation in the hCD83ext should be prevented to display consistent biological activity. Cys^{129} is the one primarily and directly involved in intermolecular disulfide bond formation. However, the impacts of Cys^{100} and Cys^{27} on intermolecular disulfide bond formation are very important, although they do not tend to participate directly.
3. The strategy of using *E. coli* Origami strain as a host to facilitate disulfide bond formation *in vivo* is proven not helpful for hCD83ext production. Aberrant disulfide

bond formation occurs *in vivo*, and subsequently aggravates the intermolecular disulfide bond formation in hCD83ext and its mutant variants. On the other hand, rational site-directed mutagenesis on cysteine residues is a more attractive approach. hC83ext mutant variant M3 with Cys100Ser was identified as an improved version of the therapeutic since it consistently displayed excellent biological activity.

4. The storage stability of hCD83ext can be altered by manipulating disulfide bond formation. Mutating Cys^{27} to Ser or Glu significantly enhanced the stability.
5. Environment significantly affects the disulfide bond formation in hCD83ext, reflected by the differences constantly demonstrated between *in vivo* and *in vitro* disulfide bond formation in hCD83ext and its mutant variants.

8.2 Recommendations

The thesis presents some new understanding of the structural properties of hCD83ext. The following is recommended for future work to further improve the performance of production.

1. The mechanism underlying the elimination of multimerization in BL21(DE3)-derived M3 was assumed to be the formation of intramolecular disulfide bond Cys^{27}-Cys^{129}. To solidly prove this assumption, effective methodology for disulfide bond assignment needs to be developed. Further, disulfide bond patterns that are possibly beneficial for biological activity development needs to be elucidated. In addition, the application of this methodology to the wild-type hCD83ext and the M1 will help to reveal the mechanism by which Cys^{27} mutation stabilizes hCD83ext.

2. hCD83ext derived by BL21 and Origami strains exhibited distinctive molecular properties commitment with *in vivo* disulfide bond formation, implying that hCD83ext might undergo a disulfide-coupled oxidative folding pathway. Therefore, reductive unfolding followed by oxidative folding may be good to carry out for hCD83ext and its mutant variants to learn more details on disulfide bond formation of hCD8ext and the role of each cysteine residue. The investigation of folding pathway might shed light on the development of effective strategies to promote consistent disulfide bond formation in hCD83ext.
3. Although oxidative cytoplasmic expression of the fusion GST-hCD83ext was not helpful in consistent disulfide bond formation, it is worthy of developing periplasmic expression or extracellular secretion of hCD83ext in *E. coli*. The native disulfide bond formation pathway might better facilitate disulfide bond formation in hCD83ext. Furthermore, a wide variety of well-established strategies can be adopted to improve the expression.

Appendix A

DNA sequencing results and alignment

The DNA sequence alignment was conducted using NCBI BLASTN program accessible at http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome.

The Query is the sequence of constructed hCD83ext mutant obtained from DNA sequencing. The Sbjct (Subject) is the sequence of wild-type hCD83ext. The designated nucleotidic mutations are highlighted in green. The mutation automatically arising during PCR cycles is highlight in yellow, which is a silent mutation.

Aa1. Alignment of the DNA sequences of hCD83ext_mut_C27S and the hCD83ext

```
Query 43  AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC 102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 378 AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC 319

Query 103 TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT 162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 318 TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT 259

Query 163 GTATGTCCCCGAGTTGCAGCTGGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 258 GTATGTCCCCGAGTTGCAGCTGGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT 199

Query 223 GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCCCTGAGGTGGTC 282
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct 198 GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCTCTGAGGTGGTC 139

Query 283 TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 138 TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT 79
```

```
Query 343 GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGAGCTCGC 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Sbjct 78  GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGAGCAAGC 19

Query 403 CACCTTCACCTCCGGCGT 420
          ||||||||||||||||||
Sbjct 18  CACCTTCACCTCCGGCGT 1
```

Aa2. Alignment of DNA sequences of hCD83ext_mut_C27E and hCD83ext

```
Query 39  AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC 98
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 378 AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC 319

Query 99  TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT 158
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 318 TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT 259

Query 159 GTATGTCCCCGAGTTGCAGCTGGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT 218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 258 GTATGTCCCCGAGTTGCAGCTGGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT 199

Query 219 GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCCCTGAGGTGGTC 278
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct 198 GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCTCTGAGGTGGTC 139

Query 279 TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT 338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 138 TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT 79

Query 339 GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGACTCAGC 398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Sbjct 78  GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGAGCAAGC 19
```

```
Query 399 CACCTTCACCTCCGGCGT 416
          ||||||||||||||||||
Sbjct 18  CACCTTCACCTCCGGCGT 1
```

Aa3. Alignment of the DNA sequences of hCD83ext_mut_C100S (M3) and hCD83ext

```
Query  44   AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC  103
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  378  AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC  319

Query  104  TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT  163
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  318  TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT  259

Query  164  GTATGTCCCCGAGTTGGAGCTCGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT  223
            ||||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||
Sbjct  258  GTATGTCCCCGAGTTGCAGCTGGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT  199

Query  224  GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCCCTGAGGTGGTC  283
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  198  GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCTCTGAGGTGGTC  139

Query  284  TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT  343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  138  TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT  79

Query  344  GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGAGCAAGC  403
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  78   GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGAGCAAGC  19

Query  404  CACCTTCACCTCCGGCGT  421
            ||||||||||||||||||
Sbjct  18   CACCTTCACCTCCGGCGT  1
```

