

3. ผลการดำเนินงานวิจัย

3.1 DNA receptor (ComEA)

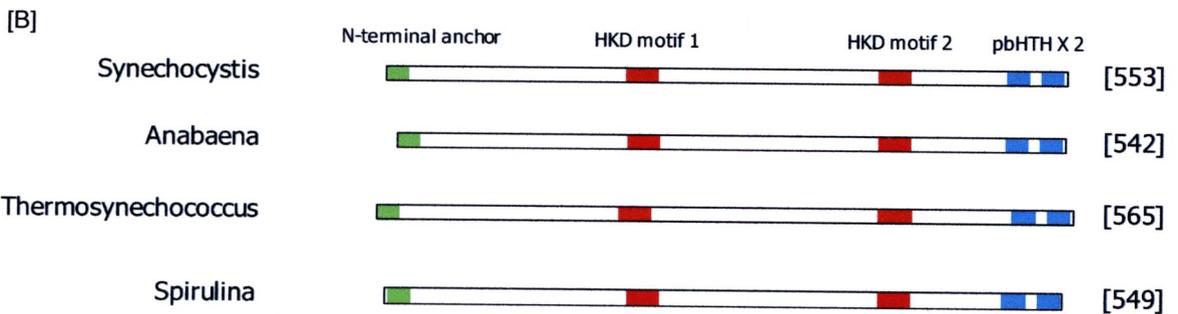
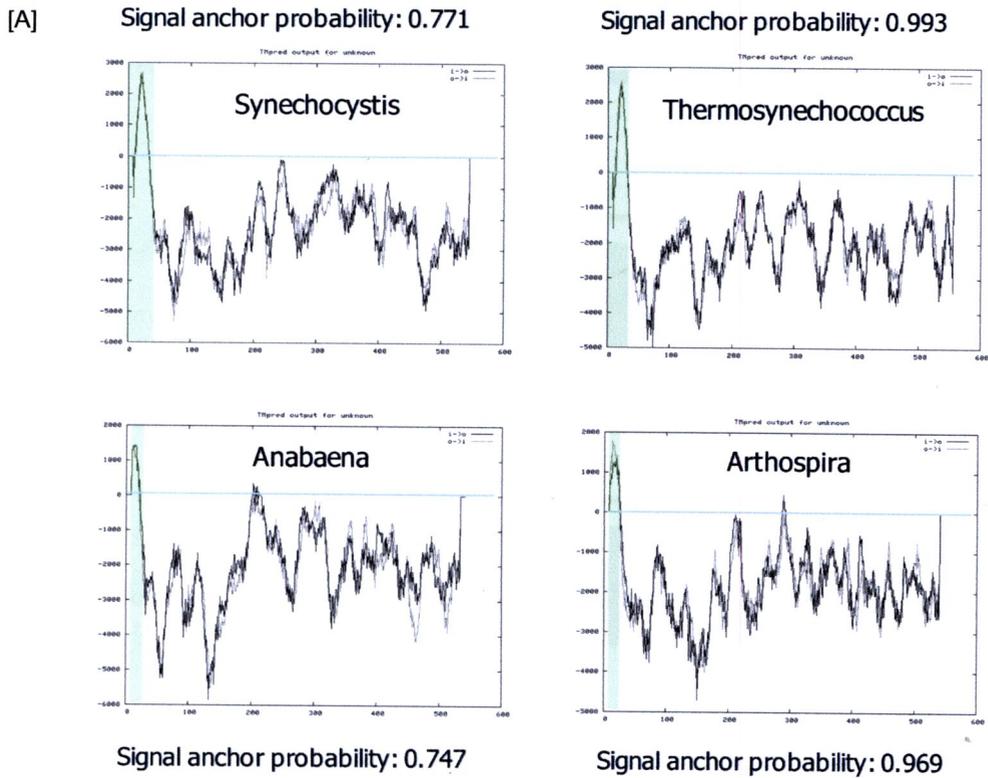
จากข้อมูลจีโนมของ *Spirulina* พบว่ายีน PY01430014V03 มีลำดับของกรดอะมิโนคล้ายคลึงกับยีน ComEA ที่พบในไซยาโนแบคทีเรียชนิดอื่นมากที่สุด (Based on sequence homology) นอกจากนั้นเมื่อทำการวิเคราะห์คุณสมบัติของยีน (Topology analysis) พบว่าทางด้านปลาย N-terminal ของเอนไซม์ PY01430014V03 ซึ่งยาวประมาณ 30 อะมิโน นั้นมีคุณสมบัติเป็น hydrophobic residue และเมื่อทำการวิเคราะห์ signal sequence ที่บริเวณดังกล่าวด้วยโปรแกรม SignalP 3.0 [Bendsten, et al., 2004] และ SIG-Pred พบว่าบริเวณดังกล่าวเป็น uncleaved signal peptide หรือ ที่รู้จักกันในชื่อของ “signal anchor” ซึ่งจะทำหน้าที่ในการแทรกผ่านเข้าไปในผนังเซลล์ (periplasmic transmembrane) และยึดเกาะอยู่กับผนังเซลล์ ซึ่งการมีคุณสมบัติดังกล่าวแสดงให้เห็นว่า PY01430014V03 นั้นเป็น transmembrane protein (ดังแสดงในรูปที่ 3.1)

นอกจากนั้นเมื่อทำการเปรียบเทียบเอนไซม์ที่เกิดจากยีน PY01430014V03 ซึ่งมีความยาว 549 อะมิโน กับเอนไซม์ ComEA ของไซยาโนแบคทีเรียอื่นๆพบว่า เอนไซม์ PY01430014V03 นั้นมีบริเวณอนุรักษ์ (Conserved regions) ที่ตรงกับ ComEA ของไซยาโนแบคทีเรียชนิดอื่น และบริเวณดังกล่าวยังเป็นบริเวณที่สำคัญและจำเป็นต่อการทำงานของเอนไซม์ ComEA อีกด้วย (ดังแสดงในรูปที่ 3.2) ซึ่งบริเวณต่างๆเหล่านั้นประกอบด้วย HKD motif 2 บริเวณ และ pbHTH (phosphate-binding helix-turn-helix) module 2 บริเวณ โดย HKD motif คือบริเวณที่มีลักษณะของ amino acid sequence เป็น HXKXXXXDXXXV (X แทนกรดอะมิโนชนิดใดก็ได้) ซึ่งในเอนไซม์ PY01430014V03 นั้นพบ HKD motif อยู่ 2 แห่งบริเวณส่วนกลางของเอนไซม์ เช่นเดียวกับที่พบในเอนไซม์ ComEA ของ *Synechocystis* ซึ่ง HKD motif นี้จะเกี่ยวข้องกับการทำหน้าที่ในการย่อยสาย Double strand DNA บริเวณผิวของเซลล์ (endonuclease domain) [Yura, et al., 1999] นอกจากนี้ที่บริเวณส่วนปลายด้าน C-terminal (ยาวประมาณ 60 อะมิโน) มี amino acid sequence ตรงกับ pbHTH module ซึ่งพบอยู่ด้วยกัน 2 แห่ง เช่นเดียวกับที่ปรากฏในเอนไซม์ ComEA ของไซยาโนแบคทีเรียอื่นๆทั้ง *Synechocystis* sp. PCC6803 (slr0197), *Anabaena* sp. PCC7120 (all3087) และ *Themosynechococcus elongatus* BP-1 (tll2339) ซึ่งบริเวณนี้มีรายงานใน *Synechocystis* ว่าทำหน้าที่ในการจับ (interact) กับ DNA [Yura, et al., 1999] ดังนั้นจึงสรุปได้ว่า pbHTH module ทำหน้าที่เป็น DNA-binding domain ในเอนไซม์ ComEA

จากการพิจารณาเปรียบเทียบลักษณะต่างๆของเอนไซม์ PY01430014V03 ของ *Spirulina* กับเอนไซม์ ComEA ของไซยาโนแบคทีเรียต่างๆนั้น สามารถสรุปได้ว่าเอนไซม์ PY01430014V03 ทำหน้าที่เป็น DNA receptor หรือ ComEA เนื่องจากคุณสมบัติต่างๆดังนี้

- (1) เป็น Transmembrane protein เนื่องจากมีบริเวณปลาย N-terminal เป็น signal anchor ที่ทำหน้าที่แทรกตัวเข้ายึดอยู่กับ membrane
- (2) มี HKD motif ที่ทำหน้าที่เป็น nuclease domain ซึ่งจำเป็นต่อการรับ DNA จากภายนอกเข้าสู่ภายในเซลล์ (DNA uptake/ DNA transport)
- (3) มีบริเวณที่ทำหน้าที่จับกับ DNA ที่จะถูกนำเข้าสู่ภายในเซลล์ (DNA-binding domain) คือ pbHTH module

และจากคุณสมบัติดังกล่าวแสดงให้เห็นว่า PY01430014V03 ของ *Spirulina* มีความเกี่ยวข้องและสำคัญต่อกระบวนการส่งผ่าน DNA จากภายนอกเข้าสู่ภายในเซลล์



