



Title	ナイロンオリゴマー分解性プラスミド p0AD2 の解析
Author(s)	加藤, 晃
Citation	大阪大学, 1994, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3094152
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

ナイロンオリゴマー分解性プラスミド pOAD 2 の解析

1994 年

加 藤 晃

目次

緒論	1
第一章 p0AD2 プラスミド全塩基配列の決定	7
第一節 緒言	
第二節 実験材料および方法	
第三節 結果	
第四節 考察	
第五節 要約	
第二章 Repeat-sequence (RS-I と RS-II) 領域の解析	24
第一節 緒言	
第二節 実験材料および方法	
第三節 結果	
第四節 考察	
第五節 要約	
第三章 p0AD2 上に存在する non-stop frame (NSF) の解析	55
第一節 緒言	
第二節 実験材料および方法	
第三節 結果	
第四節 考察	
第五節 要約	
総括	70
参考文献	73
謝辞	78

生物が生命現象を営む上で、遺伝子産物であるタンパク質や酵素は、非常に重要な物質であり、その数や種類は非常に多い。例えば、大腸菌で約 4000、酵母で約 7000、線虫で約 15000、人に至っては 5万-10万のタンパク質や酵素に関する遺伝子が存在すると言われている。世の中に存在しているタンパク質すべてについて相同性検索を行い、グループに分け、系統樹を作成したとする (Fig. 1)。その結果、その中には分岐の多いグループもあれば、分岐の少ないグループもあるはずである。言い換えると、相同性を持つもの、進化的に関係のあるものが多いグループもあれば、少ないグループもある。

アミノ酸が他のアミノ酸に置換する速度は、遺伝子座や分子の場所で違うらしいが、すべてのタンパク質でその速度が一定であったとすると、現在グループ内で多くの変化を持つタンパク質、すなわち、分岐が多く、系統樹上で進化距離が長いタンパク質のグループほど古い過去に生まれたタンパク質で、分岐が少なく、進化距離が短いタンパク質のグループほど、新しく生まれたタンパク質となる。また、現在分岐も少なく、進化距離が短いタンパク質のグループも未来においては変異が蓄積され、分岐が増え、多くの変化を持った集団になると考えられる (Fig. 1)。逆に言うと、現在分岐が多く、進化距離の長いタンパク質のグループも過去においては、分岐が少なく、進化距離の短いものどうしのグループであったはずである。もっと遡ると他のタンパク質とは系統樹がかけない1つのタンパク質、すなわちそのグループ内の先祖型タンパク質に行き着くはずである。そのタンパク質が出現した時点では、その他のタンパク質と相同性を持たないタンパク質である。1つのグループに属するタンパク質や酵素は、ある時生まれ、変異を蓄積し、現在のグループ内での分布になっていると考えられる。

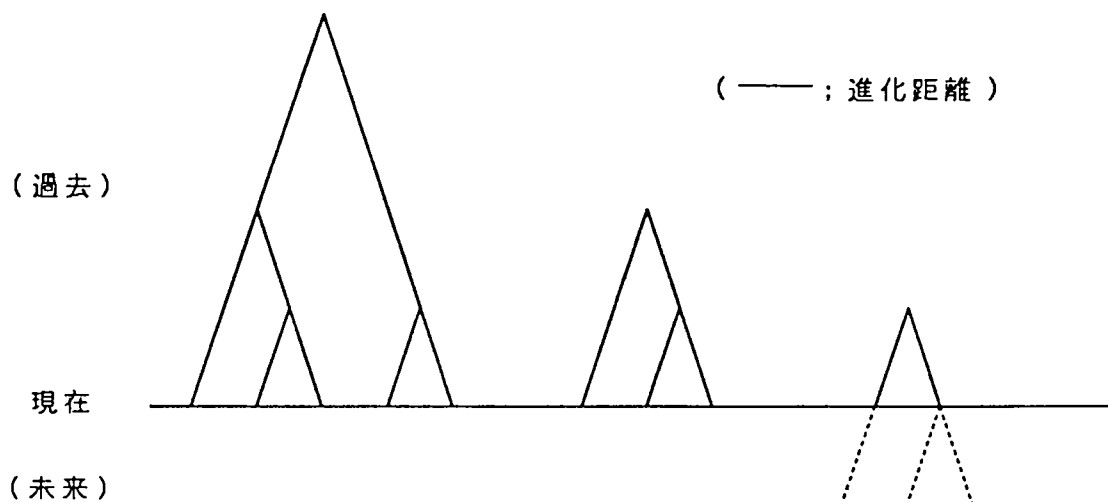


Fig. 1 Phylogenetic tree of enzymes.

先祖型タンパク質（新しいタンパク質）を、他のタンパク質と進化的に関係のないタンパク質、具体的には相同性の無いものと定義する。では、新しいタンパク質はどうやって出現したのだろうか。それは、分子進化学上の興味のみならず、ある生物種、例えば微生物種的能力を人為的に改変し、発展させる上でも重要である。自然界には、特殊環境下において特異な能力を持った微生物が数多く存在する。抗生物質耐性能、難分解性物質（非天然型化合物）資化性能などがその一例である。微生物がどのようにして新しい能力を獲得したのかということも興味深い。

Flavobacterium sp. KI72 株も、非天然型化合物であるナイロンオリゴマーを分解し、単一の炭素源ならに窒素源として資化する¹⁾。このナイロンオリゴマーは、6-ナイロンが ϵ -カプロラクタムの開環重合によって合成される過程の副産物として生成し、オリゴマーの段階で以後の重合が停止した直鎖状オリゴマーや、分子末端が環化した環状オリゴマーの混合物である。*Flavobacterium* sp. KI72 株は、ナイロンオリゴマーに対して Fig. 2 に模式的に示すように、作用様式の異なる3種類の酵素、6-aminohexanoate-cyclic-dimer hydrolase (EI), exotype 6-aminohexanoate-oligomer hydrolase (EII), および endotype 6-aminohexanoate-oligomer hydrolase (EIII) を持っている (Table 1)。EI は、6-aminohexanoate (Ahx) の環状2量体 (6-aminohexanoate linear dimer; Ald) のみに作用するが、重合度3以上の環状あるいは直鎖状オリゴマーには活性を示さない²⁾。EII は、2量体以上の直鎖状オリゴマーをアミノ末端側からエキソ型で1分子ずつ分解するが、環状オリゴマーには作用しない³⁾。一方、EIII は、環状ならびに直鎖状オリゴマー (3量体以上) をエンド型で分解する⁴⁾。いずれの酵素活性も検討した約 60種類の天然型ペプチドには活性を示さず、ナイロンオリゴマーに特異性の高い酵素である。また、EI, EII, および EIII 各酵素間には、免疫学的類縁性は認められず、アミノ酸配列においても相同性は認められない。この EII と EIII 酵素は、相同性を示す酵素をデータベース (Gen Bank ver. 71) 上で検出することができない。EI 酵素は *Pseudomonas savastanoi* 由来の indoleacetamide hydrolase (IAAH)⁵⁾ とアミノ末端側半分の 237アミノ酸残基領域で 33% の相同性を示すが、カルボキシル末端側では相同性は認められない。EI 酵素には、微弱ながらインドールアセトアミドの加水分解活性が検出されることから、EI 酵素ならびに IAAH 酵素のアミノ末端部分は、共通の祖先に由来すると考えられる。

KI72 株の EI, EII, および EIII 3酵素 (それぞれ F-EI, F-EII, および F-EIII) をコードする遺伝子 (それぞれ F-*nylA*, F-*nylB*, および F-*nylC*) は、本菌株が保有する3種類のプラスミド (pOAD1, pOAD2, および pOAD3) の中で、pOAD2 上にコードされており^{4, 6)}、同じプラスミド上に、F-EII 類似酵素 F-EII' をコードする領域 (F-*nylB'*) もある⁷⁾。F-EII と F-EII' とは互いによく似ているが (アミノ酸配列で 88% の相同性)、F-EII' の 6-aminohexanoate dimer (Ald) に対する活性は、F-EII の 1% 以下である。*Pseudomonas* sp. NK87 株もナイロンオリゴマーを資化する細菌であり、EI と EII 酵素に対応する酵素 (それぞれ P-EI と P-EII) を持っている⁸⁾。この NK87 株は、6種類のプラスミドを保持しており、P-EI 遺伝子 (P-*nylA*) は pNAD2 上に、P-EII 遺伝子 (P-*nylB*) は pNAD6 上にコードされている⁸⁾。F-EII と P-EII の相同性は、アミ

substrate		Specific activity (mol NH ₂ -/mol·s)		
		EI	EII	EIII
Ahx-linear oligomers	AL2	<1.0×10 ⁻³	0.66	2.7×10 ⁻⁴
	AL3	<1.0×10 ⁻³	0.53	0.067
	AL4	<1.0×10 ⁻³	0.40	0.26
	AL5	<1.0×10 ⁻³	0.17	0.29
Ahx-cyclic oligomers	AC1	<1.0×10 ⁻³	<1.0×10 ⁻⁴	<1.0×10 ⁻⁴
	AC2	2.4	<1.0×10 ⁻⁴	<1.0×10 ⁻⁴
	AC4	<1.0×10 ⁻³	<1.0×10 ⁻⁴	0.22

Table 1 Specific activity of F-EI, F-EII, and EIII enzymes toward various nylon oligomers. Increase of amino group were analyzed by TNBS method. Specific activities were expressed as turn over number per subunit. Ahx: 6-aminohexanoate, AL2: Ahx linear dimer, AL3: Ahx linear trimer, AL4: Ahx linear tetramer, AL5: Ahx linear pentamer, AC1: Ahx cyclic dimer, AC2: Ahx cyclic tetramer, AC4: Ahx cyclic octamer

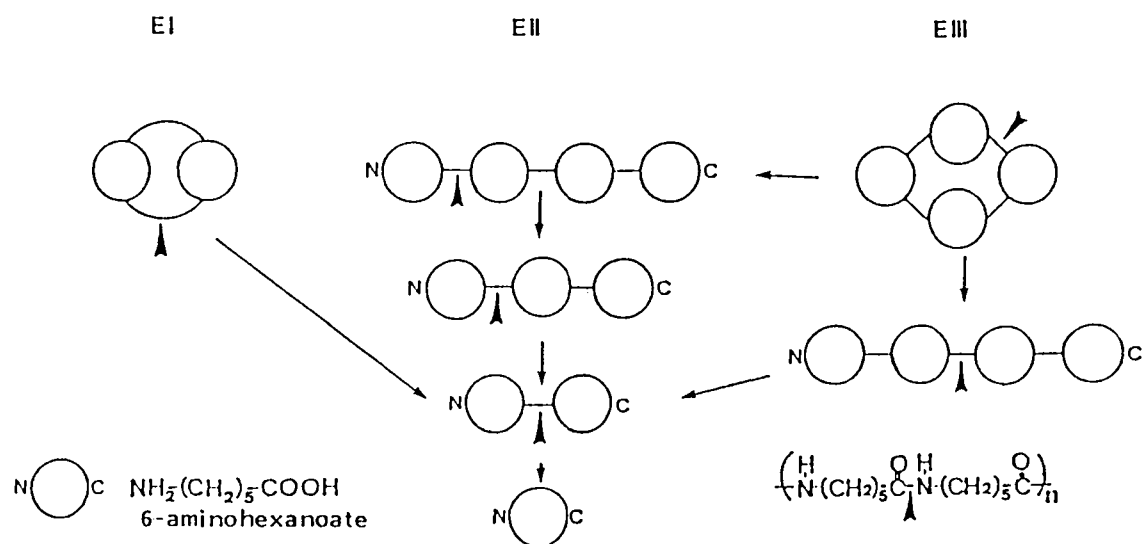


Fig. 2 Model of degradation of nylon oligomers by 6-aminohexanoate-cyclic-dimer hydrolase (EI), exotype 6-aminohexanoate-oligomer hydrolase (EII) and endotype 6-aminohexanoate-oligomer hydrolase (EIII). Arrowheads indicate the hydrolysis position by these enzymes.

ノ酸配列で 37%、F-EII' と P-EII のそれは 35% であり、これらの遺伝子は、同じ先祖型遺伝子から進化したと考えられる。F-nylB と P-nylB の進化距離を Kimura の方法⁹⁾により、単位時間あたり、アミノ酸座位あたりの置換率を 9×10^{-9} として計算すると、少なくとも 1.4×10^8 年前に2つの遺伝子は、共通の祖先から分かれたことになる¹⁰⁾。

興味深いことに、F-nylB では、対応する相補鎖上に終止コドンが存在しない領域 (non-stop frame; 以後 NSF と略す) が存在する¹⁰⁾ (Fig. 3)。その他の EII 遺伝子 (F-nylB' と P-nylB) や EIII 遺伝子 (F-nylC) についても、対応する相補鎖に NSF が存在している¹⁰⁾ (Fig. 3)。これら遺伝子の相補鎖に NSF が存在する確率を、各遺伝子の塩基配列から相補鎖に現れる終止コドン数の期待値を求め、この値から計算すると、0.0001-0.0080 と非常に小さい。もともと nylB 遺伝子ファミリーの先祖型遺伝子が、その相補鎖に NSF を持っていたとしても、分岐後アミノ酸残基置換を蓄積する間、その NSF に終止コドンが出現せず、各 nylB 遺伝子に NSF が受け継がれる確率は、わずか 0.007 でしかない。この数値は進化過程において、アンチセンス鎖上の NSF が維持されていることを示唆している。また、NSF はいずれもアンチセンス鎖上で、センス鎖と同じ読み枠 (frame 0) のみに存在しており、他の読み枠 (frame+1 と frame-1) では期待される頻度で終止コドンが出現している¹⁰⁾ (Fig. 3)。そして、NSF のアミノ酸組成およびコドン使用頻度は、センス鎖の酵素のアミノ酸組成およびコドン使用頻度と似ている。Gen Bank データベース上に登録されている細菌の遺伝子について、センス鎖上の読み取り枠 (open reading frame; 以後 ORF と略す) を長さにより分け、それぞれのアンチセンス鎖上に NSF が存在する割合を調べた (Fig. 4)。アンチセンス鎖の3つの読み枠 (CA-F1, CA-F2, および CA-F3) と終止コドンの使用頻度より予想される値 (Expect) とを比較すると、CA-F2 と CA-F3 は予想値と同様の挙動を示すのに対し、センス鎖の ORF に対応するアンチセンス鎖の読み枠である CA-F1 は、CA-F2 とか CA-F3 に対して、また予想されるよりも多くの、そして長い NSF が存在することが明らかになった。このことは、センス鎖の ORF に対応するアンチセンス鎖において、NSF を維持する機構が存在することを示唆している。

酵素が創造されるには、既存の遺伝子が転移し、配列が変わり、更に突然変異等によって、新たな活性を持った酵素となることも考えられるが、これでは、全く新しいアミノ酸配列を持った酵素となるには、長い年月が必要であり、また、他のタンパク質とも進化的な関係を持つこととなる。遺伝子として使われていない配列から遺伝子が生まれるならば、その遺伝子がコードするタンパク質は、全く新しい配列を持ったタンパク質となる。新しい遺伝子が誕生するためには、読み取り枠が開放されていることが重要である。各 nylB (F-nylB, F-nylB', および P-nylB) が共通して NSF を持っていることから、長い年月の間、終止コドンが生まれず開放されたままになっているナイロンオリゴマー分解酵素遺伝子の相補鎖に存在する NSF は、新規遺伝子が発生する1つの場所ではないかと考えられる。nylB, nylB', および nylC 遺伝子の相補鎖に存在する NSF 由来のアミノ酸配列は、コドン使用頻度が、調べた範囲ではセンス鎖由来のアミノ酸と大きくずれていないことから、もしこの NSF がタンパク質を産生したならば、タンパク質として機能する

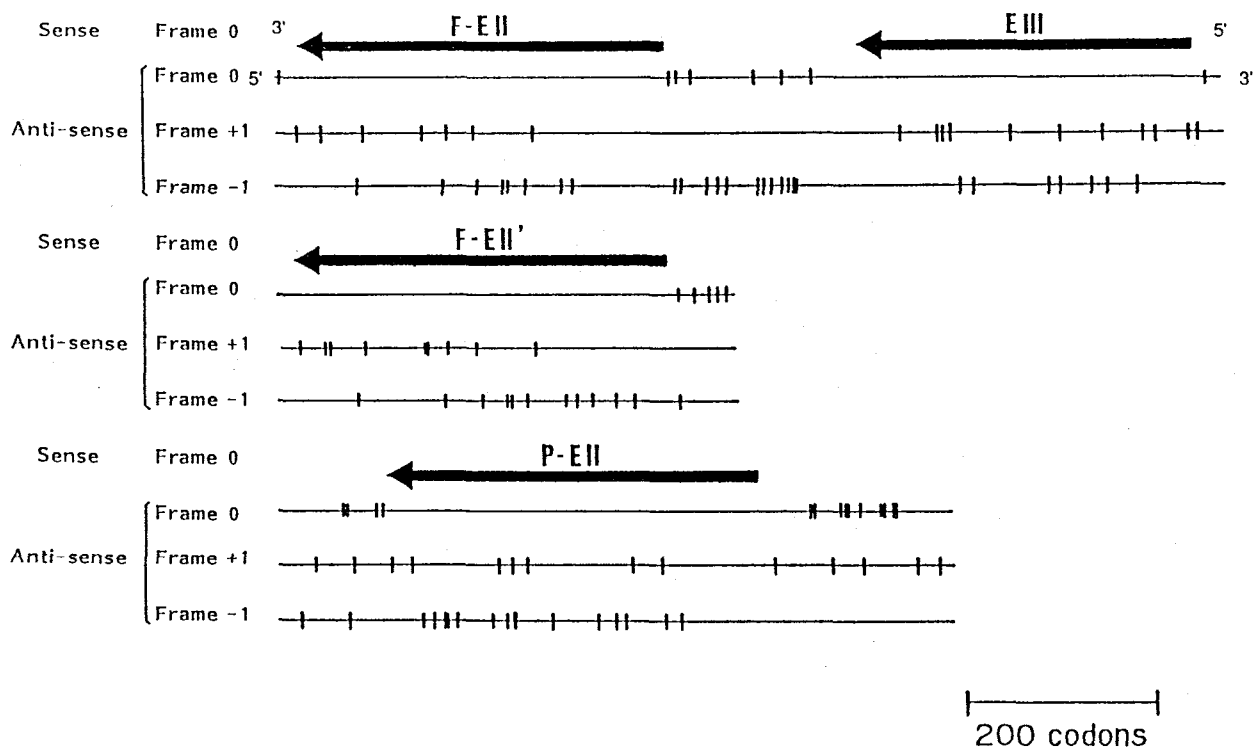


Fig. 3 Distribution of the stop codons on the antisense strands of the *nylB* gene family. The three horizontal lines (from 5' to 3') for each antisense strand of the F-*nylB* and F-*nylC*, F-*nylB'*, or P-*nylB* including flanking regions show the three kinds of coding frames of each sequence: the top lines correspond to the native coding frames of the *nylB* family and the *nylC*, and the middle and bottom lines correspond to frames shifted 1 base forward and backward, respectively, from the top frame. The vertical lines on the these horizontal lines show the position of stop codons. Arrows show the position of the ORFs on the sense strands for F-EII and EIII, F-EII', and P-EII.

可能性が高く、また、これら NSF 由来のアミノ酸配列は、データベース上にある遺伝子由来のアミノ酸配列と相同性がなく、現在までに配列のわかっているタンパク質と進化的に関係のないタンパク質である。

薬剤耐性、難分解性物質資化性といった能力をつかさどる遺伝子は、染色体外因子であるプラスミドが支配する場合が多い。非天然型化合物であるナイロンオリゴマーを単一炭素源および窒素源として生育可能な *Flavobacterium* sp. KI72 株もナイロンオリゴマーの分解に関与する 3 種類の酵素をコードする遺伝子を持っているが、これら遺伝子も本菌株の保持する p0AD2 プラスミドが担っている。微生物が特異な能力を新たに獲得するには、プラスミドが重要な役割を果たしていると考え、特異な基質であるナイロンオリゴマーに作用し、また、遺伝子の相補鎖に NSF が存在する *nylB* および *nylC* を担うプラスミド p0AD2 の理解を深めるために、全塩基配列の決定を行った。

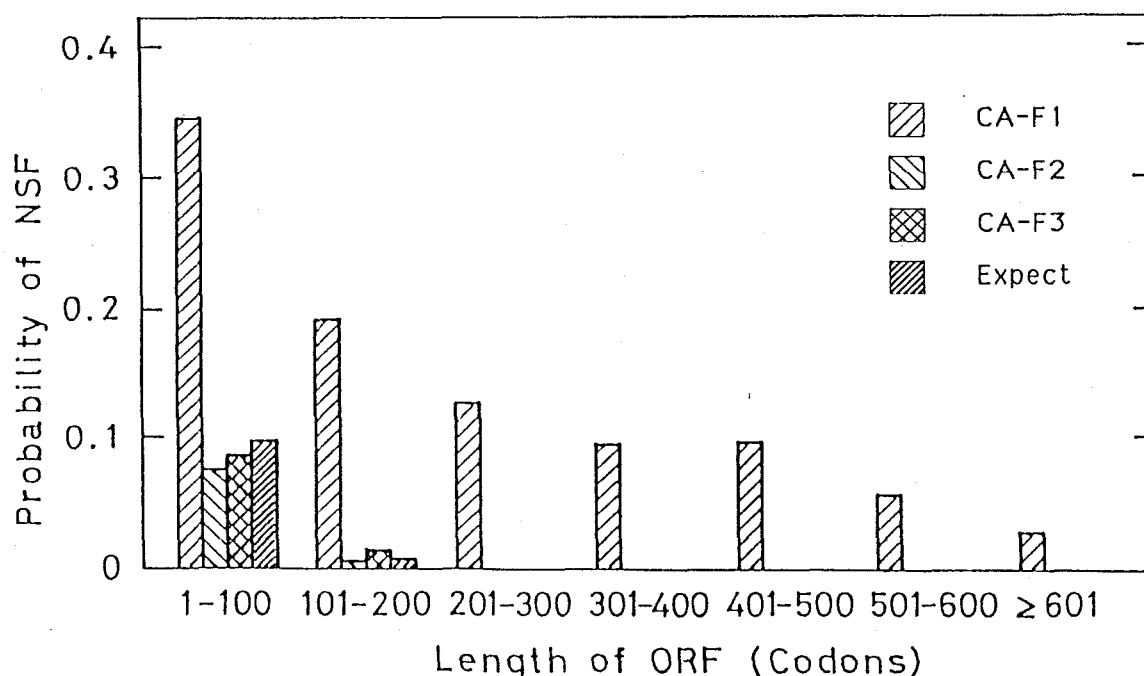


Fig. 4 Probability of NSF on the antisense strand of the ORF. CA-F1 correspond to the native coding frames of ORFs, CA-F2 and CA-F3 correspond to frames shifted 1 base forward and backward, respectively, from the CA-F1. Expect shows the calculated value from codon usage.

