



저작자표시-동일조건변경허락 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이차적 저작물을 작성할 수 있습니다.
- 이 저작물을 영리 목적으로 이용할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



동일조건변경허락. 귀하가 이 저작물을 개작, 변형 또는 가공했을 경우에는, 이 저작물과 동일한 이용허락조건하에서만 배포할 수 있습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

ABSTRACT

Structural basis for target search and recognition in protein complexes probed by NMR spectroscopy

Taekyung Yu

Department of Agricultural Biotechnology

The Graduate School

Seoul National University

The protein–protein interaction is important in many biological processes, including enzyme catalysis, immune response, and cell signaling. It is important to understand the structural basis for the protein interactions, since there is a direct relationship between the structure and the biological function in cell. NMR

spectroscopy has become a powerful analytical tool for the protein structure and dynamics at the molecular level and emphasizing the application in biomolecular systems. Many proteins have been studied on the determination of protein structure and the biological functions using NMR. In this thesis, I demonstrate that the target search process of a protein complex can be visualized by NMR paramagnetic relaxation enhancement (PRE), and that the specific target recognition can be achieved by unusual coupling of binding and unfolding.

In Chapter I, the target search pathway via encounter complex ensemble is characterized between the N-terminal domain of enzyme I (EIN) and the heat-stable histidine phosphocarrier (HPr) of bacterial phosphotransferase system (PTS). The encounter complex ensemble represents short-lived and lowly-populated nonspecific complexes that quickly relax into the final specific complex. It has been difficult to experimentally detect the encounter complexes, but the recent application of PRE was enabled to prove the characterization of transient encounter complexes during protein-protein association. I employed rational mutations of EIN based on PRE data, which potentially perturbed the target search pathway, and examined how the mutations altered the encounter complex formation and also the equilibrium binding constant. If encounter complexes at specific region influences the protein association, they are classified as productive encounter, otherwise non-productive. Mapping the region of productive encounter complexes on the surface of EIN, I visually demonstrate the allowed and disallowed target search pathways between EIN and HPr.

In Chapter II, the complex structure of fibronectin extradomain B (EDB) and a specific binding peptide (aptide, APT) was determined by NMR spectroscopy.

EDB is a prominent marker of tumor angiogenesis, and APT is a small peptide (26 aa) with a β -hairpin scaffold and two engineered target binding arms. APT recognizes EDB via an unusual β strand replacement mechanism, in which APT permanently unfolds the entire C-terminal β strand of EDB and forms a new intermolecular β sheet within the complex. Thus the interaction interface of EDB is not located on the surface, but buried inside of the protein. Unfolding of EDB exposes the hydrophobic binding interface that is specifically recognized by APT. The unique binding mode of coupled binding and unfolding will broaden our understanding of the diverse protein interactions.

Keyword: aptide, encounter complex, fibronectin extradomain B, paramagnetic relaxation enhancement, phosphotransferase system,

Student Number: 2010-31344

CONTENTS

ABSTRACT	i
CONTENTS	iv
LIST OF FIGURES	ix
LIST OF TABLES	xviii
ABBREVIATION	xx

PART I. VISUALIZATION OF PROTEIN TARGET SEARCH

PATHWAYS BY NMR PARAMAGNETIC

RELAXATION ENHANCEMENT

1. Abstract	2
2. Introduction	4
2.1. Phosphotransferase system (PTS)	4
2. 2. Paramagnetic relaxation enhancement (PRE)	6
2. 3. Encounter complex between EIN and HPr	8
3. Materials and Methods	11

3. 1. Cloning	11
3. 2. Protein overexpression	11
3. 2. 1. Luria bertani medium	11
3. 2. 2. Isotope labeling medium	11
3. 2. 3. Selective labeling	12
3. 3. Purification	15
3. 3. 1. EIN and EIN mutants	15
3. 3. 2. HPr	15
3. 3. 3. Paramagnetic spin labeling	16
3. 4. NMR experiment	19
3. 5. Isothermal titration calorimetry (ITC)	19
3. 6. Circular dichroism (CD) spectroscopy	20
4. Results	21
4.1. Mutation of EIN	21
4. 2. Analysis of secondary structure using circular dichroism (CD) spectroscopy	25
4. 3. Binding affinity between EINs and HPr	27
4. 4. Backbone assignment of EIN mut3	30
4. 5. Investigation of conformation and dynamics by NMR	33
4. 6. Visualization of encounter complex in EINs–HPr	38
4. 7. Productive and non-productive encounter complexes between EIN and HPr	44

5. Discussion	47
6. Conclusion	51

PART II. STRUCTURAL BASIS FOR TARGET

RECOGNITION OF A β -HAIRPIN APTAMER-LIKE

PEPTIDE

1. Abstract	53
2. Introduction	54
2. 1. Structure of fibronectin	54
2. 2. Alternatively spliced fibronectin isoformes	55
2. 3. Structure of the EDB domain	58
2. 4. β -hairpin aptamer like peptide (Aptide)	60
3. Materials and Methods	62
3.1. Cloning	62
3.2. Culture of EDBs and APT _{EDB}	62
3.3. Purification of EDBs and APT _{EDB}	63
3.3.1 EDB and EDB mutants	63

3.3.2 APT _{EDB}	64
3. 4. NMR spectroscopy	67
3. 5. Structure calculations	67
3. 6. Isothermal titration calorimetry (ITC)	68
4. Results	69
4.1. Binding between EDB and APT _{EDB}	69
4. 2. Determination of complex structure	74
4. 3. Description on EDB in complex structure	83
4. 4. Proofs upon unfolding of β strand G of EDB	85
4. 5. Description on APT _{EDB} in complex structure	92
4. 6. Binding interface between EDB and APT _{EDB}	96
4. 7. Thermodynamics on EDB and APT _{EDB}	102
4. 8. Relationship between EDB and FN8	106
5. Discussion	109
6. Conclusion	111
III. References	112
IV. SUMMARY IN KOREAN	120
V. ACKNOWLEDGEMENT	126

LIST OF FIGURES

Part I

Figure 1. (a) Transport of sugar by the PEP-dependent phosphotransferase system (PTS). HPr and EI are cytoplasmic proteins common to all sugars transported. EIIA and EIIB are proteins specific for glucose. (b) Ribbon diagrams illustrating the EIN–HPr complex. Helixes (red), β -strands (yellow), and domains were annotated. 7

Figure 2. Illustration of protein–protein association on a reaction coordinate diagram. Two interacting proteins are shown in blue and orange in a cartoon representation. Non-specific encounter complexes fall into stereospecific complex through the productive target search pathway. Several non-specific encounter complexes do not reach the complex and dissociate to the free states. 10

Figure 3. Purification procedure of EIN and EIN mutants. Purity of proteins was confirmed by SDS-PAGE. 17

Figure 4. Purification procedure of EIN and EIN mutants. Purity of proteins was confirmed by SDS-PAGE. HPr mutants (E5C, E25, and E32C) were used to buffers

contained 2 mM 2-Mercaptoethanol for reducing the disulfide bonds. 18

Figure 5. Mutation sites of the five EIN mutants. Shown upper are color-coded ribbon representation of EIN (helix: yellow and β sheet: purple) –HPr (cyan) complex. The mutation sites introduced to EIN are shown with space-filling models (red). 24

Figure 6. CD spectra of the EIN and EIN mutants. All CD spectra were collected at 25°C using a JASCO J-801 CD spectrometer. 26

Figure 7. Comparison of 2D ^1H - ^{15}N TROSY correlation spectrum of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled EINmut3 and [^{15}N -Lys]-EIN mut3 (red). EIN mut3 recorded at 900 MHz and 37°C. The selectively labeled EINmut3 are annotated. 31

Figure 8. Comparison of 2D ^1H - ^{15}N TROSY correlation spectrum of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled EIN (black) and EINmut3 (red). The backbone amide resonances with chemical shift changes are annotated by the residue name in blue, and those without chemical shift changes are annotated in black. 2D ^1H - ^{15}N TROSY NMR spectroscopy was recorded at 900 MHz and 37°C in 20 mM Tris-HCl, pH 7.4. 32

Figure 9. Chemical shift perturbation (CSP) of (a) EINmut3, (b) EINmut3a, and (c) EINmut2 compared to EIN. Weighted average $^1\text{H}_\text{N}/^{15}\text{N}$ chemical shift perturbation ($\Delta\delta = ([(\Delta\delta_\text{HN})^2 + (\Delta\delta_\text{N})^2/25]/2)^{1/2}$) was plotted as a function of residue number.

..... 35

Figure 10. Ratio of R_2 relaxation parameters of amide protons (a) EIN mut3, (b) EIN mut3a, and (c) EIN mut2 compared to EIN as a function of residue number (R_2 ratio = $R_2_{\text{EINmut}} / R_2_{\text{EIN}}$). 36

Figure 11. RDC analysis of (a) EIN mut3, (b) EIN mut3a, and (c) EIN mut2 in complex with HPr. Shown is a comparison of the observed and calculated $^1D_{\text{NH}}$ RDCs obtained by singular value decomposition (SVD) with the NMR coordinates of the EIN–HPr complex (PDB ID 3EZA). The values of RDC R-factors and the correlation coefficients are shown. RDCs from the residues in helices α_2 and α_2' are shown in red, and Lys81 and Lys86 of the mutation sites are annotated in (a). 37

Figure 12. Mass Spectrum data of EDTA– Mn^{+2} Conjugated HPr at E5C. The measured spinlabeled HPr molecular weight was 9503.7 (calculated spinlabeled HPr: 9495.58). Mass spectrum was collected in 20mM Tris-HCl, pH 7.4 using electron spray ionization (ESI) Mass spectroscopy. 41

Figure 13. Comparison of the observed and calculated intermolecular PREs for the a) EIN–HPr (E5C), b) EIN mut3–HPr (E5C), c) EIN mut3a–HPr (E5C), and d) EIN mut2–HPr (E5C) complexes. The observed intermolecular $^1\text{H}_\text{N}$ - Γ_2 rates (red circles) and calculated values (black line) are shown with error bars. PREs arising

from the specific complex structure are highlighted in the purple boxes, and PREs attributable to the nonspecific encounter complex ensemble are highlighted in the yellow boxes. Differences between the PRE profiles of the EIN and EIN mutants are shown in the dashed red boxes. 42

Figure 14. The PRE profiles are color-coded in the solution structure of the EIN–HPr complex (PDB ID 3EZA) for the (a) EIN–HPr (E5C), and (b) EIN mut3–HPr (E5C) complexes. EIN is shown in gray and HPr is shown in red. The mutation sites in the EIN mut3 (E81K and E86K) are shown as a space-filling model representation in cyan. Intermolecular PREs are color-coded on the structure in blue or yellow in the same manner as the PRE profiles. The expansions illustrate the mutation site disappeared on PREs. 43

Figure 15. 2D ^1H - ^{15}N HSQC spectra of ^{15}N -HPr titrated with a) EIN, b) EINmut3, and c) EINmut2. HPr (0.5 mM) was titrated with stoichiometric ratios of EIN (mut) protein of 0 (black), 0.1 (red), 0.2 (green), 0.3 (blue), 0.5 (magenta), 1 (cyan), and 2 (violet). The progressive chemical-shift perturbation of Gln51 is characteristic of the fast side of intermediate exchange. 46

Figure 16. A model of target search pathways of the productive and non-productive encounter complexes between EIN (yellow) and HPr (red). Below panel is illustrated on the model of target search pathway between EIN (yellow square) and HPr (red circle). 50

Part II

Figure 1. Representation of a subunit of human Fibronectin. FN subunit is comprised on a series of repeating units of three different types (type I, type II, type III). Two subunits are connected by two disulfide bonds at their C-term. Three repeats can be either inserted or omitted in the molecule by a mechanism of alternative splicing: EDB, EDA, IIICS. Separate domains have been identified that contain binding activities for other components, such as heparin, DNA, fibrin, and etc., of the ECM. 57

Figure 2. Structure of the EDB domain of fibronectin. The EDB structure represents ribbon drawing of one of the 20 energy-minimized conformers of human EDB (PDB code, 2FNB). The two β -sheets are comprised of three and four β -strands, respectively. The β -strands are annotated A to G following the nomenclature used for other FN type-III domains. 59

Figure 3. Structure of Trpzip and APT_{EDB}. a) Representative structure of trpzip is consisted of two tryptophan–tryptophan cross pairs (PDB code: 1LE0). b) APT is comprised of tryptophan zipper (arrow) and two ligand binding regions (circle). 61

Figure 4. Purification procedure of EDB and EDB mutants. Purity of proteins was confirmed by SDS-PAGE. 65

Figure 5. Purification procedure of APT and APT mutants. Purity of proteins was confirmed by SDS-PAGE.	66
Figure 6. Comparison of ^1H - ^{15}N HSQC spectrum of EDB-APT _{EDB} (27-mer amino acid, black) and -APT _{EDB} (26-mer amino acid, red) recorded at 600 MHz and 25°C in 20 mM NaPi, pH 6.0.	71
Figure 7. ^1H - ^{15}N HSQC and chemical shift perturbation (CSP) between human EDB and APT _{EDB} . (a) Comparison of ^1H - ^{15}N HSQC spectra of free ^{15}N -EDB (black) and the complex of ^{15}N -EDB with APT _{EDB} (red). Residues with large chemical shift changes are annotated. (b) The weighted average chemical shift perturbation of backbone amide resonances of EDB as a function of residue number upon complex formation with APT _{EDB}	73
Figure 8. Comparison of 2D ^1H - ^{15}N HSQC spectrum of complex structures of EDB, [^{15}N -Leu]-EDB (red), and [^{15}N -Ile]-EDB (blue). The residues of selectively labeled EDB are annotated.	76
Figure 9. Ramachandran plot of 20 lowest energy NMR structure of EDB-APT _{EDB} complex.	82
Figure 10. (a) The complex structure as a ribbon diagram with EDB in yellow and	

APT _{EDB} in green. (b) Comparison between free (blue) and complex (green) EDB. The C-term β strand G was annotated.	84
Figure 11. ¹ H- ¹⁵ N heteronuclear NOE data as a function of the residue number of EDB (a) in the free state, and (b) in complex with APT.	87
Figure 12. Superimposed ¹ H- ¹⁵ N HSQC spectra of ¹⁵ N-EDB (<i>black</i>) and ¹⁵ N- EDB ₅₋₈₃ (<i>red</i>) in complex with APT. The missing resonances of ¹⁵ N-EDB ₅₋₈₃ due to the truncation are annotated by residue types and numbers.	88
Figure 13. Superimposed ¹ H- ¹⁵ N HSQC spectra of ¹⁵ N-APT in complex with EDB (<i>black</i>) and EDB ₅₋₈₃ (<i>red</i>). The assignments for backbone amides of APT are annotated in the spectra, and the side chain H ϵ resonances of four tryptophan residues are annotated by asterisks.	89
Figure 14. Predicted secondary structure of free (a) and complex (b) EDB. The positive bars means a β sheet, and the negative bars is a helix. The height of the bars reflects the probability of the neural network secondary structure prediction.	91
Figure 15. Comparison of ¹ H- ¹⁵ N HSQC spectrum between free (black) and complex (red) APT _{EDB}	94

Figure 16. ^1H - ^{15}N heteronuclear NOE as a function of the residue number of APT in complex with EDB. The resonance for Ser202 of APT could not be detected in the presence of ^1H saturation, indicating that its ^1H - ^{15}N heteronuclear NOE was close to zero. 95

Figure 17. The interface between EDB and APT_{EDB} . Interfacial residues of EDB in red and APT in blue with the backbone are shown as ribbon model and the side chains shown as sticks. Interaction interface residues of APT_{EDB} are annotated (yellow). 99

Figure 18. (a)The intermolecular hydrogen bonds between β strand F and G. (b) the intramolecular backbone hydrogen bonds between β strands F and β strand II of APT in the complex. The hydrogen bonds between β strands are denoted by vertical bars (red). 101

Figure 19. Schematic illustration of coupled unfolding and binding during the interaction between EDB and APT_{EDB} 105

Figure 20. Superimposed ^1H - ^{15}N HSQC spectra of ^{15}N -FN (black) and ^{15}N -FN in complex with APT_{EDB} (red). 107

Figure 21. Comparison of ^1H - ^{15}N HSQC spectrum between EDB linked with FN8 and EDB unlinked with FN8. (a) ^1H - ^{15}N HSQC spectrum of $\text{EDB-FN8}:\text{APT}_{\text{EDB}}$

(black), EDB (red), and FN8 (blue). EDB and FN8 were chemical shift change against EDB-FN8 due to loop interaction between EDB and FN8. (b) Comparison of ^1H - ^{15}N HSQC spectrum of EDB-FN8:APT_{EDB} (black), EDB:APT_{EDB} (red), and FN8:APT_{EDB} (blue). EDB:APT_{EDB}-FN8 spectra was similar to two spectrum combined with EDB:APT_{EDB} and FN8:APT_{EDB}. 108

LIST OF TABLES

Part I

Table 1. The composition of media.	13
Table 2. Unlabeled amino acid mixture except for selective labeling amino acid.	14
Table 3. Residues that were mutated in each EIN mutant, the positions of the mutation sites in the EIN structure, the spin-labeled HPr used to design the EIN mutants, the product yield of EIN mutants.	23
Table 4. Thermodynamic parameters for the interaction between EIN (or EIN mutants) and HPr.	29

Part II

Table 1. APT _{EDB} mutants. The mutated residues were marked on red.	72
Table 2. Backbone ¹ H, ¹⁵ N, and ¹³ C α , and ¹³ C β , ¹ H β chemical shifts of complex EDB.	

.....	77
Table3. Backbone ^1H , ^{15}N , and $^{13}\text{C}\alpha$, and $^{13}\text{C}\beta$, $^1\text{H}\beta$ chemical shifts of complex APT_{EDB}	80
Table 4. Restraints and structural statistics.	81
Table 5. Chemical shift index (CSI) of free and complex EDB. The residues appear the length of β strand.	90
Table 6. Thermodynamic parameters for the interaction between EDB and APT_{EDB} (APT_{EDB} mutants). Equilibrium dissociation constants (K_D), binding free energy (ΔG), binding enthalpy (ΔH), and binding entropy (ΔS) were obtained from ITC at 25°C	100
Table 5. Thermodynamic parameters for the interaction between EDB (or EDB mutants) and APT, and between EDB–FN8 and APT. Equilibrium dissociation constants (K_D), binding free energy (ΔG), binding enthalpy (ΔH), and binding entropy (ΔS) were obtained from ITC at 25°C	91

ABBREVIATION

CBCACONH	Correlation spectroscopy for $H^N-N^H-C\alpha^i-C\beta^i$
CD	Circular dichroism
CSI	Chemical shift index
CSP	Chemical shift perturbation
DEAE	Diethylaminoethyl
EDB	Extradomain B
EDTA	Ethlyenediaminetetraacetic acid
EIN	N-terminal domain of enzyme I
FN	Fibronectin
HBHACONH	Correlation spectroscopy for $H^N-N^H-H^\alpha-H^\beta$
HNCACB	Correlation spectroscopy for $C\alpha^{i-1}-C\beta^{i-1}-H^N-N^H-C\alpha^i-C\beta^i$
HPr	Histidine phosphocarrier
HSQC	Heteronuclear single quantum coherence spectroscopy
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
OD	Optical density
PMSF	Phenylmethanesulfonyl fluoride

PRE	Paramagnetic relaxation enhancement
PTS	Phosphotransferase system
RDC	Residual dipolar coupling
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SVD	Singular value decomposition
TOCSY	Total correlation spectroscopy
TROSY	Transverse relaxation optimized spectroscopy

PART I.

**VISUALIZATION OF TARGET SEARCH
PATHWAYS BY USING NMR PARAMAGNETIC
RELAXATION ENHANCEMENT**

1. Abstract

In general, specific protein–protein recognition proceeds via a two–step target search process. The first process is a weak association via diffusion–controlled intermolecular collisions, followed by the formation of an ensemble of short–lived, encounter complexes. Secondly, two–dimensional search by translational and rotational rearrangement between the partner proteins results in the specific complex formation. Here I investigated the target search pathways between the N–terminal domain of enzyme I (EIN) and the heat–stable histidine phosphocarrier protein (HPr) from bacterial phosphotransferase system (PTS) using NMR paramagnetic relaxation enhancement (PRE). The target search process was examined how mutations away from the binding interface alter the equilibrium binding and reshape the distribution and/or population of the encounter complex ensemble. Probing the target search pathway was initiated on rational mutations, and which were analyzed on secondary structures, equilibrium dissociation constants, and three–dimensional structures to visualize the encounter complex ensemble. I show that encounter complexes are heterogeneous in their contributions toward the specific complex formation, which can be visually monitored and estimated in a quantitative manner. The encounter complexes related on association or/and dissociation were visualized via PRE. For such results, I propose the target search pathway model related on the productive and non-productive encounter complex during the conformation of the stereospecific

complex. When a protein interacts with multiple partner proteins, the hot-spot regions that form productive or nonproductive encounter complexes can vary and sometimes may even switch their roles to discriminate different partner proteins. The exploration of target search pathways could eventually be applied to the manipulation of equilibrium binding in protein-protein or protein-ligand complexes without affecting their binding interface.

2. Introduction

2.1. Phosphotransferase system (PTS)

Phosphoenol pyruvate (PEP) group translocation, also known as the PTS, is a definite method used by bacteria for sugar uptake where the source of energy is from PEP. It is known as multi-component system involved in enzymes of the plasma membrane and cytoplasm of *Escherichia coli*. The PTS is involved in transporting many sugars such as glucose, mannose, fructose and cellobiose into *E. coli* cells. The consumed sugar transfers the phosphoryl group to the plasma membrane and cytoplasm using several proteins that are enzyme I (EI), phosphocarrier protein (HPr) and enzyme II (EII) to a conserved histidine residue [1].

In the process of glucose PTS transport specific of enteric bacteria, PEP transfers its phosphoryl to a histidine residue on EI. EI in turn transfers the phosphate to HPr. From HPr the phosphoryl is transferred to EIIA. EIIA is specific for glucose and it further transfers the phosphoryl group to EIIB in adjacent membrane on one side of it. Finally, EIIB phosphorylates glucose as it crosses the plasma membrane through the transmembrane enzyme II C (EIIC), forming glucose-6-phosphate (Figure 1a). The HPr is common to the PTS systems of the other substrates mentioned earlier, as is the upstream EI. Proteins downstream of HPr tend to vary between the different sugars. The transfer of a phosphate group to the substrate once it has been imported through the membrane transporter prevents the transporter from

recognizing the substrate again, thus maintaining a concentration gradient that favours further import of the substrate through the transporter.

The EI is a homodimers (~64 kD) that needs Mg^{2+} for phosphorylation by PEP. Each subunit contains three domains: PEP binds to the C-terminal domain and HPr binds to the N-terminal domain. A central domain, tethered to the N-term domain by two associated linkers and to the C-term domain by a long α -helix, comprises a phosphorylatable residue His189. In the EI structure, His189 is phosphorylated and oriented for directly phosphotransfer to/from the PEP/HPr. Therefore, the structure typifies an enzyme intermediate after phosphotransfer from PEP and prior to a conformational transition that brings phosphoryl group of His189 to the phosphoryl group acceptor (His15 of HPr).

EIN comprises two subdomains which are α and α/β domains (Figure 1b). The α domain (residues 33–143) is a four helix bundle (H1, H2/H2', H3 and H4). The α/β domain (residues 1–20 and 148–230) comprises a β -sandwich, formed by a four-stranded parallel β -sheet (β 1– β 4) and a three-stranded antiparallel β -sheet (β 1, β 5, β 6), as well as three short helices (H5–H7). There also is a long C-terminal helix (H8) that serves as a linker to the C-terminal domain of EI. HPr comprises three helices and a four stranded antiparallel β -sheet. The structures of EIN and HPr in the complex are very similar to those in the free state. The interaction surface on EIN comprises virtually the entire helix 2, the N-terminal end of helix 2', the C-terminal half of helix 3 and the N-terminal half of helix 4. The interaction surface on HPr comprises helix 1, helix 2, and the turn proceeding helix 1 and the beginning of the loop following helix 2.

Enzyme II contains five classes that are glucose–sucrose, mannitol–fructose, mannose–sorbose, lactose–cellobiose, and glucitol [2]. In the different sugar classes, the EIIA and EIIB domains have specifically its characters on sequence or three–dimensional structure. The bacterial PEP: sugar PTS system provides a typical example for understanding protein–protein interactions and the factors governing their specificity. Thus, HPr searches the specific binding interface of enzyme I and the various sugar–specific IIA domains, even though all these target proteins involved in the PTS system are structurally dissimilar.

2. 2. Paramagnetic relaxation enhancement (PRE)

Paramagnetic Relaxation Enhancement may be first introduced to in the 1950s. PRE is increasingly becoming a NMR technique to provide information of long–range distance that can complement nuclear overhauser effect restraints (limited to distances of less than 6 Å). Paramagnetic NMR has come out as an alternative, orthogonal, and influential tool to study dynamic processes. The PRE generating from unpaired electrons with an isotropic g–tensor (such as nitroxide or EDTA–Mn²⁺) has proved to be a particularly versatile tool. And it can provide information about structural conditions of dynamic process due to the $\langle r^{-6} \rangle$ distance dependence of the PRE between the unpaired electron and the nucleus. The PRE effect is very huge, because of the large magnetic moment of an unpaired electron, allowing distances up to 35 Å to be detected. Providing the paramagnetic center–nuclei distances are shorter in the minor species than in the major one, the mark of the minor species will be apparent in the observed PRE rates, thus

enabling structural information on the minor species to be obtained.

2. 3. Encounter complex between EIN and HPr

For forming a protein complex, proteins undergo at least a two-step process, where an encounter complex precedes the final, well-defined stereospecific complex. Generally, it is presumed that the encounter state exists only short-lived and leads quickly to the final complex. However, NMR spectroscopy has argued that the weak and low existing encounter state represents a significant fraction of the population or can even be the dominant form upon formation of stereospecific complex. Recently, the encounter state was visualized by a paramagnetic NMR technique for complexes in which it represents a small fraction of the population. In weak protein complexes, the fraction of time spent in the encounter complex may be largest, because a balance between specificity and fast dissociation has to be found.

Transient encounter complexes have first been proved using the EIN and the phosphocarrier HPr protein existed on the bacterial PTS system. It reveals that the distribution of nonspecific encounter complexes between the two proteins depends on the electrostatic surface potentials of the interacting proteins through the results of ensemble-simulated annealing refinement against experimental PREs. The nonspecific encounter complexes are formed by weak non-specific electrostatic interactions, eventually falling into a narrow energy funnel (Figure 2) that leads directly to the stereospecific complex characterized by an array of complementary van Waals and electrostatic interaction^[3, 4]. Encounter complexes can be productive

if they make their way to a final stereospecific complex, or non-productive if they dissociate back to the free individual proteins.

In this study, I employed rational mutations based on several criteria upon the previous experimental PREs to distinguish between productive and non-productive encounter complexes and to probe a possible target search pathway. That is to say, potential hot-spot residues for encounter complex formation detected by PRE are mutated, and the changes of equilibrium dissociation constants (K_D) are measured by the mutants via ITC. If the mutation alters the equilibrium binding, intermolecular PREs show how the altered equilibrium binding constants are related with changes in the encounter complex formation.