รูปที่ 3.1: (A) hydropathy plot ของเอนไซม์ ComEA จาก *Synechocystis* slr0179, *Anabaena* all3087, *Thermosynechococcus* tll2339 และ PY01430014V03 ของ *Spirulina* จากผลการวิเคราะห์ด้วยโปรแกรม TMpred [Hofmann and Stoffel, 1993] และ signal peptide probability score ด้วยผลการวิเคราะห์ด้วยโปรแกรม SignalP (B) แสดงการเปรียบเทียบโครงสร้างของเอนไซม์ PY01430014V03 ของ *Spirulina* เปรียบเทียบกับไซยาโนแบคทีเรียชนิดอื่น โดยกล่องสีเขียวแสดงแทนบริเวณที่เป็น putative signal anchor (transmembrane region) กล่องสีแดงแสดงแทนบริเวณที่เป็น HKD motif และกล่องสีน้ำเงิน แสดงแทนบริเวณที่เป็น pbHTH module

HKD motif (1)

<i>Synechocystis</i>	190	DGSKGSGLMHHKFMVLDNKTVIIVTGANFTPSDQ	slr0197	
<i>Anabaena</i>	181	DGSAGSKLMHHKFFVVDNRIVIVTGANFTLSDV	all13087	(A)
<i>Thermosynech.</i>	191	DGSKGSMLMHHKFIIVIDNRQVIATTANFTLSDV	t112339	
<i>Arthrospira</i>	191	DGSKGSGLMHHKFFVIVDQNTVIIVTGANFTTSDI	PY01430014V03	

HKD motif (2)

<i>Synechocystis</i>	397	RGDKLHHKFAALVDIDDTVITGSHNWSPAANHN	slr0197	
<i>Anabaena</i>	388	EGDLLLHHKFSVIDNQIVITGSHNWSAANYG	all13087	(B)
<i>Thermosynech.</i>	405	EGDKLHHKYGVVDDRIVIVGSHNWSAANRG	t112339	
<i>Arthrospira</i>	397	EGDRLHHKFAVIDDQIVITGSHNWSRAANHS	PY01430014V03	

pbHTH (1)

pbHTH (2)

<i>Synechocystis</i>	501	QEELES LP GIGPSTAEKIIISTRE-[7]-EDLGRVKGIGPSKIEGLRGKVTW		
<i>Anabaena</i>	490	LAELETLP GV GKKAQKIIIRQ-[7]-QDIDKVPGISPMMIENWEGRIQF		(C)
<i>Thermosynech.</i>	513	AAELETLP GV GSKI AA IIIRARE-[7]-ADIDAVPGVGP PK LLERLRDRVTW		
<i>Arthrospira</i>	497	QPELES LP GIGPGLAQR II EARS-[7]-DDVTRVSGIG PK TIERWENLV TW		

รูปที่ 3.2: การวิเคราะห์เปรียบเทียบระหว่าง ComEA ของไซยาโนแบคทีเรีย โดยบริเวณที่เป็นแถบสีดำ และสีเทาแสดงตำแหน่งของกรดอะมิโนที่เป็น identity และ similar กันตามลำดับ (A) แสดงบริเวณ HKD motif ตำแหน่งที่ 1 (B) แสดงบริเวณ HKD motif ตำแหน่งที่ 2 และ (C) แสดงบริเวณที่พบ pbHTH modules ทั้ง 2 ตำแหน่ง โดยตัวเลขที่อยู่ในวงเล็บแสดงถึงจำนวนของกรดอะมิโนที่อยู่ระหว่างpbHTH module ทั้ง 2 ตำแหน่ง

3.2 Membrane channel protein (ComEC)

Membrane channel protein (ComEC) เป็นหนึ่งในโปรตีนที่มีบทบาทสำคัญอย่างมากต่อการส่งผ่าน DNA จากภายนอกเข้าสู่ภายในเซลล์ เนื่องจากทำหน้าที่ในการสร้างช่องทางที่ DNA จะผ่านเข้าสู่เซลล์ที่เรียกว่า "Aqueous channel" [Chen and Dubnau, 2004] ซึ่งจากการวิเคราะห์หา ComEC ในจีโนมของ *Spirulina* พบว่ามีเพียงหนึ่งยีนที่มี amino acid sequence ใกล้เคียงกับ ComEC ของไซยาโนแบคทีเรียชนิดอื่น (เช่น *Synechocystis* sll1929, *Anabaena* all1983 และ *Thermosynechococcus* thl1702) คือโปรตีนที่ได้จากยีน PY01540002V03

และจากการวิเคราะห์ Membrane topology ของโปรตีน PY01540002V03 ดังแสดงในรูปที่ 3.3 และ 3.4 พบว่าบนสายโปรตีนนี้ประกอบด้วยบริเวณที่มีคุณสมบัติเป็น hydrophobic อยู่มากทั้งบริเวณส่วนปลาย N-terminal และบริเวณส่วนกลาง (central region) ของสายโปรตีน ซึ่งบริเวณส่วนกลางนี้ถูกเรียกว่า "Competence domain" ซึ่งเป็นบริเวณสำคัญที่ทำหน้าที่ในการฟอร์มตัวเป็นช่องทาง (aqueous pore) เพื่อให้ DNA ผ่านเข้าสู่ภายในเซลล์ได้ ในขณะที่ hydrophobic region ที่ปลาย N-terminal (N-terminal loop) จะทำหน้าที่ในการช่วยรักษาเสถียรภาพ (stability) ของโครงสร้างของโปรตีน ComEC เอาไว้ [Draskovic and Dubnau, 2005] และจากข้อมูลเหล่านี้จึงสามารถเชื่อได้ว่าโปรตีน PY01540002V03 ทำหน้าที่เป็น membrane channel protein ที่ทำงานร่วมกับ DNA receptor protein (ComEA) ในการนำ DNA จากภายนอกเข้าสู่ภายในเซลล์

TM1

sll1929	234	P L L I S S I V L E A K A V N L P T E L R N L F Q S V G L S E F L A A S G Y Q V S L I V G T V L L
all1983	266	E P I V S A M V L G S K A V D L P Y D I R D L E V Q A G L A A L A A S G F Q T S I I I S V I L C
t111702	237	H E A V V S G M V L G N R A V A I P F E V R D S E R R A G L S A L A A S G F H T A I L L A V V L T
PY01540002V03	263	K E T I L S A M V L G R R A V D I S P E I S D R E V R V G E A T L A A S G F H V S L L I G V V L A

TM2

sll1929	284	I G K K A G V S S K L A I A V G L I T L A F Y L G L T G L E P S V I F R A S L L W V G V M A A I A S E
all1983	316	L T K R A -- K K V T Q V S L G A L A I I F L S T G F Q P A V I E A V I M G F A L V G L A L D
t111702	287	A T R P L -- P Q R W R S G L G A G V L V F F A F L T G F A P S A M F A A L M G L A G L V A I V N G
PY01540002V0	313	L T A R W -- S P K L R L V F S A T I L L Y V G L T G G S A S V L R A A L M G L A A L V G V M K

		TM3	TM4
sll1929	334	Q K L N T V G A L L L I A T V M L V N F Y W S Q L G F Q L S F L A T L G I V I M V P P L Q Q R I	
all1983	364	R E V K Q I G S L L L A A T I L I V F N P W I W L G F Q L S F L A T L G I V I S V E A I T S F I	
t111702	285	Q K G Q P L V I L L A T T A M L I Y N F W I E I I G F Q L S F L A T L G L I I S A Q E I S D R I	
PY01540002V03	361	R Q V R P L A V L A V A V F L L W N E L L Q L G F F E S F L A T L G L L T V P P L M R S I	

		TM5	TM6
sll1929	384	D F L P G K I A G V I A V P L P A S V T T E I L L Q C F G Y F I P A A L P L N L I T E L L L A L I	
all1983	414	G W I P P A I A S L I S P L A A T I T T E I Q L F V E G V V P A Y S L L L N I T T P L I S I I	
t111702	335	H W I P T P L R S P L A V A L A A T L V L E I S L A I E C T I P V Y G L P A V I A N F L I L I	
PY01540002V03	411	Q W L P P V I A S C I A V P I P A T V I L E I Q I Y V E N L V S P F S I F V N L V V E F I T V I	

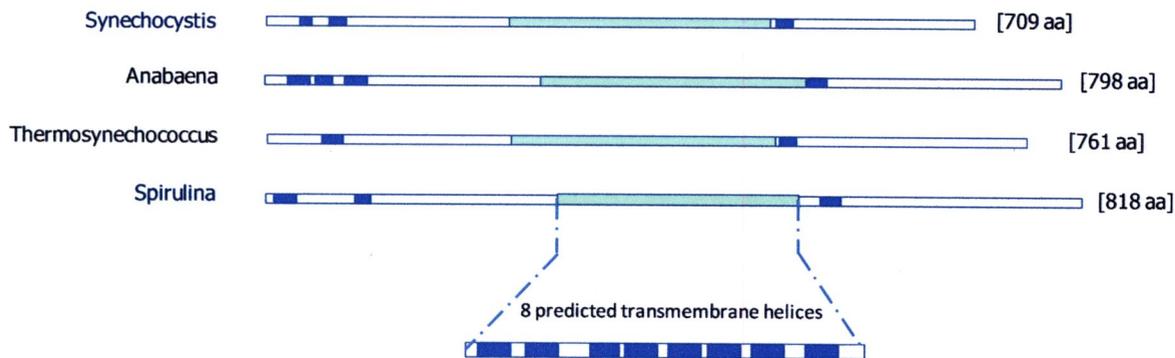
TM7

sll1929	434	S L G G M V S A L C A L I F P P L G S A T G W L A F F A R W L M A I A D Q F Q L E -- T V A V G
all1983	464	S I G G I I S A M F A L L S G A G S F V A G F L H Y P T D W I K L V E I F S Q L F C N S L I V G
t111702	385	T V G G F I S A T V G F L P P L G S A L A W L L S Y T A L L W L I E T I G R W E G A V I A L G
PY01540002V03	461	T L G G M L T S V A I L S P W A A A L A T A L Y Y L E G I I Y I V E F F N D L P G N S I A V G