Aa4. Alignment of the sequenced DNA of hCD83ext_mut_C27S_C100S (M13) and the DNA of hCD83ext

```
Query 43  AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC 102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 378 AATCTCCGCTCTGTATTTCTTAAAAGTCTCTTCTTTACGCTGTGCAGGGCATCCTGTCAC 319

Query 103 TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT 162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 318 TCTCAAGATCACCTTGCCACTTAGGTTTCTCTGCCCATCCGGGTCCTGCAGAGTGCACCT 259

Query 163 GTATGTCCCCGAGTTGGAGCTCGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT 222
          ||||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||
Sbjct 258 GTATGTCCCCGAGTTGCAGCTGGTAGTGTTTCGGATCTTCAGGGAATAGGGCCTTTCATT 199

Query 223 GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCCCTGAGGTGGTC 282
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct 198 GGGGGCGTCGAAAGAACCATTTTGCCCCTTCTGATGATAGTGCTGTCCTCTGAGGTGGTC 139

Query 283 TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 138 TTCCTGGGGTGTCTCCATCCTCTCTTCACCACCCTCCAATAACTTGACCCAGGAGACCGT 79

Query 343 GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGAGCTCGC 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Sbjct 78  GTAGGGAACCTGCGGATCCCAGGGGGCGGTGCAGGGCAAGTCCACATCTTCGGAGCAAGC 19

Query 403 CACCTTCACCTCCGGCGT 420
          ||||||||||||||||||
Sbjct 18  CACCTTCACCTCCGGCGT 1
```

Aa5. Alignment of the DNA sequences of hCD83ext_mut_C27S_C100S_C129S (M135) and hCD83ext

```
Query 1   ACGCCGGAGGTGAAGGTGGCTTGCTCCGAAGATGTGGACTTGCCCTGCACCGCCCCCTGG 60
          ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Sbjct 420 ACGCCGGAGGTGAAGGTGGCGAGCTCCGAAGATGTGGACTTGCCCTGCACCGCCCCCTGG 361

Query 61  GATCCGCAGGTTCCCTACACGGTCTCCTGGGTCAAGTTATTGGAGGGTGGTGAAGAGAGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 360 GATCCGCAGGTTCCCTACACGGTCTCCTGGGTCAAGTTATTGGAGGGTGGTGAAGAGAGG 301

Query 121 ATGGAGACACCCCAGGAAGACCACCTCAGAGGACAGCACTATCATCAGAAGGGGCAAAAT 180
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct 300 ATGGAGACACCCCAGGAAGACCACCTCAGGGGACAGCACTATCATCAGAAGGGGCAAAAT 241

Query 181 GGTTCTTTCGACGCCCCCAATGAAAGGCCCTATTCCCTGAAGATCCGAAACACTACCAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct 240 GGTTCTTTCGACGCCCCCAATGAAAGGCCCTATTCCCTGAAGATCCGAAACACTACGAGC 181

Query 241 TGCAACTCGGGGACATACAGGTGCACTCTGCAGGACCCGGATGGGCAGAGAAACCTAAGT 300
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 180 TCCAACTCGGGGACATACAGGTGCACTCTGCAGGACCCGGATGGGCAGAGAAACCTAAGT 121

Query 301 GGCAAGGTGATCTTGAGAGTGACAGGATGCCCTGCACAGCGTAAAGAAGAGACTTTTAAG 360
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct 120 GGCAAGGTGATCTTGAGAGTGACAGGATCCCCTGCACAGCGTAAAGAAGAGACTTTTAAG 61
Query 361 AAATACAGAGCGGAGATT 378
          ||||||||||||||||||
Sbjct 60  AAATACAGAGCGGAGATT 43
```

Appendix B

Metropolis-Hastings algorithm used for parameter estimation of dimerization reactions of mutant variants hCD83ext_mut_C27E and hCD83ext_mut_C27S

Kinetic modeling often requires parameter estimates by some statistical methods. Metropolis-Hastings algorithm, one of Markov Chain Monte Carlo (MCMC) methods, has been found to be a very powerful tool, especially when some prior knowledge about the parameters is available. This method is applicable to any set of observations that conform to a Markov chain. The algorithm has been employed to solve difficult, non-linear parameter estimation problems arising in various disciplines. It involves one parameter at a time sampling and iteration to find the posterior parameter distribution.

Let us assume that $y_1, y_2, y_3 \ldots y_{n-1}, y_n$ are independent observations that have been collected in a temporal or physical space. These observations are said to form a Markov chain if the following conditional probability is obeyed:

$$P(y_n | y_1, y_2, \ldots y_{n-2}, y_{n-1}) = P(y_n | y_{n-1}) \quad \text{(Ab-1)}$$

Thus the probability of obtaining the current observation (i.e. observation "n") is conditional on only the previous (i.e. penultimate) observation ("n-1"). A number of data sets resulting from natural processes such as growth, death, reaction products, etc., are in fact Markov

chain observations. The Metropolis-Hastings algorithm is then fully incorporated into the LAKEVIEW program.

A set of parameters θ (i.e. θ_1, θ_2, θ_3 etc.) that needs to be estimated given the set of observations, y. The Bayesian inference of model parameters θ given the observed data involves the joint posterior probability distribution conditional on the data P(θ | y). According to Bayes' theorem, the joint probability of θ given y is:

$$P(\theta \mid y) \propto P(\theta)\, P(y \mid \theta) \qquad \text{(Ab-2)}$$

where:

P(θ) = prior distribution of the parameters

P(y | θ)= likelihood function conditional on θ.

