

MASS SPECTROMETRY BASED METHODS FOR STUDIES OF  
COMPONENT ASSEMBLY AND STRUCTURE/DYNAMICS OF  
THIAMIN DIPHOSPHATE DEPENDENT ENZYMES

By

JIEYU ZHOU

A dissertation submitted to the

Graduate School - Newark

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Chemistry

Written under the direction of

Professor Frank Jordan

And approved by

---

---

---

---

---

Newark, New Jersey

October, 2016

## **ABSTRACT OF THE DISSERTATION**

### **MASS SPECTROMETRY BASED METHODS FOR STUDIES OF COMPONENT ASSEMBLY AND STRUCTURE/DYNAMICS OF THIAMIN DIPHOSPHATE DEPENDENT ENZYMES**

by JIEYU ZHOU

Research Advisor: Professor Frank Jordan

For the metabolism of a cell, most of the proteins function as complexes rather than by themselves. Understanding not only how proteins behave in isolation but also how they recognize their binding partners is therefore critical for understanding the functions of proteins. In recent decades, our understanding of protein-protein or protein-ligand interactions have moved beyond rigid binding to conformational changes upon binding, multistep ordered assembly, and structural fluctuations occurring within fully assembled complexes. In this thesis, the Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS) and Chemical Cross-Linking Mass Spectrometry (CX-MS) methods were employed to study the assembly of human 2-oxoglutarate dehydrogenase multienzyme complex (OGDHc). Our experiments revealed the binding loci, which is essential to the assembly of human OGDHc. Our findings provide a new insight to understand the structure, organization and assembly of the human OGDHc. Moreover, we have used HDX-MS to explore the protein-ligand interactions of 1-deoxy-D-xylulose 5-phosphate synthase (DXPS). The unique peptides exhibiting EX1-type HDX kinetics found in this thiamin diphosphate (ThDP) dependent enzyme enable us to understand how DXPS structure and dynamics respond to different ligands. A better picture of ligand induced DXPS

conformational changes can be provided to help elucidate the mechanism of catalysis, and enable structure-based inhibitor design.

## ACKNOWLEDGEMENTS

First of all, I wish to acknowledge my research advisor Dr. Frank Jordan for his support, mentorship throughout the course of this journey. Working with him has been a great and wonderful experience. His guidance and encouragement helped me during the research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study.

I would like to thank my thesis committee members (Dr. Phillip Huskey, Dr. John Sheridan and Dr. Jeehiun Katherine Lee) for their valuable input during the course of my research and for their constructive remarks during my pre-oral presentation, dissertation defense, and for their critical review of this thesis.

I would also like to thank my group members: Dr. Natalia S. Nemeria, Dr. Junjie Wang for the enormous help in project development and mass spectrometry work. I would like to thank our collaborators Dr. Caren L. Freel Meyers and Ms. Alicia DeColli for providing the DXPS used in this study. I would like to thank Dr. Roman Brukh for help in Mass Spectrometry, and Dr. Lazaros Kakalis for help in protein NMR. I would like to thank all my friends and lab mates in Rutgers University.

Finally, I would like to thank my family and good friends for all their support throughout the years. My gratitude is beyond words.

## **DEDICATION**

I would like to dedicate this thesis to my parents and my wife for providing me motivation and support.

I would also like to dedicate this thesis to my son, Daniel Zhou, who has grown into a wonderful 6 years old boy in spite of his father spending a lot of time with research.

## TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION .....	ii
ACKNOWLEDGEMENTS .....	iv
DEDICATION .....	v
TABLE OF CONTENTS .....	vi
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
LIST OF EQUATIONS.....	xvi
CHAPTER 1. Introduction.....	1
1.1 Protein structure and dynamics.....	1
1.2 Methods for studying protein structure and dynamics.....	2
1.2.1 Optical methods.....	2
1.2.2 X-ray crystallography.....	2
1.2.3 NMR Spectroscopy.....	3
1.2.4 Cryo-electron microscopy.....	4
1.3 Mass spectrometry based methods.....	4
1.3.1 Hydrogen Deuterium Exchange Mass Spectrometry.....	4
1.3.2 Chemical Cross-linking with Mass Spectrometry.....	14
1.4 Scope of Thesis.....	19

CHAPTER 2. Human 2-oxoglutarate dehydrogenase multienzyme complex and the interaction between its E1o and E2o components.....	24
2.1 Introduction.....	24
2.2 E1o-E2o Interactions in Human OGDHc.....	27
2.2.1 Introduction.....	27
2.3 Materials and Methods.....	29
2.3.1 Materials.....	29
2.3.2 E1o Specific Activity.....	29
2.3.3 Measurement of OGDHc Activity.....	29
2.3.4 Interaction of OGDHc components Monitored by Size-exclusion Chromatography.....	30
2.3.5 Densitometric analysis of OGDHc components on SDS-PAGE gels.....	31
2.3.6 Isothermal Titration Calorimetry.....	31
2.3.7 Uniform <sup>15</sup> N labeling of human E1o.....	32
2.3.8 Two-dimensional <sup>1</sup> H- <sup>15</sup> N HSQC NMR.....	32
2.3.9 Hydrogen/deuterium exchange mass spectrometry.....	33
2.3.10 Bioinformatics methods.....	35
2.4 Results and Discussion.....	35

2.4.1 The activity of human E1o and the overall activity of <i>in vitro</i> reconstituted OGDHc.....	35
2.4.2 Two E2o domain constructs.....	36
2.4.3 Investigation of a potential E1o to E2o interaction with gel filtration chromatography and ITC.....	37
2.4.4 HDX-MS experiments.....	41
2.4.5 NMR evidence on the interaction of E1o with E2o.....	57
2.5 Conclusions.....	59
CHAPTER 3. E1o-E3, and E2o-E3 Interactions in Human OGDHc.....	65
3.1 Introduction.....	65
3.2 Material and methods.....	65
3.2.1 Size-exclusion chromatography.....	66
3.2.2 Isothermal titration calorimetry.....	66
3.2.3 Densitometric analysis of OGDHc components on SDS-PAGE gels.....	67
3.2.4 HDX-MS analysis.....	67
3.2.5 Two-dimensional <sup>1</sup> H- <sup>15</sup> N HSQC NMR.....	68
3.3 Results and Discussion.....	68
3.3.1 Investigation of the interaction of E1o/E3, and E2o/E3 by gel filtration and ITC methods.....	68
3.3.2 Densitometry study of OGDHc - mixing of all three components.....	71

3.3.3 HDX-MS experiment.....	75
3.3.4 HDX-MS analysis of the binding between E1o and E3.....	75
3.3.5 E2o/E3 interaction from OGDHc.....	81
3.4 Conclusion.....	85
CHAPTER 4 Study of OGDHc assembly using chemical crosslinking coupled with mass spectrometry.....	90
4.1 Introduction.....	90
4.2 Material and methods.....	91
4.2.1 Chemical cross-linking mass spectrometry.....	91
4.2.2 Fractionation of cross-linked peptides by size exclusion chromatography.....	92
4.2.3 Avidin agarose beads enrichment.....	93
4.2.4 Liquid Chromatography-Mass Spectrometry.....	93
4.2.5 Data processing.....	93
4.3 Result and Discussion.....	94
4.3.1 Strategies for identifying cross-linked peptides.....	94
4.3.2 OGDHc assembly profiling by CX-MS method.....	100
4.4 Conclusion.....	106
4.5 Future work and recommendations.....	106
CHAPTER 5 DXP synthase structure/dynamics studied with HDX-MS.....	110

5.1 Introduction.....	110
5.2 Material and methods.....	111
5.2.1 Sample preparation for HDX.....	112
5.2.2 LC-MS method and data processing.....	112
5.3 Results and Discussion.....	113
5.3.1 Overview of HDX patterns.....	113
5.3.2 EX1 and EX2 kinetics.....	114
5.3.3 HDX-MS detects local structural dynamics of DXPS.....	116
5.3.4 DXPS binds and responds differently to substrates, substrate analogue and product.....	118
5.4 Conclusion.....	120
Appendix A. Protein purification.....	138
A.1 Human E1o Purification.....	138
A.2 Human E2o N-didomain Purification.....	141
A.3 Protocol for purification of human E2o protein.....	143
A.4 Protein Sequences .....	145
List of Publications.....	150

## LIST OF TABLES

Table 2.1 HDX-MS sequence coverage.....	43
Table 2.2 Peptides from human E1o.....	43
Table 2.3 Peptides from human E2o.....	46
Table 3.1 Absolute and relative quantification of E1o, E2o, and E3 subunits in OGDHc.....	73
Table 3.2 Peptides from human E3.....	80
Table 4.1 LC-MS analysis of tryptic peptides derived from OGDHc cross-linking.....	105
Table 5.1 DXPS peptides resulting from pepsin digestion chosen for HDX-MS study.	132
Table 5.2 Rate constant of unfolding and the unfolding half-life for the four DXPS peptides displaying EX1-type HDX kinetics.....	133

## LIST OF FIGURES

Figure 1.1 Three backbone amide hydrogen exchange mechanisms.....	6
Figure 1.2 The pH dependence of exchange rate constants of different types of hydrogens occurring in peptides and proteins.....	8
Figure 1.3 Appearance of MS spectra for the EX1 and EX2 exchange kinetics during protein HDX-MS.....	11
Figure 1.4 General procedure used in HDX-MS.....	12
Figure 1.5 Four operation modes in on-line sample digestion by proteases.....	13
Figure 1.6 Typical analysis of a cross-linked sample by shotgun proteomics.....	15
Figure 1.7 Protein cross-linking using homo-bifunctional NHS-esters.....	17
Figure 1.8 Nomenclature of common products of chemical cross-linking reactions. Intra-molecular crosslink, inter-molecular crosslink, mono-link, and loop-link.....	18
Figure 2.1 The domain structure of E2o and two truncated E2o domain constructs.....	28
Figure 2.2 ITC and GFC experiments probing the interaction of E1o and E2o.....	39
Figure 2.3 ITC measurements of the binding of E2o Core-domain to E1o.....	40
Figure 2.4 Comparative HDX-MS analysis of the interaction of E1o and E2o.....	50
Figure 2.5 Deuterium uptake difference plots.....	51
Figure 2.6 Predicted secondary structure of E1o N-terminal 1-152.....	52
Figure 2.7 Comparative HDX-MS analysis of E2o in the absence and presence of E1o .	54

Figure 2.8 Sequence alignment of known binding domains with the linker-core region peptide of E2o identified from HDX-MS.....	56
Figure 2.9 <sup>1</sup> H- <sup>15</sup> N TROSY-HSQC spectra of E1o.....	58
Figure 3.1 Gel filtration study of possible interaction between E1o and E3, and E2o and E3.....	70
Figure 3.2 ITC of E1o complexed with E3, and E2o complexed with E3.....	71
Figure 3.3 Gel filtration of human OGDHc reconstituted from E1o, E2o and E3.....	74
Figure 3.4 Comparative HDX-MS analysis of the interaction of E1o and E3.....	78
Figure 3.5 Comparative HDX-MS analysis of the interaction of E1o and E3.....	79
Figure 3.6 Comparative HDX-MS analysis of the interaction of E2o with E3.....	82
Figure 3.7 Comparative HDX-MS analysis of the interaction of E3 with E2o.....	83
Figure 3.8 <sup>1</sup> H- <sup>15</sup> N TROSY-HSQC spectra of E2o N-didomain.....	84
Figure 4.1 The NHS ester cross-linking reaction.....	91
Figure 4.2 Two cross-linkers used in this study.....	94
Figure 4.3. Monitoring the OGDHc cross-linking reaction by SDS-PAGE.....	95
Figure 4.4 General analytical strategy for biotinylated cross-linker enrichment by chemical cross-linking and mass spectrometry.....	98
Figure 4.5 Enrichment of cross-linked peptides with SEC method.....	99
Figure 4.6 Schematic for CX-MS data acquisition and data analysis pipeline.....	102

Figure 4.7 Inter-cross-links within the OGDHc obtained with BS3 H <sub>12</sub> /D <sub>12</sub> cross-linker.....	103
Figure 4.8 Inter-(green) and intra-cross-linking (purple) map within the OGDHc components obtained with BS3 H <sub>12</sub> /D <sub>12</sub> cross-linker.....	104
Figure 5.1 DXPS catalyzed carboligation of pyruvate and D-GAP.....	111
Figure 5.2 Deuterium uptake plotted on the crystal structure of <i>E. coli</i> DXP synthase..	112
Figure 5.3 Free state DXPS deuterium uptake change in 30 min.....	124
Figure 5.4 DXPS peptic peptide 249-262 displaced a typical EX2 kinetics mass envelope.....	125
Figure 5.5 Difference plot showing deuterium incorporation changes of peptic fragments of DXPS in the absence and presence of ligands (deuterons exchanged in the absence of DXPS minus deuterons exchanged in the presence of DXPS) .....	126
Figure 5.6 Complete view of HDX-MS spectra of peptide 42-56 in its ThDP-bound state and in other ligand induced states.....	128
Figure 5.7 Complete view of HDX-MS spectra of peptide 51-58 in its ThDP-bound state and in other ligand induced states.....	129
Figure 5.8 Complete view of HDX-MS spectra of peptide 183-199 in its free state and in other ligand induced states.....	130
Figure 5.9 Complete view of HDX-MS spectra of peptide 278-298 in its free state and in other ligand induced states.....	131

Figure 5.10 Example of calculation of half-life of unfolding.....134

## LIST OF EQUATIONS

Equation 1.1.....	7
Equation 1.2.....	9
Equation 5.1.....	115

## CHAPTER 1. Introduction

### 1.1 Protein structure and dynamics

The primary protein structure refers to the sequence of amino acids and the location of disulfide bonds. Protein structures are also classified by their secondary structure. Secondary structures are formed by intramolecular and sometimes intermolecular hydrogen bonding of amino acid residues. The two common types of secondary structure are  $\alpha$ -helix and  $\beta$ -sheet. The spatial arrangement of secondary structures results in the formation of protein tertiary structure. The tertiary structure of a protein is the protein's geometric shape, and it is determined by a variety of non-covalent interactions provided by amino acids side chains. A number of tertiary structure elements may fold into a quaternary structure. Quaternary structure includes organization from simple dimers to large homo-oligomers and complexes with defined or variable numbers of subunits [1].

Proteins are large molecules, which can catalyze complex chemical and biological reactions, can change their activity or shape in response to metabolic signals or messages from outside the cell, and are able to recognize a broad range of substances in all biological processes. Although static structures are known for many proteins, the functions of proteins are governed ultimately by their dynamic character [2]. As we study the protein-protein or protein-ligand interactions, an important issue is the extent of conformational changes that occur between the free and bound states. For the metabolism of a cell, most of the proteins function as complexes rather than by themselves. Understanding not only how proteins

behave in isolation but also how they recognize their binding partners is therefore critical for understanding the functions of proteins.

## **1.2 Methods for studying protein structure and dynamics**

### **1.2.1 Optical methods**

Optical methods are widely used in protein structure and protein-protein interaction characterizations. The most commonly used spectroscopic techniques to study protein structure are ultraviolet visible (UV/Vis) [3], fluorescence [4], and circular dichroism (CD) spectroscopy [5]. The optical methods provide a readily accessible method to examine protein structure and protein-protein/ligand interactions. These methods are sensitive, and can enable studies of proteins under physiological conditions. These methods can be valuable in trying to understand the overall structure of target proteins, but are relatively uninformative about structural details of the underlying representations.

### **1.2.2 X-ray crystallography**

X-ray crystallography is essentially a form of very high-resolution microscopy. It enables us to visualize protein structures at the atomic level and thus enhances our understanding of protein function. To date, the primary source of structural information for protein-ligand complexes is X-ray crystallography [6-8]. We can study how proteins interact with other molecules, how they undergo conformational changes, and how they perform catalysis in the case of enzymes. However, this method does have some limitation

in the study protein structure and dynamics. First, because not all proteins are amenable to crystallization, crystal growth remains a challenge for X-ray crystallography [9]. Second, the atomic coordinates derived from X-ray crystallography are usually accompanied by displacement parameters known as the B-factors. The B-factor measures primarily the molecular disorder in the crystal and can be used as an indirect estimate of the dynamics of the protein [10]. Unfortunately for those interested in protein flexibility and dynamics, the X-ray structures of proteins are modelled as a single conformation in most cases, hence the dynamic information is therefore unavailable.

### **1.2.3 NMR Spectroscopy**

NMR spectroscopy is a powerful tool for biologists interested in the structure, dynamics and interactions of biological macromolecules. In contrast to most other methods, NMR spectroscopy studies chemical properties by studying individual nuclei. NMR spectroscopy can be applied to structure determination of proteins in the size range between 5 and 50 kDa [11, 12]. For many proteins in this size range structure determination is relatively easy. However, there are many examples of failures in protein structure determinations due to problems of aggregation, dynamics and reduced solubility. The application of NMR spectroscopy to routine protein structure determination is limited, as it requires large quantities of soluble, multiply isotopically labelled protein, considerable time and is limited to comparatively small proteins [13, 14]. Solid-state NMR spectroscopy is being applied to protein structural work and removes the limitation of protein size but is still in the developing stage for total structure determination [15, 16].

### **1.2.4 Cryo-electron microscopy**

Cryo-EM is a form of transmission electron microscopy where the sample is studied at an unstained, frozen-hydrated state. This method provides structural information with improved validity, and can even provide sufficient information to build structural models at atomic resolution. Although the method requires very small amounts of a sample in comparison to NMR, and does not require crystals that are a must for X-ray crystallography, one of the major drawbacks of this method is the relatively low resolution of the structure obtained from cryo-EM, as most of the cryo-EM structures' resolution is still well above 10 Å [17, 18]. We can certainly learn a lot from structures that have been solved at low resolutions, However, the structural details were indistinct at such resolution, and some of the structural assignments are questionable.

Although cryo-EM of biological macromolecules has made important advances in the past few years (such as the atomic cryo-EM structure of human  $\gamma$ -secretase and group IIA intron) [19], the level of current technical performance is still well below what the physics of electron scattering would allow.

## **1.3 Mass spectrometry based methods**

### **1.3.1 Hydrogen Deuterium Exchange Mass Spectrometry**

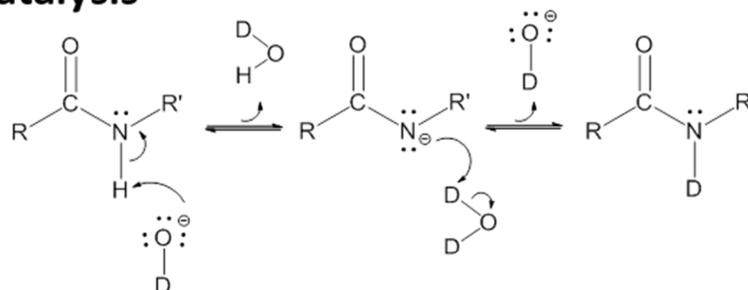
Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) is a powerful tool for probing the higher order structures and dynamics of proteins [20]. HDX-MS was

developed in the early 1990s and has undergone rapid growth in the recent decade [21]. This method takes advantage of the three types of hydrogens present in proteins (see Figure 1.1): those in carbon-hydrogen bonds, those in side-chain groups attached to electronegative atom, and those in amide functional groups (also called backbone hydrogens). The exchange rates of hydrogens in carbon-hydrogen bonds are too slow to observe, and those of side-chain hydrogens (e.g., OH, COOH) are so fast that they back-exchange rapidly when the reaction is quenched in H<sub>2</sub>O-based solution, and the exchange is not registered. Only the protein backbone amide-type hydrogens are useful for reporting protein structure and dynamics because their exchange rates are measurable and reflect hydrogen bonding and solvent accessibility. As the mass of hydrogen is 1.0078 Da and the mass of deuterium is 2.0141 Da, deuterated proteins will have a larger mass than non-deuterated proteins.

Backbone amide hydrogen exchange can occur through three distinct mechanisms: base-catalyzed exchange, acid-catalyzed N-protonation and acid-catalyzed O-protonation exchange [22]. At physiological conditions, the base-catalyzed exchange mechanism is the most important mechanism, which dominates protein or peptide backbone amide hydrogen exchange processes. The base-catalyzed mechanism was proposed by Berger et al. in 1959 based on the proteolysis and ionization experiments on N-methylacetamide [23]. Exchange begins by abstraction of the amide proton by a hydroxide ion and forms an imidate anion (Figure 1.1). The imidate anion is then reprotonated to complete the hydrogen exchange. On the acidic pH side, the acid-catalyzed hydrogen exchange may occur by two distinct mechanisms [24]. The first mechanism is called N-protonation, in which the amide nitrogen is protonated first followed by deprotonation to restore its neutral state. The other

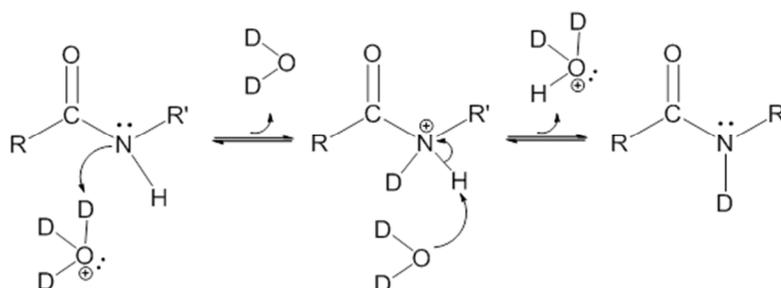
acid-catalyzed mechanism is called the O-protonation mechanism [24]. In this mechanism, the carbonyl oxygen is protonated first, the acidified amide proton is then removed by a water molecule to produce the imidic acid intermediate. Through reversing these steps, this intermediate picks a proton from solvent and returns to the amide state.

### Base catalysis

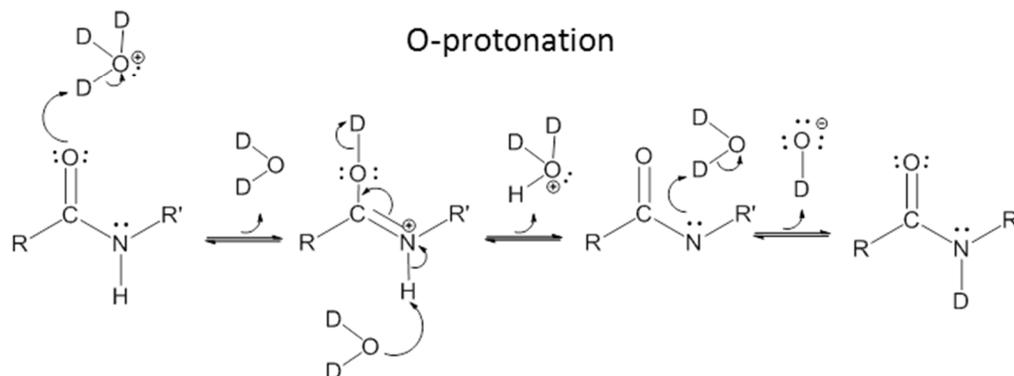


### Acid catalysis

#### N-protonation



#### O-protonation



**Figure 1.1 Three backbone amide hydrogen exchange mechanisms**

HDX-MS can provide very useful information such as: where and how fast a certain region of protein has changed, how the conformation changes upon binding, pathways of protein folding and unfolding, and protein stability under various conditions.

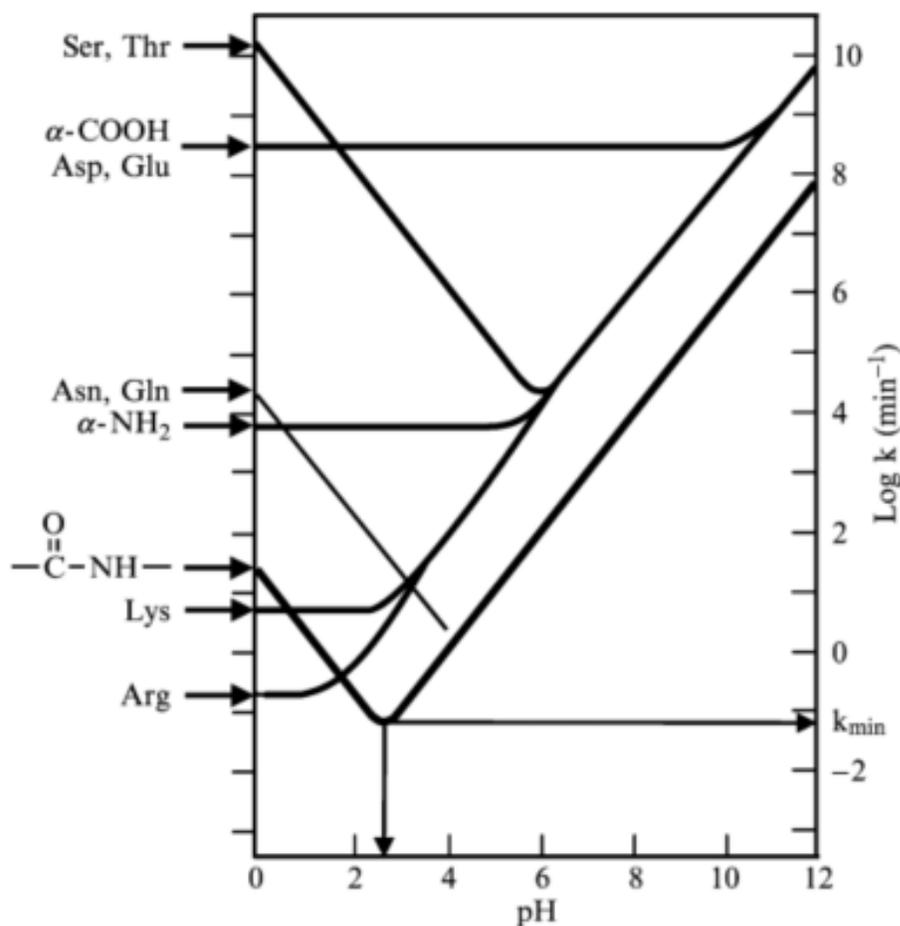
## Factors Affecting Amide Hydrogen Exchange Rate

### pH Effects

Among all the factors related to HDX, the rate and the pH can greatly affect hydrogen exchange rates, and without this pH dependency, it would be impossible to measure hydrogen exchange with mass spectrometry. As discussed above, backbone amide hydrogen exchange is a base-catalyzed and acid-catalyzed reaction. Therefore, the hydrogen exchange rate constant ( $k_{ex}$ ) can be described as a sum of the base-catalyzed ( $k_{OH^-}$ ), acid-catalyzed ( $k_{H^+}$ ), and water catalyzed ( $k_{water}$ ) contributions [25].

$$k_{ex} = k_{OH^-}[OH^-] + k_{H^+}[H^+] + k_{water} \quad \text{Equation 1.1}$$

An exchange study of poly-DL-alanine has yielded a quantitative description of pH effects on the backbone amide hydrogen exchange rate [26-28]. The pH dependence of several types of hydrogens found in peptides and proteins is shown in Figure 1.2. It is clear that the log ( $k_{ex,NH}$ ) versus pH plot has a pronounced minimum at pH 2.5-3.0. Each unit change in pH from the  $pH_{min}$  will result in an approximately 10-fold increase of hydrogen exchange rate. Therefore, the hydrogen exchange rate can be minimized during MS analysis, when the entire system pH is adjusted to 2.5.



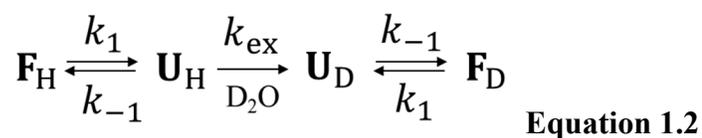
**Figure 1.2** The pH dependence of exchange rate constants of different types of hydrogens present in peptides and proteins [29].

### Temperature Effects

The hydrogen exchange rate is also strongly temperature dependent. Since the hydrogen exchange is an acid and base-catalyzed reaction, the exchange rate increases approximately three-fold for each 10 °C increase [24]. The amide hydrogen exchange rate can therefore be reduced by approximately 10-fold when the exchange reaction temperature is reduced from 20 °C to 0 °C.

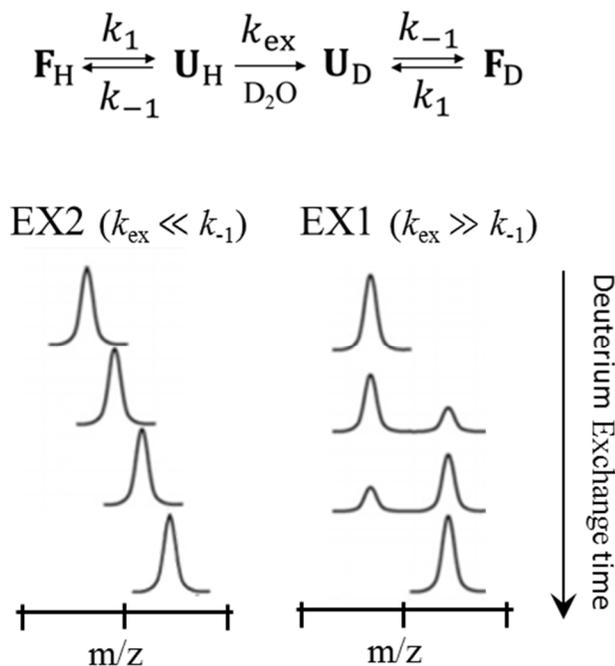
## EX1 and EX2 kinetics

Two exchange mechanisms have been proposed to explain hydrogen exchange under physiological conditions in proteins. The first mechanism suggests that exchange may occur from the folded form, the other suggests partial or global unfolding first, followed by exchange [30-32]. Under physiological conditions, the unfolding is the predominant mechanism for HD exchange in proteins.



As illustrated in equation 1.2, there are three steps involved in backbone amide hydrogen exchange in proteins. In order to exchange, the folded conformation  $\mathbf{F}_H$  is converted to the unfolded conformation  $\mathbf{U}_H$ , and exposes backbone amide hydrogens for exchange with D from  $\text{D}_2\text{O}$  solvent. The rate constants  $k_1$ ,  $k_{-1}$ , and  $k_{\text{ex}}$  describe the unfolding, refolding, and hydrogen exchange rate constants, respectively. There are two kinetics schemes, EX1 and EX2, to describe hydrogen exchange in this case. When the exchange is governed by EX1 kinetics, the exchange rate is much faster than the refolding rate ( $k_{\text{ex}} \gg k_{-1}$ ). Under these conditions, all of the amide hydrogens exchange with deuterium in the unfolded state before refolding occurs. As a result, EX1 kinetics gives rise to two distinct and separated mass envelopes, the lower mass envelope represents the folded state, and the

higher mass envelope represents the unfolded state (see Figure 1.3). In EX2 kinetics, the protein refolding rate is much faster than the exchange rate ( $k_{\text{ex}} \ll k_{-1}$ ), therefore the unfolding has to occur multiple times before a successful exchange reaction takes place. EX2 kinetics leads to a single isotopic distribution gradually shifting to a higher m/z range over time (see Figure 1.3). The majority of native proteins are very stable under physiological conditions, therefore proteins must make many brief visits to a partially unfolded conformation to undergo exchange, and the exchange follows EX2 kinetics. A few proteins exhibit EX1 kinetics under physiological conditions. The occurrence of EX1 kinetics under physiological conditions can provide important information not only for understanding the nature of protein structural dynamics in solution but also for the development of binding assays that can exploit the stability of the proteins. However, EX1 kinetics are often observed when proteins are under denaturing conditions, or as a reflection of experimental error, such as carryover between injections. To avoid an erroneous assignment of EX1 kinetics, in our studies the protein's stability was carefully examined and carryover was avoided by extensively washing the pepsin and C18 columns between runs.



**Figure 1.3 Appearance of MS spectra for the EX1 and EX2 exchange kinetics during protein HDX-MS**

### General Workflow of HDX-MS Experiments

The general procedure used in HDX-MS is illustrated in Figure 1.4. The typical HDX-MS experiment contains four steps: 1) the intact protein or protein complex undergoes exchange in  $\text{D}_2\text{O}$  buffer under physiological conditions for certain length of time; 2) the exchanged protein sample is quenched by decreasing the pH to 2.5 and the temperature to  $0\text{ }^\circ\text{C}$ ; 3) the protein is digested with pepsin or other acid active protease under quench conditions; 4) The resulting peptides are retained on a C8 or C18 trapping column, and are subjected to LC-MS analysis.

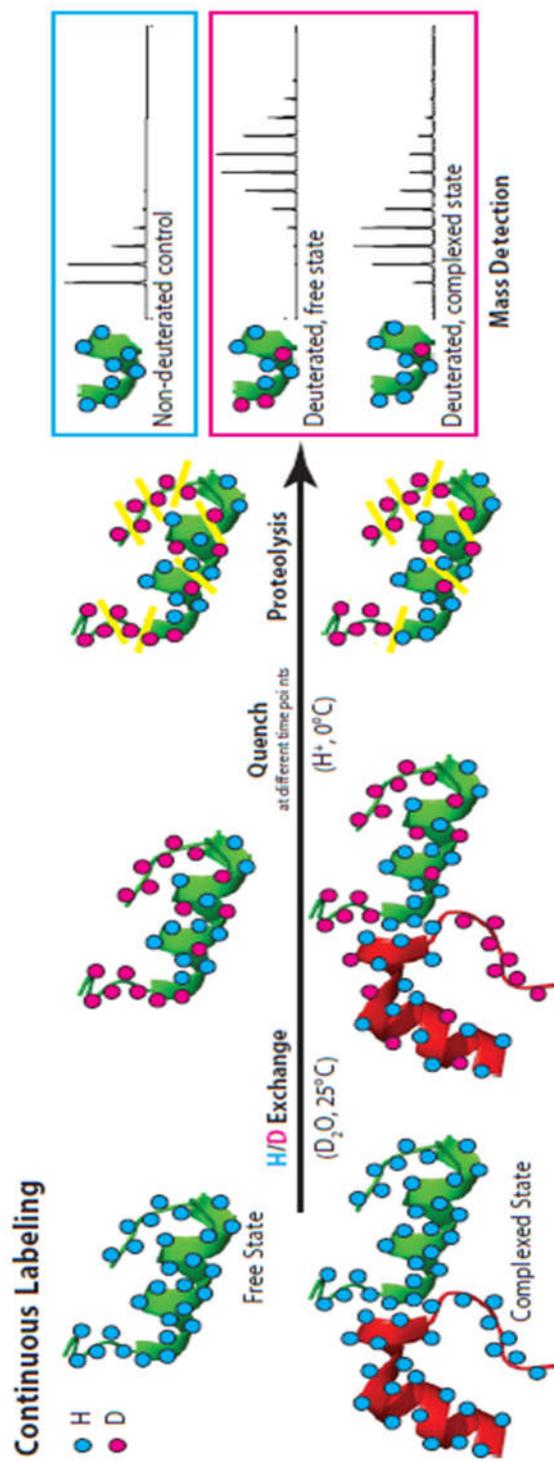
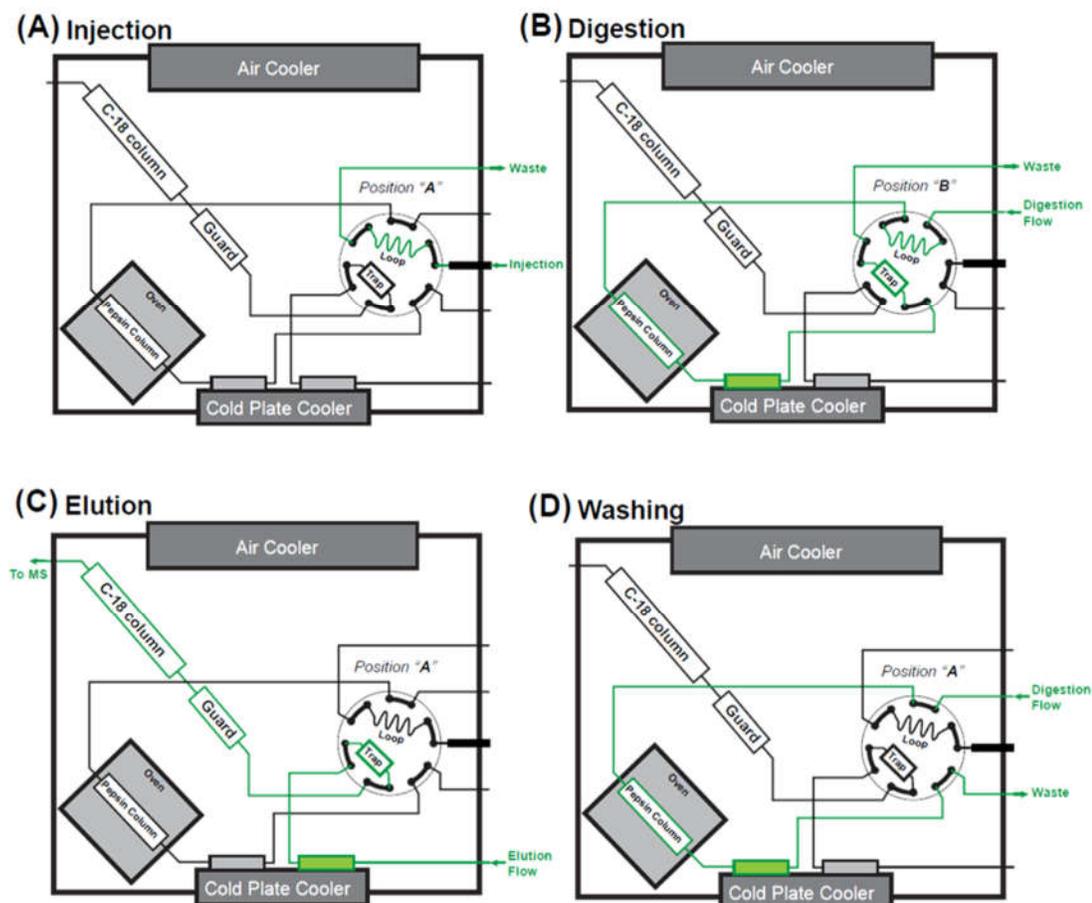


Figure 1.4 General procedure used in HDX-MS

In our lab, the HDX device built by Dr. Junjie Wang has included four operation modes to successfully conduct on-line sample digestion and separation with a general back exchange rate below 20 % [33]. Figure 1.5 illustrates the process of the four operation modes.