TM8

sll1929	482	E V G L G V V V I M Y G A F C I I M A F
all1983	514	S I S T W L L A I Y G L V F L V I V
t111702	435	S L G W L Q V I V L G L M L I V N L S
PY01540002V03	511	S L Y M A L I G L Y T I N I I C L L

รูปที่ 3.3: แสดงผลการเปรียบเทียบ amino acid sequence บริเวณที่เป็น competence domain ของโปรตีน ComEC ของไซยาโนแบคทีเรียต่างๆ ประกอบด้วย PY01540002V03 ของ *Spirulina*, *Synechocystis* sll1929, *Anabaena* all1983 และ *Thermosynechococcus* t11702 โดยตำแหน่งที่เป็น identical amino acid และ similar amino acid แสดงด้วยแถบสีดำและสีเทาตามลำดับ และกล่องสีแดงแสดงบริเวณที่มีคุณสมบัติเป็น hydrophobic region



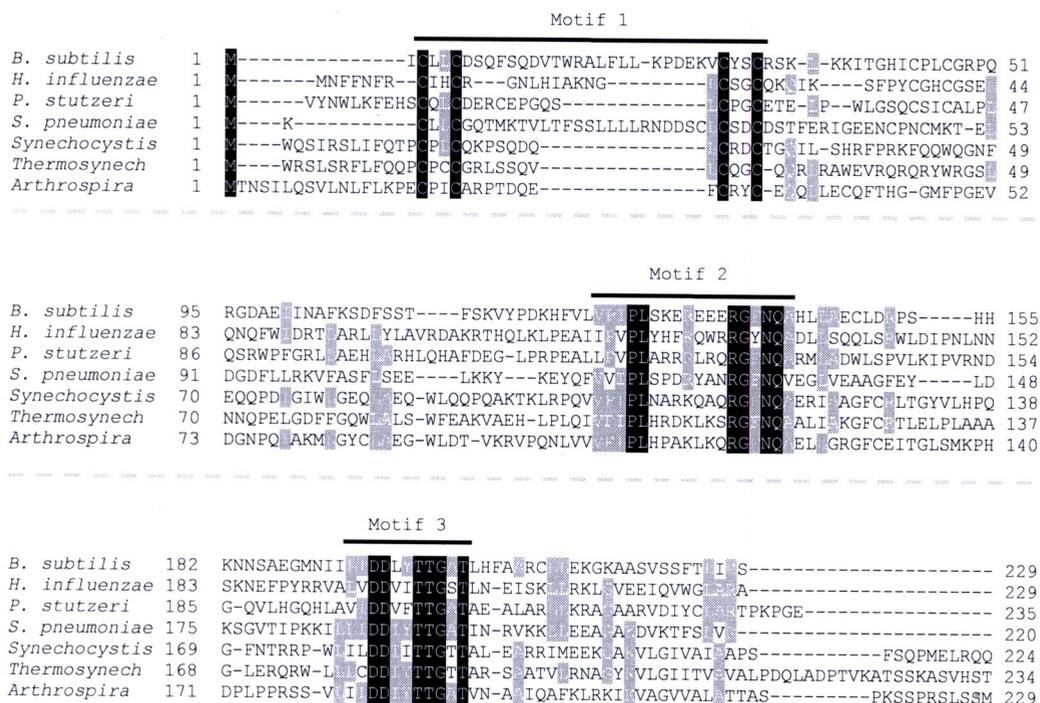
รูปที่ 3.4: แสดงการเปรียบเทียบโครงสร้างของโปรตีน PY01540002V03 ของ *Spirulina* เทียบกับโปรตีน ComEC ของไซยาโนแบคทีเรียชนิดอื่น โดยสีเขียวแสดงบริเวณที่เป็น competence domain และสีน้ำเงินแสดงบริเวณที่เป็น transmembrane region ที่วิเคราะห์เทียบกับฐานข้อมูลของ SMART [Letunic, et al., 2006] และ Pfam [Finn, et al., 2006] ซึ่งพบว่า competence domain ของ *Spirulina* ประกอบด้วย 8 transmembrane helices ที่จะทำการฟอร์มตัวเป็นช่อง (pore) สำหรับให้ DNA จากภายนอกผ่านเข้าสู่ภายในเซลล์

3.3 The ATP-binding protein (ComFA/ComF)

ในกระบวนการส่งผ่านยีนจากภายนอกเข้าสู่ภายในเซลล์ สายของ DNA ที่จะสามารถส่งผ่านเข้าสู่ cytoplasm ได้ นั้น จะต้องมิลักษณะเป็นสายเดี่ยว (single strand) ดังนั้นก่อนที่สาย DNA จะผ่าน pore channel ได้จึงจำเป็นต้องมีเอนไซม์มาทำหน้าที่ย่อยสาย DNA สายคู่ (double strand) ให้อยู่ในรูปของ DNA สายเดี่ยวเสียก่อน ซึ่งเอนไซม์ที่รับหน้าที่ดังกล่าวคือ ComFA หรือ ComF

เมื่อทำการเปรียบเทียบข้อมูลระดับ Amino acid sequence พบว่าในจีโนมของ *Spirulina* มีอยู่เพียง 1 โปรตีน คือ PY00380001V03 ซึ่งมีความคล้ายคลึงกับ amino acid sequence ของโปรตีน ComF ในระดับที่ไม่สูงนัก ดังแสดงไว้ในตารางที่ 3.1 แต่อย่างไรก็ตามเมื่อทำการเปรียบเทียบหาบริเวณอนุรักษ์ระหว่างโปรตีนในกลุ่มนี้ ด้วยเทคนิค multiple alignment (โดยอาศัยโปรแกรม clustalW) พบว่ามีบริเวณอนุรักษ์อยู่ 3 บริเวณด้วยกัน โดยบริเวณที่หนึ่งอยู่ส่วนปลายทางด้าน N-terminal คือ putative tetracysteine motif (4cys) ซึ่งมีลักษณะเป็น CxxC--CxxC motif ซึ่งเป็นลักษณะที่มักปรากฏเสมอในโปรตีนที่ทำหน้าที่เกี่ยวข้องกับกระบวนการส่งถ่ายยีนจากภายนอกเข้าสู่เซลล์ (DNA transfer) และกระบวนการที่เกี่ยวข้องกับ type-4 piliation นอกจากนี้ยังมีรายงานว่า 4cys motif นี้ยังเกี่ยวข้องกับการจับกับ metal ion อีกด้วย [Possot and Pugsley, 1997] [Xia, et al., 2005] บริเวณที่สอง คือ (PLxxxxxxRGxNQ) ซึ่งยังไม่มีรายงานในฐานข้อมูลทางด้าน motif และ domain แต่จากการศึกษาใน *Synechocystis* คาดว่าบริเวณนี้จะเกี่ยวข้องกับการจับกับสารตั้งต้น (Substrate selectivity และ Catalytic stability) [Nakasuki, et al., 2006] และในบริเวณที่สามคือ PRT domain signature (ตรงกับข้อมูล PF00156 ของฐานข้อมูล Pfam และ PS00103 ของฐานข้อมูล Prosite) หรือ phosphoribosyltransferase (PRTase) enzymes โดย domain นี้มักเกี่ยวข้องกับการกระบวนการทางชีวเคมีภายในเซลล์เป็นจำนวนมากโดยเฉพาะกระบวนการที่เกี่ยวข้องกับ nucleotide [Craig and Eakin, 2000; Sinha and Smith, 2001] รวมทั้งการทำงานของ PRT domain นี้ยังจำเป็นต้องมี metal ions มาเกี่ยวข้องด้วย [Sinha and Smith, 2001; Eriksen, et al., 2002] ยิ่งไปกว่านั้นยังพบว่า PRT domain นั้นทำหน้าที่ในการจับกับ DNA โดยตรงอีกด้วย [Sinha, et al., 2003]

จากข้อมูลทั้งหมดจึงสามารถสรุปได้ว่าโปรตีน PY00380001V03 ของ *Spirulina* น่าจะเป็นโปรตีน ComF และทำหน้าที่เกี่ยวข้องโดยตรงต่อกระบวนการส่งถ่ายยีนจากภายนอกเซลล์เข้าสู่ภายในเซลล์



รูปที่ 3.5: ผลการเปรียบเทียบ amino acid sequence ของโปรตีน ComF ระหว่าง *Spirulina* และไซยาโนแบคทีเรียชนิดต่างๆ เช่น *Synechocystis sp.* PCC 6803, *B. subtilis*, *H. influenzae*, *P. stutzeri*, and *S. pneumonia*