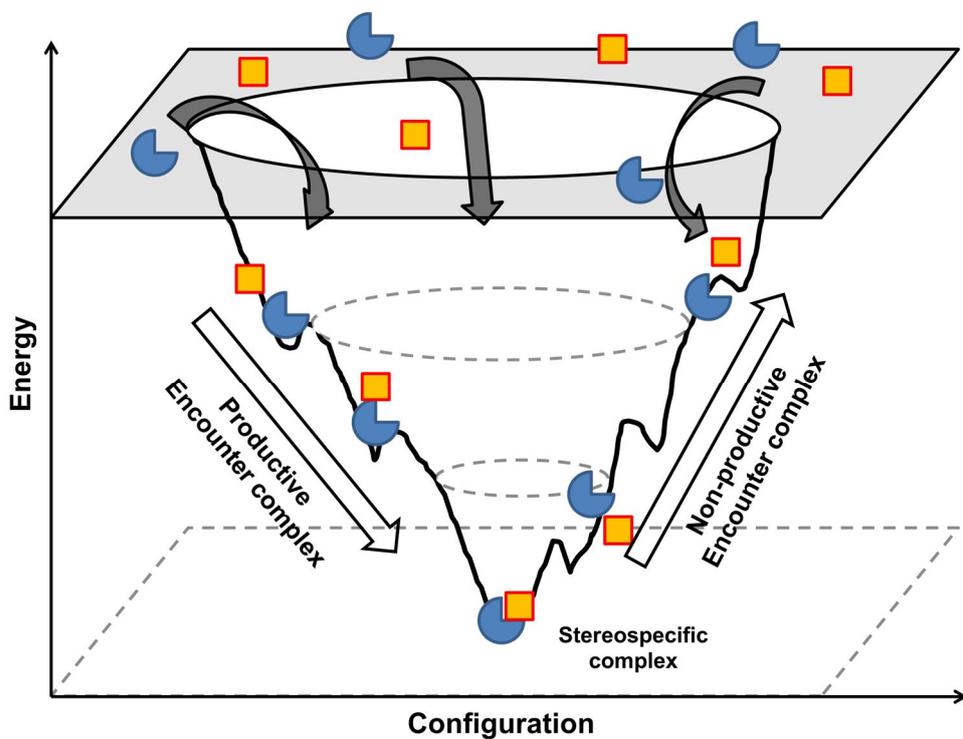


Figure 2. Illustration of protein–protein association on a reaction coordinate diagram. Two interacting proteins are shown in blue and orange in a cartoon representation. Non–specific encounter complexes fall into stereospecific complex through the productive target search pathway. Several non–specific encounter complexes do not reach the complex and dissociate to the free states.

3. Materials and Methods

3. 1. Cloning

EIN (residues 1–249) and HPr (residues 1–85) were cloned into a pET11a vector (Novagen) without tags. Selected residues of EIN or HPr were mutated using the QuickChange kit (Agilent Technologies), and the new constructs were verified by DNA sequencing.

3. 2. Protein overexpression

3. 2. 1. Luria bertani medium

The expression of the EIN, EIN mutants, and HPr domain was performed using the expression vector pET11a in *Escherichia. coli* BL21 cells (DE3, Invitrogen) using LB medium (Table 1). The cells transformed by the expression vector were cultured at 37°C to an OD₆₀₀ of 0.6–0.8, and the protein overexpression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, followed by induction for 4 h at 37°C and harvested by centrifugation. The cell separated on medium was storage on -80°C

3. 2. 2. Isotope labeling medium

Isotope labeling proteins have been useful for increasing the sensitivity of NMR experiments. Uniformly labeled *E. coli* [²H; ¹³C; ¹⁵N]–EIN (1–249) and

histidine-containing phosphocarrier protein (HPr) at natural isotopic abundance were expressed in *E. coli* cells (BL21 Star (DE3)) grown in M9 minimal media (Table 1) at 37°C until an OD₆₀₀ ≈ 0.6–0.8. Cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37°C for approximately 8 h and then harvested by centrifugation. If the cell was below OD₆₀₀: 1.0 during 8 h, the cell was cultured at 25°C for overnight.

3. 2. 3. Selective labeling

Selective labeling culture method has been considered as an efficient method to simplify the NMR spectra. In this method, it make only the desired signals appear (¹³C- or ¹⁵N-labeling), or undesired ones disappear (²H-labeling) in the NMR spectrum. This facilitates the analysis of the large molecules and resonance assignments. In this study, ¹⁵N-isoleucine, leucine, and lysine were performed on the selective labeling upon EIN. For selective labeling, a medium containing the appropriate unlabeled amino acid mixture (table 2) was prepared, which is based on M9 medium. The labeled amino acid (25mg) was filtered on sterile filter (0.22 um) and added to medium after the induction. The each selective labeling medium were cultured at 37°C to an OD₆₀₀ of 0.8, and protein overexpression was induced by adding IPTG to a final concentration of 1 mM, followed by induction for 4 h at 37°C and harvested by centrifugation.

Table 1. The composition of media.

Medium	Composition (per 1 L)
LB	10 g Tryptone, 5 g Yeast extract, 10 g NaCl
M9 minimal medium (¹⁵ N or ¹³ C/ ¹⁵ N)	10 g K ₂ HPO ₄ , 13 g K ₂ HPO ₄ , 9 g Na ₂ HPO ₄ , 2.4 g K ₂ PO ₄ , 1 g ¹⁵ NH ₄ Cl, 2 g U- ¹³ C ₆ Glucose (or 5 g Glucose), 10 mM MgCl ₂ , 0.1 mM Thiamine, Trace element, 0.2 mM CaCl ₂ , 50 mg Carbenicillin
* M9 minimal medium for 1 L D ₂ O (² H/ ¹⁵ N or ² H/ ¹⁵ N/ ¹³ C)	10 g K ₂ HPO ₄ , 13 g K ₂ HPO ₄ , 9 g Na ₂ HPO ₄ , 2.4 g K ₂ PO ₄ , 1 g ¹⁵ NH ₄ Cl, 2 g U- ¹² C-U- ² H-D-Glucose (or U- ¹³ C-U- ² H-D-glucose), 952 mg MgCl ₂ , 34 mg Thiamine, Trace element, 100 mg Carbenicillin

* All components are dissolved on D₂O and filtered on sterile filter (0.22 um), transferring to an autoclave flask.

Table 2. Unlabeled amino acid mixture except for selective labeling amino acid

Amino acid	Amount of amino acid (g)	Method of sterilization
Alanine	0.50	Autoclave
Arginine	0.40	Autoclave
Aspartic acid	0.40	Filter
Asparagine	0.40	Filter
Cysteine	0.05	Filter
Glutamine	0.40	Filter
Glutamic acid	0.65	Filter
Glycine	0.55	Autoclave
Histidine	0.10	Autoclave
Isoleucine	0.23	Autoclave
Leucine	0.23	Autoclave
Lysine-HCl	0.42	Autoclave
Methionine	0.25	Autoclave
Phenylalanine	0.13	Autoclave
Proline	0.10	Autoclave
Serine	2.10	Autoclave
Threonine	0.23	Autoclave
Tyrosine	0.17	Filter
Tryptophan	0.05	Filter
Valine	0.23	Autoclave

- Filter was used on sterile filter (0.22 μ m).

3. 3. Purification

3. 3. 1. EIN and EIN mutants

The cells were harvested by centrifugation, and resuspended in buffer [20 mM Tris–HCl (pH 7.4), 200 mM NaCl, 1 mM phenylmethanesulfonyl fluoride (PMSF)], and disrupted by emulsiflex. The lysate was centrifuged to remove insoluble debris. The supernatant was loaded on a HiPrep DEAE column equilibrated with binding buffer [20 mM Tris–HCl (pH 7.4)] and was eluted with a gradient concentration of NaCl. The solutions containing protein were separated by a HiLoad Superdex 75 column [20 mM Tris–HCl (pH 7.4), 200 mM NaCl] and finally were purified by Mono Q column. The EIN and EIN mutants were analyzed by SDS–polyacrylamide gel electrophoresis to confirm sample mass and sample purity (Figure 3).

3. 3. 2. HPr

The HPr protein also was purified on identical methods to the purification steps of EIN protein. The cell containing the HPr protein was resuspended in 50 ml of 50 mM Tris, pH 7.4, 200 mM NaCl, and 1 mM PMSF. The suspension was lysed by three passages through a emulsiflex and centrifuged at 40,000 g for 20 min. The supernatant fraction was filtered and loaded onto a DEAE anion exchange column, and the protein was eluted with a 400–ml gradient of 1 M NaCl. The fractions containing the protein were confirmed by SDS–PAGE and purified by gel filtration on a Sephadex 75 column equilibrated with 20 mM Tris, pH 7.4, 200 mM NaCl. Relevant fractions were exchanged to buffer [20 mM Tris (pH 7.4)], and finally

purified on a monoQ anion exchange column (Figure 4).

3. 3. 3. Paramagnetic spin labeling

Immediately prior to conjugation with the EDTA-Mn²⁺ tag, the E5C-HPr protein was passed through a desalting column (HiPrep 26/10 Desalting; GE Healthcare) to remove residual reducing agent. The final buffer of E5C-HPr was 20 mM Tris-HCl, pH 7.4 and 2 mM β -mercaptoethanol and every buffers chelated using chelex 100 for removing metal ions in the purification buffer was used on purification steps. The eluted protein through the desalting column was collected into a solution containing 10 mg of N-[S-(2-Pyridylthio) cysteaminy] ethylenediamine-N,N,N',N'-tetraacetic acid (#P996250; Toronto Research Chemicals, Inc.) and Mn²⁺ (Sigma Aldrich M-1787, a protein to metal ratio of approximately 1:5), and incubated at room temperature for 3 h. Unconjugated HPr was separated from the tagged protein by a monoQ anion exchange column. To remove contaminating divalent ions, HPr was incubated overnight at room temperature with 50 mM EDTA. The excess EDTA was removed by exchanging with a high salt buffer (500 mM NaCl) using the Amicon ultracentrifugal concentrator. Excess metal ion was removed by exchange with the high salt buffer, followed by exchange into the NMR buffer with 20 mM Tris-HCl, pH 7.4. The protein was >99% pure by SDS-PAGE. Electrospray mass spectrometry was used to determine the E5C-HPr conjugated with EDTA-Mn²⁺.

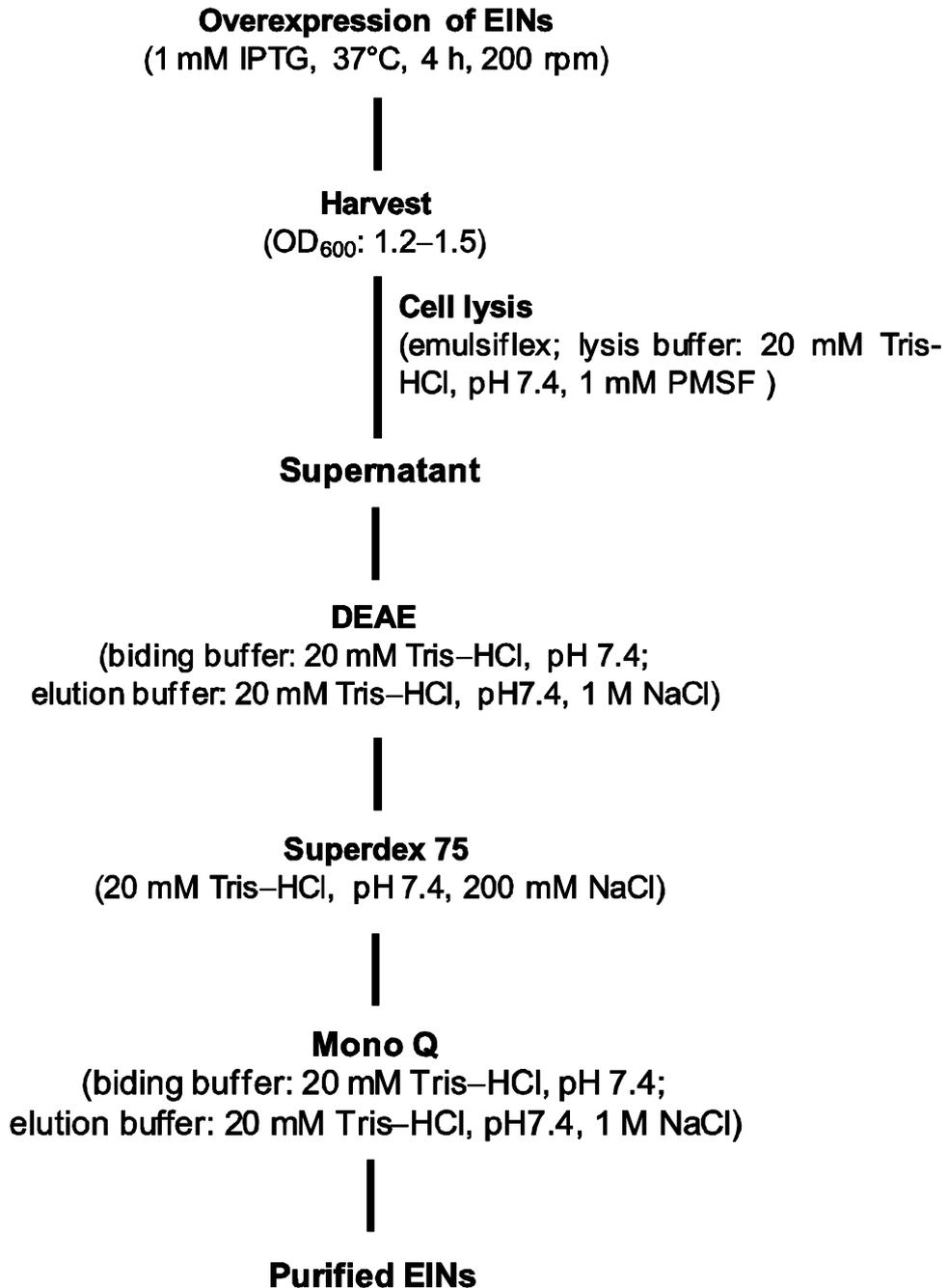


Figure 3. Purification procedure of EIN and EIN mutants. Purity of proteins was confirmed by SDS-PAGE.

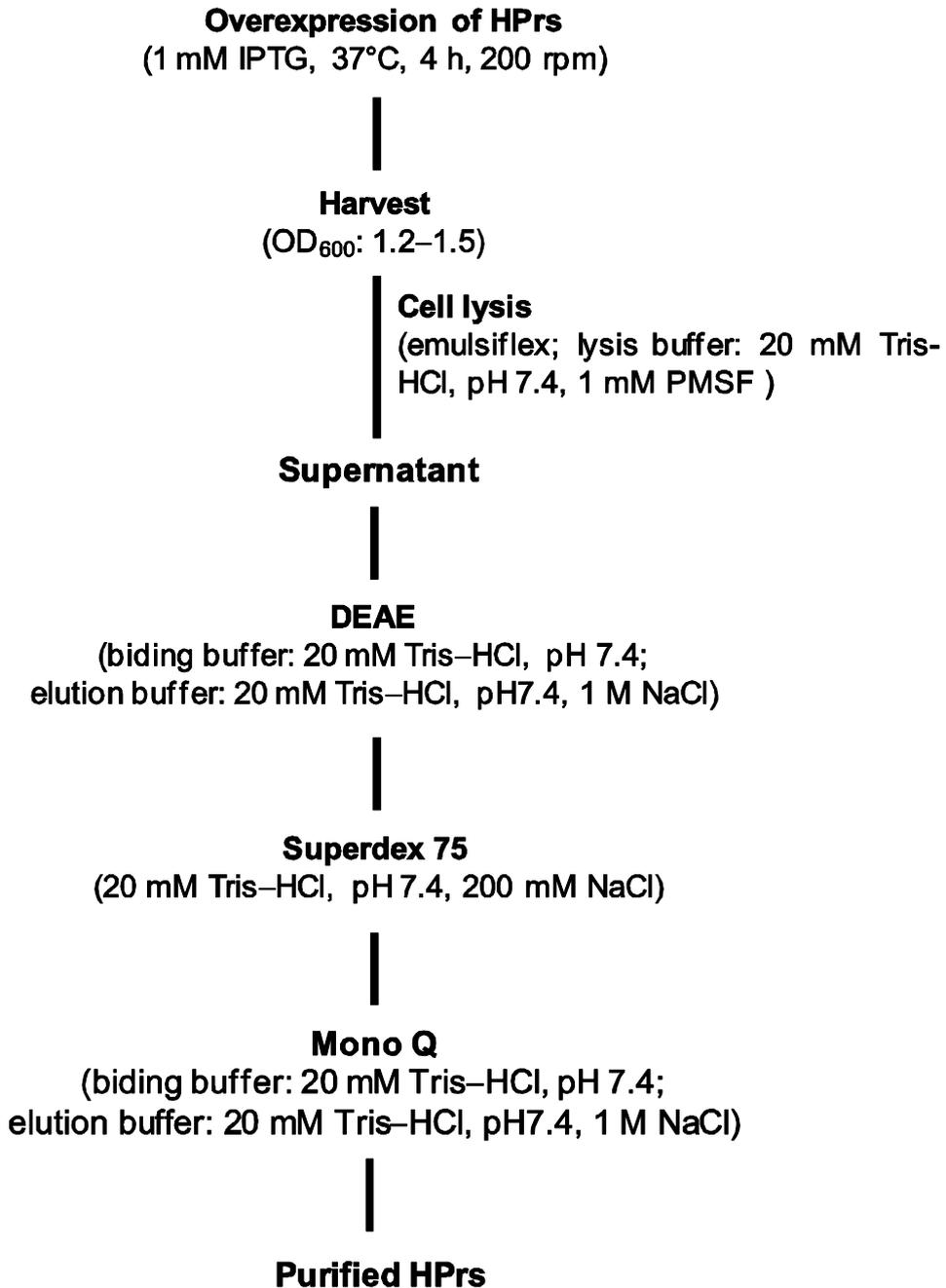


Figure 4. Purification procedure of EIN and EIN mutants. Purity of proteins was confirmed by SDS-PAGE. HPr mutants (E5C, E25, and E32C) were used to buffers contained 2 mM 2-Mercaptoethanol for reducing the disulfide bonds.

3. 4. NMR experiment

NMR spectroscopy. NMR spectra were recorded at 37°C on Bruker 800 and 900 MHz spectrometers equipped with a z-shielded gradient triple resonance cryoprobe. Spectra were processed using NMRPipe1 and analyzed using the program PIPP^[5] or NMRView.^[6] Sequential assignment of EINmut3 was performed using three-dimensional triple resonance through-bond scalar correlation experiments (HNCACB and CBCACONH). $^1D_{NH}$ RDCs were measured by taking the difference in $^1J_{NH}$ scalar couplings in aligned and isotropic media at 37 °C. The alignment media employed were phage pfl (11 mg/ml; ASLA Biotech). Singular value decomposition (SVD) analysis of RDCs was carried out using Xplor-NIH^[4]. $^1HN-R_2$ rates for U- $[^2H/^{15}N]$ -EIN and U- $[^2H/^{15}N]$ -EINmut3 in the presence of HPr (paramagnetic or diamagnetic) were obtained from a two time point measurement (12 μ s and 27 ms, NT = 27 ms) recorded in an interleaved manner using a transverse relaxation optimized (TROSY) version of the $^1H-^{15}N$ heteronuclear single quantum correlation based pulse scheme, where the relaxation delay is incorporated in the first insensitive nuclei enhanced by polarization transfer step ^[7]. The $^1H_N-\Gamma_2$ rates are given by the difference in $^1H_N-R_2$ rates between the paramagnetic and diamagnetic samples.

3. 5. Isothermal titration calorimetry (ITC)

The equilibrium dissociation constants between EIN (or EIN mutants) and HPr were measured in 20 mM Tris-HCl buffer, pH 7.4, at 37°C using an iTC200 calorimeter (GE Healthcare). 0.1 mM EIN (or EIN mutants) was placed in the cell

and titrated with 1 mM HPr in the syringe. Twenty consecutive 2 μ L aliquots of proteins were titrated into the cell. The duration of each injection was 4 s, and injections were made at intervals of 150 s. Data analysis was performed using the Origin software provided with the instrument.

3. 6. Circular dichroism (CD) spectroscopy

CD was conducted at 25 °C using a with ChirascanTM–plus CD spectrometer. Wave scans were acquired by sampling data at 1 nm intervals between 190 and 260 nm for far UV CD measurement. Far UV CD spectroscopy was carried out using 10 μ M of EIN and EIN mutants in 20 mM Tris–HCl, pH 7.4, using a 1–mm quartz cuvette. Each far UV CD spectrum was obtained from an average of three scans and the results were presented as mean residue ellipticity (deg·cm²/dmol) at each wavelength.

4. Results

4.1. Rational mutation of EIN

I designed mutants that could obstruct the encounter complex formation between EIN and HPr. If the equilibrium constant based on the specific complex formation is changed due to the mutation, unformed the encounter complex against mutation would have the possibility of containing a critical role on protein–protein association and/or dissociation. For selecting mutation residues, it is necessary to analysis the binding module of EIN and HPr. Previous studies showed that EIN largely employed its negatively charged surfaces to interact with the positively charged surfaces of HPr in the encounter complex ensemble ^[3]. Helices α_2 , α_2' , α_3 , and α_4 of EIN and the helices α_1 and α_2 of HPr were directly involved in forming EIN–HPr complex ^[8]. And I would have to identify PRE profiles of EIN–spinlabeled HPrs (E5C, E25C, and E32C) for designing EIN mutants. There were regions with intermolecular $^1\text{H}_\text{N}$ – T_2 rates by E5C, which are residues 110–137 and 59–97 of EIN. $^1\text{H}_\text{N}$ – T_2 rate of E25C and E32C were existed on residues 23–37, 50–92, 105–124, 183–189, and 241–249 of EIN and residues 20–71, 73–140, 184–189, and 232–249 of EIN, respectively ^[3]. Therefore, I designed EIN mutants in which surface–exposed negatively charged aspartate or glutamate residues were replaced by positively charged lysine residues to perturb the electrostatic interaction. The EIN mutation sites were selected using the following criteria:

- 1) Residues that exhibited PREs attributable to the encounter complexes, as identified in PRE profiles using spinlabeled HPr at positions E5C, E25C, or E32C ^[3,9]
- 2) Residues that were located away from the binding interface with HPr.

Five EIN mutants were designed as follows: EIN mut1 (Asp119→Lys, Asp120→Lys, and Glu121→Lys) was designed based on the PRE profile of the EIN–E25C complex; EIN mut2 (Glu116→Lys and Glu117→Lys) and EIN mut3 (Glu81→Lys, Glu86→Lys) were designed based on the PRE profile of the EIN–E5C complex; EIN mut4 (Asp34→Lys, Asp37→Lys, and Glu41→Lys) and EIN mut5 (Asp63→Lys and Asp70→Lys) were designed based on the PRE profile of the EIN–E32C complex (Table 3). EIN mut1 included mutations in the loop between the $\alpha 3$ and $\alpha 4$ helices, and the other four EIN mutants included mutations as following: EIN mut2–helix $\alpha 3$; EIN mut3–helices $\alpha 2$, and $\alpha 2'$; EIN mut4–helix $\alpha 1$; EIN mut5–helices $\alpha 1$, and $\alpha 2$ (Table 3). The positions of the mutation sites in each EIN mutant are shown in the three–dimensional structure of the EIN–HPr complex in Figure 5.

EIN and EIN mutants were purified by the purification steps (Figure 3), and obtaining purified EIN samples (table 3). HPr and HPr mutants (E5C, E25C, and E32C) also purified by the purification steps (figure 4). HPr was obtained for 480 mg/ml, and the HPr mutants yielded range from 132–157 mg/ml. These samples were used on ITC, CD, and NMR experiments.

Table 3. Residues that were mutated in each EIN mutant, the positions of the mutation sites in the EIN structure, the spin-labeled HPr used to design the EIN mutants, the product yield of EIN mutants.

EIN mutants	Positions of the mutation sites in the EIN structure	Spin-labeled HPr	Protein yield (mg/ml)
EIN mut1 (D119K, D120K, E121K)	loop between helices $\alpha 3$ and $\alpha 4$	E25C	24
EIN mut2 (E116K, E117K)	helix $\alpha 3$	E5C	14
EIN mut3 (E81K, E86K)	helices $\alpha 2$ and $\alpha 2'$	E5C	50
EIN mut3a (E81Q, E86Q)	helices $\alpha 2$ and $\alpha 2'$	E5C	112
EIN mut3b (E81A, E86A)	helices $\alpha 2$ and $\alpha 2'$	E5C	138
EIN mut4 (D34K, D37K, E41K)	helix $\alpha 1$	E32C	184
EIN mut5 (E63K, E70K)	helices $\alpha 1$ and $\alpha 2$	E32C	47

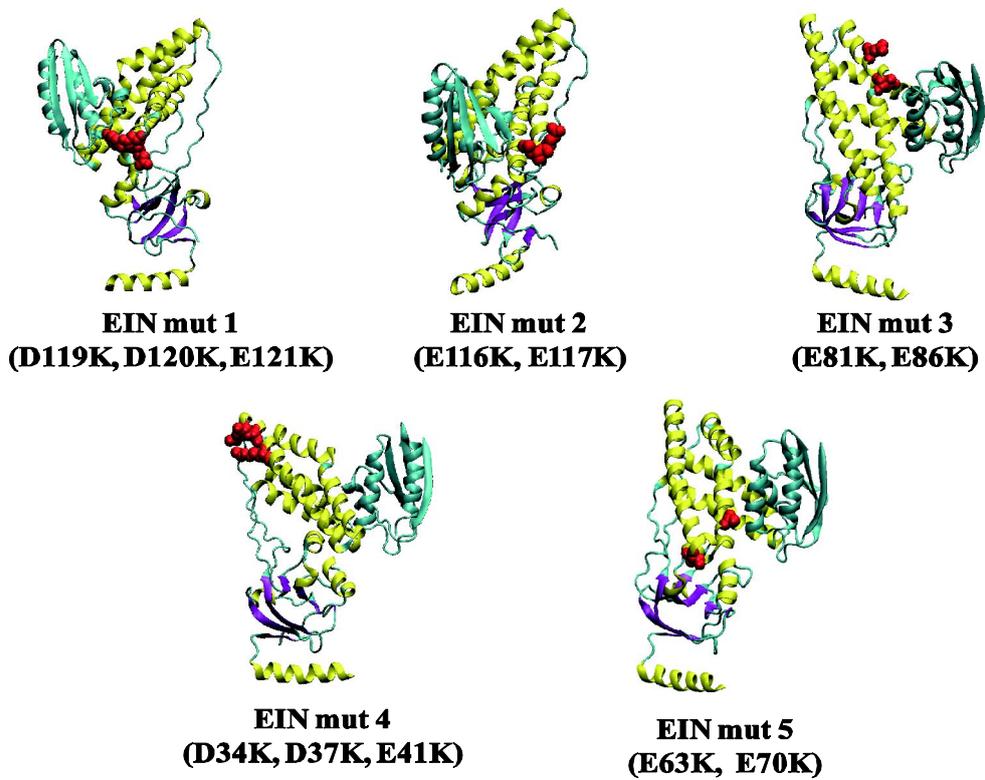


Figure 5. Mutation sites of the five EIN mutants. EIN–HPr complex is color–coded ribbon representation of EIN (helix: yellow and β sheet: purple) –HPr (cyan) complex. The mutation sites introduced to EIN are shown with space–filling models (red).