第一章 pOAD2 プラスミド全塩基配列の決定

第一節 緒言

Flavobacterium sp. KI72 株は、ナイロンオリゴマーを単一の炭素源および窒素源として生育可能な細菌である。ナイロンオリゴマーの分解に関与する3種類の酵素 EI, EII, および EIII のうち、EII と EIII 両酵素のアミノ酸配列は、データベースに登録されているいかなるタンパク質とも相同性を示さず、また、遺伝子の対応する相補鎖には終止コドンが存在しない (NSF がある)。このことは非常に希にしか起こらないことである。これらの酵素をコードする遺伝子は、本菌株の保持する3種類のプラスミドのうち、pOAD2 上に存在する。この pOAD2 プラスミドが、どのような構造を持っているかを調べる目的で、本プラスミドの全塩基配列決定を行った。

第二節 実験材料および方法

1 使用菌株、プラスミドおよびファージ

Flavobacterium sp. KI723T1 株⁶⁾を pOAD2 プラスミドの調製用菌株として用いた。形質転換の宿主としては、*Escherichia coli* JM109 株¹¹⁾ {*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ (*lac-proAB*) F' [*traD36 proAB⁺ lacI^a lacZ* Δ 15]}}, XL1-Blue¹²⁾ {*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi lac⁻* F' [*proAB⁺ lacI^a lacZ* Δ 15 Tn10(*tet^r*)]}を使用した。

プラスミドベクターとして、pBR322¹³⁾、pBluescript II SK^{14,15)} および pUC18, 19¹⁶⁾、また、塩基配列の決定に M13 mp18, 19¹⁷⁾を用いた。

2 培地

E. coli 用の培地として、LB 培地 (1% Tryptone, 0.5% yeast extract, 0.5% NaCl), 2 x TY 培地 (1.6% Tryptone, 1% yeast extract, 0.5% NaCl), H 培地 (1% Tryptone, 0.5% NaCl) を用い^{11,18)}、必要に応じて LB 培地と 2 x TY 培地にはアンピシリンを 50 μ g/ml となるように加えた。この他に寒天平板培地には、必要に応じて Isopropylthio- β -D-galactoside (IPTG) および 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-gal) をそれぞれ 0.1 mM, 0.005% となるように加えた。*Flavobacterium* sp. KI723T1 株の培地としては、環状ナイロンオリゴマー培地 (環状ナイロンオリゴマー 0.8%, KH₂PO₄ 0.3%, K₂HPO₄ 0.1%, yeast extract 0.05%, NaCl 0.5% MgSO₄ 1 mM, FeCl₃ 20 μ M) を用いた⁶⁾。

3 *Flavobacterium* sp. KI723T1 株からのプラスミド DNA の調製

KI723T1 株からの pOAD2 調製は、Negoro らの方法⁶⁾を改変して行った。KI723T1 株が保持するその他のプラスミド (pOAD11 と pOAD3) 上には、制限酵素 *Xba*I の認識部位が存在するが、

pOAD2 上には存在しない。このことを利用し、粗プラスミド溶液を XbaI で処理した後、CsCl-エチジウムブロミド平衡密度勾配遠心分離法により精製を行った。XbaI 処理により pOAD2 のみが、covalently-closed-form として分離できる。

4 大腸菌からのプラスミド DNA の調製および大腸菌の形質転換

大腸菌からのプラスミド DNA の調製は、Birnboim と Doly のアルカリ-SDS 変性法¹⁹⁾に従い、CsCl-エチジウムブロミド平衡密度勾配遠心法で精製を行った。少量の精製プラスミド DNA を調製する場合には、20 ml の 2 x TY 培地 (50 μ g/ml のアンピシリンを含む) に一白金耳目的プラスミドを保持する大腸菌を植菌し、37°C で一晩振とう培養を行った。遠心分離によって 2 本のサンプルチューブに集菌した。これを 100 μ l の溶液 I (25 mM Tris-HCl pH 8.0, 50 mM グルコース, 10 mM EDTA) に懸濁後、200 μ l の溶液 II (0.2 N NaOH, 1% SDS) を加え、穏やかに攪はんした。次いで、150 μ l の溶液 III (3 M 酢酸カリウム pH 4.8) を加え、よく攪はんしたのち、フェノール/クロロホルム抽出を行い、得られた上澄液から DNA をエタノール沈澱により回収し、粗プラスミド標品を得た。これを 50 μ l の TE 緩衝液 (10 mM Tris-HCl pH 7.5, 1 mM EDTA) に懸濁し、終濃度 10 μ g/ml になるように RNase A を加えて 37°C 30分間放置した。30 μ l の PEG 溶液 [20% polyethylene glycol-6000 (PEG6000), 2.5 M NaCl] を加え、氷上で 1時間放置した。遠心分離後、沈澱物を 70% エタノールで洗浄したのち乾固させることにより精製プラスミド標品を得た。

少量のプラスミド DNA を簡便に調製する場合には、上記の粗プラスミド標品を用いた他、He らの方法²⁰⁾を用いた。サンプルチューブに集菌後、100 μ l の TELT 溶液 [2.5 M LiCl, 50 mM Tris-HCl pH 8.0, 62.5 mM EDTA, 4% (v/v) Triton X-100] に懸濁し、等量のフェノール/クロロホルムを加え、15秒間激しく攪はんした。遠心分離により得られた上澄液から DNA をエタノールにより沈澱させ、その沈澱物を乾固することによりプラスミド標品を得た。

プラスミドや M13 ファージによる大腸菌の形質転換は、Cohen らの方法²¹⁾に従い、15% グリセロールを含む 50 mM CaCl₂ 溶液を用いて調製したコンピテント細胞を用いた。

5 ファージ一本鎖 DNA の調製

5 ml の 2 x TY 培地に宿主大腸菌の一晩培養液を 1% およびファージプラークを植菌し、37°C 5~7時間激しく振とう培養した。培養液をサンプルチューブに移し、遠心分離を 2 回を行い、菌体を完全に除去し、上澄液 1 ml に対し、PEG 溶液 (20% PEG 6000, 2.5 M NaCl) を 200 μ l 加え、室温で 15分間放置した。遠心分離を 2 回を行い、上澄液を完全に除去した後、150 μ l の TE 緩衝液に懸濁し、フェノール抽出を 2 回、フェノール/クロロホルム抽出を 1 回を行い、エタノール沈澱後、沈澱物を乾固させ、ファージ一本鎖 DNA 標品とした。

6 アガロースゲル・ポリアクリルアミドゲル電気泳動、およびゲルからの DNA の回収

制限酵素処理した DNA 断片を 0.8~1.5% アガロースゲル電気泳動、あるいは 4~6% のポリアクリルアミドゲル電気泳動により分析を行った。電気泳動用緩衝液は、アガロースゲルの場合、40 mM トリス-酢酸緩衝液 pH 8.0, 2 mM EDTA、ポリアクリルアミドの場合 50 mM トリス-ホウ酸緩衝液 pH 8.3, 1 mM EDTA をそれぞれ使用した。泳動後、エチジウムブロミド溶液 (1 μ g/ml) にゲルを浸し、DNA を染色したのち、UV 照射により DNA の検出を行った。アガロースゲルからの DNA 断片の回収は Vogelstein と Gillespie の方法²²⁾を用いた。一連の操作は、Bio101 社の Gene clean キットを用い、その使用書に従った。ポリアクリルアミドゲルからの DNA 断片の回収は、Maxam と Gilbert の方法²³⁾に従った。

7 欠失変異体の作製

ExoIII と Mung Been ヌクレアーゼを用いた一連の欠失変異体の作製方法は、Stratagene Cloning System 社の方法に従った。

8 塩基配列の決定

DNA 塩基配列の決定は、一本鎖、または、二本鎖 DNA を鋳型として、ジデオキシシーケンス法^{11, 17)}で行った。一連の操作は、USB 社の Sequenase ver. 2.0 キットの使用書に従った。

9 その他の実験材料および方法

制限酵素および修飾酵素などは、宝酒造製あるいは東洋紡績製のものを主に使用し、実験操作については、各々の使用書などに従った。

第三節 結果

1 pOAD2 全塩基配列決定のための欠失変異体作製

pOAD2 は、HindIII 完全分解により 6 つの断片 (A, B, C, D, E, および F 断片) にわかれるが (Fig. 1-1)、その HindIII 部分分解物を pBR322 に組み込んだハイブリッドプラスミド (pNDH5, pNDH6, pNDH10, および pNDH16)²⁴⁾ が得られており (Fig. 1-2)、これらを pOAD2 全塩基配列決定のための試料とした。各 HindIII 断片を pBluescript II SK⁺ に組み込み、制限酵素地図を作製した。次に、制限酵素地図に従い、500-2000 bp 程度の断片を pBluescript II SK⁺ に組み込んだ。再度正確な制限酵素地図を作製し、ExoIII と Mung Been ヌクレアーゼを用いた欠失変異体が取得可能な制限酵素部位の組み合わせで、M13 mp18 または M13 mp19 に組み込み、欠失変異体を作製した。最終的には約 100 bp づつ縮小化された欠失変異体を作製した。その一例を C 断片について Fig. 1-3A, 3B, および 3C に示す。

2 塩基配列の決定

各小断片の塩基配列は、M13 フェージ一本鎖 DNA を用いたジデオキシ法により、両鎖とも決定した。その一例を C 断片の HindIII-PstI 1.61 kb について、Fig. 1-3E に示す。小断片間の連結部位については、連結部位を含む領域を適当な制限酵素を用いてサブクローニングを行い、少なくとも一方向の塩基配列を決定し、小断片間の連結部位に見落とした領域がないことを確認した。その一例を C 断片について Fig. 1-3D に示す。

HindIII 断片間の連結部位については、まず、KI723T1 株から抽出した pOAD2 を EcoRI で消化して得られた 24 kb と 15 kb の断片を pBR322 に組み込み、pNDE1 (24 kb 断片; pOAD2 上での位置が 8.5-31.5 kb) と pNDE2 (15 kb 断片; pOAD2 上での位置が 34-43.6/0-5 kb) (Fig. 1-1) を作成した。pNDE1 から 0.8 kb の SalI 断片 (A-F 連結部位を含む), 0.6 kb の SalI 断片

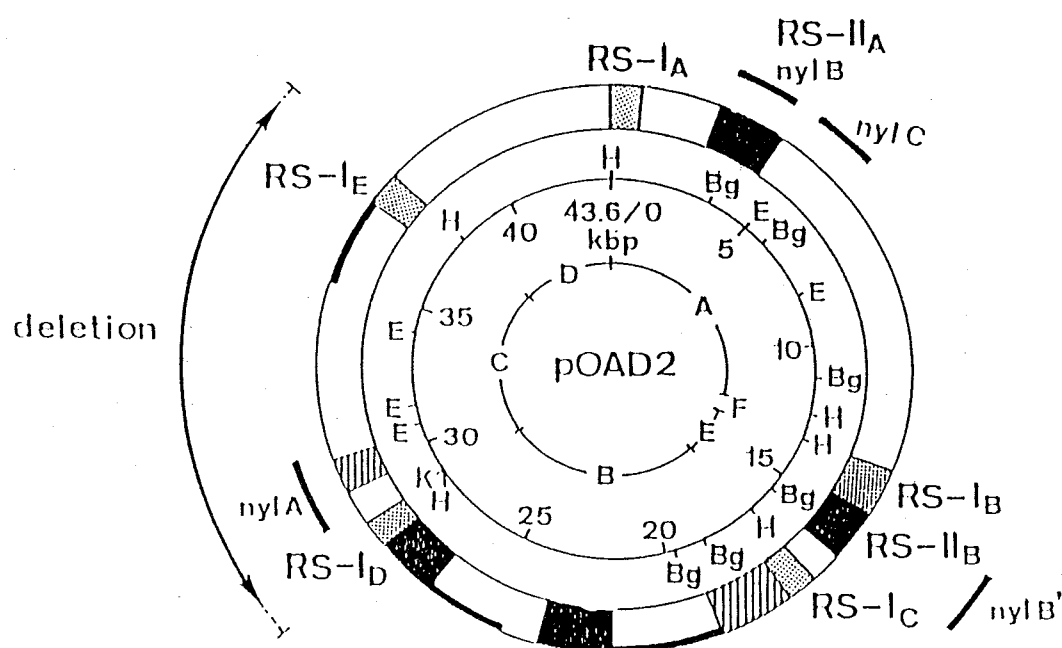


Fig. 1-1 Structural and functional map of pOAD2 plasmid. The nylon oligomer degradation genes were at the following loci: the nylA (29.1-30.6 kb region on pOAD2 map), nylB (2.7-4.1 kb region), nylB' (14.7-16.1 kb region), and nylC (4.2-5.3 kb region). RS-I and RS-II are repeated sequences identified by southern hybridization experiments. RS-I appeared 5 times on the pOAD2 (RS-IA, 0-0.8 kb; RS-IB, 13.5-14.7 kb; RS-IC, 16.8-17.4 kb; RS-ID, 28.2-28.8 kb; RS-IE, 37.1-37.9 kb region on the map). The nylB and nylB' genes are included in RS-IIA and RS-IIB-regions, respectively. E, Bg, H, and K represent restriction sites for EcoRI, BglIII, HindIII, and KpnI, respectively. A, B, C, D, E, and F represent fragments generated by HindIII. Arrow indicates the region deleted in pOAD2.

(F-E 連結部位を含む), 0.7 kb の *Nru*I 断片 (E-B 連結部位を含む) そして 2.0 kb の *Cla*I-*Kpn*I 断片 (B-C 連結部位を含む) をそれぞれ M13 をベクターとしてサブクローニングを行い、少なくとも一方向の塩基配列を決定し、連結部位に見落とした領域がないことを確認した。同様に、pNDE2 から 1.0 kb の *Nru*I 断片 (C-D 連結部位を含む) と 2.5 kb の *Pvu*II-*Sph*I 断片 (D-A 連結部位を含む) をそれぞれ M13 をベクターとしてサブクローニングを行い、確認した。

3 pOAD2、初発試料および小断片間の整合性

pOAD2 と初発試料 (pNDH5, pNDH6, pNDH10, および pNDH16)、また初発試料と小断片を適当な制限酵素を用いて分解し、アガロースゲル電気泳動で生じた断片の長さを比較した。各試料間に整合性の崩れは見られず、大腸菌宿主 (JM109) 内や遺伝子操作中に、組換え、挿入、欠失などが起きていないことを確認した。

4 pOAD2 の全塩基配列

以上の情報を整理し、pOAD2 の全塩基配列を決定した (Fig. 1-4)。pOAD2 の全長は、45,519 bp であった。その G+C 含量は 66.6% であり、300 bp 以上の ORF が 157 個認められ、600 bp 以上にわたり、終止コドン (TAA, TAG, および TGA) の存在しない NSF¹⁰⁾ が 71 個認められた。各 NSF について、その長さ と pOAD2 上での位置を Fig. 1-5 に示す。太線太数字で示した NSF は時計回りの方向 (外側からフレーム 1, 2, および 3)、細線白抜き数字で示した NSF は反時計回りの方向 (内側からフレーム 4, 6, および 5) である。フレーム 1 と 4, 2 と 6, そして 3 と 5 はそれぞれ相補鎖の関係である。領域により、NSF の数と長さに片寄りが見られた。0-18 kb の領

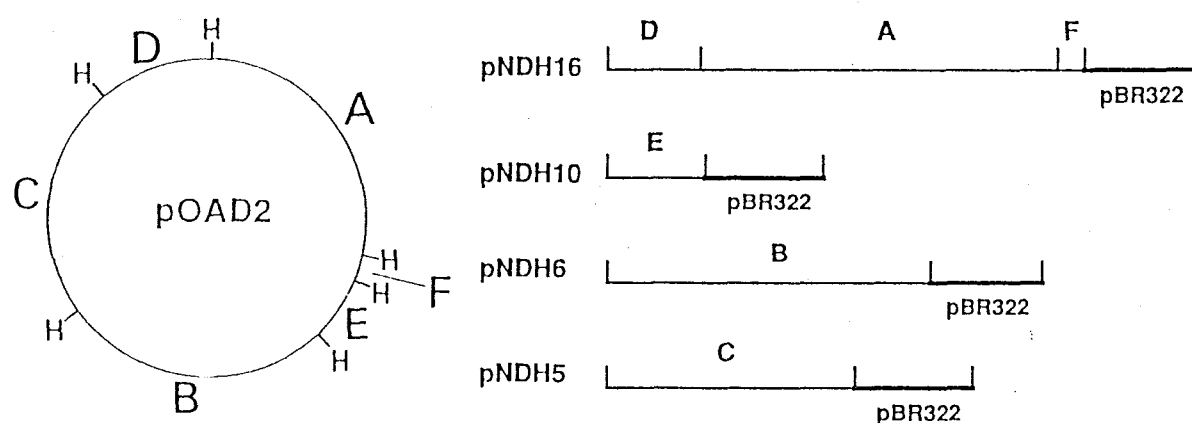


Fig. 1-2 Primary *Hind*III fragments from pOAD2 for sequencing. The positions of *Hind*III restriction sites on pOAD2 map (left panel). Primary fragments for base sequencing (right panel). Plasmids pNDH16, pNDH10, pNDH6, and pNDH5 were composed of pBR322 and various *Hind*III fragments of pOAD2.

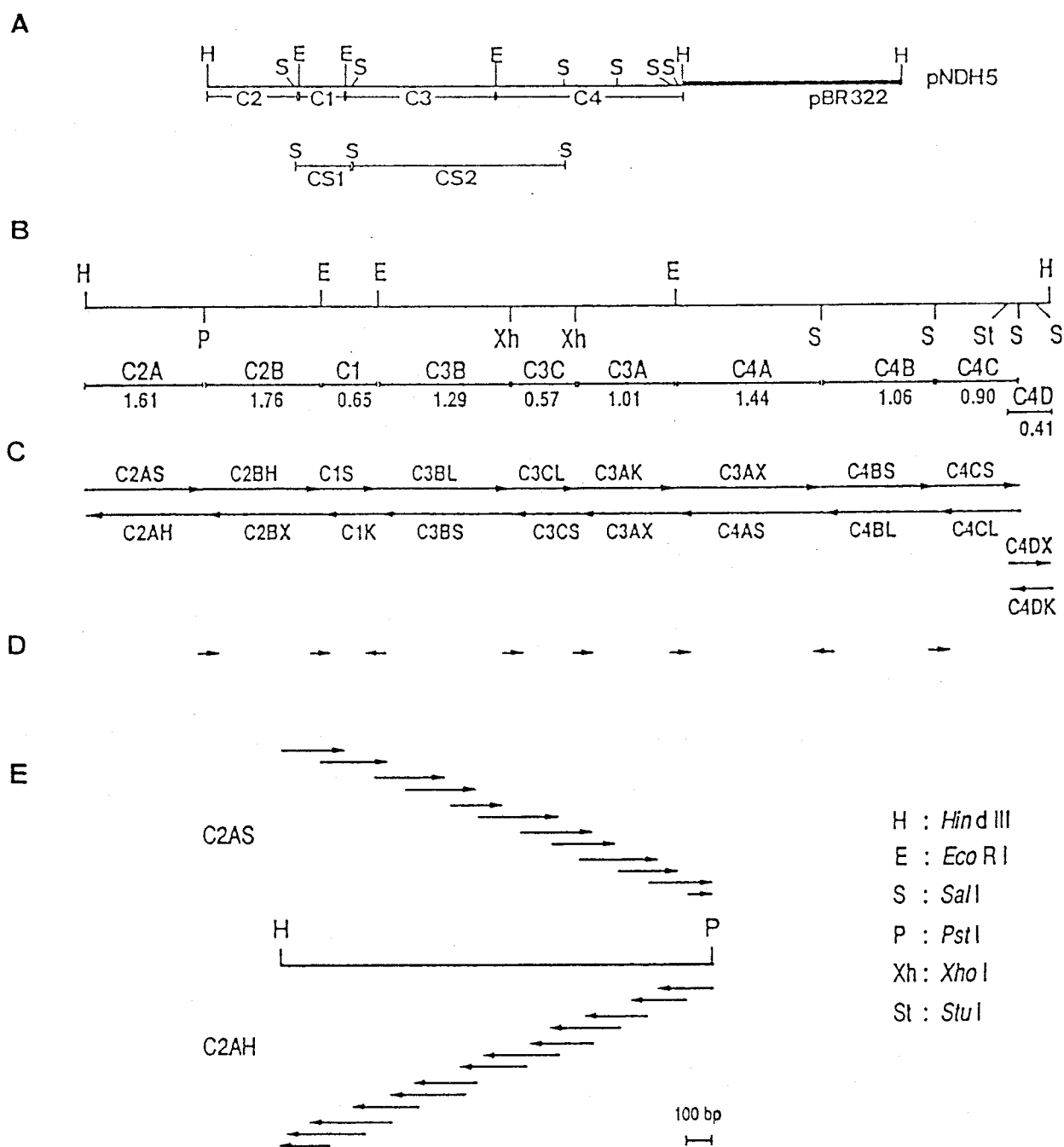


Fig. 1-3 Sequence strategy of *Hind*III C fragment of pOAD2. A) Restriction map of pNDH5. The fragments indicated were cloned into pBluescriptII SK⁺ vector. B) The fragments indicated were cloned into SK⁺ vector. C) The fragments indicated were cloned into M13 vectors and sequenced in the direction shown by the arrows. D) Connection of the fragments. E) Sequence strategy of C2A fragment. The deletion mutants made by *Not*II/*Mun*I have been sequenced. Arrows indicate direction and extent of DNA sequencing.

AAGCTTGAGC	ATGCTTGGCG	GAGATTGGAC	GGACGGAAACG	ATGACGGATT	TCAAGTGGCG	CCATTTCCAG	GGTGATGTGA	80
TCTGTGGGC	GGTGCCTGG	TATTGTGCT	ATCCGATCAG	CTATCGCGAC	CTTGAGGAAA	TGCTGGCGGA	ACGCGGCATT	160
TCGGTCGACC	ATACGACGAT	CTATCGCTGG	GTCCAGTGCT	ACGCCCCGGA	GATGGAGAAG	CGGCTGCGCT	GGTTCTGGCG	240
GCGTGGCTTT	GATCCGAGCT	GGCGCCTGGA	TGAAACCTAC	GTCAAGGTGC	GGGGCAAGTG	GACCTACCTG	TACCGGGCAG	320
TCGACAAGCG	GGGCGACACG	ATCGATTCT	ACCTGTGCGC	GACCCGACGC	GCCAAGGCAG	CGAAGCGGTT	CCTGGGCAAG	400
GCCCTGCGAG	GCCTGAAGCA	CTGGGAAAAG	CCTGCCACGC	TCAATACCGA	CAAAGCGCCG	AGCTATGGTG	CAGCGATCAC	480
CGAATTGAAG	CGCGAAGGAA	AGCTGGACCG	GGAGACGGCC	CACCGGCAGG	TGAAGTATCT	CAATAACGTG	ATCGAGGCCG	560
ATCACGGAAG	GCTCAAGATA	CTGATCAAGC	CGGTGCGCGG	TTTCAAATCG	ATCCCCACGG	CCTATGCCAC	GATCAAGGGA	640
TTCGAAGTCA	TGCGAGCCCT	GCGCAAAGGA	CAGGCTCGCC	CCTGGTGCCT	GCAGCCCGGC	ATCAGGGGCG	AGGTGCGCCT	720
TGTGGAGAGA	GCTTTTGGCA	TTGGGCCCTC	GGCGCTGACG	GAGGCCATGG	GCATGCTCAA	CCACCATTTC	GCAGCAGCCG	800
CCTGATCGGC	GCAGAGCGAC	AGCCTACCTC	TGACTGCCGC	CAATCTTTGC	AACAGAGCCC	CGACGACTCC	GACATCGACG	880
ATGACGACAC	GGTGCCTCC	GGCGGGCCGC	AGGCGCGCCG	GATCCCGAGC	TTCTCGATC	CGGATCGCCC	GATGGATCCC	960
GTGCGCATGT	CCGAGACCGG	CAGCCTGGCG	GAGCACCTCG	GGTTGCGCCG	GAAGAACGAC	GACCAGACGG	ATGTGGGGCC	1040
GGTGCATGTA	GGCGCGGACT	GGTGCTCGGC	GAGGTCTGGA	ACGGCCTGCG	CGCAATATG	TCGGTGGTCA	TCTCCGTGGT	1120
GCTGGTGACC	TTCTGTGTCG	TCACCTTCGT	GGGGGCGGCG	ATCCTGATGC	AGCTGCAGAT	CCAGCAGATG	AAGACGTTCT	1200
GGTACGACCG	CGCGCAGGTG	GCCGTCTACC	TCTGCACGGA	CTACGACCAG	AGTTCGACCT	GCGCCGGCGA	GGACGCGGGC	1280
GAGGACGAGA	TCGCGGCCCT	CGAGGAGGCG	CTGCAGTCGG	ACACGCTGCG	CCCCTACGTC	GACGACTTCT	TCTTCGTGAC	1360
GCACGAGCAG	GCCTACAAGG	AGTTCTGTGA	GCAGTTCGAG	GGCAACCCGA	TCGTCGACAT	CACGAAGCCG	GAGCAGCTCA	1440
ACCAGACCTT	CTGGATCAAG	CTCAAGGATC	CCTCGCGCTC	CGACATCATC	ACGGAGACGT	TCTCGGGCAT	CCCCGGCGTG	1520
CAGAGCGTCT	CGGACCAGCG	CAGCCTGCTC	GATCGGATCT	TCTGTTCCT	GGGGGTGGCG	AGCTACACGG	CGATCGTCAT	1600
CGCGGGGCTC	ATGCTGGTGG	CGGCCATGCT	GCTGATCTCG	ACCACGATCA	GGCTCTCGGC	GTTCTCGCGC	CGCGCGGAGA	1680
TCGGGATCAT	GCCTCTCGTG	GGCGCGTCGA	ACCGGTTTCAT	CCAGACCCCC	TTCATATTGG	AGGGCATCAT	CGCCGCGCTC	1760
TTGCGCGCCG	TGCTCGCGAG	CGCCGCGTCG	GTGGCGATCG	TGCGCTTCTT	CGTGCAGGGA	TTCTCGCCC	GGGAGGTGCC	1840
GTTACAGAGT	TACATCACGG	TGGAGCAGGC	GCTCGTGGTG	CCGCCGATCC	TCGTGCTGGT	GGGAGTGGTG	CTCTCGGCGA	1920
TCGCGCGGAA	GATCGCGATC	ACCAGATACC	TGCGGGTGTA	GCGCAGAGCG	TCCCCGGGGC	GGGCCGCGGA	TCCGGATCTC	2000
GGGGCCGCGG	CTCGTCGAAA	CCGAGCCCAT	CGTGCGAGCC	CGAGCCCTCC	GCGCCGATA	ATCCGTGCGG	GTTGGAGCAA	2080
TCGGCTCGGT	TCGGGCGCCT	CGAGGCCGGG	CGGCTCGGCG	TGGCGGGCGG	AACGGCGGCG	GTGTGCGATA	AGCTGGTTCT	2160
TCGCGCGTGA	GCGCGGACCC	CGAACCACCC	CGGCCGACAG	AAGGATGTGA	CATGAGCCCG	AAGGAGACCG	GCGAGAAGCT	2240
GATCGCCTCC	AACAAGAAGG	CCCGGCACGA	GTACCACATC	CTCGACACCT	ACGAGGCGGG	CCTGGTGCTC	ACGGGCACGG	2320
AGGTGAAGTC	GCTGCGCATG	GGACGCGCGT	CGCTCGTCGA	CGGATACGTG	TTCATCGAGC	GCGGCGAGGC	CTGGCTCGAC	2400
GCCGCCTACA	TCCCCGAGTA	CCTCAACGGC	TCGTGGACGA	ACCACGCGCC	GCGCCGCAAG	CGCAAGCTGC	TGCTGCACCG	2480
CCAGCAGATC	GACAAGCTCT	ACCAGAAGAC	CCGCGAGGGC	GGCATGACGA	TCGTGCCGCT	GCGCCTGTAC	TTCTTCGACG	2560
GCCGCGCGAA	GGTCGAGATC	GCCCTCGCGA	AGGGCAAGAA	GGAGTACGAC	AAGCGCCAGA	CCATGCGCGA	GCGGCAGGAC	2640
AAGCGCGAGG	CCGAACGGGC	GATGCGCACC	CGCAACCGGA	TGGGGGAGTA	GGCGCTCCCT	CGGGTGTCCC	CGCGTCGCGG	2720
GCGGATCCTG	CGACGCGCCT	GACGCGGTGG	CGCCACCGCG	TCAGGACGTC	GCAGGATCCA	CCGGCGCGGT	GCCGACGCGG	2800
GTCACGCGCG	CTACACCGCG	TCGAGGGCAC	GGCTGACGTC	GAGCAGGATC	CCGCTCTGCA	GCCCGTGCCA	GTGCGGGGTG	2880
TCGGGATCGG	GCCACGACGA	GAGCTTGACG	ATCACGGAGT	CGGTGCGCGG	ATCGAGCCAG	AGGTTCTGGC	CGTGATGCC	2960
GATGCCGCTC	ACGTTGCCCG	GCTCGTTGCC	CGTGCAACAC	CACCTGGCGG	TGTAGCTGCC	CTCGGGGAAT	GCGGAGGTGA	3040
AACCTCTGTC	GGTCATGGCT	TCGCGGGAGC	CGCCGCGCCAG	CACGCTTTCC	ACCCAGCCCT	GCGATACGAC	CCGTCCGCGG	3120
GGAGCGACGC	CGCCGTCGAG	CATCATGCGG	CCCACGCGTG	CGAGATCCCG	CGCGGTGCAG	GAGACGCCCC	CGTTCCGCGAA	3200
GCCGAAGCCG	GTCTGGTTCGA	CCGTGATGGT	CGCATCCCGA	TCGGCGTCGA	GCTTCGCCCA	CAGGTACGTG	GAGAGCGCTT	3280
CGACGTAGCG	CAGACCGGTG	ACCGCTCGA	CGATCCAGGC	GAGCACGTCG	GTGTTGCGCG	AGCAGTACTG	GAACTCGCGG	3360
GTGCCGCGCT	CGCCGCGGAG	GGTGGTGAGG	AACTCGTAGG	TGTCGGCGGG	GTCCCCGTCT	CGCCGCGTGC	GCCAGCCGGC	3440
GGAGCGATCG	TGGGTCTGCA	CCTCCGAGGC	CGGATCGACG	TAGTCTCTGT	TGTAGTCGAT	CGAGATCTGC	ATGTCGAGCA	3520
CCTGCGACAC	GGAGGGCCCG	TCGTAGACGG	AGCCCGCGAG	CTCGGTGACA	TACTCGGTGA	CGGGCTGCGC	GGGATCGATG	3600
CGCCCCCTCGT	CGATCAGCGC	GCCGACGACC	GTGCCGCACA	GCGACTTCGA	GACGCTCATC	AGCAGGTGAC	GGTCGTGCGG	3680
TGCCAAACCC	GCCCCGTAGT	ACTCGGCGAG	GACCTCGGAG	CCGCGCAGCA	CGAGGAATGC	GTGCGTGCAG	GTCTCTCGA	3760
GCCGCTGCTC	GAGATCGGGG	AGCCGCGTCG	CGAGCGCGTC	GAGCCGACAG	ACGGGCTCCG	CGGGCTGCGC	CGGGTCCGCG	3840
CGGGAGACCG	CCGCCGTGGG	CAGCAGCTCG	CCAGGCGGGG	GCAAGGCCCA	GCGGTTGTGC	GGGGCCTCCT	GCCAGCTGTC	3920
GAGTGTGCGC	TCCCCGGCCG	CGGCTCCGGG	ATACCTGGCG	GGGTGTGGC	CGGTGGAACG	TGCGTTTCATC	GAGTGTCTCC	4000
GAAGGTGTAG	CCGCGGCGGG	TGAAGAGATA	GTAGAGGCCA	CCGCTGACGA	TCATCGCGAC	GGGGGCGGAG	AGATCGATGT	4080
AGCCCATGGC	CTCGGCGACC	GCGCCGCTCC	AGAAATCGGT	GGTGAGGAAG	AAGGAGGCGG	CAGCGCCGCC	GATGATCCCC	4160
GCGGCGAGGC	CGGAATGCT	GAAGCCGCGG	GTGTACCACT	ACTTCGCCCC	CCGGCGCTCG	TCGAAGAGAT	CGAGGCCATC	4240
GTATCGGTTG	CGGCGCAGCC	AGATGTCCAT	GCAGTAGACG	GTGATGCCCC	GAGCCGAGAT	GAGGATCAGC	ACCTGCAGCA	4320
TGATGTTTAC	CGCGTCGAGC	AGGCTGGTGC	ACATGACCAG	GATGATGGTG	AAGAAGGTGC	CGAGCACGCC	GATGATGATG	4400
GCCGAGGGGA	TGCGCCGGAT	CGGGATGCCG	ATCGACTGGA	ACGCCATGCT	GCGGTTGTAG	GTGGTCATGC	CGTCGACCGC	4480
GTGCGGGATC	CCCTCCCCGA	AGCGGCTCCC	GATCTCGGCG	CCCCGGCTTC	CCGGATCCGT	GCCCGTCACC	GTGCGGGTGA	4560
TGATCTCCAC	AGCGTTCTCT	TCCTGTTCCG	GGGCCGGGCC	CCTACTTGCC	GGCCTCGAGG	ACGGCGTCCC	ACATCACCTC	4640
GGAGGCGATC	GCGCCGAGCG	CGGTGCGGTT	CACCGACAGC	CGCCCGCGCG	ACGACCCCGG	GGTCGTGCGC	AGATCGATCT	4720
CGTCGGTGGT	GACGGCGAAG	AGCGTGTGCG	CGTCCATGTC	GGTGTGGAAC	GGCTGGATGC	CGGGTGCAT	CGAACTGTGC	4800