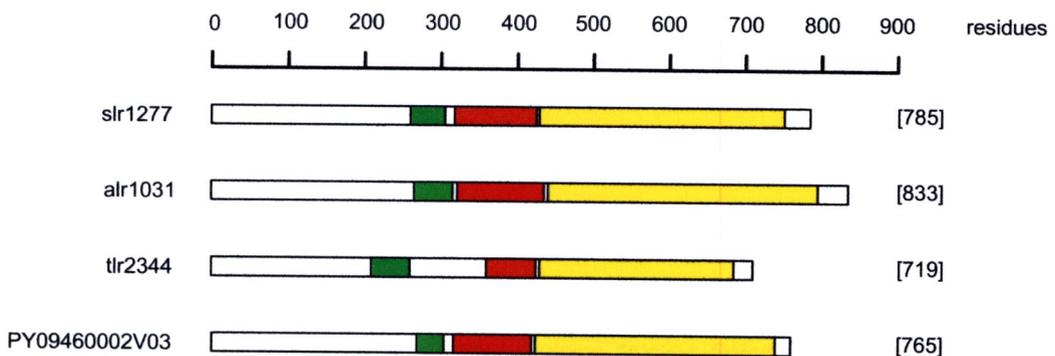
ตารางที่ 3.1: แสดงค่าความคล้ายคลึงกันระหว่าง amino acid sequence ของโปรตีนกลุ่ม ComF ระหว่าง *Spirulina platensis* strain C1, *Synechocystis sp.* strain PCC 6803, *Thermosynechococcus elongatus* strain BP-1, *B. subtilis*, *H. influenzae*, *P. stutzeri* และ *S. pneumonia* โดยค่าที่แสดงในตารางเป็นค่า identity score และ (similarity score) ซึ่งเป็นผลมาจากวิเคราะห์ด้วยโปรแกรม Blastp ของ NCBI

	<i>Synechocystis</i> slr0388	<i>T. elongatus</i> tlr1336	<i>B. subtilis</i> ComFC	<i>H. influenzae</i> com101A	<i>P. stutzeri</i> ComF	<i>S. pneumoniae</i> ComFC
<i>S. platensis</i> C1 PY00380001V03	43% (58%)	39%(53%)	28% (44%)	35% (59%)	27% (49%)	28% (41%)
<i>Synechocystis</i> slr0388	-	40% (57%)	26% (43%)	30% (47%)	28% (40%)	30% (42%)
<i>T. elongatus</i> tlr1336	-	-	27% (43%)	28% (47%)	35% (55%)	26% (40%)
<i>B. subtilis</i> ComFC	-	-	-	23% (39%)	26% (43%)	34% (52%)
<i>H. influenzae</i> com101A	-	-	-	-	35% (57%)	24% (43%)
<i>P. stutzeri</i> ComF	-	-	-	-	-	24% (41%)

3.4 Secretin (PilQ)

โปรตีน secretin PilQ เป็นโปรตีนที่สำคัญในกระบวนการ type 4 pilus เนื่องจากเป็นโปรตีนที่เป็นองค์ประกอบของโครงสร้าง Pilus ที่ถูกสร้างไว้ที่ผิวเซลล์ (cell surface) โดยจะจับตัวกันในลักษณะแบบ dodecameric cage-like macromolecular complex [Collins, et al., 2001; Bayan, et al., 2006; Frye, et al., 2006] และโครงสร้างดังกล่าวมีความเกี่ยวข้องกับกระบวนการส่งถ่ายยีนในขั้นตอนแรกๆ เนื่องจากโครงสร้างของ pilus จะทำให้เกิดช่องว่างสำหรับให้ DNA ผ่านผนังเซลล์ภายนอก (outer-membrane channel) เข้ามาได้ [Assalkhou, et al., 2007] ใน *Spirulina* นั้นพบเพียงยีนเดียวที่มี amino acid sequence คล้ายกับ secretin PilQ ของไซยาโนแบคทีเรียชนิดอื่น (*Synechocystis* slr1277, *Anabaena* alr1031 และ *Thermosynechococcus* tlr234) คือยีน PY09460002V03

นอกจากนี้เมื่อทำการวิเคราะห์ enzyme topology ของเอนไซม์ที่ได้จากยีน PY09460002V03 เทียบกับ secretin PilQ ของไซยาโนแบคทีเรียชนิดอื่นๆ พบว่าทั้งหมดมีลักษณะทาง topology เหมือนกัน ดังแสดงในรูปที่ 3.6 คือประกอบด้วย 3 domains สำคัญ ได้แก่ (1) STN domain (PF07660), (2) Secretin_N domain (PF03958) และ (3) Secretin domain (PF00263) โดย STN domain นั้นพบอยู่บริเวณปลาย N-terminal ทำหน้าที่ในการเลือกจับกับ substrate (the uptake of selective substrates) สำหรับโปรตีนที่เกี่ยวข้องกับกระบวนการ secretion [Chen, et al., 2004] ในขณะที่อีกสอง domains ทำหน้าที่เกี่ยวข้องโดยตรงกับการสร้างช่องที่ cell membrane โดยจะมาเชื่อมต่อกันเป็นโครงสร้างขนาดใหญ่ (oligomerization) ที่เรียกว่า "pilus" [Brok, et al., 1999; Bitter, et al., 2007] โดย DNA จากภายนอกจะสามารถผ่านเข้ามาทางช่องนี้ โดยอาศัยกลไกที่เซลล์ทำการหด pilus เข้ามา (retracting pili) ซึ่งจะทำให้ DNA ที่เกาะอยู่กับส่วนของเซลล์ (DNA-binding components) เช่น PilQ เคลื่อนที่เข้าสู่ภายในเซลล์ได้ผ่านทางช่องที่เปิดนี้ได้ [Assalkhou, et al., 2007]

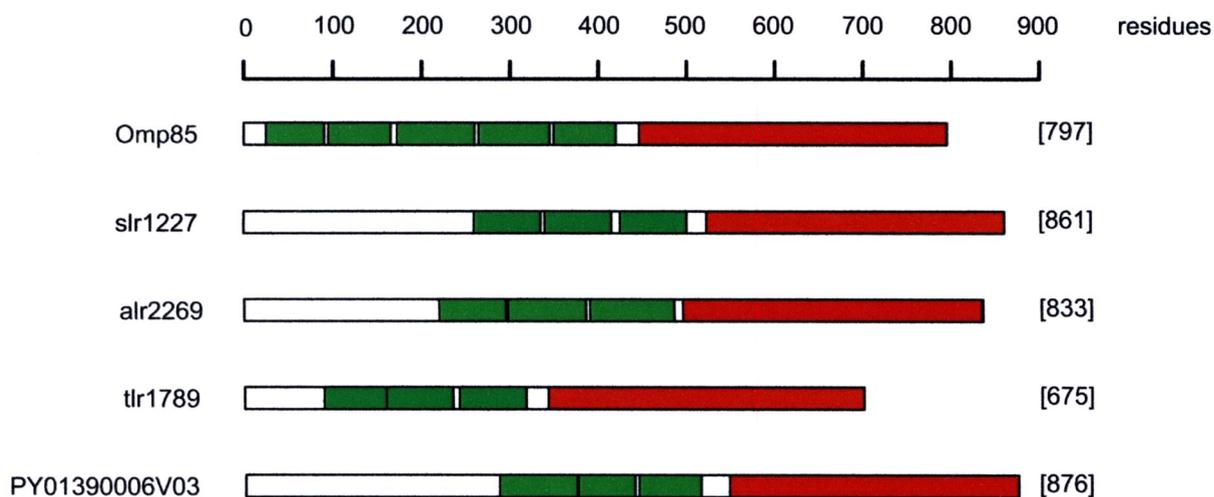


รูปที่ 3.6: แสดงการเปรียบเทียบโครงสร้างของ secretin proteins ระหว่าง *Spirulina* และไซยาโนแบคทีเรียต่างๆ โดยกล่องสีเขียวแสดง STN domain กล่องสีแดงแสดง secretin_N domain และกล่องสีเหลืองแสดง secretin domain

3.5 Chloroplast outer envelope membrane protein (Omp85) หรือ Pilot protein

Pilot protein เป็น outer membrane lipoprotein เป็นโปรตีนที่มีบทบาทสำคัญอย่างมากต่อกระบวนการ oligimerization ของ PilQ เพื่อสร้างไปเป็น pilus ดังนั้นถ้าเซลล์ขาด Omp85 ไป จะส่งผลให้เกิดการสะสมโปรตีน PilQ ในรูปของโมเลกุลเดี่ยวแทน (monomer) และจะไม่สามารถสร้างเป็น pilus ที่สมบูรณ์ได้ [Voulhoux, et al., 2003; Wolfgang, et al., 2000; Burghot, et al., 2004; Collins, et al., 2004]