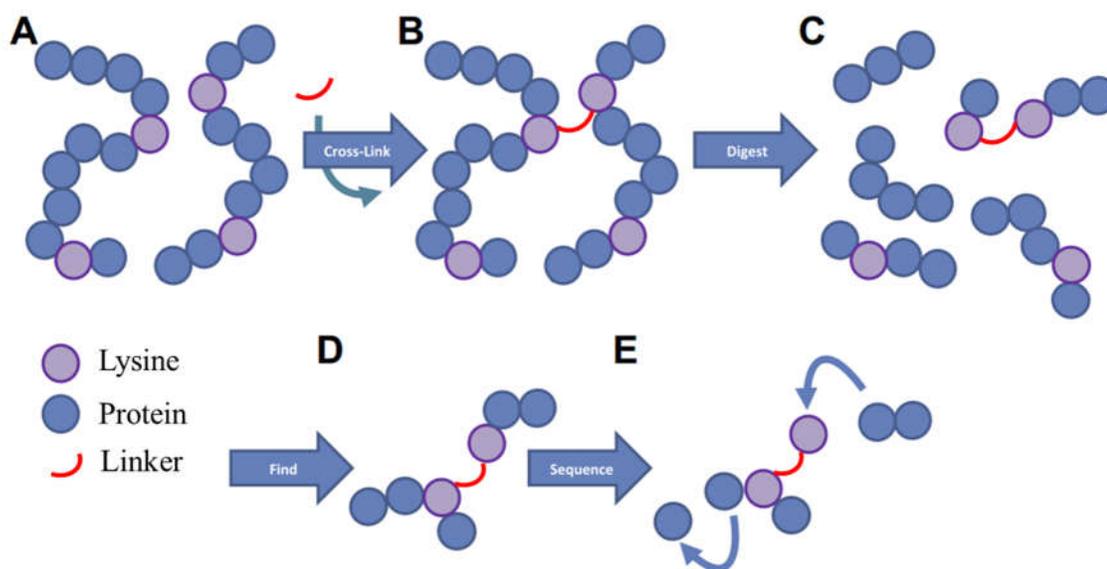


**Figure 1.5 Four operation modes in on-line sample digestion**

A) Injection mode, the HDX quenched protein sample is injected into a pre-cooled 20  $\mu$ l loop; B) Digestion mode, the protein sample is carried by digestion flow (0.1 % formic acid in H<sub>2</sub>O) to a pepsin immobilized column pre-heated to 15 °C, and the resulting peptides were immediately cooled down to 0 °C, desalted and enriched on a pre-cooled C18 peptide trap; C) Elution mode, the elution flow (acetonitrile and water gradient wash with 0.1 % formic acid) elute all trapped peptides to C18 column for separation and mass spectrometric analysis; D) Washing mode, between sample to sample injection, the pepsin column was flushed and regenerated to avoid carry over.

### **1.3.2 Chemical Cross-linking with Mass Spectrometry**

Theoretically, chemical cross-linking of proteins and protein complexes followed by mass spectrometric analysis (CX-MS) can build up a set of structurally defined interactions by covalently connecting two amino acids within one protein or different proteins. The location of the identified cross-links can suggest a distance constraint on the location of the respective amino acids and allows drawing conclusions on the distance geometries of proteins or protein complexes [34]. Figure 1.6 illustrates a typical analysis of a cross-linked sample with the CX-MS method.



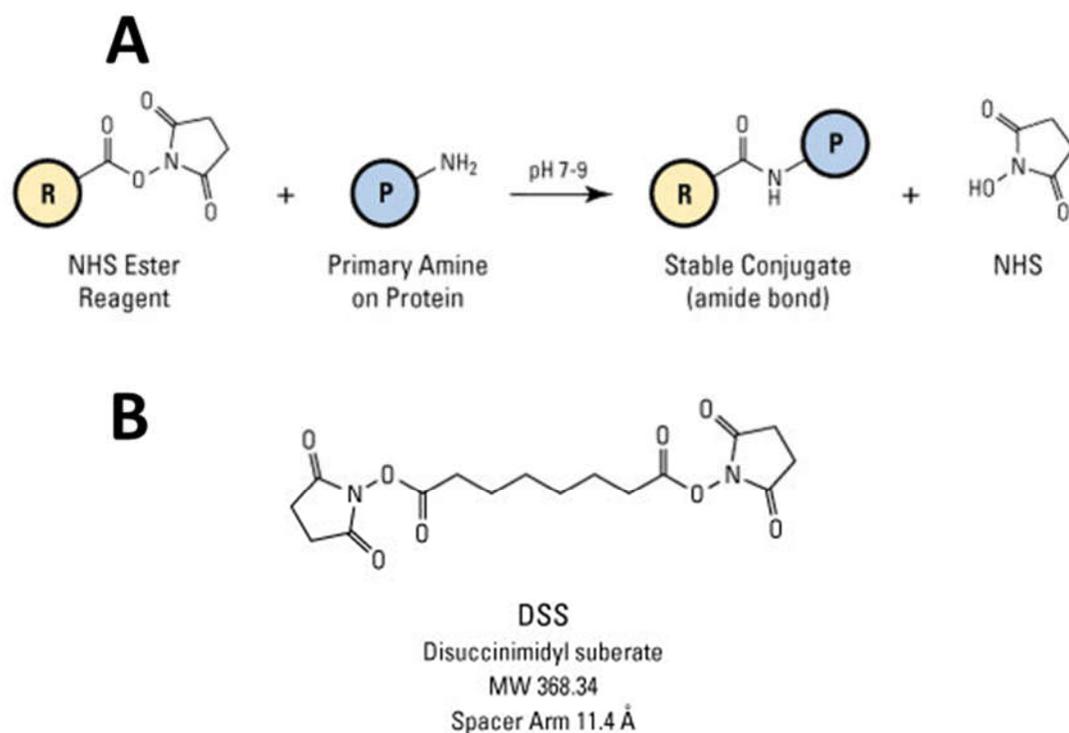
**Figure 1.6 The typical analysis of a cross-linked sample by shotgun proteomics**

(A) The protein or proteins are incubated with a residue specific cross-linking reagent. (B) Residues within the range of the cross-linking reagent are then covalently bonded and transient interactions are stabilized. (C) The protein is then digested by a specific protease to form peptides. (D) Data-dependent acquisition is used to identify peptides as they elute from an HPLC directly coupled to the mass spectrometer. (E) The identified peptides are then fragmented to provide sequence specific information.

Protein cross-linking with mass spectrometry as a powerful method to characterize protein structural information and protein-protein interactions was introduced by Young et al. in 2000 [35]. The folding of bovine basic fibroblast growth factor (FGF)-2, and its tertiary structure was probed with a lysine-specific cross-linking agent, bis(sulfosuccinimidyl) suberate (BS3). The cross-links were identified on line by ESI-TOF, and further confirmed by MALDI-TOF. In this early research 15 cross-links that did not bridge the adjacent lysines were identified, and interatomic distance constraints were all consistent with the tertiary structure of FGF-2.

In recent years, the application of chemical cross-linking with mass spectrometry to study protein structures and protein-protein interactions has rapidly expanded [36-38]. New developments in cross-linker design [39, 40], cross-linked peptide enrichment [36, 41], cross-link identification and data analysis [36, 42] have had an impact on a variety of biological disciplines.

With the continuing effort to study larger and more complex systems, the application of CX-MS provided a perfect tool to achieve these goals. The uniqueness of this method is the ability to apply it at, or near physiological conditions with very small amounts of proteins. Protein cross-linking is commonly undertaken using homobifunctional NHS-esters (see Figure 1.7 A) which bridges two spatially proximate lysine residues within one peptide or between two peptides. The range of this interaction is often controlled by varying the cross-linker length of the space between the two functional groups (see Figure 1.7 B).

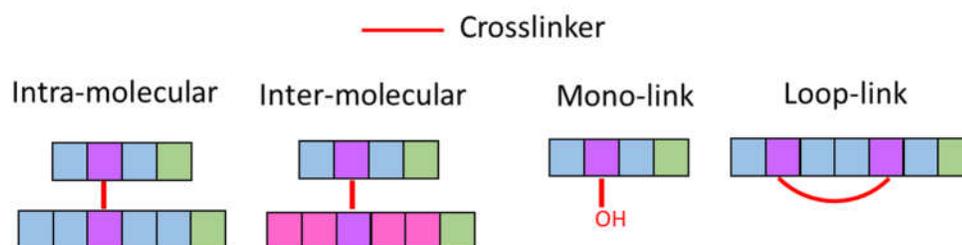


**Figure 1.7 Cross-linking using homo-bifunctional NHS-esters**

### Cross linker types

The intent of protein chemical cross-linking is to covalently bond two spatially proximate amino acids within two different peptides. However, during the cross-linking reactions, other side reactions can occur and result in unexpected cross-linked products. Such different products of the cross-linking reaction are summarized in Figure 1.8. The intra-molecular and inter-molecular crosslinks are the crosslinking products resulting from two peptides from the same protein or two different proteins, respectively. The mono-link products are produced when only one end of a bifunctional cross-linker reacts with the protein, the other end does not make contact with any cross-linkable amino acids, and is

subsequently hydrolyzed or quenched by quench reagent. The loop-link product is produced when no proteolytic site exists between two cross-linked amino acids within one peptide. Although only inter-molecular crosslinks can provide direct information on protein-protein interactions, other crosslinks can provide insight into protein structure, location of surface residues and solvent accessible residues.



**Figure 1.8 Nomenclature of common products of chemical cross-linking reactions.**

**Intra-molecular crosslink, inter-molecular crosslink, mono-link, and loop-link.**

## 1.4 Scope of Thesis

Proteins are generally thought to maintain certain structures based on their primary amino acids sequence, however, proteins are not strictly static structures but are dynamic systems that undergo fluctuation and equilibration between two or more conformations. Understanding the structure and dynamics of proteins is essential to explore their function and activity. The goal of this study is to apply various techniques specially HDX-MS (along with CX-MS) for (1) Investigation of interactions and assembly among the three components in the human 2-oxoglutarate dehydrogenase multienzyme complex; (2) A study of the conformational dynamics of 1-deoxy-xyluose 5-phosphate synthase (DXPS), a thiamin diphosphate dependent enzyme, upon interaction with its two substrates, a substrate analogue and its product.

## References

1. Berg, M.J., Tymoczko, L.J., and Stryer, L. Section 3.5 Quaternary Structure: Polypeptide Chains Can Assemble Into Multisubunit Structures. *Biochemistry* (5. ed., 4. print. ed.). New York, NY
2. Henzler-Wildman K., and Kern, D. (2007) Dynamic personalities of proteins. *Nature* *450*, 964-972.
3. Xiao, Z.Y., and Yin, J.H. (2013) Spectroscopic studies on bilirubin aggregate at liquid/liquid interface. *Anal. Bioanal. Chem.* *405*, 2723-2728.
4. Martinière, A., Lavagi, I., Nageswaran, G., Rolfe, D.J., Maneta-Peyret, L., Luu, D.T., Botchway, S.W., Webb, S.E., Mongrand, S., Maurel, C., Martin-Fernandez, M.L., Kleine-Vehn, J., Friml, J., Moreau, P., and Runions, J. (2012) Cell wall constrains lateral diffusion of plant plasma-membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 12805-12810.
5. Nemeria, N., Tittmann, K., Joseph, E., Zhou, L., Vazquez-Coll, M.B., Arjunan, P., Hübner, G., Furey, W. and Jordan, F. (2005) Glutamate 636 of the Escherichia coli pyruvate dehydrogenase-E1 participates in active center communication and behaves as an engineered acetolactate synthase with unusual stereoselectivity. *J. Biol. Chem.* *280*, 21473-21482.
6. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958) A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* *181*, 662-666.
7. Frauenfelder, H., Petsko, G. A., and Tsernoglou, D. (1979) Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* *280*, 558-563.
8. Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991) The Energy Landscape and Motions of Proteins. *Science* *254*, 1598-1603.
9. Frauenfelder, H., Chen, G., Berendzen, J., Fenimore, P. W., Jansson, H., McMahon, B. H., Stroe, I. R., Swenson, J., and Young, R. D. (2009) A unified model of protein dynamics. *Proc. Natl. Acad. Sci. U.S.A.* *106*, 5129-5134.
10. DePristo, M. A., de Bakker, P. I., and Blundell, T. L. (2004) Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. *Structure* *12*, 831-838.
11. Brünger A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M.,

- Simonson, T., and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* *54*, 905-921.
12. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* *273*, 283-298.
  13. Lange, O.F., Rossi, P., Sgourakis, N.G., Song, Y., Lee, H.W., Aramini, J.M., Ertekin, A., Xiao, R., Acton, T.B., Montelione, G.T., and Baker, D. (2012) Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *Proc. Nat.l Acad. Sci. U. S. A.* *109*, 10873-10878.
  14. Opella, S.J., and Marassi, F.M. (2004) Structure determination of membrane proteins by NMR spectroscopy. *Chem Rev.* *104*, 3587-3606.
  15. Ravera, E., Cerofolini, L., Martelli, T., Louka, A., Fragai, M., and Luchinat, C. (2016) (1)H-detected solid-state NMR of proteins entrapped in bioinspired silica: a new tool for biomaterials characterization. *Sci Rep.* *6*, 27851./
  16. Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* *420*, 98-102.
  17. Gabashvili, I.S., Agrawal, R.K., Spahn, C.M., Grassucci, R.A., Svergun, D.I., Frank, J., and Penczek, P. (2000) Solution structure of the E. coli 70S ribosome at 11.5 Å resolution. *Cell* *100*, 537-549.
  18. Allen, G.S., Zavialov, A., Gursky, R., Ehrenberg, M., and Frank, J. (2005) The cryo-EM structure of a translation initiation complex from Escherichia coli. *Cell* *121*, 703-712.
  19. Bai, X.C., Yan, C., Yang, G., Lu, P., Ma, D., Sun, L., Zhou, R., Scheres, S.H., and Shi, Y. (2015) An atomic structure of human  $\gamma$ -secretase. *Nature* *525*, 212-217.
  20. Wales, T.E., and Engen, J.R. (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.* *25*, 158-170.
  21. Katta, V., and Chait, B.T. (1991) Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* *5*, 214-217.
  22. Dempsey, C.E. (1986) pH dependence of hydrogen exchange from backbone peptide amides in apamin. *Biochemistry* *25*, 3904-3911.

23. Berger, A., Loewenstein, A., and Meibom, S. (1959) Nuclear magnetic resonance study of the protolysis and ionization of N-methylacetamide. *J. Am. Chem. Soc.* *81*, 62–67.
24. Englander, S.W., Sosnick, T.R., Englander, J.J., and Mayne, L. (1996) Mechanisms and uses of hydrogen exchange. *Curr. Opin. Struct. Biol.* *6*, 18–23.
25. Konermann, L., Pan, J., and Liu, Y.H. (2011) Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* *40*, 1224-1234.
26. Englander, S.W., Downer, N.W., and Teitelbaum, H. (1972) Hydrogen exchange. *Annu. Rev. Biochem.* *41*, 903-924.
27. Englander, S.W., and Staley, R. (1969) Measurement of the free and the H-bonded amides of myoglobin. *J. Mol. Biol.* *45*, 277-295.
28. Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* *17*, 75-86.
29. Maier, C.S., and Deinzer, M.L. (2005) Protein conformations, interactions, and H/D exchange. *Methods Enzymol.* *402*, 312-360.
30. Woodward, C., Simon, I., and Tuchsén, E. (1982) Hydrogen exchange and the dynamic structure of proteins. *Mol. Cell Biochem.* *48*, 135–160.
31. Englander, S.W., and Kallenbach, N.R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* *16*, 521–655.
32. Morgan, C.R., and Engen, J.R. (2009) Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr. Protoc. Protein Sci.* CHAPTER: Unit–17.617
33. Wang, J. (2014) Interaction of the components in human and escherichia coli pyruvate dehydrogenase multienzyme complexes (PhD dissertation). Rutgers University-Newark. Retrieved from <http://dx.doi.org/doi:10.7282/T3FX77J5>
34. Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* *25*, 663-682.
35. Young, M.M., Tang, N., Hempel, J.C., Oshiro, C.M., Taylor, E.W., Kuntz, I.D., Gibson, B.W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. USA.* *97*, 5802–5806.

36. Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L.N., Beck, M., Schmidt, A., Mueller, M., and Aebersold, R. (2008) Identification of cross-linked peptides from large sequence databases. *Nat. Methods.* *5*, 315–318.
37. Yan, F., Che, F.Y., Rykunov, D., Nieves, E., Fiser, A., Weiss, L.M., and Hogue, Angeletti R. (2009) Nonprotein based enrichment method to analyze peptide cross-linking in protein complexes. *Anal. Chem.* *81*, 7149–7159.
38. Zhang, H., Tang, X., Munske, G.R., Tolic, N., Anderson, G.A., and Bruce, J.E. (2009) Identification of protein–protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. *Mol. Cell. Proteomics.* *8*, 409–420.
39. Petrotchenko, E.V., Olkhovik, V.K., and Borchers, C.H. (2005) Isotopically coded cleavable cross-linker for studying protein–protein interaction and protein complexes. *Mol. Cell. Proteomics.* *4*, 1167–1179.
40. Petrotchenko, E.V., Serpa, J.J., and Borchers, C.H. (2010) Use of a combination of isotopically coded cross-linkers and isotopically coded N-terminal modification reagents for selective identification of inter-peptide crosslinks. *Anal. Chem.* *82*, 817–823.
41. Dowell, J. A., Frost, D. C., Zhang, J., and Li, L. (2008) Comparison of two-dimensional fractionation techniques for shotgun proteomics. *Anal. Chem.* *80*, 6715–6723.
42. Anderson, G. A., Tolic, N., Tang, X., Zheng, C., and Bruce, J. E. (2007) Informatics strategies for large-scale novel cross-linking analysis. *J. Proteome Res.* *6*, 3412–3421.

## **CHAPTER 2. Human 2-oxoglutarate dehydrogenase multienzyme complex and the interaction between its E1 $\alpha$ and E2 $\alpha$ components**

### **2.1 Introduction**

The 2-oxoglutarate dehydrogenase multienzyme complex is a rate-limiting enzyme in the citric acid cycle and produces succinyl-CoA by oxidative decarboxylation of 2-oxoglutarate. Deficiencies of OGDHc are likely to impair brain energy metabolism and therefore brain function, and lead to manifestations of brain disease such as infantile lactic acidosis, psychomotor retardation in childhood, Parkinson's disease, and Alzheimer's disease [1].

The superfamily of 2-oxo acid dehydrogenase multienzyme complexes contains three members: pyruvate dehydrogenase complex (PDHc), 2-oxoglutarate dehydrogenase complex (OGDHc), and branched-chain 2-oxoacid dehydrogenase complex (BCOADHc), and also includes the glycine cleavage system. Each complex consists of multiple copies of three enzyme components, termed E1, E2 and E3, which catalyze consecutive steps in the oxidative decarboxylation of a 2-oxo acid and the reductive acylation of the lipoamidated E2 component followed by formation of acyl-CoA, with overall concomitant reduction of NAD<sup>+</sup> to NADH and the release of CO<sub>2</sub>. The E1 component is a 2-oxo acid dehydrogenase with thiamin diphosphate (ThDP) as cofactor, the E2 component is a dihydrolipoamide acyltransferase and has variable numbers of tandem lipoyl domains, each carrying a covalently attached lipoic acid as cofactor, and E3 is a dihydrolipoamide

dehydrogenase containing FAD as a cofactor. These complexes have molecular masses of several million Daltons and self-assemble in the mammalian body [2-4].

To assemble the complexes, in general, the oligomeric E2 component forms the core of the multienzyme complex, to which the E1 and E3 (sometimes called peripheral) components are non-covalently attached. The mammalian PDHc has a molecular mass of approximately  $9.5 \times 10^6$  Da [5]. To the 48-meric E2p core (this thesis will adopt the *p* suffix for pyruvate and *o* for 2-oxoglutarate), 20-30 copies of heterotetrameric E1p, 12 copies of homodimeric E3 and 12 copies of monomeric E3-binding protein (E3BP) are non-covalently attached [6]. The E1p and E3 components bind independently to a subunit-binding domain (PSBD) located on E2p and E3BP, respectively [6-8]. Mammalian PDHc requires a separate protein E3BP to tether E3 to the complex, the mammalian BCOADHc has a 24-meric cubic E2b core, to which E1b and E3 bind simultaneously at the same binding domain of each E2b (note that there are multiple copies of each component) [10-12]. In contrast, the mammalian OGDHc differs significantly from the general assembly mode as the sequence alignment of the E2o could not identify any apparent PSBD [13, 14]. In addition, there is no extra-component such as E3BP which could provide PSBD in this complex. Hence, as to how this complex is assembled, and whether indeed the mammalian OGDHc has a new unprecedented mode of assembly, is still an active research topic.

Previous studies have highlighted that protein-protein interactions between the E1o and E3 components may be important in the assembly of OGDHc [15-17], but the putative interaction sites have not been identified. Owing to earlier difficulties in producing active full-length human E1o or E2o, the assembly of human OGDHc could not be studied. With

the success of the Jordan lab in producing active full length E1o and E2o by recombinant methods [18], such studies could now be undertaken as reported in this thesis.

The studies described herein were performed to determine interactions among the three components of human OGDHc. We present *in vitro* reconstituted binary sub-complex studies between pairs of components, as well as a study of *in vitro* reconstituted entire complex of human OGDHc using H/D exchange MS (HDX-MS) and chemical crosslinking MS (CX-MS) methods. The HDX-MS method showed that, although lacking an obvious PSBD on E2o, E1o rigidly associates with E2o through the E1o N-terminal region and the linker-core region of E2o. CX-MS studies on the E1o/E2o sub-complex and on the intact OGDHc further validate the nature of the association between E1o and E2o at these two specific regions. With the same methods, studies of the E1o/E3 sub-complex show that the N-terminal region of E1o associates with E3, however, the binding is much weaker than that in the E1o/E2o sub-complex, as proved with ITC and gel filtration methods. Moreover, we showed that E2o and E3 both interact with the same peptides located at the N-terminal region of E1o. These results implied that the E1o N-terminus is pivotal for mediating assembly of the OGDHc. This, together with the fact that the affinity of E1o/E2o sub-complex was much stronger than that of E1o/E3 sub-complex, has led to much speculation that OGDHc assembly is initiated by the self-integration of E1o/E2o to form a sub-complex and subsequently promoting binding of E3 with E1o thus forming a stable complex.

## 2.2 E1o-E2o Interactions in Human OGDHc

### 2.2.1 Introduction

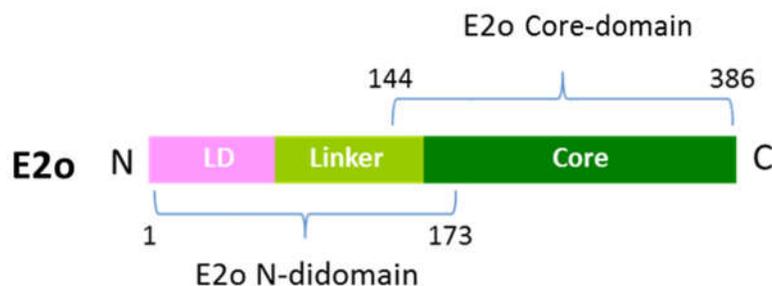
As discussed previously, the organization of the family of 2-oxo acid dehydrogenase complexes is based on the self-assembly of its E2 components to form a 60-meric icosahedral E2p-PDHc core, or a 24-meric cube E2o-OGDHc core or a 60-meric pentagonal dodecahedron E2b-BCOADC core, and then multiple copies of E1 and E3 components are attached non-covalently. Multiple sequence alignment attempts in human E2o showed that there is no obvious E1o binding or E3 binding domain in E2o [13, 14]. Previous studies have shown that selective tryptic proteolysis of bovine heart OGDHc leads to its rapid inactivation, and cleavage of the first 77 residues from the N-terminus of bovine OGDHc also leads to dissociation of the E3 component [17].

Due to the inability to produce full-length active E1o in soluble form, Lindsay's group used three N-terminal E1o constructs (E1o<sub>1-60</sub>, E1o<sub>1-90</sub>, and E1o<sub>1-153</sub>) to study the assembly of human OGDHc [19]. Although the N-terminal region of E1o failed to interact with E2o post-translationally according to gel filtration and affinity chromatography methods, the co-expressed GST fusion N-terminal region of E1o and His-tagged E2o were able to integrate together as a stable complex. These observations raised the possibility that, like other 2-oxo acid dehydrogenase complexes, E1o can bind with E2o.

For the research planned for this chapter, a full-length soluble and active E1o enzyme was overexpressed and purified in our laboratory. Initially, the size-exclusion chromatography (SEC) and isothermal titration calorimetry (ITC) methods were employed to investigate the interaction between E1o and E2o. Next, the HDX-MS method was used

to map the binding loci of E1o and E2o at a peptide-specific resolution. In addition, two E2o constructs (E2o N-didomain (E2o 1-173) and Core-domain (E2o 144-386)) were independently over-expressed and purified (Figure 2.1). The HDX-MS analysis of E1o interaction with the two E2o constructs revealed a new binding locus on E2o. Finally, two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR experiments were conducted to confirm the direct interaction between E1o and E2o.

This chapter not only provides evidence for direct binding between E1o and E2o, which is essential to the assembly of human OGDHc, but also reveals the dynamic nature of the N-terminal region of E1o. This provides a new insight to understand the structure, organization and assembly of the human OGDHc.



**Figure 2.1** The domain structure of E2o and two truncated E2o constructs.

E2o N-didomain (residues 1-173), and E2o Core-domain (residues 144-386) have a 30 amino acids overlapping sequence from the linker-core region.

## **2.3 Materials and Methods**

### **2.3.1 Materials**

Deuterium oxide (D<sub>2</sub>O) and <sup>15</sup>NH<sub>4</sub>Cl (99%) were from Cambridge Isotope Laboratories. ThDP was from Affymetrix. All other fine chemicals were from Sigma-Aldrich. Expression and purification of full-length E1o, E2o, E2o N-didomain, and E2o Core-domain were described in Appendix A.

### **2.3.2 Measurement of E1o Specific Activity**

The E1o specific activity was measured in the reaction medium with the external oxidizing agent 2,6-dichlorophenolindophenol (DCPIP) [18] at 600 nm in the following reaction medium contained in 1 ml: 50 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.0), 0.50 mM ThDP, 1.0 mM MgCl<sub>2</sub>, 2 mM 2-oxoglutarate (OG), and DCPIP (0.08 mM) at 37 °C. The reaction was initiated by the addition of 0.01– 0.015 mg of E1o. Steady-state velocities were taken from the linear portion of the progress curve recorded at 600 nm. One unit of activity is defined as the amount of reduced DCPIP produced ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg E1o}^{-1}$ ).

### **2.3.3 Measurement of OGDHc Activity**

Overall activity was measured as previously described [18]. For reconstitution of OGDHc from its individual components, E1o (0.13 mg; 4.6  $\mu\text{M}$  subunits) was incubated for 30 min with E2o (0.65 mg; 60  $\mu\text{M}$  subunits) and E3 (0.65 mg; 52  $\mu\text{M}$  subunits) with a mass ratio ( $\mu\text{g}/\mu\text{g}/\mu\text{g}$ ) of 1:5:5 in 0.25 ml of 50 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.5) containing 0.50 mM

ThDP, 1.0 mM MgCl<sub>2</sub>, and 0.15 M NaCl at 25 °C. A 10 µl aliquot of the reaction mixture was withdrawn after a 1-h incubation to start the reaction. The stoichiometry of components used for activity measurement was chosen from kinetic experiments where, at constant concentrations of E1o subunits (0.046 µM) and E3 subunits (0.52 µM), the concentration of E2o subunits was varied (0.059– 0.80 µM), leading to  $S_{0.5, E2o} = 0.18 \pm 0.08$  µM. In a similar experiment with varied concentrations of E3 subunits (0.026– 0.53 µM) and constant concentrations of E1o (0.046 µM) and E2o (0.24 µM) subunits, an  $S_{0.5, E3}$  of  $0.17 \pm 0.02$  µM was calculated. The reaction medium contained the following in 1.0 ml: 0.10 M Tris-HCl (pH 7.5), 0.50 mM ThDP, 2.0 mM MgCl<sub>2</sub>, 2.0 mM OG, 2.0 mM DTT, 5 mM NAD<sup>+</sup>, and 0.1-0.2 mM CoA at 37 °C. The reaction was initiated by the addition of CoA and OGDHc. Steady-state velocities were taken from the linear portion of the progress curve. One unit of activity is defined as the amount of NADH produced ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg E1o}^{-1}$ ).

#### **2.3.4 Interaction of OGDHc components Monitored by Size-exclusion Chromatography**

Size-exclusion chromatography experiments were performed using a Varian ProStar HPLC system with a UV detector. A Yarra 3u SEC-3000 column with a 20 µl sample loop was used at a flow rate of 1 ml/min. The column was equilibrated with 50 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.5) containing 0.15 M NaCl, 0.5 mM ThDP, 1 mM MgCl<sub>2</sub>, and calibrated with the standards ( $M_r$  in parenthesis): thyroglobulin (669,000), ferritin (440,000), catalase (232,000), aldolase (158,000), bovine serum albumin (67,000), and ovalbumin (43,000). The protein samples were eluted at a flow rate of 1 ml/min and monitored at 280 nm. All

protein samples were incubated at various molar ratios with running buffer at room temperature for 1 h. Samples were centrifuged at  $17,530 \times g$  for 5 min to remove any precipitates before being applied to the column.

### **2.3.5 Densitometric analysis of OGDHc components on SDS-PAGE gels**

SDS-PAGE gels were digitalized using a Molecular Imager® Gel Doc™ XR+ System and analyzed with Image Lab software (Bio-Rad Laboratories Inc.). Densitometry analyses were performed using lane-based background subtraction, followed by measurement of the area under the peaks. The optical density was then used for statistical analysis. Each sample was analyzed on triplicated gels.

### **2.3.6 Isothermal Titration Calorimetry**

ITC measurements were performed with a VP-ITC microcalorimeter (MicroCal, Northampton, MA). Titrations were carried out in a buffer containing 50 mM potassium phosphate (pH 7.5), 100 mM KCl, 0.5 mM ThDP and 2 mM MgCl<sub>2</sub> at 15 °C. Prior to the experiments, both protein samples were dialyzed exhaustively against this buffer. To measure E1o binding to E2o, 450 μM of the E2o in the syringe was injected into the cell containing 25 μM of E1o (all based on subunit concentration). A one-site binding model was used for curve fitting, and the binding constant ( $1.0/K_d$ ) and other thermodynamic parameters were calculated with ORIGIN 5.0 software. provided by MicroCal.

### 2.3.7 Uniform $^{15}\text{N}$ labeling of human E1o

A single colony on the LB plate was inoculated in 20 mL LB medium containing 50  $\mu\text{g}/\text{ml}$  of ampicillin, and grown at 37 °C overnight. The next day, cells were collected by centrifugation (10 min, 4 °C) and washed with M9 minimal medium supplemented with 1g/liter of [ $^{15}\text{N}$ ]  $\text{NH}_4\text{Cl}$  to remove the traces of LB medium. The cells were collected by centrifugation, dissolved in 20 ml of the M9 medium, and inoculated into 800 ml of the M9 minimal medium supplemented with [ $^{15}\text{N}$ ]  $\text{NH}_4\text{Cl}$ . Cells were grown to an  $OD_{600}$  of 0.5-0.6 at 37 °C, then the temperature was lowered to 20 °C, and protein expression was induced by adding 0.50 mM IPTG for 16 h. Cells were collected by centrifugation at 2236  $\times g$  for 10 min and washed with 50 mM  $\text{KH}_2\text{PO}_4$  (pH 7.5) containing 100 mM NaCl and stored at 20 °C.

The  $^{15}\text{N}$ -labeled proteins were purified using Ni-Sepharose 6 Fast flow column with 10 to 300 mM imidazole gradient in 50 mM  $\text{K}_2\text{HPO}_4$  (pH 7.5) containing 500 mM NaCl, 0.5 mM ThDP, and 1.0 mM  $\text{MgCl}_2$ . The purified  $^{15}\text{N}$ -labeled proteins were dialyzed against 50 mM  $\text{KH}_2\text{PO}_4$  (pH 7.5) containing 150 mM NaCl, 0.50 mM ThDP, 1.0 mM  $\text{MgCl}_2$ , concentrated using a Centriprep 30 (Millipore) concentrating unit, and stored at 80 °C.

### 2.3.8 Two-dimensional $^1\text{H}$ - $^{15}\text{N}$ HSQC NMR

The NMR experiments were performed on a Varian INOVA 600 MHz spectrometer at 25 °C. For the HSQC NMR experiment, the 200  $\mu\text{M}$   $^{15}\text{N}$  labeled human E1o was exchanged to a buffer containing 50 mM  $\text{K}_2\text{HPO}_4$  (pH 7.5), 100 mM NaCl, 0.5 mM ThDP,

and 1.0 mM MgCl<sub>2</sub>, and 7 % D<sub>2</sub>O was added for a lock signal before the NMR spectrum was recorded.

### **2.3.9 Hydrogen/deuterium exchange mass spectrometry**

The H/D exchange analysis was conducted on an HPLC system (1200 series HPLC, Agilent) interfaced to a Fourier transform-mass spectrometer (Apex-ultra 70 hybrid Fourier transform-mass spectrometer, Bruker Daltonics), equipped with an in-house built refrigeration system as described previously [20, 21].

Prior to H/D exchange, E1o was exchanged into 10 mM potassium phosphate (pH 7.5) with 100 mM NaCl, 0.5 mM ThDP, and 1 mM MgCl<sub>2</sub>; E2o, E2o N-didomain, and E2o Core-domain were exchanged into 10 mM potassium phosphate (pH 7.5) with 100 mM NaCl; E3 was exchanged into 10 mM potassium phosphate (pH 7.5) with 100 mM NaCl and 2 μM FAD. All the binary complex study samples were prepared by mixing equal volume of 160 μM of each protein to a final concentration of 80 μM each. The samples were incubated at 20 °C for 1 h prior to proceeding to the H/D exchange experiments. The H/D exchange experiments were initiated by mixing 15 μl of the protein samples with 285 μl of D<sub>2</sub>O buffer resulting in a final concentration of 95 % D<sub>2</sub>O at pH 7.5. D<sub>2</sub>O buffer was prepared the same way as protein buffer except 99.9% D<sub>2</sub>O was used to dissolve the buffer components. The samples were incubated at 20 °C for an additional 20 s and 1, 3, 10, 30, 90, 270 min. The deuterium exchange was quenched by rapidly mixing 30 μl aliquots of the exchange reaction with 36 μl of ice-cold quench buffer (0.2 M potassium phosphate, 3 M guanidine hydrochloride, pH 2.1) to acidify the final sample pH to 2.5. The samples

were immediately frozen in liquid nitrogen and stored at  $-80\text{ }^{\circ}\text{C}$  before analysis. Undeuterated samples were generated following the same procedure except that protein samples were diluted into aqueous buffer and incubated for 5 min followed by the quench process. All experiments were run in triplicates.

The frozen deuterated sample was quickly thawed and loaded with an ice-cold syringe into a  $20\text{ }\mu\text{l}$  sample loop inside the refrigeration system. The protein sample (40 pmol) was carried by a  $0.2\text{ ml/min}$  digestion flow (0.1% formic acid) into an immobilized pepsin column (Poroszyme Immobilized Pepsin Cartridge,  $2.1 \times 30\text{ mm}$ , Applied Biosystems) and digested at  $15\text{ }^{\circ}\text{C}$  for 1 min. The resultant peptides were immediately cooled down to  $0\text{ }^{\circ}\text{C}$  through a heat exchanger and were concentrated and desalted on a peptide trap (Michrom Peptide MacroTrap,  $3 \times 8\text{ mm}$ ). The peptides were eluted and separated in 15 min through a reversed-phase C18 HPLC column (Agilent Poroshell 300SB-C18,  $2.1 \times 75\text{ mm}$ ) at a flow rate of  $0.2\text{ ml/min}$  with a  $0\text{ }^{\circ}\text{C}$  2–40% acetonitrile gradient containing 0.1% formic acid. ESI-Fourier transform-mass spectrometry (ESI-FTMS) measurements began 5 min after the initiation of the elution process and lasted 10 min. The time from initiation of digestion to elution of the last peptide was less than 20 min. Bruker Daltonics DataAnalysis 4.0 was used for spectrum analysis and data treatment. Peptides were identified from undeuterated samples by a customized program DXgest, which matches experimental peptide mass with theoretically generated peptic peptide mass by using statistical data for the pepsin cleavage pattern under H/D exchange conditions. Mass tolerance was set at 1.0 ppm. H/D exchange data for each individual peptide at various time points were processed using the program HX-Express to determine deuterium uptake and peak width. No correction was applied for back exchange. The number of

backbone amides, peptide coverage, and H/D heat map were generated by using MSTools. Microsoft Excel was used to produce time-dependent deuterium uptake plot, butterfly plot, and difference plot.

### 2.3.10 Bioinformatics methods

Alignment of multiple sequences was carried out using the Clustal Omega program (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) with default settings.

The secondary structure of E1o N-terminal region was calculated and predicted with JPred program (<http://www.compbio.dundee.ac.uk/jpred/>). To predict the possible 3D structure of the E1o N-terminal region, I-TASSER (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) and SWISS-MODEL (<http://swissmodel.expasy.org/>) were employed.

## 2.4 Results and Discussion

### 2.4.1 The activity of human E1o and the overall activity of *in vitro* reconstituted OGDHc

In the E1o-specific activity assay, the enamine intermediate produced on E1o from OG is oxidized by DCPIP with the formation of succinate and (reduced) DCPIPH<sub>2</sub>. By monitoring the reduction of the external oxidizing agent 2,6-DCPIP at 600 nm, an E1o specific activity of  $1.88 \pm 0.04 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg E1o}^{-1}$  ( $k_{\text{cat}} = 7.12 \text{ s}^{-1}$ ) was obtained under

our experimental conditions. This result is similar to the  $1.78 \pm 0.40 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg E1o}^{-1}$  reported by our group earlier [18].

The overall OGDHc activity of our *in vitro* reconstituted OGDHc was  $6.54 \pm 0.4 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg E1o}^{-1}$  ( $k_{\text{cat}} = 24.42 \text{ s}^{-1}$ ). This activity result is also consistent with our previously reported  $7.86 \pm 0.74 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg E1o}^{-1}$  ( $k_{\text{cat}} = 30 \text{ s}^{-1}$ ). Thus, our full length E1o was considered as a fully functioning enzyme. A detailed study of full length E1o protein was reported earlier from our group [18].