4. 2. Analysis of secondary structure using circular dichroism (CD) spectroscopy

Prior to isothermal titration calorimetry (ITC), PRE, and other NMR experiments, I performed the circular dichroism (CD) experiment to confirm the secondary structure of the EIN mutants. The CD spectroscopy is one of the most general and basic tools to study protein folding. In proteins, secondary structures such as α -helices and β -sheets are chiral, and thus absorb circularly polarized light. I obtained the CD spectra of EIN and EIN mutants to identify whether the introduced mutations perturbed the original secondary structures of EIN. The CD spectrum of EIN showed that the well-folded structure appeared negatively strong ellipticity value by α -helices and β -sheets of EIN at wavelengths of 208 nm and 220 nm. The CD spectra of EIN mut1, however, showed negatively the strong value at the wavelength of 201 nm, which the secondary structure of EIN mut1 was unordered. It indicated that the secondary structures were significantly disrupted by the mutation of EIN mut1, whereas the other four mutants maintained the secondary structure profiles similar to that of EIN (Figure 6). In the EIN mut1, mutations at Asp119→Lys, Asp120→Lys, and Glu121→Lys most likely disrupted the salt bridges between the carboxylic acid side chains and the neighboring ϵ -amino group (NH_3^+) of Lys124, which could be deleterious to the proper folding of the protein. Therefore, well-ordered mutants except for EIN mut1 proceeded on further studies such as ITC and NMR experiments.

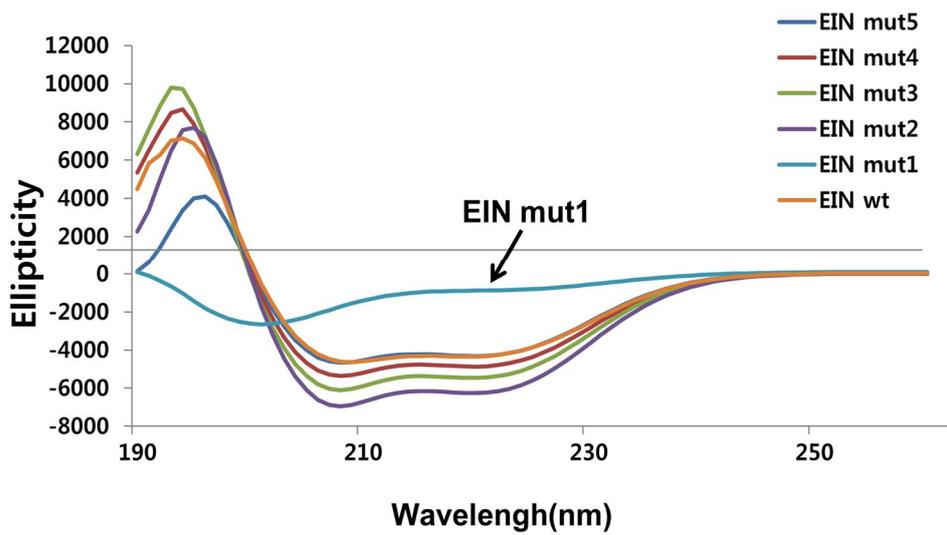


Figure 6. The CD spectra of the EIN and EIN mutants. All CD spectra were collected at 25°C using a JASCO J-801 CD spectrometer.

4. 3. Binding affinity between EIN mutants and HPr

For analyzing influences on the equilibrium constant by the mutation, I examined the equilibrium binding constant between EIN (or EIN mutants) and HPr using ITC. In previous reports, equilibrium dissociation constant (K_D) between EIN and HPr was 3.7 μM in 20 mM Tris-HCl, pH 7.4, at 40°C using a high-precision VP-ITC calorimetry system.^[9] I, however, was measured to be 4.8 μM in the identical buffer at 37°C. The subtly different K_D would be generated because of the temperature condition. EIN mutants generally showed a reduced binding affinity for HPr, but the magnitude of the reduction varied among the mutants (Table 4). It means that the encounter complexes obstructed by the mutations are related to the equilibrium constant upon the specific complex formation, varying the degree of influence on forming the stereospecific complex. EIN mut2 and EIN mut4 showed the approximately twofold decrease of K_D (9.9 and 8.4 μM , respectively) for HPr, while EIN mut5 exhibited a similar binding affinity (K_D : 5.2 μM) for HPr to that of EIN (Table 4). EIN mut3 showed the largest reduction among the binding affinity constants of EIN mutants for HPr, with a K_D of 51.0 μM , a factor of ten higher than that of the EIN-HPr complex. In the stereospecific complex of EIN and HPr characterized by van Waals and electrostatic interactions, I further investigated EIN mut3 variants where the negatively charged residues were replaced by neutral residues instead of the lysines (Glu81→Lys and Glu86→Lys), because the mutation that changed negatively charged aspartate to positively charged lysine was artificial alteration. When Glu81 and Glu86 were replaced by isosteric glutamine (EIN mut3a) or alanine residues (EIN mut3b), the K_D values were 13.9

uM and 16.5 uM in the identical condition, respectively; hence the neutral mutations made an intermediate reduction of the equilibrium binding. The binding affinities of EIN mut3a and 3b showed approximately threefold increase compared to EIN mut3. It indicated that the encounter complexes on the mutation sites of EIN mut3 were related on specific binding of EIN–HPr complex. Therefore, I progressed on the PRE experiment for visualizing the disappeared encounter complex ensemble upon the mutation

Table 4. Thermodynamic parameters for the interaction between EIN (or EIN mutants) and HPr.

	K_D ^[a] (μM)	ΔG ^[b] (kcal/mol)	ΔH ^[c] (kcal/mol)	ΔS ^[d] (cal/mol/deg)
EIN	4.8 ± 0.7	-7.6 ± 0.09	3.7 ± 0.1	36.3 ± 0.4
EIN mut2 (E116K, E117K)	9.9 ± 0.9	-7.1 ± 0.06	3.5 ± 0.1	34.3 ± 0.4
EIN mut3 (E81K, E86K)	51.0 ± 0.7	-6.1 ± 0.08	5.7 ± 0.4	37.9 ± 1.3
EIN mut3a (E81Q, E86Q)	13.9 ± 1.8	-6.9 ± 0.08	4.1 ± 0.2	36.3 ± 0.7
EIN mut3b (E81A, E86A)	16.5 ± 2.0	-6.8 ± 0.07	4.3 ± 0.1	37.8 ± 0.4
EIN mut4 (D34K, D37K, E41K)	8.4 ± 1.0	-7.2 ± 0.07	4.2 ± 0.1	36.9 ± 0.4
EIN mut5 (E63K, E70K)	5.2 ± 0.4	-7.6 ± 0.05	4.2 ± 0.1	37.8 ± 0.4

[a] Equilibrium dissociation constant; [b] binding free energy; [c] binding enthalpy; [d] binding entropy.

4. 4. Backbone assignment of EIN mut3

I determined the backbone assignment of EIN and EIN mutants for analyzing the PRE rates. I performed the three-dimensional triple-resonance through-bond scalar correlation HNCACB and CBCA(CO)NH experiments using 0.3 mM $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled EIN (or EIN mutants) $-0.45\text{ mM }^{12}\text{C}, ^{14}\text{N}$ -HPr complex for the backbone assignment in 20 mM Tris-HCl (pH 7.4). The peaks of three dimensional NMR spectra were picked, and clustered by the spin-system patterns and matched using the PIPP^[5] program. Sequential assignment was achieved by verifying and linking of the individual spin-system pattern obtained from the triple NMR experiments.

For confirming the mutation residues of EIN mut3 (E81K and E86K), the lysine residues were carefully assigned by 2D ^1H - ^{15}N transverse relaxation optimized spectroscopy (TROSY) correlation spectrum of [^{15}N -Lys]-EIN mut3, which were prepared using amino acid selective isotope labeling with ^{15}N -lysine as a precursor (Figure 7). The approximately 80 % of backbone assignment upon EIN mut3-HPr was confirmed using the 2D and 3D-NMR spectroscopy. The ^1H - ^{15}N TROSY spectrum showed that large chemical shift changes were reasonably on residues around the mutation sites at helix $\alpha 1$ (Figure 8). It means that the mutation sites of EIN mut3 do not alter the secondary structure and folding of EIN mut3. However, it is necessary to confirm the EIN conformational change. I proceeded the next experiment such as chemical shift perturbation (CSP), residual dipolar coupling (RDC), and R_2 relaxation.

4. 5. Investigation of conformation and dynamics by NMR

I examined R_2 relaxation parameters and residual dipolar couplings (RDCs) of amide groups for EINmut3, EINmut3a, and EINmut2. And for mapping detail chemical shift changes, I performed chemical shift perturbation (CSP) experiment against EIN mut3, EIN mut3a, and EIN mut2 (Figure 9). The CSP of EIN mut3 and EIN mut3a were mostly localized to the mutation sites (residue 81 and 86) and its close vicinity that existed from residue 46 to residue 52 on helix α_1 . The chemical shift perturbations upon EIN mut2 were placed on the mutation sites (residue 116 and 117 at helix α_1) (Figure 9). It indicates that the mutation sites of EIN mut3, 3a, and mut2 did not change other secondary structure and protein folding of the EIN mutants. Also, the ratio of the R_2 relaxation parameters of amide protons between individual mutants and EIN were uniform and close to unity, indicating that the dynamics of the protein are minimally perturbed by the mutations (Figure 10).

Finally, to assess whether the mutations affected the folding of the backbone in mutants, I proceeded to measure backbone amid RDCs on EIN mutants in a liquid crystalline medium of phage *pfl* (11 mg/ml). RDCs of fixed bond vectors, such as the backbone N–H bond vector, are dependent on the orientation of the bond vectors relative to the alignment tensor and provide a very sensitive indicator of changes in relative domain orientations.^[10, 11] The magnitude and orientation of the alignment tensors obtained by single value decomposition (SVD) to obtain a best fit between the observed RDCs, and those calculated from the atomic coordinates of EIN (Protein Data Bank code 3EZA)^[8] were the same within experimental error. In addition, the RDC R-factors (R_f) obtained from EIN mut3, EIN mut3a, and EIN

mut2 were 15.4 %, 15.1 %, and 15.1 %, respectively (Figure 11). The agreement between the measured and calculated RDCs from the residues in helices α_2 and α_2' , in particular, mutated Qlu81 and Qlu86, was as good as the agreement over the rest of the structure. It indicated that both the local and global structures remained unperturbed. Namely, the structure and dynamics in whole protein, as well as specific binding interface, were not changed in EIN mut3 EIN mut3a, and EIN mut2. As a result, I could confirm the EIN mutants conformation similar to EIN conformation via CSP, R_2 relaxation, and RDC.

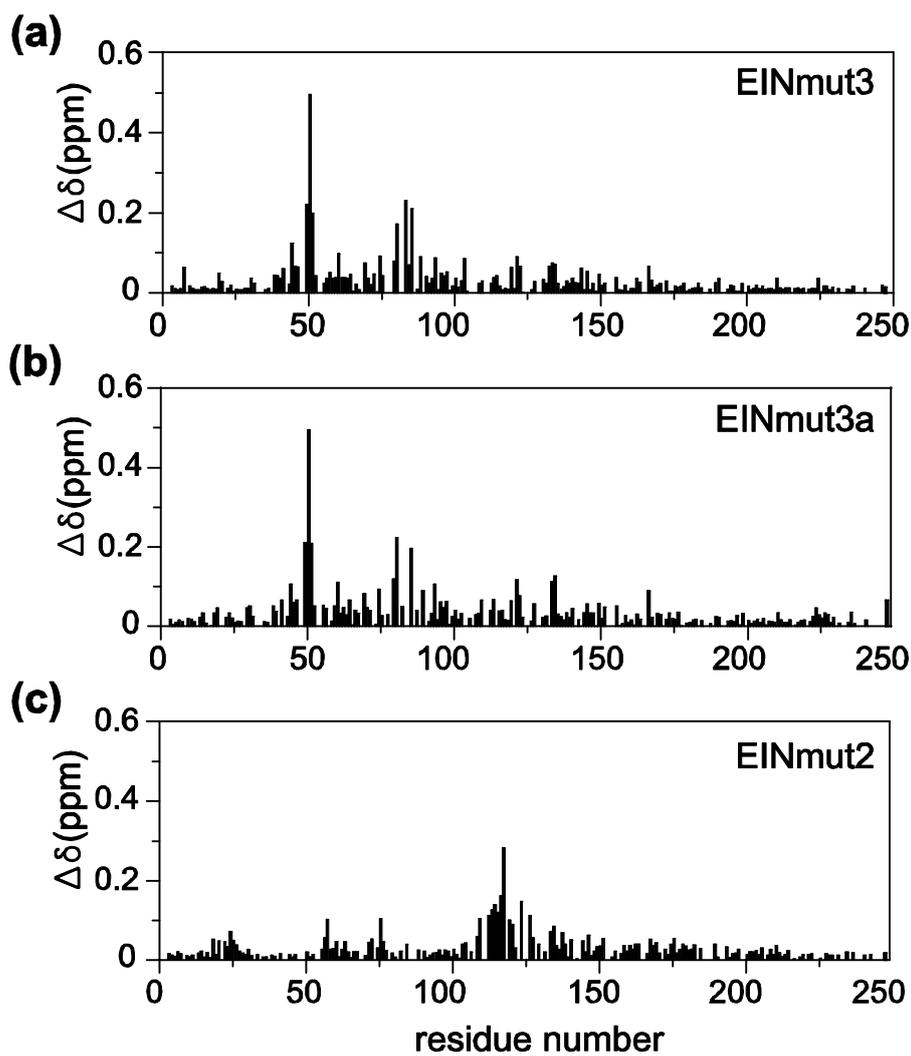


Figure 9. Chemical shift perturbation (CSP) of (a) EINmut3, (b) EINmut3a, and (c) EINmut2 compared to EIN. Weighted average $^1\text{H}_\text{N}/^{15}\text{N}$ chemical shift perturbation ($\Delta\delta = [(\Delta\delta_{\text{HN}})^2 + (\Delta\delta_{\text{N}})^2/25]/2)^{1/2}$) was plotted as a function of residue number.

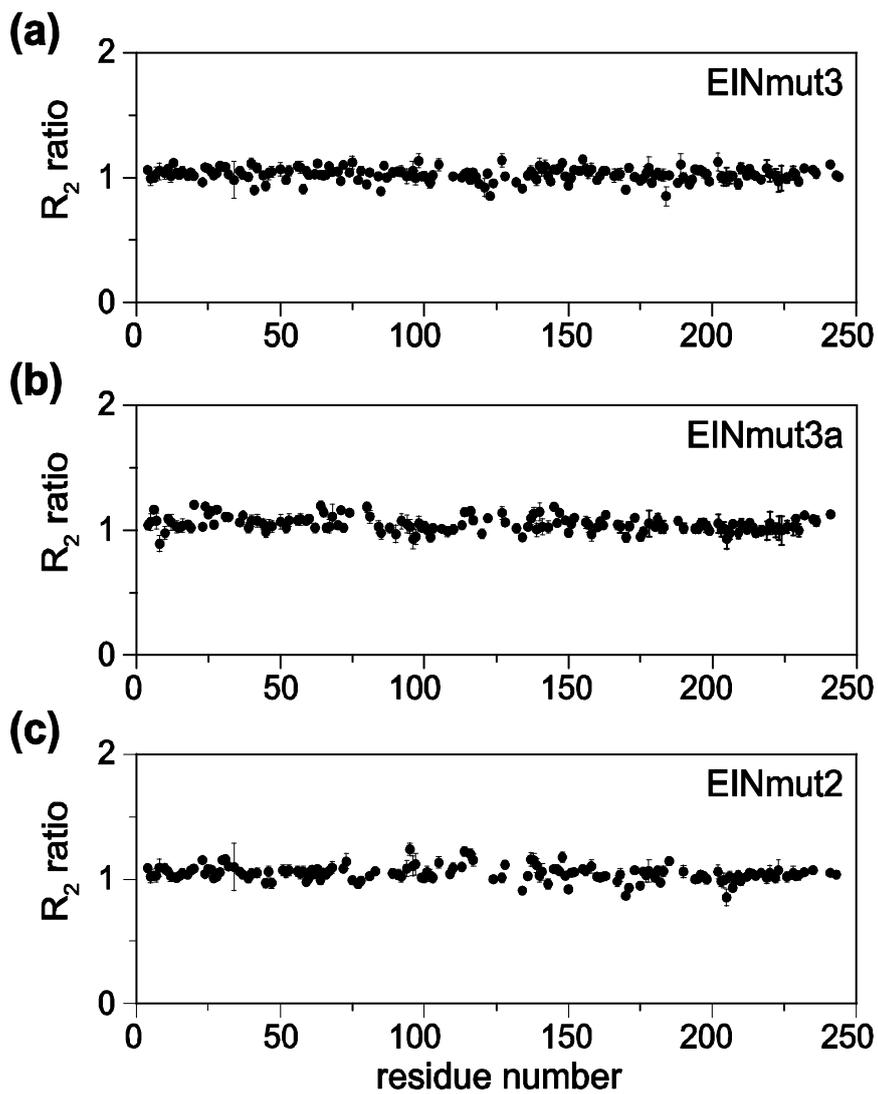


Figure 10. Ratio of R_2 relaxation parameters of amide protons (a) EIN mut3, (b) EIN mut3a, and (c) EIN mut2 compared to EIN as a function of residue number (R_2 ratio = $R_2_{\text{EINmut}} / R_2_{\text{EIN}}$).

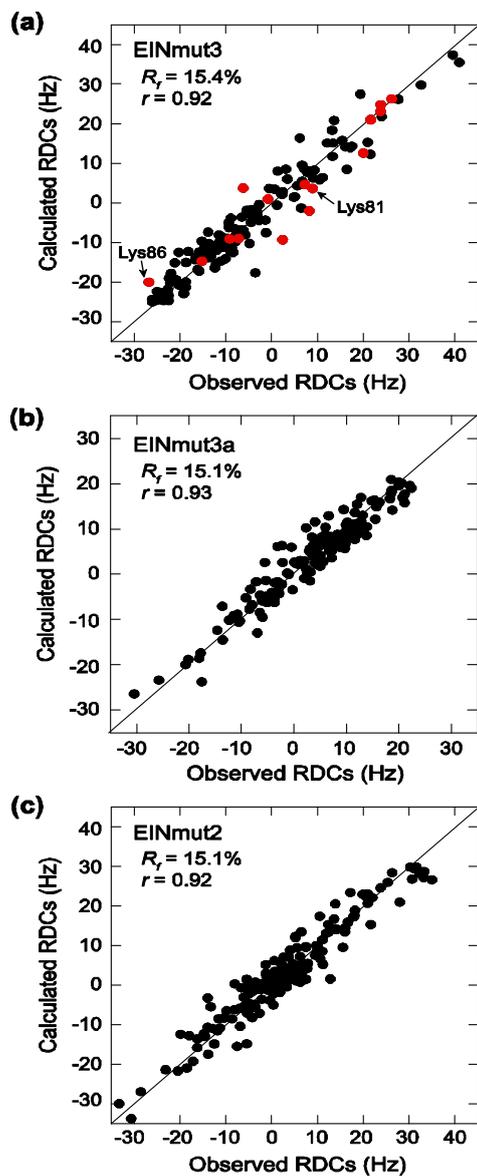


Figure 11. RDC analysis of (a) EIN mut3, (b) EIN mut3a, and (c) EIN mut2 in complex with HPr. It is a comparison of the observed and calculated $^1D_{NH}$ RDCs obtained by singular value decomposition (SVD) with the NMR coordinates of the EIN–HPr complex (PDB ID 3EZA). The values of RDC R-factors and the correlation coefficients are shown. RDCs from the residues in helices α_2 and α_2' are shown in red, and Lys81 and Lys86 of the mutation sites are annotated in (a).

4. 6. Visualization of encounter complex in EINs–HPr

The EIN mutants examined in this study were designed in view of the PREs associated with the encounter complex ensemble. Also, the mutation sites were selected such that they would not interfere with the binding interface in the stereospecific complex. In other words, I found mutation sites to change encounter complex ensemble and equilibrium constant without altering binding interface and conformational structure change of EIN mutants. The fact that the EIN mutants displayed a varying degree of reduction in their equilibrium binding with HPr suggests that the encounter complexes detected by PRE contribute to the specific complex formation to differing degrees, depending on their locations. Mutations that perturb the formation of productive encounter complexes likely have a larger impact on the equilibrium binding than those that perturb the formation of non-productive encounter complexes. To investigate how these mutations affected the formation of the encounter complex, I measured the intermolecular PREs of EIN, EIN mut3, EIN mut3a, and EIN mut2 in complex with E5C. Therefore, I would understand the target search process through the productive and/or non-productive encounter complexes by measured PREs. The NMR sample concentration used for the PRE measurement was not increased beyond 0.5 mM, since unwanted PRE effects arising from transient. Random collisions (solvent PRE effect) may become significant at a higher concentration.^[12] The NMR samples were used to 0.3 mM $^2\text{H}/^{15}\text{N}$ -labeled EIN (or EIN mutants) and 0.5 mM HPr in 20 mM Tris–HCl, pH 7.4. HPr was spinlabeled by conjugating EDTA– Mn^{2+} to HPr by way of disulfide linkage to the surface-engineered cysteine residue at

position E5C. The spinlabeled HPr (E5C) was prepared by the method (chapter 3.3.3). The spinlabeled HPr was confirmed using SDS–PAGE and ESI–Mass spectroscopy (Figure 12). I confirmed whether EDTA and Mn^{+2} were conjugated on HPr (measured spinlabeled HPr molecular weight: 9503.7).

The PRE rates measured in this concentrated sample were in good agreement with those obtained using the dilute sample. The intermolecular backbone 1H_N PRE rates (Γ_2) were obtained by measuring the differences between the R_2 relaxation rates of the paramagnetic and diamagnetic samples at 37°C using a TROSY–based pulse scheme at a 1H frequency of 800 MHz.^[7] Experimental PREs of EIN, EIN mut3, and EIN mut3a are shown as red circles, and calculated PREs from the EIN–E5C complex structure are shown as black lines in Figure 13a–c. Large PREs from residues 115 to residue 127 were predicted from the stereospecific complex structure, and PREs outside of this region were attributed to nonspecific encounter complexes. Measured large PREs between residue 115 and 127 on EIN, EIN mut3, and EIN mut3a complexes were consistent with the calculated 1H_N – Γ_2 rates. By contrast, PREs attributed to the nonspecific encounter–complex ensemble had clearly different PREs profiles between the EIN and EIN mutants. Large PREs appeared between residues 62 and 89 in the EIN–Hpr (E5C) complex, whereas large PREs showed up only between residues 62 and 75 in the EIN mut3–HPr (E5C) complex. Notably, the PREs attributable to the encounter complex ensemble largely disappeared near the mutation sites in EIN mut3 (residues 81 and 86), as highlighted by yellow dashed boxes in the PRE profiles (Figure 13a, b). EIN mut3a exhibited PREs attributable to the encounter complexes smaller than EIN but larger

than EINmut3 near the mutation site (Figure 13c). I note that the intermediate magnitude of the PREs of EIN mut3a correlates with its moderate reduction in binding affinity for HPr compared to EIN mut3. This correlation suggests that encounter complex formation near the mutation site is mainly driven by electrostatic interactions, and is also directly linked to the specific complex formation. The three-dimensional structure of the EIN–HPr complex color-coded by PRE profiles clearly demonstrates that the encounter complexes mostly disappeared at the mutation site (Figure 14). I also measured PREs from the EIN mut2–HPr (E5C) complex, and the PREs predicted from the stereospecific complex appeared similar between the EIN mut2–HPr (E5C) and EIN–HPr (E5C) complexes (Figure 13d). On the other hand, PREs attributable to the nonspecific encounter complex ensemble appeared in a wider region (residues 47–89) than EIN (residues 62–89), and also the magnitude of PREs in this region was slightly larger in EIN mut2. The PREs from the EIN mut2–HPr (E5C) complex were reproducible from separately prepared samples. This indicates that the mutation in EIN mut2 apparently affected the encounter complex formation in a region remote from the mutation site. Because the equilibrium binding of EIN mut2–HPr was twice as weak as that of EIN–HPr, I suppose that the emerging encounter complexes in EIN mut2 did not promote the equilibrium binding. In fact, PREs in EIN mut2 mainly increased on helix α_1 , the mutation of which showed little change in the equilibrium binding, as shown in the EIN mut5–HPr complex (Table 4). It is notable that surface-charge mutations can influence the encounter complex formation directly at the mutation site (e.g. EIN mut3), or indirectly at a remote site

(e.g. EIN mut2).

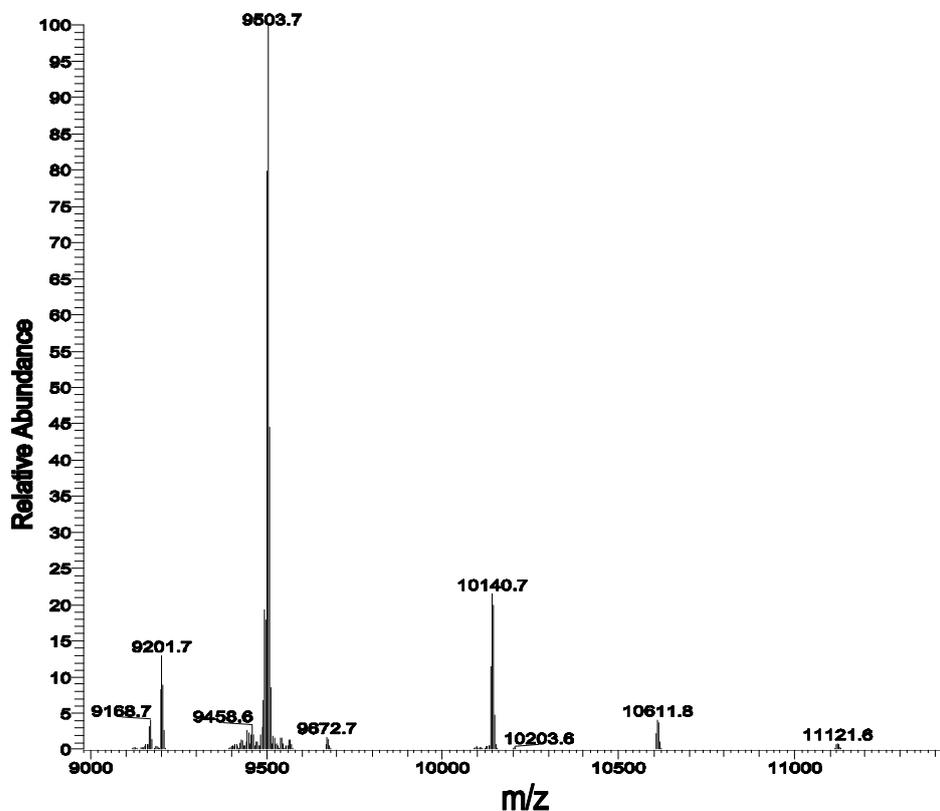


Figure 12. Mass Spectrum data of EDTA-Mn⁺² Conjugated HPr at E5C. The measured spinlabeled HPr molecular weight was 9503.7 (calculated spinlabeled HPr: 9495.58). Mass spectrum was collected in 20mM Tris-HCl, pH 7.4 using electron spray ionization (ESI) Mass spectroscopy.