จากการเปรียบเทียบข้อมูลระดับ amino acid sequence โดยใช้โปรตีน Omp85 จาก *Neisseria meningitidis* Z2491 เป็นต้นแบบในการค้นหาจากจีโนมของ *Spirulina platensis* นั้นพบยีนที่เป็น chloroplast outer envelope membrane protein ที่คาดว่าจะทำหน้าที่เป็น pilot protein คือ โปรตีนที่ได้จากยีน PY01390006V03 ยิ่งกว่านั้นยังพบว่าโปรตีนนี้ประกอบด้วย 2 domains ซึ่งตรงกับบริเวณอนุรักษ์ (conserved regions) ของกลุ่มโปรตีน Omp85 ของไซยาโนแบคทีเรียอีกด้วย ดังแสดงในรูปที่ 3.7 โดยทั้ง 2 domains นี้คือ (1) NH2-terminal α/β periplasmic domain (PF07244) ที่เกี่ยวข้องกับกระบวนการ "secretion" [Voulhoux, et al., 2003] และ (2) COOH-terminal "surface antigen" domain (PF01103) [Loosmore, et al., 1997] โดยบนบริเวณ domain นี้มีโครงสร้างเป็น β -strands ที่จะทำการพับตัว (folding) ไปเป็นโครงสร้างแบบที่เรียกว่า β -barrel ซึ่งจะทำหน้าที่แทรกเข้าไปในผนังเซลล์เพื่อในช่วยการขนส่ง (transport) ผ่านผนังเซลล์เข้ามา [Genevrois, et al., 2003; Voulhoux, et al., 2003]



รูปที่ 3.7: แสดงการเปรียบเทียบโครงสร้างของ pilot proteins ระหว่าง *Spirulina* และไซยาโนแบคทีเรียต่างๆ โดยกล่องสีเขียวแสดง NH2-terminal α/β periplasmic domain ซึ่งแต่ละสปีชีส์จะมีจำนวน domain นี้แตกต่างกันไป และกล่องสีแดงแสดง surface antigen domain.

3.6 Pilus retraction protein (PiIT)

โปรตีน PiIT เป็นโปรตีนที่อยู่ในกลุ่มของ NTPase ซึ่งเกี่ยวข้องกับทั้ง type 2 และ 4 secretion systems รวมถึงยังมีความสำคัญต่อกระบวนการตอบสนองต่อแสง และกระบวนการสร้างแรงเคลื่อนตัว (pilus retraction force) ไปตามพื้นผิวต่างๆ โดยทั่วไปแล้วในจีโนมของไซยาโนแบคทีเรียมักพบ PiIT โปรตีนอยู่ 2 ยีนด้วยกัน คือ PiIT1 และ PiIT2 ซึ่งทั้ง 2 ยีนนี้เป็นยีนที่จำเป็นต่อการทำงานของ type 4 pilus system เนื่องจากว่า ถ้าขาดยีนเหล่านี้แล้วจะส่งผลให้ (1) เซลล์ไม่สามารถเคลื่อนที่ได้ (non-motile) (2) เกิดการสร้าง pilus ออกมาเป็นจำนวนมาก (overproducing pili) เนื่องจาก pilus ที่สร้างมาไม่สามารถทำงานได้ และ (3) มีการแสดงออกของยีนที่เกี่ยวข้องกับการสร้าง pilus มากกว่าสภาวะปกติ [Bhaya, et al., 2000; Okamoto and Ohmori, 2002]

จากข้อมูลจีโนมของ *S. platensis* พบยีนที่คาดว่าจะหน้าที่เป็น Pilus retraction protein (ซึ่งเกี่ยวข้องกับการหดตัว และสลายตัวของโครงสร้าง pilus) หรือ PiIT อยู่ 3 ยีนด้วยกันอาศัยความคล้ายคลึงกันของ amino acid sequence เทียบกับไซยาโนแบคทีเรียชนิดอื่น (based on sequence homology) โดยทั้ง 3 ยีน คือ PY11770002V03 (PiIT1), PY06840007V03 (PiIT2a) และ PY00970023V03 (PiIT2b) ซึ่งทั้ง 3 ยีนนี้ยังพบ 4 consensus patterns ดังแสดงในรูปที่ 3.8 ซึ่งพบได้ในกลุ่มของโปรตีน PiIT ประกอบด้วย

(1) Nucleotide-binding Walker box motif มีลักษณะเป็น GxxGxGK[S/T] พบได้ในโปรตีนกลุ่ม NTPase และ PiIT [Walker, et al., 1982; Friedich, et al., 2002] ซึ่งเป็นบริเวณที่ส่งผลให้เกิดการหดตัวของ pilus [Bhaya, et al., 2000; Nakasuki, et al., 2007]

(2) E signature เกี่ยวข้องกับการทำงานของ type 2 และ 4 systems [Patenge, et al., 2001]

(3) Aspartate-containing sequence motif (Asp box)

(4) Histidine-containing sequence motif (His box)

ซึ่ง Asp และ His boxes มักจะพบในโปรตีนกลุ่ม nucleotide-binding proteins และมีบทบาทสำคัญต่อการทำงานของโปรตีนที่เกี่ยวข้องกับกระบวนการ type 2 และ 4 secretion กระบวนการ type 4 pilus assembly และกระบวนการนำ DNA จากภายนอกเข้าสู่เซลล์อีกด้วย [Planet, et al., 2001; Hardendorf, et al., 2002; Satyshur, et al., 2007; Duggan, et al., 2007]

s1r0161-PiIT1		-----
a112443-PiIT1		-----
t110121-PiIT1		-----
PY11770002V03-PiIT1		-----
s111533-PiIT2	1	MNQPPRPSVPP-----LPPLPNN---PAGRPAASSIRQESMAPETQFMPAP-PSIGQ
a113329-PiIT2	1	MTEQSPSTSHP--VASRNIPVPPP---PLPPPPSLTQRQHTQTLDMSVNRSANQ
t110554-PiIT2	1	MTNPQSSPPPPRPVPHSPMPPPRPMGAAVPPMPQPQMAPPAPPPPPQGMPPRPMPP
PY06840007V03-PiIT2a	1	MTQSRPPAPP-----RPPSPPA---PPRPPAP-----GAQK
PY00970023V03-PiIT2b	1	MAHSTQS-----
s1r0161-PiIT1	1	-----MALEY-----MIEDLMEQLVEMGGSDMHIQAGAPV
a112443-PiIT1	1	-----MEM-----MIEDLMEQMIEMGGSDMHLISAGLPP
t110121-PiIT1	1	-----MEL-----MIEDLMEQVVANGGSDLHISAGIPP
PY11770002V03-PiIT1	1	-----MRGL-----MIEDVLESLVEQGGSDMHIQAGAPI
s111533-PiIT2	48	PQPQHRPPTPNLPTSPAPSHANLGRSPGQ---PSLEELIRSAFDQGFSDVHLGIGEV
a113329-PiIT2	55	PSPPPPSAHRPGTPPISTP---ASDSP---LTLAQIIRKEAFDNGYSDIHLGVGEIP
t110554-PiIT2	61	QPGMPRPPMPQPAMPQVAVATQPMPTPAARISGPTMEQLVREAFEHGYSDIHLGVGEAP
PY06840007V03-PiIT2a	30	SSPPAPAAQASPVAVGITSRGP-GPSPGQ---PTLREIVQRADELGASDVHLGVNETP
PY00970023V03-PiIT2b	8	PSPENLPPKPTVVRARQRQSTN--VCPSN-----SIKDMVKDAFVKKASDHIIRVGHTT

slr0161-PilT1	31	YFRVSGKLEPIN--EEVLTQPESQKLI FMSLNNSQRKELEQNWELDCSYGVKGLARFRIN
all12443-PilT1	29	YFRISGKLTPIG--EEVLTADQCQLIFMSLNNTQRKLTLEQTWELDCSYGVKGLARFRVN
t110121-PilT1	29	YFRISGKLTQPM--YEPLTPEQCQLIFMSLNNTQRKHLEQNWELDCSYGVRGLARFRVN
PY11770002V03-PilT1	30	FYRVSGKLTQPQFGETLSAEVQVLI FQMLNMQRQKLEQNWELDCAYGVKGLARFRVN
sl11533-PilT2	104	RMRDRGEMVPLR--YEPTDQTFMSWLREILTEDEIQRFKRELEFDGAKQYD-FARFRIN
all13329-PilT2	108	RYRNRGEINPTE--YPETDKATFMSWLREIMTEPEIQRFQEHLEFDGATQYE-FARVRIN
t110554-PilT2	121	RFRDRGRLEITN--YPVTDEETFNYWLNELLTPQQIEQFRTHLEYDGAYQYEGLCRVRIN
PY06840007V03-PilT2a	85	RLRYRGDMTPQED-WPVTDLNSFMSWLREVLTDDEIHIQFQENLDFDGAADLG-FVFRIRIS
PY00970023V03-PilT2b	61	RYRIQGE MVR IKD-QVKVTP EIFEQYLS EILAPHQRQQFAEMKELDTAIFYPGIVRCRIN