The calculation procedure is based on the Gibbs sampling version of the Metropolis-Hastings algorithm. The Metropolis-Hastings algorithm for parameter estimation can be summarized as follows:

a) Assign prior probability distribution $J(\theta_n \mid \theta_{n-1})$ for each parameter to be estimated.
b) Obtain $\theta_{(0)}$ as the initial realization of the parameter vector.
c) Sample the proposed distribution $J(\theta_n \mid \theta_{n-1})$ to obtain parameter value θ*.
d) Calculate the decision variable (r).

$$r = \frac{P(\theta^* \mid y) / J(\theta^* \mid \theta_{n-1})}{P(\theta_{n-1} \mid y) / J(\theta_{n-1} \mid \theta^*)} \qquad \text{(Ab-3)}$$

set $\theta_n = \theta^*$ with probability min (r, 1)

$\theta_n = \theta_{n-1}$ otherwise

e) Continue obtaining realizations of θ_n for n = 1, 2, 3, N until $P(\theta_n \mid y)$ converges to $P(\theta \mid y)$.

In the LAKEVIEW model, each parameter to be estimated is assigned a triangular *a priori* distribution, $J(\theta_n|\theta_{n-1})$, containing the maximum, mean (most likely) and the minimum possible values. The mean value is established either by graphical means or "eyeballing" the experimental data. The obvious advantage of choosing a symmetric prior distribution such as the triangular distribution is that $J(\theta^*|\theta_{n-1}) = J(\theta_{n-1}|\theta^*)$. Consequently "*r*" reduces to:

$$r = \frac{P(\theta^* \mid y)}{P(\theta_{n-1} \mid y)} \qquad \text{(Ab-4)}$$

The parameter distribution space is sampled and a given parameter value is assigned by a Monte Carlo draw. Using the sampled value, the predicted concentrations are computed by integrating the differential equations. The predictions are then compared with measurements by calculating the normalized sum of squares between the observations and predictions. In LAKEVIEW, the normalized squared differences are calculated between predictions and observations for the concentration of monomer. The parameter sampling procedure is

repeated a number of times. Both the new parameter and its value are chosen by random draw using the Monte Carlo technique.

The objective of the parameter estimation is to minimize the sum of squared differences between predicted and observed dependent variables. A "benchmark" conditional model probability is calculated after substituting the expected parameter values $\theta_{mean} = E(\theta_{Ii})$ into the differential equations and integrating to give the predicted dependent variable $y(\theta_{mean},t)$:

$$p\left(y|\theta_{mean}\right) \propto e\exp - \frac{\Sigma\left[y|\left(\theta_{mean}, t_k\right) - y(t_k)\right]^2}{2\, n_t\, \sigma^2} \qquad \text{(Ab-5)}$$

where:

$p(y\|\theta_{,mean})$	=	conditional "baseline" model probability
$y(\theta_{mean},t_k)$	=	model prediction using expected parameter values at time t_k
$y(t_k)$	=	observation taken at time t_k
n_t	=	total number of observations
σ^2	=	variance

Using Bayesian inference, the posterior probability is given by:

$$P(\theta_{mean} \mid y) \quad \propto p(\theta_{mean}) \times p(y|\theta_{mean}) \qquad \text{(Ab-6)}$$

where:

$P(\theta_{mean} \mid y)$ = "benchmark" predictive model probability for a given set of observations, y_k

$p(\theta_{,mean})$ = the probability of selecting parameter set $\theta_{,mean}$

It is evident that minimizing the sum of squared differences between predictions and observations implies maximizing the probability $P(\theta_{mean} \mid y)$. The posterior probability is re-calculated after each parameter selection.

The sampling space consists of *n* trials and a maximum of *m* iterations per trial. This leads to a maximum of *n* x *m* parameter selection. The number of trials (the value of "*n*") should be at least 100 resulting in a set (θ_1, θ_2, θ_3 etc.) of 100 parameter estimates. The number of iterations is based on the number of parameters to be estimated. As a rule of thumb, the number of draws should be about 10 for each parameter. Thus, the total number of iterations per trial, *m*, should be about tenfold the number of parameters to be estimated. The first trial commences with the calculation of the benchmark (or baseline probability) with the expected parameter values, $E(\theta_i)$-s. Then the selection of the parameter, say parameter "i" and its value (θ_i) is performed by random draw as described above. Model probability with the selected value, θ_i is given as:

$$P(\theta_i \mid y) = p(\theta i \mid \delta_{i,} \theta) \times p(y \mid \theta_i, \theta) \qquad \text{(Ab-7)}$$

where:

$$p\left(y|\theta_i,\theta\right) \propto e\exp - \frac{\Sigma\left[y|(\theta_i,\theta,t_k) - y(t_k)\right]^2}{2\,n_t\,\sigma^2} \qquad \text{(Ab-8)}$$

where:

θ_I = most recently drawn parameter value

It should be noted that the Metropolis-Hastings algorithm is quite flexible and the posterior probability function (equation (Ab-5)) can also be defined as:

$$p\left(y|\theta_i,\theta\right) \propto e\exp - \frac{\Sigma\left[\ln\dfrac{y|(\theta_i,\theta,t_k)}{y(t_k)}\right]^2}{2\,n_t\,\sigma^2} \qquad \text{(Ab-9)}$$

If the probability min{r, 1} for equation (Ab-4) is satisfied with respect to the benchmark value in trial 1, then $P(\theta_i|\ y)$ becomes a "first" estimate of a "global probability maximum", $P(\theta^g\,|y)$.