#### **2.4.2 The two E2o constructs**

The current understanding of the general organization of the 2-oxo acid dehydrogenase complex family (PDHc, OGDHc, and BCOADHc) recognizes a distinction between mammalian and bacterial PDHc and BCOADHc, but is less clear about the differences between human and bacterial OGDHc [22-25]. The E2 components contain one or more tandem lipoyl domains (each lipoyl domain is usually 80 amino acids in length) at their N-termini joined to a flexible linker region, which provides binding sites for the E1 and E3 components [26]. The C-terminal domain of the E2 components is called the core or catalytic domain and has the acyltransferase active sites, and is responsible for the formation of the core structures in different complexes (Figure 2.1). Although the sequence alignment of E2o found no obvious E1o binding or E3 binding sites in E2o, but as found in other E2 components, E2o also consists of lipoyl domain, flexible linker and catalytic core domain. Since the domain organization of E2o is similar to that of other E2 enzymes,

there can be a unique sequence still at the linker region which could be responsible for binding the E1o or E3.

With this hypothesis in mind, we constructed two truncated E2o proteins, one called E2o N-domain (residues 1-173) containing the N-terminal lipoyl domain, linker and the first twenty amino acids from the N-terminus of the core domain; and the second called E2o Core-domain (residues 144-386) containing the last ten amino acids from the linker and core domain of E2o. The two truncated E2o proteins would enable us to examine the response of various domains to its binding partner, and to derive more information on the assembly of the entire complex.

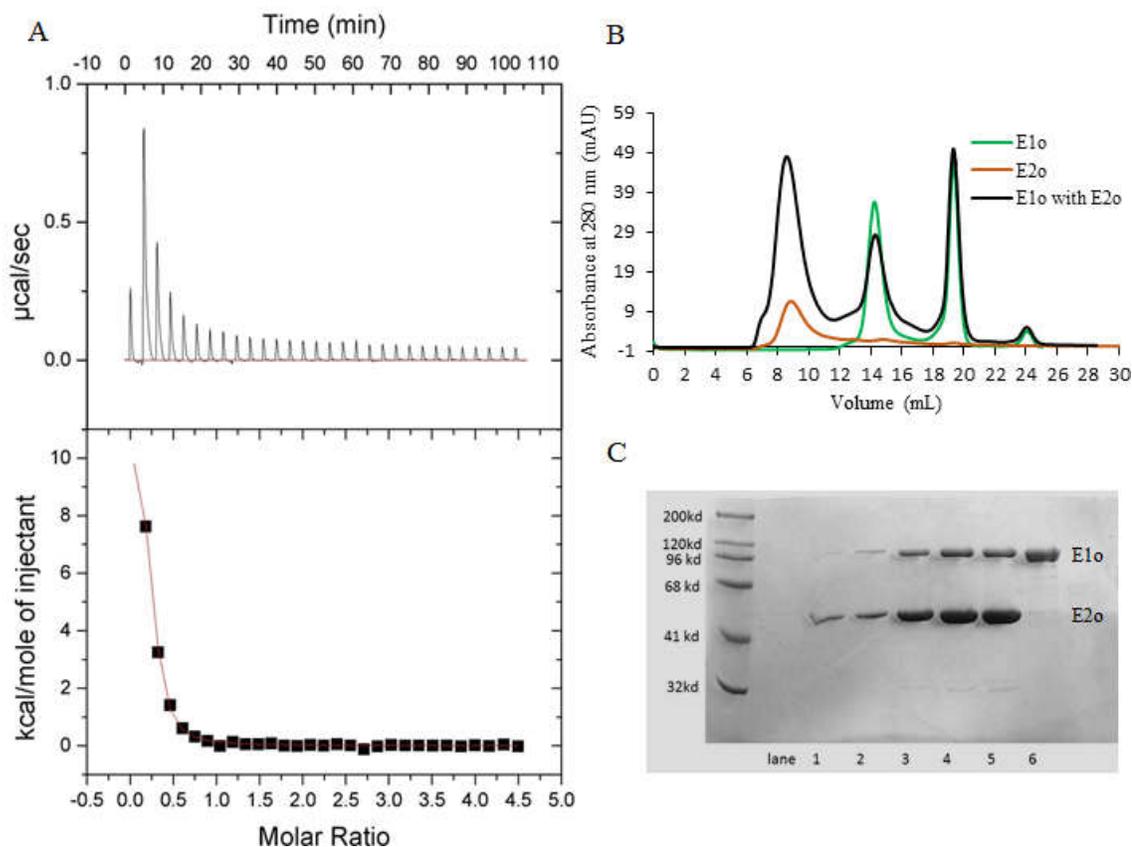
### **2.4.3 Investigation of a potential E1o to E2o interaction with gel filtration chromatography and ITC**

The direct interaction between E1o and E2o was initially investigated using gel filtration and ITC. Both enzymes were individually over-expressed and purified as described in the Methods section. After incubating E1o and E2o in 2:1 subunit ratio for 1 h, the mixture was quickly passed through a Superose 6 10/300 GL gel filtration column. Two separate peaks of eluted protein were observed (Figure 2.2 B). The fractions from both peaks were applied to SDS-PAGE (Figure 2.2 C), and revealed that the first peak eluting at the void volume contained both E2o and E1o, whereas the second peak contained only E1o, indicating that E1o and E2o bind each other forming a sub-complex. This result was also confirmed by ITC experiments, where the E1o dimer was titrated into the reaction cell containing the E2o, a binding isotherm was obtained (Figure 2.2 A) that could be fit to

an association constant and a calculated  $K_d$  of  $0.95 \pm 0.10 \mu\text{M}$  for the E1o/E2o interaction. However, under the same gel filtration experimental conditions, we did not get clear evidence for the interaction of the two truncated E2o proteins with E1o.

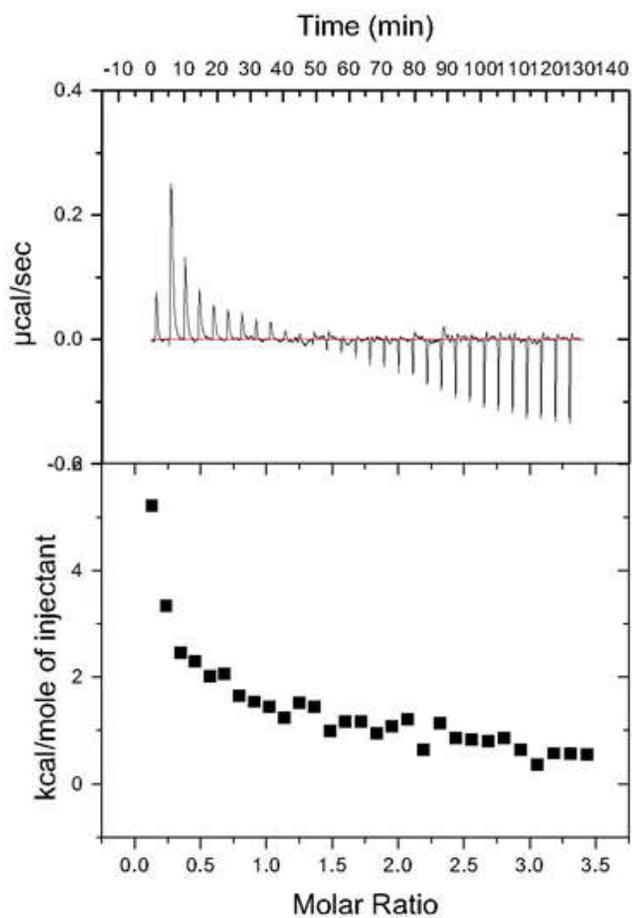
In parallel ITC studies, the E2o Core-domain was titrated into the E1o in the reaction cell (Figure 2.3). Significantly, the binding of the E2o Core-domain to E1o was much weaker than that of the entire E2o. A similar result was obtained when titrating E2o N-didomain into E1o. In addition, no clear new peak was identified from the gel filtration experiments when either truncated E2o protein was mixed with E1o.

Taken together, the gel filtration and ITC experiments established a clear association between E1o and E2o. With the ITC method, both truncated E2o proteins still displayed a weak interaction with E1o, however, the binding was greatly diminished, indicating that the linker-core region on E2o is possibly involved in the locus of interaction with E1o.



**Figure 2.2 ITC and GFC experiments probing the interaction of E1o and E2o.**

(A) ITC measurements of E1o binding to the E2o. The top panel is the thermogram that results from titrating the E1o dimer into the cell containing E2o. The bottom panel is the data fitting result as calculated by ORIGIN 5.0 with a single binding site model. (B) E1o and E2o were pre-incubated in a 2:1 subunit molar ratio for 1 h at room temperature followed by separation on a Superose 6 10/300 GL column. The E2o eluted at the void volume (orange line) indicating that it is a pre-assembled 24- or 60-meric enzyme. (C) SDS-PAGE analysis of the eluted protein peaks (10 % gel with Coomassie Brilliant Blue staining). Lane 1, small amount of aggregated E2o from 6-7 ml fraction (black line); Lane 2-5, fractions from 7 to 12 (black line). Lane 6, fraction from 13-15 ml fraction (black line).



**Figure 2.3 ITC measurements of the binding of the E2o Core-domain to E1o.**

The top panel is the raw data that results from titrating the E2o Core-domain (trimer) into the cell containing E1o dimer.

#### **2.4.4 HDX-MS experiments**

Protein secondary and tertiary structures can be affected when the protein interacts or binds with other proteins or small molecules. The additional interaction or binding partner to a protein may produce changes in the conformation and dynamics, not just in the vicinity of the binding site, but can also lead to global effects. These conformational and dynamic changes are mostly related to protein function or enzyme activity. Here, we used the HDX-MS to study the interactions of E1o with E2o and two truncated E2o proteins. This method allowed us to monitor both conformational changes of both enzymes at peptide resolution, and identify the regions affecting binding to each other of both the E1o and E2o.

#### **Overview of E1o and E2o HDX behaviors**

As the first step to understand the interaction of E1o with E2o, HDX-MS experiments were conducted with full-length proteins. The time dependent HDX-MS experiments of E1o and E2o were carried out at 20 s, 1, 3, 10, 30, 90, 270 min. On-line digestion by pepsin followed by LC-MS analysis under our HDX conditions resulted in 66 peptides for E1o, and 19 peptides for E2o, many of them partially overlapping, providing 94 % sequence coverage for E1o, and 66.75 % sequence coverage for E2o (Tables 2.1-2.3). The high sequence coverage on E1o, including the N-terminal flexible region 1-77, which was missing from previous crystallographic study of a bacterial E1o, provided access to study the structural changes throughout the protein during its assembly. Although E2o has a good sequence coverage on its catalytic core domain, information about many peptides

was missing from E2o lipoyl domain and linker region. The low sequence coverage on these two regions of E2o is not surprising as our sequence analysis on lipoyl domain and linker region showed, they tend to digest poorly with pepsin as they are relatively depleted of hydrophobic residues and enriched in proline (26 prolines in these two regions) and charged residues [27]. To overcome the low sequence coverage problem on E2o, later we used two truncated E2o proteins to examine the role of these two regions in the interaction with E1o or E3.

A summary of all peptide data on E1o and E2o is provided in the deuterium uptake curves in Fig. 2.4. The following discussion will focus on that subset of peptides which displayed statistically significant difference in deuterium uptake upon binding.

**Table 2.1 HDX-MS sequence coverage**

	E1o	E2o	E2o N-didomain	E2o Core-domain	E3
sequence coverage	94%	66.75%	55.44%	100%	90.93%

**Table 2.2 Peptides identified from pepsin digestion of human E1o.**

Peptides			Monoisotopic Mass		Error (ppm)
			[M+H] <sup>+</sup> (Da)		
No.	Position	Sequence	Experimental	Theoretical	
1	2-8	SAPVAAE	644.3248	644.3250	-0.2
2	9-17	PFLSGTSSN	909.4309	909.4312	-0.3
3	18-22	YVEEM	670.2756	670.2753	0.5
4	22-32	MYCAWLENPKS	1341.5956	1341.5966	-0.8
5	34-47	HKSWDIFFRNTNAG	1692.8260	1692.8241	1.1
6	41-65	FRNTNAGAPPGTAYQSPLPLSRGSL	2572.3259	2572.3267	-0.3
7	66-74	AAVAHAQSL	867.4678	867.4683	-0.5
8	75-84	VEAQPNVDKL	1112.5938	1112.5946	-0.7
9	85-89	VEDHL	612.2986	612.2988	-0.2
10	90-94	AVQSL	517.2980	517.2980	0.0
11	95-114	IRAYQIRGHHVAQLDPLGIL	2270.2893	2270.2880	0.6
12	115-135	DADLDSSVPADIISSTDKLGFL	2166.0460	2166.0448	0.5
13	136-143	YGLDESDL	911.3996	911.3993	0.3
14	144-151	DKVFHLPT	956.5211	956.5200	1.1
15	153-161	TFIGGQESA	909.4305	909.4312	-0.8
16	163-187	PLREIIRLEMAYCQHIGVEFMFIN	3078.5865	3078.5838	0.9
17	186-207	INDLEQCQWIRQKFETPGIMQF	2724.3279	2724.3273	0.2
18	207-216	FTNEEKRTLL	1250.6742	1250.6739	0.2
19	220-228	VRSTRFEEF	1170.5917	1170.5902	1.3

20	229-245	LQRKWSSEKRFGLGCE	2053.0306	2053.0284	1.1
21	246-264	VLIPALKTIIDKSSSENGVD	2012.1267	2012.1274	-0.3
22	265-291	YVIMGMPHRGRNLNVLNVIRKELEQIF	3196.7582	3196.7598	-0.5
23	292-326	CQFDSKLEAADEGSGDVKYHLGMYHRRINRVTDNR	4080.9581	4080.9515	1.6
24	327-331	ITLSL	546.3501	546.3497	0.7
25	332-338	VANPSHL	737.3941	737.3941	0.0
26	339-354	EAADPVVMGKTKAEQF	1720.8570	1720.8574	-0.3
27	355-375	YCGDTEGKKVMSILLHGDAAF	2255.0829	2255.0835	-0.3
28	376-385	AGQGIVYETF	1084.5312	1084.5310	0.3
29	386-406	HLSDLPSYTTHTGTVHVVVNNQ	2318.1550	2318.1524	1.1
30	407-434	IGFTTDPRMARSSPYPTDVARVVNAPIF	3078.5865	3078.5830	1.2
31	435-445	HVNSDDPEAVM	1213.5153	1213.5154	-0.1
32	446-453	YVCKVAAE	882.4385	882.4390	-0.5
33	454-466	WRSTFHKDVVVDL	1601.8450	1601.8435	1.0
34	467-481	VCYRRNGHNEMDEPM	1850.7719	1850.7731	-0.7
35	482-487	FTQPLM	736.3700	736.3698	0.2
36	488-504	YKQIRKQKPVLQKYAEL	2133.2548	2133.2543	0.2
37	505-528	LVSQGVVNQPEYEEEISKYDKICE	2799.3407	2799.3393	0.5
38	529-565	EAFARSKDEKILHIKHWLDSPPWPGFFTLDGQPRMSC	4330.1352	4330.1325	0.6
39	566-573	PSTGLTED	819.3732	819.3731	0.2
40	574-589	ILTHIGNVASSVPVEN	1649.8873	1649.8857	1.0
41	590-596	FTIHGGL	744.4050	744.4039	1.4
42	597-616	SRILKTRGEMVKNRTVDWAL	2373.3207	2373.3183	1.0
43	620-625	MAFGSL	625.3010	625.3014	-0.7
44	626-638	LKEGIHIRLSGQD	1465.8143	1465.8122	1.5
		VERGTFSHRHHVLHDQNVDKRTCIPMNLWPNQAPYT			
45	639-678	VCN	4750.2876	4750.2837	0.8
46	682-687	SEYGVL	667.3298	667.3297	0.1
47	688-693	GFELGF	669.3250	669.3243	1.1
48	694-701	AMASPNAL	774.3812	774.3815	-0.3
49	702-708	VLWEAQF	892.4561	892.4563	-0.3
50	715-734	AQCIIDQFICPGQAKWVRQN	2318.1550	2318.1533	0.8
51	736-759	IVLLPHGMGMEGPEHSSARPERF	2660.3449	2660.3436	0.5
52	760-770	LQMCNDDPDVL	1262.5412	1262.5392	1.6
53	771-778	PDLKEANF	933.4684	933.4676	0.8

54	779-783	DINQL	602.3142	602.3144	-0.3
55	784-788	YDCNW	700.2408	700.2395	1.8
		VVVNCSTPGNFFHVLRRQILLPFRKPLIIFTPKSLLRHPE			
56	789-833	ARSSF	5230.9478	5230.9455	0.4
57	834-842	DEMLPGTHF	1046.4593	1046.4612	-1.7
58	843-864	QRVIPEDGPAAQNPENVKRLLF	2491.3439	2491.3416	0.9
59	865-887	CTGKVYYDLTRERKARDMVGQVA	2659.3447	2659.3443	0.1
60	888-902	ITRIEQLSPFPDLL	1788.9894	1788.9894	0.0
61	903-914	LKEVQKYPNAEL	1431.7835	1431.7842	-0.5
62	915-928	AWCQEEHKNQGYD	1770.7188	1770.7177	0.7
63	929-945	YVKPRLRTTISRAKPVW	2071.2293	2071.2287	0.3
64	946-967	YAGRDPAAPATGNKKTHLTEL	2282.1902	2282.1888	0.6
65	968-975	QRLLDTAF	963.5263	963.5258	0.6
66	976-992	DLDFVKNFSLEHHHHHH	2149.0136	2149.0111	1.2

---

**Table 2.3 Peptides identified from pepsin digestion of human E2o.**

Peptides			Monoisotopic Mass		Error (ppm)
			[M+H] <sup>+</sup> (Da)		
No.	Position	Sequence	Experimental	Theoretical	
1	11-22	AESVTEGDVRWEKAVGDTVAED	2363.0979	2363.0997	-0.8
2	58-74	LLVPDGGKVEGGTPLFT	1699.9265	1699.9265	-0.8
3	103-132	AAVPPAAPIPTQMPPVSPSQPPSGKPVS	2899.5352	2899.5386	-1.2
4	180-185	LTFNE	724.3504	724.3512	-1.2
5	186-193	IDMSNIQE	949.4283	949.4295	-1.3
6	194-212	MRARHKEAFLKKHNLKLG	2324.3281	2324.3285	-0.1
7	216-222	FVKASAF	769.4240	769.4243	-0.4
8	223-231	ALQEQPVVN	997.5296	997.5313	-1.7
9	232-240	AVIDDTTKE	991.4935	991.4942	-0.7
10	241-248	VVYRDYID	1042.5207	1042.5204	0.3
11	249-261	ISVAVATPRGLVV	1281.7888	1281.7889	-0.1
12	262-271	PVIRNVEAMN	1142.5976	1142.5987	-0.9
13	272-291	FADIERTITELGEKARKNEL	2333.2448	2333.2459	-0.5
14	292-300	AIEDMDGGT	908.3656	908.3665	-1.0
15	301-312	TISNGGVFGSL	1051.5407	1051.5416	-0.9
16	313-332	FGTPIINPPQSAILGMHGIF	2110.1135	2110.1154	-0.9
17	333-349	DRPVAIGGKVEVRPMMY	1918.0019	1918.0037	-0.9
18	350-367	VALTYDHRLIDGREAVTF	2076.0869	2076.0873	-0.2
19	368-382	LRKIKAAVEDPRVLL	1721.0791	1721.0796	-0.3

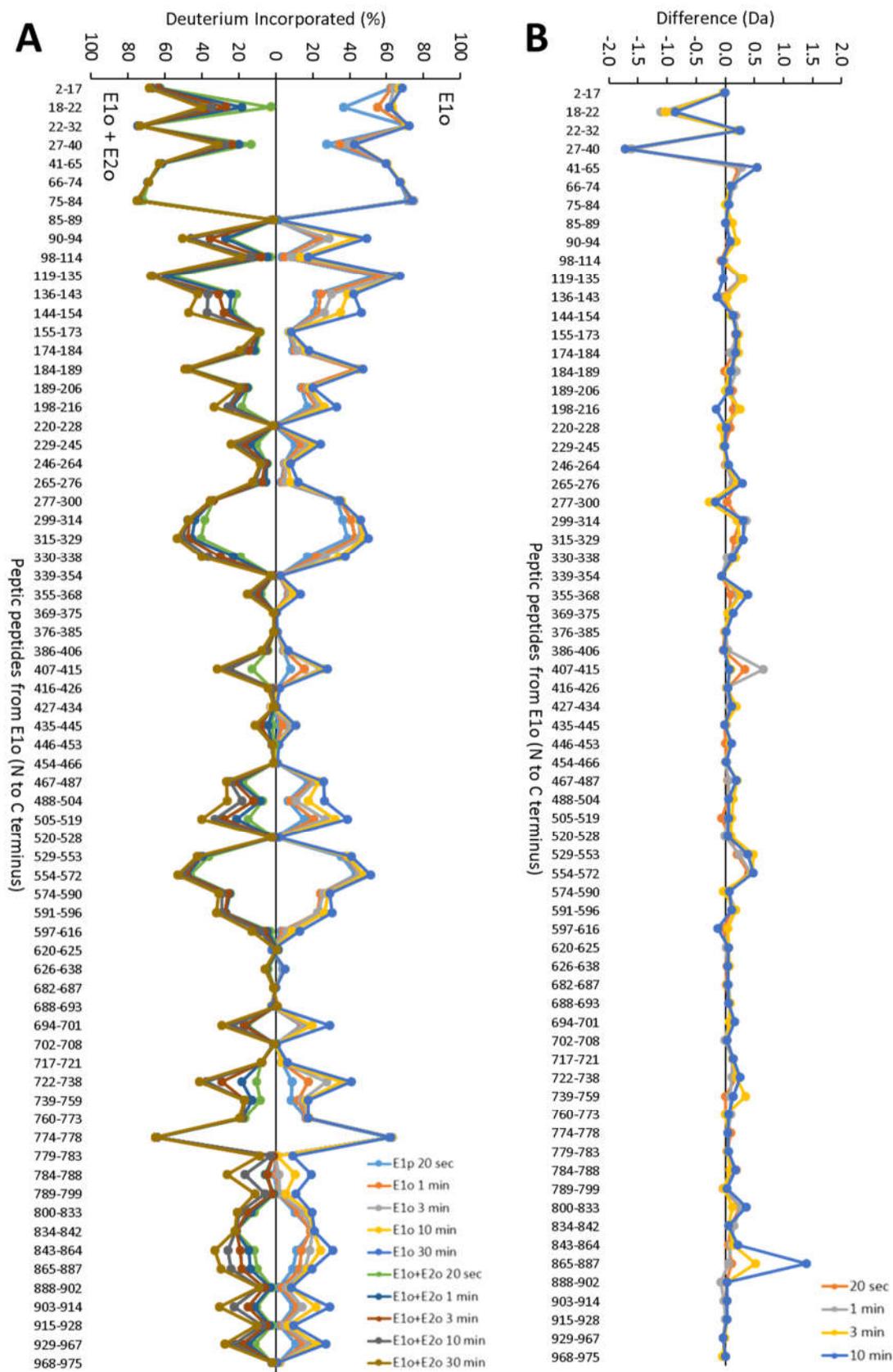
### **Deuterium uptake changes in E1o on complexation with E2o.**

Structural changes of the E1o with two truncated E2o proteins upon formation of binary complexes were determined with HDX-MS. The levels of deuterium uptake of uncomplexed proteins were compared with its mixtures with its binding partner after 0, 0.33, 1, 3, 10, 30, 90, and 270 min.

Our HDX-MS results indicate that the interaction of E1o with E2o did not induce large global conformational changes in E1o (Figure 2.4 B). However, local backbone amide proton perturbations were observed. Two peptides 18-22 and 27-40 were the unique regions which experienced significant deuterium uptake retardation. Surprisingly, there are also some peptides that show insignificant increased deuterium uptake upon interaction with E2o. The higher deuteration level peptides are mostly located at the ThDP binding and Mg binding regions present in all ThDP enzymes (residues 301-338, 407-415, and 529-572). These small perturbations in the regions around the ThDP and Mg binding sites could be related mechanistically to perturbations in the E1o catalytic center in response to direct binding to E2o.

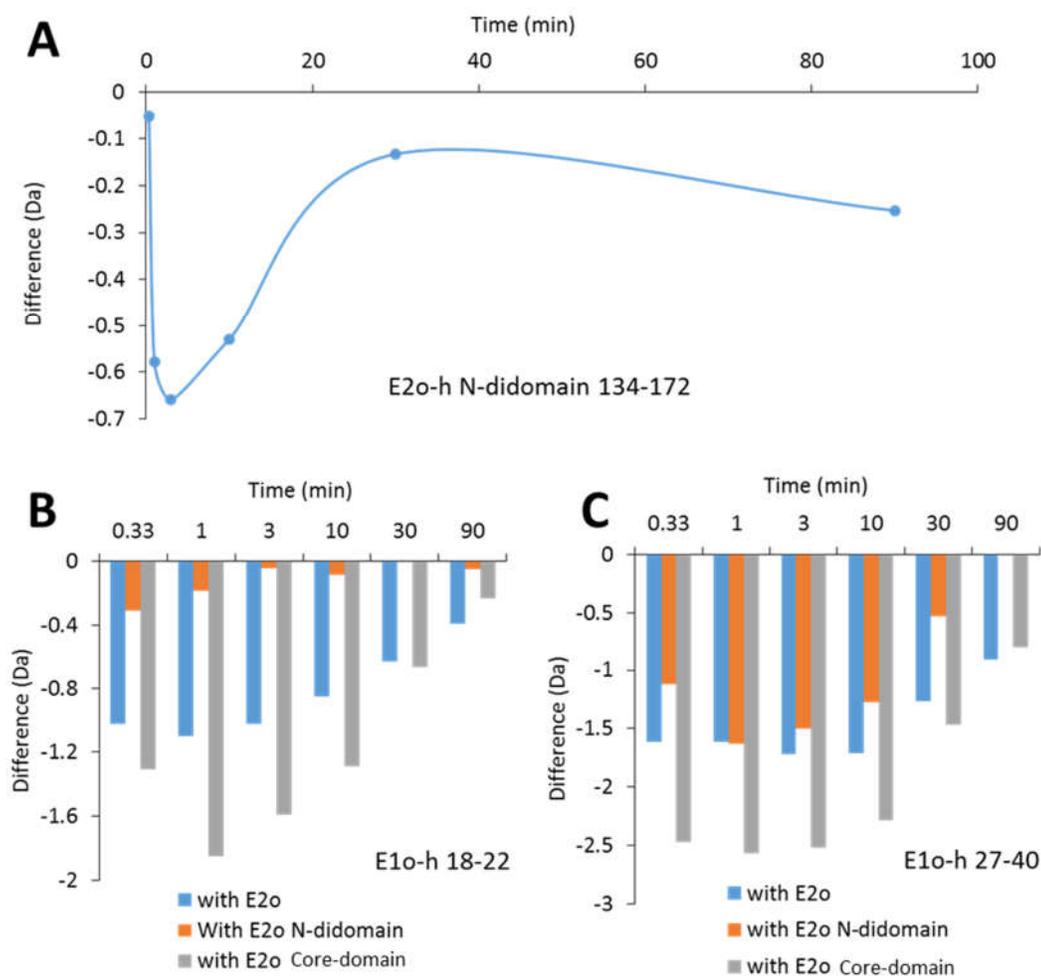
Two peptides in the N-terminal domain of E1o displayed a decrease in deuterium uptake when incubated with E2o (Figure 2.4), E2o N-domain, and E2o Core-domain at all time points (Figure 2.5 B and 2.5 C). Thus, the peptides 18-22 and 27-40 on E1o were always protected by E2o and its two truncated proteins. These regions therefore represented the most likely candidates for binding to E2o. The HDX kinetics for peptides 18-22 and 27-40 show a decreased rate of deuterium incorporation in the E1o/E2o complex compared to free E1o consistent with the loss of conformational dynamics upon interaction with E2o (Figure 2.4 B).

This result, together with the predicted secondary structure of the E1o N-terminal region (Figure 2.6), suggested that upon binding to E2o, the predicted two helical regions 18-22 and 27-40 in the N-terminal region of E1o tend to face to the E2o, while the two peptides 22-32 and 41-65 next to the two helical regions will bind to the opposite face of the E2o resulting in increased deuterium uptake due to the small-scale conformational change. A previous crystallographic study of *E. coli* E1o reported that it could not be crystallized unless the N-terminal region (first 77 N-terminal amino acids) was removed [28]. This X-ray study suggested that the N-terminal region of E1o is highly dynamic or with natively-disordered periphery. The HDX-MS results showed that the first 84 amino acids from the N-terminal region of E1o are highly solvent accessible, with more than 60 % deuteration within 1 min except the peptide 27-40. These data also indicated significant flexibility of these regions with transient opening of hydrogen bonds.



**Figure 2.4 Comparative HDX-MS analysis of the interaction of E1o and E2o**

(A) A butterfly plot representing average relative deuterium incorporation percentage (y axis) (deuterons exchanged/maximum exchangeable amides  $\times$  100%) of peptic fragments from E1o (x axis, listed from N to C terminus) in the absence of E2o (top) versus in the presence of E2o (bottom) based on three independent experiments. (B) Difference plot showing deuterium incorporation changes of peptic fragments of E1o in the absence and presence of E2o (deuterons exchanged in the absence of E2o minus deuterons exchanged in the presence of E2o).



**Figure 2.5 Deuterium uptake difference plots.**

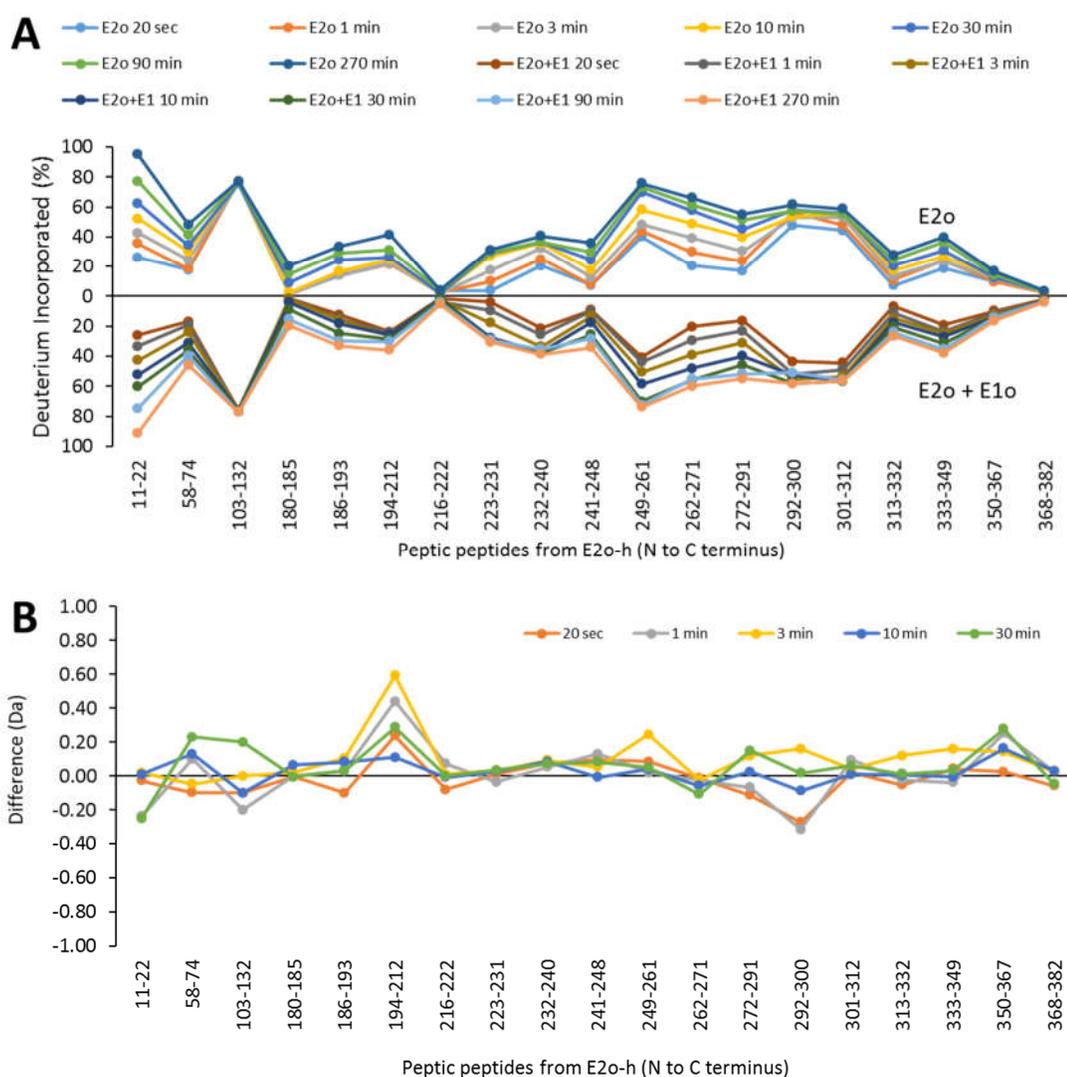
(A) Difference plot of E2o N-didomain 134-172 upon its incubation with E1o. (B) E1o 18-22 deuterium uptake difference when incubated with E2o, E2o N-didomain and E2o core. (C) E1o 27-40 deuterium uptake difference when incubated with E2o, E2o N-didomain and E2o core.



**Deuterium uptake change in E2o on interaction with E1o.**

On the E2o partner, upon binding to E1o, no significant decrease in deuterium uptake was observed (Figure 2.7). However, due to a lack of pepsin cleavage site in the linker-core region on E2o, this region was not well covered in the HDX-MS experiments.

Although we could not obtain high sequence coverage for E2o, the core domain of E2o was very well covered under our experimental conditions. A butterfly plot of E2o (Figure 2.7 A) revealed several regions, which were highly protected in the core domain. Residues 180-185, 216-222, and 368-382 had almost no deuterium uptake on the time scale of exchange used. Difference plot of E2o (Figure 2.7 B) revealed changes of deuterium incorporation of each peptide upon interaction with E1o. As shown in Figure 2.7, there was no significant decrease in deuteration on the E2o core domain. A Pfam database search on E2o showed that the residues 153-384 are in a highly conserved region belonging to the 2-oxo acid dehydrogenase acyltransferase (catalytic domain) family. Together with the HDX-MS results, the E2o core domain is unlikely to be the region which could directly bind to E1o.



**Figure 2.7 Comparative HDX-MS analysis of E2o in the absence and presence of E1o.**

A, butterfly plot representing average relative deuterium incorporation percentage of peptic fragments from E2o in the absence and presence of E1o. B, difference plot showing deuterium incorporation changes of peptic fragments of E2o in the absence and presence of E2o (deuterons exchanged in the presence of E1o minus deuterons exchanged in the absence of E1o).

Our preliminary GFC and ITC results suggested that the linker-core region on E2o contains the binding site for E1o. To further investigate the role of this region in the association of E1o and E2o, the two truncated E2o proteins, which divide this region into two parts, were used for the next HDX-MS study.

Upon binding to E1o, the C-terminal peptide 134-172 in the E2o N-didomain revealed a significant decrease in deuterium uptake (Figure 2.5 A); however, this deuterium uptake retardation quickly faded away after 3 min exchange time. This weaker protection suggested the involvement of this region in E1o/E2o interactions. In addition, this result agreed with the previous GFC and ITC results (Figure 2.2 and 2.3) suggesting much weaker binding when examining the interaction of the truncated E2o proteins with E1o. Similar results were obtained when determining the interaction of the E2o Core-domain with E1o. Upon binding to E1o, the N-terminal peptide 1-33 of the E2o Core-domain also displayed a significant decrease in deuterium uptake level. Residues 134-172 from E2o N-didomain and residues 1-33 from E2o Core-domain overlap significantly (both peptides contain the same 30 amino acids AEAGAGKGLRSEHREKMNRMRQRIAQRLKE). Since this is the only region that displayed significant decrease in deuterium uptake level upon E1o binding, this region could be critical for binding to E1o. A sequence alignment of this region to some known binding domain sequences shows that a number of residues from this region are well aligned to a part of known binding domains (Figure 2.8). This finding suggested that E2o shares some, but not all sequence features of other PSBDs that are thought important for E1 binding. In addition, in contrast to other E2's, the binding locus of E2o also involved some residues from its core domain. The different position of this binding locus on E2o might be related to its unique selectivity and specificity of interaction.

CLUSTAL O(1.2.1) multiple sequence alignment

```

Odo2_human      -----AEAGAGKGLRSEHREKMNRMRQRIAQRLKE
Odp2_Ecoli      ATPLIRRLAREFGVNLAKVKGTGRKGRIIFEDVQAYVKEAI-----
Odb2_human      ATPAVRRLAMENNIKLSVVGSGKDGRILKEDILNYLEKQT-----
Odp2_human      VSPLAKKLAVEKGIDLTQVKGTGPDGRITKDKIDSFV-----
Odp2_Bacst      AMPSVRKYAREKGVDIRLVQGTGKNGRVLKEDIDAFLAGGA-----
Odb2_Psepu      ASPAVRKRALDAGIELRYVHGSGPAGRIIHEDLDAFMS-----
Odo2_Ecoli      LSPAIRRLLAEHNLDAISAIKGTGVGGRLTREDVEKHLAKAP-----
Odo2_Azovi      LSPAARKIAEENAIADSITGTGKGGRVTKEDAVAAAE-----
                *:*  *  :.:

```

**Figure 2.8** Sequence alignment of known binding domains with the linker-core region peptide of E2o identified from HDX-MS.

