

INTERACTION OF THE COMPONENTS WITHIN PYRUVATE
DEHYDROGENASE COMPLEX, 2-OXOGLUTARATE DEHYDROGENASE
COMPLEX AND 1-DEOXY-D-XYLULOSE 5-PHOSPHATE SYNTHASE

By

LUYING YANG

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

May, 2018

©2018

Luying Yang

ALL RIGHTS RESERVED

ABSTRACT OF THE DISSERTATION

INTERACTION OF THE COMPONENTS WITHIN PYRUVATE DEHYDROGENASE COMPLEX, 2-OXOGLUTARATE DEHYDROGENASE COMPLEX AND 1-DEOXY-D-XYLULOSE 5-PHOSPHATE SYNTHASE

by LUYING YANG

Research Advisor: Professor Frank Jordan

Thiamin diphosphate (ThDP)-dependent enzymes [pyruvate dehydrogenase complex (PDHc), 2-oxoglutarate dehydrogenase complex (2-OGDHc) and 1-deoxy-D-xylulose 5-phosphate synthase (DXPS)] form a diverse class of proteins, which play critical roles in catalyzing a vast variety of metabolic reactions, especially the formation of carbon-carbon bonds. Many of their functions rely on complexes rather than on individual components. Understanding how proteins behave in isolation as well as how they recognize their binding partners is therefore critical for understanding the functions and regulation 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 for most experiments two methods were utilized: (a) External fluorophores introduced via site-specific covalent labeling to measure dissociation constants (K_d) to establish the strength of interaction; and (b) Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS) to study the interactions within PDHc, 2-OGDHc and DXPS.

First, we designed a fluorescence titration experiment to measure the dissociation constants between different components within PDHc, to quantitatively study the role that each domain plays in interaction with its regulatory enzyme PDKs. We also conducted a series of HDX-MS experiments to investigate the interaction pattern between the first component of PDHc E1p and PDKs in Chapter 3. In Chapter 4, we mainly focus on the binary complex interactions within 2-OGDHc. By using an external site-specifically introduced fluorophore tag, we measured the dissociation constants between different components and with FRET experiments, we also could identify principal interaction loci. In Chapter 5, we used the HDX-MS method for a study on the dynamics of DXPS. Three of all the peptic peptides derived from DXPS displayed the very unique EX1-type hydrogen/deuterium kinetics, which enabled us to understand how some of the DXPS domains respond to different components and identified two likely slowly interconverting conformation in the active center of the enzyme, a closed and an open one, suggesting mechanistic information about this bisubstrate ThDP enzyme. In the Appendix, we report some preliminary data related to the phosphorylation loop in the PDHc E1p component. It shows that peptides located on phosphorylation loop also display EX1-kinetics, which might be a critical property for the catalysis and regulation of thiamin-dependent enzymes.

All studies in this thesis are aimed at a better understanding of mechanisms in ThDP catalysis, and we hope that these fundamental research subjects could guide the development of future structure-based inhibitors and perhaps cure some diseases.

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. Richard Mendelsohn 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. Jieyu Zhou for the enormous help in project development and mass spectrometry work. 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 Chemistry Department.

Finally, I would like to thank my parents 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 for providing me encouragement and support.

TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION	ii
ACKNOWLEDGEMENTS	iii
DEDICATION	iv
TABLE OF CONTENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xii
LIST OF EQUATIONS	xvii
CHAPTER 1. Introduction.....	1
1.1 Components and organization of the mammalian PDHc	1
1.2 Regulation and diseases relate to PDKs	3
1.3 Inhibitors of PDKs	8
1.4 Objectives of this thesis	9
CHAPTER 2. Site specifically introduced external fluorophore to study protein-protein interactions within PDHc	15
2.1 Introduction	15
2.2 Site-specific labeling method to study protein-protein interactions within PDHc ..	17
2.3 Results and discussion of E2·E3BP-derived proteins interacting with PDKs	21
2.4 Results of doubly substituted L1L2S variants interacting with PDKs and E1p	30
2.5 Conclusions	37

CHAPTER 3. Using H/D exchange mass spectrometry to study the interaction patterns between E1p and four isoforms of PDKs	41
3.1 Introduction	41
3.2 Material and methods	44
3.2.1 Sample preparation for HDX-MS	44
3.2.2 LC-MS method and data processing	45
3.3 Results and Discussion	46
3.3.1 Overview of HDX patterns on E1p	46
3.3.2 Effects of the PDK isoforms on the E1p structure	49
3.3.3 Effects of PDKs on E1p phosphorylation	53
3.3.4 Effects of E1p on PDKs	56
3.3.5 Ligands effects on four isoforms of PDKs	61
CHAPTER 4 Study of protein-protein interactions within the human 2-oxoglutarate dehydrogenase complex using site-specifically introduced external fluorophores	79
4.1 Introduction	79
4.2 Material and methods	81
4.2.1 Labeling of lipoamide in the LDo with a dansyl group for fluorescence studies	81
4.2.2 Labeling E2o ¹⁻¹⁷³ and E2o ¹⁴⁴⁻³⁸⁶ with N-(1-pyrene)maleimide for fluorescence studies	82
4.2.3 Fluorescence spectroscopy titration studies	82

4.2.4 Measurement of E1o-specific activity	83
4.2.5 Fluorescence study of the binding of thiochrome diphosphate to E1o	83
4.2.6 FRET experiments	83
4.2.7 Labeling E2o ¹⁻¹⁷³ , E2o ¹⁴⁴⁻³⁸⁶ with Alexa Fluor® 350 C5 Maleimide dye for FRET studies	84
4.3 Results and Discussion	85
4.3.1 E1o-E2o interactions	85
4.3.2 E1o-E3 interactions	90
4.3.3 E2o-E3 interactions	95
4.4 Conclusions	99
CHAPTER 5 Conformational dynamics of 1-Deoxy-D-xylulose 5-phosphate synthase upon ligand binding revealed by H/D exchange mass spectrometry	
5.1 Introduction	104
5.2 Material and methods	107
5.2.1 Sample preparation for HDX	107
5.2.2 LC-MS method and data processing	108
5.3 Results and Discussion	109
5.3.1 Mechanisms of EX1 and EX2 kinetics	109
5.3.2 Overview of HDX patterns on DXPS	110
5.3.3 HDX-MS detects local structural dynamics of DXPS	112

5.3.4 DXPS binds and responds differently to substrates, substrate analogue and product	122
5.4 Conclusions	125
Appendix A	133
A.1. Development of a bifunctional cross-linker to cross link E1p with lipoylated E2p-derived proteins	133
A.2. ^{19}F NMR studies of loop dynamics in E1p.....	135
A.3. EPR studies of loop dynamics in E1p	140
A.4. HDX-MS studies of loop dynamics in E1p	142
Appendix B	153

LIST OF TABLES

Table 1.1 PDKs and associated pathological conditions	6
Table 2.1 K_d values for binding of the E2·E3BP-derived domains to PDK1-PDK4 as detected by fluorescence spectroscopy	27
Table 2.2 Peptides mass checked by LC-MS	32
Table 2.3 K_d values for binding of the L1L2S mutants to PDK1-PDK4 and E1p as detected by fluorescence spectroscopy	36
Table 3.1 Deuteration level change of selected peptides on addition of ATP to the four PDK isozymes based on sequence alignment	65
Table 3.2 Peptides that show significant deuteration level change in PDK1 on addition of pyruvate	66
Table 3.3 Peptides that show a significant deuteration level change in PDK3 on addition of pyruvate	68
Table 3.4 Peptides identified from pepsin digestion of the E1p α -subunit	70
Table 3.5 Peptides identified from pepsin digestion of the E1p β -subunit	71
Table 3.6 Peptides identified from pepsin digestion of PDK1 ^{rat}	72
Table 3.7 Peptides identified from pepsin digestion of PDK2 ^{rat}	73
Table 3.8 Peptides identified from pepsin digestion of PDK3 ^{human}	75
Table 3.9 Peptides identified from pepsin digestion of PDK4 ^{rat}	76
Table 4.1 Protein modification checked by FT-MS	98

Table 5.1 Rate constant of unfolding and the unfolding half-life for the four peptides displaying EX1 kinetics	124
Table 5.2 Peptides resulting from pepsin digestion chosen for HDX-MS study	127
Table A.2.1: Peptides after trypsin digestion for the detection of E1p modification	140

LIST OF FIGURES

Figure 1.1 Structure of pyruvate dehydrogenase complex	2
Figure 1.2 Catalytic cycle of the mammalian pyruvate dehydrogenase complex	3
Figure 1.3 Tissue distribution of four isoforms of PDKs	5
Figure 1.4 Structures of PDKs inhibitors	8
Figure 2.1 Domain structures of E2·E3BP- derived proteins	16
Figure 2.2 Using DANS-As to chemically modify LDs	19
Figure 2.3 An example of chemical modification monitored by ESI-MS	21
Figure 2.4 Fluorescence titration of DANS-As labeled L2S by PDK1	23
Figure 2.5 Fluorescence titration of DANS-As labeled L3S' by PDK2.....	24
Figure 2.6 Fluorescence binding curves for DANS-As-labeled E2 E3BP derived lipoyl domains on titration by PDK isoforms	25
Figure 2.7 Fluorescence titration of dapoxyl-labeled L3S' by PDK2	26
Figure 2.8 Fluorescence binding curves for dapoxyl-labeled L3S' on titration by PDK2 and PDK4.....	27
Figure 2.9 Fluorescence binding curves for DANS-As-labeled E2·E3BP derived lipoyl domains on titration by E1p	30
Figure 2.10 Domain structures of L1L2S doubly substituted variants	31
Figure 2.11 Fluorescence titration of DANS- labeled L1L2S (K46A/E209K) by four isoforms of PDKs	33

Figure 2.12 Fluorescence binding curves for DANS labeled L1L2S doubly substituted variant (K46A/E209K) on titration by PDK isoforms	34
Figure 2.13 Fluorescence titration of DANS- labeled L1L2S (K46A/E209K) by E1p	35
Figure 2.14 Fluorescence binding curves of DANS-As-L1L2S (K46A/E209K) titrated by E1p	36
Figure 3.1 Topology of PDK2	42
Figure 3.2 Experimental design	45
Figure 3.3 Deuterium incorporation pattern of E1p	47
Figure 3.4 Effect of complexation with PDK isoforms on deuterium exchange in E1p ..	49
Figure 3.5 Effects of four PDK isozymes on H/D exchange in E1p	51
Figure 3.6 Peptides of human E1p involved in conformational change upon complexation with PDKs	52
Figure 3.7 Phosphorylation sites on the E1p- α subunit	53
Figure 3.8 PO_4^{3-} incorporation ratios on peptide $^{258}\text{YRYHGHSMKSDPGVS}^{271}$ during 30 min reaction	54
Figure 3.9 PO_4^{3-} incorporation on peptide $^{192}\text{ICENNRYGMGTSVE}^{205}$ during 30 min reaction with PDK1	55
Figure 3.10 Crystal structure of four PDK isozymes	56
Figure 3.11 Deuteration level change in PDK1 on interaction with E1p	57
Figure 3.12 Deuteration level change in PDK2 on interaction with E1p	58

Figure 3.13 Residues of PDK2 displaying significant $\Delta\Delta D$ on binding to E1p in HDX-MS analysis	59
Figure 3.14 Deuteration level change in PDK3 on interaction with E1p	60
Figure 3.15 Deuteration level change in PDK4 on interaction with E1p	61
Figure 3.16 Difference plot showing the changes in deuterium incorporation	63
Figure 3.17 Crystal structure of human PDK2 containing ATP	64
Figure 3.18 Crystal structure of PDK2 in the presence of DCA	67
Figure 3.19 Peptides that show significant deuteration change on PDK1 when incubated with pyruvate and ATP	67
Figure 3.20 Peptic peptides from PDK3 that show significant deuteration change on incubation with pyruvate and ATP	69
Figure 4.1 Domain structure of E2o and E2o constructs	81
Figure 4.2 Labeling of E2o ¹⁻¹⁷³ and E2o ¹⁴⁴⁻³⁸⁶ with N-(1-pyrene)maleimide	82
Figure 4.3. Structure of Alexa Fluor® 350 C5 Maleimide dye	85
Figure 4.4 Fluorescence titration of DANS- labeled LDo by E1o	87
Figure 4.5 Fluorescence titration of Pyrene- labeled E2o ¹⁻¹⁷³ and Pyrene-labeled E2o ¹⁴⁴⁻³⁸⁶ by E1o	88
Figure 4.6 Fluorescence titration curves for DANS-As- labeled E2o ¹⁻⁹⁵ and E2o ¹⁻¹⁷³ and for pyrene labeled E2o ¹⁻¹⁷³ and E2o ¹⁴⁴⁻³⁸⁶ proteins by E1o	89
Figure 4.7 Structural formulae of thiamin and thiochrome diphosphate	90
Figure 4.8 Fluorescence titration of E1o by thiochrome diphosphate	92

Figure 4.9 A plot of excitation and emission spectrum of FAD and TCDP at a concentration of 10 μ M each	94
Figure 4.10 Measurements of energy transfer between E1o bound thiochrome diphosphate and E3 bound FAD	95
Figure 4.11 A plot of excitation and emission spectrum of FAD and Alexa350 with a concentration of 10 μ M	97
Figure 4.12 Measurements of energy transfer between $E2o^{1-173}$ -Alexa350 and E3 bound FAD	97
Figure 4.13 Summary of binary component interactions on the basis of fluorescence ..	100
Figure 5.1 Average deuterium uptake by key peptides of E. coli DXPS	111
Figure 5.2 Mass spectrum of peptides 183-199 at the 1-min HDX time point in the presence of GAP	112
Figure 5.3 DXPS deuterium uptake change over 30 min	114
Figure 5.4 Spatial proximity of all regions displaying EX1 kinetic behavior in the DXPS structure	115
Figure 5.5 Complete time course of HDX-MS spectra of peptides 183-199 in its ThDP-bound and ligand-bound states	117
Figure 5.6 Complete time course of HDX-MS spectra of peptides 42-56 in its ThDP-bound and ligand-bound states	118
Figure 5.7 Complete time course of HDX-MS spectra of peptides 51-58 in its ThDP-bound and ligand-bound states	119

Figure 5.8 Complete view of HDX-MS spectra of peptide 186-196 in its free state and in other ligand induced states	120
Figure 5.9 Complete view of HDX-MS spectra of peptide 278-298 in its free state and in other ligand induced states	121
Figure A.1.1: Synthetic procedure of BRAO.....	133
Figure A.1.2: LC-MS of cross-linked E1p-LD1 after pepsin digestion	135
Figure A.2.1: ^{31}P NMR results of ThDP bounded phos-S-E1p	137
Figure A.2.2: LC-MS result after trypsin digestion of modified E1p-MS 2, 3 sample ...	139
Figure A.3.1: Site-directed labelling of MTSL onto E1p- α subunit.....	141
Figure A.3.2: LC-MS peak of “YHGHSMSPGVSYR-MTSL” ($\text{C}_{77}\text{H}_{112}\text{N}_{22}\text{O}_{25}\text{S}_3\text{P}$) ..	141
Figure A.3.3: EPR spectrum of E1p-MTSL	142
Figure A.4.1: Phosphorylated peptide detected by LC-MS	144
Figure A.4.2: Structures of the Phosphorylation loops in Wild-type and Phospho-S1-E1p	146
Figure A.4.3: Complete view of HDX-MS spectra of peptide 192-205 of E1p- in its ThDP-bound state and in other ligand induced states	148
Figure A.4.4: Complete view of HDX-MS spectra of peptide 162-177 of E1p- in its ThDP-bound state and in other ligand induced states	149
Figure A.4.5: Peptides from E1p- that show bimodal distribution in HDX-MS	149
Figure A.4.6: Half-life calculation of the unfolded form of E1p- α 192-205	150

LIST OF EQUATIONS

Equation 1.1	6
Equation 2.1	20
Equation 4.1	83
Equation 4.2	83
Equation 4.3	91
Equation 4.4	91
Equation 4.5	92
Equation 5.1	109

CHAPTER 1. Introduction

The human pyruvate dehydrogenase complex (PDHc), which governs the overall carbohydrate oxidation, is a key enzyme at the entry of the citric acid cycle in every cell that contains mitochondria. It catalyzes the irreversible oxidative decarboxylation of pyruvate to produce acetyl-CoA, an important intermediate that participates in many biosynthetic pathways such as those of carbohydrate and lipid metabolism. PDHc is also responsible for regulating the rate of oxidative phosphorylation and subsequent ATP generation from carbohydrate derived carbon sources. During well fed state, PDHc is active and promotes glucose oxidation, while in the fasting state, PDHc activity would be turned off and the three-carbon compound —pyruvate could be preserved for sustaining gluconeogenesis. The activity of PDHc is not only important in glucose production but it is also related to the transition from glucose oxidation to fatty acid oxidation as proposed by the Randle cycle [1]. Randle and his colleagues conducted a series of experiments in cardiac and skeletal muscle, showing that increased fatty acid oxidation enhances the ratio of [acetyl-CoA]/[CoA] and [NADH]/[NAD⁺], both of which would inhibit PDHc activity by activating the pyruvate dehydrogenase kinases (PDKs).

1.1 Components and organization of the mammalian PDHc

The mammalian pyruvate dehydrogenase complex consists of three catalytic components: pyruvate dehydrogenase (E1p), dihydrolipoyl transacetylase (E2p) and dihydrolipoyl dehydrogenase (E3). It also contains a unique E3-binding protein (E3BP) and two

regulatory enzymes: pyruvate dehydrogenase kinases (PDK, four human isoforms) and pyruvate dehydrogenase phosphatase (PDP, two human isoforms). The complex assembles together using 48 copies of E2p and 12 copies of E3BP forming a E2p.E3BP subcomplex core, to which the E1p component binds to E2p and the E3 component binds to E3BP.

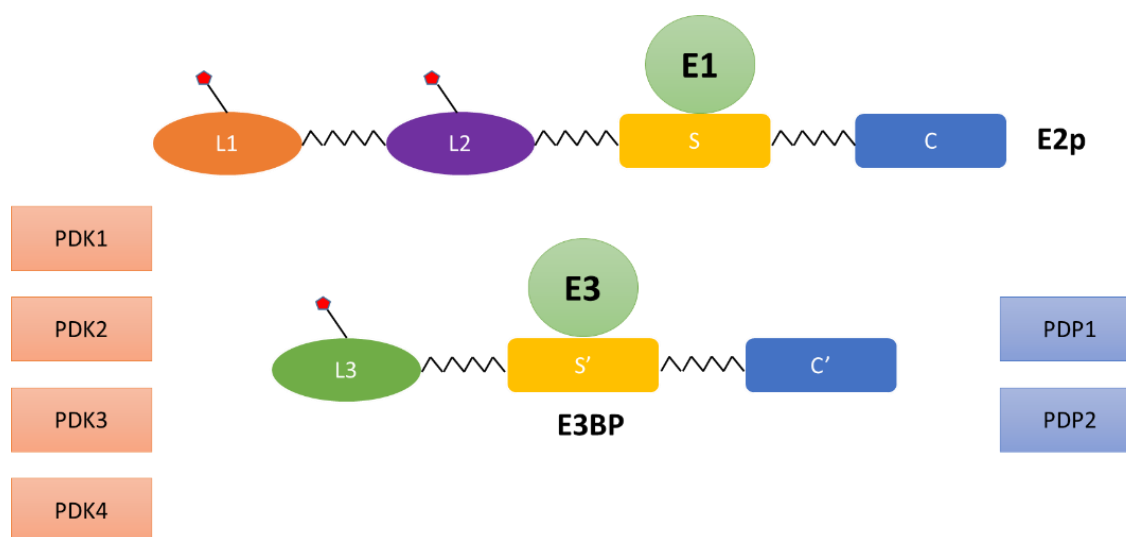


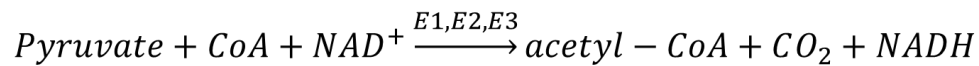
Figure 1.1: Structure of pyruvate dehydrogenase complex

E2p has a multidomain structure: an outer lipoyl domain (L1), an inner lipoyl domain (L2), a subunit-binding domain (S) and an inner core domain (core). The structure of E3BP is quite similar to that of E2p, while instead of two lipoyl domains, it only has one (Figure 2) [2], and a catalytically incompetent core domain.

The E1 component in PDHc catalyzes the decarboxylation of pyruvate with the auxiliary of thiamin diphosphate (ThDP), the hydroxyethyl group derived from pyruvate is then transferred to lipoamide derived from E2. The hydroxyethyl carbanion is oxidized to an acetyl group as the lipoamide disulfide is reduced. E2 also catalyzes the transfer of acetyl

groups to form acetyl-CoA. The E3 component then reoxidizes the dihydrolipoamide by its tightly-bound FAD coenzyme. Finally, FADH₂ is reoxidized to FAD by NAD⁺, yielding NADH [3].

The overall reaction is shown below.



equation 1.1

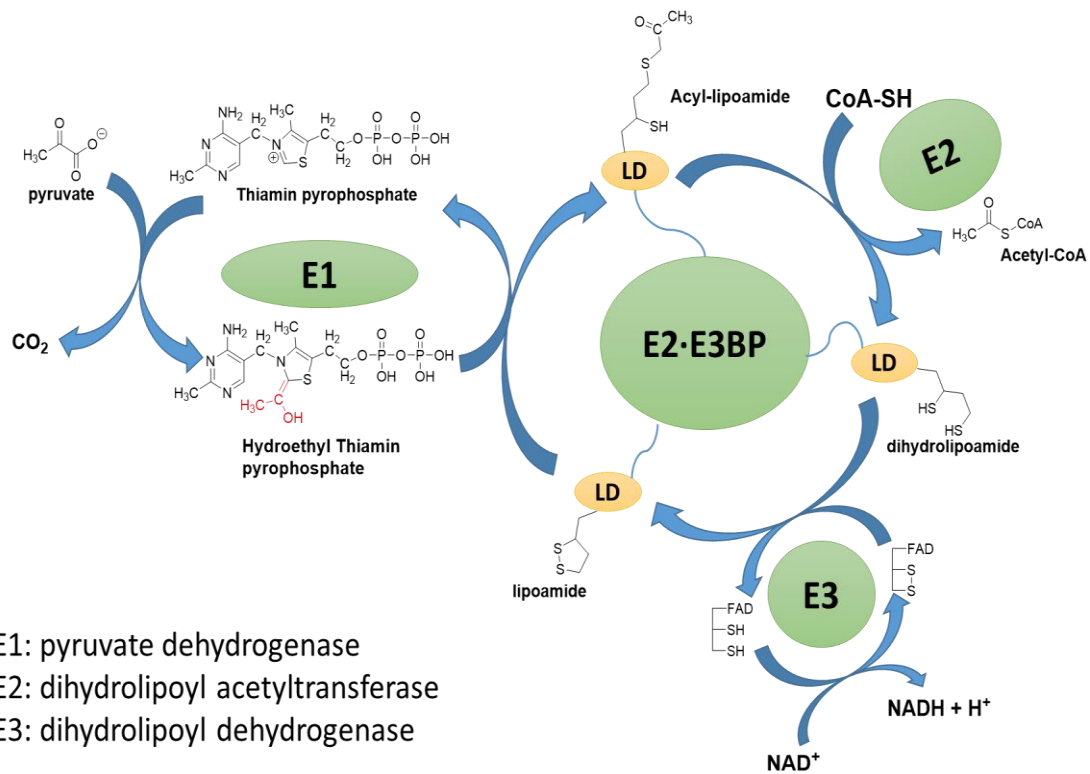


Figure 1.2: Catalytic cycle of the mammalian pyruvate dehydrogenase complex

1.2 Regulation and diseases relate to PDKs

PDHc regulation is accomplished by the combined reactions of two phosphatases (PDP1 and PDP2) and four kinases (PDK1-4). Through phosphorylation and dephosphorylation on the specific sites located on the E1p α subunit, PDKs and PDPs can shut down or recover, respectively, the activities of PDHc. The multiple isoforms of the PDKs and the PDPs are distinguished by differences in tissue distribution, specific activities toward the phosphorylation sites, kinetic properties, and sensitivity to regulatory molecules [4-5]. Regulation of PDHc is achieved by site-specific phosphorylation of the α subunits of the heterotetrameric ($\alpha_2\beta_2$) E1p component at three different sites — Ser264 (site 1), Ser271 (site 2) and Ser203 (site 3). The three sites were phosphorylated *in vivo* to a different extent, with maximum phosphorylation of site 1 and less phosphorylation of sites 2 and 3. The four PDK isozymes have different activity and phosphorylation rates at each site. At site 1 from fastest to slowest, $\text{PDK2} > \text{PDK4} \approx \text{PDK1} > \text{PDK3}$; for site 2 it is, $\text{PDK3} > \text{PDK4} > \text{PDK2} > \text{PDK1}$. Only PDK1 can phosphorylate site 3 [3]. Dephosphorylation of phospho-E1 by PDP's recovers full PDHc activity.

The existence of three potential phosphorylation sites together with four isoforms of PDKs provides the possibility of delicate control towards metabolism, although the detailed mechanism is still unclear.

The structure of PDKs can be divided into two distinct domains, the N- and C-terminal domains. The N-terminal domain consists of eight α -helices, of which four form a bundle-like structure creating the core. The sequence in the N-terminal domain of the four isoforms of PDKs is poorly conserved. In contrast, the C-terminal domain containing the phosphoryl transfer catalytic site is highly conserved among the four PDKs. They

share four conserved motifs, N-box (Glu-X-X-Lys-Asn-X-X-X-Ala); G1-box (Asp-X-Gly-X-Gly); G2-box (Gly-X-Gly-X-Gly); and G3-box (Gly-X-Gly-Thr) [4], which form a unique ATP-binding fold. This fold includes a common structural element known as the “ATP lid,” whose conformational change is coupled to both ATP hydrolysis and protein-protein interactions [5].

The tissue distribution of the four isoforms is also different. PDK1 is expressed predominantly in the heart [6] and pancreatic islets [9], while PDK2 has a relatively wide distribution and is present in liver, heart, skeletal muscle, etc. with lower amounts in spleen and lung [10]. PDK3 mainly exists in testis, kidney and brain [6], and PDK4 is expressed in heart, skeletal muscle and liver [11-12].

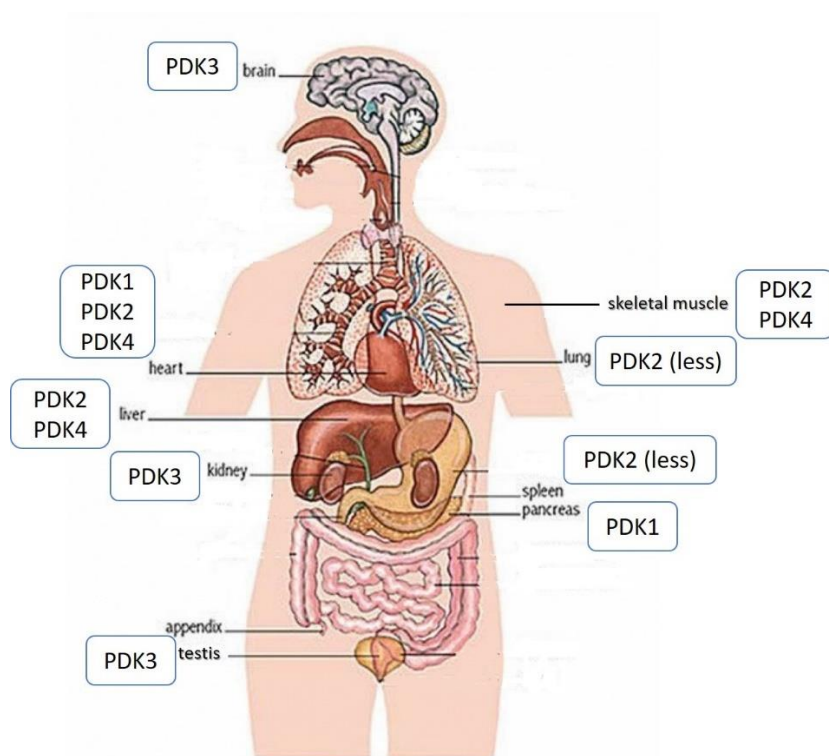


Figure 1.3: Tissue distribution of four isoforms of PDKs.

Transcription of PDKs is regulated by a lot of hormones including insulin, glucocorticoids, thyroid hormone and fatty acids. The abnormal increased expression of PDKs is related to a variety of diseases such as diabetes [13-15], heart disease [16-17], and fatty liver [18].

High level of PDK1 expression has been discovered in some tumor cells, which rely on aerobic glycolysis to generate energy. Instead of turning pyruvate to acetyl-CoA, the increasing transcription of PDK1 suppresses the activity of PDHc, thereby attenuating ROS (reactive oxygen species) production. The results of increasing glycolysis and ATP levels may suggest shunt pyruvate toward lactate production. [39]

PDK2 and PDK4 play a critical role in regulation of PDHc to maintain glucose homeostasis. Glucocorticoids [19] and free fatty acids [20-21] can stimulate PDK4 expression, while insulin can suppress it. Insulin has also been shown to repress PDK2 expression [20].

Failure of insulin to down-regulate the expression of PDK4 could result in type 2 diabetes. Overexpression of PDK4 in heart has shown to relate to heart disease for increasing fatty acid oxidation and decreasing glycolysis.

Table 1.1: PDKs and associated pathological conditions [22].

<i>PDK isoforms</i>	<i>Associated pathological conditions</i>
<i>PDK1</i>	Glioblastoma
	Brain aging

<i>PDK2</i>	Type 2 diabetes
	Brain aging
	Glioblastoma
	Ovarian cancer
<i>PDK3</i>	Charcot-Marie-Tooth neuropathy
<i>PDK4</i>	Type 2 diabetes
	Hemochromatosis
	Glucocorticoid excess, e.g., Cushing syndrome
	Cardiac hypertrophy
	Dilated cardiomyopathy
	Angiotensin II-induced heart failure
	Right ventricular hypertrophy and pulmonary hypertension
	Statin-induced myopathy
	Disuse osteoporosis
	Ovarian cancer
	Anoikis and tumor metastasis

The short-term control of PDKs is largely influenced by the organization and catalytic function of mammalian PDHc. The products during the five-step reaction of PDHc can exert some effects on the activity of PDKs, with high $[NADH]/[NAD^+]$ and $[acetyl-CoA]/[CoA]$ ratios enhancing the activity of PDKs [23-24], while adequate concentration of substrates pyruvate, NAD^+ and CoA could inhibit PDKs [24].

The phosphorylation of three target serine residues display different effects within the complex. Phosphorylation at site 1 is the fastest and could shut down the activity of PDHc. Phosphorylation of site 2 is more about promoting the inactivation and suppressing the reactivation via PDPs. The role of site 3 is mostly in preventing the reactivation [25]. Site 1 is positioned in the substrate channel leading to the E1p active center. One of the explanations of inactivation by phosphorylation is that introducing a negatively charged and bulky group would prevent the lipoyl domain of E2 from visiting the active site and picking up intermediates. This is supported by the results that phosphorylation of site 1 has nearly no impact on decarboxylation but severely attenuates the reductive acetylation reaction [26]. Site 2 and Site 3, which are not located near the active site, may affect PDHc activity by decreasing the binding ability towards ThDP [27] or disordering the conformation of the loop [28].

It has been widely believed that lipoyl domains play a pivotal role in modulation of PDK activities, though different PDKs may have different preferences towards the lipoyl domains located on E2 and E3BP. The structure of PDK3 bound to the L2 domain showed that binding to the inner lipoyl domain induced a conformational change of PDK3, ordering the C-terminal tails of the dimer to form a ‘crossover’, thereby enlarge the active site cleft to facilitate the release of ADP [29].

1.3 Inhibitors of PDKs

There are three types of inhibitors that have been mostly used and tested: AZD7545, dichloroacetate (DCA), and radicicol. DCA an analog of pyruvate, is one of the classic inhibitors for PDK isoforms. DCA inhibits the activity of PDKs by binding to the

allosteric site in the N-terminal domain. However, the specificity of DCA is low and it needs high doses for therapeutic effects [30-31].

Disrupting the interaction between PDKs and the inner lipoyl domain, thereby attenuating the recruitment of PDK1 to the complex, led to the development of AZD7545 [32]. This group of inhibitors also includes Pfz-3 and Nov3r [32-33].

Radicalol represents a class of ATP-competitive inhibitors, however, the low specificity limits its usage [34-35].

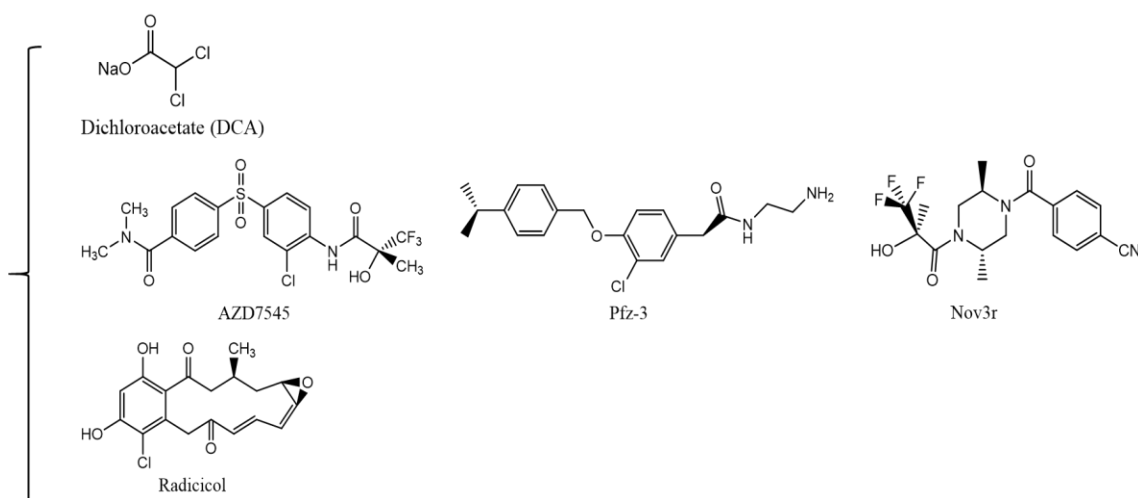


Figure 1.4: Structures of PDKs inhibitors

Additional inhibitors with higher selectivity across kinase family and sustained inhibitory effect, some forming covalent bonds and their mechanisms have been studied. Geng's group had investigated covalent modification at Cys240 of PDK1 with JX06, which could induce a conformational change through Van der Waal forces to block the access of ATP to its binding pocket and gave an IC_{50} value of 49 nM [36].

1.4 Objectives of this thesis

A recent study showed that PDHc regulation through PDK1 promotes aerobic glycolysis rather than pyruvate oxidation to oxidative phosphorylation in mitochondria (a phenomenon termed the *Warburg effect*), and plays a crucial role in glioblastoma multiforme, the most common and most aggressive malignant primary brain tumor in humans [37]. In addition, PDHc was also suggested to be a key player in oncogene-induced senescence. The oncogene-induced senescence was accompanied by simultaneous suppression of PDK1 and induction of PDP2. The resulting combined activation of PDHc enhanced the use of pyruvate in the tricarboxylic acid cycle, causing increased respiration and redox stress [38]. An elucidation of interactions between PDKs with E1 and E2p·E3BP components, and identification of interaction region between PDKs and components of PDHc can provide new insights into the long standing question of how human PDHc is assembled, and what kind of conformational changes human PDHc is experiencing through the interactions between PDKs and PDHc components.

Given that there is no X-ray or NMR structure for the intact E2 (from any source), alternative approaches must be developed to study the questions posed in this research.

The studies described herein were performed to determine the protein-protein interactions among six components of human pyruvate dehydrogenase complex. We determine the thermodynamics of binding of PDKs to other components of the PDHc by site-specific introduction of fluorophores onto specific domains of E2p·E3BP. In addition to that, we use the HDX-MS method to study the *in vitro* reconstitution of binary complexes of E1p and PDKs as well as the ligand effects on E1p and the PDK isozymes.

The author hopes that the experiments conducted in this thesis will help us to understand the activation mechanisms of PDKs and direct rational drug design of PDK isoform-specific inhibitors.

References

- [1] Randle P J, Garland P B, Hales C N, Newsholme E A. The glucose fatty-acid cycle its role in insulin sensitivity and the metabolic disturbances of diabetes mellitus[J]. The Lancet, 1963, 281(7285): 785-789.
- [2] Patel M S, Korotchkina L G. The biochemistry of the pyruvate dehydrogenase complex*[J]. Biochemistry and Molecular Biology Education, 2003, 31(1): 5-15.
- [3] Patel M S, Roche T E. Molecular biology and biochemistry of pyruvate dehydrogenase complexes[J]. The FASEB Journal, 1990, 4(14): 3224-3233.
- [4] Steussy C N, Popov K M, Bowker-Kinley M M, Sloan R B, Harris R A, Hamilton J A. Structure of Pyruvate Dehydrogenase Kinase A Novel Folding Pattern for A Serine Protein Kinase[J]. Journal of Biological Chemistry, 2001, 276(40): 37443-37450.
- [5] Wynn R M, Kato M, Chuang J L, Tso S C, Li J, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. Journal of Biological Chemistry, 2008, 283(37): 25305-25315.

- [6] Bowker-Kinley M M, Davis I W, Wu P, Harris R A, Popov K M. Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex[J]. *Biochemical Journal*, 1998, 329(1): 191-196.
- [7] Huang B, Gudi R, Wu P, Harris R A, Hamilton J, Popov K M. Isoenzymes of pyruvate dehydrogenase phosphatase DNA-derived amino acid sequences, expression, and regulation[J]. *Journal of Biological Chemistry*, 1998, 273(28): 17680-17688.
- [8] Di R, Feng Q, Chang Z, Luan Q, Zhang Y, Huang J, Li X, Yang Z. PDK1 plays a critical role in regulating cardiac function in mice and human[J]. *Chinese Medical Journal (English Edition)*, 2010, 123(17): 2358.
- [9] Sugden M C, Bulmer K, Augustine D, Holness M J. Selective modification of pyruvate dehydrogenase kinase isoform expression in rat pancreatic islets elicited by starvation and activation of peroxisome proliferator-activated receptor- α [J]. *Diabetes*, 2001, 50(12): 2729-2736.
- [10] Holness M J, Sugden M C. Regulation of pyruvate dehydrogenase complex activity by reversible phosphorylation[J]. 2003.1143-1151.
- [11] Pengfei Wu, Juichi Sato, Yu Zhao, Jaskiewicz J, Popov K M, Harris R A. Starvation and diabetes increase the amount of pyruvate dehydrogenase kinase isoenzyme 4 in rat heart[J]. *Biochemical Journal*, 1998, 329(1): 197-201.
- [12] Wu P, Inskeep K, Bowker-Kinley M M, Popov K M, Harris R A. Mechanism responsible for inactivation of skeletal muscle pyruvate dehydrogenase complex in starvation and diabetes[J]. *Diabetes*, 1999, 48(8): 1593-1599.
- [13] Kim Y I, Lee F N, Choi W S, Youn J H. Insulin regulation of skeletal muscle PDK4 mRNA expression is impaired in acute insulin-resistant states[J]. *Diabetes*, 2006, 55(8): 2311-2317.
- [14] Pehleman T L, Peters S J, Heigenhauser G J, Spriet L L. Enzymatic regulation of glucose disposal in human skeletal muscle after a high-fat, low-carbohydrate diet[J]. *Journal of applied physiology*, 2005, 98(1): 100-107.

- [15] Peters S J, Harris R A, Wu P, Pehleman T L, Heigenhauser G J, Spriet L L. Human skeletal muscle PDH kinase activity and isoform expression during a 3-day high-fat/low-carbohydrate diet[J]. *American Journal of Physiology-Endocrinology And Metabolism*, 2001, 281(6): E1151-E1158.
- [16] Kong S W, Bodyak N, Yue P, Liu Z, Brown J, Izumo S, Kang P M. Genetic expression profiles during physiological and pathological cardiac hypertrophy and heart failure in rats[J]. *Physiological Genomics*, 2005, 21(1): 34-42.
- [17] Zhao G, Jeoung N H, Burgess S C, Rosaaen-Stowe K A, Inagaki T, Latif S, Shelton J M, McAnally J, Bassel-Duby R, Harris R A, Richardson J A, Klierer S A. Overexpression of pyruvate dehydrogenase kinase 4 in heart perturbs metabolism and exacerbates calcineurin-induced cardiomyopathy[J]. *American Journal of Physiology-Heart and Circulatory Physiology*, 2008, 294(2): H936-H943.
- [18] Hwang B, Jeoung N H, Harris R A. Pyruvate dehydrogenase kinase isoenzyme 4 (PDHK4) deficiency attenuates the long-term negative effects of a high-saturated fat diet[J]. *Biochemical Journal*, 2009, 423(2): 243-252.
- [19] Kwon H S, Huang B, Unterman T G, Harris R A. Protein kinase B- α inhibits human pyruvate dehydrogenase kinase-4 gene induction by dexamethasone through inactivation of FOXO transcription factors[J]. *Diabetes*, 2004, 53(4): 899-910.
- [20] Huang B, Wu P, Bowker-Kinley M M, Harris R A. Regulation of pyruvate dehydrogenase kinase expression by peroxisome proliferator-activated receptor- α ligands, glucocorticoids, and insulin[J]. *Diabetes*, 2002, 51(2): 276-283.
- [21] Wu P, Peters J M, Harris R A. Adaptive increase in pyruvate dehydrogenase kinase 4 during starvation is mediated by peroxisome proliferator-activated receptor α [J]. *Biochemical and biophysical research communications*, 2001, 287(2): 391-396.
- [22] Lee I K. The role of pyruvate dehydrogenase kinase in diabetes and obesity[J]. *Diabetes & metabolism journal*, 2014, 38(3): 181-186.

- [23] Ravindran S, Radke G A, Guest J R, Roche T E. Lipoyl domain-based mechanism for the integrated feedback control of the pyruvate dehydrogenase complex by enhancement of pyruvate dehydrogenase kinase activity[J]. *Journal of Biological Chemistry*, 1996, 271(2): 653-662.
- [24] Behal R H, Buxton D B, Robertson J G, Olson M S. Regulation of the pyruvate dehydrogenase multienzyme complex[J]. *Annual review of nutrition*, 1993, 13(1): 497-520.
- [25] Kolobova E, Tuganova A, Boulatnikov I, Popov K M. Regulation of pyruvate dehydrogenase activity through phosphorylation at multiple sites[J]. *Biochemical Journal*, 2001, 358(1): 69-77.
- [26] Ciszak E M, Korotchkina L G, Dominiak P M, Sidhu S, Patel M S. Structural basis for flip-flop action of thiamin pyrophosphate-dependent enzymes revealed by human pyruvate dehydrogenase[J]. *Journal of Biological Chemistry*, 2003, 278(23): 21240-21246.
- [27] Korotchkina L G, Patel M S. Probing the mechanism of inactivation of human pyruvate dehydrogenase by phosphorylation of three sites[J]. *Journal of Biological Chemistry*, 2001, 276(8): 5731-5738.
- [28] Kato M, Wynn R M, Chuang J L, Tso S C, Machius M, Li J, Chuang D T. Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops[J]. *Structure*, 2008, 16(12): 1849-1859.
- [29] Kato M, Chuang J L, Tso S C, Wynn R M, Chuang D T. Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex[J]. *The EMBO journal*, 2005, 24(10): 1763-1774.
- [30] Whitehouse S, Cooper R H, Randle P J. Mechanism of activation of pyruvate dehydrogenase by dichloroacetate and other halogenated carboxylic acids[J]. *Biochemical Journal*, 1974, 141(3): 761-774.

- [31] Bersin R M, Stacpoole P W. Dichloroacetate as metabolic therapy for myocardial ischemia and failure[J]. American heart journal, 1997, 134(5): 841-855.
- [32] Morrell J A, Orme J, Butlin R J, Roche R M, Mayers E K. AZD7545 is a selective inhibitor of pyruvate dehydrogenase kinase 2[J]. 2003.
- [33] Kato M, Li J, Chuang J L, Chuang D T. Distinct structural mechanisms for inhibition of pyruvate dehydrogenase kinase isoforms by AZD7545, dichloroacetate, and radicicol[J]. Structure, 2007, 15(8): 992-1004.
- [34] Roe S M, Prodromou C, O'Brien R, et al. Structural basis for inhibition of the Hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin[J]. Journal of medicinal chemistry, 1999, 42(2): 260-266.
- [35] Sun W, Xie Z, Liu Y, Zhao D, Wu Z, Zhang D, Lv H, Tang S, Jin N, Jiang H, Tan M, Ding J, Luo C, Li J, Huang M, Geng M. JX06 selectively inhibits pyruvate dehydrogenase kinase PDK1 by a covalent cysteine modification[J]. Cancer research, 2015, 75(22): 4923-4936.
- [36] Velpula, K. K., Bhasin, A., Asuthkar, S., Tsung, A. J. (2013) Combined targeting of PDK1 and EGFR triggers regression of glioblastoma by reversing the Warburg effect. Cancer Res. Oct 22.
- [38] Hur, H., Xuan, Y., Kim, Y. B., Lee, G., Shim, W., Yun, J., Ham I. H., Han, S. U. (2013) Expression of pyruvate dehydrogenase kinase-1 in gastric cancer as a potential therapeutic target. Int. J. Oncol. 42, 44-54.
- [39] Kim J, Tchernyshyov I, Semenza G L, Dang C V. HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia[J]. Cell metabolism, 2006, 3(3): 177-185.

CHAPTER 2. Site specifically introduced external fluorophore to study protein-protein interactions within PDHc

2.1 Introduction

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^+ to NADH and the release of CO_2 [1-2].

The E2 component in those complexes is a dihydrolipoamide acyltransferase and has variable numbers of tandem lipoyl domains, each carrying a covalently attached lipoic acid as cofactor.

The E2p component of PDHc has a multi-domain structure, comprising from the N-terminal end: two tandem lipoyl domains, the outer (L1) and inner lipoyl domain (L2) approximately 9 kDa each, a peripheral subunit-binding domain (S, 4kDa) and the acetyltransferase or catalytic domain (Core, 28kDa), all separated by 25-30 amino acid-long flexible linkers.

The E3BP is composed of three linker-connected domains, similar but not identical to those in E2, a single lipoyl domain (L3), a variant of S, called S' to which E3 binds, and a catalytic domain C' that is incompetent to produce acetyl-CoA.

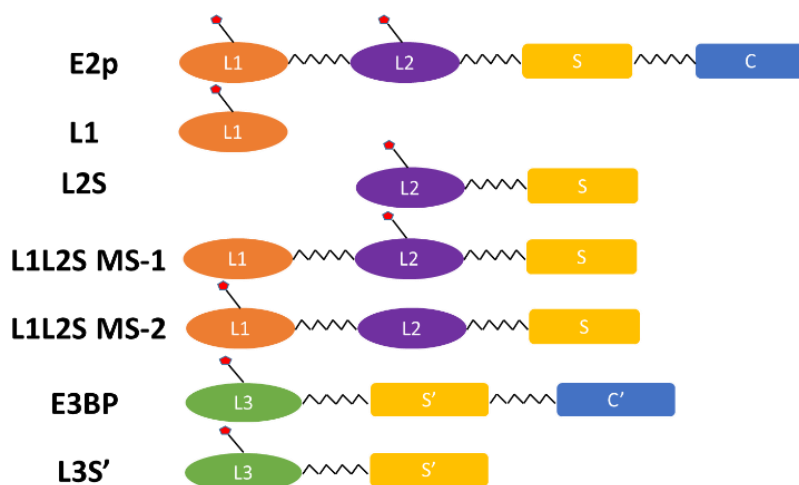


Figure 2.1: Domain structures of E2-E3BP- derived proteins

Trivalent arsenicals were used for potent irreversible inhibition of PDHc and many other proteins for many years, since they can form high-affinity ring structures with closely spaced dithiols [3-4].

The cyclic dithiolarsinites are markedly more stable than the non-cyclic products formed from trivalent arsenicals and monothiols because of entropic considerations [5].

1-Dimethylamino-naphthalene-5-sulfonyl chloride (DANS-Cl) also known as dansyl chloride is a useful fluorescent tag, which has been widely reacted with the lysine amino group of enzymes in attempts to understand the chemical nature of the active centers. Dansyl chloride is non-fluorescent until it reacts with amines. The fluorescence properties

of the resulting dansyl amides are environmentally sensitive and have large Stoke shifts [6-8].

In order to study the interaction between lipoyl domains and PDKs, we developed a method for *in situ* derivatization of the lipoyl domain. First of all, we use p-arsanilic acid (NH_2PhAsO) to react with dansyl chloride, forming a fluorescent tag which contains a trivalent arsenoxide end. Then, the fluorophore is specifically attached to a lipoyl domain by the reaction of reduced dihydrolipoyl groups with the trivalent arsenoxide end of DANS-As. Finally, the change in fluorescence is monitored when different PDKs are titrated with lipoyl domains bearing the fluorophore.

2.2 Site-specific labeling method to study protein-protein interactions within PDHc

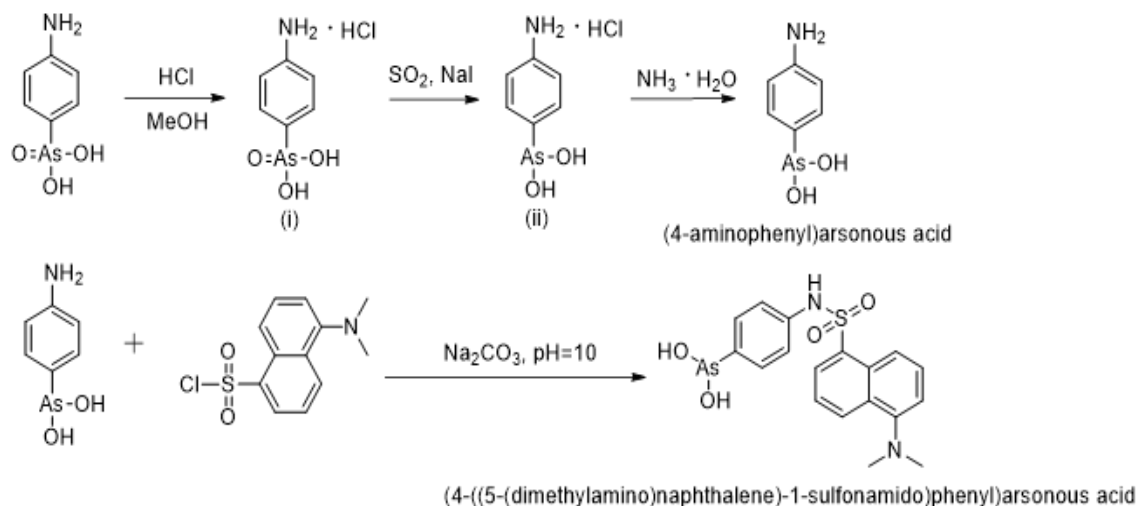
2.2.1 Materials and Methods

Recombinant E2•E3BP core was overexpressed in *E. coli* cells and was purified as described previously (Hiromasa et al., 2004). The following recombinant C-terminally truncated proteins of the E2•E3BP core available from the Patel group were used in this study: L1, containing the outer lipoyl domain and linker (residues 1–98); the L2S didomain containing the inner lipoyl domain (L2), second hinge region, peripheral subunit-binding domain (S), and third hinge region (residues 128–330); the L1L2S tridomain, which comprises L1, L2, both hinge regions, and the subunit-binding domain (S, residues 1–330); the L3S' didomain of E3BP (residues 1–230), all expressed from pET28b in *E. coli* BL21 (DE3) cells and purified using Ni Sepharose 6 Fast Flow column [9]. Two singly substituted variants of L1L2S were used in this study: L1L2S-MS1

(Lys46Ala substitution in the outer lipoyl domain with inner lipoyl domain available for modification), and L1L2S-MS2 (Lys173Ala substitution in the inner lipoyl domain with outer lipoyl domain available for modification). We also constructed four doubly substituted variants of L1L2S to study the function of specific residues on L1L2S: E35A/K173A, E35K/K173A, K46A/E209A, K46A/E209K.

Recombinant rat PDK1, rat PDK2, human PDK3 and rat PDK4 were from the Patel group. The PDK isozymes were overexpressed and purified individually from *E. coli* BL21(DE3) cells transformed with pPDK expression vector using protocols reported in the literature [10-13]. Analysis of the deduced amino acid sequences of human PDK1 and PDK2 reported in the literature revealed high identity between human and rat PDK2 (96%) and between human and rat PDK1 (93%). PDK3 shares 68% and 67% identity with PDK1 and PDK2, respectively [14].

The synthesis of DANS-As is described in Scheme 1. Briefly, p-arsanilic acid (10.85 g, 50 mmol) was acidified by hydrochloric acid and then reduced to (4-aminophenyl) arsenous acid using sulfur dioxide with sodium iodide as a catalyst [15]. Sulfur dioxide was bubbled into the solution for 2 h, then the mixture was cooled in an ice-water bath, and a pale yellow precipitate resulted. The precipitate (5 g) was dissolved in 15 mL of NH₄OH and stirred for 30 min. Once the solution was cooled in an ice-water bath, a white layer of crystals of (4-aminophenyl)arsenous acid precipitated. Next, (4-aminophenyl) arsenous acid (149 mg, 0.742 mmol) was reacted with dansyl chloride (100 mg, 0.371 mmol) at pH 10 for 8–12 h at room temperature producing (4-((5-(dimethylamino)naphthalene)-1-sulfonamido)phenyl) arsenous acid (DANS-As).



Scheme 2.1: Synthesis of (4-(5-(dimethylamino)naphthalene)-1-sulfonamido)phenyl)-arsenous acid (DANS-As).

In vitro lipoylation of the E2·E3BP derived proteins. The E2·E3BP derived proteins with the lipoyl domains were lipoylated *in vitro* using *E. coli* lipoyl protein ligase as reported by us recently [16]. Lipoylation was confirmed by FT-MS using the electrospray ionization sampling method.

Reaction of DANS-As with the source of the lipoyl domains (L1, L2S, L1L2S, L1L2S-MS1, L1L2S-MS2, doubly substituted L1L2S and L3S'). First, the lipoylated source of the lipoyl domain (50 μM) in 30 mM KH₂PO₄ (pH 7.5) containing 0.15 M NaCl (sample buffer) was reduced by incubation for 5 min with TCEP (100 μM) at room temperature, resulting in fully reduced lipoyl domains according to FT-MS. Second, the reduced lipoyl domains were then reacted with 150 μM DANS-for 2 h at room temperature, leading to chemically modified lipoyl domains. The excess of TCEP and DANS-As were removed

from the reaction mixture with a 10K MWCO centrifugal filter unit (Vivaspin 500, 10K MWCO).

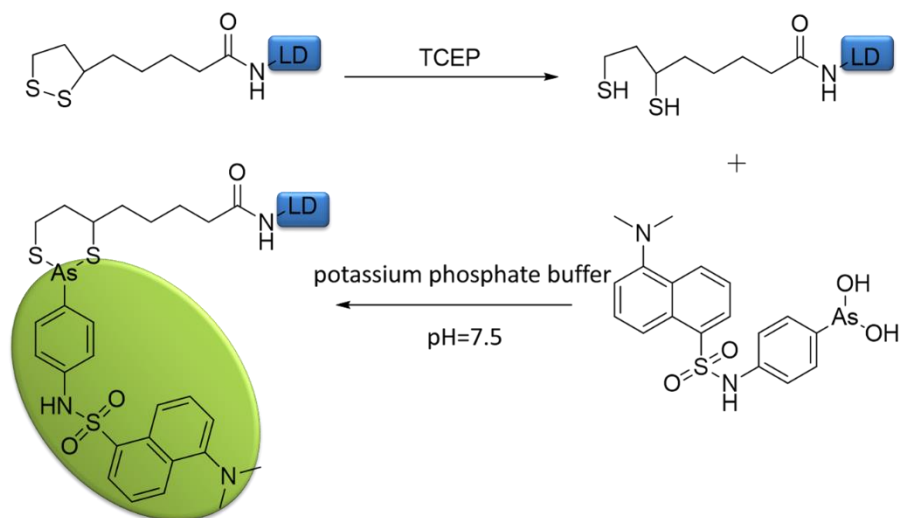


Figure 2.2: Using DANS-As to chemically modify LDs.

L3S' was modified by using Dapoxyl carboxylic succinimidyl ester. L3S' (50 μ M) was incubated with 100 μ M Dapoxyl carboxylic succinimidyl ester in sample buffer (30 mM KH_2PO_4 , 0.15 M NaCl, pH=7.5) at room temperature for 2 h. The reaction mixture was dialyzed against the sample buffer with a centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the extra Dapoxyl carboxylic succinimidyl ester.

Modification of the lipoyl domains was confirmed by FT-MS, for which each modified lipoyl domain containing protein was diluted to 1 μ M concentration in a solution containing 50:50:0.1 (v/v/v) of methanol, water and formic acid.

Modification of the lipoyl domains from doubly substituted L1L2S variants was confirmed by LC-MS. DANS-modified protein (50 ng in 20 μ L) was added to a 20 mM

sodium acetate buffer (pH=4.5) with a 50% slurry of immobilized pepsin. Then, the mixture was incubated at 37°C for 4 h with high speed shaking. The digested reaction mixture was separated from the immobilized pepsin by centrifugation, then, lyophilization. The lyophilized material was dissolved in 50 µL of 5% acetonitrile and LC-MS was run on a Phenomenex Luna C18 column.

Fluorescence Spectroscopy. For the fluorescence titration of the modified lipoyl domains by PDK isozymes, the DANS-modified lipoyl domain (2 µM) or Dapoxyl modified lipoyl domain (2 µM) in 30 mM KH₂PO₄ (pH7.5) were titrated by PDK1 (0.78-37.9 µM), or PDK2 (0.18-4.38 µM), or PDK3 (0.24-3.86 µM), or PDK4 (0.24-4.20 µM).

Fluorescence spectra were recorded at 25 °C using a Cary Eclipse spectrometer. The excitation wavelength was 338 nm, and the emission spectra were recorded in the range of 425-600 nm in 3 mL quartz cuvettes. Data were fitted to a Hill equation,

$$\Delta F / \Delta F_{max} = [PDK]^n / (S_{0.5}^n + [PDK]^n) \quad \text{Equation 2.1}$$

where $\Delta F / \Delta F_{max}$ is a relative fluorescence; $\Delta F = F_{max} - F_i$, where F_{max} is a maximum fluorescence intensity reached on titration by PDK and F_i is a fluorescence intensity at a given concentration of PDK; $\Delta F_{max} = F_{max} - F_o$ where F_o is the initial fluorescence before addition of PDK; $S_{0.5}$, is the concentration at half saturation. n is the Hill coefficient. For $n=1$, the value of $S_{0.5}$ is equal to K_d .

2.3 Results and discussion of E2-E3BP-derived proteins interacting with PDKs

The consecutive reactions of lipoylation, reduction, then attachment of the external fluorophore were monitored by mass spectrometry. An example of L1 is shown as figure.

On reduction of the lipoamide tethered to the L1, its mass was increased from 11972.09 Da to 11974.11 Da (the theoretical mass of the lipoylated L1 is 11971.3 Da). After attaching the DANS-As, the L1 mass was further increased to 12372.09 Da corresponding to the mass of an added DANS-As group.

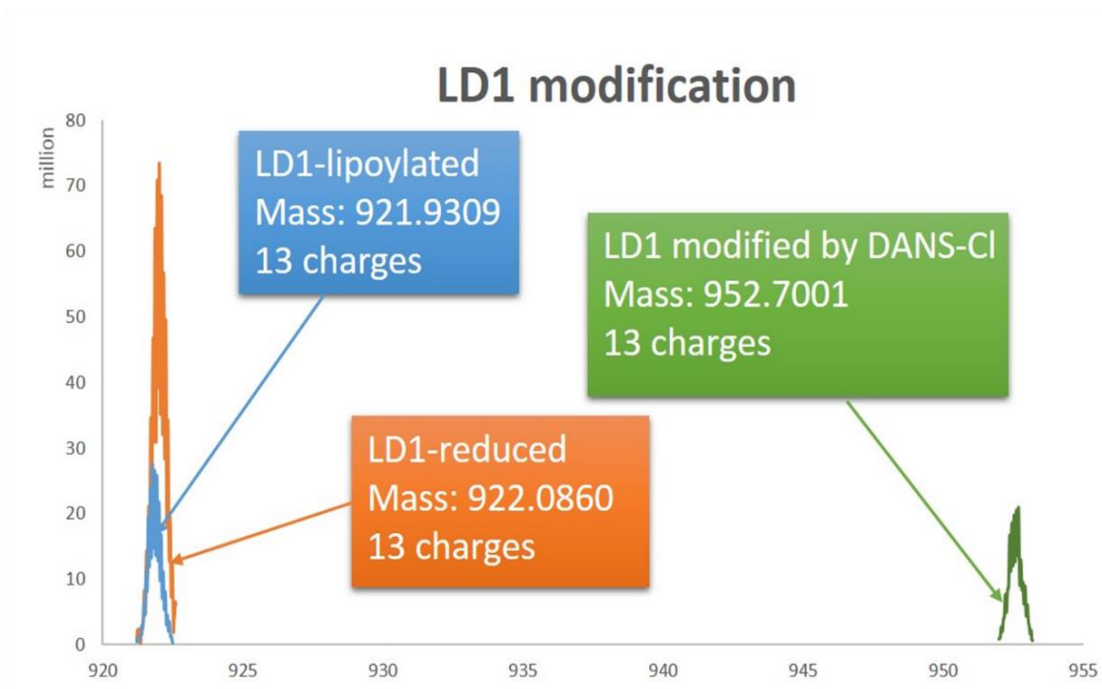


Figure 2.3: An example of chemical modification monitored by ESI-MS. The lipoamide from L1 can be reduced to the dihydrolipoyl group by TCEP, as shown on the mass spectrum: the peak with 13 charges increased from 921.9309 Da to 922.0860 Da. After attaching the DANS-As to L1, the peak increased from 922.0860 Da to 952.7001 Da, corresponding to the mass of adding the DANS-As group. The labeling efficiency during this process was approximately 75%.

Typical changes in the fluorescence of the DANS-As-labeled L2S and DANS-As-labeled L3S' on titration by PDK1 and PDK2 are shown in Figs. 2.4 and 2.5. It is evident that with the dansyl group attached to L2S, binding to PDK1 enhanced its fluorescence intensity. Similar Fluorescence intensity changes were observed for DANS-As-labeled L1, L2S, L1L2S-ML1, L1L2S-ML2 on titration by each of the four PDK isozymes. Attachment of the DANS-As group to L3S', in contrast, led to DANS fluorescence quenching rather than enhancement when interact with PDK2 and PDK3. These data suggest different binding environment for DANS-As-labeled L2S and L3S' on binding the PDK isoforms. The fluorescence enhancement suggests that dansyl group is transferred to a more hydrophobic environment, while the fluorescence quenching suggests that the dansyl group is transferred to a relatively more hydrophilic environment [17].

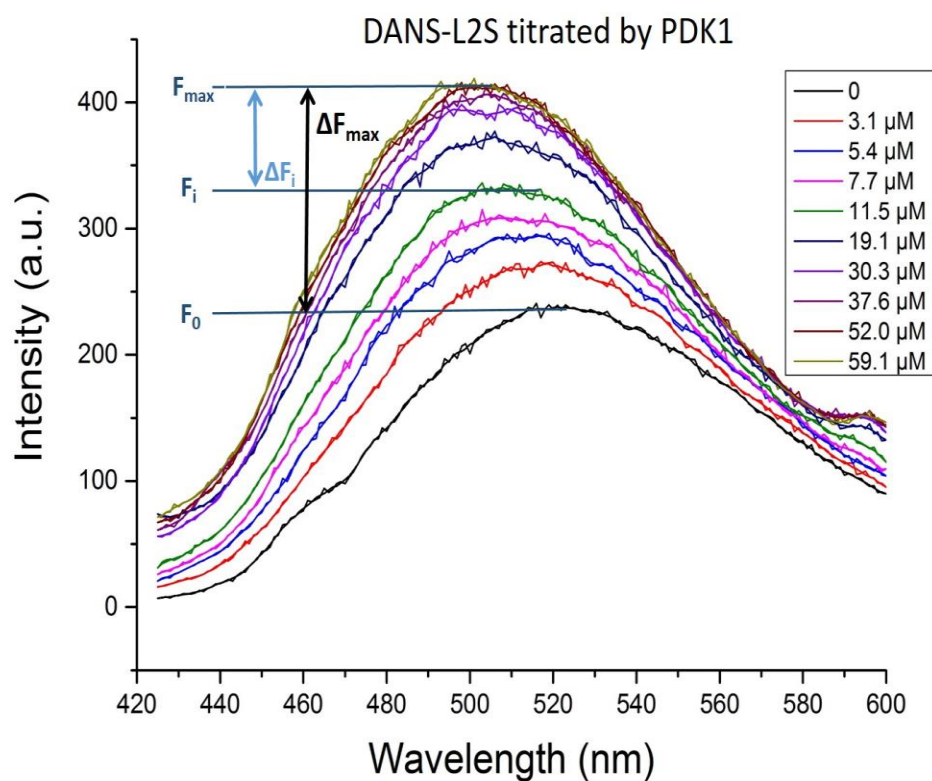


Figure 2.4: Fluorescence titration of DANS-As labeled L2S by PDK1. Enhancement of fluorescence of DANS-As-L2S upon PDK1 binding. DANS-As-L2S (1.5 μM) in 30 mM KH_2PO_4 (pH=7.5) was titrated by PDK1 (3.1-59.1 μM)

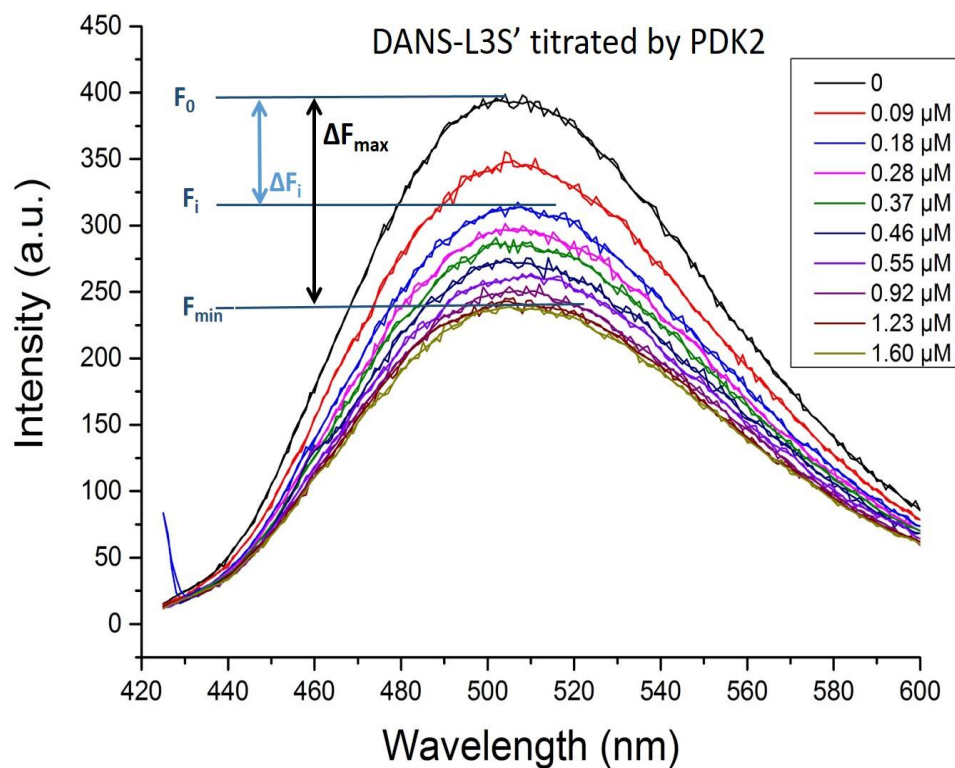


Figure 2.5: Fluorescence titration of DANS-As labeled L3S' by PDK2. Quenching of fluorescence of DANS-As-L3S' upon PDK2 binding. DANS-As-L3S' (1.5 μM) in 30 mM KH_2PO_4 (pH=7.5) was titrated by PDK2 (0.09-1.60 μM)

The fluorescence titration curves fitted with a Hill equation are summarized in the Figure below.

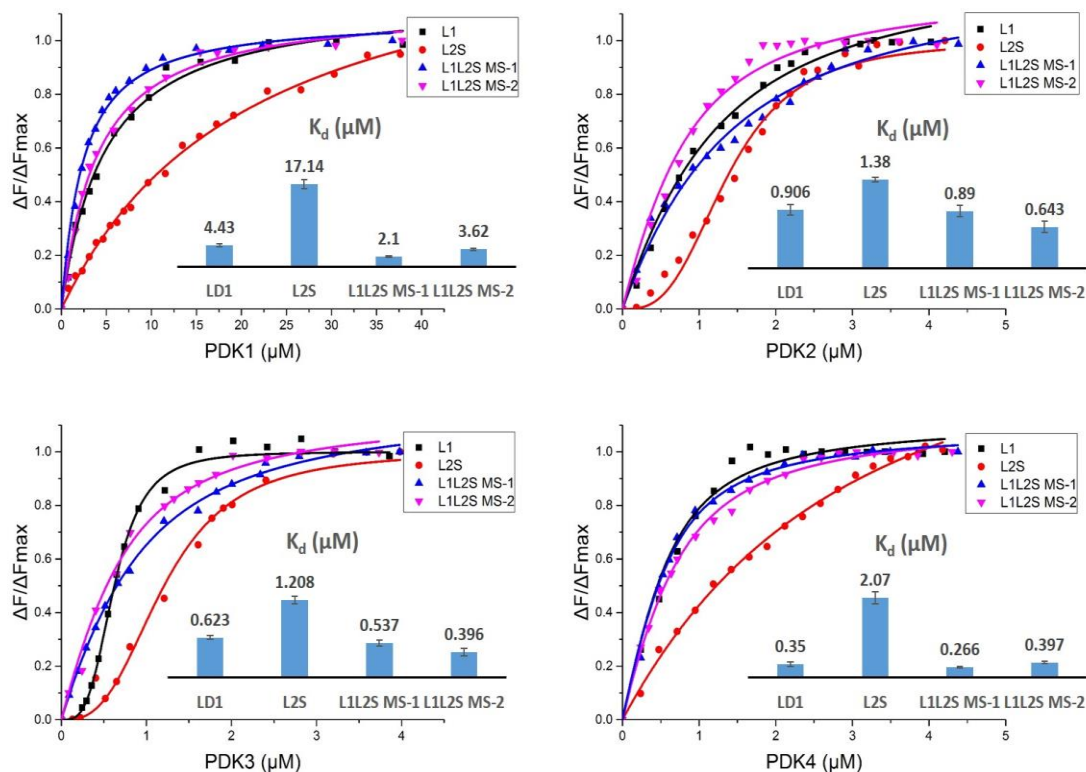


Figure 2.6: Fluorescence binding curves for DANS-As-labeled E2 E3BP derived lipoyl domains on titration by PDK isoforms. The DANS-As-labeled L1 (1.0 μM), or L2S (1.5 μM) or L1L2S MS-1 (2.5 μM) or L1L2S MS-2 (3.0 μM) in 30 mM KH_2PO_4 (pH=7.5) was titrated by PDK1 (0.78-37.93 μM) or PDK2 (0.18-4.38 μM) or PDK3 (0.24-3.86 μM) or PDK4 (0.24-4.2 μM) at room temperature. The excitation wavelength was 338 nm and the emission spectra were recorded in the 425-600 nm range. In all cases the fluorescence intensity of the DANS-labeled lipoyl domains was enhanced on PDKs binding.

A second fluorophore (the dapoxyl group) was also introduced to L3S', this time non-specifically to lysine residues of the L3S'. As seen in Figure 2.7, fluorescence

enhancement resulted from the interaction of dapoxyl-labeled L3S' with PDK2 and PDK4.

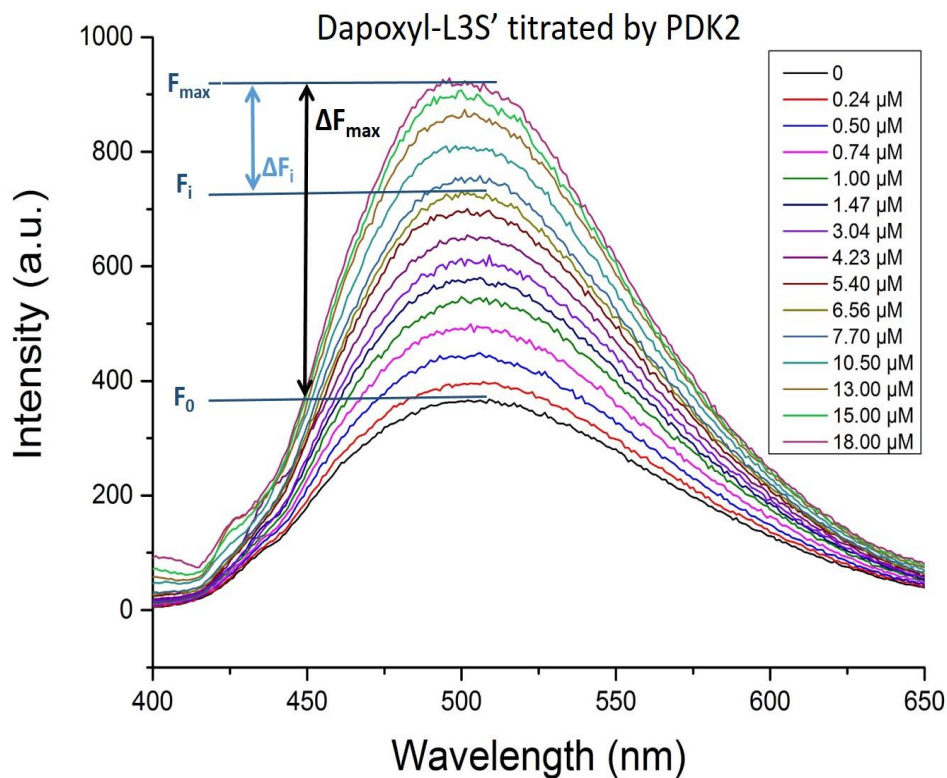


Figure 2.7: Fluorescence titration of dapoxyl-labeled L3S' by PDK2. The dapoxyl-L3S' (3 μM) in 30 mM KH₂PO₄ (pH=7.5) was titrated by PDK2 (0.24-18.0 μM) that was accompanied by the enhancement of the dapoxyl-L3S' fluorescence intensity. The PDK4 behaves similarly.

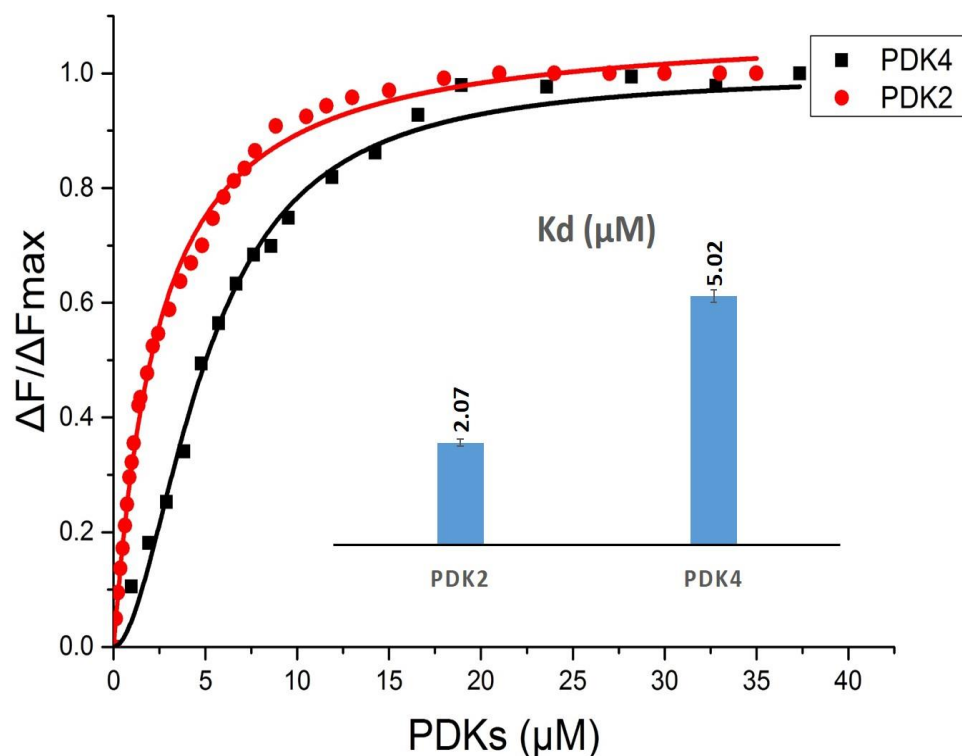


Figure 2.8: Fluorescence binding curves for dapoxyl-labeled L3S' on titration by PDK2 and PDK4. The K_d values for PDK2-L3S' and PDK4-L3S' complexes are 2.07 μM and 5.02 μM , respectively. Excitation wavelength was at 327 nm and the emission spectra were recorded in the 400-650 nm range. No changes in fluorescence intensity of dapoxyl-L3S' was observed on titration by PDK1 or PDK3.

The calculated values of K_d are presented in Table 2.1.

K_d (μM)	DANS-L1	DANS-L2S	DANS-L1L2S MS1	DANS-L1L2S MS2	DANS-L3S'	Dapoxyl-L3S'
PDK1	4.43 ± 0.35	17.14 ± 0.94	2.10 ± 0.12	3.62 ± 0.25	No binding	No binding
PDK2	0.91 ± 0.08	1.38 ± 0.03	0.89 ± 0.09	0.64 ± 0.09	0.025^a	2.07 ± 0.06

PDK3	0.62 ± 0.03	1.21 ± 0.06	0.54 ± 0.05	0.396 ± 0.06	0.103^a	No binding
PDK4	0.35 ± 0.06	2.07 ± 0.16	0.27 ± 0.02	0.397 ± 0.03	5.93 ± 1.06	5.02 ± 0.13

Table 2.1: K_d values for binding of the E2-E3BP-derived domains to PDK1-PDK4 as

detected by fluorescence spectroscopy. On binding of PDK2 and PDK3 to DANS-L3S', quenching of the DANS-L3S' fluorescence was observed while in others enhancement resulted.

As reported in Table 2.1, among the four PDK isoforms, the interaction between PDK1 and different DANS-As-labeled lipoyl domains is the weakest, which might explain why PDK1 is the only kinase that can phosphorylate all three sites of the E1 α subunit (15). It appears that PDK2 and PDK3 interact with the DANS-As-labeled lipoyl domains similarly, with both DANS-As-L1 and DANS-As-L2S contributing significantly to binding. The K_d values of the PDK2-DANS-As-L1 complex (0.91 μ M) and PDK2-DANS-As-L1L2S-MS1 complex (0.89 μ M) are nearly identical, as are the values for the PDK3-DANS-As-L1 (0.62 μ M) and PDK3-DANS-As-L1L2S-MS1 (0.54 μ M) complexes, suggesting that the lipoamide region in L2 does not contribute to binding of PDK2 or PDK3. The PDK4 has clear preference for L1 ($K_{d,L1} = 0.35$ μ M; $K_{d,L2S} = 2.1$ μ M), similarly to PDK2 and PDK3.

We also wished to investigate the interaction of the PDK isoforms with L3S' labeled with both the DANS-As and the dapoxyl group. The PDK1 displayed no binding with either labeling, while PDK4 showed weak binding (Table 2.1). Analysis of the titration data indicated stronger binding to PDK2 ($K_d = 2.1$ μ M for dapoxyl-L3S' and 0.025 μ M for

DANS-As-L3S') than to PDK4 ($K_d \sim 5.0 \mu\text{M}$ with both labeled L3S'). The PDK3 isoform produced a signal only with L3S'.

In comparison with our data, an earlier report on the interaction of PDK4 with individual L1 and L2 domains was unable to identify these interactions by enthalpy changes using ITC, that study suggested weak binding. A K_d of $5 \mu\text{M}$ was reported for L3 binding to PDK4 [18], a value that correlates well with values of $K_d = 5.93 \mu\text{M}$ (DANS-As-L3S') and $K_d = 5.02 \mu\text{M}$ (dapoxyl-L3S') determined by fluorescence spectroscopy in Table 1. On the basis of the K_d values presented in Table 2.1, it became evident that L1 and L3S' also participate in the interaction with some PDK isoforms, not recognized before. This conclusion is supported by a comparison of the K_d values in Table 2.1 with those reported earlier, and using different approaches to calculate binding constants. Values of K_d of $\sim 175 \mu\text{M}$ (oxidized L2) and $K_d \sim 130 \mu\text{M}$ (reduced L2) for the PDK2-L2 complex were obtained by analytical ultracentrifugation [19]. On conjugation of L2 to glutathione-S transferase, the PDK2-L2 complex appeared to be stronger with $K_d \sim 3 \mu\text{M}$ (oxidized L2), and $K_d \sim 0.4 \mu\text{M}$ (reduced L2), as compared with $K_d \sim 22 \mu\text{M}$ (L1) and $K_d \sim 35 \mu\text{M}$ (L3), where L1 and L3 were also conjugated to glutathione-S transferase [20]. A value of $K_d \sim 10 \mu\text{M}$ was calculated for the PDK1-L2 and PDK2-L2 complexes using gel filtration chromatography, while the same method gave the following relative affinities for complexation with L2: PDK3 > PDK1 = PDK2 > PDK4 (46?). A value of $K_d = 1.17 \pm 0.23 \mu\text{M}$ was obtained by ITC for the PDK3-L2 complex [21].

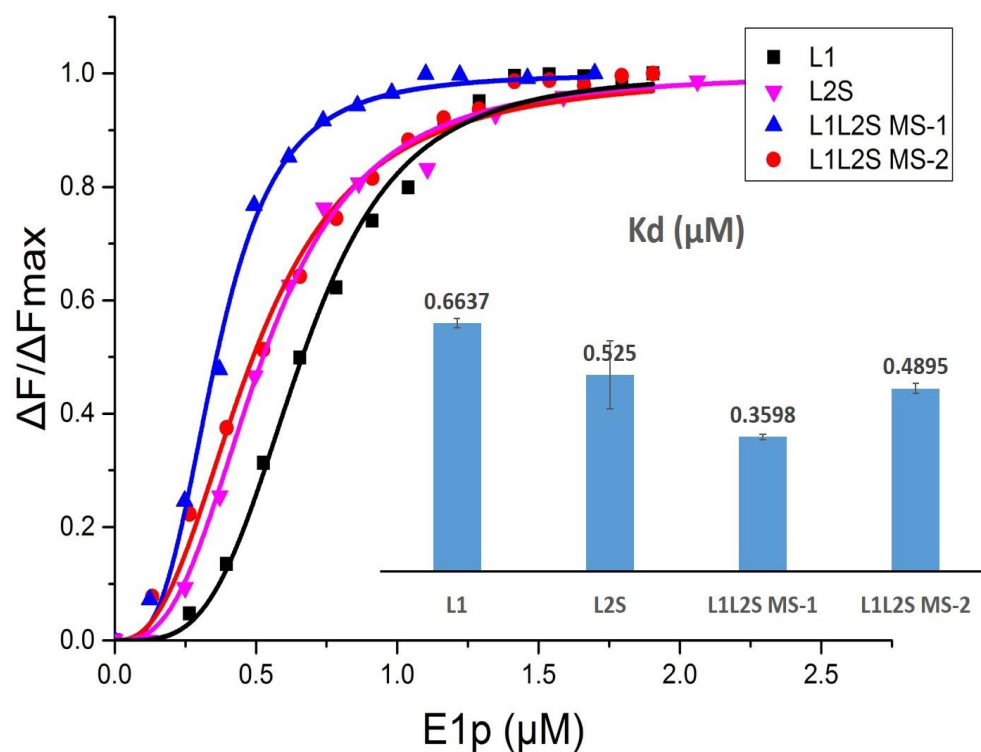


Figure 2.9: Fluorescence binding curves for DANS-As-labeled E2·E3BP derived lipoyl domains on titration by E1p.

We used the same method to study the interaction between E1p and different lipoyl domains. The results showed that E1p prefers L2S over L1. Chuang's group has published a K_d value of $1.29 \pm 0.40 \mu\text{M}$ for E1p-L2 by using ITC [22]. Our value was lower than his, since our protein construct not only included the L2 domain, but also the subunit-binding domain.

2.4 Results of doubly substituted L1L2S variants interacting with PDKs and E1p

Previously, two complementary methods had been used to identify the ‘hot spot’ in interaction loci between PDK1 (PDK2) and the E2·E3BP-derived domains in our group. Some residues have been selected and the site-directed mutagenesis method was used to validate the function of these sites. Four doubly substituted variants of L1L2S have been constructed, they are L1L2S (E35A/K173A, E35K/K173A, K46A/E209A, K46A/E209K). Here, we utilize the same method we have described before, of introducing the DANS-As fluorophore onto the lipoyl groups of L1L2S doubly substituted variants, and then study the dissociation constants between these mutants and PDK isoforms.

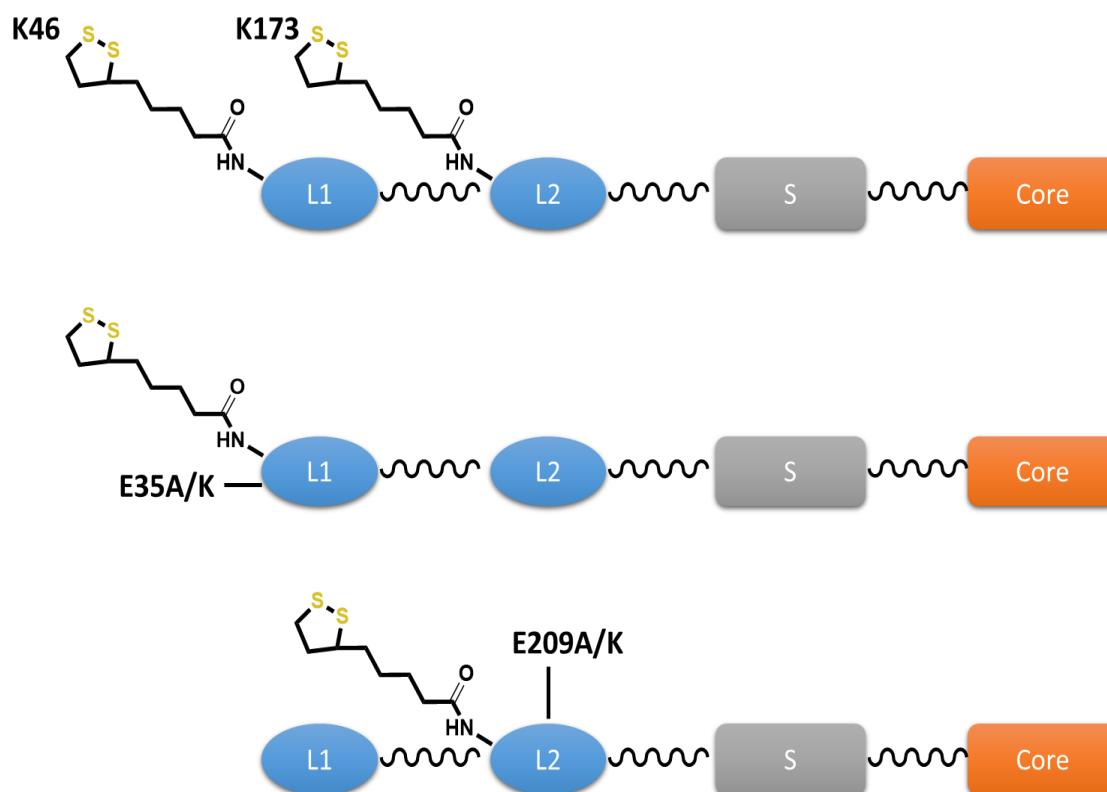


Figure 2.10: Domain structures of L1L2S doubly substituted variants (E35A/K173A, E35K/K173A, K46A/E209A, K46A/E209K).