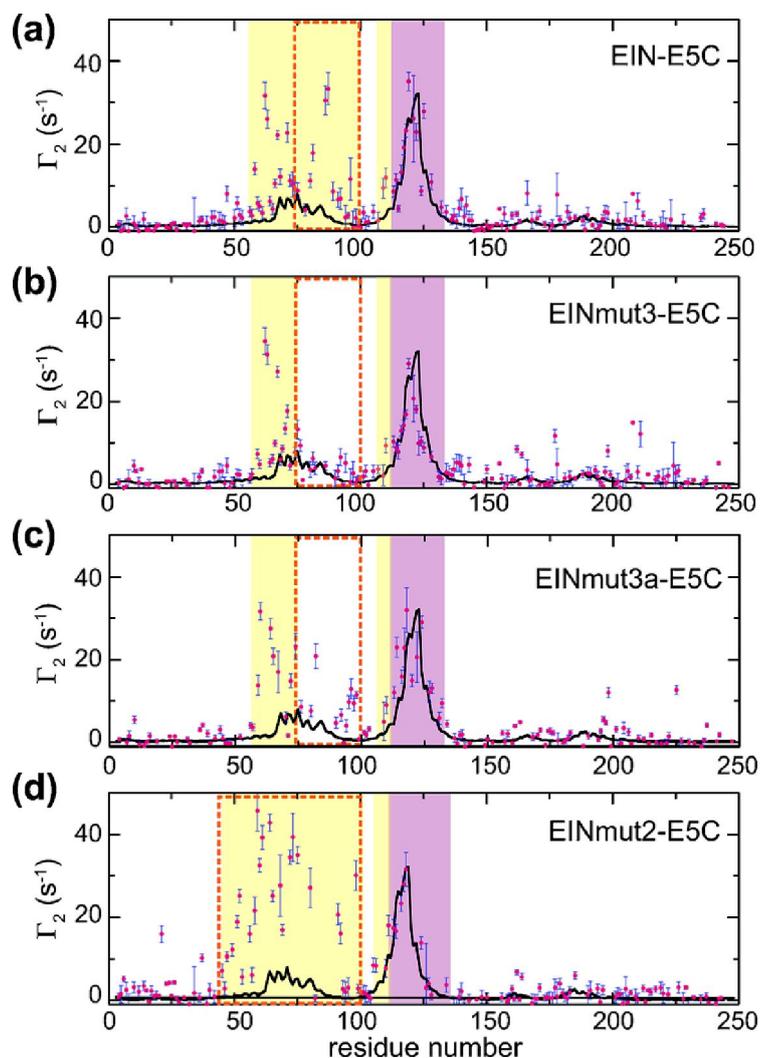


Figure 13. Comparison of the observed and calculated intermolecular PREs for the a) EIN-HPr (E5C), b) EIN mut3-HPr (E5C), c) EIN mut3a-HPr (E5C), and d) EIN mut2-HPr (E5C) complexes. The observed intermolecular $^1\text{H}_\text{N}$ - Γ_2 rates (red circles) and calculated values (black line) are shown with error bars. PREs arising from the specific complex structure are highlighted in the purple boxes, and PREs attributable to the nonspecific encounter complex ensemble are highlighted in the yellow boxes. Differences between the PRE profiles of the EIN and EIN mutants are shown in the dashed red boxes.

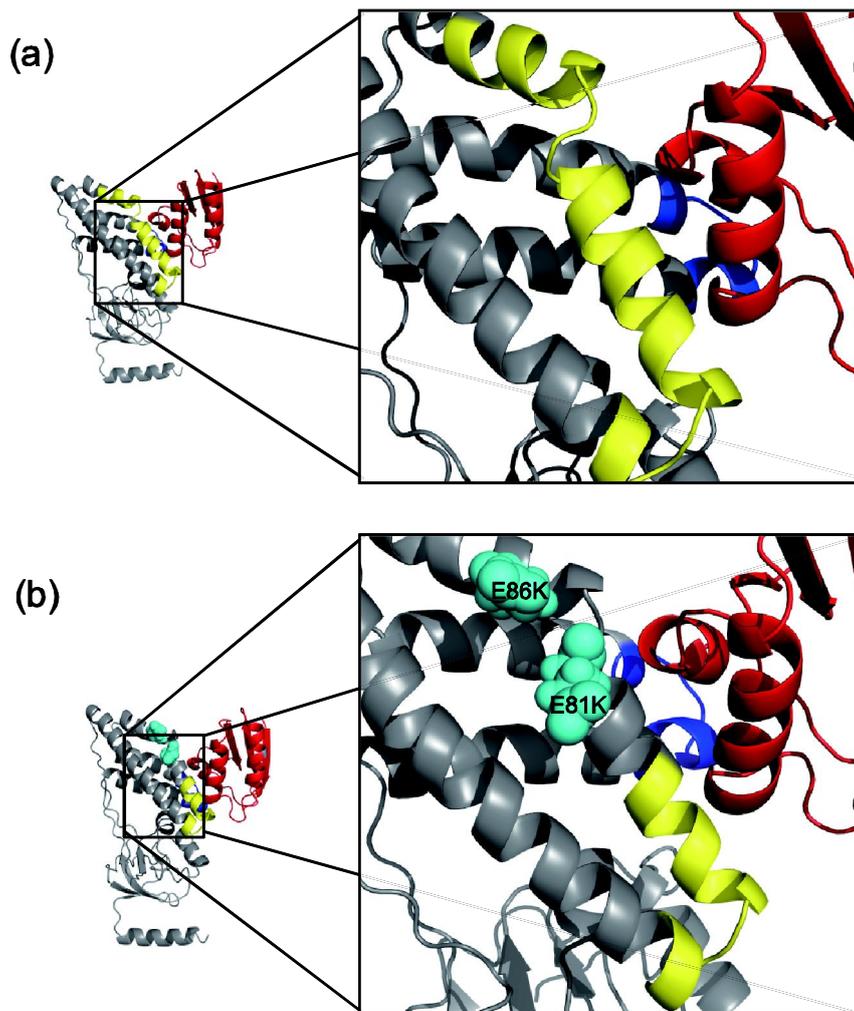


Figure 14. The PRE profiles are color-coded in the solution structure of the EIN-HPr complex (PDB ID 3EZA) for the (a) EIN-HPr (E5C), and (b) EIN mut3-HPr (E5C) complexes. EIN is shown in gray and HPr is shown in red. The mutation sites in the EIN mut3 (E81K and E86K) are shown as a space-filling model representation in cyan. Intermolecular PREs are color-coded on the structure in blue or yellow in the same manner as the PRE profiles. The expansions illustrate the mutation site disappeared on PREs.

4. 7. Productive and non-productive encounter complexes

between EIN and HPr

I further investigated whether the reduced equilibrium binding of EIN mutants resulted from changes in the association^[10] or dissociation (k_{off}) rate constants by qualitative line-shape analysis of ^1H - ^{15}N correlation spectra during titration. When I titrated ^{15}N -HPr with EIN, the sidechain amide resonance of Gln51 showed a continual shift with line broadening, which was maximal when the fractional population of the complex was approximately 0.3 (Figure 15a). This indicates an exchange rate (k_{ex}) on the fast side of intermediate exchange, suggesting that $k_{\text{ex}} \approx 7500 \text{ s}^{-1}$ ($2\Delta\omega_{\text{HN}} < k_{\text{ex}} < 3\Delta\omega_{\text{HN}}$).^[13] For a second-order exchange, $k_{\text{ex}} = k_{\text{off}} + k_{\text{on}}[\text{EIN}] = k_{\text{off}} + k_{\text{off}}[\text{EIN-HPr}]/[\text{HPr}] = k_{\text{off}}(1 + P_{\text{EIN-HPr}}/P_{\text{HPr}})$, where $[\text{EIN}]$, $[\text{HPr}]$, and $[\text{EIN-HPr}]$ are the concentrations of free EIN, free HPr, and EIN-HPr complex at equilibrium, respectively, and $P_{\text{EIN-HPr}}$ and P_{HPr} are the fractional population of EIN-HPr complex, and free HPr, respectively. From this equation, k_{off} is estimated as 4400 s^{-1} , and k_{on} is approximately $9.2 \times 10^8 \text{ M}^{-1}\text{s}^{-1}$, which is of the same order of magnitude as the phosphoryl transfer reaction rate constant between Enzyme I and HPr from kinetic measurements.^[14] Because HPr exhibited a similar line-broadening profile during the titration with EIN, EINmut3, and EINmut2, the dissociation rate constants would be comparable between these protein complexes (Figure 15a-c). Hence, the lower degree of binding (higher K_D) of the EIN mutants most likely resulted from the decrease of the association rate constants.

The data presented in this study suggest that the encounter complexes probed by the intermolecular PRE measurements play distinct roles in protein-protein

association. Based on our current results, I propose a model of possible target search pathways during the association between EIN and HPr. The mutation sites of EINmut3 and EINmut4 are located above the specific binding interface, and those of EINmut2 and EINmut5 are at either side of the specific binding interface. The target search pathways in our model are as follows: HPr forms an encounter complex with EIN at a region above the specific binding interface and slides down toward the specific binding site. In addition, HPr can form an encounter complex with EIN at the side of the specific binding interface, but lateral sliding toward the specific binding interface is less efficient. In particular, sliding from the mutation site of EIN mut5 is mostly restricted possibly owing to steric hindrance or less favorable electrostatic interactions guiding the encounters to the specific binding site. Figure 12 summarizes the target search pathways by way of productive encounter complexes, and also unsuccessful attempts of non-productive encounter complexes. Note that the electrostatic surface potential of EIN alone does not discriminate between the interfaces for productive and non-productive encounter complexes.

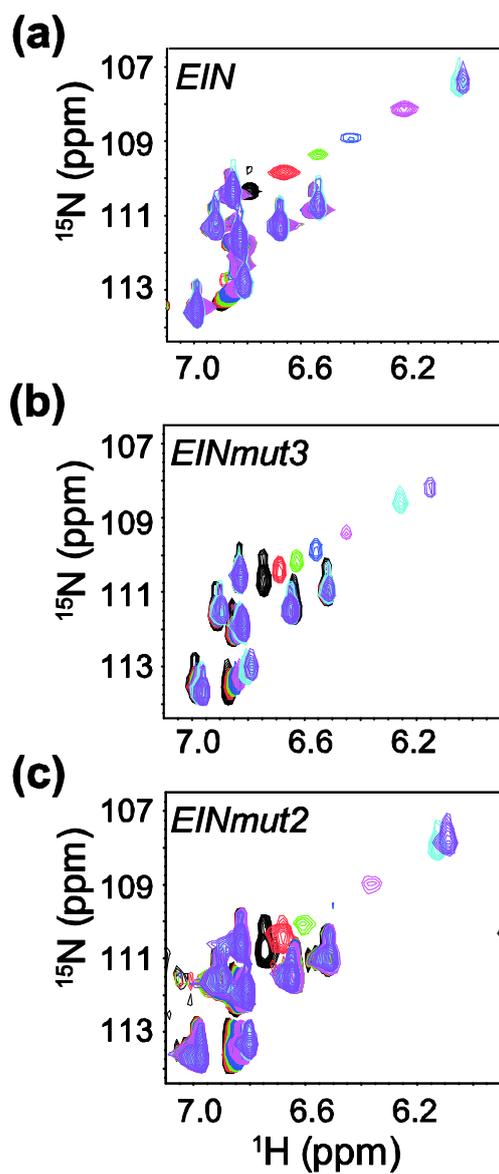


Figure 15. 2D ^1H - ^{15}N HSQC spectra of ^{15}N -HPr titrated with a) EIN, b) EInmut3, and c) EInmut2. HPr (0.5 mM) was titrated with stoichiometric ratios of EIn (mut) protein of 0 (black), 0.1 (red), 0.2 (green), 0.3 (blue), 0.5 (magenta), 1 (cyan), and 2 (violet). The progressive chemical-shift perturbation of Gln51 is characteristic of the fast side of intermediate exchange.

5. Discussions

In study, the encounter complexes probed by the intermolecular PRE measurements and several experiments played important roles in protein–protein association. The designed mutants upon the criteria were cloned, cultured, and purified in our system. And the mutants were estimated on secondary structure and three–dimensional conformation via CD and NMR spectroscopy. For probing the target search process, I selected four mutants except for EIN mut1. Specially, the equilibrium dissociation constant (K_D) of EIN mut3 was measured to the largest reduced K_D value via ITC. I confirmed that EIN mut3 perturbed the encounter complex related on the formation of stereospecific complex by using PRE.

Based on current results, I propose a model of possible target search pathways during the association. The mutation sites of EIN mut3 are placed above the specific binding interface, and those of EIN mut2 and EIN mut5 are at either side of the specific binding interface. The model of the target search pathways are that HPr forms an encounter complex with EIN at a region above the specific binding interface and slides down toward the specific binding site. In addition, HPr can form an encounter complex with EIN at the side of the specific binding interface, but lateral sliding toward the specific binding interface is less efficient. In particular, sliding from the mutation site of EINmut5 is mostly restricted possibly owing to steric hindrance or less favorable electrostatic interactions guiding the encounters to the specific binding site. Figure 16 summarizes the target search

pathways by way of productive encounter complexes, and also unsuccessful attempts of non-productive encounter complexes. Note that the electrostatic surface potential of EIN alone does not discriminate between the interfaces for productive and non-productive encounter complexes.

I infer that initial encounters are driven by long-range electrostatic attractions, but a successful journey from the encounter to the specific binding site requires favorable short-range interactions along the pathway as well. The distinct nature of the encounter complexes has been predicted using Brownian dynamics simulations and kinetic measurements of several protein-protein complexes ^[15]. Because PRE was first used to visualize the encounter complex ensemble, efforts have been made to characterize the physicochemical nature of the encounter complexes and also to understand the roles of the different encounter complexes in a specific complex formation ^[4]. I showed herein that the differential contributions of encounter complexes toward the specific complex formation can now be quantitatively evaluated by rationally designed mutants, and changes in the encounter complex formation of the mutants can be visually monitored. Further, experimental PREs combined with the thermodynamic parameters of equilibrium binding determine how much a particular hot-spot region for encounter complex formation contributes to the specific binding. For example, diminished encounter complexes of EIN mut3 (charge reversal) and EIN mut3a (charge neutralization) accounted for the free energy differences of binding, as much as approximately 1.5 kcal/mol and 0.7 kcal/mol, respectively, compared to EIN. The biological and evolutionary roles of non-productive encounter complexes are yet uncertain. When a protein interacts

with multiple partner proteins, the hot-spot regions that form productive or nonproductive encounter complexes can vary and sometimes may even switch their roles to discriminate different partner proteins. The exploration of target search pathways could eventually be applied to the manipulation of equilibrium binding in protein-protein or protein-ligand complexes without affecting their binding interface.

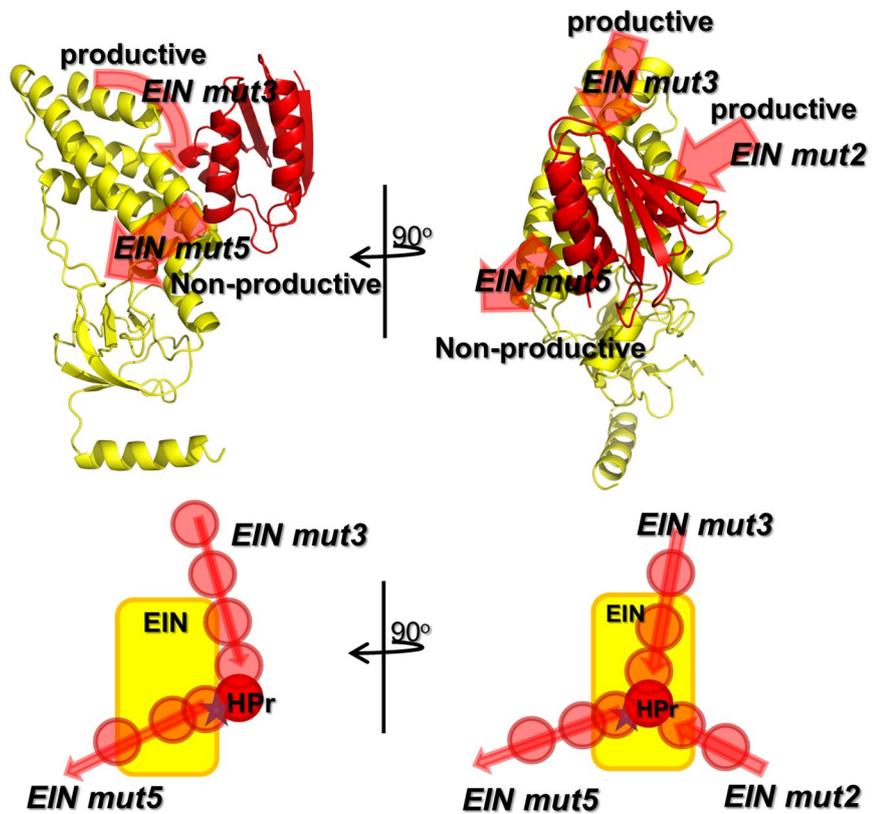


Figure 16. A model of target search pathways of the productive and non-productive encounter complexes between EIN (yellow) and HPr (red). Below panel is illustrated on the model of target search pathway between EIN (yellow square) and HPr (red circle).

6. Conclusion

Protein structure have been investigated by X-ray and NMR methods and provided an insight into the functional mechanism. But the study of a target search process during the complex formation mostly fall behind. Lowly populated, transient encounter complexes appear during target search process, and NMR paramagnetic relaxation enhancement has been applied to detect the encounter complexes. In this study, I designed rational mutation that obstructs the target search process. I performed experiments on PRE and thermodynamics to model the target search pathway, and examined how they alter the population and distribution of the encounter complex. I proposed the model of the target search pathway via these experiments. Based on the model of the target search pathway, it would enhance the understanding and controlling protein-protein interaction. Comprehending the protein-protein interaction helps biological functions of proteins to understand in detail. Also, the model will be supply new idea on protein complex and utilized on therapeutic application.

PART II.

**STRUCTURAL BASIS FOR TARGET
RECOGNITION OF
 β -HAIRPIN APTAMER-LIKE PEPTIDE**

1. Abstract

The fibronectin extradomain B (EDB), a naturally occurring marker of tissue remodeling and angiogenesis, is expressed in the majority of aggressive solid human tumors. EDB has been screened to yield a small aptamer-like peptide aptide (APT_{EDB}) of high affinity. APT is based on a tryptophan zipper scaffold with variable target-binding regions, which can recognize diverse target proteins with high specificity. Here I determined the complex structure of EDB- APT_{EDB} by NMR spectroscopy and investigated the binding mechanism by isothermal titration calorimetry (ITC). The complex structure showed unfolding of a complete C-terminal β strand of EDB, which is replaced by a β strand to APT_{EDB} . The C-terminal β strand of EDB remained unfolded upon APT_{EDB} interaction, exposing large hydrophobic inner surface of EDB for the binding interface. β strand unfolding of EDB likely alters the fibronectin domain structure and the downstream signaling associated with tumor angiogenesis, promising APT_{EDB} as a potential angiogenic modulator.

2. Introduction

2. 1. Structure of fibronectin

Fibronectin ^[16] is one of the largest multi-domain proteins for which domain structure and molecular interactions have been studied in detail.^[17] A soluble form of FN is found in the plasma (300 µg/ml) and an insoluble form is in the extracellular matrix (ECM), to which most of the FN activities in the body have been ascribed: binding to cell surfaces and various compounds, including collagen, fibrin, heparin, DNA, and actin (Figure 1). FN are involved in a number of important functions like wound healing, cell adhesion, blood coagulation, cell differentiation and migration, maintenance of the cellular cytoskeleton, hemostasis and thrombosis, wound healing, oncogenic transformation as well as in tumor metastasis migration ^[18]. These essential functions are underlined by the fact that mouse embryos which are homozygous for the loss of the FN gene have been observed to die at about 8.5 days of gestation. Embryonal death is associated with widespread defects in mesoderm-derived structures, including absence of somites and notochord, as well as developmental defects in the heart and vascular system.^[19] In a soluble form fibronectin (pFN) is abundant in plasma and other body fluids, while insoluble fibronectin (cFN) is a major component of the extracellular matrix (ECM). FN usually exists as a dimer of two nearly identical ~250 kDa (≈ 2500 amino acids) subunits, covalently linked by a pair of disulfide bonds near the C-terminal end of the subunits.^[20-22] Each monomer comprises 3 types of repeating

units, termed typeI (ca. 40 amino acid Residues), typeII (ca. 60 amino acid residues) and typeIII (ca. 90 amino residues), arranged like beads on a string: 12 typeI repeats, 2 typeII repeats and 15–17 typeIII repeats.^[17, 23] These modular repeating units are resistant to proteolysis and contain binding sites for extracellular matrix proteins such as collagen and thrombospondin, cell–surface receptors such as integrins, circulating blood proteins such as fibrin and glycosaminoglycans such as heparin and chondroitin sulphate. FNs are glycoproteins containing up to 9% carbohydrate, depending on the cell source. Glycosylation sites are found predominantly within the typeIII repeats and the collagen–binding domain.

2. 2. Alternatively spliced fibronectin isoforms

Although FNs are generated from the single FN gene, the protein from the FN gene can exist in multiple forms which result from alternative splicing of its primary RNA transcript.^[24, 25] This genetic variation that leads to 20 different isoforms in human FN, by that means generating FNs with different properties of solubility, cell adhesive and ligand–binding, provides cells with the possibility to alter the composition of the ECM in a tissue–specific mode. Alternative splicing generates in three parts of the RNA transcript. Namely, Exon utilization or omission brings about either inclusion or skipping of two typeIII repeats, extradomain B (EDB) and extradomain A (EDA). The EDB is introduced between FN typeIII repeats III7 and III8 (FN7 and FN8), and the EDA inserted between FN typeIII repeats FN11 and FN12. This character of splicing happens in many animals having a spinal column (Figure 1). These extrodomains A and B are almost

always existed on plasma–fibronectin and ECM–fibronectin. A third region of splicing generates at a non–homologous stretch termed V region or IIICS which is segment connected on typeIII. The structural fluctuations at this locale are a lot more complicated and species dependent. In human, five dissimilar V–region variants have been observed. Other modes of splicing have been found in cartilage, where the whole V–region together with III15 and I10 is losing. This form is an unusual monomeric configuration as well as occasionally a homodimer. The splicing design of FN pre–mRNA is changed leading to an increased expression of FN isoforms including sequences of the EDA, EDB, and IIICS. Therefore, these variants are critical due to in transformed cells and malignancies. ^[26-35]

2. 3. Structure of the EDB domain

The three dimensional structure of the alternatively spliced-in EDB domain, (91 amino acids) has been solved by NMR spectroscopy in solution (PDB code, 2FNB).^[36] The EDB structure has two antiparallel β sheets similar to a β sandwich (Figure 2). One β sheet is organized by three β strands (A, B and E), and the other is constituted to four β strands (C, C', F and G). The two β sheets enfold a hydrophobic core of 25 amino acid sidechains. In the whole structure, EDB protein is acidic and includes only two positively charged residues. The most of negatively charged residues are distributed on the protein surface. One portion of residues Ile35, Phe54, Ile78 and Leu80 forms a solvent-exposed hydrophobic patch.^[36] Crystal and solution structures of typeIII domains of FN have been studied previously. The crystal structure of a FN7-10 fragment (PDB code, 1FNF)^[23] exhibits a rod-like molecule with variable relationships between the individual pairs of adjacent domains. These contain the loops located at the N and C termini of each domain. Structural comparisons of EDB with FN7 and 8 shows that differences upon loop structures cause critically structural differences in the relative orientation whether FN7 and FN8 containing EDB or without EDB.^[23]

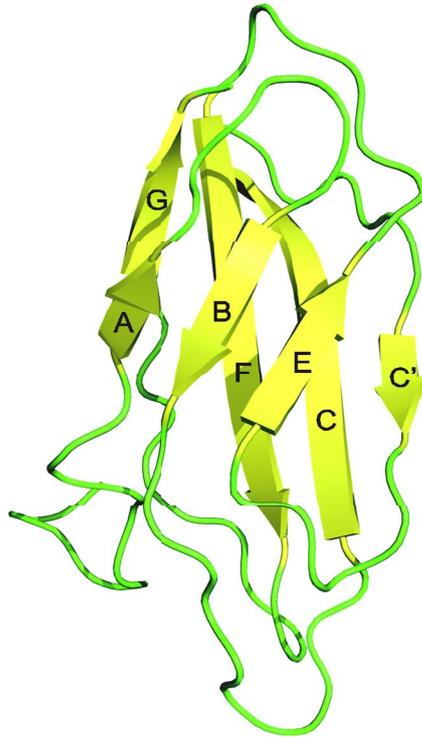


Figure 2. Structure of the EDB domain of fibronectin. The EDB structure represents ribbon drawing of one of the 20 energy-minimized conformers of human EDB (PDB code, 2FNB). The two β -sheets are comprised of three and four β -strands, respectively. The β -strands are annotated A to G following the nomenclature used for other FN typeIII domains.

2. 4. β -hairpin aptamer like peptide (Aptide)

Antibodies have been widely utilized in the field of biopharmaceutical and biomedical applications due to their intrinsic high affinity and specificity toward various targets. However, poor tissue penetration owing to their large size, undesired effectors functions, immunogenicity, costly recombinant production in mammalian cells, and complex intellectual property barriers have led researchers to seek alternatives to antibodies. Aptamer comprises nucleic acids and proteins that have potential biotechnological applications as an alternative to the immunoglobulin from various biochemical assays. Compared to antibodies, aptamer are much smaller in size but possess similar binding affinity and recognition specificity to their target proteins both *in vitro* and *in vivo*.

Phage display was used to screen and isolate the specific aptide for EDB that showed high affinity and specificity.^[37] After the screening and affinity maturation, aptide with nanomolar binding affinity against EDB *in vitro* and specific binding *in vivo* have been obtained. β -hairpin aptamer like peptide (Aptide) are novel class of peptide based molecules that provide a general scaffold of high affinity and specificity against various targets. The aptide comprises a β -hairpin scaffold and two target binding regions. The hairpin scaffold consists of a small tryptophan zipper motif (12 amino acids, SWTWENGKWTWK) that forms a highly stable β -hairpin structure (Figure 3a). The β -hairpin conformation is stabilized by two tryptophan-tryptophan cross-strand pairs that make an edge-to-face interaction. The target binding regions consist of six amino acids (N-term: SSSPIQ and C-term: IIRLEQ) at each end of the hairpin scaffold (Figure 3b).