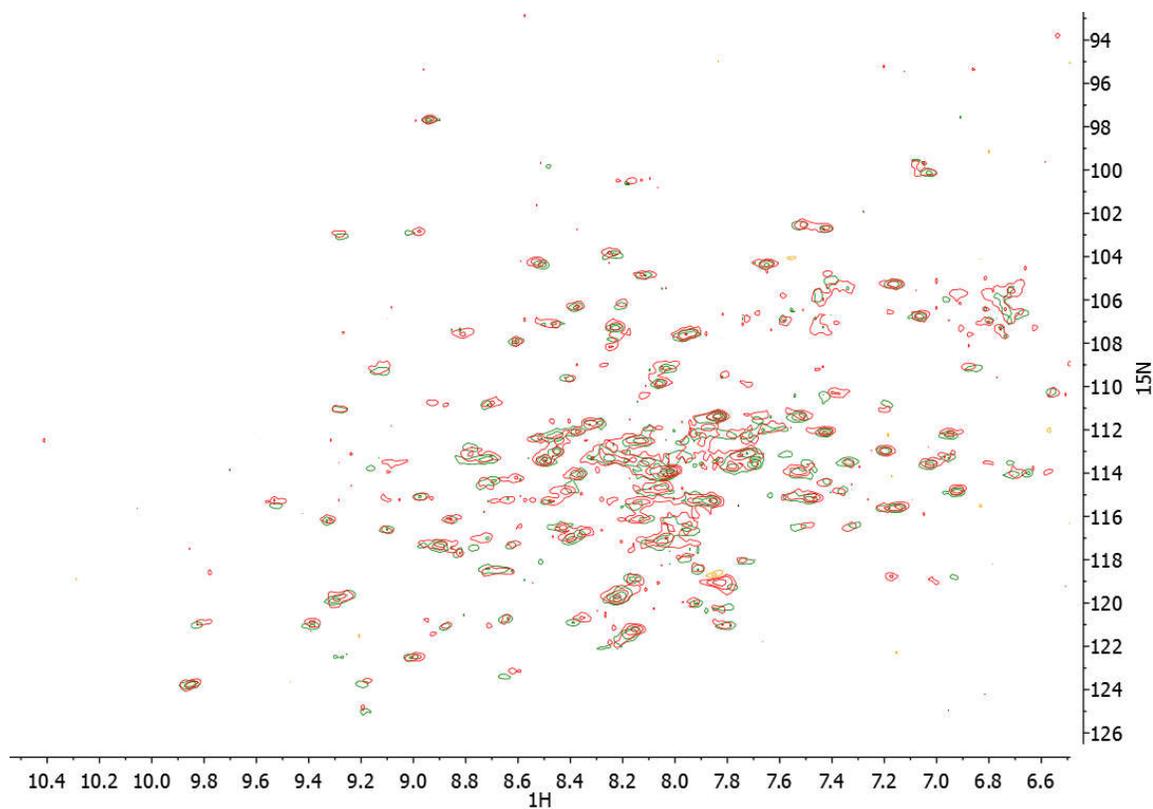
Odp2, pyruvate dehydrogenase (PDH); Odo2, 2-oxoglutarate dehydrogenase (OGDH); Odb2, branched chain 2-oxo acid dehydrogenase (BCDH); Bacst, *Bacillus stearothermophilus*; Azovi, *Azotobacter vinelandii*; Ecoli, *Escherichia coli*; Psepu, *Pseudomonas putida*. Boxed residues are possibly involved in binding to E1o. Our sequence alignment result is in good agreement with previous sequence analysis, and the boxed residues are assigned residues in known binding domains, which are possibly involved in binding to E1o [28].

These HDX-MS experiments demonstrated formation of an E1o/E2o sub-complex, the HDX-MS data indicated the binding region on each protein, and further confirmed the flexibility of the N-terminal region in E1o. The observation that the linker-core region of E2o is the critical locus interacting with E1o does raise the question whether this region on E2o also serves as a binding site for E3? Is there a new mode of assembly, which can differentiate OGHDC assembly from PDHC and BCOADHC assembly?

#### **2.4.5 NMR evidence for the interaction of E1o with E2o.**

To further confirm the direct interaction between E1o and E2o, we have performed a standard  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiment. In this method, the magnetization is transferred from hydrogen to attached  $^{15}\text{N}$  nuclei via the J-coupling, and only those protons correlated with  $^{15}\text{N}$  can be detected. The protein  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectrum mainly reflects changes in backbone amide groups, however, the Trp, Asn, and Gln sidechain N-H groups are also visible. The spectrum is more like a fingerprint for a  $^{15}\text{N}$  labeled protein. By simply comparing the chemical shift of each spot, one can assess whether a protein-protein interaction is likely to be present.

For this purpose, we have conducted  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments using  $^{15}\text{N}$  isotope labeled E1o with the experimental conditions described under Methods. The superimposed  $^{15}\text{N}$  E1o NMR spectrum collected in the presence and absence of E2o is shown in Figure 2.9. Upon binding with E2o, multiple differences were observed on the  $^{15}\text{N}$  E1o NMR spectrum. This is clear evidence for the binding of E1o and E2o. While at this level of analysis (no sequence-specific assignments were made), these NMR results do not provide information about the locus of interaction on E1o, the evidence for the existence of interaction with E2o is supported.



**Figure 2.9**  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra of E1o in the absence and presence of E2o.

Superimposed spectra of uniformly  $^{15}\text{N}$ -labeled E1o in green, and uniformly  $^{15}\text{N}$ -labeled E1o with E2o in red. The  $^{15}\text{N}$  labeled E1o was run at 200  $\mu\text{M}$  in a buffer containing 50 mM  $\text{K}_2\text{HPO}_4$  (pH 7.5), 100 mM NaCl, 0.5 mM ThDP, and 1.0 mM  $\text{MgCl}_2$ , and 7 %  $\text{D}_2\text{O}$  was added for a lock signal before the NMR spectrum was recorded. The  $^{15}\text{N}$  labeled E1o (200  $\mu\text{M}$ ) with E2o (300  $\mu\text{M}$ ) sample was run at the same conditions as the  $^{15}\text{N}$  labeled E1o.

## 2.5 Conclusions

Since sequence alignment of E2o indicated the absence of any obvious E1o and E3 binding domains, the assembly and integration of mammalian OGDHc is still under debate and speculation [13, 17]. Previous selective proteolysis studies [29] and immunological analysis [17] suggested that the mammalian E1o N-terminal region has limited sequence and structural similarity with human E3BP, which is important in maintaining the structural integrity of OGDHc.

Due to the difficulties in generating a functional recombinant full-length E1o, a GST fused to the E1o N-terminal region was used to study the interaction with E2o and E3 [19]. It was found that the post-translationally expressed E1o N-terminal region was unable to associate with E2o, and the association was only observed when these two proteins were expressed co-translationally.

In this study, the full length active E1o was successfully expressed, purified, and characterized. The interaction between human E1o and E2o was first characterized with gel filtration and ITC methods. A very stable E1o/E2o sub-complex was detected, however, the interactions of E1o with the two truncated E2o proteins are likely to be considerably weaker than with the full length E2o. The HDX-MS analysis of intact components, as well as of the two truncated E2o proteins enabled identification of the sub-complex binding loci. Two peptides 18-22 and 27-40 from the E1o N-terminal region were identified as the direct binding regions to E2o. The HDX-MS results also showed that conformational changes occur around the ThDP and  $Mg^{2+}$  binding regions of E1o upon binding to E2o. These conformational changes might reflect the more favored conformation for the pre-

decarboxylation state promoting substrate binding and product release, which was induced by the lipoyl domain on E2o.

On the E2o partner of the binary complex, while we did not obtain high sequence coverage on full length E2o, the two truncated E2o proteins allowed us to recover some of the important regions which are critical for E1o binding. Our observations strongly suggest that the linker-core region (residues 134-172) on E2o is essential for E1o binding. Sequence alignments revealed that a sequence located in this region displays some similarities to the known binding domain sequences; this finding suggests that the E1o/E2o interaction might share some common features with more ‘conventional’ E1/E2 binding loci. Even though there are some similarities in sequence, the E1o/E2o interaction should be regarded as a different type of binding in terms of its unique E1o binding region, and the extended binding sequence to the core domain of E2o. Lastly, our  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR experiments also confirmed the interaction between E1o and E2o.

In summary, our current findings provide the first evidence that the E1o N-terminal region residues 18-22 and 27-40 are responsible to E2o binding, and that the E2o linker-core region is critical for E1o binding. These findings are also consistent with earlier studies in mammalian OGDHc suggesting the E1o N-terminal region to be important for complex structural integrity [17]. Moreover, the E1o/E2o sub-complex was found to be very stable, suggesting that the entire E1o/E2o might serve as a core structure of the OGDHc. Importantly, the identified binding domain sequence on both enzymes are different from previous known binding domain sequences, suggesting a novel binding model for the human E1o/E2o complex.

In future work, it will be informative to explore the linker-core binding region on E2o. We have used the two truncated E2o proteins to discover the linker-core region for E1o binding, however, this approach might impair the structural integrity of this region. Given the fact that, compared with full length E2o, both truncated E2o proteins exhibited weaker interactions with E1o, it is possible that both truncated proteins lose some important structural features, which are also important in binding. Moreover, further experiments on HDX-MS using other pepsin-like enzymes to obtain a better sequence coverage on E2o, especially of the lipoyl domain and linker region, will be very helpful to better understand the assembly of OGDHc.

## References

1. Gibson, G. E., Park, L. C. H., Sheu, K. F. R., Blass, J. P., and Calingasan, N. Y. (2000) The  $\alpha$ -ketoglutarate dehydrogenase complex in neurodegeneration. *Neurochem. Int.* *36*, 97–112.
2. Yeaman, S. J. (1986) The mammalian 2-oxoacid dehydrogenases: a complex family. *Trends Biochem. Sci.* *11*, 293-296.
3. Randle, P. J. (1983) Mitochondrial 2-oxoacid dehydrogenase complexes of animal tissues. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *302*, 47-57.
4. Hackert, M. L., R. M. Oliver and Reed, L. J. (1983) Evidence for a multiple random coupling mechanism in the alpha-ketoglutarate dehydrogenase multienzyme complex of *Escherichia coli*: a computer model analysis. *Proc. Natl. Acad. Sci. U. S. A.* *80*, 2226-30.
5. Behal, R.H., DeBuysere, M.S., Demeler, B., Hansen, J.C., and Olson, M.S. (1994) Pyruvate dehydrogenase multienzyme complex. Characterization of assembly intermediates by sedimentation velocity analysis. *J Biol Chem.* *269*, 31372-31377.
6. Hiromasa, Y., T. Fujisawa, Y. Aso and Roche, T. E. (2004) Organization of the Cores of the Mammalian Pyruvate Dehydrogenase Complex Formed by E2 and E2 Plus the E3-binding Protein and Their Capacities to Bind the E1 and E3 Components. *J. Biol. Chem.* *279*, 6921-6933.
7. Oliver, R. M. and Reed L. J. (1982) Multienzyme complexes. *Electron Microscopy of Proteins.* J.R. Harris. *2*, 1-48.
8. Wagenknecht, T., R. Grassucci and Schaak, D. (1990) Cryoelectron microscopy of frozen-hydrated alpha-ketoacid dehydrogenase complexes from *Escherichia coli*. *J. Biol. Chem.* *265*, 22402-22408.
9. Perham, R. N. (2000) Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* *69*, 961-1004.
10. AEvarsson, A., Chuang, J.L., Wynn, R.M., Turley, S., Chuang, D.T., and Hol, W.G. (2000) Crystal structure of human branched-chain alpha-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease. *Structure* *8*, 277-291.
11. Reed, L.J., and Hackert, M.L. (1990) Structure-function relationships in dihydrolipoamide acyltransferases. *J. Biol. Chem.* *265*, 8971–8974.

12. Perham, R. N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry*. *30*, 8501-8512.
13. Bradford, A. P., A. Aitken, F. Beg, K. G. Cook and Yeaman S. J. (1987) Amino acid sequence surrounding the lipoic acid cofactor of bovine kidney 2-oxoglutarate dehydrogenase complex. *FEBS Lett*. *222*, 211-214.
14. Nakano, K., Takase, C., Sakamoto, T., Nakagawa, S., Inazawa, J., Ohta, S., and Matuda, S. (1994) Isolation, Characterization and Structural Organization of the Gene and Pseudogene for the Dihydrolipoamide Succinyltransferase Component of the Human 2-Oxoglutarate Dehydrogenase Complex. *Eur. J. Biochem*. *224*, 179-189.
15. Kresze, G. B., Ronft, H., Dietl B., and Steber, L. (1981) Limited proteolysis of 2-oxoglutarate dehydrogenase multienzyme complex from bovine kidney. *FEBS Lett*. *127*, 157-160.
16. Rice, J. E., Dunbar B., and Lindsay, J. G. (1992) Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. *The EMBO J*. *11*, 3229-3235.
17. McCartney, R. G., Rice, J. E., Sanderson, S. J., Bunik, V., Lindsay, H. and Lindsay, J. G. (1998) Subunit interactions in the mammalian alpha-ketoglutarate dehydrogenase complex. Evidence for direct association of the alpha-ketoglutarate dehydrogenase and dihydrolipoamide dehydrogenase components. *J. Biol. Chem*. *273*, 24158-24164.
18. Nemeria, N.S., Ambrus, A., Patel, H., Gerfen, G., Adam-Vizi, V., Tretter, L., Zhou, J., Wang, J., and Jordan, F. (2014) Human 2-oxoglutarate dehydrogenase complex E1 component forms a thiamin-derived radical by aerobic oxidation of the enamine intermediate. *J. Biol. Chem*. *289*, 29859-29873.
19. Al-Alaway, Adel Ibrahim Ahmad (2013) Subunit organisation and assembly of the 2-Oxoglutarate dehydrogenase multienzyme complex (OGDC). PhD thesis, University of Glasgow.
20. Chandrasekhar, K., Wang, J., Arjunan, P., Sax, M., Park, Y.H., Nemeria, N.S., Kumaran, S., Song, J., Jordan, F., and Furey, W. (2013) Insight to the interaction of the dihydrolipoamide acetyltransferase (E2) core with the peripheral components in the Escherichia coli pyruvate dehydrogenase complex via multifaceted structural approaches *J. Biol. Chem*. *288*, 15402-15417.
21. Wang, J., Kumaran, S., Zhou, J., Nemeria, N.S., Tao, H., Kakalis, L., Park, Y.H., Birkaya, B., Patel, M.S., and Jordan, F. (2015) Elucidation of the interaction loci of

- the human pyruvate dehydrogenase complex E2·E3BP core with pyruvate dehydrogenase kinase 1 and kinase 2 by H/D exchange mass spectrometry and nuclear magnetic resonance. *Biochemistry* *54*, 69-82.
22. Vijayakrishnan, S., S. M. Kelly, R. J. C. Gilbert, P. Callow, D. Bhella, T. Forsyth, J. G. Lindsay and Byron, O. (2010) Solution Structure and Characterisation of the Human Pyruvate Dehydrogenase Complex Core Assembly. *J. Mol. Biol.* *399*, 71-93.
  23. Mattevi, A., G. Obmolova, E. Schulze, K. Kalk, A. Westphal, A. de Kok and Hol, W. (1992) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science*. *255*, 1544-1550.
  24. Wallis, N. G., M. D. Allen, R. W. Broadhurst, I. A. Lessard and Perham, R. N. (1996) Recognition of a surface loop of the lipoyl domain underlies substrate channelling in the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* *263*, 463-474.
  25. Frank, R. A., J. V. Pratap, X. Y. Pei, R. N. Perham and Luisi, B. F. (2005) The molecular origins of specificity in the assembly of a multienzyme complex. *Structure*. *13*, 1119-1130.
  26. Angelides, K. J., Akiyama, S. K., and Hammes, G. G. (1979) Subunit stoichiometry and molecular weight of the pyruvate dehydrogenase multienzyme complex from *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* *76*, 3279–3283.
  27. Rey, M., Yang, M., Burns, K.M., Yu, Y., Lees-Miller, S.P., and Schriemer, D. C. (2013) Nepenthesin from monkey cups for hydrogen/deuterium exchange mass spectrometry. *Mol. Cell Proteomics*. *12*, 464-472.
  28. Mande, S.S., Sarfaty, S., Allen, M.D., Perham, R.N., and Hol, W.G. (1996) Protein-protein interactions in the pyruvate dehydrogenase multienzyme complex: dihydrolipoamide dehydrogenase complexed with the binding domain of dihydrolipoamide acetyltransferase. *Structure*. *4*, 277-286.
  29. Frank, R. A. W., Price, A. J., Northrop, F. D., Perham, R. N., and Luisi, B. F. (2007) Crystal Structure of the E1 Component of the *Escherichia coli* 2-Oxoglutarate Dehydrogenase Multienzyme Complex. *J. Mol. Biol.* *368*, 639-651.

## CHAPTER 3. E1o-E3, and E2o-E3 Interactions in Human OGDHc

### 3.1 Introduction

A previous study on the limited proteolysis of mammalian E1o with trypsin led to the dissociation of OGDHc into its individual E2, E3, and truncated E1 components, and rapid inactivation of the entire complex as a result of a single site cleavage near the E1o N-terminus [1-3]. The dissociation of OGDHc caused by the deletion of the N-terminal region of E1o underlines the importance of this region in maintaining the integrity of the OGDHc. Moreover, N-terminal sequence analysis of mammalian E1o indicated that sequences located at the beginning of the N-terminal region (the first 20 amino acids from Bovine heart E1o) display limited similarity to corresponding sequences in mammalian E2p and bovine heart E3BP, suggesting that the N-terminal region of E1o may be involved in interacting with E3 [2, 4].

The main goal of this chapter is to investigate the ability of the human E1o and E2o to associate with E3, and to map binding interaction loci in the binary complexes. Initially, gel filtration and ITC methods were used to characterize the interactions. Moreover, a gel based densitometric analysis was performed of *in vitro* reconstituted OGDHc. Following up on these experimental results, the E1o/E3, and E2o/E3 interactions were carefully examined with the HDX-MS method at peptide level resolution. In addition, <sup>1</sup>H-<sup>15</sup>N HSQC NMR was also used to investigate the E1o/E3 and E2o/E3 interactions.

### 3.2 Material and methods

Deuterium oxide (D<sub>2</sub>O) was from Cambridge Isotope Laboratories. ThDP was from Affymetrix. All other fine chemicals were from Sigma-Aldrich. Expression and purification of human E1o and E3 are described in Appendix A.

### 3.2.1 Size-exclusion chromatography

Size-exclusion chromatography experiments were performed using a Varian ProStar HPLC system with a UV detector. A Yarra 3u SEC-3000 column with a 20  $\mu$ l sample loop was used at a flow rate of 1 ml/min. The column was equilibrated with 50 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.5) containing 0.15 M NaCl, 0.5 mM ThDP, 1 mM MgCl<sub>2</sub>, and calibrated with the standards (*M<sub>r</sub>*): thyroglobulin (669,000), ferritin (440,000), catalase (232,000), aldolase (158,000), bovine serum albumin (67,000), and ovalbumin (43,000). The protein samples were eluted at a flow rate of 1 ml/min and monitored at 280 nm. All protein samples were incubated at various molar ratios with running buffer at room temperature for 1 h. Samples were centrifuged at 17,530  $\times$  g for 5 min to remove any precipitates before being applied to the column.

### 3.2.2 Isothermal titration calorimetry

ITC measurements were performed with a VP-ITC microcalorimeter (MicroCal, Northampton, MA). Titrations were carried out in a buffer contains 50 mM potassium phosphate (pH 7.5), 100 mM KCl, 0.5 mM ThDP, 2 mM MgCl<sub>2</sub> and 2  $\mu$ M FAD at 15 °C. Prior to the experiments, both protein samples were dialyzed exhaustively against this buffer. To measure E1o/E3 interaction, 300  $\mu$ M of E3 in the syringe was injected into the

cell containing 30  $\mu\text{M}$  of E1o (all based on subunit concentration). To measure the E2o/E3 interaction, 300  $\mu\text{M}$  of E3 in the syringe was injected into the cell containing 30  $\mu\text{M}$  of E2o. A single site binding model was used for curve fitting, and the dissociation constant ( $K_d$ ) and other thermodynamic parameters were calculated with ORIGIN 5.0 software provided by MicroCal.

### **3.2.3 Densitometric analysis of OGDHc components on SDS-PAGE gels**

SDS-PAGE Gels were digitized using a Molecular Imager® Gel Doc™ XR+ System and analyzed with Image Lab software (Bio-Rad Laboratories Inc.). Densitometric analyses were performed using lane-based background subtraction, followed by measurement of the area under the peaks. The optical densities were then used for statistical analysis. Each sample was analyzed on triplicated gels.

### **3.2.4 HDX-MS analysis**

Prior to H/D exchange, the E1o, E2o and E3 components were exchanged into a 10 mM potassium phosphate buffer (pH 7.5) with 100 mM NaCl, 0.5 mM ThDP, 1 mM  $\text{MgCl}_2$ , and 2  $\mu\text{M}$  FAD. All samples were prepared the same as described in Chapter 2, except the E1o/E3 mixture sample was prepared by mixing equal volume of 160  $\mu\text{M}$  E1o with 80  $\mu\text{M}$  E3 to make the final concentrations of E1o and E3 equal 80  $\mu\text{M}$  and 40 $\mu\text{M}$ , respectively. Other experimental handling and data analysis were conducted in the same way as described in Chapter 2.

### **3.2.5 Two-dimensional $^1\text{H}$ - $^{15}\text{N}$ HSQC NMR**

The NMR experiments were performed on a Varian INOVA 600 MHz spectrometer at 25 °C. For the HSQC NMR experiment, the 200  $\mu\text{M}$   $^{15}\text{N}$  labeled human E1o was exchanged to a buffer containing 50 mM  $\text{K}_2\text{HPO}_4$  (pH 7.5), 100 mM NaCl, 0.5 mM ThDP, and 1.0 mM  $\text{MgCl}_2$ , 2  $\mu\text{M}$  FAD, and 7 %  $\text{D}_2\text{O}$  was added before the NMR spectrum was recorded.

## **3.3 Results and Discussion**

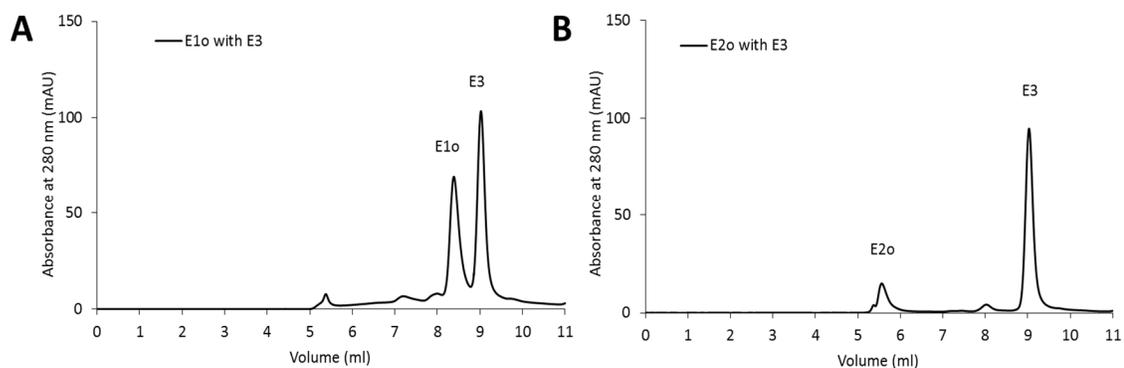
### **3.3.1 Investigation of the interaction of E1o/E3, and E2o/E3 by gel filtration and ITC methods**

The possibility of direct interaction between E1o with E3 and E2o with E3 was investigated using the same methods as for the study of E1o binding with E2o. The E1o, E2o and E3 components were individually expressed and purified according to methods described in Appendix A. E1o/E3 and E2o/E3 mixtures were prepared by mixing each enzyme in a 1:1 molar ratio and passing through a Yarra 3u SEC-3000 analytical column attached to Varian ProStar HPLC system with UV detector. Two separated peaks were observed from the E1o/E3 mixture, and there was no new peak identified. SDS-PAGE analysis of the fractions from 7 to 10 ml showed that the first peak eluting at 8.3 ml was the E1o homo-dimer, and the second peak eluting at 9.1 ml was the E3 homo-dimer (Figure 3.1). The SDS-PAGE analysis of fractions from 7 to 10 ml did not indicate a clear association between E1o and E3 (data not shown). Similarly to the E1o/E3 sample, on the

E2o/E3 sample, the E2o eluting at or near the void volume at 5.7 ml contained only E2o, presumably as a larger oligomer, perhaps a 24-mer.

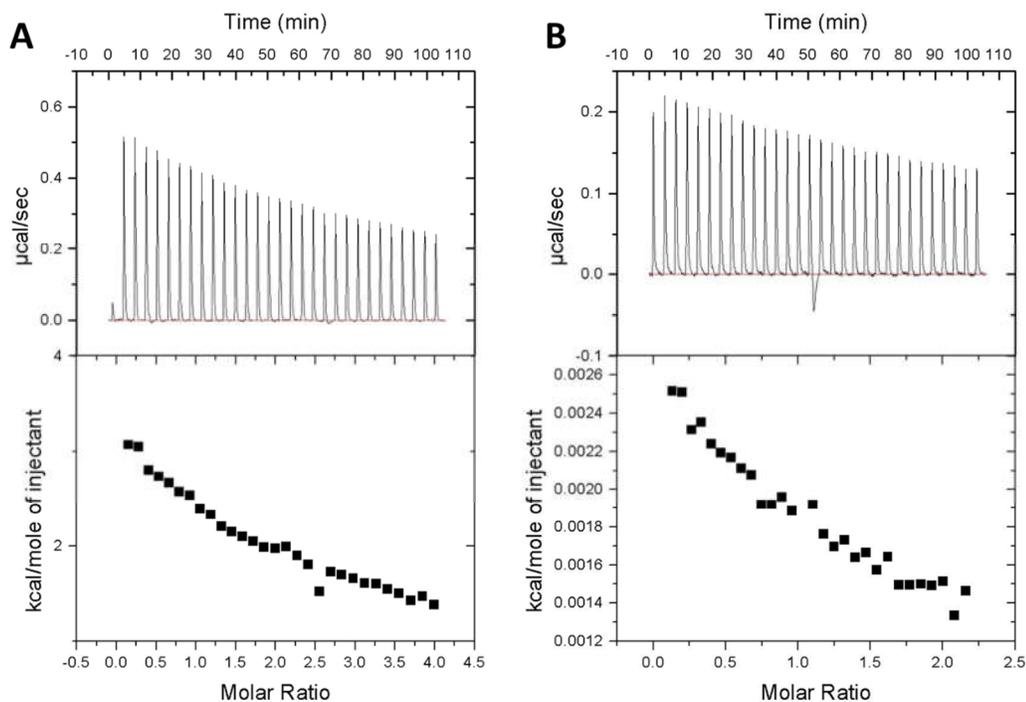
With the ITC experiments, the E3 (300  $\mu$ M) was injected in 10  $\mu$ l aliquots into the mixing cell containing 30  $\mu$ M E1o at 15 °C. A decreasing trend was observed on the endothermic signal, however, the association between E1o and E3 appeared to be very weak (Figure 3.2 A). A possible interaction between the E2o and E3 was also examined in the same manner. However, no significant binding was apparent (Figure 3.2 B).

Although the GPC and ITC experiments did not give a clear picture on the binding of E1o to E3 or of E2o to E3, some details on the interaction of the other two components with E3 became apparent. For example, while there might be an interaction between E2o and E3 according to the ITC data, the interaction between E1o and E3 appeared stronger than the interaction between E2o and E3 (Figure 3.2). The weak interaction signal obtained from the ITC experiments on mixing E2o and E3, could result from a transient binding between these two components. A transient binding occurs through a relatively large domain surface, and a very small surface section of another protein. A transient binding is usually unstable and happens in a short time duration [5-7]. Given the fact that, from the catalytic mechanism, the reduced lipoic acid from lipoyl domain on E2o has to be re-oxidized by visiting the FAD binding site on E3, the weak interaction signal detected with ITC can be the result from the time to time visit of lipoyl domain to the FAD binding site. On the other hand, the relatively stronger interaction of E1o with E3 observed in the ITC experiments is more likely associated with the organization and assembly of the OGDH complex.



**Figure 3.1 Gel filtration study of possible interaction between E1o and E3, and E2o and E3.**

A) A 1:1 molar ratio mixture of E1o and E3 was passed through a Yarra 3u SEC-3000 analytical column. Two peaks were observed, E1o eluted at 8.3 ml, while E3 eluted at 9.1 ml. Fractions from 7 to 10 ml were collected for SDS-PAGE analysis. B) A 1:1 molar ratio mixture of E2o and E3 was passed through a Yarra 3u SEC-3000 analytical column. Two peaks were observed, E2o eluted at the void volume 5.7 ml, while E3 eluted at 9.1 ml. Fractions from 5 to 6 ml were collected for SDS-PAGE analysis.



**Figure 3.2 ITC of E1o with E3, and E2o with E3.**

A) Raw data obtained from ITC measurements of E3 interaction with E1o. The top panel is the thermogram that results from titrating the E3 dimer into the cell containing E1o at 15 °C. The bottom panel is the heat change data result as calculated by ORIGIN 5.0; B) The same ITC experimental conditions were used for measuring E3 interaction with E2o.

### 3.3.2 Densitometry study of OGDHc - mixing of all three components.

The individually purified human E2o was incubated with molar excess of E1o and E3 (E1o, E2o, and E3 were incubated in a 5:1:5 subunit molar ratio). After 30 min incubation at room temperature, the mixture was passed through a Yarra 3 µm SEC-3000 column to remove the unbound E1o and E3. Three separated peaks were observed (Figure 3.3 A), and the corresponding fractions were subjected to SDS-PAGE and the gels were

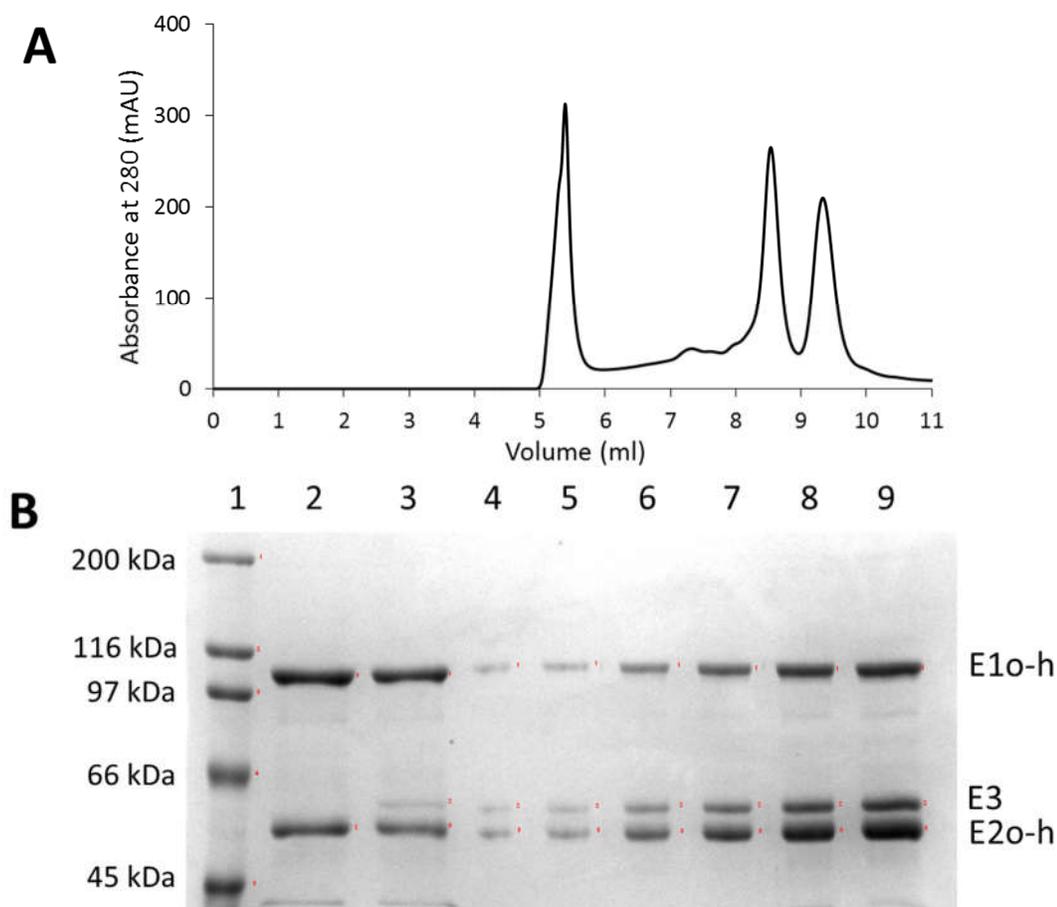
stained with Coomassie Brilliant Blue. It was found that the first peak that eluted at or close to the void volume contained E1<sub>o</sub>, E2<sub>o</sub> and E3, the second peak eluted at 8.6 min contained only E1<sub>o</sub>, the third peak eluted at 9.3 min contained only E3. This result indicated that the relatively stable E1<sub>o</sub>/E2<sub>o</sub> sub-complex can provide a better association to E3.

Absolute and relative quantification were determined by comparing the staining density between the bands or by comparing with the standard curves obtained with individual protein standards. Results from the two quantification methods are in good agreement. The subunit component ratio of E1<sub>o</sub>:E2<sub>o</sub>:E3 is 4:6:1. The detailed data and a SDS-PAGE gel sample are listed in Table 3.1.

It was reported that the overall maximal stoichiometry of *in vitro* reconstituted human PDHc for E1<sub>p</sub>:E2<sub>p</sub>:E3BP:E3 is 40:40:20:20 [8]. While based on the purified native mammalian PDHc, the E2<sub>p</sub>/E3BP core binds 20-30 copies of E1<sub>p</sub> and 6-12 copies of E3 [9]. The BCOADHc is organized around 24 E2<sub>b</sub> subunits, 12 E1<sub>b</sub> α<sub>2</sub>β<sub>2</sub> tetramers, and 6 E3 homodimers [10]. Here, using an *in vitro* reconstituted OGDHc, we obtained densitometry evidence that the E1<sub>o</sub>:E2<sub>o</sub>:E3 associated in a 4:6:1 subunit ratio. Comparing to other 2-oxo acid dehydrogenase complexes, as a rate limiting multienzyme complex in the citric acid cycle, OGDHc seems to contain fewer E3 components than other members of the family of mammalian 2-oxo acid dehydrogenase complexes [8-10]. Notably, our gel filtration results suggested that the interaction between the E1<sub>o</sub> and E3 can be strengthened by incorporation of E2<sub>o</sub>.

**Table 3.1 Absolute and relative quantification of E1o, E2o, and E3 subunits in OGDHc.**

<b>Absolute Quantification</b>					
	#1	#2	#3	Average molar ratio	
E1o subunit moles	9.8	10.6	10.2	10.2	
E2o subunit moles	11.87	14.6	13.8	13.4	
E3 subunit moles	2.8	2.6	2.8	2.7	
<b>Relative Quantification</b>					
	#1	#2	#3	Average protein amount	molar ratio
E1o subunit protein amount	1	1	1	1	10
E2o subunit protein amount	0.59	0.59	0.72	0.633	16.4
E3 subunit protein amount	0.12	0.14	0.12	0.127	2.7



**Figure 3.3 Gel filtration of human OGDHc reconstituted from E1o, E2o and E3**

A) E1o, E2o and E3 were pre-incubated for 30 min at room temperature followed by separation on a Yarra 3  $\mu$ m SEC-3000 column. E1o, E2o, and E3 were mixed in a 5:1:5 subunits ratio. B) SDS-PAGE (10%) analysis of eluted protein peak stained with Coomassie Brilliant Blue. Lane 1, molecular mass markers; lane 2, the E1o/E2o sub-complex previously purified with the same gel filtration condition; lane 3, the first eluted peak; lanes 4-9, the pre-mixed known amount of individual E1o, E2o, and E3 proteins for relative quantification. The presence of E1o, E2o, and E3 at lane 3 affirms formation of OGDHc.

### 3.3.3 HDX-MS experiment

To examine the possible direct binding between E1o and E3, and E2o and E3, a time-dependent HDX-MS study was carried out at time points from 0 to 270 min.

#### **Overview of the HDX behavior of E1o, E2o, and E3 HDX.**

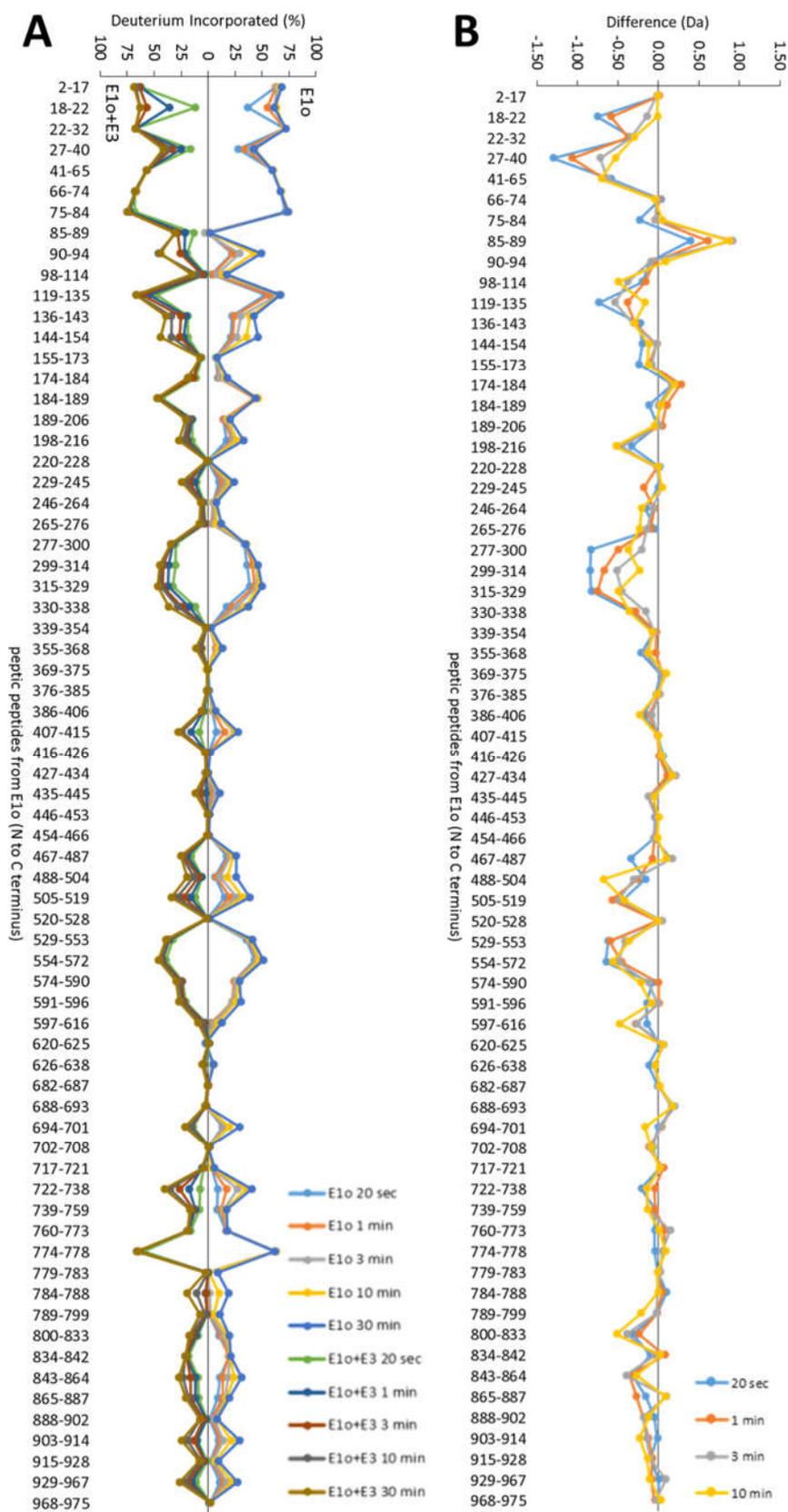
As a first step to understand the interaction of E3 with other components of OGDHc, HDX-MS experiments were conducted with full-length proteins. The time dependent HDX-MS experiments were carried out at 20 s, 1, 3, 10, 30, 90 and 270 min. The E1o and E2o components displayed the same overall digestion behavior as reported in Chapter 2. On-line digestion by pepsin followed by LC-MS analysis under our HDX conditions resulted in 25 peptides for E3, many of them partially overlapping, providing 90.9 % sequence coverage for E3. The high sequence coverage on E3 provided access to study the structural changes throughout the protein during its assembly. Compared to the E1o and E2o components, E3 deuterium uptake is relatively low. (< 60% shown in Figure 3.5). The low deuterium uptake in E3 might be related to a tightly organized protein structure. A summary of peptic peptide data of E3 is provided in the deuterium uptake curves in Figure 3.5 and Table 3.2. The following discussion will focus on the subset of peptides, which displayed deuterium uptake difference upon binding.