First of all, pepsin digestion and then LC-MS were used to confirm the site-directed mutagenesis peptide sequence as well as the site-specific introduction of the DANS-As label.

Table 2.2: Peptides mass checked by LC-MS.

	peptides	Theoretical mass	Mass detected by FT-MS
K46	VETD K ATVGF+DANS	827.8022 ²⁺	827.8061 ²⁺
E35A	EKKEGD A INKGDL	708.8701 ²⁺	708.8721 ²⁺
K173A	IETD A ATIGF	1037.5150 ¹⁺	1037.5212 ¹⁺
E35K	EKKEGD K INKGDL	737.3990 ²⁺	737.4006 ²⁺
K46A	IAEVETD A ATVGF	661.8274 ²⁺	661.8309 ²⁺

Among the four double mutants (*E35A/K173A*, *E35K/K173A*, *K46A/E209A*, *K46A/E209K*) we tried, only *K46A/E209K* showed fluorescence change when titrated by PDKs. The spectra displaying fluorescence enhancement are displayed in Figure 2.10.

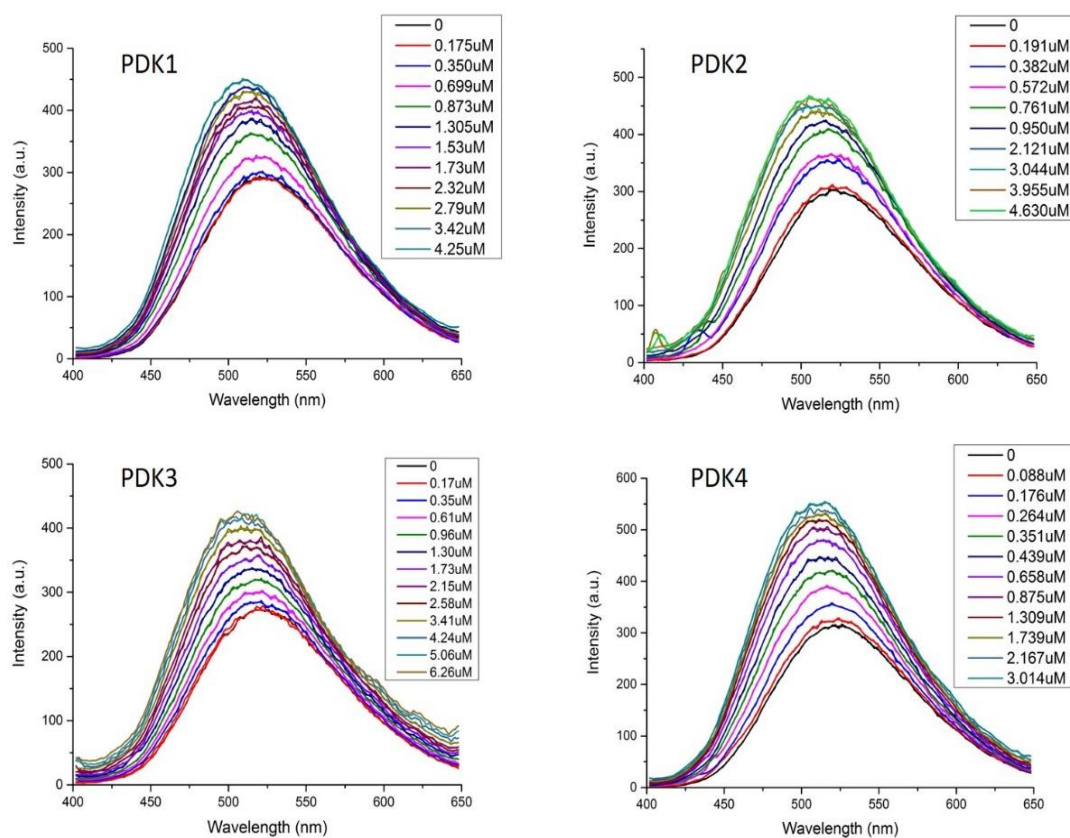


Figure 2.11: Fluorescence titration of DANS- labeled L1L2S (K46A/E209K) by four isoforms of PDKs. Enhancement of fluorescence of the DANS- L1L2S on PDKs binding.

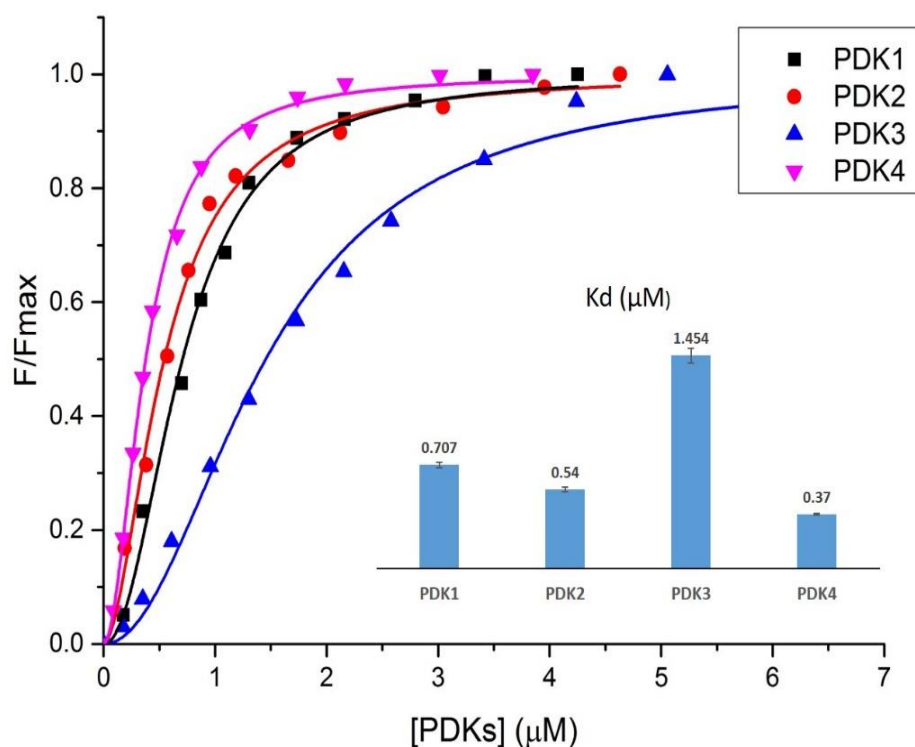


Figure 2.12: Fluorescence binding curves for DANS labeled L1L2S doubly substituted variant (K46A/E209K) on titration by PDK isoforms. In all cases the fluorescence intensity of the DANS-labeled lipoyl domains was enhanced on PDKs binding. The DANS-modified K46A/E209K (1.0 μM) in 30 mM KH_2PO_4 (pH 7.5) was titrated by PDKs at room temperature. The excitation wavelength was 338 nm and the emission spectra were recorded in the 400 - 600 nm range. The fluorescence titration curves were fitted by using the Hill equation (Equation 2.1 in Experimental procedures), as some of the curves displayed significant sigmoidicity.

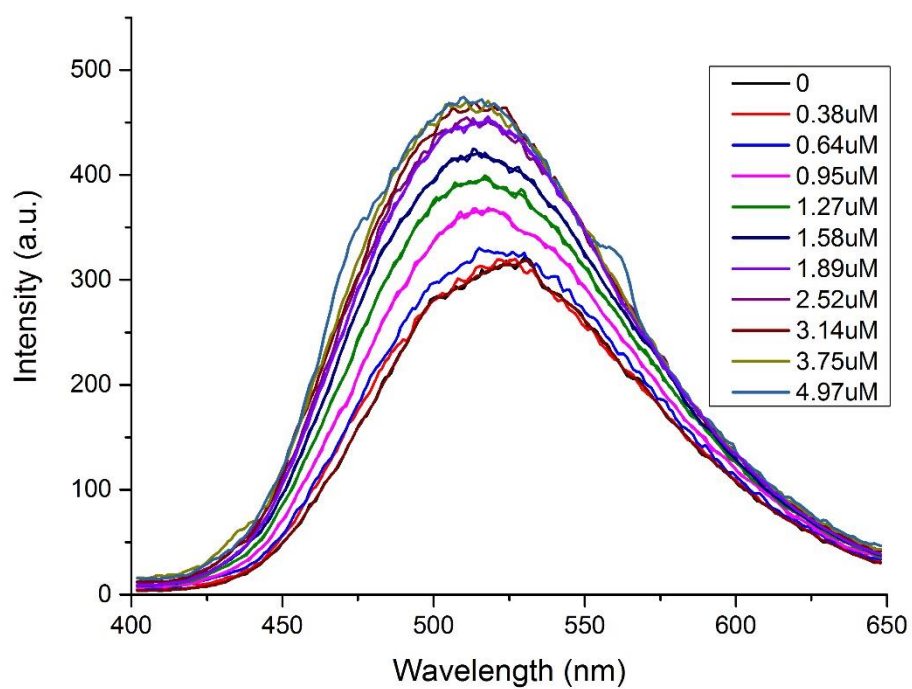


Figure 2.13: Fluorescence titration of DANS- labeled L1L2S (K46A/E209K) by E1p. Enhancement of fluorescence of the DANS- L1L2S on E1p binding.

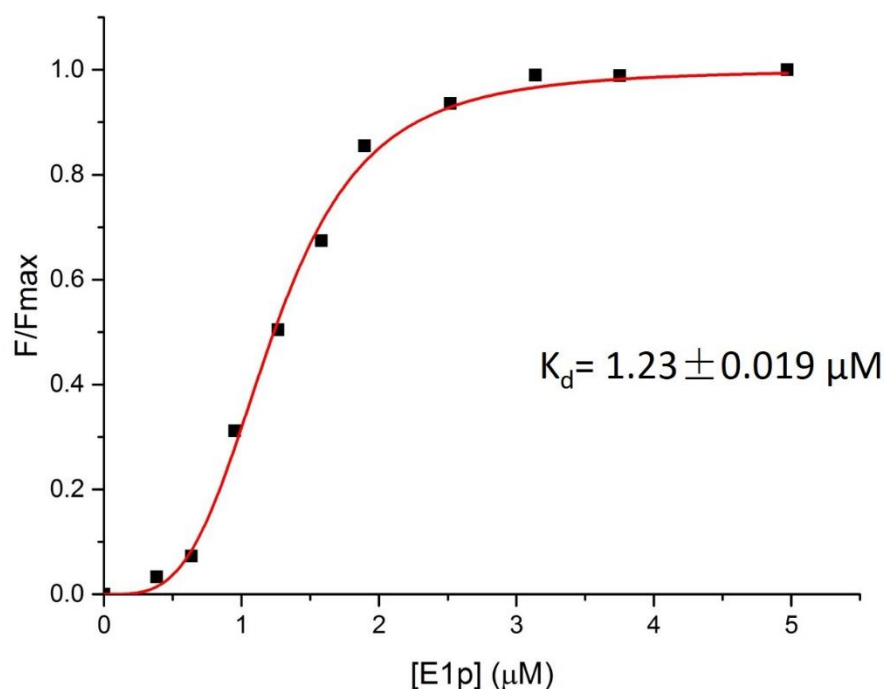


Figure 2.14: Fluorescence binding curves of DANS-As-L1L2S (K46A/E209K)

titrated by E1p. Binding of E1p enhanced the fluorescence intensity and the enhancement was used to measure the dissociation constants. The titration experiment was conducted in 30 mM KH₂PO₄ (pH 7.5) at room temperature. Excitation wavelength at 338 nm, the spectrum was recorded in the range of 400-600 nm.

/μM	K46A	K173A	K46A/E209K
PDK1	2.10±0.12	3.62±0.025	0.707±0.020
PDK2	0.89±0.09	0.643±0.029	0.54±0.017
PDK3	0.537±0.05	0.396±0.056	1.454±0.050
PDK4	0.266±0.02	0.397±0.029	0.37±0.005

E1p	0.36±0.007	0.49±0.013	1.23±0.019
------------	------------	------------	------------

Table 2.3: K_d values for binding of the L1L2S mutants to PDK1-PDK4 and E1p as detected by fluorescence spectroscopy. K_{ds} are in μ Molar units.

Among the four double variants (*E35A/K173A*, *E35K/K173A*, *K46A/E209A*, *K46A/E209K*) we tried, only *K46A/E209K* showed fluorescence change when titrated by both PDKs and E1p. And compared with the K_d values with those of single variants (*K46A* and *K173A*), the additional substitution on E209K site has an effect of increasing the ability of L1L2S binding with PDK1. It doesn't have much influence on binding with PDK2 and PDK4, however, it indeed weakens the ability of L1L2S binding with PDK3 and E1p.

2.5 Conclusions

It had been claimed earlier that only the inner lipoyl domain plays an important role in activation of PDKs, while other components such as the outer lipoyl domain, the substrate binding domain and also the third lipoyl domain derived from E3BP are all involved in interaction with PDKs and E1p. We here showed that the substitution of some additional residues such as residue E209 also impairs or enhances the binding ability towards PDKs.

References

- [1] Yeaman, S. J. (1986) The mammalian 2-oxoacid dehydrogenases: a complex family. *Trends Biochem. Sci.* 11, 293-296.
- [2] Randle, P. J. (1983) Mitochondrial 2-oxoacid dehydrogenase complexes of animal tissues. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 302, 47-57.
- [3] Park D, Don A S, Massamiri T, Karwa A, Warner B, MacDonald J, Hemenway C, Naik A, Kuan K, Dilda P, Wong J W.H., Camphausen K, Chinen L, Dyszlewski M, Hogg P. Noninvasive imaging of cell death using an Hsp90 ligand [J]. *Journal of the American Chemical Society*, 2011, 133(9): 2832-2835.
- [4] Donoghue N, Yam P T W, Jiang X M, Hogg P J. Presence of closely spaced protein thiols on the surface of mammalian cells [J]. *Protein Science*, 2000, 9(12): 2436-2445.
- [5] Dilda P J, Decollogne S, Weerakoon L, Norris M D, Haber M, Allen J D, Hogg P J. Optimization of the antitumor efficacy of a synthetic mitochondrial toxin by increasing the residence time in the cytosol [J]. *Journal of medicinal chemistry*, 2009, 52(20): 6209-6216.
- [6] Yamada H, Yamahara A, Yasuda S, Abe M, Oguri K, Fukushima S, Ikeda-Wada S. Dansyl chloride derivatization of methamphetamine: a method with advantages for screening and analysis of methamphetamine in urine[J]. *Journal of analytical toxicology*, 2002, 26(1): 17-22.
- [7] Bartzatt R. Fluorescent labeling of drugs and simple organic compounds containing amine functional groups, utilizing dansyl chloride in Na₂CO₃ buffer [J]. *Journal of pharmacological and toxicological methods*, 2001, 45(3): 247-253.

- [8] Sharma V, Wang Q, Lawrence D S. Peptide-based fluorescent sensors of protein kinase activity: design and applications [J]. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 2008, 1784(1): 94-99.
- [9] Patel M S, Korotchkina L G, Sidhu S. Interaction of E1 and E3 components with the core proteins of the human pyruvate dehydrogenase complex[J]. *Journal of Molecular Catalysis B: Enzymatic*, 2009, 61(1): 2-6.
- [10] Korotchkina L G, Patel M S. Probing the mechanism of inactivation of human pyruvate dehydrogenase by phosphorylation of three sites[J]. *Journal of Biological Chemistry*, 2001, 276(8): 5731-5738.
- [11] Baker J C, Yan X, Peng T, Kasten S, Roche T E. Marked differences between two isoforms of human pyruvate dehydrogenase kinase[J]. *Journal of Biological Chemistry*, 2000, 275(21): 15773-15781.
- [12] Bowker-Kinley M M, Davis I W, Wu P, Harris R A, Popov K M. Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex[J]. *Biochemical Journal*, 1998, 329(1): 191-196.
- [13] Wynn R M, Kato M, Chuang J L, Tso S C, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. *Journal of Biological Chemistry*, 2008, 283(37): 25305-25315.
- [14] Gudi R, Melissa M B K, Kedishvili N Y, Zhao Y, Popov K M. Diversity of the pyruvate dehydrogenase kinase gene family in humans[J]. *Journal of Biological Chemistry*, 1995, 270(48): 28989-28994.
- [15] Stevenson K J, Hale G, Perham R N. Inhibition of pyruvate dehydrogenase multienzyme complex from *Escherichia coli* with mono-and bifunctional arsenoxides[J]. *Biochemistry*, 1978, 17(11): 2189-2192.
- [16] Balakrishnan A, Nemeria N S, Chakraborty S, Kakalis L, Jordan F. Determination of pre-steady-state rate constants on the *Escherichia coli* pyruvate dehydrogenase complex

reveals that loop movement controls the rate-limiting step[J]. *Journal of the American Chemical Society*, 2012, 134(45): 18644-18655.

[17] Schuldiner S, Weil R, Robertson D E, Kaback H R. Microenvironment of the binding site in the lac carrier protein[J]. *Proceedings of the National Academy of Sciences*, 1977, 74(5): 1851-1854.

[18] Wynn R M, Kato M, Chuang J L, Tso S C, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. *Journal of Biological Chemistry*, 2008, 283(37): 25305-25315.

[19] Hiromasa Y, Roche T E. Facilitated interaction between the pyruvate dehydrogenase kinase isoform 2 and the dihydrolipoyl acetyltransferase[J]. *Journal of Biological Chemistry*, 2003, 278(36): 33681-33693.

[20] Crewe C, Schafer C, Lee I, Kinter M, Szweda L I. Regulation of Pyruvate Dehydrogenase Kinase 4 in the Heart through Degradation by the Lon Protease in Response to Mitochondrial Substrate Availability[J]. *Journal of Biological Chemistry*, 2017, 292(1): 305-312.

[21] Tuganova A, Boulatnikov I, Popov K M. Interaction between the individual isoenzymes of pyruvate dehydrogenase kinase and the inner lipoyl-bearing domain of transacetylase component of pyruvate dehydrogenase complex[J]. *Biochemical Journal*, 2002, 366(1): 129-136.

[22] Kato, M., Wynn, R. M., Chuang, J. L., Tso, S.-C., Machius, M., Li, J., and Chuang, D. T. (2008) Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops. *Structure*, **16**, 1849-1859.

CHAPTER 3. Using H/D Exchange Mass Spectrometry to Study the Interaction Patterns between E1p and Four Isoforms of PDKs

3.1 Introduction

The human pyruvate dehydrogenase complex (PDHc) provides the key link between glycolysis and the citric acid cycle. This large complex comprises four principal proteins: The E1 (pyruvate dehydrogenase) component carries out pyruvate decarboxylation and reductive acetylation of the E2 component (dihydrolipoamide acetyl transferase), reducing its covalently attached lipoamide to dihydrolipoamide and producing acetylCoA; the E3 component (dihydrolipoamide dehydrogenase) then reoxidizes dihydrolipoamide to lipoamide with concomitant formation of an equivalent of NADH from NAD^+ ; finally, the E3 binding protein (E3BP), acts as a bridge between the E2 and E3 components. A key to redox chemistry in such complexes is provided by lipoyl domains, two in the E2 domain (L1 or outer, and L2 or inner) and one in E3BP (L3). Primary regulation of the human PDHc is by covalent phosphorylation (turning off complex activity and NADH production) by four pyruvate dehydrogenase kinases (PDK1-4) and the activity could be recovered by two pyruvate dehydrogenase phosphatases (PDP1-2). Regulation of PDHc is achieved by site-specific phosphorylation of the α subunits of the heterotetrameric ($\alpha_2\beta_2$) E1p component at three different sites: Ser264 (site 1), Ser271 (site 2) and Ser203 (site 3). The three sites were phosphorylated *in vivo* to different extents, with maximum phosphorylation of site 1 and lesser phosphorylation of sites 2 and 3. The four PDK isozymes have different activity and phosphorylation rates at each site. At site 1 from fastest to slowest, $\text{PDK2} > \text{PDK4} \approx$

PDK1 > PDK3. For site 2, PDK3 > PDK4 > PDK2 > PDK1. Only PDK1 can phosphorylate site 3 [1]. Dephosphorylation of phosphoryl-E1 by PDP's recovers full PDHc activity.

The structure of PDKs comprises distinct N- and C-terminal domains. The N-terminal domain consists of eight α -helices, of which four form a bundle-like structure, the core. The sequences of the N-terminal domains of the four PDK isoforms are poorly conserved. In contrast, the C-terminal domain containing the phosphoryl transfer catalytic site is highly conserved among the four PDKs. They share four conserved motifs: (i) the N-box (Glu-X-X-Lys-Asn-X-X-X-Ala); (ii) the G1-box (Asp-X-Gly-X-Gly); (iii) the G2-box (Gly-X-Gly-X-Gly); and (iv) the G3-box (Gly-X-Gly-Thr) [2], all four together form a unique ATP-binding fold. This fold includes a common structural element known as the "ATP lid," whose conformational change is coupled to both ATP hydrolysis and protein-protein interactions [3].

Unlike the canonical ATP binding fold (in a cleft between a smaller N-terminal lobe of β -sheets and a larger C-terminal lobe of α -helices [4] shared among most of the hundreds of protein kinases in mammals, the ATP-binding fold in its K domain (ATP binds to one catalytic domain consisting of tightly aggregated alternating β -sheets and α -helices), indicates unique mechanisms for PDK activation and phosphorylation.

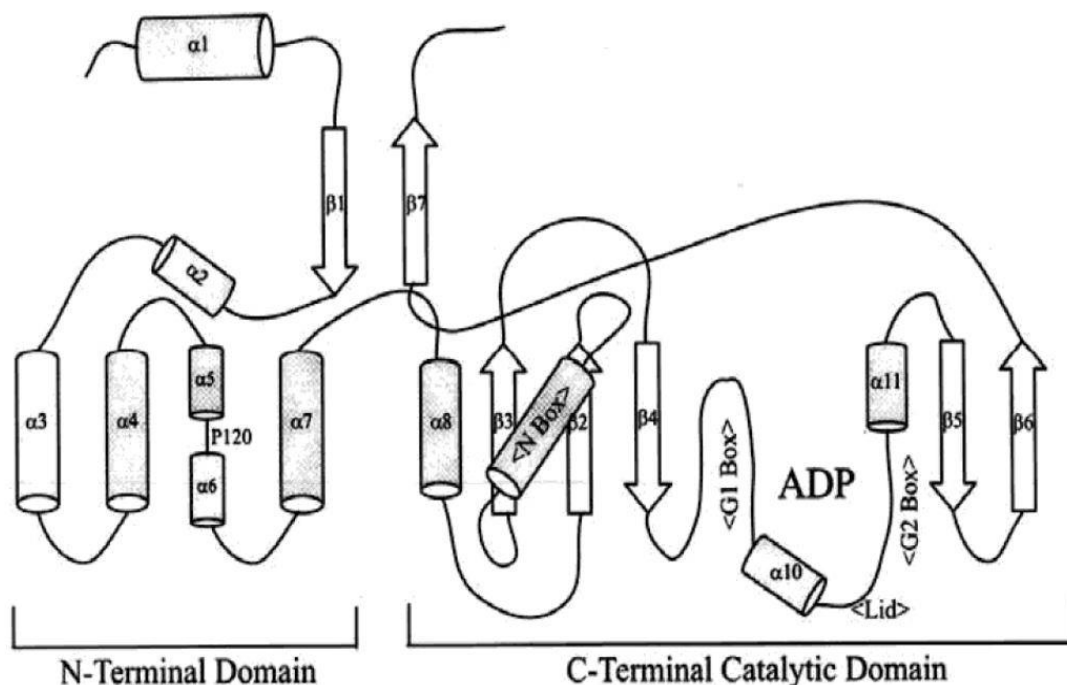


Figure 3.1: Topology of PDK2 [3].

It has been widely believed that lipoyl domains play a pivotal role in modulation of PDK activities, though different PDKs may have different preferences towards the lipoyl domains originating from E2 and E3BP. The structure of PDK3 bound to L2 domain showed that binding to the inner lipoyl domain induced a conformational change of PDK3, ordering the C-terminal tails of the dimer to form a ‘crossover’, thereby enlarge the active site cleft to facilitate the release of ADP [5]. As a result, most of the studies were focused on the interaction between E2·E3BP-derived proteins and PDKs. To date, only crystal structures of individual PDKs and some of their sub-complexes with L2 are available [5-9]. No structure with any E1p component is available.

In fact, PDKs can phosphorylate E1p without any source of E2·E3BP-derived proteins, at varying rates, some only slowly, others needing no activation from E2·E3BP. It is also

known that each of the four isozymes of PDKs has a preference to certain phosphorylation sites, however, the structural origin of this specific recognition between the PDK and the phosphorylation sites of E1p is still not understood.

Here, we have carried out a series of experiments using the hydrogen/deuterium exchange detected by mass spectrometry (HDX-MS) method to study the direct interactions between E1p and PDKs, trying to understand the differences among the four PDKs when forming complexes with E1p as well as to differentiate the four PDK isoforms by their unique pattern and provide some guides in a rational drug design of PDK isoform-specific inhibitors.

3.2 Materials and Methods

Deuterium oxide (D_2O) was from Cambridge Isotope Laboratories. All other fine chemicals were from Sigma-Aldrich. The E1p and the four PDK isoforms were purified and characterized as described in the Supplementary section.

3.2.1 Sample preparation for HDX-MS.

Prior to H/D exchange, the E1p and PDK were exchanged into buffer containing 50 mM KH_2PO_4 (pH=7.5), 50 mM KCl, 0.5 mM ThDP, and 1 mM $MgCl_2$. The E1p (80 μM) was incubated alone or with equal aubunit concentration of PDK. containing 5 mM ATP at 25 °C for 30 min prior to H/D exchange. The PDK isoform (80 μM) by itself or in the presence of 5 mM ATP were prepared as control. The HDX experiments were initiated by mixing 15 μl of the protein samples with 285 μl of D_2O buffer (E1p exchange buffer prepared in 99.9% D_2O) to yield a final concentration of 95 % D_2O at pH 7.5. The H/D exchange reactions were incubated at 25 °C for 30 s, 1, 3, 5, 10 and 30 min, and then

quenched by rapidly mixing with an equal volume 40 μL of ice-cold quench buffer (99% trifluoroacetic acid, 3 M guanidine hydrochloride, pH 1.6) to reduce the final sample pH to 2.5. The samples were immediately frozen in liquid nitrogen and stored at $-80\text{ }^{\circ}\text{C}$ until 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.

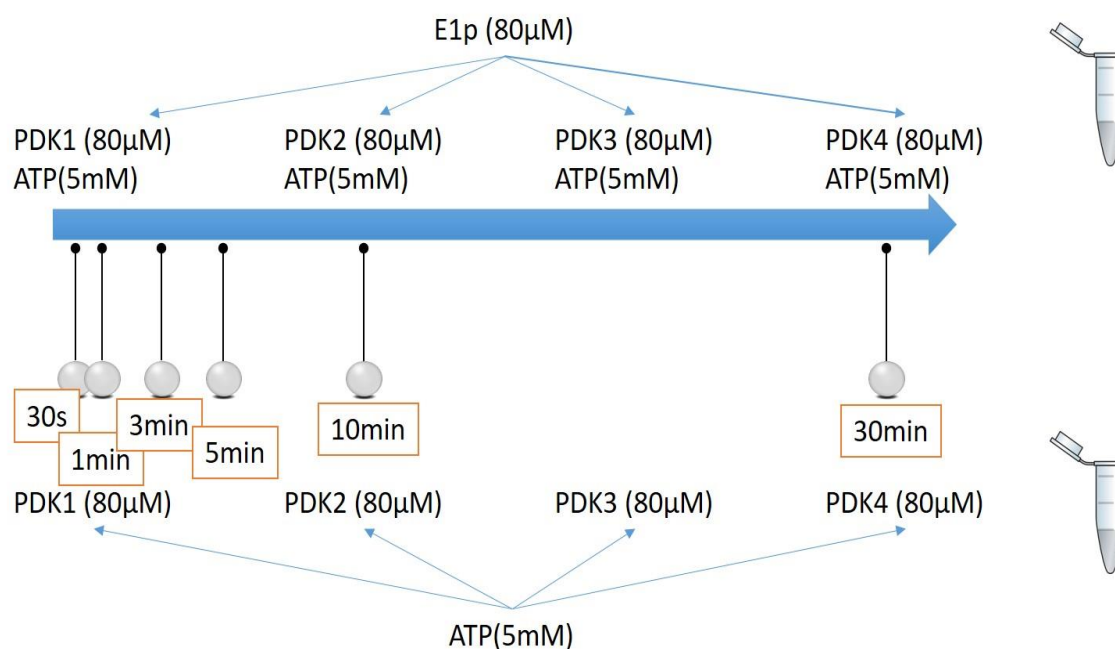


Figure 3.2: Experimental design. This series of experiments used free E1p and free PDK isoform as controls. Then the change of E1p and PDK upon complex formation was monitored in the presence of ATP.

3.2.2 LC-MS method and data processing

Frozen deuterated samples were quickly thawed and loaded with an ice-cold syringe into a 20 μL sample loop inside the refrigeration system. The protein sample (40 pmol E1p)

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 over a 15 min period 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 un-deuterated 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 2.0 ppm.

3.3 Results and Discussion

3.3.1 Overview of HDX patterns on E1p

This is the first application of HDX-MS to study the direct interaction between E1p and all four PDK isoforms. The time dependence of hydrogen-deuterium exchange of the backbone amide protons of E1p and PDKs was studied over a 30 min time course (30 s, 1, 3, 5, 10 and 30 min). On-line digestion by pepsin followed by LC-MS analysis under

the selected HDX conditions yielded 24 peptides from E1p- α with 78% sequence coverage and 18 peptides from E1p- β with 89% sequence coverage, many of which were partially overlapping (Table 3.1 and 3.2)

The deuterium uptake percentage changes over the 30 min's time course and the average deuterium incorporation level of E1p at 30 min is illustrated in Fig. 3.3

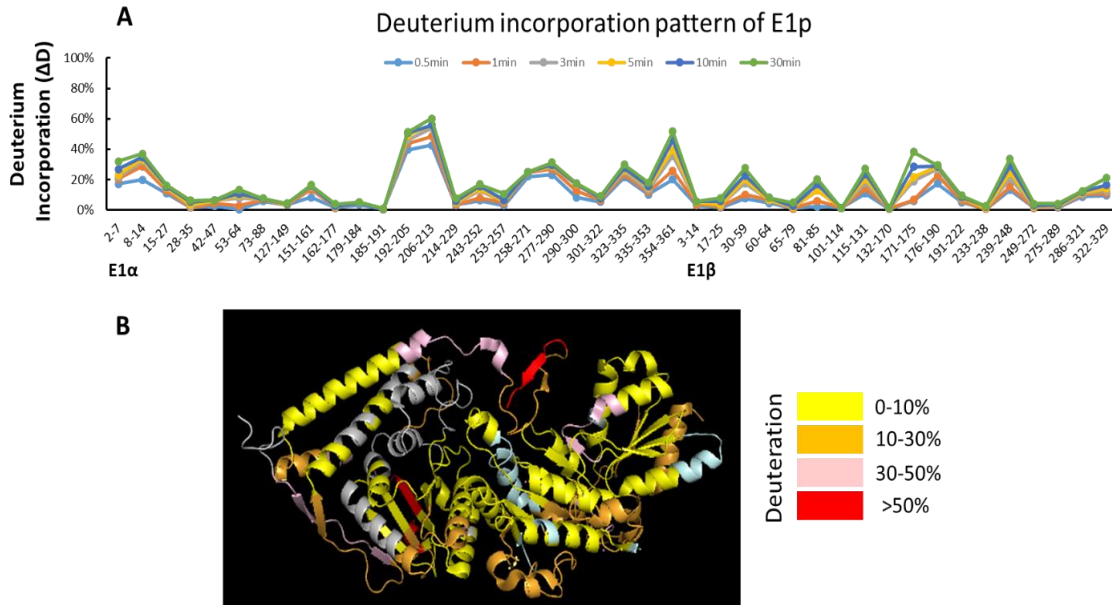


Figure 3.3: Deuterium incorporation pattern of E1p. **A.** The D uptake percentage changes of E1p over the 30 minutes. The percentage of D incorporation (without back exchange correction) of each peptide was calculated from the equation $\Delta D\% = \Delta D / (\max D \times 1.0063 \times 0.948) \times 100\%$, where 1.0063 is the atomic mass difference between D and H and 0.948 represents the fractional D_2O content of the labeling reaction mixture. **B.** The average percentage of D incorporation at the 30 min time point is mapped according to color (see legend) onto the crystal structure of *E1p*. The peptides that were not identified in E1p- α and E1p- β are shown in grey and light blue respectively.

The following information was revealed from an analysis of the HDX-MS data.

(1) From the peptic digestion, two peptides from E1p- α , ²⁵⁸YRYHGHSMKSDPGVS²⁷¹ and ¹⁹²ICENNRYGMGTSVE²⁰⁵, containing the three phosphorylation sites (Site 1-Ser264; Site 2-Ser271 and Site 3-Ser203), were identified.

(2) The overall rate of deuterium incorporation was low on this enzyme; many regions underwent less than 30% deuteration after a 30 min exchange. The deuterium uptake level of E1p- α is largely divided, with some regions showing deuterium uptake level below 10% while others higher than 30%. E1p- β exhibits relatively stable deuterium uptake.

(3) The deuterium incorporation level for peptides 127-149, 151-161 in E1p- α and 60-64, 65-79 in E1p- β are low. On the basis of the crystal structure (Ciszak et al.), a probable explanation is that in order to form the $\alpha_2\beta_2$ heterotetramer, an alpha subunit and beta subunit associate through predominantly hydrophobic contacts by permitting a pair of helices related to ThDP binding, 140-155 of the diphosphate binding domain from E1p-alpha and 59-72 of the pyrimidine-binding domain from E1p-beta, to pack together tightly.

(4) In contrast, peptides 192-205, 258-271 and 277-290 experienced relatively high deuterium incorporation. Peptides 258-271 and 277-290 are located on the conserved phosphorylation loop A (E1p- α 259-282), which forms one wall of the E1p active-site channel and helps anchor the ThDP to the active sites. On the other hand, peptide 192-

205 is located on the adjacent Ph-loop B (E1p- α 198-205) that provides coordination to a Mg^{2+} ion chelated by the diphosphate group of ThDP.

(5) In the wild-type E1p structure, Ser264 and Ser266 in the Ph-loop A form a hydrogen-bond network with two neighboring water molecules to form an ordered loop. And E1p- β' Tyr33 also participates in this H-bond network, as reflected by a 27.5% deuterium uptake level for peptide E1p- β 30-59.

3.3.2 Effects of the PDK isoforms on the E1p structure

Previous studies from this laboratory had focused on PDK activation by E2·E3BP or E2·E3BP-derived proteins, since it was believed that L2 of the E2·E3BP core is a docking site for association of PDKs with PDC, in turn responsible for activating the kinase and providing regulation. But there were also several reports showing that some PDKs are capable of phosphorylating free E1p without any help from E2·E3BP. Here, we use HDX-MS to study the influence of different PDKs exerted on E1p in the absence of E2·E3BP.

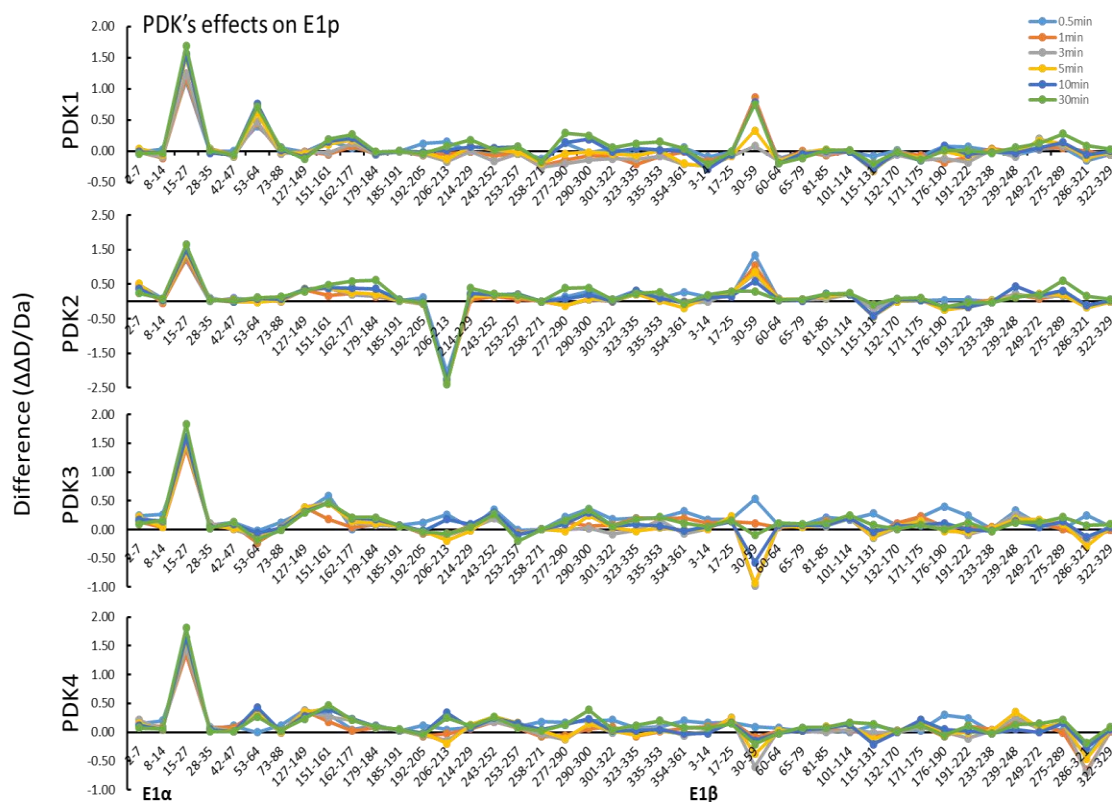


Figure 3.4: Effect of complexation with PDK isoforms on deuterium exchange in E1p.

A difference plot showing the changes in deuterium incorporation ($\Delta\Delta D$, y axis, deuterons exchanged of E1p in the presence of the corresponding PDK and ATP minus deuterons exchanged in the free form of E1p).

Upon incubation of E1p with PDKs and ATP, the N-terminal region of E1p- α (residues 15-27) has undergone a conformational change irrespective of the PDK selected. With PDK1, the most significant HDX-MS changes were observed in the N-terminal region of both E1p- α and E1p- β , which experienced a $\Delta\Delta D$ change around 1.7 Da for peptide E1p- α 15-27, a change of 0.7 Da for peptide E1p- α 53-64 and of 0.75 Da for peptide E1p- β 30-59. We could not identify any potential binding sites for PDK1 on E1p with the

current sequence coverage. But, it's clear that PDK1 induces a conformational change on E1p, especially on the N-terminal regions of both the E1p- α and E1p- β subunits.

The interaction pattern between E1p and PDK2 is slightly different from that with PDK1. Except for the higher deuterium incorporation level of peptide E1p- α 15-27 ($\Delta\Delta D = 1.65$ Da), there is also a potential binding site of PDK2 on peptide E1p- α 206-213 ($\Delta\Delta D = -2.39$ Da), which is adjacent to phosphorylation site 3. The effects of PDK2 on the N terminal of E1p- α subunit is less significant compared to that of PDK1, since the deuterium uptake level of peptide E1p- α 53-64 doesn't change in the presence of PDK2 while the same peptide displays a $\Delta\Delta D$ increase of 0.75 Da.

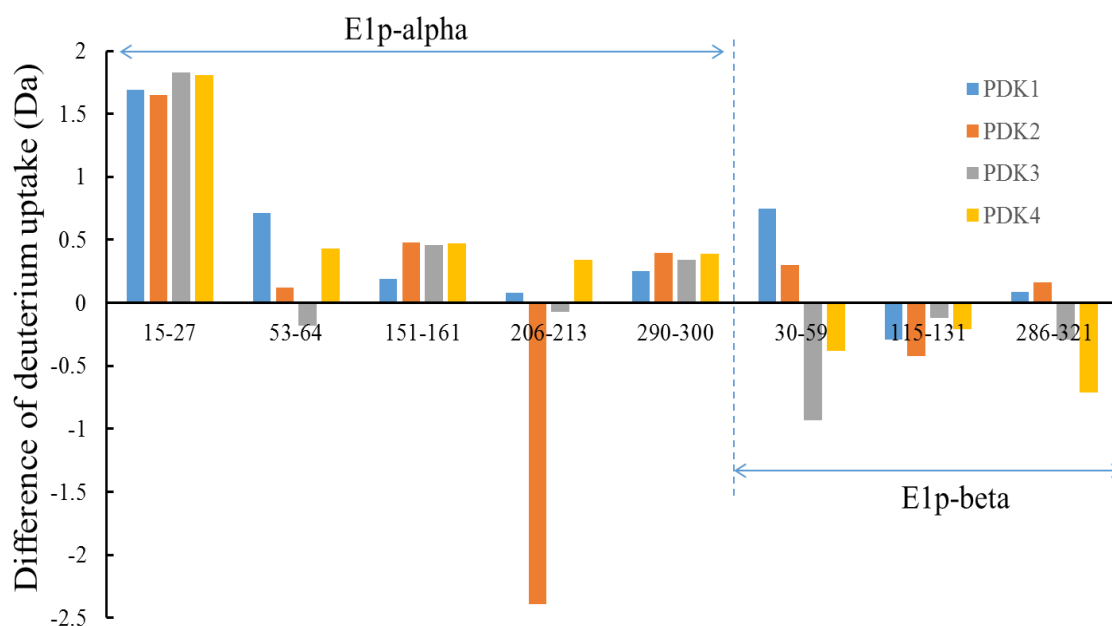


Figure 3.5: Effects of four PDK isozymes on H/D exchange in E1p. Selected peptides from E1p that show significant deuteriation incorporation change upon complexation with PDKs. The interaction patterns of E1p with PDK3 and E1p with PDK4 are alike. Both of

them exhibit a large positive $\Delta\Delta D$ change for peptide E1p- α 15-27 with 1.83 Da and 1.81 Da respectively. E1p may undergo a small conformational change in the region of E1p- α 127-161 when incubated with PDK3 or PDK4. The interaction that occurs at the N-terminal region of E1p- β might be transient, since the deuterium incorporation level of peptide E1p- β 30-59 decreased in the course of 30 min. There are some minor differences in the effects exerted by PDK3 and PDK4. In the presence of PDK4, the peptide of E1p- α 53-64 experienced a $\Delta\Delta D$ change of 0.4 Da, while in the presence of PDK3, this peptide showed a slight decrease with $\Delta\Delta D$ of -0.18 Da. In addition, the PDK4 also induced some transient effects on the C-terminal region of E1p- β 268-321.

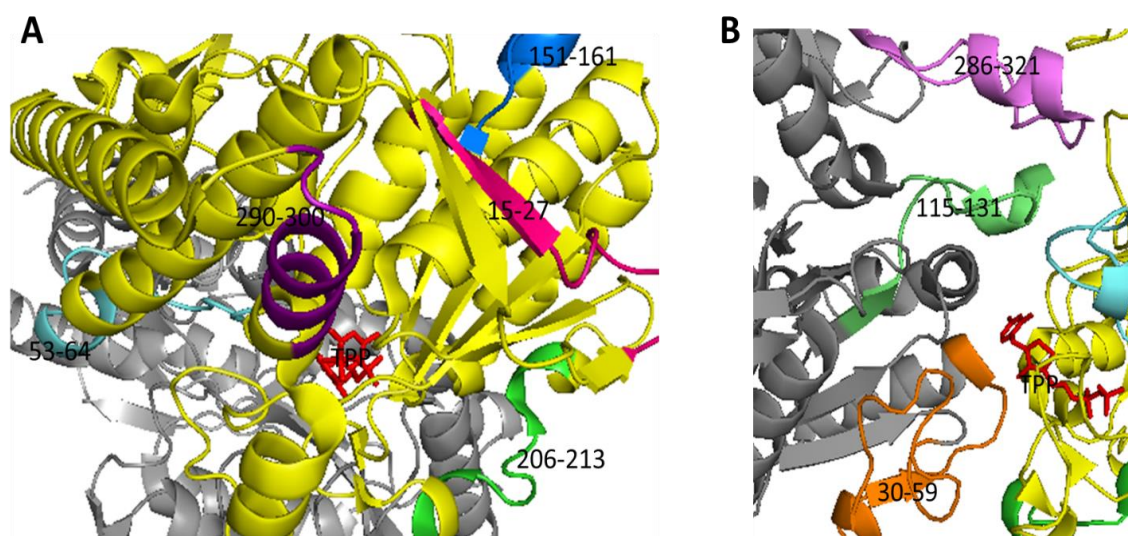


Figure 3.6: Peptides of human E1p involved in conformational change upon complexation with PDKs. A. Five peptides from the E1p- α subunit that show significant $\Delta\Delta D$ change upon complexation with PDKs are highlighted in colors: 15-27 (pink); 53-64 (cyan); 151-161 (blue); 206-213 (green); 290-300 (purple). B. Three peptides from the E1p- β subunit that show significant $\Delta\Delta D$ change upon complexation with PDKs are highlighted in colors: 30-59 (orange); 115-131 (lime); 286-321 (magenta).

The binding with PDKs causes significant conformational change of the loops at the entrance of the ThDP active center, which explains the significant $\Delta\Delta D$ in two peptides (E1p- α 53-64 and E1p- β 30-59, lying at the entrance). Although the peptide E1p- β 30-59 has no direct contact with the ThDP cofactor, it is effectively connected to the active center regions of E1p- α , by hydrogen bonds of $\alpha Q172$ - $\beta E59$, $\alpha Q172$ - $\beta I57$, and a salt bridge of $\alpha R206$ - $\beta D54$, by which it could induce conformational change of peptide E1p- α 206-213, once it changes itself. All of these conformational changes may induce the E1p- α subunit

to expose the phosphorylation sites, especially site 1 (the major phosphorylation site), which is located at the substrate channel, to solvent and facilitate the phosphorylation reaction.

3.3.3 Effects of PDKs on E1p phosphorylation

That pyruvate dehydrogenase kinase (PDK) can inactivate PDC by phosphorylation was first discovered in the late 1960s. (Linn et al.,) The reversible phosphorylation occurs on three specific serine residues on E1p- α subunit at site 1, Ser264; site 2, Ser271; and site 3, Ser203. The four PDK isoforms have different specificity and different rates toward the three phosphorylation sites. All four PDKs phosphorylate site 1 and site 2, while only PDK1 could phosphorylate site 3.

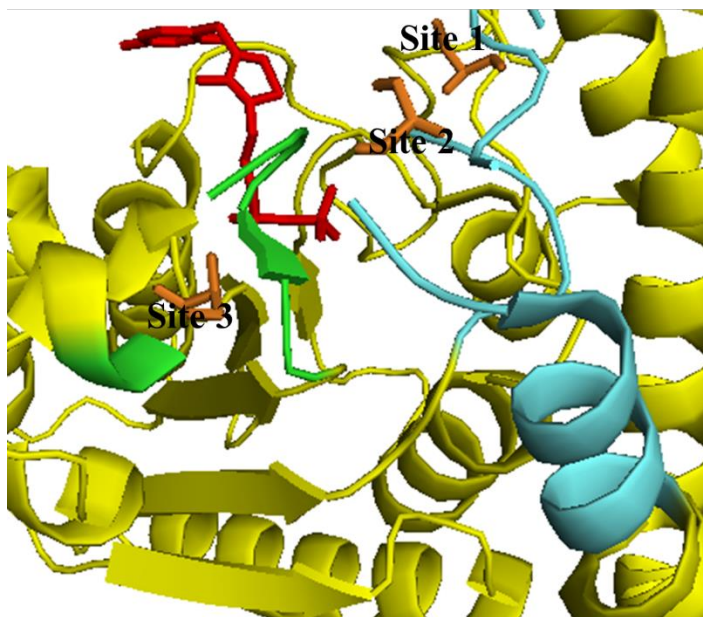


Figure 3.7: Phosphorylation sites on the E1p- α subunit. Site 1 (Ser264) and Site 2 (Ser271) are located on Ph-loop A (259-282), colored in *cyan*. Site 3 (Ser203) is located on Ph-loop B (198-205), colored in *lime*.

Here, we used MS to detect the PO_4^{3-} incorporation by analyzing the two peptides: $^{258}\text{YRYHGHSMKSDPGVS}^{271}$ and $^{192}\text{ICENNRYGMGTSVE}^{205}$. Since the peptide $^{258}\text{YRYHGHSMKSDPGVS}^{271}$ contains both sites 1 and 2, we cannot differentiate these two sites without further fragmentation. The ratio is calculated from (peak intensities of phosphorylated peptide) / (peak intensities of phosphorylated peptide + peak intensities of original peptide).

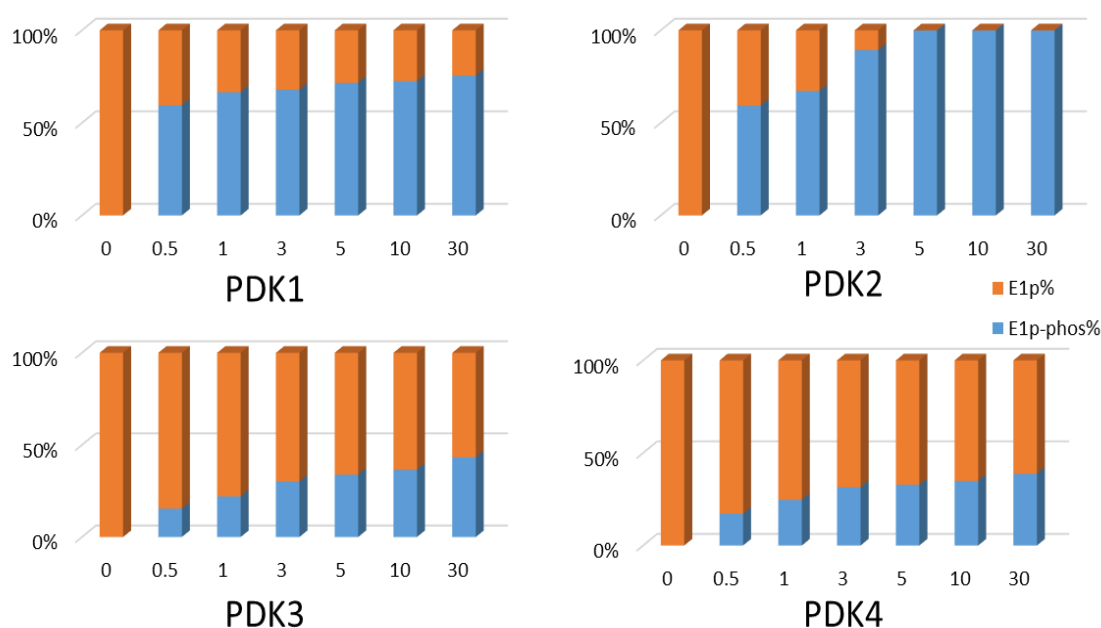


Figure 3.8: PO_4^{3-} incorporation ratios on peptide $^{258}\text{YRYHGHSMKSDPGVS}^{271}$ during 30 min reaction. The *blue* represents the fraction of phosphorylated peptide and *orange* represents the fraction of original peptide remaining.

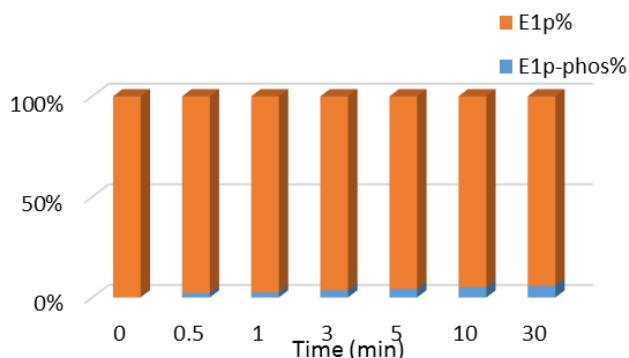


Figure 3.9: PO_4^{3-} incorporation on peptide $^{192}\text{ICENNRYGMGTSVE}^{205}$ during 30 min reaction with PDK1. The *blue* represents the fraction of phosphorylated peptide and *orange* represents the fraction of original peptide remaining.

Most of these MS results are in agreement with the activity data from the Patel group [10]. PDK2 has the highest activity in the absence of E2·E3BP, which can almost fully phosphorylate E1p within 5 min. Even though PDK1 can phosphorylate all three sites, the activity towards site 1 is much higher than towards site 3, in accord with a small percentage (5% percent) phosphorylation at site 3 after 30 min. Here, we reported that without E2·E3BP, PDK3 and PDK4 can still phosphorylate E1p to an extent of approximately 40%. While the activity measured by Patel's group for PDK3 was 12-times lower than for PDK4. Recently, we also reported PDK activities assessed by measuring the overall PDHc activity (NADH production). It was shown that 210 min incubation of E1p with PDK1 led to retention of 90% of the PDC activity remaining, while 90 min incubation with PDK2 led to retention of only 15% activity remaining. PDK3 and PDK4 by themselves exhibited little inactivation of PDC: after 60 min incubation, 95% of PDC activity was retained [11].

3.3.4 Effects of E1p on PDKs

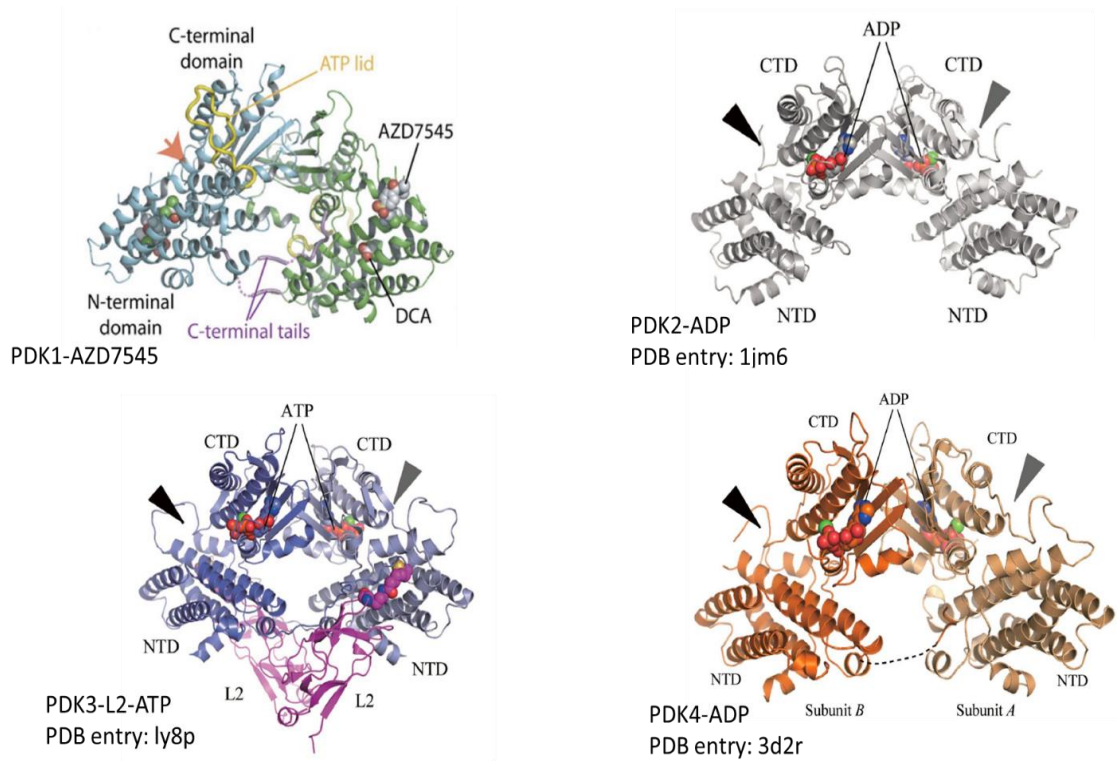


Figure 3.10: Crystal structure of four PDK isozymes [5-9].

PDK1: The sequence coverage for PDK1 is as high as 99.4% (see digestion map in Table 3.6).

The direct interaction between E1p and PDK1 didn't induce global conformational changes in PDK1. However, there are some regions that experienced significant deuterium uptake decrease (peptides spanning residues 66-77, 88-101 and 395-417), and a small increased deuterium uptake in the G1 box region (peptide spanning residues 311-333).

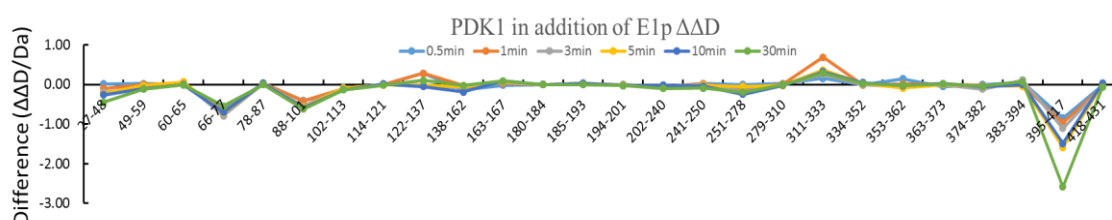


Figure 3.11: Deuteration level change in PDK1 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

It appears that the PDK1 interacts with E1p through its $\alpha 2$ to $\alpha 4$ helices ($\alpha 2$: 59-67; $\alpha 3$: 72-94; $\alpha 4$: 99-102), and at the same time, reorganizes the C-terminal tails. As suggested by the crystal structure of the PDK3-L2 complex, the fully ordered “crosstail” configuration could help freeze the open conformation of PDK3, in which the active cleft will be wider, facilitating the exchange of ATP/ADP [6]. The interaction between PDK1 and E1p may result in the same change by forcing the active cleft to become wider by ordering the C-terminal tails, leading to an increase in deuterium uptake in peptide 311-333, while decreasing the deuterium uptake in peptide 396-417.

PDK2: The sequence coverage of PDK2 is 99.3% (see digestion map in Table 3.7).

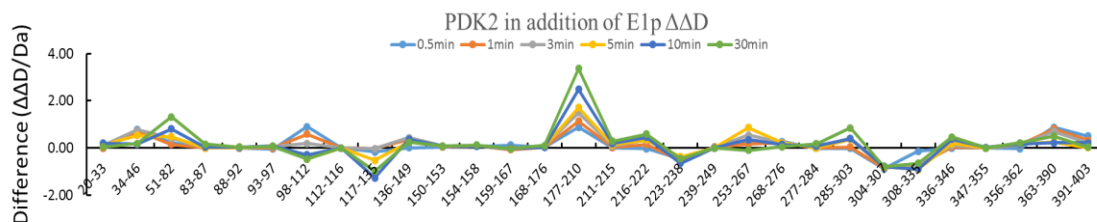


Figure 3.12: Deuteration level change in PDK2 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

The interaction between E1p and PDK2 induces a relatively larger conformational change on PDK2. According to the crystal structure of PDK2, peptides 168-210 together with 117-136 and 308-335 form a cleft between the catalytic domain and the regulatory domain with the ADP/ATP cofactor on one end. This time, it's more likely that PDK2 binds with E1p through this cleft, since the gate helix (residues 117-136), which comprises helices $\alpha 6$ and $\alpha 7$ ($\alpha 6$: residues 106-122; $\alpha 7$: residues 126-141), peptide 177-210, which includes β strand 1 and the beginning of the $\alpha 10$ helix, all experienced a significant conformational change. In addition, the N-terminal region and peptides containing the ATP lid and G-box also experience some deuterium uptake change during complexation.

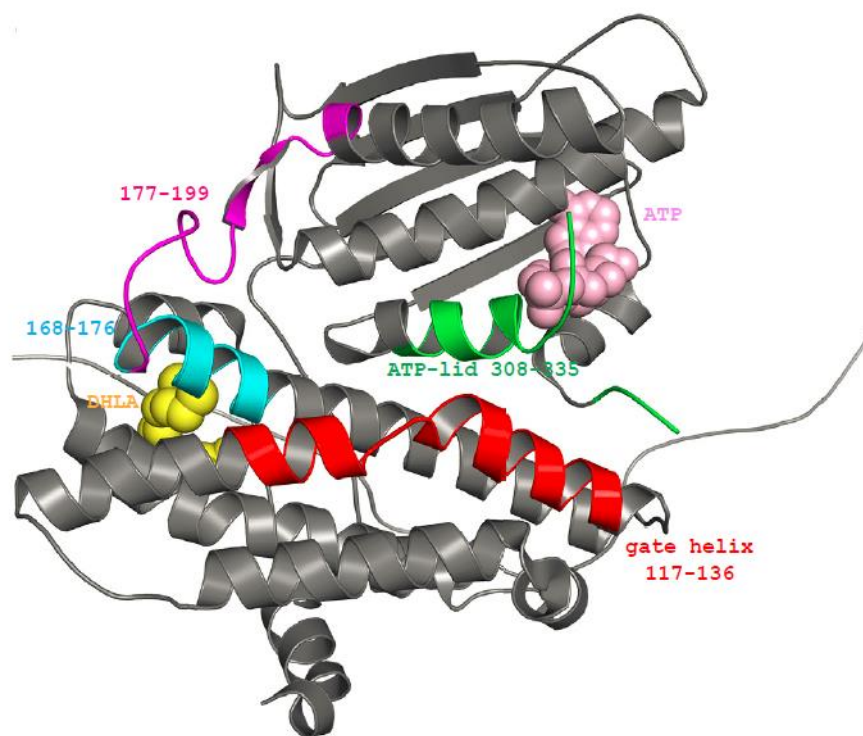


Figure 3.13: Residues of PDK2 displaying significant $\Delta\Delta D$ on binding to E1p in HDX-MS analysis. Includes gate helix (residues 117-136, in *red*), ATP-lid (residues 308-335, in *green*), and another region whose function remains unknown (residues 168-199, in *cyan* and *magenta*).

PDK3: The sequence coverage of PDK3 is 65.5% (see digestion map in Table 3.8).

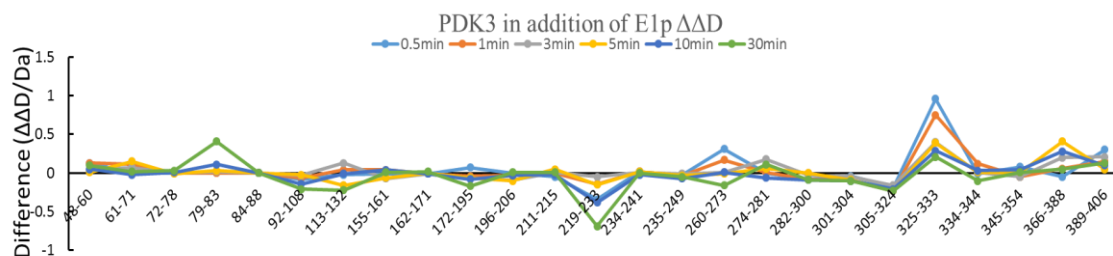


Figure 3.14: Deuteration level change in PDK3 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

The binding mode between PDK3 and E1p is similar to that between PDK2 and E1p. The corresponding region 172-195, gate helix 113-132 together with G2 box 325-333 form a cleft involved in binding to E1p. The interactions between PDK3 and E1p shield the peptide 219-233 from solvent, which includes a β sheet and a hinge. It also induces the conformational change on the entire ATP binding fold. Vassilyev's group proposed a model of the E1p-PDK3 interaction based on computational modelling, which suggested that the peptide bearing phosphorylation sites was docked into the cleft of PDK3. The model also suggested that the peptide underwent a significant conformational change from a loop to a helix during the complexation [8]. However, the HDX-MS data showed that the peptides containing the phosphorylation sites on E1p indeed undergo some deuteration level change upon binding with PDK3, the change is not statistically significant to support the hypothesis that it experienced a conformational change from a loop to a helix.

PDK4: The sequence coverage of PDK4 is 70.6% (see digestion map in Table 3.9).

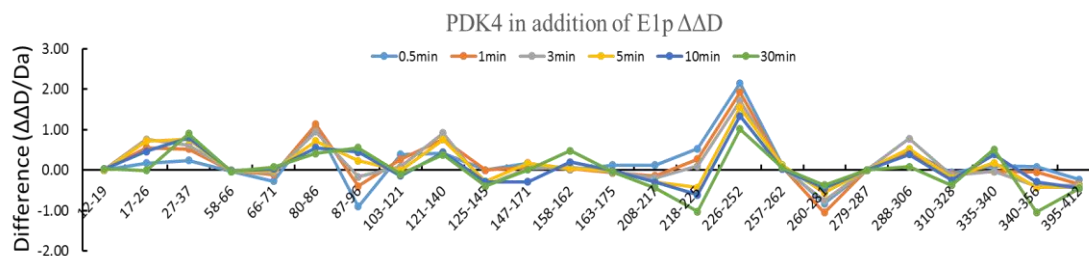


Figure 3.15: Deuteration level change in PDK4 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

Upon binding of E1p to PDK4, many regions on PDK4 were observed to undergo change in deuterium uptake. With the current sequence coverage, we couldn't identify the likely interaction site on PDK4. The peptides 226-252 that experience the most significant increase in deuterium incorporation contain β strand 2, β strand 3 and part of the α 11 helix, which suggest that these regions become more solvent accessible in the PDK4-E1p complex compared with PDK4 alone.

However, all significant decreases in deuterium uptake were only observed at or before the 3 min H/D exchange time points. These weak protections or the weak lock effects on conformations suggested a loose association between E1p and PDK4.

The crystal structure of PDK4 shows that PDK4 adopts a unique conformation which is intrinsically open without L2 binding. The wider cleft created by the open conformation facilitates recruitment of ATP and release of ADP [9]. This might be the reason why many

more peptides from PDK4 display a change in deuterium incorporation during ATP binding, since the structure is more flexible compared to the other three PDK isozymes.

3.3.5 Ligands effects on four isoforms of PDKs

Pyruvate dehydrogenase kinases act to inactivate pyruvate dehydrogenase complex (PDHc) by phosphorylating it using ATP as phosphoryl donor. The highly conserved sequences among the four PDKs contain the phosphoryl transfer catalytic site, which form a unique ATP-binding fold. This fold includes a common structural element known as the “ATP lid,” whose conformational change is coupled to both ATP hydrolysis and protein-protein interactions.

It is already known that higher ratio of ADP/ATP together with high concentration of pyruvate would attenuate the activities of PDKs. Some crystal structures also established DCA (pyruvate analog)/pyruvate binding sites in the N-terminal domain.

Fluorescence quenching experiments done by Roche group together with the mutation of Trp383 on ATP/ADP binding site of PDK2 even confirms that the interaction of pyruvate or DCA would have pronounced effects on the ATP/ADP binding region. [12]

Here, we used the HDX-MS method to study the effects of ATP and pyruvate on PDKs, from which we hope that it could help us understand the mechanism of inactivation of PDKs.

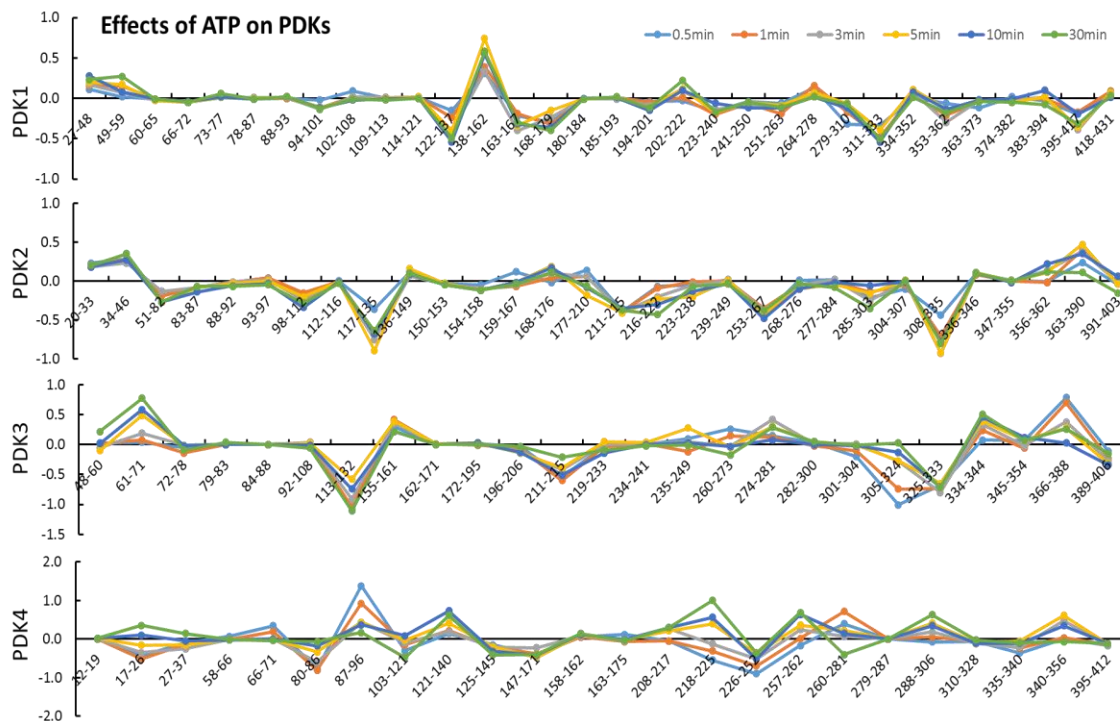


Figure 3.16: Difference plot showing the changes in deuterium incorporation ($\Delta\Delta D$, y axis, deuterons exchanged of corresponding PDKs in the presence of ATP minus deuterons exchanged in the free form of corresponding PDKs).

In the PDKs, the conserved G1-box is involved in binding the adenine moiety of the nucleotide. In PDK1, a 0.3 Da decrease in deuterium uptake was observed in the peptide sequence $^{311}\text{TVKMS}\underline{\text{DRGGGVPLRKIDRLFN}}^{331}$ containing the conserved G1-box motif (DXGXG) upon ATP binding (residues from the G1-box are underlined). A 0.3 Da decrease in ΔD was found in the corresponding $^{285}\text{SIKMS}\underline{\text{DRGGGVPLRKIERL}}^{303}$ peptide in PDK2. Peptide $^{282}\text{SIKISDL}\underline{\text{GGGVPLRKIDRL}}^{300}$ in PDK3 didn't show significant ΔD change, while the peptide in PDK4 $^{288}\text{TIKIS}\underline{\text{DRGGGVPLRKIDRL}}^{306}$ showed a 0.6 increase in ΔD .

The G2-box is responsible for stabilization of phosphate groups of the nucleotide.

Peptides in PDK2 and PDK3 containing the central Gly residues of the G2-box motif (GXGXG) displayed reduced ΔD values of 0.8 Da and 1.0 Da respectively.

In addition to that, gate helix also undergoes some conformational changes. PDK2 (residues 117-135), PDK3 (residues 113-132) experienced a decrease in ΔD , while PDK1(residue 138-162) and PDK4 (121-140) experienced an increase in ΔD .

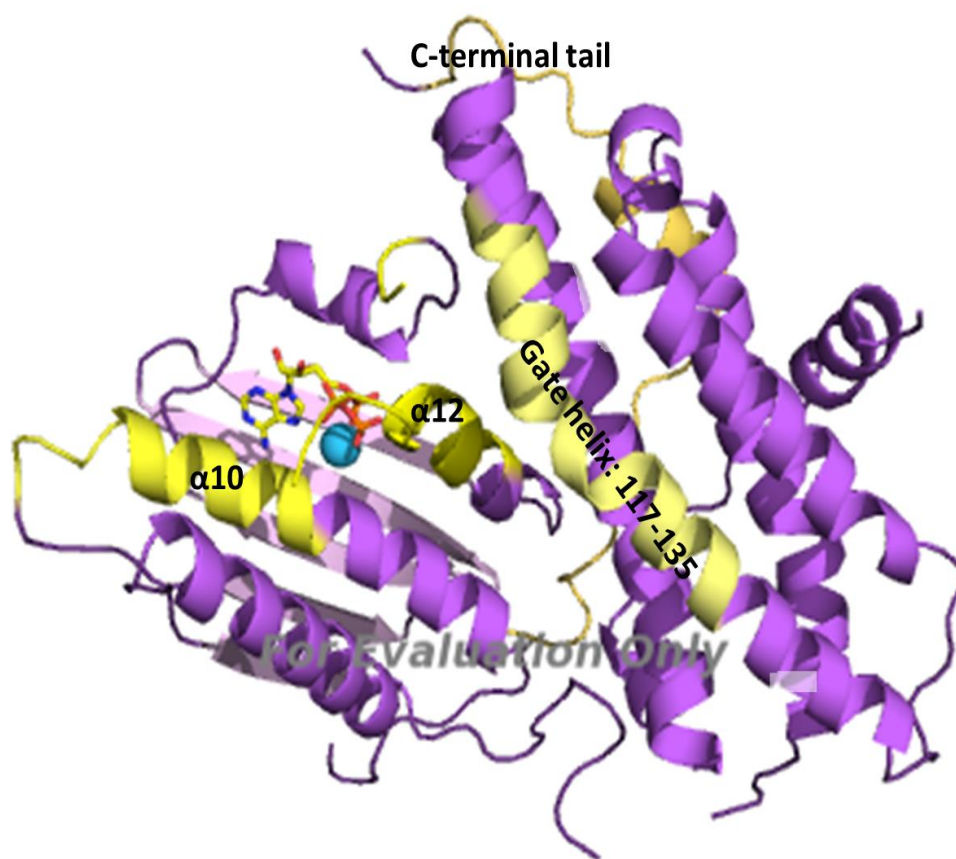


Figure 3.17: Crystal structure of human PDK2 containing ATP. The chain is colored by purple. Residues experienced significant deuteration change is colored by yellow. PDB: 2bu2

The structure of PDK2 shows that the ATP binding pocket is located between the central β -sheet and adjacent layer of helices. Based on the HDX-MS data, the binding of ATP will simultaneously induce the conformational changes on the gate helix and reorder the C-terminal tail.

Additional changes were also observed in other peptides, which are summarized in Table 3.1.

Table 3.1: Deuteration level change of selected peptides on addition of ATP to the four PDK isozymes based on sequence alignment.

	Peptide	Sequences	$\Delta\Delta D$ (Da)
PDK1	88-101	ANIMKEISLLPDNL	-0.11
PDK2	51-82	TFLRQELPVRLANIMKEINLLPDRVLSTPSVQ	-0.3
PDK3	61-71	MREVNLLPDNL	0.7
PDK4	Missing	Missing	Missing
PDK1	122-137	LDFKDKSAEDAKTIYE	-0.5
PDK2	98-112	FLDKDPEDHRTLSQF	-0.3
PDK3	92-108	LEYENKSPEDPQVLDNF	-0.05
PDK4	103-121	HEKSPEDQKVLSDFVDTLV	-0.4
PDK1	138-162	FTD TVIRIRNRHNDVIPTMAQGVTE	0.55
PDK2	117-135	VTIRNRHNDVVPTMAQGV	-0.7
PDK3	113-132	IKVRNRHNDVVPTMAQGVIE	-1.1
PDK4	121-140	VKVRNRHNVVPTMAQGILE	0.6
PDK1	279-310	FKNAMRATMEHHADKGVYPPIQVHVTLG EEDL	-0.2
PDK2	253-267	FKNAMRATVESHESS	-0.4
PDK3	Missing	Missing	Missing
PDK4	260-281	MRATVEHQENRPFLTPVEATVV	0.7
PDK1	311-333	TVKM SDRGGGVPLRKIDRLFN YM	-0.5
PDK2	285-303	SIKM SDRGGGVPLRKIERL	-0.3
PDK3	282-300	SIKISDLGGGVPLRKIDRL	0.05
PDK4	288-306	TIKISDRGGGVPLRKIDRL	0.6
PDK1	334-352	YSTAPRPRVETSRAPLAG	0.01
PDK2	308-335	YSTAPTQPGTGGTPLAGFGYGLPISRL	-0.8
PDK3	305-324	YSTAPRPSLEPTRAAPLAGF	-1.0
PDK4	310-328	TYSTAPTVM DNSRNAPLA	-0.05
PDK1	395-417	IERLPVYNKAAWKHYRTNHEADD	-0.3
PDK2	363-390	LSTD SVERLPVYNKSAWRHYQTIQEAGD	0.4
PDK3	366-388	FERLPVFNKSAWRHYKTTPEADD	0.8
PDK4	Missing	Missing	Missing

From Table 3.1, it can be seen that even peptides that have the same functions among PDKs, on addition of the same ligand, such as ATP, will still have different effects on them.

According to the HDX-MS data, binding of pyruvate to PDK1 will induce a large conformational change especially on the N-terminus and nucleotide-binding site located at the C-terminus of PDK1. The peptides that show significant deuteration level increase are listed below.

Table 3.2: Peptides that show significant deuteration level change in PDK1 on addition of pyruvate.

Peptide	Sequence	$\Delta\Delta D$ (Da)
27-48	ASDSASGSGPASESGVPGQVDF	1.2
49-59	YARFSPSPLSM	2.2
122-137	LDFKDKSAEDAKTIYE	0.6
163-167	YKESF	0.7
251-278	EELNAKSPGQPIQVVVV	0.6
334-352	YSTAPRPRVETSRAVPLAG	0.7
363-373	YAQYFQGDLKL	0.8
395-417	IERLPVYNKAAWKHYRTNHEADD	1.4

There were no significant negative $\Delta\Delta D$ changes observed on peptides that are related to interaction with pyruvate. The reason could be that the DCA-binding site is located inside the helix bundle in the N-terminal domain of PDK1 and is completely shielded from solvent in the apo-PDK1 structure. As a result, the pyruvate entered into a pocket that originally has low solvent accessibility, in other words, it will not induce much protective effect during hydrogen deuterium exchange.

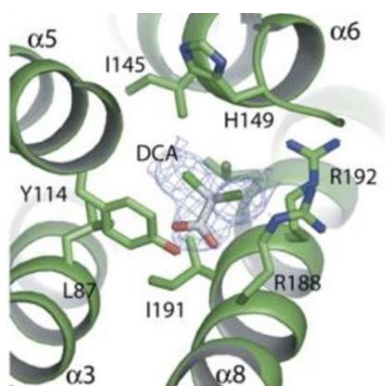


Figure 3.18: Crystal structure of PDK2 in the presence of DCA. [6]

The effects of pyruvate on PDK1 are the opposite from those exerted by ATP. The interaction of ATP with PDK1 will shield most of the peptides on the C-terminal part, which is responsible for ATP binding; in contrast the interaction between pyruvate and PDK1 will convert those PDK1 regions to a more solvent accessible conformation. This result is in consistent with what Roche and coworkers had concluded from fluorescence quenching experiments. Addition of pyruvate into buffer containing PDK2 could quench the fluorescence of Trp383 (the equivalent Trp415 in PDK1), for which the intensity would be lower if the microenvironment surrounding the tryptophan is more hydrophilic.

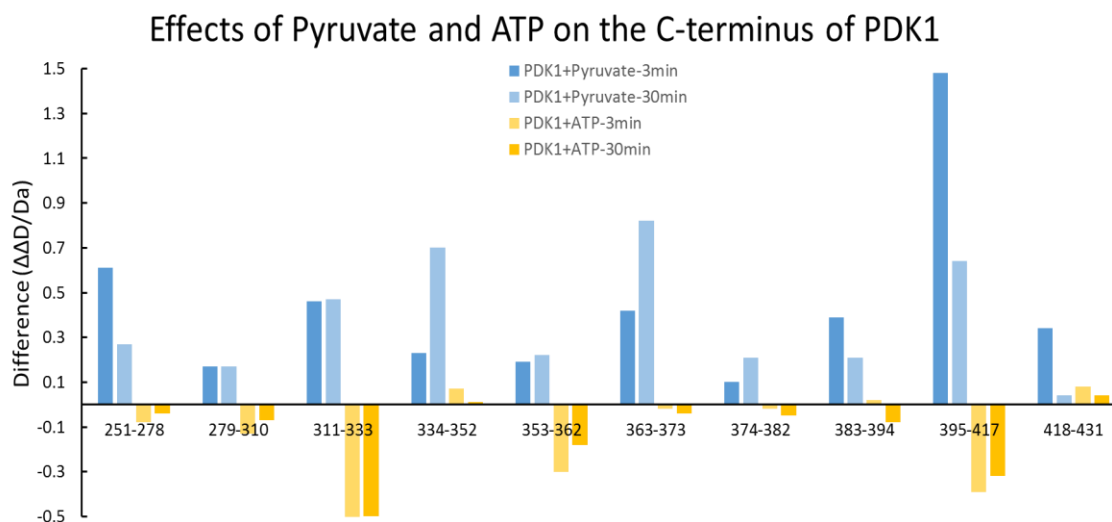











Figure 3.19: Peptides that show significant deuteration change on PDK1 when incubated with pyruvate and ATP.

A similar observation could be made from the effect of pyruvate binding to PDK3. There is no peptide displaying significant decrease in deuterium incorporation when pyruvate interacts with PDK3. Based on the similarities among PDKs, the reason might be the same: the pyruvate is entirely enveloped by a pocket that originally has low solvent accessibility.

Table 3.3: Peptides that show a significant deuteration level change in PDK3 on addition of pyruvate.

Peptide	Sequence	$\Delta\Delta D$ (Da)	
113-132	IKVRNRHNDVVPTMAQGVIE		0.4
235-249	VYVPSHLFHMLFELF		0.35
274-281	VTLGKEDL		0.4
282-300	SIKISDLGGGVPLRKIDRL		0.7
301-304	FN YM		-0.4
305-324	YSTAPRPSLEPTRAAPLAGF		0.5
325-333	GYGLPISRL		-0.5
366-388	FERLPVFNKSARHRYKTTPEADI		0.8
389-406	WSNPSSSEPRDASKYKAKQ		0.7

When comparing the effects of pyruvate and ATP on PDK3, it was apparent that most of the peptides that show significant deuteration change upon binding to pyruvate also display change when interacting with ATP, though some showed the same pattern, others the opposite. The data of PDK1 interacts with pyruvate and that of with ATP suggests that in the presence of pyruvate, PDK1 might adopt a conformation that is totally unsuitable for ATP binding, as the peptides that experiencing exposure to the solvent in the presence of pyruvate would be shielded from solvent if ATP presents instead. However, the interaction between PDK3 and pyruvate or ATP doesn't exhibit that much differences. These differences may be able to explain why PDK1 is more sensitive to pyruvate/DCA inhibition, while PDK3 is nearly insensitive. [13]

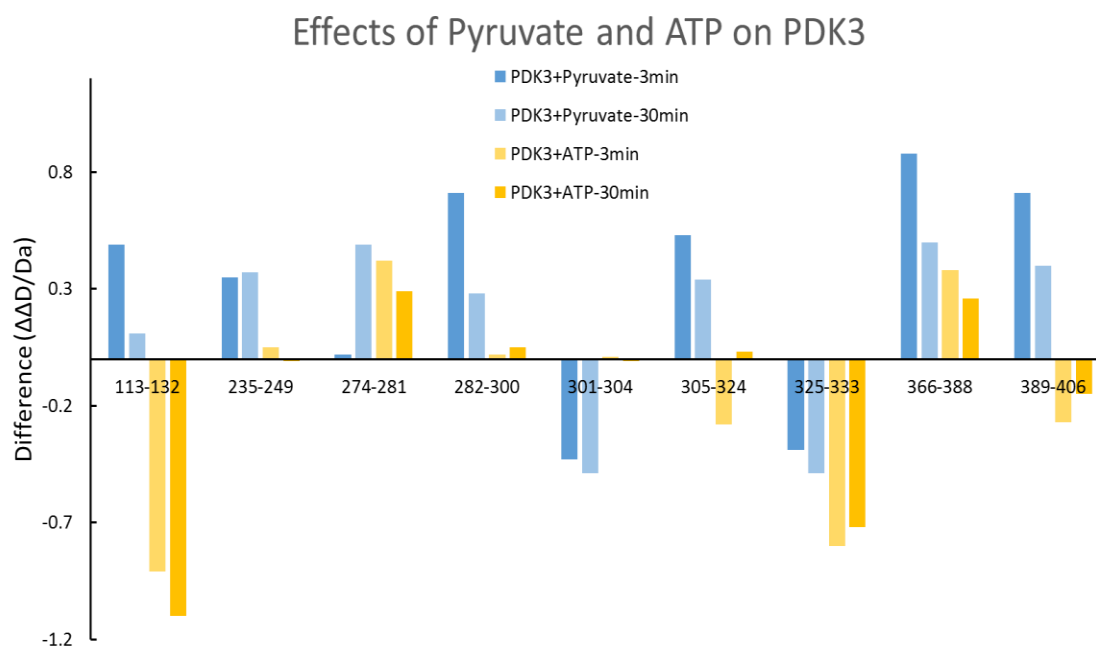


Figure 3.20: Peptic peptides from PDK3 that show significant deuteration change on incubation with pyruvate and ATP.

Both interaction patterns between PDKs and pyruvate show that pyruvate binds to the allosteric site located on the N-terminal domain, which will induce a large conformational change on the ATP binding site.