In applying the Metropolis-Hastings algorithm for parameter estimation in the present case, 9,000 to 30,000 trials were performed using each set of data. The first 1/3 of the trials (3,000 – 10,000) were used for burn-in. During this burn-in period model variance was estimated in the following manner:

$$\sigma^2 \approx s^2 = \frac{\Sigma\left[y|(\theta_i,\theta,t_k) - y(t_k)\right]^2}{n_t} \qquad \text{(Ab-10)}$$

Parameter acceptance and storage begins after this initial trial period. After storing the accepted parameter values, the parameters are reset to their expectations ($E(\theta)$) and a new trial begins. Alternatively, the selected parameter values are replaced and the draw may continue until the number of trials, n, is exceeded. At this point the parameters are also reset to $E(\theta)$-s and a new trial begins. In subsequent trials, if the probability $P(\theta_i| y)$ is greater than the benchmark probability, but less than the current global probability, then the drawn parameter θ_i is retained and is used in subsequent iterations within the trial as a starter value. If, however, the probability $P(\theta_I | y)$ is less than the benchmark probability then θ_i is replaced by previous estimate. An accelerated search technique allows the parameters corresponding to interim $P(\theta |y)$ to become the benchmark parameters (new $E(\theta_i)$-s by centering the triangular distribution on the best current estimate of θ^g) after a certain number of unsuccessful trials (i.e. trials in which the number of iterations exceeds "m"). This is performed once after a number of "training" trials (1/3 of the total number of trials). In this case, the "bins" for storing the accepted θ-s is established after the initial number of "training" trials. If needed, the triangular parameter range (i.e. θ_{max} and θ_{min}) is adjusted to maintain symmetry. This adjustment becomes necessary whenever the current best parameter value approaches but cannot be less than zero.

A difficult problem remains the need to estimate the marginal posterior density for the parameter set θ:

$$P(\theta| y(t_k)\} = \int p\{\theta| \sigma,\delta, y(t_k)) \, d\delta \qquad \text{(Ab-11)}$$

It is well known that the posterior density for θ possesses a multivariate normal distribution, but the performance of the integration analytically is far too difficult. An acceptable empirical solution is to save the set of θ-s after the training trials. The argument is that, if the maximum trial number "*n*" is chosen large enough, then the parameters have converged (or nearly so) to their best estimate of the variance may be calculated empirically as follows:

$$\sigma_i^2 = \frac{\sum_{l} (\theta_{i,j} - \theta_{i,mean})^2}{\ell - 1} \qquad \text{(Ab-12)}$$

where:

σ_i^2 = variance of parameter "i"

$\theta_{i,j}$ = the parameter value of type "i" in the j th data set (η-k <= j <=η)

$\theta_{i,mean}$ = the average value of parameter "i"

ℓ = total number accepted of parameter sets in nxm-*l to "nxm"* number of tries

References

[1] F. Baneyx, M. Mujacic, Recombinant protein folding and misfolding in *Escherichia coli*. Nat. Biotechnol. 22 (2004) 1399-1408.

[2] A. Marchler-Bauer, S.H. Bryant, CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32 (2004) W327-W331.

[3] G. Walsh, Biopharmaceutical benchmarks 2006. Nat. Biotechnol. 24 (2006) 769-776.

[4] M. Lechmann, D. Krooshoop, D. Dudziak, E. Kremmer, C. Kuhnt, C.G. Figdor, G. Schuler, A. Steinkasserer, The extracellular domain of CD83 inhibits dendritic cell-mediated T cell stimulation and binds to a ligand on dendritic cells. J. Exp. Med. 194 (2001) 1813-1821.

[5] E. Zinser, M. Lechmann, A. Golka, M.B. Lutz, A. Steinkasserer, Prevention and treatment of experimental autoimmune encephalomyelitis by soluble CD83. J. Exp. Med. 200 (2004) 345-351.

[6] A. Munoz, Press Release-Durham, NC-April 26, 2006-Argos Therapeutics, Beckman Coulter sign license agreement. CD83 licensure to advance work on autoimmune disorders, transplant rejection therapy. http://www.argostherapeutics.com/pdfs/CD83argos_beckman_press_release.pdf (2006).

[7] Y. Xu, L. Zhang, W. Yao, S.S. Yedahalli, S. Brand, M. Moo-Young, C.P. Chou, Bioprocess development for production, purification, and structural characterization of recombinant hCD83ext as a potential therapeutic protein. Protein Expr. Purif. 65 (2009) 92-99.

[8] M. Lechmann, E. Kremmer, H. Sticht, A. Steinkasserer, Overexpression, purification, and biochemical characterization of the extracellular human CD83 domain and generation of monoclonal antibodies. Protein Expr. Purif. 24 (2002) 445-452.

[9] M. Lechmann, N. Kotzor, E. Zinser, A.T. Prechtel, H. Sticht, A. Steinkasserer, CD83 is a dimer: Comparative analysis of monomeric and dimeric isoforms. Biochem. Biophys. Res. Commun. 329 (2005) 132-139.