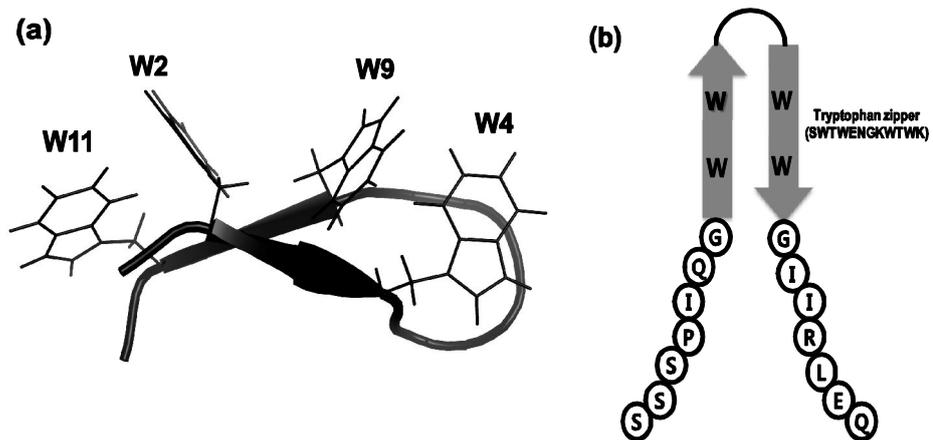


Figure 3. Structure of Trpzip and APT_{EDB}. a) Representative structure of trpzip is consisted of two tryptophan–tryptophan cross pairs (PDB code: 1LE0). b) APT is comprised of tryptophan zipper (arrow) and two ligand binding regions (circle).

3. Materials and Methods

3.1. Cloning

EDB (5–95) and APT_{EDB} (amino acid sequence: SSSPIQG SWTWENGKWTWK GIIRLEQ) were cloned into a modified pET32a vector to express the fusion proteins with a His₆-tag. For separating the fusion protein on the EDB protein, a thrombin cleavage site (LVPRGS) was used. EDB mutations for C-term truncations were introduced using PCR primers that defined the domain lengths, and the new constructs were confirmed by DNA sequencing. The FN8 domain was obtained by gene synthesis and ligated to the pET32a digested with NcoI and XhoI sites

3.2. Culture of EDBs and APT_{EDB}

The expression of the EDB, EDB mutants, and APT_{EDB} domain was performed using the expression vector pET32a in *Escherichia. coli* BL21 cells (DE3, Invitrogen) using LB or minimal medium (with ¹⁵NH₄Cl and/or ¹³C₆-glucose as the sole nitrogen or carbon sources, respectively) (table 1). ¹⁵N-isoleucine, leucine, and glycine were performed on the selective labeling upon EDB. For selective labeling, a medium containing the appropriate unlabeled amino acid mixture (table 2) was prepared, which is based on M9 medium. The cells transformed by the expression vector were cultured at 37°C to an OD₆₀₀ of 0.8, and protein

overexpression was induced by adding isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, followed by induction for 4 h at 37°C and harvested by centrifugation.

3.3. Purification of EDBs and APT_{EDB}

3.3.1 EDB and EDB mutants

The cells were harvested by centrifugation, and resuspended in buffer [20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 1 mM phenylmethanesulfonyl fluoride (PMSF)], and disrupted by emulsiflex and centrifuged at 50,000 x *g* for 20 min. The lysate was centrifuged to remove insoluble debris. The supernatant was loaded on a HisTrap HP column equilibrated with binding buffer [20 mM Tris-HCl, 200 mM NaCl (pH 7.4)] and was eluted with a gradient of imidazole (15–500 mM). The fusion protein buffer was then exchanged against 20 mM Tris, pH 8.0, and 200 mM NaCl, and the thioredoxin- and His₆-tags cleaved using thrombin at room temperature during at least 4 h. After the digestion, thrombin was inactivated by 2 mM of benzamidine, and the cleaved thioredoxin-His₆ tag was removed by loading the reaction onto the HisTrap HP column. The solutions containing protein were separated by a HiLoad Superdex 75 column [20 mM Tris-HCl (pH 7.4), 200 mM NaCl] and finally were purified by Mono Q column. The EDB and EDB mutants were analyzed by SDS-polyacrylamide gel electrophoresis to confirm sample mass and sample purity (Figure 4).

3.3.2 APT_{EDB}

The APT_{EDB} also was purified on similar methods to the purification steps of EDB protein. The cell containing the APT_{EDB} was resuspended in 50 ml of 20 mM Tris, pH 7.4, 200 mM NaCl, and 1 mM PMSF. The suspension was lysed by three passages through a emulsiflex and centrifuged at 40,000 g for 20 min. The supernatant fraction was loaded onto a HisTrap HP column and eluted with a 100-ml gradient of 500 mM imidazole. The fractions containing the APT_{EDB} were confirmed by SDS-PAGE and cleaved using thrombin at identical condition with EDB. The cleaved thioredoxin-His₆ tag was removed using the HisTrap HP column (Figure 5). Further Purification of APT_{EDB} did not process the size exclusion chromatography, because APT_{EDB} had a tendency to bind nonspecifically to the packing materials in the column.

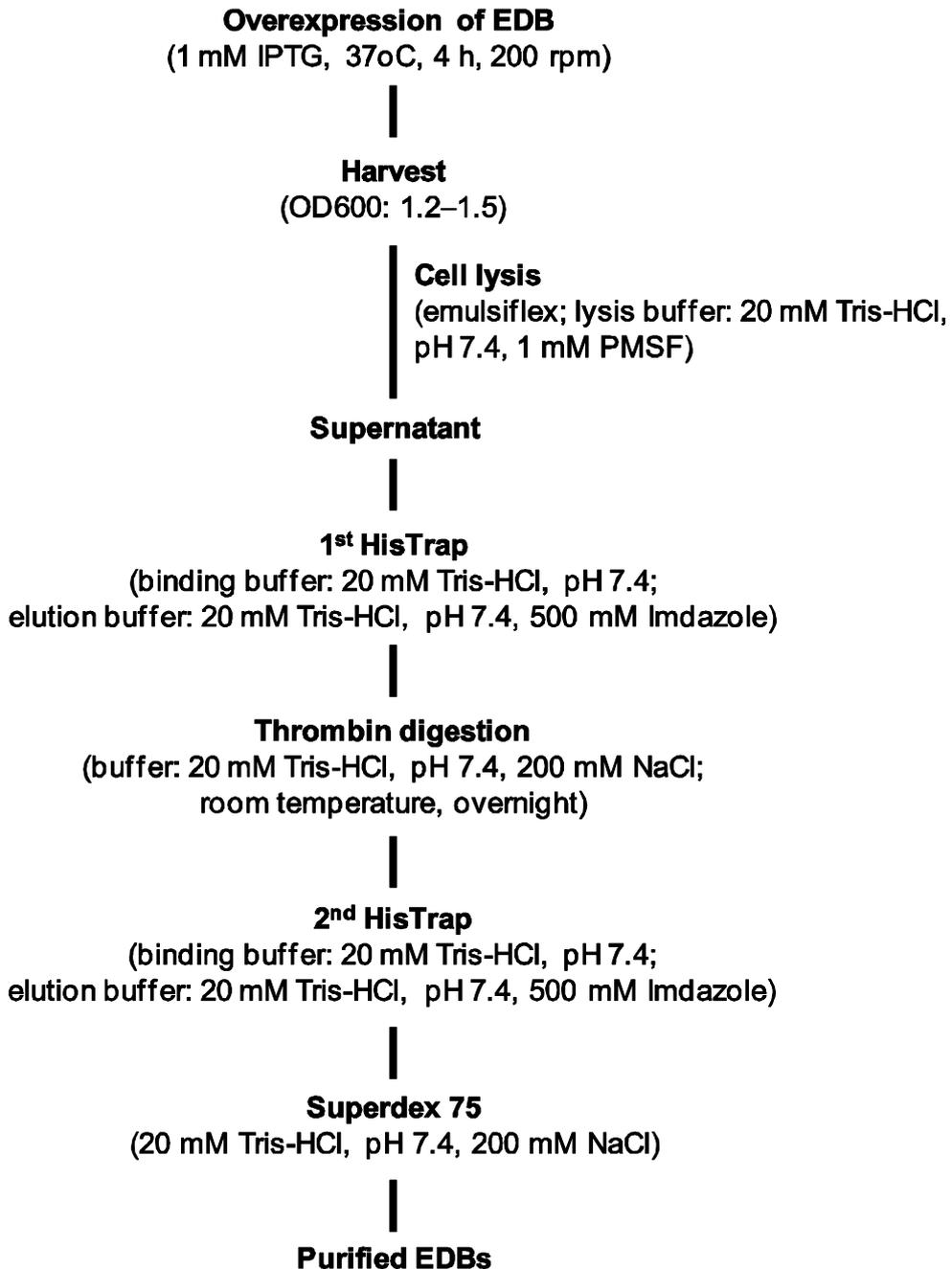


Figure 4. Purification procedure of EDB and EDB mutants. Purity of proteins was confirmed by SDS-PAGE.

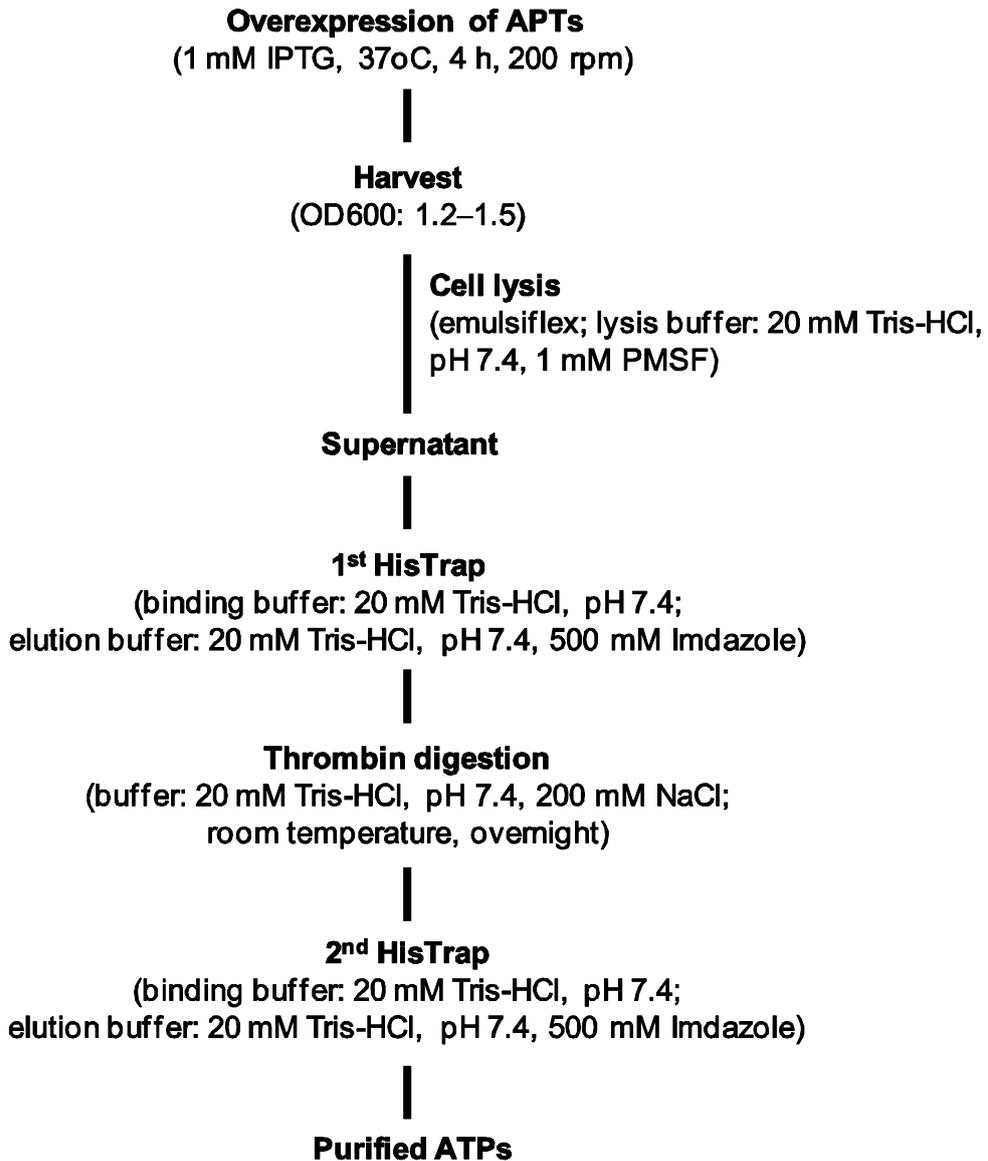


Figure 5. Purification procedure of APT and APT mutants. Proteins purity are confirmed by SDS-PAGE.

3. 4. NMR spectroscopy.

NMR samples contained 0.3 mM of ^{13}C , ^{15}N -EDB with 0.45 mM of unlabeled APT_{EDB} , or 0.3 mM of ^{13}C , ^{15}N - APT_{EDB} with 0.45 mM of unlabeled EDB in 20 mM sodium phosphate buffer, pH 6.0. NMR spectra were recorded at 25°C on Bruker Avance 600 MHz, 800 MHz and 900 MHz spectrometers equipped with an x,y,z -shielded gradient triple resonance probe. Sequential assignment of ^1H , ^{15}N , and ^{13}C resonances was achieved by three-dimensional triple resonance through-bond scalar correlation CBCACONH, HNCACB, HBHA(CO)NH, HCCH-TOCSY, and ^{15}N -TOCSY-HSQC experiments. Three-dimensional ^{13}C -separated NOESY, ^{15}N -separated NOESY, and ^{13}C -separated/ ^{12}C -filtered NOESY experiments were obtained using the mixing time of 120 ms. Residual $^1\text{D}_{\text{NH}}$ dipolar couplings were obtained by taking the difference in the J splitting values measured in oriented (15 mg/ml of *pf1* phage alignment medium) and isotropic EDB: APT_{EDB} complex using 2D in-phase/antiphase ^1H - ^{15}N HSQC spectrum ^[10]. NMR spectra were processed using the NMRPipe program ^[38], and analyzed using PIPP ^[5] and NMRView programs ^[6]. NOE experiments obtained

3. 5. Structure calculations.

Interproton distance restraints were derived from the NOE spectra and classified into distance ranges according to the peak intensity. ϕ/ψ torsion angle restraints were derived from backbone chemical shifts using the program TALOS+ ^[39]. Structures were calculated by simulated annealing in torsion angle space using the Xplor-NIH program ^[38]. The target function for simulated annealing included a

covalent geometry, a quadratic van der Waals repulsion potential ^[5], square-well potentials for interproton distance and torsion angle restraints ^[6], hydrogen bonding, RDC restraints ^[40], harmonic potentials for ¹³C α /¹³C β chemical shift restraints ^[41], a multidimensional torsion angle database potential of mean force ^[42], and a radius of gyration term ^[43]. The radius of gyration represents a weak overall packing potential, and residues 10–80 of EDB and residues 204–225 of APT were selected for the term excluding disordered regions. Structures were displayed using the VMD–XPLOR software ^[44].

3. 6. Isothermal titration calorimetry (ITC)

ITC experiments were performed at 25°C using the ITC₂₀₀ microcalorimeter (GE Healthcare). 0.8 mM EDB or EDB truncation mutants were titrated into 100 μ M APT in 20 mM sodium phosphate, pH 6.0. Twenty consecutive 2 μ L aliquots of proteins were titrated into the cell. The duration of each injection was 4 s, and injections were made at intervals of 150 s. The heats associated with the dilution were subtracted from the measured heats of binding. ITC titration data were analyzed with the Origin version 7.0 program provided with the instrument.

4. Results

4.1. Binding between EDB and APT_{EDB}

For over-expression of APT_{EDB} in *E. coli* cell, I used the sequence of APT_{EDB} (amino acids sequence: GSSSPIQGSWTWENGKWTWKGIIRLEQ) with two minor modifications, which was different from original sequence of APT_{EDB} optimized in previous screening.^[45] The first modification was that N-term cysteine residue of the original sequence was mutated to a serine residue to prevent the formation of an intermolecular disulfide bond. The second, an extra N-term glycine residue was left after removal of the Trx- and His₆-tag using the thrombin enzyme (thrombin cleavage site: LVPRGS). The extra residue can generate differences on solubility, binding affinity and complex structure. However, both 26-mer and 27-mer APT_{EDB} showed identical dissociation constant (K_D) (26-mer: 240 nM and 27-mer: 230 nM, respectively) and ¹H-¹⁵N HSQC spectrum on the complex formation (Figure 6). The mutants of APT_{EDB} also were designed to confirm the binding interface and mechanism (Table 1). The expressed EDB sequence (residues 5-95) also contained two extra N-terminal residues (Gly-Ser) after reacting the thrombin enzyme.

For estimating binding interface of EDB upon APT_{EDB}, I performed a NMR titration experiment that was adding the APT_{EDB} on the ¹⁵N-EDB. The ¹H-¹⁵N HSQC spectra of NMR titration showed that the EDB-APT_{EDB} complex was in

slow exchange on the chemical shift time scale with a dissociation constant (K_D) of 230 nM and a 1:1 binding stoichiometry determined from isothermal titration calorimetry (ITC). The most of peaks of ^1H - ^{15}N HSQC spectra disappeared and new peaks appeared simultaneously, during NMR titration until the EDB and APT_{EDB} became equal concentration. The backbone amide resonances of EDB showed extensive chemical shift perturbations upon complex formation with APT_{EDB} , implying a large conformational change upon the formation of the EDB- APT_{EDB} complex (Figure 7a). Thus, I could not assign the back bone as well as confirming the binding interface of EDB upon APT_{EDB} . The largest chemical shift perturbation (CSP) was observed in the β strand C', where the maximal $\Delta\omega_{\text{H}}$ was ~ 7500 Hz for Gly60, and the maximal $\Delta\omega_{\text{N}}$ was ~ 4600 Hz for Ser57. In similar with ^1H - ^{15}N HSQC, most of the residues showed chemical shift perturbations. Specially, the residues in β strands F and G of EDB exhibited generally chemical shift perturbations (Figure 7b). It indicated that EDB conformation was changed by binding with APT_{EDB} .

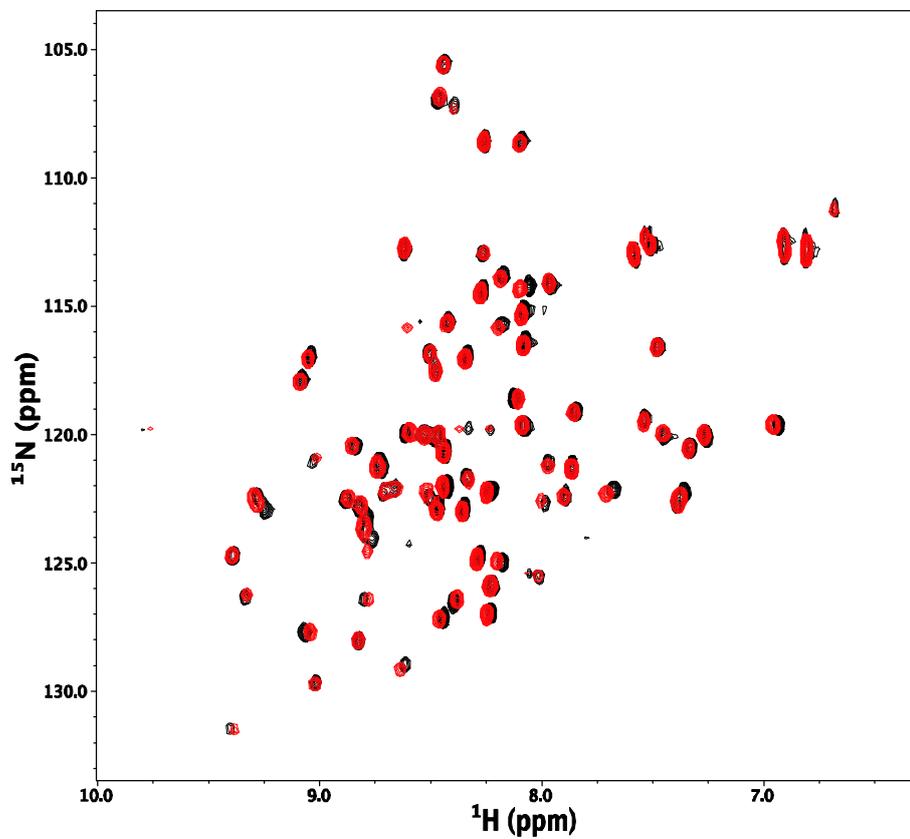


Figure 6. Comparison of ^1H - ^{15}N HSQC spectrum of EDB-APT_{EDB} (27-mer amino acid, black) and APT_{EDB} (26-mer amino acid, red) recorded at 600 MHz and 25°C in 20 mM NaPi, pH 6.0.

Table 1. APT_{E_{DB}} mutants. The mutated residues were marked on red.

	Amino acids	Sequence	Solubility
APT _{E_{DB}} mut1	27	GSSSPIQGSWTWENGKWTWKG GGRGEQ	○
APT _{E_{DB}} mut2	27	GGGG PIQGSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut3	27	GSSSP GG GSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut4	27	GSSSPIQGSWTWENGKWTWKGII GLGQ	○
APT _{E_{DB}} mut5	27	GGGGGGG GSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut6	27	GSSSPIQGSWTWENGKWTWKG GGGGGG	○
APT _{E_{DB}} mut7	27	GGGG PIQGSWTWENGKWTWKG GGGGG	○
APT _{E_{DB}} mut8	27	GGGGGGG GSWTWENGKWTWKG GGGGGG	○
APT _{E_{DB}} mut9	27	GSSSPIQGSWT V ENK V TWKGIIIRLEQ	Low solubility
APT _{E_{DB}} mut10	27	GSSSPIQGSWT V ENKWT V KGIIRLEQ	Low solubility
APT _{E_{DB}} mut11	27	GSSSPIQGS V TWENGKWT V KGIIRLEQ	○
APT _{E_{DB}} mut12	27	GSSSPIQGSWTWENG Q WTWKGIIIRLEQ	○
APT _{E_{DB}} mut13	27	GSSSPIQGSWTWENG Q WTW Q GIIRLEQ	○
APT _{E_{DB}} mut14	27	GGGGGGG GSWTWENG Q WTWKG GGGGGG	○
APT _{E_{DB}} mut15	27	GSSSPIQGSWTWENGKWTWKG I RLEQ	○
APT _{E_{DB}} mut16	27	GSSSPIQGSWTWENGKWTWKGII G LEQ	Low solubility
APT _{E_{DB}} mut17	27	GSSSPIQGSWTWENGKWTWKGII RGEQ	○
APT _{E_{DB}} mut18	26	SSSPIQGSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut19	26	GGGGGG GSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut20	26	SSSPIQGSWTWENGKWTWKG GGGGGG	○
APT _{E_{DB}} mut21	26	GGGGGG GSWTWENGKWTWKG GGGGGG	○
APT _{E_{DB}} mut22	26	SSSPIQGSWTWENGKWTWKG I RLEQ	○
APT _{E_{DB}} mut23	26	SSSPIQGSWTWENGKWTWKGII G LEQ	○
APT _{E_{DB}} mut24	26	SSSPIQGSWTWENGKWTWKGII RGEQ	○
APT _{E_{DB}} mut25	24	SPIQGSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut26	25	SSPIQGSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut27	28	GGSSSPIQGSWTWENGKWTWKGIIIRLEQ	○
APT _{E_{DB}} mut28	29	GGSSSPIQGSWTWENGKWTWKGIIIRLEQ	○

- ○: Soluble APT mutants in 20 mM NaPi, pH 6.0.

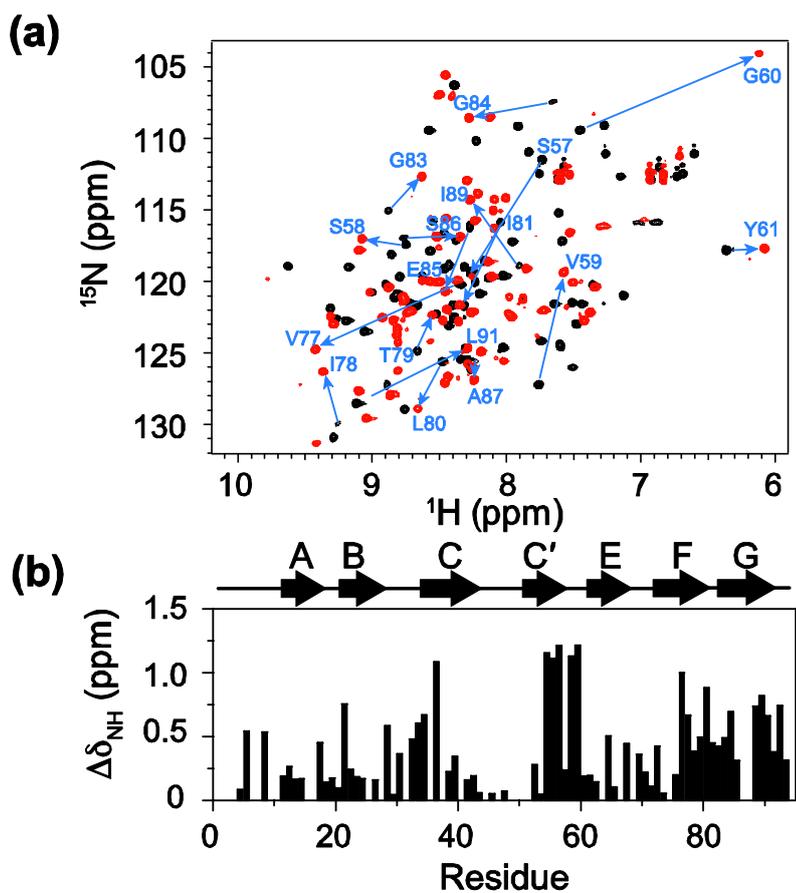


Figure 7. ^1H - ^{15}N HSQC and chemical shift perturbation (CSP) between human EDB and APT_{EDB} . (a) Comparison of ^1H - ^{15}N HSQC spectra of free ^{15}N -EDB (black) and the complex of ^{15}N -EDB with APT_{EDB} (red). Residues with large chemical shift changes are annotated. (b) The weighted average chemical shift perturbation of backbone amide resonances of EDB as a function of residue number upon complex formation with APT_{EDB} .

4. 2. Determination of complex structure

The NMR experiments for the assignment and determination of EDB–APT_{EDB} complex were conducted using samples containing 0.3 mM of the labeled EDB and 0.45 mM of the unlabeled APT_{EDB}, to ensure that the labeled protein was predominantly in the complex state (> 99 %). I performed three–dimensional heteronuclear correlation NMR experiments, that were CBCACONH, HNCACB, HBHA(CO)NH, HCCH–TOCSY, and ¹⁵N–TOCSY–HSQC, using ¹³C,¹⁵N–EDB:¹²C,¹⁴N–APT_{EDB} and ¹²C,¹⁴N–EDB:¹³C,¹⁵N–APT_{EDB} samples to obtain backbone and side chain assignment.