#### **3.3.4 HDX-MS analysis of the binding between E1o and E3**

Upon binding of E1o to E3, many regions on both E1o and E3 were observed to undergo change in deuterium uptake. On the E1o, the significant changes are located in the N-terminal region and ThDP-binding fold (Figure 3.4). In particular, residues 18-22, 27-

40, and 119-135 from the N-terminal region of E1o, and residues 277-329, 488-519, and 529-572 from the ThDP-binding fold of E1o, displayed a decrease in deuterium uptake. However, all the significant decrease in deuterium uptake was only observed at or before the 3 min exchange time points. These weak protections suggested a loose association between E1o and E3, which also explained why no strong association of E1o with E3 could be observed from the GFC and ITC experiments.

Our HDX-MS results showed that E1o has a wide range of regions which displayed deuterium uptake difference upon binding to E3. Notably, two peptides (18-22 and 27-40) from the N-terminal region of E1o displayed the most significant decrease in deuterium uptake while associated with E3. These two peptides were also found to be significantly protected in the E1o/E2o sub-complex. In addition, the adjacent peptide 85-89 located in the N-terminal region of E1o, which belongs to the third predicted alpha helix showed an increase in deuterium uptake (Figure 2.6). These findings suggested that once E1o binds to E3, the three predicted alpha helices in the E1o N-terminal region will experience a conformational change which is different from the conformational change while binding to E2o. As suggested by the stability of these two sub-complexes, the N-terminal region of the E1o/E2o sub-complex is likely to have a more stable overall structure than the E1o/E3 sub-complex.

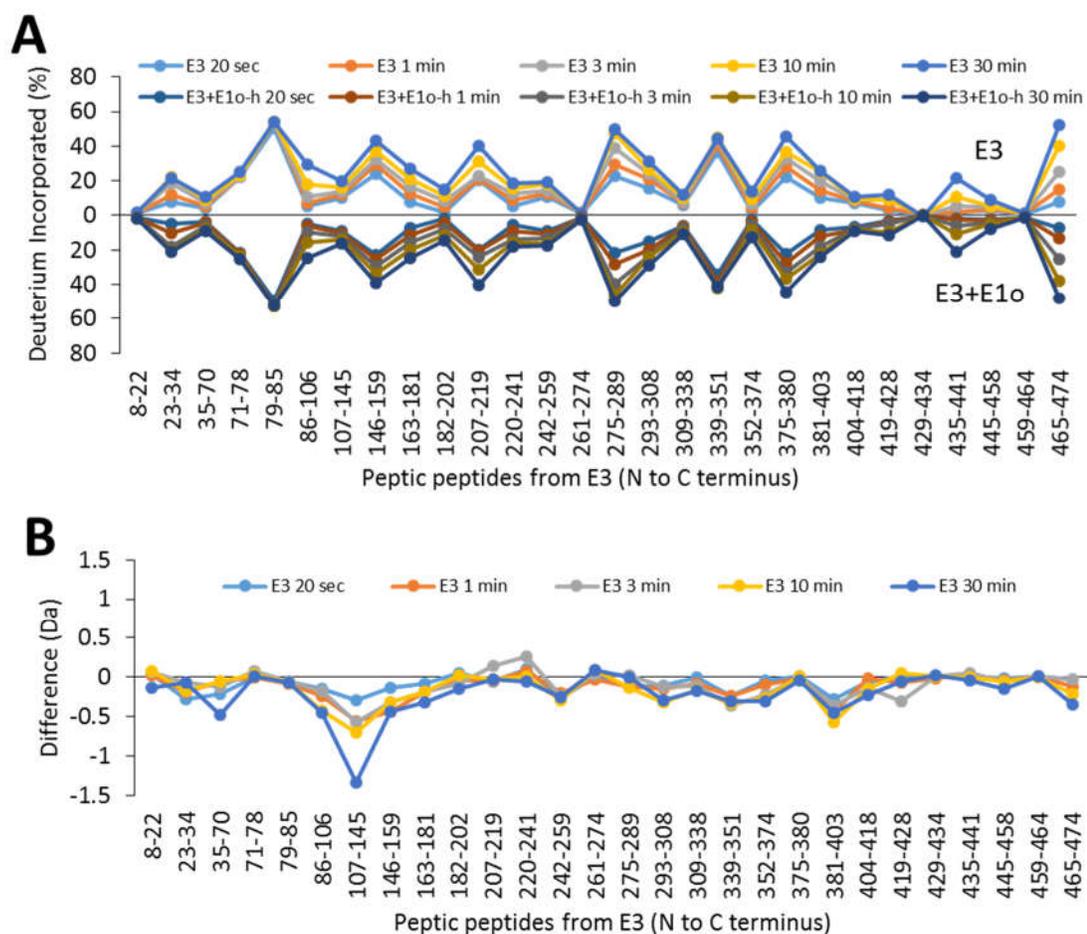


**Figure 3.4 Comparative HDX-MS analysis of the interaction of E1o and E3**

A) A butterfly plot representing average relative deuterium incorporation percentage (y axis) (deuterons exchanged/maximum exchangeable amides  $\times$  100%) of peptic fragments from E1o (x axis, listed from N to C terminus) in the absence of E3 (top) versus in the presence of E3 (bottom) based on three independent experiments. B) Difference plot showing deuterium incorporation changes of peptic fragments of E1o in the absence and presence of E3 (deuterons exchanged in the absence of E3 minus deuterons exchanged in the presence of E3).

On the E3 partner, upon binding to E1o, two peptides 35-70 and 107-145 in the FAD-binding domain of E3, and one peptide 381-403 from the interface domain of E3 showed the most decrease in deuterium uptake (Figure 3.5). Previous HDX-MS data on human E3 [11] showed that the same peptides from E3 also had the most significant decrease when binding to the E2·E3BP sub-complex. This result suggested that there is a similar binding mode in the E1o/E3 association and the E2·E3BP/E3 association.

This, together with the fact that E1o N-terminal region showed limited similarity to the corresponding sequence in E3BP suggested that the N-terminal region of E1o, especially the two peptides 18-22 and 27-40, is the new binding region responsible for the assembly of OGDHc.



**Figure 3.5 Comparative HDX-MS analysis of the interaction of E1o and E3**

A) A butterfly plot representing average relative deuterium incorporation percentage (y axis) (deuterons exchanged/maximum exchangeable amides  $\times$  100%) of peptic fragments from E3 (x axis, listed from N to C terminus) in the absence of E1o (top) versus in the presence of E1o (bottom) based on three independent experiments. B) Difference plot showing deuterium incorporation changes of peptic fragments of E3 in the absence and presence of E1o (deuterons exchanged in the absence of E1o minus deuterons exchanged in the presence of E1o).

**Table 3.2 Peptic peptides examined from human E3.**

Peptides			Monoisotopic [M+H] <sup>+</sup> (Da)	Mass	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	)
1	8-22	DVTVIGSGPGGYVAA	1362.6935	1362.6900	2.6
2	23-34	IKAAQLGFKTVC	1278.7239	1278.7239	0.0
3	35-70	EKNETLGGTCLNVGCIPSKALLNNSHYHMAHGKD	3812.7952	3812.7962	0.3
4	71-78	FASRGIEM	910.4453	910.4451	0.2
5	79-85	SEVRL	603.3464	603.3461	0.5
6	86-106	DKMMEQKSTAVKALTGGIAHL	2229.1723	2229.1730	-0.3
7	107-145	FKQNKVVHVNGYGKITGKNQVTATKADGGTQVIDTK NIL	4185.2797	4185.2779	0.4
8	146-159	IATGSEVTPFPGIT	1389.7285	1389.7260	1.8
9	163-181	DTIVSSTGALSLLKKVPEKM	2004.1040	2004.1046	-0.3
10	182-202	VVIGAGVIGVELGSVWQRLGA	2080.1903	2080.1913	-0.5
11	207-219	VEFLGHVGGVGID	1298.6724	1298.6739	-1.2
12	220-241	MEISKNFQRILKQGFKFLNT	2698.4906	2698.4861	1.7
13	242-259	KVTGATKKSDBGKIDVSIE	1876.0379	1876.0386	-0.4
14	261-274	ASGGKAEVITCDVL	1362.6935	1362.6933	0.1
15	275-289	LVCIGRRPFTKNLGL	1686.9881	1686.9836	2.7
16	293-308	GIELDPRGRIPVNTRF	1840.0173	1840.0188	-0.8
17	309-338	QTKIPNIYAIGDVVAGPMLAHKAEDEGIIC	3166.6249	3166.6275	-0.8
18	339-351	VEGMAGGAVHIDY	1318.6095	1318.6096	-0.1
19	352-374	NCVPSVIYTHPEVAWVGKSEEQQL	2585.2676	2585.2705	-1.1
20	381-403	YKVGKFPFAANSRAKTNADTDGM	2489.2241	2489.2242	0.0
21	404-418	VKILGQKSTDRVLGA	1584.9432	1584.9432	0.0

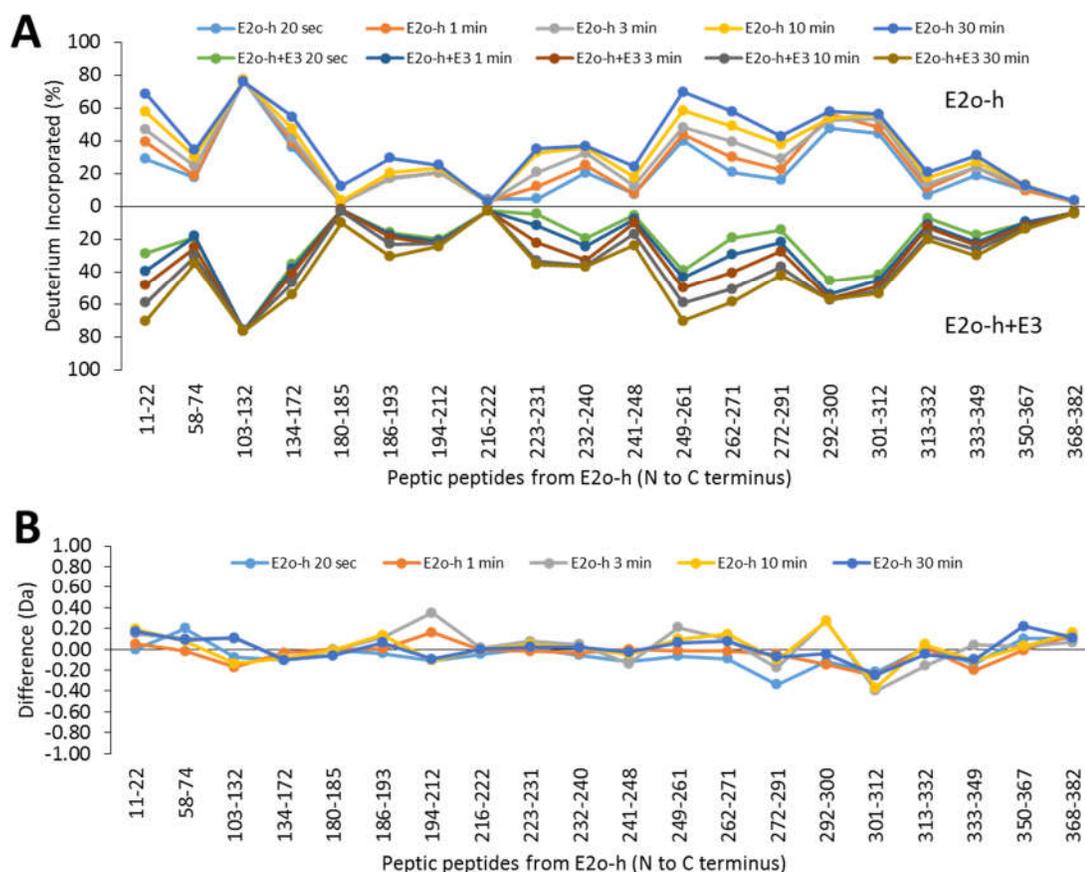
22	419-428	HILGPGAGEM	981.4821	981.4822	-0.1
23	429-434	VNEAAL	616.3298	616.3301	-0.5
24	435-441	ALEYGAS	710.3353	710.3355	-0.3
25	445-458	IARVCHAHPTLSEA	1504.7683	1504.7689	-0.4
26	459-464	FREANL	749.3931	749.3941	-1.3
25	465-474	AASFGKSINF	1041.5357	1041.5364	-0.6

---

### 3.3.5 E2o/E3 interaction from OGDHc.

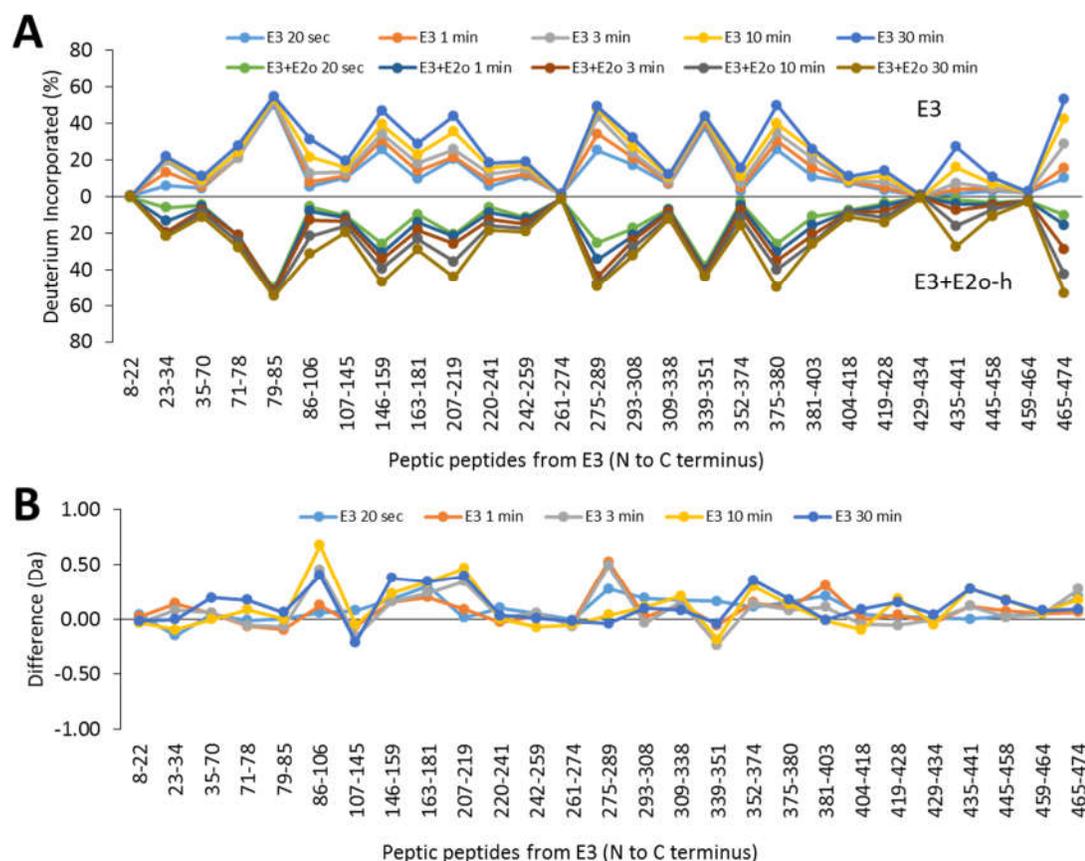
To check the possible binding between E2o and E3, the HDX-MS was used to investigate interactions of the two components (Figure 3.6 and 3.7). There was no clear evidence showing direct binding between these two proteins. On the E2o, no significant deuterium uptake change was observed while incubated with E3. On the E3 partner, some modest (statistically insignificant) increase of deuteration was observed on the FAD-binding domain and NAD<sup>+</sup> binding domain of E3. Given the fact that during the catalytic cycle of OGDHc, the reduced lipoic acid covalently attached to E2o is being re-oxidized by FAD on E3, the increase of deuteration in the FAD and NAD<sup>+</sup> binding regions could result from the transient visit of the lipoyl domain on E2o to the FAD binding domain on E3.

The two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC NMR experiments on the <sup>15</sup>N labeled E2o N-domain incubated with E3 further suggested that the lipoyl domain from E2o can visit E3, from time to time to induce the small changes on E3 (Figure 3.8).



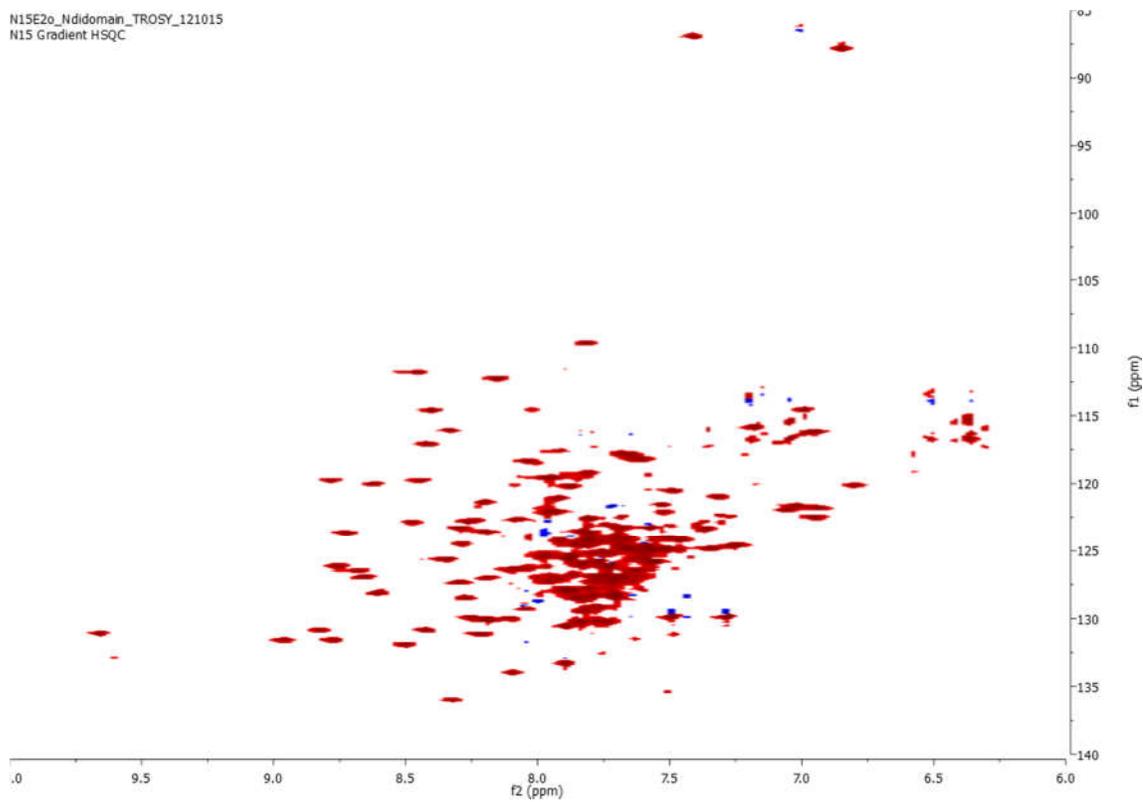
**Figure 3.6 Comparative HDX-MS analysis of the interaction of E2o with E3**

A) A butterfly plot representing average relative deuterium incorporation percentage (y axis) (deuterons exchanged/maximum exchangeable amides  $\times 100\%$ ) of peptic fragments from E2o (x axis, listed from N to C terminus) in the absence of E3 (top) versus in the presence of E3 (bottom) based on three independent experiments. B) Difference plot showing deuterium incorporation changes of peptic fragments of E2o in the absence and presence of E3 (deuterons exchanged in the absence of E3 minus deuterons exchanged in the presence of E3). Peptide 134-172 was generated from the experiments on the N-domain E2o.



**Figure 3.7 Comparative HDX-MS analysis of the interaction of E3 with E2o**

A) A butterfly plot representing average relative deuterium incorporation percentage (y axis) (deuterons exchanged/maximum exchangeable amides  $\times$  100%) of peptic fragments from E3 (x axis, listed from N to C terminus) in the absence of E2o (top) versus in the presence of E2o (bottom) based on three independent experiments. B) Difference plot showing deuterium incorporation changes of peptic fragments of E3 in the absence and presence of E2o (deuterons exchanged in the absence of E2o minus deuterons exchanged in the presence of E3).



**Figure 3.8  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra of E2o N-didomain**

Superimposed spectra of uniformly  $^{15}\text{N}$ -labeled E2o N-didomain in red, and uniformly  $^{15}\text{N}$ -labeled E2o N-didomain with E3 in blue.

### 3.4 Conclusion

The data presented in this chapter demonstrated that human E1 $\alpha$  and E2 $\alpha$  can both interact with E3. ITC experimental data showed that the interaction between E1 $\alpha$  and E3 is much stronger than the interaction between E2 $\alpha$  and E3. Although a E1 $\alpha$ /E3 sub-complex was not clearly observed in gel filtration experiments, formation of the entire complex was detected from analytical SEC column when E2 $\alpha$  was incubated with excess amount of E1 $\alpha$  and E3. This observation indicated that the E1 $\alpha$ /E2 $\alpha$  sub-complex can provide a better binding to E3 than could E1 $\alpha$  alone. In other words, E2 $\alpha$  can promote the binding between E1 $\alpha$  and E3. Moreover, a densitometric method was used to study the composition of OGDHc. With absolute and relative quantification methods, the overall ratio of an *in vitro* reconstituted OGDHc for E1 $\alpha$ :E2 $\alpha$ :E3 is 4:6:1 (subunit ratio). Compared to other family members of mammalian 2-oxo acid dehydrogenase complexes, OGDHc seems to contain fewer E3 components. One possible explanation is that in the mammalian cell, E3 was shared by three closely related multienzyme complexes (PDHc, OGDHc, and BCOADHc). As a rate limiting multienzyme complex in the citric acid cycle, E3 might bind loosely in OGDHc, the dissociation of E3 from OGDHc could happen while it passes through the SEC column.

To further explore the binding loci between E1 $\alpha$  and E3, the HDX-MS method was used. Upon binding to E3, difference in deuterium uptake behavior was observed across wide regions on E1 $\alpha$ . Interestingly, supporting the relatively weak binding between E1 $\alpha$  and E3, all significant deuterium exchange protections were only observed during less than 3 min exposure to D<sub>2</sub>O on E1 $\alpha$ . Notably, the same two peptides (18-22 and 27-40) from the N-terminal region of E1 $\alpha$  previously found critical in formation of the E1 $\alpha$ /E2 $\alpha$  sub-

complex also displayed the most significant deuterium exchange protection while associated with E3. Although the overall N-terminal region of E1o was protected from deuterium exchange when associated with E2o or E3, the different deuterium uptake behaviors suggested that the N-terminal region of E1o can respond differently to different binding partners. Our data suggest that E2o and E3 might be able to induce different conformational changes in this unstructured region in E1o.

Given the fact that the N-terminal region of E1o displays a limited similarity to corresponding sequences in E3BP, the residues 18-40 in E1o might be the unique peripheral-subunit binding domain (PSBD) in OGDHc. Similarly to the PSBD in bacterial E2p and E2b enzyme, in OGDHc assembly, E2o and E3 bind in a mutually exclusive fashion via this new PSBD in the N-terminal region of the E1o subunits.

While E3 binds to E1o, E3 displayed a E2E3BP-binding like HDX behavior. A previous study showed that both E3 and the E3 binding domain on E3BP appear to be rigid structural units that associate as a consequence of their surface electrostatic and hydrogen-bonding compatibility [8, 12]. Similar surface electrostatic and hydrogen-bonding might be the major force involved in the association of E1o with E3. In addition, since the association of E1o and E3 can induce a wide range of HDX difference on both proteins, a space with a wide surface area might be required for the association of E1o to E3.

Lastly, there was no clear evidence indicating direct binding between E2o and E3. Although HDX-MS showed some changes of low statistical significance on E2o, it was likely caused by the visit of lipoyl domain on E2o to E3 due to enzyme catalytic reasons.

As stated above, our data confirm a role for the N-terminal region of E1o in E3 binding. These findings are in good agreement with previous biochemical and

immunological studies on mammalian OGDHc [3]. The residues 18-40 on the N-terminal region of E1 $\alpha$  identified by HDX-MS can serve as a unique PSBD for the assembly of OGDHc.

## References

1. Kresze, G. B., Ronft, H., Dietl, B., and Steber, L. (1981) Limited proteolysis of 2-oxoglutarate dehydrogenase multienzyme complex from bovine kidney. *FEBS Lett.* *127*, 157-160.
2. Rice, J. E., Dunbar, B., and Lindsay, J. G. (1992) Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. *The EMBO journal.* *11*, 3229-35.
3. McCartney, R. G., Rice, J. E., Sanderson, S. J., Bunik, V., Lindsay H., and Lindsay, J. G. (1998) Subunit interactions in the mammalian alpha-ketoglutarate dehydrogenase complex. Evidence for direct association of the alpha-ketoglutarate dehydrogenase and dihydrolipoamide dehydrogenase components. *J. Biol. Chem.* *273*, 24158-24164.
4. Rice, J. E., and Lindsay, J. G. (1991) Evidence for a protein X-like domain at the N-terminus of the E1 component of the mammalian 2-oxoglutarate dehydrogenase complex. *Biochem. Soc. Trans.* *19*, 403S.
5. Iwahara, J., and Clore, G.M. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* *440*, 1227-1230.
6. Rothe, M., Gruber, T., Gröger, S., Balbach, J., Saalwächter, K., and Roos, M. (2016) Transient binding accounts for apparent violation of the generalized Stokes-Einstein relation in crowded protein solutions. *Phys. Chem. Chem. Phys.* *18*, 18006-18014.
7. Tonddast-Navaei, S., and Stan, G. (2013) Mechanism of transient binding and release of substrate protein during the allosteric cycle of the p97 nanomachine. *J. Am. Chem. Soc.* *135*, 14627-14636.
8. Brautigam, C. A., Wynn, R. M., Chuang, J. L., and Chuang, D. T. (2009) Subunit and catalytic component stoichiometries of an in vitro reconstituted human pyruvate dehydrogenase complex. *J. Biol. Chem.* *284*, 13086-13098.
9. Perham, R. N. (2000) Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* *69*, 961-1004.
10. Machius, M., Wynn, R.M., Chuang, J.L., Li, J., Kluger, R., Yu, D., Tomchick, D.R., Brautigam, C.A., and Chuang, D.T. (2006) A versatile conformational switch regulates reactivity in human branched-chain alpha-ketoacid dehydrogenase. *Structure* *14*, 287-298.

11. Wang, J. (2014) Interaction of the components in human and escherichia coli pyruvate dehydrogenase multienzyme complexes (PhD dissertation). Rutgers University-Newark. Retrieved from <http://dx.doi.org/doi:10.7282/T3FX77J5>
12. Wang, J., Kumaran, S., Zhou, J., Nemeria, N.S., Tao, H., Kakalis, L., Park, Y.H., Birkaya, B., Patel, M.S., and Jordan, F. (2015) Elucidation of the interaction loci of the human pyruvate dehydrogenase complex E2·E3BP core with pyruvate dehydrogenase kinase 1 and kinase 2 by H/D exchange mass spectrometry and nuclear magnetic resonance. *Biochemistry* 54, 69-82.

## **CHAPTER 4. The OGDHc assembly studied with chemical crosslinking coupled with mass spectrometry.**

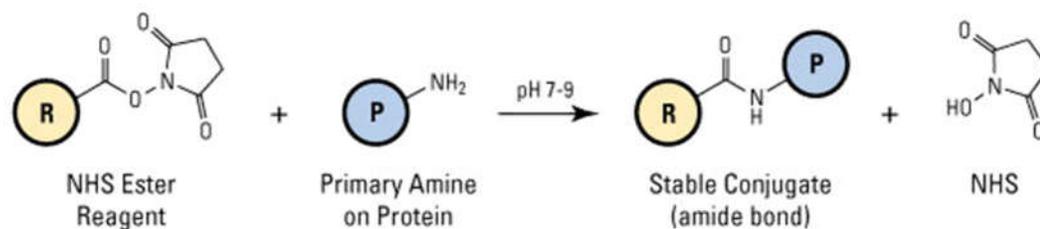
### **4.1 Introduction**

The association of proteins plays an important role in protein function and regulation. Structural analysis of protein complexes can provide a way to look at the interactions between two or more proteins in a complicated biological system. With the continuing interest to study larger and more complex systems, instead of using traditional protein structure studies, methods such as X-ray crystallography and NMR, a mass spectrometry based method can be used to determine the protein structure in a large complex. The application of chemical cross-linking coupled with mass spectrometry (CX-MS) provided a perfect tool to achieve these goals [1-3]. Compared to NMR and X-ray crystallography methods, CX-MS requires significantly less protein material, and the method is applicable to study protein-protein interactions at, or near physiological conditions.

Protein cross-linking is commonly undertaken with homo-bifunctional NHS-esters. This kind of cross-linker can bridge two lysine residues that are close to each other. The range of this cross-linking reaction can be controlled by varying the length of the space between the two NHS-ester groups (Figure 4.1).

In this chapter, the most popular cross-linking reagent BS3 (Figure 4.2) was used to study the interaction and assembly of OGDHc. To successfully identify the cross-linked

peptides, various approaches have been attempted. The CX-MS data presented here well agreed with HDX-MS results from previous chapters.



**Figure 4.1 The NHS ester cross-linking reaction.**

## 4.2 Materials and methods.

The 1:1 ratio premixed stable isotope labelled cross-linker BisSulfoSuccinimidylSuberate (BS3-H<sub>12</sub>/D<sub>12</sub>) and biotinylated cross-linker CyanurBiotinDimercaptoPropionylSulfoSuccinimide (CBDPSS-H<sub>8</sub>/D<sub>8</sub>) were from Creative Molecules Inc. Sequencing grade modified trypsin was from Promega. Monomeric avidin agarose was from ThermoFisher Scientific. All other fine chemicals were from Sigma-Aldrich.

### 4.2.1 Chemical cross-linking mass spectrometry

The purified E1o, E2o, and E3 were mixed in 2:1:1 ratio and incubated in 50 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.5) with 200 mM NaCl, 1 mM MgCl<sub>2</sub>, and 0.5 mM ThDP for one hour at 25 °C. The incubation mixture was then clarified from any precipitate that might have

formed during incubation, and adjusted to a final protein concentration close to 1.5 mg/mL for cross-linking.

The cross-linking reaction was initiated by adding BS3-H<sub>12</sub>/D<sub>12</sub> into 300 µl of E1o/E2o complex solution to a final concentration of 0.3 mM. The final cross-linker reagent concentration was optimized based on small scale pilot experiments with different levels of cross-linker excess, followed by SDS-PAGE to get most of the individual E1o, E2o, and E3 converted into high molecular weight bands. The cross-linking sample was incubated at 25 °C for 30 min, and quenched by adding ammonium bicarbonate to a final concentration of 50 mM and incubated at room temperature for an additional 15 min. The cross-linked proteins were then dried in vacuum, and denatured with 8 M urea for 1 h at 60 °C. The urea solution was then diluted to a concentration of 1 M with 50 mM ammonium bicarbonate, and digested overnight with trypsin at a 1:50 ratio at 37 °C. The resulting peptides were purified using 50 mg SepPak cartridges (Waters) according to the manufacturer's manual. The eluted peptide mixtures were vacuum dried and re-suspended in 50 µl 0.1 % formic acid, 5 % acetonitrile for LC-MS analysis, or further fractionated with a Superdex peptide 10/300 GL column, and submitted to LC-MS analysis.

#### **4.2.2 Fractionation of cross-linked peptides by size exclusion chromatography**

The peptide mixture resulting from solid-phase extraction (SPE) was vacuum dried and re-suspended in 200 µl SEC mobile phase (water/acetonitrile/formic acid, 70:30:0.1). The re-suspended peptides were then separated with a GE Healthcare Superdex peptide 10/300 GL column, and monitored by UV absorption at 214 nm. Peptides in the mass range from

1046 to 6511 Dalton were collected in 12 fractions (the column was pre-calibrated with aprotinin mass 6511 Da, and angiotensin II mass 1046 Da). All fractions were vacuum dried and re-suspended in 40  $\mu$ l 0.1 % formic acid, 5 % acetonitrile for LC-MS analysis.

#### **4.2.3 Avidin agarose beads enrichment.**

After the SepPak cartridge desalting and enriching step, the protein sample cross-linked with biotinylated cross-linker was further enriched with monomeric avidin agarose according to the manufacturer's manual. The final eluted peptide mixtures were vacuum dried and re-suspended for LC-MS analysis.

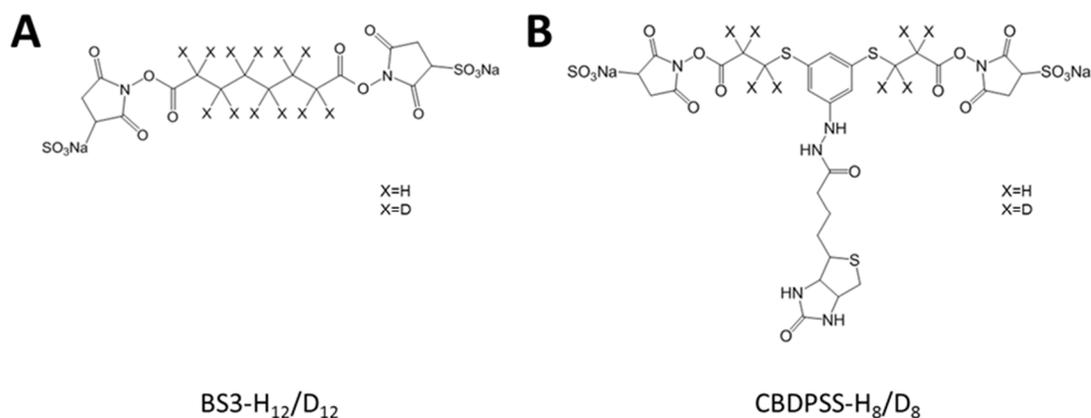
#### **4.2.4 Liquid Chromatography-Mass Spectrometry.**

LC-MS analysis was carried out on a Agilent 1200 series HPLC system connected to a Fourier transform-mass spectrometer (Apex-ultra 70 hybrid Fourier transform-mass spectrometer, Bruker Daltonics). The recovered peptide was directly injected into a Kinetex 30 cm  $\times$  2.1 mm I.D. C18 column. Peptides were separated at a flow rate of 0.1 ml/min using a linear gradient going from 5 % acetonitrile with 0.1 % formic acid to 50 % acetonitrile with 0.1 % formic acid in 70 min.

#### **4.2.5 Data processing**

The raw MS files were processed into mass lists using Bruker Daltonics DataAnalysis 4.0. The exported peak list files were analyzed using the ICC-CLASS

software. The mass lists of each fraction were searched for H<sub>12</sub>/D<sub>12</sub> doublets using the DX ESI LCMS program of ICC\_CLASS. The obtained H<sub>12</sub>/D<sub>12</sub> doublets list was then searched with DXMS program, and the search parameters were as follows: Maximum number of miss cleavages is 2, mass tolerance is  $\pm 1.5$  ppm. All identified cross-linking peptides were then manually validated.