[10] G. Walsh, Biopharmaceuticals: biochemistry and biotechnology, John Wiley & sons, Ltd, Chichester, 2003.

[11] J.R. Swartz, Advances in *Escherichia coli* production of therapeutic proteins. Curr. Opin. Biotechnol. 12 (2001) 195-201.

[12] R.R. Burgess, Refolding Solubilized Inclusion Body Proteins, Guide to Protein Purification, Second Edition, ELSEVIER ACADEMIC PRESS INC, San Diego, 2009, pp. 259-282.

[13] E.D. Clark, Protein refolding for industrial processes. Curr. Opin. Biotechnol. 12 (2001) 202-207.

[14] G. Walsh, R. Jefferis, Post-translational modifications in the context of therapeutic proteins. Nat. Biotechnol. 24 (2006) 1241-1252.

[15] F.M. Wurm, Production of recombinant protein therapeutics in cultivated mammalian cells. Nat. Biotechnol. 22 (2004) 1393-1398.

[16] T.U. Gerngross, Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. Nat. Biotechnol. 22 (2004) 1409-1414.

[17] F.R. Schmidt, Recombinant expression systems in the pharmaceutical industry. Appl. Microbiol. Biotechnol. 65 (2004) 363-372.

[18] W.W. Metcalf, W.H. Jiang, B.L. Wanner, Use of the Rep Technique for Allele Replacement to Construct New Escherichia-Coli Hosts for Maintenance of R6k-Gamma Origin Plasmids at Different Copy Numbers. Gene 138 (1994) 1-7.

[19] S.C. Makrides, Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiol. Rev. 60 (1996) 512-538.

[20] J. Arnau, C. Lauritzen, G.E. Petersen, J. Pedersen, Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr. Purif. 48 (2006) 1-13.

[21] J.H. Choi, S.Y. Lee, Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl. Microbiol. Biotechnol. 64 (2004) 625-635.

[22] A.P. Pugsley, The complete gereral secretory pathway in gram-negative bacteria. Microbiol. Rev. 57 (1993) 50-108.

[23] K.J. Jeong, S.Y. Lee, Excretion of human beta-endorphin into culture medium by using outer membrane protein F as a fusion partner in recombinant Escherichia coli. Appl. Environ. Microbiol. 68 (2002) 4979-4985.

[24] E.W.M. Wan, F. Baneyx, TolAIII co-overexpression facilitates the recovery of periplasmic recombinant proteins into the growth medium of Escherichia coli. Protein Expr. Purif. 14 (1998) 13-22.

[25] L.A. Fernandez, I. Sola, L. Enjuanes, V. de Lorenzo, Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system. Appl. Environ. Microbiol. 66 (2000) 5024-+.

[26] Y.Y. Li, C.X. Chen, B.U. von Specht, H.P. Hahn, Cloning and hemolysin-mediated secretory expression of a codon-optimized synthetic human interleukin-6 gene in Escherichia coli. Protein Expr. Purif. 25 (2002) 437-447.

[27] S.Y. Lee, High cell density culture of *Escherichia coli*. Trends Biotechnol. 14 (1996) 98-105.

[28] C. Dian, S. Eshaghi, T. Urbig, S. McSweeney, A. Heijbel, G. Salbert, D. Birse, Strategies for the purification and on-column cleavage of glutathione-S-transferase fusion target proteins. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 769 (2002) 133-144.

[29] T.M. Przybycien, N.S. Pujar, L.M. Steele, Alternative bioseparation operations: life beyond packed-bed chromatography. Curr. Opin. Biotechnol. 15 (2004) 469-478.

[30] S.Y. Patro, E. Freund, C.B. S., Protein formulation and fill-finish operations. Biotechnol. Ann. Rev. 8 (2002) 55-84.

[31] J.L. Cleland, M.F. Powell, S.J. Shire, The development of stable protein formulations - a close look at protein aggregation, deamidation and oxidation. Crit. Rev. Ther. Drug Carrier Syst. 10 (1993) 307-377.

[32] V. Sluzky, J.A. Tamada, A.M. Klibanov, R. Langer, Kinetics of Insulin Aggregation in Aqueous-Solutions Upon Agitation in the Presence of Hydrophobic Surfaces. Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 9377-9381.

[33] L. Jorgensen, S. Hostrup, E.H. Moeller, H. Grohganz, Recent trends in stabilising peptides and proteins in pharmaceutical formulation - considerations in the choice of excipients. Expert Opin Drug Deliv 6 (2009) 1219-1230.

[34] R. Krishnamurthy, M.C. Manning, The stability factor: importance in formulation development. Curr. Pharm. Biotechnol. 3 (2002) 361-371.

[35] T.E. Creighton, Disulphide bonds between cysteine residues, in: T.E. Creighton (Ed.) Protein structure: a practical approach, Oxford University Press, Oxford, 1989, pp. 155-167.

[36] N. Narayanan, C.P. Chou, Physiological Improvement to Enhance *Escherichia coli* Cell-Surface Display via Reducing Extracytoplasmic Stress. Biotechnol. Prog. 24 (2008) 293 -301.

[37] N. Narayanan, C.P. Chou, Periplasmic chaperone FkpA reduces extracytoplasmic stress response and improves cell-surface display on Escherichia coli. Enzyme Microb. Technol. 42 (2008) 506-513.