Fig. 1-4 Nucleotide sequence of pOAD2

ACCTGCTTGG	CGAACTGGTT	CAGCTCGACG	GGGCTCATCC	GCACGTTTCGT	GACGATCGCG	CTGATCGTGG	TGTTGCCGGC	4880
CTCGGTGACG	GGCGGGACCT	GCTCGGCGAA	CGCCTCCTGG	TAGTCGAACA	CCGGGTGGCG	CCGGACCCCG	GTCTGCCGCT	4960
CGTAGTTGCC	GCGCACCAAC	GTGCCGCGCG	GGTCCACGAT	CACACCGACC	GGTTTCGGCA	CGACGACGGC	GAGGATCGCG	5040
ACGTGCGCGA	GACGACGGAA	CGCCGCGCCC	TGCCCGGTGA	TCTCGGTGCG	GTCCCAGTCC	ACCTTGCCCG	CGGACGCGCT	5120
CATGCCCCGC	CCCCCGCGCC	CCTGCGGGAA	CTCACCGGGA	ACGGCGAATT	CGAGCGCCGC	GCGGCGGAGC	GCCTTGTCGG	5200
GGTAGACCGC	GGTGGAGCGC	GCCGAGAACT	CGTAGATGAC	CGCCGACGAC	ACCAGCTGGA	GCTCGGCGAA	GCCGGTGCGA	5280
TGCTCGAGGC	GTTCCAGGAG	CGCGTCGCTC	ACCCCGGCGC	CCGCCTCGAG	CCCGTAGCCG	GCTCCGCGCG	CCAGGCAGAT	5360
GGCGTGTTG	AAGTCGTAGC	CGCCGAGAG	CCCCACCGCC	CCGCCCGCGC	CGTCCACCGC	GGTGCGGGCG	CCGGCGGGGA	5440
TGTGGATCAC	GGTCGCACCG	GTGGGCCCCCT	CCTCGTAGTG	CGCCGCGCGC	ATGCTGACCC	CGGGGAAGTC	GAACCGCAGC	5520
ATCTCGCGGC	CCGTGCTCCT	GACCGCGCGC	AGATCGAAGG	CGTCGTTGCC	CGGACCCCGG	AAGACCGGAG	GGCCGGCCAG	5600
CCGGGGTGCG	GGATCGACCG	CGATCCCGCC	GTCGATGTGC	GTGAGTGCGT	GGACCGGTGT	CGTATTGATA	GGGCCTCCAA	5680
GAGTGTCTGC	GAGTAGGGAT	GCCGGGGGTT	CCGGAAGAAC	TCGTGCGTGC	CGGCGGTCTC	GACGAGATCT	CCCCGGAACA	5760
GCACGGCGAC	GCGATCGGCG	ACGCGCTGCA	CCGCGCCGAG	ATCGTGCGAG	ATGAAGAGGG	CGCCGAAGCG	GTGCTCCGCC	5840
TGCAGTTCCG	GGATGAGCGC	GAGGATCTGC	ATCTGCACGG	CGACGTCGAG	CGCGCTCAG	GCCTCGTCGA	AGACGATGAA	5920
GCGCGGCTCG	GTCACGAGCG	CCCGGGCCAC	CGAGACGCGC	TGGCGCTGTC	CTCCGAGAG	CTCGTGGGGG	TAGCGCTCGG	6000
CGAAGGCGGG	GTTTCAGCCCC	ACCTGCTCGA	GCATCTCCGT	CACGCGCTGC	CGCTGCTCGG	AGCGGCCGAG	GTCGCGTCGC	6080
GCCACGGTGA	GCGGTTTCGG	CACCGACTCC	CCGACCTTGC	GCCGGGGGTT	GAGCGACCAG	TGCGGATGCT	GCAGCACCGC	6160
CTGCAGCTTG	CCGGCGTTGG	CCCTGCGCGA	CCGGGTGAAG	GGGCTCCCGA	GCAGCAGCAC	CTCGCCGGAG	TCGGGCCGCT	6240
TGAGGCCAG	GGCGACCGAG	GCGGTCTGTG	TCTTGCCCGA	TCCGGAAGTC	CCCACCGACC	CCAGGTCTCT	GCCCGGGGCG	6320
ACCTCGAGCG	TCACCGCGTT	CATGCGCGCG	ACGGTCCGCC	GCTTCGCTCC	CGCCCCGCTC	GTGAACGCGA	CGCTGACATC	6400
GCGCAGCTCC	AGCAGGGGGG	TCGTCTTCGC	AGACATCATT	CCGTCTCCTT	CGCCGCGAGA	TCCGCGCGGC	CGGGCATGGC	6480
CGCGACCAGC	GTTCTCGTGT	AGTCCGCCCG	CGGCTCGTGC	AGCACCCGCG	CGGTCCGGCC	CTCCTCGACG	ATACGCCCTT	6560
GGTGCATCAC	GACGACCCCG	TCGCACCAGC	GGCGCACCGC	GCCGAGGTGC	TGGCTGATCC	AGATGAGGCT	CGCGCCGGAG	6640
CGCTGCACCA	GGTCGACGAT	GAGGGTCAGG	ATCTCGTTCG	GCACCTGTGC	GTGAGCGGAT	GCCGTGGGCT	CGTCGGCGAT	6720
CAGGATCTCG	GGTCGCGCGC	CCATGGTCAG	CGCGATCGCG	ATGCGCTGCG	CCATCCCGCC	CGACACCTCT	TGCGGGTAGA	6800
GCCTGAGCAC	CGCTCGGGCG	TCGAGGATCC	TGACGCTCTC	GAGCAGTTGC	GCGAGCGACC	CCGGGTCTGAG	CGGCCCGCCG	6880
AGCGGCCCTCA	GTGCGAGTTC	GAGCTGGCGG	CGGATCCGCA	TGCTGGGGTC	GAGCGAGGCG	ATGGGATCCT	GCGGGATGAA	6960
ACCCAGCGAC	TCGCGTCGCA	GCGCCCGTAT	CGCATCGGGC	GCGAGGTCCG	CGAGATCCCT	GCCGAGCAGC	TGCACCGACC	7040
CGGAGGTAC	GTGGCAGGCC	CGCGGCAGCA	TGCGGCCGAT	CGCCTGGGAC	ACGGTCGACT	TGCCGGATCC	CGACTCGCCG	7120
ACGATCGCGA	GCAGTTTCGCC	GCGGGCGACA	TCGAGATCGA	TCCCGTCGAG	CGCTCGGGTC	ACGCCGAGAC	GGGTGCGGTA	7200
CTCGGCGCTG	AATCCCGTGA	GCGAGACCAC	CGGGGCAGCG	GGGGCGTTTC	CGGTGATACG	AGGCTCCTTG	TCTCGCGCGG	7280
ATCGAGCCGA	TCGCGCATCG	TGTCGCGCAT	GAGGTTGATG	GAGAGGATCG	CGACGACGAG	CACCGCGGCC	GGGATGAACA	7360
GCAGCCAGGG	CGCTGACGAG	ATGTAGCTGC	CGCCCTCCTG	GATCAGCAGG	CCCAGCGACG	AGCGGGGCGG	CTGCACCCCG	7440
TAGCCGAGGA	AGCTCAGTCC	TCCCTCGACG	AGGATGCCGA	CCGACAGCGC	GTACGTCCCC	TGCACCGCGA	TGGTGCCGCT	7520
GACGTTTCGG	AGCAGGTGCC	GGAAGAGGAT	CACCGGGGCC	GACACGCCGC	TGATCTTCGC	GGTGGTCACG	AAGTCGCGCT	7600
TGGCGACCGA	CATCGCCGCC	GCGCTCACCA	TGCGGGTCAT	GAGCGGCACC	GTGACGAGCA	CGATGCTGGA	GAGCACCGCG	7680
GCCTGGCCGG	GGCCACGACG	CGCGGCGATG	AGGATCGCGA	GCACGATCGA	CGGGAAGGAG	TAGAGGACGT	CGCTCAGGCG	7760
CATGATGACT	TCGCGCGCCC	AGCCGCCCCG	GTAGCCGGCG	GATAGGCCGA	GGATCACGCT	GATGACCGCG	GTGCACACGA	7840
CGGCGAGGGT	GGAGAGGACG	AGCGTGGTGC	GCGTCCCTTC	GAGCAGGCGT	GCGAGCACCG	AGCGGCCGAG	GGCGTCCGTG	7920
CCGAGCGGGT	ACGCCGCCGA	CGGCGGGCTG	AGCCGGGTGC	CGACGATCTT	CGCCGGGTTC	CCGGCGAAGC	GGATGACGGC	8000
GGCGGCGACC	GCGACGAGCA	CGATGAGGCC	GAGCGCGCAG	GCGGCGGCGA	TGAAGAGGGC	GTCGCGGGGC	GGGCGGGTCT	8080
TGACCGCGCG	CTTCTTCTTG	GAGAGCGGGA	GGGAGTAGGT	GGTCATGATC	CTGCCTTCCC	GGTGACGGCG	ACGCGCGGGT	8160
CGAGCGCGCT	CGTGATGAGG	TCGAGCACGA	GGCTCAGCAC	GACGAAGACG	GCGGCGGCGA	AGAGCAGCGT	GGCCTGCACG	8240
ACCGCGAAGT	CCCGCGGGCC	GAGCCCGGTG	ACCATGAACG	ATCCGAGACC	GGGCAGGTTG	AACAGGGTCT	CGATGATGAC	8320
GGCGCCGCGG	AGCAGGTAGG	CGGTGATGGT	CCCGAGCAGG	GTCAGGATCG	GGATGCCGGC	GTTGCGGAGC	ACATGGTGGC	8400
GCACGATGTG	CCGGGTCTGC	TCCCGCGCTG	CCACGGCGGC	CGTGATGTGC	GGTTCGACGA	GCACTCCCAT	GACGGCGTGC	8480
CGCGTCTGTC	GCGCGGTCTC	CGCGACGACG	AAGACGGAGA	GCACAGGGC	GGGAGCAGG	AGCGCCCCGA	GCGTCGGCCA	8560
GAATCCGGTG	CCGATGCTCG	CGAACC CGCC	GACCTTCAGG	CCGAGCGAGA	ACTGCGAGAA	GAGGAAGATC	ACGATGCTGC	8640
CGAGCACGAA	TTCCGCGACG	CTGATGCCGA	TGCTCGACAC	GACGCGGCCG	ACGGCACCGC	CCTGCTTCGC	TCCGGACCTC	8720
ACTCCCGTCC	AGATCCCGAG	CGGAACCCCG	ACGAGCAGCG	TCACGGCCAT	CGCCATGATC	GTGAGCGTGC	CGGTACCCCG	8800
GATGCGCTCG	AGCATCTCGC	CGGAGACGGA	GCTCTGCGTG	ACGTAGGAGG	TGCCGAGCTC	GCCCCGTGAG	ACCTGTCCGA	8880
TCCAGGCGAA	GTACTGCACG	AAGATCGGCT	GGTTAGAGCC	GTACGCCCTC	GCGAGCAGCT	GCCGTTCCCT	GGCGGAGGCG	8960
AGCGGGCCGA	GGATGATGTC	CTCGTAGCTG	CCGGGCATGA	GACGGATCGC	GACGAAGATG	AGCACGGAGA	CACCGAGCAG	9040
GGTGATGAGC	GCGCTCATCA	CCCGCCCGGC	GGACCAGCGA	AGGATCTGCA	TCATCCGGAA	CTACTCCCCG	GCCATCCGGT	9120
AGTCGACGAT	GTTGCGGAGG	ATGTTGCCGT	ATCCCTCGTT	GCTCTGAATC	GTCGGGCTCA	CGGCCGTGGT	GTCGAAGCCG	9200
ATGATGGACG	GCCGGTGAC	GAGCGGCACC	ATCTCGCGCT	ACTCGTCCGC	GGAAAGCGAC	AGCCGGGTCA	GCACCTCGGC	9280
GCGCGCATCG	GGATCGGTTT	CGCTCGCGCC	CTCGGCGATG	AGCGCGTTGA	GATCCTCGTG	CACGCCCGTG	AAGCCGACGT	9360
TGAAGAAGGC	CTGCTCGGGG	TTCCACCACT	TGGTCACCAT	CGACGGGTGC	GAGTAGCCCG	CGAACCAGCT	GATCGACAGG	9440
TCGAAGTCGC	CCGGCTGGGC	CGAGAAGACC	CGCTCGTTGT	AGGTCGCGGT	GTCGACCTGC	TCGATCCGGA	CGGTGACGCC	9520
GATCTCCGCG	AGCTGCTGCT	GCACGAGCTG	GGCGATCTGC	GCCATGACGG	GTTCCAAGGT	GTTGACGATG	AGGTCGAGCT	9600

CGACATCGCC	CGCGCCCGTC	TCCTCGATGG	TCGCCCTCGC	GTCTTCGATC	GCGGCCTGCT	CCGACGGCAG	CTCGTCGACC	9680
AGGCAGGCGC	CGGGCAGGTT	CGAGGGGGTG	ATGCCGGTGG	GGTTCGTTTC	GCCGGCGAAG	ACGATGTCGG	CGATCGCCTG	9760
GCGGTCGATC	CGCGTGTTGA	TCGCGAAGCG	GACGTCCTGG	TCGTTTCAGCG	GGGAATCGGG	CTTCTGCGAG	TTGACGAGGA	9840
CGTAGTAGAG	GTGCTCTGCG	GTCTGCTTGA	CGATCTGGAG	GCCCTCGTTG	CCGCCGAGCT	GCGACAGCGA	GTCGATGTTG	9920
ACGAAGTTCG	CCATCCCGGC	GCTGCCTTCG	CGGAGCGCCG	CCTGCCTGGT	CGACTCCTGA	TCCACGATCT	GCAGTTCGAG	10000
GGTCTGCACG	GCGAGGTTGT	CGCCGCCGTG	CCAGTCGGGG	TTCCGCGTGA	AGGTCCACGA	CTCGTCTGCT	CGGTGGTCGG	10080
CGACGAGGTA	CGGGCCGGTG	CCGAGCATCG	TCGTGGTGAG	GTGACGCTG	CCGTCTGTTA	CCTCTGTCAT	CGGGAGGATC	10160
GCGGCCGCGC	TGTTCCGCGAG	CGCGGCGAGG	AACGGCGCGT	ACGGGCTGCC	GAGCTTCACG	GTGACCTGAT	CGTCGCCGGT	10240
GGCTTCGACG	GACTCGACGG	GGCCCATCTG	GCCGGACCAC	ACGCTCTGAC	TCGCCACCTG	GCGCTCGAGG	GTGCCGACGA	10320
CGTCGTTCGAC	GGTGAGCGCG	CGGCCGTTCG	AGAACGTCGC	GGTGGTGTCG	ATCGTGAAAG	CGTAGGTCGT	GTCGTCCGGC	10400
TGCTCCCGAG	AGGTGGCGAG	CTTCGGCTGG	ATCTCGAAGT	TCTCGTCCGT	GGTCACGAGC	GTCTCGTAGA	CCAGCCCCAT	10480
GATCATCCAC	GAGCGGTTCG	TCCGCGCGGT	CTGGGGGTTC	AGCCCGTCGG	CCTCGCTCGT	GTCGTAGTCC	AGCTGCACGG	10560
TGAGGGCGCG	CGTGCCCGTG	AGCGCCTCGT	CGGCGACCGC	GAGGAAGCCC	GGCGCGGCTT	CCGTCTGCTC	GGTGCCGCCC	10640
GGCGCCGAGC	CGGGCGCGGC	CGAGGAGGTG	AGGACGCACG	CGAGGCTCAG	CCGATCGCGG	GCGCCGCGCG	CCGCCCGGGA	10720
GCGGGTTCGT	TGGTTGATCA	ACACCGTGCT	GACTCCTTCG	ACAGTGGTGG	TGGCCCGGAC	GCGCGCGATG	CGCTCCGGCC	10800
GCGGCGCGCC	GCCGATGGGC	GCGCCTTGGA	TGATCACGGT	ACATACATAA	ATCGTTTTAC	CTCACGACGG	TAGCAGAGTA	10880
AAACGATTTA	TGTAAGATG	CACGGCACCC	ACCGCCCCAC	GAGATCCTCG	TCTCGACCGA	CTCGTCATCC	GGCGCATCGG	10960
CGGATGCTCC	GGGCTCCCG	GCTCCCGGTC	CTCCGCCGCT	GAGGCGAGTC	GCGCCCATGC	GCGCTGGTCA	GGAGGCGTCG	11040
AGATCCCGCG	AGAATCGTGC	ATCTCCCGAC	GGCGCGGAGT	CGTCGGGGAC	GGGGCGACGC	AGCCGGCAGG	CGATCGGGAG	11120
AGGGCGCTCG	GTGCCCCCGG	CGGCGACAGC	CCGCCGGGTG	CCCGGGAGGC	GCCCGGCCGG	CCGTTCCCCC	GCGCCCGGGG	11200
GGAGCCCTGC	GGCCCGTGCC	CGATGGGCGG	GCGCACGCGG	GCCGGTGCCC	GCGCTCACA	CCGTGAGCGT	GCGCCACTCC	11280
CCCGTCTCGG	CCTCGGTGGG	CATGCCGTCG	AGGATCCGGA	TCGTGCGGGC	CTCCGGGTTC	GTGATGAGCG	TCGCGAGGGT	11360
GGCCACCGG	TCCCCGTAGC	CCTTGACAT	GTCGGCGCGG	CAGGTGAGCG	GCGGCTCGCC	CTCCTCGGTC	AGCAGCACCC	11440
GCACCAGGTC	CCCCGAGGCG	ACCGGCAGGC	CCTCGGCGAG	GCGCGCGCGC	ACGAGCGCGA	GGCGCTCCGA	GGAGTCGGGG	11520
TCGTAGATCT	CGGACTTCTG	CGCCGCGAGC	GCGCTCTCGT	GCTGGAAGTG	ATTGGTGCGC	TGCACCGAGC	CCCGACCGGC	11600
CTCGATCAGC	AACACCCCCG	CGGGGCTCAT	CTCGACCGAC	ACTGCGCGCT	GCGGCTCGAG	CATCGTGAAC	GCCGAGGACG	11680
ACGCGATCCG	GCGCTCGCGG	ATCAGCTCGA	TGGCCTCGTC	GACGGAACGG	CACTCCGAGA	GCACCACGCT	CGAGAGCACG	11760
TGCATGGGCA	CACCGTCGGG	GCGATCCTCG	CGGTGCCCGA	GGATGTTGAA	GTGCAGGGCC	AGGCCGGCCT	CGTTACGCGC	11840
GATCTTCGAG	AGGATGCCCT	GCTCGGTTCG	GCCGGCGTAG	CGCAGCCCGG	GGCCGGAGAC	CGTGTGCGTG	TGCCAGAACG	11920
GATCGAGCTC	GACGTGCCAG	TCCCAGGTCT	GCACGCCGTA	GCGGTGGTTC	CCGATCTCGG	CCGTGAGCGT	CGAGCACTCG	12000
CTCGACGGCT	TCTCGCCGAG	GGCGATGATC	TCGGTGCGGG	CGTTGAGCGC	GACGACCTGC	GAGAGCTCGA	GCGCCGGCCC	12080
CTCGGCGATC	CCCTCGAACT	GCTCGAGCAG	CCCCGGCCGC	CAGCCGCGCA	GCGCGGAGAC	GACCCGCGCG	ACGCCGTCGC	12160
GCTCGCGCTC	CTCGTCCACG	CCCAGCACGC	GGAACAGCTC	GCGCTACTTC	GTGTAGGCGC	CCCCGAGGCT	TCCGCGCAGC	12240
GCGGCGCCCC	GGCTGAGGCC	CAGTTCCGCG	GGCGTGCGGC	CCGCGACCTC	GAGGTGCCGG	AGTCGGTGGC	CGGGTCGCGG	12320
ACGATCTGCA	GCAGGGCCGG	CACCTGGTCG	GTGATCCGCA	CCTCCTGGTA	CTCGCTGACG	TCCTGCGGGA	TGTTATCGCA	12400
GATCCGGTAC	GTGCGGCTCT	CGTCGAGGGA	GCCCAGGACG	CCGTGCGCCT	CGGAGACGAT	CTTCGTGAAC	GCCCCGGCGG	12480
GGTCGGGCGA	GCCCAAGTTC	ACCTCGGTGT	TCGGCTTCGG	CTCGGGTCTC	TCCGGCTGCT	GGTCGCCCGG	CAGGTTCCGA	12560
TCCCACCTGA	CCGAGACGGT	CGCCGTGTTT	GTGAGGGCGC	CGTTGTGCGC	CGAGATCGCG	CCCCGACGCA	CCTTACCAC	12640
AACGTCCACG	GTGAGCGCCT	GCCCGCGGTT	CGCCTGCACG	GTGCGCTGAT	CCAGCGCCCA	GGTCAGCGTC	TGGCCGGCGA	12720
CCGTGGCGTT	CACCCCGGCG	AGCCCCGTGT	CGGACACGTA	CTCCAGCTCG	GCCGGCAGTG	CGTCCGTCAG	GGTGACGCCC	12800
GCGTAGTCGG	ACACGTCTCT	CGGGAGGGTG	AAGCCGAGGC	GGTAGGTCGC	GGTGTCCGGC	GAGGTGCGCG	ACTTGTGAT	12880
GAGCTTCAGG	TCGGCGCCCG	AGCCCTGCGC	GTCTTGAAC	ACGCTGTCCG	GGTCGATGGG	CTCGGGGTAG	GAGACCACCA	12960
CGTTCTCGGG	GTGCGGGTCG	GGGTGCGCCG	GATCCTGCGC	ATCCTGGATG	GGCGTCACCG	CCCACGAGCC	GGTGTTTCGT	13040
AGCGTCCCGT	TGTCGCTGTC	ATCGATCCCG	TCGCCCGCGA	CCGCGAGGGT	CATGGTGAGG	GTCTTGCCGG	CGAGGGCCGC	13120
CGCGTTCGAC	GCGTCGGTGA	GGTCGACGGT	CCACCGGGTG	GAGGCCGCGG	TGAAGGTGCC	GAGGCCGGTG	ACGTCGGCGC	13200
CGTCCGCGTC	CACGCTCGGC	GTCCCTCGA	CCGCGAGGCC	GGCCGGCAGC	GTATCGATCA	GCGTGAACGA	CGCGTAGTTG	13280
GCGACGTTCG	CCGGAAGCTT	CGCCGACACC	TGGAACGTGA	CCCGCTCAGC	CGCGTGATTC	AGCAGCCCTC	CGTCCCATG	13360
GGGCTTGCCC	GCTCGGCCCG	TCGACAGCGC	CTTCGTGAGC	CCGGCTGCGC	CGGGCGTTCG	CCGCGCTTCC	ACCTCGGTGT	13440
TCGTGGCCCG	CGGGAGGTTC	TCGGTCGAGG	CGTCCGAGAA	GCTGACCGTC	GTGTTTCGCG	TGTTCTGTAT	CACCCCGTTG	13520
TCGGCGGCGG	TCGCGTTCGC	TGCCGCCCTG	ACGGGGATCG	TCAGGGTGAT	CGTCTTCCCC	GCCGCGTCCG	CCAGTTGCGC	13600
CGGCGTGTC	AGCGTCCACG	TCACGGTGTT	ACCGGACCGG	GACGCGGTGC	CGGACGCCGT	CGTCGGGGAA	CCGACGATCT	13680
GCAACGTGCT	CGGCACGGCG	TCCGTGCGCA	CGATCGACGC	GTACTCCGAC	ACGTCCGCCG	GCACGTTTCAT	GGAGATCGTG	13760
TACTCGGCCG	TCTCTCCAG	GCTGTTGAGG	ATCCCGTCGC	CGTCCGACTT	CACCTTCGTG	AACCCGGACG	GCGCGTCCGG	13840
CAGCCCGATC	TCGAGTTCGG	TGTTCTGTGG	GGGATCGGGG	TCCGCCGGCG	GGTCCGGCAT	CGACGGGTTC	TAGTCGACGT	13920
GCAGGGCGGC	GGTGTTTCGT	ACCGCGCCGT	TGTCCTGCGA	CGTCGCGGCC	GCCGACAGCT	TCACCGTCAC	GTCCACGGAC	14000
ACGGTGGATC	CCGCGGCCCG	GGCGATCTGC	GCCTGGGTCA	GCGTCCAGGA	CAGGTCTGCT	CCGTTACAGG	TCGCCCGGGC	14080
ACCCGCCAGC	CCCGAATCGG	ACTGATACGC	CAACTCGTCC	GGGAGCGGTG	CCGACAGCGT	GAGCCCTCGA	TACCCCGAGA	14160
CATCCGACGG	CAGATCAAAC	GACAACCGGT	ACGTGGCCGT	GTGCTCGTTC	TTGTTGATCA	GTTTCAGATC	CGCACCCGTG	14240
CCCAGGGCGT	CCTTGCCGAA	GGAGGGCGGG	GCTTGCAGAG	TGGCCCGCAC	GGTGATCTTG	CGAGGGCTCT	GTTGCAAAAA	14320
TCGTGAAGCT	TGAGCATGCT	TGGCGGAGAT	TGGACGGACG	GAACGATGAC	GGATTTCAAG	TGGCGCCATT	TCCAGGGTGA	14400