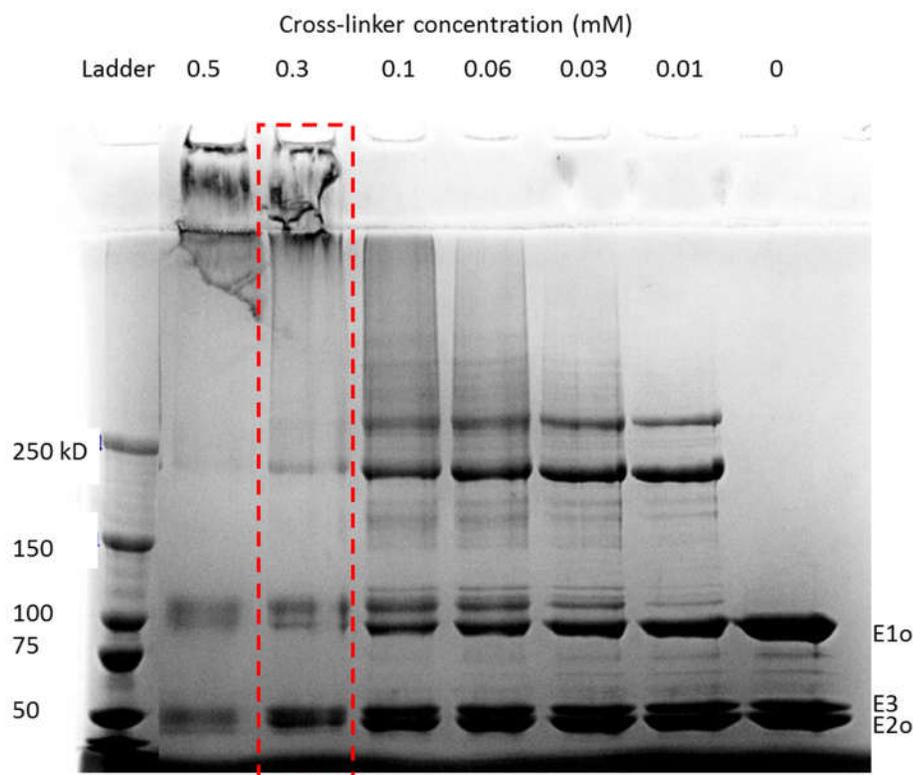
**Figure 4.2 Two cross-linkers used in this study.**

## 4.3 Result and Discussion

### 4.3.1 Strategies for identifying cross-linked peptides

The cross-linking reaction was performed with a stable isotope labeled cross-linker bis(sulfosuccinimidyl) suberate (BS3 H<sub>12</sub>/D<sub>12</sub>). As described by Seebacher et al. [4], using a mixture of the light form (H<sub>12</sub>) and the heavy form (D<sub>12</sub>) of BS3 can facilitate the detection of low abundant cross-linked products. Initially, the reaction conditions were optimized to use the lowest useful cross-linker concentration to avoid perturbation on protein structure and non-specific cross-linking, and SDS-gel was used to monitor the

cross-linking process. Different levels of cross-linker excess were tried to ensure an optimal cross-linking conditions. As expected, with the increase of cross-linker concentration, the individual protein bands were getting broader and the higher molecular weight bands, stand for cross-linked proteins, appeared on the top of the gel (Figure 4.3). Generally, at a very high cross-linker concentration, the protein monomer bands decrease notably, and only the higher mass protein bands appear on the top of the gel. Under that cross-linking conditions, the non-specific cross-linking could contribute to the overall cross-linking reaction, so this degree of cross-linking should be avoided [5, 6]. Our SDS-PAGE result suggested an optimum BS3 cross-linker concentration of 0.3 mM as highlighted in the red frame in Figure 4.3.



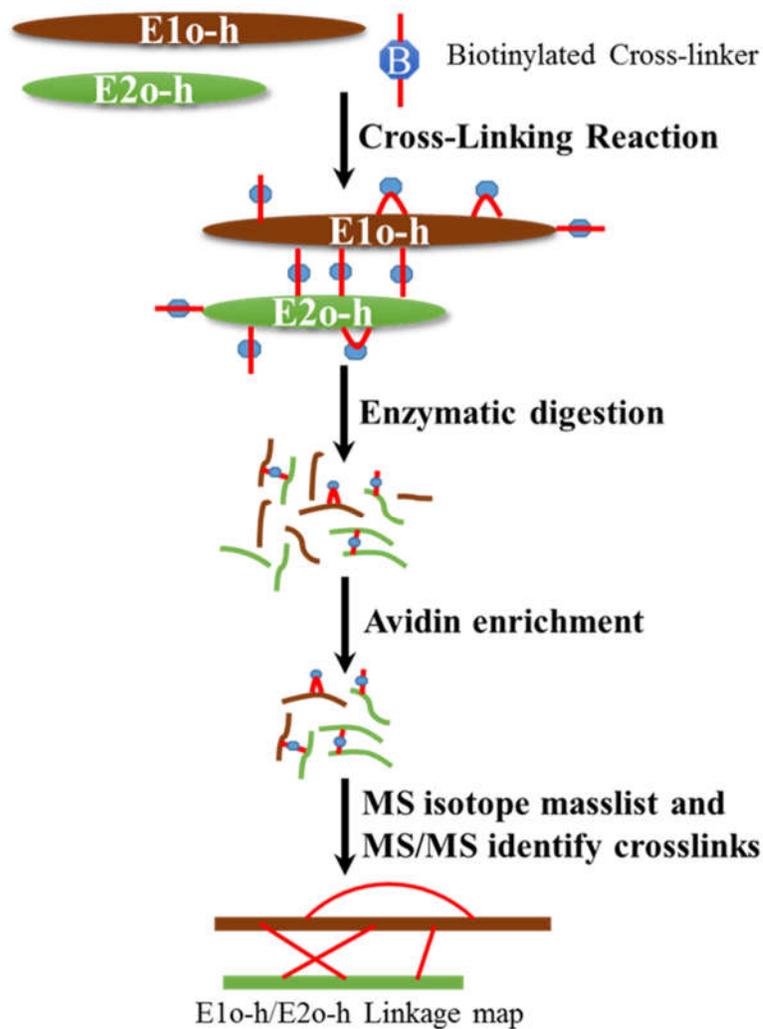
**Figure 4.3. Monitoring the OGDHc cross-linking reaction by SDS-PAGE.**

At very low cross-linker concentrations, most of the E1o, E2o, and E3 monomers are still un-cross-linked. Increasing the cross-linker concentration, more cross-linked bands begin to appear. At 0.5 mM BS3 concentration, most of protein monomers are gone, and very high mass products become dominant.

As mentioned in Chapter 1, a low yield of specific cross-linked peptides was one of the common problems in CX-MS, which often results in the loss of cross-linking information. Moreover, the cross-linked peptides are often difficult to detect because the signal intensity for these peptides is generally far lower than that for non-modified peptides [7], and the co-eluted non-modified peptides can mask the cross-linked peptides by ion suppression mechanisms in the use of LC-ESI-MS [8].

To overcome these practical challenges, different approaches have been developed to enrich cross-linked peptides from the digested peptide mixtures. The most popular method is a SDS-gel based enrichment [9-11]. Cross-linked and un-treated protein control samples were excised from gels, and processed in parallel to achieve the detection of cross-linked peptides. However, this method is mostly applied to protein complex mass lower than ~ 1 million Dalton. Because of the super large molecular weight of the pre-assembled E2o component, the cross-linked OGDHc hardly migrated to 3 % SDS-gel, hence the classical SDS-gel enrichment method was impractical for our OGDHc samples. In order to get insight to the proximity map in OGDHc, two other enrichment methods were tested in this study.

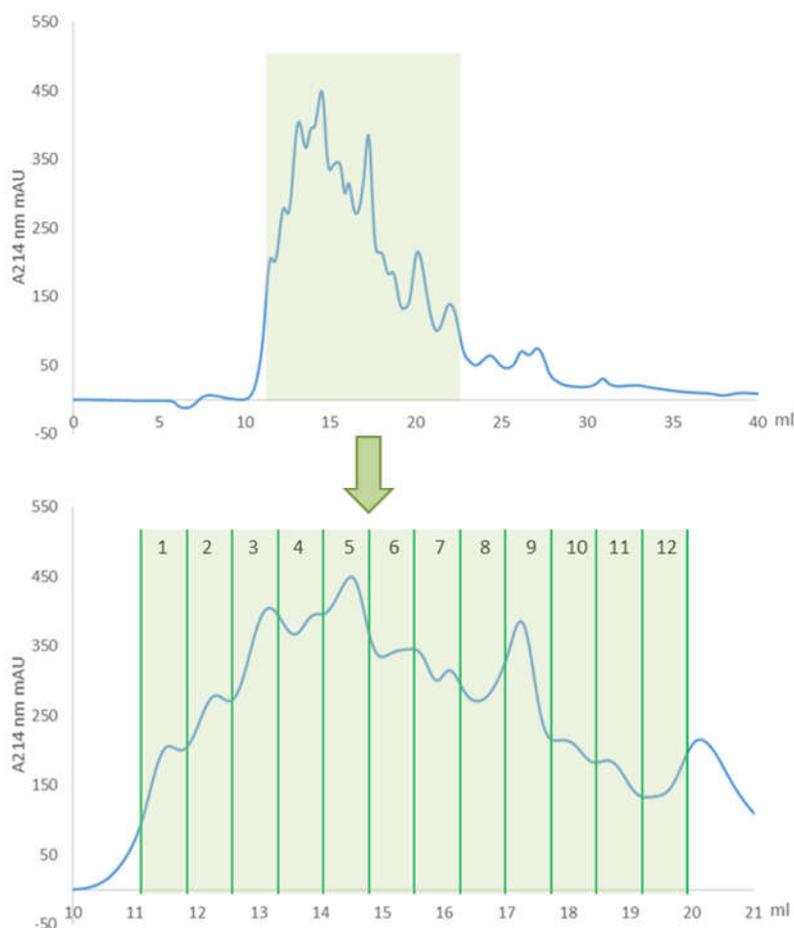
First, we have attempted to use the affinity tag method to enrich cross-linked peptides (Figure 4.4). Several research groups have introduced biotin as the affinity group, and isolation of modified peptides by avidin affinity beads [12-14]. However, we have found that the biotinylated cross-linker (CBDPSS) was not working well for OGDHc. Even at a very high cross-linker concentration (1 mM CBDPSS), the individual proteins were barely cross-linked. Given the fact that most of the biotinylated cross-linked peptides were identified at unstructured domains of the protein complex, or with small synthetic peptides [12, 15], we suspect that by introducing a bulky biotin group to the cross-linker, the cross-linker accessibility to rigid regions is decreased, and it is expected to adversely affect cross-linking in a large and tightly assembled complex.



**Figure 4.4 General analytical strategy for biotinylated cross-linker enrichment by chemical cross-linking and mass spectrometry**

In addition to the SDS-gel and affinity tag enrichment methods, a size-exclusion chromatography (SEC) was also used for enrichment purposes. Enrichment of cross-linked peptides by SEC has been evaluated and applied to many protein cross-linking studies

before [5, 16]. Because of the increase in molecular weight upon cross-linking, most of the cross-linked peptides should have a mass above 2000 Da. Therefore, cross-linked peptides will be eluted earlier, and can be enriched to some degree. Figure 4.5 shows a typical SEC elution profile of a peptide mixture pre-purified by SPE. Peptides in the mass range from 1046 to 6511 were collected in 12 fractions, each fraction was vacuum dried and submitted to LC-MS analysis. As expected, the maximum number of cross-linked peptides was identified from fractions 3 to 9, and no cross-linked peptides were found from fraction 12.



**Figure 4.5 Enrichment of cross-linked peptides with SEC method.**

Tryptic peptides resulting from a sample of cross-linked OGDHc were passed through a Superdex Peptide 10/300 GL column, 12 fractions from 11 to 20 ml were collected for LC-MS identification.

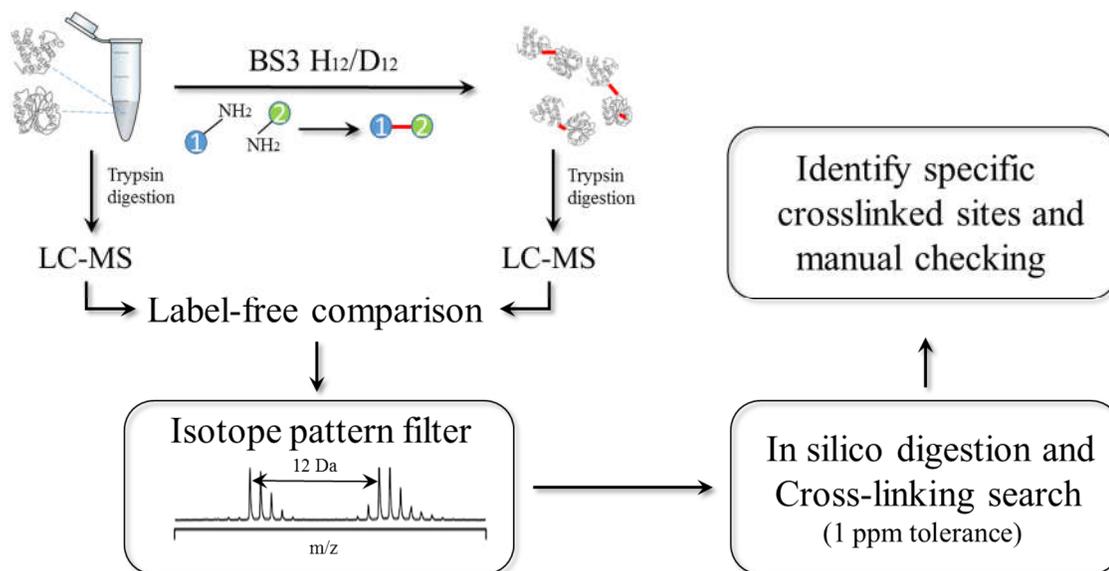
#### **4.3.2 OGDHc assembly profiling by the CX-MS method.**

Figure 4.6 shows our method used for the CX-MS data acquisition and data analysis in this chapter. Our LC-MS analysis shows that a total of 70 unique cross-links were identified (Table 4.1). Among them 45 were identified as inter-cross-links, 25 were identified as intra-cross-links. The linkage maps have been plotted from these results to visualize these inter and intra-cross-links (Figure 4.7 and 4.8).

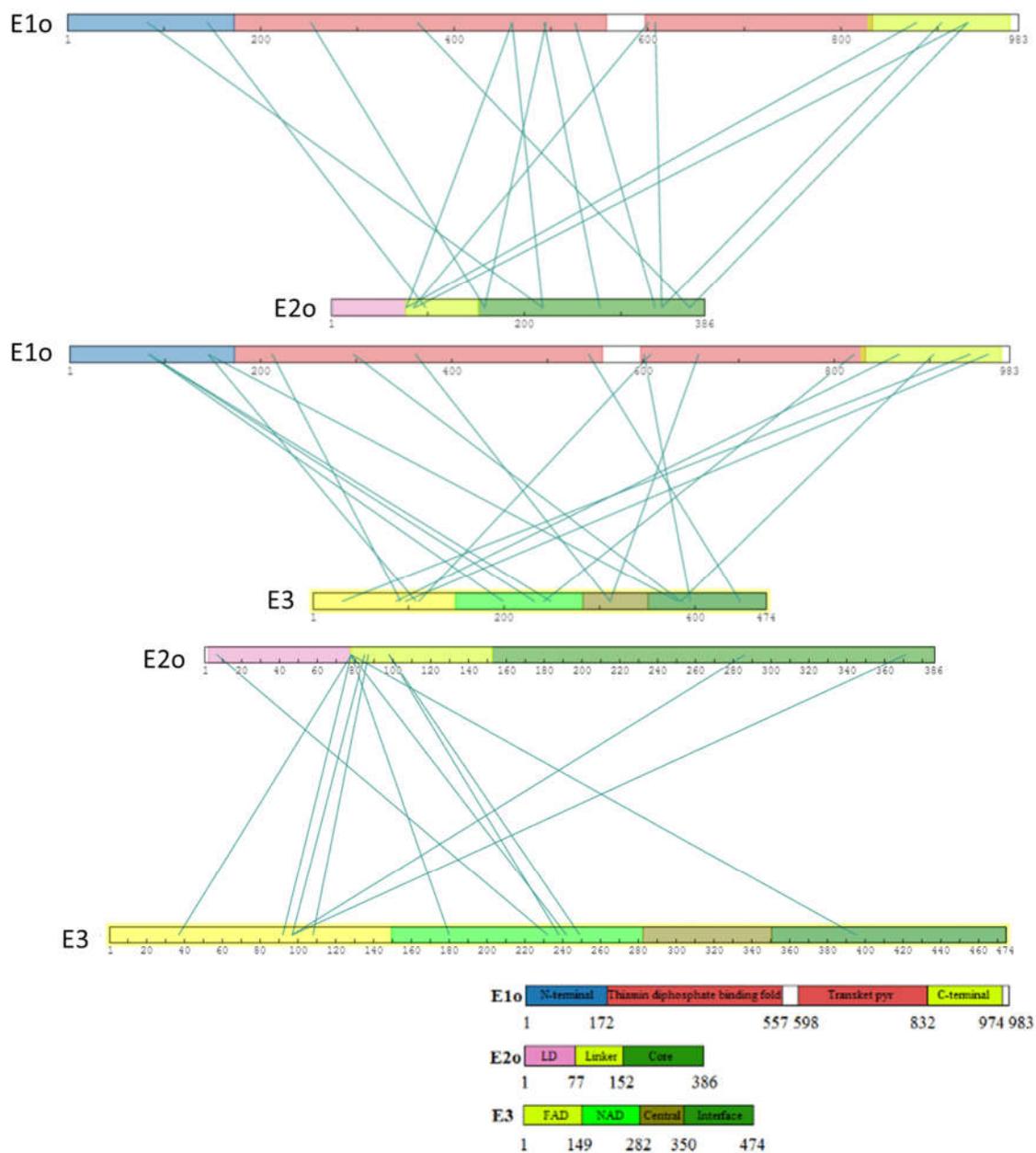
The cross-linking maps clearly highlighted several areas with many linkages that have been detected. In particular, most of the inter-crosslinking sites on E2o were identified from the end of the lipoyl domain, in good agreement with the mechanism of OGDHc catalysis, which serves as a swinging arm to shuffle reaction intermediates between active sites. Extensive cross-linking was observed among E1o, E2o and E3, suggesting a strong network of connections among all components of OGDHc. Similar to our HDX-MS results, the cross-links detected from multiple regions on E1o to both E2o and E3 established the central role of E1o in OGDHc assembly. Importantly, lysine K83 and K145 from the E1o N-terminal region are cross-linked with lysine K108 from the E3 FAD domain, lysine K199, K232 and K249 from E3 NAD binding domain, and lysine K382 from E3 interface domain. As identified with the HDX-MS method, upon E1o/E3 association, the most significant changes were observed at the N-terminal region of E1o and FAD and interface

domain in E3. The vicinity of these lysine residues identified with CX-MS method agreed well with our HDX-MS data, which strongly suggests that the close of E1o N-terminal region is close to a large surface area of E3 required for E1o/E3 sub-complex formation.

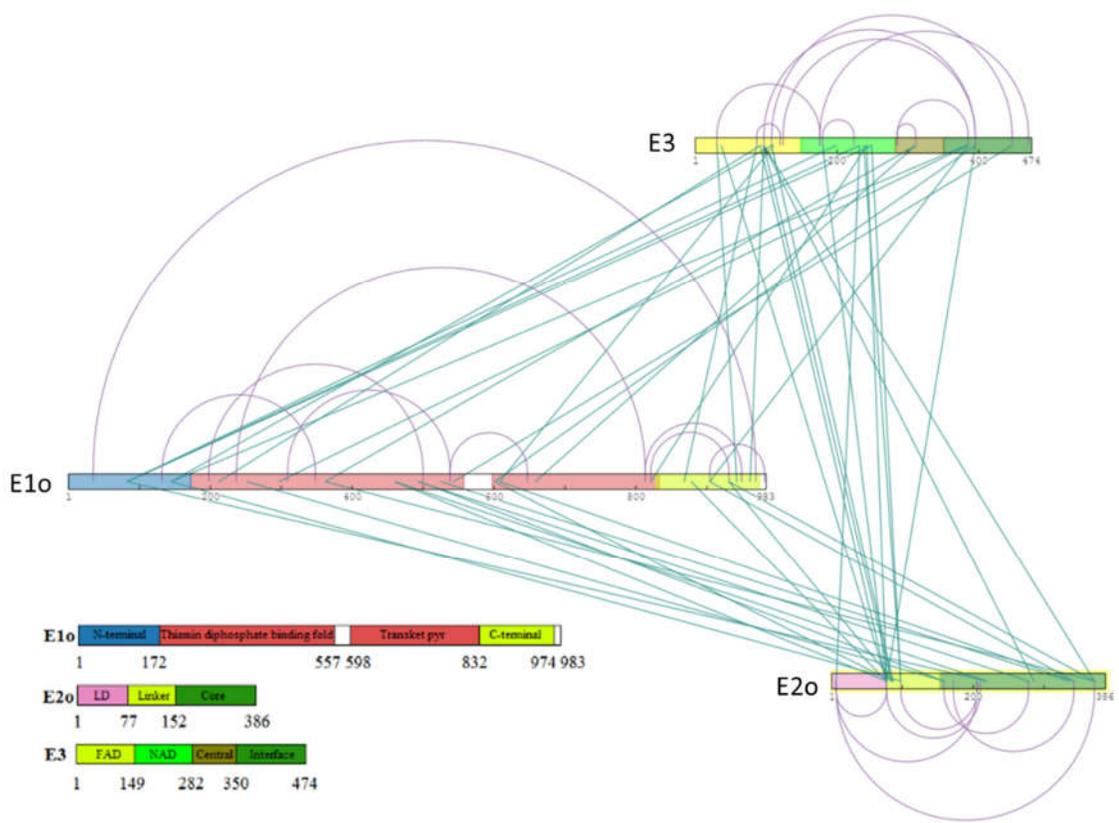
Although many crosslinks were identified between E1o and E2o, no cross-linking was observed between the N-terminal region of E1o and the linker-core region on E2o. The absence of observed cross-linked peptides between these previously well characterized two interaction regions from the E1o/E2o sub-complex might be very interesting to exploit in the future. Here, two reasons were proposed to explain the absence of cross-linking between these two regions. First, the lysine distance on these two regions is longer than 11.4 Å (the spacer arm length of cross-linker BS3). Considering there are only two lysines among the first 80 amino acids from E1o N-terminal region, it is highly possible that a longer distance between two lysines can prevent the cross-linking reaction. Second, E1o N-terminal is lacks trypsin digestion sites (four cutting sites on E1o N-terminal are K31, K35, R 42, and R62). In addition, trypsin is bound to miss cutting the cross-linker modified lysine residues. As a result, the mass of a crosslinked peptide containing this region could be very large. Since the MS signal intensity of large peptides can be suppressed by co-eluted small peptides, it is possible that the MS signal of this kind of cross-linked peptide is too low to confidently identify.



**Figure 4.6** Schematics for the CX-MS data acquisition and data analysis pipeline.



**Figure 4.7 Inter-cross-links within the OGDHc obtained with BS3 H<sub>12</sub>/D<sub>12</sub> cross-linker**



**Figure 4.8 Inter-(green) and intra-cross-linking (purple) map within the OGDHc components obtained with BS3 H12/D12 cross-linker**

Table 4.1 LC-MS analysis of tryptic peptides derived from OGDHc cross-linking

Protein1	Protein2	LinkPos1	LinkPos2	Observed mass	Theoretical mass	Error (ppm)
E3	E3	97	395	1603.8333	1603.832	-0.9
E3	E3	108	447	3720.9294	3720.9324	0.8
E3	E3	176	470	4157.2499	4157.2473	-0.6
F3	F3	285	311	2764.5206	2764.5224	0.7
E3	E3	180	224	2116.1937	2116.1919	-0.9
E3	E3	285	385	3527.9021	3527.9014	-0.2
L3	L3	31	176	2145.2017	2145.1993	-1.1
E3	E3	124	395	2406.3131	2406.3145	0.6
E3	E3	87	120	2613.3722	2613.371	-0.4
E2o	E3	286	97	3706.0483	3706.0444	-1
E2o	E2o	78	210	5741.0793	5741.078	0.2
E2o	E3	78	395	3301.7439	3301.7466	0.8
E2o	E3	371	97	2228.2795	2228.2807	0.5
E2o	E3	87	108	3781.0892	3781.0917	0.7
E2o	E2o	159	278	1704.8853	1704.8842	0.6
E2o	E2o	172	342	4761.4162	4761.42	0.8
F2o	F2o	98	205	1866.0394	1866.0377	-0.9
E2o	E2o	7	206	1978.0429	1978.0432	0.1
E2o	E2o	7	78	2082.0935	2082.0905	-1.4
F2o	F2o	7	371	2302.7107	2302.7117	0.4
E2o	E3	98	249	1664.8862	1664.8859	-0.2
E2o	E3	78	92	1745.9304	1745.9294	0.6
L2o	L3	98	238	1881.0492	1881.0486	-0.3
E2o	E3	78	242	2240.3306	2240.3342	1.6
E2o	E3	85	97	2736.5439	2736.5412	-1
E2o	E3	78	37	2743.5072	2743.5068	-0.1
E2o	E3	7	232	2845.5069	2845.5035	-1.2
L2o	L3	78	395	3301.7475	3301.7466	-0.3
E2o	E3	78	180	3502.9773	3502.9824	1.4
E1o	E2o	525	335	1203.6624	1203.6625	0.1
F1o	F2o	252	159	2303.7142	2303.7143	0.1
E1o	E2o	460	78	2545.3762	2545.3778	0.6
E1o	E2o	460	219	3685.884	3685.8866	0.7
E1o	E3	904	385	2100.1998	2100.1996	-0.1
E1o	E3	608	111	2342.2676	2342.2655	-0.9
F1o	F1o	813	949	2792.3086	2792.308	-0.3
E1o	E1o	309	538	2323.2055	2323.2033	-1
E1o	E2o	878	78	1312.7594	1312.7589	0.4
E1o	E3	212	92	3670.8649	3670.8647	-0.1
E1o	E1o	132	348	5712.9295	5712.9322	0.5
E1o	E3	362	311	2816.3914	2816.3929	0.5
L1o	L1o	237	813	2761.5405	2761.5405	0
E1o	E2o	931	87	6068.1728	6068.1748	0.3
L1o	L3	601	395	2101.1581	2101.1579	-0.1
F1o	F3	961	97	2760.5527	2760.5524	-0.1
E1o	E3	297	385	2817.3948	2817.3922	-0.9
F1o	F1o	904	981	3677.999	3677.9973	-0.5
E1o	E2o	362	371	2817.4493	2817.4497	0.1
E1o	E3	868	87	2431.2964	2431.2947	-0.7
E1o	E1o	538	647	2026.0113	2026.0106	-0.3
E1o	E1o	500	198	2392.2128	2392.211	-0.7
E1o	E2o	601	85	1927.1039	1927.105	0.6
E1o	E2o	931	371	1928.0298	1928.0282	-0.8
E1o	E2o	608	342	2352.3171	2352.3186	0.6
E1o	E2o	904	342	2922.6575	2922.6596	0.7
L1o	L3	658	311	2002.0145	2002.0146	0.1
E1o	E3	821	242	3157.8473	3157.8465	-0.3
E1o	E3	543	447	3677.9304	3677.9266	-1
E1o	E1o	821	931	3797.1053	3797.1019	0.9
L1o	L2o	495	159	1654.9323	1654.9314	-0.5
E1o	E2o	493	278	2055.1906	2055.1926	1
E1o	E3	942	31	1647.944	1647.9434	-0.4
F1o	F1o	35	969	2004.0673	2004.0694	1
E1o	E2o	145	98	2602.2905	2602.2928	0.9
E1o	E2o	83	219	3241.691	3241.6931	0.7
E1o	E3	145	108	2736.4148	2736.4136	-0.4
E1o	E3	83	232	2743.4991	2743.4994	0.1
F1o	F3	145	382	2760.3542	2760.3507	-1.3
E1o	E3	83	249	2776.4512	2776.4481	-1.1
E1o	E3	83	199	5784.0513	5784.0434	-1.4

#### **4.4 Conclusion**

A previous crystallography study of a bacterial E1o has indicated that it cannot be crystallized unless the N-terminal region (first 77 amino acids) was removed [17-19]. This study suggests that the N-terminal region of E1o is a highly dynamic or natively-disordered periphery.

The CX-MS method can identify the cross-linked proteins and reveal precisely which residues were involved in the cross-linking [5, 6]. The cross-linked peptide could provide information regarding the distance between cross-linked residues, which would increase the resolution of the method for protein structure determination.

Although we did not further examine the information from CX-MS results, several important findings have been discovered. First, the sample enrichment step is critical in our current CX-MS method. With SEC enrichment we identified 70 cross-linked peptides from OGDHc. Second, the CX-MS experiments revealing a network of connections among all components of OGDHc. The association of E1o to E2o and E3 played an important role in OGDHc assembly. Last, combined with the other techniques, the spatial proximity information provided by chemical cross-linking can help to reveal the structure of protein components in OGDHc.

#### **4.5 Future work and recommendations**

Although a very low mass tolerance (1.5 ppm) and isotope pattern filter were used to identify the cross-linked peptides, sequencing these cross-linked peptides could unambiguously identify the cross-linked peptide candidates. Moreover, various situations

could prevent the cross-linking reaction and identification of cross-linked peptides. For example there is no lysine in a particular region, or there is a lack of trypsin cutting sites. Various types of cross-linkers or proteases could be used to obtain a more comprehensive cross-linking map. More systematic data processing should also be done to not only look at the cross-linked peptides, but also the dead end peptides to help understand the potential problems in CX-MS experiments. A more accurate conclusion can be drawn when using a combination of CX-MS, HDX-MS and other complementary data.

### **Structural modeling**

Although the structure of human E1 $\alpha$  and E2 $\alpha$  is unknown, several domains comprising these two proteins are available or can be predicted by homology modeling of closely related sequences. The CX-MS results can be used to help structural modeling with MODELER and docking tools, and the results from structural modeling will be important to understand how these OGDHc components assemble through the E1 $\alpha$  N-terminal region, and furthermore the results can provide a framework to move forward to answer questions such as: Can the assembly of OGDHc result in different phenotypes on the E1 $\alpha$  N-terminal region? How do these interactions regulate the overall activity of this complex? Are there more potential interactions among the three components which contribute to the stability of this complex?

## Reference

1. Huang, B.X., Kim, H.Y., and Dass, C. (2004) Probing three-dimensional structure of bovine serum albumin by chemical cross-linking and mass spectrometry. *J. Am. Soc. Mass Spectrom.* *15*, 1237-1247.
2. Schulz, D.M., Ihling, C., Clore, G.M., and Sinz, A. (2004) Mapping the topology and determination of a low-resolution three-dimensional structure of the calmodulin-melittin complex by chemical cross-linking and high-resolution FTICRMS: direct demonstration of multiple binding modes. *Biochemistry* *43*, 4703-4715.
3. Leitner, A., Joachimiak, L.A., Unverdorben, P., Walzthoeni, T., Frydman, J., Förster, F., and Aebersold, R. (2014) Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 9455-9460.
4. Seebacher, J., Mallick, P., Zhang, N., Eddes, J.S., Aebersold, R., and Gelb, M.H. (2006) Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing. *J. Proteome. Res.* *5*, 2270-2282.
5. Leitner, A., Walzthoeni, T., and Aebersold, R. (2014) Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat. Protoc.* *9*, 120-137.
6. Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M., Cramer, P., and Rappsilber, J. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* *29*, 717-726.
7. Chowdhury, S.M., Du, X., Tolić, N., Wu, S., Moore, R.J., Mayer, M.U., Smith, R.D., and Adkins, J.N. (2009) Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry. *Anal. Chem.* *81*, 5524-5532.
8. Nadeau, O.W., Wyckoff, G.J., Paschall, J.E., Artigues, A., Sage, J., Villar, M.T., and Carlson, G.M. (2008) CrossSearch, a user-friendly search engine for detecting chemically cross-linked peptides in conjugated proteins. *Mol. Cell. Proteomics.* *7*, 739-749.
9. Kukacka, Z., Rosulek, M., Strohal, M., Kavan, D., and Novak, P. (2015) Mapping protein structural changes by quantitative cross-linking. *Methods* *89*, 112-120.
10. Schmidt, C., and Robinson, C.V. (2014) A comparative cross-linking strategy to probe conformational changes in protein complexes. *Nat. Protoc.* *9*, 2224-2236.

11. Hu, J., Wang, Y., Zhang, X., Lloyd, J.R., Li, J.H., Karpiak, J., Costanzi, S. and Wess, J. (2010) Structural basis of G protein-coupled receptor-G protein interactions. *Nat. Chem. Biol.* *6*, 541-548.
12. Petrotchenko, E.V., Serpa, J.J., Hardie, D.B., Berjanskii, M., Suriyamongkol, B.P., Wishart, D.S., and Borchers, C.H. (2012) Use of proteinase K nonspecific digestion for selective and comprehensive identification of interpeptide cross-links: application to prion proteins. *Mol. Cell. Proteomics.* *11*, M111.013524.
13. Hurst, G.B., Lankford, T.K., and Kennel, S.J. (2004) Mass spectrometric detection of affinity purified crosslinked peptides. *J. Am. Soc. Mass Spectrom.* *15*, 832-839.
14. Luo, J., Fishburn, J., Hahn, S., and Ranish, J. (2012) An integrated chemical cross-linking and mass spectrometry approach to study protein complex architecture and function. *Mol Cell. Proteomics.* *11*, M111.008318.
15. Petrotchenko, E.V., Serpa, J.J., and Borchers, C.H. (2011) An isotopically-coded CID-cleavable biotinylated crosslinker for structural proteomics. *Mol. Cell. Proteomics.* *10*, M110.001420.
16. Seifert, F.U., Lammens, K., Stoehr, G., Kessler, B., and Hopfner, K.P. (2016) Structural mechanism of ATP-dependent DNA binding and DNA end bridging by eukaryotic Rad50. *EMBO J.* *35*, 759-772.
17. Kresze, G. B., Ronft, H., Dietl, B., and Steber, L. (1981) Limited proteolysis of 2-oxoglutarate dehydrogenase multienzyme complex from bovine kidney. *FEBS Lett.* *127*, 157-160.
18. Rice, J. E., Dunbar, B., and Lindsay, J. G. (1992) Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. *The EMBO journal.* *11*, 3229-35.
19. McCartney, R. G., Rice, J. E., Sanderson, S. J., Bunik, V., Lindsay H., and Lindsay, J. G. (1998) Subunit interactions in the mammalian alpha-ketoglutarate dehydrogenase complex. Evidence for direct association of the alpha-ketoglutarate dehydrogenase and dihydrolipoamide dehydrogenase components. *J. Biol. Chem.* *273*, 24158-24164.

## CHAPTER 5. DXP synthase structure and dynamics studied with HDX-MS.

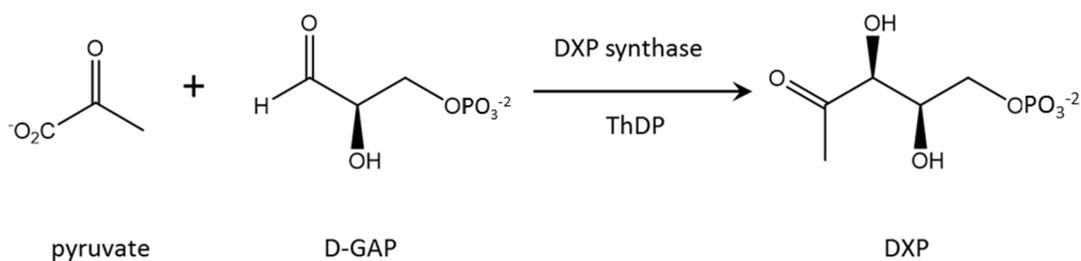
### 5.1 Introduction

The biosynthesis of isoprenoids such as steroid hormones, cholesterol, and coenzyme Q is very important in all living cells [1]. The unique methylerythritol phosphate (MEP) pathway for isoprenoid biosynthesis is found to be essential in several human pathogens, including mycobacterium tuberculosis [2]. Because the MEP pathway is absent in mammals, this makes it an attractive target for the development of new anti-infective drugs. 1-Deoxy-D-xylulose 5-phosphate synthase (DXPS) is the first enzyme in the MEP pathway, and it catalyzes the formation of DXP from D-glyceraldehyde 3-phosphate (D-GAP) and pyruvate (Figure 5.1). DXPS is a ThDP dependent enzyme which exhibits a limited sequence homology to the E1p subunit from PDHc and transketolase [3].

Protein-ligand interactions are involved in all of enzymatic catalysis and characterization of such interactions is essential to enzymatic activity, and the regulation of biological function. Understanding the modulation of protein-ligand interactions is important for the elucidation of biological processes and is considered a promising strategy towards specific inhibitor development. Most small molecule ligands usually bind to well-defined cavities of enzymes (substrate binding pocket). Although several papers have been published on the mechanism and inhibition of DXPS [2, 4-7], little information was known regarding the substrate binding pocket. The *E. coli* DXPS could be crystallized only after *in situ* proteolysis by a fungal protease [3]. A X-ray crystal structure of the remaining parts of the *E. coli* DXPS was lacking residues 183-238 and 292-317. The poorly defined *E. coli*

DXPS crystal structure resulted in ambiguities in assessing the active site and the domain organization.

In this chapter, a hydrogen-deuterium exchange mass spectrometric (HDX-MS) method was used to explore DXPS structure and dynamics upon incubating with its two substrates. In parallel with these experiments, the pyruvate structural analogue methylacetylphosphonate (MAP) and the product DXP were also used for structural characterization of DXPS. By understanding how DXPS structure and dynamics respond to different ligands, a better picture of ligand induced DXPS conformational changes can be provided to help elucidate the mechanism of catalysis, and enable structure-based inhibitor design.



**Figure 5.1 DXPS catalyzed carboligation of pyruvate and D-GAP.**

## 5.2 Materials and methods

Deuterium oxide ( $D_2O$ ) was from Cambridge Isotope Laboratories. All other fine chemicals were from Sigma-Aldrich. The *E. coli* DXPS was provided by Dr. Meyers' group from the John Hopkins University School of Medicine.

### 5.2.1 Sample preparation for HDX.

Prior to H/D exchange, the DXPS was exchanged into 50 mM HEPES (pH 8.0) buffer with 50 mM NaCl, 0.2 mM ThDP, and 1 mM MgCl<sub>2</sub>, and then the protein concentration was adjusted to 80 μM. The concentrated DXPS (80 μM) was incubated with or without pyruvate, methylacetylphosphonate (MAP), D-glyceraldehyde-3-phosphate (D-GAP), or 1-deoxy-D-xylulose 5-phosphate (DXP) at a final concentration of 1 mM for 20 min at 25 °C prior to initiation of the HDX experiments. The HDX experiments were initiated by mixing 15 μl of the protein samples with 285 μl of D<sub>2</sub>O buffer to yield a final concentration of 95 % D<sub>2</sub>O at pH 8.0. D<sub>2</sub>O buffer was prepared the same way as DXPS exchange buffer except 99.9% D<sub>2</sub>O was used to dissolve the buffer components. The samples were incubated at 25 °C for 20 s, 1, 2, 3, 5, 10 and 30 min, and then quenched by rapidly mixing with an equivalent of ice-cold quench buffer (trifluoroacetic acid, 2 M guanidine hydrochloride, pH 1.4) to reduce the final sample pH to 2.5. The samples were immediately frozen in liquid nitrogen and stored at -80 °C before analysis. Un-deuterated samples were generated following the same procedure except that protein samples were diluted into aqueous buffer and incubated for 3 min followed by the quenching process.