[38] S. Kleist, G. Miksch, B. Hitzmann, M. Arndt, K. Friehs, E. Flaschel, Optimization of the extracellular production of a bacterial phytase with Escherichia coli by using different fed-batch fermentation strategies. Appl. Microbiol. Biotechnol. 61 (2003) 456-462.

[39] S.D. Zhou, L.P. Yomano, A.Z. Saleh, F.C. Davis, H.C. Aldrich, L.O. Ingram, Enhancement of expression and apparent secretion of Erwinia chrysanthemi endoglucanase (encoded by celZ) in Escherichia coli B. Appl. Environ. Microbiol. 65 (1999) 2439-2445.

[40] Z.B. Fu, S.B. Ab Hamid, C. Nyonya, A. Razak, M. Basri, A.B. Salleh, R.N.Z. Abd Rahman, Secretory expression in Escherichia coli and single-step purification of a heat-stable alkaline protease. Protein Expr. Purif. 28 (2003) 63-68.

[41] C.B. Anfinsen, Principles That Govern Folding of Protein Chains. Science 181 (1973) 223-230.

[42] T.A. Klink, K.J. Woycechowsky, K.M. Taylor, R.T. Raines, Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. Eur. J. Biochem. 267 (2000) 566-572.

[43] G. Bulaj, Formation of disulfide bonds in proteins and peptides. Biotechnol. Adv. 23 (2005) 87-92.

[44] P.J. Flory, Theory of Elastic Mechanisms in Fibrous Proteins. J. Am. Chem. Soc. 78 (1956) 5222-5234.

[45] S.F. Betz, Disulfide Bonds and the Stability of Globular-Proteins. Protein Sci. 2 (1993) 1551-1558.

[46] A.J. Doig, D.H. Williams, Is the Hydrophobic Effect Stabilizing or Destabilizing in Proteins - the Contribution of Disulfide Bonds to Protein Stability. J. Mol. Biol. 217 (1991) 389-398.

[47] J. Messens, J.F. Collet, Pathways of disulfide bond formation in Escherichia coli. Int. J. Biochem. Cell Biol. 38 (2006) 1050-1062.

[48] H. Bruggemann, G. Gottschalk, Insights in metabolism and toxin production from the complete genome sequence of Clostridium tetani. Anaerobe 10 (2004) 53-68.

[49] P.A. Kosen, C.B. Marks, A.M. Falick, S. Anderson, I.D. Kuntz, Disulfide Bond-Coupled Folding of Bovine Pancreatic Trypsin-Inhibitor Derivatives Missing One or 2 Disulfide Bonds. Biochemistry 31 (1992) 5705-5717.

[50] T. Creighton, Disulfide-Coupled Protein-Folding Pathways. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 348 (1995) 5-10.

[51] N. Darby, T.E. Creighton, Probing protein folding and stability using disulfide bonds. Mol. Biotechnol. 7 (1997) 57-77.

[52] G. Bulaj, D.P. Goldenberg, Early events in the disulfide-coupled folding of BPTI. Protein Sci. 8 (1999) 1825-1842.

[53] G. Bulaj, R.E. Koehn, D.P. Goldenberg, Alteration of the disulfide-coupled folding pathway of BPTI by circular permutation. Protein Sci. 13 (2004) 1182-1196.

[54] Z.S. Qiao, Z.Y. Guo, Y.M. Feng, Putative disulfide-forming pathway of porcine insulin precursor during its refolding in vitro. Biochemistry 40 (2001) 2662-2668.

[55] Z.S. Qiao, C.Y. Min, Q.X. Hua, M.A. Weiss, Y.M. Feng, In vitro refolding of human proinsulin - Kinetic intermediates, putative disulfide-forming pathway, folding initiation site, and potential role of C-peptide in folding process. J. Biol. Chem. 278 (2003) 17800-17809.

[56] W.J. Wedemeyer, E. Welker, M. Narayan, H.A. Scheraga, Disulfide bonds and protein folding. Biochemistry 39 (2000) 4207-4216.

[57] Z.Y. Guo, X.Y. Jia, Y.M. Feng, Replacement of the interchain disulfide bridge-forming amino acids A7 and B7 by glutamate impairs the structure and activity of insulin. Biol. Chem. 385 (2004) 1171-1175.

[58] Q.X. Hua, J.P. Mayer, W.H. Jia, J.W. Zhang, M.A. Weiss, The folding nucleus of the insulin superfamily - A flexible peptide model foreshadows the native state. J. Biol. Chem. 281 (2006) 28131-28142.

[59] J. Bass, G. Chiu, Y. Argon, D.F. Steiner, Folding of insulin receptor monomers is facilitated by the molecular chaperones calnexin and calreticulin and impaired by rapid dimerization. J. Cell Biol. 141 (1998) 637-646.

[60] N.S. Dangoria, M.L. DeLay, D.J. Kingsbury, J.P. Mear, B. Uchanska-Ziegler, A. Ziegler, R.A. Colbert, HLA-B27 misfolding is associated with aberrant intermolecular disulfide bond formation (dimerization) in the endoplasmic reticulum. J. Biol. Chem. 277 (2002) 23459-23468.

[61] J. Niwa, S. Yamada, S. Ishigaki, J. Sone, M. Takahashi, M. Katsuno, F. Tanaka, M. Doyu, G. Sobue, Disulfide bond mediates aggregation, toxicity, and ubiquitylation of familial amyotrophic lateral sclerosis-linked mutant SOD1. J. Biol. Chem. 282 (2007) 28087-28095.