Table 3.4: Peptides identified from pepsin digestion of the E1p -subunit.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	2-7	ANDATF	638.2790	638.2780	1.7
2	8-14	EIKKCDL	848.4562	848.4546	1.9
3	15-27	HRLEEGPPVTTVL	1447.7924	1447.7904	1.4
4	28-35	TREDGLKY	981.5001	981.5000	0.1
5	42-47	VRRMEL	803.4552	803.4556	-0.3
6	53-64	YKQKIIRGFCHL	1505.8412	1505.8410	0.1
7	73-88	VGLEAGINPTDHLITA	1620.8606	1620.8592	0.4
8	127-149	YAKNFYGGNGIVGAQVPLGAGIA	2237.1720	2237.1713	0.3
9	151-161	ACKYNGKDEVK	1567.7760	1567.7751	0.6
10	162-177	LTLYGDGAANQGQIFE	1696.8190	1696.8177	0.8
11	179-184	YNMAAL	682.3236	682.3229	0.9
12	185-191	WKLPCIF	906.4919	906.4906	1.5
13	192-205	ICENNRYGMGTSVE	1572.6792	1572.6781	0.9
14	206-213	RAAASTDY	854.4008	854.4003	0.2
15	214-229	YKRGFIPGLRVDGMD	2001.9851	2001.9854	0.1
16	243-252	YCRSGKGPIIL	1093.5823	1093.5834	1.0
17	253-257	MELQT	621.2921	621.2912	1.4
18	258-271	YRYHGHSMSPGVS	1592.6911	1592.6927	1.0
19	277-290	EIQEVRSKSDPIML	1644.8633	1644.8625	0.5
20	291-300	LKDRMVNSNL	1189.6376	1189.6358	1.5
21	301-322	ASVEELKEIDVEVRKEIEDAAQ	2500.2790	2500.2777	0.5
22	323-334	FATADPEPPLEEL	1428.6907	1428.6893	1.0
23	335-353	LGYHIYSSDPPFEVRGANQ	2150.0322	2150.0301	1.0
24	354-361	WIKFKSVS	994.5737	994.5720	1.7

Table 3.5: Peptides identified from pepsin digestion of the E1p β -subunit.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	3-14	VTVRDAINQGMD	1318.6436	1318.6420	1.2
2	17-25	LERDEKVFL	1148.6325	1148.6310	1.3
3	30-59	VAQYDGAYKVSRLWKKYGDKRIIDTPISE	3456.8296	3456.8274	0.6
4	60-64	MGFAG	482.2072	482.2068	0.9
5	65-79	IAVGAAMAGLRPICE	1471.7760	1471.7775	1.0
6	81-85	MTFNF	659.2866	659.2858	1.1
7	101-114	YYMSGGLQPVPVIF	1570.7985	1570.7974	0.9
8	115-131	RGPNGASAGVAAQHSQC	1610.7465	1610.7452	0.8
9	132-170	FAAWYGHCPGLKVVSPWNSEDAKGLIKSAIRDNNPVVVL	4251.2190	4251.2172	0.4
10	171-175	ENELM	635.2708	635.2705	0.2
11	176-190	YGVPFEPPEAQSKD	1710.8019	1710.8010	0.5
12	191-222	FLPIGKAKIERQGTHITVVSHSRPVGHCLEA	3493.9213	3493.9256	1.2
13	233-238	CEVINM	708.3058	708.3055	0.1
14	239-248	RTIRPMDMET	1249.6049	1249.6028	1.7
15	249-272	IEASVMKTNHLVTVEGGWPQFGVG	2314.1672	2314.1649	1.0
16	275-289	ICARIMEGPAFNFLD	1696.8197	1696.8186	0.6
17	286-321	NFLDAPAVRVTGADVPMKYKILEDNSIPQVKDIIF	3957.0872	3957.0830	1.1
18	322-329	AIKKTLNI	900.5892	900.5877	1.7

Table 3.6: Peptides identified from pepsin digestion of PDK1^{rat}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Error (ppm)	
No.	Position	Sequence	Experimental	Theoretical	
1	27-48	ASDSAGSGPASESGVPGQVDF	2008.8730	2008.8742	0.6
2	49-59	YARFSPSPLSM	1255.6140	1255.6157	1.4
3	60-65	MKQFLDF	797.4192	797.4187	-0.6
4	66-77	GSVNACEKTSFM	1273.5551	1273.5573	1.7
5	78-87	FLRQELPVRL	1270.7630	1270.7624	-0.5
6	88-101	ANIMKEISLLPDNL	1570.8509	1570.8518	0.6
7	102-113	LRTPSVQLVQSW	1413.7849	1413.7847	-0.1
8	114-121	YIQLQEL	993.5251	993.5258	0.7
9	122-137	LDFKDKSAEDAKTIYE	1872.9226	1872.9260	1.8
10	138-162	FTDTVIRIRNRHNDVIPTMAQGVTE	2883.4894	2883.4901	0.2
11	163-167	YKESF	673.3192	673.3188	-0.6
12	168-179	GVDPVTSQNVQY	1306.6274	1306.6266	-0.9
13	180-184	FLDRF	697.3668	697.3674	0.9
14	185-193	YMSRISIRM	1156.5965	1156.5972	0.6
15	194-201	LLNQHSLL	937.5465	937.5478	1.4
16	202-240	VVEVIKDGyenARRLCDLFGGKGSPSHRKHIGSINPNCd	4281.1404	4281.1432	0.6
17	241-250	YYVNSPELEL	1226.5939	1226.5941	0.2
18	251-278	EELNAKSPGQPIQVVYVPSHLYHMFEL	3224.6449	3224.6452	0.1
19	279-310	FKNAMRATMEHHADKGVYPPIQVHVTLGEEDL	3633.7941	3633.7943	0.1
20	311-333	TVKMSDRGGGVPLRKIDRLFNyM	2653.4065	2653.4057	-0.3
21	334-352	YSTAPRPRVETSRAVPLAG	2175.1669	2175.1672	0.1
22	353-362	FGYGLPISRL	1122.6306	1122.6322	1.4
23	363-373	YAQYFQGDLKL	1345.6787	1345.6803	1.2
24	374-382	YSLEGYGTd	1004.4207	1004.4209	0.2
25	383-394	AVIYIKALSTES	1294.7253	1294.7274	1.6
26	395-417	IERLPVYNKAawkHYRTNHEADd	2826.4070	2826.4107	1.3
27	418-431	WCVPSREPkdMTTF	1549.7138	1549.7142	0.2

Table 3.7: Peptides identified from pepsin digestion of PDK2^{rat}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Theoretical	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	20-33	IEHFSKFSPSPLSM	1606.7952	1606.7934	1.1
2	34-46	KQFLDFGSSNACE	1445.6369	1445.6366	0.2
3	51-82	TFLRQELPVRLANIMKEINLLPDRVLSTPSVQ	3691.0751	3691.0727	0.6
4	83-87	LVQSW	632.3411	632.3402	1.4
5	88-92	YVQSL	609.3247	609.3243	0.7
6	93-97	LDIME	620.2968	620.2960	1.3
7	98-112	FLDKDPEDHRTLSQF	1847.8927	1847.8923	0.2
8	112-116	FTDAL	566.2831	566.2821	1.8
9	117-135	VTIRNRHNDVVPTMAQGVL	2120.1401	2120.1393	0.4
10	136-149	EYKDTYGDDPVSQ	1630.6875	1630.6867	0.5
11	150-153	NIQY	537.2668	537.2667	0.1
12	154-158	FLDRF	697.3682	697.3668	2.0
13	159-167	YLSRISIRM	1138.6411	1138.6401	0.9
14	168-176	LINQHTLIF	1098.6310	1098.6306	0.4
15	177-210	DGSTNPAHPKHIGSIDPNCVSDVVKDAYDMAKL	3581.7003	3581.6999	0.1
16	211-215	LCDKY	641.2966	641.2963	0.5
17	216-222	YMASPDL	796.3551	796.3546	0.6
18	223-238	EIQEVNATNATQPIHM	1795.8662	1795.8643	1.1
19	239-249	VYVPSHLYHML	1358.6928	1358.6926	0.1
20	253-267	FKNAMRATVESHESS	1693.7977	1693.7962	0.9
21	268-276	LTLPIKIM	1025.6433	1025.6428	0.5
22	277-284	VALGEEDL	845.4270	845.4251	2.2
23	285-303	SIKMSDRGGGVPLRKIERL	2112.2077	2112.2070	0.3
24	304-307	FSYM	547.2222	547.2221	0.1
25	308-335	YSTAPTPQPGTGGTPLAGFGYGLPISRL	2776.4321	2776.4305	0.6
26	336-346	YAKYFQGDQL	1345.6791	1345.6787	0.3
27	347-355	FSMEGFQTD	990.3881	990.3873	0.8
28	356-362	AVIYLKA	777.4881	777.4869	1.5

29	363-390	LSTDSVERLPVYNKSAWRHYQTIQEAGD	3263.6087	3263.6080	0.2
30	391-403	WCVPSTEPKNTST	1449.6684	1449.6679	0.3

Table 3.8: Peptides identified from pepsin digestion of PDK3^{human}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Theoretical	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	48-60	FLRKELPVRLANT	1556.9301	1556.9271	1.9
2	61-71	MREVNLLPDNL	1313.6902	1313.6882	1.4
3	72-78	LNRPSVG	742.4213	742.4206	0.9
4	79-83	LVQSW	632.3416	632.3402	2.1
5	84-88	YMQSF	675.2818	675.2807	1.7
6	92-108	LEYENKSPEDPQVLDF	2036.9476	2036.9447	1.4
7	113-132	IKVRNRHNDVVPTMAQGVIE	2276.2322	2276.2292	1.3
8	155-161	YTNRISF	900.4587	900.4574	1.4
9	162-171	RMLINQHTLL	1238.7053	1238.7038	1.2
10	172-195	FGGDTNPVHPKHIGSIDPTCNVAD	2491.1692	2491.1670	0.9
11	196-206	VVKDAYETAKM	1254.6419	1254.6399	1.6
12	211-215	YYLVA	628.3353	628.3341	2.0
13	219-233	EVEEFNAKAPDKPIQ	1714.8671	1714.8646	1.5
14	234-241	VVYVPSHL	913.5154	913.5142	1.4
15	235-249	VYVPSHLFHMLFELF	1878.9666	1878.9662	0.1
16	260-273	YEDRKEGYPAVKTL	1668.8609	1668.8592	1.0
17	274-281	VTLGKEDL	874.4888	874.4880	0.9
18	282-300	SIKISDLGGGVPLRKIDRL	2037.2212	2037.2179	1.6
19	301-304	FNYM	574.2340	574.2330	1.7
20	305-324	YSTAPRPSLEPTRAAPLAGF	3058.6530	3058.6473	1.9
21	325-333	GYGLPISRL	975.5639	975.5622	1.7
22	334-344	YARYFQGDLKL	1373.7240	1373.7212	2.0
23	345-354	YSMEGVGTDA	1029.4203	1029.4194	0.9
24	366-388	FERLPVFNKSAWRHYKTTPEADD	2807.3937	2807.3900	1.3
25	389-406	WSNPSSEPRDASKYKAKQ	2079.0275	2079.0254	1.0

Table 3.9: Peptides identified from pepsin digestion of PDK4^{rat}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Theoretical	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	12-19	SSLGNAGL	718.3738	718.3730	1.1
2	17-26	AGLVPREVEL	1082.6221	1082.6204	1.6
3	27-37	FSRYSPSPLSM	1271.6092	1271.6089	0.2
4	58-66	RQELPVRLA	1081.6492	1081.6477	1.4
5	66-71	ANILKE	687.4042	687.4036	0.9
6	80-86	VNTPSVQ	744.3891	744.3886	0.7
7	87-96	LVKSWYIQL	1236.6997	1236.6987	0.8
8	103-121	HEKSPEDQKVLSDFVDTLV	2186.0991	2186.0976	0.7
9	121-140	VKVRNRHHNVPTMAQGILE	2298.2627	2298.2612	1.8
10	125-145	TMAQGVIEYKEKFGDFPFIST	2408.1855	2408.1843	0.6
11	147-171	VDPVTNQNLQYFLDRFYMNRISTRM	3121.5377	3121.5346	1.0
12	158-162	FLDRF	697.3674	697.3668	0.9
13	163-175	YMNRISTRMLMNQ	1657.7990	1657.7971	0.9
14	208-217	FECAKMLCDQ	1187.4901	1187.4894	1.6
15	218-225	YYLTSPEL	985.4882	985.4877	1.2
16	226-252	KLTQVNGKFPGQPIHIVYVPSHLHHML	3090.6841	3090.6822	0.6
17	257-262	KNAMRA	690.3722	690.3716	0.9
18	260-281	MRATVEHQENRPFLTPVEATVV	2524.2994	2524.2977	0.7
19	279-287	TVVLGKEDL	973.5568	973.5564	0.4
20	288-306	TIKISDRGGGVPLRITDRL	2067.2041	2067.2033	0.4
21	310-328	TYSTAPTPVMDNSRNAPLA	2005.9661	2005.9648	0.6
22	335-340	PISRLY	748.4355	748.4352	0.4
23	340-356	YAKYFQGDLNLYMSGY	2019.9162	2019.9157	0.2
24	395-412	WCIPSKPKNLSKEKLAV	2070.1441	2070.1416	1.2

References

- [1] Roche, T. E., Baker, J. C., Yan, X., Hiromasa, Y., Gong, X., Peng, T., Dong, J., Turkan, A., and Kasten, S. A. (2001) Distinct regulatory properties of pyruvate dehydrogenase kinase and phosphatase isoforms. *Prog. Nucleic Acid Res. Mol. Biol.* **70**, 33-75.
- [2] Steussy C N, Popov K M, Bowker-Kinley M M, Sloan R B, Harris R A, Hamilton J A. Structure of Pyruvate Dehydrogenase Kinase novel folding pattern for a serine protein [J]. *Journal of Biological Chemistry*, 2001, 276(40): 37443-37450.
- [3] Wynn R M, Kato M, Chuang J L, Tso S C, Li J, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. *Journal of Biological Chemistry*, 2008, 283(37): 25305-25315.
- [4] Ubersax, J.A. and J.E. Ferrell Jr, Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol*, 2007. 8(7): p. 530-541.
- [5] Kato M, Chuang J L, Tso S C, Wynn R M, Chuang D T. Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex[J]. *The EMBO journal*, 2005, 24(10): 1763-1774.
- [6] Kato M, Li J, Chuang J L, Chuang D T, Distinct structural mechanisms for inhibition of pyruvate dehydrogenase kinase isoforms by AZD7545, dichloroacetate, and radicicol. *Structure*, 2007. 15(8): p. 992-1004.
- [7] Knoechel T R, Tucker A D, Robinson C M, Phillips C, Taylor W, Bungay P J, Kasten S A, Roche T E, Brown D G, Regulatory Roles of the N-Terminal Domain Based on Crystal Structures of Human Pyruvate Dehydrogenase Kinase 2 Containing Physiological and Synthetic Ligands. *Biochemistry*, 2005. 45(2): p. 402-415

- [8] Devedjiev, Y., C.N. Steussy, and D.G. Vassilyev, Crystal structure of an asymmetric complex of pyruvate dehydrogenase kinase 3 with lipoyl domain 2 and its biological implications. *J Mol Biol*, 2007. 370(3): p. 407-16.
- [9] Kukimoto-Niino M, Tokmakov A, Terada T, Ohbayashi N, Fujimoto T, Gomi S, Shiromizu I, Kawamoto M, Matsusue T, Shirouzu M, Yokoyama S, Inhibitor-bound structures of human pyruvate dehydrogenase kinase 4. *Acta Crystallogr D Biol Crystallogr*, 2011. 67(Pt 9): p. 763-73.
- [10] Patel M S, Roche T E. Molecular biology and biochemistry of pyruvate dehydrogenase complexes[J]. *The FASEB Journal*, 1990, 4(14): 3224-3233.
- [11] Guevara E L, Yang L, Birkaya B, Zhou J, Nemeria N S, Patel M S, Jordan F. Global view of cognate kinase activation by the human pyruvate dehydrogenase complex[J]. *Scientific Reports*, 2017, 7.
- [12] Hiromasa Y, Hu L, Roche T E. Ligand-induced effects on pyruvate dehydrogenase kinase isoform 2[J]. *Journal of Biological Chemistry*, 2006, 281(18): 12568-12579.
- [13] Roche T E, Hiromasa Y. Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer[J]. *Cellular and molecular life sciences*, 2007, 64(7): 830-849.

CHAPTER 4. Study of protein-protein interactions within the human 2-oxoglutarate dehydrogenase complex using site-specifically introduced external fluorophores

Abbreviation used in this chapter: the postscript *o* is for oxoglutarate, while *p* is for pyruvate.

4.1 Introduction

Site-specific labeling provides biochemists a handle for target detection, quantification and analysis, which is essential for biological research, diagnostics, and therapeutics. Many labeling methods have been developed, including covalent modification of amino acid side chains, covalent labeling by enzymes, conjugation to unnatural amino acids, and affinity labeling [1-6].

One of the most common specific covalent coupling methods is the conjugation of a moiety bearing a thiol-reactive group to a solvent-accessible cysteine. This method affords the site-specific labeling of a protein at a unique engineered (or native) surface cysteine. This cysteine could be specifically labeled with dye derivatives of haloacetyl compounds, disulfide, or maleimides. The maleimide is by far the most prevalent functional group to be coupled to a cysteine because the coupling reaction is highly specific and efficient [7].

The protein must be maintained in a reduced form prior to the maleimide-cysteine coupling reaction to prevent the formation of disulfide bridges and inactivation of the cysteines. Reducing agents such as dithiothreitol (DTT) and tris[2-

carboxyethyl]phosphine (TCEP) can be used for this purpose. These reducing agents must be removed before conjugation so that their thiol groups will not compete with target thiols in proteins. Also, the labeling reaction must be performed shortly after the elimination of reducing agents to prevent reoxidation of the target thiols, resulting in decreased labeling efficiency [7].

Fluorescence spectral properties can provide a variety of molecular details of proteins, such as the solvent exposure of residues, the existence of protein conformers, the rate of rotational diffusion of a protein, and the distance between sites on a protein. This method is so sensitive that only a few micromoles or even nanomoles of sample is needed during the measurements. In addition to that, the fluorescence spectrum, and the phenomenon it reports on, is also responsive to the changes in solvent polarity, pH, temperature or the addition of quenching agents, which make it possible to study more complicated cases [8-12].

2-Oxoglutarate dehydrogenase complex (OGDHc) belongs to the superfamily of 2-oxo acid dehydrogenase multienzyme complexes. However, the E2o component from OGDHc differs from that from PDHc significantly as the sequence alignment of the E2o could not identify any apparent peripheral subunit-binding domain (S) [13-14].

In addition, there is no extra-component such as E3BP which could provide PSBD in this complex. The structure of E2o is composed of a lipoyl domain (LDo), a linker and a catalytic domain (Core). Hence, as to how this complex is assembled, and whether indeed the mammalian OGDHc has a new unprecedented mode of assembly, are still open questions.

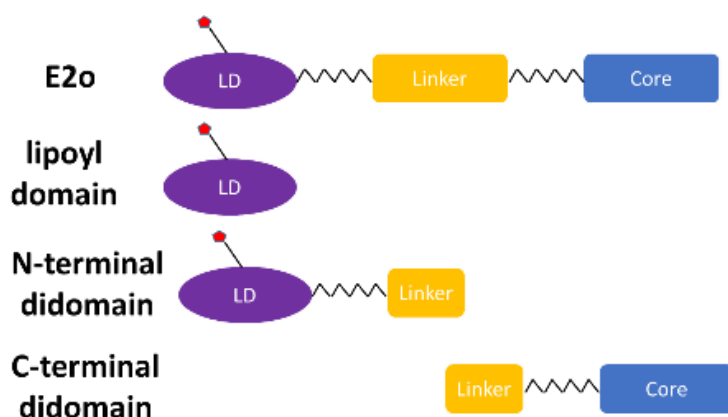


Figure 4.1: Domain structure of E2o and E2o constructs. E2o¹⁻¹⁷³ (N-terminal didomain) and E2o¹⁴⁴⁻³⁸⁶ (C-terminal didomain) have 30 amino acid overlapping sequence.

4.2 Materials and Methods

N-(1-Pyrene)maleimide was purchased from Ana Spec; dansyl chloride and p-arsanilic acid were from Sigma Aldrich. All the chemicals were used without further purification.

4.2.1 Labeling of lipamide in the LDo with a dansyl group for fluorescence studies.

The lipoylated LDo (50 μ M) was incubated with 100 μ M TCEP in sample buffer (30 mM KH₂PO₄, 0.15 M NaCl, pH 7.5) at room temperature for 5 min, resulting in fully reduced lipoyl domains, which were then reacted with 150 μ M DANS-As at room temperature for 2 h, forming chemically modified lipoyl domains. The reaction mixture was dialyzed against the sample buffer with a centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the surplus TCEP and DANS-As. The product was diluted to 1 μ M with the ESI buffer (a 50:50:0.1 v/v/v mixture of methanol, water and formic acid), and examined by

FT-ESI-MS in the positive ion mode. The procedure has been published recently by this group [15].

4.2.2 Labeling $E2o^{1-173}$ and $E2o^{144-386}$ with *N*-(1-pyrene)maleimide for fluorescence

studies. The $E2o^{1-173}$ or $E2o^{144-386}$ (50 μ M), each containing a single cysteine (*for E2o¹⁻¹⁷³ it's Cys38; for E2o¹⁴⁴⁻³⁸⁶ it's Cys179*), was incubated with 100 μ M TCEP in sample buffer (30 mM KH_2PO_4 , 0.15 M NaCl, pH=7.5) at room temperature for 5 min, then reacted with 100 μ M *N*-(1-pyrene)maleimide at room temperature for 2 h, forming a protein bearing a single pyrene fluorophore. The reaction mixture was dialyzed against sample buffer with a centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the surplus TCEP and *N*-(1-pyrene)maleimide.

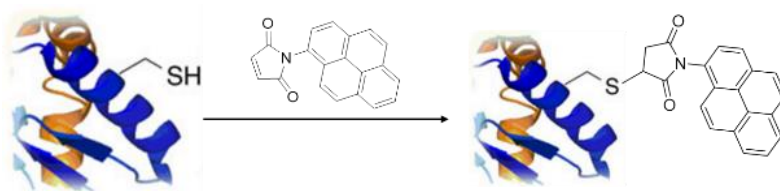


Figure 4.2: Labeling of $E2o^{1-173}$ and $E2o^{144-386}$ with *N*-(1-pyrene)maleimide.

4.2.3 Fluorescence spectroscopy titration studies. All fluorescence spectra were recorded on a Varian Cary Eclipse fluorescence spectrophotometer. The titration was conducted in 30 mM KH_2PO_4 (pH 7.5). A stock solution of chemically modified lipoyl domains was added to 30 mM KH_2PO_4 (pH 7.5) to reach a final concentration of 1-2 μ M. The excitation wavelength for the DANS-As modified proteins was 338 nm, and the emission spectra were recorded from 400-650 nm. The maximum emission intensities at 520 nm were used for calculating binding parameters.

For N-(1-pyrene)maleimide labeled proteins (E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶), the excitation wavelength was 340 nm, and recorded emission spectra from 350-600 nm. The maximum emission intensities at 374 nm were used to calculate binding parameters.

All data sets are normalized ($\Delta F/\Delta F_{\max}$) to their maximum fluorescence change.

Equilibrium dissociation constants (K_d) were obtained by using the Hill equation to fit data.

$$\Delta F / \Delta F_{\max} = [E1o]^n / (K_A^n + [E1o]^n) \quad \text{Equation 4.1}$$

4.2.4 Measurement of E1o-specific activity. The E1o-specific activity was measured with the external oxidizing agent 2,6-dichlorophenolindophenol (DCPIP) [18] at 600 nm in the following reaction medium contained in 1 ml: 50 mM KH₂PO₄ (pH 7.0), 0.50 mM ThDP, 1.0 mM MgCl₂, 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 ($\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg E1o}^{-1}$).

4.2.5 Fluorescence study of the binding of thiochrome diphosphate to E1o. Thiochrome diphosphate (TCDP) is a fluorescent analogue of ThDP and can act as a competitive inhibitor for ThDP dependent enzymes. It is known that the intrinsic protein fluorescence due to tryptophan can be quenched by the binding of ThDP and ThDP analogues. The fluorescence quenching resulting from addition of TCDP exhibited saturation behavior, which can be fitted to Equation xx to estimate the K_d value.

$$\Delta F / \Delta F_{max} = [TCDP]^n / (S_{0.5}^n + [TCDP]^n) \quad \text{Equation 4.2}$$

where $\Delta F / \Delta F_{max}$ is a relative fluorescence; $\Delta F = F_{max} - F_i$, where F_{max} is a maximum fluorescence intensity reached on titration by TCDP and F_i is a fluorescence intensity at a given concentration of TCDP; $\Delta F_{max} = F_{max} - F_o$, where F_o is the initial fluorescence before addition of TCDP; $S_{0.5}$, is the concentration of ligand at half saturation; n is the Hill coefficient. For $n=1$, the value of $S_{0.5}$ is equal to K_d .

4.2.6 FRET experiments. The energy transfer from E1o-TCDP to E3 were conducted on a Varian Cary Eclipse fluorescence spectrophotometer. The efficiency of energy transfer was obtained from the excitation spectrum of the energy acceptor—TCDP with the concentrations of E1o and E3 both at 5 μ M concentration in 50 mM KH_2PO_4 (pH=7.0). Also, the spectrum of E1o bound thiochrome and hE3 bound FAD at the same concentration in the same buffer were recorded separately.

4.2.7 Labeling $E2o^{1-173}$, $E2o^{144-386}$ with Alexa Fluor® 350 C5 Maleimide dye for FRET studies.

The truncated E2o proteins $E2o^{1-173}$ and $E2o^{144-386}$ (50 μ M; each contain only one cysteine), were incubated with 100 μ M TCEP in sample buffer (30 mM KH_2PO_4 , 0.15 M NaCl, pH=7.5) at room temperature for 5 min, then reacted with 100 μ M Alexa Fluor® 350 C5 Maleimide dye at room temperature for 2 h, forming a protein bearing a single Alexa350 fluorophore. The reaction mixture was dialyzed against sample buffer with a

centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the surplus TCEP and Alexa Fluor® 350 C5 Maleimide dye.

The fluorescence measurements were the same as described above. 5 μM of E2o¹⁻¹⁷³-Alexa350 (or of E2o¹⁴⁴⁻³⁸⁶-Alexa350) and E3 were incubated in 50 mM KH₂PO₄ (pH=7.0) at room temperature for 5 min, then the excitation wavelength was set at 350 nm, and the spectrum was recorded. The hE3 was gradually titrated into the mixture, and the spectrum was recorded after the addition of each aliquot. The spectrum of E2o¹⁻¹⁷³-Alexa350 (and of E2o¹⁴⁴⁻³⁸⁶-Alexa350) were recorded at the same concentration in the same buffer as control. The emission intensity of the donor (Alexa350) remained the same, during all additions of E3. The results indicated no energy transfer between E2o and E3, irrespective of whether the Alexa350 label was attached to E2o¹⁻¹⁷³ or E2o¹⁴⁴⁻³⁸⁶.

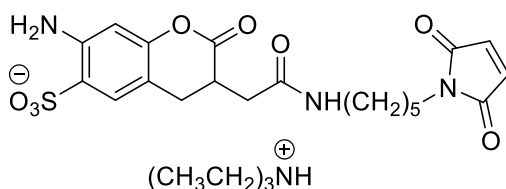


Figure 4.3: Structure of Alexa Fluor® 350 C5 Maleimide dye

4.3 Results and Discussion

Two approaches were developed to introduce an external fluorophore site-specifically onto the E2o proteins. In the first approach, the lipoamide carrying proteins, E2o¹⁻⁹⁵ (lipoyl domain, LDo, comprising lipoyl domain and part of the linker) and the E2o¹⁻¹⁷³ were dansylated. The lipoamide was first reduced to dihydrolipoamide, then labeled with

a dansyl group which had been tethered to the amino end of 4-aminophenylarseneoxide, resulting from a reaction of the trivalent arsenoxide with the reduced dihydrolipoamide [15]. In the second approach, N-(1-pyrene)-maleimide was used to label the unique cysteines of E2o¹⁻¹⁷³ (at Cys38) and E2o¹⁴⁴⁻³⁸⁶ (at Cys179) proteins.

4.3.1 E1o-E2o interactions

Fluorescence Spectroscopy Studies.

Titration of DANS-As labeled E2o¹⁻⁹⁵ by E1o resulted in quenching of the fluorescence of dansylamide, as also observed on titration of DANS-As labeled E2o¹⁻¹⁷³, suggesting that on interaction with E1o, a more hydrophilic microenvironment around the fluorophore was created according to data from the literature [16]. The calculated values of K_{ds} were: 0.135 μ M (with E2o¹⁻⁹⁵), and 0.038 μ M (with E2o¹⁻¹⁷³), providing the first hint that the linker region present in the latter could be an important determinant for E1o-E2o interaction.

With the N-(1-pyrene) maleimide labeled E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ proteins, on excitation at 340 nm, three major fluorescence emission bands were in evidence (374, 394 and 416 nm). The maximum emission intensities at 374 nm were used to calculate binding parameters. In the spectrum of E2o¹⁻¹⁷³, there is an additional broad peak centered around 470 nm, which is likely due to excimer formation. Pyrene-thiol conjugates often have long fluorescence lifetimes, which can allow adjacent pyrene rings within 6-10Å of each other to form excited-state dimers (excimers) that emit at longer wavelength than the

excited monomeric fluorophore [17]. This may be the result of unselective simultaneous labeling at both the single cysteine thiol and the reduced dihydrolipoyl group, located at residues 38 and 44, respectively.

Addition of E1o to pyrene-labeled E2o¹⁻¹⁷³ quenched the pyrene fluorescence ($K_d = 0.041 \mu\text{M}$), while addition of E1o to pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ enhanced the pyrene fluorescence intensity without any blue shift or red shift of the peak positions ($K_d = 0.06 \mu\text{M}$, Figure 4.5, 4.6). These results suggested that interactions of E1o with the E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ induce different changes in the microenvironment around the pyrene fluorophore: interaction between E1o and fluorophore labeled E2o¹⁻¹⁷³ leads the pyrene fluorophore to a relatively more hydrophilic, while interaction of E1o with labeled E2o¹⁴⁴⁻³⁸⁶ leads the pyrene fluorophore to a more hydrophobic environment. The calculated K_d 's are in good accord when the interaction with the same protein is measured with the two different fluorophores. The fluorescence binding experiments enabled us to conclude that the linker region in E2o located between LDo and CDo domains participates in the interaction with E1o.

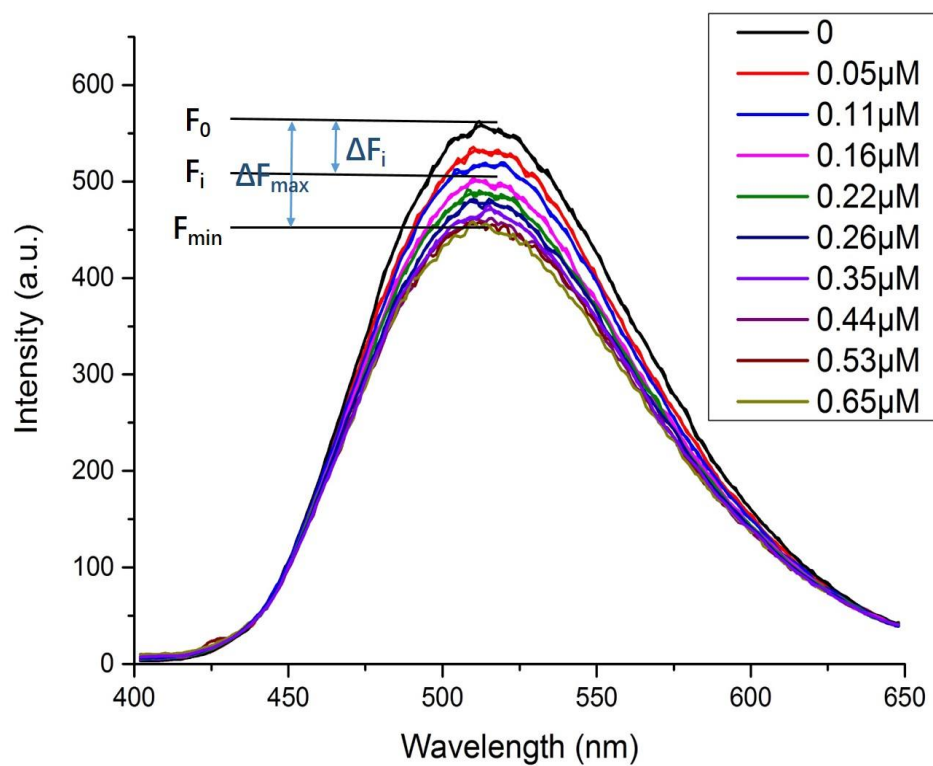
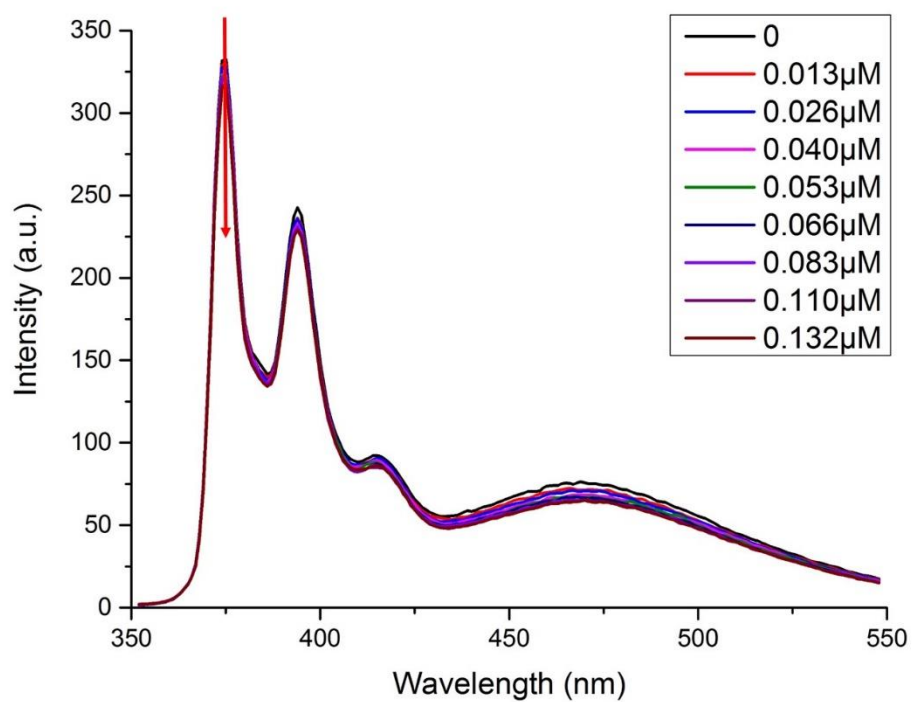


Figure 4.4. Fluorescence titration of DANS- labeled LDo by E1o. Quenching of the fluorescence of DANS-LDo on E1o binding. DANS-LDo (1.0 μM) in 30 mM KH_2PO_4 (pH7.5) was titrated by E1o (0.05-0.65 μM).



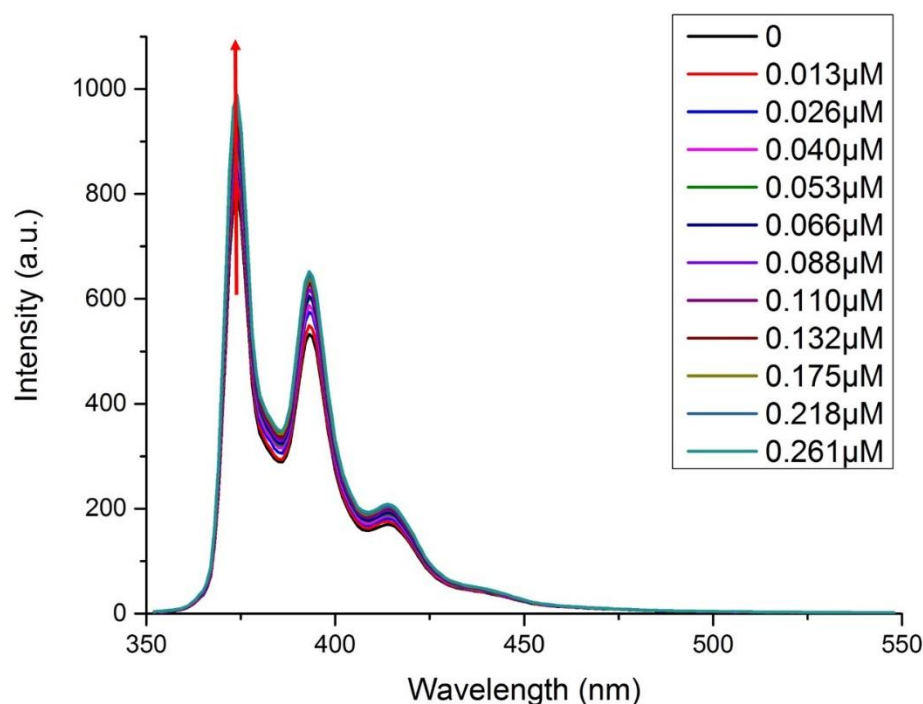


Figure 4.5. Fluorescence titration of Pyrene- labeled E2o¹⁻¹⁷³ and Pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ by E1o. (Top) Quenching of the fluorescence of Pyrene- labeled E2o¹⁻¹⁷³ on E1o binding. Pyrene- labeled E2o¹⁻¹⁷³ (1.5 μM) in 30 mM KH₂PO₄ (pH 7.5) was titrated by E1o (0.013-0.261 μM). (Bottom) Enhancement of fluorescence of the Pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ on E1o binding. Pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ (1.5 μM) in 30 mM KH₂PO₄ (pH7.5) was titrated by E1o (0.013-0.261 μM).

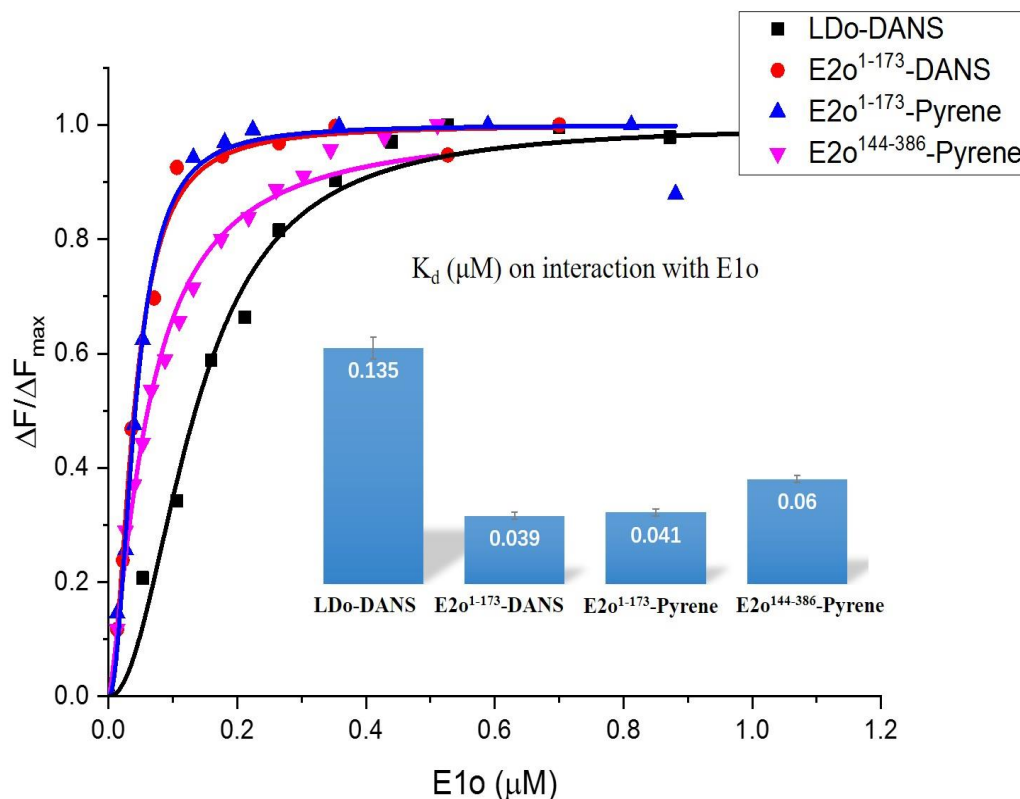


Figure 4.6. Fluorescence titration curves for DANS-As- labeled E2o¹⁻⁹⁵ and E2o¹⁻¹⁷³ and for pyrene labeled E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ proteins by E1o. In all cases the fluorescence intensity of the fluorophore-labeled E2o protein was quenched on titration by E1o. In a typical experiment, the fluorophore-labeled E2o protein (1.0-2.0 μM concentration of subunits) in 30 mM KH₂PO₄ (pH 7.5) was titrated by E1o (0.01-1.5 μM) at room temperature. The excitation wavelength was 338 nm for DANS and 340 nm for pyrene labeled E2o proteins. The fluorescence titration curves were fitted by using the Hill equation (equation 4.1). Inset: The K_d values for E1o interaction with E2o truncated proteins are presented.

4.3.2 E1o-E3 interactions

In view of the much weaker interactions between these two components, we designed a fluorescence resonance energy transfer (FRET) experiment to identify any potential interaction in this complex. Thiochrome diphosphate (TCDP, binds tightly and specifically to the E1o active center), a fluorescent analogue of ThDP was used as a donor, while the fluorescent FAD on E3 would act as energy acceptor, as reported by Hammes and coworkers on the *E. coli* PDHc).

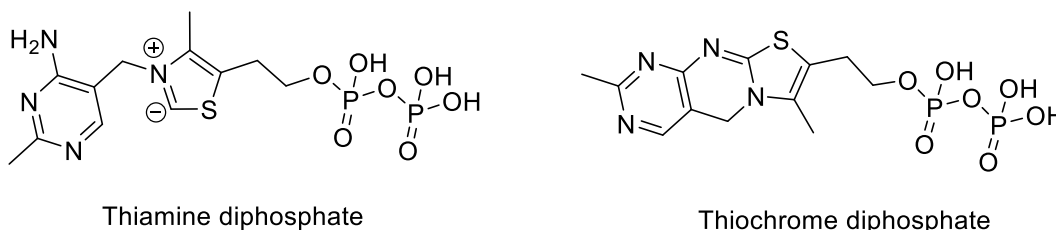


Figure 4.7. Structural formulae of thiamin and thiochrome diphosphate

First, we removed most of the ThDP from E1o. According to the E1o-specific activities, only 17% of ThDP still remained on the protein after passing it through a desalting column followed by extensive dialysis. Flavins have a broad absorption range with four structure-less peaks centered at 446, 375, 265 and 220 nm, all of which possess high molar extinction coefficients. In aqueous solution, flavins exhibit a bright yellow fluorescence at 520 nm with a quantum yield of 0.26 in pH 7.25 buffer (19-20). TCDP has a maximal absorption at 367 nm with a molar extinction coefficient equal to 20600 M⁻¹cm⁻¹. The maximal emission of TCDP is centered at 450 nm [21].

The efficiency of energy transfer is related to the distance between donor and acceptor molecules, which can be treated by the Forster equation [20].

$$E = r^{-6}/(r^{-6} + R_0^{-6}) \quad \text{Equation 4.3}$$

R_0 also called the Forster radius which is defined as the distance of separation at which the energy transfer efficiency is 50%.

$$R_0 = (9.79 \times 10^3)(JK^2Qn^{-4})^{1/6} \quad \text{Equation 4.4}$$

In this equation, K^2 is a dipole-dipole orientation factor, Q represents the quantum yield of the donor molecule, n is the refractive index of the medium, and J is the integral of spectral overlap of the donor and acceptor. According to studies by the Hammes group, the measured efficiency of energy transfer of $\sim 7.8\%$ could be estimated, corresponding to a distance of $\sim 45 \text{ \AA}$ with a possible range of 30-60 \AA between the thiochrome diphosphate and FAD binding sites in *E. coli* PDHc [17]. To estimate the efficiency of energy transfer, individual spectra of the E1o-bound thiochrome diphosphate and E3-bound FAD were recorded (Fig. 11, right). On excitation at 367 nm, the emission maximum at 450 nm (E1o-bound thiochrome diphosphate) and at 520 nm (E3-bound FAD) were detected and were in a good accord with that for free thiochrome diphosphate and FAD (Fig. 11, left), and with data reported in the literature. At pH 7.25 in water, FAD exhibits fluorescence with an emission maximum centered at 520 nm and quantum yield of 0.26 (30, 31), while thiochrome diphosphate has emission maximum centered at 450 nm (29) (Fig. 11, left). On mixing of the E1o-bound thiochrome-diphosphate and of E3-bound-FAD, the fluorescence intensity at 450 nm was quenched and at 520 nm was enhanced. Using the value of R_0 estimated by the Hammes group, and the observed change in fluorescence intensity at 450 nm (from 1027.7 a.u. to 906.4 a.u.), the efficiency of energy transfer was calculated to be 11.8% according to equation 4.5, corresponding to

a distance of approximately 28-56 Å between the ThDP binding site on E1o and FAD binding site on E3.

$$E = I - \tau_{DA} / \tau_A = I - F_{DA} / F_A \quad \text{Equation 4.5}$$

In equation 4.5, τ_{DA} and τ_A are the donor (thiochrome diphosphate) fluorescence lifetime in the presence and absence of an acceptor (FAD) respectively; F_{DA} and F_D are the donor fluorescence intensities with and without an acceptor, respectively. This result attests to interactions between E1o and E3, more specifically, an interaction between E1o ThDP binding site and FAD binding site on E3.

Quenching of the E1o intrinsic fluorescence by TCDP provides a K_d . The analogue TCDP has been known as a competitive inhibitor for ThDP dependent enzymes. It has also been reported that the intrinsic protein fluorescence due to tryptophan can be quenched by the binding of ThDP and ThDP analogues. The fluorescence quenching resulting from addition of TCDP exhibited saturation behavior, which can be fitted to Equation 2 to estimate the K_d value.

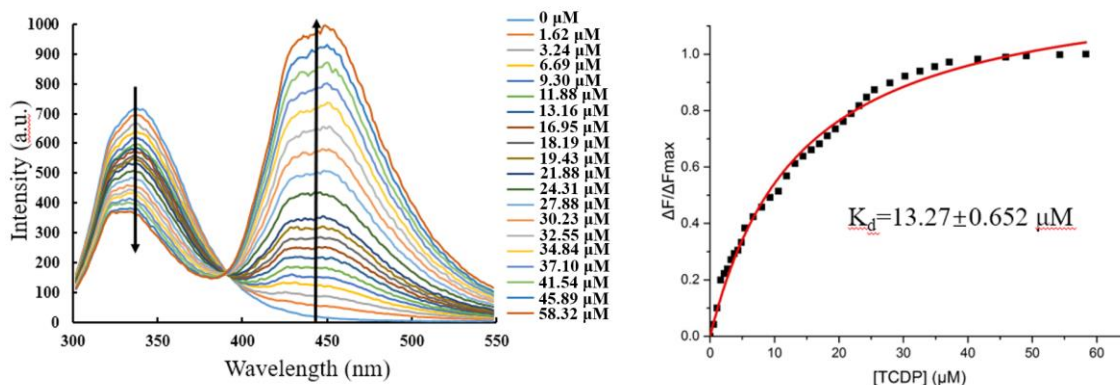


Figure 4.8: Fluorescence titration of E1o by thiochrome diphosphate. (Left) The E1o (3 μ M) in 50 mM KH_2PO_4 (pH 7.0) was titrated by thiochrome diphosphate in the indicated range of concentrations. A sample of E1o with 17% of tightly-bound ThDP was used in the experiment. The excitation wavelength was 280 nm due to E1o tryptophan fluorescence that could be quenched by ThDP and its analogues. On addition of thiochrome diphosphate, the intrinsic tryptophan fluorescence at 348 nm was gradually quenched and the fluorescence at 450 nm due to thiochrome diphosphate binding was gradually built up. (Right) Plot of the relative fluorescence changes at 348 nm on added thiochrome diphosphate at the indicated concentration. The trace is the regression fit line to eq. $\Delta F / \Delta F_{\text{max}} = [\text{TCDP}]^n / (S_{0.5}^n + [\text{TCDP}]^n)$, where $\Delta F / \Delta F_{\text{max}}$ is a relative fluorescence; $S_{0.5}$, is the concentration of TCDP at half saturation and n is the Hill coefficient.

The energy transfer from E1o-TCDP to E3 was determined on a Varian Cary Eclipse fluorescence spectrophotometer. The efficiency of energy transfer was obtained from the excitation spectrum of the energy acceptor-TCDP with the concentrations of E1o and E3 both at 5 μ M concentration in 50 mM KH_2PO_4 (pH 7.0). Also, separate spectra were recorded for E1o-bound TCDP, and for E3-bound FAD at the same concentration in the same buffer. This result confirms the interactions between E1o and E3, more specifically, an interaction between the ThDP binding site from E1o with the FAD binding site on E3. This result is also in accord with HDX-MS data, from which, upon binding to E1o, two peptides with residues 35-70 and 107-145 in the FAD-binding domain of E3, and one

peptide with residues 381-403 from the interface domain of E3 displayed the largest decrease in deuterium uptake.

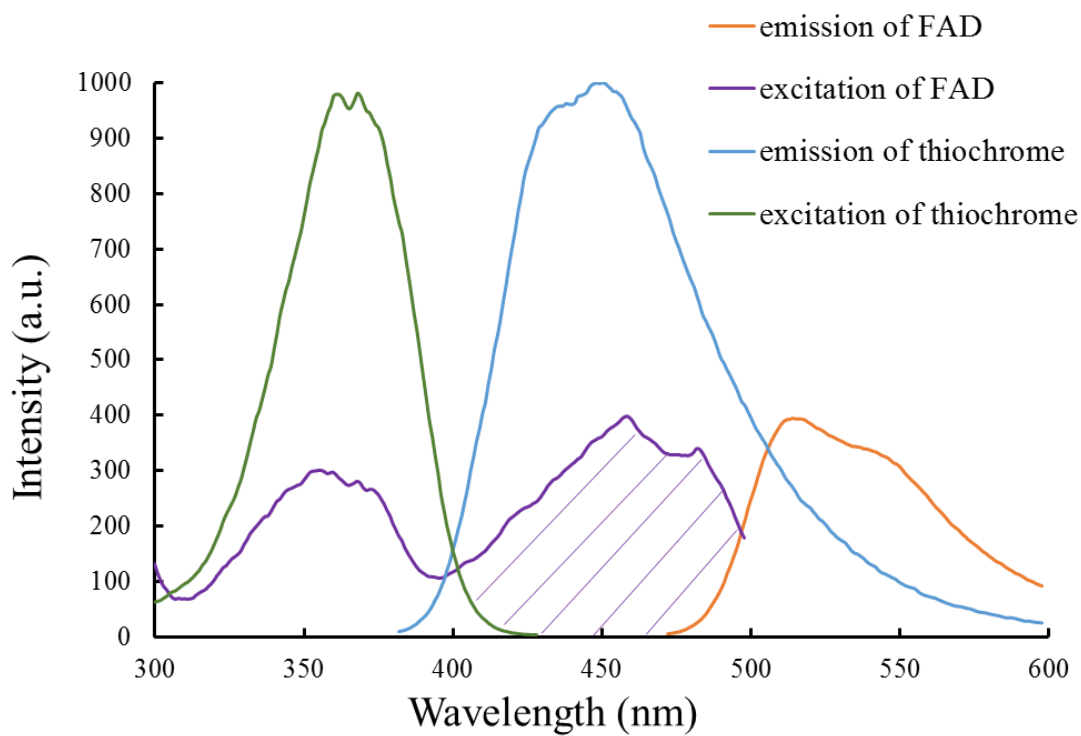


Figure 4.9: A plot of excitation and emission spectrum of FAD and TCDP at a concentration of 10 μM each. The spectral overlap is shown in shadow.

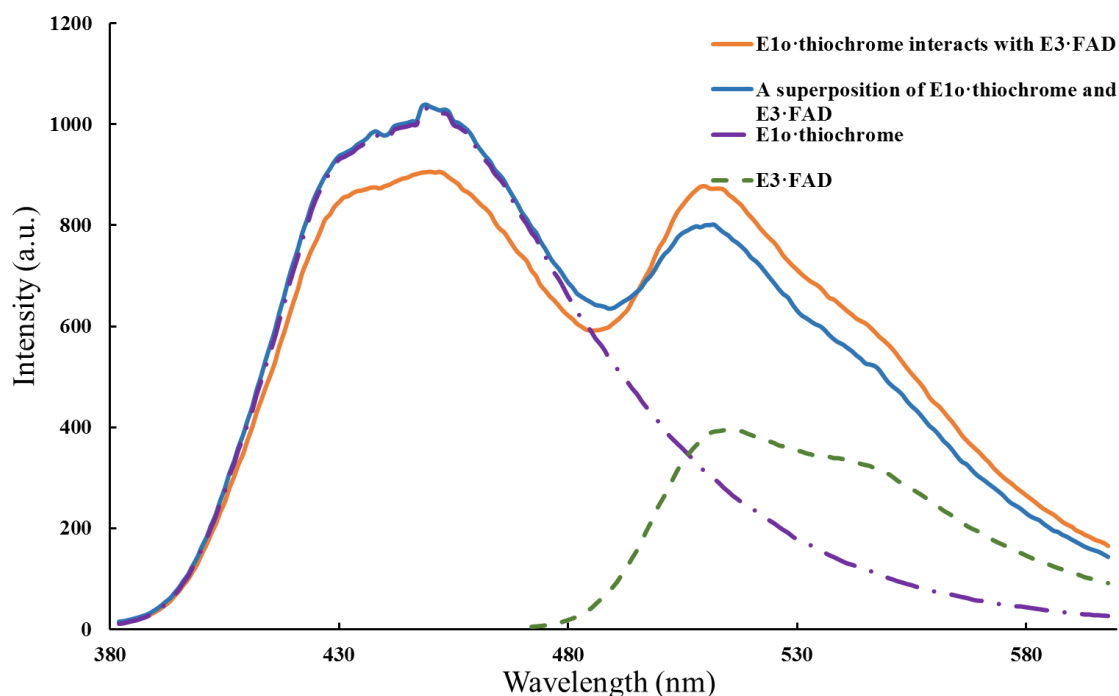


Figure 4.10: Measurements of energy transfer between E1o bound thiochrome diphosphate and E3 bound FAD. The emission spectra of E1o-thiochrome diphosphate (5 μ M) and of E3-FAD (5 μ M) recorded individually (dashed lines) and superimposed (in yellow) in 50 mM KH_2PO_4 (pH 7.0) at the excitation wavelength of 365 nm. In blue, the emission spectrum of a mixture of E1o-thiochrome diphosphate and E3-FAD recorded in 50 mM KH_2PO_4 (pH 7.0) at 5 μ M concentration of each protein.

4.3.3 E2o-E3 interactions

To test potential interaction in the E2o-E3 sub-complex by FRET, we used E2o-truncated proteins where the specific cysteines were labeled with Alexa Fluor®350 C5 maleimide for donor and FAD on E3 as potential acceptor. The Cys³⁸ located near the lipoylated Lys⁴⁴ in E2o¹⁻¹⁷³ di-domain, and Cys¹⁷⁹ in E2o⁴⁴⁻³⁸⁶ core domain were each labeled with the Alexa Fluor® 350 C5 maleimide dye. Energy transfer experiments were carried out

similarly to that reported above and data for Cys³⁸-labeled E2o¹⁻¹⁷³ are presented in Fig. 12. The spectrum of Cys³⁸-labeled E2o¹⁻¹⁷³ by itself displayed an emission maximum at 450 nm due to the labeled Cys³⁸. On addition of FAD-E3 at different concentrations (5-25 μ M) an emission maximum at 520 nm due to E3-bound FAD developed with no changes in the emission maximum due to labeled Cys³⁸, indicating that no energy transfer occurred between Cys³⁸-labeled E2o¹⁻¹⁷³ and E3-bound FAD (Figure 4.12). The most likely explanation is that the distance between the two chromophores is out of the FRET range ($>60\text{\AA}$), suggesting that neither the fluorophore on the E2o lipoyl domain , nor the one on the E2o core domain come in close proximity to the FAD binding site on E3. The reasons for not being able to observe energy transfer could be complex, since based on the equations that $E = r^{-6}/(r^{-6} + R_0^{-6})$ and $R_0 = (9.79 \times 10^3)(JK^2Qn^{-4})^{1/6}$, the efficiency of energy transfer is not only related to the inverse sixth power of the distance between the groups, but is also related to the mutual orientation of the groups, the refractive index of the solvent, and the overlap of the emission spectrum of the energy donor and the absorption spectrum of the energy acceptor. A likely explanation is that the distance between the two chromophores is out of the FRET range ($>60\text{\AA}$), suggesting that neither the fluorophore on the lipoyl domain, nor the one on E2o core domain is in close proximity to the FAD binding site on E3.

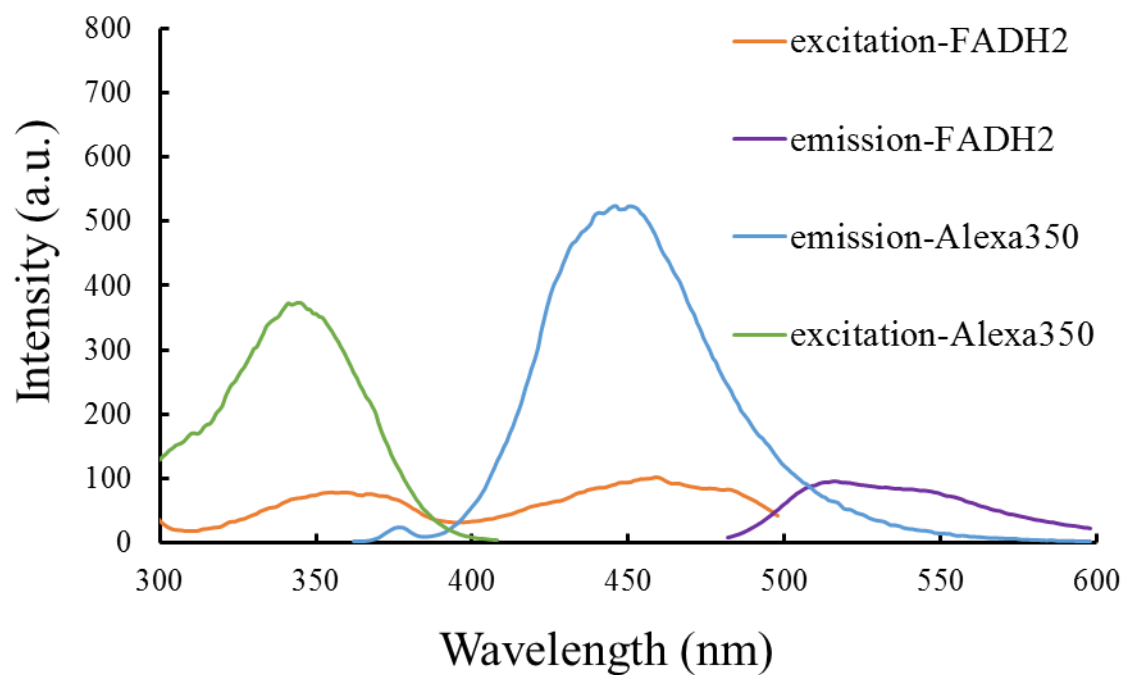


Figure 4.11: A plot of excitation and emission spectrum of FAD and Alexa350 with a concentration of 10 μM .

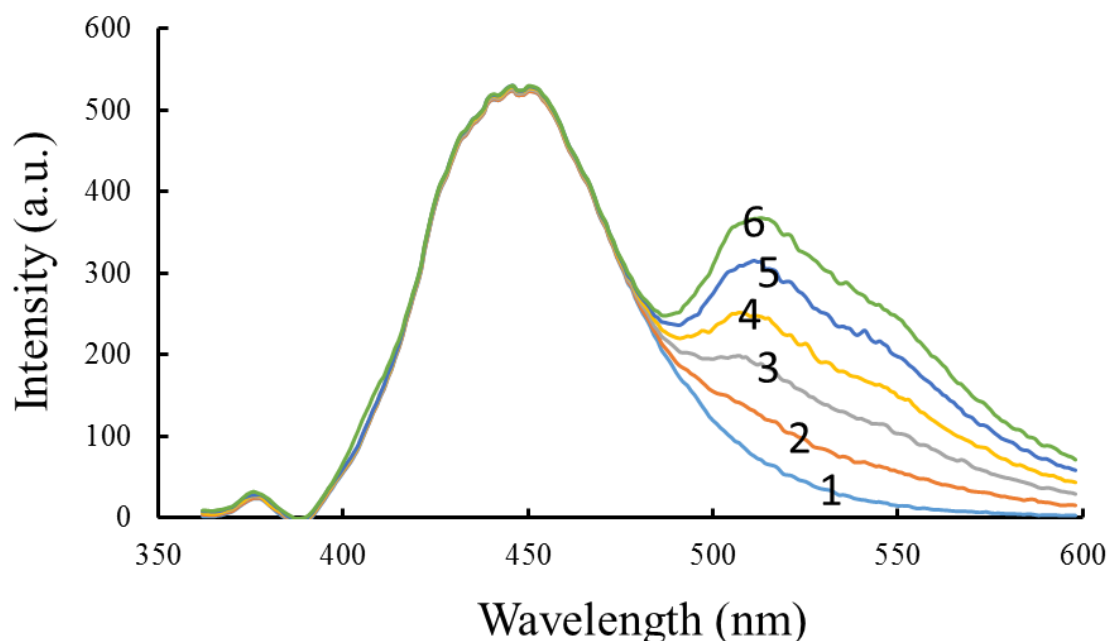


Figure 4.12: Measurements of energy transfer between $E2o^{1-173}$ -Alexa350 and E3 bound FAD. The excitation wavelength is 350 nm and the spectrum are recorded from 360 nm to 600 nm. The concentration of $E2o^{1-173}$ is 5 μ M in 50 mM KH_2PO_4 (pH=7.0). 1: the spectrum of $E2o^{1-173}$ -Alexa350 alone; 2: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 5 μ M E3; 3: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 10 μ M E3; 4: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 15 μ M E3; 5: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 20 μ M E3; 6: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 25 μ M E3. The emission intensity of $E2o^{1-173}$ -Alexa350 stay all the same, which means there is no energy transfer between Alexa350 fluorophore and FAD. (the results of $E2o^{144-386}$ -Alexa350 with E3 are the same, not shown).

Table 4.1: Protein modification checked by FT-MS

proteins	Peptides	Theoretical mass	Mass detected by FT-MS
LDo	Whole protein (lipoylated)	11246.2	938.1415 ¹²⁺
	Whole protein +TCEP	11248.2	938.3098 ¹²⁺
	Whole protein + DANS-As	11646.2	971.5631 ¹²⁺
E2o ¹⁴⁴⁻³⁸⁶	EAQNTCAMLTTFNEIDMSNIQEMR	931.0792 ³⁺	931.0831 ³⁺
	EAQNTCAMLTTFNEIDMSNIQEMR + N-(1-pyrene)maleimide	1030.1055 ³⁺	1030.1078 ³⁺
E2o ¹⁻¹⁷³	AVGDTVAEDEVVCEIETDK (lipoylated)	1105.9827 ²⁺	1105.9838 ²⁺
	AVGDTVAEDEVVCEIETDK (lipoylated) + N-(1-pyrene)maleimide	836.6838 ³⁺	836.6859 ³⁺
	AVGDTVAEDEVVCEIETDK (lipoylated) + 2* N-(1-pyrene)maleimide	935.7101 ³⁺	935.7142 ³⁺
E2o ¹⁻¹⁷³	Whole protein (lipoylated)		1003.9740 ²⁰⁺
	Whole protein + TCEP		1004.0759 ²⁰⁺
	Whole protein + DANS-As		1024.1036 ²⁰⁺

4.4 Conclusions from the fluorescence studies

Sub-micromolar K_{ds} were determined for binding E1o to E2o labeled with the dansyl group at dihydrolipoate, or to the two E2o didomains labeled with pyrene at their single cysteines. The fluorescence studies as well as the FRET experiments threw light on the loci of interactions.

By contrast, no evidence could be obtained for a stable E2o-E3 complex according to the FRET studies, suggesting that the FRET donor and acceptor are further than 60 Å from

each other. In view of the obligatory interaction of the dihydrolipoyl domain of E2o with the FAD binding site of E3, this result probably suggests a weak, perhaps short-lived interaction.

Perhaps counterintuitively, the E1o-E3 interactions falls somewhere in between the other two binary complexes. There is indeed a FRET observed between the TCDP site on E1o and the FAD site at E3. According to the HDX-MS results, the large distance may signal that the FAD site of E3 is near the amino terminal domain of E1o (the binding locus identified on E1o for both E2o and E3), but the current information is too low in resolution to decide on the likely locus from this experiment.

The lack of response of the TCDP fluorescence to addition of E2o proteins suggests that the strongest E2o binding sites on E1o are not at or near the ThDP binding loci, consistent with the HDX-MS studies suggesting that the new peripheral subunit binding site of E2o interacts with E1o at its N-terminal region, quite distant from the loci of ThDP and Mg^{+2} binding.

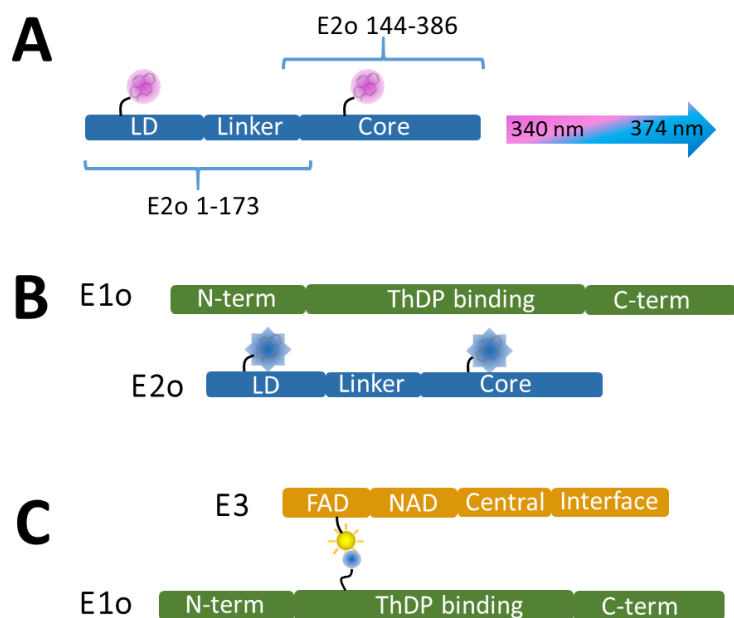


Figure 4.13: Summary of binary component interactions on the basis of fluorescence. (A) Two E2o constructs modified at the single cysteine on E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶. (B) Fluorescence titration experiments showed that E1o interacts with two E2o constructs through N-(1-Pyrene) maleimide modified single thiol group on E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ (C) FRET experiments proved that FAD binding region on E3 is close to the ThDP binding region of E1o.

References

- [1] Kapanidis A N, Weiss S. Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules[J]. The Journal of chemical physics, 2002, 117(24): 10953-10964.
- [2] Michalet X, Weiss S, Jäger M. Single-molecule fluorescence studies of protein folding and conformational dynamics[J]. Chemical reviews, 2006, 106(5): 1785.
- [3] Holmes K L, Lantz L M. Protein labeling with fluorescent probes[J]. Methods in cell biology, 2001, 63: 185-204.
- [4] Gentle I E, De Souza D P, Baca M. Direct production of proteins with N-terminal cysteine for site-specific conjugation[J]. Bioconjugate chemistry, 2004, 15(3): 658-663.
- [5] Lin C W, Ting A Y. Transglutaminase-catalyzed site-specific conjugation of small-molecule probes to proteins in vitro and on the surface of living cells[J]. Journal of the American Chemical Society, 2006, 128(14): 4542.
- [6] Jäger M, Nir E, Weiss S. Site-specific labeling of proteins for single-molecule FRET by combining chemical and enzymatic modification[J]. Protein Science, 2006, 15(3): 640-646.
- [7] Kim Y, Ho S O, Gassman N R, Korlann Y, Landorf E V, Collart F R, Weiss S. Efficient site-specific labeling of proteins via cysteines[J]. Bioconjugate chemistry, 2008, 19(3): 786.

- [8] R.F. Chen and H. Edelhoch, Eds, Biochemical Fluorescence: Concepts, Vol. 1 and 2, Dekker, New York, 1975
- [9] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, Plenum, New York, 1983
- [10] D.M. Jameson and G.D. Reinhart, Eds., Fluorescent Biomolecules, Plenum, New York, 1989
- [11] G.G. Hammes, in Protein-Protein Interactions, C. Frieden and L.W. Nichol, Eds., Wiley, New York, 1981, pp. 257-287
- [12] G. Weber, in Spectroscopic Approaches to Biomolecular Conformation, D.W. Urry, Ed., American Med. Assn., Chicago, 1970, pp. 23-31
- [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] Guevara, E.L., Yang, L., Birkaya, B., Zhou, J., Nemeria, N.S., Patel, M.S. and Jordan, F. (2017) Global view of cognate kinase activation by the human pyruvate dehydrogenase complex. Sci Rep. 10.1038/srep42760.

[16] Schuldiner S, Weil R, Robertson D E, Kaback H R. Microenvironment of the binding site in the lac carrier protein[J]. Proceedings of the National Academy of Sciences, 1977, 74(5): 1851-1854.

[17] Han M K, Lin P, Paek D, Harvey J J, Fuor E, Knutson J R. Fluorescence studies of pyrene maleimide-labeled translin: excimer fluorescence indicates subunits associate in a tail-to-tail configuration to form octamer[J]. Biochemistry, 2002, 41(10): 3468-3476.

[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] Weber G. Fluorescence of riboflavin and flavin-adenine dinucleotide[J]. Biochemical Journal, 1950, 47(1): 114.

[20] Heelis P F. The photophysical and photochemical properties of flavins (isoalloxazines)[J]. Chemical Society Reviews, 1982, 11(1): 15-39.

[21] Moe Jr O A, Lerner D A, Hammes G G. Fluorescence energy transfer between the thiamine diphosphate and flavine adenine dinucleotide binding sites in the pyruvate dehydrogenase multienzyme complex[J]. Biochemistry, 1974, 13(12): 2552-2557.

CHAPTER 5. Conformational Dynamics of 1-Deoxy-D-xylulose 5-phosphate Synthase Upon Ligand Binding Revealed by H/D Exchange Mass Spectrometry

5.1 Introduction

Due to the increasing spread of antibiotic resistance, there is a critical need to investigate novel antibiotic targets. The enzyme 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) is an attractive potential anti-infective target because it catalyzes formation of a branch point metabolite that is essential in human pathogen metabolism, yet is an enzyme not utilized by humans [1-3]. DXP, the product of DXPS synthase, is a key intermediate in the biosynthesis of ThDP, PLP and isoprenoids [1-3], all of which are essential for bacterial growth and survival; thus, inhibition of DXPS should result in widespread toxicity within a bacterial pathogen.

DXPS catalyzes the thiamin diphosphate (ThDP)-dependent decarboxylation of pyruvate (Py) and subsequent carboligation to D-glyceraldehyde 3-phosphate (GAP), reminiscent of the reactions of two essential mammalian enzymes, the human pyruvate dehydrogenase E1 component (PDHc-E1) and transketolase (TK). Development of probes targeting DXPS is therefore challenging, and few selective inhibitors of this enzyme are reported [4-11]. However, despite its similarities to other ThDP-dependent enzymes, previous mechanistic and structural studies have uncovered several distinguishing features of DXPS [12-17]. Unlike the classical ping-pong mechanism carried out by other ThDP-dependent enzymes, DXPS catalyzes a random sequential,

preferred order mechanism which is unprecedented in ThDP-dependent enzymology [12-16]. The first enzyme-bound tetrahedral pre-decarboxylation intermediate, C2-lactyl/ThDP (LThDP), is uniquely stable on DXPS in the absence of GAP, and ternary complex formation upon binding of GAP is required to trigger decarboxylation of LThDP [12, 14]. Accordingly, DXPS possesses a large active site, nearly twice the volume of TK and PDH-E1 active sites, presumably to accommodate ternary complex formation [9]. Lastly, DXPS is characterized by an unusual structure compared to other homodimeric enzymes in this class, with its active site located at the interface of domains I and II on each monomer, in contrast to other ThDP-dependent enzymes whose active sites are found at the dimer interface [17]. Taken together, the unique structural and mechanistic features of DXPS provide a basis for the selective inhibition of this enzyme, and new inhibitor classes are now emerging as potential tools to study this unique antimicrobial target [5, 10, 11].

Despite these advances, gaps remain in our understanding of the factors underlying the novel mechanism of DXPS. Structural studies of DXPS are scarce, limiting the potential to understand critical links between DXPS structure and mechanism that are needed to guide selective inhibitor design. We and others have hypothesized that binding of GAP triggers a conformational change that causes the destabilization and subsequent decarboxylation of LThDP [15, 16]. Evidence for GAP induced conformational changes has been reported on both the *Escherichia coli* and *Plasmodium falciparum* enzymes in pre-steady state CD [15] and fluorescence binding experiments, respectively [16]. While these studies lend support for the conformational dynamics of DXPS, they report

indirectly on potential conformational changes through the unusual behavior of mechanistic and binding data in the presence of GAP and offer little information about specific regions of conformational flexibility. Currently, the most direct evidence for the conformational flexibility of DXPS may come from the incomplete crystal structures of DXPS from *Deinococcus radiodurans* and *E. coli*. In the report by Xiang, et al. [17], in situ proteolysis was required for the successful crystallization of *E. coli* DXPS.

Interestingly, one of the cleaved *E. coli* DXPS segments overlaps with residues that are not observable in the *D. radiodurans* structure, suggesting these regions are conformationally flexible [17]. Structural studies of other ThDP-dependent enzymes suggest they may also undergo conformational change along their reaction coordinates [18-20]. However, the intriguing evidence for conformational dynamics of DXPS, in conjunction with its unprecedented mechanism, suggests that DXPS can exist in distinct conformations that are potential targets for selective inhibition. Investigating DXPS structure is a necessary step toward identifying targetable conformations. Unfortunately, structural studies of DXPS structure are severely limited by the potential conformational flexibility of the enzyme which impedes classic crystallographic studies. Therefore, there is a pressing need to develop techniques to directly probe the dynamics of DXPS along its reaction coordinate, which could reveal unique conformations that are starting points for selective inhibitor design.

The goal of this study was to probe the conformational dynamics of *E. coli* DXPS in the presence and absence of GAP, DXP, and a stable Py mimic, methyl acetylphosphonate (MAP), using hydrogen-deuterium exchange mass spectrometry (HDX-MS). Our results

reveal, for the first time, the conformational dynamics of three previously uncharacterized segments of *E. coli* DXPS supporting the utility of this technique to study DXPS structure and mechanism simultaneously. In the absence of any ligands, these regions of DXPS display a bimodal pattern of deuterium uptake, implying EX1 kinetic behavior, in addition to EX2 kinetics, a finding for a protein in its native, ligand-free state that is increasingly prevalent in the recent literature. [22-25]. Although the rates of deuterium uptake are unaffected by ligand binding, the rates of unfolding in the EX1 regions are ligand-dependent, suggesting that the closed conformation of DXPS in the presence of MAP is distinct from the conformational changes induced by GAP and DXP. The observation of conformational changes induced on DXPS by GAP binding provides the highest resolution data collected to date to support the role of conformational dynamics in GAP-induced LThDP decarboxylation. Collectively these results provide the first direct evidence for the conformational changes of DXPS in response to substrate, product, and substrate analog binding, the latter of which offers insights into conformation changes induced by inhibitors based on this substrate analog scaffold. This report also constitutes the first example of a ThDP-dependent enzyme that displays EX1 kinetics in the flexible regions of the protein as observed by HDX-MS, adding to the short list of proteins that display EX1 kinetic behavior under physiological conditions, and providing crucial insight into the novel mechanism of DXPS.

5.2 Materials and methods

Deuterium oxide (D₂O) 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 MgCl₂, 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 D₂O buffer to yield a final concentration of 95 % D₂O at pH 8.0. D₂O buffer was prepared the same way as DXPS exchange buffer except 99.9% D₂O 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

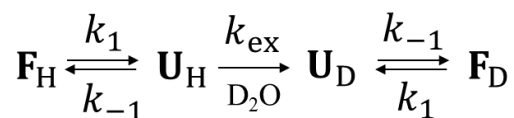
The frozen deuterated sample was quickly thawed and loaded with an ice-cold syringe into a 20- μ l sample loop inside the refrigeration system. The protein sample (40 pmol) was carried by a 0.2 ml/min digestion flow (0.1% formic acid) into an immobilized pepsin column (Poroszyme Immobilized Pepsin Cartridge, 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 Mechanisms of EX1 and EX2 kinetics

Two mechanisms have been proposed to explain hydrogen/deuterium exchange under physiological conditions in proteins. The first one suggests that exchange may occur from the folded form, the second one suggests that partial local unfolding occurs before

exchange can take place [26, 27]. Under physiological conditions, the second mechanism (Equation 5.1) appears to be the predominant pathway for H/D 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 . This unfolding process exposes backbone amide hydrogens which can subsequently undergo exchange with D_2O . The rate constants k_1 , k_{-1} , and k_{ex} describe the kinetics of unfolding, refolding, and hydrogen exchange steps, respectively [26]. The hydrogen exchange in this model can be described by two different kinetic schemes, EX1 and EX2. When the exchange is governed by EX1 kinetics, the unfolding rate is much faster than the refolding rate ($k_1 \gg k_{-1}$). Under these conditions, all of the amide hydrogens exchange with deuterium in the unfolded (open) state before refolding occurs. As a result, EX1 kinetics gives rise to two distinct mass envelopes. The lower mass envelope represents the folded (closed, \mathbf{F}_H) undeuterated state, and the higher mass envelope the unfolded (open, \mathbf{U}_D), deuterated state. In EX2 kinetics, the protein refolding rate is much faster than the unfolding rate ($k_1 \ll k_{-1}$), therefore the unfolding has to occur multiple times before a successful exchange reaction takes place. EX2 kinetics leads to a single isotopic distribution, which gradually shifts to a higher m/z range over time. While the majority of native proteins behave according to EX2 kinetics under physiological conditions, there are increasing examples reporting EX1 kinetic behavior in

the absence of ligand. The EX1 behavior must be confirmed by rigorous control experiments, thoroughly described by Engen and coworkers [28]. Under our experimental conditions, we can indeed attribute our observations to EX1 behavior in the three regions of DXPS identified. The observation of EX1 H/D exchange kinetics on DXPS under physiological conditions could provide important insights into the conformational dynamics of DXPS along its reaction coordinate.

5.3.2 Overview of HDX patterns on DXPS

The time dependence of hydrogen-deuterium exchange of the backbone amide protons of DXPS was studied over a 30 min time course (20 s, 1, 2, 3, 5, 10 and 30 min). On-line digestion by pepsin followed by LC-MS analysis under the selected HDX conditions yielded 56 peptides, many of which were partially overlapping, providing 93.7% sequence coverage (Table 5.1). The redundancy of peptides detected is very important because it provides an internal control for consistency of sequence assignment. Throughout the entire protein, overlapping peptides display similar HDX kinetics. While most of the overlapping peptides of DXPS exclusively display EX2 exchange kinetics, there are three distinct regions of the protein in which each of the overlapping peptides are characterized by EX1 kinetics (results reported below). These observations serve as an additional control to prevent false assignment of exchange behavior. Importantly, the high sequence coverage achieved in this study allowed us, for the first time, to monitor structural changes throughout the entire protein including peptides 183-238 and 292-317, which are not observed in the only available crystal structure of *E. coli* DXPS [17]. This new structural

information could offer novel mechanistic insights to guide the design of selective DXPS inhibitors.

The average deuterium uptake percentage by DXPS at 30 min is illustrated in Fig. 2. The peptides spanning residues 183-238 and 292-317 are represented schematically as they are important for this study, but are missing from the published structure of DXPS (see Figs.5.1). The overall rate of deuterium incorporation was low on this enzyme; many regions underwent less than 35% deuteration after a 30 min exchange (see Figs.5.1). The deuterium incorporation pattern clearly indicates that the greatest extent of H/D exchange occurs in domain I, suggesting that this is the most dynamic of the three domains of the DXPS monomer. In contrast, the areas with the least extent of deuterium incorporation are located in domains II and III at the dimer interface. These findings are intriguing as the crystal structure of *E. coli* DXPS has revealed that its active site is located between domains I and II, in contrast to other homodimeric ThDP-dependent enzymes in which the active sites reside at the dimer interface [17].

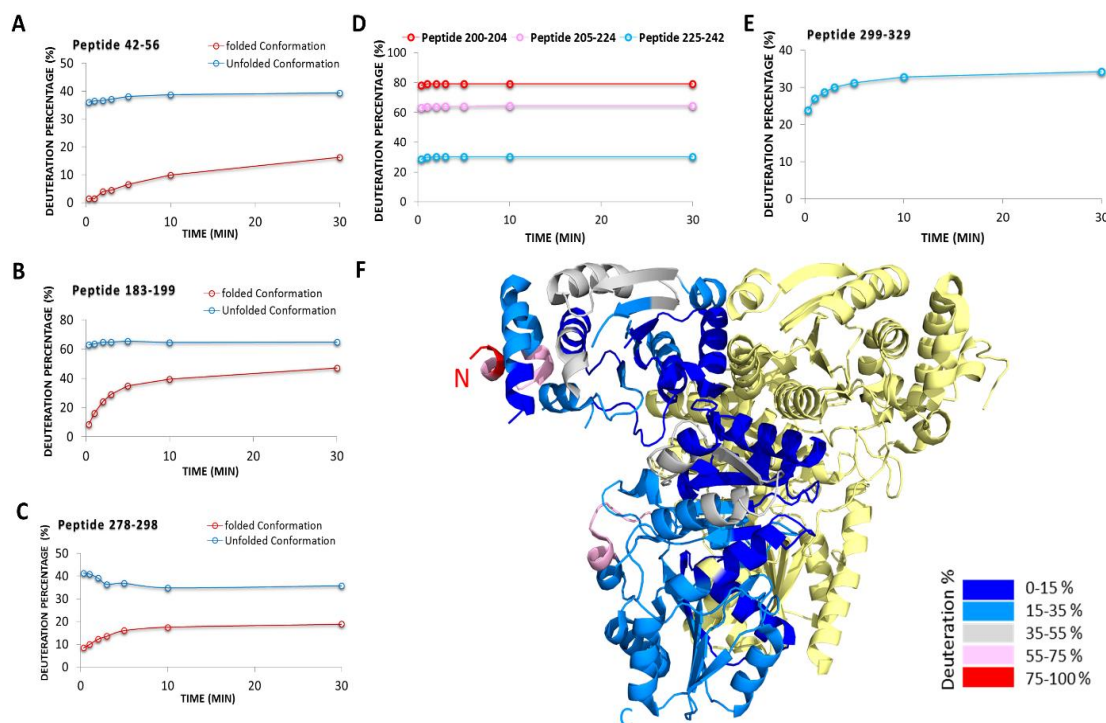


Figure 5.1: Average deuterium uptake by key peptides of *E. coli* DXPS. (A-C) Time course of deuterium uptake by peptides exhibiting a bimodal isotopic distribution. The centroid masses of lower (red traces) and higher (blue traces) mass envelopes were fit and calculated individually. Since the centroid mass of lower mass envelope also changes with increasing exposure time, the displayed kinetics is likely a mixture of EX1 and EX2 kinetics. (D and E) The deuterium uptake plots for regions comprising residues 183-238 and 292-317, which are missing from the crystal structure [18]. (F) The average percentage of deuterium incorporation at the 30 min time point is mapped onto the crystal structure of *E. coli* DXPS. For clarity, percentage deuterium uptake is represented for one DXPS monomer, while the other monomer is shown in yellow.

5.3.3 HDX-MS detects local structural dynamics of DXPS.

First, the conformational flexibility of full length DXPS in the absence of ligands was examined. Deuterium uptake for most peptides identified under our HDX conditions revealed a single isotopic distribution that increased in mass over time, suggesting that these regions display common EX2 kinetic behavior. Interestingly, three regions (42-58, 183-199, and 278-298) displayed a bimodal isotopic distribution, consistent with EX1 kinetics (Fig. 5.2). Because peptides 42-56 and 51-58 have overlapping residues, they will be referred to as peptide 42-58 in the following discussion. Addition of the ligands selected for this study had little effect on the deuterium incorporation by the peptides of DXPS that were characterized by EX2 kinetics.

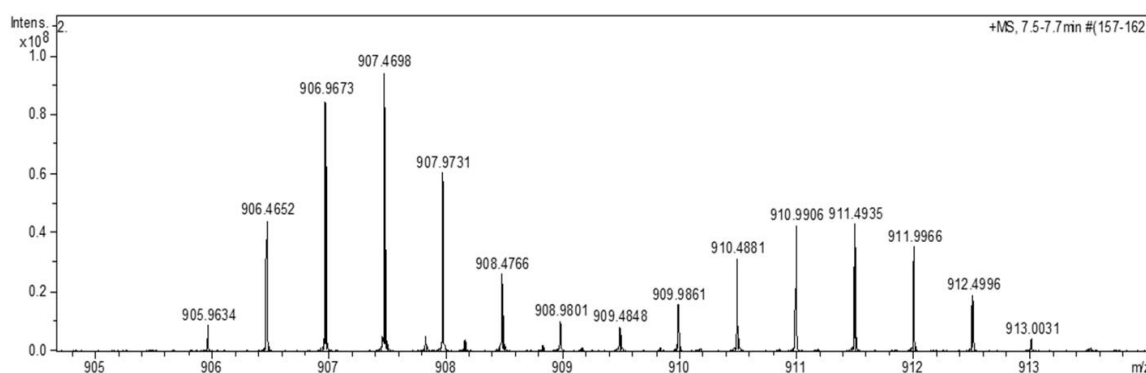


Figure 5.2: Mass spectrum of peptides 183-199 at the 1-min HDX time point in the presence of GAP.

Regions that were affected by the addition of ligands were those that display EX1 kinetics. Notably, unlike most of the reported bimodal patterns which are induced upon adding a binding partner [29, 30], the bimodal patterns of these three regions are displayed by DXPS even in the absence of bound ligands. Although residues 183-199 and 292-298 (part of the 278-298 peptide) are missing from the crystal structure, the positions of the adjacent N-

terminal residues of these peptides (which are observed in the crystal structure) suggest that the three regions displaying EX1 kinetics are close to each other. (Fig. 5.1). Given their close spatial proximity to one another and the active site, it is possible that the conformational changes accounting for the observed EX1 kinetics occur simultaneously in all three regions (Fig. 5.1).

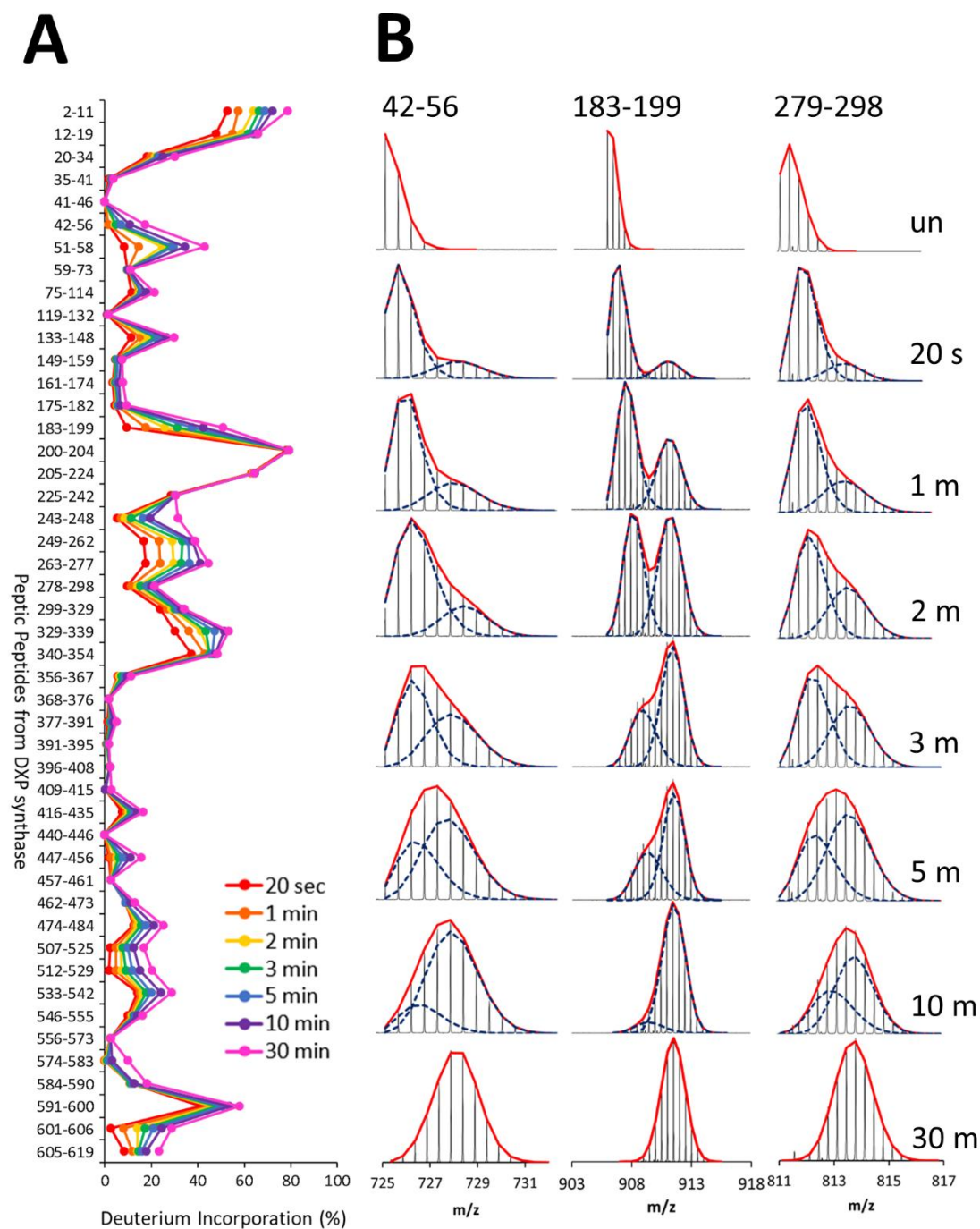


Figure 5.3: DXPS deuterium uptake change over 30 min. (A) Deuterium incorporation plot of full-length *E. coli* DXPS across seven different time points (20s and 1, 2, 3, 5, 10

and 30 min). (B) HDX-MS spectra of selected peptides (42-56, 183-199, and 278-298) charge envelopes (red lines) at increasing deuterium exchange time shows the presence of two isotopic distributions. Gaussian deconvolution of the mass spectrum reveals the presence of two isotopic distributions (blue dot lines).