The peaks of three dimensional NMR spectra were picked, and the spin–system patterns were clustered and matched using the PIPP ^[5] and the NMRView (One Moon Scientific, Inc.) programs. Sequential assignment was achieved by verifying and linking of the individual spin–system pattern obtained from the triple NMR experiments. The peaks in the nearest neighborhood could be combined into a new cluster, when the ¹H and ¹⁵N chemical shifts represent a unique spin system. The peaks from ¹H–¹⁵N HSQC spectroscopy were combined with inter– and intra–residue ¹³C chemical shift obtained from CBCACONH and HNCACB, respectively. The chemical shifts of proton and carbon at sidechain were assigned to HCCH–TOCSY, and ¹⁵N–TOCSY–HSQC experiments (Table 2 and 3). I also confirmed the overlapping and ambiguous peaks of isoleusine, leusine and glycine using amino acid selectively labeled [¹⁵N–Leu], [¹⁵N–Ile], and [¹⁵N–Gly]–EDB samples (Figure 8).

The calculation of EDB–APT_{EDB} complex applied three–dimensional

^{13}C -separated NOE, ^{15}N -separated NOE, and ^{13}C -separated/ ^{12}C -filtered NOE restraints as well as RDCs of EDB and APT_{EDB} measured in the *pfl* phage alignment medium (15 mg/ml). The complex structure was determined using 1,907 NMR restraints including 1,610 experimental NOE restraints (EDB: 1145, APT_{EDB} : 297, and intermolecular: 168), 225 dihedral angle restraints, and 72 backbone $^1\text{D}_{\text{NH}}$ RDCs. Structures were calculated by simulated annealing using the cyana and the Xplor-NIH program. A summary of the structural statistics of the EDB- APT_{EDB} complex is provided in Table 4. The rmsd of twenty resultant structures with the lowest energy was 0.47 Å for backbone atoms and 0.79 Å for heavy atoms in the structured regions comprised of residues 10–81 of EDB and 204–225 of APT_{EDB} . PROCHECK-NMR analysis showed that 90.6 and 9.4 % of the backbone angles lay in regions of Ramachandran space (Figure 9) classified as most favored and additionally allowed, respectively.

Table2. Backbone ^1H , ^{15}N , and $^{13}\text{C}_\alpha$, and $^{13}\text{C}_\beta$, $^1\text{H}_\beta$ chemical shifts of complex EDB.

Residue	^1H	^{15}N	$^{13}\text{C}_\alpha$	$^1\text{H}_\alpha$	$^{13}\text{C}_\beta$	$^1\text{H}_\beta$
S4			58.29	4.40	63.61	3.80
E5	8.45	121.96	55.47	4.39	31.95	2.00
V6	8.27	122.17	56.38		29.50	
L9	7.77	123.83	54.82		41.93	
D11			53.06	4.73	37.99	2.61
L12	7.42	122.73	55.06	4.65	43.39	
S13	9.02	120.76	56.88		65.57	3.71
F14	8.37	119.92	61.03	5.51	39.93	2.77
V15	9.10	117.82	59.07	4.67	36.28	2.13
D16	8.52	116.85	54.95	4.18	38.65	2.66
I17	8.09	119.00	63.05	4.14	37.66	1.85
T18	8.74	122.16	60.57	4.60	70.60	4.91
D19	8.45	115.60	54.91	4.47	40.23	2.75
S20	8.12	108.50	57.66	4.62	65.77	4.10
S21	7.53	116.59	56.96	5.23	66.61	3.81
I22	8.21	113.85	60.48	4.48	42.19	1.57
G23	8.30	112.95	44.11	3.88		
L24	8.82	123.74	53.93	4.82	42.47	1.58
R25	8.57	119.98	54.04	4.90	33.15	1.54
W26	7.96	122.48	56.79	4.87	31.09	2.97
T27	9.78	119.82	60.69		68.17	
P28			63.26	4.47	31.98	2.28
L29	8.57	124.22	54.4	4.49	43.23	1.74
N30	8.63	119.95	53.84	4.70	38.18	2.86
S31	8.09	114.29	57.50		63.48	
S32			58.51	4.55	63.69	4.01
T33	8.00	114.16	62.99	4.22	69.22	
I34	7.98	121.09	61.92		38.08	
I35	8.81	126.23	61.86	4.22	38.26	1.71

G36			45.81	4.00		
Y37	7.34	120.40	57.93		38.95	
R38			54.50	5.15	31.53	
I39	9.04	129.57	60.00	5.05	40.02	1.45
T40	8.88	120.43	60.91	3.99	70.84	5.22
V41			61.37	5.03	33.8	
V42	9.10	127.67	60.64	4.65	35.93	2.01
A43	9.42	131.34	52.24	4.41	18.40	1.30
A44	8.26	126.04	53.47	4.12	18.37	1.317
G45			45.77	4.12		
E46	8.11	119.65	55.81	4.47	30.52	2.17
G47			45.91			
I48	7.44	120.01	58.50		39.4	
P49			63.59	4.53	31.88	2.30
I50	8.72	122.12	63.31		38.02	
E52			55.27	5.35	32.2	1.90
D53	8.85	122.75	52.95	5.00	44.77	2.57
F54	8.77	121.10	56.42	5.38	41.42	3.02
V55	8.29	125.76	59.25	4.32	35.04	1.66
D56	8.46	127.10	55.15	4.56	42.49	2.93
S57	8.26	119.63	60.63	4.33	62.68	4.10
S58	9.08	117.03	60.64	4.40	63.38	4.05
V59	7.57	119.35	65.55	3.79	31.80	2.27
G60	6.12	104.07	44.47	3.39		
Y61	6.08	117.69	56.69	5.03	41.16	2.75
Y62	8.44	126.65	58.73	4.44	42.76	3.23
T63	7.38	122.16	61.61	4.78	69.61	3.62
V64	8.86	127.94	61.74	3.66	31.21	1.94
T65	7.29	116.15	59.53	4.64	70.88	4.21
G66	8.46	105.56	45.86	4.00		
L67	7.86	119.08	53.36	4.18	41.02	1.10
E68	8.50	120.03	53.27		30.83	
P69			62.74	4.48	31.41	2.14
G70	8.50	106.94	45.83	4.06		

I71	7.51	120.08	58.72	4.08	39.70	1.54
D72	8.20	124.92	54.35	4.94	41.01	2.47
Y73	9.31	122.44	57.34	5.07	40.39	2.81
D74	9.29	122.98	54.17	5.22	42.04	2.71
I75	8.81	124.27	58.24	5.47	40.16	1.73
S76	8.76	121.30	56.97	5.81	65.53	3.77
V77	9.42	124.74	61.19	4.76	32.69	2.00
I78	9.37	126.30	61.16	4.71	40.59	1.89
T79	8.55	122.35	64.26	4.28	69.13	4.20
L80	8.66	128.90	54.06	4.73	40.97	1.39
I81	8.35	121.64	59.30	4.87	40.21	1.933
N82	8.82	123.24	53.61	4.64	38.80	2.91
G83	8.63	112.68	45.42	4.00		
G84	8.28	108.56	45.19	3.98		
E85	8.46	120.69	56.67	4.32	29.70	2.06
S86	8.35	116.86	57.95	4.43	63.56	3.83
A87	8.24	126.89	50.58	4.58	17.79	1.34
P88			62.96	4.471	31.46	2.28
T89	8.27	114.28	62.32	4.34	69.48	4.21
T90	8.09	116.29	61.89	4.33	69.50	4.20
L91	8.29	124.66	55.37	4.41	41.60	1.62
T92	8.10	115.04	61.89	4.30	69.50	4.18
Q93	8.36	122.80	55.72	4.35	29.05	2.10
Q94	8.48	122.72	55.78	4.40	29.10	2.14
T95	7.88	121.21	63.17		70.40	

Table3. Backbone ^1H , ^{15}N , and $^{13}\text{C}_\alpha$, and $^{13}\text{C}_\beta$, $^1\text{H}_\beta$ chemical shifts of complex APT_{EDB} .

Residue	^1H	^{15}N	$^{13}\text{C}_\alpha$	$^1\text{H}_\alpha$	$^{13}\text{C}_\beta$	$^1\text{H}_\beta$
S201			58.43	4.58	64.12	3.91
S202	8.52	118.00	58.30	4.56	63.89	3.89
S203	8.36	118.16	56.09		63.92	3.90
P204			63.95	4.63	32.66	2.41
I205	7.64	117.53	59.75	4.71	39.74	1.50
Q206	8.27	125.25	54.63	4.79	31.75	2.15
G207	8.13	111.33	46.28	3.79		
S208	6.99	109.86	57.29	4.55	65.83	3.82
W209	8.65	122.32	58.02	5.08	31.00	3.17
T210	9.84	118.41	61.13	4.91	71.67	4.15
W211	8.62	128.47	55.85	4.71	28.72	2.86
E212	8.22	126.98	55.45	4.37	32.59	1.78
N213	9.22	124.48	53.90	4.18	37.48	2.91
G214	7.67	101.87	45.96	3.75		
K215	6.82	117.09	54.81	5.23	35.92	1.66
W216	9.27	125.33	58.37	5.35	30.31	3.48
T217	10.01	125.74	61.39	4.92	71.61	4.15
W218	9.67	135.30	57.92	4.30	28.17	2.34
K219	7.55	126.23	54.36	5.24	35.32	1.63
G220	8.502	111.11	45.92	4.29		
I221	8.49	120.66	59.07	5.46	41.43	1.67
I222	9.21	125.74	59.25	5.51	41.72	1.66
R223	8.30	126.29	55.72	4.74	32.36	1.96
L224	9.98	127.92	54.49	4.51	41.31	1.58
E225	8.62	121.37	55.54	4.52	32.20	2.08
Q226	8.22	125.88	57.79	4.13	30.60	1.93

Table 4. Restraints and structural statistics

Nonredundant NOEs	
EDB	1145
APT	297
Intermolecular	168
Dihedral angles, $\phi / \psi / \chi$	98 / 101 / 26
Hydrogen bonds	29
Total number of restraints	
All residues (1–95, 201–226)	1864
Structured region (10–81, 204–225)	1767 (18.8 per residue)
rms deviation from experimental restraints*	
Distances (Å) (1610)	0.049 ± 0.003
Torsion angles (°) (225)	0.55 ± 0.08
Residual dipolar coupling <i>R</i> -factor (%) ^{*†}	
¹ D _{NH} (%) (72)	0.53 ± 0.07
rms deviation from idealized covalent geometry*	
Bonds (Å)	0.004 ± 0
Angles (°)	0.50 ± 0.03
Improper (°)	0.41 ± 0.01
Coordinate precision (Å) ^{*‡}	
Backbone	0.47 ± 0.06
Heavy atoms	0.79 ± 0.11
Ramachandran statistics (%) ^{*‡§}	
Most favorable regions	90.6
Allowed regions	9.4

* For ensemble of 20 structures

† The magnitudes of axial and rhombic components of the alignment tensor were –12.8 Hz and 0.41, respectively.

‡ Residues 10–81 of EDB and 204–225 of APT

§ Calculated using the program PROCHECK^[46]

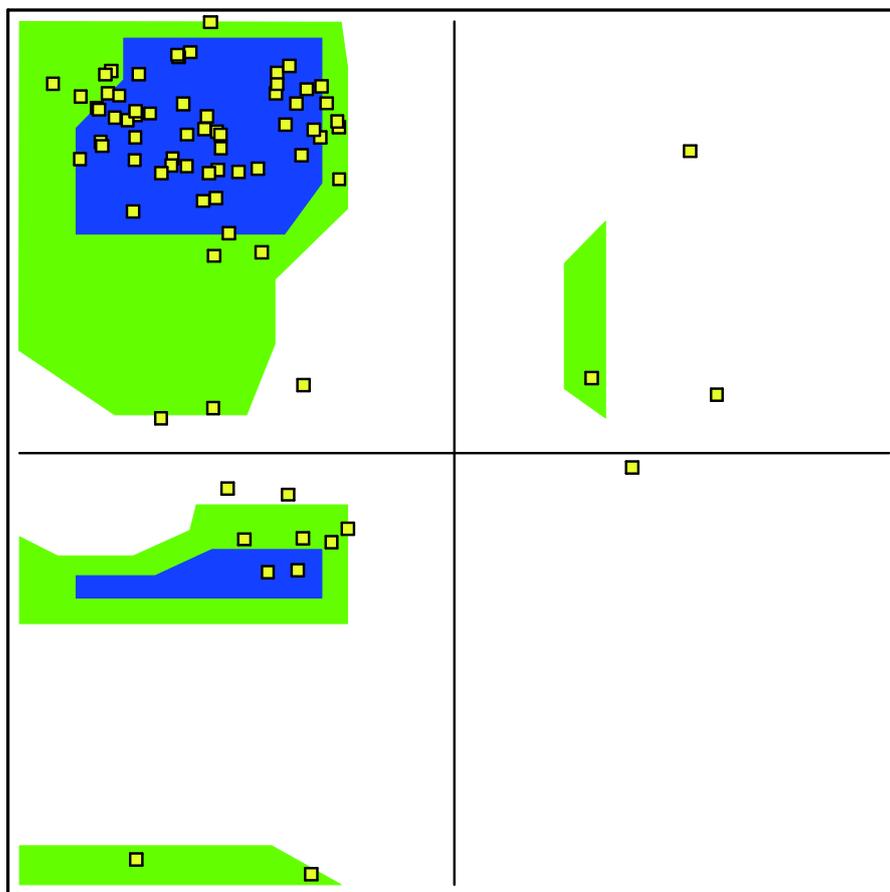


Figure 9. Ramachandran plot of 20 lowest energy NMR structure of EDB-APTEDB complex. The yellow squares are residues of EDB (10-85).

4. 3. Description on EDB in complex structure

EDB includes seven β strands forming a β sandwich structure (A, B, and E strands forming one β sheet, and C, C', F, and G forming another β sheet) (Figure 2), which is well conserved between fibronectin type III repeats ^[23]. EDB in the complex structure with APT_{EDB} also maintained a similar β sandwich fold except for the C-term β strand. The complex EDB comprised six β strands (β strand A, residues 12–17; B, residues 20–25; C, residues 36–43; C', residues 52–55; E, residues 60–66; F, residues 72–78), which were similar to the length of secondary structure of free EDB (Figure 10a). There was extraordinary structural difference between free and complex EDB, which the C-term β strand G of free EDB completely unfolded upon complex formation with APT_{EDB} (Figure 10b). In other words, C-term β strand G stayed on the disordered secondary structure in the EDB–APT_{EDB} complex. Also, secondary structure lengths of complex EDB were more shot than free EDB after binding with APT_{EDB}. APT_{EDB} unzipped β strand G of EDB upon binding, and alternatively formed a new intermolecular β sheet using β strand II region of APT_{EDB} instead of the β strand G of EDB. Unfolding β strand G of EDB in the complex structure was provoked to losing the intramolecular hydrogen bonding network between β strands F and G in free EDB. EDB–APT_{EDB} complex showed that the six β strands of EDB were well ordered to keep the β sandwich structure. The complex EDB structure was connected with APT_{EDB} to form an antiparallel β strand (Figure 10).

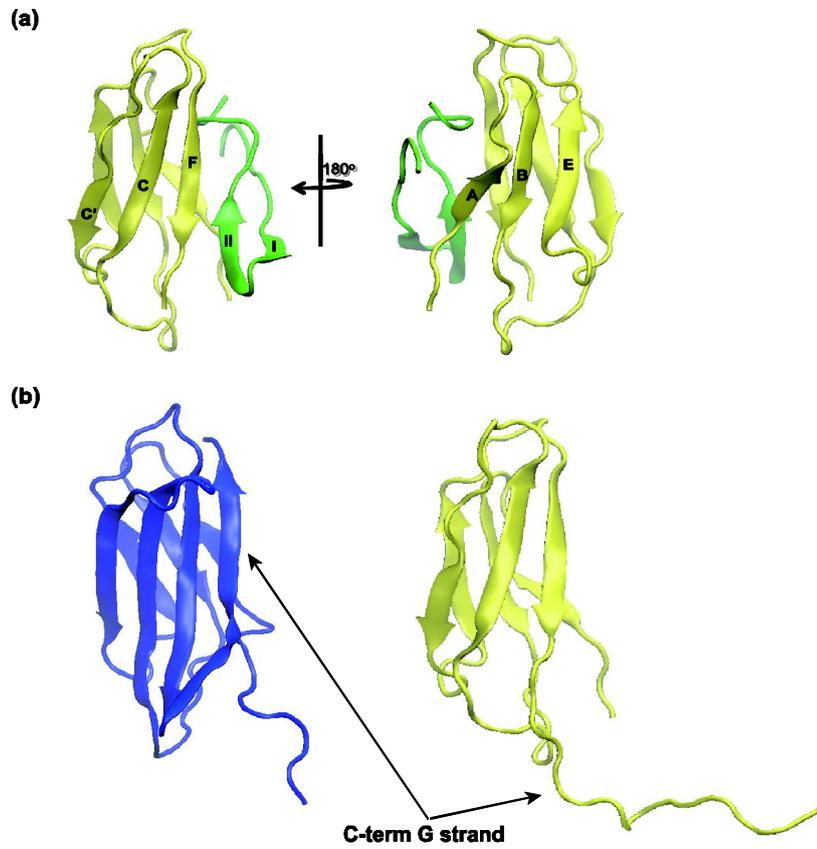


Figure 10. (a) The complex structure as a ribbon diagram with EDB in yellow and APT_{EDB} in green (PDB cord, 2MNU). (b) Comparison between free (blue) and complex (green) EDB. The C-term β strand G was annotated.

4. 4. Proofs upon unfolding of β strand G of EDB

Unfolding of β strand G of EDB upon binding to APT_{EDB} like as showed calculated $\text{EDB-APT}_{\text{EDB}}$ complex was supported by several experimental evidences such as heteronuclear $^1\text{H-}^{15}\text{N}$ NOE, chemical shift index (CSI) [47], TALOS+ [39], and ^{15}N R_2 relaxation.

First, heteronuclear $^1\text{H-}^{15}\text{N}$ NOE data between free EDB and the $\text{EDB-APT}_{\text{EDB}}$ complex indicated a pronounced difference in the β strand G region (Figure 11). The backbone heteronuclear $^1\text{H-}^{15}\text{N}$ NOE provides information about the motion of individual N-H bond vectors. Those that go through fast motion (pico and nano second time scale) and the overall tumbling of the molecules showed a decreased comparative NOE intensity on the average observed for the majority of the residues. For instance, decreased values are usually found at both N- and C-term ends of the protein. Therefore, large $^1\text{H-}^{15}\text{N}$ NOE values (> 0.8) were observed for free EDB in β strand G along with the other six β strands, and also the $^1\text{H-}^{15}\text{N}$ NOE values of other secondary structures except for β strand G in complex EDB were similar with free EDB. However, $\text{EDB-APT}_{\text{EDB}}$ showed significantly reduced $^1\text{H-}^{15}\text{N}$ NOE values (< 0.4) in the β strand G region (residues 84-93). It clearly indicated that β strand G becomes disordered upon binding with APT_{EDB} .

Second, I designed the truncated EDB_{5-83} mutant (residues 5-83) for removing β strand G. If β strand G is unfolded by APT_{EDB} , both EDB and EDB_{5-83} would be observed similar to $^1\text{H-}^{15}\text{N}$ HSQC spectra. The figure 12 showed a remarkably similar $^1\text{H-}^{15}\text{N}$ HSQC spectrum to that of free EDB. Backbone amide group

chemical shifts in the ^1H - ^{15}N HSQC spectrum were mostly identical except for disappeared amino acids (truncated residues 84–95). Also, I reversely monitored the ^1H - ^{15}N HSQC spectra of ^{15}N -APT_{EDB} upon complex formation with EDB and EDB_{5–83}. EDB-APT_{EDB} and EDB_{5–83}-APT_{EDB} complexes showed indistinguishable ^1H - ^{15}N HSQC spectra (Figure 13). Therefore, similar chemical shift perturbation profiles in the ^1H - ^{15}N HSQC spectra of complex EDB and APT_{EDB} supported the unfolding of β strand G upon binding. It also indicated that β strand G of EDB acted as a flexible tail without significant interactions with APT_{EDB} in the complex. Third, secondary structure predictions from the chemical shift index ^[47] (Table 5) and TALOS⁺ ^[39] (Figure 14) showed fully coherent with an unfolding case of β strand G in the EDB-APT_{EDB} complex. ^{15}N R_2 relaxation rates of backbone amides in β strand G were also greatly reduced when EDB bound to APT, which is typically observed in disordered tail regions. Therefore, I could confirm the determined EDB-APT_{EDB} complex structure, especially the unfolded β strand G, through supported NMR experiments.

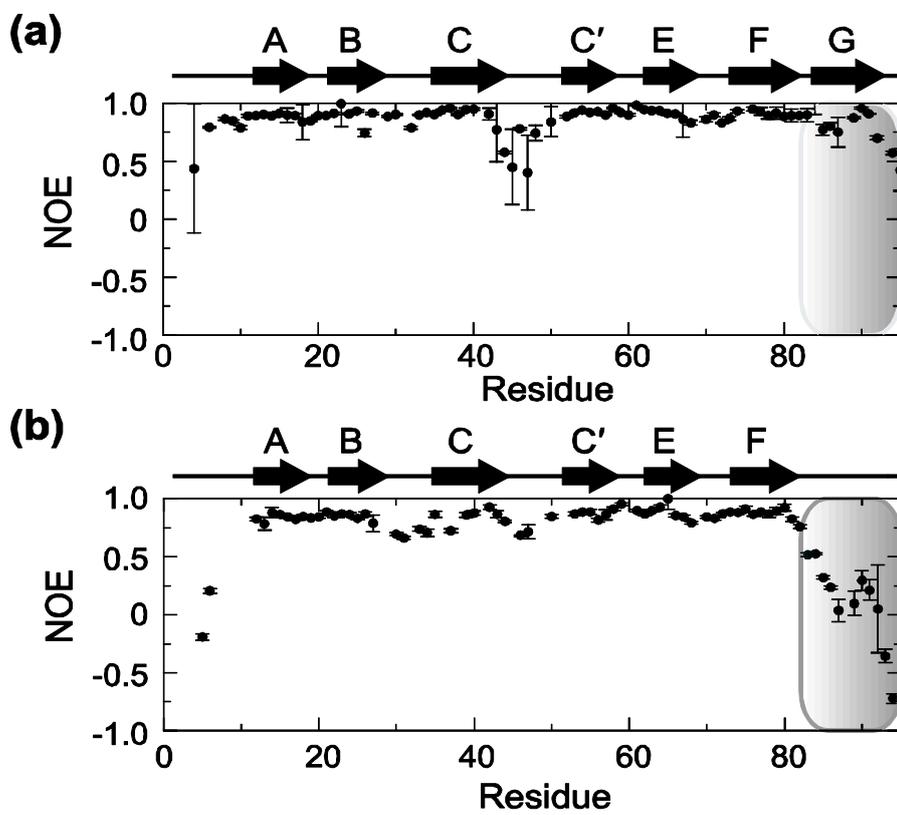


Figure 11. ^1H - ^{15}N heteronuclear NOE data as a function of the residue number of EDB (a) in the free state, and (b) in complex with APT.

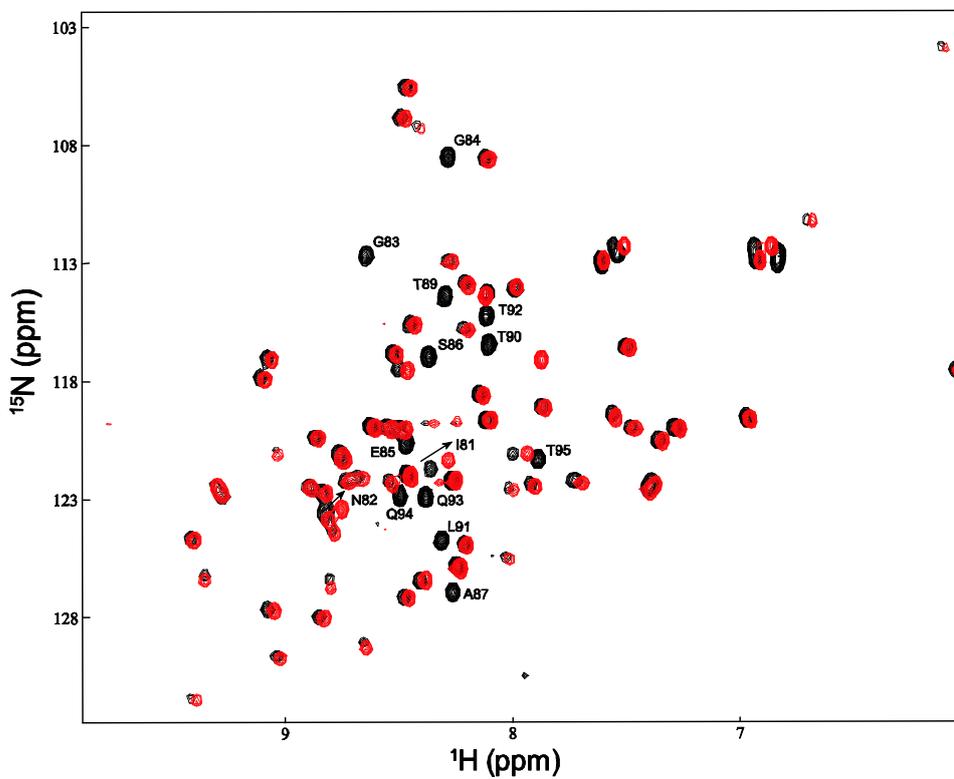


Figure 12. Superimposed ^1H - ^{15}N HSQC spectra of ^{15}N -EDB (black) and ^{15}N -EDB₅₋₈₃ (red) in complex with APT. The missing resonances of ^{15}N -EDB₅₋₈₃ due to the truncation are annotated by residue types and numbers.