[62] L.J. Wang, H.X. Deng, G. Grisotti, H. Zhai, T. Siddique, R.P. Roos, Wild-type SOD1 overexpression accelerates disease onset of a G85R SOD1 mouse. Hum. Mol. Genet. 18 (2009) 1642-1651.

[63] A. Vankimmenade, M.W. Bond, J.H. Schumacher, C. Laquoi, R.A. Kastelein, Expression, Renaturation and Purification of Recombinant Human Interleukin-4 from Escherichia-Coli. Eur. J. Biochem. 173 (1988) 109-114.

[64] P. Sarmientos, M. Duchesne, P. Denefle, J. Boiziau, N. Fromage, N. Delporte, F. Parker, Y. Lelievre, J.F. Mayaux, T. Cartwright, Synthesis and Purification of Active Human-Tissue Plasminogen-Activator from Escherichia-Coli. Bio-Technology 7 (1989) 495-501.

[65] M.P. Chadwick, F.E.B. May, B.R. Westley, Production and Comparison of Mature Single-Domain Trefoil Peptides Pnr-2/Ps2 Cys(58) and Pnr-2/Ps2 Ser(58). Biochem. J. 308 (1995) 1001-1007.

[66] M.M. Altamirano, C. Garcia, L.D. Possani, A.R. Fersht, Oxidative refolding chromatography: folding of the scorpion toxin Cn5. Nat. Biotechnol. 17 (1999) 187-191.

[67] F. Vinci, M. Ruoppolo, P. Pucci, R.B. Freedman, G. Marino, Early intermediates in the PDI-assisted folding of ribonuclease A. Protein Sci. 9 (2000) 525-535.

[68] J. Winter, P. Klappa, R.B. Freedman, H. Lilie, R. Rudolph, Catalytic activity and chaperone function of human protein-disulfide isomerase are required for the efficient refolding of proinsulin. J. Biol. Chem. 277 (2002) 310-317.

[69] G. Yin, J.R. Swartz, Enhancing multiple disulfide bonded protein folding in a cell-free system. Biotechnol. Bioeng. 86 (2004) 188-195.

[70] K.J. Jeong, S.Y. Lee, Secretory production of human leptin in Escherichia coli. Biotechnol. Bioeng. 67 (2000) 398-407.

[71] J. Qiu, J.R. Swartz, G. Georgiou, Expression of active human tissue-type plasminogen activator in Escherichia coli. Appl. Environ. Microbiol. 64 (1998) 4891-4896.

[72] D. Sandee, S. Tungpradabkul, Y. Kurokawa, K. Fukui, M. Takagi, Combination of Dsb coexpression and an addition of sorbitol markedly enhanced soluble expression of single-chain Fv in *Escherichia coli*. Biotechnol. Bioeng. 91 (2005) 418-424.

[73] P.H. Bessette, F. Aslund, J. Beckwith, G. Georgiou, Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 13703-13708.

[74] A.I. Derman, W.A. Prinz, D. Belin, J. Beckwith, Mutations That Allow Disulfide Bond Formation in the Cytoplasm of Escherichia-Coli. Science 262 (1993) 1744-1747.

[75] W.A. Prinz, F. Aslund, A. Holmgren, J. Beckwith, The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm. J. Biol. Chem. 272 (1997) 15661-15667.

[76] R. Levy, R. Weiss, G. Chen, B.L. Iverson, G. Georgiou, Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli* trxB gor mutants via the coexpression of molecular chaperones. Protein Expr. Purif. 23 (2001) 338-347.

[77] M.W. Larsen, U.T. Bornscheuer, K. Hult, Expression of Candida antarctica lipase B in Pichia pastoris and various Escherichia coli systems. Protein Expr. Purif. 62 (2008) 90-97.

[78] E.A. Walker, A.M. Clark, M. Hewison, J.P. Ride, P.M. Stewart, Functional expression, characterization, and purification of the catalytic domain of human 11-beta-hydroxysteroid dehydrogenase type 1. J. Biol. Chem. 276 (2001) 21343-21350.

[79] T.A. Phillips, R.A. Vanbogelen, F.C. Neidhardt, Lon gene product of *Escherichia coli* is a heat-shock protein. J. Bacteriol. 159 (1984) 283-287.

[80] H. Towbin, T. Staehelin, J. Gordon, Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. U. S. A. 76 (1979) 4350-4354.

[81] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, USA, 1989.

[82] N. Sreerama, R.W. Woody, Circular dichroism of peptides and proteins, in: N. Berova, K. Nakanishi, R.W. Woody (Eds.) Circular dichroism: principles and applications, John wiley & Sons, Inc., New York, 2000, pp. 601-620.

[83] N. Sreerama, R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal. Biochem. 287 (2000) 252-260.

[84] E. Zinser, M. Lechmann, A. Golka, B. Hock, A. Steinkasserer, Determination of the inhibitory activity and biological half-live of soluble CD83: Comparison of wild type and mutant isoforms. Immunobiology 211 (2006) 449-453.

[85] C.T. Chung, R.H. Miller, Preparation and storage of competent *Escherichia coli* cells. Methods Enzymol. 218 (1993) 621-627.

[86] G.L. Ellman, Tissue Sulfhydryl Groups. Arch. Biochem. Biophys. 82 (1959) 70-77.

[87] P.S. Stewart, A review of experimental measurements of effective diffusive permeabilities and effective diffusion coefficients in biofilms. Biotechnol. Bioeng. 59 (1998) 261-272.