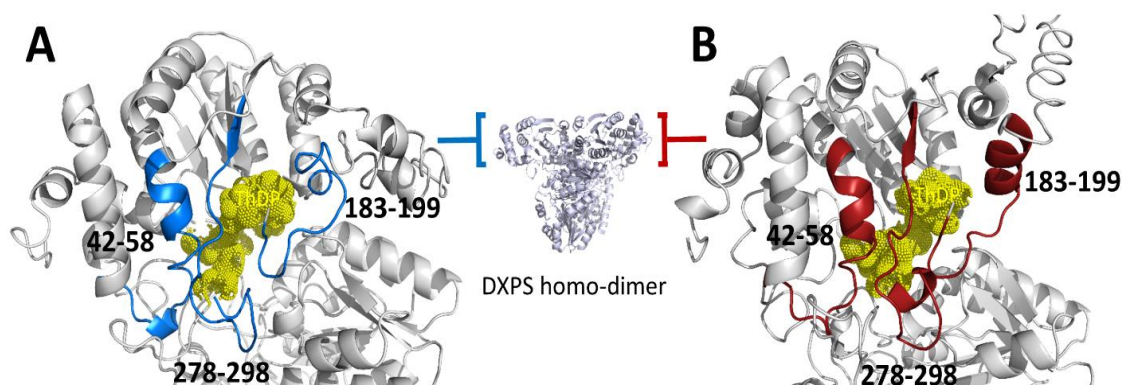


Figure 5.4: Spatial proximity of all regions displaying EX1 kinetic behavior in the DXPS structure. The position of all regions displaying EX1 kinetics on domain I of *E. coli* DXPS indicated the spatial proximity of these regions near the ThDP binding site of DXPS. (A) The predicted open state of all EX1 kinetic regions shown in blue. Yellow dot spheres indicate the ThDP binding pocket. All EX1 kinetic regions contain multiple residues, which are involved in ThDP binding. (B) Closed state of all EX1 kinetic regions shown in red. The open and closed structures were calculated with I-TASSER. *E. coli* DXPS structure (Protein Data Bank ID code 2o1s) was assigned as a restraint and template to guide I-TASSER modeling.

To further characterize the EX1 kinetics displayed by several peptides in our initial experiments, we carried out an H/D exchange time course study over 30 min (20 s, 1, 2, 3, 5, 10, and 30 min). Because of the relatively low degree of deuteration at 20 s (< 10 %)

and high degree at 30 min (> 65 %), peptides were observed at 20 s and 30 min, and additional shorter or longer time points were not deemed necessary, hence were not acquired for this study.

We considered two possible explanations for our observation of the bimodal distribution that could result from EX1 kinetics on DXPS, in which two distinct conformations interconvert during the HDX incubation period. In this case, it appears that the equilibration rate between closed and open conformations of the corresponding peptide is slower than the HDX rate, which indicates that both conformations are fairly stable under physiological conditions. Since DXPS is naturally present as a homo-dimer, a plausible explanation for the EX1 kinetic behavior of the three regions is that the DXPS homo-dimer is asymmetric, suggesting that the two subunits of 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 are in a disordered region in the DXPS structures from *E. coli* and *D. radiodurans* [17], 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. Interestingly, DXPS is the first ThDP-dependent enzyme to display this behavior, suggesting that this conformational flexibility may have an important role in the unique mechanism catalyzed by DXPS.

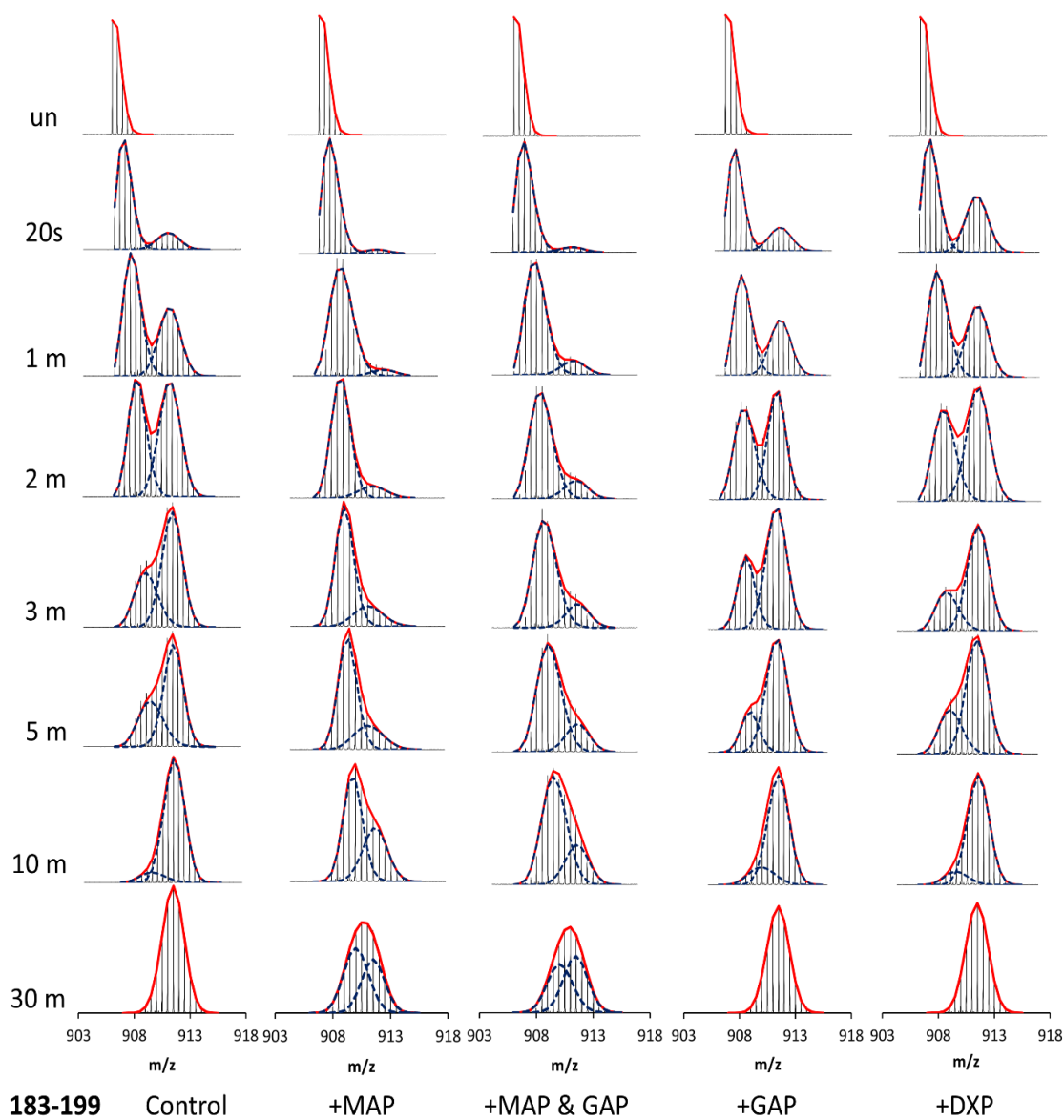


Figure 5.5: Complete time course of HDX-MS spectra of peptides 183-199 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

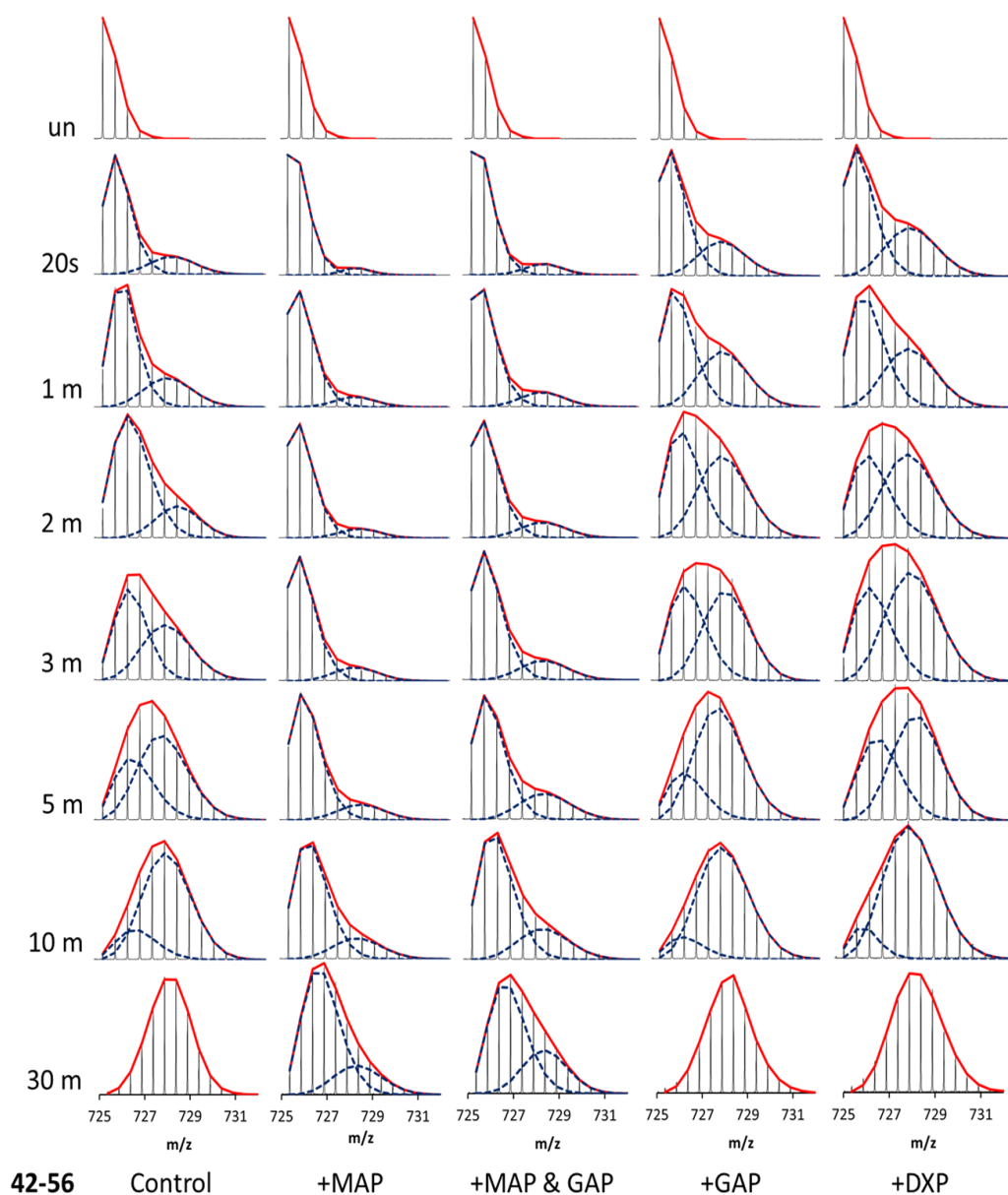


Figure 5.6: Complete time course of HDX-MS spectra of peptides 42-56 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

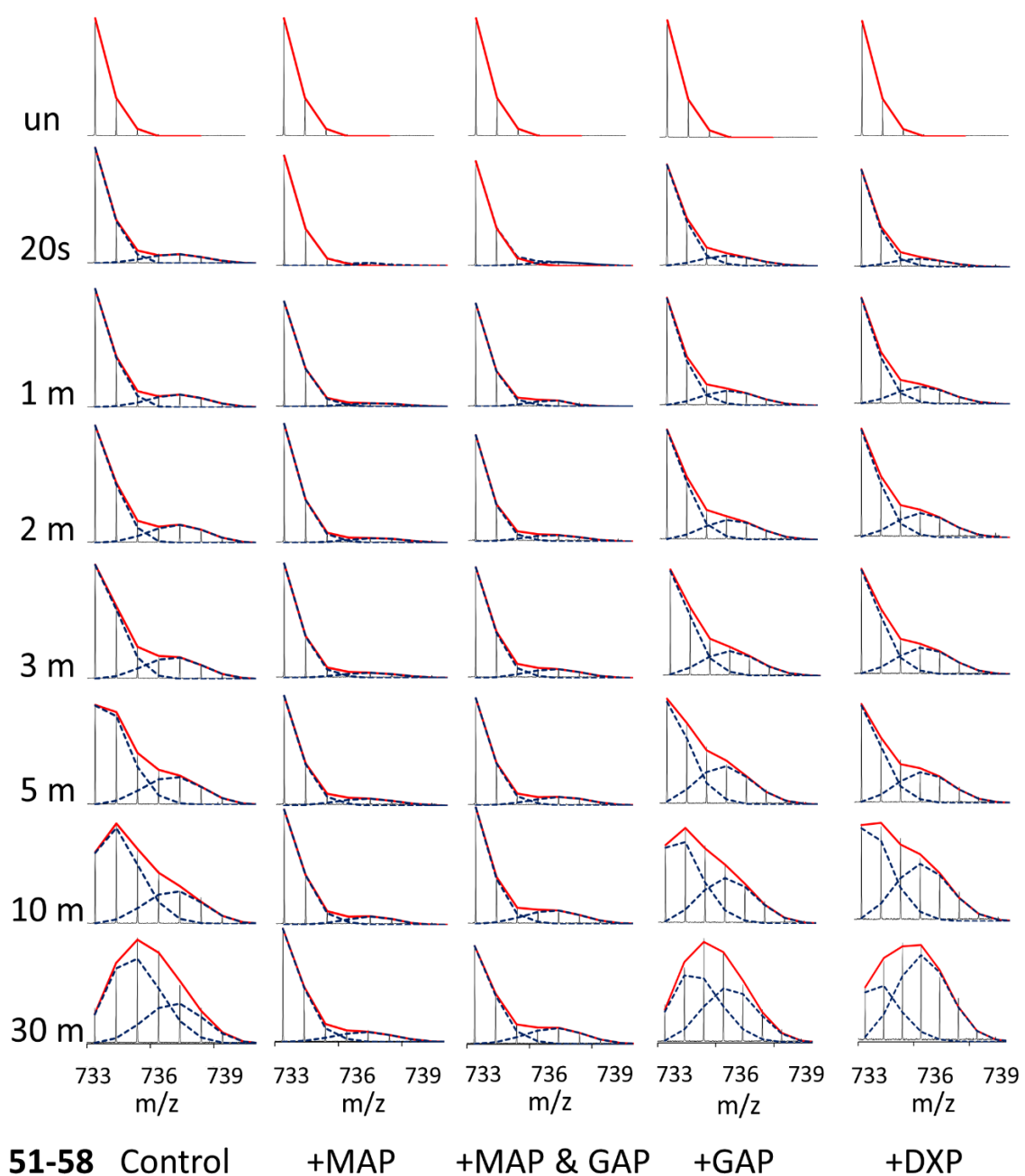


Figure 5.7: Complete time course of HDX-MS spectra of peptides 51-58 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines

indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

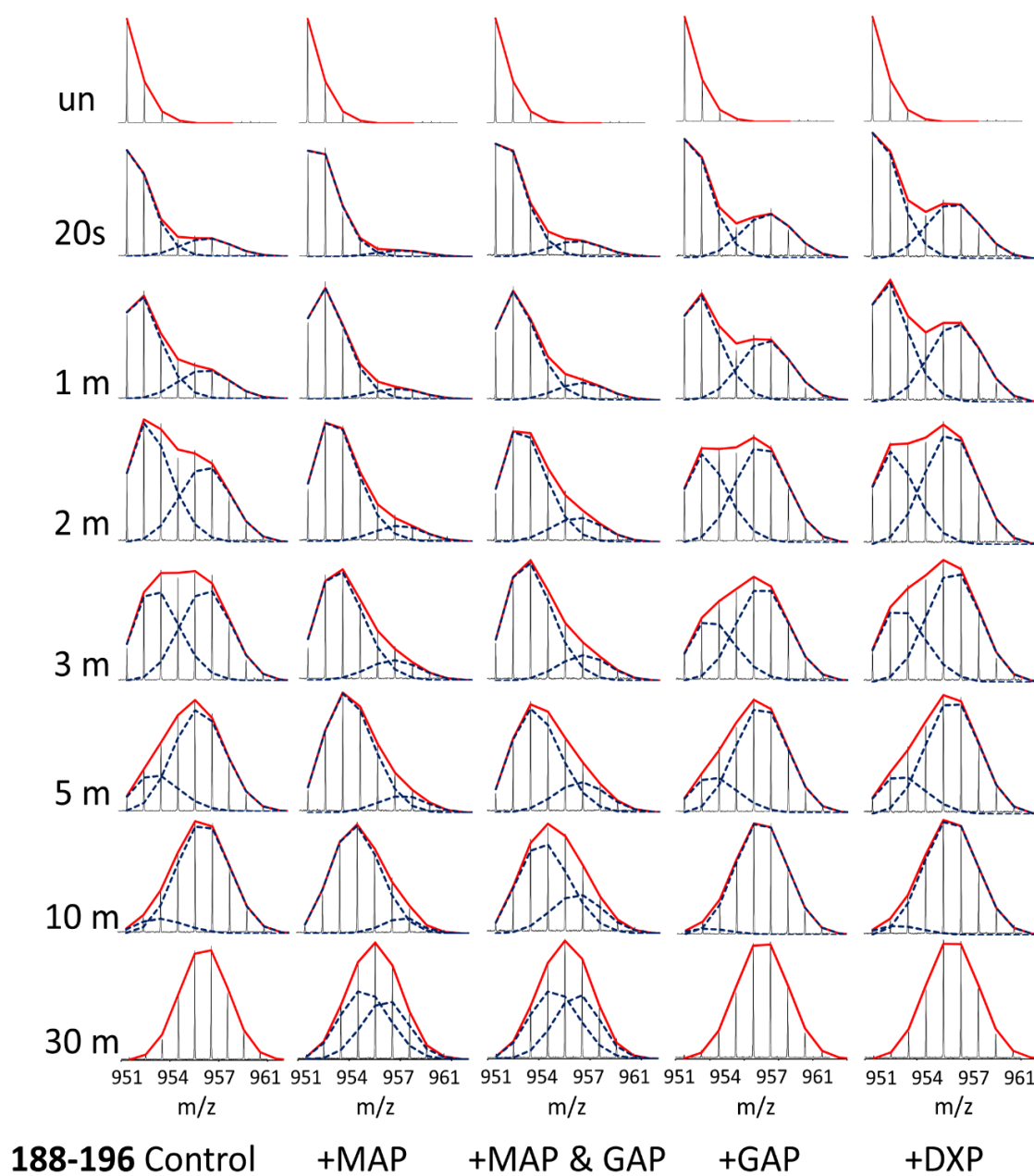


Figure 5.8: Complete time course of HDX-MS spectra of peptides 186-196 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

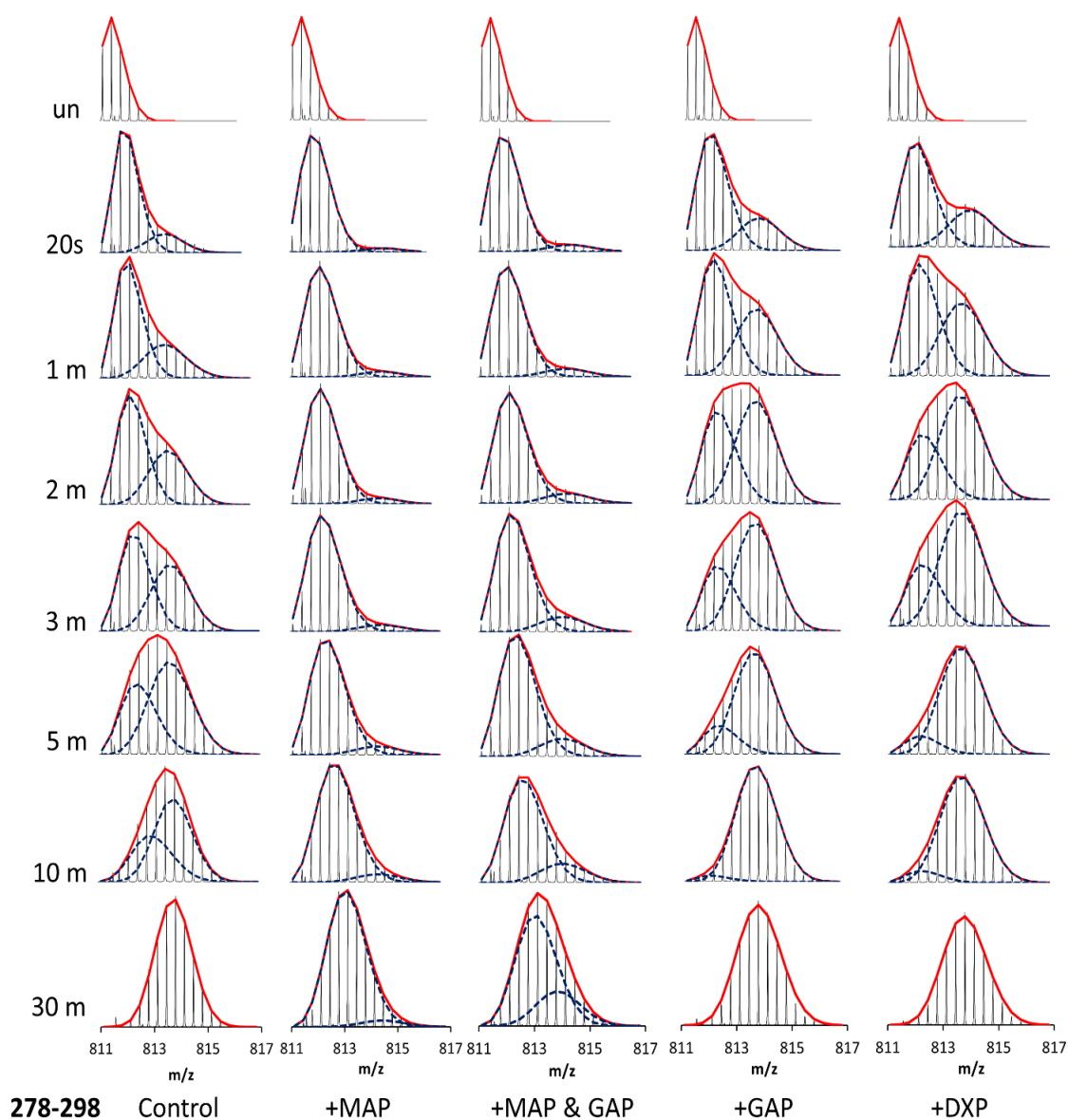


Figure 5.9: Complete time course of HDX-MS spectra of peptides 278-298 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

Careful examination of the three regions displaying EX1 kinetics in the absence of ligand revealed several interesting behaviors. First, these three regions all adopted mixed EX1/EX2 kinetics. In addition to the distinct bimodal EX1 signature described above, the lower mass envelope of each peptide gradually moved to the high mass range over time (see Figs. 5.5-5.9). As these peptides are long, they may contain a portion that has undergone exchange via the EX1 mechanism 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 the 20 s time point indicated a 4 Da difference in peptide 42-56 (30.8% of residues), 8 Da in peptide 183-199 (53.5% of residues) and 6 Da difference in peptide 278-298 (35.3% of the residues) being involved in the unfolding event (see Figs. 5.1).

5.3.4 DXPS binds and responds differently to substrate, substrate analogue and product

MAP is a stable mimic of Py and known inhibitor of ThDP-dependent enzymes that use Py as substrate [13]. Upon binding of MAP to DXPS, a stable C2 α -phosphonolactylthiamin diphosphate (PLThDP) intermediate is formed and observed by CD. Addition of MAP (1 mM), and formation of the pre-decarboxylation intermediate mimic on DXPS prior to H/D exchange significantly slows down the unfolding of all three EX1 signature regions as mentioned above (residues 42-58, 183-199, and 278-298). The slowdown is apparent in the HDX-MS spectra (Fig. 5.5-5.9), wherein a bimodal distribution is observed but with a greater contribution from the lower mass envelope relative to DXPS in the absence of ligand. This phenomenon can be quantified by the “slowdown factor” (24), which is calculated using the rate constant (k_u) or half-life ($t_{1/2}$) for unfolding. The slowdown factor is simply the $t_{1/2}$ of unfolding in the presence of ligand divided by the $t_{1/2}$ of unfolding in the absence of ligand. A summary of the unfolding rate constants, half-lives and slowdown factors of each of the peptides exhibiting an EX1 signature is presented in Table 5.1. It is clear that MAP gives rise to the most pronounced slowdown factor. A likely explanation for this observation is that the formation of the stable PLThDP intermediate in the presence of MAP [31, 32], effectively traps the enzyme in a pre-decarboxylation-like state that exists in the closed conformation (illustrated by the lower mass envelope in the EX1 kinetic profiles). As the PLThDP intermediate is a mimic of LThDP formed upon binding of pyruvate [13, 31, 32], it is conceivable that the pre-decarboxylation LThDP intermediate could also exist in the closed conformation. This is a key finding that warrants further investigation as it is currently unknown how DXPS stabilizes LThDP in the active site. The mass envelope distribution difference in peptides displaying EX1 kinetics, suggests that binding of MAP could induce large-scale conformational changes around the active

site of DXPS, and favor the closed conformation. Such conformational changes induced by this substrate analog give insights into the mechanism of inhibition by the alkylacetylphosphonates inhibitor class based on this substrate analog which has significant implications for future inhibitor design.

The HDX behaviors of DXPS in the presence of GAP, GAP + MAP, or DXP were also examined. Upon binding of GAP or DXP to DXPS, the regions 42-58, 183-199, and 278-298 display unique unfolding events as shown in Fig. 5.5-5.9. The increased abundance of the higher mass envelope of the bimodal distribution suggests that both GAP and DXP can induce conformational changes in the regions displaying EX1 behavior, resulting in the formation of an open conformation. As shown in Table 2, the half-life of unfolding is slightly decreased on all peptides compared to DXPS in the absence of ligand; however, these rates of unfolding are still significantly greater than those observed in the presence of MAP. We next examined the role of GAP to induce a conformational rearrangement of DXPS subsequent to MAP binding, to model events in GAP-induced LThDP decarboxylation. The HDX-MS behavior of DXPS with MAP bound was compared in the absence and presence of GAP. All peptides consistently exhibited EX1 kinetics profiles, as induced by MAP, even in the presence of GAP. However, we observed a slight increase in the isotopic distribution favoring the higher mass envelope of all EX1 kinetics peptides in the presence of GAP compared to DXPS in the presence of MAP. (Fig. 5.5-5.9). Accordingly, the rates of unfolding of the three EX1 kinetics peptides were found to be about 2 to 3 times faster than with MAP alone (Table 5.1).

Table 5.1 Rate constant of unfolding and the unfolding half-life for the four peptides displaying EX1 kinetics.

Peptide	42-56			51-58		
	k (min ⁻¹)	t _{1/2} (min)	slowdown factor	k (min ⁻¹)	t _{1/2} (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 ⁻¹)	t _{1/2} (min)	slowdown factor	k (min ⁻¹)	t _{1/2} (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

Importantly, these results provide additional evidence for the unique ability of GAP to trigger conformation changes on DXPS both in the absence and presence of a donor substrate mimic. This is consistent with a model (Fig. 5.3) in which LThDP is stabilized in the closed conformation prior to binding of GAP. Upon binding of GAP to the LThDP-DXPS complex (mimicked by the PLThDP-DPXS complex), DXPS undergoes an unfolding event, implying the possibility that GAP induces an open conformation to trigger decarboxylation of LThDP. Further, an unfolded conformation is favored on DXPS in the presence of the product, consistent with an open state during product release. Taken together, the results suggest that the 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 picture of the dynamics of the DXPS structure in response to the binding of different ligands that is the most detailed to date. The high sequence coverage of DXPS in our experiments (93.7%) enabled us to study the flexibility of the entire enzyme, including the two regions (residues 183-238, and 292-317), which have never been observed by X-ray crystallography.

HDX-MS data showed that domain I of DXPS (residues 1-319) displays unusual conformational flexibility that appears to play important roles in substrate recognition. Three regions (residues 42-58, 183-199, and 278-298) near the active center displayed an EX1 signature in both ligand-free and ligand-associated states (Fig. 5.2), and the half-lives of unfolding of these peptides were different depending upon the identity of the ligand-DXPS complex (Table 5.1). In the presence of MAP, DXPS appears to favor a closed conformation in those EX1-displaying peptides, and favors the unfolded conformation in the presence of D-GAP or DXP. Thus, we propose that the closed conformation around the active center of DXPS is critical for stabilization of the LThDP-DXPS complex, while the open conformation may play a role in the GAP-triggered decarboxylation of LThDP and in product release. Intriguingly, the EX1 exchange kinetics displayed by DXPS appears to be a unique characteristic of this enzyme compared to other enzymes in its class as this behavior has not been reported for any other ThDP-dependent enzyme to date. Although MAP induced-conformational changes have also been observed on the *E. coli* PDH-E1 component [32], the specific dynamics reported in this study in the presence of the PLThDP intermediate on DXPS are novel. These intriguing findings support a distinct mechanism of DXPS. The discovery of MAP-induced conformational changes also suggests the

potential utility of this HDX MS method for the characterization of conformational effects of inhibitor binding to DXPS. Inhibitors that are selective for DXPS over other ThDP-dependent enzymes could induce unique conformational changes on this enzyme resulting in stabilization of a native intermediate conformation or by promoting an unproductive form of the enzyme. Given the difficulty of investigating DXPS structure via traditional methods, HDX MS may be useful for the characterization of DXPS inhibitor mechanisms toward the design of more selective and potent compounds.

This study has marked an important step forward in our understanding of the novel mechanism of DXPS. Our findings reveal that several regions in domain I of DXPS (residues 42-58, 183-199, and 278-298), some of which have not been previously characterized structurally, play an important role in substrate recognition and enzymatic catalysis. Given that the unique H/D exchange behavior and conformational flexibility of DXPS appear to be distinct characteristics of this enzyme, it is possible that these dynamics are major factors underlying the novel mechanism of DXPS. These findings have revealed a path for further investigation of the structural and functional roles of the regions of DXPS displaying EX1 behavior in order to gain a deeper understanding of DXPS catalysis. Moreover, a better understanding of the structure and dynamics around the DXPS active center may also lead to design of more potent and selective inhibitors of this important enzyme target in bacterial metabolism. Importantly, the work presented here illustrates the value of HDX-MS methods as a useful tool for mapping conformational changes of flexible regions in DXPS, which are so far undetectable by traditional crystallographic methods.

Thus, this study serves as a model for the investigation of other relatively large, highly dynamic protein structures, which are difficult to capture by X-ray or NMR methods.

Table 5.2 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	LRLPKESLPKL
948.5634	948.5625	1.0	35	41	LRRYLLD
650.3108	650.3104	0.6	41	46	DSVSR
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	LTVALHYVYNTPFDQ
1396.6553	1396.6532	1.5	64	74	HYVYNTPFDQL
4636.4686	4636.4549	3.0	75	114	IWDVGHQAYPHKILTGRDRKIGTIRQKGGHLPFPWRGESE
1259.6217	1259.6226	-0.8	119	132	SVGHSSTISAGIG
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	MSISENVGALNNHQAQL
1263.6801	1263.6804	-0.3	188	199	NVGALNNHQAQL
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	YSSLREGGKKVFSGVPPIKE
1727.9805	1727.9803	0.1	209	224	REGGKKVFSGVPPIKE
2021.1577	2021.1576	0.1	225	242	LLKRTEHIKGMVVPGL
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	LHIMTKKGRGYEPAEKDPIF
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	YTGYYHNDGPSA
1187.6657	1187.6644	1.1	474	484	VRYPRGNAVGV
1975.0549	1975.0569	-1.0	507	525	AILNFGTLMPEAAKVAESL
1815.9510	1815.9521	-0.6	512	529	GTLMPAAKVAESLNATL
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	MAHRKPPVPL
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

References

- [1] Rodriguez-Concepcion M. (2014) The MEP pathway: a new target for the development of herbicides, antibiotics, and antimalarial drugs. *Curr Pharm Des.* 10, 2391-2400.
- [2] Du Q, Wang H, and Xie J. (2011) Thiamin (Vitamin B₁) biosynthesis and regulation: a rich source of anti-microbial drug targets? *Int J of Biol Sci.* 7(10), 41-52.
- [3] Laber B, Maurer W, Scharf S, Stepusin K, and Schmidt FS. (1999) Vitamin B₆ biosynthesis: formation of pyridoxine 5'-phosphate from 4-(phosphohydroxy)-L-threonine and 1-deoxy-D-xylulose-5-phosphosphate by PdxA and PdxJ protein. *FEBS Lett.* 449, 45-8.
- [4] Matsue Y, Mizuno H, Tomita T, Asami T, Nishiyama M, and Kuzuyama T. (2010) The herbicide ketoclofazone inhibits 1-deoxy-D-xylulose 5-phosphate synthase in the 2-C-methyl-D-erythritol 4-phosphate pathway and shows antibacterial activity against *Haemophilus influenzae*. *J Antibiot.* 63(10), 583–588.
- [5] Smith JM, Vierling RJ, and Freel Meyers C. (2012) Selective inhibition of *E. coli* 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *MedChemComm.* 3(65), 65-7.
- [6] Hayashi D, Kato N, Kuzuyama T, Sato Y, and Ohkanda J. (2013) Antimicrobial *N*-(2-chlorobenzyl)-substituted hydroxamate is an inhibitor of 1-deoxy-D-xylulose 5-phosphate synthase. *ChemComm.* 49, 5535–5537.
- [7] Witschel M, Rühl F, Niggeweg R, and Newton T. (2013) In search of new herbicidal inhibitors of the non-mevalonate pathway. *Pest Manage Sci.* 69(5), 559–563.

- [8] Masini T, Pilger J, Kroezen BS, Illarionov B, Lottmann P, Fischer M, Griesinger C, and Hirsch AKH. (2014) *De novo* fragment-based design of inhibitors of DXS guided by spin-diffusion-based NMR spectroscopy. *Chem Sci.* 5, 3543–3551.
- [9] Morris, F., Vierling, R.J., Boucher, L., Bosch, J., and Freel Meyers, C. L. (2013) DXP synthase-catalyzed C-N bond formation: Nitroso substrate specificity studies guide selective inhibitor design. *ChemBioChem.* 14, 1309-1315.
- [10] Smith JM, Warrington NV, Vierling RJ, Khum ML, Anderson WF, Koppisch AT, and Freel Meyers CL. (2014) Targeting DXP synthase in human pathogens: enzyme inhibition and antimicrobial activity of butylacetylphosphonate. *J Antibiotics.* 67, 77-83.
- [11] Bartee D, Morris F, Al-khouja A, and Freel Meyers CL. (2015) Hydroxybenzaldoximes are D-GAP-competitive inhibitors of *E. coli* 1-deoxy-D-xylulose 5-phosphate synthase. *ChemBioChem.* 16, 1771-81.
- [12] Eubanks LM and Poulter CD. (2003) *Rhodobacter capsulatus* 1-deoxy-D-xylulose 5-phosphate synthase: Steady-state kinetics and substrate binding. *Biochemistry.* 42(4), 1140-9.
- [13] Brammer LA, Smith JM, Wade H, and Meyers CF. (2011) 1-deoxy-D-xylulose-5-phosphate synthase catalyzes a novel random sequential mechanism. *J Biol Chem.* 286(42), 36522-31.
- [14] Patel H, Natalia NS, Brammer LA, Freel Meyers CL, 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. *JACS.* 134, 18374-79.

- [15] Brammer Basta LA, Patel H, Kakalis L, Jordan F, and Freel Meyers CL. (2014) Defining critical residues for substrate binding to 1-deoxy-D-xylulose-5-phosphate synthase- active site substitutions stabilize the predecarboxylation intermediate C2 α -lactylthiamin diphosphate. *FEBS J.* 281(12), 2820-37.
- [16] 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.
- [17] Xiang S, Usunow G, Lange G, Busch M, Tong L. (2007) Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *J Biol Chem.* 282, 2676-82
- [18] Wille G, Meyer D, Steinmetz A, Hinze E, Golbik R, and Tittmann K. (2006) The catalytic cycle of a thiamin diphosphate enzyme examined by cocrystallography. *Nat Chem Biol.* 2(6), 324-8.
- [19] Meyer D, Neumann P, Parthier C, Friedemann R, Nemeria N, Jordan F, and Tittmann K. (2010) Double duty for a conserved glutamate in pyruvate decarboxylase: evidence of the participation in stereoelectronically controlled decarboxylation and in protonation of the nascent carbanion/enamine intermediate. *Biochemistry.* 49(37), 8197-8212.
- [20] Fiedler E, Throell S, Sandalova T, Golbik R, Konig S, and Schneider G. (2002) Snapshot of a key intermediate in enzymatic thiamin catalysis: crystal structure of the α -carbanion of (α,β -dihydroxyethyl)-thiamin diphosphate in the active site of transketolase from *Saccharomyces cerevisiae*. *Biochemistry.* 99(2), 591-5.
- [21] Nemeria NS, Shome B, DeColli AA, Heflin K, Begley TP, Meyers CF, Jordan F. (2016) Competence of thiamin diphosphate-dependent enzymes with 2'-methoxythiamin

diphosphate derived from Bacimethrin, a naturally occurring thiamin anti-vitamin. *Biochemistry* 55(7), 1135-48.

[22] Weis DD, Wales TE, Engen JR, Hotchko M, Ten Eyck LF. (2006) Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J Am Soc Mass Spectrom.* 17(11): 1498-509.20

[23] 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.

[24] Woodward, C., Simon, I., and Tuchsén, E. (1982) Hydrogen exchange and the dynamic structure of proteins. *Mol. Cell. Biochem.* 48, 135-160.

[25] Englander, S.W., and Kallenbach, N.R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* 16, 521-655.

[26] Chance, M. (2008) *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics*. Hoboken, New Jersey: John Wiley & Sons, Inc.

[27] 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.

[28] 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.

- [29] 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
- [30] 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.
- [31] 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.
- [32] Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5, 725-738.
- [33] Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12.

Appendix A

A.1. Development of a bifunctional cross-linker to cross link E1p with lipoylated E2p-derived proteins

Crystal structure determination is always the most direct way to see how two proteins interact with each other. However, there is still no crystal structure of any E1p-LD complexes. We hypothesized that covalent linking of E1p with lipoyl domains may assist the crystallization, to finally define the interaction between the E1p and E2p components.

We've already shown that trivalent arsenicals can form high-affinity ring structures with the dihydrolipoyl group. E1p can be nearly fully phosphorylated by PDK2 in situ. Using ATP- γ -S instead of ATP, the thiophosphoryl group could be introduced onto E1p, containing a reactive thiol nucleophile for further derivatization.

Bromoacetylaminophenyl arsenoxide (BRAO) [1] is a bifunctional linker that has a trivalent arsenoxide at one end and an acetylbromide that could selectively react with the thiol group on the other end. A mixture of E1p and LD would be incubated, followed by adding BRAO to initiate alkylation of thiol of the thiophosphoryl group on E1p. Then

add TCEP to reduce the lipoamide of LDs to dihydrolipoamide, which can react with the trivalent arsenoxide group, resulting in crosslinked E1p-LD.

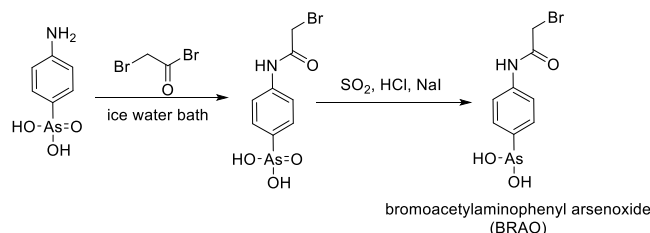


Figure A.1.1: Synthetic procedure of BRAO

The route to synthesize BRAO is shown above. p-Arsanilic acid (10 g) was added in discrete portions to a well stirred solution of sodium carbonate (30 g) in water (150 mL). Bromoacetyl bromide (10 mL) was diluted with CH_2Cl_2 (50 mL). Then the acylating solution was added into the alkaline arsenate solution at a rate of 2 drops per second. After stirring at ambient temperature for 2 h, the organic layer was separated and discarded. The aqueous solution was carefully acidified with 98% (w/w) sulfuric acid to pH 4 forming a white precipitate that was collected by suction filtration. The white precipitate (10 g) was suspended in 75 mL of methanol with constant stirring in a two-necked 250 mL round bottomed flask and 75 mL of HBr (48%) was added. Addition of a small amount of NaI turned the solution brown immediately. SO_2 was bubbled into the mixture for 2 to 3 h, yielding a white precipitate which was the final product BRAO.

LC-MS was used to check the formation of cross-linked product. 20 ng of cross-linked protein was denatured by adding 100 μL of 8 M urea, incubating at 37 $^\circ\text{C}$ for 30 min. Then TCEP was added to a final conc. of 6 mM, and the mixture was incubated at 37 $^\circ\text{C}$ for another 30 min. Next idoacetamide was added to a final conc. of 15 mM, the mixture

was incubated in the dark for 30 min, then the same amount of DTT solution (final conc. of 15 mM) was added to quench the reaction. Finally, 55 mM ammonium bicarbonate was added to dilute urea to 1 M, then add 70 ng (35 times of protein) trypsin to digest the sample at 37 °C overnight. The sample was desalted by a Sep-Pak tC18 cartridge. A sample was dissolved in 30 µL of 5% acetonitrile solution and LC-MS was run on a Phenomenex Luna C18 column.

The LC-MS results are shown below. Only a tiny amount of cross-linked peptide was identified.

To make the method usable for biophysical studies, the method would have to be modified/optimized.

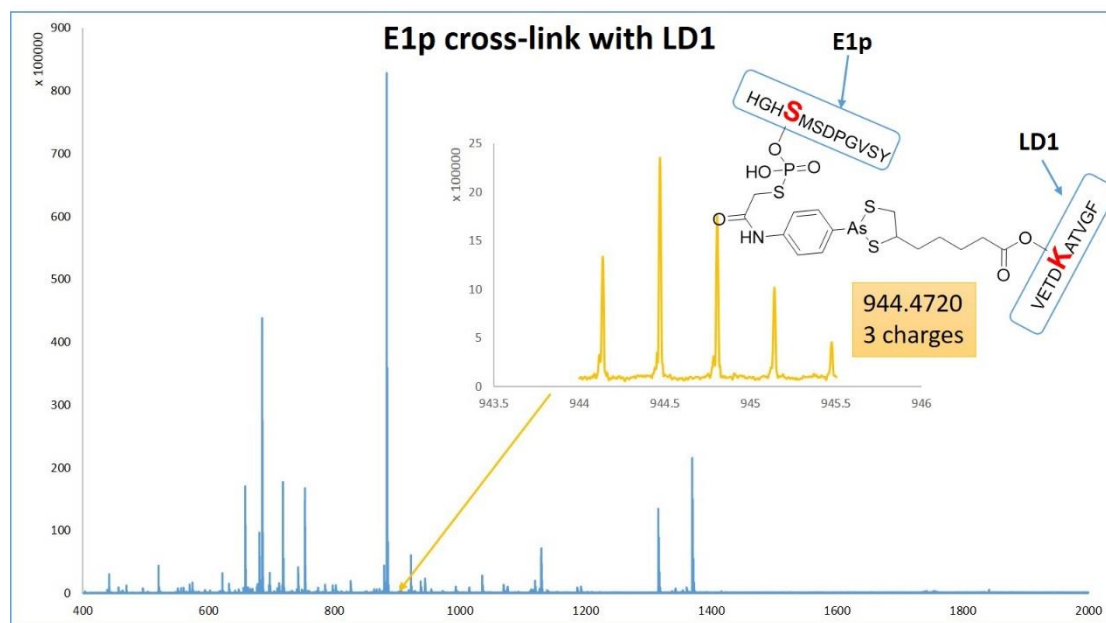


Figure A.1.2: LC-MS of cross-linked E1p-LD1 after pepsin digestion.

The reason why the efficiency of this reaction is so low is still unknown. Maybe the cross-linker would change the conformation of protein, bury the other functional end into pockets and make the other reactive end unavailable.

A.2. ^{19}F NMR studies of loop dynamics in E1p

The fluorine nucleus is a spin $-1/2$ species, which exists in 100% natural abundance and possesses a magnetogyric ratio that is 83% that of proton [10]. The large magnetogyric ratio translates into both high sensitivity in ^{19}F NMR spectroscopy, and strong dipolar couplings. Also, ^{19}F NMR possesses the advantage of inherent sensitivity of the chemical shift to the local environment. The fluorine chemical shift is primarily influenced by a large paramagnetic term, originating from an unpaired valence electron making it highly sensitive to local van der Waals interactions and electrostatic fields. [2] That's why ^{19}F NMR studies of proteins is so powerful in providing unique insight into biologically relevant phenomena such as conformational fluctuations, folding and unfolding, binding and catalysis. [3-4]

Another advantage of ^{19}F NMR is that there is no fluorine found in natural proteins, as a result, a protein containing a ^{19}F -label will exhibit NMR peaks due only to the label. By using a site specific labeling method, we can incorporate a fluorine probe into the active center, and study protein conformational changes.

The phenomenon that phosphorylation on specific serine residues of the $^h\text{E1p}$ α subunit can inactivate PDC was discovered in the late 1960s. [8] The three phosphorylation sites are Ser264: site 1; Ser271: site 2; and Ser203: site 3. Based on crystal structures of human E1p, site 1 and site 2 are located in the same loop A (Ph-loop A, residues 259 to

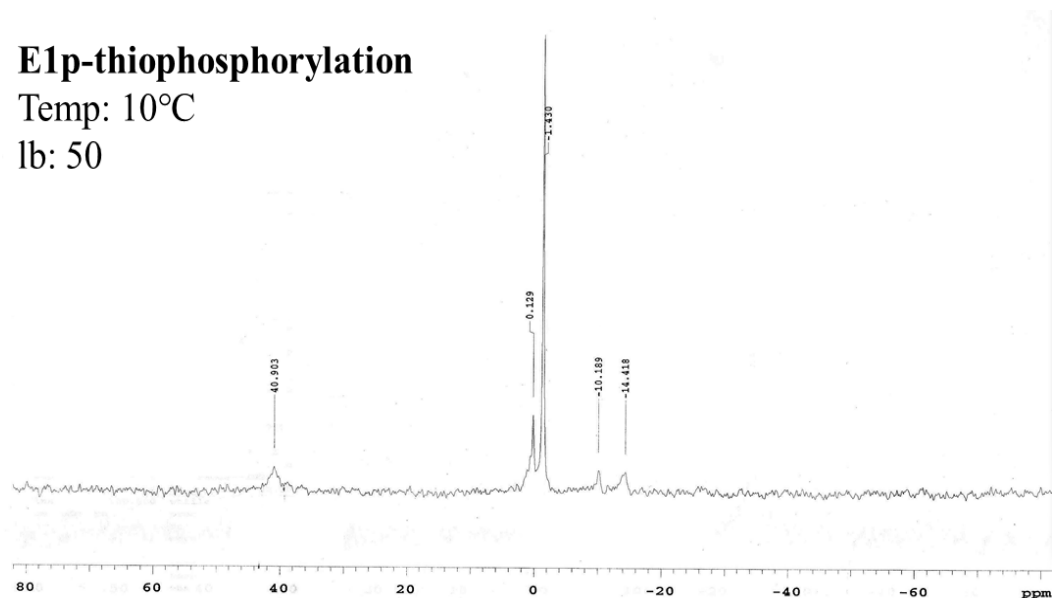
282) and site 3 is located in a relatively short loop called loop B (Ph-loop B, residues 198 to 205). [5-6]

It has been suggested that loop A is responsible for forming the E1p active-site channel and helping anchor the cofactor ThDP to the active site, while loop B provides coordination to a Mg^{2+} ion chelated by the diphosphate group of ThDP. In other words, the binding of ThDP co-factor needs ordered structure of both loops. [9]

A E1p-thiophosphorylation

Temp: 10°C

lb: 50

**B E1p-thio-MTSL**

Temp: 10°C

lb: 50

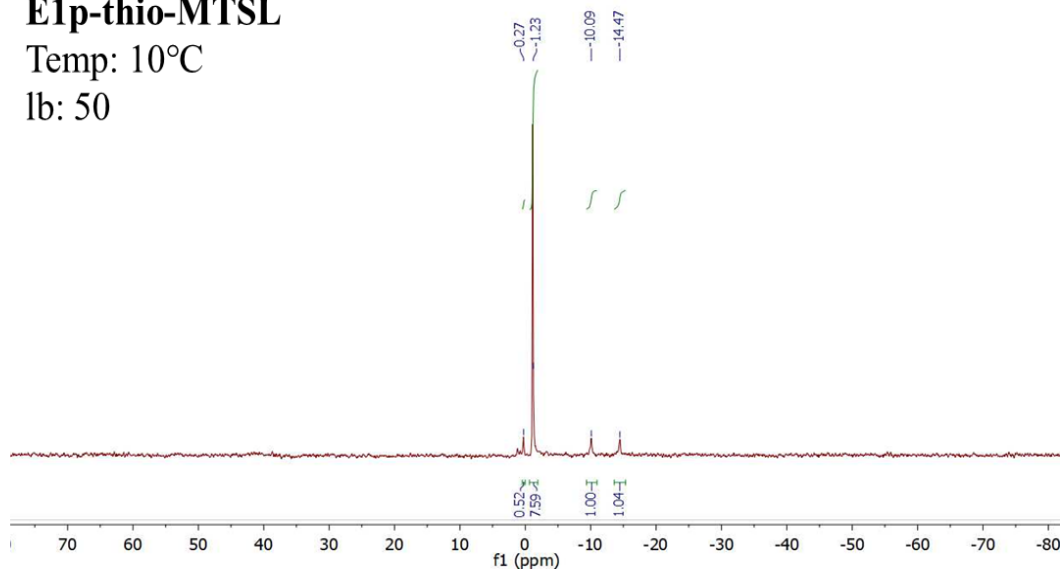


Figure A.2.1: ^{31}P NMR results of ThDP bounded phos-S-E1p. All experiments were conducted on a Bruker 600 NMR at 10 °C overnight. The line broadening of all spectra are 50. (A). NMR spectra of thiophosphorylated wild-type E1p. The broad peak located on 40.893 ppm is from P directed linked to the thio group. The peaks of -10.199 and -14.428 ppm originate from ThDP's α -P and β -P, respectively. (B). NMR spectra of MTSL labeled thiophosphorylated wild-type E1p. The only difference between the two

spectra is that the broad peak located on 40 ppm disappeared, since the unpaired electron from MTSL radical will negatively impact protein peak via paramagnetic relaxation and therefore decrease resolution.

According to the crystal structures of S1-E1p (containing Mg-ThDP) and phospho-S1-E1p (containing Mn-ThDP) determined by Chuang's group in 2008, incorporation of phosphoryl group will disrupt the conformations of both Ph-loop A and Ph-loop B. Phosphorylation on Ser264 can pose a steric clash, which eliminates the essential hydrogen-bonding network that is very important for stabilizing the loop conformation. [9] Consequently, the order-to-disorder transition of the two phosphorylation loops in E1p led to disruption of substrate channeling in PDC, which would inactivate the E1p.

It was already shown in the Jordan group by Sachin Kale that in *Escherichia coli* pyruvate dehydrogenase complex, the loop dynamics of ^{13}C -E1p active center promotes the rate of covalent addition of substrate to the enzyme bound ThDP. [7] It's reasonable to hypothesize whether the human E1p loop serves the same role.

As there are 12 cysteines in the ^1H -E1p α subunit, using cysteine thio-group directed labeling would lead to ambiguous results. Instead, we have at our disposal three different E1p variants: E1p-MS 2,3, in which only site 1 is available for phosphorylation; E1p-MS 1,3, which only site 2 is available; and E1p-MS 1,2, in which only site 3 is available. We would therefore use ATP- γ -S, an analog of ATP, to introduce a thiophosphoryl group into the Ph-loop of a single Ser loop variant, then incorporate an ^{19}F NMR probe via a trifluoroacetyl group onto the more reactive thiophosphoryl group to study the dynamics of the loops.

Materials and Methods

- (i) Three ^hE1p variants, E1p-MS2,3 (S264, [S271A/S203A]), E1p-MS1,3 (S271, [S264A/S203A]), E1p-MS1,2 (S203, [S264A/S271A]). The leading serine is the one available for phosphorylation, and those in the bracket are unavailable.
- (ii) The E1p variant (50 μ M) was incubated with 5 μ M of PDK1 at room temperature for 10 min. Then ATP- γ -S was added to a final concentration of 500 μ M to initiate the thiophosphorylation reaction and the reaction was carried out at 4 °C for overnight. Next, excess reagents were removed by dialysis against the sample buffer with a centrifugal filter unit (Vivaspin 500, 30K MWCO). 20 ng sample was saved for trypsin digestion, then LC-MS was run to analyze the data.
- (iii) Next, 50 μ M of thiophosphorylated E1p variant was treated with 150 μ M of trifluoroacetyl bromide at room temperature for 2 h. Excess reagent was removed by buffer exchange with a centrifugal filter unit (Vivaspin 500, 30K MWCO). 20 ng sample was saved for trypsin digestion, then LC-MS was run to analyze the data.

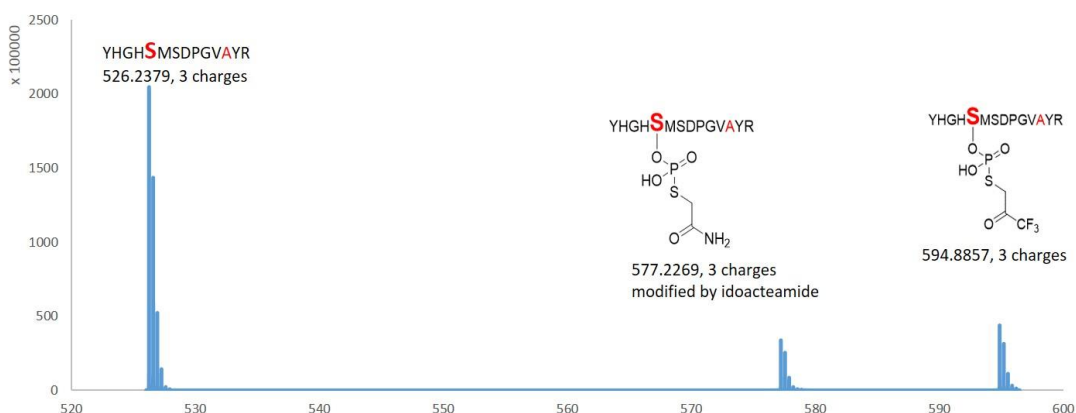


Figure A.2.2: LC-MS result after trypsin digestion of modified E1p-MS 2, 3 sample.

20 ng of protein were denatured by adding 100 μ L of 8 M urea, incubating at 37 °C for 30

min. Then TCEP was added to a final concentration of 6 mM, and the mixture was incubated for another 30 min at 37 °C. Next, iodoacetamide was added to a final conc. of 15 mM, the mixture was incubated in the dark for 30 min, then the same amount of DTT solution (final conc. of 15 mM) was added to quench the reaction. Finally, 55 mM ammonium bicarbonate was added to dilute the urea to 1 M, then trypsin was added (70 ng, 35 times the concentration of protein) to digest the sample at 37 °C overnight. The sample was desalted by a Sep-Pak tC18 cartridge. The sample was dissolved in 30 µL of 5% acetonitrile solution and the LC-MS was run on a Phenomenex Luna 5u C18 column.

Table A.2.1: Peptides after trypsin digestion for the detection of E1p modification

Sequence	Modification	Chemical Composition	Most Abundant Isotopic Mass			Error
			[M + nH] ⁿ⁺ (Da)			(ppm)
			Charges	Experimental	Theoretical	
²⁶⁰ YHGHSMSPGVAYR ²⁷³	None	C ₆₈ H ₉₇ N ₂₁ O ₂₁ S	2+	788.8543	788.8517	3.3
	Thiophosphorylation (iodoacetamide)	C ₆₈ H ₉₇ N ₂₁ O ₂₁ SPO ₂ SCH ₂ CONH ₂	2+	865.3373	865.3342	3.6
	Bromo-trifluoroacetone	C ₆₈ H ₉₇ N ₂₁ O ₂₁ SPO ₂ SCH ₂ COCF ₃	2+	891.8257	891.8224	3.7

So far, we used LC-MS to prove that it's possible to attach the trifluoroacetyl group to the E1p variant with, an efficiency of labeling of approximately 30%. We are going to work on improving the efficiency of labeling and prepare a larger scale sample for ¹⁹F NMR.

A.3. EPR studies of loop dynamics in E1p

Site-directed spin labelling (SDSL) combined with electron paramagnetic resonance (EPR) spectroscopy is a very powerful technique used widely for studying the structural properties and dynamical processes of biological systems. [11-12]

Spin labelling provides a relatively non-intrusive technique to study the structure, dynamics, and conformational changes of biomacromolecules.

By far the most popular nitroxide spin label is MTSL (Figure 10). MTSL is highly selective for thiol groups, and is commonly attached to cysteine residues in proteins. The MTSL side chain (commonly known as R1) also has a minimum impact on the secondary and tertiary structure of proteins.

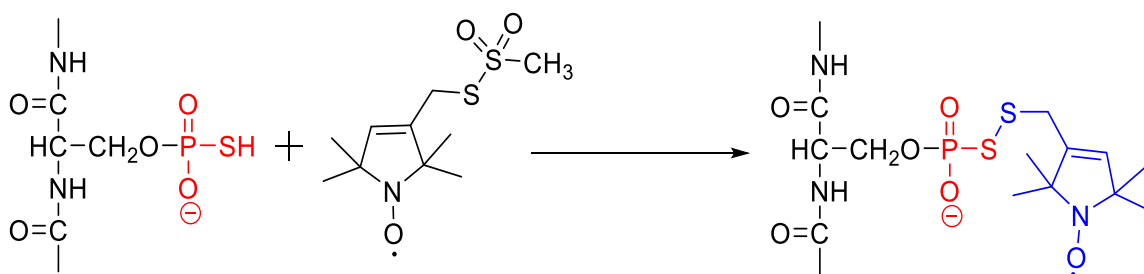


Figure A.3.1: Site-directed labelling of MTSL onto E1p- α subunit.

We also use LC-MS to monitor the modification.

Peptide of E1p-MTSL

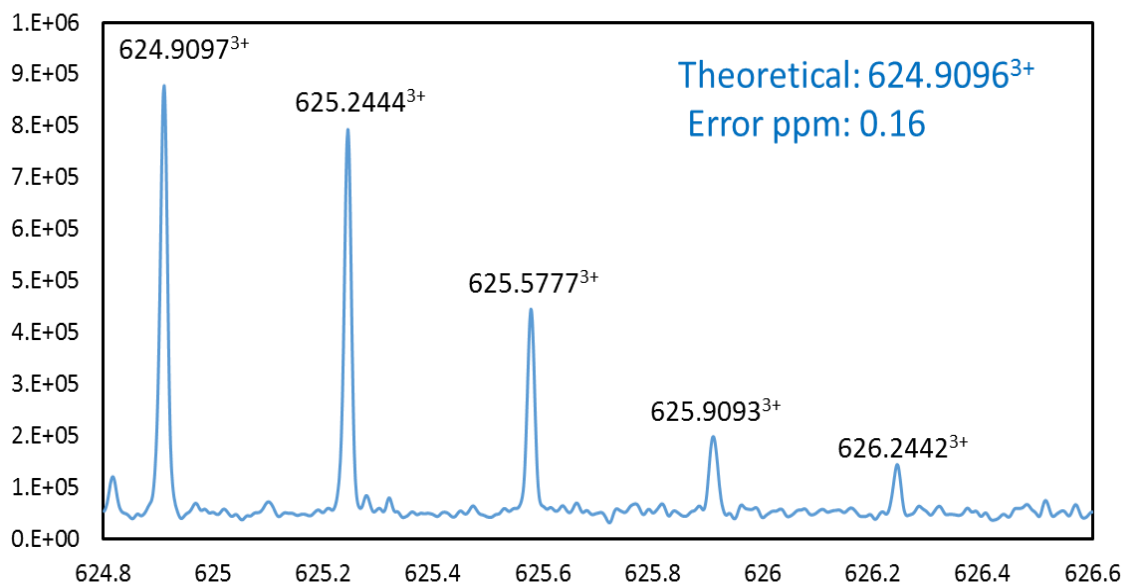


Figure A.3.2: LC-MS peak of “YHGHMSDPGVSYR-MTSL” (C₇₇H₁₁₂N₂₂O₂₅S₃P).

The theoretical mass of this peptide is 624.9096³⁺, the mass detected by FT-MS is 624.9097³⁺.

The EPR spectrum was measured with 7.5 mg/ml (96 μM per subunit) E1p-MTSL sample dissolved in 50 mM Tris (pH=7.5) containing 1 mM MgCl₂, 0.5 mM ThDP.

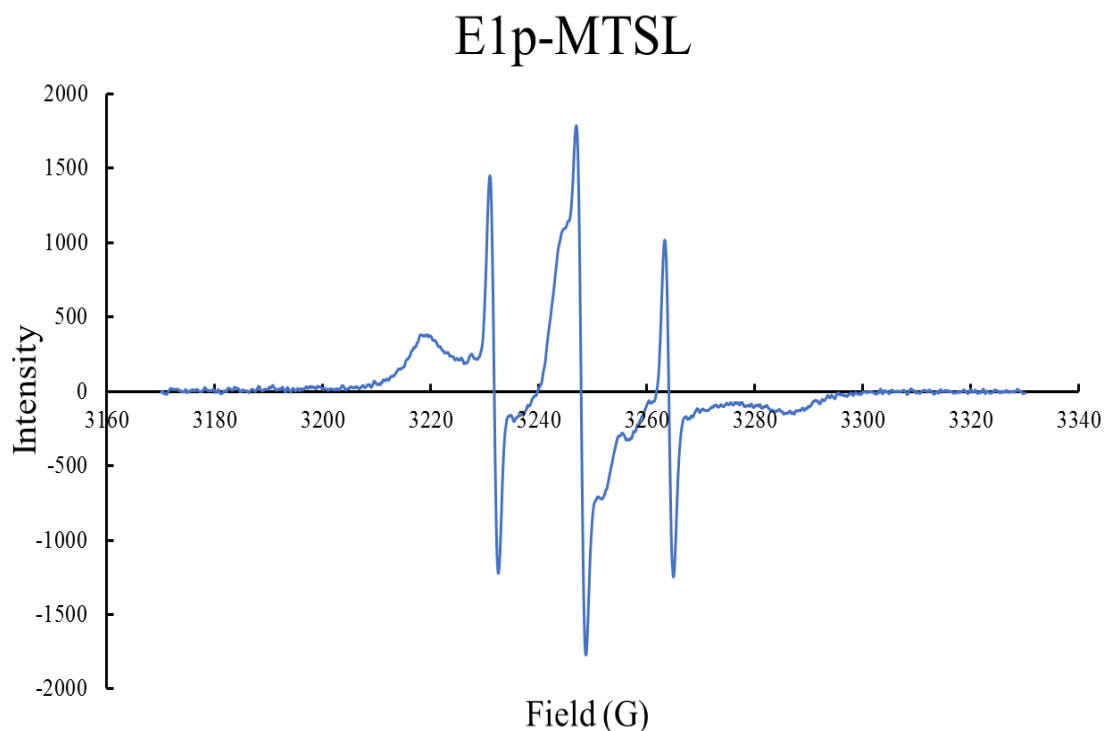


Figure A.3.3: EPR spectrum of E1p-MTSL. There are two forms of the component measured at room temperature. The broad peaks represent the slow component, which is dominant; the sharp peaks corresponds to the fast component, which is less than 5%.

A.4. HDX-MS studies of loop dynamics in E1p

Materials and methods

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

Sample preparation for HDX.

Prior to H/D exchange, the E1p, L1 or L2S protein was exchanged into 50 mM KH_2PO_4 (pH 7.5) with 100 mM NaCl, 0.2 mM ThDP, and 1 mM MgCl_2 , and then the protein concentration was adjusted to 160 μM and 80 μM . The concentrated E1p (80 μM) was incubated with or without pyruvate, methylacetylphosphonate (MAP) at a final concentration of 1 mM for 30 min at 25 °C prior to initiation of the HDX experiments. The concentrated E1p (160 μM) was incubated with equal amount of L1 or L2S (both 160 μM) for 30 min at 25 °C prior to initiation of the HDX experiments. E1p-phos is prepared separately. 50 μM of E1p was incubated with 5 μM of PDK2 at room temperature for 30 min. Then ATP was added to a final concentration of 500 μM to initiate the phosphorylation reaction and the reaction was carried out at 4 °C for overnight. Next day, excess reagents were removed by dialysis against the sample buffer with a centrifugal filter unit (Vivaspin 500, 30K MWCO). 20 ng sample was saved for trypsin digestion, then LC-MS was run to monitor the phosphorylation reaction.

The HDX experiments were initiated by mixing 15 μl of the protein samples with 285 μl of D_2O buffer to yield a final concentration of 95 % D_2O at pH 7.5. D_2O buffer was prepared the same way as E1p exchange buffer except 99.9% D_2O was used to dissolve the buffer components. The samples were incubated at 25 °C for 30 s, 1, 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.

Results and Discussions

The phosphorylation reaction is checked by LC-MS.

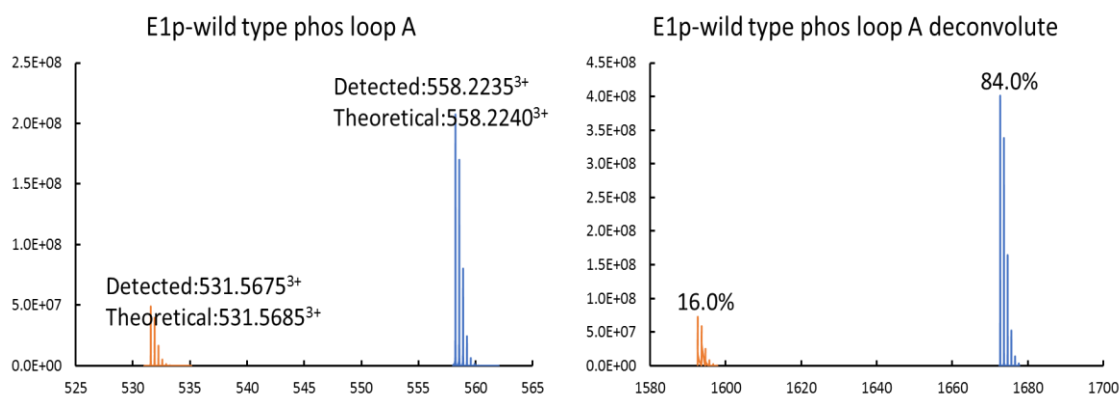


Figure A.4.1: Phosphorylated peptide detected by LC-MS. The peak of unmodified peptide is colored in orange, and the peak of phosphorylated peptide is colored in blue.

The ratio of unmodified peptides and phosphorylated peptides is calculated based on their intensities after deconvolution. 16% peptide remained unmodified according to this estimate.

Based on the intensity of the original peptide and phosphorylated peptides, we can estimate that 84% of the protein that was used in the HDX-MS experiments had been phosphorylated.

Among the 42 peptides we analyzed (24 peptides from E1p- α with 78% sequence coverage and 18 peptides from E1p- β with 89% sequence coverage), only two display EX1 kinetics. Residues 192-205 and 162-177 from E1p- α around the active center

displayed an EX1 signature in both ligand-free and ligand-associated states. The half-life of conformational change between folding and unfolding is dependent on ligands. ThDP-bound E1p without any ligand intrinsically adopts two conformations, which can only be seen within 1 min. MAP, pyruvate and L2S can freeze the folded (closed) conformation of E1p, while L1 and E1p-phos are prone to adopt the unfolded (open) conformation.

Both peptides displaying EX1 kinetics are involved in binding ThDP, peptide of 192-205 even contains the phosphorylation site 3 (Ser 203). It has been suggested that loop A is responsible for forming the E1p active-site channel and helping anchor the cofactor ThDP to the active site, while loop B, which contains site 3 provides coordination to a Mg^{2+} ion chelated by the diphosphate group of ThDP.

According to the crystal structures of wild type E1p (containing Mg-ThDP) and phospho-S1-E1p (containing Mn-ThDP) determined by Chuang's group in 2008, incorporation of phosphoryl group will disrupt the conformations of both Ph-loop A and Ph-loop B.

Phosphorylation on Ser264 can create a steric clash, which eliminates the essential hydrogen-bonding network that is very important for stabilizing the loop conformation. Consequently, the order-to-disorder transition of the two phosphorylation loops in E1p led to disruption of substrate channeling in PDC, which would inactivate the E1p.

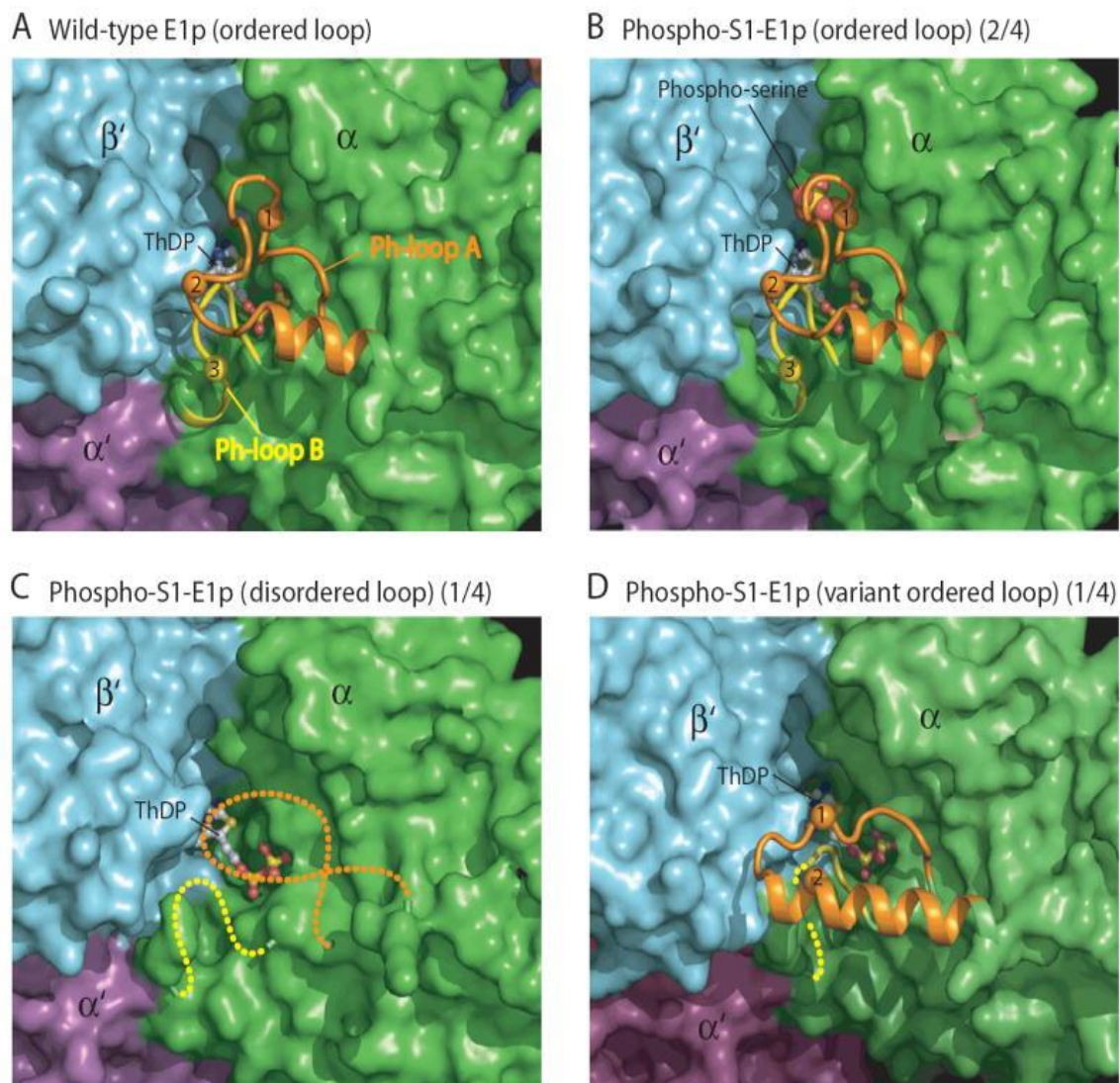


Figure A.4.2: Structures of the Phosphorylation loops in Wild-type and Phospho-S1-E1p. The structures of Ph-loops in E1p- α subunits are shown as ribbon models against other parts of the protein in surface representation. Ph-loop A (residues from 259- α to 282- α) is in orange, and Ph-loop B (from 198- α to 205- α) in yellow. The three phosphorylation sites are indicated by spheres and labeled according to the site numbers. The bound ThDP is shown as a ball-and-stick model. (A) Fully ordered Ph-loops in wild-

type E1p. (B) Wild-type-like ordered Ph-loops in phospho-S1-E1p with bound Mn-ThDP. Two of the four E1p- α subunits in the asymmetric unit (depicted as 2/4 in the figure caption) exhibit this conformation, which is maintained through interactions with a symmetry-related molecule. (C) Completely disordered Ph-loops in phospho-S1-E1p containing the bound Mn-ThDP. One of the four E1p- α subunits in the asymmetric unit (1/4) has this conformation. No symmetry-related molecule is present near the Ph-loops. (D) Variant ordered conformation of Ph-loop A and partially ordered Ph-loop B in the remaining E1p- α subunits (1/4). This conformation is also maintained by interactions with a symmetry-related molecule.

On the basis of the results, one could hypothesize that substrates prefer the folded (closed) state of E1p, while the phosphorylation promotes the E1p to transition from the folded state to unfolded state, which impedes the ability of E1p to bind substrate, as a result, blocking the entire reaction. Inner lipoyl domain (L2S) could help E1p adopt the conformation that is favored by substrate. MAP can form a stable intermediate that stays on the protein, which may partially explain why MAP can lock the folded (closed) conformation much better than any other ligands.

192-205: ICENRYGMGTSVE

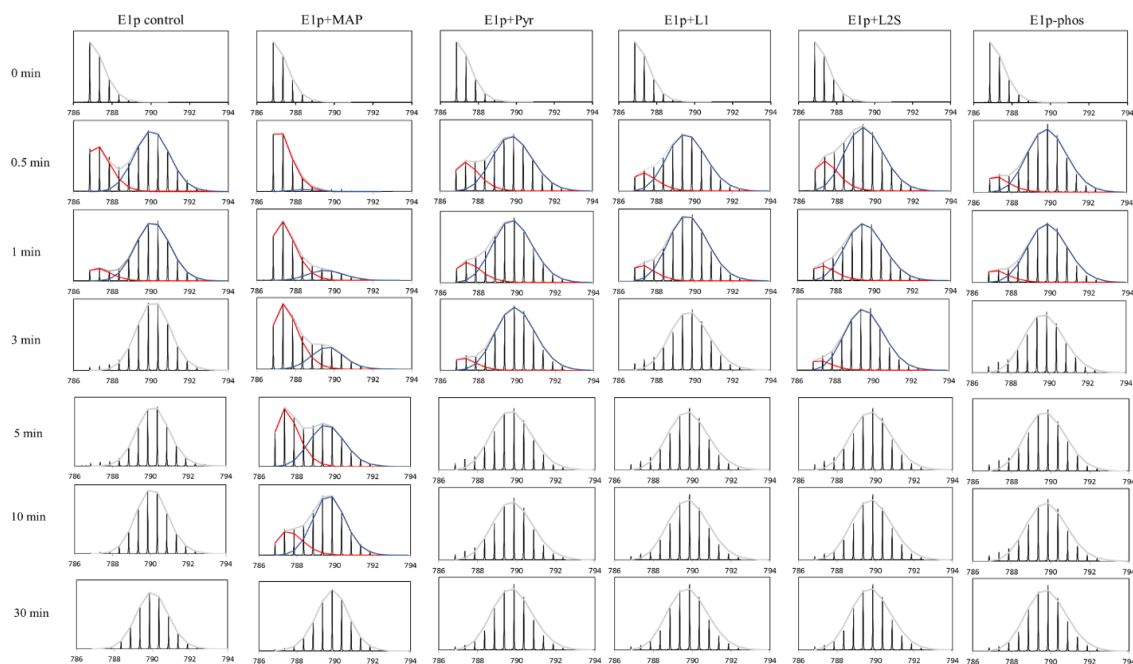


Figure A.4.3: Complete view of HDX-MS spectra of peptide 192-205 of E1p- α in its ThDP-bound state and in other ligand induced states.

162-177: LTLYGDGAANQGQIFE

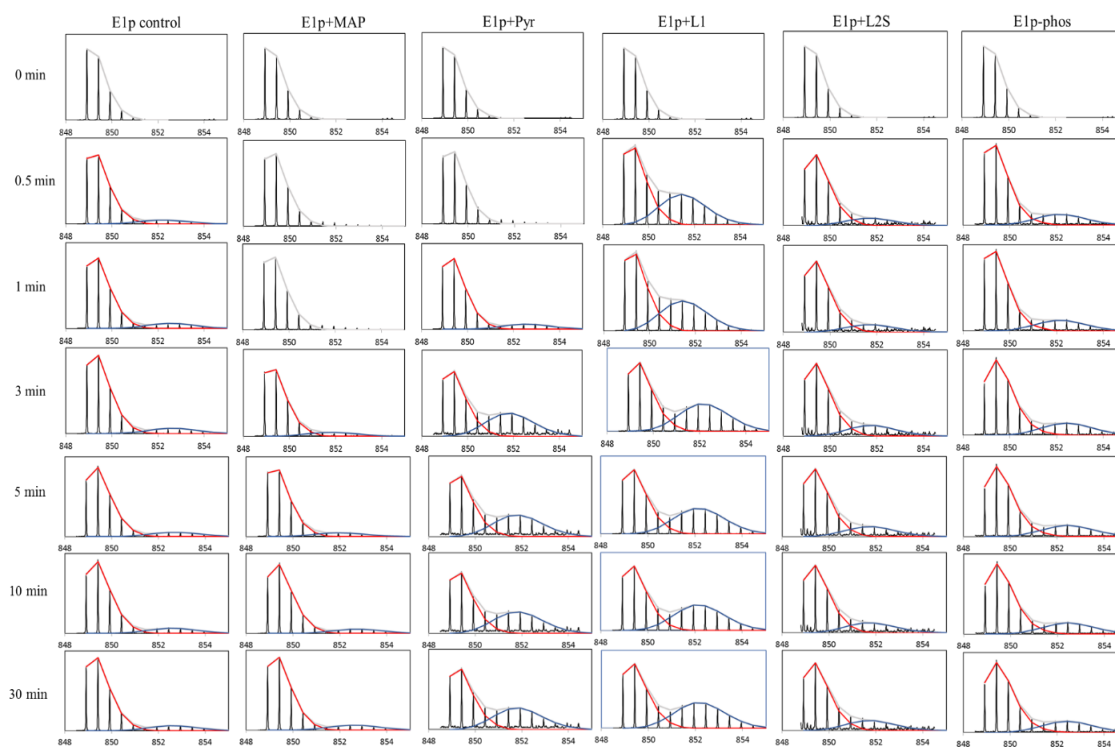


Figure A.4.4: Complete view of HDX-MS spectra of peptide 162-177 of E1p- α in its ThDP-bound state and in other ligand induced states.

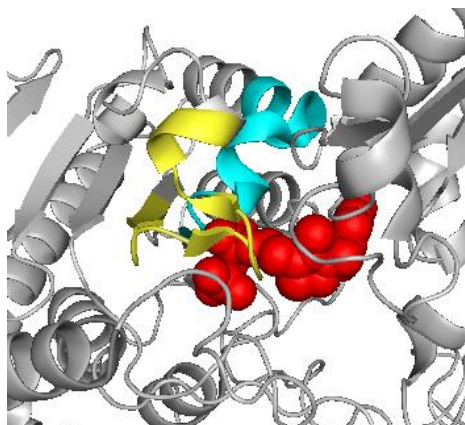


Figure A.4.5: Peptides from E1p- α that show bimodal distribution in HDX-MS.

ThDP is colored by red, peptide 192-205 from E1p- α is colored by yellow, peptide 162-177 from E1p- α is colored by cyan. Both of peptides displaying EX1 kinetics are involved in ThDP binding.

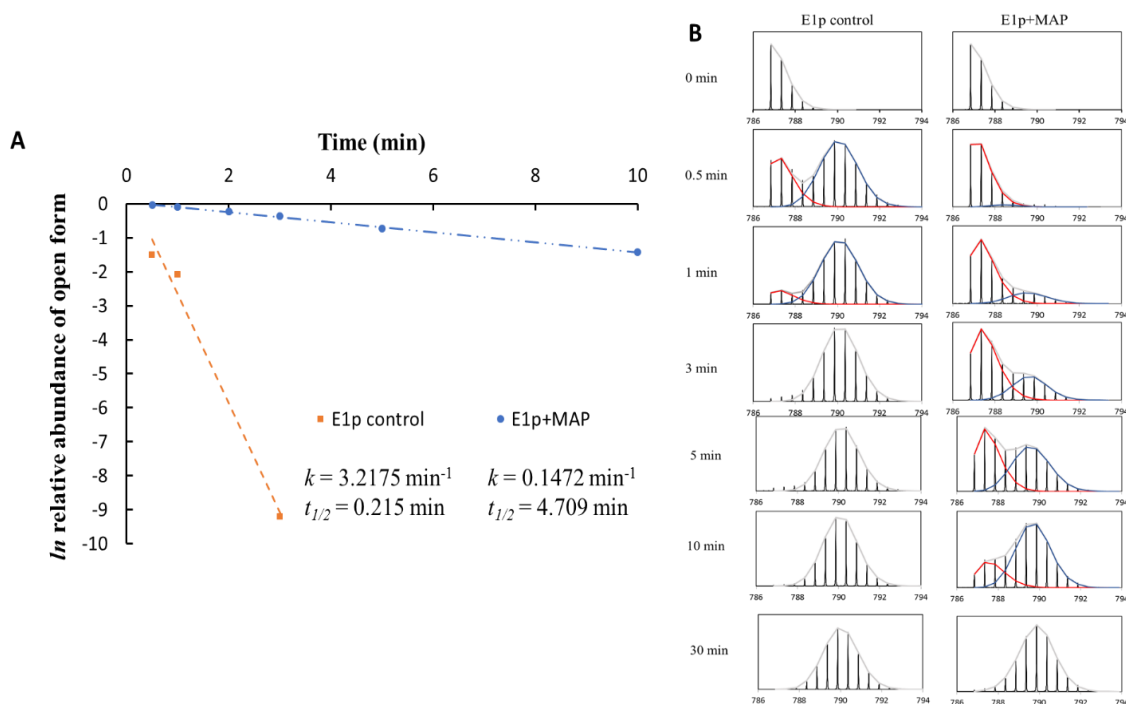


Figure A.4.6: A. Half-life calculation of the unfolded form of E1p- α 192-205. The slope of the line fit to the data yields the first order rate constant, and based on that number is calculated the half-life of unfolding. **B. HDX-MS spectra of peptide 192-205 that display EX1 kinetics in the free state as well as in the presence of MAP,** a clear bimodal isotopic distribution was observed.

References

- [1] Donoghue, N., Yan, P. T. W., Jiang, X.-M., Hogg, P. J. (2000) Presence of closely spaced protein thiols on the surface of mammalian cells. *Protein Science*, 9, 2436-2445
- [2] Gerig J T. Fluorine NMR of proteins[J]. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 1994, 26: 293-370.
- [3] Horng J C, Raleigh D P. Φ -Values beyond the ribosomally encoded amino acids: kinetic and thermodynamic consequences of incorporating trifluoromethyl amino acids in a globular protein[J]. *Journal of the American Chemical Society*, 2003, 125(31): 9286-9287.
- [4] Li H, Frieden C. Fluorine-19 NMR studies on the acid state of the intestinal fatty acid binding protein[J]. *Biochemistry*, 2006, 45(20): 6272-6278.

- [5] Ciszak, E. M., Korotchkina, L. G., Dominiak, P. M., Sidhu, S., and Patel, M. S. (2003) Structural basis for flip-flop action of thiamin pyrophosphate-dependent enzymes revealed by human pyruvate dehydrogenase. *J. Biol. Chem.*, 278, 21240-21246
- [6] Seifert, F., Ciszak, E., Korotchkina, L., Golbik, R., Spinka, M., Dominiak, P., Sidhu, S., Brauer, J. Patel. M. S., and Tittmann, K. (2007) Phosphorylation of serine 264 impedes active site accessibility in the E1 component of the human pyruvate dehydrogenase multienzyme complex. *Biochemistry*, 46, 6277-6287
- [7] Kale S, Ulas G, Song J, Brudvig G W, Furey W, Jordan F. Efficient coupling of catalysis and dynamics in the E1 component of Escherichia coli pyruvate dehydrogenase multienzyme complex[J]. *Proceedings of the National Academy of Sciences*, 2008, 105(4): 1158-1163.
- [8] Linn T C, Pettit F H, Reed L J. α -Keto acid dehydrogenase complexes, X. Regulation of the activity of the pyruvate dehydrogenase complex from beef kidney mitochondria by phosphorylation and dephosphorylation[J]. *Proceedings of the National Academy of Sciences*, 1969, 62(1): 234-241.
- [9] Kato, M., Wynn, R. M., Chuang, J. L., Tso, S.-C., Machius, M., Li, J., and Chuang, D. T. (2008) Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops. *Structure*, **16**, 1849-1859

[10] Kitevski-LeBlanc J L, Prosser R S. Current applications of ^{19}F NMR to studies of protein structure and dynamics[J]. Progress in nuclear magnetic resonance spectroscopy, 2012, 62: 1-33.

[11] Hoff, A.J., Advanced EPR: Applications in Biology and Biochemistry. 1989.

[12] Jeschke, G., Conformational dynamics and distribution of nitroxide spin labels. Progress in

Nuclear Magnetic Resonance Spectroscopy, 2013. 72: p. 42-60.

[13] Woodward, C., Simon, I., and Tuchsén, E. (1982) Hydrogen exchange and the dynamic structure of proteins. Mol. Cell. Biochem. 48, 135-160.

[14] Englander, S.W., and Kallenbach, N.R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. Q. Rev. Biophys. 16, 521-655.

[15] Chance, M. (2008) Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics. Hoboken, New Jersey: John Wiley & Sons, Inc.

Appendix B

Proteins expression and purification

- (i) **Expression and purification of E1p and E1p loop variants.** Cells containing the pET-28b-E1p plasmid will be grown at 37 °C in LB medium supplemented with 35 µg/mL of kanamycin and 100 µg/mL of thiamin chloride. E1p expression will be induced by 0.5 mM IPTG at 25 °C for 16 h. The E1p will be purified using Ni-NTA agarose column equilibrated with 50 mM KH_2PO_4 , 300 mM NaCl and 30 mM imidazole (pH= 7.5). E1p will be eluted from the column with 50 mM KH_2PO_4 ,

300 mM NaCl and 200 mM imidazole (pH= 7.5). Protein will be dialyzed against 50 mM KH_2PO_4 , 300 mM NaCl, 1 mM MgCl_2 , 0.5 mM ThDP, 0.5 mM benzamidine hydrochloride (pH= 7.5) buffer.