### 5.2.2 LC-MS method and data processing

The frozen deuterated sample was quickly thawed and loaded with an ice-cold syringe into a 20-μl sample loop inside the refrigeration system. The protein sample (40 pmol) was carried by a 0.2 ml/min digestion flow (0.1% formic acid) into an immobilized pepsin column (Poroszyme Immobilized Pepsin Cartridge, 2.1 × 30 mm, Applied

Biosystems) and digested at 15 °C for 30 s. The resultant peptides were immediately cooled to 0 °C through a heat exchanger and were concentrated and desalted on a peptide trap (Michrom Peptide MacroTrap, 3 × 8 mm). The peptides were eluted and separated in 15 min through a reversed-phase C18 HPLC column (Agilent Poroshell 300SB-C18, 2.1 × 75 mm) at a flow rate of 0.2 ml/min at 0 °C using a 2– 40% acetonitrile gradient containing 0.1% formic acid. ESI-Fourier transform-mass spectrometry (FT-MS) measurements began 5 min after the initiation of the elution process and lasted for 10 min. The time from initiation of digestion to elution of the last peptide was less than 20 min. Bruker Daltonics DataAnalysis 4.0 was used for spectrum analysis and data treatment. Peptides were identified from undeuterated samples by a customized program DXgest, which matches experimental peptide mass with theoretically generated peptic peptide mass by using statistical data for the pepsin cleavage pattern under HDX conditions. Mass tolerance was set at 1.0 ppm. The bimodal EX1 kinetics MS data were deconvoluted with HX-Express2 [8].

## **5.3 Results and Discussion**

### **5.3.1 Overview of HDX patterns**

The time dependence of hydrogen-deuterium exchange of the backbone amide protons of DXPS was carried out at 20 s, 1, 2, 3, 5, 10 and 30 min. On-line digestion by pepsin followed by LC-MS analysis under our HDX conditions resulted in 56 peptides, many of them partially overlapping, providing 93.7% sequence coverage (Table 5.1). The availability of redundant peptides is very important, because it provides an internal

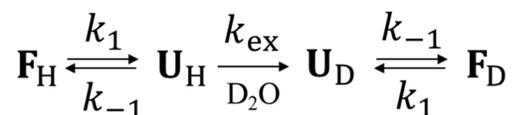
consistency control. In this study, the peptides from DXPS displayed two different types of HDX kinetics, EX1 and EX2. The overlapped peptides could help to assure that peptides representing similar protein regions exhibit comparable HDX kinetics. The high sequence coverage, including the regions 183-238 and 292-317, which were missing from the published *E. coli* DXPS structure [3], provided access to structural changes throughout the protein.

The average deuterium uptake percentage at the 30 min time point is mapped onto the crystal structure of DXPS (PDB: 2o1s). Since two regions of importance to this study (183-238, and 292-317) are missing from the published crystal structure, these regions are illustrated schematically (Figure 5.2). The overall deuterium incorporation rate is relatively low on this enzyme, many regions underwent less than 35% deuteration after 30 min exchange time (see Figure 5.2 and Figure 5.3 A). The deuterium incorporation pattern clearly indicated that the greatest extent of incorporation occurs in domain I, suggesting that this is the most dynamic domain among the three domains in DXPS. The areas with the least extent of deuterium incorporation are located at the DXPS dimer interface, according to the crystal structure. These highly protected regions are located mainly in domains II and III.

### **5.3.2 EX1 and EX2 kinetics to interpret HDX experiments.**

Two exchange mechanisms have been proposed to explain hydrogen exchange under physiological conditions in proteins. The first mechanism suggests that exchange may occur from the folded form, the other suggests partial or global unfolding first,

followed by exchange [9-11]. Under physiological conditions, the unfolding is the predominant mechanism for HD exchange in proteins.



### Equation 5.1

As illustrated in equation 5.1, there are three steps involved in backbone amide hydrogen exchange in proteins. In order to exchange H for D, the folded conformation  $\mathbf{F}_H$  is converted to the unfolded conformation  $\mathbf{U}_H$ , and exposes backbone amide hydrogens for exchange with D from  $\text{D}_2\text{O}$  solvent. The rate constants  $k_1$ ,  $k_{-1}$ , and  $k_{\text{ex}}$  describe the unfolding, refolding, and hydrogen exchange steps, respectively. There are two kinetics schemes, EX1 and EX2, to describe hydrogen exchange in this case. When the exchange is governed by EX1 kinetics, the unfolding rate is much faster than the exchange rate ( $k_{\text{ex}} \gg k_{-1}$ ). Under these conditions, all of the amide hydrogens exchange with deuterium in the unfolded state before refolding occurs. As a result, EX1 kinetics gives rise to two distinct and separated mass envelopes, the lower mass envelope represents the folded state, and the higher mass envelope represents the unfolded state. In EX2 kinetics, the protein refolding rate is much faster than the exchange rate ( $k_{\text{ex}} \ll k_{-1}$ ), therefore the unfolding has to occur multiple times before a successful exchange reaction takes place. EX2 kinetics leads to a single isotopic distribution gradually shifting to a higher m/z range over time. The majority of native proteins follow EX2 kinetics under physiological conditions. EX1 kinetics is often observed when proteins are under denaturing conditions or as a reflection of experimental error, such as carryover between injections. To avoid an erroneous assignment of EX1

kinetics, in this study protein stability was carefully examined and carryover was avoided by extensively washing the pepsin and C18 columns between runs. The observation, hence occurrence of EX1 kinetics under physiological conditions can provide important insights into the nature of DXPS structural dynamics in solution with and without its substrates, substrate analogue or product.

### **5.3.3 HDX-MS detects local structural dynamics of DXPS.**

We first investigated the conformational flexibility of full length DXPS (Figure 5.2 and 5.3). Deuterium uptake for most peptides identified under the HDX conditions used displayed a single isotopic distribution (EX2 kinetics), increasing in mass over time. However, four peptides (i.e., peptides 42-56, 51-58, 183-199, and 278-298; see Figure 5.3 B) displayed a bimodal isotopic distribution (consistent with EX1 kinetics). Because peptides 42-56 and 51-58 share largely overlapping residues, they were grouped into residue 42-58 in the following discussion. Upon incubating with small molecule ligands, there was little deuterium uptake difference observed on all EX2 peptides in DXPS (Figure 5.4). A significant deuterium exchange difference was observed only on those peptides displaying EX1 behavior.

Notably, unlike most of the reported bimodal patterns, which were induced upon adding a binding partner [12, 13], the bimodal pattern of these four peptides from DXPS was already apparent in the absence of any substrates bound. Although residues 183-199 and 278-298 were missing from the crystal structure, the positions of the N-terminal residues of these peptides affirmed the spatial proximity of these four peptides. This result

indicated that all peptides displaying EX1 kinetics are proximal to each other, and the conformational change was triggered simultaneously in all four peptides.

To properly characterize EX1 kinetics with our prescreening experiments, seven time points (20 s, 1, 2, 3, 5, 10, and 30 min) were investigated. Because of the relatively low degree (< 10 %) and high degree (> 65 %) of deuteration of EX1 kinetics, peptides were observed at 20 s and 30 min respectively, additional shorter or longer time points were not acquired for this study.

A bimodal distribution could result from EX1 kinetics [14], in which two distinct conformations interconvert during the HDX incubation period. In our case, there were two conformations (folded and unfolded) in these peptides. It appears that the equilibration rate between folded and unfolded conformations of the corresponding peptide is slower than the HDX rate, which indicates that both conformations were fairly stable under physiological conditions. Since DXPS is naturally present as a homo-dimer, a plausible explanation for the EX1 kinetics of the three peptides is that the DXPS homo-dimer is asymmetric, suggesting that the two subunits of the DXPS are in different conformational states. Although this could explain the EX1 kinetics found in this study, it cannot explain the changes of relative abundance of the two isotopic distributions over time. Given the observation that residues 183-238 and 292-317 were either missing or in a disordered region from *E. coli* and *D. radiodurans* DXPS crystal structures, an alternative explanation for the observation of the EX1 kinetics is that the dynamic nature of these two regions may be responsible for the localized folding/unfolding event.

After carefully examining the four peptides displaying EX1 behavior, some interesting further properties were revealed: First, these four peptides all adopted a mixed

EX1/EX2 kinetics. In addition to the distinct EX1 signature, the lower mass envelope of each peptides gradual moved to the high mass range over time (see Figure 5.3 B and Figure 5.6-5.9). As these peptides are all long peptides, they may contain a portion that has undergone exchange via EX1 and a portion that has exchanged via the EX2 mechanism. Second, the higher mass envelope of all peptides was maximally labeled (40% deuteration for 42-56, 63% deuteration for 183-199, and 36% deuteration for 278-298) by the 20 s time point. The centroid mass difference of higher and lower mass envelopes at 20 s time point indicated 4, 8 and 6 residues in peptides 42-56, 183-199, and 278-298, suggesting that 30.8%, 53.5%, and 35.3% of the residues in the peptides, respectively, were involved in the unfolding event (see Figure 5.2 C-E).

### **5.3.4 DXPS binds and responds differently to substrates, substrate analogue and product**

Results of the DXPS saturated with pyruvate (1 mM) or MAP (1 mM, a mimic of pyruvate and known inhibitor of ThDP-dependent enzymes) revealed no significant exchange rate changes. However, in the presence of pyruvate or MAP, there was a significant slowdown of unfolding in the kinetics of all three EX1 signature peptides as mentioned above (residues 42-58, 183-199, and 278-298). The slowdown was apparent in the HDX-MS spectra (see Figure. 5.6-5.9 and Table 5.2), wherein a bimodal distribution was observed, but the higher mass envelope of the distribution (unfolded conformation) was shifted to the lower mass envelope (folded conformation). In addition, by calculating the rate constant ( $k_u$ ) or half-life ( $t_{1/2}$ ) for unfolding, the slowdown of protein dynamics could be quantified by calculating a “slowdown factor” [13]. The slowdown factor is

simply the  $t_{1/2}$  of unfolding in the presence of pyruvate or MAP divided by the  $t_{1/2}$  of unfolding without pyruvate or MAP. An example of the calculation can be gleaned from Figure 5.10, where in the presence of pyruvate, peptide 183-199 unfolding slowed from a  $t_{1/2}$  of 3 min in the free state to a  $t_{1/2}$  of 4.9 min in the pyruvate-bound state. The slowdown factor for the case of peptide 183-199 + pyruvate was therefore  $4.9/3 = 1.6$ . All of the unfolding rate constants, half-life and slowdown factors of each of the three peptides with EX1 signature are listed in Table 5.2. It is clear that compared with pyruvate, MAP has the most pronounced slowdown factor. A plausible explanation for this observation could be that as an analogue of pyruvate, MAP can bind in the large DXPS active site and react with bound ThDP to form a stable C2 $\alpha$ -phosphonolactyl-thiamin diphosphate (PLThDP) intermediate, effectively trapping the enzyme at a pre-decarboxylation complex state [15], and the pre-decarboxylation complex state could be the folded conformation (the lower mass envelope in these EX1 kinetics profiles). The shifting of EX1 kinetics data for these three peptides suggests that binding of pyruvate or MAP could induce large-scale conformational changes around the active site of DXPS, and favor the folded conformation. We therefore suggest that the folded (lower mass) conformation in these three peptides is the more favorable conformation for the decarboxylation step in the catalytic sequence of DXP synthase.

Upon binding D-GAP or DXP to DXPS, residues 42-58, 183-199, and 278-298 displayed unique unfolding events as shown in Figures 5.6-5.9. The increased higher mass envelope of the bimodal distribution suggests that D-GAP or DXPS can induce those peptides to adopt an unfolded conformation. As shown in Table 5.2, the half-life of unfolding was slightly decreased on all peptides. We next examined the role of D-GAP in

driving the sequential conformational rearrangement of DXPS subsequent to MAP recognition. HDX-MS results of DXPS with MAP bound were compared in the absence and presence and of D-GAP. All peptides consistently exhibited EX1 kinetics profiles, as induced by MAP, even in the presence of D-GAP. However, we observed a slight increase on the higher mass envelope part of all EX1 kinetics peptides in the presence of D-GAP (Figure 5.6-5.9). Although the overall isotopic distribution was more similar to the profile induced by MAP, the unfolding of the three EX1 kinetics peptides were found to be about 2 to 3 times faster than with MAP alone (Table 5.2). Together, these results suggest that the unfolded conformation in the three peptides displaying EX1 kinetics is more favorable at the carbonylation step of the DXPS catalytic sequence. The results presented suggest that those three peptides displaying EX1 kinetics play a central role in the response to the different stages of DXPS catalysis.

#### **5.4 Conclusion**

Our HDX-MS data reveal a much more detailed picture of DXPS structure and dynamics in response to binding different ligands than published before. The high sequence coverage of DXPS (93.7%) enabled us to study the flexibility of the entire enzyme, including the two regions (residues 183-238, and 292-317), which were never observed before.

HDX-MS data showed that domain I of DXPS (residues 1-319) has unusual dynamics/flexibility and that this is related to substrate recognition. Several regions (residue 42-58, 183-199, and 278-298) around the active center displayed an EX1 signature

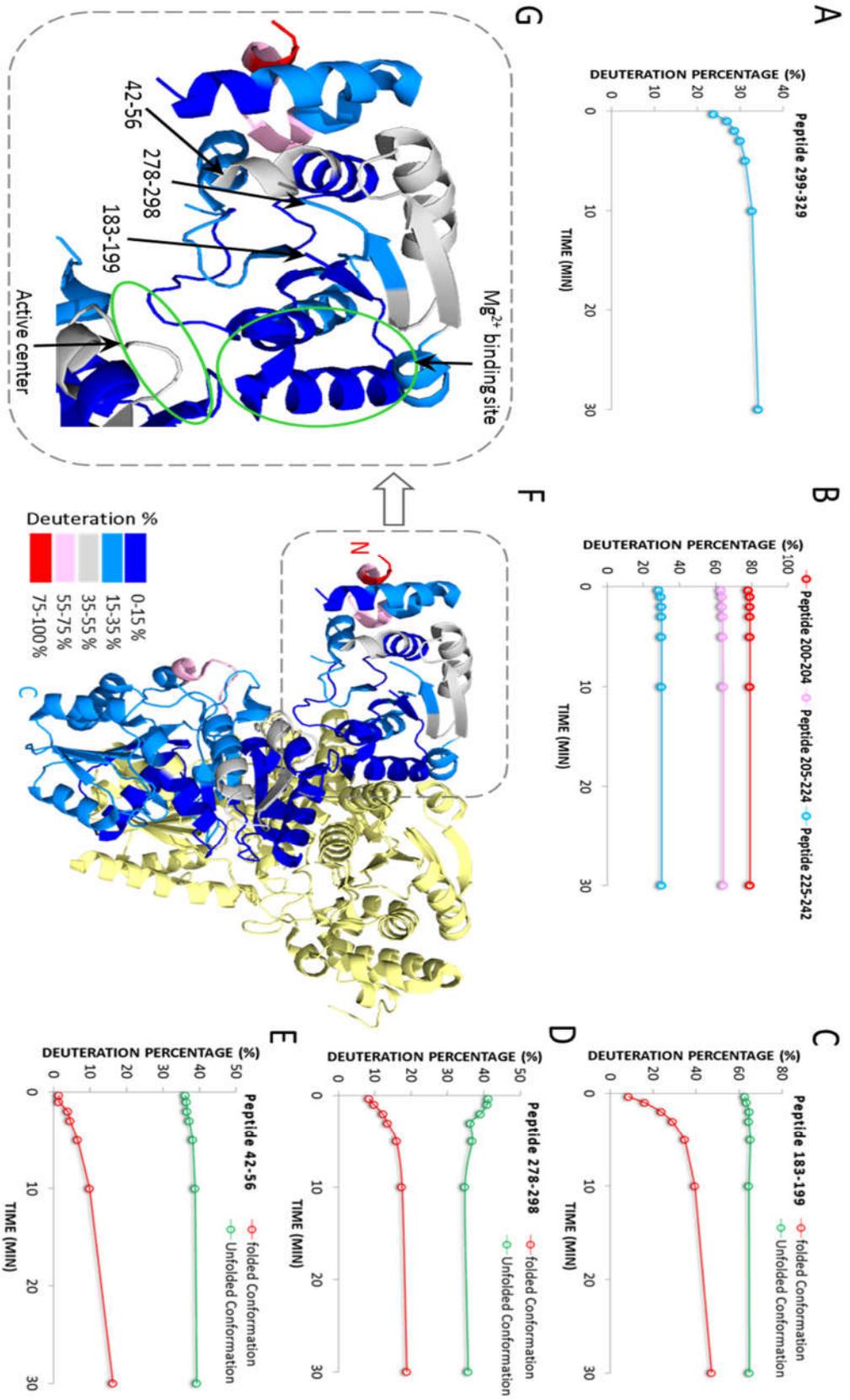
in both ligand-free and ligand-associated states (Figure 5.2). The half-life of unfolding of those peptides displaying EX1 kinetics were different under the two conditions, and those differences were also ligand-dependent (Table 5.2). The EX1 exchange kinetics characterized in DXPS is so far unique, it has not been reported for any other ThDP dependent enzyme.

Unlike DXPS, the E1p component from the pyruvate dehydrogenase multienzyme complex (PDHc), a ThDP dependent homologous enzyme to DXPS, displayed no conformational changes while associated with pyruvate substrate [16, 17]. The DXPS complexed with pyruvate or MAP favors the folded conformation in those EX1-displaying peptides, and favors the unfolded conformation while incubated with D-GAP or DXP. Thus, we propose that the folded conformation around the active center of DXPS is critical for formation of the pre-decarboxylation complex, while the unfolded conformation is favored in the carbonylation step of the DXPS catalytic sequence and subsequently in product release.

Our findings highlighted that the disordered regions in domain I of DXPS (residues 42-58, 183-199, and 278-298) play an important role in substrate recognition and enzymatic catalysis. Further understanding of the structural and functional consequences of the active site may provide a way to regulate DXPS activity. Moreover, a better understanding of the structure and dynamics around the DXPS active center could lead to design of selective inhibitors.

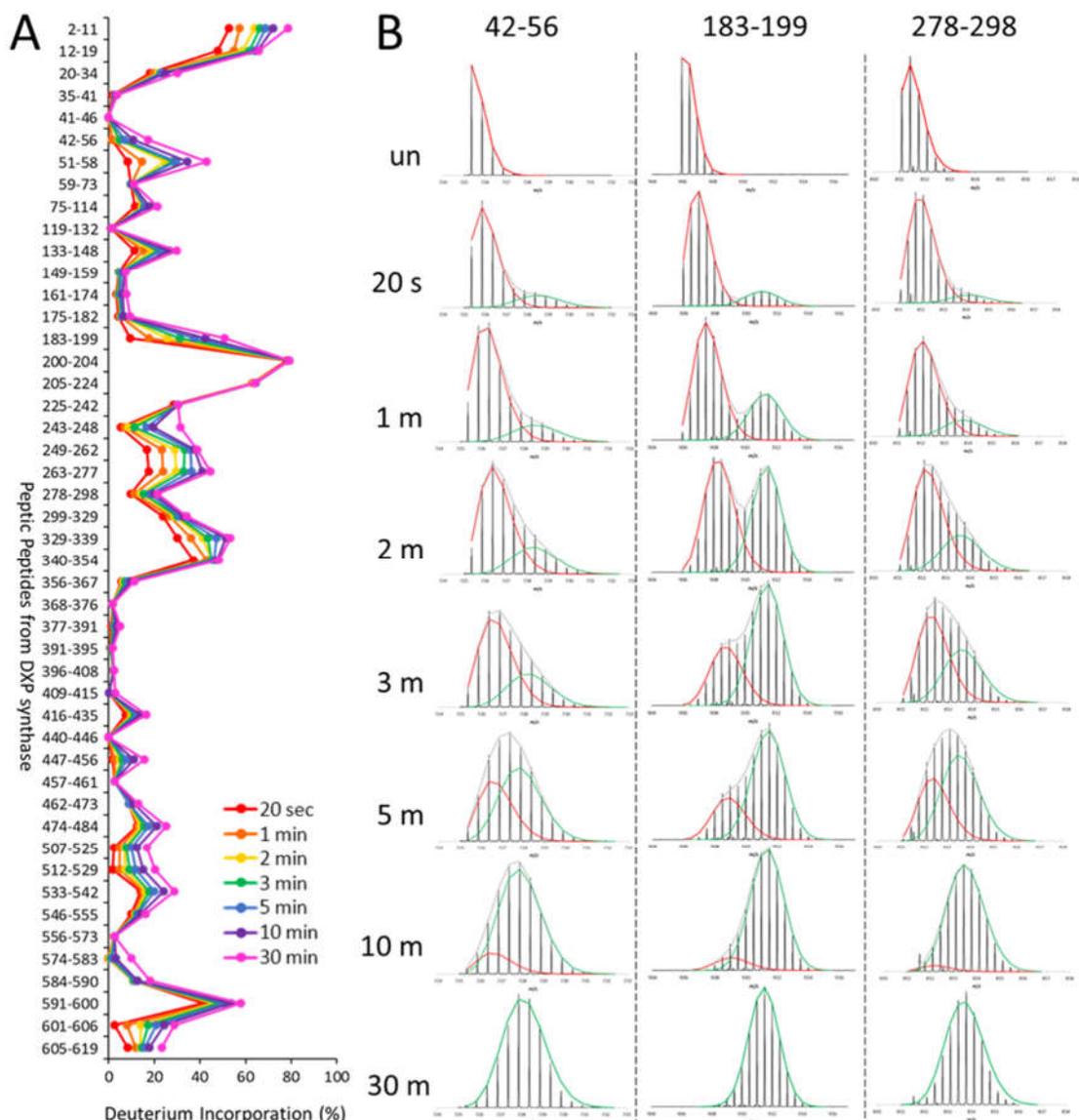
Finally, the time-dependence of the HDX-MS presented shows that HDX-MS is a valuable tool for determining the conformational changes of flexible region in DXPS, and

this study can serve as a model for studies of other highly dynamic protein structures which are difficult to study by X-ray or NMR methods.



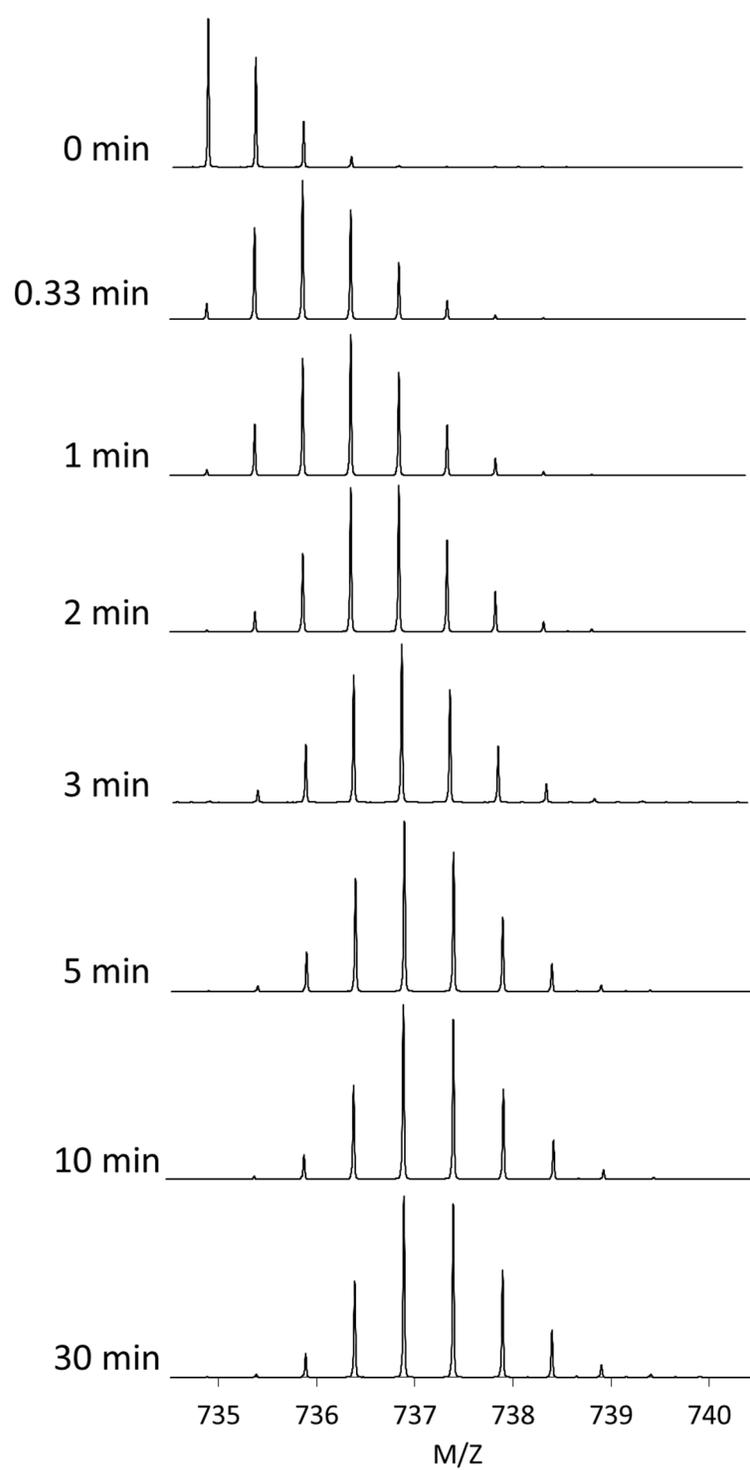
**Figure 5.2 Deuterium uptake plotted on the crystal structure of *E. coli* DXP synthase**

(A-E) The deuterium uptake plots for regions of 183-238 and 292-317, which were missing from the crystal structure. (C-E) Peptides exhibiting a bimodal isotopic distribution were plotted through 30 min. The centroid mass of lower (red trace) and higher (green trace) mass envelopes were fit and calculated individually. Since the centroid mass of lower mass envelope also changes with increasing exposure time, the displayed kinetics are likely a mixture of EX1 and EX2 kinetics. (F) The average percentage of deuterium incorporation at 30 min time point is mapped onto the crystal structure of *E. coli* DXP synthase. The color is according to the color code bar in the figure. Only one copy of the DXP synthase dimer was colored, the other copy of the dimer was colored in pale yellow as background. (G) Deuteration percentage color-coded in domain I of *E. coli* DXP synthase. The position of all EX1 kinetics peptides indicated the spatial proximity of these peptides, and they are right in front of  $Mg^{2+}$  binding site and active center of DXP synthase. The folded/unfolded conformations of these peptides could play an important role in substrate or intermediate recognition at different stages of DXP synthase catalysis.

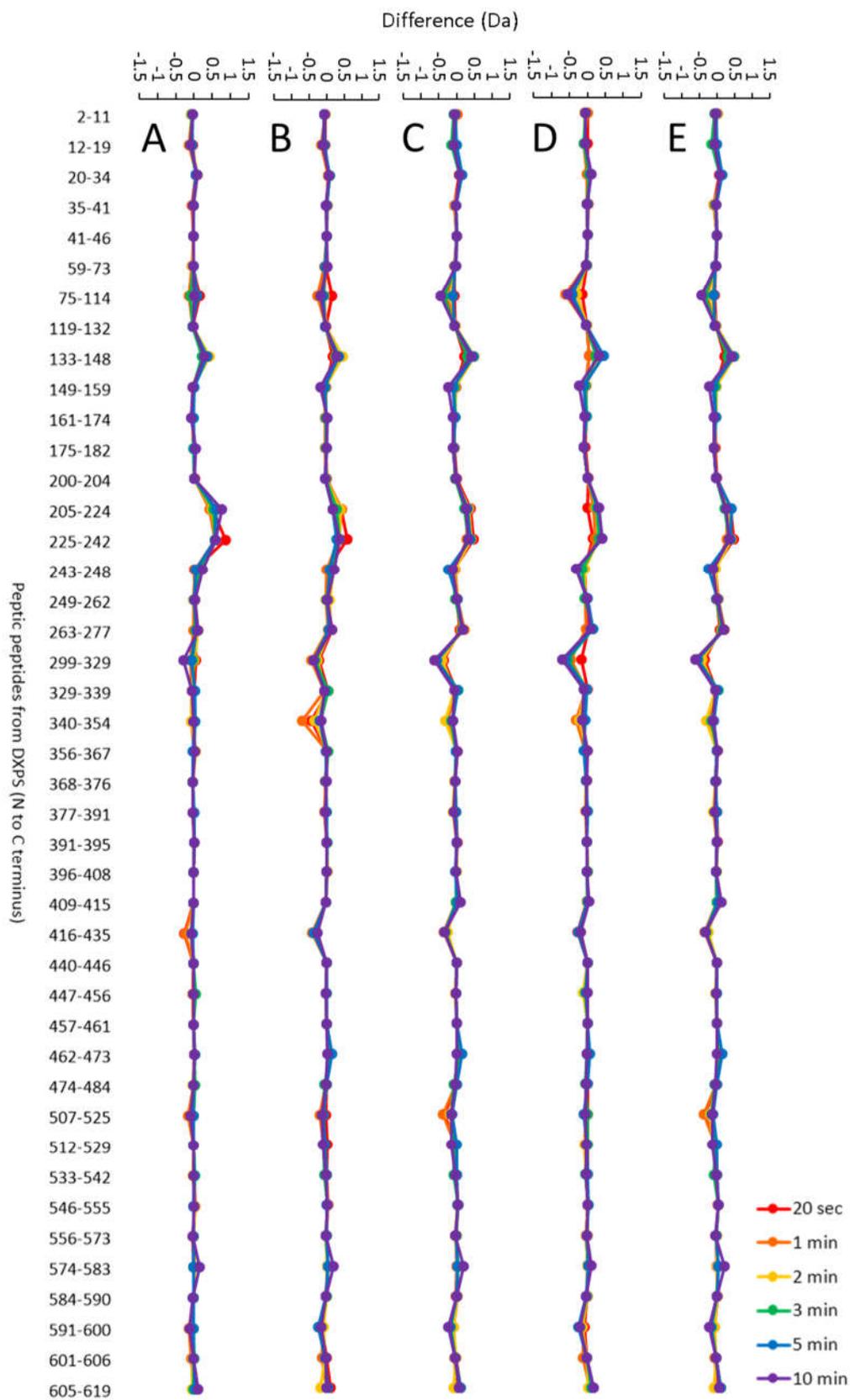


**Figure 5.3 Free state DXPS deuterium uptake change in 30 min.**

(A) Deuterium incorporation plot of full length *E. coli* DXPS across seven different time points (20 s, 1, 2, 3, 5, 10, and 30 min). (B) HDX-MS spectra of three peptides displaying EX1 kinetics (peptides 42-56, 183-199, and 278-298), a clear bimodal isotopic distribution was observed.

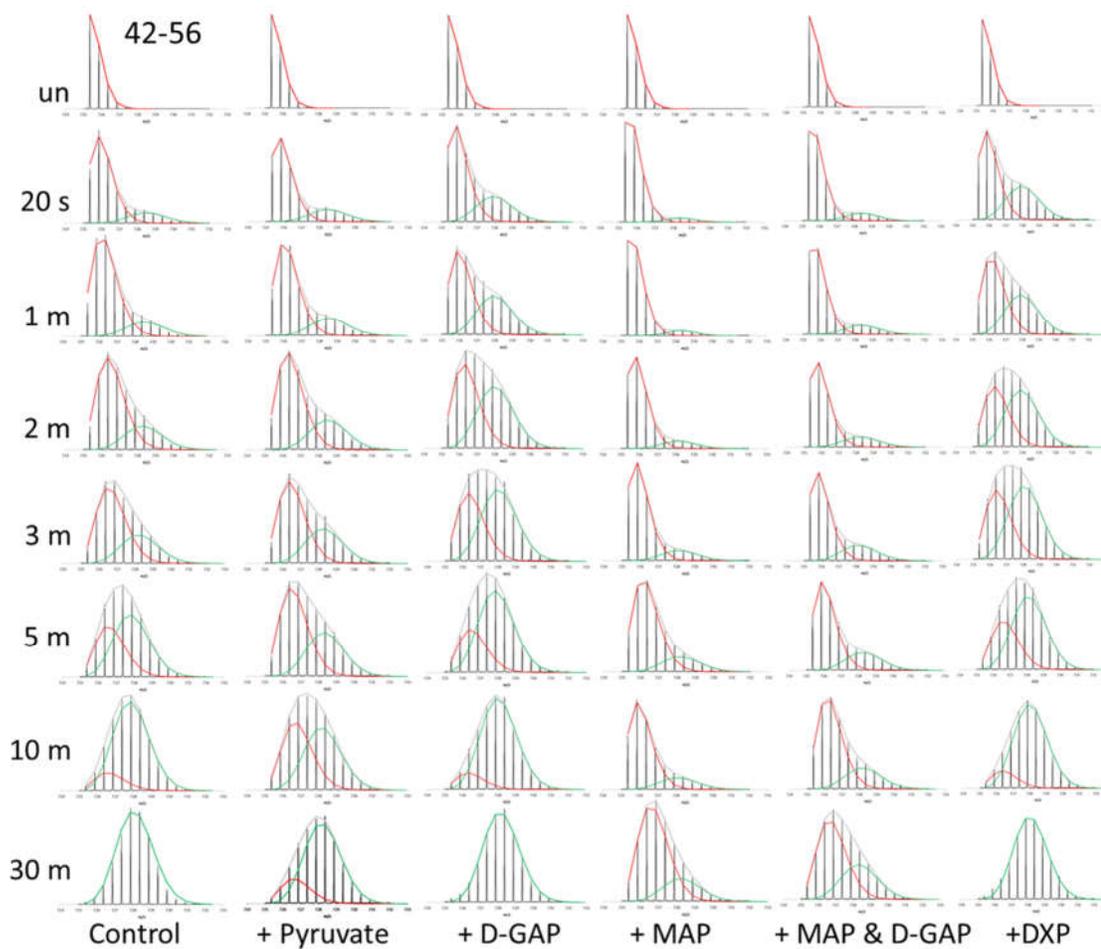


**Figure 5.4** DXPS peptic peptide 249-262 displayed a typical EX2 kinetics mass envelope.

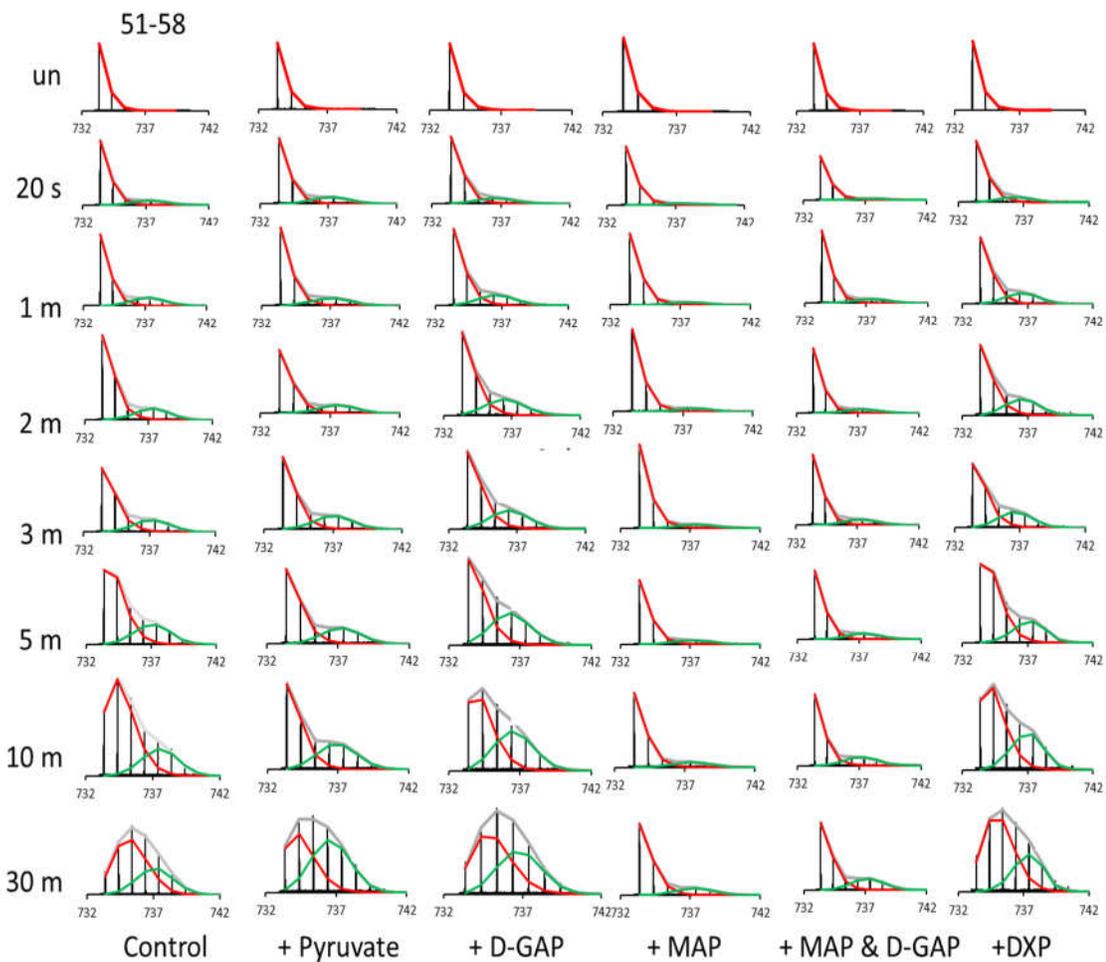


**Figure 5.5 Difference plot showing deuterium incorporation changes of peptic fragments of DXPS in the absence and presence of ligands (deuterons exchanged in the absence of DXPS minus deuterons exchanged in the presence of DXPS)**

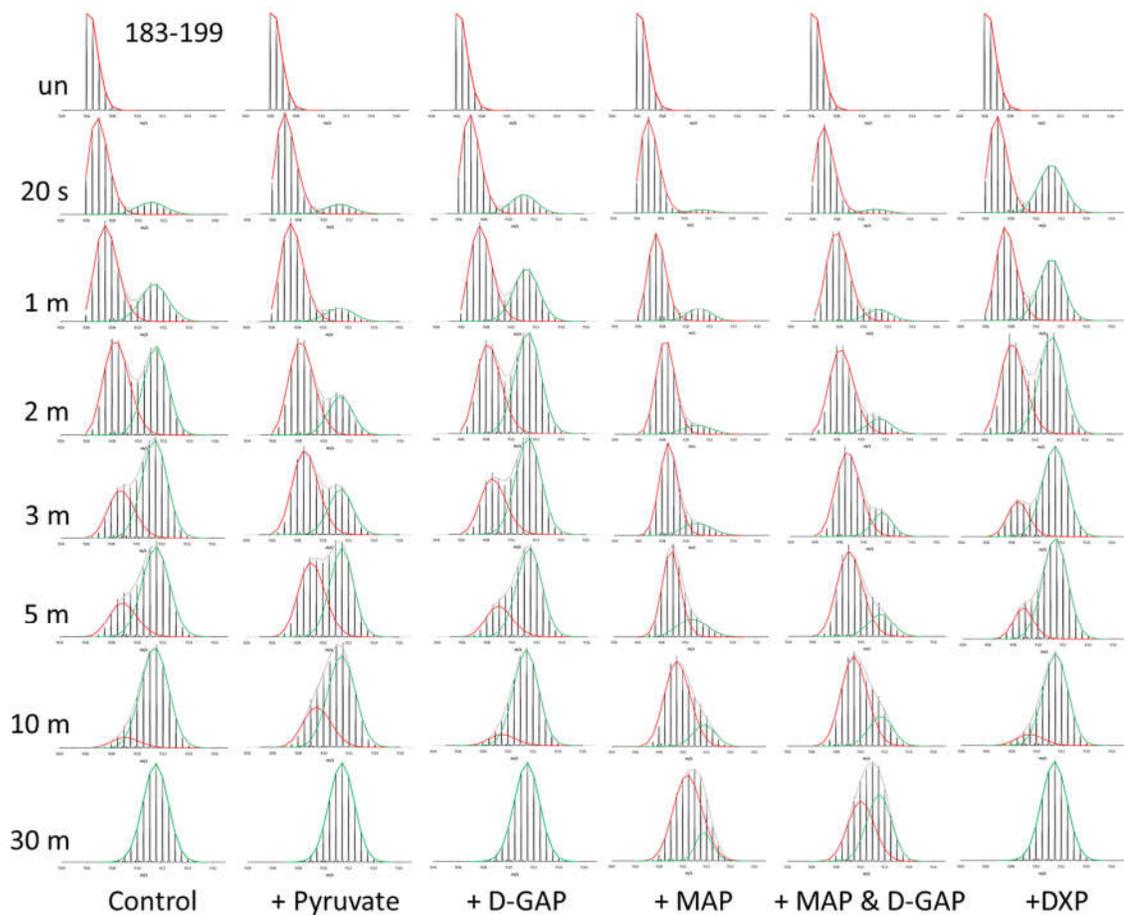
Three regions with EX1 kinetics (residues 42-58, 183-199, and 279-298) displayed a much stronger deuterium uptake difference than the remaining peptides, which were not included in this plot. The peptides displaying EX1 kinetics with difference plots in Figure 5.2 C-E. A) DXPS incubated with DXP; B) DXPS incubated with D-GAP; C) DXPS incubated with MAP and D-GAP; D) DXPS incubated with MAP; E) DXPS incubated with pyruvate.