[88] M.A. Andrade, P. Chacon, J.J. Merelo, F. Moran, Evaluation of Secondary Structure of Proteins from Uv Circular-Dichroism Spectra Using an Unsupervised Learning Neural-Network. Protein Eng. 6 (1993) 383-390.

[89] L. Zhang, M. Moo-Young, C.P. Chou, Molecular manipulation associated with disulfide bond formation to enhance the stability of recombinant therapeutic protein. Protein Expr. Purif. in press (2010).

[90] A. Holmgren, Tryptophan Fluorescence Study of Conformational Transitions of Oxidized and Reduced Form of Thioredoxin. J. Biol. Chem. 247 (1972) 1992-1998.

[91] C.N. Pace, G.R. Grimsley, J.A. Thomson, B.J. Barnett, Conformational Stability and Activity of Ribonuclease-T1 with Zero, One, and 2 Intact Disulfide Bonds. J. Biol. Chem. 263 (1988) 11820-11825.

[92] T. Arakawa, S.J. Prestrelski, L.O. Narhi, T.C. Boone, W.C. Kenney, Cysteine-17 of Recombinant Human Granulocyte-Colony-Stimulating Factor Is Partially Solvent-Exposed. J. Protein Chem. 12 (1993) 525-531.

[93] K. Chakraborty, S. Thakurela, R.S. Prajapati, S. Indu, P.S.S. Ali, C. Ramakrishnan, R. Varadarajan, Protein stabilization by introduction of cross-strand disulfides. Biochemistry 44 (2005) 14638-14646.

[94] J.F. Culajay, S.I. Blaber, A. Khurana, M. Blaber, Thermodynamic characterization of mutants of human fibroblast growth factor 1 with an increased physiological half-life. Biochemistry 39 (2000) 7153-7158.

[95] E. Hsu, T. Osslund, R. Nybo, B.L. Chen, W.C. Kenney, C.F. Morris, T. Arakawa, L.O. Narhi, Enhanced stability of recombinant keratinocyte growth factor by mutagenesis. Protein Eng. Des. Sel. 19 (2006) 147-153.

[96] J.X. Liu, A. Escher, Improved assay sensitivity of an engineered secreted Renilla luciferase. Gene 237 (1999) 153-159.

[97] J. Mansfeld, G. Vriend, B.W. Dijkstra, O.R. Veltman, B. VandenBurg, G. Venema, R. UlbrichHofmann, V.G.H. Eijsink, Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. J. Biol. Chem. 272 (1997) 11152-11156.

[98] S.A. Marshall, G.A. Lazar, A.J. Chirino, J.R. Desjarlais, Rational design and engineering of therapeutic proteins. Drug Discovery Today 8 (2003) 212-221.

[99] L. Zhang, M. Moo-Young, C.P. Chou, Effect of aberrant disulfide bond formation on protein conformation and molecular property of recombinant therapeutics. Pure Appl. Chem. 82 (2010) 149-159.

[100] J. Xu, M. Li, D. Kim, Y. Xu, RAPTOR: optimal protein threading by linear programming. J Bioinform Comput Biol 1 (2003) 95-117.

[101] L. Zhang, N. Narayanan, S.R. Brand, C.A. Nicolette, M. Baroja, J. Arp, H. Wang, M. Moo-Young, C.P. Chou, Structural Identification of Recombinant Human CD83 Mutant Variant as a Potent Therapeutic Protein. Protein Expr. and Purif. 73 (2010) 140-146.

[102] S. Chib, E. Greenberg, Understanding the Metropolis-Hastings Algorithm. Am. Stat. 49 (1995) 327-335.

[103] A. Lobley, B.A. Wallace, Dichroweb: A website for the analysis of protein secondary structure from circular dichroism spectra. Biophys. J. 80 (2001) 1570.

[104] A. Lobley, L. Whitmore, B.A. Wallace, DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. Bioinformatics 18 (2002) 211-212.

[105] L. Zhang, M. Moo-Young, C.P. Chou, Stability Improvement of a Therapeutic Protein by Reducing Agent Pretreatment. Chin. J. Biotechnol. 24 (2008) 2142-2143.

[106] P. Bour, T.A. Keiderling, Structure, spectra and the effects of twisting of beta-sheet peptides. A density functional theory study. J. Mol. Struct. THEOCHEM 675 (2004) 95-105.

[107] C. Chothia, Conformation of Twisted Beta-Pleated Sheets in Proteins. J. Mol. Biol. 75 (1973) 295-302.

[108] C.M. Santiveri, J. Santoro, M. Rico, M.A. Jimenez, Factors involved in the stability of isolated beta-sheets: Turn sequence, beta-sheet twisting, and hydrophobic surface burial. Protein Sci. 13 (2004) 1134-1147.

[109] A.S. Ladokhin, S. Jayasinghe, S.H. White, How to measure and analyze tryptophan fluorescence in membranes properly, and why bother? Anal. Biochem. 285 (2000) 235-245.

[110] S.V. Sikorskaya, A.V. Ignatenko, S.N. Cherenkevich, Certain Relationships of Formation of the Products of Ozonization of Tryptophan. J. Appl. Chem. USSR 57 (1984) 1910-1914.

[111] D.M. Kim, J.R. Swartz, Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of Escherichia coli. Biotechnol. Bioeng. 85 (2004) 122-129.