- (ii) **Expression and purification of E2p-derived proteins.** The following E2p-derived proteins have been expressed and purified: the outer lipoyl domain LD1, the L2S didomain comprising inner lipoyl domain and subunit-binding domain, L1L2S tridomain, which comprises L1, L2, both hinge regions, and the subunit-binding domain, L1L2Sm1, in which the Lys132 has been converted to alanine, designated as K132A; L1L2Sm2, in which the Lys259 has been converted to alanine, designated as K259A. E2p-derived proteins will be expressed from pET-28b vector. Cells containing the plasmid will be grown at 37 °C in LB medium supplemented with 35 µg/mL of kanamycin and 0.3 mM lipoic acid to ensure full lipoylation of the lipoyl domains. The purification steps are the same as reported for E1p.
- (iii) **Expression and purification of human E2·E3BP protein.** Cells containing the plasmid will be grown at 37 °C in LB medium supplemented with 35 µg/mL of Chloramphenicol and 0.3 mM lipoic acid to ensure full lipoylation of the lipoyl domains. The cells were dissolved in 50-60 ml of sonication buffer [50 mM KH_2PO_4 (pH 7.5), 0.3 M NaCl, 5 mM MgCl_2 , 0.5% (v/v) Triton X100] by gently stirring with a glass rod, then 1 protease inhibitor cocktail tablet was added. Lysozyme was added to 0.6 mg/mL and the solution was incubated on ice for 20 min. Next, 1000 units of each DNase and Nuclease were added and the solution was incubated on ice for another 20-60 min. The solution was next sonicated for 6

min with 10s pulse “on” and 30s pulse “off” program. The lysate was clarified by centrifugation twice for 30 min at $23700 \times g$ at 4°C .

PEG precipitation of the E2·E3BP protein. PEG 8000 (50% w/v) solution was added dropwise to the clarified lysate to 1% (v/v). The precipitated protein was removed by centrifugation at $18\,500 \times g$ for 20 min. Next, 5 % (v/v) of PEG 8000 solution was added dropwise to the clarified supernatant and the mixture was incubated for 15 min. The precipitate was collected by centrifugation at $18500 \times g$ for 20 min. The precipitate was dissolved in approximately 10 ml of buffer A, and clarified by centrifugation. The supernatant was loaded onto a Sephacryl S 300HR column equilibrated with buffer A. The cloudy fractions (4 mL) (caution: *the first two cloudy fractions were not collected because these might contain some aggregated proteins*) were collected, combined and mixed with polyethylenimine (M_w 60,000) to a final concentration of 0.05% and the protein was precipitated by centrifugation at $23700 \times g$ for 30 min. The *E2·E3BP* was pelleted by two-step ultracentrifugation: (i) The impurities and a small amount of *E2·E3BP* were pelleted at $100,000 \times g$ for 4 h, at 4°C . The resulting supernatant was combined and concentrated to approximately 8 ml with a 50 K concentration unit. (ii) The concentrated solution (8 ml) was centrifuged at $250\,000 \times g$ for 4 h at 4°C to pellet the *E2·E3BP*. The pellets were resuspended in a small volume of buffer containing 0.4 M potassium chloride and left overnight. The suspension was clarified by centrifugation and stored at -80°C . The sonication buffer was 50 mM potassium phosphate (pH 7.5), 0.5 mM EDTA, 150 mM KCl, 1 mM β -mercaptoethanol, 0.1% Pluronic-F68. Buffer A was 50 mM potassium phosphate

(pH 7.5), 0.5 mM EDTA, 200 mM potassium chloride, 1 mM β -mercaptoethanol, 0.1% Pluronic-F68.

- (iv) **Expression and purification of PDKs.** Cells are grown on LB medium supplemented with 35 μ g/mL of kanamycin. Protein expression will be induced by 0.5 mM IPTG at 25 °C for 16 h. Dissolve cells in 60 mL 50 mM KH_2PO_4 (pH=7.5) containing 300 mM NaCl, 10 mM imidazole, 1 mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100. Add lysozyme to 0.6 mg/mL and incubate on ice for 30 min. Add 1000 units of DNase and nuclease, incubate on ice for an additional hour. Then, do the sonication. PDK1,2 were purified by using Ni-NTA agarose column equilibrated with 50 mM KH_2PO_4 , 300 mM NaCl, 30 mM imidazole, 1mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100 (pH= 7.5) buffer. PDK3,4 were equilibrated with 50 mM KH_2PO_4 , 300 mM NaCl, 20 mM imidazole, 1 mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100 (pH= 7.5) buffer. PDKs was eluted from the column with 50 mM KH_2PO_4 , 300 mM NaCl and 200 mM imidazole, 1 mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100 (pH= 7.5). Proteins were dialyzed against 50 mM KH_2PO_4 , 300 mM NaCl, 1 mM MgCl_2 , 1 mM DTT, 1 mM benzamidine hydrochloride (pH= 7.5) buffer.
- (v) **Expression and purification of human E1o.** The same method was used to purify E1o, E3, and E2o¹⁴⁴⁻³⁸⁶ (C-terminal didomain). Cells were dissolved in 50-60 mL of sonication buffer while stirring with a glass rod, then 1 protease inhibitor cocktail tablet was added. Lysozyme was added to 1 mg/ml and the solution was incubated on ice for 20 min. Next 1000 units of DNase and Nuclease each were added and

the solution was further incubated on ice for another 20-60 min. The solution was next sonicated for 6 min, with the 10s pulse “on” and 30s pulse “off” program. The solution was clarified by centrifugation three times for 30 min at $23700 \times g$ at 4 °C. The clarified His₆-tagged human E1o (50 mL) dissolved in 50 mM potassium phosphate (pH 7.4), with 500 mM NaCl, 30 mM imidazole, 2 mM MgCl₂, 0.5 mM ThDP (binding buffer) was filtered through a 0.45 µm filter then applied to a High Performance Ni Sepharose column with a 6 mL bed volume (for a 6 mL volume Ni column, the pressure is approximately 0.2 mbar at 1 ml/min flow rate, and 2.6 mbar at 2 ml/min flow rate). The column was eluted with a solution of 50 mM sodium phosphate (pH 7.4), 500 mM NaCl, 500 mM imidazole, 2 mM MgCl₂, 0.5 mM ThDP with the following gradient: a 2CV linear gradient of 30 to 77 mM imidazole, followed by a 20 CV linear gradient of 77 to 218 mM imidazole, followed finally by a push with 500 mM imidazole. Typically, 40 mg human E1o protein could be purified from 4 L LB culture media.

List of Publications

Publications at Rutgers University

1. Zhou J, Yang L, DeColli A, Meyers C F, Nemeria N S, Jordan F. Conformational dynamics of 1-deoxy-d-xylulose 5-phosphate synthase on ligand binding revealed by H/D exchange MS[J]. PNAS, 2017: 201619981.
2. Guevara E L, Yang L, Birkaya B, Zhou J, Nemeria N S, Patel M S, Jordan F. Global view of cognate kinase activation by the human pyruvate dehydrogenase complex[J]. Scientific Reports, 2017, 7. (Co-first author)
3. Nemeria N S, Gerfen G, Nareddy P R, Yang L, Zhang X, Szostak M, Jordan F. The mitochondrial 2-oxoadipate and 2-oxoglutarate dehydrogenase complexes share their E2 and E3 components for their function and both generate reactive oxygen species[J]. Free Radical Biology and Medicine, 2018, 115: 136-145.

Publications at Wuhan University

1. Li H, Yang L, Tang R, Hou Y, Yang Y, Wang H, Han H, Qin J, Li Q, Li Z. Organic dyes incorporating N-functionalized pyrrole as conjugated bridge for dye-sensitized solar cells: Convenient synthesis, additional withdrawing group on the π -bridge and the suppressed aggregation[J]. Dyes and Pigments, 2013, 99(3): 863-870.

CHAPTER 1. Introduction

The human pyruvate dehydrogenase complex (PDHc), which governs the overall carbohydrate oxidation, is a key enzyme at the entry of the citric acid cycle in every cell that contains mitochondria. It catalyzes the irreversible oxidative decarboxylation of pyruvate to produce acetyl-CoA, an important intermediate that participates in many biosynthetic pathways such as those of carbohydrate and lipid metabolism. PDHc is also responsible for regulating the rate of oxidative phosphorylation and subsequent ATP generation from carbohydrate derived carbon sources. During well fed state, PDHc is active and promotes glucose oxidation, while in the fasting state, PDHc activity would be turned off and the three-carbon compound —pyruvate could be preserved for sustaining gluconeogenesis. The activity of PDHc is not only important in glucose production but it is also related to the transition from glucose oxidation to fatty acid oxidation as proposed by the Randle cycle [1]. Randle and his colleagues conducted a series of experiments in cardiac and skeletal muscle, showing that increased fatty acid oxidation enhances the ratio of [acetyl-CoA]/[CoA] and [NADH]/[NAD⁺], both of which would inhibit PDHc activity by activating the pyruvate dehydrogenase kinases (PDKs).

1.1 Components and organization of the mammalian PDHc

The mammalian pyruvate dehydrogenase complex consists of three catalytic components: pyruvate dehydrogenase (E1p), dihydrolipoyl transacetylase (E2p) and dihydrolipoyl dehydrogenase (E3). It also contains a unique E3-binding protein (E3BP) and two

regulatory enzymes: pyruvate dehydrogenase kinases (PDK, four human isoforms) and pyruvate dehydrogenase phosphatase (PDP, two human isoforms). The complex assembles together using 48 copies of E2p and 12 copies of E3BP forming a E2p.E3BP subcomplex core, to which the E1p component binds to E2p and the E3 component binds to E3BP.

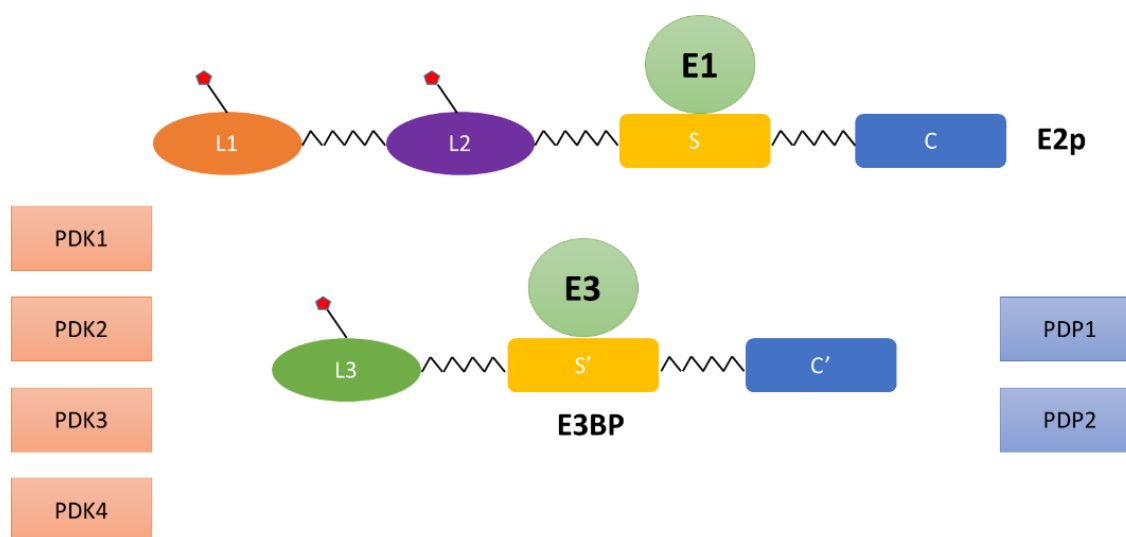


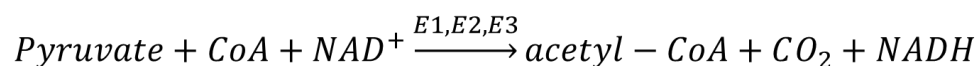
Figure 1.1: Structure of pyruvate dehydrogenase complex

E2p has a multidomain structure: an outer lipoyl domain (L1), an inner lipoyl domain (L2), a subunit-binding domain (S) and an inner core domain (core). The structure of E3BP is quite similar to that of E2p, while instead of two lipoyl domains, it only has one (Figure 2) [2], and a catalytically incompetent core domain.

The E1 component in PDHc catalyzes the decarboxylation of pyruvate with the auxiliary of thiamin diphosphate (ThDP), the hydroxyethyl group derived from pyruvate is then transferred to lipoamide derived from E2. The hydroxyethyl carbanion is oxidized to an acetyl group as the lipoamide disulfide is reduced. E2 also catalyzes the transfer of acetyl

groups to form acetyl-CoA. The E3 component then reoxidizes the dihydrolipoamide by its tightly-bound FAD coenzyme. Finally, FADH₂ is reoxidized to FAD by NAD⁺, yielding NADH [3].

The overall reaction is shown below.



equation 1.1

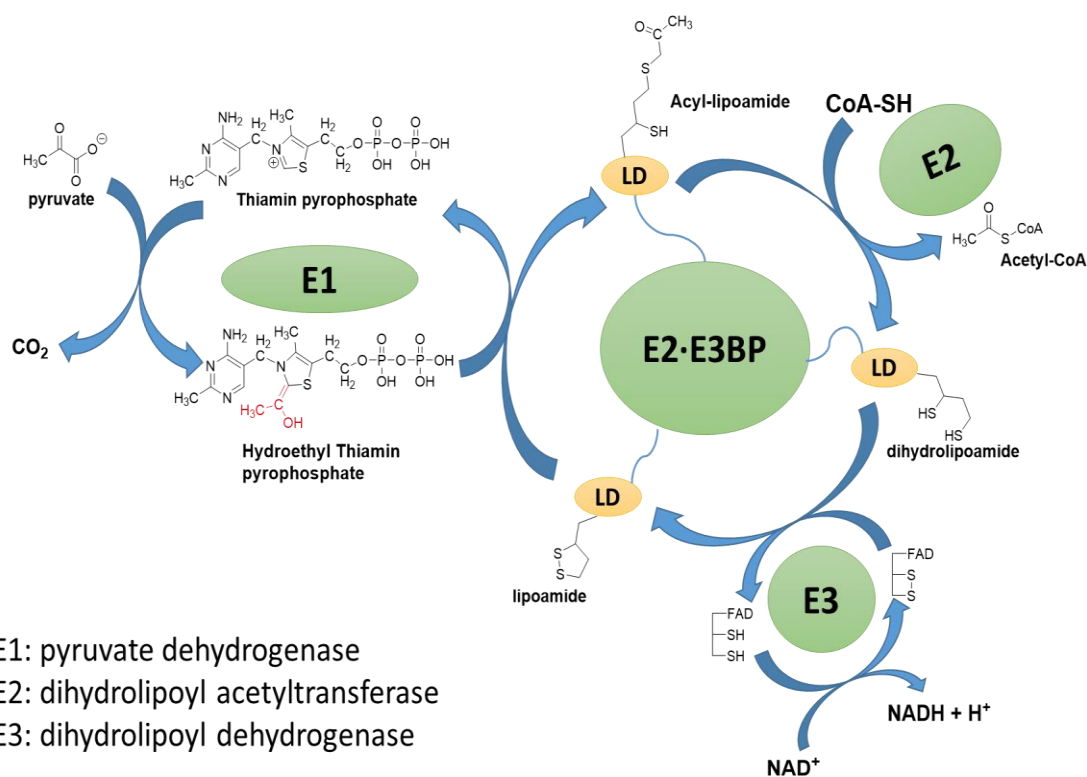


Figure 1.2: Catalytic cycle of the mammalian pyruvate dehydrogenase complex

1.2 Regulation and diseases relate to PDKs

PDHc regulation is accomplished by the combined reactions of two phosphatases (PDP1 and PDP2) and four kinases (PDK1-4). Through phosphorylation and dephosphorylation on the specific sites located on the E1p α subunit, PDKs and PDPs can shut down or recover, respectively, the activities of PDHc. The multiple isoforms of the PDKs and the PDPs are distinguished by differences in tissue distribution, specific activities toward the phosphorylation sites, kinetic properties, and sensitivity to regulatory molecules [4-5]. Regulation of PDHc is achieved by site-specific phosphorylation of the α subunits of the heterotetrameric ($\alpha_2\beta_2$) E1p component at three different sites — Ser264 (site 1), Ser271 (site 2) and Ser203 (site 3). The three sites were phosphorylated *in vivo* to a different extent, with maximum phosphorylation of site 1 and less phosphorylation of sites 2 and 3. The four PDK isozymes have different activity and phosphorylation rates at each site. At site 1 from fastest to slowest, $\text{PDK2} > \text{PDK4} \approx \text{PDK1} > \text{PDK3}$; for site 2 it is, $\text{PDK3} > \text{PDK4} > \text{PDK2} > \text{PDK1}$. Only PDK1 can phosphorylate site 3 [3]. Dephosphorylation of phospho-E1 by PDP's recovers full PDHc activity.

The existence of three potential phosphorylation sites together with four isoforms of PDKs provides the possibility of delicate control towards metabolism, although the detailed mechanism is still unclear.

The structure of PDKs can be divided into two distinct domains, the N- and C-terminal domains. The N-terminal domain consists of eight α -helices, of which four form a bundle-like structure creating the core. The sequence in the N-terminal domain of the four isoforms of PDKs is poorly conserved. In contrast, the C-terminal domain containing the phosphoryl transfer catalytic site is highly conserved among the four PDKs. They

share four conserved motifs, N-box (Glu-X-X-Lys-Asn-X-X-X-Ala); G1-box (Asp-X-Gly-X-Gly); G2-box (Gly-X-Gly-X-Gly); and G3-box (Gly-X-Gly-Thr) [4], which form a unique ATP-binding fold. This fold includes a common structural element known as the “ATP lid,” whose conformational change is coupled to both ATP hydrolysis and protein-protein interactions [5].

The tissue distribution of the four isoforms is also different. PDK1 is expressed predominantly in the heart [6] and pancreatic islets [9], while PDK2 has a relatively wide distribution and is present in liver, heart, skeletal muscle, etc. with lower amounts in spleen and lung [10]. PDK3 mainly exists in testis, kidney and brain [6], and PDK4 is expressed in heart, skeletal muscle and liver [11-12].

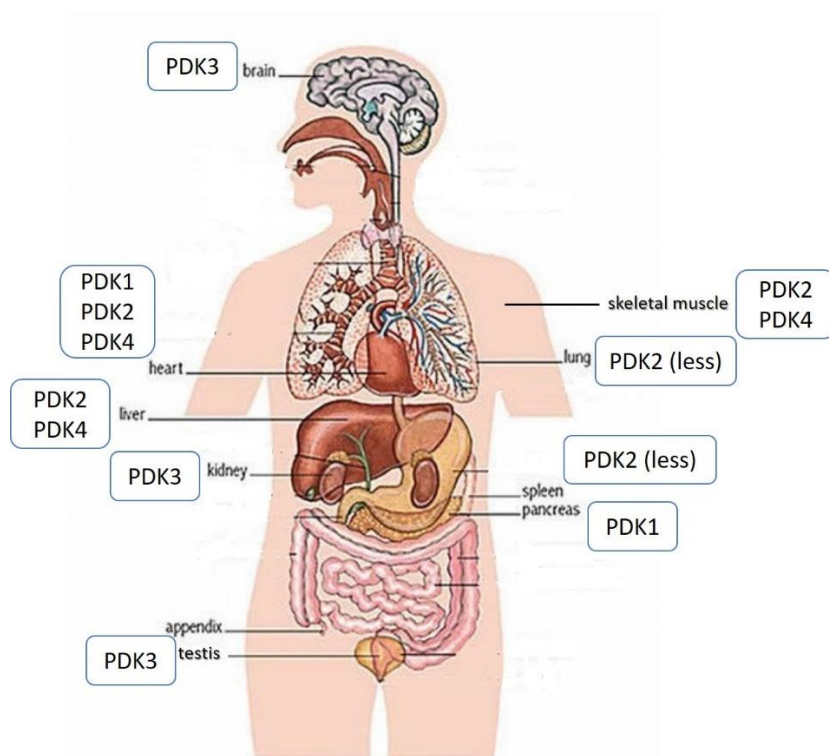


Figure 1.3: Tissue distribution of four isoforms of PDKs.

Transcription of PDKs is regulated by a lot of hormones including insulin, glucocorticoids, thyroid hormone and fatty acids. The abnormal increased expression of PDKs is related to a variety of diseases such as diabetes [13-15], heart disease [16-17], and fatty liver [18].

High level of PDK1 expression has been discovered in some tumor cells, which rely on aerobic glycolysis to generate energy. Instead of turning pyruvate to acetyl-CoA, the increasing transcription of PDK1 suppresses the activity of PDHc, thereby attenuating ROS (reactive oxygen species) production. The results of increasing glycolysis and ATP levels may suggest shunt pyruvate toward lactate production. [39]

PDK2 and PDK4 play a critical role in regulation of PDHc to maintain glucose homeostasis. Glucocorticoids [19] and free fatty acids [20-21] can stimulate PDK4 expression, while insulin can suppress it. Insulin has also been shown to repress PDK2 expression [20].

Failure of insulin to down-regulate the expression of PDK4 could result in type 2 diabetes. Overexpression of PDK4 in heart has shown to relate to heart disease for increasing fatty acid oxidation and decreasing glycolysis.

Table 1.1: PDKs and associated pathological conditions [22].

<i>PDK isoforms</i>	<i>Associated pathological conditions</i>
<i>PDK1</i>	Glioblastoma
	Brain aging

<i>PDK2</i>	Type 2 diabetes
	Brain aging
	Glioblastoma
	Ovarian cancer
<i>PDK3</i>	Charcot-Marie-Tooth neuropathy
<i>PDK4</i>	Type 2 diabetes
	Hemochromatosis
	Glucocorticoid excess, e.g., Cushing syndrome
	Cardiac hypertrophy
	Dilated cardiomyopathy
	Angiotensin II-induced heart failure
	Right ventricular hypertrophy and pulmonary hypertension
	Statin-induced myopathy
	Disuse osteoporosis
	Ovarian cancer
	Anoikis and tumor metastasis

The short-term control of PDKs is largely influenced by the organization and catalytic function of mammalian PDHc. The products during the five-step reaction of PDHc can exert some effects on the activity of PDKs, with high $[NADH]/[NAD^+]$ and $[acetyl-CoA]/[CoA]$ ratios enhancing the activity of PDKs [23-24], while adequate concentration of substrates pyruvate, NAD^+ and CoA could inhibit PDKs [24].

The phosphorylation of three target serine residues display different effects within the complex. Phosphorylation at site 1 is the fastest and could shut down the activity of PDHc. Phosphorylation of site 2 is more about promoting the inactivation and suppressing the reactivation via PDPs. The role of site 3 is mostly in preventing the reactivation [25]. Site 1 is positioned in the substrate channel leading to the E1p active center. One of the explanations of inactivation by phosphorylation is that introducing a negatively charged and bulky group would prevent the lipoyl domain of E2 from visiting the active site and picking up intermediates. This is supported by the results that phosphorylation of site 1 has nearly no impact on decarboxylation but severely attenuates the reductive acetylation reaction [26]. Site 2 and Site 3, which are not located near the active site, may affect PDHc activity by decreasing the binding ability towards ThDP [27] or disordering the conformation of the loop [28].

It has been widely believed that lipoyl domains play a pivotal role in modulation of PDK activities, though different PDKs may have different preferences towards the lipoyl domains located on E2 and E3BP. The structure of PDK3 bound to the L2 domain showed that binding to the inner lipoyl domain induced a conformational change of PDK3, ordering the C-terminal tails of the dimer to form a ‘crossover’, thereby enlarge the active site cleft to facilitate the release of ADP [29].

1.3 Inhibitors of PDKs

There are three types of inhibitors that have been mostly used and tested: AZD7545, dichloroacetate (DCA), and radicicol. DCA an analog of pyruvate, is one of the classic inhibitors for PDK isoforms. DCA inhibits the activity of PDKs by binding to the

allosteric site in the N-terminal domain. However, the specificity of DCA is low and it needs high doses for therapeutic effects [30-31].

Disrupting the interaction between PDKs and the inner lipoyl domain, thereby attenuating the recruitment of PDK1 to the complex, led to the development of AZD7545 [32]. This group of inhibitors also includes Pfz-3 and Nov3r [32-33].

Radicalol represents a class of ATP-competitive inhibitors, however, the low specificity limits its usage [34-35].

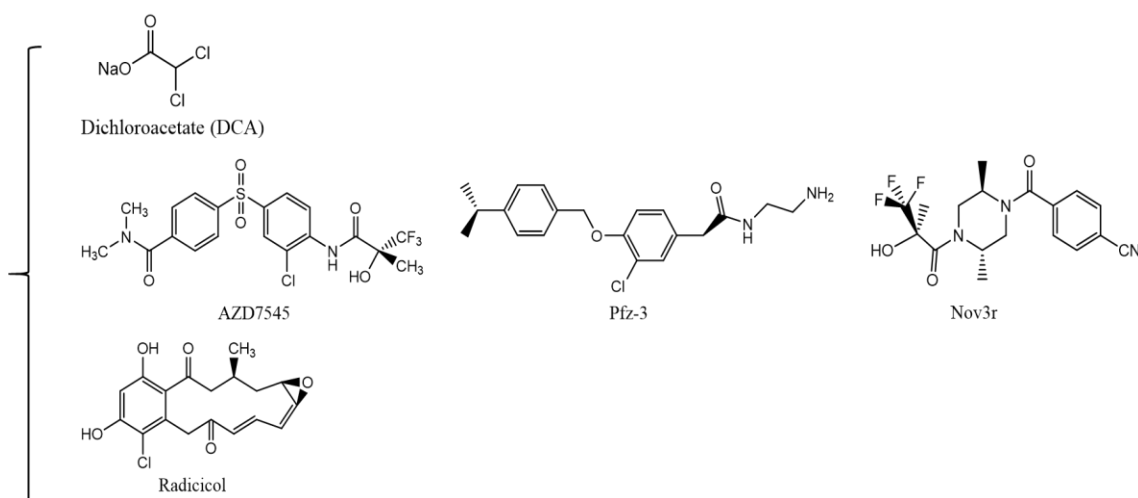


Figure 1.4: Structures of PDKs inhibitors

Additional inhibitors with higher selectivity across kinase family and sustained inhibitory effect, some forming covalent bonds and their mechanisms have been studied. Geng's group had investigated covalent modification at Cys240 of PDK1 with JX06, which could induce a conformational change through Van der Waal forces to block the access of ATP to its binding pocket and gave an IC_{50} value of 49 nM [36].

1.4 Objectives of this thesis

A recent study showed that PDHc regulation through PDK1 promotes aerobic glycolysis rather than pyruvate oxidation to oxidative phosphorylation in mitochondria (a phenomenon termed the *Warburg effect*), and plays a crucial role in glioblastoma multiforme, the most common and most aggressive malignant primary brain tumor in humans [37]. In addition, PDHc was also suggested to be a key player in oncogene-induced senescence. The oncogene-induced senescence was accompanied by simultaneous suppression of PDK1 and induction of PDP2. The resulting combined activation of PDHc enhanced the use of pyruvate in the tricarboxylic acid cycle, causing increased respiration and redox stress [38]. An elucidation of interactions between PDKs with E1 and E2p·E3BP components, and identification of interaction region between PDKs and components of PDHc can provide new insights into the long standing question of how human PDHc is assembled, and what kind of conformational changes human PDHc is experiencing through the interactions between PDKs and PDHc components.

Given that there is no X-ray or NMR structure for the intact E2 (from any source), alternative approaches must be developed to study the questions posed in this research.

The studies described herein were performed to determine the protein-protein interactions among six components of human pyruvate dehydrogenase complex. We determine the thermodynamics of binding of PDKs to other components of the PDHc by site-specific introduction of fluorophores onto specific domains of E2p·E3BP. In addition to that, we use the HDX-MS method to study the *in vitro* reconstitution of binary complexes of E1p and PDKs as well as the ligand effects on E1p and the PDK isozymes.

The author hopes that the experiments conducted in this thesis will help us to understand the activation mechanisms of PDKs and direct rational drug design of PDK isoform-specific inhibitors.

References

- [1] Randle P J, Garland P B, Hales C N, Newsholme E A. The glucose fatty-acid cycle its role in insulin sensitivity and the metabolic disturbances of diabetes mellitus[J]. The Lancet, 1963, 281(7285): 785-789.
- [2] Patel M S, Korotchkina L G. The biochemistry of the pyruvate dehydrogenase complex*[J]. Biochemistry and Molecular Biology Education, 2003, 31(1): 5-15.
- [3] Patel M S, Roche T E. Molecular biology and biochemistry of pyruvate dehydrogenase complexes[J]. The FASEB Journal, 1990, 4(14): 3224-3233.
- [4] Steussy C N, Popov K M, Bowker-Kinley M M, Sloan R B, Harris R A, Hamilton J A. Structure of Pyruvate Dehydrogenase Kinase A Novel Folding Pattern for A Serine Protein Kinase[J]. Journal of Biological Chemistry, 2001, 276(40): 37443-37450.
- [5] Wynn R M, Kato M, Chuang J L, Tso S C, Li J, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. Journal of Biological Chemistry, 2008, 283(37): 25305-25315.

- [6] Bowker-Kinley M M, Davis I W, Wu P, Harris R A, Popov K M. Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex[J]. *Biochemical Journal*, 1998, 329(1): 191-196.
- [7] Huang B, Gudi R, Wu P, Harris R A, Hamilton J, Popov K M. Isoenzymes of pyruvate dehydrogenase phosphatase DNA-derived amino acid sequences, expression, and regulation[J]. *Journal of Biological Chemistry*, 1998, 273(28): 17680-17688.
- [8] Di R, Feng Q, Chang Z, Luan Q, Zhang Y, Huang J, Li X, Yang Z. PDK1 plays a critical role in regulating cardiac function in mice and human[J]. *Chinese Medical Journal (English Edition)*, 2010, 123(17): 2358.
- [9] Sugden M C, Bulmer K, Augustine D, Holness M J. Selective modification of pyruvate dehydrogenase kinase isoform expression in rat pancreatic islets elicited by starvation and activation of peroxisome proliferator-activated receptor- α [J]. *Diabetes*, 2001, 50(12): 2729-2736.
- [10] Holness M J, Sugden M C. Regulation of pyruvate dehydrogenase complex activity by reversible phosphorylation[J]. 2003.1143-1151.
- [11] Pengfei Wu, Juichi Sato, Yu Zhao, Jaskiewicz J, Popov K M, Harris R A. Starvation and diabetes increase the amount of pyruvate dehydrogenase kinase isoenzyme 4 in rat heart[J]. *Biochemical Journal*, 1998, 329(1): 197-201.
- [12] Wu P, Inskeep K, Bowker-Kinley M M, Popov K M, Harris R A. Mechanism responsible for inactivation of skeletal muscle pyruvate dehydrogenase complex in starvation and diabetes[J]. *Diabetes*, 1999, 48(8): 1593-1599.
- [13] Kim Y I, Lee F N, Choi W S, Youn J H. Insulin regulation of skeletal muscle PDK4 mRNA expression is impaired in acute insulin-resistant states[J]. *Diabetes*, 2006, 55(8): 2311-2317.
- [14] Pehleman T L, Peters S J, Heigenhauser G J, Spriet L L. Enzymatic regulation of glucose disposal in human skeletal muscle after a high-fat, low-carbohydrate diet[J]. *Journal of applied physiology*, 2005, 98(1): 100-107.

- [15] Peters S J, Harris R A, Wu P, Pehleman T L, Heigenhauser G J, Spriet L L. Human skeletal muscle PDH kinase activity and isoform expression during a 3-day high-fat/low-carbohydrate diet[J]. *American Journal of Physiology-Endocrinology And Metabolism*, 2001, 281(6): E1151-E1158.
- [16] Kong S W, Bodyak N, Yue P, Liu Z, Brown J, Izumo S, Kang P M. Genetic expression profiles during physiological and pathological cardiac hypertrophy and heart failure in rats[J]. *Physiological Genomics*, 2005, 21(1): 34-42.
- [17] Zhao G, Jeoung N H, Burgess S C, Rosaaen-Stowe K A, Inagaki T, Latif S, Shelton J M, McAnally J, Bassel-Duby R, Harris R A, Richardson J A, Klierer S A. Overexpression of pyruvate dehydrogenase kinase 4 in heart perturbs metabolism and exacerbates calcineurin-induced cardiomyopathy[J]. *American Journal of Physiology-Heart and Circulatory Physiology*, 2008, 294(2): H936-H943.
- [18] Hwang B, Jeoung N H, Harris R A. Pyruvate dehydrogenase kinase isoenzyme 4 (PDHK4) deficiency attenuates the long-term negative effects of a high-saturated fat diet[J]. *Biochemical Journal*, 2009, 423(2): 243-252.
- [19] Kwon H S, Huang B, Unterman T G, Harris R A. Protein kinase B- α inhibits human pyruvate dehydrogenase kinase-4 gene induction by dexamethasone through inactivation of FOXO transcription factors[J]. *Diabetes*, 2004, 53(4): 899-910.
- [20] Huang B, Wu P, Bowker-Kinley M M, Harris R A. Regulation of pyruvate dehydrogenase kinase expression by peroxisome proliferator-activated receptor- α ligands, glucocorticoids, and insulin[J]. *Diabetes*, 2002, 51(2): 276-283.
- [21] Wu P, Peters J M, Harris R A. Adaptive increase in pyruvate dehydrogenase kinase 4 during starvation is mediated by peroxisome proliferator-activated receptor α [J]. *Biochemical and biophysical research communications*, 2001, 287(2): 391-396.
- [22] Lee I K. The role of pyruvate dehydrogenase kinase in diabetes and obesity[J]. *Diabetes & metabolism journal*, 2014, 38(3): 181-186.

- [23] Ravindran S, Radke G A, Guest J R, Roche T E. Lipoyl domain-based mechanism for the integrated feedback control of the pyruvate dehydrogenase complex by enhancement of pyruvate dehydrogenase kinase activity[J]. *Journal of Biological Chemistry*, 1996, 271(2): 653-662.
- [24] Behal R H, Buxton D B, Robertson J G, Olson M S. Regulation of the pyruvate dehydrogenase multienzyme complex[J]. *Annual review of nutrition*, 1993, 13(1): 497-520.
- [25] Kolobova E, Tuganova A, Boulatnikov I, Popov K M. Regulation of pyruvate dehydrogenase activity through phosphorylation at multiple sites[J]. *Biochemical Journal*, 2001, 358(1): 69-77.
- [26] Ciszak E M, Korotchkina L G, Dominiak P M, Sidhu S, Patel M S. Structural basis for flip-flop action of thiamin pyrophosphate-dependent enzymes revealed by human pyruvate dehydrogenase[J]. *Journal of Biological Chemistry*, 2003, 278(23): 21240-21246.
- [27] Korotchkina L G, Patel M S. Probing the mechanism of inactivation of human pyruvate dehydrogenase by phosphorylation of three sites[J]. *Journal of Biological Chemistry*, 2001, 276(8): 5731-5738.
- [28] Kato M, Wynn R M, Chuang J L, Tso S C, Machius M, Li J, Chuang D T. Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops[J]. *Structure*, 2008, 16(12): 1849-1859.
- [29] Kato M, Chuang J L, Tso S C, Wynn R M, Chuang D T. Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex[J]. *The EMBO journal*, 2005, 24(10): 1763-1774.
- [30] Whitehouse S, Cooper R H, Randle P J. Mechanism of activation of pyruvate dehydrogenase by dichloroacetate and other halogenated carboxylic acids[J]. *Biochemical Journal*, 1974, 141(3): 761-774.

- [31] Bersin R M, Stacpoole P W. Dichloroacetate as metabolic therapy for myocardial ischemia and failure[J]. American heart journal, 1997, 134(5): 841-855.
- [32] Morrell J A, Orme J, Butlin R J, Roche R M, Mayers E K. AZD7545 is a selective inhibitor of pyruvate dehydrogenase kinase 2[J]. 2003.
- [33] Kato M, Li J, Chuang J L, Chuang D T. Distinct structural mechanisms for inhibition of pyruvate dehydrogenase kinase isoforms by AZD7545, dichloroacetate, and radicicol[J]. Structure, 2007, 15(8): 992-1004.
- [34] Roe S M, Prodromou C, O'Brien R, et al. Structural basis for inhibition of the Hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin[J]. Journal of medicinal chemistry, 1999, 42(2): 260-266.
- [35] Sun W, Xie Z, Liu Y, Zhao D, Wu Z, Zhang D, Lv H, Tang S, Jin N, Jiang H, Tan M, Ding J, Luo C, Li J, Huang M, Geng M. JX06 selectively inhibits pyruvate dehydrogenase kinase PDK1 by a covalent cysteine modification[J]. Cancer research, 2015, 75(22): 4923-4936.
- [36] Velpula, K. K., Bhasin, A., Asuthkar, S., Tsung, A. J. (2013) Combined targeting of PDK1 and EGFR triggers regression of glioblastoma by reversing the Warburg effect. Cancer Res. Oct 22.
- [38] Hur, H., Xuan, Y., Kim, Y. B., Lee, G., Shim, W., Yun, J., Ham I. H., Han, S. U. (2013) Expression of pyruvate dehydrogenase kinase-1 in gastric cancer as a potential therapeutic target. Int. J. Oncol. 42, 44-54.
- [39] Kim J, Tchernyshyov I, Semenza G L, Dang C V. HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia[J]. Cell metabolism, 2006, 3(3): 177-185.

CHAPTER 2. Site specifically introduced external fluorophore to study protein-protein interactions within PDHc

2.1 Introduction

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^+ to NADH and the release of CO_2 [1-2].

The E2 component in those complexes is a dihydrolipoamide acyltransferase and has variable numbers of tandem lipoyl domains, each carrying a covalently attached lipoic acid as cofactor.

The E2p component of PDHc has a multi-domain structure, comprising from the N-terminal end: two tandem lipoyl domains, the outer (L1) and inner lipoyl domain (L2) approximately 9 kDa each, a peripheral subunit-binding domain (S, 4kDa) and the acetyltransferase or catalytic domain (Core, 28kDa), all separated by 25-30 amino acid-long flexible linkers.

The E3BP is composed of three linker-connected domains, similar but not identical to those in E2, a single lipoyl domain (L3), a variant of S, called S' to which E3 binds, and a catalytic domain C' that is incompetent to produce acetyl-CoA.

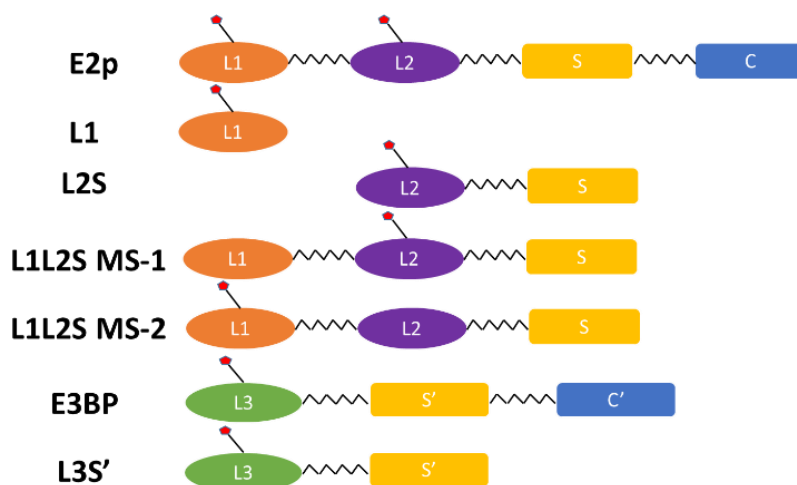


Figure 2.1: Domain structures of E2-E3BP- derived proteins

Trivalent arsenicals were used for potent irreversible inhibition of PDHc and many other proteins for many years, since they can form high-affinity ring structures with closely spaced dithiols [3-4].

The cyclic dithiolarsinites are markedly more stable than the non-cyclic products formed from trivalent arsenicals and monothiols because of entropic considerations [5].

1-Dimethylamino-naphthalene-5-sulfonyl chloride (DANS-Cl) also known as dansyl chloride is a useful fluorescent tag, which has been widely reacted with the lysine amino group of enzymes in attempts to understand the chemical nature of the active centers. Dansyl chloride is non-fluorescent until it reacts with amines. The fluorescence properties

of the resulting dansyl amides are environmentally sensitive and have large Stoke shifts [6-8].

In order to study the interaction between lipoyl domains and PDKs, we developed a method for *in situ* derivatization of the lipoyl domain. First of all, we use p-arsanilic acid (NH_2PhAsO) to react with dansyl chloride, forming a fluorescent tag which contains a trivalent arsenoxide end. Then, the fluorophore is specifically attached to a lipoyl domain by the reaction of reduced dihydrolipoyl groups with the trivalent arsenoxide end of DANS-As. Finally, the change in fluorescence is monitored when different PDKs are titrated with lipoyl domains bearing the fluorophore.

2.2 Site-specific labeling method to study protein-protein interactions within PDHc

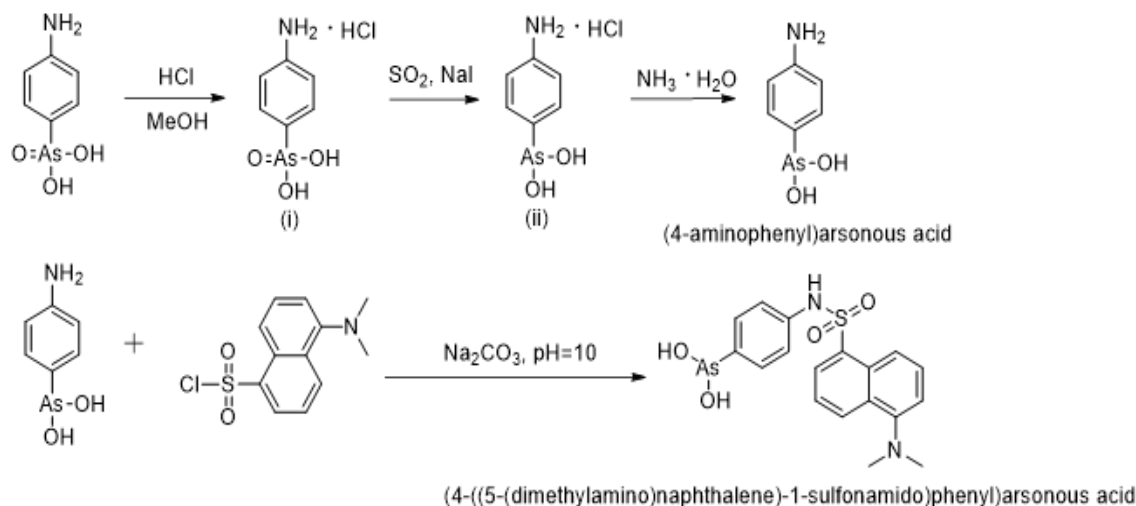
2.2.1 Materials and Methods

Recombinant E2•E3BP core was overexpressed in *E. coli* cells and was purified as described previously (Hiromasa et al., 2004). The following recombinant C-terminally truncated proteins of the E2•E3BP core available from the Patel group were used in this study: L1, containing the outer lipoyl domain and linker (residues 1–98); the L2S didomain containing the inner lipoyl domain (L2), second hinge region, peripheral subunit-binding domain (S), and third hinge region (residues 128–330); the L1L2S tridomain, which comprises L1, L2, both hinge regions, and the subunit-binding domain (S, residues 1–330); the L3S' didomain of E3BP (residues 1–230), all expressed from pET28b in *E. coli* BL21 (DE3) cells and purified using Ni Sepharose 6 Fast Flow column [9]. Two singly substituted variants of L1L2S were used in this study: L1L2S-MS1

(Lys46Ala substitution in the outer lipoyl domain with inner lipoyl domain available for modification), and L1L2S-MS2 (Lys173Ala substitution in the inner lipoyl domain with outer lipoyl domain available for modification). We also constructed four doubly substituted variants of L1L2S to study the function of specific residues on L1L2S: E35A/K173A, E35K/K173A, K46A/E209A, K46A/E209K.

Recombinant rat PDK1, rat PDK2, human PDK3 and rat PDK4 were from the Patel group. The PDK isozymes were overexpressed and purified individually from *E. coli* BL21(DE3) cells transformed with pPDK expression vector using protocols reported in the literature [10-13]. Analysis of the deduced amino acid sequences of human PDK1 and PDK2 reported in the literature revealed high identity between human and rat PDK2 (96%) and between human and rat PDK1 (93%). PDK3 shares 68% and 67% identity with PDK1 and PDK2, respectively [14].

The synthesis of DANS-As is described in Scheme 1. Briefly, p-arsanilic acid (10.85 g, 50 mmol) was acidified by hydrochloric acid and then reduced to (4-aminophenyl) arsenous acid using sulfur dioxide with sodium iodide as a catalyst [15]. Sulfur dioxide was bubbled into the solution for 2 h, then the mixture was cooled in an ice-water bath, and a pale yellow precipitate resulted. The precipitate (5 g) was dissolved in 15 mL of NH₄OH and stirred for 30 min. Once the solution was cooled in an ice-water bath, a white layer of crystals of (4-aminophenyl)arsenous acid precipitated. Next, (4-aminophenyl) arsenous acid (149 mg, 0.742 mmol) was reacted with dansyl chloride (100 mg, 0.371 mmol) at pH 10 for 8–12 h at room temperature producing (4-((5-(dimethylamino)naphthalene)-1-sulfonamido)phenyl) arsenous acid (DANS-As).



Scheme 2.1: Synthesis of (4-(5-(dimethylamino)naphthalene)-1-sulfonamido)phenyl)-arsenous acid (DANS-As).

In vitro lipoylation of the E2·E3BP derived proteins. The E2·E3BP derived proteins with the lipoyl domains were lipoylated *in vitro* using *E. coli* lipoyl protein ligase as reported by us recently [16]. Lipoylation was confirmed by FT-MS using the electrospray ionization sampling method.

Reaction of DANS-As with the source of the lipoyl domains (L1, L2S, L1L2S, L1L2S-MS1, L1L2S-MS2, doubly substituted L1L2S and L3S'). First, the lipoylated source of the lipoyl domain (50 μ M) in 30 mM KH_2PO_4 (pH 7.5) containing 0.15 M NaCl (sample buffer) was reduced by incubation for 5 min with TCEP (100 μ M) at room temperature, resulting in fully reduced lipoyl domains according to FT-MS. Second, the reduced lipoyl domains were then reacted with 150 μ M DANS-for 2 h at room temperature, leading to chemically modified lipoyl domains. The excess of TCEP and DANS-As were removed

from the reaction mixture with a 10K MWCO centrifugal filter unit (Vivaspin 500, 10K MWCO).

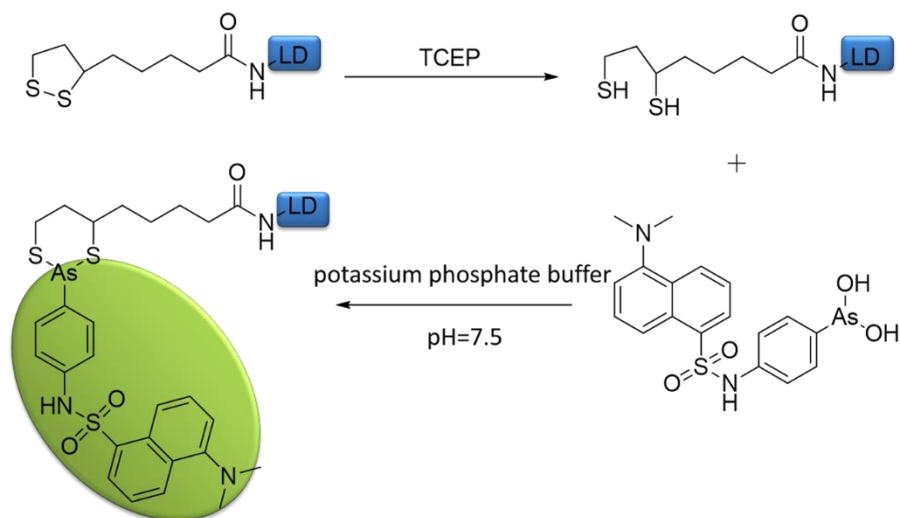


Figure 2.2: Using DANS-As to chemically modify LDs.

L3S' was modified by using Dapoxyl carboxylic succinimidyl ester. L3S' (50 μ M) was incubated with 100 μ M Dapoxyl carboxylic succinimidyl ester in sample buffer (30 mM KH_2PO_4 , 0.15 M NaCl, pH=7.5) at room temperature for 2 h. The reaction mixture was dialyzed against the sample buffer with a centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the extra Dapoxyl carboxylic succinimidyl ester.

Modification of the lipoyl domains was confirmed by FT-MS, for which each modified lipoyl domain containing protein was diluted to 1 μ M concentration in a solution containing 50:50:0.1 (v/v/v) of methanol, water and formic acid.

Modification of the lipoyl domains from doubly substituted L1L2S variants was confirmed by LC-MS. DANS-modified protein (50 ng in 20 μ L) was added to a 20 mM

sodium acetate buffer (pH=4.5) with a 50% slurry of immobilized pepsin. Then, the mixture was incubated at 37°C for 4 h with high speed shaking. The digested reaction mixture was separated from the immobilized pepsin by centrifugation, then, lyophilization. The lyophilized material was dissolved in 50 µL of 5% acetonitrile and LC-MS was run on a Phenomenex Luna C18 column.

Fluorescence Spectroscopy. For the fluorescence titration of the modified lipoyl domains by PDK isozymes, the DANS-modified lipoyl domain (2 µM) or Dapoxyl modified lipoyl domain (2 µM) in 30 mM KH₂PO₄ (pH7.5) were titrated by PDK1 (0.78-37.9 µM), or PDK2 (0.18-4.38 µM), or PDK3 (0.24-3.86 µM), or PDK4 (0.24-4.20 µM).

Fluorescence spectra were recorded at 25 °C using a Cary Eclipse spectrometer. The excitation wavelength was 338 nm, and the emission spectra were recorded in the range of 425-600 nm in 3 mL quartz cuvettes. Data were fitted to a Hill equation,

$$\Delta F / \Delta F_{max} = [PDK]^n / (S_{0.5}^n + [PDK]^n) \quad \text{Equation 2.1}$$

where $\Delta F / \Delta F_{max}$ is a relative fluorescence; $\Delta F = F_{max} - F_i$, where F_{max} is a maximum fluorescence intensity reached on titration by PDK and F_i is a fluorescence intensity at a given concentration of PDK; $\Delta F_{max} = F_{max} - F_o$ where F_o is the initial fluorescence before addition of PDK; $S_{0.5}$, is the concentration at half saturation. n is the Hill coefficient. For $n=1$, the value of $S_{0.5}$ is equal to K_d .

2.3 Results and discussion of E2-E3BP-derived proteins interacting with PDKs

The consecutive reactions of lipoylation, reduction, then attachment of the external fluorophore were monitored by mass spectrometry. An example of L1 is shown as figure.

On reduction of the lipoamide tethered to the L1, its mass was increased from 11972.09 Da to 11974.11 Da (the theoretical mass of the lipoylated L1 is 11971.3 Da). After attaching the DANS-As, the L1 mass was further increased to 12372.09 Da corresponding to the mass of an added DANS-As group.

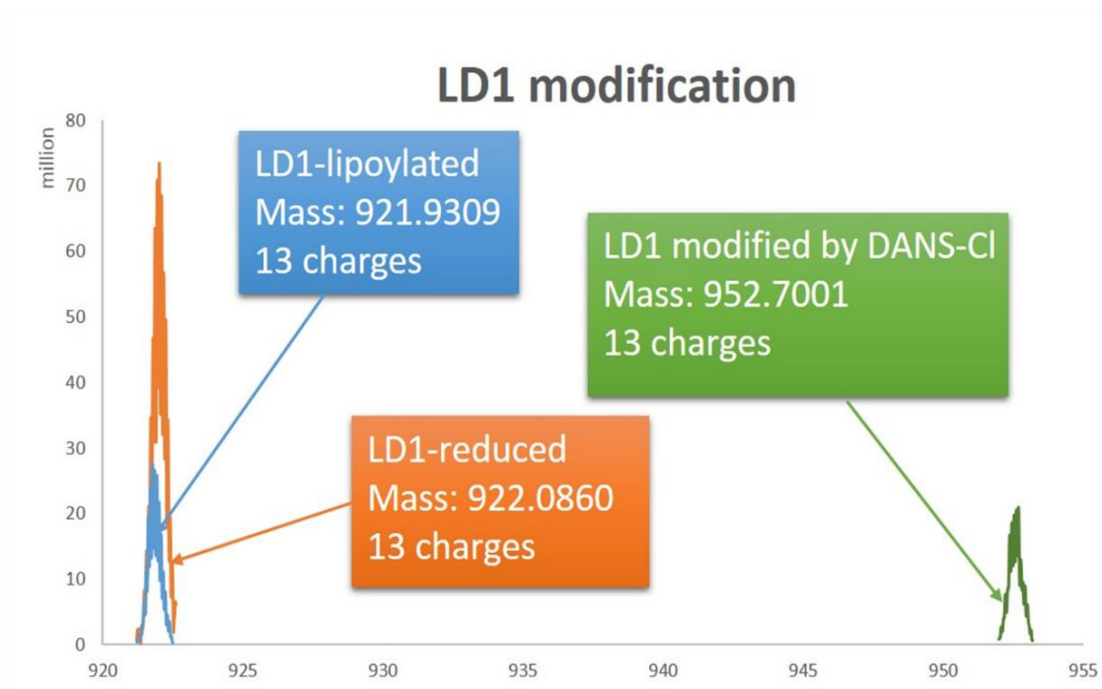


Figure 2.3: An example of chemical modification monitored by ESI-MS. The lipoamide from L1 can be reduced to the dihydrolipoyl group by TCEP, as shown on the mass spectrum: the peak with 13 charges increased from 921.9309 Da to 922.0860 Da. After attaching the DANS-As to L1, the peak increased from 922.0860 Da to 952.7001 Da, corresponding to the mass of adding the DANS-As group. The labeling efficiency during this process was approximately 75%.

Typical changes in the fluorescence of the DANS-As-labeled L2S and DANS-As-labeled L3S' on titration by PDK1 and PDK2 are shown in Figs. 2.4 and 2.5. It is evident that with the dansyl group attached to L2S, binding to PDK1 enhanced its fluorescence intensity. Similar Fluorescence intensity changes were observed for DANS-As-labeled L1, L2S, L1L2S-ML1, L1L2S-ML2 on titration by each of the four PDK isozymes. Attachment of the DANS-As group to L3S', in contrast, led to DANS fluorescence quenching rather than enhancement when interact with PDK2 and PDK3. These data suggest different binding environment for DANS-As-labeled L2S and L3S' on binding the PDK isoforms. The fluorescence enhancement suggests that dansyl group is transferred to a more hydrophobic environment, while the fluorescence quenching suggests that the dansyl group is transferred to a relatively more hydrophilic environment [17].

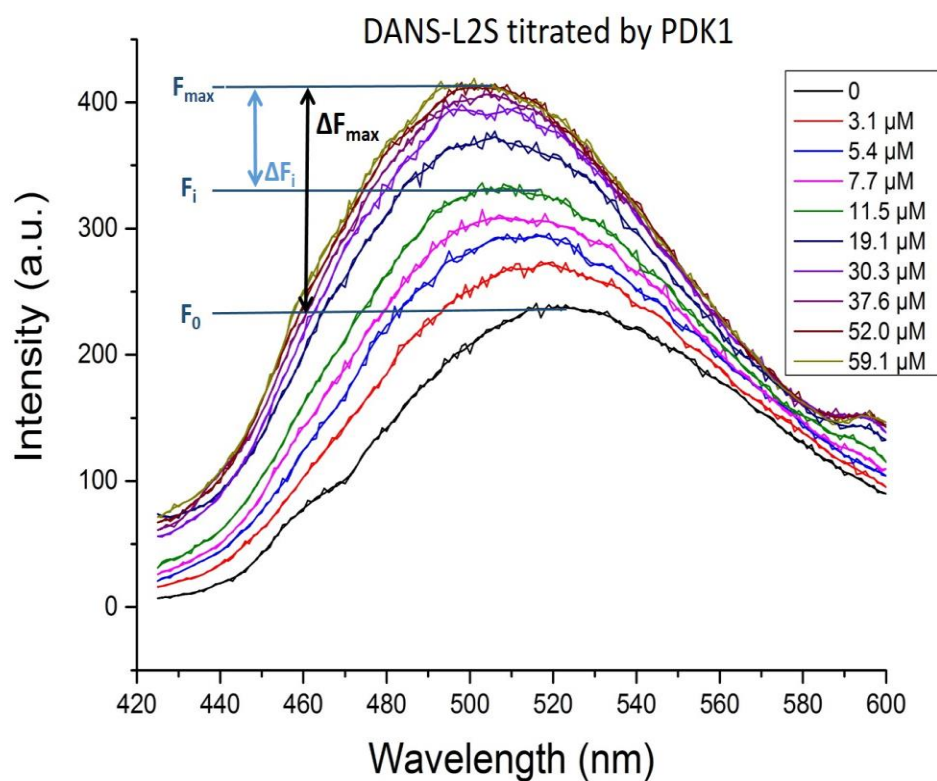


Figure 2.4: Fluorescence titration of DANS-As labeled L2S by PDK1. Enhancement of fluorescence of DANS-As-L2S upon PDK1 binding. DANS-As-L2S (1.5 μM) in 30 mM KH_2PO_4 (pH=7.5) was titrated by PDK1 (3.1-59.1 μM)

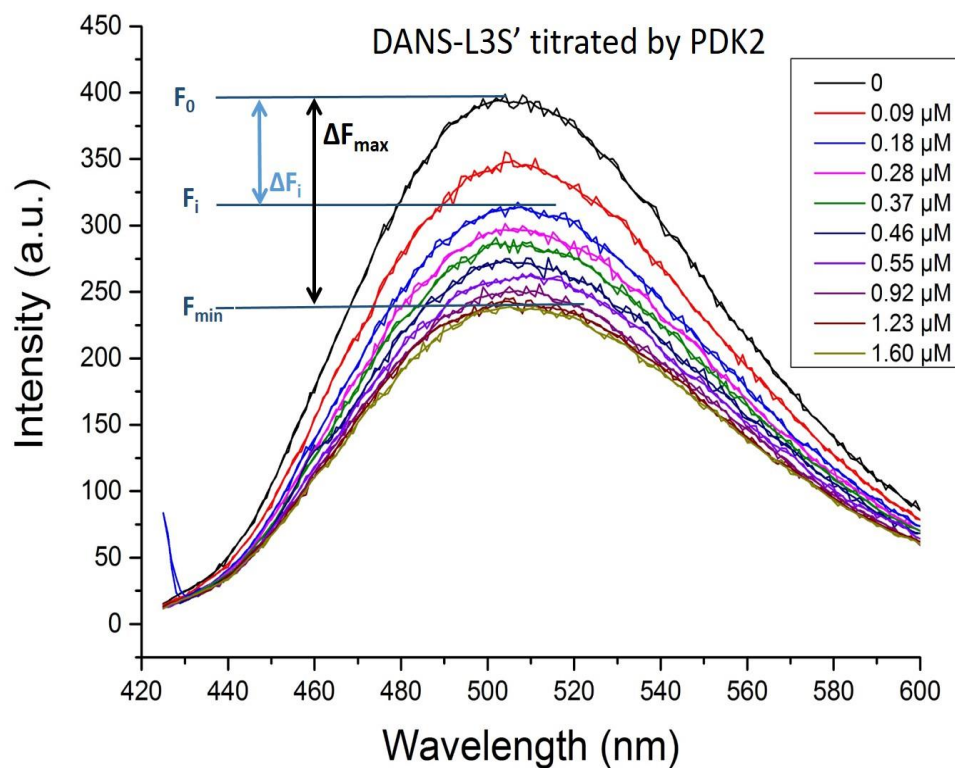


Figure 2.5: Fluorescence titration of DANS-As labeled L3S' by PDK2. Quenching of fluorescence of DANS-As-L3S' upon PDK2 binding. DANS-As-L3S' (1.5 μM) in 30 mM KH_2PO_4 (pH=7.5) was titrated by PDK2 (0.09-1.60 μM)

The fluorescence titration curves fitted with a Hill equation are summarized in the Figure below.

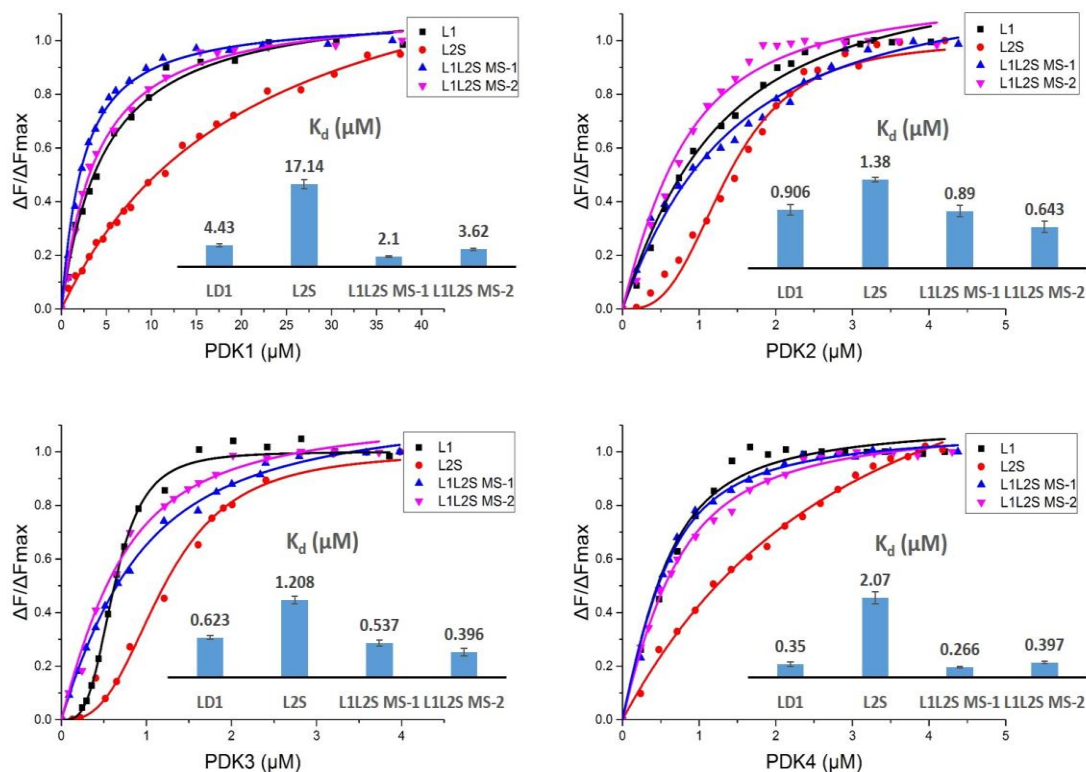


Figure 2.6: Fluorescence binding curves for DANS-As-labeled E2 E3BP derived lipoyl domains on titration by PDK isoforms. The DANS-As-labeled L1 (1.0 μM), or L2S (1.5 μM) or L1L2S MS-1 (2.5 μM) or L1L2S MS-2 (3.0 μM) in 30 mM KH_2PO_4 (pH=7.5) was titrated by PDK1 (0.78-37.93 μM) or PDK2 (0.18-4.38 μM) or PDK3 (0.24-3.86 μM) or PDK4 (0.24-4.2 μM) at room temperature. The excitation wavelength was 338 nm and the emission spectra were recorded in the 425-600 nm range. In all cases the fluorescence intensity of the DANS-labeled lipoyl domains was enhanced on PDKs binding.

A second fluorophore (the dapoxyl group) was also introduced to L3S', this time non-specifically to lysine residues of the L3S'. As seen in Figure 2.7, fluorescence

enhancement resulted from the interaction of dapoxyl-labeled L3S' with PDK2 and PDK4.

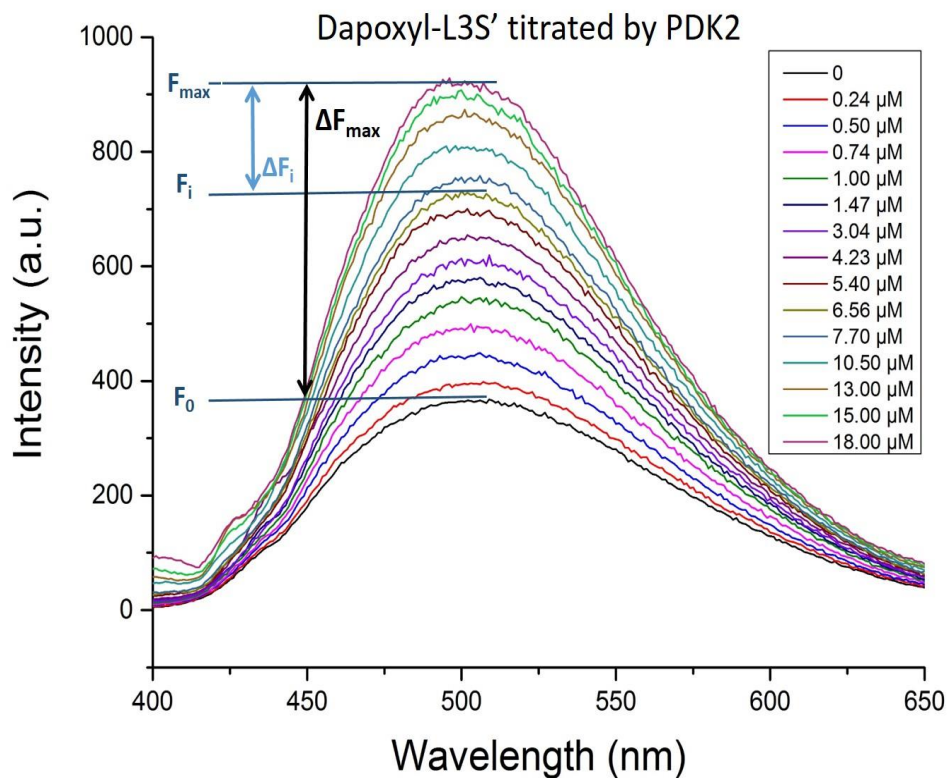


Figure 2.7: Fluorescence titration of dapoxyl-labeled L3S' by PDK2. The dapoxyl-L3S' (3 μM) in 30 mM KH₂PO₄ (pH=7.5) was titrated by PDK2 (0.24-18.0 μM) that was accompanied by the enhancement of the dapoxyl-L3S' fluorescence intensity. The PDK4 behaves similarly.

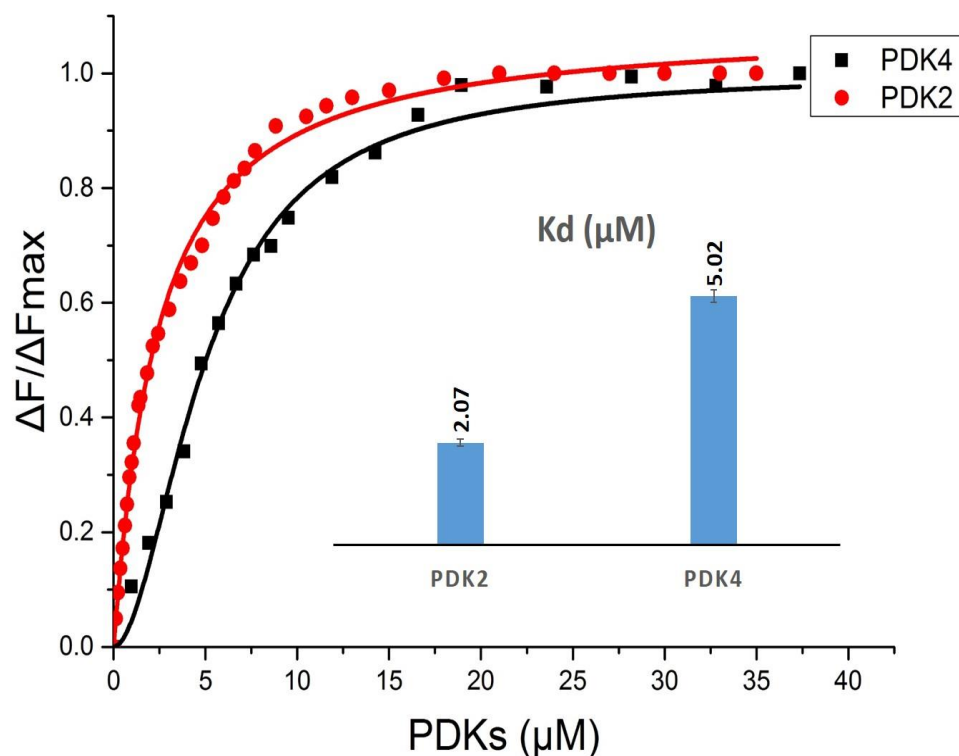


Figure 2.8: Fluorescence binding curves for dapoxyl-labeled L3S' on titration by PDK2 and PDK4. The K_d values for PDK2-L3S' and PDK4-L3S' complexes are 2.07 μM and 5.02 μM , respectively. Excitation wavelength was at 327 nm and the emission spectra were recorded in the 400-650 nm range. No changes in fluorescence intensity of dapoxyl-L3S' was observed on titration by PDK1 or PDK3.

The calculated values of K_d are presented in Table 2.1.

K_d (μM)	DANS-L1	DANS-L2S	DANS-L1L2S MS1	DANS-L1L2S MS2	DANS-L3S'	Dapoxyl-L3S'
PDK1	4.43 ± 0.35	17.14 ± 0.94	2.10 ± 0.12	3.62 ± 0.25	No binding	No binding
PDK2	0.91 ± 0.08	1.38 ± 0.03	0.89 ± 0.09	0.64 ± 0.09	0.025^a	2.07 ± 0.06

PDK3	0.62 ± 0.03	1.21 ± 0.06	0.54 ± 0.05	0.396 ± 0.06	0.103^a	No binding
PDK4	0.35 ± 0.06	2.07 ± 0.16	0.27 ± 0.02	0.397 ± 0.03	5.93 ± 1.06	5.02 ± 0.13

Table 2.1: K_d values for binding of the E2-E3BP-derived domains to PDK1-PDK4 as

detected by fluorescence spectroscopy. On binding of PDK2 and PDK3 to DANS-L3S', quenching of the DANS-L3S' fluorescence was observed while in others enhancement resulted.

As reported in Table 2.1, among the four PDK isoforms, the interaction between PDK1 and different DANS-As-labeled lipoyl domains is the weakest, which might explain why PDK1 is the only kinase that can phosphorylate all three sites of the E1 α subunit (15). It appears that PDK2 and PDK3 interact with the DANS-As-labeled lipoyl domains similarly, with both DANS-As-L1 and DANS-As-L2S contributing significantly to binding. The K_d values of the PDK2-DANS-As-L1 complex ($0.91 \mu\text{M}$) and PDK2-DANS-As-L1L2S-MS1 complex ($0.89 \mu\text{M}$) are nearly identical, as are the values for the PDK3-DANS-As-L1 ($0.62 \mu\text{M}$) and PDK3-DANS-As-L1L2S-MS1 ($0.54 \mu\text{M}$) complexes, suggesting that the lipoamide region in L2 does not contribute to binding of PDK2 or PDK3. The PDK4 has clear preference for L1 ($K_{d,L1} = 0.35 \mu\text{M}$; $K_{d,L2S} = 2.1 \mu\text{M}$), similarly to PDK2 and PDK3.

We also wished to investigate the interaction of the PDK isoforms with L3S' labeled with both the DANS-As and the dapoxyl group. The PDK1 displayed no binding with either labeling, while PDK4 showed weak binding (Table 2.1). Analysis of the titration data indicated stronger binding to PDK2 ($K_d = 2.1 \mu\text{M}$ for dapoxyl-L3S' and $0.025 \mu\text{M}$ for

DANS-As-L3S') than to PDK4 ($K_d \sim 5.0 \mu\text{M}$ with both labeled L3S'). The PDK3 isoform produced a signal only with L3S'.

In comparison with our data, an earlier report on the interaction of PDK4 with individual L1 and L2 domains was unable to identify these interactions by enthalpy changes using ITC, that study suggested weak binding. A K_d of $5 \mu\text{M}$ was reported for L3 binding to PDK4 [18], a value that correlates well with values of $K_d = 5.93 \mu\text{M}$ (DANS-As-L3S') and $K_d = 5.02 \mu\text{M}$ (dapoxyl-L3S') determined by fluorescence spectroscopy in Table 1. On the basis of the K_d values presented in Table 2.1, it became evident that L1 and L3S' also participate in the interaction with some PDK isoforms, not recognized before. This conclusion is supported by a comparison of the K_d values in Table 2.1 with those reported earlier, and using different approaches to calculate binding constants. Values of K_d of $\sim 175 \mu\text{M}$ (oxidized L2) and $K_d \sim 130 \mu\text{M}$ (reduced L2) for the PDK2-L2 complex were obtained by analytical ultracentrifugation [19]. On conjugation of L2 to glutathione-S transferase, the PDK2-L2 complex appeared to be stronger with $K_d \sim 3 \mu\text{M}$ (oxidized L2), and $K_d \sim 0.4 \mu\text{M}$ (reduced L2), as compared with $K_d \sim 22 \mu\text{M}$ (L1) and $K_d \sim 35 \mu\text{M}$ (L3), where L1 and L3 were also conjugated to glutathione-S transferase [20]. A value of $K_d \sim 10 \mu\text{M}$ was calculated for the PDK1-L2 and PDK2-L2 complexes using gel filtration chromatography, while the same method gave the following relative affinities for complexation with L2: PDK3 > PDK1 = PDK2 > PDK4 (46?). A value of $K_d = 1.17 \pm 0.23 \mu\text{M}$ was obtained by ITC for the PDK3-L2 complex [21].

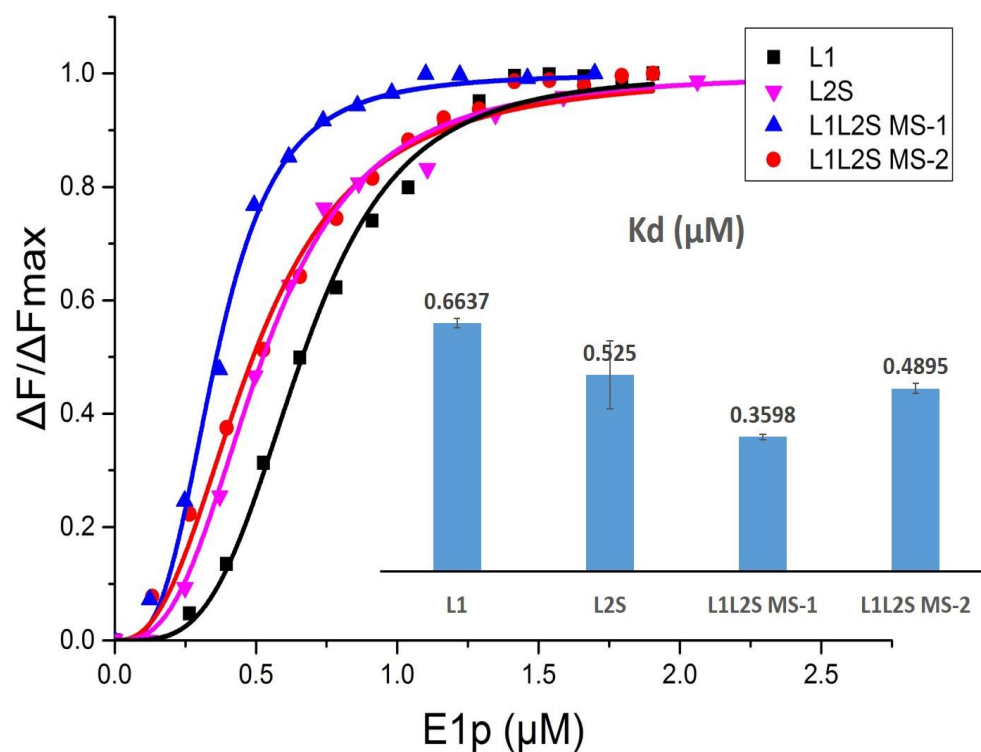


Figure 2.9: Fluorescence binding curves for DANS-As-labeled E2·E3BP derived lipoyl domains on titration by E1p.

We used the same method to study the interaction between E1p and different lipoyl domains. The results showed that E1p prefers L2S over L1. Chuang's group has published a K_d value of 1.29 ± 0.40 μM for E1p-L2 by using ITC [22]. Our value was lower than his, since our protein construct not only included the L2 domain, but also the subunit-binding domain.

2.4 Results of doubly substituted L1L2S variants interacting with PDKs and E1p

Previously, two complementary methods had been used to identify the ‘hot spot’ in interaction loci between PDK1 (PDK2) and the E2·E3BP-derived domains in our group. Some residues have been selected and the site-directed mutagenesis method was used to validate the function of these sites. Four doubly substituted variants of L1L2S have been constructed, they are L1L2S (E35A/K173A, E35K/K173A, K46A/E209A, K46A/E209K). Here, we utilize the same method we have described before, of introducing the DANS-As fluorophore onto the lipoyl groups of L1L2S doubly substituted variants, and then study the dissociation constants between these mutants and PDK isoforms.

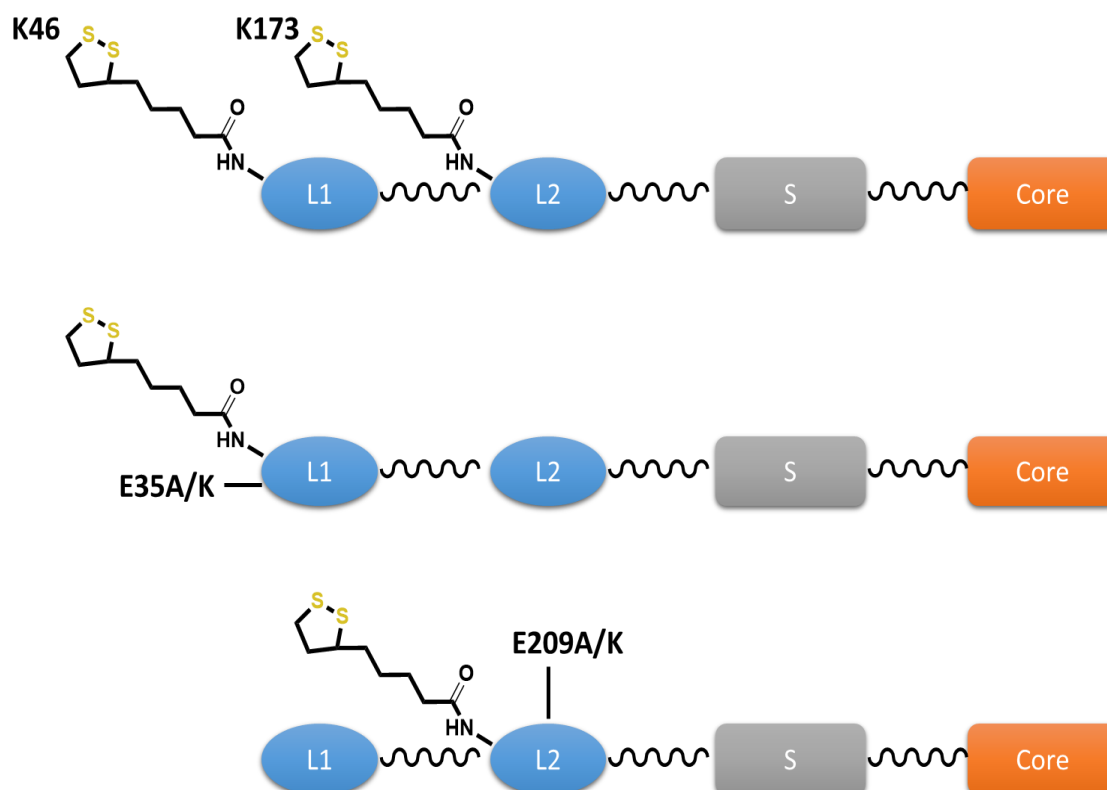


Figure 2.10: Domain structures of L1L2S doubly substituted variants (E35A/K173A, E35K/K173A, K46A/E209A, K46A/E209K).

First of all, pepsin digestion and then LC-MS were used to confirm the site-directed mutagenesis peptide sequence as well as the site-specific introduction of the DANS-As label.

Table 2.2: Peptides mass checked by LC-MS.

	peptides	Theoretical mass	Mass detected by FT-MS
K46	VETD K ATVGF+DANS	827.8022 ²⁺	827.8061 ²⁺
E35A	EKKEGD A INKGDL	708.8701 ²⁺	708.8721 ²⁺
K173A	IETD A ATIGF	1037.5150 ¹⁺	1037.5212 ¹⁺
E35K	EKKEGD K INKGDL	737.3990 ²⁺	737.4006 ²⁺
K46A	IAEVETD A ATVGF	661.8274 ²⁺	661.8309 ²⁺

Among the four double mutants (*E35A/K173A*, *E35K/K173A*, *K46A/E209A*, *K46A/E209K*) we tried, only *K46A/E209K* showed fluorescence change when titrated by PDKs. The spectra displaying fluorescence enhancement are displayed in Figure 2.10.

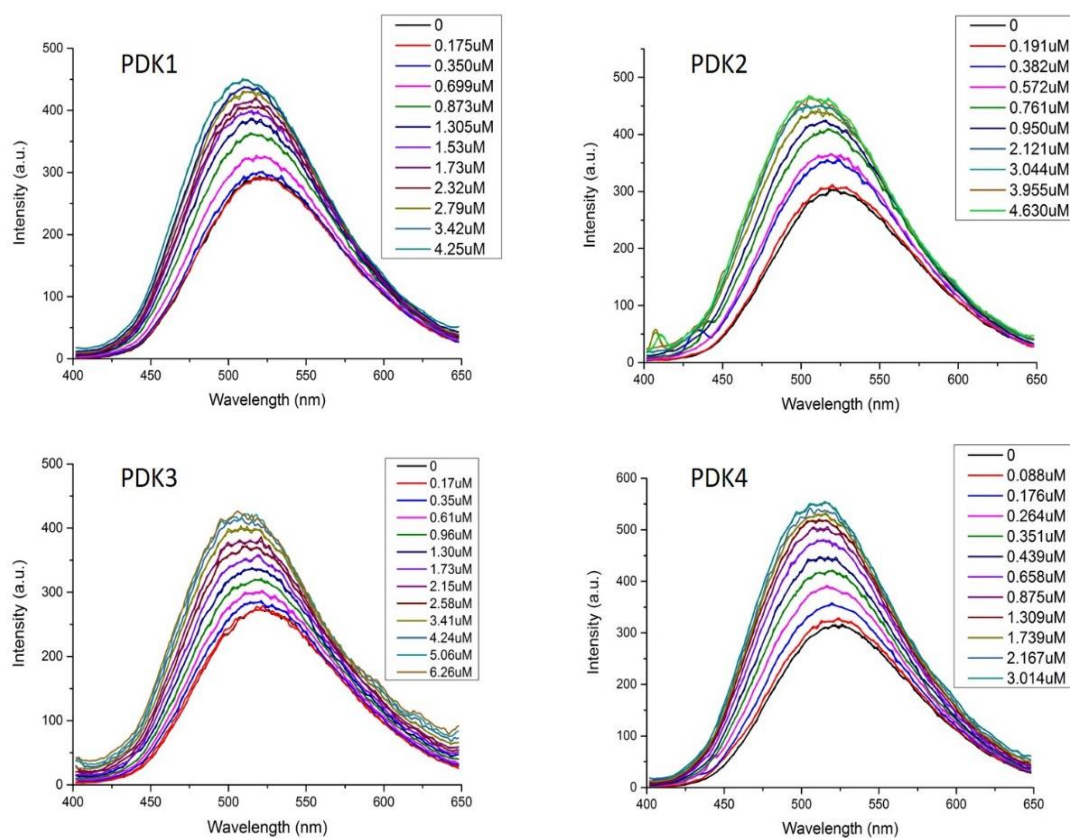


Figure 2.11: Fluorescence titration of DANS- labeled L1L2S (K46A/E209K) by four isoforms of PDKs. Enhancement of fluorescence of the DANS- L1L2S on PDKs binding.

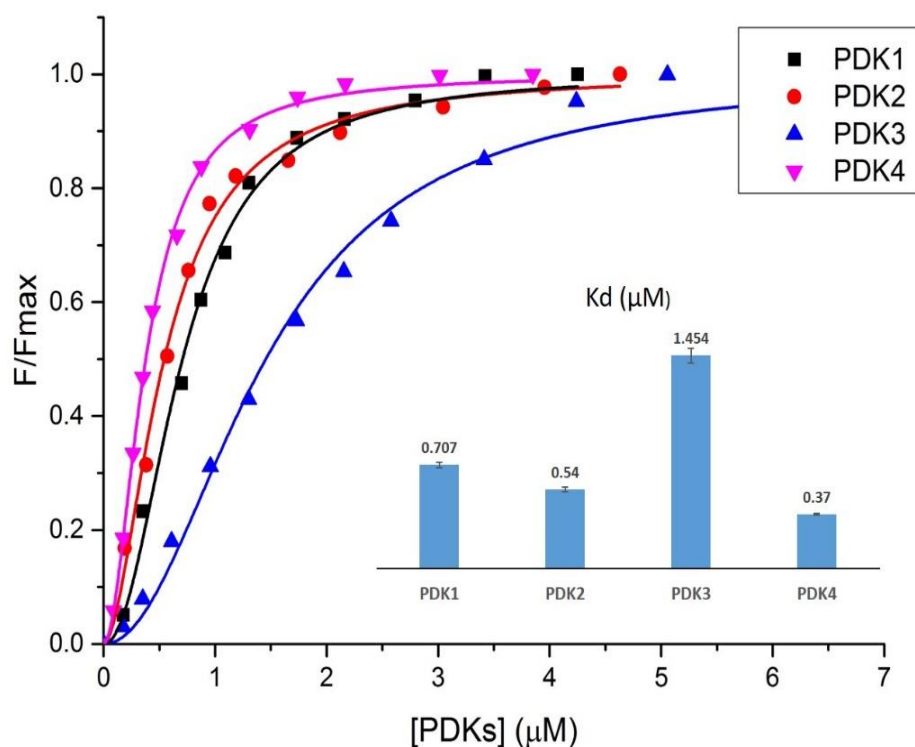


Figure 2.12: Fluorescence binding curves for DANS labeled L1L2S doubly substituted variant (K46A/E209K) on titration by PDK isoforms. In all cases the fluorescence intensity of the DANS-labeled lipoyl domains was enhanced on PDKs binding. The DANS-modified K46A/E209K (1.0 μM) in 30 mM KH_2PO_4 (pH 7.5) was titrated by PDKs at room temperature. The excitation wavelength was 338 nm and the emission spectra were recorded in the 400 - 600 nm range. The fluorescence titration curves were fitted by using the Hill equation (Equation 2.1 in Experimental procedures), as some of the curves displayed significant sigmoidicity.

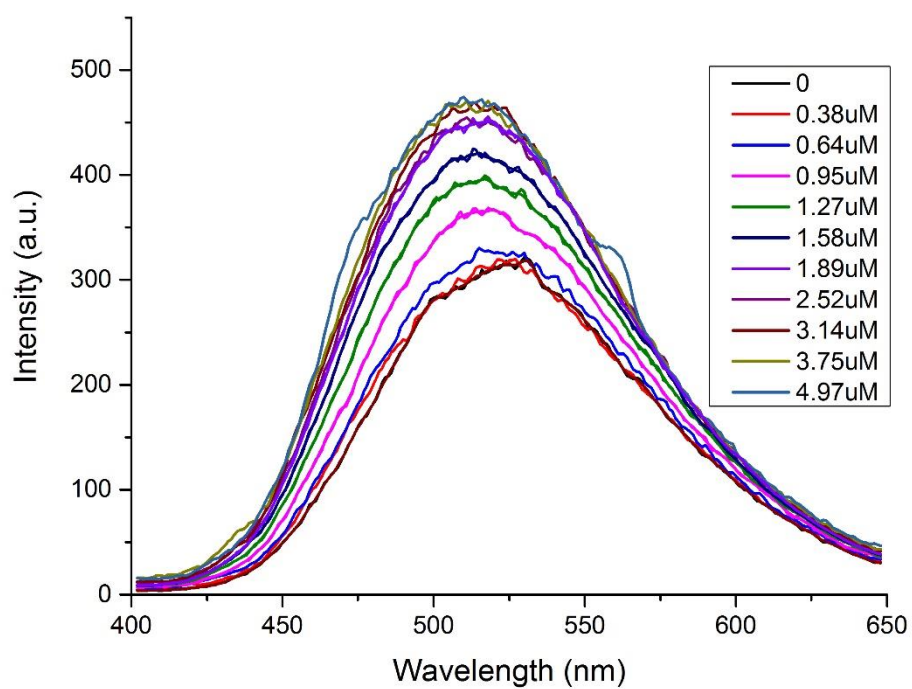


Figure 2.13: Fluorescence titration of DANS- labeled L1L2S (K46A/E209K) by E1p. Enhancement of fluorescence of the DANS- L1L2S on E1p binding.