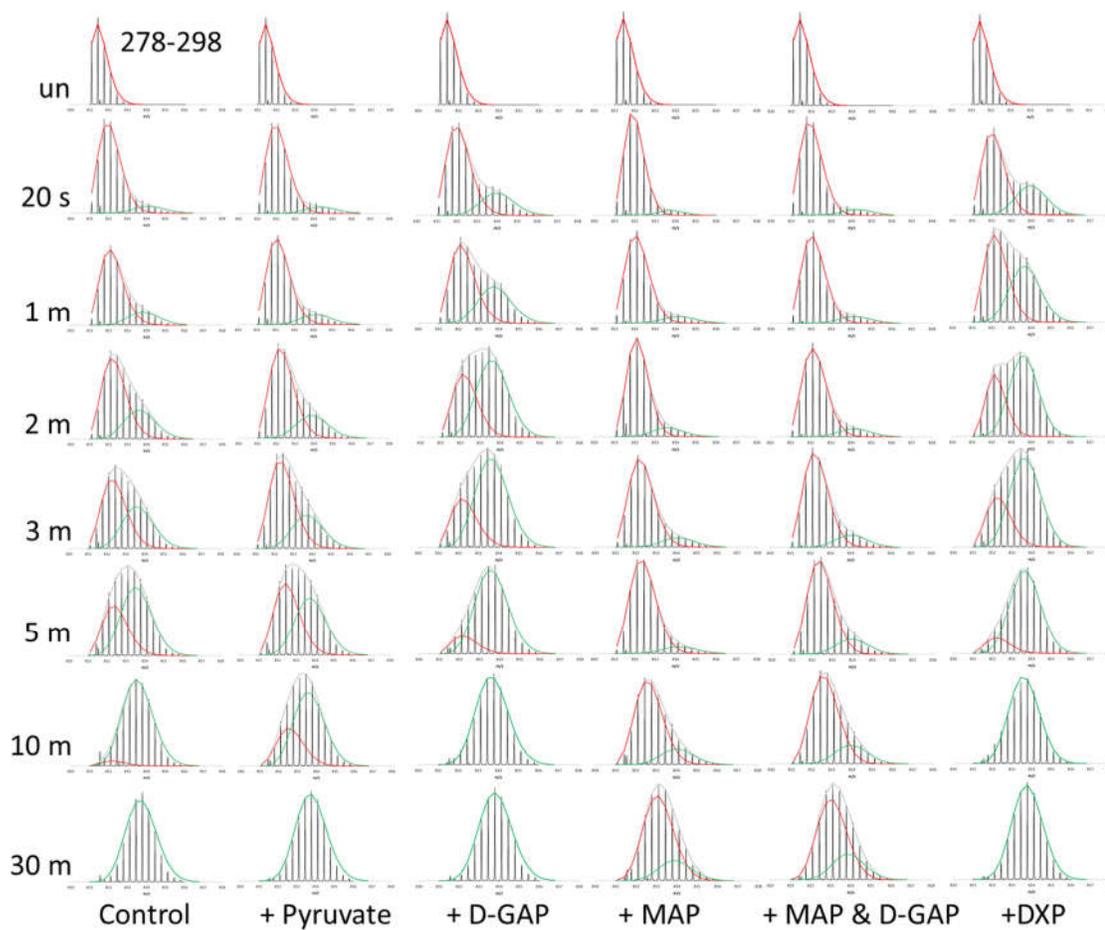
**Figure 5.6 Complete view of HDX-MS spectra of peptide 42-56 in its ThDP-bound state and in other ligand induced states.**



**Figure 5.7 Complete view of HDX-MS spectra of peptide 51-58 in its ThDP-bound state and in other ligand-bound states.**



**Figure 5.8 Complete view of HDX-MS spectra of peptide 183-199 in its free state and in other ligand-induced states.**



**Figure 5.9 Complete view of HDX-MS spectra of peptide 278-298 in its free state and in other ligand-induced states.**

**Table 5.1 Peptides resulting from pepsin digestion chosen for HDX-MS study**

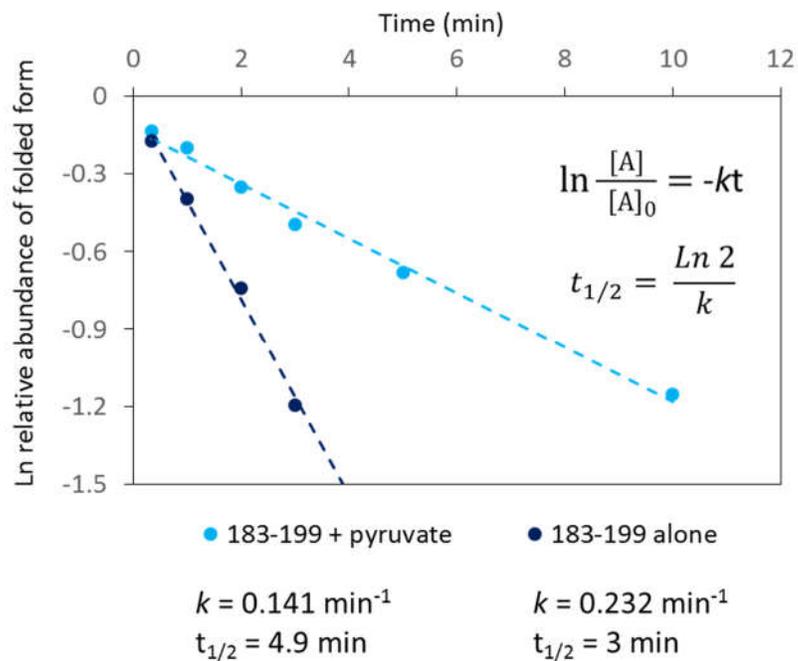
Theroretical Mass	Observed Mass	Error (ppm)	Start	End	Sequence
1154.6088	1154.6092	-0.4	2	11	SFDIAKYPTL
862.4142	862.4153	-1.2	12	19	ALVDSTQE
975.4998	975.4993	0.5	12	20	ALVDSTQEL
1406.9104	1406.9093	0.7	20	31	LRLLPKESLPKL
948.5634	948.5625	1.0	35	41	LRRYLLD
650.3108	650.3104	0.6	41	46	DSVSRS
1449.7081	1449.7081	0.0	42	56	SVSRSSGHFASGLGT
733.3730	733.3727	0.5	51	58	ASGLGTVE
1780.8903	1780.8905	-0.1	59	73	LTVLHYVYNTPFQDQ
1396.6553	1396.6532	1.5	64	74	HYVYNTPFQDL
4636.4686	4636.4549	3.0	75	114	IWDVGHQAYPHKILTGRDRKIGTIRQKGGGLHPPFWRGESE
1259.6217	1259.6226	-0.8	119	132	SVGHSSTSISAGIG
1744.9483	1744.9487	-0.2	133	148	IAVAAEKEGKNRRTVC
1004.5076	1004.5081	-0.5	149	159	VIGDGAITAGM
1603.6991	1603.6992	0.0	161	174	FEAMNHAGDIRPDM
929.4933	929.4938	-0.5	175	182	LVILNDNE
1810.9137	1810.9116	1.1	183	199	MSISENVGALNNHLAQL
1263.6801	1263.6804	-0.3	188	199	NVGALNNHLAQL
951.5009	951.5007	0.2	188	196	NVGALNNHL
517.3343	517.3344	-0.3	200	204	LSGKL
2178.1937	2178.1917	0.9	205	224	YSSLREGGKVFSGVPPIKE
1727.9805	1727.9803	0.1	209	224	REGGKVFSGVPPIKE
2021.1577	2021.1576	0.1	225	242	LLKRTEEHKGMVVPGTL
741.3450	741.3454	-0.5	243	248	FEELGF
1468.7439	1468.7431	0.6	249	262	NYIGPVDGHDVLGL
1761.9675	1761.9680	-0.3	263	277	ITTLKNMRDLKGPQF
2431.2815	2431.2802	0.5	278	298	LHIMTKKGRGYEPAEKDPITF
3277.6338	3277.6350	-0.4	299	329	HAVPKFDPSSGCLPKSSGGLPSYSKIFGDWL
1205.6189	1205.6194	-0.4	329	339	LCETAAKDNKL
1579.7267	1579.7277	-0.7	340	354	MAITPAMREGSGMVE
1500.7599	1500.7594	0.3	356	367	SRKFPDRYFDVA
1015.5211	1015.5207	0.4	368	376	IAEQHAVTF
1413.8466	1413.8464	0.1	377	391	AAGLAIGGYKPIVAI
988.5826	988.5826	0.0	381	390	AIGGYKPIVA
630.3133	630.3134	0.0	391	395	IYSTF
1527.7915	1527.7914	0.0	396	408	LQRAYDQVLHDVA
810.5458	810.5448	1.3	409	415	IQKLPVL
2031.0030	2031.0043	-0.6	416	435	FAIDRAGIVGADGQTHQGAF
610.3085	610.3083	0.3	436	440	DLSYL
861.4329	861.4321	0.9	440	446	LRCIPEM
1134.4968	1134.4983	-1.4	447	456	VIMTPSDENE
650.3108	650.3113	-0.7	457	461	CRQML
1344.5491	1344.5491	0.0	462	473	YTGYYNDGPSA
1187.6657	1187.6644	1.1	474	484	VRYPGRNAVGV
1975.0549	1975.0569	-1.0	507	525	AILNFGTLMPEAAKVAESL
1815.9510	1815.9521	-0.6	512	529	GTLMPPEAAKVAESLNATL
1187.6783	1187.6783	0.0	533	542	RFVKPLDEAL
1029.5034	1029.5034	0.0	546	555	MAASHEALVT
1745.8379	1745.8374	0.2	556	573	VEENAIMGGAGSGVNEVL
1147.6769	1147.6768	0.1	574	583	MAHRKPVPVL
775.3978	775.3985	-0.9	584	590	NIGLPDF
1179.5340	1179.5351	-0.9	591	600	FIPQGTQEEM
658.3885	658.3883	0.3	601	606	RAELGL
1573.8421	1573.8407	0.9	605	619	GLDAAGMEAKIKAWL

Peptide	42-56			51-58		
	k (min <sup>-1</sup> )	t1/2(min)	slowdown factor	k (min <sup>-1</sup> )	t1/2(min)	slowdown factor
Free state	0.116	6		0.046	15.2	
With Pyruvate	0.08	8.7	1.5	0.0279	24.8	1.6
With MAP	0.003	223.6	37.3	0.0136	51.0	3.4
With MAP+GAP	0.008	85.5	14.2	0.0143	48.5	3.2
With GAP	0.147	4.7	0.8	0.0694	10.0	0.7
With DXP	0.146	4.7	0.8	0.0533	13.0	0.9

Peptide	183-199			278-298		
	k (min <sup>-1</sup> )	t1/2(min)	slowdown factor	k (min <sup>-1</sup> )	t1/2(min)	slowdown factor
Free state	0.232	3		0.19	3.7	
With Pyruvate	0.141	4.9	1.6	0.102	6.8	1.8
With MAP	0.014	51.3	17.1	0.006	117.5	31.8
With MAP+GAP	0.027	25.3	8.4	0.019	36.9	10
With GAP	0.255	2.7	0.9	0.348	2	0.5
With DXP	0.304	2.3	0.8	0.32	2	0.5

**Table 5.2 Rate constant of unfolding and the unfolding half-life for the four peptides displaying EX1 kinetics**



**Figure 5.10 Example of calculation of half-life of unfolding**

The slope of the line fit to the data yields the first order rate constant, and half-life of unfolding was then calculated with the equations shown in this figure.

## References

1. Morris, F., Vierling, R.J., Boucher, L., Bosch, J., and Freel Meyers, C. L. (2013) DXP synthase-catalyzed C-N bond formatin: Nitroso substrate specificity studies guide selective inhibitor design. *ChemBioChem*. *14*, 1309-1315.
2. Smith, J.M., Warrington, N.V., Vierling, R.J., Kuhn, M.L., Anderson, W.F., Koppisch, A.T., and Freel Meyers, C.L. (2013) Targeting DXP synthase in human pathogens: enzyme inhibition and antimicrobial activity of butylacetylphosphonate. *J. Antibiotics* *67*, 77-83.
3. Xiang S, Usunow G, Lange G, Busch M, Tong L. Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. (2007) *J Biol Chem*. *282*, 2676-82
4. Bartee, D., Morris, F., Al-Khouja, A., and Freel Meyers, C.L. (2015) Hydroxybenzaldoximes Are D-GAP-Competitive Inhibitors of E. coli 1-Deoxy-D-Xylulose-5-Phosphate Synthase. *Chembiochem*. *16*, 1771-1781.
5. Murkin, A.S., Manning, K.A., Kholodar, S.A. (2014) Mechanism and inhibition of 1-deoxy-D-xylulose-5-phosphate reductoisomerase. *Bioorg Chem*. *57*, 171-185.
6. Banerjee, A., Wu, Y., Banerjee, R., Li, Y., Yan, H., and Sharkey, T.D. (2013) Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *J. Biol. Chem*. *288*, 16926-16936.
7. Smith, J.M., Vierling, R.J., and Meyers, C.F. (2012) Selective inhibition of E. coli 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *Medchemcomm*. *3*, 65-67.
8. Guttman, M., Weis, D.D., Engen, J.R., and Lee, K.K. (2012) Analysis of overlapped and noisy Hydrogen/Deuterium exchange data. *J. Amer. Soc. Mass Spectrom*. *24*, 1906-1912.
9. Woodward, C., Simon, I., and Tuchsén, E. (1982) Hydrogen exchange and the dynamic structure of proteins. *Mol. Cell. Biochem*. *48*, 135-160.
10. Englander, S.W., and Kallenbach, N.R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys*. *16*, 521-655.
11. Chance, M. (2008) *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics*. Hoboken, New Jersey: John Wiley & Sons, Inc.

12. Zhang, Q., Chen, J., Kuwajima, K., Zhang, H.M., Xian, F., Young, N.L., and Marshall, A.G. (2013) Nucleotide-induced conformational changes of tetradecameric GroEL mapped by H/D exchange monitored by FT-ICR mass spectrometry. *Sci. Rep.* *3*, 1247-1253.
13. Zheng, J., Yong, H.Y., Panutdaporn, N., Liu, C., Tang, K., and Luo, D. (2015) High-resolution HDX-MS reveals distinct mechanisms of RNA recognition and activation by RIG-I and MDA5. *Nucleic Acids Res.* *43*, 1216-1230.
14. Engen, J.R., Wales, T.E., Chen, S., Marzluff, E.M., Hassell, K.M., Weis, D.D., and Smithgall, T.E. (2013) Partial cooperative unfolding in proteins as observed by hydrogen exchange mass spectrometry. *Int. Rev. Phys. Chem.* *32*, 96-127
15. Patel, H., Nemeria, N.S., Brammer, L.A., Freel Meyers, C.L., and Jordan, F. (2012) Observation of thiamin-bound intermediates and microscopic rate constants for their interconversion on 1-deoxy-D-xylulose 5-phosphate synthase: 600-fold rate acceleration of pyruvate decarboxylation by D-glyceraldehyde-3-phosphate. *J. Am. Chem. Soc.* *134*, 18374-18379.
16. Chandrasekhar, K., Wang, J., Arjunan, P., Sax, M., Park, Y.H., Nemeria, N.S., Kumaran, S., Song, J., Jordan, F., and Furey, W. (2013) Insight to the interaction of the dihydrolipoamide acetyltransferase (E2) core with the peripheral components in the Escherichia coli pyruvate dehydrogenase complex via multifaceted structural approaches *J. Biol. Chem.* *288*, 15402-15417.
17. Wang, J., Kumaran, S., Zhou, J., Nemeria, N.S., Tao, H., Kakalis, L., Park, Y.H., Birkaya, B., Patel, M.S., and Jordan, F. (2015) Elucidation of the interaction loci of the human pyruvate dehydrogenase complex E2·E3BP core with pyruvate dehydrogenase kinase 1 and kinase 2 by H/D exchange mass spectrometry and nuclear magnetic resonance. *Biochemistry* *54*, 69-82.

## Appendix A. Protein purification

### A.1 Human E1o Purification

This method was used to purify E1o, E3, and E2o Core-domain.

#### Sample preparation for Ni column

Dissolve cells in 50-60 ml of sonication buffer with the glass rod, add 1 cocktail protease inhibitor tablet. Add lysozyme to 1 mg/ml and incubate on ice for 20 mins. Add 1000 units of DNase and Nuclease and incubate on ice for another 20 to 60 mins. Sonication for 6 min, 10s pulse “on” and 30s pulse “off”. Centrifugation for 30 min at  $23710 \times g$  at 4 °C three times to clarify the lysate.

#### Column separation

Ni column should always works under 0.35 mbar. For 6 ml volume Ni column, pressure is about 0.2 mbar at 1 ml/min flow rate, and pressure is about 2.6 mbar at 2 ml/min flow rate.

Column: Ni column 6 ml bed volume

Media: Ni Sepharose High Performance

Sample: Histidine-tagged human E1o protein, E. coli extract

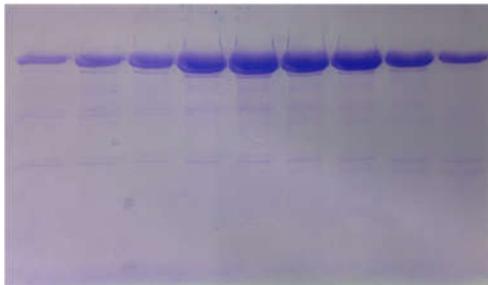
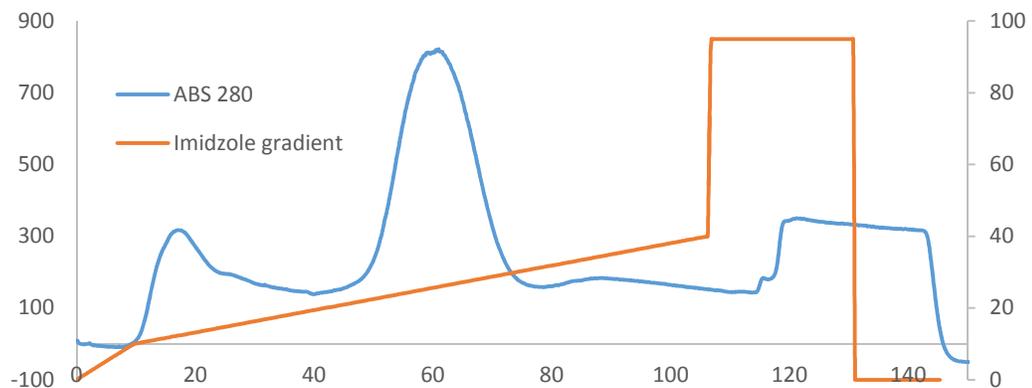
Sample volume: 50 ml (clarified extract), flited through 0.45 µm filter

Binding buffer: 50 mM potassium phosphate, 500 mM NaCl, 30 mM imidazole, 2 mM MgCl<sub>2</sub>, 0.5 mM ThDP, pH 7.4

Elution buffer: 50 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, 2 mM MgCl<sub>2</sub>, 0.5 mM ThDP, pH 7.4

Gradient: The first step, 2CV linear gradient of 30 to 77 mM imidazole, the second step, 20 CV linear gradient of 77 to 218 mM imidazole followed by a push with 500 mM imidazole

Result: 40 mg human E1o protein can be purified from 4 L LB culture media



**Figure A.1 Human E1o fractions form 52 to 70 ml**

Sample storage

The eluted protein sample was overnight dialyzed, and concentrated for -80 °C storage.

Sonication buffer

50 mM  $\text{KH}_2\text{PO}_4$  (pH 7.5), 0.5 M NaCl, 30 mM Imidazole, 5 mM  $\text{MgCl}_2$ , 2 mM ThDP

Dialysis buffer

50 mM  $\text{KH}_2\text{PO}_4$  (pH 7.5), 0.2 M NaCl, 2 mM  $\text{MgCl}_2$ , 0.2 mM ThDP

## A.2 Human E2o N-didomain Purification

### Sample preparation for Ni column

Dissolve cells in sonication buffer, final volume is about 50 mL, add 1 cocktail protease inhibitor tablet or add PMSF to 1mM final concentration. Add lysozyme to 1 mg/ml and incubate on ice for 15 mins. Add 1000 units of DNase and incubate on ice for another 30 mins. Sonication 10 mins, 15s pulse “on” and 30s pulse “off”. Centrifugation for 30 mins at  $23710 \times g$  at 4 C three times to clarify the lysate. Add Streptomyces sulfate 0.8g / 100 mL, centrifugation for 30 mins at  $23710 \times g$  at 4 C.

### Column separation

Column: Ni column 5 ml bed volume

Media: Ni Sepharose fast flow

Sample: Histidine-tagged human E2o didomain protein, E. coli extract

Sample volume: 50 ml (clarified extract)

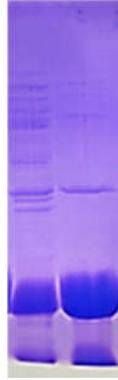
Washing buffer 1: 50 mM potassium phosphate, 300 mM NaCl, 0.1% Triton X100, pH 7.5

Washing buffer 2: 50 mM potassium phosphate, 300 mM NaCl, 0.1% Triton X100, 20 mM imidazole, pH 7.5

Elution buffer: 50 mM sodium phosphate, 300 mM NaCl, 200 mM imidazole, Triton X100, pH 7.5

Load sample to Ni column, then wash the column with 500 mL washing buffer 1, 3 to 4 CV washing buffer 2, and elute with Elution buffer.

Result: 30 mg human E2o didomain E194 protein can be purified from 3 L LB culture media



**Figure A.2 Purified Human E2o N-didomain**

Left lane protein eluted from washing buffer 2; Right lane purified N-didomain

Sample storage

The eluted protein sample was overnight dialyzed, and concentrated for -80 °C storage.

Sonication buffer

50 mM  $\text{KH}_2\text{PO}_4$  (pH 7.5), 0.3 M NaCl, 5 mM  $\text{MgCl}_2$ , 0.5% (v/v) Triton X100

Dialysis buffer

50 mM  $\text{KH}_2\text{PO}_4$  (pH 7.5), 0.2 M NaCl

### A.3 Protocol for purification of human E2o protein

#### Sample preparation for PEG precipitation

Dissolve cells in 50-60 ml of sonication buffer with the glass rod, add 1 cocktail protease inhibitor tablet. Add lysozyme to 0.6 mg/ml and incubate on ice for 20 mins. Add 1000 units of DNase and Nuclease and incubate on ice for another 20 to 60 mins. Sonication 6 mins, 10s pulse “on” and 30s pulse “off”. Centrifugation for 30 mins at  $23710 \times g$ , at 4 C twice to clarify the lysate.

#### PEG precipitation

Add PEG 8000 (50% w/v) solution dropwise to the clarified lysate to 1% (v/v). The precipitated protein was removed by centrifugation at  $18459 \times g$ , 20 mins. Add additional 5 % (v/v) of PEG 8000 solution dropwise to the clarified supernatant and incubate for 15 mins. This time the precipitate was collected at  $18459 \times g$ , 20 mins. The collected precipitate was dissolved in approximately 10 ml of buffer A, and clarified by centrifugation. The supernatant was load to a Sephacryl S 300HR column equilibrated with buffer A. The cloudy fractions (4 ml) (*do not collect the first two cloudy fractions because these might contain some of aggregated proteins*) were collected, combined and mixed with polyethylenimine (Mw 60,000) to a final level of 0.05% and the precipitate was removed by centrifugation at  $23710 \times g$ , 30 mins.

#### Ultracentrifugation

1. The human E2o was pelleted by two step ultracentrifugation

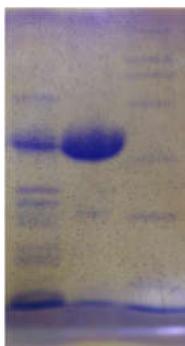
2. Use  $100,000 \times g$  for 4 hours at  $4\text{ }^{\circ}\text{C}$  to pellet the impurities and small amount of E2o. The result supernatant was combined and concentrated to about 8 ml with a 50 K concentration unit.
3. Use  $250,000 \times g$  for 4 hours at  $4\text{ }^{\circ}\text{C}$  to pellet E2o form 8 ml concentrated solution. The pellets were resuspended in a small volume of buffer containing 0.4 M potassium chloride and left overnight.
4. The suspension was clarified by centrifugation and stored at  $-80\text{ }^{\circ}\text{C}$ .

#### Sonication buffer

50 mM potassium phosphate pH 7.5, 0.5 mM EDTA, 150 mM potassium chloride, 1 mM  $\beta$ -mercaptoethanol, 0.1% pluronics-F68

#### Buffer A

50 mM potassium phosphate (pH 7.5), 0.5 mM EDTA, 200 mM potassium chloride, 1 mM  $\beta$ -mercaptoethanol, 0.1% pluronics-F68



**Figure A.3 Purified Human E2o**

Right lane: protein standard ladder, Middle lane: purified human E2o ( $100,000 \times g$ ), Left lane: pellets from first ultracentrifugation ( $250,000 \times g$ )

## A.4 Protein Sequences

Human Elo

```

      10      20      30      40      50      60
MSAPVAAEPF LSGTSSNYVE EMYCAWLENP KSVHKSWDIF FRNTNAGAPP GTAYQSPLPL

      70      80      90     100     110     120
SRGSLAAVAH AQSLVEAQPN VDKLVEDHLA VQSLIRAYQI RGHHVAQLDP LGILDADLDS

     130     140     150     160     170     180
SVPADIISST DKLGFYGLDE SDLDKVFHLP TTTFIGGQES ALPLREIIRR LEMAYCQHIG

     190     200     210     220     230     240
VEFMFINDLE QCQWIRQKFE TPGIMQFTNE EKRTLLARLV RSTRFEEFLQ RKWSSEKRFQ

     250     260     270     280     290     300
LEGCEVLIPA LKTIIDKSSE NGVDYVIMGM PHRGRLNVLA NVIRKELEQI FCQFDSKLEA

     310     320     330     340     350     360
ADEGSGDVKY HLGMYHRRIN RVTDRNITLS LVANPSHLEA ADPVVMGKTK AEQFYCGDTE

     370     380     390     400     410     420
GKKVMSILLH GDAAFAGQGI VYETFHLSLSD PSYTTHTGTVH VVVNNQIGFT TDPRMARSSP

     430     440     450     460     470     480
YPTDVARVVN APIFHVNSDD PEAVMYVCKV AAERSTFHK DVVVDLVCYR RNGHNEMDEP

     490     500     510     520     530     540
MFTQPLMYKQ IRKQKPVLQK YAELLVSQGV VNQPEYEEEI SKYDKICEEA FARSKDEKIL

     550     560     570     580     590     600
HIKHWLDSPW PGFFTLDGQP RSMSCPSTGL TEDILTHIGN VASSVPVENF TIHGGLSRIL

     610     620     630     640     650     660
KTRGEMVKNR TVDWALAEYM AFGSLLKEGI HIRLSGQDVE RGTFSHRHHV LHDQNVDKRT

     670     680     690     700     710     720
CIPMNHLWPN QAPYTVCNSS LSEYGVLGFE LGFAMASPNA LVLWEAQFGD FHNTAQCIID

     730     740     750     760     770     780
QFICPGQAKW VRQNGIVLLL PHGMEGMGPE HSSARPERFL QMCNDDPDVL PDLKEANFDI

     790     800     810     820     830     840
NQLYDCNWVV VNCSTPGNFF HVLRRQILLP FRKPLIIFTP KSLLRHPEAR SSFDEMLPGT

     850     860     870     880     890     900
HFQRVIPEDG PAAQNPENVK RLLFCTGKVY YDLTRERKAR DMVGQVAITR IEQLSPFPFD

     910     920     930     940     950     960
LLLKEVQKYP NAELAWCQEE HKNQGYDYV KPRLRTTISR AKPVVYAGRQ PAAAPATGNK

     970     980     990
KTHLTELQRL LDTAFDLDFV KNFSLEHHHH HH

```

## Human E2o

10            20            30            40            50            60  
 GSSHHHHHS SGLVPRGSHM DDLVTVKTPA FAESVTEGDV RWEKAVGDTV AEDEVVCEIE  
  
70            80            90            100            110            120  
 TDKTSVQVPS PANGVIEALL VPDGGKVEGG TPLFTLRKTG AAPAKAKPAE APAAAAPKAE  
  
130            140            150            160            170            180  
 PTAAAVPPPA APIPTQMPPV PSPSQPPSGK PVSAVKPTVA PPLAEAGAGK GLRSEHREKM  
  
190            200            210            220            230            240  
 NRMQRARIAQR LKEAQNTCAM LTTFNEIDMS NIQEMRARHK EAFLKKNLKG LGFMSAFVKA  
  
250            260            270            280            290            300  
 SAFALQEQPV VNAVIDDTTK EVVYRDYIDI SVAVATPRGL VVPVIRNVEA MNFADIERTI  
  
310            320            330            340            350            360  
 TELGEKARKN ELAIEDMDGG TFTISNGGVF GSLFGTPIIN PPQSAILGMH GIFDRPVAIG  
  
370            380            390            400  
 GKVEVRPMMY VALTYDHRLI DGREAVTFLR KIKAAVEDPR VLLLDL

## Human E3

10            20            30            40            50            60  
 RGSHHHHHG SADQPIDADV TVIGSGPGGY VAAIKAAQLG FKTVCIEKNE TLGGTCLNVG  
  
70            80            90            100            110            120  
 CIPSKALLNN SHYYHMAHGK DFASRGIEMS EVRLNLDKMM EQKSTAVKAL TGGIAHLFKQ  
  
130            140            150            160            170            180  
 NKVVHVNGYG KITGKNQVTA TKADGGTQVI DTKNILIATG SEVTPFPGIT IDEDTIVSST  
  
190            200            210            220            230            240  
 GALSLKKVPE KVVVIGAGVI GVELGSVWQR LGADVTAVEF LGHVGGVGID MEISKNFQRI  
  
250            260            270            280            290            300  
 LQKQGFKFKL NTKVTGATKK SDGKIDVSIE AASGGKAEVI TCDVLLVCIG RRPFTKNLGL  
  
310            320            330            340            350            360  
 EELGIELDPR GRIPVNTRFQ TKIPNIYAIG DVVAGPMLAH KAEDGEIICV EGMAGGAVHI  
  
370            380            390            400            410            420  
 DYNCPVSVIY THPEVAWVGK SEEQLKEEGI EYKVGKFPFA ANSRAKTNAD TDGMVKILGQ  
  
430            440            450            460            470            480  
 KSTDRVLGAH ILGPGAGEMV NEAALALEYG ASCEDIARVC HAHPTLSEAF REANLAASFG

KSINF

## E2o N-didomain

10            20            30            40            50            60  
 MGSSHHHHHH SSGLVPRGSH MDDLVTVKTP AFAESVTEGD VRWEKAVGDT VAEDEVVCEI  
  
70            80            90            100            110            120  
 ETDKTSVQVP SPANGVIEAL LVPDGGKVEG GTPLFLLRKT GAAPAKAKPA EAPAAAAPKA  
  
130            140            150            160            170            180  
 EPTAAAVPPP AAPIPTQMPP VSPSQPPSG KPVS AVKPTV APPLAEAGAG KGLRSEHREK  
  
190  
 MNRMRQRIAQ RLKE

## E2o Core domain

10            20            30            40            50            60  
 MAEAGAGKGL RSEHREKMNR MRQRIAQRLK EAQNTCAML T FNEIDMSNI QEMRARHKEA  
  
70            80            90            100            110            120  
 FLKKHNLKLG FMSAFVKASA FALQEQPVVN AVIDDTTKEV VYRDYIDISV AVATPRGLVV  
  
130            140            150            160            170            180  
 PVIRNVEAMN FADIERTITE LG EKARKNEL AIEDMDGGTF TISNGGVFGS LFGTPIINPP  
  
190            200            210            220            230            240  
 QSAILGMHGI FDRPVAIGGK VEV RPMMYVA LTYDHRLIDG REAVTFLRKI KAAVEDPRVL  
  
250            260  
 LLDLEASENL YFQGLEHHHH HH

## E. coli DXP synthase

10            20            30            40            50            60  
 MSFDIAKYPT LALVDSTQEL RLLPKESLPK LCDELRRYLL DSVSRSSGHF ASGLGTVELT

70            80            90            100           110           120  
 VALHYVYNTP FDQLIWDVGH QAYPHKILTG RRDKIGTIRQ KGGLHPFPWR GESEYDVLSV

130           140           150           160           170           180  
 GHSSTISISAG IGIAVAAEKE GKNRRTVCVI GDGAITAGMA FEAMNHAGDI RPDMLVILND

190           200           210           220           230           240  
 NEMSISENVG ALNNHLAQLL SGKLYSSLRE GGKKVFSGVP PIKELLKRTE EHIKGMVVP

250           260           270           280           290           300  
 TLFEELGFNY IGPVDGHDVL GLITTLKNMR DLKGPQFLHI MTKKGRGYEP AEKDPITFHA

310           320           330           340           350           360  
 VPKFDPSSGC LPKSSGGLPS YSKIFGDWLC ETAAKDNKLM AITPAMREGS GMVEFSRKFP

370           380           390           400           410           420  
 DRYFDVAIAE QHAVTFAAGL AIGGYKPIVA IYSTFLQRAY DQVLHDVAIQ KLPVLFAIDR

430           440           450           460           470           480  
 AGIVGADGQT HQGAFDLSYL RCIPEMVIMT PSDENECRQM LYTGYHYNDG PSAVRYPRGN

490           500           510           520           530           540  
 AVGVELTPLE KLPIGKGIVK RRGEKLAILN FGTLMPEAAK VAESLNATLV DMRFVKPLDE

550           560           570           580           590           600  
 ALILEMAASH EALVTVEENA IMGGAGSGVN EVLMAHRKPV PVLNIGLPDF FIPQGTQEEM

610           620  
 RAEGLDAAG MEAKIKAWLA

## List of Publications

### Publications at Rutgers University

1. Wang J, Kumaran S, Zhou J, Nemeria NS, Tao H, Kakalis L, Park YH, Birkaya B, Patel MS, and Jordan F. Elucidation of the interaction loci of the human pyruvate dehydrogenase complex E2·E3BP core with pyruvate dehydrogenase kinase 1 and kinase 2 by H/D exchange mass spectrometry and nuclear magnetic resonance. *Biochemistry*. 2015 54(1):69-82.
2. Nemeria NS, Ambrus A, Patel H, Gerfen G, Adam-Vizi V, Tretter L, Zhou J, Wang J, and Jordan F. Human 2-oxoglutarate dehydrogenase complex E1 component forms a thiamin-derived radical by aerobic oxidation of the enamine intermediate. *J Biol Chem*. 2014 289(43):29859-73.

### Publications at East China University of Science and Technology

1. Zhou J, Xu X, and Wang Y. Competitive immunoassay for clenbuterol using capillary electrophoresis with laser-induced fluorescence detection. *J Chromatogr B*. 2007 848(2):226-31.
2. Zhou J, Zhang Z, Xu X, Lu Y, and Wei D. Chiral separation of hesperidin by capillary electrophoresis. *Chinese Journal of Food Science* 2006 27(5):202-205
3. Zhou J, Xu X, and Ren Y. Separation and Determination of Atomoxetine Enantiomers by Capillary Electrophoresis. *The Fourth National Pharmaceutical Engineering Academic Conference* 2005 11:179-187. ISBN7-5628-1808-8/R·14