TGTGATCCTG	TGGGCGGTGC	GCTGGTATTG	TCGCTATCCG	ATCAGCTATC	GCGACCTTGA	GGAAATGCTG	GCGGAACGCG	14480
GCATTTTCGGT	CGACCATACG	ACGATCTATC	GCTGGGTCCA	GTGCTACGCC	CCGGAGATGG	AGAAGCGGGT	GCGCTGGTTC	14560
TGGCGGCGTG	GCTTTGATCC	GAGCTGGCGC	CTGGATGAAA	CCTACGTCAA	GGTGCGGGGC	AAGTGGAACCT	ACCTGTACCG	14640
GGCAGTCGAC	AAGCGGGGCG	ACACGATCGA	TTTCTACCTG	TCGCCGACCC	GCAGCGCCAA	GGCAGCGAAG	CGGTTCTTGG	14720
GCAAGGCCCT	GCGAGGCCCTG	AAGCACTGGG	AAAAGCCTGC	CACGCTCAAT	ACCGACAAAG	CGCCGAGCTA	TGGTGCAGCG	14800
ATCACCGAAT	TGAAGCGCGA	AGGAAAGCTG	GACCGGGAGA	CGGCCACCCG	GCAGGTGAAG	TATCTCAATA	ACGTGATCGA	14880
GGCCGATCAC	GGAAAGCTCA	AGATACTGAT	CAAGCCGGTG	CGCGTTTCA	AATCGATCCC	CACGGCCTAT	GCCACGATCA	14960
AGGGATTCGA	AGTCATGCGA	GCCCTGCGCA	AAGGACAGGC	TCGCCCTTGG	TGCCTGCAGC	CCGGCATCAG	GGGCGAGGTG	15040
CGCCTTGTTG	AGAGAGCTTT	TGGCATTGGG	CCCTCGGCGC	TGACGGAGGC	CATGGGCATG	CTCAACCACC	ATTTCTGCAGC	15120
AGCCGCCTGA	TCGGCGCAGA	GCGACAGCCT	ACCTCTGACT	GCCGCCAATC	TTTGCAACAG	AGCCCTCAAT	ACCGACAAAG	15200
CGCCGAGCTA	TGGTGCAGCG	ATCACCGAAT	TGAAGCGCGA	AGGAAAGCTG	GACCGGGAGA	CGGCCACCCG	GCAGGTGAAG	15280
TATCTCAATA	ACGTGATCGA	GGCCGATCAC	GGAAAGCTCA	AGATACTGAT	CAAGCCGGTG	CGCGTTTCA	AATCGATCCC	15360
CACGGCCTAT	GCCACGATCA	AGGGAATTCGA	AGCTTCGCGA	GCCCTGCGCA	AAGGACAGGC	TCGCCCTTGG	TGCTGCAGC	15440
CCGGCATCAG	GGGCGAGGTG	CGCCTTGTTG	AGAGAGCTTT	TGGCATTGGG	CCCTCGGCGC	TGACGGAGGC	CATGGGCATG	15520
CTCAACCACC	ATTTCTGCAGC	AGCCGCCTGA	TCGGCGCAGA	GCGACAGCCT	ACCTCTGACT	GCCGCCAATC	TTTGCAACAG	15600
AGCCCCGCAA	CCGATACGAT	GGCCTCGATC	TCTTCGATGA	GCGCCCGGGG	GCCAAGTACT	GGTACACCGG	TGGCTTCAGC	15680
ATTCCCGGCC	TCGCCCGGGG	GATCATCGGC	GGCGCAGCCG	CCTCCTTCTT	CCTCACCACC	GACTTCTGGA	GCGGCGCGAT	15760
CGCCGAGGCC	ATGGGCTACA	TCGATCTCTC	CGTCCCGTT	GCGATGATCG	TCAGCGGCGG	ACTGTACTAC	CTCTTCACCC	15840
GCCGCGGCCA	CACCTTCGGA	GACGCTCGAT	GAACACGCCC	ACCACCGGCT	CGCACCCCGC	CCGATACCCC	GACGCCGCCG	15920
CGGGTGAACC	GACGCTCGAC	AGCTGGCAGG	AGCCCCCGCA	CAACCCTGCG	GCCTTCGCTC	ACCTCGGCGA	GATGGTGCCC	16000
TCGGCAGCGG	TCTCGCGGCG	CCCGGTGAAC	GCGCCCGGTC	ATGCCCTCGC	GCGGCTCGGC	GCGATCGCGG	CGCAGCTTCC	16080
CGACCTCGAA	CAGCGCCTCG	AACAGACCTA	CACCGACGCC	TTCTCTGTGC	TGCGCGGAAC	GGAGGTCGTG	GCCGAGTACT	16160
ACCGCGCCGG	CTTCGCGCCC	GACGACCTGC	ACCTGCTGAT	GAGCGTCTGT	AAGTCGCTGT	GCGGCACGGT	CTCGGAGCA	16240
CTGGTTCGACG	AGGGGCGCAT	GATCCCGCG	CAGCCCGTCA	CAGAGTATGT	ACCCGAGCTC	GCGGCTCCG	TCTACGACGG	16320
GCCCTCCGTG	CTGCAGGTGC	TCGACATGCA	GATCTCGATC	GACTACAACG	AGGACTACGT	CGATCCCGCA	TCCGAGGTGC	16400
AGACCCACGG	TCGCTCGGCC	GGCTGGCGCA	CGCGCCGCCA	CGGCGACCCG	GCCGACACCT	ACGAGTTTCT	CACCACCTTC	16480
CGCGGCGACG	GCAGCACCGG	CGAGTTCCAG	TACTGCTCGG	CGAACACCGA	CGTACTCGCC	TGGATCGTCG	AACGGGTCAC	16560
CGGTCTGCGC	TACGTGGAAG	CGCTCTCCAC	GTACCTGTGG	GCGAAGCTCG	ACGCCGACCG	CGACGCTACC	ATCACCGTCG	16640
ACACGACGGG	TTTCGGCTTC	GCGCACGGCG	GCGTCTCTTG	CACGGCGCGG	GATCTCGCGC	GCGTGGGGCG	CATGATGCTC	16720
GACGGCGGCG	TCGCCCCCGG	TGGCCGCGTG	GTCTCCGAGG	ACTGGGTGCG	TCGCGTACTC	GCCGGCGGAT	CGCACGAGGC	16800
GATGACCGAC	AAGGGCTTCA	CGAACACCTT	CCCCGACGGC	AGCTACACCC	GCCAGTGGTG	GTGCACGGGC	AACGAGCGCG	16880
GCAACGTGAG	CGGCATCGGC	ATCCACGGGC	AGAATCTGTG	GCTCGACCCG	CTCACCGACT	CCGTGATCGT	CAAGCTGTCT	16960
TCGTGGCCCC	ATCCGACAC	CGAGCACTGG	CACCGGCTGC	AGAACGGGAT	CCTGCTCGAC	GTCAGCCGCG	CGCTCGACGC	17040
AGGTAGGCAA	CGGTGCTCC	GCTCGCGGCG	GGCCTCCCC	GCCGGTTGAG	CGAGGAGCGC	AGCGACGAGT	CGAAACCCGA	17120
CCGCCCCGGG	TTTCGACTCG	CTCCGCTCGC	TCAACCAGCG	GGGGCCCGCT	CAACCGGCGG	GCGCCCGCGG	CTCGCCCGGG	17200
CCGACTGAAT	CGCACCGGGT	CGCGGGCGCT	ACTCCCCCAG	CCGGTTCTTG	CTGCGCATGG	CCCGCTCGGC	CTCGCGCTTG	17280
TCCTGCCGTT	CGCGCAGCGT	CTGCCGCTTG	TCGTATTCCCT	TCTTGCCCTT	CGCGAGGGCG	ATCTCGACCT	TCGCGCGGCC	17360
GTCCAGGAAG	TACAGGCGCA	GCGGCACGAT	CGTCATGCCG	CCCTCGCGGG	TCTTCTGGTA	GAGCTTGTCG	ATCTGCTGGC	17440
GATGCAGCAG	CAGCTTGCGC	TTGCGGCGCG	GCGCGTGGTT	GGTCCACGAC	CCGTGAGGT	ACTCGGGGAT	GTACGCCGCG	17520
TCGAGCCAGG	CCTCGCCGCG	CTCGATGAAG	ACGTAGCCGT	CGACGAGCGA	CGCGCGCCCC	ATGCGCAGCG	ATTTACCTTC	17600
GCTGCCGGTG	AGCACCAAGC	CCGCTCGTGA	GGTGTGAGG	ATGTTGTACT	CGTGCCGGGC	CTTCTTGTTC	GAGGCGATCA	17680
GCTTCTCGCC	GGTCTCCTTG	GGCATGTAC	ATCCTTTCTG	CCGTGCGGCT	GGTGCGGGAT	CCGCGCTCAC	GCGCGAGGAT	17760
CCAGCCTACC	GCACCTTCGC	CGCTAGACCC	GCAGATACCG	CGTGATCGCG	ATCTTCGCGG	CGATCGCCGA	GAGGATCACG	17840
CCCACGAGGG	CTCTGTTGCA	AAAATCGTGA	AGCTTGAGCA	TGCTTGCGCG	AGATTGGACG	GACGGAACGA	TGACGGATTT	17920
CAAGTGGCGC	CATTTCCAGG	GTGATGTGAT	CCTGTGGGCG	GTGCGCTGCT	ATTGTCGCTA	TCCGATCAGC	TATCGCGACC	18000
TTGAGGAAAT	GCTGGCGGAA	CGCGGCATTT	CGGTGACCA	TACGACGATC	TATCGCTGGG	TCCAGTGCTA	CGCCCCGGAG	18080
ATGGAGAAGC	GGCTGCCGTG	GTCTGGCGG	CGTGGCTTTG	ATCCGAGCTG	GCGCCTGGAT	GAACCTACG	TCAAGGTGCG	18160
GGGCAAGTGG	ACCTACCTGT	ACCGGGCAGT	CGACAAGCGG	GGCGACACGA	TCGATTCTTA	CCTGTGCGCG	ACCCGACAGC	18240
CCAAGGCAGC	GAAGCGGTTT	CTGGGCAAGG	CCCTGCGAGG	CCTGAAGCAC	TGGGAAAAGC	CTGCCACGCT	CAATACCGAC	18320
AAAGCGCCGA	GCTATGGTGC	AGCGATCACC	GAATTGAAGC	GCGAAGGAAA	GCTGGACCGG	GAGACGGCCC	ACCGGCAGGT	18400
GAAGTATCTC	AATAACGTGA	TCGAGGCCGA	TCACGGAAAG	CTCAAGATAC	TGATCAAGCC	GGTGCGCGGT	TTCAAATCGA	18480
TCCCCACGGC	CTATGCCACG	ATCAAGGGAT	TCGAAGTCAT	GCGAGCCCTG	CGCAAAGGAC	AGGCTCGCCC	CTGGTGCCTG	18560
CAGCCCGGCA	TCAGGGGCGA	GGTGCGCCTT	GTGGAGAGAG	CTTTTGGCAT	TGGGCCCTCG	GCGCTGACGG	AGGCCATGGG	18640
CATGCTCAAC	CACCATTTCTG	CAGCAGCCGC	CTGATCGGCG	CAGAGCGACA	GCCTACCTCT	GACTGCCGCC	AATCTTTGCA	18720
ACAGAGCCCT	GGGGATCACG	CCGGTATCGT	GCAACGCGCT	GCGCAATCTG	GTTCTTTGCT	TCGTTCTGTC	ATGCTGCGCG	18800
ATCAGACATT	CTCGGACTCC	TTCGCCGTCT	CATACCAGCG	CACACCGCTG	TTGATCGAGT	GCGCCGAGCC	GATCGCCCCG	18880
TATGCTTCCG	CGAGCGCGAT	TGCAGCACTG	TGCAGAAACG	GCACGACGTT	TTCTGGGGCA	TACTTTTCCC	CGTCGAGGCC	18960
GTTGTAGTGG	GTGCGCTCGA	TGACGCCCTG	TGCGCACTGT	GCGAGCTGCT	GGGAGACCTG	CTGCAAGTGA	TCCACGGTGA	19040
GCCACAGTTC	GGCGAGTAGG	TTGTATGACT	CGCTCGGGTT	GCGAACGGTG	CGCGTCGAGT	GCGCGAGCGT	ACGTGTCACT	19120
GCTTCGACCT	CGCGTGCGAT	TTGAGAGTGT	CGCGTTTCTT	CGTTGATAGC	CATTGGTCTG	TCAATCCTTC	GTTGTGGGGT	19200

CGACGCCGAT	TTCATCCGCG	TCTGCGGTCG	GGCGTCCGAG	CGCAGCACTA	GTGACGTCTT	CGATGTTTGT	CACGCCGGTC	19280
GCATCCATGA	GTGCGCGTCG	GGCTTCGGCA	GCGGCGCGT	CCCCCTGCAA	CGTGACGCCG	CCTGGCTTCA	CGGTACAGTT	19360
GCCGCGCTCC	GCGCGCTCGG	CCAGATCCTC	ATATCCGCGG	CCCTTGGCCT	CGGTGTACAG	CGCGTAGAGG	ATCGCGAATT	19440
GCTGGGGGTC	GTGACGAGT	AACGTAGTGA	TGCGCTGCAC	GGTCTCCCTG	GCGATCGGGA	AGCCATCGTT	GAGCAGGTAC	19520
CAGCACGCGC	CGATGGCTTG	TTCTCGTTGC	TGCTCGGTGA	GGGGTTCGAA	CAGCTCGGGC	CAGCGTGCCT	GGAAGTCGGT	19600
ATCGGTGTTT	ATCACGTGAT	CCTTTCGCGG	ACCTCGGACA	GTGGGCGGTT	ATTGAGAGGG	GCGGCCTCGG	GAGATGGAGC	19680
CGGGGTGCAC	GGGCTCGAAC	GGCGCTCCCC	CGTTGAACTC	CCGCTCTAGC	TGTTCCACGT	GCTGCATGAT	GATCTGGTTC	19760
TGCATGTCGG	TGAGGGTCTG	CCATCCGTAC	GTGTCGCGGC	CGCGTAGTA	CGCGGCGCGG	ATGCGGCCCT	GCTGGTCTTC	19840
ATCCGCGTAG	TAATTCATCT	TGCGGCTTCC	CTTTGAGGGC	TGCGCCTTCG	GCCGCGTGCG	GGGCGCTACG	GGCGCAGCGG	19920
GTGCTGGAC	TGACGAGATC	TGCGCCGCTG	GCTCTGACGC	CGGGTGAGTG	GTTACTGGCG	TGAGCCCTGC	CAGGCTCGAT	20000
GTGCGCTCT	TGCGTTCAGT	GCTCATCGTT	CGTTCCTTCC	GGTCAGTTCG	GCGATGTAAC	CGGCGTAGAT	CTTCGCGAGT	20080
TGCGCGCGGT	CGGTGTCGCC	TTCATCGAGG	CCGCGTGATT	CTTCGGTGGC	GTGCGCGATC	GCCTGGCGCT	TGGGTACTGG	20160
CGGCGCGAGC	AGCGGCAGGC	CTCGGGTTTC	CGCTGCGGCC	GTGATTTCCT	CGGCCCAATG	GGTTCCGCGG	ATGGTGGCAG	20240
CTTCGTGCTG	GTTGAGCAGC	ATTCTGCGA	CGCGCAAGCG	GGGGTTGTAT	CTCTGGTTGA	CTACCGTGAT	GGTATCGAGC	20320
AGCTTTTGGC	GGCCGTTGGT	CGACGAGAGC	TTGGTCTGAG	TGACACGAGG	TGCGCCGTCA	GCGCGCTGTA	GAGCGTTGGT	20400
GACGAGCTGA	TGAGGGCTCG	GCCACAGTAC	GATCAAGATC	AGGTGCTAGC	GGCCGCGTAC	GGGCTCTAGC	ACTTTCCGCA	20480
TGCGTGCTTC	GCGGCCGCGG	CCGCGCATCA	CGAGCTGGTC	GCGGACAATA	CCGAGTGTTT	CGCCTGTGGT	GGGCACGACC	20560
TGACGCGCTT	CCCAGATTGC	TGGCAGCATG	ACCTCTTCGA	GGGTTTCGCT	GGCGTATTTC	CTGAGCACGT	CAGCGAGGCC	20640
GGGTTCTTCC	TCCCCGACAC	CCTCGGTGAG	CACCTGGGTG	ATGTTGCCCT	GGGGTCCGCG	GTGACGACG	AGCAGCGGCA	20720
GGCCGTGCTT	CACCGCGCGG	CGGGTGAGGT	GGTACAGCGT	GGTTCGATTG	CCACGCGCGC	CTTCTGTGGT	GCAGGCGGTC	20800
AGGATGACGG	CCTCACGGGT	TGCGCTCATG	GTTCCAACTC	TCTATCTCTA	ATGCTCTCTA	TAGACCTATT	TATTGGGAGG	20880
GCCGCGGGTT	GCGGCCCTCC	CTGGGGGCGA	TCGTGAGTCG	CCCTGTCCGA	TCTTCGAAAC	GGTGACGCGC	TGGTTTCGGA	20960
GGCTGACGCC	TACCTCATCT	GCTCGCACCT	CGTGCTGCAC	GCGCTCCTCG	CCGCCCTCGG	GCTGTGACGA	GGTGACGCGG	21040
TAATCGCCGG	AGAAGGTGAC	GCGAATGTTG	CCGGAGATCT	CGGCTGCCTC	GACGAGCTGT	TCCGCCTGCT	GGCCGGTCAC	21120
GTTCACTGAG	TAGAAGATCG	TGGGGCCGTC	GTACCATTCG	CCGGTGTCTC	TGTCCCTGAG	TCGATCAGAG	ACGGCGACGC	21200
GGGCGTAGGT	GTACGGCGTG	TCCTTGTCTG	TGCGGCGAAG	GGTTGGGGCT	TCGACGAGGT	TTCTGTGTGG	GGTGATGTAG	21280
CTCATTCGCG	TTCTTTTCGT	GTGCGGGTAC	TGGTCGATGT	AGTTCTATTA	TACCTAATGC	CTCTATTAGG	GACAATAGAG	21360
GCATTAGTGT	TTCAGGATTC	GTTGCGGAGC	ATGATGGTTA	TGACGGGTTT	TGCGGTGTCT	CCGGGGCCGA	TGTGCAGAGC	21440
CAGCTCGATC	GGCTCGGGCT	GGTTCCAATC	GTGGCCGCGG	TTCTCGTTGG	GCACGCGCAG	CACCTGGAAG	GTACGCGGCT	21520
TACTGCCGGC	CGTTCTCGAT	GCGGCATGGC	CTGCGAGGTA	GAGAATATCC	CAGAGCCGGC	CGCTTTCGTC	CTGCGGCTCG	21600
GGCCTGTCTT	CATCCACAGC	GACCACCTCG	GCCCACGCTT	CTCGCGTGAG	CGCGACGGGT	ACACGGAATC	GCGGCTCTTG	21680
CGCCATCTCG	GTAACGTGCA	CGAGCGCTCC	GTCATCGAGC	GCCTGGGCTC	GGGTGTAGCT	AGCGATAATG	GGGCCGAAGG	21760
CCTCGTGTGC	AGTCATGGTG	TCTCCATTGT	CTCTATTAGG	GATAATGCCT	CTATTGATGG	CTGGGTTACT	CGCCGAGGCG	21840
TGCGGCGAGC	TGTTGATGTA	GGTCGATTTC	GAAGCCGGCC	CAGTGCTCGT	GCTCGTTGAC	GAATACGACG	GGGGCGCTGA	21920
GGTAGCCGAG	CTGCTTCACG	TACTCCGTGG	CGGTGTCGTT	CTCGCTGAGG	TCGACGACGT	CGTACAGGAT	GCCGTTCCGC	22000
TCGAGTGCGC	GGTAGCTGCG	GGTGCACTGC	ACGCATCCGG	GCTTGGTGTA	GACCGTCAGG	CTGTGCGTGT	TCATGGTGCT	22080
CATTCTCTCT	TCACCCCTCA	AATTTTTTTG	CCGTTGATGG	CACCTGTGTC	CAGCGCGAGC	GCGGCAGCGG	AGCGCAACCG	22160
TAGTGACGCA	CCCGCACGGC	CGAGCGCAGC	GAGAGACGGG	CGAGAGCGTC	GCGTCAAGGA	AGGCCGCGCA	GCGGCCGACC	22240
GGACGGATGC	GTGAAGCAGA	GCGGTGCGTA	TGCTGCCGAA	GGCATCAACG	GCAATAAATT	GGGCGCTTGC	GCCTACTGCG	22320
GCCGGCTTGC	GCGCCGTCGT	CATGTTCCGG	CCGCTCCGTC	GGCTCGAACT	GGGGGAGTGC	GAGGGGGTGC	CACCTTCGCA	22400
CCGCGCCCCG	TGAGGGGCGC	GGTTTCGGTT	CACCGGGTGG	CCCATGTCCA	GGCGGCCTCG	ACGAGTGCGG	ATAGGTCTGC	22480
GTGCGTGTGC	TGCGCCCTCG	CGTGGGGGTT	CCAGCTCTCT	GCTTGGTAGC	GCTCGGCGCA	CCAGTGCCAG	CCGATGACTT	22560
CGCGGGTGTC	GTGCTCCTCG	GGCTTGTGCG	TGCGCTCGGT	GATCGCGAAT	CGGCGGGTCT	TGTCGTGCGC	GCGGTAGAGC	22640
ACCAGGTGCC	CGGCACCGAA	TGCGCTGGGG	TCTTCTCTGT	TCAGCTCGAC	CACGCGGACC	GATTGCGTGA	TGGTGGCCAG	22720
CTCGTAGCGG	ATGCCCTGGC	CTGCGCTGCG	GTGGTGGCTG	TCGCTGAGT	CGTGGGTGTA	GCTGTGCTGG	TTGATCCACT	22800
CGGCCGCGAG	GTGCGGCGCG	TCGTGTGCGG	TGAGCAGATC	CGCGATCGCC	TCGGTTTCGC	CGCATGTGAG	TCTGGGGGCG	22880
AGGTGCGCGG	CCATCCCTTC	ATCCAGTAG	GCGATGAATG	GTTGAGTTCG	CTGTTGCTGG	TCGGTGTTGT	TCATCTCTAA	22960
TGCCCTCATT	GTCTCTATTG	GGGATAATGC	CTCTATTAGT	TGCTGGTTTC	GTTGTCTCGG	GTGAGCGTGT	CAACGAACGT	23040
TGTGAAGTCG	GGGCCGCTAA	CGGTTGCGGT	GTGCGCAAGC	GCGTATCCGG	CGCGGGCGTG	GTGGGCGAGG	TCTTCTTCGC	23120
GCTCGCCGAG	CTTGCGCAAC	TGCGCTGAA	TCTCGGTGTC	GTCGACGACG	GGCACGAACG	TGGTGATGGT	GGTGATGCTC	23200
TGCATGGGTT	CTCTTCTCTT	CATTTTTTTG	CCGTTGATGG	CACGTAGTGC	CACAGCGAGC	GCGGCAGCGG	AGTGACAGGAG	23280
TAGTGCTGGC	CCTGCGCGGC	CGAGCGGAGC	GAGAGACGGG	CGAGGCCGGC	GCGTCAAGGA	AGGCCGCGCA	GCGGCTGCCC	23360
GGACGGATGC	GCGTAGCGGA	GCGGTTGCTA	CAGTGCCGAA	GGCATCAACG	GCAATAAATT	GGGCGCTTGC	GCCTGCTGCG	23440
GCCGGCTTGC	GCGCCGCGGT	CTTGCTCGGG	CCGCTCCGTC	GGTTCGAGCT	GGGGGAGCGC	GAGGGGGTGC	CACCTTCGCG	23520
TCGCGCCCCG	GCGGGGTGCG	ACTGGGGCCA	CGCTCTGCGC	GCGGCCCTTT	ATGCCTCTAT	TGGGGATAAT	GCCTCTATT	23600
TGCTGCGGGG	CGGTGGTCTT	GCACTTCGGC	CCAGGCGTAG	GCCAGATCCC	ATCCGGCTTC	AACCTGGATG	CGCACGGTCT	23680
TGGGGTTGGC	GCGCTGCACG	AGTGACCAGG	TGCCGCGGAT	CTTACGCGCG	TGCGCTTTCT	TGACGGTTTC	GCGGCTGTAG	23760
TTTGTGGCCT	GGCCGGTGGC	GATCTGGTTC	GCGCGGATCT	GCTGCCAGTA	GGCGAGCTGG	TCGCGCAGCT	CGTCGAGCTG	23840
GGGCGCGTAG	TGCTCGGCGC	GCTTGACCTG	CTGCTCTGCG	GTGGCGGGCT	TGTATCCGGT	TTGCGGGTGC	TAGACGTGCG	23920
CCGAGATGCT	GCGCTCGACT	CGGCGAATGT	CGGCGGTGAT	TTTCTCGATG	CGGTTCCGCA	CGGTCACGGG	CTTGTAGCGG	24000