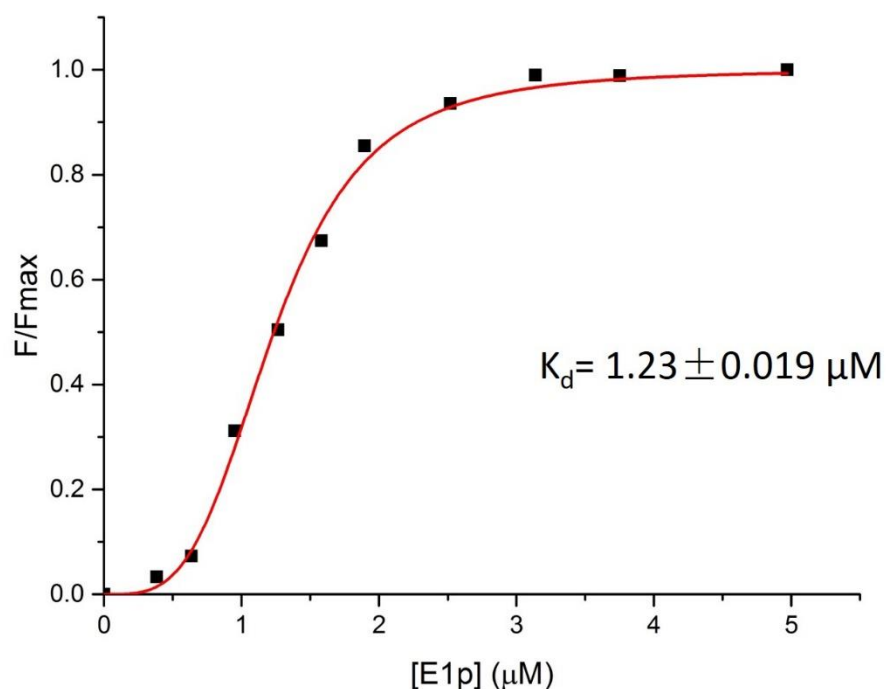


Figure 2.14: Fluorescence binding curves of DANS-As-L1L2S (K46A/E209K)

titrated by E1p. Binding of E1p enhanced the fluorescence intensity and the enhancement was used to measure the dissociation constants. The titration experiment was conducted in 30 mM KH₂PO₄ (pH 7.5) at room temperature. Excitation wavelength at 338 nm, the spectrum was recorded in the range of 400-600 nm.

/μM	K46A	K173A	K46A/E209K
PDK1	2.10±0.12	3.62±0.025	0.707±0.020
PDK2	0.89±0.09	0.643±0.029	0.54±0.017
PDK3	0.537±0.05	0.396±0.056	1.454±0.050
PDK4	0.266±0.02	0.397±0.029	0.37±0.005

E1p	0.36±0.007	0.49±0.013	1.23±0.019
------------	------------	------------	------------

Table 2.3: K_d values for binding of the L1L2S mutants to PDK1-PDK4 and E1p as detected by fluorescence spectroscopy. K_{ds} are in μ Molar units.

Among the four double variants (*E35A/K173A*, *E35K/K173A*, *K46A/E209A*, *K46A/E209K*) we tried, only *K46A/E209K* showed fluorescence change when titrated by both PDKs and E1p. And compared with the K_d values with those of single variants (*K46A* and *K173A*), the additional substitution on E209K site has an effect of increasing the ability of L1L2S binding with PDK1. It doesn't have much influence on binding with PDK2 and PDK4, however, it indeed weakens the ability of L1L2S binding with PDK3 and E1p.

2.5 Conclusions

It had been claimed earlier that only the inner lipoyl domain plays an important role in activation of PDKs, while other components such as the outer lipoyl domain, the substrate binding domain and also the third lipoyl domain derived from E3BP are all involved in interaction with PDKs and E1p. We here showed that the substitution of some additional residues such as residue E209 also impairs or enhances the binding ability towards PDKs.

References

- [1] Yeaman, S. J. (1986) The mammalian 2-oxoacid dehydrogenases: a complex family. *Trends Biochem. Sci.* 11, 293-296.
- [2] Randle, P. J. (1983) Mitochondrial 2-oxoacid dehydrogenase complexes of animal tissues. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 302, 47-57.
- [3] Park D, Don A S, Massamiri T, Karwa A, Warner B, MacDonald J, Hemenway C, Naik A, Kuan K, Dilda P, Wong J W.H., Camphausen K, Chinen L, Dyszlewski M, Hogg P. Noninvasive imaging of cell death using an Hsp90 ligand [J]. *Journal of the American Chemical Society*, 2011, 133(9): 2832-2835.
- [4] Donoghue N, Yam P T W, Jiang X M, Hogg P J. Presence of closely spaced protein thiols on the surface of mammalian cells [J]. *Protein Science*, 2000, 9(12): 2436-2445.
- [5] Dilda P J, Decollogne S, Weerakoon L, Norris M D, Haber M, Allen J D, Hogg P J. Optimization of the antitumor efficacy of a synthetic mitochondrial toxin by increasing the residence time in the cytosol [J]. *Journal of medicinal chemistry*, 2009, 52(20): 6209-6216.
- [6] Yamada H, Yamahara A, Yasuda S, Abe M, Oguri K, Fukushima S, Ikeda-Wada S. Dansyl chloride derivatization of methamphetamine: a method with advantages for screening and analysis of methamphetamine in urine[J]. *Journal of analytical toxicology*, 2002, 26(1): 17-22.
- [7] Bartzatt R. Fluorescent labeling of drugs and simple organic compounds containing amine functional groups, utilizing dansyl chloride in Na₂CO₃ buffer [J]. *Journal of pharmacological and toxicological methods*, 2001, 45(3): 247-253.

- [8] Sharma V, Wang Q, Lawrence D S. Peptide-based fluorescent sensors of protein kinase activity: design and applications [J]. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 2008, 1784(1): 94-99.
- [9] Patel M S, Korotchkina L G, Sidhu S. Interaction of E1 and E3 components with the core proteins of the human pyruvate dehydrogenase complex[J]. *Journal of Molecular Catalysis B: Enzymatic*, 2009, 61(1): 2-6.
- [10] Korotchkina L G, Patel M S. Probing the mechanism of inactivation of human pyruvate dehydrogenase by phosphorylation of three sites[J]. *Journal of Biological Chemistry*, 2001, 276(8): 5731-5738.
- [11] Baker J C, Yan X, Peng T, Kasten S, Roche T E. Marked differences between two isoforms of human pyruvate dehydrogenase kinase[J]. *Journal of Biological Chemistry*, 2000, 275(21): 15773-15781.
- [12] Bowker-Kinley M M, Davis I W, Wu P, Harris R A, Popov K M. Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex[J]. *Biochemical Journal*, 1998, 329(1): 191-196.
- [13] Wynn R M, Kato M, Chuang J L, Tso S C, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. *Journal of Biological Chemistry*, 2008, 283(37): 25305-25315.
- [14] Gudi R, Melissa M B K, Kedishvili N Y, Zhao Y, Popov K M. Diversity of the pyruvate dehydrogenase kinase gene family in humans[J]. *Journal of Biological Chemistry*, 1995, 270(48): 28989-28994.
- [15] Stevenson K J, Hale G, Perham R N. Inhibition of pyruvate dehydrogenase multienzyme complex from *Escherichia coli* with mono-and bifunctional arsenoxides[J]. *Biochemistry*, 1978, 17(11): 2189-2192.
- [16] Balakrishnan A, Nemeria N S, Chakraborty S, Kakalis L, Jordan F. Determination of pre-steady-state rate constants on the *Escherichia coli* pyruvate dehydrogenase complex

reveals that loop movement controls the rate-limiting step[J]. *Journal of the American Chemical Society*, 2012, 134(45): 18644-18655.

[17] Schuldiner S, Weil R, Robertson D E, Kaback H R. Microenvironment of the binding site in the lac carrier protein[J]. *Proceedings of the National Academy of Sciences*, 1977, 74(5): 1851-1854.

[18] Wynn R M, Kato M, Chuang J L, Tso S C, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. *Journal of Biological Chemistry*, 2008, 283(37): 25305-25315.

[19] Hiromasa Y, Roche T E. Facilitated interaction between the pyruvate dehydrogenase kinase isoform 2 and the dihydrolipoyl acetyltransferase[J]. *Journal of Biological Chemistry*, 2003, 278(36): 33681-33693.

[20] Crewe C, Schafer C, Lee I, Kinter M, Szweda L I. Regulation of Pyruvate Dehydrogenase Kinase 4 in the Heart through Degradation by the Lon Protease in Response to Mitochondrial Substrate Availability[J]. *Journal of Biological Chemistry*, 2017, 292(1): 305-312.

[21] Tuganova A, Boulatnikov I, Popov K M. Interaction between the individual isoenzymes of pyruvate dehydrogenase kinase and the inner lipoyl-bearing domain of transacetylase component of pyruvate dehydrogenase complex[J]. *Biochemical Journal*, 2002, 366(1): 129-136.

[22] Kato, M., Wynn, R. M., Chuang, J. L., Tso, S.-C., Machius, M., Li, J., and Chuang, D. T. (2008) Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops. *Structure*, **16**, 1849-1859.

CHAPTER 3. Using H/D Exchange Mass Spectrometry to Study the Interaction Patterns between E1p and Four Isoforms of PDKs

3.1 Introduction

The human pyruvate dehydrogenase complex (PDHc) provides the key link between glycolysis and the citric acid cycle. This large complex comprises four principal proteins: The E1 (pyruvate dehydrogenase) component carries out pyruvate decarboxylation and reductive acetylation of the E2 component (dihydrolipoamide acetyl transferase), reducing its covalently attached lipoamide to dihydrolipoamide and producing acetylCoA; the E3 component (dihydrolipoamide dehydrogenase) then reoxidizes dihydrolipoamide to lipoamide with concomitant formation of an equivalent of NADH from NAD^+ ; finally, the E3 binding protein (E3BP), acts as a bridge between the E2 and E3 components. A key to redox chemistry in such complexes is provided by lipoyl domains, two in the E2 domain (L1 or outer, and L2 or inner) and one in E3BP (L3). Primary regulation of the human PDHc is by covalent phosphorylation (turning off complex activity and NADH production) by four pyruvate dehydrogenase kinases (PDK1-4) and the activity could be recovered by two pyruvate dehydrogenase phosphatases (PDP1-2). Regulation of PDHc is achieved by site-specific phosphorylation of the α subunits of the heterotetrameric ($\alpha_2\beta_2$) E1p component at three different sites: Ser264 (site 1), Ser271 (site 2) and Ser203 (site 3). The three sites were phosphorylated *in vivo* to different extents, with maximum phosphorylation of site 1 and lesser phosphorylation of sites 2 and 3. The four PDK isozymes have different activity and phosphorylation rates at each site. At site 1 from fastest to slowest, $\text{PDK2} > \text{PDK4} \approx$

PDK1 > PDK3. For site 2, PDK3 > PDK4 > PDK2 > PDK1. Only PDK1 can phosphorylate site 3 [1]. Dephosphorylation of phosphoryl-E1 by PDP's recovers full PDHc activity.

The structure of PDKs comprises distinct N- and C-terminal domains. The N-terminal domain consists of eight α -helices, of which four form a bundle-like structure, the core. The sequences of the N-terminal domains of the four PDK isoforms are poorly conserved. In contrast, the C-terminal domain containing the phosphoryl transfer catalytic site is highly conserved among the four PDKs. They share four conserved motifs: (i) the N-box (Glu-X-X-Lys-Asn-X-X-X-Ala); (ii) the G1-box (Asp-X-Gly-X-Gly); (iii) the G2-box (Gly-X-Gly-X-Gly); and (iv) the G3-box (Gly-X-Gly-Thr) [2], all four together form a unique ATP-binding fold. This fold includes a common structural element known as the "ATP lid," whose conformational change is coupled to both ATP hydrolysis and protein-protein interactions [3].

Unlike the canonical ATP binding fold (in a cleft between a smaller N-terminal lobe of β -sheets and a larger C-terminal lobe of α -helices [4] shared among most of the hundreds of protein kinases in mammals, the ATP-binding fold in its K domain (ATP binds to one catalytic domain consisting of tightly aggregated alternating β -sheets and α -helices), indicates unique mechanisms for PDK activation and phosphorylation.

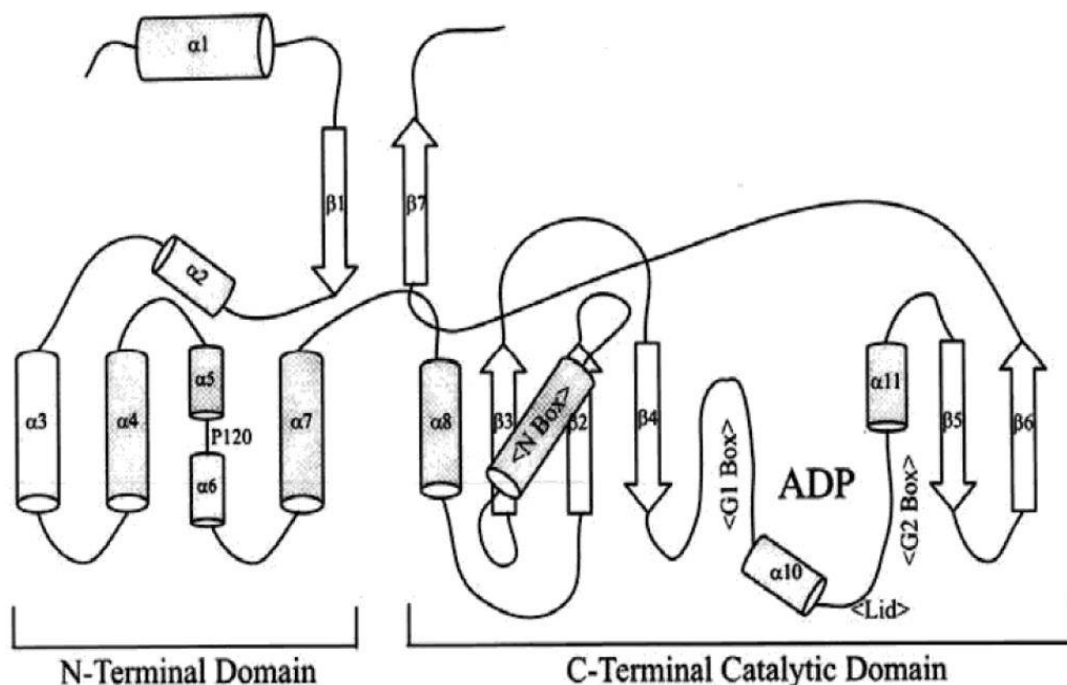


Figure 3.1: Topology of PDK2 [3].

It has been widely believed that lipoyl domains play a pivotal role in modulation of PDK activities, though different PDKs may have different preferences towards the lipoyl domains originating from E2 and E3BP. The structure of PDK3 bound to L2 domain showed that binding to the inner lipoyl domain induced a conformational change of PDK3, ordering the C-terminal tails of the dimer to form a ‘crossover’, thereby enlarge the active site cleft to facilitate the release of ADP [5]. As a result, most of the studies were focused on the interaction between E2·E3BP-derived proteins and PDKs. To date, only crystal structures of individual PDKs and some of their sub-complexes with L2 are available [5-9]. No structure with any E1p component is available.

In fact, PDKs can phosphorylate E1p without any source of E2·E3BP-derived proteins, at varying rates, some only slowly, others needing no activation from E2·E3BP. It is also

known that each of the four isozymes of PDKs has a preference to certain phosphorylation sites, however, the structural origin of this specific recognition between the PDK and the phosphorylation sites of E1p is still not understood.

Here, we have carried out a series of experiments using the hydrogen/deuterium exchange detected by mass spectrometry (HDX-MS) method to study the direct interactions between E1p and PDKs, trying to understand the differences among the four PDKs when forming complexes with E1p as well as to differentiate the four PDK isoforms by their unique pattern and provide some guides in a rational drug design of PDK isoform-specific inhibitors.

3.2 Materials and Methods

Deuterium oxide (D_2O) was from Cambridge Isotope Laboratories. All other fine chemicals were from Sigma-Aldrich. The E1p and the four PDK isoforms were purified and characterized as described in the Supplementary section.

3.2.1 Sample preparation for HDX-MS.

Prior to H/D exchange, the E1p and PDK were exchanged into buffer containing 50 mM KH_2PO_4 (pH=7.5), 50 mM KCl, 0.5 mM ThDP, and 1 mM $MgCl_2$. The E1p (80 μM) was incubated alone or with equal aubunit concentration of PDK. containing 5 mM ATP at 25 °C for 30 min prior to H/D exchange. The PDK isoform (80 μM) by itself or in the presence of 5 mM ATP were prepared as control. The HDX experiments were initiated by mixing 15 μl of the protein samples with 285 μl of D_2O buffer (E1p exchange buffer prepared in 99.9% D_2O) to yield a final concentration of 95 % D_2O at pH 7.5. The H/D exchange reactions were incubated at 25 °C for 30 s, 1, 3, 5, 10 and 30 min, and then

quenched by rapidly mixing with an equal volume 40 μ L of ice-cold quench buffer (99% trifluoroacetic acid, 3 M guanidine hydrochloride, pH 1.6) to reduce the final sample pH to 2.5. The samples were immediately frozen in liquid nitrogen and stored at -80 $^{\circ}$ C until 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.

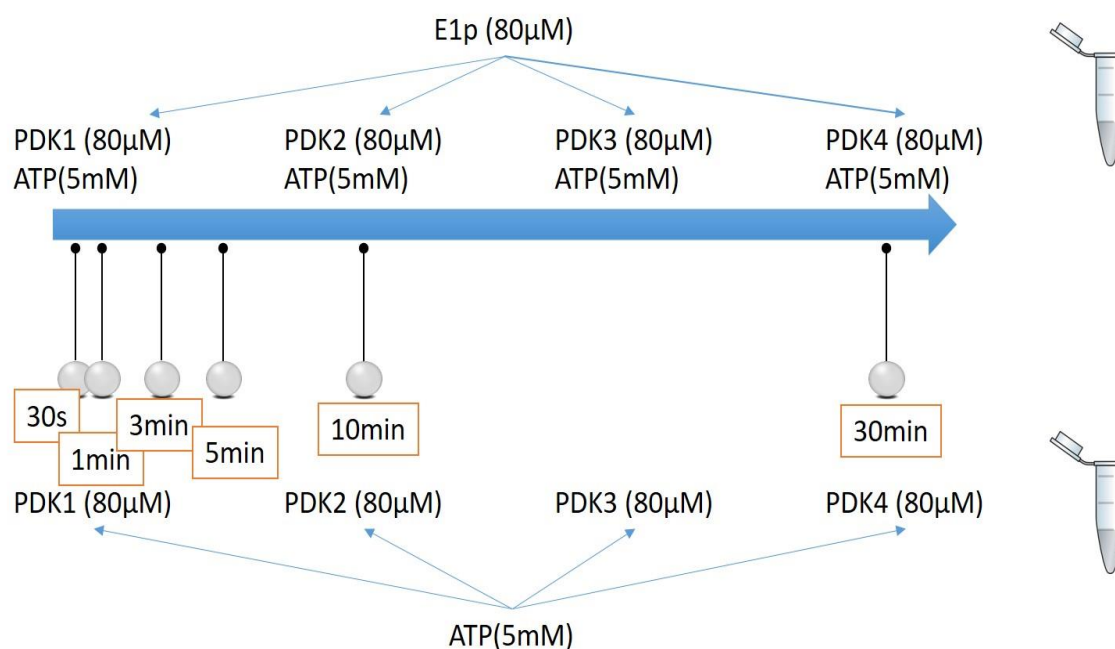


Figure 3.2: Experimental design. This series of experiments used free E1p and free PDK isoform as controls. Then the change of E1p and PDK upon complex formation was monitored in the presence of ATP.

3.2.2 LC-MS method and data processing

Frozen deuterated samples were quickly thawed and loaded with an ice-cold syringe into a 20 μ L sample loop inside the refrigeration system. The protein sample (40 pmol E1p)

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 over a 15 min period 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 un-deuterated 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 2.0 ppm.

3.3 Results and Discussion

3.3.1 Overview of HDX patterns on E1p

This is the first application of HDX-MS to study the direct interaction between E1p and all four PDK isoforms. The time dependence of hydrogen-deuterium exchange of the backbone amide protons of E1p and PDKs was studied over a 30 min time course (30 s, 1, 3, 5, 10 and 30 min). On-line digestion by pepsin followed by LC-MS analysis under

the selected HDX conditions yielded 24 peptides from E1p- α with 78% sequence coverage and 18 peptides from E1p- β with 89% sequence coverage, many of which were partially overlapping (Table 3.1 and 3.2)

The deuterium uptake percentage changes over the 30 min's time course and the average deuterium incorporation level of E1p at 30 min is illustrated in Fig. 3.3

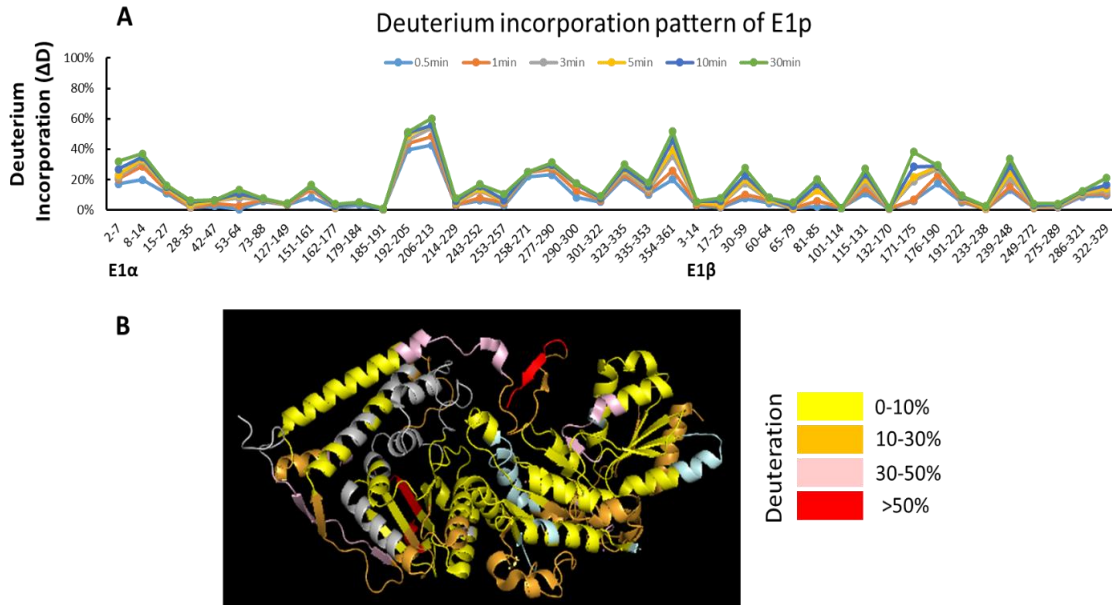


Figure 3.3: Deuterium incorporation pattern of E1p. **A.** The D uptake percentage changes of E1p over the 30 minutes. The percentage of D incorporation (without back exchange correction) of each peptide was calculated from the equation $\Delta D\% = \Delta D / (\max D \times 1.0063 \times 0.948) \times 100\%$, where 1.0063 is the atomic mass difference between D and H and 0.948 represents the fractional D₂O content of the labeling reaction mixture. **B.** The average percentage of D incorporation at the 30 min time point is mapped according to color (see legend) onto the crystal structure of *E1p*. The peptides that were not identified in E1p- α and E1p- β are shown in grey and light blue respectively.

The following information was revealed from an analysis of the HDX-MS data.

(1) From the peptic digestion, two peptides from E1p- α , ²⁵⁸YRYHGHSMKSDPGVS²⁷¹ and ¹⁹²ICENNRYGMGTSVE²⁰⁵, containing the three phosphorylation sites (Site 1-Ser264; Site 2-Ser271 and Site 3-Ser203), were identified.

(2) The overall rate of deuterium incorporation was low on this enzyme; many regions underwent less than 30% deuteration after a 30 min exchange. The deuterium uptake level of E1p- α is largely divided, with some regions showing deuterium uptake level below 10% while others higher than 30%. E1p- β exhibits relatively stable deuterium uptake.

(3) The deuterium incorporation level for peptides 127-149, 151-161 in E1p- α and 60-64, 65-79 in E1p- β are low. On the basis of the crystal structure (Ciszak et al.), a probable explanation is that in order to form the $\alpha_2\beta_2$ heterotetramer, an alpha subunit and beta subunit associate through predominantly hydrophobic contacts by permitting a pair of helices related to ThDP binding, 140-155 of the diphosphate binding domain from E1p-alpha and 59-72 of the pyrimidine-binding domain from E1p-beta, to pack together tightly.

(4) In contrast, peptides 192-205, 258-271 and 277-290 experienced relatively high deuterium incorporation. Peptides 258-271 and 277-290 are located on the conserved phosphorylation loop A (E1p- α 259-282), which forms one wall of the E1p active-site channel and helps anchor the ThDP to the active sites. On the other hand, peptide 192-

205 is located on the adjacent Ph-loop B (E1p- α 198-205) that provides coordination to a Mg^{2+} ion chelated by the diphosphate group of ThDP.

(5) In the wild-type E1p structure, Ser264 and Ser266 in the Ph-loop A form a hydrogen-bond network with two neighboring water molecules to form an ordered loop. And E1p- β' Tyr33 also participates in this H-bond network, as reflected by a 27.5% deuterium uptake level for peptide E1p- β 30-59.

3.3.2 Effects of the PDK isoforms on the E1p structure

Previous studies from this laboratory had focused on PDK activation by E2·E3BP or E2·E3BP-derived proteins, since it was believed that L2 of the E2·E3BP core is a docking site for association of PDKs with PDC, in turn responsible for activating the kinase and providing regulation. But there were also several reports showing that some PDKs are capable of phosphorylating free E1p without any help from E2·E3BP. Here, we use HDX-MS to study the influence of different PDKs exerted on E1p in the absence of E2·E3BP.

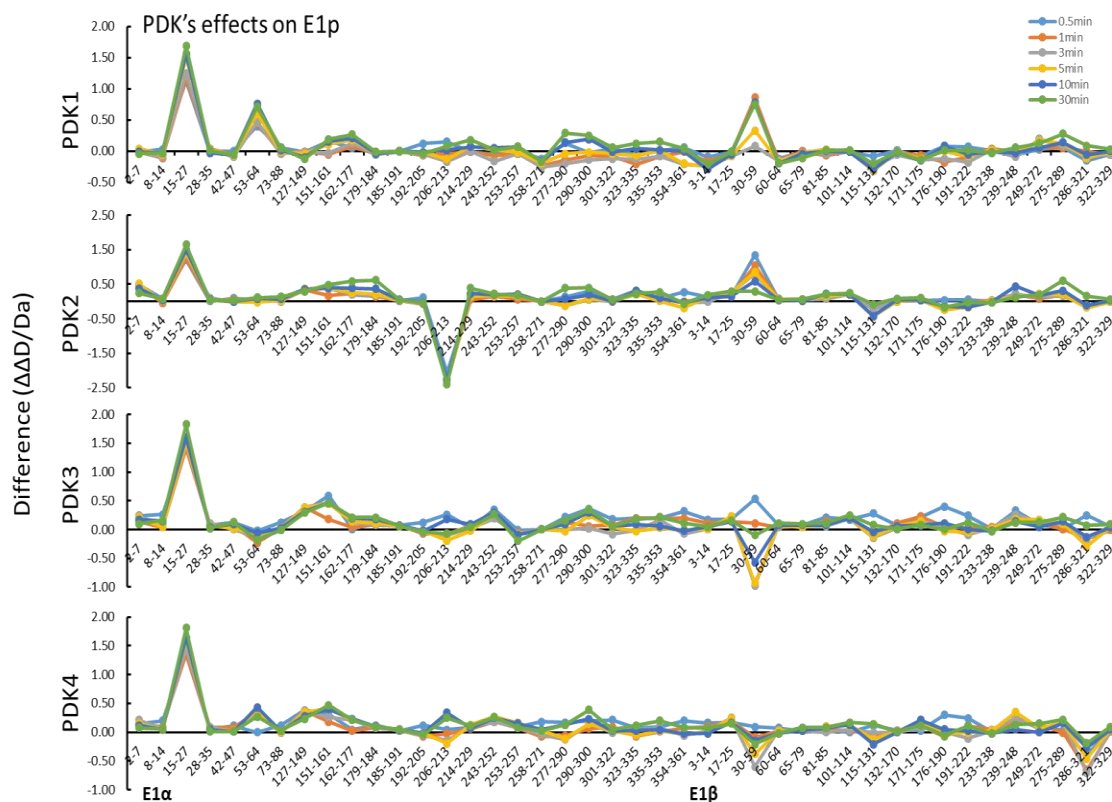


Figure 3.4: Effect of complexation with PDK isoforms on deuterium exchange in E1p.

A difference plot showing the changes in deuterium incorporation ($\Delta\Delta D$, y axis, deuterons exchanged of E1p in the presence of the corresponding PDK and ATP minus deuterons exchanged in the free form of E1p).

Upon incubation of E1p with PDKs and ATP, the N-terminal region of E1p- α (residues 15-27) has undergone a conformational change irrespective of the PDK selected. With PDK1, the most significant HDX-MS changes were observed in the N-terminal region of both E1p- α and E1p- β , which experienced a $\Delta\Delta D$ change around 1.7 Da for peptide E1p- α 15-27, a change of 0.7 Da for peptide E1p- α 53-64 and of 0.75 Da for peptide E1p- β 30-59. We could not identify any potential binding sites for PDK1 on E1p with the

current sequence coverage. But, it's clear that PDK1 induces a conformational change on E1p, especially on the N-terminal regions of both the E1p- α and E1p- β subunits.

The interaction pattern between E1p and PDK2 is slightly different from that with PDK1. Except for the higher deuterium incorporation level of peptide E1p- α 15-27 ($\Delta\Delta D = 1.65$ Da), there is also a potential binding site of PDK2 on peptide E1p- α 206-213 ($\Delta\Delta D = -2.39$ Da), which is adjacent to phosphorylation site 3. The effects of PDK2 on the N terminal of E1p- α subunit is less significant compared to that of PDK1, since the deuterium uptake level of peptide E1p- α 53-64 doesn't change in the presence of PDK2 while the same peptide displays a $\Delta\Delta D$ increase of 0.75 Da.

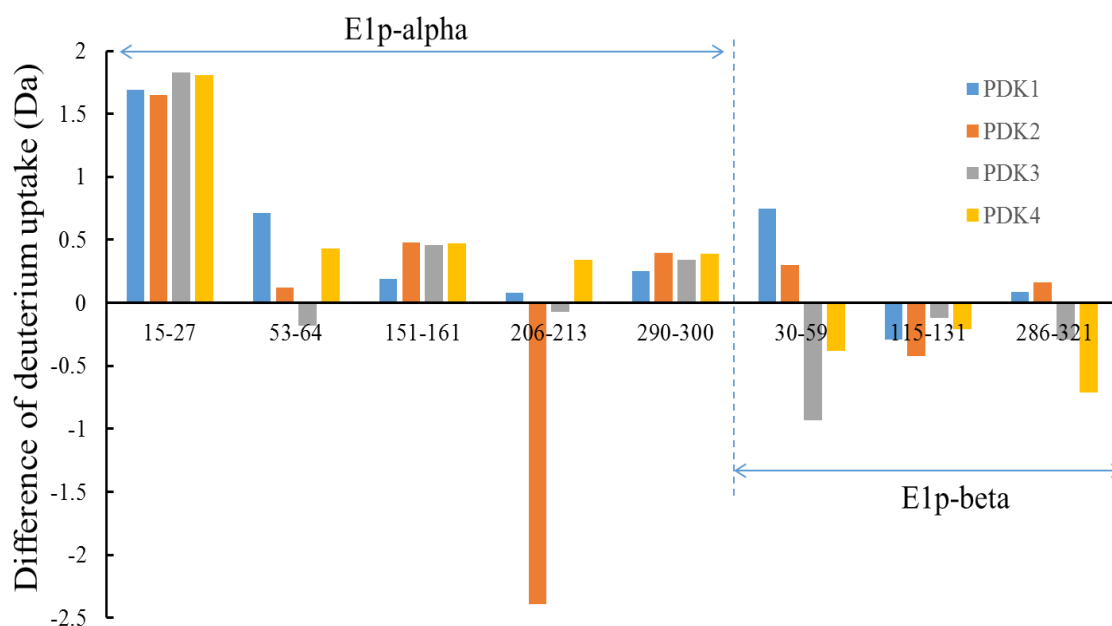


Figure 3.5: Effects of four PDK isozymes on H/D exchange in E1p. Selected peptides from E1p that show significant deuteriation incorporation change upon complexation with PDKs. The interaction patterns of E1p with PDK3 and E1p with PDK4 are alike. Both of

them exhibit a large positive $\Delta\Delta D$ change for peptide E1p- α 15-27 with 1.83 Da and 1.81 Da respectively. E1p may undergo a small conformational change in the region of E1p- α 127-161 when incubated with PDK3 or PDK4. The interaction that occurs at the N-terminal region of E1p- β might be transient, since the deuterium incorporation level of peptide E1p- β 30-59 decreased in the course of 30 min. There are some minor differences in the effects exerted by PDK3 and PDK4. In the presence of PDK4, the peptide of E1p- α 53-64 experienced a $\Delta\Delta D$ change of 0.4 Da, while in the presence of PDK3, this peptide showed a slight decrease with $\Delta\Delta D$ of -0.18 Da. In addition, the PDK4 also induced some transient effects on the C-terminal region of E1p- β 268-321.

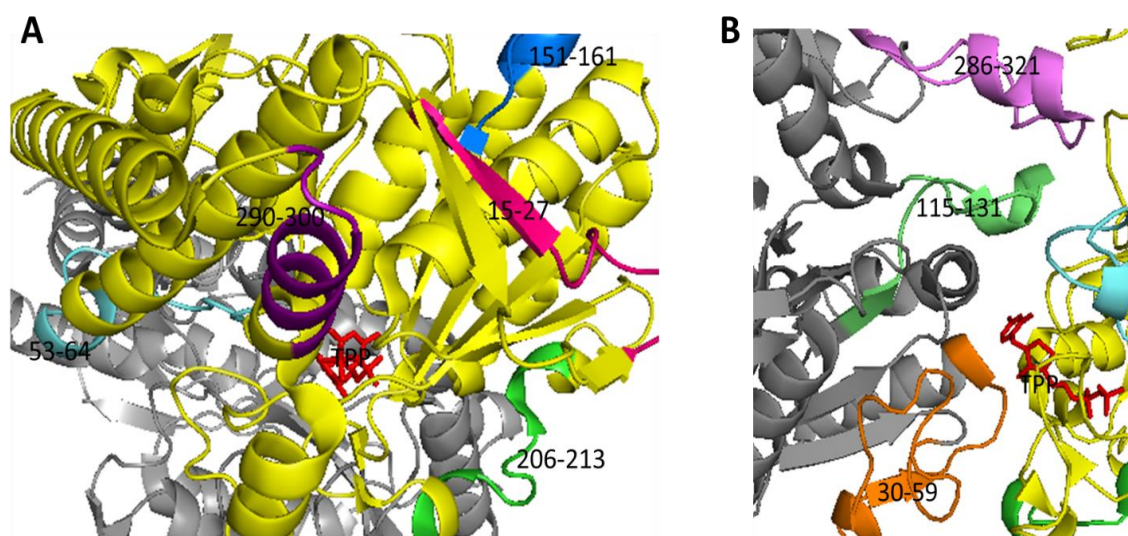


Figure 3.6: Peptides of human E1p involved in conformational change upon complexation with PDKs. A. Five peptides from the E1p- α subunit that show significant $\Delta\Delta D$ change upon complexation with PDKs are highlighted in colors: 15-27 (pink); 53-64 (cyan); 151-161 (blue); 206-213 (green); 290-300 (purple). B. Three peptides from the E1p- β subunit that show significant $\Delta\Delta D$ change upon complexation with PDKs are highlighted in colors: 30-59 (orange); 115-131 (lime); 286-321 (magenta).

The binding with PDKs causes significant conformational change of the loops at the entrance of the ThDP active center, which explains the significant $\Delta\Delta D$ in two peptides (E1p- α 53-64 and E1p- β 30-59, lying at the entrance). Although the peptide E1p- β 30-59 has no direct contact with the ThDP cofactor, it is effectively connected to the active center regions of E1p- α , by hydrogen bonds of $\alpha Q172$ - $\beta E59$, $\alpha Q172$ - $\beta I57$, and a salt bridge of $\alpha R206$ - $\beta D54$, by which it could induce conformational change of peptide E1p- α 206-213, once it changes itself. All of these conformational changes may induce the E1p- α subunit

to expose the phosphorylation sites, especially site 1 (the major phosphorylation site), which is located at the substrate channel, to solvent and facilitate the phosphorylation reaction.

3.3.3 Effects of PDKs on E1p phosphorylation

That pyruvate dehydrogenase kinase (PDK) can inactivate PDC by phosphorylation was first discovered in the late 1960s. (Linn et al.,) The reversible phosphorylation occurs on three specific serine residues on E1p- α subunit at site 1, Ser264; site 2, Ser271; and site 3, Ser203. The four PDK isoforms have different specificity and different rates toward the three phosphorylation sites. All four PDKs phosphorylate site 1 and site 2, while only PDK1 could phosphorylate site 3.

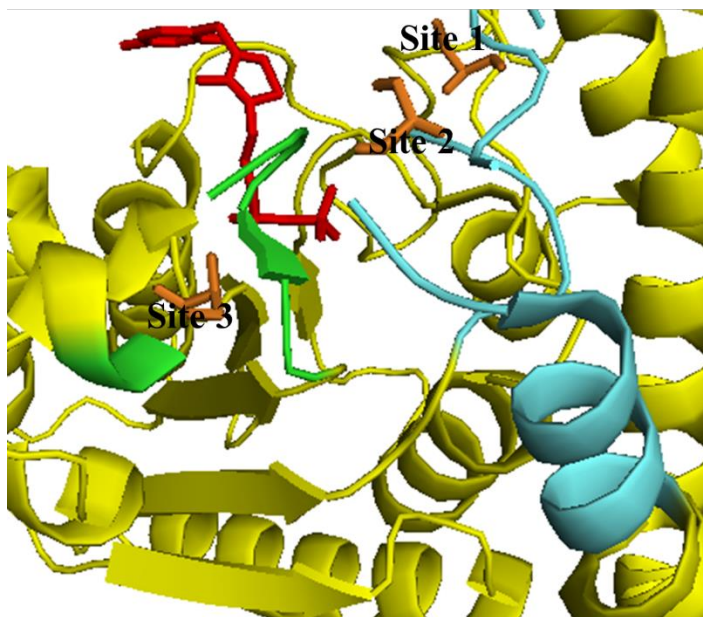


Figure 3.7: Phosphorylation sites on the E1p- α subunit. Site 1 (Ser264) and Site 2 (Ser271) are located on Ph-loop A (259-282), colored in *cyan*. Site 3 (Ser203) is located on Ph-loop B (198-205), colored in *lime*.

Here, we used MS to detect the PO_4^{3-} incorporation by analyzing the two peptides: $^{258}\text{YRYHGHSMKSDPGVS}^{271}$ and $^{192}\text{ICENNRYGMGTSVE}^{205}$. Since the peptide $^{258}\text{YRYHGHSMKSDPGVS}^{271}$ contains both sites 1 and 2, we cannot differentiate these two sites without further fragmentation. The ratio is calculated from (peak intensities of phosphorylated peptide) / (peak intensities of phosphorylated peptide + peak intensities of original peptide).

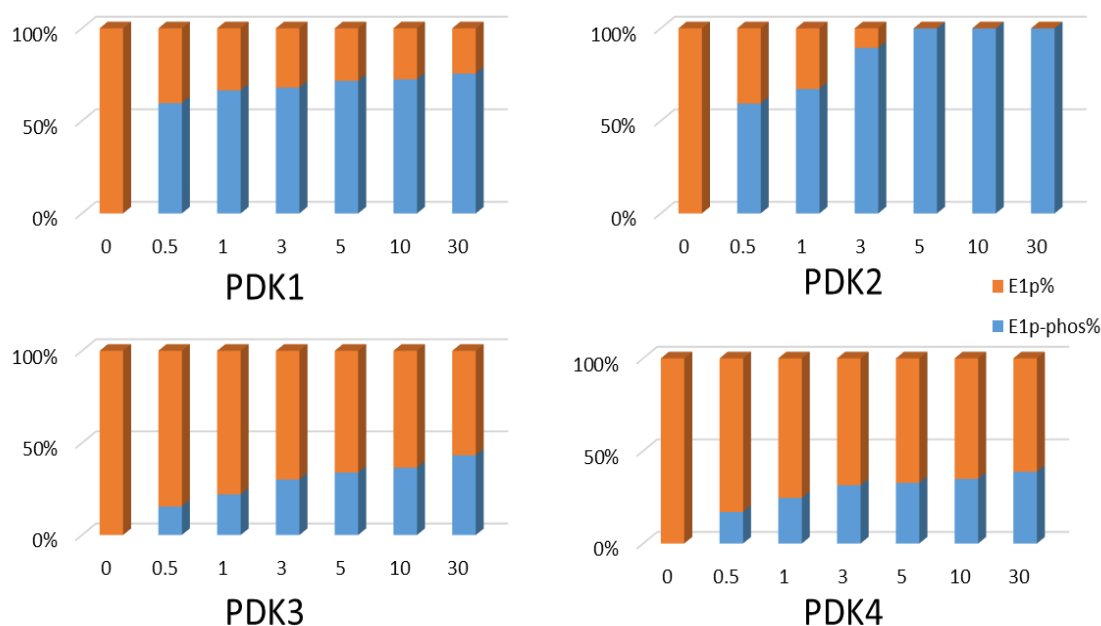


Figure 3.8: PO_4^{3-} incorporation ratios on peptide $^{258}\text{YRYHGHSMKSDPGVS}^{271}$ during 30 min reaction. The *blue* represents the fraction of phosphorylated peptide and *orange* represents the fraction of original peptide remaining.

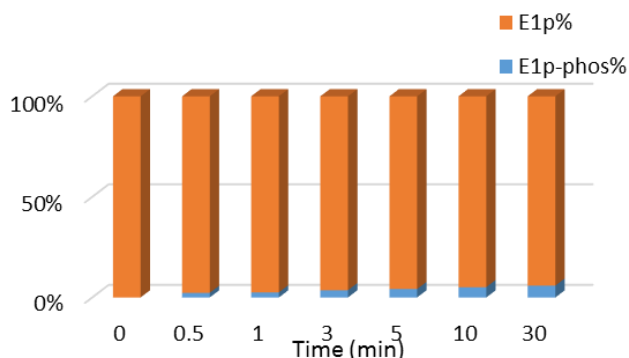


Figure 3.9: PO_4^{3-} incorporation on peptide $^{192}\text{ICENNRYGMGTSVE}^{205}$ during 30 min reaction with PDK1. The *blue* represents the fraction of phosphorylated peptide and *orange* represents the fraction of original peptide remaining.

Most of these MS results are in agreement with the activity data from the Patel group [10]. PDK2 has the highest activity in the absence of E2·E3BP, which can almost fully phosphorylate E1p within 5 min. Even though PDK1 can phosphorylate all three sites, the activity towards site 1 is much higher than towards site 3, in accord with a small percentage (5% percent) phosphorylation at site 3 after 30 min. Here, we reported that without E2·E3BP, PDK3 and PDK4 can still phosphorylate E1p to an extent of approximately 40%. While the activity measured by Patel's group for PDK3 was 12-times lower than for PDK4. Recently, we also reported PDK activities assessed by measuring the overall PDHc activity (NADH production). It was shown that 210 min incubation of E1p with PDK1 led to retention of 90% of the PDC activity remaining, while 90 min incubation with PDK2 led to retention of only 15% activity remaining. PDK3 and PDK4 by themselves exhibited little inactivation of PDC: after 60 min incubation, 95% of PDC activity was retained [11].

3.3.4 Effects of E1p on PDKs

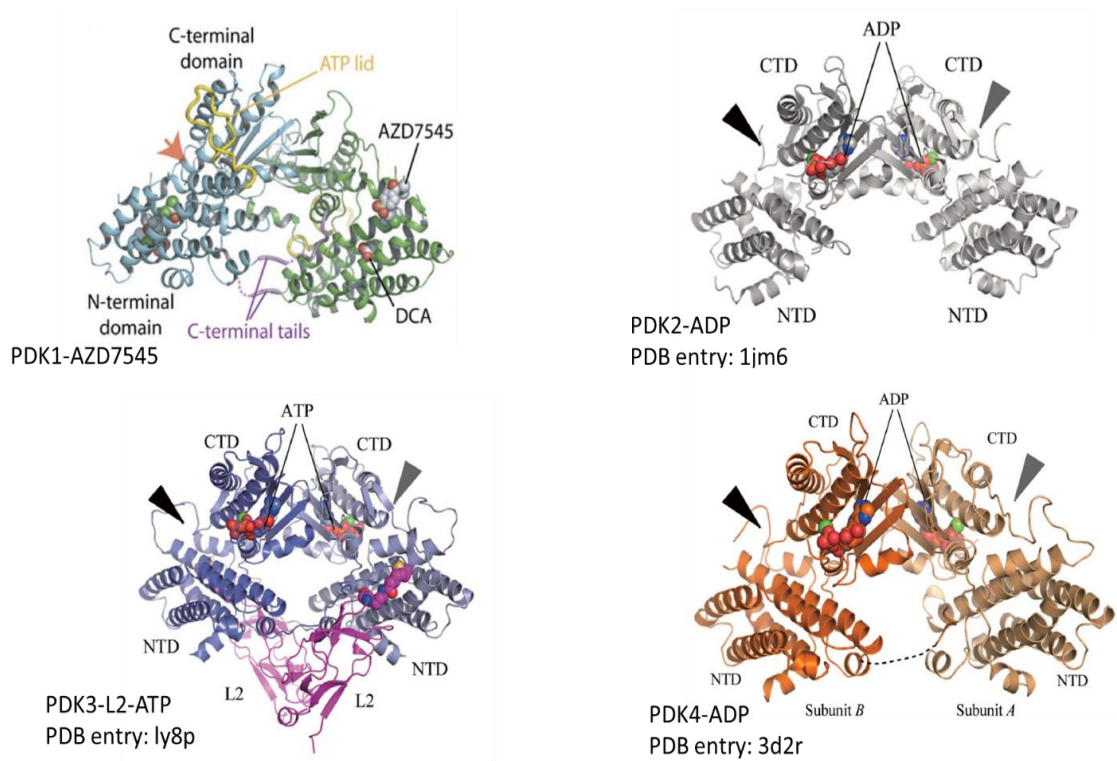


Figure 3.10: Crystal structure of four PDK isozymes [5-9].

PDK1: The sequence coverage for PDK1 is as high as 99.4% (see digestion map in Table 3.6).

The direct interaction between E1p and PDK1 didn't induce global conformational changes in PDK1. However, there are some regions that experienced significant deuterium uptake decrease (peptides spanning residues 66-77, 88-101 and 395-417), and a small increased deuterium uptake in the G1 box region (peptide spanning residues 311-333).

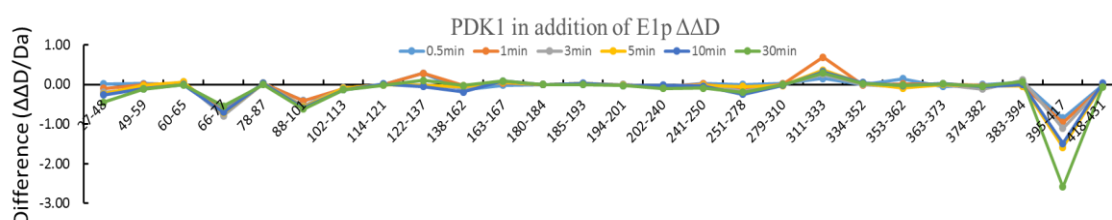


Figure 3.11: Deuteration level change in PDK1 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

It appears that the PDK1 interacts with E1p through its $\alpha 2$ to $\alpha 4$ helices ($\alpha 2$: 59-67; $\alpha 3$: 72-94; $\alpha 4$: 99-102), and at the same time, reorganizes the C-terminal tails. As suggested by the crystal structure of the PDK3-L2 complex, the fully ordered “crosstail” configuration could help freeze the open conformation of PDK3, in which the active cleft will be wider, facilitating the exchange of ATP/ADP [6]. The interaction between PDK1 and E1p may result in the same change by forcing the active cleft to become wider by ordering the C-terminal tails, leading to an increase in deuterium uptake in peptide 311-333, while decreasing the deuterium uptake in peptide 396-417.

PDK2: The sequence coverage of PDK2 is 99.3% (see digestion map in Table 3.7).

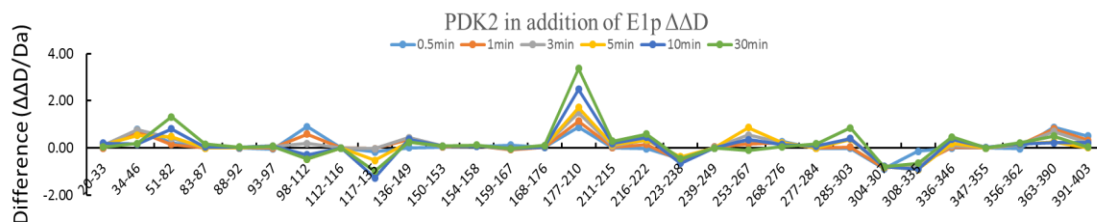


Figure 3.12: Deuteration level change in PDK2 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

The interaction between E1p and PDK2 induces a relatively larger conformational change on PDK2. According to the crystal structure of PDK2, peptides 168-210 together with 117-136 and 308-335 form a cleft between the catalytic domain and the regulatory domain with the ADP/ATP cofactor on one end. This time, it's more likely that PDK2 binds with E1p through this cleft, since the gate helix (residues 117-136), which comprises helices $\alpha 6$ and $\alpha 7$ ($\alpha 6$: residues 106-122; $\alpha 7$: residues 126-141), peptide 177-210, which includes β strand 1 and the beginning of the $\alpha 10$ helix, all experienced a significant conformational change. In addition, the N-terminal region and peptides containing the ATP lid and G-box also experience some deuterium uptake change during complexation.

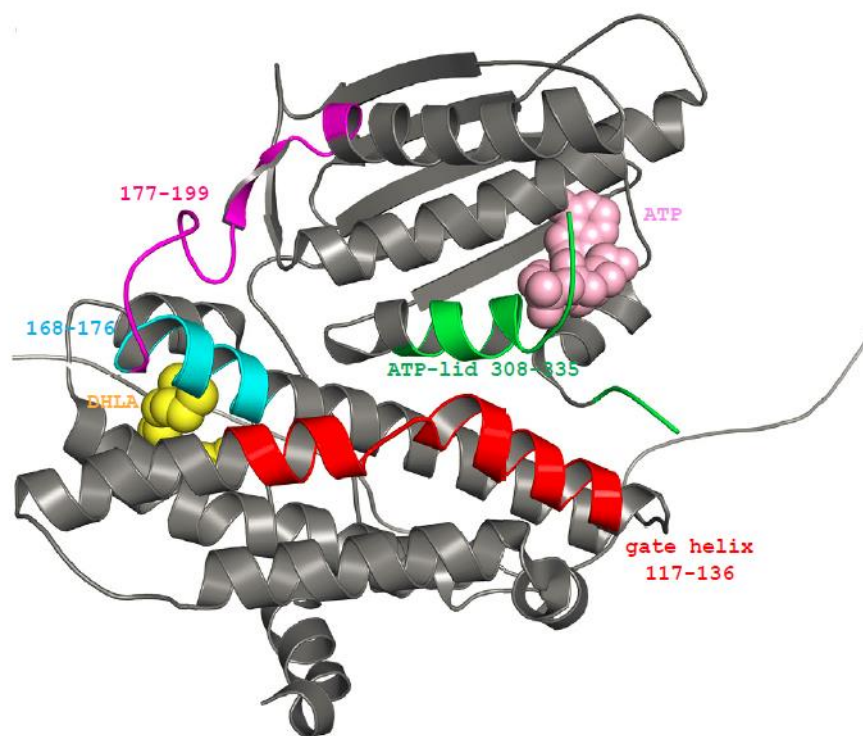


Figure 3.13: Residues of PDK2 displaying significant $\Delta\Delta D$ on binding to E1p in HDX-MS analysis. Includes gate helix (residues 117-136, in *red*), ATP-lid (residues 308-335, in *green*), and another region whose function remains unknown (residues 168-199, in *cyan* and *magenta*).

PDK3: The sequence coverage of PDK3 is 65.5% (see digestion map in Table 3.8).

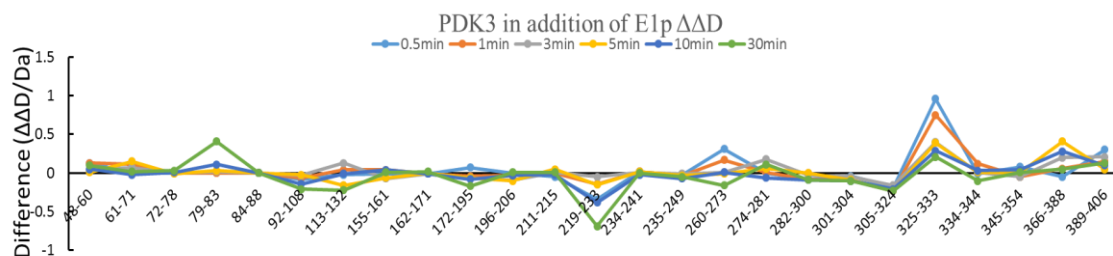


Figure 3.14: Deuteration level change in PDK3 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

The binding mode between PDK3 and E1p is similar to that between PDK2 and E1p. The corresponding region 172-195, gate helix 113-132 together with G2 box 325-333 form a cleft involved in binding to E1p. The interactions between PDK3 and E1p shield the peptide 219-233 from solvent, which includes a β sheet and a hinge. It also induces the conformational change on the entire ATP binding fold. Vassilyev's group proposed a model of the E1p-PDK3 interaction based on computational modelling, which suggested that the peptide bearing phosphorylation sites was docked into the cleft of PDK3. The model also suggested that the peptide underwent a significant conformational change from a loop to a helix during the complexation [8]. However, the HDX-MS data showed that the peptides containing the phosphorylation sites on E1p indeed undergo some deuteration level change upon binding with PDK3, the change is not statistically significant to support the hypothesis that it experienced a conformational change from a loop to a helix.

PDK4: The sequence coverage of PDK4 is 70.6% (see digestion map in Table 3.9).

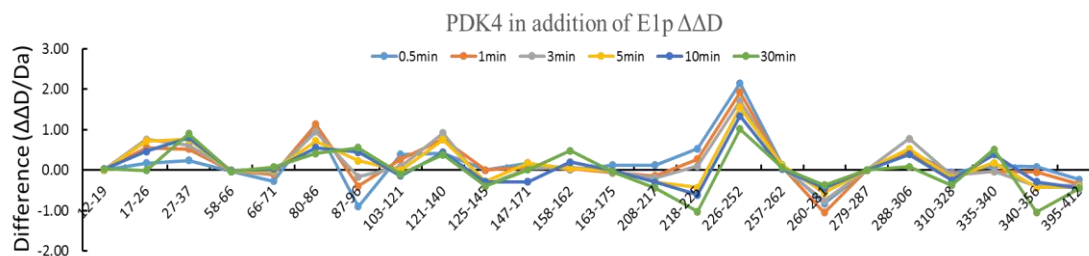


Figure 3.15: Deuteration level change in PDK4 on interaction with E1p. Difference plot showing deuterium incorporation changes of peptic fragments (deuterons exchanged in complexed state minus deuterons exchanged in free state).

Upon binding of E1p to PDK4, many regions on PDK4 were observed to undergo change in deuterium uptake. With the current sequence coverage, we couldn't identify the likely interaction site on PDK4. The peptides 226-252 that experience the most significant increase in deuterium incorporation contain β strand 2, β strand 3 and part of the α 11 helix, which suggest that these regions become more solvent accessible in the PDK4-E1p complex compared with PDK4 alone.

However, all significant decreases in deuterium uptake were only observed at or before the 3 min H/D exchange time points. These weak protections or the weak lock effects on conformations suggested a loose association between E1p and PDK4.

The crystal structure of PDK4 shows that PDK4 adopts a unique conformation which is intrinsically open without L2 binding. The wider cleft created by the open conformation facilitates recruitment of ATP and release of ADP [9]. This might be the reason why many

more peptides from PDK4 display a change in deuterium incorporation during ATP binding, since the structure is more flexible compared to the other three PDK isozymes.

3.3.5 Ligands effects on four isoforms of PDKs

Pyruvate dehydrogenase kinases act to inactivate pyruvate dehydrogenase complex (PDHc) by phosphorylating it using ATP as phosphoryl donor. The highly conserved sequences among the four PDKs contain the phosphoryl transfer catalytic site, which form a unique ATP-binding fold. This fold includes a common structural element known as the “ATP lid,” whose conformational change is coupled to both ATP hydrolysis and protein-protein interactions.

It is already known that higher ratio of ADP/ATP together with high concentration of pyruvate would attenuate the activities of PDKs. Some crystal structures also established DCA (pyruvate analog)/pyruvate binding sites in the N-terminal domain.

Fluorescence quenching experiments done by Roche group together with the mutation of Trp383 on ATP/ADP binding site of PDK2 even confirms that the interaction of pyruvate or DCA would have pronounced effects on the ATP/ADP binding region. [12]

Here, we used the HDX-MS method to study the effects of ATP and pyruvate on PDKs, from which we hope that it could help us understand the mechanism of inactivation of PDKs.

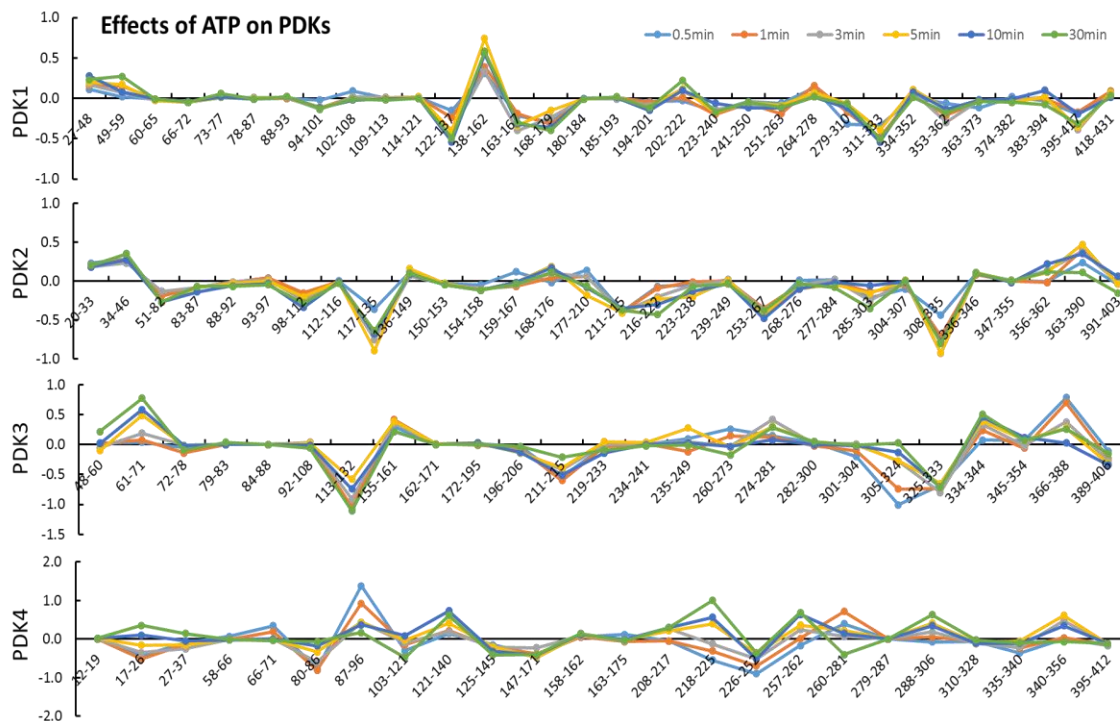


Figure 3.16: Difference plot showing the changes in deuterium incorporation ($\Delta\Delta D$, y axis, deuterons exchanged of corresponding PDKs in the presence of ATP minus deuterons exchanged in the free form of corresponding PDKs).

In the PDKs, the conserved G1-box is involved in binding the adenine moiety of the nucleotide. In PDK1, a 0.3 Da decrease in deuterium uptake was observed in the peptide sequence $^{311}\text{TVKMS}\underline{\text{DRGGGVPLRKIDRLFN}}^{331}$ containing the conserved G1-box motif (DXGXG) upon ATP binding (residues from the G1-box are underlined). A 0.3 Da decrease in ΔD was found in the corresponding $^{285}\text{SIKMS}\underline{\text{DRGGGVPLRKIERL}}^{303}$ peptide in PDK2. Peptide $^{282}\text{SIKISDL}\underline{\text{GGGVPLRKIDRL}}^{300}$ in PDK3 didn't show significant ΔD change, while the peptide in PDK4 $^{288}\text{TIKIS}\underline{\text{DRGGGVPLRKIDRL}}^{306}$ showed a 0.6 increase in ΔD .

The G2-box is responsible for stabilization of phosphate groups of the nucleotide.

Peptides in PDK2 and PDK3 containing the central Gly residues of the G2-box motif (GXGXG) displayed reduced ΔD values of 0.8 Da and 1.0 Da respectively.

In addition to that, gate helix also undergoes some conformational changes. PDK2 (residues 117-135), PDK3 (residues 113-132) experienced a decrease in ΔD , while PDK1(residue 138-162) and PDK4 (121-140) experienced an increase in ΔD .

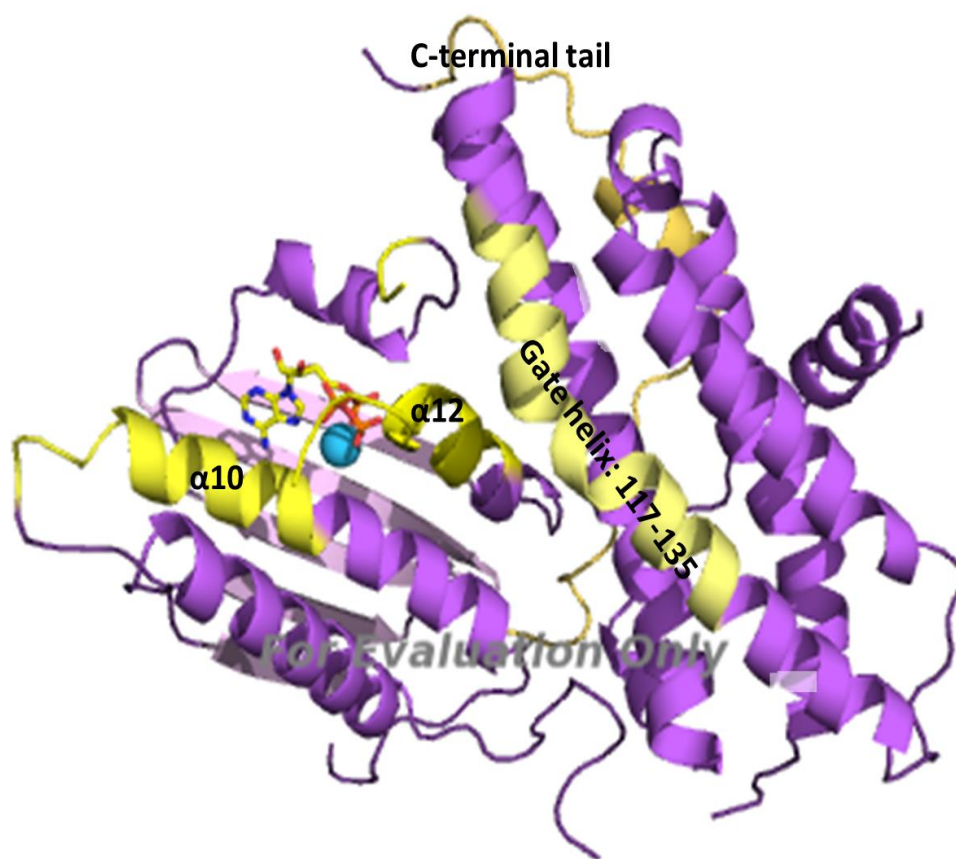


Figure 3.17: Crystal structure of human PDK2 containing ATP. The chain is colored by purple. Residues experienced significant deuteration change is colored by yellow. PDB: 2bu2

The structure of PDK2 shows that the ATP binding pocket is located between the central β -sheet and adjacent layer of helices. Based on the HDX-MS data, the binding of ATP will simultaneously induce the conformational changes on the gate helix and reorder the C-terminal tail.

Additional changes were also observed in other peptides, which are summarized in Table 3.1.

Table 3.1: Deuteration level change of selected peptides on addition of ATP to the four PDK isozymes based on sequence alignment.

	Peptide	Sequences	$\Delta\Delta D$ (Da)
PDK1	88-101	ANIMKEISLLPDNL	-0.11
PDK2	51-82	TFLRQELPVRLANIMKEINLLPDRVLSTPSVQ	-0.3
PDK3	61-71	MREVNLLPDNL	0.7
PDK4	Missing	Missing	Missing
PDK1	122-137	LDFKDKSAEDAKTIYE	-0.5
PDK2	98-112	FLDKDPEDHRTLSTQF	-0.3
PDK3	92-108	LEYENKSPEDPQVLDFN	-0.05
PDK4	103-121	HEKSPEDQKVLSDVFDTLV	-0.4
PDK1	138-162	FTDVTIRIRNRHNDVIPTMAQGVTE	0.55
PDK2	117-135	VTIRNRHNDVVPTMAQGV	-0.7
PDK3	113-132	IKVRNRHNDVVPTMAQGVIE	-1.1
PDK4	121-140	VKVRNRHNVVPTMAQGILE	0.6
PDK1	279-310	FKNAMRATMEHHADKGVYPPIQVHVTLGEDL	-0.2
PDK2	253-267	FKNAMRATVESHESS	-0.4
PDK3	Missing	Missing	Missing
PDK4	260-281	MRATVEHQENRPFLTPVEATVV	0.7
PDK1	311-333	TVKMSDRGGGVPLRKIDRLFNYM	-0.5
PDK2	285-303	SIKMSDRGGGVPLRKIERL	-0.3
PDK3	282-300	SIKISDLGGGVPLRKIDRL	0.05
PDK4	288-306	TIKISDRGGGVPLRKIDRL	0.6
PDK1	334-352	YSTAPRPRVETSRVPLAG	0.01
PDK2	308-335	YSTAPTQPGTGGTPLAGFGYGLPISRL	-0.8
PDK3	305-324	YSTAPRPSLEPTRAAPLAGF	-1.0
PDK4	310-328	TYSTAPTVMDSNRNAPLA	-0.05
PDK1	395-417	IERLPVYNKAAWKHYRTNHEADD	-0.3
PDK2	363-390	LSTDSEVERLPVYNKSAWRHYQTIQEAGD	0.4
PDK3	366-388	FERLPVFNKSAWRHYKTTPEADD	0.8
PDK4	Missing	Missing	Missing

From Table 3.1, it can be seen that even peptides that have the same functions among PDKs, on addition of the same ligand, such as ATP, will still have different effects on them.

According to the HDX-MS data, binding of pyruvate to PDK1 will induce a large conformational change especially on the N-terminus and nucleotide-binding site located at the C-terminus of PDK1. The peptides that show significant deuteration level increase are listed below.

Table 3.2: Peptides that show significant deuteration level change in PDK1 on addition of pyruvate.

Peptide	Sequence	$\Delta\Delta D$ (Da)
27-48	ASDSASGSGPASESGVPGQVDF	1.2
49-59	YARFSPSPLSM	2.2
122-137	LDFKDKSAEDAKTIYE	0.6
163-167	YKESF	0.7
251-278	EELNAKSPGQPIQVVVV	0.6
334-352	YSTAPRPRVETSRAVPLAG	0.7
363-373	YAQYFQGDLKL	0.8
395-417	IERLPVYNKAAWKHYRTNHEADD	1.4

There were no significant negative $\Delta\Delta D$ changes observed on peptides that are related to interaction with pyruvate. The reason could be that the DCA-binding site is located inside the helix bundle in the N-terminal domain of PDK1 and is completely shielded from solvent in the apo-PDK1 structure. As a result, the pyruvate entered into a pocket that originally has low solvent accessibility, in other words, it will not induce much protective effect during hydrogen deuterium exchange.

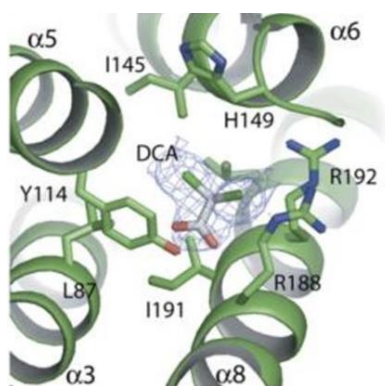


Figure 3.18: Crystal structure of PDK2 in the presence of DCA. [6]

The effects of pyruvate on PDK1 are the opposite from those exerted by ATP. The interaction of ATP with PDK1 will shield most of the peptides on the C-terminal part, which is responsible for ATP binding; in contrast the interaction between pyruvate and PDK1 will convert those PDK1 regions to a more solvent accessible conformation. This result is in consistent with what Roche and coworkers had concluded from fluorescence quenching experiments. Addition of pyruvate into buffer containing PDK2 could quench the fluorescence of Trp383 (the equivalent Trp415 in PDK1), for which the intensity would be lower if the microenvironment surrounding the tryptophan is more hydrophilic.

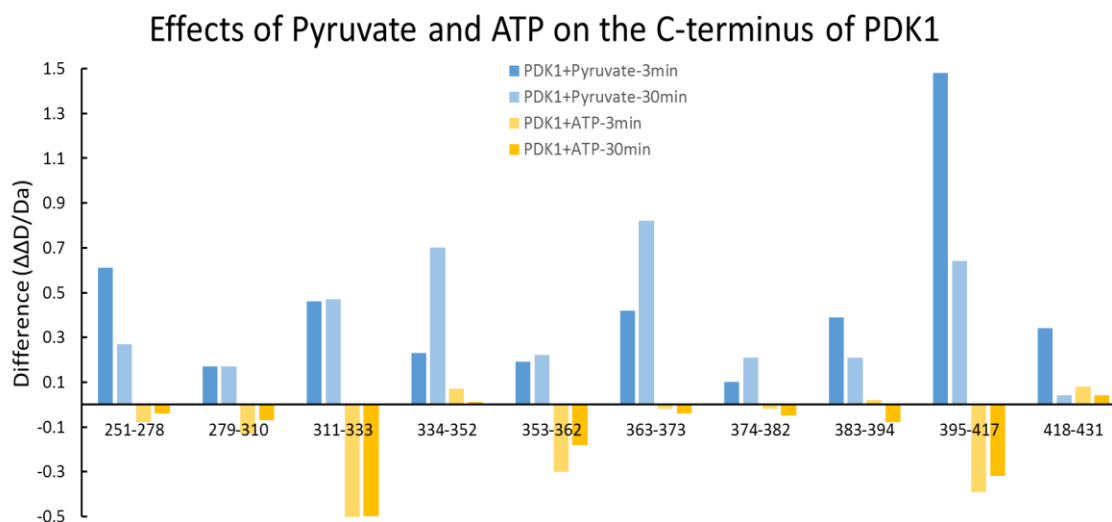











Figure 3.19: Peptides that show significant deuteration change on PDK1 when incubated with pyruvate and ATP.

A similar observation could be made from the effect of pyruvate binding to PDK3. There is no peptide displaying significant decrease in deuterium incorporation when pyruvate interacts with PDK3. Based on the similarities among PDKs, the reason might be the same: the pyruvate is entirely enveloped by a pocket that originally has low solvent accessibility.

Table 3.3: Peptides that show a significant deuteration level change in PDK3 on addition of pyruvate.

Peptide	Sequence	$\Delta\Delta D$ (Da)	
113-132	IKVRNRHNDVVPTMAQGVIE		0.4
235-249	VYVPSHLFHMLFELF		0.35
274-281	VTLGKEDL		0.4
282-300	SIKISDLGGGVPLRKIDRL		0.7
301-304	FN YM		-0.4
305-324	YSTAPRPSLEPTRAAPLAGF		0.5
325-333	GYGLPISRL		-0.5
366-388	FERLPVFNKSAWRHYKTTPEADI		0.8
389-406	WSNPSSSEPRDASKYKAKQ		0.7

When comparing the effects of pyruvate and ATP on PDK3, it was apparent that most of the peptides that show significant deuteration change upon binding to pyruvate also display change when interacting with ATP, though some showed the same pattern, others the opposite. The data of PDK1 interacts with pyruvate and that of with ATP suggests that in the presence of pyruvate, PDK1 might adopt a conformation that is totally unsuitable for ATP binding, as the peptides that experiencing exposure to the solvent in the presence of pyruvate would be shielded from solvent if ATP presents instead. However, the interaction between PDK3 and pyruvate or ATP doesn't exhibit that much differences. These differences may be able to explain why PDK1 is more sensitive to pyruvate/DCA inhibition, while PDK3 is nearly insensitive. [13]

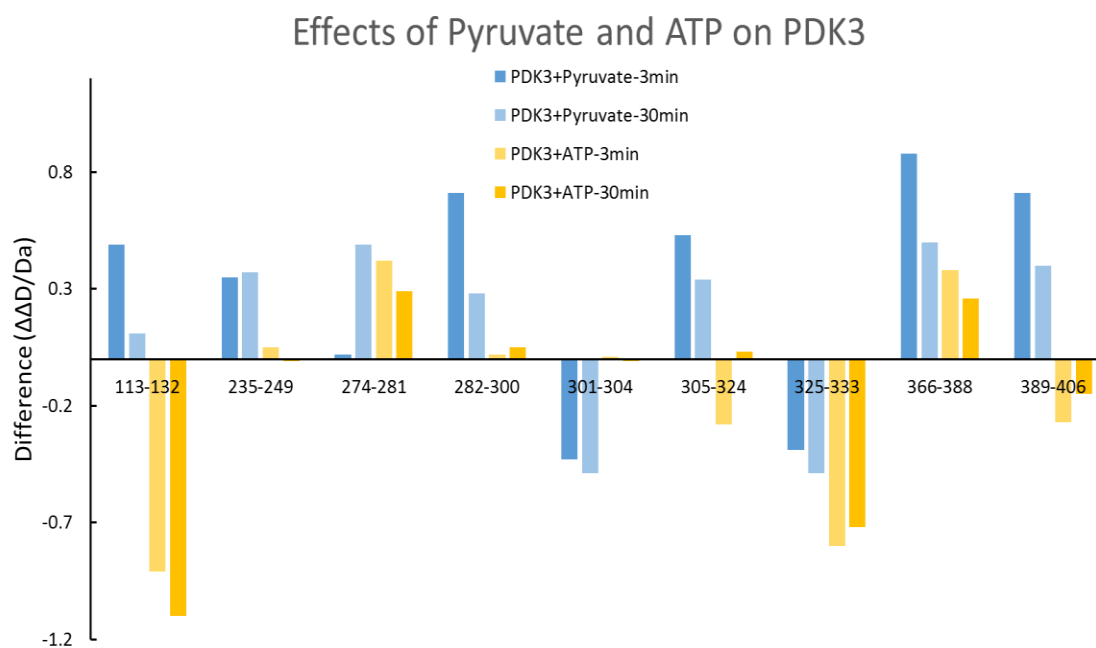


Figure 3.20: Peptic peptides from PDK3 that show significant deuteration change on incubation with pyruvate and ATP.

Both interaction patterns between PDKs and pyruvate show that pyruvate binds to the allosteric site located on the N-terminal domain, which will induce a large conformational change on the ATP binding site.

Table 3.4: Peptides identified from pepsin digestion of the E1p -subunit.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	2-7	ANDATF	638.2790	638.2780	1.7
2	8-14	EIKKCDL	848.4562	848.4546	1.9
3	15-27	HRLEEGPPVTTVL	1447.7924	1447.7904	1.4
4	28-35	TREDGLKY	981.5001	981.5000	0.1
5	42-47	VRRMEL	803.4552	803.4556	-0.3
6	53-64	YKQKIIRGFCHL	1505.8412	1505.8410	0.1
7	73-88	VGLEAGINPTDHLITA	1620.8606	1620.8592	0.4
8	127-149	YAKNFYGGNGIVGAQVPLGAGIA	2237.1720	2237.1713	0.3
9	151-161	ACKYNGKDEVK	1567.7760	1567.7751	0.6
10	162-177	LTLYGDGAANQGQIFE	1696.8190	1696.8177	0.8
11	179-184	YNMAAL	682.3236	682.3229	0.9
12	185-191	WKLPCIF	906.4919	906.4906	1.5
13	192-205	ICENNRYGMGTSVE	1572.6792	1572.6781	0.9
14	206-213	RAAASTDY	854.4008	854.4003	0.2
15	214-229	YKRGFIPGLRVDGMD	2001.9851	2001.9854	0.1
16	243-252	YCRSGKGPIIL	1093.5823	1093.5834	1.0
17	253-257	MELQT	621.2921	621.2912	1.4
18	258-271	YRYHGHSMSPGVS	1592.6911	1592.6927	1.0
19	277-290	EIQEVRSKSDPIML	1644.8633	1644.8625	0.5
20	291-300	LKDRMVNSNL	1189.6376	1189.6358	1.5
21	301-322	ASVEELKEIDVEVRKEIEDAAQ	2500.2790	2500.2777	0.5
22	323-334	FATADPEPPLEEL	1428.6907	1428.6893	1.0
23	335-353	LGYHIYSSDPPFEVRGANQ	2150.0322	2150.0301	1.0
24	354-361	WIKFKSVS	994.5737	994.5720	1.7

Table 3.5: Peptides identified from pepsin digestion of the E1p β -subunit.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Theoretical	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	3-14	VTVRDAINQGMD	1318.6436	1318.6420	1.2
2	17-25	LERDEKVFL	1148.6325	1148.6310	1.3
3	30-59	VAQYDGAYKVSRLWKKYGDKRIIDTPISE	3456.8296	3456.8274	0.6
4	60-64	MGFAG	482.2072	482.2068	0.9
5	65-79	IAVGAAMAGLRPICE	1471.7760	1471.7775	1.0
6	81-85	MTFNF	659.2866	659.2858	1.1
7	101-114	YYMSGGLQPVPVIF	1570.7985	1570.7974	0.9
8	115-131	RGPNGASAGVAAQHSQC	1610.7465	1610.7452	0.8
9	132-170	FAAWYGHCPGLKVVSPWNSEDAKGLIKSAIRDNNPVVVL	4251.2190	4251.2172	0.4
10	171-175	ENELM	635.2708	635.2705	0.2
11	176-190	YGVPFEPPEAQSKD	1710.8019	1710.8010	0.5
12	191-222	FLPIGKAKIERQGTHITVVSHSRPVGHCLEA	3493.9213	3493.9256	1.2
13	233-238	CEVINM	708.3058	708.3055	0.1
14	239-248	RTIRPMDMET	1249.6049	1249.6028	1.7
15	249-272	IEASVMKTNHLVTVEGGWPQFGVG	2314.1672	2314.1649	1.0
16	275-289	ICARIMEGPAFNFLD	1696.8197	1696.8186	0.6
17	286-321	NFLDAPAVRVTGADVPMKYKILEDNSIPQVKDIIF	3957.0872	3957.0830	1.1
18	322-329	AIKKTLNI	900.5892	900.5877	1.7

Table 3.6: Peptides identified from pepsin digestion of PDK1^{rat}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Error (ppm)	
No.	Position	Sequence	Experimental	Theoretical	
1	27-48	ASDSAGSGPASESGVPGQVDF	2008.8730	2008.8742	0.6
2	49-59	YARFSPSPLSM	1255.6140	1255.6157	1.4
3	60-65	MKQFLDF	797.4192	797.4187	-0.6
4	66-77	GSVNACEKTSFM	1273.5551	1273.5573	1.7
5	78-87	FLRQELPVRL	1270.7630	1270.7624	-0.5
6	88-101	ANIMKEISLLPDNL	1570.8509	1570.8518	0.6
7	102-113	LRTPSVQLVQSW	1413.7849	1413.7847	-0.1
8	114-121	YIQLQEL	993.5251	993.5258	0.7
9	122-137	LDFKDKSAEDAKTIYE	1872.9226	1872.9260	1.8
10	138-162	FTDTVIRNRHNDVIPTMAQGVTE	2883.4894	2883.4901	0.2
11	163-167	YKESF	673.3192	673.3188	-0.6
12	168-179	GVDPVTSQNVQY	1306.6274	1306.6266	-0.9
13	180-184	FLDRF	697.3668	697.3674	0.9
14	185-193	YMSRISIRM	1156.5965	1156.5972	0.6
15	194-201	LLNQHSLL	937.5465	937.5478	1.4
16	202-240	VVEVIKDGyenARRLCDLFGGKGSPSHRKHIGSINPNCd	4281.1404	4281.1432	0.6
17	241-250	YYVNSPELEL	1226.5939	1226.5941	0.2
18	251-278	EELNAKSPGQPIQVVYVPSHLYHMFEL	3224.6449	3224.6452	0.1
19	279-310	FKNAMRATMEHHADKGVYPPIQVHVTLGEEDL	3633.7941	3633.7943	0.1
20	311-333	TVKMSDRGGGVPLRKIDRLFNYM	2653.4065	2653.4057	-0.3
21	334-352	YSTAPRPRVETSRAVPLAG	2175.1669	2175.1672	0.1
22	353-362	FGYGLPISRL	1122.6306	1122.6322	1.4
23	363-373	YAQYFQGDLKL	1345.6787	1345.6803	1.2
24	374-382	YSLEGYGTd	1004.4207	1004.4209	0.2
25	383-394	AVIYIKALSTES	1294.7253	1294.7274	1.6
26	395-417	IERLPVYNKAawkHYRTNHEADd	2826.4070	2826.4107	1.3
27	418-431	WCVPSREPkdMTTF	1549.7138	1549.7142	0.2

Table 3.7: Peptides identified from pepsin digestion of PDK2^{rat}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Theoretical	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	20-33	IEHFSKFSPSPLSM	1606.7952	1606.7934	1.1
2	34-46	KQFLDFGSSNACE	1445.6369	1445.6366	0.2
3	51-82	TFLRQELPVRLANIMKEINLLPDRVLSTPSVQ	3691.0751	3691.0727	0.6
4	83-87	LVQSW	632.3411	632.3402	1.4
5	88-92	YVQSL	609.3247	609.3243	0.7
6	93-97	LDIME	620.2968	620.2960	1.3
7	98-112	FLDKDPEDHRTLSQF	1847.8927	1847.8923	0.2
8	112-116	FTDAL	566.2831	566.2821	1.8
9	117-135	VTIRNRHNDVVPTMAQGVL	2120.1401	2120.1393	0.4
10	136-149	EYKDTYGDDPVSQ	1630.6875	1630.6867	0.5
11	150-153	NIQY	537.2668	537.2667	0.1
12	154-158	FLDRF	697.3682	697.3668	2.0
13	159-167	YLSRISIRM	1138.6411	1138.6401	0.9
14	168-176	LINQHTLIF	1098.6310	1098.6306	0.4
15	177-210	DGSTNPAHPKHIGSIDPNCVSDVVKDAYDMAKL	3581.7003	3581.6999	0.1
16	211-215	LCDKY	641.2966	641.2963	0.5
17	216-222	YMASPDL	796.3551	796.3546	0.6
18	223-238	EIQEVNATNATQPIHM	1795.8662	1795.8643	1.1
19	239-249	VYVPSHLYHML	1358.6928	1358.6926	0.1
20	253-267	FKNAMRATVESHESS	1693.7977	1693.7962	0.9
21	268-276	LTLPIKIM	1025.6433	1025.6428	0.5
22	277-284	VALGEEDL	845.4270	845.4251	2.2
23	285-303	SIKMSDRGGGVPLRKIERL	2112.2077	2112.2070	0.3
24	304-307	FSYM	547.2222	547.2221	0.1
25	308-335	YSTAPTPQPGTGGTPLAGFGYGLPISRL	2776.4321	2776.4305	0.6
26	336-346	YAKYFQGDQL	1345.6791	1345.6787	0.3
27	347-355	FSMEGFQTD	990.3881	990.3873	0.8
28	356-362	AVIYLKA	777.4881	777.4869	1.5

29	363-390	LSTDSVERLPVYNKSAWRHYQTIQEAGD	3263.6087	3263.6080	0.2
30	391-403	WCVPSTEPKNTST	1449.6684	1449.6679	0.3

Table 3.8: Peptides identified from pepsin digestion of PDK3^{human}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Theoretical	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	48-60	FLRKELPVRLANT	1556.9301	1556.9271	1.9
2	61-71	MREVNLLPDNL	1313.6902	1313.6882	1.4
3	72-78	LNRPSVG	742.4213	742.4206	0.9
4	79-83	LVQSW	632.3416	632.3402	2.1
5	84-88	YMQSF	675.2818	675.2807	1.7
6	92-108	LEYENKSPEDPQVLDF	2036.9476	2036.9447	1.4
7	113-132	IKVRNRHNDVVPTMAQGVIE	2276.2322	2276.2292	1.3
8	155-161	YTNRISF	900.4587	900.4574	1.4
9	162-171	RMLINQHTLL	1238.7053	1238.7038	1.2
10	172-195	FGGDTNPVHPKHIGSIDPTCNVAD	2491.1692	2491.1670	0.9
11	196-206	VVKDAYETAKM	1254.6419	1254.6399	1.6
12	211-215	YYLVA	628.3353	628.3341	2.0
13	219-233	EVEEFNAKAPDKPIQ	1714.8671	1714.8646	1.5
14	234-241	VVYVPSHL	913.5154	913.5142	1.4
15	235-249	VYVPSHLFHMLFELF	1878.9666	1878.9662	0.1
16	260-273	YEDRKEGYPAVKTL	1668.8609	1668.8592	1.0
17	274-281	VTLGKEDL	874.4888	874.4880	0.9
18	282-300	SIKISDLGGGVPLRKIDRL	2037.2212	2037.2179	1.6
19	301-304	FNYM	574.2340	574.2330	1.7
20	305-324	YSTAPRPSLEPTRAAPLAGF	3058.6530	3058.6473	1.9
21	325-333	GYGLPISRL	975.5639	975.5622	1.7
22	334-344	YARYFQGDLKL	1373.7240	1373.7212	2.0
23	345-354	YSMEGVGTDA	1029.4203	1029.4194	0.9
24	366-388	FERLPVFNKSAWRHYKTTPEADD	2807.3937	2807.3900	1.3
25	389-406	WSNPSSEPRDASKYKAKQ	2079.0275	2079.0254	1.0

Table 3.9: Peptides identified from pepsin digestion of PDK4^{rat}.