Walker box

slr0161-PilT1	89	VYKERGCYAAALRALSSKIPNFEQGLPNI VREMAERPRGLILVTGQTSGKTTTLLAAIL
all12443-PilT1	87	VYKERGAYAAALRALSSKIPNFEKGLPVDVREMCDKPRGLILVTGPTSGKTTTLLAAMI
t110121-PilT1	87	VYKDRGTYAAALRALSSKIPNFEQGLPNI VREMSESRPRGLILVTGPTSGKTTTLLAAMI
PY11770002V03-PilT1	90	VYRERGCWAAALRALASQIPNADKIGVPEQLRDLTERPRGMVLVTGQTSGKTTTMAALL
sl11533-PilT2	161	VLESLLGSAMVLRRIIPKILTLTEQRLPDLILRVS DSHKGLILVTGPTSGSKSTTMAAMV
all13329-PilT2	165	VFDSLKGYSMLVLRILPKLISIDQRLPVPVFRDVCVTHKGLILVTGPTSGSKSTTMAAMI
t110554-PilT2	179	IFVALKGPAMVLRILIPVKILTLEQLNLPVFKDLCHYHKLILVTGPTSGSKSTTMAAMV
PY06840007V03-PilT2a	143	IFDSLSGPAMVLRILIGATILTLEQLKLEVFKKICNYHKLMLITGPTSGSKSTTMAAMI
PY00970023V03-PilT2b	120	CFESLTGGAMVLRILINKLIPSIDQLRLEVLKNIISYPHGLILVTGPTSGSKSTTMAAML

Asp Box

Protein E signature

slr0161-PilT1	149	DLINRTRAEHILTIEDPIEFVYFPNVRSLFHQRQKGEDTKSEANALRAALREDPDIILVGE
all12443-PilT1	147	DLINRRTKAEHILTVEDPIEFVYEPKSLVHQRLGEDTKSEANALKAALREDPDIILVGE
t110121-PilT1	147	DLINRTRAEHILTIEDPIEFVYEPKSLIHQRVGEDTKSEANALRAALREDPDIILVGE
PY11770002V03-PilT1	150	DLINRTRSEHILTVEDPIEFVYFPNKSIFHQRQKGEDTKSEANALKAALREDPDIILVGE
sl11533-PilT2	221	DYINKEHAKHILSIEDPIEFVHQSRRLIRQREVGMHTLEFDNALKASLREDPDIILIGE
all13329-PilT2	225	DYINREMPKHILITIEDPVEFVHQSRKSLIKQREVGMHTKFDNALKAALREDPDIILVGE
t110554-PilT2	239	DYINTEMQKHILSIEDPIEFVHQSRRLIRQREVGHILTKFDNALKASLREDPDIILIGE
PY06840007V03-PilT2a	203	DYINKNYAYHILITIEDPVEFVHQSRKSLIKHREVRGHTLKFNALKALRQDPDMMLVGE
PY00970023V03-PilT2b	180	RYLNENARKHIVTIEDPIEFVHASRQCLISQTEVGLHLEFHHALRAVLEDPDVIILIGE

His Box

slr0161-PilT1	209	LRDLETIALAITAAETGHLVFGTLHTNSAAGTIDRMLDVFPAHQQAIRAMLSNSLAVF
all12443-PilT1	207	MRDLETISLAISAAETGHLVFGTLHTSSASQTVDRIIDVFPHEKQTVRVQLSNSLVAVF
t110121-PilT1	207	MRDLETIQLAITAAETGHLVFGTLHTSSAAQTVDRMVDVFPPEQQQIRVQLSNSLVAVF
PY11770002V03-PilT1	210	MRDLETIQLAISAAETGHLVFGTLHTNSASATVDRMLDVFPPICQPQVRAQMSGSLVGV
sl11533-PilT2	281	MRDRSTVNTALKAQTGHLVMGTLHTNSAIKTLERILTYLTAEEOPAMRTAIAESLVAIV
all13329-PilT2	285	MRDKETVNTALKAQTGHLVMGTLHTNSAVKTIERILNLYSGEODAMARIAAESLVAVI
t110554-PilT2	299	MRDRETVNTALKAQTGHLVFGTLHTNSAVKTIERILNLYNPDECGPMRQVAESLVAVI
PY06840007V03-PilT2a	263	IRDKETMQAIAKAASTGHLVAGTLHTNSAIKTLTRVLDMFSAEEOLSIKTALGETLVAII
PY00970023V03-PilT2b	240	MRDRTVNTALQASQTGHLVLTGLHTNRNAINSVNRLNLYNVDEOPAVRIQICDSLVAVI

slr0161-PilT1	269	AQNLVKKKSPKPGFGRALVQBEIMVIIPAIANLIREGKAAQIYSAIQTGAKLGMQTMEOGQ
all12443-PilT1	267	SQTLVPKKNPKPGEYGRVMAQETIMIIIPAINLIREGKTSQIYSAIQTGKGLGMQTMLEKV
t110121-PilT1	267	SQTLVPKKNPKPGEFRIMVQETIMVVIPAINLIREGKTSQIYSAIQTGKGLGMQTMLEKV
PY11770002V03-PilT1	270	SQNLVPRIGG-----GRCAAMEIMLNTIPAMGNLIREGKTSQIYSAIQTMGAKLGMQTMEMS
sl11533-PilT2	341	AQGLCRTTDGK----RAGFHDILINTEVKDYIRQGGYDEIAELMEDGEYDGMITANQS
all13329-PilT2	345	AQGLCRTTDGK----RAAFHDILINTEAIKEWIKDGGYDEISELMKQASFDGMVMTNQS
t110554-PilT2	359	AQALVRTTDGK----RAAIHEIMINTDAIKDYILRGEVEEIEAIPQCTYDGMCTMNNQC
PY06840007V03-PilT2a	323	AQLLCKTTDGK----RAAFHDILINTDVIKEYTMKDQYEEINQIIMLKDTYEGMTMNR
PY00970023V03-PilT2b	300	AQLLIPVTDNK----RAAVHDIILVNTIPAMKDYLLKGDDEESATHLMVNDTYEGMLINQA

slr0161-PilT1	359	LATLVVSGVISLEGLAKSGKPELQRLIGGMTPVAAKRR-----
all12443-PilT1	327	LADYKSGTISFBAAMSKTSKPEIQRLIGPSAPPAGAKAGAVKAH----
t110121-PilT1	327	LADYYRAGIITYBAAMAKSSRQDELQRLIGSGTAAVAR-----
PY11770002V03-PilT1	325	LAKLYQEGKATWBAAMSKSSKPELERLIGPQPKAGKQGVKA-----
sl11533-PilT2	396	LFNLYQEGRITEVALEMAPPVRNEMAMRLRGR-----
all13329-PilT2	400	LLNLYQEGRITEBTALEMSPTPNEMAQFLRGRV-----
t110554-PilT2	414	LYELYEAGRIDEETAIENSPKPNEMAQILRGR-----
PY06840007V03-PilT2a	378	LYELYQEGRITEICLDVSPFPNEMSQMLRGRV-----
PY00970023V03-PilT2b	355	IFEQVTLGRITMEBAEKSSPDAADLDRFRMAGLDSSGVAREFRQKSLF

รูปที่ 3.8: แสดงผลการเปรียบเทียบ amino acid sequence ของโปรตีน PilT ในกลุ่มไซยานแบคทีเรีย โดยแถบเส้นที่บ่งชี้แสดงบริเวณที่เป็น Walker box, Asp box, Protein E signature และ His box ตามลำดับ

3.7 Pilus Biogenesis protein (PilB)

PilB เป็นหนึ่งในโปรตีนกลุ่ม ATPase เช่นเดียวกับ PilT แต่มีหน้าที่การทำงานในกระบวนการ type 4 pili biogenesis ที่ตรงข้ามกัน (antagonistically function) คือ PilT ใช้พลังงานในรูปของ ATP ในการหดและคลาย pilus (retraction) ในขณะที่ PilB จะนำพลังงานในรูปของ ATP เช่นกันมาใช้ในการยืดสายของ pilus (extension by polymerization) [Wall and Kaiser, 1999] ดังนั้นจึงพบ conserved region ที่ทำหน้าที่จับกับ ATP คือ บริเวณ “walker box A” [Turner, et al., 1993] และจากข้อมูลของจีโนม *Spirulina* พบยีนที่สร้างโปรตีน PilB 1 ยีน คือ PY09480047V03 ซึ่งเป็นโปรตีนที่มี amino acid sequence ใกล้เคียงกับโปรตีน PilB ของไซยาโนแบคทีเรียอื่นๆ เช่น *Synechocystis* slr0063, *Anabaena* all2444 และ *Thermosynechococcus* th0122 โดยโปรตีน PilB เหล่านี้ จะมีบริเวณอนุรักษ์อยู่ด้วยกัน 2 บริเวณ (domains) คือ (1) GSPII_E domain (“Type II/IV secretion system protein”; PF00437) และ (2) GSPII_E N-terminal domain (PF05157) ดังแสดงในรูปที่ 3.9 ซึ่งทั้งสองบริเวณนี้มีบทบาทและหน้าที่เกี่ยวข้องโดยตรงกับกระบวนการ Type IV pilus biogenesis และ DNA transfer system

```

16      ...ATPFGNKLIQSGYIDPERLQQAIQDSR-GGKPLTEVLESIT
57      GQPLPPELARHYKKQQLFELKIVYGVEAFDLEIDQISVAELEETKKLETSLETCKRHRF
117     VEWKIIIESEPTPSVMMAMVDPDNLSALDELRNQILRPKGFKLQRRVITQEDYDKLITAYG
177     EEEVKKRQVEAAREEERKQLQFGSEFEDVELVEASEDMG--DMNLDAQLQEAETAPVIKA
235     VNVILARALSEGVSVDIHVEPQEHDMRIRMRKDGVLQEFYFS-FPKKVISAITSRFKIADL
294     DIAERRQAQDGRIRRIYQGRTVDFRVNSLPSRYGEKIVLRILDSSSTQLGLDKLISDQET
354     LEMVRETASRPFGILVLTGPTGSGKSTSLYSVLAERNDEGINISTAEDPIEYALPGITQC
          *****
414     QVIREKDLTFANILRAFLRQDPDVLVGETRDLLETAKTAIEAALTGHLVLTTLHTNDAAG
474     AIARLDEMGEVPEFMVSAALLGIIAQRLMRRVCT...