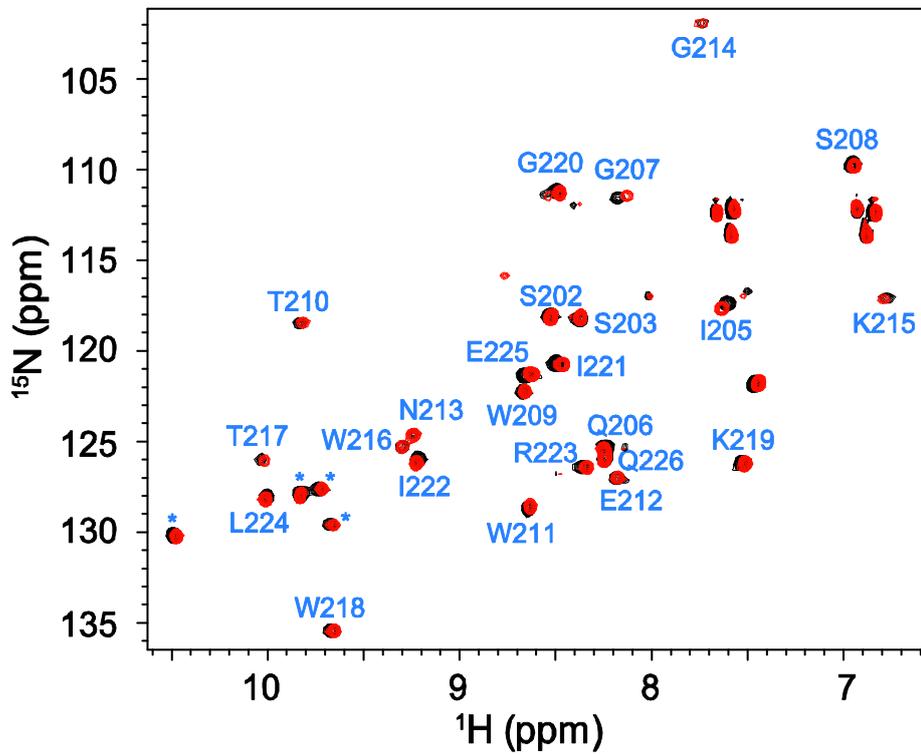


Figure 13. Superimposed ^1H - ^{15}N HSQC spectra of ^{15}N -APT in complex with EDB (black) and EDB₅₋₈₃ (red). The assignments for backbone amides of APT are annotated in the spectra, and the side chain H ϵ resonances of four tryptophan residues are annotated by asterisks.

Table 5. Chemical shift index (CSI) of free and complex EDB. The residues appear the length of β strand.

β strand	Residues of free EDB	Residues of complex EDB
A	9 – 15	5 – 15
B	20 – 26	20 – 28
C	33 – 42	34 – 42
C'	51 – 55	51 – 55
E	61 – 65	62 – 64
F	74 – 81	72 – 80
G	90 – 94	*

* Random coil region of complex EDB

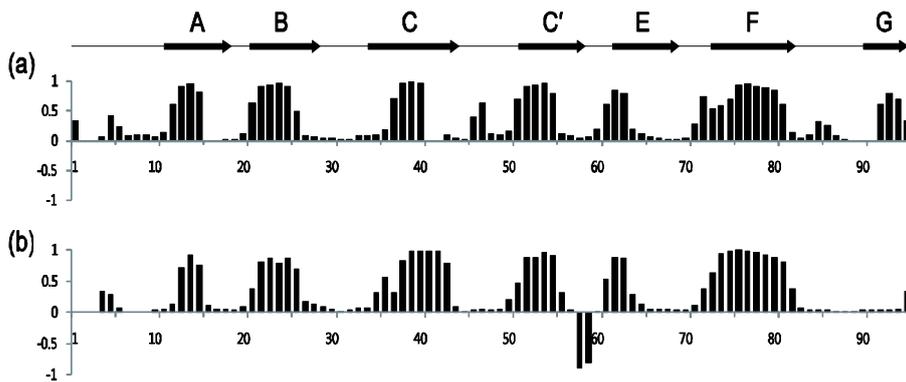


Figure 14. Predicted secondary structure of free (a) and complex (b) EDB. The positive bars means a β sheet, and the negative bars is a helix. The height of the bars reflects the probability of the neural network secondary structure prediction.

4. 5. Description on APT_{EDB} in complex structure

APT_{EDB} is comprised of a twelve-residue β -hairpin structure in its central region (SWTWENGKWTWK) with target binding arms at the N- and C-term of the hairpin (Figure 18b). Target binding regions comprises six variable residues (N-term residues: SSSPIQ and C-term residues: IIRLEQ) associated with the β hairpin via a glycine linker residue. In previous study, the β -hairpin structure has been reported to form a highly stable ($T_m = 72^\circ\text{C}$) trpzip scaffold. Cross-strand interactions between two pairs of tryptophan residues (one pair: W2-W4 and another pair: W9-W11) stabilize the trpzip scaffold^[48]. Even if free APT_{EDB} was observed to maintain the stable β -hairpin scaffold in its center, any conformation in the target binding arms did not presume^[45]. Also, ~ 50 peaks were observed on ^1H - ^{15}N HSQC spectrum, indicating that free APT_{EDB} existed on approximately two states (Figure 15).

APT_{EDB} in the EDB-APT_{EDB} complex maintained a similar β hairpin scaffold with two β strands (β strand I, residues 208-211; β strand II, residues 216-219), with the tryptophan residues participating in characteristic edge-to-face packing interactions, like as the trpzip scaffold^[48]. The target binding regions maintained a well-defined conformation, where the C-term region was more closely approached to EDB than the N-term region (Figure 10a). The C-term target binding region was observed to be highly ordered except for the Gln226 from the heteronuclear ^1H - ^{15}N NOE data, whereas the N-term target binding region from Ser201 to Ser203 was less ordered in the EDB-APT_{EDB} complex (Figure 16). The

well-ordered N- and C-term target binding arms were bent at Pro204 and Arg223, respectively, because locating on the cleft between the two layers of β sheets of EDB.

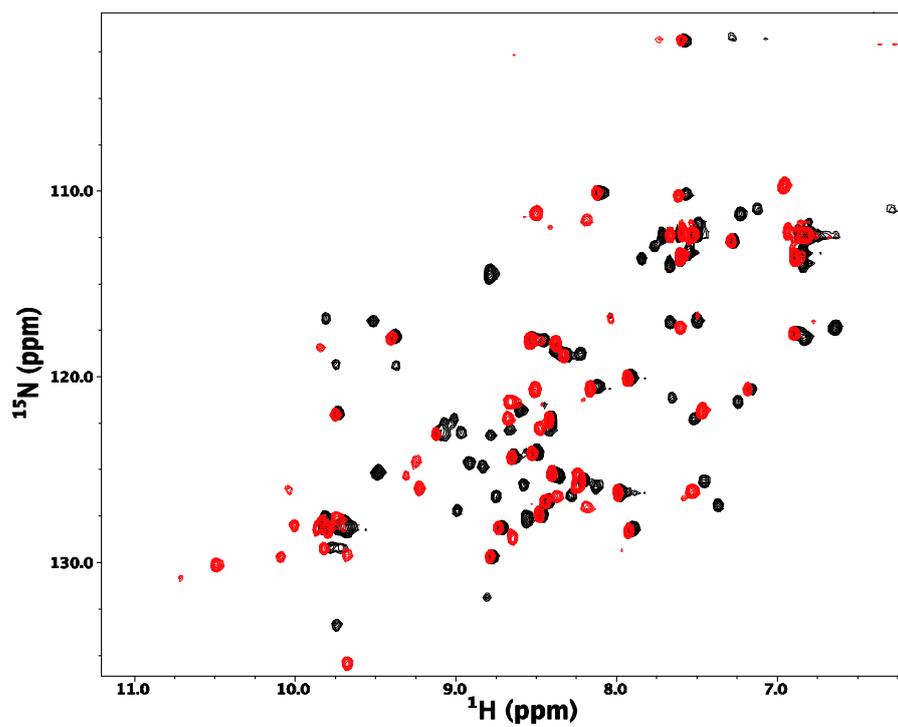


Figure 15. Comparison of ^1H - ^{15}N HSQC spectrum between free (black) and complex (red) APT_EDB.

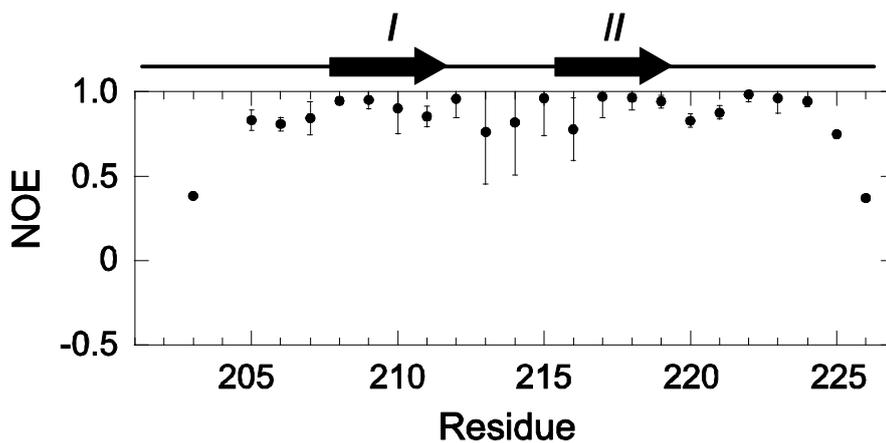


Figure 16. ^1H - ^{15}N heteronuclear NOE as a function of the residue number of APT in complex with EDB. The resonance for Ser202 of APT could not be detected in the presence of ^1H saturation, indicating that its ^1H - ^{15}N heteronuclear NOE was close to zero.

4. 6. Binding interface between EDB and APT_{EDB}

The binding interface was an elongated ellipsoid primarily formed by β strands A and F of EDB after unfolding the β strand G (Figure 17). The interaction interface was relatively flat and buries $\sim 670 \text{ \AA}^2$ of EDB and $\sim 840 \text{ \AA}^2$ of APT_{EDB}. The interaction interface of EDB comprised Leu12, Phe14, Ile17 (β strand A), Asp19 (loop AB), Ile22 (β strand B), Glu68 (loop EF), Asp72, Tyr73, Asp74, Tyr75, Ser76, Val77, Ile78, and Thr79 (β strand F). The interaction interface of APT comprised Pro204, Ile205 (N-terminal target binding region), Trp209 (β strand I), Trp216, Thr217, Trp218, Lys219 (β strand II), Ile222, Arg223, and Leu224 (C-terminal target binding region) (Figure 17).

Unfolding of β strand G exposed a wide hydrophobic core region that provided key interaction surfaces for APT_{EDB}. The hydrophobic patch formed by Phe14, Ile17, and Ile22 of EDB was exposed by unfolding β strand G, and provided the interface for Pro204, Ile205, and Leu224 of APT. Leu12 and Val77 of EDB were similarly exposed to provide the hydrophobic interface for Trp209 and Trp218 of APT. In addition, the backbone and aliphatic groups of Asp72, Tyr73, Asp74, and Ile75 in β strand F of EDB participated in van der Waals interactions with Ile222. Finally, the backbone carbonyl groups of Asp19, Glu68, and Ile71 were in close contact with the guanidinium side chain of Arg223 via electrostatic and hydrogen bonding interactions. The key molecular interactions from the target binding arms of APT were thus provided by Ile222, Arg223, and Leu224 from the C-terminal region, and Pro204 and Ile205 from the N-terminal region.

When β strand G of EDB was unfolded and opened to the concealed

hydrophobic area on EDB, the stable complex EDB–APT_{EDB} structure was formed by inducing the hydrophobic interaction with I222, R223 and L224. I designed three APT_{EDB} mutants (mut15: Ile222→Gly, mut16: Arg223→G, and mut17: Leu224→Gly) to find the key residue. The equilibrium dissociation constants of the mutants were reduced 17–20 folds; APT_{EDB} mut15 and 17 (APT_{EDB} mut3 precipitation) were K_D 4.7 μ M and 3.9 μ M, respectively (Table 6). In ¹H–¹⁵N HSQC spectrum, APT_{EDB} mut15 did not form the stable complex structure, whereas APT_{EDB} mut17 showed a similar ¹H–¹⁵N HSQC spectrum with EDB–APT_{EDB} complex. As a result, the L224 among the three residues plays important role in stabilizing the EDB–APT_{EDB} complex.

The interaction interface presented a hydrogen bond network between the backbone amide groups of β strand F of EDB and β strand II of APT to form an intermolecular β sheet of β strand II. Hydrophobic and electrostatic interactions between side chain groups further stabilized the interaction of EDB–APT_{EDB} complex. Specifically, the amide group of Thr79 forms a hydrogen bond with the carbonyl group of Trp216, and the amide and carbonyl groups of Val77 form a pair of hydrogen bonds with the carbonyl and amide groups of Trp218 (Figure 18b). By contrast, the amide group of Thr79 in the EDB did not participate in the hydrogen bond, because Ala87 in the opposite strand protruded to form a β -bulge structure (Figure 18a). The hydrogen bonds of Thr89 with Val77 were replaced by Trp218 of APT_{EDB} in the complex. In determined complex structure, I conformed that the C-terminal target binding arm was more important than that at the N-term. In previous study, It is reasonable because both the APT_{EDB} derivatives with N- and

C-term deletions showed dramatic decreases of their binding affinities, which dropped from 65 nM to 12 and 592 nM, respectively^[45]. Therefore, the interactions of EDB-APT_{EDB} complex were comprised of intermolecular hydrogen bonds through β -hairpin scaffold of APT_{EDB} and hydrophobic and electrostatic interactions through the target binding regions.

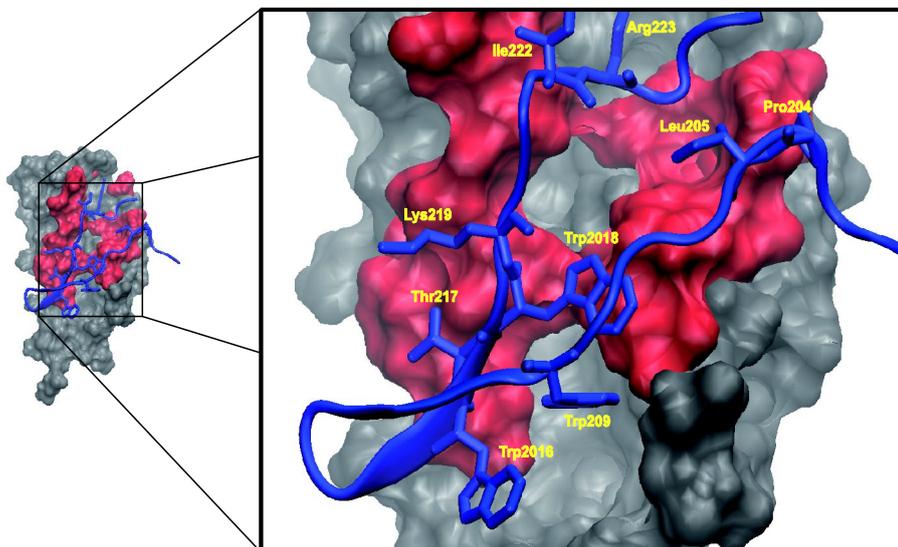


Figure 17. The interface between EDB and APT_{EDB}. Interfacial residues of EDB in red and APT in blue with the backbone are shown as ribbon model and the side chains shown as sticks. Interaction interface residues of APT_{EDB} are annotated (yellow).

Table 6. Thermodynamic parameters for the interaction between EDB and APT_{EDB} (APT_{EDB} mutants). Equilibrium dissociation constants (K_D), binding free energy (ΔG), binding enthalpy (ΔH), and binding entropy (ΔS) were obtained from ITC at 25°C.

	Kd (nM)	ΔS (cal/mol/deg)	ΔH (kcal/mol)	ΔG (kcal/mol)
APT_{EDB}	230 ± 36	39.6 ± 0.32	2.8 ± 0.03	-9.0 ± 0.09
APT_{EDB} mut18	240 ± 44	36.9 ± 0.38	2.0 ± 0.02	-9.0 ± 0.11
APT_{EDB} mut2	500 ± 250	35.1 ± 1.05	1.9 ± 0.09	-8.6 ± 0.3
APT_{EDB} mut3	1500 ± 200	46.3 ± 0.40	5.9 ± 0.09	-7.9 ± 0.08
APT_{EDB} mut5	230 ± 36	39.6 ± 0.32	2.8 ± 0.03	-9.0 ± 0.09
APT_{EDB} mut6	2800 ± 560	24.8 ± 0.42	1.2 ± 0.04	-6.2 ± 0.12
APT_{EDB} mut15	4700 ± 140	28.5 ± 0.66	1.2 ± 0.08	-7.3 ± 0.18
APT_{EDB} mut17	3900 ± 500	34.8 ± 0.36	3.0 ± 0.07	-7.4 ± 0.08

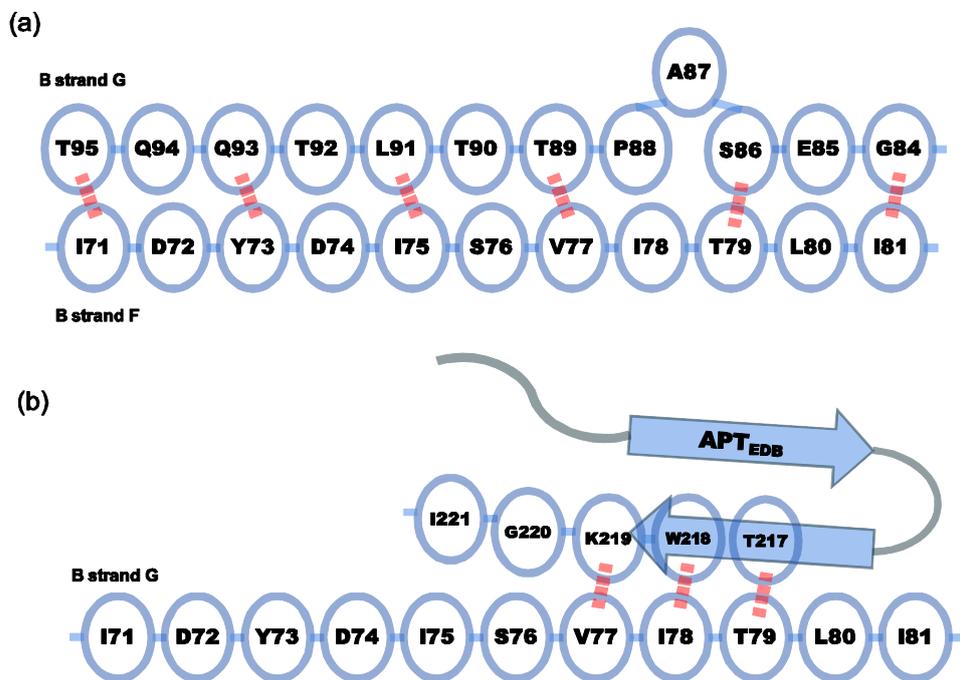


Figure 18. (a)The intermolecular hydrogen bonds between β strand F and G. (b) the intramolecular backbone hydrogen bonds between β strands F and β strand II of APT in the complex. The hydrogen bonds between β strands are denoted by vertical bars (red).

4. 7. Thermodynamics on EDB and APT_{EDB}

I probed the binding thermodynamics of EDB and APT_{EDB} to understand the binding mechanism, which was unfolding and binding sequentially, using a series of C-term truncation mutants (residues 5–94, 5–92, 5–90, 5–86, 5–83) of EDB. The mutants of the truncated C-term were analyzed on the structural stability and similarity using the ¹H–¹⁵N HSQC spectrum. The ¹H–¹⁵N HSQC spectra of the complex structures of each EDB mutants with APT_{EDB} showed that their NMR spectra were same, differing only in the lengths of the unfolded C-terminal tail region. On the other hand, the truncated mutants showed variations on thermodynamic profiles. The EDB–APT_{EDB} complex formation was predominantly driven by a large entropic gain ($-T\Delta S = -11.8$ kcal/mol), which was consistent with the observed β strand G unfolding (Table 7). The enthalpic change was unfavorable in the EDB–APT complex formation, implying the energetic cost to pull β strand G apart. C-terminal truncations decreased the entropic gains and also relieved the enthalpic burdens for unfolding. Notably, the deletion of a single C-terminal residue dramatically changed the ΔH and $-T\Delta S$ profiles, whereas further truncations resulted in smaller changes. This observation indicates that the integrity of β strand G critically depends on the presence of C-terminal residues. β strand G harbors a β -bulge region in the middle, which may destabilize the inter-strand hydrogen bonding network and make β strand G susceptible to unfolding. The β -bulge region in the β strand G was designed to three mutants that were insertion of Glycine (between Ala 87 and Pro 88) and deletion of Ala 87 or Pro 88. However, the β strand G of mutants was already unfolded in the free state on the ¹H–¹⁵N

HSQC spectrum. It did not provide the information of unfolding of the β strand G upon the β -bulge region.

I also examined if an isolated β -hairpin peptide (12 aa) without the target binding regions could interact with EDB. The β -hairpin peptide caused significant line-broadening of ^{15}N -EDB, and the residues that exhibited large chemical shift perturbations by APT were mainly broadened out. Peaks of residues on the β strand F and F-G loop (L67, G70, I71, D72, Y73, D74, I78, T79, I81, G83, G84, E85, S86, A87, T89, and T90) also were disappeared. The intermediate exchange indicates that the β -hairpin peptide bound to EDB, but did not form a stable complex due to the lack of target binding arms for further stabilizing interactions.

The N-term and C-terms of APT_{EDB} were replaced on glycine. The APT_{EDB} mut6 (C-term glycine) showed the similar ^1H - ^{15}N HSQC spectrum with the EDB: β -hairpin peptide complex, whereas the APT_{EDB} mut5 (N-term glycine) showed the similar ^1H - ^{15}N HSQC spectrum with the EDB- APT_{EDB} complex. It means that the β hairpin structure of APT_{EDB} is interacted with the β strand F, G and F-G loop, and the stable complex is formed by C-term binding region. I insist that the target recognition of APT proceeds in two steps. First, β hairpin of APT binds to EDB and unfolds β strand G. Second, the target binding arms of APT find the specific interaction surfaces of EDB to form a stable complex (Figure 19). Given that the β hairpin of APT replaces the β -bulge region of EDB, it is possible that APT first recognizes the β -bulge region to begin unfolding of EDB (Figure 18a).

Table 7. Thermodynamic parameters for the interaction between EDB (or EDB mutants) and APT, and between EDB–FN8 and APT. Equilibrium dissociation constants (K_D), binding free energy (ΔG), binding enthalpy (ΔH), and binding entropy (ΔS) were obtained from ITC at 25°C.

	K_D (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (cal/mol/deg)
EDB	230 ± 36	−9.0 ± 0.09	2.8 ± 0.03	39.6 ± 0.32
EDB₅₋₉₄	170 ± 33	−9.2 ± 0.11	−2.1 ± 0.03	23.8 ± 0.38
EDB₅₋₉₂	100 ± 40	−9.6 ± 0.24	−3.8 ± 0.06	19.3 ± 0.83
EDB₅₋₉₀	91 ± 24	−9.6 ± 0.18	−4.4 ± 0.05	17.4 ± 0.63
EDB₅₋₈₆	110 ± 37	−9.5 ± 0.20	−4.7 ± 0.06	16.1 ± 0.70
EDB₅₋₈₃	120 ± 32	−9.5 ± 0.16	−4.6 ± 0.05	16.3 ± 0.56
EDB–FN8	3200 ± 300	−7.5 ± 0.06	6.6 ± 0.1	47.3 ± 0.39

* Equilibrium dissociation constant (K_D) was not detected in ITC because of low binding affinity.

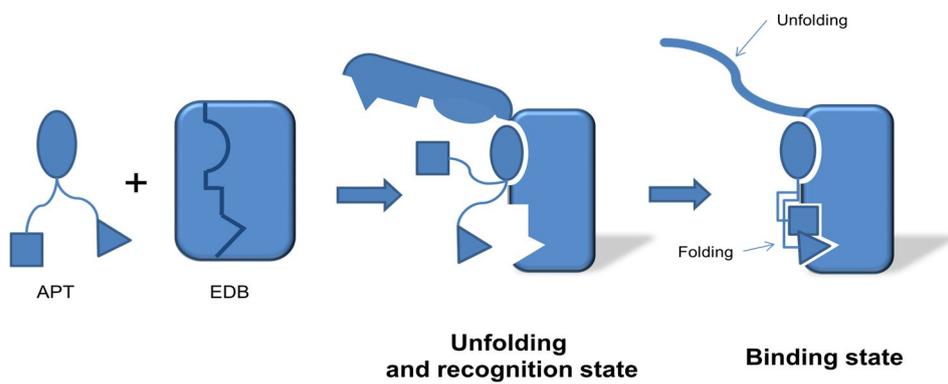


Figure 19. Schematic illustration of coupled unfolding and binding during the interaction between EDB and APT_{EDB} .

4. 8. Relationship between EDB and FN8

In cellular FN, the EDB is introduced between FN type III repeats FN7 and FN8. Individual FN domains are tightly interacted with each other via their loop regions to form an extended structure ^[23], identifying also FN7–EDB–FN8 domain. For conforming changes such as binding affinity and unfolding conformation of EDB–APT_{EDB} complex, EDB–FN8 was cloned. Before further study, I noted that FN8 did not interact with APT_{EDB} via ¹H–¹⁵N HSQC spectrum (Figure 19). The binding affinity of the EDB–FN8:APT_{EDB} complex reduced by 14–fold compared to that of EDB–APT_{EDB} (K_D : 230 nM) and the K_D was measured as 3.2 μ M (Table 5). The reduced binding was caused by an extra energetic cost to break the loop interactions between EDB and FN8. The ¹H–¹⁵N HSQC spectrum of free ¹⁵N–EDB–FN8 showed chemical shift changes from the overlaid HSQC spectra of separately expressed individual domains. The chemical shift perturbation resulted from the loop interactions indicating that the two domains were associated. The HSQC spectrum of the ¹⁵N–EDB–FN8:APT complex was remarkably similar to the spectra of individual complexes (complex EDB, FN8, and EDB–FN8), indicating that EDB in complex with APT was associated with FN8 (Figure 20). The results showed that APT_{EDB} binds to EDB–FN8 and unfolds β strand G of EDB regardless of FN8 domain. The unfolding perturbs the loop interaction between EDB and FN8, and disengages the two domains in the complex.

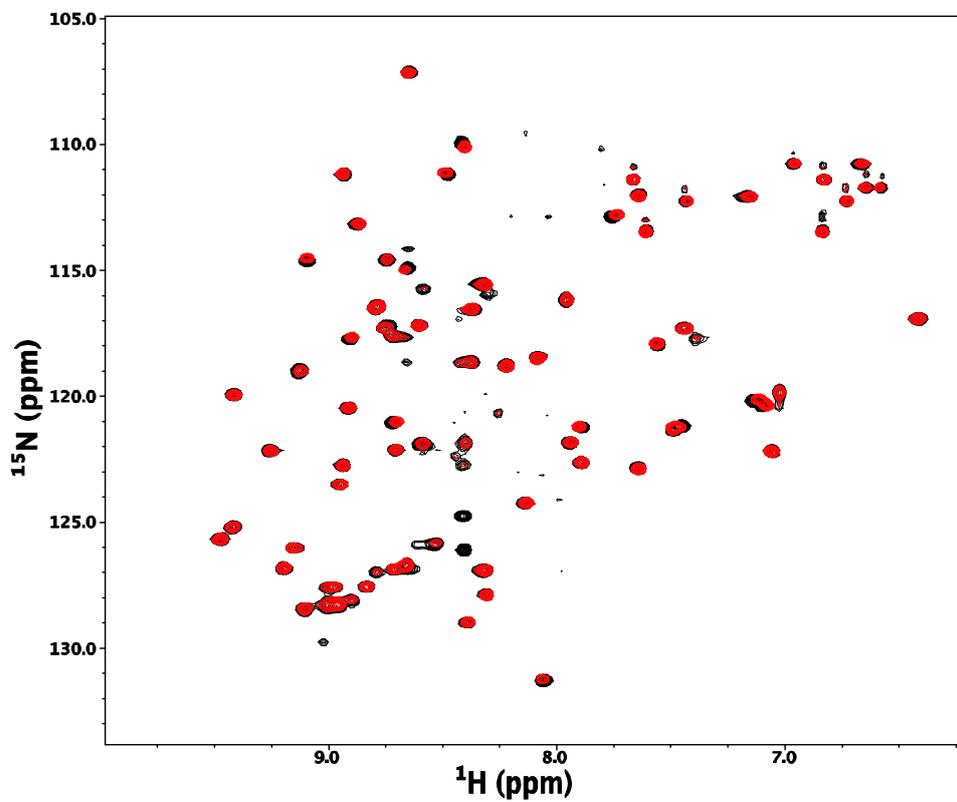


Figure 20. Superimposed ^1H - ^{15}N HSQC spectra of ^{15}N -FN (black) and ^{15}N -FN in complex with APT_{EDB} (red).