GCGCTCGTGG	TGGCGGCCGC	GACGCGGGCG	GGCCGTGCGG	TTTCCTTGCC	CTCGGCGTCG	GCTGCGATCG	CCTGCCCCAT	24080
GCGGCGGTGT	GCCTTCTCGA	TGGCGTTGCG	GTGGCGCTGC	TCGGAGTGGT	GGCCGATCTT	GATGGGCTCG	CCACCCCTCG	24160
GGAGCGCGTT	ACCGGCCGCG	TCTGCCCGGT	TCCATGCCTG	GTGTGCGTCT	GCGCTCTTGC	GCAGCGCCTT	CGCTTCGAGC	24240
GCGTCGCGCG	GGTTGGCCTG	GCGCATGGCC	TTGCCTGCTT	CGACTTCTCC	GGTGCTGCGG	GTGGTGCGGT	CGATTTTCGAG	24320
GGCCAGCGCG	TAGCCTGCGG	TTTCGAGCGC	TGAGGCGGTG	CGGGTGATGA	CGTAATCCTT	GGGCGAGCGG	TCGCGGCTGT	24400
TCGGGACGTA	CCAGGCGGAG	ATGCTGCGGC	CCCAGCGCCA	CCCATTGTGT	TTGAGGATCT	GCGCGGTGCC	GTGCGCGCGC	24480
TCGGTGCCCT	CGATGATGGT	TCCGGCTTCG	TGCGTGTTGG	TGATGGTCAG	CATGTCAATG	TCTCCATTGT	CTCTATTGGG	24560
GATAATGCCT	CTATTTGGTG	GGCTTGTTCG	GTGTTCCCGG	TTCTCCCTCG	ATTTCTTTT	GCCGTTGATG	GCACCTGTG	24640
CCAGCGCGAG	CGCGGCAGCG	GAGCGCAACC	GTAGTGACGC	ACCCGACGCG	CCGAGCGCAG	CGAGAGACGG	GCGAGAGCGT	24720
CGCGTCAAGG	AAGGCCGCGC	AGCGGCCGAC	CGGACGGATG	CGTGAAGCAG	AGCGGTGCGT	ATGCTGCCGA	AGGCATCAAC	24800
GGCAAAGAA	TGAGCGCGTC	AGCGCGGGGG	AGTGCGGCGG	CGTAGCGGCC	TGCTCGAACT	CGGCCGCGCT	CGGTGCGCGG	24880
CCGTGCGCGG	CGGATGCGGC	CGGTTACGCT	CGACGAGCTG	CGGGAGGCCT	GCGGCCGACC	ATGACGGTGC	TCGGCCGTCC	24960
ACACTGGGTG	CAGGTCAGCC	CTGCGATAGC	TGAGCTGTGC	CCATTGTGCG	GCGGGCGAGC	CGAATAGTCG	TCGTGCGCAG	25040
GTGGTTCGGT	GCGTGACGCT	GCTCCACGTG	TGGCGCAGCA	CAACGGGTGC	GGTGGGGCTG	CCGCGCGATC	CGCAAGCGATC	25120
ACTTGCCCTC	GAATGGGGGT	TTGGTTGCGG	TGGAGCTGTG	CAGAGCAGCT	ATGCGGCGAG	GTGCTAGAAC	TGGGGGAGTA	25200
GGCCGGTTCG	GATGGCGATT	TTGGCGGCGT	CGTGATCGGG	GCGGCCGTCG	GGAGTGCGGG	GGTAGGCGGG	ATAGCGATCC	25280
CATGCGCGGC	CGAGTACGAA	TGCTTGCTGT	TCGTGTCGGG	CTGGGCGTTT	GTTCCCATTC	TTGCCGGGGC	TGCTGAGTCT	25360
GAGCACTTCT	GCTTGCCACC	ATCCCCAATG	CACGCGCTCG	ACGGTGTAAG	CGTCGGCGCG	TGCAGCCAGG	CGGCCGTCGA	25440
CGCCGAGCTG	GACGGCTACG	CGGTGCGGCT	CGTCCGGGTC	TGTCGACGTC	CAGCCGGCAC	GAGTGAAGTG	TACGAGTCCG	25520
TGAGCTGCGA	GCGGTTGGAG	GGCGTGTGTA	GCGCTGAGCA	TCGTTTGAGC	GGTGGCCAGG	GCGACGCTCT	CAAGTGTGAG	25600
GGGGAGCGAG	AGTCGAGCGT	AGGTGTTTTCC	AGCGTGTAAG	CCGAGAGCCC	TAGTGACAGT	GAAGACGTCG	TGGCGTGCGC	25680
GCTGCAATCG	GGTGGTCAGG	GTGCTGATGA	GATAGTCCCG	GTGGGCGGCC	GCCGCCCTTG	CGGGGCGAGC	GCCCGCTTGT	25760
GACCGGGCAG	TTTCCGAGTC	GCTGTGGATA	GCGTTTTGGG	GGTCGATGGT	CCAGTGCGCT	CCGTTGGGTC	CGTCGGCGCG	25840
GCGGGTACGG	GTGATCCATC	CGCTTTCGCG	GAGGCGCAGC	AGCGCCGTAG	GCGCGGTTTC	TCGGCCGACG	CCTGCTAGGT	25920
CACCAAGGCG	GCGGGTATCG	GCCTCCACAG	AGGCGCTGAC	GGCCTGTAGC	GCGAGGAGGC	AGAGGGTATC	GAAGATGCGA	26000
CGGTCTGCGG	GGCCGCCGCC	GCGCGTCCAG	CGTCTACGCG	AGGCGTCGGC	GCGGGTCTGC	ACCTGGTCTGA	CGTGCTCGGC	26080
GATCGCCTGG	GCGCGGGGCT	CGAACGTAGG	GTCTTCTCCG	ACCTGGCGGG	GCGTCGTGGC	CACGCGGCGC	ACTGCGCGGG	26160
TCCATTGTGC	GCGGAGAACG	CCTGTGAGG	AGTGCGAGCC	GTGCGAGGGG	CGGGGCGGGC	GCTGCGCGGA	GCCCTTGCTG	26240
CGACGCTCGG	TGCGGGCGTG	CTCTAGGCCG	GGTGCGGTGG	GTAGGAGCGC	TGCGACGTCG	GCGTAGCGCC	ACCTGGCGGC	26320
CGCGGCACCG	ATGAGCACGC	GCCAGAGCAC	GGCTGAGGCG	TCGGTGTCGC	GGGTGACGTC	CTGGTGACAG	GCTTCGGCGC	26400
TTCCGCTGG	AAGGGGCCCG	CGTGGGCCCG	GGAGGTACAC	CTGCCCGGTC	TCGTGACGCG	GCGAGCGGGC	GGCTTCCGAG	26480
CGGGCAGCGG	CGCTGCTGTC	GGTCGGGTAG	TCGGTGTTGA	GCTGTGAGGC	GAGCTGCACG	AGCTGCGCGG	GGGTGGTCTG	26560
GGCGGTGGTG	AGTGCGGCGA	CGTCGCCGCT	GAGCACGCGG	GAGTGCCCGC	CGAGTCGGTG	GGGTGCTCCT	GGGGGCCGCA	26640
CTGAGCCCGT	CGACGGGTTG	CTGATGGGCG	TGCGGTCCAG	GCTCGGGCAG	GCGCGTGCGA	ATGAGATGCC	GAGGTCTTTT	26720
ACGAGCTGCG	CGTCTGCGGG	CTCTGCGAGC	GCGATCCATC	CGTGCGGGCG	GCCGGTCGCG	CCGGAATCGC	AAATCACGTG	26800
CTCGATGCCG	AACCGTGTGA	GGTATTCGGA	GATAGTCCGG	GCGTCTTCGG	GCGCCGCGGC	CGAGCCTGCT	TTCCGCTCGA	26880
GGTCGAATGC	GATCAGGTGG	AATCGGCCGT	CGCGGTGCGG	GAGCATGACT	GCCACCGGTC	CTACGGGGGC	AGGGCCAGTG	26960
ATCGCCCGCG	TGGAGGGATA	GAGGTTCAGT	GTGCTGCCCT	CGACGTCTTC	GACGGCGACG	GCGACGCGGG	GCGCGGGGCT	27040
CAGCTCAGCG	GTGAGCGCCC	ACGCCACGCC	GATAGACGTA	CATATAGGAC	TAAGCGGTTC	TGATGGTACG	ATGGTGCCAC	27120
CTCCTTCAAG	GCAGGCAAAA	CCCCTCGGGG	CTTTTGATTG	ACCTGAACCG	TTCCGGTCGCC	AAACTTCTGC	GGCTCAGGTA	27200
CTGACTTCAG	ATCCCCCAC	TCCCCGAGT	GGGGGATCA	CTCATTTACG	CGGTGTCGAT	AGGCAACTCC	CCGTCTTTCA	27280
GCTCTGGTTC	GGGCCACAC	GCTGGCACCC	CACGATCAGT	AAGCGGCGTG	TTCCGCGACCA	TCTGCTCGAA	CATCACTGCT	27360
GACGACACAT	TGGAGTGCTC	AGCGATGCGG	TCCAGTCCGT	CTTTGCTGTA	GCGCTCGATG	ACGTACCCGA	TCGAAACCGG	27440
GTGACCATTT	GAGCCTCGGG	GCCGTGACG	AACTGCCATG	CTCGAAACGA	TAGCACTGTT	CCTAGGATTA	ACTCAGCGAC	27520
ACGCTACGAA	GTTGATTTCG	GCATGTGACG	TCGCCGAGCA	CAAGCGTGGT	TCCTGGTTGG	ACGACGTAGG	GGCGTCCTGG	27600
CTGGAAGATC	ATGGGCGGCT	CGGTGGCGGT	CTGTACGCCG	TTGCCGCTGT	GGTGGTCCGT	GACGAGGATC	TGGCCTGCGT	27680
TGTCGACGCT	GACGAGTGGC	TGGGTGCGGG	AGAGCATGCG	GCCGGGGCTG	TGTACGGCGA	TCGGCCGGGC	GCCGTGACG	27760
GGCTCGGGGC	GTCGAGTCAT	TGCGCGGCTG	TCCGAGTCCG	CGGGTTCGTG	GGCCGTGCTG	AGGAAGAGGC	GACGCTCCGG	27840
CCGGGTGGCC	GGCGCTGCGG	TGCCCGTGGG	CGGCGTCTGC	TCGATAGCGT	GCTCGGCGGG	CGGGCCGACT	ACCTGCACGT	27920
CGACGCGATG	CGGGCCGCGG	ATGCCGCACT	GCTCGCACTT	TGCCGCGGTA	ATGGGCTGGG	TGTGCCGGCA	GACTCGGCAG	28000
GGGCCGTTGT	GGACGGTGAG	CCCGTCTGCG	GGCGGGATCT	TGCCGGTGTC	ATCGATGAGC	TGTACGTAGC	CCTCGGGCCG	28080
GATGGCGAGT	TGCCAGGTCT	GGGTGTCGTC	GGTGACCTCG	AACCGTACGG	GGCGGCGGAG	CTGCTGCGCG	ATGACTGCGC	28160
AGCGGGCGAT	CATGCCGGTG	GCGAGCTCGT	CGACGCTGGC	AGCGTGCCAG	AGTCGATCGG	TCTTGTGATG	GGTGAGCACC	28240
CCGGACCCAT	CGGGGCGGAC	GGTCGCGCTG	ATGCGCGGCC	AGTCCGCTAC	CCGGTGGGTT	TTATCGGGGG	TGGTCATAGT	28320
CCGCGTGCGA	GCGCAGCGCC	GGCGGCGAGC	CAGGCGCGCT	GTGTGCGGGG	TCGGAGGGAG	CCGTAGCGCA	GCACCCCATC	28400
GACCATTTGCC	GGGTGCTGGG	GGATCGTGAT	GGCGTCGCGT	GCGATCGCGC	GGTAGCCGTC	TGCGACACGG	GCGGCGTTGT	28480
TGGCCTTGGG	GTCGGCCTGG	TTACAGACGA	CAACGGCGTT	TTTGGCGAGC	TGCGCGGTGC	GCTCGTCGCG	CTCGGCGAGG	28560
GCTTCGAGGA	GAAGCGCGCC	GGCCTCTGCG	TGATCGTTCG	GAGTCGTAGT	CGCGACCACG	AGCTGGTCCG	TGTGGTTCGAT	28640
CATGCGCCGC	CACATCGGGT	CTGACTCGTC	GTTGCCGCTG	TCGATAAAGA	TCAGCCGGTA	ATACTTGCTG	GCCACCTGGT	28720
GAATGCGGTC	GACGTCGCGG	GCGTCGACGC	GCTGCGCGCT	CGCCAGCTCC	ATCGGCTTGG	AGCGCAGCAC	GTGCTAGCGG	28800

TCGCGCGTCT	GGTGGTGCAC	GAAGTGGGCG	AGGTCGGGCTG	ACTGAGCACC	CGTGCTCAGC	AGCCGGTCGG	TCTGGGGGAG	28880
AAGGTCAAGC	AACGTGGCTT	CGTGC GGCCC	CTGCTCGGTC	CGCCATCCGA	GCGTGCCCTT	CGTCTGGTTG	TTATCCACG	28960
TGAGGACGCC	GGCACCGCCG	TAGCGAGCGA	ACACGGCGGC	CATGAGGACC	GTTGTTGGGG	GCTTCGATGC	TCCGCCCTTG	29040
CCGTTGACGA	TGGCGATTGT	GCGGGGGCCG	GGCCAGTGCT	GGCTTACCTC	GTGGACGTCG	GCGCGCTCGG	GCGGCTCGGC	29120
GTCGCTCGGG	CTCATGCGGA	GCCCCACGCG	GGAGGGCTCT	GTTGCAAAGA	TTGGCGGCAG	TCAGAGGTAG	GCTGTGCTCT	29200
TGCGCCGATC	AGGCGGCTGC	TGCGAAATGG	TGGTTGAGCA	TGCCCATGGC	CTCCGTCAGC	GCCGAGGGCC	CAATGCCAAA	29280
AGCTCTCTCC	ACAAGGCGCA	CCTCGCCCTT	GATGCCGGGC	TGCAGGCACC	AGGGGCGAGC	CTGTCTTTTG	GCGAGGGCTC	29360
GCATGACTTC	GAATCCCTTG	ATCGTGGCAT	AGGCCGTGGG	GATCGATTTC	AAACCGCGCA	CCGGCTTGAT	CAGTATCTTG	29440
AGCTTTCCGT	GATCGGCTTC	GATCACGTTA	TTGAGATACT	TCACCTGCCG	GTGGGCCGTC	TCCCGGTCCA	GCTTTCCTTC	29520
GCGCTTCAAT	TCGGTGATCG	CTGCACCATA	GCTCGGCGCT	TTGTCCGTAT	TGAGCGTGGC	AGGCTTTTCC	CAGTGCTTCA	29600
GGCCTCGCAG	GGCCTTGCCC	AGGAACCGCT	TCGCTGCCTT	GGCGCTGCGG	GTCGGCGACA	GGTAGAAATC	GATCGTGTCT	29680
CCCCGCTTGT	CGACTGCCCG	GTACAGGTAG	GTCCACTTGC	CCCGCACCTT	GACGTAGGTT	TCATCCAGGC	GCCAGCTCGG	29760
ATCAAAGCCA	CGCCGCCAGA	ACCAGCGCAG	CCGCTTCTCC	ATCTCCGGGG	CGTAGCACTG	GACCCAGCGA	TAGATCGTCT	29840
TATGGTCGAC	CGAAATGCCG	CGTTCGCCCA	GCATTTCTCT	AAGGTCGCGA	TAGCTGATCG	GATAGCGACA	ATACCGACGC	29920
ACCGCCACAC	GGATCACATC	ACCCTGGAAA	TGGCGCCACT	TGAAAATCCGT	CATCGTTCCG	TCCGTCCAAT	CTCCGCCAAG	30000
CATGCTCAAG	CTTACAGATT	TTTGCAACAG	AGCCGTCTGA	CGAATGGGCC	GAGGGTCGTC	GTTACCTCGG	TCTTGACGTT	30080
CTCGCGAGGT	GCCGCTCTAC	GCTCGTGGGT	ACCGACACTA	TCGGCCAGCA	CCATGAGCTC	CCCGAGCTCG	GGGCCGCGAG	30160
ATGAGCGCCC	CCACTCTCTG	GCTGTACTCC	GACACCGTGA	GTGAGCTCGT	CGCGAGCATC	TGCAATCACA	ACGAAAACCTA	30240
GCCCCACAT	CCCTGAAGGA	TCACGAAACG	GTTACACCAC	TCCACGGGAC	TTGACCCGCC	GGAGCTGTCT	CGGTAGGTTA	30320
TTGTCCACGC	GGCTTAGACC	AGCAGAAGGA	GCGCGACAGT	GAGCAAGGTG	GACCTTTGGC	AGGATGCGAC	GGCCAGGGCC	30400
GAACTCGTGC	GCTCGGGAGA	GATAAGCCGG	ACCGAACTGC	TGGAGGCGAC	GATCGCTCAT	GTCCAAGCTG	TGAACCCAGA	30480
GATCAATGCC	GTCATCATCC	CGCTCTTTGA	GAAGGCGCGG	CGTGAGAGCG	AGTTGGCCTC	CGGGCCCTTC	GCTGGCGTGC	30560
CCTATCTTCT	GAAGGACCTC	ACCGTCGTGA	GCCAAGGCGA	TATCAACACG	TCCAGCATCA	AAGGCATGAA	AGAGTCCGGC	30640
TACAGAGCCG	ACCACGATGC	CTATTTCTGT	CAGCGGATGC	GCGCGGCGGG	GTTCTGTTTG	CTCGGCAAAA	CGAATACACC	30720
GGAGATGGGC	AATCAGGTAA	CGACGGAGCC	CGAAGCGTGC	GGCGCCACTC	GGAAACCTTG	GAACCTCGGA	CGGTCGGTTG	30800
GCGGATGGAG	CGGCGGCTCA	GGGGCTGCGG	TAGCCGCGGC	TCTGTACCCG	GTTGCGCACG	GCAATGACGC	GGCAGGTTCC	30880
GTGCCCATAC	CTGCCAGCGT	CTGCGGCGTC	GTCGGCCTCA	AGCCGACGCG	TGGGCGGATA	TCCCCAGGCC	CCTTGGTCAC	30960
TGATTGAGAC	AACGTCGAG	GAGCCGCCCA	TGAAGGCCTC	TTTGCGCGAA	GCGTGCGGGA	CATCGCGGCG	TTACTGGACG	31040
TCGTCAGCGG	CCATCGGCCG	GGCGACACGT	TCTGCGCACC	GACAGCGTCA	CGGCCCTACG	CCCAGGGCAT	AAGCGAGAAT	31120
CCGGGCGAGT	TCCGCGTGGG	GGTCTTGACC	CACAACCCCG	TTGGCGACTT	CGCTCTGGAT	CCGGAATGCG	CCGCGGCGGC	31200
ACGGGGCGCC	GCTGCGGCGC	TTGCCGCGCT	CGGCCATGAC	GTGAACGACG	CCTACCCAGA	GCGCTCGGCG	GATCGGTCTT	31280
TCCTCAAGGA	CTACTCGACG	ATTTGCGACG	TCGCGATCGC	GCGAGAGATC	GAACGCAACG	GAGAACTGAT	CGGCCGGCCG	31360
CTGACGGAGG	ATGACGTTGA	GTGGACCAGT	TGGGAAATGG	TCAAACGCGC	AGACCAAGTC	ACGGGGCGCG	CCTTCGCCGC	31440
GTGCGTGAGC	GAGCTTCGAT	ACTACGCTGG	GAAGGTGGAG	CGTTGGTGGG	AGGCGGGTTG	GGACCTCCTG	ATCTTGCCGA	31520
CTGTAACGCG	TCAGACGCCG	GAAATTGGGG	AACTCATGCT	CGCCAAGGGA	ACCGACTTGG	AAGGGCGTCA	GTCCGCCTTT	31600
ATCTCGGGGC	GTCTGCAGAT	GCTGGCCTTC	ACCGTTCCCT	TCAACGTTTC	GGGGCAGCCG	GCTATCTCCC	TCCCCATCGG	31680
CATGCTTAGC	GATGGGATGC	CGATTGGGGT	GCAGATCGTC	GCGGCTTATG	GACGCGAGGA	CCTCCTCCTT	CAGGTGGCCG	31760
CGCAGTTGGA	AGGGGCGCTG	CCATGGGTCT	CTCGGCGGCC	CCAAATTGCTG	AACCCGTCGC	GGAAGATCCC	AGCGGCCTGA	31840
TGCTTGCGGG	AGATGGGGTC	CGGCCGACAG	ATACGGGCGC	GGCCAGGTTT	TCGTCCCCAA	CTGGCACCCC	ACACCGCCTT	31920
AGTGAAGGTC	CGCCTCGGGG	ACGCCCCACA	GCCGGATTGA	GGCCACTCTT	TTCTGTGCTG	GAACCCCGGC	CTGTCGTCTG	32000
CTGCTTGTTA	GCTCGCACCA	CGGTCGAGCT	TAGCATCAGG	TACTGAGTGT	CTCAATCCTT	CGTCAGTTCT	TCGATCAGAC	32080
GCTCCACAC	CCGGCCGCCG	CCCACCGACT	GAAGAGCTTG	CCGTATTCCA	CTTGCCGTAC	CGCTCGAGCA	TGCGCCGGCC	32160
AGATCGAGCC	GGTCGTGGGG	CTCTTCCAGT	CGACGGCAGG	CAAGAGGGGG	CTGGCGCGCA	TGGTGCCCCA	GGGTTACAGC	32240
GATCCACACG	TTGCGGGCCT	CCCGGCTCCA	TCTTTCTCTC	GTGCGGTGGG	CGTTGCGGGG	CCTGCTCGCT	CGAAGGGGTC	32320
GGTCGAGCCC	TGGGCTCGGC	CGAGGGGGAG	CATTGTATGG	GGGAATTCTT	GGTCAAGGTC	CGACACCTGC	CGGGCCGATC	32400
AGCGATCACA	GGATCGGGCA	CGTCGCCGAT	CGCCTTCGTT	CCGGGGTGTC	CTTCTGGGAC	GCCGACCTCC	AGCGCACCCG	32480
CGCCTACGAC	GATGAGGTCT	ACGAAATGCT	GCGGAAGTTC	CCCGGCCGCT	GAGACGCTGC	GTCTTCGAAA	GGGGAAGGGA	32560
GGGCGAGGGT	TAGGTGAGTT	TGGCTGCTTG	GCTGAGCTTC	GTTTGATGAT	CTCCCCACCT	TGAGGCCGCT	TGAGGCCGCT	32640
GAGCGCGAGC	CTTCTTTCTT	CGAGCCCCAC	CAGCCCGGAG	ACGCGCTGAC	CTACCGCGAA	GAGTTTCCCG	ATTTCTGACTT	32720
CGACGTGCCG	TCCGTCTGTC	CCGGCGCTTG	GGACTTCGAA	GACACCTCCT	GGCGCCAGGA	CGTCTGCCCC	TCATTCTGCT	32800
GCGACGTGTT	CCGCCGTGTT	ATCGACTTCA	CCGATCCCCA	GAAGCGCGAG	CGGCCACGCG	GCCCAGCGGT	CCGCAGTTCT	32880
TCATGGTCTG	CGAGGACGTG	ACGATGCTCG	AAAGCAACAG	CTGGGACGAC	GTCGTGGCCT	TTGTGCAAGA	CGGCAAGCGT	32960
CACTTTCCCA	ACTTCCCTCA	ACCGAAGGGC	CCCGCATGAG	CGATCGAATT	CTGATCATCG	AGCCTTGAC	CCCCGACACC	33040
CCGGTCCGCT	CCAGCGACGA	CGCCTGCGGA	TGGGGCTGCG	TCGCTAGCGA	GCTTCGCCCC	ATCACCGAGG	GCGTCTGCG	33120
GCCCGGGGAG	CCCGTGCCGG	CCGGCGACTG	CCCCAAGTGC	GGGTGCGACG	CCTTCCCGAT	CCGTGCGGCG	GACCGGGTGC	33200
GCGCCAAGGC	GCGCGTATTC	GCCGCGGTCT	CTGTTGAGAT	TATTTTCGGA	GAAGGCACGG	ATAGGTCGAC	CCAGCGCGCC	33280
ATCGACCTGT	CGCGCGAGGC	CTTCAAGCC	GCCGCGATCA	CACGCGGCGT	CGCAGAGAGC	CTCGACCTTC	ACCCCCAGCC	33360
CGCCGAGGCG	GCGTGATGGA	GACGAAACCA	ATGATCCGCG	AACTCACAGC	GGAAGTCCGT	TTGCAATCCG	GCATCGACGC	33440
CCTGGTCGCC	GGAAAGCGGT	ACGCCTTCCT	AGCCATCGTC	GGCGACCGCT	ACGTCTGCTG	CCTTGGCGCG	GCGGTGGCTG	33520
ATGAACCGGG	CTATTATCCC	ATTCCCAGCA	TTGGGCGCAT	TCCGACAGCA	TGCTTGAGAT	GCAAAAGCAC	GCTGATGCGC	33600