Peptides			Monoisotopic [M+H] ⁺ (Da)	Mass Theoretical	Error (ppm)
No.	Position	Sequence	Experimental	Theoretical	
1	12-19	SSLGNAGL	718.3738	718.3730	1.1
2	17-26	AGLVPREVEL	1082.6221	1082.6204	1.6
3	27-37	FSRYSPSPLSM	1271.6092	1271.6089	0.2
4	58-66	RQELPVRLA	1081.6492	1081.6477	1.4
5	66-71	ANILKE	687.4042	687.4036	0.9
6	80-86	VNTPSVQ	744.3891	744.3886	0.7
7	87-96	LVKSWYIQL	1236.6997	1236.6987	0.8
8	103-121	HEKSPEDQKVLSDFVDTLV	2186.0991	2186.0976	0.7
9	121-140	VKVRNRHHNVPTMAQGILE	2298.2627	2298.2612	1.8
10	125-145	TMAQGVIEYKEKFGDFPFIST	2408.1855	2408.1843	0.6
11	147-171	VDPVTNQNLQYFLDRFYMNRISTRM	3121.5377	3121.5346	1.0
12	158-162	FLDRF	697.3674	697.3668	0.9
13	163-175	YMNRISTRMLMNQ	1657.7990	1657.7971	0.9
14	208-217	FECAKMLCDQ	1187.4901	1187.4894	1.6
15	218-225	YYLTSPEL	985.4882	985.4877	1.2
16	226-252	KLTQVNGKFPGQPIHIVYVPSHLHHML	3090.6841	3090.6822	0.6
17	257-262	KNAMRA	690.3722	690.3716	0.9
18	260-281	MRATVEHQENRPFLTPVEATVV	2524.2994	2524.2977	0.7
19	279-287	TVVLGKEDL	973.5568	973.5564	0.4
20	288-306	TIKISDRGGGVPLRITDRL	2067.2041	2067.2033	0.4
21	310-328	TYSTAPTPVMDNSRNAPLA	2005.9661	2005.9648	0.6
22	335-340	PISRLY	748.4355	748.4352	0.4
23	340-356	YAKYFQGDLNLYMSGY	2019.9162	2019.9157	0.2
24	395-412	WCIPSKPKNLSKEKLAV	2070.1441	2070.1416	1.2

References

- [1] Roche, T. E., Baker, J. C., Yan, X., Hiromasa, Y., Gong, X., Peng, T., Dong, J., Turkan, A., and Kasten, S. A. (2001) Distinct regulatory properties of pyruvate dehydrogenase kinase and phosphatase isoforms. *Prog. Nucleic Acid Res. Mol. Biol.* **70**, 33-75.
- [2] Steussy C N, Popov K M, Bowker-Kinley M M, Sloan R B, Harris R A, Hamilton J A. Structure of Pyruvate Dehydrogenase Kinase novel folding pattern for a serine protein [J]. *Journal of Biological Chemistry*, 2001, 276(40): 37443-37450.
- [3] Wynn R M, Kato M, Chuang J L, Tso S C, Li J, Chuang D T. Pyruvate dehydrogenase kinase-4 structures reveal a metastable open conformation fostering robust core-free basal activity[J]. *Journal of Biological Chemistry*, 2008, 283(37): 25305-25315.
- [4] Ubersax, J.A. and J.E. Ferrell Jr, Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol*, 2007. 8(7): p. 530-541.
- [5] Kato M, Chuang J L, Tso S C, Wynn R M, Chuang D T. Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex[J]. *The EMBO journal*, 2005, 24(10): 1763-1774.
- [6] Kato M, Li J, Chuang J L, Chuang D T, Distinct structural mechanisms for inhibition of pyruvate dehydrogenase kinase isoforms by AZD7545, dichloroacetate, and radicicol. *Structure*, 2007. 15(8): p. 992-1004.
- [7] Knoechel T R, Tucker A D, Robinson C M, Phillips C, Taylor W, Bungay P J, Kasten S A, Roche T E, Brown D G, Regulatory Roles of the N-Terminal Domain Based on Crystal Structures of Human Pyruvate Dehydrogenase Kinase 2 Containing Physiological and Synthetic Ligands. *Biochemistry*, 2005. 45(2): p. 402-415

- [8] Devedjiev, Y., C.N. Steussy, and D.G. Vassilyev, Crystal structure of an asymmetric complex of pyruvate dehydrogenase kinase 3 with lipoyl domain 2 and its biological implications. *J Mol Biol*, 2007. 370(3): p. 407-16.
- [9] Kukimoto-Niino M, Tokmakov A, Terada T, Ohbayashi N, Fujimoto T, Gomi S, Shiromizu I, Kawamoto M, Matsusue T, Shirouzu M, Yokoyama S, Inhibitor-bound structures of human pyruvate dehydrogenase kinase 4. *Acta Crystallogr D Biol Crystallogr*, 2011. 67(Pt 9): p. 763-73.
- [10] Patel M S, Roche T E. Molecular biology and biochemistry of pyruvate dehydrogenase complexes[J]. *The FASEB Journal*, 1990, 4(14): 3224-3233.
- [11] Guevara E L, Yang L, Birkaya B, Zhou J, Nemeria N S, Patel M S, Jordan F. Global view of cognate kinase activation by the human pyruvate dehydrogenase complex[J]. *Scientific Reports*, 2017, 7.
- [12] Hiromasa Y, Hu L, Roche T E. Ligand-induced effects on pyruvate dehydrogenase kinase isoform 2[J]. *Journal of Biological Chemistry*, 2006, 281(18): 12568-12579.
- [13] Roche T E, Hiromasa Y. Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer[J]. *Cellular and molecular life sciences*, 2007, 64(7): 830-849.

CHAPTER 4. Study of protein-protein interactions within the human 2-oxoglutarate dehydrogenase complex using site-specifically introduced external fluorophores

Abbreviation used in this chapter: the postscript *o* is for oxoglutarate, while *p* is for pyruvate.

4.1 Introduction

Site-specific labeling provides biochemists a handle for target detection, quantification and analysis, which is essential for biological research, diagnostics, and therapeutics. Many labeling methods have been developed, including covalent modification of amino acid side chains, covalent labeling by enzymes, conjugation to unnatural amino acids, and affinity labeling [1-6].

One of the most common specific covalent coupling methods is the conjugation of a moiety bearing a thiol-reactive group to a solvent-accessible cysteine. This method affords the site-specific labeling of a protein at a unique engineered (or native) surface cysteine. This cysteine could be specifically labeled with dye derivatives of haloacetyl compounds, disulfide, or maleimides. The maleimide is by far the most prevalent functional group to be coupled to a cysteine because the coupling reaction is highly specific and efficient [7].

The protein must be maintained in a reduced form prior to the maleimide-cysteine coupling reaction to prevent the formation of disulfide bridges and inactivation of the cysteines. Reducing agents such as dithiothreitol (DTT) and tris[2-

carboxyethyl]phosphine (TCEP) can be used for this purpose. These reducing agents must be removed before conjugation so that their thiol groups will not compete with target thiols in proteins. Also, the labeling reaction must be performed shortly after the elimination of reducing agents to prevent reoxidation of the target thiols, resulting in decreased labeling efficiency [7].

Fluorescence spectral properties can provide a variety of molecular details of proteins, such as the solvent exposure of residues, the existence of protein conformers, the rate of rotational diffusion of a protein, and the distance between sites on a protein. This method is so sensitive that only a few micromoles or even nanomoles of sample is needed during the measurements. In addition to that, the fluorescence spectrum, and the phenomenon it reports on, is also responsive to the changes in solvent polarity, pH, temperature or the addition of quenching agents, which make it possible to study more complicated cases [8-12].

2-Oxoglutarate dehydrogenase complex (OGDHc) belongs to the superfamily of 2-oxo acid dehydrogenase multienzyme complexes. However, the E2o component from OGDHc differs from that from PDHc significantly as the sequence alignment of the E2o could not identify any apparent peripheral subunit-binding domain (S) [13-14].

In addition, there is no extra-component such as E3BP which could provide PSBD in this complex. The structure of E2o is composed of a lipoyl domain (LDo), a linker and a catalytic domain (Core). Hence, as to how this complex is assembled, and whether indeed the mammalian OGDHc has a new unprecedented mode of assembly, are still open questions.

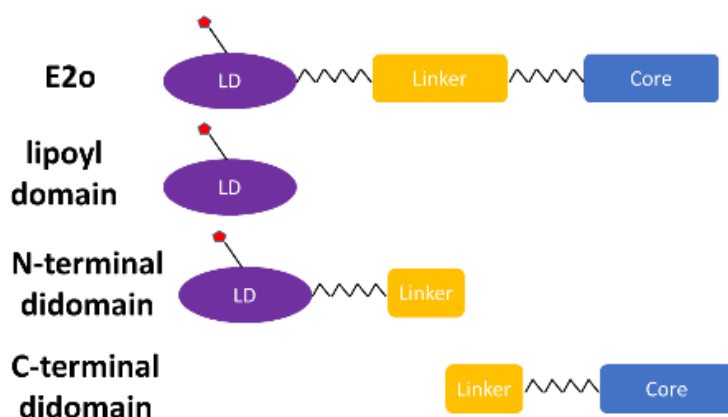


Figure 4.1: Domain structure of E2o and E2o constructs. E2o¹⁻¹⁷³ (N-terminal didomain) and E2o¹⁴⁴⁻³⁸⁶ (C-terminal didomain) have 30 amino acid overlapping sequence.

4.2 Materials and Methods

N-(1-Pyrene)maleimide was purchased from Ana Spec; dansyl chloride and p-arsanilic acid were from Sigma Aldrich. All the chemicals were used without further purification.

4.2.1 Labeling of lipamide in the LDo with a dansyl group for fluorescence studies.

The lipoylated LDo (50 μ M) was incubated with 100 μ M TCEP in sample buffer (30 mM KH₂PO₄, 0.15 M NaCl, pH 7.5) at room temperature for 5 min, resulting in fully reduced lipoyl domains, which were then reacted with 150 μ M DANS-As at room temperature for 2 h, forming chemically modified lipoyl domains. The reaction mixture was dialyzed against the sample buffer with a centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the surplus TCEP and DANS-As. The product was diluted to 1 μ M with the ESI buffer (a 50:50:0.1 v/v/v mixture of methanol, water and formic acid), and examined by

FT-ESI-MS in the positive ion mode. The procedure has been published recently by this group [15].

4.2.2 Labeling $E2o^{1-173}$ and $E2o^{144-386}$ with *N*-(1-pyrene)maleimide for fluorescence

studies. The $E2o^{1-173}$ or $E2o^{144-386}$ (50 μ M), each containing a single cysteine (*for E2o¹⁻¹⁷³ it's Cys38; for E2o¹⁴⁴⁻³⁸⁶ it's Cys179*), was incubated with 100 μ M TCEP in sample buffer (30 mM KH_2PO_4 , 0.15 M NaCl, pH=7.5) at room temperature for 5 min, then reacted with 100 μ M *N*-(1-pyrene)maleimide at room temperature for 2 h, forming a protein bearing a single pyrene fluorophore. The reaction mixture was dialyzed against sample buffer with a centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the surplus TCEP and *N*-(1-pyrene)maleimide.

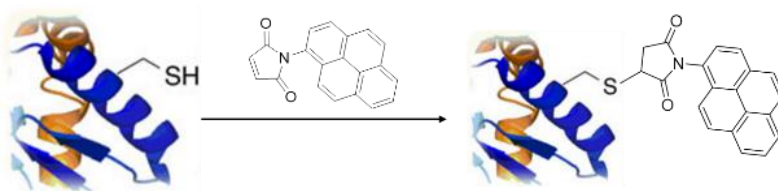


Figure 4.2: Labeling of $E2o^{1-173}$ and $E2o^{144-386}$ with *N*-(1-pyrene)maleimide.

4.2.3 Fluorescence spectroscopy titration studies. All fluorescence spectra were recorded on a Varian Cary Eclipse fluorescence spectrophotometer. The titration was conducted in 30 mM KH_2PO_4 (pH 7.5). A stock solution of chemically modified lipoyl domains was added to 30 mM KH_2PO_4 (pH 7.5) to reach a final concentration of 1-2 μ M. The excitation wavelength for the DANS-As modified proteins was 338 nm, and the emission spectra were recorded from 400-650 nm. The maximum emission intensities at 520 nm were used for calculating binding parameters.

For N-(1-pyrene)maleimide labeled proteins (E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶), the excitation wavelength was 340 nm, and recorded emission spectra from 350-600 nm. The maximum emission intensities at 374 nm were used to calculate binding parameters.

All data sets are normalized ($\Delta F/\Delta F_{\max}$) to their maximum fluorescence change.

Equilibrium dissociation constants (K_d) were obtained by using the Hill equation to fit data.

$$\Delta F / \Delta F_{\max} = [E1o]^n / (K_A^n + [E1o]^n) \quad \text{Equation 4.1}$$

4.2.4 Measurement of E1o-specific activity. The E1o-specific activity was measured with the external oxidizing agent 2,6-dichlorophenolindophenol (DCPIP) [18] at 600 nm in the following reaction medium contained in 1 ml: 50 mM KH₂PO₄ (pH 7.0), 0.50 mM ThDP, 1.0 mM MgCl₂, 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 ($\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg E1o}^{-1}$).

4.2.5 Fluorescence study of the binding of thiochrome diphosphate to E1o. Thiochrome diphosphate (TCDP) is a fluorescent analogue of ThDP and can act as a competitive inhibitor for ThDP dependent enzymes. It is known that the intrinsic protein fluorescence due to tryptophan can be quenched by the binding of ThDP and ThDP analogues. The fluorescence quenching resulting from addition of TCDP exhibited saturation behavior, which can be fitted to Equation xx to estimate the K_d value.

$$\Delta F / \Delta F_{max} = [TCDP]^n / (S_{0.5}^n + [TCDP]^n) \quad \text{Equation 4.2}$$

where $\Delta F / \Delta F_{max}$ is a relative fluorescence; $\Delta F = F_{max} - F_i$, where F_{max} is a maximum fluorescence intensity reached on titration by TCDP and F_i is a fluorescence intensity at a given concentration of TCDP; $\Delta F_{max} = F_{max} - F_o$, where F_o is the initial fluorescence before addition of TCDP; $S_{0.5}$, is the concentration of ligand at half saturation; n is the Hill coefficient. For $n=1$, the value of $S_{0.5}$ is equal to K_d .

4.2.6 FRET experiments. The energy transfer from E1o-TCDP to E3 were conducted on a Varian Cary Eclipse fluorescence spectrophotometer. The efficiency of energy transfer was obtained from the excitation spectrum of the energy acceptor—TCDP with the concentrations of E1o and E3 both at 5 μ M concentration in 50 mM KH_2PO_4 (pH=7.0). Also, the spectrum of E1o bound thiochrome and hE3 bound FAD at the same concentration in the same buffer were recorded separately.

4.2.7 Labeling $E2o^{1-173}$, $E2o^{144-386}$ with Alexa Fluor® 350 C5 Maleimide dye for FRET studies.

The truncated E2o proteins $E2o^{1-173}$ and $E2o^{144-386}$ (50 μ M; each contain only one cysteine), were incubated with 100 μ M TCEP in sample buffer (30 mM KH_2PO_4 , 0.15 M NaCl, pH=7.5) at room temperature for 5 min, then reacted with 100 μ M Alexa Fluor® 350 C5 Maleimide dye at room temperature for 2 h, forming a protein bearing a single Alexa350 fluorophore. The reaction mixture was dialyzed against sample buffer with a

centrifugal filter unit (Vivaspin 500, 10K MWCO) to remove the surplus TCEP and Alexa Fluor® 350 C5 Maleimide dye.

The fluorescence measurements were the same as described above. 5 μM of E2o¹⁻¹⁷³-Alexa350 (or of E2o¹⁴⁴⁻³⁸⁶-Alexa350) and E3 were incubated in 50 mM KH₂PO₄ (pH=7.0) at room temperature for 5 min, then the excitation wavelength was set at 350 nm, and the spectrum was recorded. The hE3 was gradually titrated into the mixture, and the spectrum was recorded after the addition of each aliquot. The spectrum of E2o¹⁻¹⁷³-Alexa350 (and of E2o¹⁴⁴⁻³⁸⁶-Alexa350) were recorded at the same concentration in the same buffer as control. The emission intensity of the donor (Alexa350) remained the same, during all additions of E3. The results indicated no energy transfer between E2o and E3, irrespective of whether the Alexa350 label was attached to E2o¹⁻¹⁷³ or E2o¹⁴⁴⁻³⁸⁶.

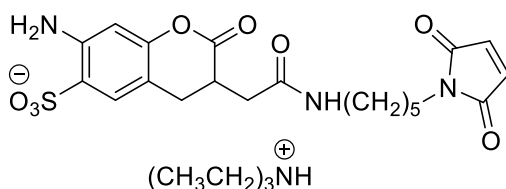


Figure 4.3: Structure of Alexa Fluor® 350 C5 Maleimide dye

4.3 Results and Discussion

Two approaches were developed to introduce an external fluorophore site-specifically onto the E2o proteins. In the first approach, the lipoamide carrying proteins, E2o¹⁻⁹⁵ (lipoyl domain, LDo, comprising lipoyl domain and part of the linker) and the E2o¹⁻¹⁷³ were dansylated. The lipoamide was first reduced to dihydrolipoamide, then labeled with

a dansyl group which had been tethered to the amino end of 4-aminophenylarseneoxide, resulting from a reaction of the trivalent arsenoxide with the reduced dihydrolipoamide [15]. In the second approach, N-(1-pyrene)-maleimide was used to label the unique cysteines of E2o¹⁻¹⁷³ (at Cys38) and E2o¹⁴⁴⁻³⁸⁶ (at Cys179) proteins.

4.3.1 E1o-E2o interactions

Fluorescence Spectroscopy Studies.

Titration of DANS-As labeled E2o¹⁻⁹⁵ by E1o resulted in quenching of the fluorescence of dansylamide, as also observed on titration of DANS-As labeled E2o¹⁻¹⁷³, suggesting that on interaction with E1o, a more hydrophilic microenvironment around the fluorophore was created according to data from the literature [16]. The calculated values of K_{ds} were: 0.135 μ M (with E2o¹⁻⁹⁵), and 0.038 μ M (with E2o¹⁻¹⁷³), providing the first hint that the linker region present in the latter could be an important determinant for E1o-E2o interaction.

With the N-(1-pyrene) maleimide labeled E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ proteins, on excitation at 340 nm, three major fluorescence emission bands were in evidence (374, 394 and 416 nm). The maximum emission intensities at 374 nm were used to calculate binding parameters. In the spectrum of E2o¹⁻¹⁷³, there is an additional broad peak centered around 470 nm, which is likely due to excimer formation. Pyrene-thiol conjugates often have long fluorescence lifetimes, which can allow adjacent pyrene rings within 6-10Å of each other to form excited-state dimers (excimers) that emit at longer wavelength than the

excited monomeric fluorophore [17]. This may be the result of unselective simultaneous labeling at both the single cysteine thiol and the reduced dihydrolipoyl group, located at residues 38 and 44, respectively.

Addition of E1o to pyrene-labeled E2o¹⁻¹⁷³ quenched the pyrene fluorescence ($K_d = 0.041 \mu\text{M}$), while addition of E1o to pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ enhanced the pyrene fluorescence intensity without any blue shift or red shift of the peak positions ($K_d = 0.06 \mu\text{M}$, Figure 4.5, 4.6). These results suggested that interactions of E1o with the E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ induce different changes in the microenvironment around the pyrene fluorophore: interaction between E1o and fluorophore labeled E2o¹⁻¹⁷³ leads the pyrene fluorophore to a relatively more hydrophilic, while interaction of E1o with labeled E2o¹⁴⁴⁻³⁸⁶ leads the pyrene fluorophore to a more hydrophobic environment. The calculated K_d 's are in good accord when the interaction with the same protein is measured with the two different fluorophores. The fluorescence binding experiments enabled us to conclude that the linker region in E2o located between LDo and CDo domains participates in the interaction with E1o.

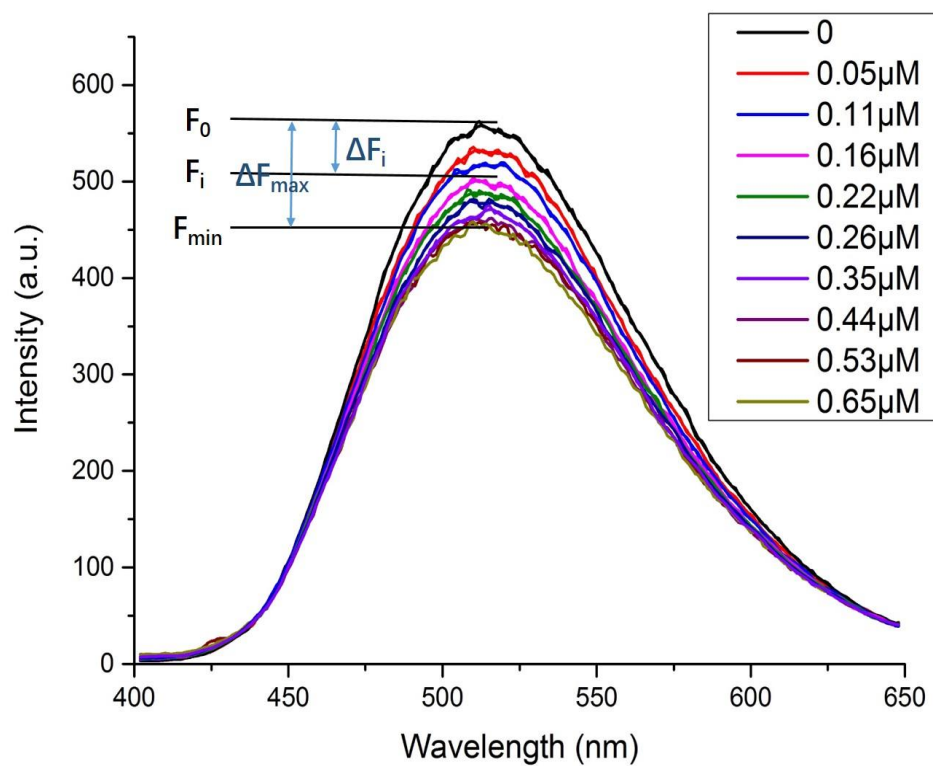
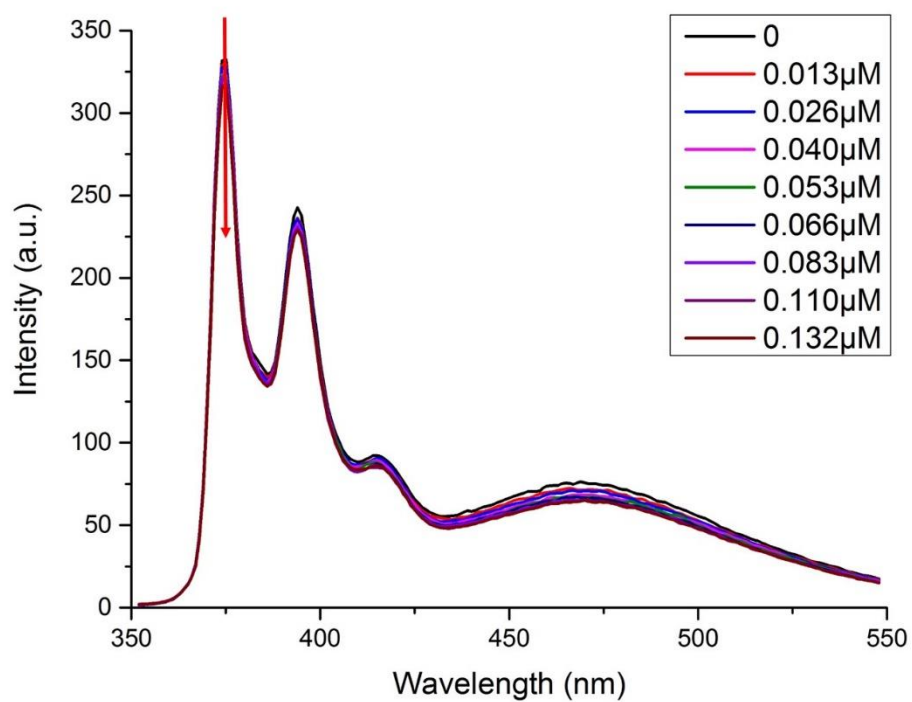


Figure 4.4. Fluorescence titration of DANS- labeled LDo by E1o. Quenching of the fluorescence of DANS-LDo on E1o binding. DANS-LDo (1.0 μM) in 30 mM KH_2PO_4 (pH7.5) was titrated by E1o (0.05-0.65 μM).



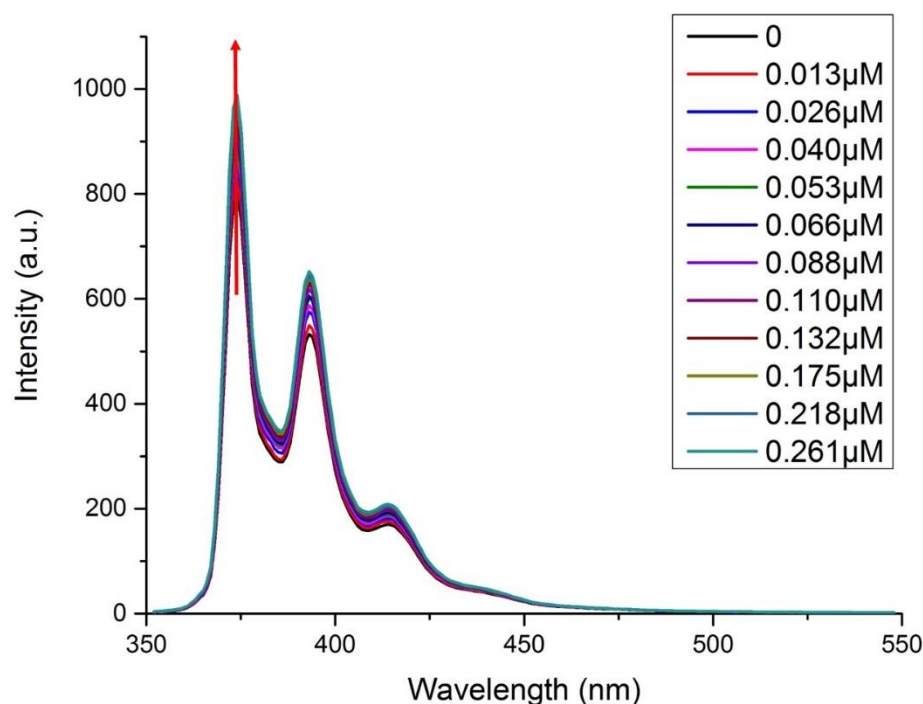


Figure 4.5. Fluorescence titration of Pyrene- labeled E2o¹⁻¹⁷³ and Pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ by E1o. (Top) Quenching of the fluorescence of Pyrene- labeled E2o¹⁻¹⁷³ on E1o binding. Pyrene- labeled E2o¹⁻¹⁷³ (1.5 μ M) in 30 mM KH₂PO₄ (pH 7.5) was titrated by E1o (0.013-0.261 μ M). (Bottom) Enhancement of fluorescence of the Pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ on E1o binding. Pyrene-labeled E2o¹⁴⁴⁻³⁸⁶ (1.5 μ M) in 30 mM KH₂PO₄ (pH7.5) was titrated by E1o (0.013-0.261 μ M).

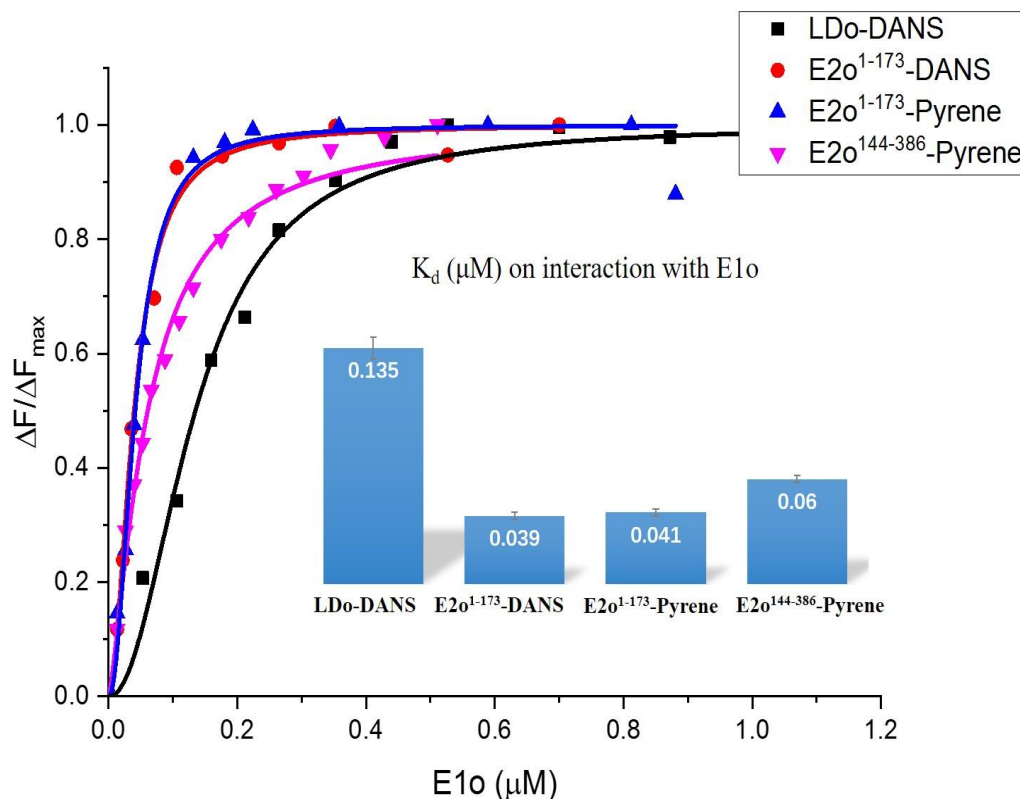


Figure 4.6. Fluorescence titration curves for DANS-As- labeled E2o¹⁻⁹⁵ and E2o¹⁻¹⁷³ and for pyrene labeled E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ proteins by E1o. In all cases the fluorescence intensity of the fluorophore-labeled E2o protein was quenched on titration by E1o. In a typical experiment, the fluorophore-labeled E2o protein (1.0-2.0 μM concentration of subunits) in 30 mM KH₂PO₄ (pH 7.5) was titrated by E1o (0.01-1.5 μM) at room temperature. The excitation wavelength was 338 nm for DANS and 340 nm for pyrene labeled E2o proteins. The fluorescence titration curves were fitted by using the Hill equation (equation 4.1). Inset: The K_d values for E1o interaction with E2o truncated proteins are presented.

4.3.2 E1o-E3 interactions

In view of the much weaker interactions between these two components, we designed a fluorescence resonance energy transfer (FRET) experiment to identify any potential interaction in this complex. Thiochrome diphosphate (TCDP, binds tightly and specifically to the E1o active center), a fluorescent analogue of ThDP was used as a donor, while the fluorescent FAD on E3 would act as energy acceptor, as reported by Hammes and coworkers on the *E. coli* PDHc).

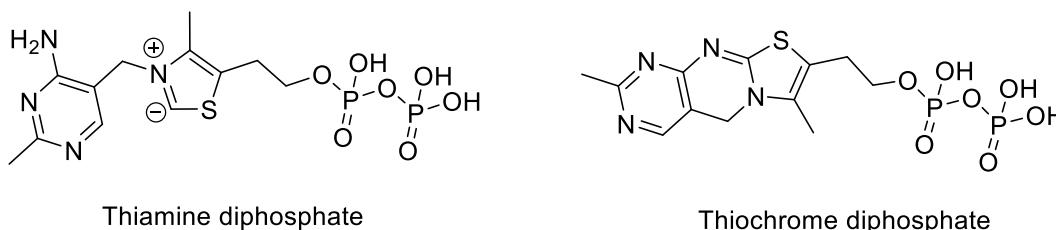


Figure 4.7. Structural formulae of thiamin and thiochrome diphosphate

First, we removed most of the ThDP from E1o. According to the E1o-specific activities, only 17% of ThDP still remained on the protein after passing it through a desalting column followed by extensive dialysis. Flavins have a broad absorption range with four structure-less peaks centered at 446, 375, 265 and 220 nm, all of which possess high molar extinction coefficients. In aqueous solution, flavins exhibit a bright yellow fluorescence at 520 nm with a quantum yield of 0.26 in pH 7.25 buffer (19-20). TCDP has a maximal absorption at 367 nm with a molar extinction coefficient equal to 20600 M⁻¹cm⁻¹. The maximal emission of TCDP is centered at 450 nm [21].

The efficiency of energy transfer is related to the distance between donor and acceptor molecules, which can be treated by the Forster equation [20].

$$E = r^{-6}/(r^{-6} + R_0^{-6}) \quad \text{Equation 4.3}$$

R_0 also called the Forster radius which is defined as the distance of separation at which the energy transfer efficiency is 50%.

$$R_0 = (9.79 \times 10^3)(JK^2Qn^{-4})^{1/6} \quad \text{Equation 4.4}$$

In this equation, K^2 is a dipole-dipole orientation factor, Q represents the quantum yield of the donor molecule, n is the refractive index of the medium, and J is the integral of spectral overlap of the donor and acceptor. According to studies by the Hammes group, the measured efficiency of energy transfer of $\sim 7.8\%$ could be estimated, corresponding to a distance of $\sim 45 \text{ \AA}$ with a possible range of 30-60 \AA between the thiochrome diphosphate and FAD binding sites in *E. coli* PDHc [17]. To estimate the efficiency of energy transfer, individual spectra of the E1o-bound thiochrome diphosphate and E3-bound FAD were recorded (Fig. 11, right). On excitation at 367 nm, the emission maximum at 450 nm (E1o-bound thiochrome diphosphate) and at 520 nm (E3-bound FAD) were detected and were in a good accord with that for free thiochrome diphosphate and FAD (Fig. 11, left), and with data reported in the literature. At pH 7.25 in water, FAD exhibits fluorescence with an emission maximum centered at 520 nm and quantum yield of 0.26 (30, 31), while thiochrome diphosphate has emission maximum centered at 450 nm (29) (Fig. 11, left). On mixing of the E1o-bound thiochrome-diphosphate and of E3-bound-FAD, the fluorescence intensity at 450 nm was quenched and at 520 nm was enhanced. Using the value of R_0 estimated by the Hammes group, and the observed change in fluorescence intensity at 450 nm (from 1027.7 a.u. to 906.4 a.u.), the efficiency of energy transfer was calculated to be 11.8% according to equation 4.5, corresponding to

a distance of approximately 28-56 Å between the ThDP binding site on E1o and FAD binding site on E3.

$$E = I - \tau_{DA} / \tau_A = I - F_{DA} / F_A \quad \text{Equation 4.5}$$

In equation 4.5, τ_{DA} and τ_A are the donor (thiochrome diphosphate) fluorescence lifetime in the presence and absence of an acceptor (FAD) respectively; F_{DA} and F_D are the donor fluorescence intensities with and without an acceptor, respectively. This result attests to interactions between E1o and E3, more specifically, an interaction between E1o ThDP binding site and FAD binding site on E3.

Quenching of the E1o intrinsic fluorescence by TCDP provides a K_d . The analogue TCDP has been known as a competitive inhibitor for ThDP dependent enzymes. It has also been reported that the intrinsic protein fluorescence due to tryptophan can be quenched by the binding of ThDP and ThDP analogues. The fluorescence quenching resulting from addition of TCDP exhibited saturation behavior, which can be fitted to Equation 2 to estimate the K_d value.

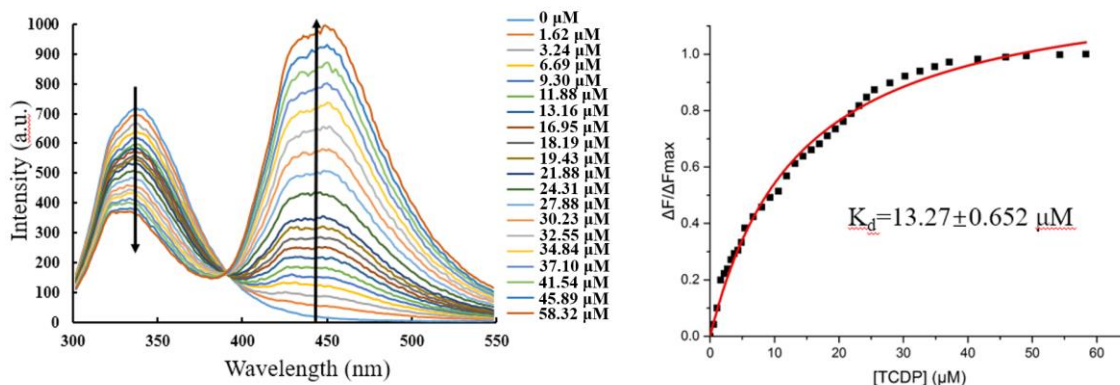


Figure 4.8: Fluorescence titration of E1o by thiochrome diphosphate. (Left) The E1o (3 μ M) in 50 mM KH_2PO_4 (pH 7.0) was titrated by thiochrome diphosphate in the indicated range of concentrations. A sample of E1o with 17% of tightly-bound ThDP was used in the experiment. The excitation wavelength was 280 nm due to E1o tryptophan fluorescence that could be quenched by ThDP and its analogues. On addition of thiochrome diphosphate, the intrinsic tryptophan fluorescence at 348 nm was gradually quenched and the fluorescence at 450 nm due to thiochrome diphosphate binding was gradually built up. (Right) Plot of the relative fluorescence changes at 348 nm on added thiochrome diphosphate at the indicated concentration. The trace is the regression fit line to eq. $\Delta F / \Delta F_{\text{max}} = [\text{TCDP}]^n / (S_{0.5}^n + [\text{TCDP}]^n)$, where $\Delta F / \Delta F_{\text{max}}$ is a relative fluorescence; $S_{0.5}$, is the concentration of TCDP at half saturation and n is the Hill coefficient.

The energy transfer from E1o-TCDP to E3 was determined on a Varian Cary Eclipse fluorescence spectrophotometer. The efficiency of energy transfer was obtained from the excitation spectrum of the energy acceptor-TCDP with the concentrations of E1o and E3 both at 5 μ M concentration in 50 mM KH_2PO_4 (pH 7.0). Also, separate spectra were recorded for E1o-bound TCDP, and for E3-bound FAD at the same concentration in the same buffer. This result confirms the interactions between E1o and E3, more specifically, an interaction between the ThDP binding site from E1o with the FAD binding site on E3. This result is also in accord with HDX-MS data, from which, upon binding to E1o, two peptides with residues 35-70 and 107-145 in the FAD-binding domain of E3, and one

peptide with residues 381-403 from the interface domain of E3 displayed the largest decrease in deuterium uptake.

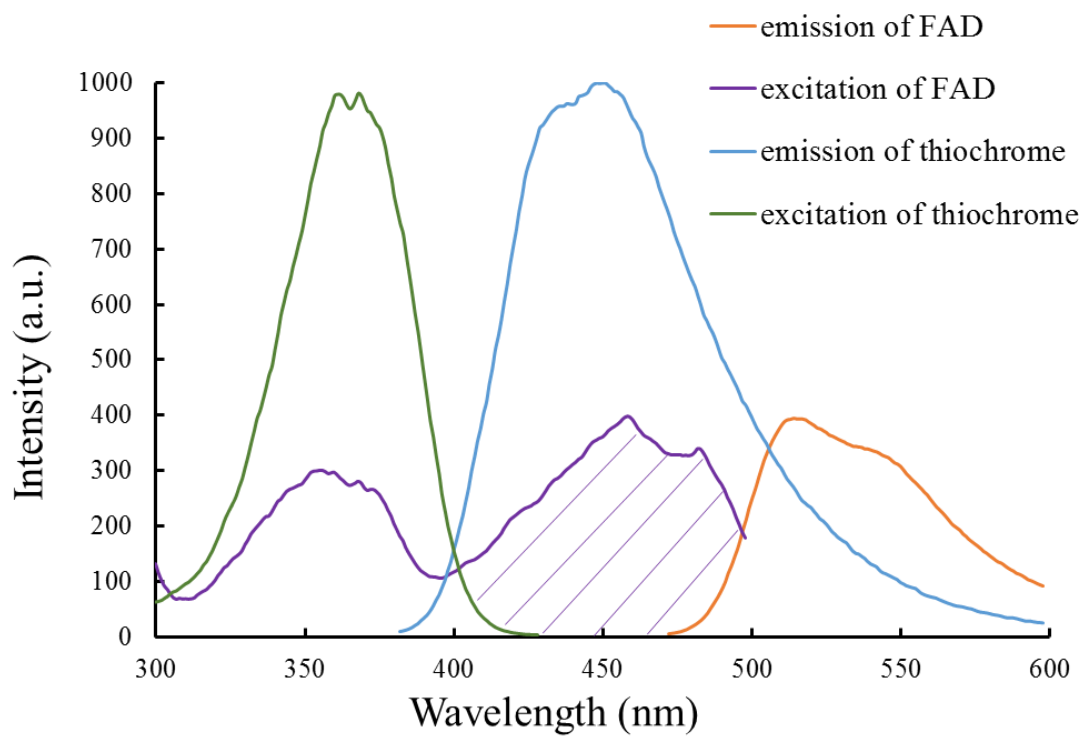


Figure 4.9: A plot of excitation and emission spectrum of FAD and TCDP at a concentration of 10 μM each. The spectral overlap is shown in shadow.

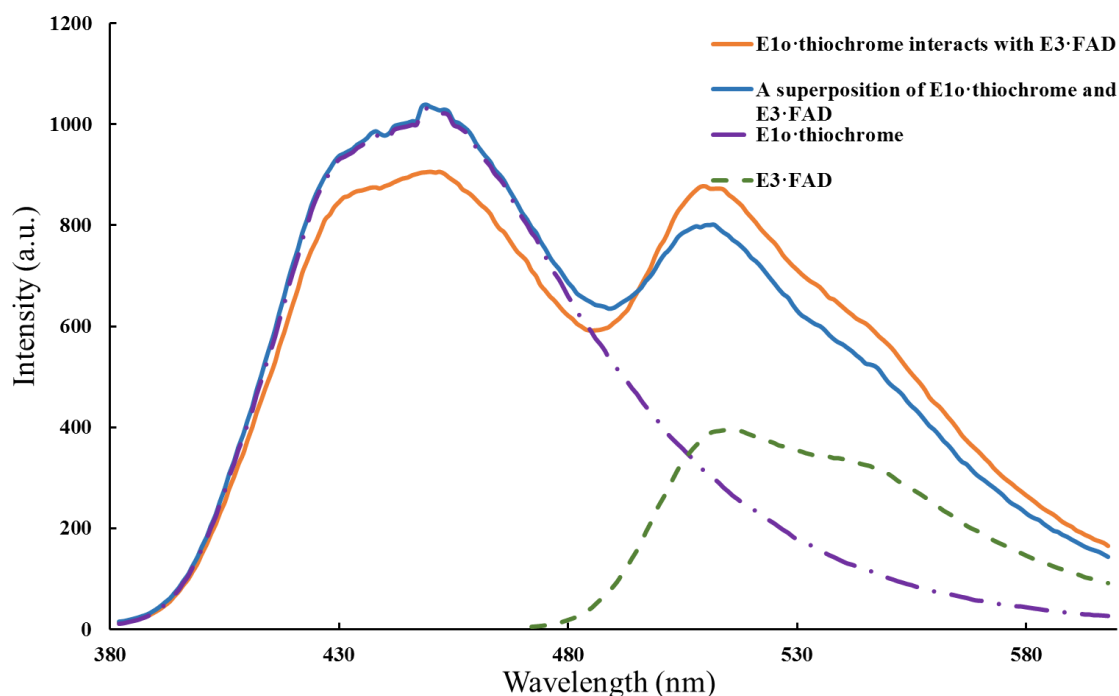


Figure 4.10: Measurements of energy transfer between E1o bound thiochrome diphosphate and E3 bound FAD. The emission spectra of E1o-thiochrome diphosphate (5 μ M) and of E3-FAD (5 μ M) recorded individually (dashed lines) and superimposed (in yellow) in 50 mM KH_2PO_4 (pH 7.0) at the excitation wavelength of 365 nm. In blue, the emission spectrum of a mixture of E1o-thiochrome diphosphate and E3-FAD recorded in 50 mM KH_2PO_4 (pH 7.0) at 5 μ M concentration of each protein.

4.3.3 E2o-E3 interactions

To test potential interaction in the E2o-E3 sub-complex by FRET, we used E2o-truncated proteins where the specific cysteines were labeled with Alexa Fluor[®]350 C5 maleimide for donor and FAD on E3 as potential acceptor. The Cys³⁸ located near the lipoylated Lys⁴⁴ in E2o¹⁻¹⁷³ di-domain, and Cys¹⁷⁹ in E2o⁴⁴⁻³⁸⁶ core domain were each labeled with the Alexa Fluor[®] 350 C5 maleimide dye. Energy transfer experiments were carried out

similarly to that reported above and data for Cys³⁸-labeled E2o¹⁻¹⁷³ are presented in Fig. 12. The spectrum of Cys³⁸-labeled E2o¹⁻¹⁷³ by itself displayed an emission maximum at 450 nm due to the labeled Cys³⁸. On addition of FAD-E3 at different concentrations (5-25 μ M) an emission maximum at 520 nm due to E3-bound FAD developed with no changes in the emission maximum due to labeled Cys³⁸, indicating that no energy transfer occurred between Cys³⁸-labeled E2o¹⁻¹⁷³ and E3-bound FAD (Figure 4.12). The most likely explanation is that the distance between the two chromophores is out of the FRET range ($>60\text{\AA}$), suggesting that neither the fluorophore on the E2o lipoyl domain , nor the one on the E2o core domain come in close proximity to the FAD binding site on E3. The reasons for not being able to observe energy transfer could be complex, since based on the equations that $E = r^{-6}/(r^{-6} + R_0^{-6})$ and $R_0 = (9.79 \times 10^3)(JK^2Qn^{-4})^{1/6}$, the efficiency of energy transfer is not only related to the inverse sixth power of the distance between the groups, but is also related to the mutual orientation of the groups, the refractive index of the solvent, and the overlap of the emission spectrum of the energy donor and the absorption spectrum of the energy acceptor. A likely explanation is that the distance between the two chromophores is out of the FRET range ($>60\text{\AA}$), suggesting that neither the fluorophore on the lipoyl domain, nor the one on E2o core domain is in close proximity to the FAD binding site on E3.

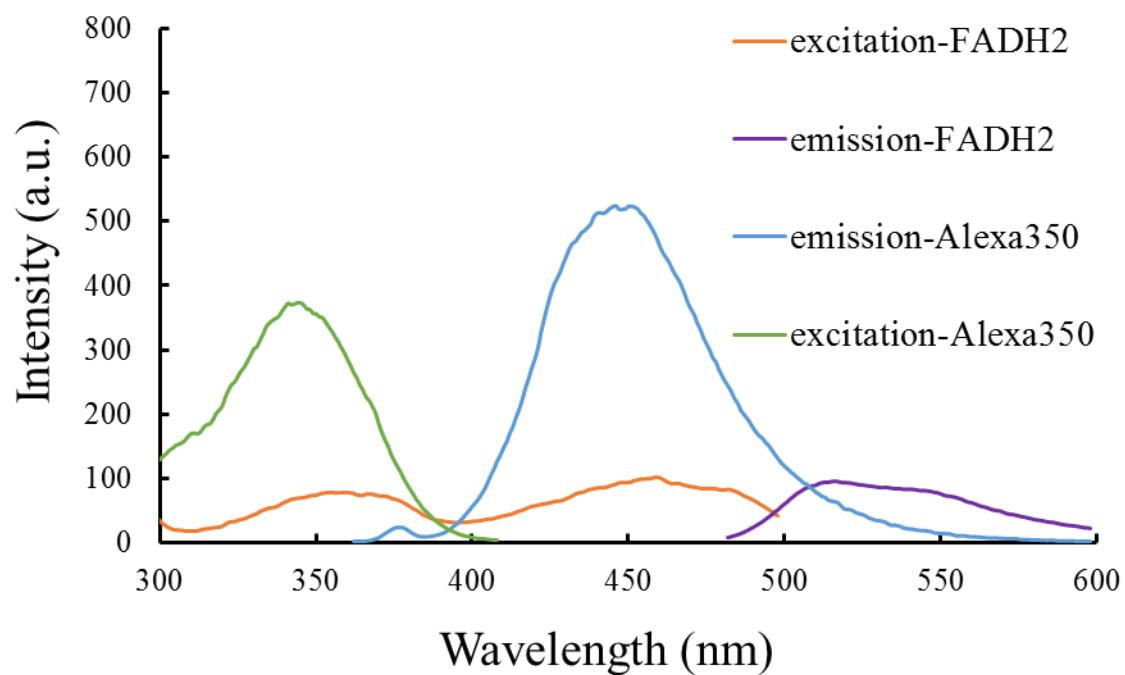


Figure 4.11: A plot of excitation and emission spectrum of FAD and Alexa350 with a concentration of 10 μM .

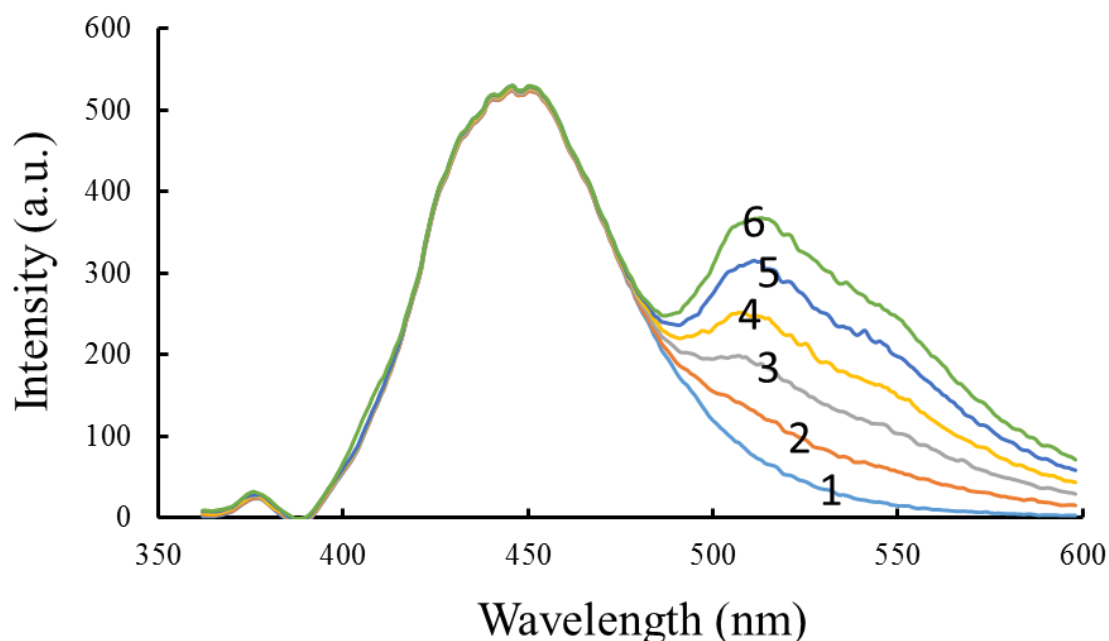


Figure 4.12: Measurements of energy transfer between $E2o^{1-173}$ -Alexa350 and E3 bound FAD. The excitation wavelength is 350 nm and the spectrum are recorded from 360 nm to 600 nm. The concentration of $E2o^{1-173}$ is 5 μ M in 50 mM KH_2PO_4 (pH=7.0). 1: the spectrum of $E2o^{1-173}$ -Alexa350 alone; 2: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 5 μ M E3; 3: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 10 μ M E3; 4: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 15 μ M E3; 5: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 20 μ M E3; 6: The mixture of 5 μ M $E2o^{1-173}$ -Alexa350 and 25 μ M E3. The emission intensity of $E2o^{1-173}$ -Alexa350 stay all the same, which means there is no energy transfer between Alexa350 fluorophore and FAD. (the results of $E2o^{144-386}$ -Alexa350 with E3 are the same, not shown).

Table 4.1: Protein modification checked by FT-MS

proteins	Peptides	Theoretical mass	Mass detected by FT-MS
LDo	Whole protein (lipoylated)	11246.2	938.1415 ¹²⁺
	Whole protein + TCEP	11248.2	938.3098 ¹²⁺
	Whole protein + DANS-As	11646.2	971.5631 ¹²⁺
E2o ¹⁴⁴⁻³⁸⁶	EAQNTCAMLTTFNEIDMSNIQEMR	931.0792 ³⁺	931.0831 ³⁺
	EAQNTCAMLTTFNEIDMSNIQEMR + N-(1-pyrene)maleimide	1030.1055 ³⁺	1030.1078 ³⁺
E2o ¹⁻¹⁷³	AVGDTVAEDEVVCEIETDK (lipoylated)	1105.9827 ²⁺	1105.9838 ²⁺
	AVGDTVAEDEVVCEIETDK (lipoylated) + N-(1-pyrene)maleimide	836.6838 ³⁺	836.6859 ³⁺
	AVGDTVAEDEVVCEIETDK (lipoylated) + 2* N-(1-pyrene)maleimide	935.7101 ³⁺	935.7142 ³⁺
E2o ¹⁻¹⁷³	Whole protein (lipoylated)		1003.9740 ²⁰⁺
	Whole protein + TCEP		1004.0759 ²⁰⁺
	Whole protein + DANS-As		1024.1036 ²⁰⁺

4.4 Conclusions from the fluorescence studies

Sub-micromolar K_{ds} were determined for binding E1o to E2o labeled with the dansyl group at dihydrolipoate, or to the two E2o didomains labeled with pyrene at their single cysteines. The fluorescence studies as well as the FRET experiments threw light on the loci of interactions.

By contrast, no evidence could be obtained for a stable E2o-E3 complex according to the FRET studies, suggesting that the FRET donor and acceptor are further than 60 Å from

each other. In view of the obligatory interaction of the dihydrolipoyl domain of E2o with the FAD binding site of E3, this result probably suggests a weak, perhaps short-lived interaction.

Perhaps counterintuitively, the E1o-E3 interactions falls somewhere in between the other two binary complexes. There is indeed a FRET observed between the TCDP site on E1o and the FAD site at E3. According to the HDX-MS results, the large distance may signal that the FAD site of E3 is near the amino terminal domain of E1o (the binding locus identified on E1o for both E2o and E3), but the current information is too low in resolution to decide on the likely locus from this experiment.

The lack of response of the TCDP fluorescence to addition of E2o proteins suggests that the strongest E2o binding sites on E1o are not at or near the ThDP binding loci, consistent with the HDX-MS studies suggesting that the new peripheral subunit binding site of E2o interacts with E1o at its N-terminal region, quite distant from the loci of ThDP and Mg^{+2} binding.

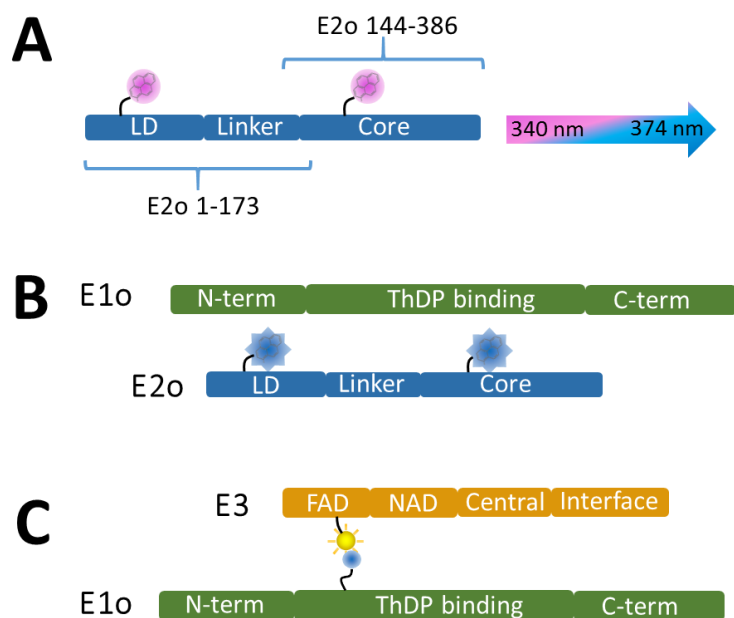


Figure 4.13: Summary of binary component interactions on the basis of fluorescence. (A) Two E2o constructs modified at the single cysteine on E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶. (B) Fluorescence titration experiments showed that E1o interacts with two E2o constructs through N-(1-Pyrene) maleimide modified single thiol group on E2o¹⁻¹⁷³ and E2o¹⁴⁴⁻³⁸⁶ (C) FRET experiments proved that FAD binding region on E3 is close to the ThDP binding region of E1o.

References

- [1] Kapanidis A N, Weiss S. Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules[J]. The Journal of chemical physics, 2002, 117(24): 10953-10964.
- [2] Michalet X, Weiss S, Jäger M. Single-molecule fluorescence studies of protein folding and conformational dynamics[J]. Chemical reviews, 2006, 106(5): 1785.
- [3] Holmes K L, Lantz L M. Protein labeling with fluorescent probes[J]. Methods in cell biology, 2001, 63: 185-204.
- [4] Gentle I E, De Souza D P, Baca M. Direct production of proteins with N-terminal cysteine for site-specific conjugation[J]. Bioconjugate chemistry, 2004, 15(3): 658-663.
- [5] Lin C W, Ting A Y. Transglutaminase-catalyzed site-specific conjugation of small-molecule probes to proteins in vitro and on the surface of living cells[J]. Journal of the American Chemical Society, 2006, 128(14): 4542.
- [6] Jäger M, Nir E, Weiss S. Site-specific labeling of proteins for single-molecule FRET by combining chemical and enzymatic modification[J]. Protein Science, 2006, 15(3): 640-646.
- [7] Kim Y, Ho S O, Gassman N R, Korlann Y, Landorf E V, Collart F R, Weiss S. Efficient site-specific labeling of proteins via cysteines[J]. Bioconjugate chemistry, 2008, 19(3): 786.

- [8] R.F. Chen and H. Edelhoch, Eds, Biochemical Fluorescence: Concepts, Vol. 1 and 2, Dekker, New York, 1975
- [9] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, Plenum, New York, 1983
- [10] D.M. Jameson and G.D. Reinhart, Eds., Fluorescent Biomolecules, Plenum, New York, 1989
- [11] G.G. Hammes, in Protein-Protein Interactions, C. Frieden and L.W. Nichol, Eds., Wiley, New York, 1981, pp. 257-287
- [12] G. Weber, in Spectroscopic Approaches to Biomolecular Conformation, D.W. Urry, Ed., American Med. Assn., Chicago, 1970, pp. 23-31
- [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] Guevara, E.L., Yang, L., Birkaya, B., Zhou, J., Nemeria, N.S., Patel, M.S. and Jordan, F. (2017) Global view of cognate kinase activation by the human pyruvate dehydrogenase complex. Sci Rep. 10.1038/srep42760.

[16] Schuldiner S, Weil R, Robertson D E, Kaback H R. Microenvironment of the binding site in the lac carrier protein[J]. Proceedings of the National Academy of Sciences, 1977, 74(5): 1851-1854.

[17] Han M K, Lin P, Paek D, Harvey J J, Fuor E, Knutson J R. Fluorescence studies of pyrene maleimide-labeled translin: excimer fluorescence indicates subunits associate in a tail-to-tail configuration to form octamer[J]. Biochemistry, 2002, 41(10): 3468-3476.

[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] Weber G. Fluorescence of riboflavin and flavin-adenine dinucleotide[J]. Biochemical Journal, 1950, 47(1): 114.

[20] Heelis P F. The photophysical and photochemical properties of flavins (isoalloxazines)[J]. Chemical Society Reviews, 1982, 11(1): 15-39.

[21] Moe Jr O A, Lerner D A, Hammes G G. Fluorescence energy transfer between the thiamine diphosphate and flavine adenine dinucleotide binding sites in the pyruvate dehydrogenase multienzyme complex[J]. Biochemistry, 1974, 13(12): 2552-2557.

CHAPTER 5. Conformational Dynamics of 1-Deoxy-D-xylulose 5-phosphate Synthase Upon Ligand Binding Revealed by H/D Exchange Mass Spectrometry

5.1 Introduction

Due to the increasing spread of antibiotic resistance, there is a critical need to investigate novel antibiotic targets. The enzyme 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) is an attractive potential anti-infective target because it catalyzes formation of a branch point metabolite that is essential in human pathogen metabolism, yet is an enzyme not utilized by humans [1-3]. DXP, the product of DXPS synthase, is a key intermediate in the biosynthesis of ThDP, PLP and isoprenoids [1-3], all of which are essential for bacterial growth and survival; thus, inhibition of DXPS should result in widespread toxicity within a bacterial pathogen.

DXPS catalyzes the thiamin diphosphate (ThDP)-dependent decarboxylation of pyruvate (Py) and subsequent carboligation to D-glyceraldehyde 3-phosphate (GAP), reminiscent of the reactions of two essential mammalian enzymes, the human pyruvate dehydrogenase E1 component (PDHc-E1) and transketolase (TK). Development of probes targeting DXPS is therefore challenging, and few selective inhibitors of this enzyme are reported [4-11]. However, despite its similarities to other ThDP-dependent enzymes, previous mechanistic and structural studies have uncovered several distinguishing features of DXPS [12-17]. Unlike the classical ping-pong mechanism carried out by other ThDP-dependent enzymes, DXPS catalyzes a random sequential,

preferred order mechanism which is unprecedented in ThDP-dependent enzymology [12-16]. The first enzyme-bound tetrahedral pre-decarboxylation intermediate, C2-lactyl/ThDP (LThDP), is uniquely stable on DXPS in the absence of GAP, and ternary complex formation upon binding of GAP is required to trigger decarboxylation of LThDP [12, 14]. Accordingly, DXPS possesses a large active site, nearly twice the volume of TK and PDH-E1 active sites, presumably to accommodate ternary complex formation [9]. Lastly, DXPS is characterized by an unusual structure compared to other homodimeric enzymes in this class, with its active site located at the interface of domains I and II on each monomer, in contrast to other ThDP-dependent enzymes whose active sites are found at the dimer interface [17]. Taken together, the unique structural and mechanistic features of DXPS provide a basis for the selective inhibition of this enzyme, and new inhibitor classes are now emerging as potential tools to study this unique antimicrobial target [5, 10, 11].

Despite these advances, gaps remain in our understanding of the factors underlying the novel mechanism of DXPS. Structural studies of DXPS are scarce, limiting the potential to understand critical links between DXPS structure and mechanism that are needed to guide selective inhibitor design. We and others have hypothesized that binding of GAP triggers a conformational change that causes the destabilization and subsequent decarboxylation of LThDP [15, 16]. Evidence for GAP induced conformational changes has been reported on both the *Escherichia coli* and *Plasmodium falciparum* enzymes in pre-steady state CD [15] and fluorescence binding experiments, respectively [16]. While these studies lend support for the conformational dynamics of DXPS, they report

indirectly on potential conformational changes through the unusual behavior of mechanistic and binding data in the presence of GAP and offer little information about specific regions of conformational flexibility. Currently, the most direct evidence for the conformational flexibility of DXPS may come from the incomplete crystal structures of DXPS from *Deinococcus radiodurans* and *E. coli*. In the report by Xiang, et al. [17], in situ proteolysis was required for the successful crystallization of *E. coli* DXPS.

Interestingly, one of the cleaved *E. coli* DXPS segments overlaps with residues that are not observable in the *D. radiodurans* structure, suggesting these regions are conformationally flexible [17]. Structural studies of other ThDP-dependent enzymes suggest they may also undergo conformational change along their reaction coordinates [18-20]. However, the intriguing evidence for conformational dynamics of DXPS, in conjunction with its unprecedented mechanism, suggests that DXPS can exist in distinct conformations that are potential targets for selective inhibition. Investigating DXPS structure is a necessary step toward identifying targetable conformations. Unfortunately, structural studies of DXPS structure are severely limited by the potential conformational flexibility of the enzyme which impedes classic crystallographic studies. Therefore, there is a pressing need to develop techniques to directly probe the dynamics of DXPS along its reaction coordinate, which could reveal unique conformations that are starting points for selective inhibitor design.

The goal of this study was to probe the conformational dynamics of *E. coli* DXPS in the presence and absence of GAP, DXP, and a stable Py mimic, methyl acetylphosphonate (MAP), using hydrogen-deuterium exchange mass spectrometry (HDX-MS). Our results

reveal, for the first time, the conformational dynamics of three previously uncharacterized segments of *E. coli* DXPS supporting the utility of this technique to study DXPS structure and mechanism simultaneously. In the absence of any ligands, these regions of DXPS display a bimodal pattern of deuterium uptake, implying EX1 kinetic behavior, in addition to EX2 kinetics, a finding for a protein in its native, ligand-free state that is increasingly prevalent in the recent literature. [22-25]. Although the rates of deuterium uptake are unaffected by ligand binding, the rates of unfolding in the EX1 regions are ligand-dependent, suggesting that the closed conformation of DXPS in the presence of MAP is distinct from the conformational changes induced by GAP and DXP. The observation of conformational changes induced on DXPS by GAP binding provides the highest resolution data collected to date to support the role of conformational dynamics in GAP-induced LThDP decarboxylation. Collectively these results provide the first direct evidence for the conformational changes of DXPS in response to substrate, product, and substrate analog binding, the latter of which offers insights into conformation changes induced by inhibitors based on this substrate analog scaffold. This report also constitutes the first example of a ThDP-dependent enzyme that displays EX1 kinetics in the flexible regions of the protein as observed by HDX-MS, adding to the short list of proteins that display EX1 kinetic behavior under physiological conditions, and providing crucial insight into the novel mechanism of DXPS.

5.2 Materials and methods

Deuterium oxide (D₂O) 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 MgCl₂, 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 D₂O buffer to yield a final concentration of 95 % D₂O at pH 8.0. D₂O buffer was prepared the same way as DXPS exchange buffer except 99.9% D₂O 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

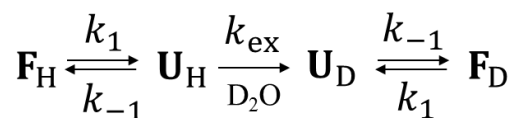
The frozen deuterated sample was quickly thawed and loaded with an ice-cold syringe into a 20- μ l sample loop inside the refrigeration system. The protein sample (40 pmol) was carried by a 0.2 ml/min digestion flow (0.1% formic acid) into an immobilized pepsin column (Poroszyme Immobilized Pepsin Cartridge, 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 Mechanisms of EX1 and EX2 kinetics

Two mechanisms have been proposed to explain hydrogen/deuterium exchange under physiological conditions in proteins. The first one suggests that exchange may occur from the folded form, the second one suggests that partial local unfolding occurs before

exchange can take place [26, 27]. Under physiological conditions, the second mechanism (Equation 5.1) appears to be the predominant pathway for H/D 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 . This unfolding process exposes backbone amide hydrogens which can subsequently undergo exchange with D_2O . The rate constants k_1 , k_{-1} , and k_{ex} describe the kinetics of unfolding, refolding, and hydrogen exchange steps, respectively [26]. The hydrogen exchange in this model can be described by two different kinetic schemes, EX1 and EX2. When the exchange is governed by EX1 kinetics, the unfolding rate is much faster than the refolding rate ($k_1 \gg k_{-1}$). Under these conditions, all of the amide hydrogens exchange with deuterium in the unfolded (open) state before refolding occurs. As a result, EX1 kinetics gives rise to two distinct mass envelopes. The lower mass envelope represents the folded (closed, \mathbf{F}_H) undeuterated state, and the higher mass envelope the unfolded (open, \mathbf{U}_D), deuterated state. In EX2 kinetics, the protein refolding rate is much faster than the unfolding rate ($k_1 \ll k_{-1}$), therefore the unfolding has to occur multiple times before a successful exchange reaction takes place. EX2 kinetics leads to a single isotopic distribution, which gradually shifts to a higher m/z range over time. While the majority of native proteins behave according to EX2 kinetics under physiological conditions, there are increasing examples reporting EX1 kinetic behavior in

the absence of ligand. The EX1 behavior must be confirmed by rigorous control experiments, thoroughly described by Engen and coworkers [28]. Under our experimental conditions, we can indeed attribute our observations to EX1 behavior in the three regions of DXPS identified. The observation of EX1 H/D exchange kinetics on DXPS under physiological conditions could provide important insights into the conformational dynamics of DXPS along its reaction coordinate.

5.3.2 Overview of HDX patterns on DXPS

The time dependence of hydrogen-deuterium exchange of the backbone amide protons of DXPS was studied over a 30 min time course (20 s, 1, 2, 3, 5, 10 and 30 min). On-line digestion by pepsin followed by LC-MS analysis under the selected HDX conditions yielded 56 peptides, many of which were partially overlapping, providing 93.7% sequence coverage (Table 5.1). The redundancy of peptides detected is very important because it provides an internal control for consistency of sequence assignment. Throughout the entire protein, overlapping peptides display similar HDX kinetics. While most of the overlapping peptides of DXPS exclusively display EX2 exchange kinetics, there are three distinct regions of the protein in which each of the overlapping peptides are characterized by EX1 kinetics (results reported below). These observations serve as an additional control to prevent false assignment of exchange behavior. Importantly, the high sequence coverage achieved in this study allowed us, for the first time, to monitor structural changes throughout the entire protein including peptides 183-238 and 292-317, which are not observed in the only available crystal structure of *E. coli* DXPS [17]. This new structural

information could offer novel mechanistic insights to guide the design of selective DXPS inhibitors.

The average deuterium uptake percentage by DXPS at 30 min is illustrated in Fig. 2. The peptides spanning residues 183-238 and 292-317 are represented schematically as they are important for this study, but are missing from the published structure of DXPS (see Figs.5.1). The overall rate of deuterium incorporation was low on this enzyme; many regions underwent less than 35% deuteration after a 30 min exchange (see Figs.5.1). The deuterium incorporation pattern clearly indicates that the greatest extent of H/D exchange occurs in domain I, suggesting that this is the most dynamic of the three domains of the DXPS monomer. In contrast, the areas with the least extent of deuterium incorporation are located in domains II and III at the dimer interface. These findings are intriguing as the crystal structure of *E. coli* DXPS has revealed that its active site is located between domains I and II, in contrast to other homodimeric ThDP-dependent enzymes in which the active sites reside at the dimer interface [17].

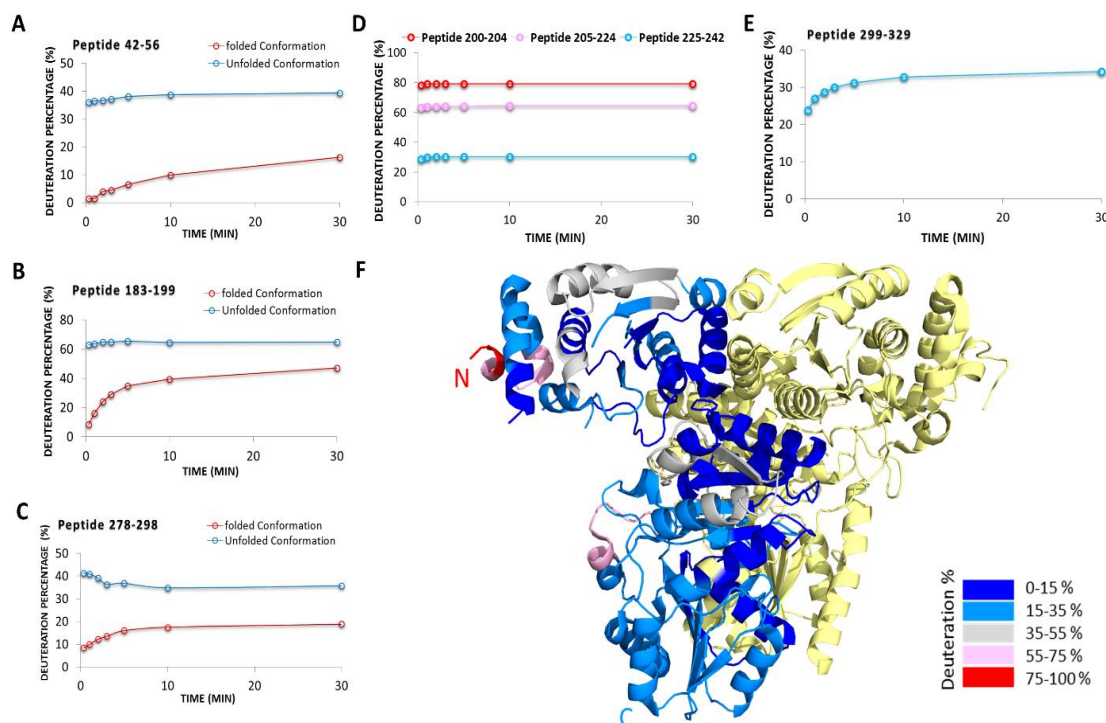


Figure 5.1: Average deuterium uptake by key peptides of *E. coli* DXPS. (A-C) Time course of deuterium uptake by peptides exhibiting a bimodal isotopic distribution. The centroid masses of lower (red traces) and higher (blue traces) mass envelopes were fit and calculated individually. Since the centroid mass of lower mass envelope also changes with increasing exposure time, the displayed kinetics is likely a mixture of EX1 and EX2 kinetics. (D and E) The deuterium uptake plots for regions comprising residues 183-238 and 292-317, which are missing from the crystal structure [18]. (F) The average percentage of deuterium incorporation at the 30 min time point is mapped onto the crystal structure of *E. coli* DXPS. For clarity, percentage deuterium uptake is represented for one DXPS monomer, while the other monomer is shown in yellow.

5.3.3 HDX-MS detects local structural dynamics of DXPS.

First, the conformational flexibility of full length DXPS in the absence of ligands was examined. Deuterium uptake for most peptides identified under our HDX conditions revealed a single isotopic distribution that increased in mass over time, suggesting that these regions display common EX2 kinetic behavior. Interestingly, three regions (42-58, 183-199, and 278-298) displayed a bimodal isotopic distribution, consistent with EX1 kinetics (Fig. 5.2). Because peptides 42-56 and 51-58 have overlapping residues, they will be referred to as peptide 42-58 in the following discussion. Addition of the ligands selected for this study had little effect on the deuterium incorporation by the peptides of DXPS that were characterized by EX2 kinetics.

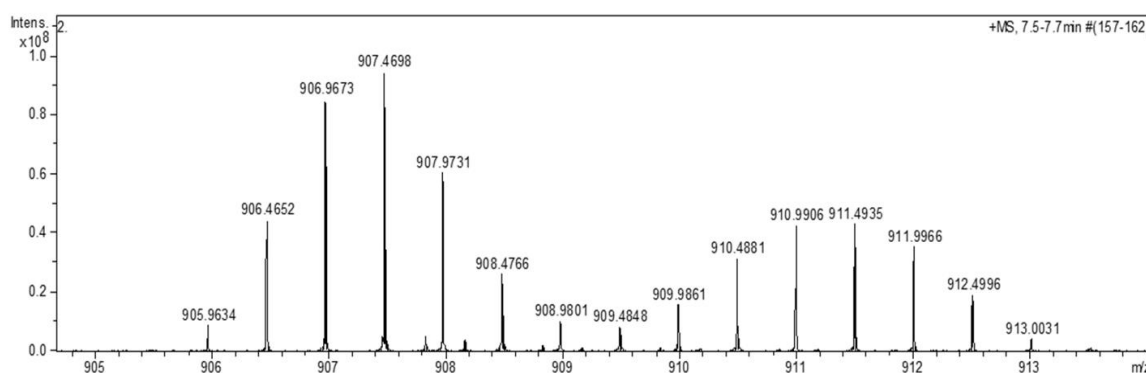


Figure 5.2: Mass spectrum of peptides 183-199 at the 1-min HDX time point in the presence of GAP.

Regions that were affected by the addition of ligands were those that display EX1 kinetics. Notably, unlike most of the reported bimodal patterns which are induced upon adding a binding partner [29, 30], the bimodal patterns of these three regions are displayed by DXPS even in the absence of bound ligands. Although residues 183-199 and 292-298 (part of the 278-298 peptide) are missing from the crystal structure, the positions of the adjacent N-

terminal residues of these peptides (which are observed in the crystal structure) suggest that the three regions displaying EX1 kinetics are close to each other. (Fig. 5.1). Given their close spatial proximity to one another and the active site, it is possible that the conformational changes accounting for the observed EX1 kinetics occur simultaneously in all three regions (Fig. 5.1).

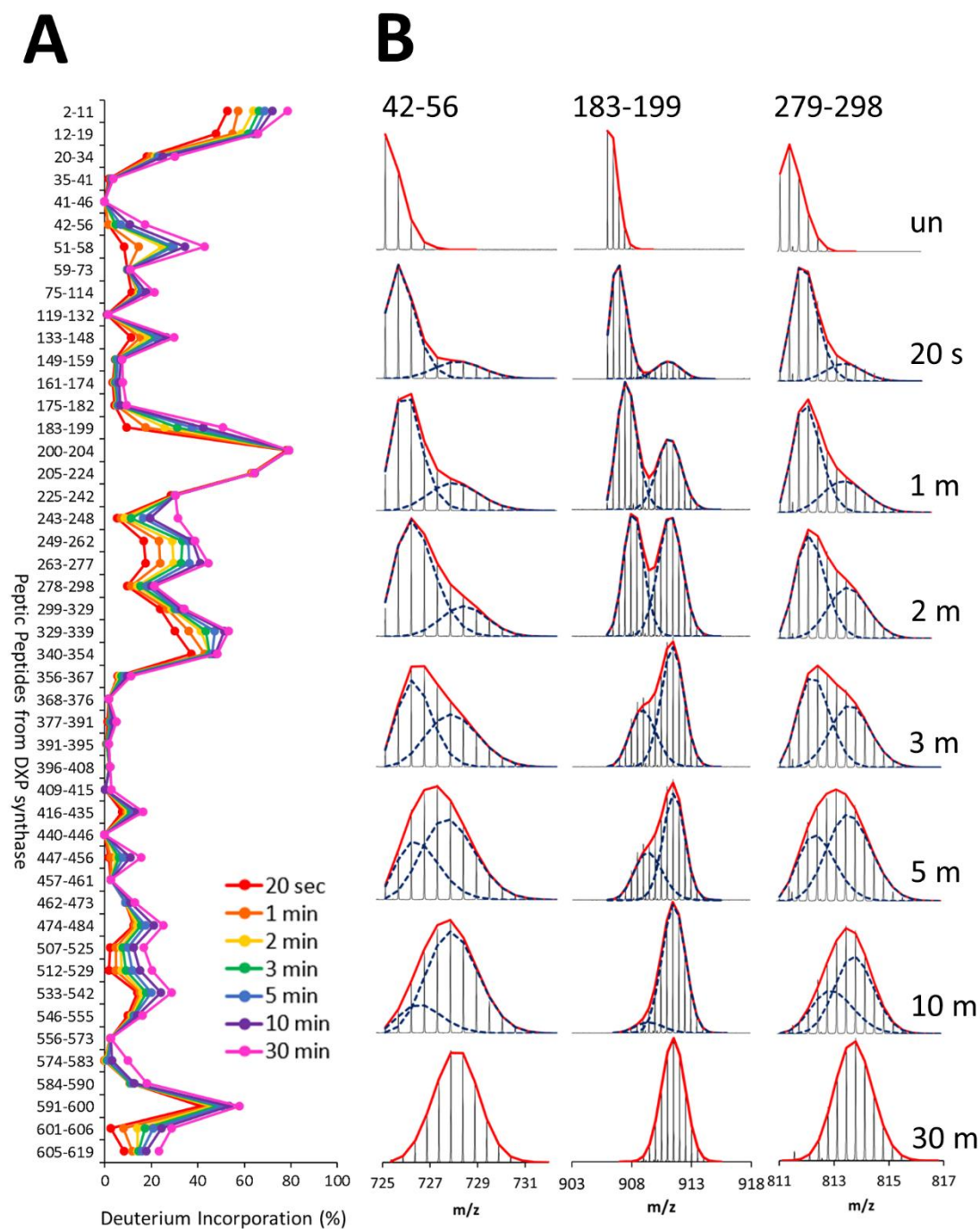


Figure 5.3: DXPS deuterium uptake change over 30 min. (A) Deuterium incorporation plot of full-length *E. coli* DXPS across seven different time points (20s and 1, 2, 3, 5, 10

and 30 min). (B) HDX-MS spectra of selected peptides (42-56, 183-199, and 278-298) charge envelopes (red lines) at increasing deuterium exchange time shows the presence of two isotopic distributions. Gaussian deconvolution of the mass spectrum reveals the presence of two isotopic distributions (blue dot lines).

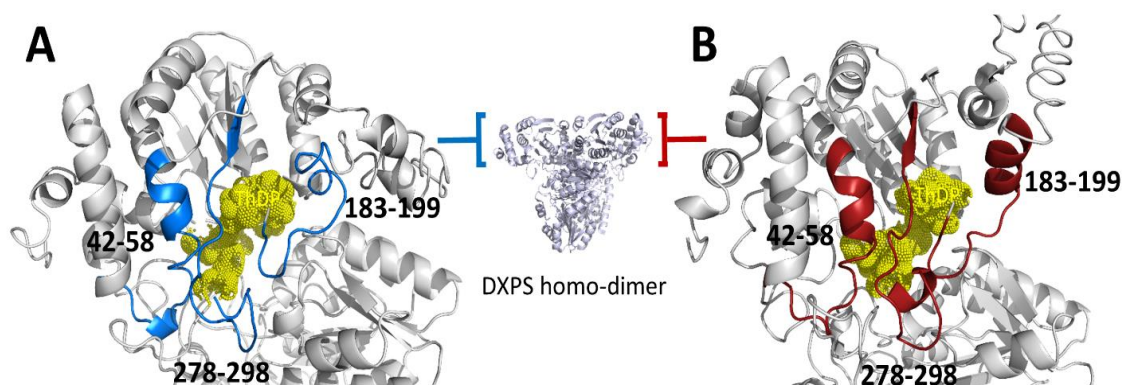


Figure 5.4: Spatial proximity of all regions displaying EX1 kinetic behavior in the DXPS structure. The position of all regions displaying EX1 kinetics on domain I of *E. coli* DXPS indicated the spatial proximity of these regions near the ThDP binding site of DXPS. (A) The predicted open state of all EX1 kinetic regions shown in blue. Yellow dot spheres indicate the ThDP binding pocket. All EX1 kinetic regions contain multiple residues, which are involved in ThDP binding. (B) Closed state of all EX1 kinetic regions shown in red. The open and closed structures were calculated with I-TASSER. *E. coli* DXPS structure (Protein Data Bank ID code 2o1s) was assigned as a restraint and template to guide I-TASSER modeling.

To further characterize the EX1 kinetics displayed by several peptides in our initial experiments, we carried out an H/D exchange time course study over 30 min (20 s, 1, 2, 3, 5, 10, and 30 min). Because of the relatively low degree of deuteration at 20 s (< 10 %)

and high degree at 30 min (> 65 %), peptides were observed at 20 s and 30 min, and additional shorter or longer time points were not deemed necessary, hence were not acquired for this study.

We considered two possible explanations for our observation of the bimodal distribution that could result from EX1 kinetics on DXPS, in which two distinct conformations interconvert during the HDX incubation period. In this case, it appears that the equilibration rate between closed and open conformations of the corresponding peptide is slower than the HDX rate, which indicates that both conformations are fairly stable under physiological conditions. Since DXPS is naturally present as a homo-dimer, a plausible explanation for the EX1 kinetic behavior of the three regions is that the DXPS homo-dimer is asymmetric, suggesting that the two subunits of 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 are in a disordered region in the DXPS structures from *E. coli* and *D. radiodurans* [17], 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. Interestingly, DXPS is the first ThDP-dependent enzyme to display this behavior, suggesting that this conformational flexibility may have an important role in the unique mechanism catalyzed by DXPS.