```

รูปที่ 3.9: แสดงบริเวณที่เป็น domains สำคัญของโปรตีน PilB ที่สร้างจากยีน PY09480047V03 ของ *S. platensis* สายพันธุ์ C1 โดยตำแหน่งของ amino acid ที่เป็นแถบสีดำ และสีเทา แสดงตำแหน่งที่เป็น identical และ similar amino acid ตามลำดับ จากผลการเปรียบเทียบ amino acid sequence จากโปรแกรม ClustalW ด้วยวิธี multiple alignment ส่วนเส้นสีแดง และสีน้ำเงินแสดงบริเวณที่เป็น conserved domain PF00437 และ PF05157 ตามลำดับ ในขณะที่เครื่องหมายดอกจันแสดงตำแหน่งที่เป็น Walker box A (GxxGxGK[S/T])

3.8 The polytopic membrane protein (PilG)

ในไซยาโนแบคทีเรีย โปรตีน PilG มีหน้าที่เป็นตัวควบคุมการตอบสนอง (response regulator) ของเซลล์ที่เกี่ยวข้องกับการเคลื่อนที่ (pilus-dependent motility) และการตอบสนองต่อแสง (phototaxis) โดยโปรตีน PilG จะช่วยให้ pilus ยึดผ่าน cytoplasmic membrane สู่ภายนอกเซลล์ผ่านทางช่องเปิดที่เกิดจากโปรตีน secretin และ pilot protein [Chen and Dubnau, 2004; Yoshihara and Ikeuchi, 2004; Collins, et al., 2007] ในจีโนมของ *S. platensis* สายพันธุ์ C1 พบว่ายีน PY05040001V03 น่าที่จะทำหน้าที่ในการสร้างโปรตีน PilG แม้ว่า amino acid sequence ของโปรตีนนี้ เทียบกับในกลุ่มของไซยาโนแบคทีเรียแล้วจะมีความเหมือนไม่มากนัก ดังแสดงในตารางที่ 3.2 แต่เมื่อทำการวิเคราะห์เปรียบเทียบแบบ multiple alignment ก็พบว่าโปรตีน PilG ของ *Spirulina* กับของไซยาโนแบคทีเรียอื่นๆ บริเวณที่เป็น conserved regions นั้นมีบทบาทและความสำคัญเกี่ยวข้องโดยตรงต่อการทำหน้าที่เป็นโปรตีน PilG ประกอบด้วย ส่วนปลาย N-terminal ยาวประมาณ 170 amino acids และส่วนปลาย C-terminal ยาวประมาณ 100 amino acids ดังแสดงไว้ในรูปที่ 3.10 ซึ่งจากการศึกษาของ Collins และคณะพบว่าบริเวณปลายทั้งด้าน N-terminal และ C-terminal ของโปรตีน PilG นี้จะมีการ folding เป็นรูปร่างคล้าย “fin” (fin-like) และรูปร่างคล้าย “cone” (cone-like) ตามลำดับ ซึ่งโครงสร้างทั้งสองส่วนนี้จะทำหน้าที่เป็นสะพานในการยึดจับกับส่วนของ pilus กับโปรตีนอื่นๆ ที่เกี่ยวข้องของกระบวนการ pilus biogenesis สามารถทำงานได้อย่างมีประสิทธิภาพ [Collins, et al., 2007] ดังนั้นจากข้อมูลที่ได้มาจึงสามารถสรุปได้ว่ายีน PY05040001V03 มีหน้าที่สร้างโปรตีน PilG ซึ่งเป็นหนึ่งในโปรตีนสำคัญสำหรับกระบวนการส่งผ่าน DNA จากภายนอกเข้าสู่ภายในเซลล์ (uptake exogenous DNA into the cell)

ตารางที่ 3.2: แสดงค่าความคล้ายคลึงกันระหว่าง amino acid sequence ของโปรตีนกลุ่ม PilG ระหว่าง *Spirulina platensis* สายพันธุ์ C1, *Synechocystis sp.* สายพันธุ์ PCC 6803, *Thermosynechococcus elongatus* สายพันธุ์ BP-1 และ *Anabaena sp.* สายพันธุ์ PCC7120 โดยค่าที่แสดงในตารางเป็นค่า identity score และ (similarity score) ซึ่งเป็นผลมาจากวิเคราะห์ด้วยโปรแกรม Blastp ของ NCBI

	<i>Synechocystis</i> slr1041	<i>Anabaena</i> all0930	<i>T. elongatus</i> tlr0345
<i>S. platensis</i> C1 PY05040001V03	41% (60%)	45% (63%)	38% (56%)
<i>Synechocystis</i> slr1041	-	40% (59%)	34% (51%)
<i>Anabaena</i> all0930	-	-	32% (54%)

* Percent identity and similarity was collected from BLAST2Sequences from the NCBI.

slr1041	1	Q T N E T L R S T I R I E I L Q F V V Y P H Q S C G R V L L D K V G ----- Y F F
all0930	1	Q S H E T L R S T I L O I E I L Q L I A P S L Y K G N Q N A K E D A E G A D Y F R N Y Q Q Y I I
tlr0345	1	E C N D T L Y T L C T L A L E Y L D D A G R T ----- L L
PY05040001V03	1	Q S N E T L R S I L O I V E V L Q M V A Y S P A A S V V V S E K Y S D R ----- L K S H C I V
slr1041	50	F V D K L L A A D Q D C G N L R Q Y R R H D I A I P L G E M E L E L A L N H I P A C Q Q S K I
all0930	61	F L N Q I I C Q E G D S - N F S I D G R Y Y R V E I G L N E Q Y I T L G K L N S S D Y L L E R N
tlr0345	39	L N H Y L V I A D R H S H S L E Q L Y G Q G I P W P S P D S L I E K A G A S W V E C W L D H C
PY05040001V03	54	C S N R I V A G E S E G - S L K R E Y Y R Y K L T E V L N E L E V P I T A V N A P G Y V V E Q N
slr1041	110	I L T A A H R V L G M V Q V V R Q N V A L G A L E P T S F A T P T H V G Q I A
all0930	120	I I N Q I A E N I Y R L V G T L L H Q E R I C Q D T A I A R T T W E S P V T Q A T Q L O E
tlr0345	99	K L S - - I H S V W R A L L R S V V V N R A W K Y S A P T T O W H N L S L T A L N D S L S H R E
PY05040001V03	113	A I T T G R S L H G M I S T L L H H C A N E M G S P L A Q M T L Q N T V P K I M K V Q E
slr1041	170	K Q Y L R T Q E D P V A I A K D P E L A S K I P A K F A Q H F R S T G A I P E Q I S Y Q K N L P A I A
all0930	180	H E L Y P Y I H P E L P I L A E R V H I N S S L P E A T M N K E Q A G K T T L O L A Y N R D I L T A
tlr0345	157	K K H E L R L D S L Q L T D I K P T G N D P A D Q W F R Q E P F V R E P I T E R L S Q G R A V T G
PY05040001V03	173	T F H H L Q S C P L I L D P D K L R Q V R P N V F E T T R A G K T E F Q I S L N R I V A T
slr1041	230	H G L Y I E S A I E I F P A H E S Q G E N G S A W E T L T D H T I H K V L L D Y A G K Q I L F T N Q -
all0930	240	V I P V V Q L W L K T Y P E P N N C Q I N - - I H I P G N H K K A I L I T K S C E T V S I Q A R E
tlr0345	217	L L L M L H Y L Q V S S W A T N G H T R T - - F P P S S W R Q S P V V V A A T V R Q V E S T Q A A E
PY05040001V03	233	A I T F L Q S F V Q L R F R P T E P S P Y P - - - R S Q F E S K V P I V I D R V R E T E L I T E H E
slr1041	289	----- N P N C V V V L Q D Q M T T L L T Q D L C T M H D L R V H A Q H L A I
all0930	298	Y E T T G Y T I P V I V L T N E E G L I F O R N V F C S S E N E L A M L Q S Q V E Y I
tlr0345	275	----- Y A T A I A H T S L I F Q N E F L A M E D A E F T L H T P R E Y T
PY05040001V03	290	----- Y A T A I G H R K S Q V E Q K D L C A M E N E T A I L S S A E T
slr1041	343	L G K E A Y L R L R M A G D T T T Q R K I S L V E S C K K S -----
all0930	357	M G S S K D K F I Q I R M A G D I P N D D M A V E K L K S S P T K G W E F G N G -----
tlr0345	327	L G S L V G W S R R R K V A G D S S T O T I T H H I G A A P P V T S L T D E Y F G D V L
PY05040001V03	342	L G G I D G F I R K R V R N T T S K R T V E N V G T A D G H S R K L E S Q L A E T F
slr1041		-----
all0930		-----
tlr0345	387	E A E S G L E T S R -----
PY05040001V03	402	S H K W P S L D S E E S Y S V