TCAACGCCGA	ACTCGGCCTC	AGCGCTCGGG	CTGCCGGGGA	GCTTGTGCGC	ACGACGTTTC	CGGCCGGGAG	GGTGAACATC	33680
CCCCAGCCTG	ACGCCGCGAG	CCTGCGAGCG	CTTCCGGGCG	GCAAGGCAGT	AATTTGCGCG	CGTAACGAAA	CCCTCGAAGC	33760
GATCCTGAAA	AATCCCCCGC	CGCGCGGGCG	TGTCATACGC	TTACACGGAG	GCAAGCTGCG	CCCACGACCG	CCGGCTGAGC	33840
GCGGGATGCC	TGCATGGCTA	GCCAGCAACA	ACACCGGCGT	TCTAGTCTCT	GTCCGTAAGT	CGTCCGGCTC	GATCAGGCTG	33920
CGGATGGAGT	CATGGAACGC	GGCGGGCGGT	GACGATTACT	ACCGGACCTT	CGATGTGGCT	TCACCGTTGA	CGTTCTACGT	34000
GGAGACGATC	CCCGTGCCCC	GCGAACTCAT	CGACCTCGAC	CTCTTGCCCT	CGCGGCTTCG	CTCCTTGGAG	CGTCGCGCCC	34080
GGGCGACACC	AACCTCCCTC	CGGGTCCACT	CAGAGCCCCC	CCTGATAGTG	GCCGACGGAG	CGGGCGGAAC	TCCAGCCGCC	34160
CAACGGAAGG	CCCGTCGTAC	GAGCGGCGGG	CCTTCTTCGT	TGTGACCACA	ACCACACCGG	AGGACCTATG	ACGGTTCGCT	34240
CTTCGATCAA	CGCCTCAAAC	CTGAGGCGTG	AACCTACGGC	CCTCGAGGAG	GCCTCGCCGA	GCAGCGTCAT	CGGCGATGTC	34320
CTCGACGCGA	TGTTCTCTGC	GGGGCGTGCC	GCACTCGGTG	CGCACGACCT	GGGCATGCCG	CAGAACGACG	GCTGGGCCGA	34400
AGCCGAAGCG	GTCATCTTCG	ATGCCGTCCG	CCGCGCCAAC	CCCGGCAAGT	TCGACTTGCC	GATCGCGCGC	GACGCGGCAG	34480
CCTGACCGAC	CTTCGCAAAC	TCCCGAGGCT	GACCGAGGCC	CCTCGCGCAA	GCGGGGGGCC	TTTTTCGGTC	GTCTTTCACC	34560
GCCCTCTGGA	GCGCTTAACC	ATGACCGACA	TCACCCCGGA	AAAGATCATC	TCCGACGCGC	CGCTGGGGGC	GCTCATTCGA	34640
TTACGTGACC	GGACGACCCA	CAGCCCGGCT	CGGTTCAAGC	GCAAGCTGTC	GGCGTGGGAG	AACGCCAACG	GGTGCGGAAC	34720
ACTGGCCGAG	GTGAACCCTG	ACCATGGCAC	GTTACGCTT	CACATGGGCG	ACCTGGGCGG	GGGTGGCGTC	ATCGTGATGC	34800
GGGTCTATCG	CACCTACGAC	CGGTCCAGCA	GCCTGGCCTT	CACCCTCGAG	CGGCAAGCCG	TCGAGGGTTC	GATCCTGTGC	34880
GCGAACGTGT	ACGGCGGACG	ATGGAAGATC	GACAAGATCA	CGGACGATGC	TGGCGGGGCC	GCATGGCTCC	GGAACAACAC	34960
CTTCGGCGAG	TTGCTTCGCG	TGGGCGCCAG	CGGCGCCCTT	AGCGTCGCAC	ACGCTCCGCG	GAGGGCCGCG	TAGCTGTGGC	35040
CGACACCGAA	CGGCTGTGTC	GTTGCTCGA	CGTGCTGAAG	TTGCAGTACC	AGTCGGAGCG	CGACGCGGCG	ATCGGCGACC	35120
CCGACCGAGA	GCACATGGCG	CCCCGTTTGG	GGCGCATGGC	CGACCTCATC	GGCGACATGC	GCTACGACCA	GGCCCTGGTC	35200
GGTCGGCTGG	CCGCGATCCT	GGCCGGCGAG	ACCATGCCGA	GGTCGCGTTT	CACCGTCACC	GCGCATCGCA	GCGGCCAGCT	35280
CGGCGAAGCG	ACCGCCGACG	TCAAAGAGGG	AGGGGAGGTC	GTTTTCTGGG	ACGATCTGGA	GGCGGCGAGG	AAGGCCGCCT	35360
CTGACCTCAA	CGACCGAACC	TCTTTTCGGT	CGAACGTCCG	CTACGTCGCC	AACCCACACC	CCCGCCGTCG	CCACTAACCC	35440
GAAGGCCCGC	AGCTTCCCAG	CGCGGGCCTT	CGTCATGTGA	ACCACTGCCC	GGAGAGGCTC	ACATGTCTCA	AGACTTCGAA	35520
CCACGCGCGC	GTCCCAACCA	CAGCCGTGCG	GGCGGCGTCC	AGAGCGCCAA	GGAAATACGC	CCCGGCATCT	GGTCCGTCGT	35600
CACGCCCCAG	CACGGCGGCT	TTTCCCTGTC	GGCCGAGCGC	AACGCGAAGG	TCGCGGATTG	CTGGCGCTCG	GACACGGGCT	35680
GGTACGAAGA	AGACTGCGAA	TGGGCCATCG	TGTGCGCCAC	CTGGCCCGAG	TTCTTCACCG	AGGTCTGGCG	CCTGCAAGCC	35760
GACGTCACGC	TGCGCAACTG	GCATCCTGAC	GGCTACGAGG	CCACCCATGG	CGTCACCTC	ACCGCGGCCA	ACAGCCATGC	35840
TGTAGCTGAG	CGCGAATTCT	GGGAGCGCCA	CATCGAGGAC	TTGCTCGTCC	GTTCCGCGTG	GGGCGATCAC	ATGGCTGGG	35920
TCCCGGAAGG	CTTCGTGGGG	GTGATCGCCG	GCATCGGGAG	ACGCCCCGTC	TGCGGCTCCC	CGCGTGAAGA	GCGATACTTC	36000
CTCGTACCTG	CGTCCGAATA	CCGGCTCGGG	AGCCACGGGT	TCGTCATCGA	CCGTGCCCGA	CACGCCGAAA	TTCCGGCGCC	36080
CACGAACCCC	CACGAACGCC	GGCAGCGGGC	CGCTTAGCAG	CCTCCATCTC	TTTGCCGCGC	CTCGCTTCTC	TCGGAGATTC	36160
CGATGACCAA	CACCCCGACC	ATCCTGACCG	GCGAGCTCGT	TGTCGCGGCC	GACCCGAACA	TCCGCTTCCG	CTCGACGAAG	36240
GACTACATCG	GGGCGATGAC	CGAGATCACG	GCGGCCAGCT	GCGCGGCCGA	CGGATCCGGC	ATCGACCTGG	AGTGGGGCGG	36320
GGCGTTCAAG	AACTTCCCGG	AGGACGGCAT	CACGCTCACC	CATGATGGGG	TCCTCTTGTA	CGAAGACGCG	TCGGCGAACA	36400
CCTATTCCGA	GGACCAAGTG	GCGGTCTGCA	CGCCGGATGG	CACGCTCATT	CCCTGGAAGG	TGGAACGCGT	CCAGCCGCTC	36480
TTCATCCGTG	ATGTCGAGTG	GGCCAAGCGC	CTGATCGTCC	CCGCCGAAGC	CGCCCTAGCG	GGCGTCGCGC	AGGGCGCTCT	36560
TCACGACAAC	CTGCAGAACG	CCGTCAACGA	GGCCGTGCTG	GCCGTCTGGC	GCCTCAGGGC	GCCGCTCCCG	GCCACGCGGA	36640
ACGACACCAC	CCATCGCTTC	GAGGTCCGCT	CTAAGGAGAC	GCGGGCGAAC	ATCGGCTGGT	TCTCCGTCGA	GCGCCACGGT	36720
CGCGAGGCGG	CCCGCGCGCA	CGCAATCGCC	CTGCTGACGC	AGTCCGGCGA	GCGCTATGAG	CTCGTCGAGC	GCGGGCGGCC	36800
GGACATCGAG	CCGGGTGAGT	TCTGGCGGTG	CGCGGAGGCG	GGGTACACCG	TGACGGAGTA	CATCGCGCCG	CTGCTGGCG	36880
GGGGGACCGA	GGTTGAGCAT	TTTGATGCGT	GCGGTCCGGA	CGGTGAGGAG	CTGGTGAGAG	GTTTCTATTG	GGCTTCTGAG	36960
CGGGTGGCCT	GGGAGGCATG	TGATAGGCAT	CGGATGTTTG	GATCAACGGC	CCCGCTAGAA	GACTAGCGGG	GCCTTATGGC	37040
CTTCGGCCAC	TCCCGACCCG	GGAAATGGAAG	ATTTGATTGT	TTGCTGATCC	TCTCCCAGCA	GCCCCGCCCG	GCGAACCGGG	37120
CGGGGCTGTT	GTCTTGCTGC	GGGACGAGGA	TCGTGTTGCG	CCCCGACGGG	CTTTTGATCT	GCGCCGGAAG	GCGGGCAGGG	37200
CCGTAGCGCG	CGCTCGGGCG	GTCATACAGC	GTCAGCAGAG	TGGATCGAAT	CCGCATCACG	TTGGACAGGG	CGCGCTGTG	37280
GGCGGTCGAG	CGCGGCGAGG	TTGCACGTCA	CCTGACGGGC	ATGCGGCCCA	AGACGCCGAT	CGCCCCGCTG	GTGGCCAGTA	37360
CGGGGAGGGT	GTCGACCCGG	TAGCCGCTGG	ATGACGCTGG	CTGGGCATTC	GGAATGGACGA	AGCGGAACGC	GATGACCGTG	37440
GCGCGGGGAG	CCGCGAGGCG	GCCGCGCGCG	GGATAGCCGG	GTTGGAAGGG	CGGCTCACGT	TGCGGAGCGG	CAGGGCTCGG	37520
GTGGGCGGGT	CTCCAGCGCG	CTGTTGGGCG	GACTTGGGGA	CGCCGGGCGG	GAGGGGGCGG	TAAGGCGGCG	TGAAAAGCCC	37600
TTCTGGGGGT	TGGCAAGACC	TTGCTCGATG	GACCTCGCCT	CCGTTTCATC	GCCCCGCGAA	GGGACGCTGC	GCTCAAAGCC	37680
CTTGCTGGGG	CGCCTGCCCG	AAGGCGCGAG	ACGGCCACGA	GCTTCGGCTC	GCCAACCTCT	CAAAGAAGGG	ATCACACCAT	37760
GCAAGTTACC	TCTCCCACTT	TCCTGACCGG	CAACCTCTCC	GACTTCGACC	AACAAGTCCT	GTTGCTGGAA	GATGTCGCGG	37840
CCTACCCAC	CGAGGACGCC	CTGACCCAGC	TCGGCACCGT	GGTCATGAAC	GAGCTTCTGG	ACGTGGTGCT	CGACACCGGC	37920
CTGGAAGATC	ACAACGCGAT	CATCGCCGAA	TCCCTCATCG	GGGCTTTCCA	CTCTGCCTG	AAGCGTCTGG	AACGCGACGC	38000
CGACCGGGCC	CGTGACAGAG	TCAACGTCTT	CAACCGCGAT	TTCGACGGTT	CGGAAATCGC	CGACTCCGAA	CTGCAGGAAG	38080
CGATGATGAA	GGGCGCGAAG	CTCGACGTCG	CCATCGCCTC	GGTGGAGTTC	ATCCGCGACG	CCGCCCACGC	CAGCTACACC	38160
GCTGCGACCG	GTGAGGTCCT	GACCGCCTGG	ACGGGCAACG	TGAAGGCCTC	CCGACGACCC	GCCGCTCAAA	TCCAAGCCCC	38240
CGATATGATC	CGCGCCAGCA	AGGCCGCCAA	GTACGCGGCA	GTGACCCCCG	GCGCACCTGT	CGTCGCCTTC	CGGGGTTTCG	38320
CCAAGGCCGA	TACCGCCGTC	GACGCCGCAC	GCATCTTCGA	TGCGCTGAAC	TGGGCCAAGA	CCACTTACCC	CGACATGGTC	38400

CTGGCCACCA	CGGGCGCGAG	CGGGCGCGAG	AAGCTGGCGG	CCAAGTGGGC	CCGCCAGAAG	GGTGTGACGA	TCATCCTCGC	38480
CGCGGCCGAC	TTCGACACCA	ACGGTCGGGC	CGCGCCCTTC	AAGGCCAACG	ACGAGATGCT	GCAGCTCGAT	CCGGTCTGCT	38560
GCCTCACCCCT	GCCTGACTCT	CTGAACCCCG	AGCGCGCGGC	CAAGAACGGC	CAGCCCTTCG	GTCCGGCGCT	GAACCTCGGT	38640
CAAAAGGCCG	AACAGTCGGG	TCGCAAGCAC	GTCAGGATCC	TCCTCAAGGC	CTAGCCGGCT	CTGTTGCAAA	GATTGGCGGC	38720
AGTCAGAGGT	AGGCTGTGCG	TCGCGCCGA	TCAGGCGGCT	GCTGCGAAAT	GGTGGTTGAG	CATGCCCATG	GCCTCCGTCA	38800
GCGCCGAGGG	CCCAATGCCA	AAAGCTCTCT	CCACAAGGCG	CACCTCGCCC	CTGATGCCGG	GCTGCAGGCA	CCAGGGGCGA	38880
GCCTGTCTTT	TGCGCAGGGC	TCGCATGACT	TCGAATCCCT	TGATCGTGGC	ATAGGCCGTG	GGGATCGATT	TGAAACCGCG	38960
CACCGGCTTG	ATCAGTATCT	TGAGCTTTCC	GTGATCGGCC	TCGATCACGT	TATTGAGATA	CTTCACCTGC	CGGTGGGCCG	39040
TCTCCCGGTC	CAGCTTTTCT	TCGCGCTTCA	ATTCCGTGAT	CGCTGCACCA	TAGCTCGGCG	CTTGTTCGGT	ATTGAGCGTG	39120
GCAGGCTTTT	CCCAGTGCTT	CAGGCCCTGC	AGGGCCTTGC	CCAGGAACCG	CTTCGCTGCC	TTGGCGCTGC	GGGTCGGCGA	39200
CAGGTAGAAA	TCGATCGTGT	CGCCCCGCTT	GTCGACTGCC	CGGTACAGGT	AGGTCCACTT	GCCCCGCACC	TTGACGTAGG	39280
TTTCATCCAG	GCGCCAGCTC	GGATCAAAGC	CACGCCGCCA	GAACCAGCGC	AGCCGCTTCT	CCATCTCCGG	GGCGTAGCAC	39360
TGGACCCAGC	GATAGATCGT	CGTATGGTCG	ACCGAAATGC	CGCGTTCGGC	CAGCATTTCC	TCAAGGTCGC	GATAGCTGAT	39440
CGGATAGCGA	CAATACCGT	GCACCGCCCA	CAGGATCACA	TCACCCGTGA	AATGGCGCCA	CTTGAAATCC	GTCACTGTTT	39520
CGTCCGTCCA	ATCTCCGCCA	AGCATGCTCA	AGCTTCACGA	TTTTTGAAC	AGAGCCGCCA	ATCTGTTGCG	GGGTGCTCG	39600
ATCCACCCAG	TCCGGCTGAT	GCACATTGCT	CAGCTCAGCG	CGCGCCGCTG	CACGTTCTGC	ATCGAGGCGG	GATTGCAGCT	39680
CGCGGTGCTC	CTGGATGCTG	TTCTCACGTG	CTCGGCGCAG	TGCCTCTCA	CGTAGGCGAG	TCATCATCTG	ACCGACCTGC	39760
GAGGTGCGCG	TCACAAGGAT	GCGTAGCTGC	CCCTCCATCG	CCTCTTCGAT	GCCGTCGCTT	TCAGCCATGA	TGGGATCTCC	39840
TCTCGTTGTG	TTCTATCGCT	CGATACCGTC	GCGGCCCGGA	AGCCTCGGCT	GAGGTGTCTT	AGGCCGCTCG	GCCGGTTGCA	39920
CTCGGTTAGG	CACCGGCGAC	CCGATTTCCG	TTGCCGGTCG	CGCCTGGCTC	GCGCGCACTC	GTTCAAGCAC	AGCTTGCGCC	40000
TCAGCATCGA	GCGGCGCGCG	CTCGATCGTC	GCCGTGGC	CCTGCTGACG	CGCGGCAGCA	TCGGTGCTCA	TCATGGGTGC	40080
CGGCAGTGCT	GCACGAACGC	GGCTTAGACG	AGCGCGCGTA	TCCTCAGCAA	GCAATCGGGC	CTGGCGCTCC	TGCGCCGCTG	40160
CTACTGATGC	CGCGTACACC	GCCTGAGTGA	GGGTCAGCAA	GGATCGGATC	ATCGCAGCCT	GCGCGACGGC	ACCCTGCCCC	40240
CCGCGCACGG	CGCTGGCCAG	CAACATCGCA	GCACCAGAGA	TAGCCACAGT	GCCGGCCTGC	TGAGGACGCA	CGGTGCGCTG	40320
GTATGTCTGA	GCGGATCGTG	ACAGCGCCTC	CGATGCCGCA	GCGAGGTGCG	CTGGCGTCTC	CTCGGTCGCG	TTGACCATG	40400
CAGCCAACGC	GCCCGCAGTC	TGTCTCGCGA	CGGTGCGCCA	TGTGCTCCGA	TCCTCGACCG	GAACCGACCG	TAGACGGTGC	40480
ACAAGAGCGG	TCAGCTCCTC	ACTGCTCCGT	GTCCACTGCT	CAGGGTTGGG	CTCCTGCGTC	TCCCGGCCGG	GAGCAACCAC	40560
GCGCCGCGCC	CGCTTCGCGG	CATTCCATTC	AGCAGCCGCT	TCGCTCGCTC	CTGTGGGCGT	ATCCGGCCAT	CCATCGCGCA	40640
ATCTCGCGAG	CGTGAGGTCA	CGGCCGAGAT	TTCCGCCGCC	GTACCAAAATC	GGCCGCTCGC	CTGCAGTCGG	ACGCTCAGCG	40720
ACCGAGTACC	CCGTGATGAC	GTCTGTACGG	CCGTACAGCT	AGCGCGGTCG	AACGACGAGC	CCTTCACGAC	GCATTCGGCG	40800
CACGAACTCG	CCCTCATCTG	CGCTTGCTTC	CGCGCATCCT	CGCACCTTCA	AGCTGAGCAA	GTATCGCGCG	TGATCTGCAC	40880
GCATCTCAGC	ACCGATCAAT	ACCTGCCGCT	GTGCACCATC	GAGTTTCGCC	CACTCGGGTG	AATCAGTGCC	GGCAGATTGG	40960
CGGTTGCGTT	CGTACTTCGA	ACGTGCGCGC	CAACGCGCTT	GTGCCTCACG	CTCAGCGGGA	TCATACCCAC	GGGTGCGCGG	41040
ATCAGCCTGC	ACCGATTCCA	GCGACTCCAA	ACCGTGCGTA	ACTTCGAGCG	CACGTGCGAC	AGTCTGCGCC	CGCTTGAGT	41120
CCATGTGAAT	GGGAGCCTTG	GTGCCGTCTT	CACGCACCAA	GTTCCAGCGG	ATGTGAACGT	GATCGTTCCC	GTCTTCGAG	41200
ACACGTGAGC	GCACGGCAAC	CCAACGGCAT	GGTGCTTGG	TTCCCTCGTG	ATCGTCGAAA	CCCATCTCAC	GCACGAAGTC	41280
GTTGGCGATC	GCATTCCACT	TTTCATCGGT	AAGCGTGCCC	TCTTCGGCGC	GCAATGACAG	TGAGCAATGC	CAGACGTGAC	41360
CGCCCTTCAC	ATCAACATCG	AAAGCCTTAC	GTGGGCGATC	CAAGTGACGC	GCGATCGCCA	GTGCAGAATC	GCGGCCAGC	41440
TCGGCATCAT	CATGCCAAGC	CATCATCGCT	GCATCCCCAG	CAACCAGGTG	AGGTTTCACTG	TGCTCGTTGG	CGCGTCTG	41520
TCCGACAAGG	TACGTATGAC	GACCGGCCAT	GCGATCGCGG	CGTGTGATGT	TCGGCATCAT	GACGATGCCT	CGCTCAGCTC	41600
ATCCACGAAT	CCGTGATGAT	GATCGGCAAC	CTCGCGCACC	TTCCGAGATG	TCGCCACCAT	CTCCGCTG	ACTGATCCAG	41680
TAGCGTTCTG	GGCCTTCGCC	ATCTGGTTCA	CGTTGTTTCG	AAGACCGCG	AGCAAACGAT	GCGTCGCGAA	CATCTGTGCG	41760
ATGACCTCGT	GACGCACCGA	TTTATTGGCC	GCAGCAGCAG	CCGCTCCCCC	AGCCAATGCC	GAGTCAACAA	GCAGCTTCGG	41840
AACGGTCACC	TTGTACTGCA	ACGCAAGCGC	AAGAAGCTGG	CCCTCTTCCT	CCGGGCTGAC	CTTCACGATG	TGCCGAGCAA	41920
CCCGACCGCC	CTTGACGTTA	CGACGTCGAG	CCTGATGGCG	CTTCGGTTCA	GCATCATCAT	GCGAATCAAC	AGACAACCCA	42000
GCACCTCCCT	CCGGCCGAGC	GCAGCGACAC	CGCAACGCGG	TGACCACACT	TCCAGTGTGC	CGCATCTCTG	ACTGAAAATC	42080
GAGGATGACG	CCCAATTAGG	TTACCTAATT	GTATAGCTTG	CTCCGTCTGG	GACGCCGAT	AGGTCGCTC	ACGGTGGCAG	42160
CCGGGAGTAG	TCGCGGCTTC	GCCGCGGCGT	GCTCGCATAT	GGTCGACGAA	GCTCGACGCG	AGCTGTCTTC	ATCCGTAGGA	42240
TGGAGACATG	GCTAAGGAAC	TCTCAATCGA	AGAAGTAATG	GACCGTGCCC	GTCGTGCTCA	GGAGGATCGC	ATCGAAGCAA	42320
TCCGAGTGCT	CGCTGAGGCG	CGGAACGAAC	TCGATACAAT	CCGTGAGGAT	GGTGCTCGTG	AGCTGGCGGA	GCTGCAAGCT	42400
CGAATCAGCG	AACGAGTGAC	GTCCGCCGAA	CGCGATGACG	TGAAGCGGTA	TAACGCGGCC	GTCAGTGCCG	GATGGAGCGC	42480
CGATGAGCTG	CGCAAGATCG	GGTTTGACAG	GCCCATAAG	AAGGCTCGCA	CGCGCAAGCG	TGCACCGCGT	AAGAACACGG	42560
TTCAGGCTTC	ATCTCGAAGC	CCCGCTGTAT	CGACGACGGC	CTCGGGTCTC	GAATCTGATT	TCAAGCTCTC	CGACGATGCT	42640
GCAACGCCCG	AGCCAGCGGG	CGTCTAGCTC	GTCGCTTTTC	GTGGGGCTTT	CCCCTCCCCC	TGAGCCCGAG	ATACTTGCAC	42720
AAGATAGCCA	CCAGTCAAGG	GCGCTTCGCG	TACTCCGCTG	ATGCGGCTGC	GCCGCACCTT	TGACAGGTGG	CTATCTTGTT	42800
CGGGGCGGCT	TTTATCAGGG	GAAGAGAAAA	GAACGGTTGG	TCTATCGTTT	GATGCCGTTT	TCTGATGCTC	GATCGGTGCG	42880
CGAATCGGC	CGCGTCGCTT	GCGCTCCTGT	GTTGCGTGTG	GCGGATCTAT	CCTTGCTGAA	ACTTGACGCA	GATAGCCTGC	42960
TCACTGACTG	CTCTACTTCG	AGCTTCCAC	CTGGCTTCGA	TCGCACAGCT	GCAACCTCTC	GAACAGGCAG	TGCCTCGCGC	43040
CCCTCACTCA	GCGATGCTCG	ATGCTCCTCG	GCGCTCGTAA	AGTGCCGCTC	ATACCGGATC	ACTCGCGGGC	ATGTATACGC	43120
AAGCTGGGCT	GCCGCATCCT	GCGACTTCGA	GCGTCCATGC	TCAGTGTGCA	ACAGCGGCGG	AGCGTAAGCG	CTCGGCACCC	43200

ATCGTCCGCC GAGTGCCTCT GCTTTGCCTG CCTCGGCTTC CTGCATCGCG TGCTGCCAGC GCTGCACGAT CCGATCCTGC 43280
 GATACCTCGA ACGGCACCTC GACGTGACG ACGGTTACCT GGTAGCCGGC TTCCGCAAGA TACGACCCAA GCTCCACTGC 43360
 TTTGTCTGGG TCGCTCAGTA CCGTGTGAC GATAATGTTT GTTCCCTGGT CGATCATCTC TGTGCGCAGC GCGCGTGCAA 43440
 GCATCGACGA TTCTTCATGC ACGAGTGAGG CAAGTTCAAG CGGATAGAAC CGCTCCCCCT GCGCCTCTAA GTCGCGCACC 43520
 GCTGGCGGTT TGATCCACTC CTCATAGCTG CCATCTGCCA TTGCTTCGCG CAGCAACAGC TTCTTGAAAT CGTCCGCGTC 43600
 AATGTTACG TAGTGGTCGC GATCCGCACC GAGCACGGTG TCAGCGACTC GCCCTTTACC GGCACCTGGC GGGCCAGCCA 43680
 GCACGAGTGC TCGACCTTCA TCTTTGCGT TCGGAGACTG GCTCACTACC GCGGCCTTGA TCGTTTCGTG CGTCGCCCCG 43760
 CGTCCGGCTT TACGCGACCC GTTGGAGCGA TACCAAGCTG GGTTCGCGGT AGTCGCGTGT GGGCCGTCCT TGTGAGTGG 43840
 TCCGCCGGCA GTAGAGAGCG CTTGGACAAC CTCGATATGC CGCTGCTCCT CTGGGGTCAC GCGGCCGTCA TCCGTGCTTC 43920
 GATGCGATCG CCAATCTCGT CGTACAGCTC GTCGTCGATC AGCCCACGGC TCAGCGCGCG GTCCAGGTCA TCGACCGAGC 44000
 CCGCAACGTC CATCAGCAGA TCGTCGAACG CATCCTTGGT CTGCGCCTGC GGCACGTA CTCCAGCGCT CAGCTCATCG 44080
 ACGAGCTGCG CGCGGTCAAT CAGCCCGGCT GCGTACCGCT CGCAAATCTC GTAAGGATCT GCGCCAGAGA AGCCTTCACG 44160
 AGGCATCGGG GTTGGCGCGG CCGCTGCAA CACCTTCTGC ACGTTCGGT GCGAGACCTT CACTAGATCG GCGATCGCTC 44240
 TCTGCCGCAA TCCACGCTGT TTCAGCCGCA GCAGACTGCG ACGGTAGTCC AGCTCTCCGA TGTTCGCAA GTGCTGCGCC 44320
 TTCCGTGACG CGACCGCCTC GACGGGAATC TCTGTGATCG TCATAGTGTC CACCTCCGCT TCCAGTCTAC CGAAAATATA 44400
 ACCATAGTTA TACTGCTATT TTGCGAGCGC GCGAATGACA CTCGATGTGC CGACTCCGAG CACGCGCCCC ATGTGGCGCT 44480
 AGCTCTGGTT CTCTGCGCGC ATCTTCTGGG CCGCGCGCAT TCGCTCGGCC GTCATCACGG TTGGGCGGCC GCCGACCTC 44560
 CCCTGACTCC GTGCGTGCTC TAAGCCAAGG CGCGTGTTCT CTCTGATCGT GTCGACACGG AGTGCGCAA AGACCAGAC 44640
 GATGCCATAG AGAGCCTTCC CCAATCGGCGT AGTGGTGTG ATGGCGGGCT CCGTGAGACT CACGATATTC AGGCCTCTAT 44720
 CGTTCAGCTC GGATAGCACC TCGATCATCA GCGCTCACT GCGGCGGAGC CTGTCGAGCT TGCGGACCTT GATCGTGTCG 44800
 CCGGGGCGCG CGGCTACGTC GATGAGCTTG CGCCACTCGG GGCGATCGGC GATGCGGCTC GATTCGCCGT GGTCGACGAA 44880
 CACGCGACTG CAGCCGGCCG CTCGAGCTC GGCTTCTGT GCGCCCGGAT CTTGCTCTTT GCGCGAGACT CGCGCGTACC 44960
 CAAATACTTC GGCCATGCCA TAAACGGTAC GCGCTCAGTA TCGGTGCGAT TCAGTTTCGG TGGGGGATTT TGGTGCCTCG 45040
 CCTCTCGCGG TGTGCGCGAG CTCGTGACG TCGCCGAGCT GCGCACCGAA AACGATCGTT ATCGGTGCGG GCTCGTCGGC 45120
 GGCAGAGCTG CATCGAGCAC CTGCCCCGATG CCGGTGCGCT CCTCGGTGTC GGCATCCGGG CCGTCCACGT CGTGA CTGAC 45200
 TGGGGATCGT CGTACCAGCG CACGACGCGG TTGGCTGCGT GCGCTTCGCG CAGCGCGGCC GAGGCGGCAT CGAACGCTGC 45280
 GGCGGCACGG TGCAGCTCTG CCGCGTGGT GACGGTGCCA GTTGCGCCGT CGCCGCGCCC GTCTTCGCGG TTGTAGTGCG 45360
 TCCCGTCGAC GACGCGCTCA TGCCAGGCCG CGAGCTGGCC GAGCATGGGT ACGTCGACGC CGACATACCT CCCGA ACTGC 45440
 GCACCGGCCG TACGCCTCAG TACAGCGTGC AGCGCGGTCG CGTCATCGAG GACGCGGGGG CTCTGTTGCA AAAATCGTG 45519