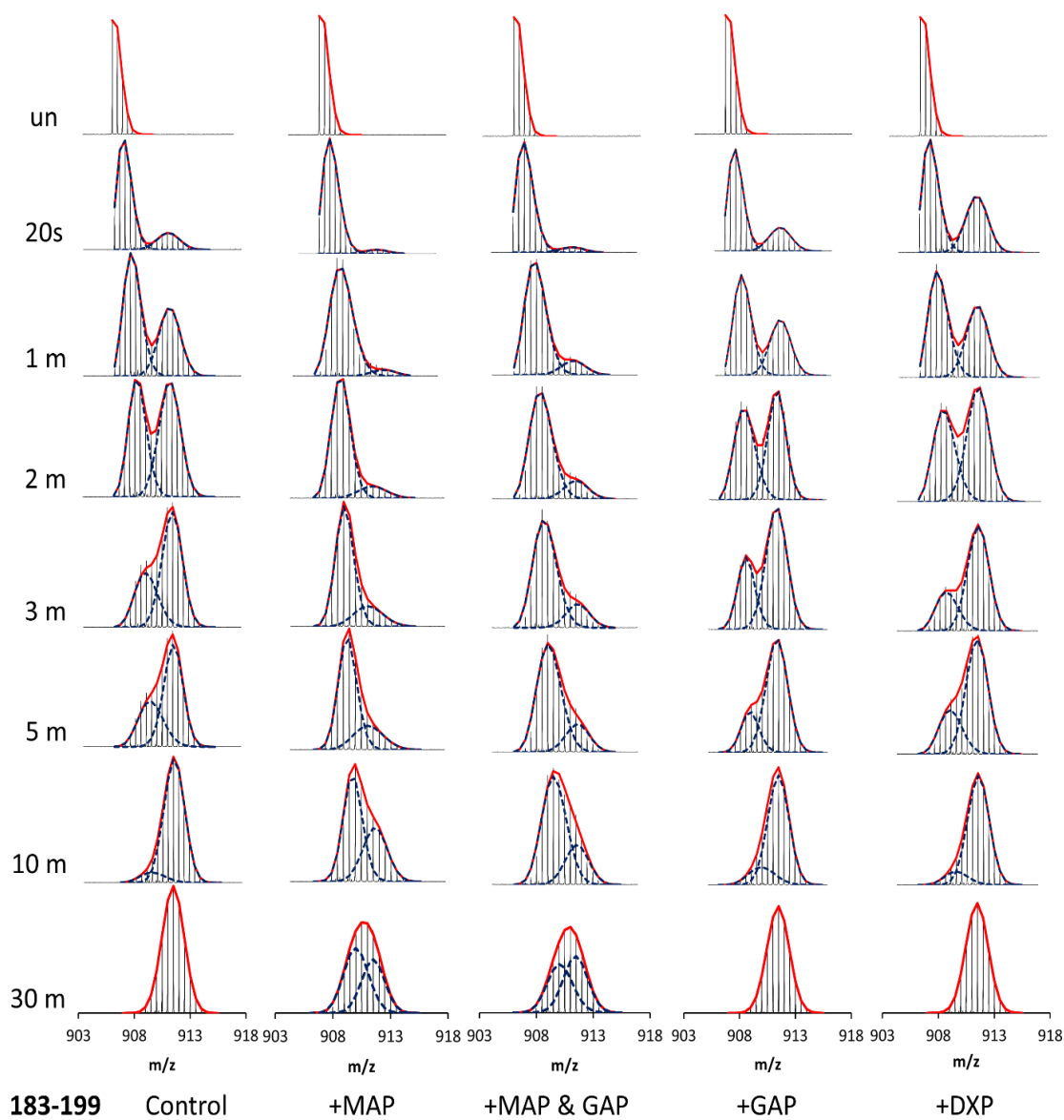


Figure 5.5: Complete time course of HDX-MS spectra of peptides 183-199 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

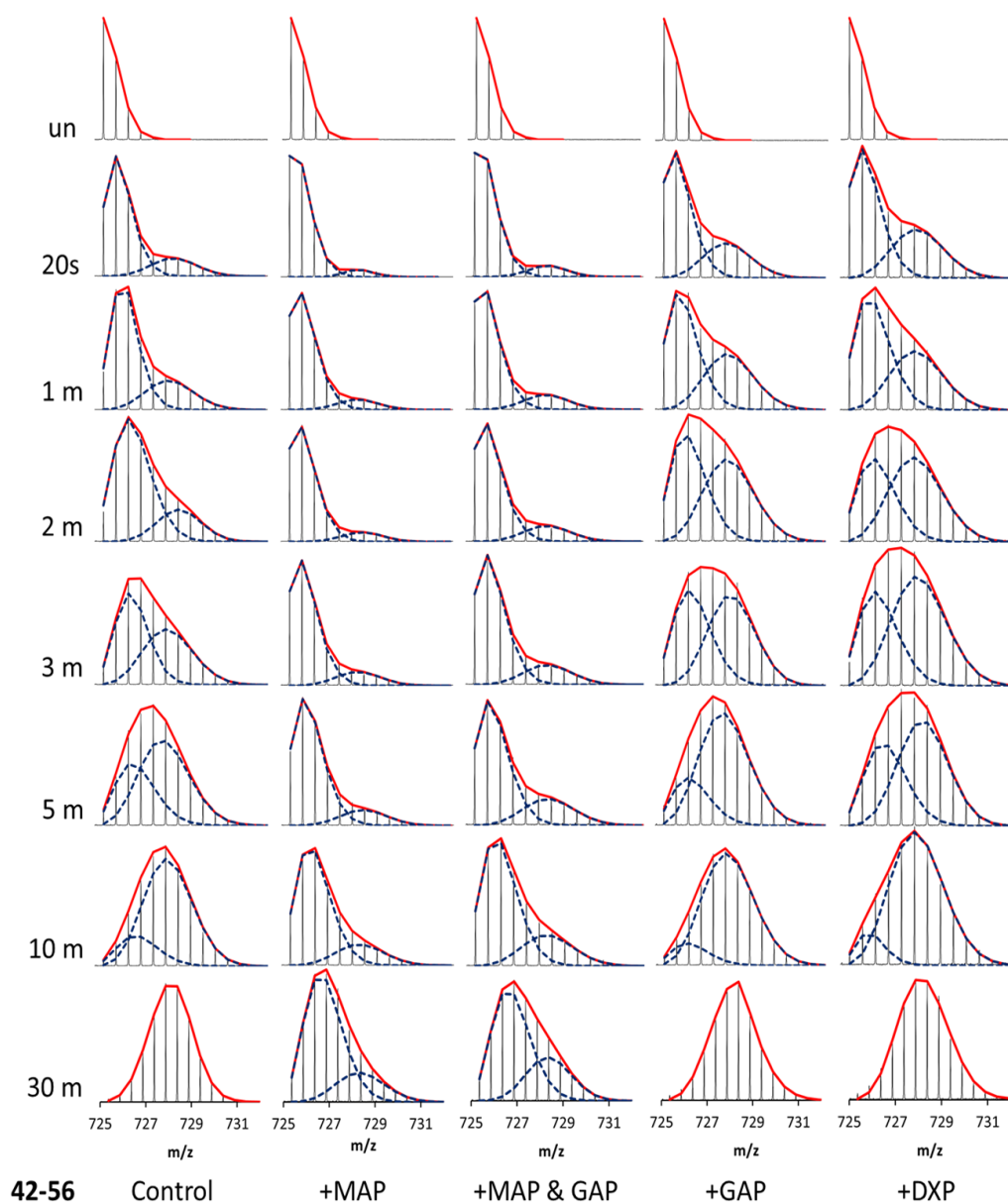


Figure 5.6: Complete time course of HDX-MS spectra of peptides 42-56 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

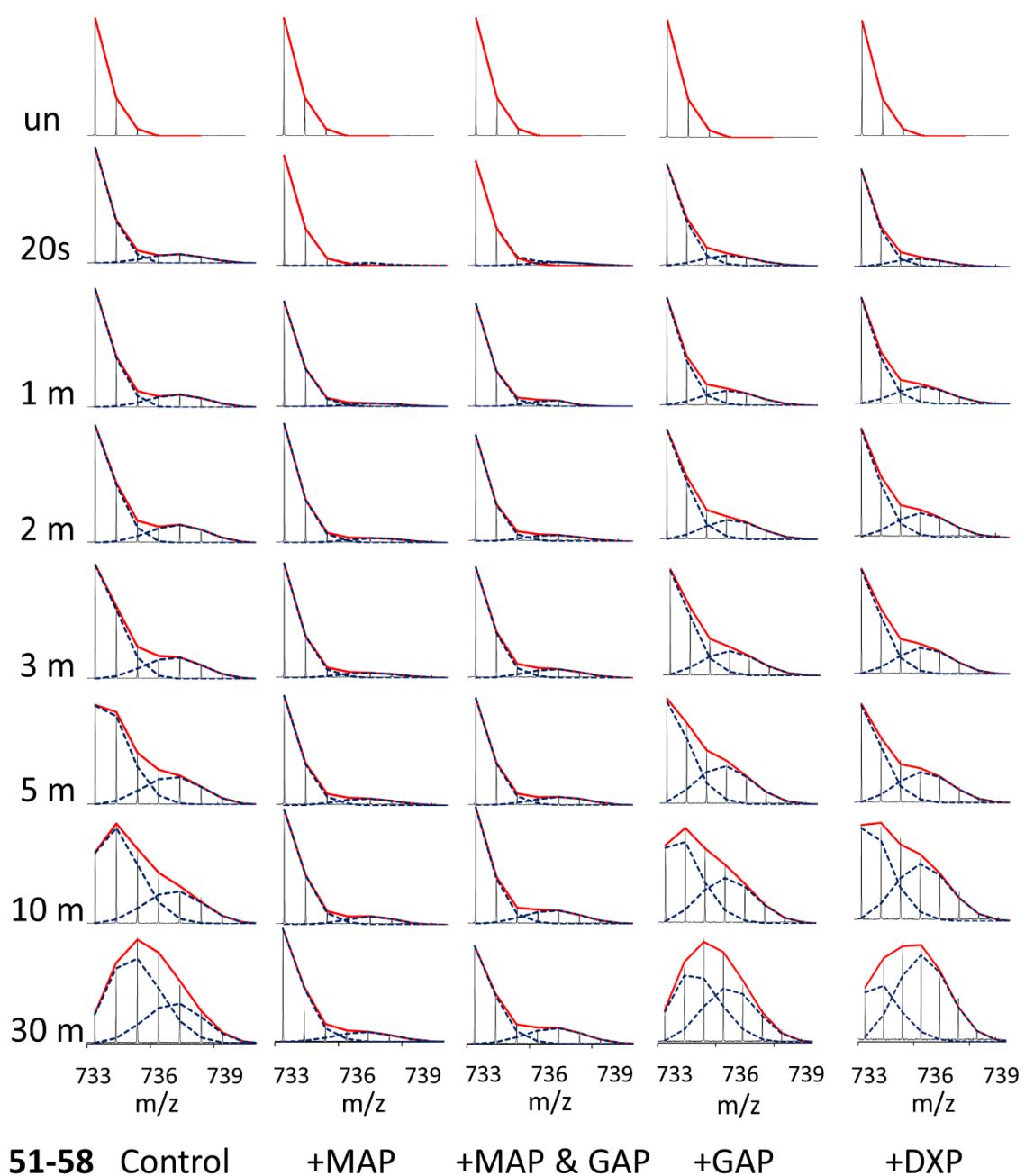


Figure 5.7: Complete time course of HDX-MS spectra of peptides 51-58 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines

indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

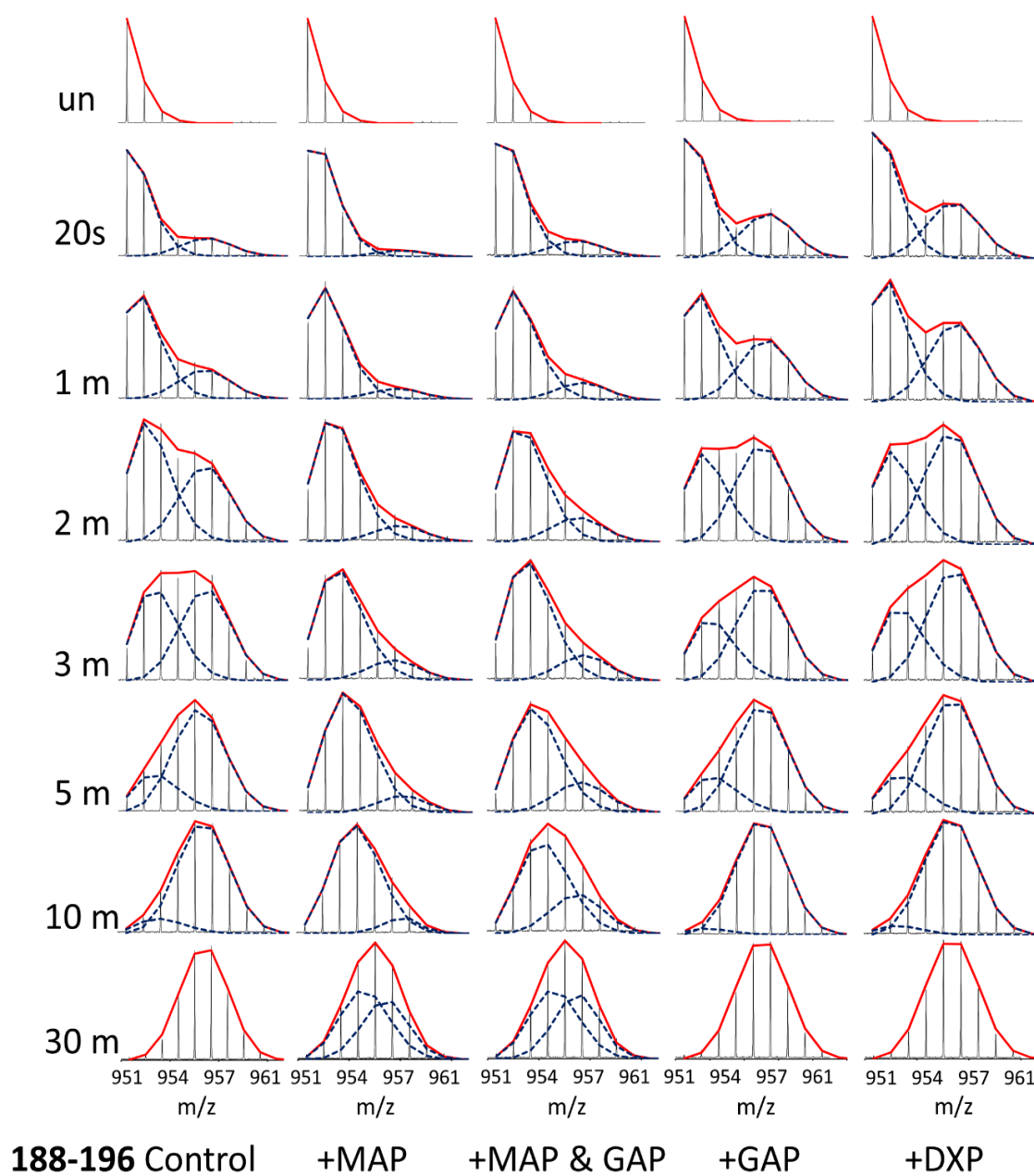


Figure 5.8: Complete time course of HDX-MS spectra of peptides 186-196 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

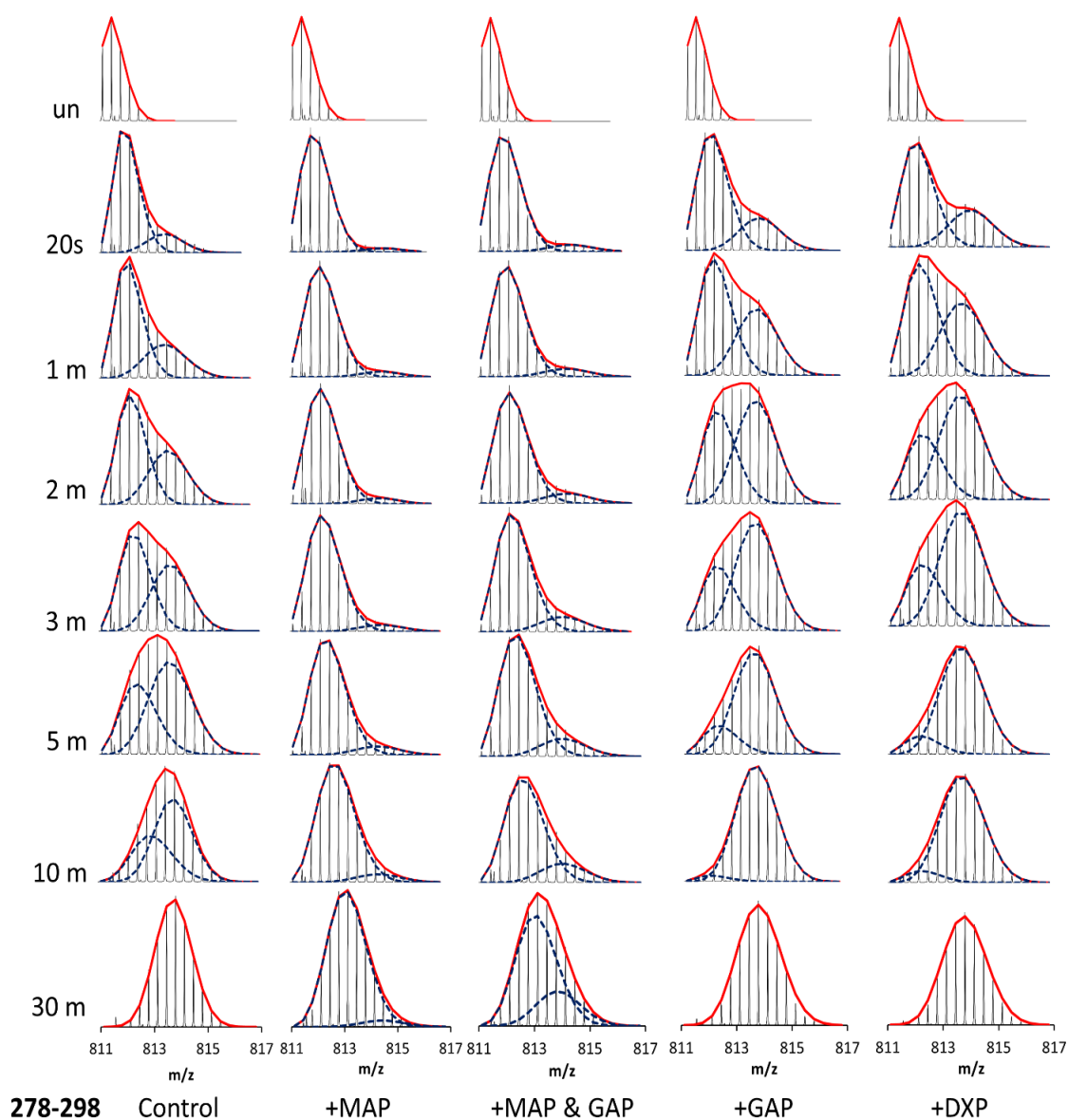


Figure 5.9: Complete time course of HDX-MS spectra of peptides 278-298 in its ThDP-bound and ligand-bound states. MS isotopic envelopes reveal the time-dependent HDX EX1 behavior of this peptide. Red lines indicate the peptide charge envelope, and blue lines indicate the binomial distribution of closed and open conformations. MAP, methyl acetylphosphonate.

Careful examination of the three regions displaying EX1 kinetics in the absence of ligand revealed several interesting behaviors. First, these three regions all adopted mixed EX1/EX2 kinetics. In addition to the distinct bimodal EX1 signature described above, the lower mass envelope of each peptide gradually moved to the high mass range over time (see Figs. 5.5-5.9). As these peptides are long, they may contain a portion that has undergone exchange via the EX1 mechanism 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 the 20 s time point indicated a 4 Da difference in peptide 42-56 (30.8% of residues), 8 Da in peptide 183-199 (53.5% of residues) and 6 Da difference in peptide 278-298 (35.3% of the residues) being involved in the unfolding event (see Figs. 5.1).

5.3.4 DXPS binds and responds differently to substrate, substrate analogue and product

MAP is a stable mimic of Py and known inhibitor of ThDP-dependent enzymes that use Py as substrate [13]. Upon binding of MAP to DXPS, a stable C2 α -phosphonolactylthiamin diphosphate (PLThDP) intermediate is formed and observed by CD. Addition of MAP (1 mM), and formation of the pre-decarboxylation intermediate mimic on DXPS prior to H/D exchange significantly slows down the unfolding of all three EX1 signature regions as mentioned above (residues 42-58, 183-199, and 278-298). The slowdown is apparent in the HDX-MS spectra (Fig. 5.5-5.9), wherein a bimodal distribution is observed but with a greater contribution from the lower mass envelope relative to DXPS in the absence of ligand. This phenomenon can be quantified by the “slowdown factor” (24), which is calculated using the rate constant (k_u) or half-life ($t_{1/2}$) for unfolding. The slowdown factor is simply the $t_{1/2}$ of unfolding in the presence of ligand divided by the $t_{1/2}$ of unfolding in the absence of ligand. A summary of the unfolding rate constants, half-lives and slowdown factors of each of the peptides exhibiting an EX1 signature is presented in Table 5.1. It is clear that MAP gives rise to the most pronounced slowdown factor. A likely explanation for this observation is that the formation of the stable PLThDP intermediate in the presence of MAP [31, 32], effectively traps the enzyme in a pre-decarboxylation-like state that exists in the closed conformation (illustrated by the lower mass envelope in the EX1 kinetic profiles). As the PLThDP intermediate is a mimic of LThDP formed upon binding of pyruvate [13, 31, 32], it is conceivable that the pre-decarboxylation LThDP intermediate could also exist in the closed conformation. This is a key finding that warrants further investigation as it is currently unknown how DXPS stabilizes LThDP in the active site. The mass envelope distribution difference in peptides displaying EX1 kinetics, suggests that binding of MAP could induce large-scale conformational changes around the active

site of DXPS, and favor the closed conformation. Such conformational changes induced by this substrate analog give insights into the mechanism of inhibition by the alkylacetylphosphonates inhibitor class based on this substrate analog which has significant implications for future inhibitor design.

The HDX behaviors of DXPS in the presence of GAP, GAP + MAP, or DXP were also examined. Upon binding of GAP or DXP to DXPS, the regions 42-58, 183-199, and 278-298 display unique unfolding events as shown in Fig. 5.5-5.9. The increased abundance of the higher mass envelope of the bimodal distribution suggests that both GAP and DXP can induce conformational changes in the regions displaying EX1 behavior, resulting in the formation of an open conformation. As shown in Table 2, the half-life of unfolding is slightly decreased on all peptides compared to DXPS in the absence of ligand; however, these rates of unfolding are still significantly greater than those observed in the presence of MAP. We next examined the role of GAP to induce a conformational rearrangement of DXPS subsequent to MAP binding, to model events in GAP-induced LThDP decarboxylation. The HDX-MS behavior of DXPS with MAP bound was compared in the absence and presence of GAP. All peptides consistently exhibited EX1 kinetics profiles, as induced by MAP, even in the presence of GAP. However, we observed a slight increase in the isotopic distribution favoring the higher mass envelope of all EX1 kinetics peptides in the presence of GAP compared to DXPS in the presence of MAP. (Fig. 5.5-5.9). Accordingly, the rates of unfolding of the three EX1 kinetics peptides were found to be about 2 to 3 times faster than with MAP alone (Table 5.1).

Table 5.1 Rate constant of unfolding and the unfolding half-life for the four peptides displaying EX1 kinetics.

Peptide	42-56			51-58		
	k (min ⁻¹)	t _{1/2} (min)	slowdown factor	k (min ⁻¹)	t _{1/2} (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 ⁻¹)	t _{1/2} (min)	slowdown factor	k (min ⁻¹)	t _{1/2} (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

Importantly, these results provide additional evidence for the unique ability of GAP to trigger conformation changes on DXPS both in the absence and presence of a donor substrate mimic. This is consistent with a model (Fig. 5.3) in which LThDP is stabilized in the closed conformation prior to binding of GAP. Upon binding of GAP to the LThDP-DXPS complex (mimicked by the PLThDP-DPXS complex), DXPS undergoes an unfolding event, implying the possibility that GAP induces an open conformation to trigger decarboxylation of LThDP. Further, an unfolded conformation is favored on DXPS in the presence of the product, consistent with an open state during product release. Taken together, the results suggest that the 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 picture of the dynamics of the DXPS structure in response to the binding of different ligands that is the most detailed to date. The high sequence coverage of DXPS in our experiments (93.7%) enabled us to study the flexibility of the entire enzyme, including the two regions (residues 183-238, and 292-317), which have never been observed by X-ray crystallography.

HDX-MS data showed that domain I of DXPS (residues 1-319) displays unusual conformational flexibility that appears to play important roles in substrate recognition. Three regions (residues 42-58, 183-199, and 278-298) near the active center displayed an EX1 signature in both ligand-free and ligand-associated states (Fig. 5.2), and the half-lives of unfolding of these peptides were different depending upon the identity of the ligand-DXPS complex (Table 5.1). In the presence of MAP, DXPS appears to favor a closed conformation in those EX1-displaying peptides, and favors the unfolded conformation in the presence of D-GAP or DXP. Thus, we propose that the closed conformation around the active center of DXPS is critical for stabilization of the LThDP-DXPS complex, while the open conformation may play a role in the GAP-triggered decarboxylation of LThDP and in product release. Intriguingly, the EX1 exchange kinetics displayed by DXPS appears to be a unique characteristic of this enzyme compared to other enzymes in its class as this behavior has not been reported for any other ThDP-dependent enzyme to date. Although MAP induced-conformational changes have also been observed on the *E. coli* PDH-E1 component [32], the specific dynamics reported in this study in the presence of the PLThDP intermediate on DXPS are novel. These intriguing findings support a distinct mechanism of DXPS. The discovery of MAP-induced conformational changes also suggests the

potential utility of this HDX MS method for the characterization of conformational effects of inhibitor binding to DXPS. Inhibitors that are selective for DXPS over other ThDP-dependent enzymes could induce unique conformational changes on this enzyme resulting in stabilization of a native intermediate conformation or by promoting an unproductive form of the enzyme. Given the difficulty of investigating DXPS structure via traditional methods, HDX MS may be useful for the characterization of DXPS inhibitor mechanisms toward the design of more selective and potent compounds.

This study has marked an important step forward in our understanding of the novel mechanism of DXPS. Our findings reveal that several regions in domain I of DXPS (residues 42-58, 183-199, and 278-298), some of which have not been previously characterized structurally, play an important role in substrate recognition and enzymatic catalysis. Given that the unique H/D exchange behavior and conformational flexibility of DXPS appear to be distinct characteristics of this enzyme, it is possible that these dynamics are major factors underlying the novel mechanism of DXPS. These findings have revealed a path for further investigation of the structural and functional roles of the regions of DXPS displaying EX1 behavior in order to gain a deeper understanding of DXPS catalysis. Moreover, a better understanding of the structure and dynamics around the DXPS active center may also lead to design of more potent and selective inhibitors of this important enzyme target in bacterial metabolism. Importantly, the work presented here illustrates the value of HDX-MS methods as a useful tool for mapping conformational changes of flexible regions in DXPS, which are so far undetectable by traditional crystallographic methods.

Thus, this study serves as a model for the investigation of other relatively large, highly dynamic protein structures, which are difficult to capture by X-ray or NMR methods.

Table 5.2 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	LRLPKESLPKL
948.5634	948.5625	1.0	35	41	LRRYLLD
650.3108	650.3104	0.6	41	46	DSVSR
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	LTVALHYVYNTPFDQ
1396.6553	1396.6532	1.5	64	74	HYVYNTPFDQL
4636.4686	4636.4549	3.0	75	114	IWDVGHQAYPHKILTGRDRKIGTIRQKGGHLPFPWRGESE
1259.6217	1259.6226	-0.8	119	132	SVGHSSTISAGIG
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	MSISENVGALNNHQAQL
1263.6801	1263.6804	-0.3	188	199	NVGALNNHQAQL
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	YSSLREGGKKVFSGVPPIKE
1727.9805	1727.9803	0.1	209	224	REGGKKVFSGVPPIKE
2021.1577	2021.1576	0.1	225	242	LLKRTEHIKGMVVPGL
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	YTGYYHNDGPSA
1187.6657	1187.6644	1.1	474	484	VRYPRGNAVGV
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	MAHRKPPVPL
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

References

- [1] Rodriguez-Concepcion M. (2014) The MEP pathway: a new target for the development of herbicides, antibiotics, and antimalarial drugs. *Curr Pharm Des.* 10, 2391-2400.
- [2] Du Q, Wang H, and Xie J. (2011) Thiamin (Vitamin B₁) biosynthesis and regulation: a rich source of anti-microbial drug targets? *Int J of Biol Sci.* 7(10), 41-52.
- [3] Laber B, Maurer W, Scharf S, Stepusin K, and Schmidt FS. (1999) Vitamin B₆ biosynthesis: formation of pyridoxine 5'-phosphate from 4-(phosphohydroxy)-L-threonine and 1-deoxy-D-xylulose-5-phosphosphate by PdxA and PdxJ protein. *FEBS Lett.* 449, 45-8.
- [4] Matsue Y, Mizuno H, Tomita T, Asami T, Nishiyama M, and Kuzuyama T. (2010) The herbicide ketoclofazone inhibits 1-deoxy-D-xylulose 5-phosphate synthase in the 2-C-methyl-D-erythritol 4-phosphate pathway and shows antibacterial activity against *Haemophilus influenzae*. *J Antibiot.* 63(10), 583–588.
- [5] Smith JM, Vierling RJ, and Freel Meyers C. (2012) Selective inhibition of *E. coli* 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *MedChemComm.* 3(65), 65-7.
- [6] Hayashi D, Kato N, Kuzuyama T, Sato Y, and Ohkanda J. (2013) Antimicrobial *N*-(2-chlorobenzyl)-substituted hydroxamate is an inhibitor of 1-deoxy-D-xylulose 5-phosphate synthase. *ChemComm.* 49, 5535–5537.
- [7] Witschel M, Rühl F, Niggeweg R, and Newton T. (2013) In search of new herbicidal inhibitors of the non-mevalonate pathway. *Pest Manage Sci.* 69(5), 559–563.

- [8] Masini T, Pilger J, Kroezen BS, Illarionov B, Lottmann P, Fischer M, Griesinger C, and Hirsch AKH. (2014) *De novo* fragment-based design of inhibitors of DXS guided by spin-diffusion-based NMR spectroscopy. *Chem Sci.* 5, 3543–3551.
- [9] Morris, F., Vierling, R.J., Boucher, L., Bosch, J., and Freel Meyers, C. L. (2013) DXP synthase-catalyzed C-N bond formation: Nitroso substrate specificity studies guide selective inhibitor design. *ChemBioChem.* 14, 1309-1315.
- [10] Smith JM, Warrington NV, Vierling RJ, Khum ML, Anderson WF, Koppisch AT, and Freel Meyers CL. (2014) Targeting DXP synthase in human pathogens: enzyme inhibition and antimicrobial activity of butylacetylphosphonate. *J Antibiotics.* 67, 77-83.
- [11] Bartee D, Morris F, Al-khouja A, and Freel Meyers CL. (2015) Hydroxybenzaldoximes are D-GAP-competitive inhibitors of *E. coli* 1-deoxy-D-xylulose 5-phosphate synthase. *ChemBioChem.* 16, 1771-81.
- [12] Eubanks LM and Poulter CD. (2003) *Rhodobacter capsulatus* 1-deoxy-D-xylulose 5-phosphate synthase: Steady-state kinetics and substrate binding. *Biochemistry.* 42(4), 1140-9.
- [13] Brammer LA, Smith JM, Wade H, and Meyers CF. (2011) 1-deoxy-D-xylulose-5-phosphate synthase catalyzes a novel random sequential mechanism. *J Biol Chem.* 286(42), 36522-31.
- [14] Patel H, Natalia NS, Brammer LA, Freel Meyers CL, 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. *JACS.* 134, 18374-79.

- [15] Brammer Basta LA, Patel H, Kakalis L, Jordan F, and Freel Meyers CL. (2014) Defining critical residues for substrate binding to 1-deoxy-D-xylulose-5-phosphate synthase- active site substitutions stabilize the predecarboxylation intermediate C2 α -lactylthiamin diphosphate. *FEBS J.* 281(12), 2820-37.
- [16] 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.
- [17] Xiang S, Usunow G, Lange G, Busch M, Tong L. (2007) Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *J Biol Chem.* 282, 2676-82
- [18] Wille G, Meyer D, Steinmetz A, Hinze E, Golbik R, and Tittmann K. (2006) The catalytic cycle of a thiamin diphosphate enzyme examined by cocrystallography. *Nat Chem Biol.* 2(6), 324-8.
- [19] Meyer D, Neumann P, Parthier C, Friedemann R, Nemeria N, Jordan F, and Tittmann K. (2010) Double duty for a conserved glutamate in pyruvate decarboxylase: evidence of the participation in stereoelectronically controlled decarboxylation and in protonation of the nascent carbanion/enamine intermediate. *Biochemistry.* 49(37), 8197-8212.
- [20] Fiedler E, Throell S, Sandalova T, Golbik R, Konig S, and Schneider G. (2002) Snapshot of a key intermediate in enzymatic thiamin catalysis: crystal structure of the α -carbanion of (α,β -dihydroxyethyl)-thiamin diphosphate in the active site of transketolase from *Saccharomyces cerevisiae*. *Biochemistry.* 99(2), 591-5.
- [21] Nemeria NS, Shome B, DeColli AA, Heflin K, Begley TP, Meyers CF, Jordan F. (2016) Competence of thiamin diphosphate-dependent enzymes with 2'-methoxythiamin

diphosphate derived from Bacimethrin, a naturally occurring thiamin anti-vitamin. *Biochemistry* 55(7), 1135-48.

[22] Weis DD, Wales TE, Engen JR, Hotchko M, Ten Eyck LF. (2006) Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J Am Soc Mass Spectrom.* 17(11): 1498-509.20

[23] 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.

[24] Woodward, C., Simon, I., and Tuchsén, E. (1982) Hydrogen exchange and the dynamic structure of proteins. *Mol. Cell. Biochem.* 48, 135-160.

[25] Englander, S.W., and Kallenbach, N.R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* 16, 521-655.

[26] Chance, M. (2008) *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics*. Hoboken, New Jersey: John Wiley & Sons, Inc.

[27] 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.

[28] 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.

- [29] 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
- [30] 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.
- [31] 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.
- [32] Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5, 725-738.
- [33] Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12.

Appendix A

A.1. Development of a bifunctional cross-linker to cross link E1p with lipoylated E2p-derived proteins

Crystal structure determination is always the most direct way to see how two proteins interact with each other. However, there is still no crystal structure of any E1p-LD complexes. We hypothesized that covalent linking of E1p with lipoyl domains may assist the crystallization, to finally define the interaction between the E1p and E2p components.

We've already shown that trivalent arsenicals can form high-affinity ring structures with the dihydrolipoyl group. E1p can be nearly fully phosphorylated by PDK2 in situ. Using ATP- γ -S instead of ATP, the thiophosphoryl group could be introduced onto E1p, containing a reactive thiol nucleophile for further derivatization.

Bromoacetylaminophenyl arsenoxide (BRAO) [1] is a bifunctional linker that has a trivalent arsenoxide at one end and an acetylbromide that could selectively react with the thiol group on the other end. A mixture of E1p and LD would be incubated, followed by adding BRAO to initiate alkylation of thiol of the thiophosphoryl group on E1p. Then

add TCEP to reduce the lipoamide of LDs to dihydrolipoamide, which can react with the trivalent arsenoxide group, resulting in crosslinked E1p-LD.

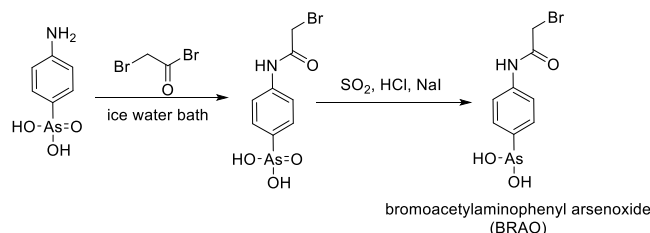


Figure A.1.1: Synthetic procedure of BRAO

The route to synthesize BRAO is shown above. p-Arsanilic acid (10 g) was added in discrete portions to a well stirred solution of sodium carbonate (30 g) in water (150 mL). Bromoacetyl bromide (10 mL) was diluted with CH_2Cl_2 (50 mL). Then the acylating solution was added into the alkaline arsenate solution at a rate of 2 drops per second. After stirring at ambient temperature for 2 h, the organic layer was separated and discarded. The aqueous solution was carefully acidified with 98% (w/w) sulfuric acid to pH 4 forming a white precipitate that was collected by suction filtration. The white precipitate (10 g) was suspended in 75 mL of methanol with constant stirring in a two-necked 250 mL round bottomed flask and 75 mL of HBr (48%) was added. Addition of a small amount of NaI turned the solution brown immediately. SO_2 was bubbled into the mixture for 2 to 3 h, yielding a white precipitate which was the final product BRAO.

LC-MS was used to check the formation of cross-linked product. 20 ng of cross-linked protein was denatured by adding 100 μL of 8 M urea, incubating at 37 $^\circ\text{C}$ for 30 min. Then TCEP was added to a final conc. of 6 mM, and the mixture was incubated at 37 $^\circ\text{C}$ for another 30 min. Next idoacetamide was added to a final conc. of 15 mM, the mixture

was incubated in the dark for 30 min, then the same amount of DTT solution (final conc. of 15 mM) was added to quench the reaction. Finally, 55 mM ammonium bicarbonate was added to dilute urea to 1 M, then add 70 ng (35 times of protein) trypsin to digest the sample at 37 °C overnight. The sample was desalted by a Sep-Pak tC18 cartridge. A sample was dissolved in 30 µL of 5% acetonitrile solution and LC-MS was run on a Phenomenex Luna C18 column.

The LC-MS results are shown below. Only a tiny amount of cross-linked peptide was identified.

To make the method usable for biophysical studies, the method would have to be modified/optimized.

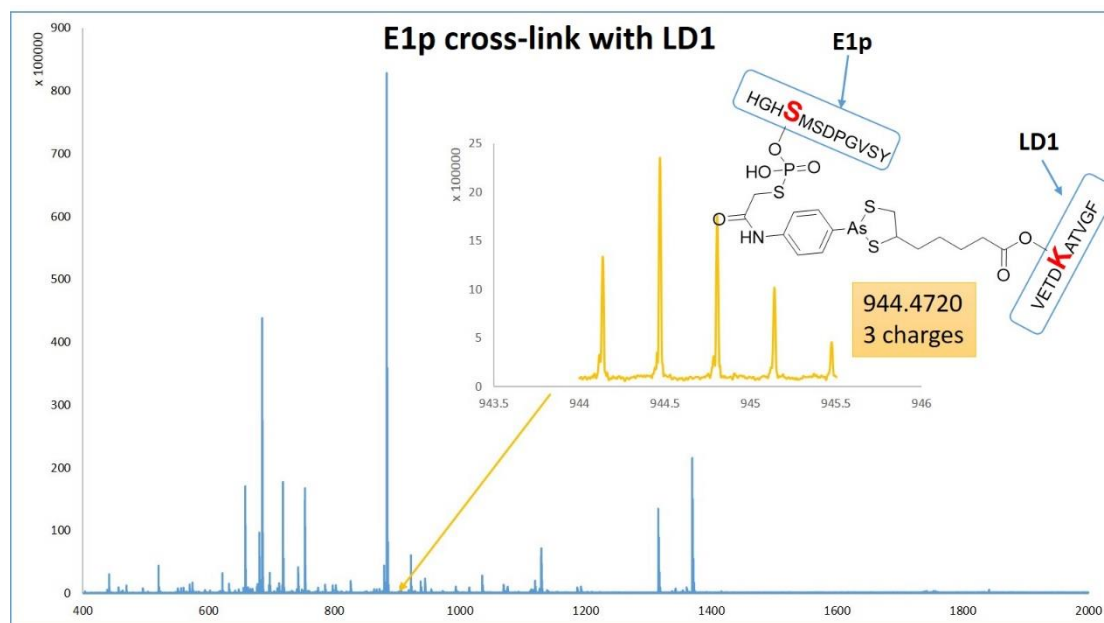


Figure A.1.2: LC-MS of cross-linked E1p-LD1 after pepsin digestion.

The reason why the efficiency of this reaction is so low is still unknown. Maybe the cross-linker would change the conformation of protein, bury the other functional end into pockets and make the other reactive end unavailable.

A.2. ^{19}F NMR studies of loop dynamics in E1p

The fluorine nucleus is a spin $-1/2$ species, which exists in 100% natural abundance and possesses a magnetogyric ratio that is 83% that of proton [10]. The large magnetogyric ratio translates into both high sensitivity in ^{19}F NMR spectroscopy, and strong dipolar couplings. Also, ^{19}F NMR possesses the advantage of inherent sensitivity of the chemical shift to the local environment. The fluorine chemical shift is primarily influenced by a large paramagnetic term, originating from an unpaired valence electron making it highly sensitive to local van der Waals interactions and electrostatic fields. [2] That's why ^{19}F NMR studies of proteins is so powerful in providing unique insight into biologically relevant phenomena such as conformational fluctuations, folding and unfolding, binding and catalysis. [3-4]

Another advantage of ^{19}F NMR is that there is no fluorine found in natural proteins, as a result, a protein containing a ^{19}F -label will exhibit NMR peaks due only to the label. By using a site specific labeling method, we can incorporate a fluorine probe into the active center, and study protein conformational changes.

The phenomenon that phosphorylation on specific serine residues of the $^h\text{E1p}$ α subunit can inactivate PDC was discovered in the late 1960s. [8] The three phosphorylation sites are Ser264: site 1; Ser271: site 2; and Ser203: site 3. Based on crystal structures of human E1p, site 1 and site 2 are located in the same loop A (Ph-loop A, residues 259 to

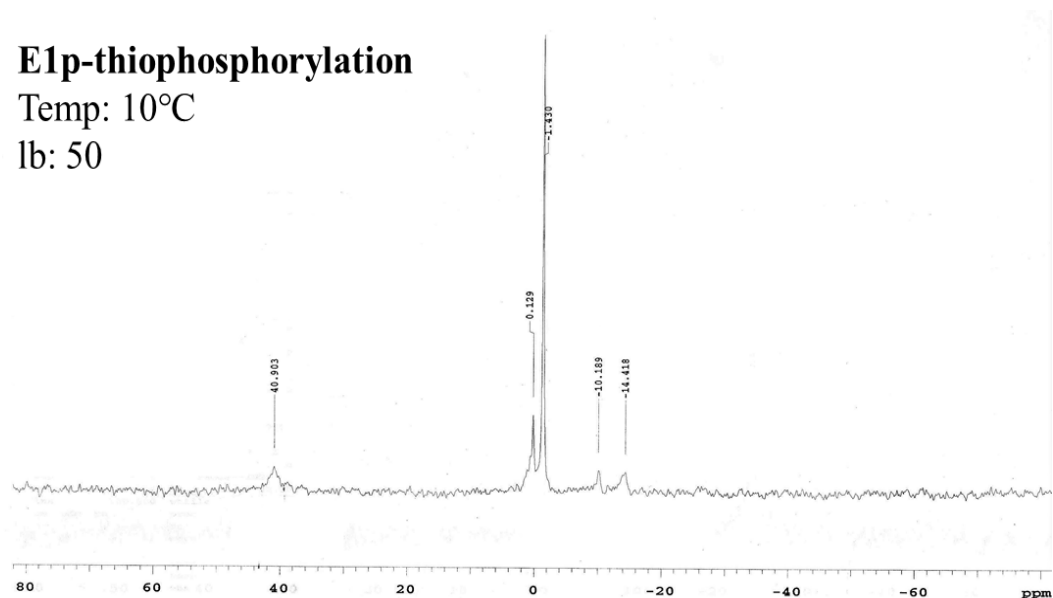
282) and site 3 is located in a relatively short loop called loop B (Ph-loop B, residues 198 to 205). [5-6]

It has been suggested that loop A is responsible for forming the Elp active-site channel and helping anchor the cofactor ThDP to the active site, while loop B provides coordination to a Mg^{2+} ion chelated by the diphosphate group of ThDP. In other words, the binding of ThDP co-factor needs ordered structure of both loops. [9]

A E1p-thiophosphorylation

Temp: 10°C

lb: 50

**B E1p-thio-MTSL**

Temp: 10°C

lb: 50

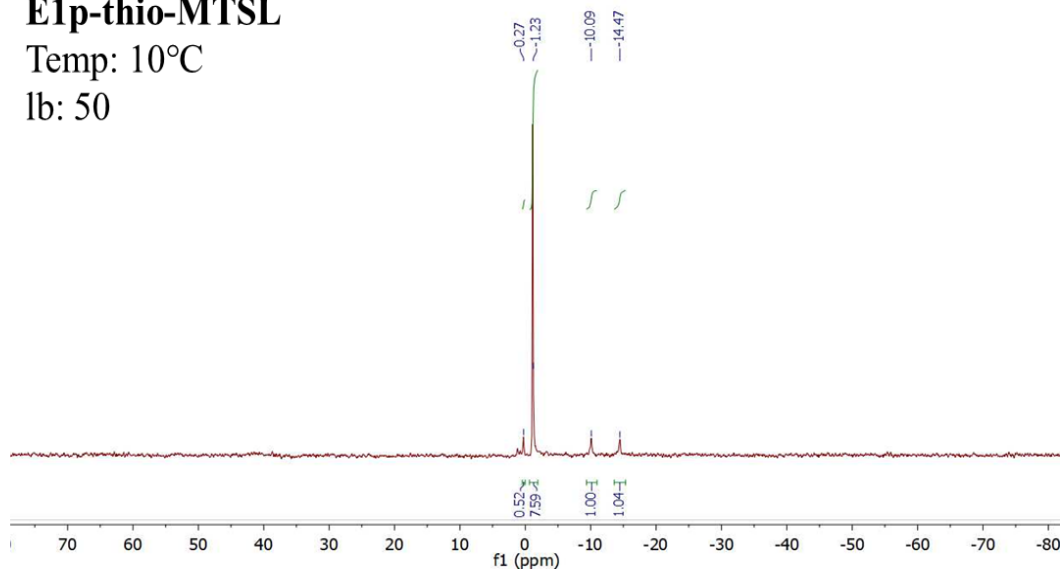


Figure A.2.1: ^{31}P NMR results of ThDP bounded phos-S-E1p. All experiments were conducted on a Bruker 600 NMR at 10 °C overnight. The line broadening of all spectra are 50. (A). NMR spectra of thiophosphorylated wild-type E1p. The broad peak located on 40.893 ppm is from P directed linked to the thio group. The peaks of -10.199 and -14.428 ppm originate from ThDP's α -P and β -P, respectively. (B). NMR spectra of MTSL labeled thiophosphorylated wild-type E1p. The only difference between the two

spectra is that the broad peak located on 40 ppm disappeared, since the unpaired electron from MTSL radical will negatively impact protein peak via paramagnetic relaxation and therefore decrease resolution.

According to the crystal structures of S1-E1p (containing Mg-ThDP) and phospho-S1-E1p (containing Mn-ThDP) determined by Chuang's group in 2008, incorporation of phosphoryl group will disrupt the conformations of both Ph-loop A and Ph-loop B. Phosphorylation on Ser264 can pose a steric clash, which eliminates the essential hydrogen-bonding network that is very important for stabilizing the loop conformation. [9] Consequently, the order-to-disorder transition of the two phosphorylation loops in E1p led to disruption of substrate channeling in PDC, which would inactivate the E1p.

It was already shown in the Jordan group by Sachin Kale that in *Escherichia coli* pyruvate dehydrogenase complex, the loop dynamics of ^{13}C -E1p active center promotes the rate of covalent addition of substrate to the enzyme bound ThDP. [7] It's reasonable to hypothesize whether the human E1p loop serves the same role.

As there are 12 cysteines in the ^1H -E1p α subunit, using cysteine thio-group directed labeling would lead to ambiguous results. Instead, we have at our disposal three different E1p variants: E1p-MS 2,3, in which only site 1 is available for phosphorylation; E1p-MS 1,3, which only site 2 is available; and E1p-MS 1,2, in which only site 3 is available. We would therefore use ATP- γ -S, an analog of ATP, to introduce a thiophosphoryl group into the Ph-loop of a single Ser loop variant, then incorporate an ^{19}F NMR probe via a trifluoroacetyl group onto the more reactive thiophosphoryl group to study the dynamics of the loops.

Materials and Methods

- (iv) Three ^hE1p variants, E1p-MS2,3 (S264, [S271A/S203A]), E1p-MS1,3 (S271, [S264A/S203A]), E1p-MS1,2 (S203, [S264A/S271A]). The leading serine is the one available for phosphorylation, and those in the bracket are unavailable.
- (v) The E1p variant (50 μ M) was incubated with 5 μ M of PDK1 at room temperature for 10 min. Then ATP- γ -S was added to a final concentration of 500 μ M to initiate the thiophosphorylation reaction and the reaction was carried out at 4 °C for overnight. Next, excess reagents were removed by dialysis against the sample buffer with a centrifugal filter unit (Vivaspin 500, 30K MWCO). 20 ng sample was saved for trypsin digestion, then LC-MS was run to analyze the data.
- (vi) Next, 50 μ M of thiophosphorylated E1p variant was treated with 150 μ M of trifluoroacetyl bromide at room temperature for 2 h. Excess reagent was removed by buffer exchange with a centrifugal filter unit (Vivaspin 500, 30K MWCO). 20 ng sample was saved for trypsin digestion, then LC-MS was run to analyze the data.

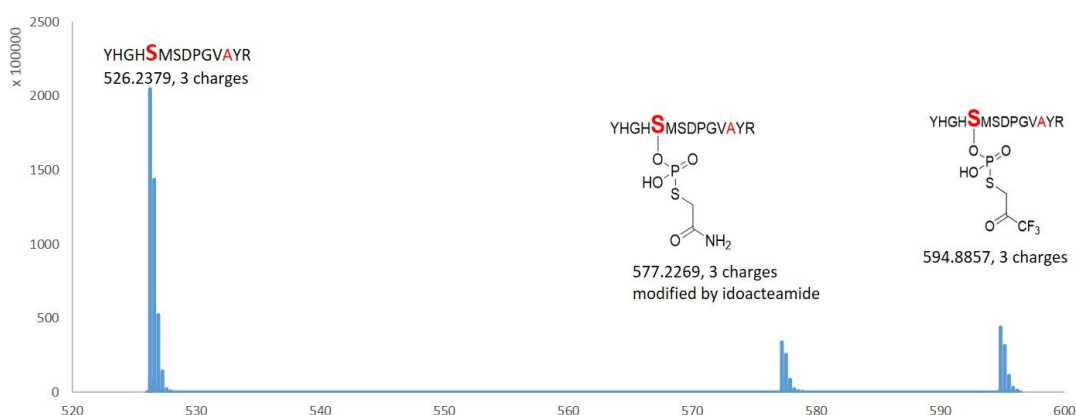


Figure A.2.2: LC-MS result after trypsin digestion of modified E1p-MS 2, 3 sample.

20 ng of protein were denatured by adding 100 μ L of 8 M urea, incubating at 37 °C for 30

min. Then TCEP was added to a final concentration of 6 mM, and the mixture was incubated for another 30 min at 37 °C. Next, iodoacetamide was added to a final conc. of 15 mM, the mixture was incubated in the dark for 30 min, then the same amount of DTT solution (final conc. of 15 mM) was added to quench the reaction. Finally, 55 mM ammonium bicarbonate was added to dilute the urea to 1 M, then trypsin was added (70 ng, 35 times the concentration of protein) to digest the sample at 37 °C overnight. The sample was desalted by a Sep-Pak tC18 cartridge. The sample was dissolved in 30 µL of 5% acetonitrile solution and the LC-MS was run on a Phenomenex Luna 5u C18 column.

Table A.2.1: Peptides after trypsin digestion for the detection of E1p modification

Sequence	Modification	Chemical Composition	Most Abundant Isotopic Mass			Error
			[M + nH] ⁿ⁺ (Da)			(ppm)
			Charges	Experimental	Theoretical	
²⁶⁰ YHGHSMSDPGVAYR ²⁷³	None	C ₆₈ H ₉₇ N ₂₁ O ₂₁ S	2+	788.8543	788.8517	3.3
	Thiophosphorylation (iodoacetamide)	C ₆₈ H ₉₇ N ₂₁ O ₂₁ SPO ₂ SCH ₂ CONH ₂	2+	865.3373	865.3342	3.6
	Bromo-trifluoroacetone	C ₆₈ H ₉₇ N ₂₁ O ₂₁ SPO ₂ SCH ₂ COCF ₃	2+	891.8257	891.8224	3.7

So far, we used LC-MS to prove that it's possible to attach the trifluoroacetyl group to the E1p variant with, an efficiency of labeling of approximately 30%. We are going to work on improving the efficiency of labeling and prepare a larger scale sample for ¹⁹F NMR.

A.3. EPR studies of loop dynamics in E1p

Site-directed spin labelling (SDSL) combined with electron paramagnetic resonance (EPR) spectroscopy is a very powerful technique used widely for studying the structural properties and dynamical processes of biological systems. [11-12]

Spin labelling provides a relatively non-intrusive technique to study the structure, dynamics, and conformational changes of biomacromolecules.

By far the most popular nitroxide spin label is MTSL (Figure 10). MTSL is highly selective for thiol groups, and is commonly attached to cysteine residues in proteins. The MTSL side chain (commonly known as R1) also has a minimum impact on the secondary and tertiary structure of proteins.

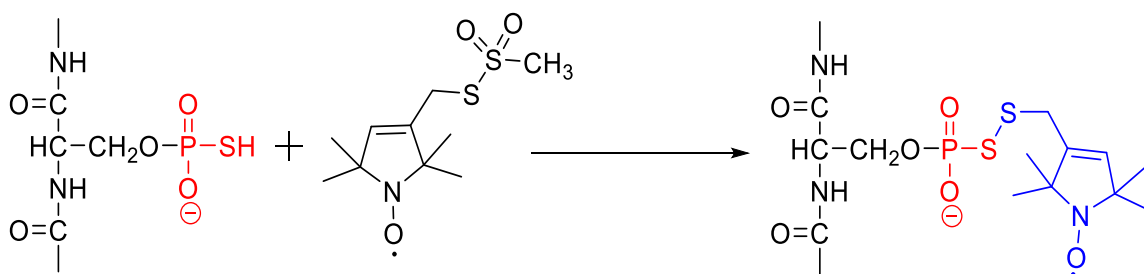


Figure A.3.1: Site-directed labelling of MTSL onto E1p- α subunit.

We also use LC-MS to monitor the modification.

Peptide of E1p-MTSL

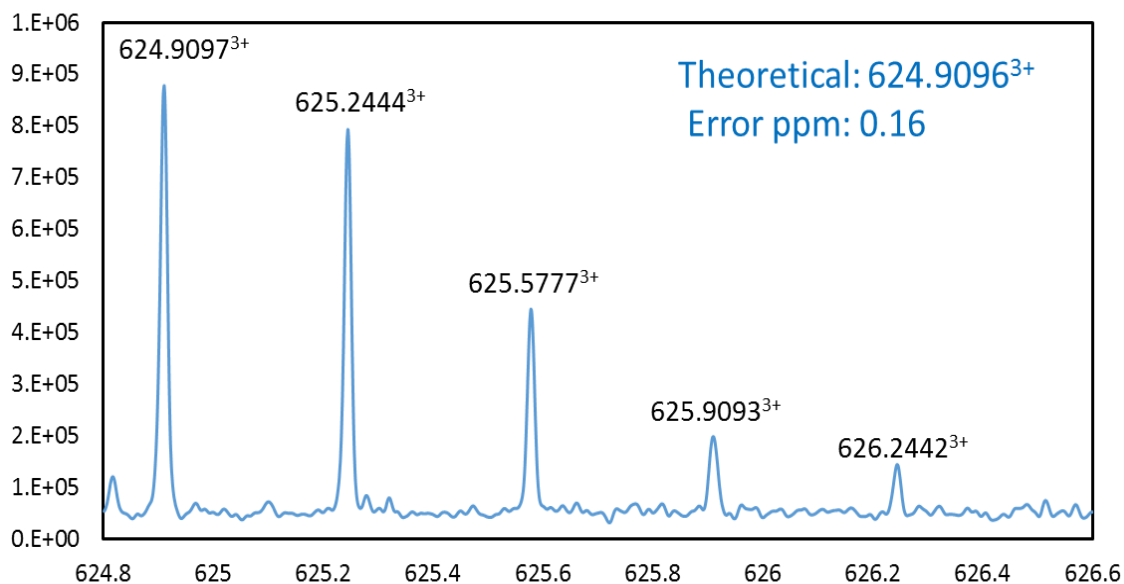


Figure A.3.2: LC-MS peak of “YHGHMSDPGVSYR-MTSL” (C₇₇H₁₁₂N₂₂O₂₅S₃P).

The theoretical mass of this peptide is 624.9096³⁺, the mass detected by FT-MS is 624.9097³⁺.

The EPR spectrum was measured with 7.5 mg/ml (96 µM per subunit) E1p-MTSL sample dissolved in 50 mM Tris (pH=7.5) containing 1 mM MgCl₂, 0.5 mM ThDP.

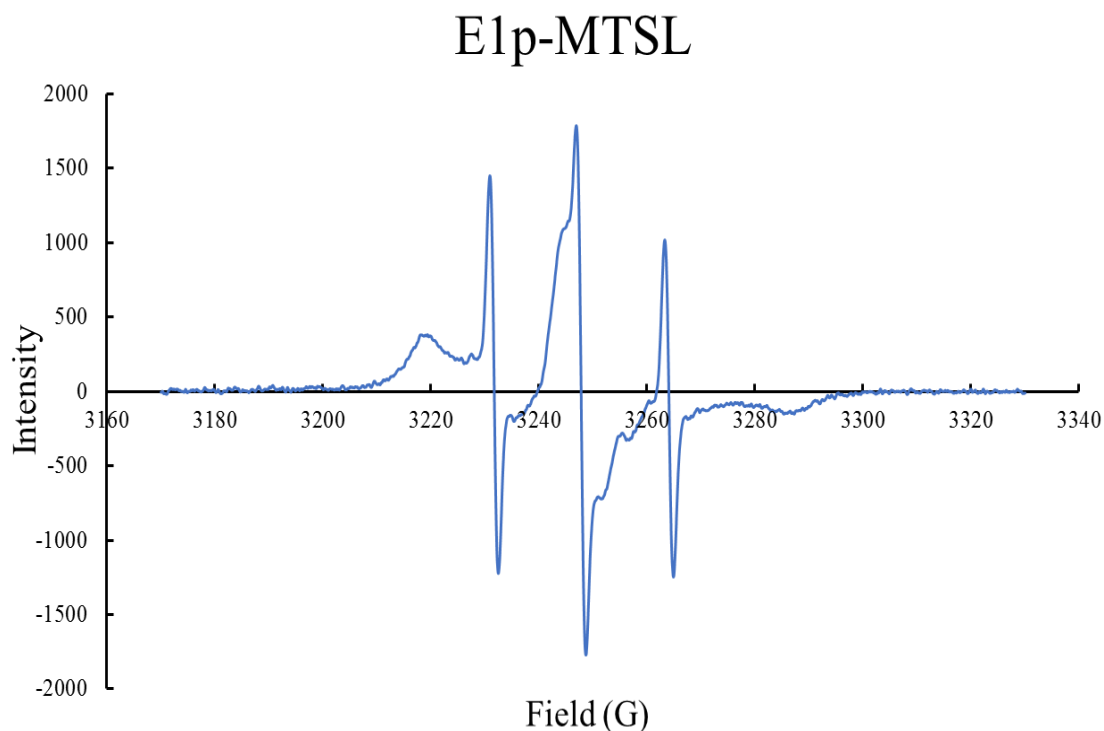


Figure A.3.3: EPR spectrum of E1p-MTSL. There are two forms of the component measured at room temperature. The broad peaks represent the slow component, which is dominant; the sharp peaks corresponds to the fast component, which is less than 5%.

A.4. HDX-MS studies of loop dynamics in E1p

Materials and methods

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

Sample preparation for HDX.

Prior to H/D exchange, the E1p, L1 or L2S protein was exchanged into 50 mM KH_2PO_4 (pH 7.5) with 100 mM NaCl, 0.2 mM ThDP, and 1 mM MgCl_2 , and then the protein concentration was adjusted to 160 μM and 80 μM . The concentrated E1p (80 μM) was incubated with or without pyruvate, methylacetylphosphonate (MAP) at a final concentration of 1 mM for 30 min at 25 °C prior to initiation of the HDX experiments. The concentrated E1p (160 μM) was incubated with equal amount of L1 or L2S (both 160 μM) for 30 min at 25 °C prior to initiation of the HDX experiments. E1p-phos is prepared separately. 50 μM of E1p was incubated with 5 μM of PDK2 at room temperature for 30 min. Then ATP was added to a final concentration of 500 μM to initiate the phosphorylation reaction and the reaction was carried out at 4 °C for overnight. Next day, excess reagents were removed by dialysis against the sample buffer with a centrifugal filter unit (Vivaspin 500, 30K MWCO). 20 ng sample was saved for trypsin digestion, then LC-MS was run to monitor the phosphorylation reaction.

The HDX experiments were initiated by mixing 15 μl of the protein samples with 285 μl of D_2O buffer to yield a final concentration of 95 % D_2O at pH 7.5. D_2O buffer was prepared the same way as E1p exchange buffer except 99.9% D_2O was used to dissolve the buffer components. The samples were incubated at 25 °C for 30 s, 1, 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.

Results and Discussions

The phosphorylation reaction is checked by LC-MS.

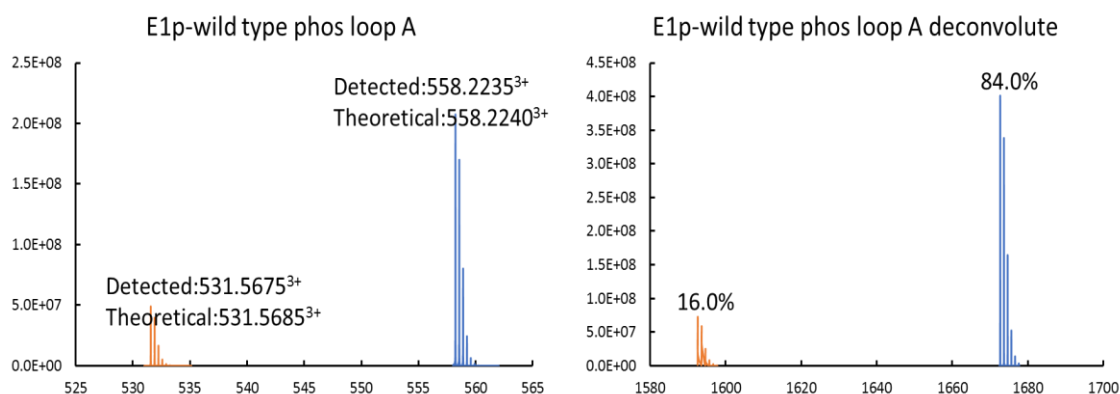


Figure A.4.1: Phosphorylated peptide detected by LC-MS. The peak of unmodified peptide is colored in orange, and the peak of phosphorylated peptide is colored in blue.

The ratio of unmodified peptides and phosphorylated peptides is calculated based on their intensities after deconvolution. 16% peptide remained unmodified according to this estimate.

Based on the intensity of the original peptide and phosphorylated peptides, we can estimate that 84% of the protein that was used in the HDX-MS experiments had been phosphorylated.

Among the 42 peptides we analyzed (24 peptides from E1p- α with 78% sequence coverage and 18 peptides from E1p- β with 89% sequence coverage), only two display EX1 kinetics. Residues 192-205 and 162-177 from E1p- α around the active center

displayed an EX1 signature in both ligand-free and ligand-associated states. The half-life of conformational change between folding and unfolding is dependent on ligands. ThDP-bound E1p without any ligand intrinsically adopts two conformations, which can only be seen within 1 min. MAP, pyruvate and L2S can freeze the folded (closed) conformation of E1p, while L1 and E1p-phos are prone to adopt the unfolded (open) conformation.

Both peptides displaying EX1 kinetics are involved in binding ThDP, peptide of 192-205 even contains the phosphorylation site 3 (Ser 203). It has been suggested that loop A is responsible for forming the E1p active-site channel and helping anchor the cofactor ThDP to the active site, while loop B, which contains site 3 provides coordination to a Mg^{2+} ion chelated by the diphosphate group of ThDP.

According to the crystal structures of wild type E1p (containing Mg-ThDP) and phospho-S1-E1p (containing Mn-ThDP) determined by Chuang's group in 2008, incorporation of phosphoryl group will disrupt the conformations of both Ph-loop A and Ph-loop B.

Phosphorylation on Ser264 can create a steric clash, which eliminates the essential hydrogen-bonding network that is very important for stabilizing the loop conformation. Consequently, the order-to-disorder transition of the two phosphorylation loops in E1p led to disruption of substrate channeling in PDC, which would inactivate the E1p.

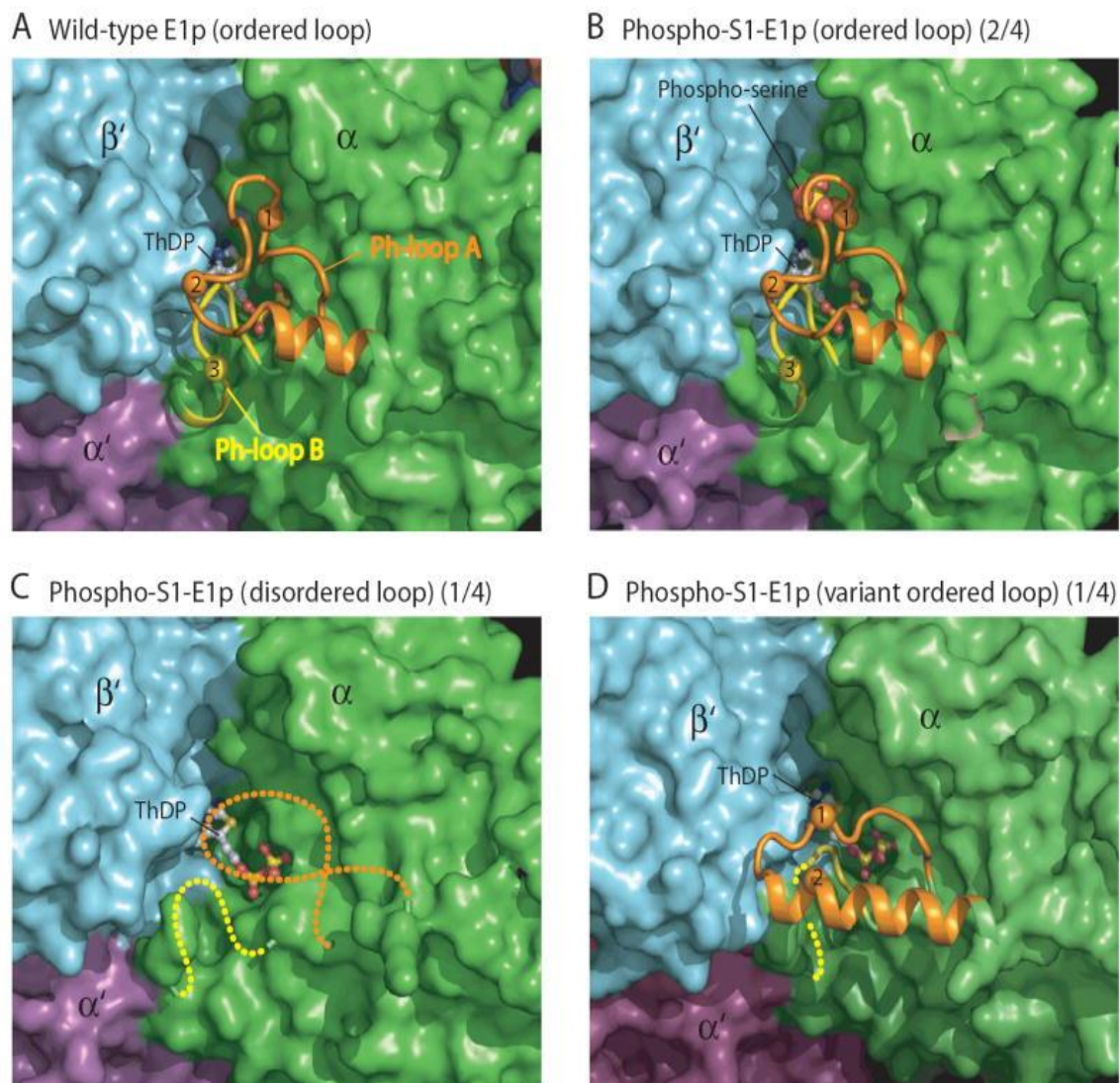


Figure A.4.2: Structures of the Phosphorylation loops in Wild-type and Phospho-S1-E1p. The structures of Ph-loops in E1p- α subunits are shown as ribbon models against other parts of the protein in surface representation. Ph-loop A (residues from 259- α to 282- α) is in orange, and Ph-loop B (from 198- α to 205- α) in yellow. The three phosphorylation sites are indicated by spheres and labeled according to the site numbers. The bound ThDP is shown as a ball-and-stick model. (A) Fully ordered Ph-loops in wild-

type E1p. (B) Wild-type-like ordered Ph-loops in phospho-S1-E1p with bound Mn-ThDP. Two of the four E1p- α subunits in the asymmetric unit (depicted as 2/4 in the figure caption) exhibit this conformation, which is maintained through interactions with a symmetry-related molecule. (C) Completely disordered Ph-loops in phospho-S1-E1p containing the bound Mn-ThDP. One of the four E1p- α subunits in the asymmetric unit (1/4) has this conformation. No symmetry-related molecule is present near the Ph-loops. (D) Variant ordered conformation of Ph-loop A and partially ordered Ph-loop B in the remaining E1p- α subunits (1/4). This conformation is also maintained by interactions with a symmetry-related molecule.

On the basis of the results, one could hypothesize that substrates prefer the folded (closed) state of E1p, while the phosphorylation promotes the E1p to transition from the folded state to unfolded state, which impedes the ability of E1p to bind substrate, as a result, blocking the entire reaction. Inner lipoyl domain (L2S) could help E1p adopt the conformation that is favored by substrate. MAP can form a stable intermediate that stays on the protein, which may partially explain why MAP can lock the folded (closed) conformation much better than any other ligands.

192-205: ICENRYGMGTSVE

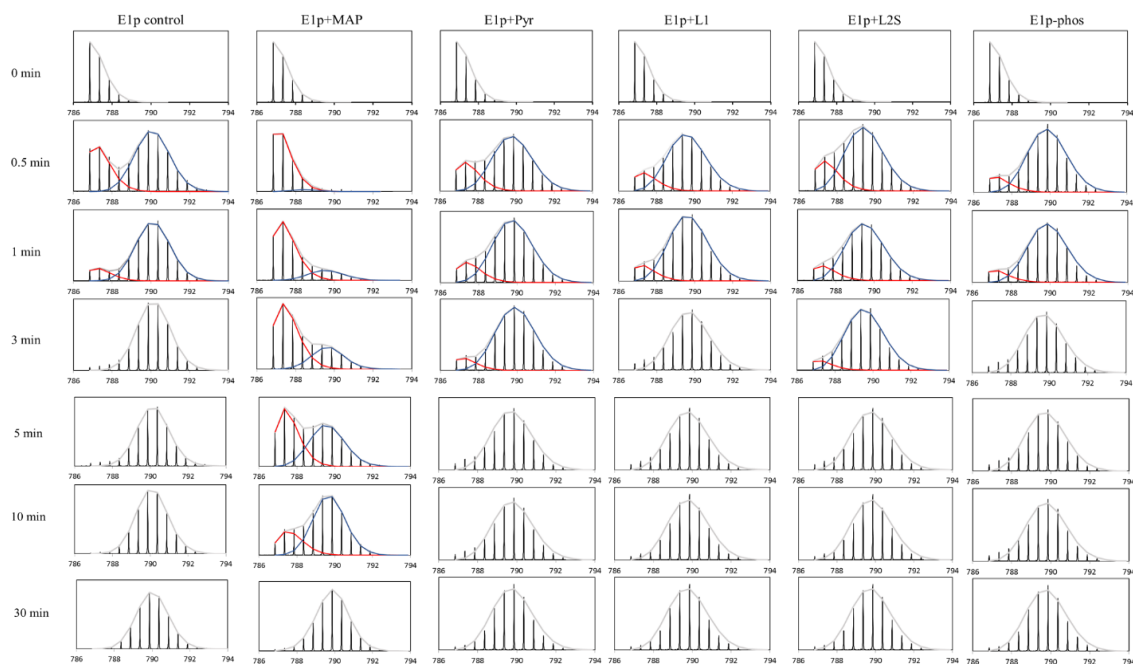


Figure A.4.3: Complete view of HDX-MS spectra of peptide 192-205 of E1p- α in its ThDP-bound state and in other ligand induced states.

162-177: LTLYGDGAANQGQIFE

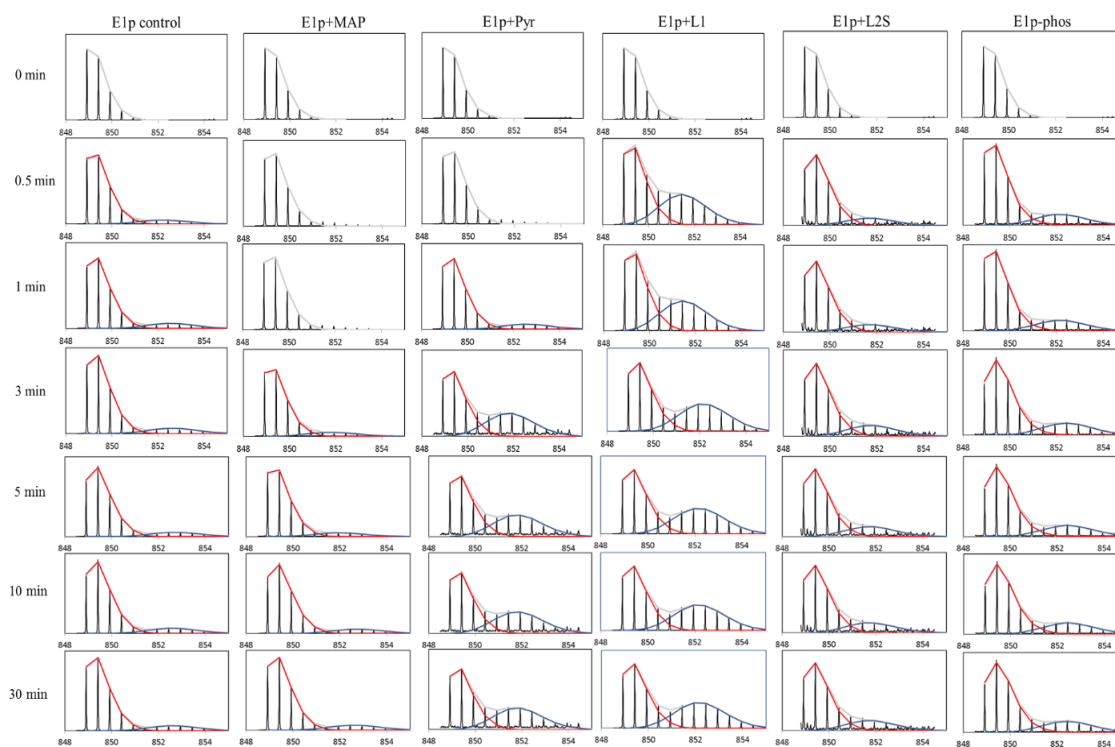


Figure A.4.4: Complete view of HDX-MS spectra of peptide 162-177 of E1p- α in its ThDP-bound state and in other ligand induced states.

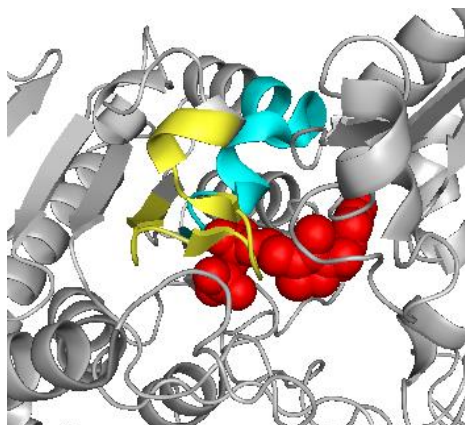


Figure A.4.5: Peptides from E1p- α that show bimodal distribution in HDX-MS.

ThDP is colored by red, peptide 192-205 from E1p- α is colored by yellow, peptide 162-177 from E1p- α is colored by cyan. Both of peptides displaying EX1 kinetics are involved in ThDP binding.

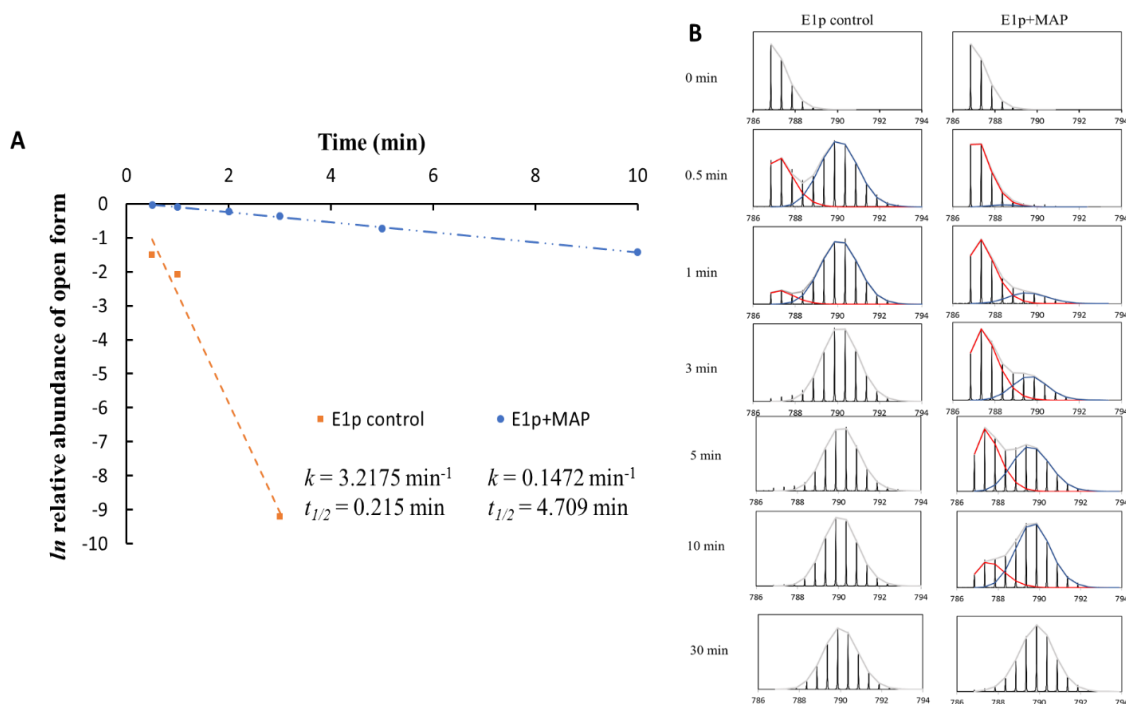


Figure A.4.6: A. Half-life calculation of the unfolded form of E1p- α 192-205. The slope of the line fit to the data yields the first order rate constant, and based on that number is calculated the half-life of unfolding. **B. HDX-MS spectra of peptide 192-205 that display EX1 kinetics in the free state as well as in the presence of MAP,** a clear bimodal isotopic distribution was observed.

References

- [1] Donoghue, N., Yan, P. T. W., Jiang, X.-M., Hogg, P. J. (2000) Presence of closely spaced protein thiols on the surface of mammalian cells. *Protein Science*, 9, 2436-2445
- [2] Gerig J T. Fluorine NMR of proteins[J]. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 1994, 26: 293-370.
- [3] Horng J C, Raleigh D P. Φ -Values beyond the ribosomally encoded amino acids: kinetic and thermodynamic consequences of incorporating trifluoromethyl amino acids in a globular protein[J]. *Journal of the American Chemical Society*, 2003, 125(31): 9286-9287.
- [4] Li H, Frieden C. Fluorine-19 NMR studies on the acid state of the intestinal fatty acid binding protein[J]. *Biochemistry*, 2006, 45(20): 6272-6278.

- [5] Ciszak, E. M., Korotchkina, L. G., Dominiak, P. M., Sidhu, S., and Patel, M. S. (2003) Structural basis for flip-flop action of thiamin pyrophosphate-dependent enzymes revealed by human pyruvate dehydrogenase. *J. Biol. Chem.*, 278, 21240-21246
- [6] Seifert, F., Ciszak, E., Korotchkina, L., Golbik, R., Spinka, M., Dominiak, P., Sidhu, S., Brauer, J. Patel. M. S., and Tittmann, K. (2007) Phosphorylation of serine 264 impedes active site accessibility in the E1 component of the human pyruvate dehydrogenase multienzyme complex. *Biochemistry*, 46, 6277-6287
- [7] Kale S, Ulas G, Song J, Brudvig G W, Furey W, Jordan F. Efficient coupling of catalysis and dynamics in the E1 component of Escherichia coli pyruvate dehydrogenase multienzyme complex[J]. *Proceedings of the National Academy of Sciences*, 2008, 105(4): 1158-1163.
- [8] Linn T C, Pettit F H, Reed L J. α -Keto acid dehydrogenase complexes, X. Regulation of the activity of the pyruvate dehydrogenase complex from beef kidney mitochondria by phosphorylation and dephosphorylation[J]. *Proceedings of the National Academy of Sciences*, 1969, 62(1): 234-241.
- [9] Kato, M., Wynn, R. M., Chuang, J. L., Tso, S.-C., Machius, M., Li, J., and Chuang, D. T. (2008) Structural basis for inactivation of the human pyruvate dehydrogenase complex by phosphorylation: role of disordered phosphorylation loops. *Structure*, **16**, 1849-1859

[10] Kitevski-LeBlanc J L, Prosser R S. Current applications of ^{19}F NMR to studies of protein structure and dynamics[J]. Progress in nuclear magnetic resonance spectroscopy, 2012, 62: 1-33.

[11] Hoff, A.J., Advanced EPR: Applications in Biology and Biochemistry. 1989.

[12] Jeschke, G., Conformational dynamics and distribution of nitroxide spin labels. Progress in

Nuclear Magnetic Resonance Spectroscopy, 2013. 72: p. 42-60.

[13] Woodward, C., Simon, I., and Tuchsén, E. (1982) Hydrogen exchange and the dynamic structure of proteins. Mol. Cell. Biochem. 48, 135-160.

[14] Englander, S.W., and Kallenbach, N.R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. Q. Rev. Biophys. 16, 521-655.

[15] Chance, M. (2008) Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics. Hoboken, New Jersey: John Wiley & Sons, Inc.

Appendix B

Proteins expression and purification

(vi) **Expression and purification of E1p and E1p loop variants.** Cells containing the pET-28b-E1p plasmid will be grown at 37 °C in LB medium supplemented with 35 µg/mL of kanamycin and 100 µg/mL of thiamin chloride. E1p expression will be induced by 0.5 mM IPTG at 25 °C for 16 h. The E1p will be purified using Ni-NTA agarose column equilibrated with 50 mM KH_2PO_4 , 300 mM NaCl and 30 mM imidazole (pH= 7.5). E1p will be eluted from the column with 50 mM KH_2PO_4 ,

300 mM NaCl and 200 mM imidazole (pH= 7.5). Protein will be dialyzed against 50 mM KH_2PO_4 , 300 mM NaCl, 1 mM MgCl_2 , 0.5 mM ThDP, 0.5 mM benzamidine hydrochloride (pH= 7.5) buffer.

- (vii) **Expression and purification of E2p-derived proteins.** The following E2p-derived proteins have been expressed and purified: the outer lipoyl domain LD1, the L2S didomain comprising inner lipoyl domain and subunit-binding domain, L1L2S tridomain, which comprises L1, L2, both hinge regions, and the subunit-binding domain, L1L2Sm1, in which the Lys132 has been converted to alanine, designated as K132A; L1L2Sm2, in which the Lys259 has been converted to alanine, designated as K259A. E2p-derived proteins will be expressed from pET-28b vector. Cells containing the plasmid will be grown at 37 °C in LB medium supplemented with 35 µg/mL of kanamycin and 0.3 mM lipoic acid to ensure full lipoylation of the lipoyl domains. The purification steps are the same as reported for E1p.
- (viii) **Expression and purification of human E2·E3BP protein.** Cells containing the plasmid will be grown at 37 °C in LB medium supplemented with 35 µg/mL of Chloramphenicol and 0.3 mM lipoic acid to ensure full lipoylation of the lipoyl domains. The cells were dissolved in 50-60 ml of sonication buffer [50 mM KH_2PO_4 (pH 7.5), 0.3 M NaCl, 5 mM MgCl_2 , 0.5% (v/v) Triton X100] by gently stirring with a glass rod, then 1 protease inhibitor cocktail tablet was added. Lysozyme was added to 0.6 mg/mL and the solution was incubated on ice for 20 min. Next, 1000 units of each DNase and Nuclease were added and the solution was incubated on ice for another 20-60 min. The solution was next sonicated for 6

min with 10s pulse “on” and 30s pulse “off” program. The lysate was clarified by centrifugation twice for 30 min at $23700 \times g$ at 4°C .

PEG precipitation of the E2·E3BP protein. PEG 8000 (50% w/v) solution was added dropwise to the clarified lysate to 1% (v/v). The precipitated protein was removed by centrifugation at $18\,500 \times g$ for 20 min. Next, 5 % (v/v) of PEG 8000 solution was added dropwise to the clarified supernatant and the mixture was incubated for 15 min. The precipitate was collected by centrifugation at $18500 \times g$ for 20 min. The precipitate was dissolved in approximately 10 ml of buffer A, and clarified by centrifugation. The supernatant was loaded onto a Sephacryl S 300HR column equilibrated with buffer A. The cloudy fractions (4 mL) (caution: *the first two cloudy fractions were not collected because these might contain some aggregated proteins*) were collected, combined and mixed with polyethylenimine (M_w 60,000) to a final concentration of 0.05% and the protein was precipitated by centrifugation at $23700 \times g$ for 30 min. The *E2·E3BP* was pelleted by two-step ultracentrifugation: (i) The impurities and a small amount of *E2·E3BP* were pelleted at $100,000 \times g$ for 4 h, at 4°C . The resulting supernatant was combined and concentrated to approximately 8 ml with a 50 K concentration unit. (ii) The concentrated solution (8 ml) was centrifuged at $250\,000 \times g$ for 4 h at 4°C to pellet the *E2·E3BP*. The pellets were resuspended in a small volume of buffer containing 0.4 M potassium chloride and left overnight. The suspension was clarified by centrifugation and stored at -80°C . The sonication buffer was 50 mM potassium phosphate (pH 7.5), 0.5 mM EDTA, 150 mM KCl, 1 mM β -mercaptoethanol, 0.1% Pluronic-F68. Buffer A was 50 mM potassium phosphate

(pH 7.5), 0.5 mM EDTA, 200 mM potassium chloride, 1 mM β -mercaptoethanol, 0.1% Pluronic-F68.

(ix) Expression and purification of PDKs. Cells are grown on LB medium supplemented with 35 μ g/mL of kanamycin. Protein expression will be induced by 0.5 mM IPTG at 25 °C for 16 h. Dissolve cells in 60 mL 50 mM KH_2PO_4 (pH=7.5) containing 300 mM NaCl, 10 mM imidazole, 1 mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100. Add lysozyme to 0.6 mg/mL and incubate on ice for 30 min. Add 1000 units of DNase and nuclease, incubate on ice for an additional hour. Then, do the sonication. PDK1,2 were purified by using Ni-NTA agarose column equilibrated with 50 mM KH_2PO_4 , 300 mM NaCl, 30 mM imidazole, 1mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100 (pH= 7.5) buffer. PDK3,4 were equilibrated with 50 mM KH_2PO_4 , 300 mM NaCl, 20 mM imidazole, 1 mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100 (pH= 7.5) buffer. PDKs was eluted from the column with 50 mM KH_2PO_4 , 300 mM NaCl and 200 mM imidazole, 1 mM MgCl_2 , 5 mM 2-mercaptoethanol and 0.5% (w/v) Triton X100 (pH= 7.5). Proteins were dialyzed against 50 mM KH_2PO_4 , 300 mM NaCl, 1 mM MgCl_2 , 1 mM DTT, 1 mM benzamidine hydrochloride (pH= 7.5) buffer.

(x) Expression and purification of human E1o. The same method was used to purify E1o, E3, and E2o¹⁴⁴⁻³⁸⁶ (C-terminal didomain). Cells were dissolved in 50-60 mL of sonication buffer while stirring with a glass rod, then 1 protease inhibitor cocktail tablet was added. Lysozyme was added to 1 mg/ml and the solution was incubated on ice for 20 min. Next 1000 units of DNase and Nuclease each were added and

the solution was further incubated on ice for another 20-60 min. The solution was next sonicated for 6 min, with the 10s pulse “on” and 30s pulse “off” program. The solution was clarified by centrifugation three times for 30 min at $23700 \times g$ at 4 °C. The clarified His₆-tagged human E1o (50 mL) dissolved in 50 mM potassium phosphate (pH 7.4), with 500 mM NaCl, 30 mM imidazole, 2 mM MgCl₂, 0.5 mM ThDP (binding buffer) was filtered through a 0.45 µm filter then applied to a High Performance Ni Sepharose column with a 6 mL bed volume (for a 6 mL volume Ni column, the pressure is approximately 0.2 mbar at 1 ml/min flow rate, and 2.6 mbar at 2 ml/min flow rate). The column was eluted with a solution of 50 mM sodium phosphate (pH 7.4), 500 mM NaCl, 500 mM imidazole, 2 mM MgCl₂, 0.5 mM ThDP with the following gradient: a 2CV linear gradient of 30 to 77 mM imidazole, followed by a 20 CV linear gradient of 77 to 218 mM imidazole, followed finally by a push with 500 mM imidazole. Typically, 40 mg human E1o protein could be purified from 4 L LB culture media.

List of Publications

Publications at Rutgers University

1. Zhou J, Yang L, DeColli A, Meyers C F, Nemeria N S, Jordan F. Conformational dynamics of 1-deoxy-d-xylulose 5-phosphate synthase on ligand binding revealed by H/D exchange MS[J]. PNAS, 2017: 201619981.
2. Guevara E L, Yang L, Birkaya B, Zhou J, Nemeria N S, Patel M S, Jordan F. Global view of cognate kinase activation by the human pyruvate dehydrogenase complex[J]. Scientific Reports, 2017, 7. (Co-first author)
3. Nemeria N S, Gerfen G, Nareddy P R, Yang L, Zhang X, Szostak M, Jordan F. The mitochondrial 2-oxoadipate and 2-oxoglutarate dehydrogenase complexes share their E2 and E3 components for their function and both generate reactive oxygen species[J]. Free Radical Biology and Medicine, 2018, 115: 136-145.

Publications at Wuhan University

1. Li H, Yang L, Tang R, Hou Y, Yang Y, Wang H, Han H, Qin J, Li Q, Li Z. Organic dyes incorporating N-functionalized pyrrole as conjugated bridge for dye-sensitized solar cells: Convenient synthesis, additional withdrawing group on the π -bridge and the suppressed aggregation[J]. Dyes and Pigments, 2013, 99(3): 863-870.