รูปที่ 3.10: แสดงการเปรียบเทียบ amino acid sequence ระหว่างโปรตีนในกลุ่ม PIIG ของ *S. platensis* สายพันธุ์ C1 เทียบกับไซยาโนแบคทีเรียต่างๆ โดยตำแหน่งของ amino acids ที่เหมือนกันทั้งหมดในไซยาโนแบคทีเรียทั้ง 4 ชนิดจะแสดงเป็นแถบสีดำ ส่วนตำแหน่งที่เหมือนกัน 3 จาก 4 ชนิดของไซยาโนแบคทีเรียจะแสดงด้วยแถบสีเทา ในขณะที่บริเวณ conserved region ที่ส่วนปลาย N-terminal (the fin-like structure) ระบุด้วยอสังหาลีสแดง และ conserved region ที่ส่วนปลาย C-terminal (the cone-like structure) ระบุด้วยอสังหาลีสน้ำเงิน

3.9 Type 4 prepilin peptidase (PiID)

โปรตีน Type 4 prepilin peptidase เป็นโปรตีนในกลุ่มของ integral membrane protein เกี่ยวข้องกับการทำงานหน้าที่ต่างๆในเซลล์ ทั้งกระบวนการสร้าง pilus (type 4 pilus formation) กระบวนการสร้างไบโอฟิล์ม (biofilm formation) กระบวนการส่งถ่ายยีน (gene transfer) และกระบวนการขับสารพิษ หรือเอนไซม์สู่ภายนอกเซลล์ (toxin and other enzyme secretion) โดยในกระบวนการ type 4 pilus system นั้น โปรตีน PiID จะทำหน้าที่ในการช่วยให้โครงสร้างของ pilus ที่ถูกสร้างขึ้นจากกระบวนการ polymerization มีการเชื่อมต่อกันอย่างถูกต้อง โดย PiID จะทำการตัดส่วนที่ไม่ถูกต้องออกไป [LaPointe, and Taylor, 2000; Szabo, et al., 2006] โดยบริเวณที่ PiID จะเข้าทำการตัดนั้นจะเป็นบริเวณที่อยู่ช่วงต้นของสายโปรตีนที่เป็นเป้าหมาย ซึ่งจะมีความยาวประมาณ 10-16 amino acids และมี consensus sequence เป็น **-Gly-∇-Phe-** ซึ่ง ∇ คือตำแหน่งที่จะถูกตัด [Storm, et al., 1993; Yoshihara, et al., 2001] ในไซยาโนแบคทีเรียโปรตีน PiID มีความสำคัญอย่างมากต่อการสร้าง pilus เนื่องจากเมื่อขาดโปรตีนนี้ไป (*piID* mutants) จะทำให้ไม่สามารถสร้าง pilus ที่มีโครงสร้างที่สมบูรณ์ได้เลย (pilus morphotypes are absent) และส่งผลให้เซลล์ไม่สามารถเคลื่อนที่ได้อีกด้วย (non-motile) [Bhaya, et al., 2000] และจากข้อมูลจีโนมของ *S. platensis* สายพันธุ์ C1 นั้นพบว่ายีน PY01350049V03 น่าที่จะทำหน้าที่เป็น type 4 prepilin peptidase โดยอาศัยข้อมูลจากการวิเคราะห์ sequence similarity นอกจากนั้นยังพบว่าโปรตีน PY01350049V03 ประกอบด้วย 2 domain คือ (1) bacterial peptidase A24 N-terminal domain (PF06750) และ (2) domain of type IV leader peptidase family (PF01478) ซึ่งเกี่ยวข้องโดยตรงกับการทำงานเป็นเอนไซม์ peptidase โดยมักพบ domain นี้ที่ปลาย N-terminal ของ prepilin peptidase [Storm, et al., 1993; LaPointe, and Taylor, 2000]



รูปที่ 3.11: แสดงผลการค้นหา domain ของโปรตีน PY01350049V03 ของ *S. platensis* สายพันธุ์ C1 โดยตำแหน่งของ PF06750 domain และ PF01478 domain แสดงด้วยรูปสี่เหลี่ยมสีเทาและขาวตามลำดับ นอกจากนั้นยังแสดงผลการวิเคราะห์เปรียบเทียบแบบ multiple alignment ด้วยโปรแกรม ClustalWระหว่าง PY01350049V03 กับโปรตีน PiID ของไซยาโนแบคทีเรียต่างๆ ประกอบด้วย *Synechocystis* sp., *Anabaena* sp. และ *T. elongatus* BP-1 โดยแถบสีดำและสีเทาแสดงตำแหน่งของ amino acids ที่เป็นตำแหน่ง identical และ similar กันตามลำดับ

3.11 ผลการเปรียบเทียบยีนที่เกี่ยวข้องกับกระบวนการส่งถ่ายยีนของไซยาโนแบคทีเรียชนิดต่าง ๆ

นอกเหนือจากการค้นหายีนที่เกี่ยวข้องกับกระบวนการส่งถ่ายยีนจากภายนอกเข้าสู่ภายในเซลล์ (DNA uptake system) แล้ว ได้ทำการวิเคราะห์เพิ่มเติม เพื่อเป็นการช่วยยืนยันอีกระดับหนึ่ง ว่ากลุ่มของยีนที่ค้นพบนั้นมีความเกี่ยวข้องโดยตรงกับความสามารถในการรับ DNA จากภายนอกเข้าสู่เซลล์ จึงได้ทำการวิเคราะห์เปรียบเทียบข้อมูลของกลุ่มยีนที่เกี่ยวข้องกับกระบวนการ DNA uptake system ระหว่าง *S. platensis* สายพันธุ์ C1 กับข้อมูลจีโนมของไซยาโนแบคทีเรียอื่น ๆ อีก 11 จีโนม ซึ่งประกอบไปด้วย *Synechocystis* sp. PCC 6803, *Anabaena* sp. PCC7120, *Thermosynechococcus elongatus* BP-1, *Synechococcus elongatus* PCC 6301, *Gleobacter* PCC 7421, *Microcystis* NIES-843, *Prochlorococcus* SS120, *Prochlorococcus* MED4, *Prochlorococcus* MIT9313, *Synechococcus* WH8102 และ *Synechococcus* CC9311

จากผลการวิเคราะห์เปรียบเทียบพบว่ามีไซยาโนแบคทีเรียเพียง 5 สายพันธุ์เท่านั้นที่มียีนที่เกี่ยวข้องกับกระบวนการ DNA uptake system อยู่ครบ (ผลแสดงไว้ในตารางที่ 3.3) ประกอบด้วย *Synechocystis* sp. PCC 6803, *Anabaena* sp. PCC7120, *Thermosynechococcus elongatus* BP-1, *Synechococcus elongatus* PCC 6301 และ *S. platensis* สายพันธุ์ C1 โดย 4 จาก 5 สายพันธุ์นี้ (ยกเว้น *S. platensis* สายพันธุ์ C1) ได้มีการวิจัยยืนยันแล้วว่าสามารถรับ DNA จากภายนอกเข้าสู่ภายในเซลล์ได้ผ่านทางกระบวนการ natural transformation system [Cyanobase] ดังนั้น *S. platensis* สายพันธุ์ C1 จึงน่าที่จะสามารถรับ DNA จากภายนอกเข้าสู่ภายในเซลล์ได้ผ่านทางกระบวนการ natural transformation system เช่นเดียวกัน

ตารางที่ 3.3: ยีนที่เกี่ยวข้องกับกระบวนการ DNA uptake system ในไซยาโนแบคทีเรียสายพันธุ์ต่าง ๆ

Symbol	Gene Description	<i>Spirulina platensis</i> C1	<i>synechocystis</i> PCC 6803	<i>Anabaena</i> PCC 7120	<i>Thermosyn</i> BP-1	<i>synechococcus</i> PCC 6301	<i>Gleobacter</i> PCC 7421	<i>Microcystis</i> NIES-843	<i>Prochlorococcus</i> SS120	<i>Prochlorococcus</i> MED4	<i>prochlorococcus</i> MIT9313	<i>synechococcus</i> WH8102	<i>synechococcus</i> CC9311
comE A	DNA receptor	1	1	1	1	1	1	1	0	0	0	1	1
comE C	Membrane channel protein	1	1	1	1	1	1	1	0	0	1	1	1
comF A	ATP-binding protein	1	1	1	1	1	1	1	0	0	1	1	1
pilA I	Prepilin protein	1	11	1	4	3	3	3	0	0	1	0	3
PilB	Pilus biogenesis protein	1	2	1	1	1	2	1	0	0	1	1	1
pilD	Prepilin peptidase	1	1	1	1	1	1	0	0	0	1	1	1
pilG	Polytopic membrane protein	1	1	1	1	1	0	0	0	0	0	0	0
pilQ	Secretin	1	1	1	1	1	1	1	0	0	1	1	1
pilT	Pilus retraction protein	3	2	2	2	2	1	2	0	0	2	0	2
pilP/omp85	Pilot protein	1	1	2	1	1	1	1	1	1	1	1	1