域には NSF が多数存在し、その長さも比較的長い、18-24kb と 40-45 kb の領域には少なく、その長さも短い。

第四節 考察

ナイロンオリゴマーの分解に関与する3つの酵素 EI, EII, および EIII を支配する pOAD2 プラスミド全塩基配列を決定した。全長が 45,519 bp であったが、当初予想されていた 43.6 kb と比べ少し大きい結果であった。両鎖ともに塩基配列を決定したが、G+C 含量が 66.6% と高いこともあり、また、大腸菌宿主における組換えや、遺伝子操作中に起こる可能性のある点突然変異のことも考え、なるべく2回ずつ両鎖で計4回読むことができるように欠失変異体を取得した。

pOAD2 に関する詳細な解析、考察は次章以降に行う。

第五節 要約

ナイロンオリゴマーの分解に関与する3酵素 EI, EII, および EIII を支配する pOAD2 プラスミドの全塩基配列を決定した。その結果、全長は 45,519 bp よりなり、その G+C 含量は 66.6% であった。300 bp 以上の ORF が 157個認められ、600 bp 以上の NSF が 71個存在していた。

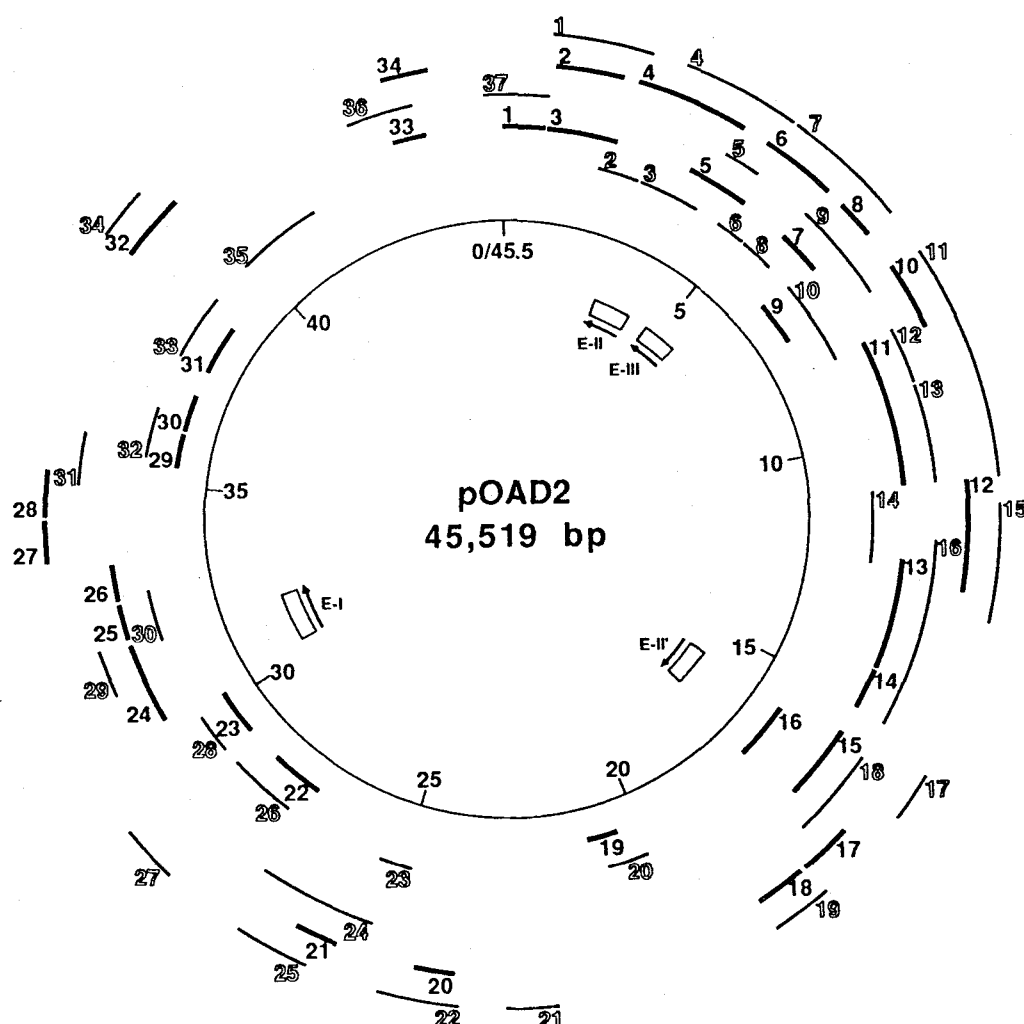


Fig. 1-5 Functional map and non-stop frames in pOAD2. Length of pOAD2 was found to be 45,519 bp. Seventy one non-stop frames (NSFs; coding regions flanked by one termination codon and its successive termination codon) with more than 600 bp were found in pOAD2. There were 34 NSFs on the three clockwise frames (shown as numbers of thick letters) and 37 NSFs on the three counterclockwise frames (shown as numbers with open letters), and these six frames are shown on the part of circle with different diameters: the most inner line, frame NS1 (Non-stop Sense frame 1); second inner line, frame NA1 (Non-stop Anti-sense frame 1); third inner line, frame NS2; fourth inner line, frame NA3; fifth inner line, frame NS3; the most outside line, frame NA2. The following pairs are on the same reading frame on the opposite orientation: NS1/NA1, NS2/NA3, and NS3/NA2.

第二章 Repeat-sequence (RS-I と RS-II) 領域の解析

第一節 緒言

pOAD2 上には、サザンハイブリダイゼーション法を用いた解析から、繰り返し配列 (repeat-sequence I; 以後 RS-I と略す) 領域が5ヶ所 (RS-IA, B, C, D, および E) に点在し、これらの領域は、相互に非常に相同性が高いことが予想されている^{24, 25)}。また、これとは別に、繰り返し配列 (repeat-sequence II; 以後 RS-II と略す) 領域も2ヶ所に認められる。RS-IIA 領域は、EII 構造遺伝子, nylB, を含み、RS-IIB 領域は EII 類似酵素 EII' をコードする nylB' 遺伝子を含むことがわかっている⁷⁾。これに加えて、RS-II 領域としては、nylB' DNA をプローブとして 55°C で弱くハイブリダイズする領域が、約 10 kb の広い範囲で観察される (Fig. 2-1)。これらの多くの繰り返し配列を有することは、このプラスミドの1つの特徴である。このような、繰り返し配列は、プラスミドの形成過程における DNA の重複と転座の結果と考えられる。そこで本章では、前章で決定した pOAD2 の塩基配列をもとに RS-I 領域を決定し、その内部構造を調べた。その結果、この領域には、細菌由来の IS 配列 (insertion sequence) が存在していた。また、この IS と他の IS との比較、Flavobacterium sp. KI72 株と同じくナイロンオリゴマー資化性菌である Pseudomonas sp. NK87 株⁸⁾における IS の有無、ならびに IS がコードするトランスポゼースの Flavobacterium sp. KI723T1 株 (pOAD2 保持株) における発現について検討した。また、RS-IIA 領域の nylB と RS-IIB 領域の nylB' の産物である EII と EII' 両酵素の差異についても検討した。

第二節 実験材料および方法

1 使用菌株およびプラスミド

Flavobacterium sp. KI723T1 株⁶⁾ (pOAD11, pOAD2, および pOAD3 保持株), KI723 株⁶⁾ (pOAD11 と pOAD3 保持株) を RNA 調製用菌株として、KI722 株⁶⁾ (pOAD1 と pOAD3 保持株), Pseudomonas sp. NK87 株⁸⁾ (pNAD1, 2, 3, 4, 5, および 6 保持株) をプラスミドならびに染色体 DNA 調製用菌株として用いた。また、E. coli C600 (r_k^- m_k^-) 株^{26, 27)} (thr-1 leuB6 thi-1 supE44 lacY tonA21 hsdM hsdR) を染色体 DNA 調製用菌株、および酵素活性測定用の宿主として用いた。形質転換の宿主としては、E. coli JM109 株¹¹⁾ と JM103 株¹⁷⁾ (thi strA supE endA sbcB15 Δ (lac-pro) F' [traD36 proAB⁺ lacI^q lacZΔM15]) を使用した。プラスミドベクターとして pUC12 19¹⁶⁾、また塩基配列の決定に M13 mp18, 19¹⁷⁾ を用いた。

2 培地

E. coli 用の培地として LB 培地, 2 x TY 培地 および H 培地を用い、必要に応じてアンピシ

リンを最終密度 50 $\mu\text{g/ml}$ となるように加えた。また、寒天平板培地には、必要に応じて IPTG と X-gal をそれぞれ 0.1 mM, 0.005% となるように加えた。*Flavobacterium* sp. KI723T1 株および *Pseudomonas* sp. NK87 株の培養には、環状ナイロンオリゴマー培地を、また、KI722 株と KI723 株の培養には LB 培地をそれぞれ用いた。

3 *Flavobacterium* sp. KI722 株からのプラスミド DNA の調製

KI722 株からのプラスミド DNA (pOAD11 と pOAD3) の調製は、Negoro らの方法⁶⁾を改変して行った。1.2 l の LB 培地に一晩培養液を 1% 植菌し、30°C 12時間往復振とう培養を行った。集菌後、TE* 緩衝液 (50 mM Tris-HCl pH 8.0, 20 mM EDTA) で洗浄した。150 ml の TE* 緩衝液

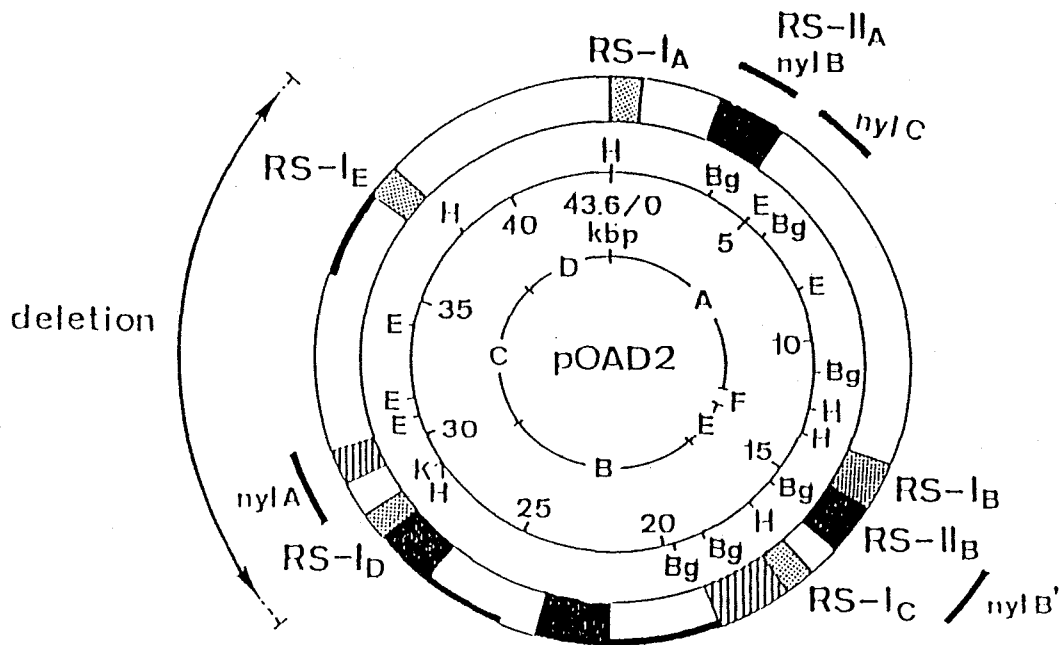


Fig. 2-1 Structural and functional map of pOAD2 plasmid. The nylon oligomer degradation genes were at the following loci: the *nylA* (29.1-30.6 kb region on pOAD2 map), *nylB* (2.7-4.1 kb region), *nylB'* (14.7-16.1 kb region), and *nylC* (4.2-5.3 kb region). RS-I and RS-II are repeated sequences identified by southern hybridization experiments. RS-I appeared 5 times on the pOAD2 (RS-I_A, 0-0.8 kb; RS-I_B, 13.5-14.7 kb; RS-I_C, 16.8-17.4 kb; RS-I_D, 28.2-28.8 kb; RS-I_E, 37.1-37.9 kb region on the map). The *nylB* and *nylB'* genes are included in RS-II_A and RS-II_B regions, respectively. E, Bg, H, and K represent restriction sites for *EcoRI*, *BglIII*, *HindIII*, and *KpnI*, respectively. A, B, C, D, E, and F represent fragments generated by *HindIII*. Arrow indicates the region deleted in pOAD2.

(20% ショ糖, 3 mg/ml リゾチームを含む) で懸濁し、30℃ 2時間保温したのち、8000 rpm 5分間で遠心、集菌を行った。その後の操作は、Birnboim と Doly のアルカリ-SDS 変性法¹⁹⁾に従い、CsCl-エチジウムブロミド平衡密度勾配遠心分離法でプラスミド DNA の精製を行った。

4 *Pseudomonas* sp. NK87 株からのプラスミド DNA の調製

NK87 株からの6種類のプラスミド DNA の調製は、Kanagawa らの方法⁸⁾と Szabo の方法²⁸⁾に従った。1 l の環状ナイロンオリゴマー培地に、一晚培養液を 1% 植菌し、37℃ 12時間培養した。集菌後、TE* 緩衝液で1回洗浄し、200 ml の TE* 緩衝液に懸濁した。これに 5 mg/ml のプロナゼ溶液を 24 ml, 20% SDS を 12 ml 加え、37℃ 2時間放置した。処理液をビーカーに移し、スターラーで攪はんしながら pH 12.4 になるまで 5 N NaOH を加え、5分間攪はんを続けた。次に、pH 8.5 になるまで 2 M Tris-HCl (pH 7.0) を加え、3分間攪はんを続けた。最終濃度が 3% となるように 30% NaCl を加え、3% NaCl 飽和フェノールで2回抽出し、クロロホルム-イソアミルアルコール抽出を1回行った。等量のエタノールを加え、DNA を沈澱させたのち、遠心分離を行い、DNA を減圧乾固させた。その後、CsCl-エチジウムブロミド平衡密度勾配遠心分離法でプラスミド DNA の精製を行った。

5 染色体 DNA の調製

E. coli C600 株, *Flavobacterium* sp. KI722 株ならびに *Pseudomonas* sp. NK87 株からの染色体 DNA の調製は、Maniatis らの方法¹⁸⁾に従った。

6 *Flavobacterium* sp. KI723T1 株と KI723 株からの RNA 調製

KI723T1 株と KI723 株からの RNA 調製は、Struhl らの方法²⁹⁾を改変して行った。40 ml の環状ナイロンオリゴマー培地または LB 培地に、それぞれの一晩培養液を 1% 植菌し、30℃ 7-9 時間振とう培養を行った。遠心分離により集菌後、TE** 緩衝液 (20 mM Tris-HCl pH 7.5, 2 mM EDTA) で1回洗浄し、集菌した。これを、4 mg/ml のリゾチームを含む 20 ml の TE**-ショ糖緩衝液 (20 mM Tris-HCl pH 7.5, 2 mM EDTA, 0.45 M ショ糖) に懸濁し、氷上に 45分間放置した。遠心分離により集菌した後、2 ml の溶菌緩衝液 (10 mM Tris-HCl pH 8.0, 70 mM NaCl, 1 mM sodiumcitrate, 1.5% SDS) 及び 60 μ l の diethylpyrocarbonate (DEPC) を加え、穏やかに攪はんし、37℃ 5分間放置した。これに DEPC 処理した飽和食塩水を加え、よく混和したのち、氷上に 10分間放置した。遠心分離によって得た上澄液に、2倍量のエタノールを加え、-80℃において RNA を沈澱させた。沈澱物を遠心分離し、70% エタノールによる洗浄ののち、減圧乾固して粗 RNA 標品とした。

7 オリゴヌクレオチドの合成ならびに 5' 末端のリン酸化

オリゴヌクレオチドの合成は、DNA 合成装置 (Applied Biosystems 381A DNA Synthesizer) に

より固相ホスホアミダイド法³⁰⁾で行った。合成終了時に末端保護基であるトリチル基は除去した状態とし、オリゴヌクレオチドの精製は行わなかった。

合成オリゴヌクレオチドは、5' 末端にリン酸基を持たないので、1 nmol のオリゴヌクレオチドを 0.1 M Tris-HCl pH 8.0, 10 mM MgCl₂, 5 mM DTT, 0.1 mM ATP, 10単位 T4 polynucleotide kinase を含む 20 μ l の溶液でリン酸化を行った。37°C 1.5時間反応させたのち、65°C で20分間加熱することにより反応を停止させた。

8 プローブの調製

オリゴヌクレオチドの 5' 末端の放射能ラベル化は、10 pmol のオリゴヌクレオチドを 50 mM Tris-HCl pH 8.0, 10 mM MgCl₂, 10 mM 2-mercaptoethanol, 10単位 T4polynucleotide kinase および等モル以上の [γ -³²P] ATP (111 TBq/mmol, Amersham) を含む 20 μ l の溶液中で 37°C 1時間反応させることによって行った¹⁸⁾。DNA 断片をプローブとして用いる時は [α -³²P] dCTP (22.2 TBq/mmol, Amersham) を使い、Amersham 社の Multiprimer DNA labeling system を用いてその説明書に従ってラベリングを行った。このキットは Feinberg と Vogelstein のランダムヘキサマープライマー法³¹⁾を基にしている。

9 ノーザンハイブリダイゼーション

ノーザンハイブリダイゼーション法は、Thomas の方法³²⁾に従った。25 μ g の粗 RNA をホルムアミドを含む 1.2% アガロースゲルで電気泳動後、20 x SSPE (3.6 M NaCl, 0.2 M NaH₂PO₄ pH 7.4, 20 mM EDTA) で、毛細管現象により Hybond-N ナイロンメンブラン (Amersham) に一晩かけトランスファーを行った。10分間の UV 照射により、RNA をメンブランに固定した。プレハイブリダイゼーション溶液 [5 x SSPE, 50% ホルムアミド, 0.1% Ficoll 400, 0.1% polyvinyl pyrrolidone, 0.5% SDS, 20 μ g/ml 変性サケ精子 DNA (Boehringer Mannheim 社)] 中で 37°C 2時間以上プレハイブリダイゼーションを行い、次いで、メンブランをハイブリダイゼーション溶液 (5 x SSPE, 50% ホルムアミド, 0.5% SDS, 20 μ g/ml 変性サケ精子 DNA) に移し、プローブを加えて、37°C で一晩反応させた。室温にて 5 x SSPE により 5分間 計2回, 42°C で 5 x SSPE により 10分間1回、65°C で 2 x SSPE により 10分間1回洗浄を行った。メンブランをサランラップで包み、X線フィルム (Fuji-RX) を -80°C にて、増感紙を用いて感光を行った。

10 溶液ハイブリダイゼーション

O'Dorovan らによる溶液ハイブリダイゼーション法³³⁾を行った。25 μ g の粗 RNA と 10 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT, 5単位 DNase I, 10単位 RNase inhibitor を含む 20 μ l の溶液を 37°C 15分間反応させた。フェノールならびにフェノール/クロロホルム抽出ののち、エタノール沈澱を行い、遠心分離により沈澱物を集め、その沈澱物を減圧乾固させた。これを 40 mM PIPES pH 6.4, 0.4 M NaCl, 1 mM EDTA, 1 pmol/ml ³²P-labeled プローブ DNA を含

む溶液 30 μ l に懸濁し、90°C 3分間加熱し、次いで、70°C 2時間ハイブリダイズを行った。これに S1 緩衝液 (40 mM PIPES pH 6.4, 0.4 M NaCl, 1 mM EDTA, 10 μ g/ml 一本鎖 DNA, 0.12 単位/ml S1 ヌクレアーゼ) を 300 μ l 加え、37°C 15分間反応し、遊離しているプローブを消化させた。反応液から 300 μ l をとり、新しいサンプルチューブに移し、10 μ l の 0.5 M EDTA を加えて反応を停止させ、更に 48 μ l の溶液 (4 M 酢酸アンモニウム, 0.1 M EDTA) と 1 ml エタノールを加えてエタノール沈澱を行った。遠心分離ののち、2.5 μ l の DEPC 水に懸濁し、等量の Stop/Loading Dye (90.3% ホルムアミド, 0.3% xylene cyanol, 0.3% bromophenol blue 0.37% EDTA pH 8.0) を加えた。90°C 2分間の熱処理後、8% 変性ポリアクリルアミドゲル (8 M 尿素を含む) で電気泳動を行い、以下は塩基配列の決定と同様の操作により X 線フィルムの感光を行った。

1.1 サザンハイブリダイゼーション

制限酵素処理した DNA 断片、または、制限酵素処理をしなかった DNA をアガロースゲル電気泳動で分画後、Southern の方法³⁴⁾に従い、ゲルから DNA をナイロンメンブラン (Hybond-N, Amersham) へ移行させた。Jeffreys と Flavell の方法³⁵⁾に従い、DNA-DNA ハイブリダイゼーションを行い、シグナルの強度に応じてハイブリダイゼーション、洗浄時の温度、洗浄溶液の組成を変えた。

1.2 粗酵素液の調製

ハイブリッドプラスミド保持株を 5 ml の LB 培地 (50 μ g/ml のアンピシリンを含む) に一白金耳植菌し、一晚 37°C で培養した。150 ml の LB 培地 (50 μ g/ml のアンピシリンを含む) に一晚培養液を 1% 植菌し、30°C で OD₆₆₀ = 1.2 まで培養し、集菌した。5 ml の緩衝液 A [10% グリセロールを含む 20 mM リン酸緩衝液 pH 7.3 (K₂HPO₄ と KH₂PO₄)] に懸濁し、超音波破碎 (20 KHz, 30 sec) を 4 回行った後、8000 rpm 5分間遠心分離を行い、その上澄液を粗酵素液とした。この方法で調製した粗酵素液は、高速液体クロマトグラフィーによる測定に用いた。

また、薄層クロマトグラフィーに用いる時は、LB 培地 (50 μ g/ml のアンピシリンを含む) で 30°C 一晚培養した菌体液 (1.5 ml) を集菌後、緩衝液 A で 1 回洗浄を行い、最後に、緩衝液 A 30 μ l に懸濁した。-80°C で凍結、室温で融解するという操作を 2 回繰り返したのち、3 μ l の 10% Triton X-100 を加えて水中で 30分間放置し、溶菌させた。12,000 rpm で 5分間遠心分離を行い、その上澄液を粗酵素液とした。

1.3 薄層クロマトグラフィーによる 6-aminohexanoate linear dimer (Ald) 分解活性の測定

1.2) で作製した 33 μ l の粗酵素液に 33 μ l の基質である 20 mM Ald (緩衝液 A で溶解) を加え、30°C 2-16時間反応させた。反応液を 1 μ l ずつ採取し、薄層 (Silica gel 60 F₂₅₄, Merck Art. 5554) に負荷し、展開溶媒 (水; n-ブタノール; 酢酸エチル; アンモニア水 = 12;

24; 4; 1.3) で展開後、生成物を 0.2% ニンヒドリン溶液 (水飽和 n-ブタノールに溶解) を用いて検出した。コントロールの試料として、5 mM Ald, 5 mM Ahx (6-aminohexanoate) をそれぞれ 1 μ l ずつ負荷した。

1.4 免疫電気拡散法^{36, 37)}

ペロナール緩衝液 [5,5-diethylbarbituric acid (barbital) 1.84 g/l, sodium 5,5-diethylbarbituric acid (sodium barbital) 10.3 g/l, pH 8.4] を 2 倍に希釈し、これにアガロースを最終濃度 1.2% となるように溶かした。ともに 50°C に保温したアガロース溶液 (19 ml) と抗 EII 血清 (1 ml) を混合後、水平台上のガラス板 (8.5 $\times$ 10.5 cm) に一様に広げた。ゲルが固まってから、陰極より 1.5 cm の位置に直径 2 mm の穴を一行にあげ、5 μ l の試料 (コントロールとして 0.1-0.4 μ g の精製 EII タンパク質を含む) を負荷したのち、20 mA で 10 時間泳動を行った。泳動用緩衝液はペロナール緩衝液を 2 倍希釈して用いた。

1.5 高速液体クロマトグラフィー (HPLC) による Ald 分解活性の測定³⁸⁾

粗酵素液と等量の基質溶液 (緩衝液 A に溶解させた 20 mM Ald) を混合し、30°C で反応させた。経時的にサンプリングした反応液を沸騰水浴中で 3 分間熱処理することにより反応を停止させた。12,000 rpm 5 分間の遠心分離により得た上澄液 10 μ l を HPLC (pre-column; Cosmosil packed column 5C₁₈-300 ϕ 4.6 $\times$ 50 mm, column; Fine pack SIL-C₁₈ ϕ 4.6 $\times$ 250 mm, reverse phase partition, elute; 10% アセトニトリル, 90% 20 mM phosphate buffer pH 7.3, flow rate; 2 ml/min) に負荷し、基質である Ald の 210 nm における吸収の減少速度から分解活性を求めた。

1.6 その他の実験材料および方法

制限酵素および修飾酵素などは、宝酒造製あるいは東洋紡績製のものを主に使用し、実験操作については、各々の使用書などに従った。プラスミド DNA の調製、ポリアクリルアミドあるいはアガロースゲル電気泳動並びにゲルからの回収、形質転換、塩基配列の決定などは、第一章、第二節に示した通りに行った。

第三節 結果

1 RS-I 領域の共通配列およびその構造

第一章の全塩基配列から pOAD2 上に点在する 5 つの RS-I 領域 (RS-IA, B, C, D, および E) の塩基配列を検討したところ、RS-IA, RS-IC, RS-ID, および RS-IE は 880 bp、RS-IB は 1300 bp からなり、RS-IA, C, D, および E の全領域と RS-IB の前半 880 bp の領域は、完全に同一の塩基配列を有していた。以後、その共通配列 (880 bp) を RS-I と称する。Fig. 2-2 に RS-I の

IR-1-L
 GGCTCTGTTGCAAAATCGTGAAGCTTGAGCATGCTTGGCGGAGATTGGACGGACGGAACGATGACGGATTTCAGTGGCGCCATTTCCTCA 90
 -35 -10
 GGGTGATGTGATCCTGTGGGCGGTGCGCTGGTATTGTGCTGCTATCCGATCAGCTATCGCGACCTTGAGGAAATGCTGGCGGAACGCGGCAT 180
 TTCGGTCGACCATAACGACGATCTATCGCTGGGTCCAGTGCTACGCCCCGAGATGGAGAAGCGGCTGCGCTGGTTCTGGCGCGTGGCTT 270
 TGATCCGAGCTGGCGCCTGGATGAAACCTACGTCAAGTGCGGGGCAAGTGGACCTACCTGTACCGGGCAGTCGACAAGCGGGGCGACAC 360
 GATCGATTTCTACCTGTGCGCCGACCCGAGCGCCAAGGAGCGAAGCGGTTCCTGGGCAAGGCCCTGCGAGGCCTGAAGCACTGGGAAAA 450
 GCCTGCCACGCTCAATACCGACAAAGCGCCGAGCTATGGTGACGATCACCGAATTGAAGCGGAAGGAAAGCTGGACCGGGAGACGGC 540
 CCACCGGCAGGTGAAGTATCTCAATAACGTGATCGAGGCCGATCAGGAAAGCTCAAGATACTGATCAAGCGGTGCGCGGTTTCAAATC 630
 IR-2-L
 GATCCCCACGGCCTATGCCAGGATCAAGGGATTGCAAGTCATGCGAGCCCTGCGCAAAGGACAGGCTCGCCCCCTGGTGCCTGCAGCCCGG 720
 IR-2-R
 CATCAGGGGCGAGGTGCGCCTTGTGGAGAGAGCTTTTGGCATTGGGCCCTCGGCGCTGACGGAGGCCATGGGCATGCTCAACCACCATT 810
 IR-1-R
 CGCAGCAGCCGCTGATCGGCGCAGAGCGACAGCCTACCTCTGACTGCGGCCAATCTTTGCAACAGAGCC 880
 -10 -35



Fig. 2-2 Nucleotide sequence and structure of RS-I. The 14 bp inverted repeats at its termini are shown by triangles (IR-1-L and IR-1-R). The 79 bp inverted repeats at its internal are shown by open triangles (IR-2-L and IR-2-R).