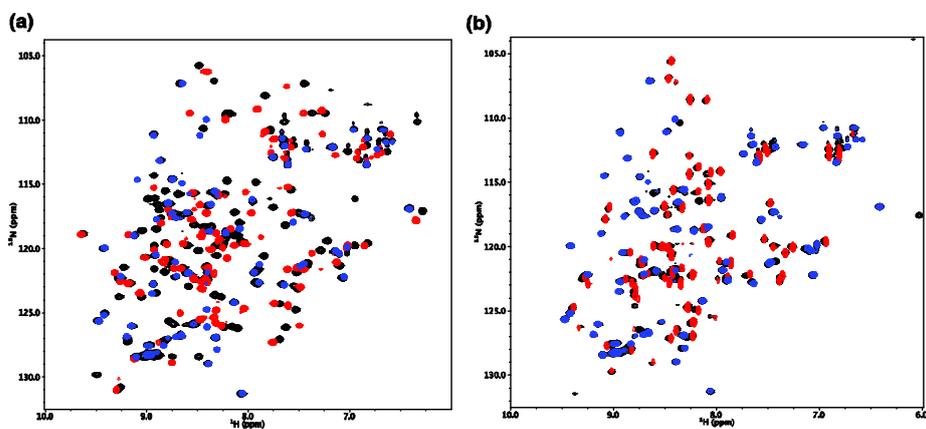


Figure 21. Comparison of ^1H - ^{15}N HSQC spectrum between EDB linked with FN8 and EDB unlinked with FN8. (a) ^1H - ^{15}N HSQC spectrum of EDB-FN8:APT_{EDB} (black), EDB (red), and FN8 (blue). EDB and FN8 were chemical shift change against EDB-FN8 due to loop interaction between EDB and FN8. (b) Comparison of ^1H - ^{15}N HSQC spectrum of EDB-FN8:APT_{EDB} (black), EDB:APT_{EDB} (red), and FN8:APT_{EDB} (blue). EDB:APT_{EDB}-FN8 spectra was similar to two spectrum combined with EDB:APT_{EDB} and FN8:APT_{EDB}.

5. Discussion

The crystal structures of FN type III domains show that individual FN domains are covalently linked ^[23]. The interdomain interfaces between FN subunits determine both the domain orientation and the loop structures, which are important in molecular interactions and biological functions of FN ^[16, 49]. For example, insertion of EDB between FN7 and FN8 alters both FN8 orientation and domain interfaces, which promotes cell adhesion and spreading ^[50].

In this study, I determined the complex structure of EDB and the specific aptide (APT_{EDB}) using NMR spectroscopy, and characterized the binding mechanism. APT_{EDB} binding triggers the unfolding of an entire β strand G of EDB, exposing and enabling interaction with a wide hydrophobic surface normally buried within EDB. The binding mechanism divides two step, unfolding and binding. First, β hairpin of APT containing four tryptophan binds to the bulge of EDB and unfolds β strand G. Therefore, the β -hairpin structure plays the important role to take off the cover (β strand G of EDB) for interacting to the hidden binding interface. Second, the target binding arms of APT find the specific binding interface of EDB and forming a stable complex. And figure 31 shows that APT binds to EDB-FN8 and lead to be unfolded β strand G of EDB, even if C-term of EDB is linked on FN8. *In vivo*, Perturbation of the domain interaction between EDB and FN8 can change the global FN structure and also loop conformations, which would influence the

molecular interactions of FN and downstream angiogenic signaling. The complex structure of EDB and APT_{EDB} demonstrates that protein–protein interactions can generate their binding interface by unfolding a secondary structural element. The complex formation between EDB and APT was a reversible process, so that association and dissociation of the protein complex was accompanied by unfolding and refolding of β strand G. The unfolding–upon–binding mechanism has several biological implications. First, it expands the diversity of known protein–protein interactions. Unfolding–upon–binding allows for a specific interaction via hidden interfaces, so that variations in hydrophobic core residues can contribute to binding specificity. Second, the stability of the secondary structure may modulate the binding affinity. Since unfolding of the secondary structure is balanced by an enthalpic cost and an entropic gain, the folding propensity of the secondary structure can offer a new option to modulate the binding affinity. Finally, characterizing the structural determinants that dictate protein unfolding and binding may help to predict and design inhibitors that target protein–protein interactions.

6. Conclusion

In this paper, small aptamer-like peptides (aptides, APT) of twenty six amino acids can recognize diverse protein targets with high affinity^[45]. Aptides selected through phage display screening against EDB have exhibited a high *in vitro* affinity and a specific tumor targeting profile in a human glioblastoma xenograft model. A clearly understanding of the structural determinants that enable aptides to achieve specific target recognition will enhance the development of low molecular weight antibody alternatives^[37, 51-53]. APT specifically targeting tumors *in vivo* have the potential application as a signal modulator for the inhibition of tumor angiogenesis^[53]. None of the antibody epitopes overlap with the binding interface for APT, suggesting that APT can be reliably used in combination with the known antibody fragments for cancer diagnosis and treatment.

Further, the complex structure of EDB-APT ensures that the β turn of APT is open to solvents, allowing for conjugation to a fluorescence probe for tumor diagnosis, or to a tumor-targeted cytotoxic agent without impairing the affinity. In particular, the lysine residue (Lys215) in the β turn is amenable to tag conjugation by canonical chemical reactions, as was demonstrated in the previous *in vivo* tumor imaging and therapy studies.

III. REFERENCE

1. van Nuland, N.A., Boelens, R., Scheek, R.M., Robillard, G.T., **1995**, High-resolution structure of the phosphorylated form of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from NMR-NOE data. *J Mol Biol.* 246(1): 180–193.
2. Robillard, G.T., Broos, J., **1999**, Structure/function studies on the bacterial carbohydrate transporters, enzymes II, of the phosphoenolpyruvate-dependent phosphotransferase system. *Biochim Biophys Acta.* 1422(2): 73–104.
3. Tang, C., Iwahara, J., Clore, G.M., **2006**, Visualization of transient encounter complexes in protein-protein association. *Nature.* 444(7117): 383–386.
4. Suh, J.Y., Tang, C., Clore, G.M., **2007**, Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement. *J Am Chem Soc.* 129(43): 12954–12955.
5. Garrett, D.S., et al., **1991**, A Common-Sense Approach to Peak Picking in 2-Dimensional, 3-Dimensional, and 4-Dimensional Spectra Using Automatic Computer-Analysis of Contour Diagrams. *J Magn Reson.* 95(1): 214–220.
6. Johnson, B.A., Blevins, R.A., **1994**, NMR View: A computer program for

- the visualization and analysis of NMR data. *J Biomol NMR*. 4(5): 603–614.
7. Iwahara, J., Tang, C., Clore, G. M., **2007**, Practical aspects of ¹H transverse paramagnetic relaxation enhancement measurements on macromolecules. *J Magn Reson*. 184(2): 185–195.
 8. Garrett, D.S., Seok, Y.J., Peterkofsky, A., Gronenborn, A.M., Clore, G.M., **1999**, Solution structure of the 40,000 Mr phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. *Nat Struct Biol*. 6(2): 166–173.
 9. Suh, J.Y., Tang, C., Clore, G.M., **2007**, Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement. *J Am Chem Soc*. 129(43): 12954–12955.
 10. Bax, A., Kontaxis, G., Tjandra, N., **2001**, Dipolar couplings in macromolecular structure determination. *Methods Enzymol*. 339: 127–174.
 11. Prestegard, J.H., Bougault, C.M., Kishore, A.I., **2004**, Residual dipolar couplings in structure determination of biomolecules. *Chem Rev*. 104(8): 3519–3540.
 12. Donaldson, L.W., Skrynnikov, N.R., Choy, W.Y., Muhandiram, D.R., Sarkar, B., Forman-Kay, J.D., Kay, L.E., **2001**, Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. *J Am Chem Soc*. 123(40): 843–847.
 13. Suh, J.Y., Iwahara, J., Clore, G.M., **2007**, Intramolecular domain-domain association/dissociation and phosphoryl transfer in the mannitol transporter

- of *Escherichia coli* are not coupled. *Proc Natl Acad Sci.* 104(9): 3153–3158.
14. Meadow, N.D., Mattoo, R.L., Savtchenko, R.S., Roseman, S., **2005**, Transient state kinetics of Enzyme I of the phosphoenolpyruvate:glycose phosphotransferase system of *Escherichia coli*: equilibrium and second-order rate constants for the phosphotransfer reactions with phosphoenolpyruvate and HPr. *Biochemistry.* 44(38): 12790–12796.
 15. Harel, M., Spaar, A., Schreiber, G., **2009**, Fruitful and futile encounters along the association reaction between proteins. *Biophys J.* 96(10): 4237–4248.
 16. Schiefner, A., Gebauer, M., Skerra, A., **2012**, Extra-domain B in oncofetal fibronectin structurally promotes fibrillar head-to-tail dimerization of extracellular matrix protein. *J Biol Chem.* 287(21): 17578–17588.
 17. Pankov, R., Yamada, K.M., **2002**, Fibronectin at a glance. *J Cell Sci.* 115(20): 3861–3863.
 18. Magnusson, M.K., Mosher, D.F., **1998**, Fibronectin: structure, assembly, and cardiovascular implications. *Arterioscler Thromb Vasc Biol.* 18(9): 1363–1370.
 19. George, E.L., Georges-Labouesse, E.N., Patel-King, R.S., Rayburn, H., Hynes, R.O., **1993**, Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. *Development.* 119(4): 1079–1091.
 20. Alitalo, K., Vaheri, A., **1982**, Pericellular matrix in malignant

- transformation. *Adv Cancer Res.* 37: 111–158.
21. Yamada, K.M., **1983**, *Cell surface interactions with extracellular materials.* *Annu Rev Biochem.* 52: 761–799.
 22. Ruoslahti, E., **1988**, Fibronectin and its receptors. *Annu Rev Biochem.* 57: 375–413.
 23. Leahy, D.J., Aukhil, I., Erickson, H.P., **1996**, 2.0 Å crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. *Cell.* 84(1): 155–164.
 24. Ffrench-Constant, C., Van de Water, L., Dvorak, H.F., Hynes, R.O., **1989**, Reappearance of an embryonic pattern of fibronectin splicing during wound healing in the adult rat. *J Cell Biol.* 109(2): 903–914.
 25. Kosmehl, H., Berndt, A., Katenkamp, D., **1996**, Molecular variants of fibronectin and laminin: structure, physiological occurrence and histopathological aspects. *Virchows Arch.* 429(6): 311–322.
 26. Castellani, P., Siri, A., Rosellini, C., Infusini, E., Borsi, L., Zardi, L., **1986**, Transformed Human-Cells Release Different Fibronectin Variants Than Do Normal-Cells. *J Cell Bio.* 103(5): 1671–1677.
 27. Borsi, L., Carnemolla, B., Castellani, P., Rosellini, C., Vecchio, D., Allemanni, G., Chang, S.E., Taylor-Papadimitriou, J., Pande, H., Zardi, L., **1987**, Monoclonal antibodies in the analysis of fibronectin isoforms generated by alternative splicing of mRNA precursors in normal and transformed human cells. *J Cell Biol.* 104(3): 595–600.
 28. Vartio, T., Laitinen, L., Närvänen, O., Cutolo, M., Thornell, L.E., Zardi, L.,

- Virtanen, I., **1987**, Differential expression of the ED sequence-containing form of cellular fibronectin in embryonic and adult human tissues. *J Cell Sci.* 88 (4): 419–430.
29. Zardi, L., Carnemolla, B., Siri, A., Petersen, T.E., Paoletta, G., Sebastio, G., Baralle, F.E., **1987**, Transformed human cells produce a new fibronectin isoform by preferential alternative splicing of a previously unobserved exon. *EMBO J.* 6(8): 2337–2342.
30. Carnemolla, B., Balza, E., Siri, A., Zardi, L., Nicotra, M.R., Bigotti, A., Natali, P.G., **1989**, A tumor-associated fibronectin isoform generated by alternative splicing of messenger RNA precursors. *J Cell Biol.* 108(3): 1139–1148.
31. Oyama, F., Hirohashi, S., Shimosato, Y., Titani, K., Sekiguchi, K., **1989**, Deregulation of alternative splicing of fibronectin pre-mRNA in malignant human liver tumors. *J Biol Chem.* 264(18): 10331–10334.
32. Oyama, F., Hirohashi, S., Shimosato, Y., Titani, K., Sekiguchi, K., **1990**, Oncodevelopmental regulation of the alternative splicing of fibronectin pre-messenger RNA in human lung tissues. *Cancer Res.* 50(4): 1075–1078.
33. Borsi, L., Balza, E., Allemanni, G., Zardi, L., **1992**, Differential expression of the fibronectin isoform containing the ED-B oncofetal domain in normal human fibroblast cell lines originating from different tissues. *Exp Cell Res.* 199(1): 98–105.
34. Castellani, P., Viale, G., Dorcaratto, A., Nicolo, G., Kaczmarek, J., Querze, G., Zardi, L., **1994**, The fibronectin isoform containing the ED-B oncofetal

- domain: a marker of angiogenesis. *Int J Cancer*. 59(5): 612–618.
35. Kaczmarek, J., Castellani, P., Nicolo, G., Spina, B., Allemanni, G., Zardi, L., **1994**, Distribution of oncofetal fibronectin isoforms in normal, hyperplastic and neoplastic human breast tissues. *Int J Cancer*. 59(1): 11–16.
 36. Fattorusso, R., Pellecchia, M., Viti, F., Neri, P., Neri, D., Wüthrich, K., **1999**, NMR structure of the human oncofoetal fibronectin ED-B domain, a specific marker for angiogenesis. *Structure*. 7(4): 381–390.
 37. Saw, P.E., Kim, S.H., Lee, I.H., Park, J.H., Yu, M.K., Lee, J.J., Kim, J.I., Jon, S.Y., **2013**, Aptide-conjugated liposome targeting tumor-associated fibronectin for glioma therapy. *J Mater Chem B*. 1(37): 4723–4726.
 38. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., Bax, A., **1995**, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 6(3): 277–293.
 39. Shen, Y., Delaglio, F., Cornilescu, G., Bax, A., **2009**, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. 44(4): 213–223.
 40. Schwieters, C.D., Kuszewski, J.J., Clore, G.M., **2006**, Using Xplor-NIH for NMR molecular structure determination. *Progress in Nuclear Magnetic Resonance Spectroscopy*. 48(1): 47–62.
 41. Nilges, M., Gronenborn, A.M., Brünger, A.T., Clore, G.M., **1988**, Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin,

- potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Eng.* 2(1): 27–38.
42. Clore, G.M., Nilges, M., Sukumaran, D.K., Brungerl, A.T., Karplus, M., Gronenborn, A.M., **1986**, The three-dimensional structure of α -1-purothionin in solution - combined use of nuclear-magnetic-resonance, distance geometry and restrained molecular-dynamics. *EMBO J.* 5(10): 2729–2735.
43. Clore, G.M., Gronenborn, A.M., Tjandra, N., **1998**, Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J Magn Reson.* 131(1): 159–162.
44. Kuszewski, J., Qin, J., Gronenborn, A.M., Clore, G.M., **2012**, The impact of direct refinement against ^{13}C alpha and ^{13}C beta chemical shifts on protein structure determination by NMR. *J Magn Reson B*, 106(1): 92–96.
45. Kim, S., Kim, D., Jung, H.H., Lee, I.H., Kim, J.I., Suh, J.Y., Jon, S., **2012**, Bio-inspired design and potential biomedical applications of a novel class of high-affinity peptides. *Angew Chem Int Ed Engl.* 51(8): 1890–1894.
46. Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., **1993**, Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J Appl Cryst.* 26: 283–291.
47. Wishart, D.S., Sykes, B.D., **1994**, The C-13 Chemical-Shift Index - a Simple Method for the Identification of Protein Secondary Structure Using C-13 Chemical-Shift Data. *J Biomol NMR.* 4(2): 171–180.
48. Cochran, A.G., Skelton, N.J., Starovasnik, M.A., **2001**, Tryptophan zippers:

- Stable, monomeric beta-hairpins. *Proc Natl Acad Sci USA*. 98(10): 5578–5583.
49. Bencharit, S., Cui, C.B., Siddiqui, A., Howard-Williams, E.L., Sondek, J., Zuobi-Hasona, K., Aukhil, I., **2007**, Structural insights into fibronectin type III domain-mediated signaling. *J Mol Biol*. 367(2): 303–309.
50. HashimotoUoshima, M., Yan, Y.Z., Schneider, G., Aukhil, I., **1997**, The alternatively spliced domains EIIIB and EIIIA of human fibronectin affect cell adhesion and spreading. *J Cell Sci*. 110: 2271–2280.
51. Kim, D., Lee, I.H., Kim, S., Choi, M., Kim, H., Ahn, S., Saw, P.E., Jeon, H., Lee, Y., Jon, S., **2014**, A Specific STAT3-Binding Peptide Exerts Antiproliferative Effects and Antitumor Activity by Inhibiting STAT3 Phosphorylation and Signaling. *Cancer Res*. 74(8): 2144–2151.
52. Kim, H., Lee, Y., Lee, I.H., Kim, S., Kim, D., Saw, P.E., Lee, J., Choi, M., Kim, Y.C., Jon, S., **2014**, Synthesis and therapeutic evaluation of an aptide-docetaxel conjugate targeting tumor-associated fibronectin. *J Control Release*. 178: 118–124.
53. Park, J., Kim, S., Saw, P.E., Lee, I.H., Yu, M.K., Kim, M., Lee, K., Kim, Y.C., Jeong, Y.Y., Jon, S., **2012**, Fibronectin extra domain B-specific aptide conjugated nanoparticles for targeted cancer imaging. *J Control Release*. 163(2): 111–118.

IV. 국문요약

Part I. 상자성 완화 현상을 통한 단백질 복합체의 표적탐색과정의 가시화

현재까지는 X-ray 와 NMR 을 통해 많은 단백질과 단백질 복합체의 구조를 밝히는 연구가 진행 되었음에도 불구하고, 생체 내에서 그 기능을 수행 할 수 있도록 하는 최종적인 복합체의 형성 과정에 대한 연구 성과는 미비한 실정이다. 최근에 실험적 결과를 통해서가 아닌 시뮬레이션을 이용한 연구에서 두 단백질이 우연히 충돌을 했을 때, 일시적인 비특이 복합체 (encounter complex)를 형성할 수 있음이 예측되었다. Encounter complex 는 안정화 시키는 상호작용이 아직 약하기 때문에 쉽게 다시 떨어질 수 있고, 일부는 상호 인력을 유지하며 완전히 떨어지기 전에 조금씩 위치를 바꾸어 가며 다른 encounter complex 로 바뀔 수 있다. 이 과정에서 일부 encounter complex 는 실제로 기능을 하는 특이 복합체를 형성해 안정화된 복합체를 이루는데, 이러한 일련의 과정을 표적탐색과정 (target search process)이라고 부른다.

중간 산물에 해당하는 encounter complex 는 불안정하여 이론적으로 그 존재가 예측되어 왔을 뿐, 실험적으로 이들을 직접 관찰하고 분석하는 일은 불가능하게 생각되어 왔다. 그러나 2006 년 처음으로 NMR 상자성 완화 (paramagnetic relaxation enhancement, PRE)라는 현상을 이용하여 순간적인 encounter complex 를 관찰하는 실험이 발표되었다. 현 논문에서는 상자성 완화 현상을 이용하여 encounter complex 에 관한 추가적인 생물적 연구인 동시에 더 나아가 target

searching pathway 에 대한 연구를 진행을 하였다. 이는 비특이적으로 복합체를 형성하는 미량으로 존재하는 encounter complex ensemble 중 특이 복합체로 진화하는 핵심적인 encounter complex 가 존재하며, 이러한 복합체를 통해 비특이적 복합체에서 특이적 복합체로 이어지는 pathway 가 존재함을 증명하기 위한 연구이다. 이와 같은 실험을 증명하기 위해서 *E. coli* 의 phosphotransferase system (PTS)을 구성하는 N-terminal domain of enzyme I (EIN)과 Histidine phosphocarrier protein (HPr)을 사용하였다.

EIN 과 HPr 의 표적탐색과정을 탐색하기 위해 EIN 의 단백질 표면의 특이 복합체를 형성하는 부분 이외에 표적탐색에서 중요한 단백질 표면에 있는 위치를 찾기 위해서 단백질을 mutants 를 제작하였다. 이 단백질의 2, 3 차 구조의 형성과 본래 EIN 단백질과 동일한 구조를 가지고 있음을 확인하기 위해서 circular dichroism (CD) spectroscopy, $^1\text{H}-^{15}\text{N}$ heteronuclear single quantum coherence (HSQC) spectroscopy, residual dipolar coupling (RDC) 측정을 통해 확인 할 수 있었다. 또한 EIN mutants 와 HPr 의 복합체 형성 affinity 를 확인하기 위해 isothermal titration calorimetry (ITC)를 사용하였으며, 최종적으로 HPr 을 통해 encounter complex 의 형성을 관찰하여 HPr 단백질이 특이 복합체를 형성하기 위한 중요한 encounter complex 를 찾고 그 복합체로 가는 pathway 를 증명 할 수 있었다.

단백질의 표적탐색과정을 이해 함으로서 이를 제어할 수 있게 된다면, 기존의 단백질 공학에 유용한 도구로 사용될 수 있을 것이다. 표적탐색과정을 보다 깊이 이해하고 제어하는 길이 열리면 단백질의 활성부위를 교란하지 않으면서 결합력을 조절하는 새로운 방법 일 것이다. 상자성 완화를 이용한 연구는 측정과 분석 과정을 구축해 놓으면 다양한 단백질 시스템에 대하여 응용이 가능하다.

Part II. Fibronectin Extradomain B 와 Aptide 의 표적 인식 메커니즘과 구조 분석에 대한 연구

현재까지 질병과 관련된 단백질 또는 질병 유발 인자를 타겟으로 하는 높은 친화력을 가지며 특이적으로 결합하는 여러 가지 화합물, 항체들이 연구 되었고, 진단과 치료 분야에서 많은 이용 가치를 가지고 활용 되었다. 특히 항체에 관한 연구가 지속적으로 이루어져 왔으나 그럼에도 불구하고 항체는 생산비용, 생체 내외 변질이라는 여러 가지 문제점이 존재한다. 하지만 일반적으로 펩타이드는 생산 단가가 낮고 반응성이 높으며 면역반응을 잘 일으키지 않으며 합성을 통해서 생산되기 때문에 변형하기 쉽고 순도가 높은 제품을 만들 수 있는 장점을 가지고 있다. 하지만 기존의 펩타이드 역시 그 단점으로 후보약물이 되기 위해서는 낮은 타겟 친화력과 특이성, 단백질 가수분해 효소에 의한 생체 내 불안정성과 같은 한계점을 극복해야 한다. 현 논문에서 언급 되어지는 펩타이드 압타머 (Aptide)는 새로운 스케폴드를 통해 특정 기질에 대한 높은 친화력과 가질 수 있도록 디자인 된 peptide 로서 현재 까지 그 성질과 복합체 형성과정에 대한 연구가 아직 전무한 상태이다.

Aptide 는 β -hairpin motif scaffold 의 trpzip 과 N- 과 C- 말단 부분에 각각 6 개의 아미노산으로 이루어진 총 26 개의 아미노산을 가진 펩타이드이다. 이는 기존의 펩타이드의 단점인 안정성을 trpzip 의 안정한 구조를 통해 극복함과 동시에 표적결합부위의 두 개의 구조로 이루어진 가변적 펩타이드를 이용하여 시너지 효과를 일으켜 타겟에 대해 매우 높은 친화력과 특이성을 확보하고 있다. 그리고 fibronectin extra-domain B (EDB), vascular endothelial growth factor (VEGF), human serum albumin (HSA)과 hexa-histidine tag 과 같은 여러 다른 타겟에 대해서 모두 nanomolar 범위의 친화력을 가진다. 특히 aggressive solid human tumors 의 바이오마커로서 alternative splicing 된

fibronectin EDB 에 대해서 aptide 는 수 nanomolar 수준의 친화력을 가지 있다. 본 논문은 EDB 단백질과 이와 친화력을 가지고 있는 aptide (APT_{EDB})의 복합체의 3 차 구조를 NMR 를 이용하여 밝혀내고, 복합체 형성 시 발생하는 EDB 와 APT_{EDB} 의 특이적인 결합 메커니즘을 규명하는데 그 연구 목적을 두고 있다.

EDB 와 APT_{EDB} 의 복합체의 3 차원 구조를 밝히기 위해서 NMR 를 통해 분석을 실시하였다. EDB 와 APT_{EDB} 의 backbone 과 각각의 residues 의 side chain 의 assignment 는 CBCACONH, HNCACB, HBHA(CO)NH, HCCH-TOCSY, CCH-TOCSY, ¹⁵N-TOCSY-HSQC 3 차원 NMR spectrum 을 이용하여 분석을 하였고, side chain 에 존재하는 aromatic ring 의 assignment 는 ¹³C-NOESY-HSQC, HDCB NMR 실험을 통해 분석을 하였다. 구조를 계산하기 위한 거리 정보는 ¹³C-separated NOESY, ¹⁵N-separated NOESY, ¹³C-separated/¹²C-filtered NOESY 실험 통해 얻었다. 이와 같은 실험을 바탕으로 단백질 복합체 구조계산은 Xplor-NIH 를 통하여 구조를 계산 하였으며, EDB 와 APT_{EDB} 복합체의 구조를 연구 하였다. 형성 시 특이적인 표적 인식 메커니즘을 가지는 것을 알 수 있었다.

EDB 와 APT_{EDB} 복합체 형성 시 일반적인 단백질 표면에 결합 부위가 존재하여 결합체를 만드는 것과 달리, 특정 2 차 구조의 unfolding 을 통해 단백질 내부에 존재하는 결합 부위가 표면으로 노출 되어 그 부위에 APT_{EDB} 가 결합한다. 이러한 결합체 형성은 현재 처음으로 보고가 되어졌다.