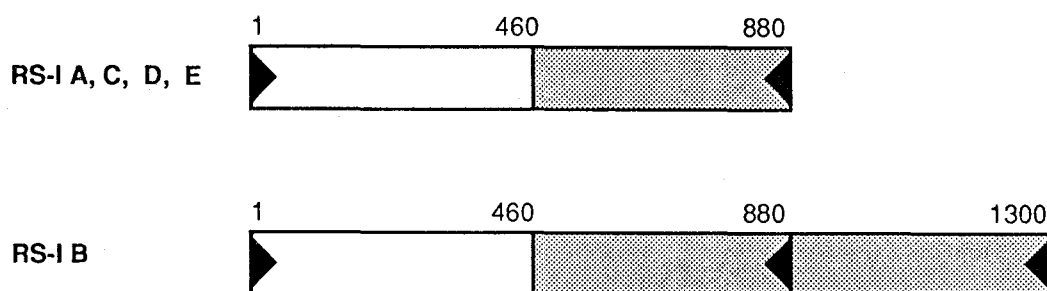


Fig. 2-3 Structural organization of RS-I (IS6100) sequence. The five RS-I regions have 880 bp of the identical sequence. The 14 bp inverted repeats at its termini are shown by triangles. In the RS-IB region, the 420 bp region at the 3' terminal of IS6100 (position 460-880 bp) was duplicated.

塩基配列とその構造の模式図を示す。880 bp の領域には 765 bp よりなる ORF があり、両端部には 14 bp よりなる完全一致の逆位繰返し配列 (inverted repeat; 以後 IR と略す) (IR-1) が存在し、内部には、78 bp の IR (IR-2) が存在していた。また、ORF の 5' 末端部には、大腸菌のプロモーター配列 (コンセンサス配列; -35 領域として TTGACA, -10 領域として TTGACA)³⁹⁾ に相当する配列が見いだされた。また、RS-IB については、RS-I の後半 420 bp の領域が繰返しした構造を有していた (Fig. 2-3)。この 420 bp の中にも変異がなく、他の RS-I の後半部分と全く同一の配列であった。5ヶ所の RS-I の pOAD2 上での方向を見ると、RS-IA, RS-IB, および RS-IC は時計回りの方向、RS-ID と RS-IE は反時計回りの方向であった (Fig. 2-4)。

2 RS-I の DNA データベースとの相同性検索

RS-I を DNA データベースと相同性検索を行った結果、RS-I は、*Mycobacterium fortuitum* 由来の挿入配列 IS6100⁴⁰⁾ と全く同一配列であった。*Mycobacterium fortuitum* と *Flavobacterium* sp. K172 は、明らかに異なる細菌であるが、これらの細菌が全く同一の配列 IS6100 を持っていることは興味深い。*M. fortuitum* において、IS6100 はトランスポゾン Tn610 の一部として存在している。Tn610 は内部に、transposase gene (tnp), sulphonamide resistance gene (sul3), そして integrase をコードすると考えられる未知の ORF (orf2M) を持ち、両端に順方向で IS6100 が位置している⁴⁰⁾ (Fig. 2-5)。

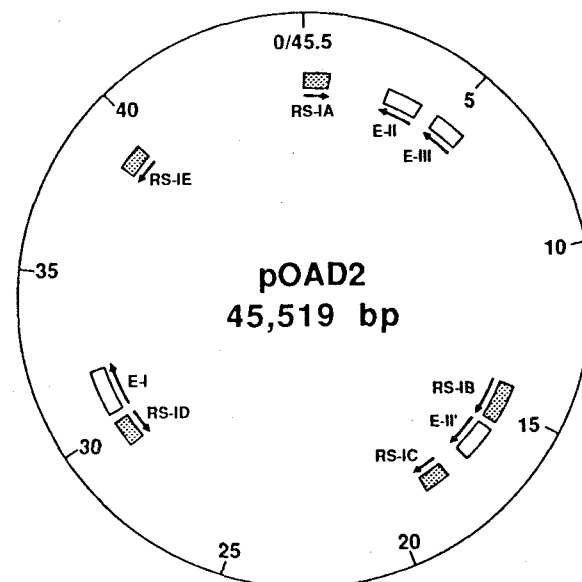


Fig. 2-4 The locations of RS-Is on pOAD2. The arrows indicate the direction of RS-Is.

3 IS6100 (RS-I) と他種 IS との相同性検策

IS6100 がコードするトランスポゼースと他の IS がコードするトランスポゼースについて、相同性と系統的關係を maximum parsimony 法⁴¹⁾により解析した。タンパク質のデータベースと IS6100 (RS-I) がコードするトランスポゼースとの相同性検索を行った結果、細菌由来の IS のコードするトランスポゼースと有意な相同性が認められた⁴²⁻⁵⁰⁾ (Table 2-1)。ここに示す 20 種のトランスポゼースの中には、互いに全く同じ配列を持つものも存在するが、その遺伝子座が異なるものは違うものとして扱った。これらトランスポゼースのアミノ酸配列のアライメントを作成した (Fig. 2-6)。IS6100 由来のトランスポゼースを含む 21 種の配列全てに同一、または、類似のアミノ酸として存在している箇所を網掛けで示している。ここで、アミノ末端部分がよく保存されていることがわかる。このアライメントをもとに系統樹を作成した (Fig. 2-7)。各トランスポゼースは、系統樹上の位置により、[2]-[7], [8]-[14], [20]-[21], および [15]-[19] の 4 グループに大別できるが、更に由来菌株のグラム染色により 2 グループに分けることができる。IS6100 のトランスポゼースはグラム陰性細菌である [15]-[19] のグループに最も近い系統的關係

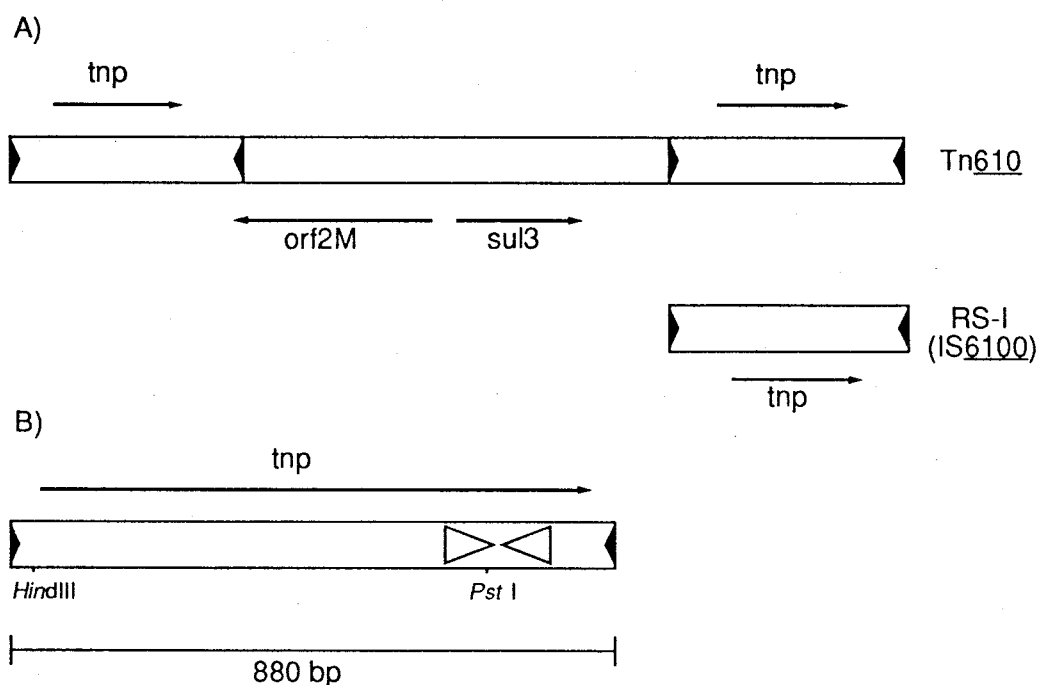


Fig. 2-5 Structural relationship of IS6100 and Tn610 sequence. (A) Structure of Tn610. The three ORFs, transposase gene (*tnp*, 765 bp), sulphonamide resistance gene (*sul3*, 849 bp) and cryptic ORF (*orf2M* deduced to be integrase gene, 909 bp), are shown by arrows. (B) Structure of RS-I (IS6100). The 14 bp inverted repeats at its termini are shown by triangles, and transposase is shown by arrow. Seventy eight base pairs inverted repeats in the transposase gene are indicated by open triangles.

係にあった。

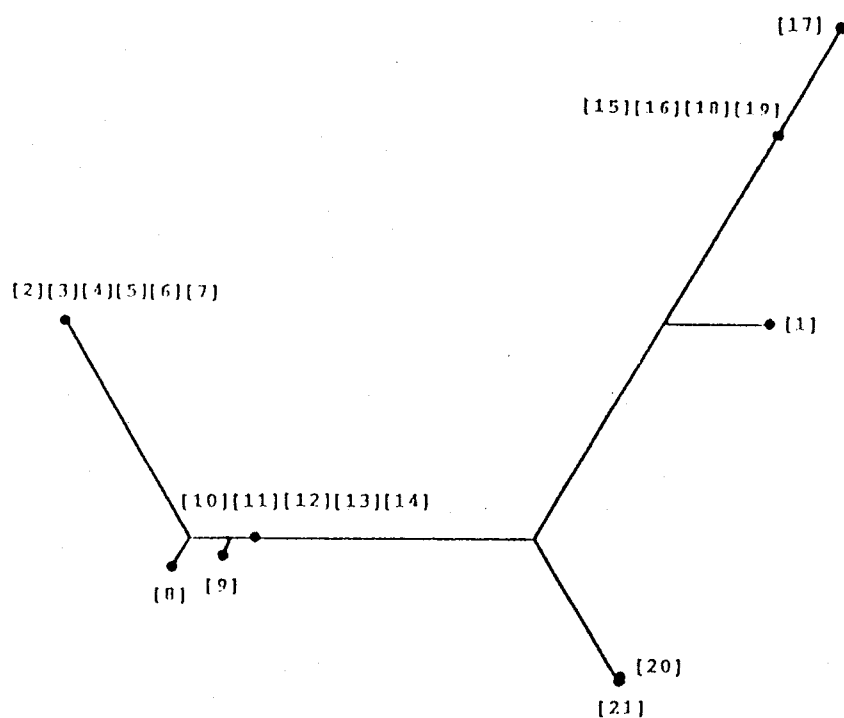
IS は、その重要な構成要素として、両端に IR を有する。Table 2-1 に示した他の IS の IR と IS6100 (RS-I) の両端の IR (IR-1) とを比較し、その相同性を調べた (Fig. 2-8)。IS6100 の IR と同じ塩基の箇所を網掛けで示した。他の IS が持つ IR の長さは 14 bp から 22 bp と多少差が見られるが、塩基配列は IS6100 の IR 全体にわたり 14 bp の範囲で 78-92% と非常に高い相同性が見られた。特に [15]-[19] のグループの IS の IR とは 1 塩基置換のみであった。以上のことにより IS6100 は、これら 20 種の IS が属する IS₆ ファミリーに位置づけられる IS であると推測される。

4 IS6100 の転写物の検出

IS6100 (RS-I) のコードするトランスポゼースが、pOAD2 を保持する *Flavobacterium* 内で発現しているかをノーザンハイブリダイゼーション法を用いて、転写レベルで調べた。KI723T1 株 (pOAD2 保持株) と KI723 株 (pOAD2 非保持株) から RNA の調製を行った。また、Table 2-2 に示すオリゴヌクレオチドを合成し、5' 末端の放射能ラベルを行い、プローブとして用いた。プローブ 1 は *nylA* (EI 遺伝子) の開始コドン Val の下流 250 bp の位置に mRNA に相補するように、プローブ 2 は *tnp* (トランスポゼース遺伝子) の開始コドン Met の下流 67 bp の位置に mRNA に相補するように、また、プローブ 3 は *tnp* の開始コドン Met の下流 732 bp の位置に mRNA

Table 2-1 Homologs of putative amino acid sequence of IS6100 transposase

Transposase		Homology (%)	Overlap (a.a)
IS 431L	<i>Staphylococcus aureus</i> plasmid pI524	41.1	214
IS 431R	<i>Staphylococcus aureus</i> plasmid pI524	41.1	214
IS 257R2	<i>Staphylococcus aureus</i> plasmid pSK1	47.6	214
IS 257R1	<i>Staphylococcus aureus</i> plasmid pSK1	47.6	214
IS 257L	<i>Staphylococcus aureus</i> plasmid pSK1	47.6	214
IS 431mec	<i>Staphylococcus aureus</i> plasmid pWDB2	42.1	214
ISS1W	<i>Lactococcus lactis</i> plasmid pWV05	44.7	213
ISS1N	<i>Lactococcus lactis</i> plasmid pSK111	45.0	213
IS 946	<i>Lactococcus lactis</i>	42.1	240
ISS1T	<i>Streptococcus lactis</i> plasmid pSK08, SK13	42.1	240
ISS1S-A	<i>Streptococcus lactis</i> plasmid pSK08, SK7-a	42.1	240
ISS1S-B	<i>Streptococcus lactis</i> plasmid pSK08, SK7-b	42.1	240
ISS1S-C	<i>Streptococcus lactis</i> plasmid pSK08, SK7	42.1	240
IS 15-R(P-22)	<i>Salmonella panama</i>	65.1	235
IS 176	<i>Salmonella typhimurium</i> plasmid NTP16	65.1	235
IS 15-R(P-21)	<i>Salmonella panama</i>	65.8	199
IS 15-ΔIV	<i>Salmonella typhimurium</i> plasmid pBP11	65.1	235
IS 240-A	<i>Bacillus thuringiensis</i> large plasmid	47.2	233
IS 240-B	<i>Bacillus thuringiensis</i> large plasmid	46.4	233



- | | |
|-------------------|--|
| [1] IS6100 | <i>Flavobacterium</i> sp.K172 plasmid pOAD2 |
| [2] IS431L | <i>Staphylococcus aureus</i> plasmid pI524 |
| [3] IS431R | <i>Staphylococcus aureus</i> plasmid pI524 |
| [4] IS257R2 | <i>Staphylococcus aureus</i> plasmid pSK1 (Tn4003) |
| [5] IS257R1 | <i>Staphylococcus aureus</i> plasmid pSK1 (Tn4003) |
| [6] IS257L | <i>Staphylococcus aureus</i> plasmid pSK1 (Tn4003) |
| [7] IS431mec | <i>Staphylococcus aureus</i> plasmid pWDB2 |
| [8] ISS1W | <i>Lactococcus lactis</i> plasmid pWV05 |
| [9] ISS1N | <i>Lactococcus lactis</i> plasmid pSK111 |
| [10] IS946 | <i>Lactococcus lactis</i> |
| [11] ISS1T | <i>Streptococcus lactis</i> plasmid pSK08,SK13 |
| [12] ISS1S-A | <i>Streptococcus lactis</i> plasmid pSK08,SK7-a |
| [13] ISS1S-B | <i>Streptococcus lactis</i> plasmid pSK08,SK7-b |
| [14] ISS1S-C | <i>Streptococcus lactis</i> plasmid pSK08,SK7 |
| [15] IS15-R(P-22) | <i>Salmonella panama</i> (Tn1525) |
| [16] IS176 | <i>Salmonella typhimurium</i> plasmid NTP16 (Tn4352) |
| [17] IS15-R(P-21) | <i>Salmonella panama</i> (Tn1525) |
| [18] IS15-ΔIV | <i>Salmonella typhimurium</i> plasmid pBP11 |
| [19] IS26 | <i>Proteus vulgaris</i> plasmid Rts1 (Tn2680) |
| [20] IS240-A | <i>Bacillus thuringiensis</i> large plasmid |
| [21] IS240-B | <i>Bacillus thuringiensis</i> large plasmid |

Fig. 2-7 Phylogenetic tree of IS. Phylogenetic relationship of the IS6100 were also analyzed by the maximum parsimony method, and the IS6100 was related to IS6 family.

[1]	IS 6100	GGCTCTGTTGCAA
[2]	IS 431L	AATTCTGGTTCTGTTGCAAAGT
[3]	IS 431R	TTCTGTTGCAAAGT
[4]	IS 257R2	GGTTCTGTTGCAAAGT
[5]	IS 257R1	GGTTCTGTTGCAAAGT
[6]	IS 257L	GGTTCTGTTGCAAAGT
[7]	IS 431mec	CGGTTCTGTTGCAAAGT
[8]	ISS 1W	GGTTCTGTTGCAAAGTTT
[9]	ISS 1N	GGTTCTGTTGCAAAGTTT
[10]	IS 946	GGTTCTGTTGCAAAGTTT
[11]	ISS 1T	GGTTCTGTTGCAAAGTTT
[12]	ISS 1S-A	GGTTCTGTTGCAAAGTTT
[13]	ISS 1S-B	GGTTCTGTTGCAAAGTTT
[14]	ISS 1S-C	GGTTCTGTTGCAAAGTTT
[15, 17]	IS 15-R	GGCACTGTTGCAA
[16]	IS 176	GGCACTGTTGCAA
[18]	IS 15-Δ IV	GGCACTGTTGCAA
[19]	IS 26	GGCACTGTTGCAA
[20]	IS 240-A	AGGTTCTGGTGCAA
[21]	IS 240-B	AAGGTTCTGGTGCAA

Fig. 2-8 Comparison of terminal inverted repeat of IS6100 and IS sequences. Terminal inverted repeat sequence of IS6100 were aligned with which of IS6 family.

Table 2-2 Syntesized oligonucleotide probes

	sequence	target
probe 1	5'-CTCTGTAGCCGGACTCTTTCATGCCTTTGATGCTGGACGT-3'	EI-mRNA
probe 2	5'-GCCAGCATTTCTCAAGGTCGCGATAGCTGATCGGATAGC-3'	Tnp-mRNA
probe 3	5'-CATGCTCAACCACCATTTTCGAGCAGCCGCTGATCGGCG-3'	anti-Tnp-mRNA

と同じ鎖で、それぞれ合成した。ノーザンハイブリダイゼーションの結果、本実験操作の陽性コントロールである EI-mRNA に対するプローブ（プローブ1）を用いると、pOAD2 保持株由来の RNA に対して強いシグナルが認められたが（Fig. 2-9, lanes 1, 2）、トランスポゼース-mRNA に対するプローブ（プローブ2）と、トランスポゼース-antisense-mRNA に対するプローブ（プローブ3）を用いた場合には、顕著なシグナルは検出されなかった（Fig. 2-9, lanes 3-6）。

次に、検出感度を上げるため、S1 マッピングの変法である溶液ハイブリダイゼーション法を用いて、転写物の検出を行った。各種プローブはノーザンハイブリダイゼーション法に用いたものと同じものを使用し、RNA については前処理したものを使用した。その結果、pOAD2 非保持株由来の RNA に対しては、プローブ1, 2, および3ともにシグナルは全く検出されなかったが（Fig. 2-10, lanes 2, 5, 8）、pOAD2 保持株由来の RNA に対しては、プローブ1だけではなく、プローブ2を使用した場合でも弱いシグナルが認められた（Fig. 2-10, lanes 3, 6）。このことから、IS6100 のトランスポゼースは pOAD2 を保持する *Flavobacterium* sp. KI723T1 株内で微弱に発

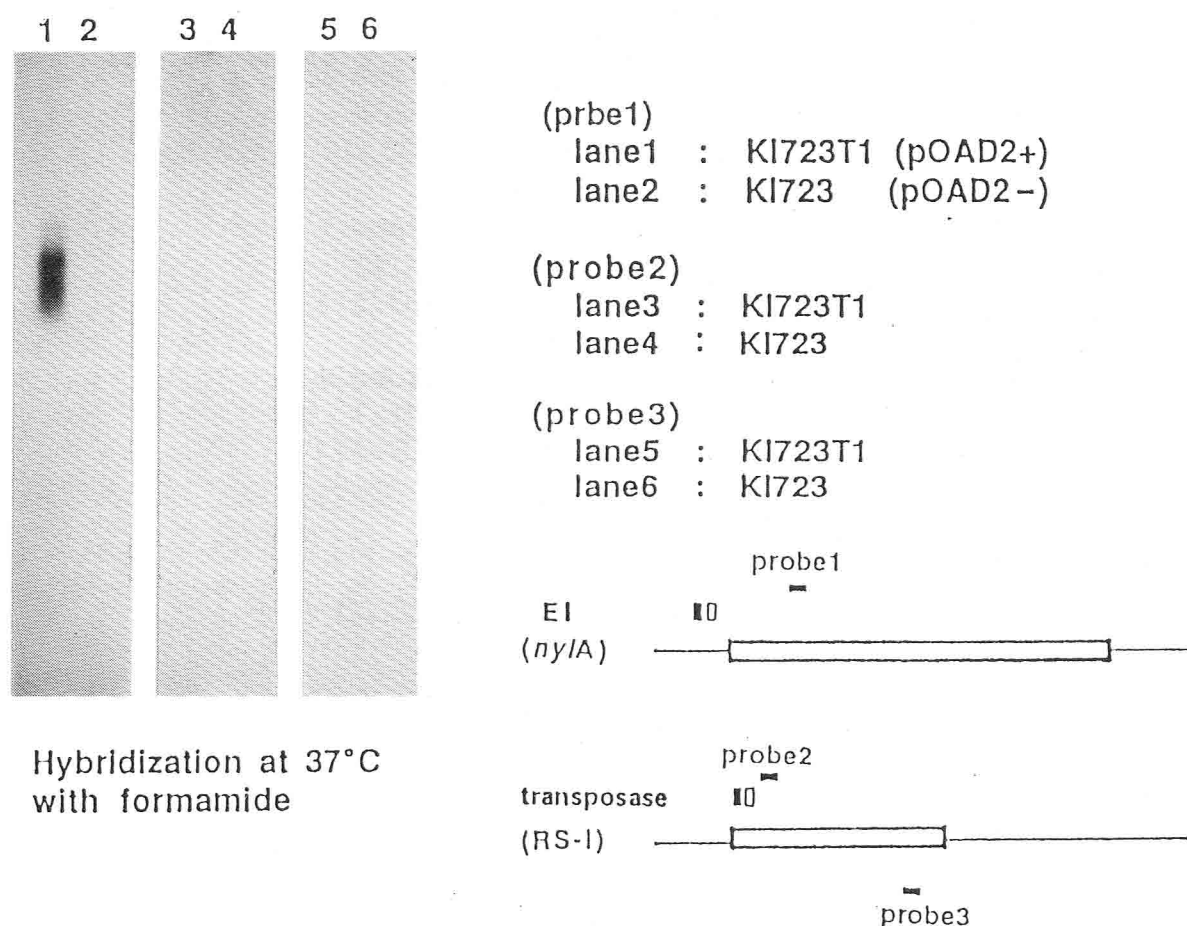


Fig. 2-9 Northern blot of RNA from *Flavobacterium* sp. KI723T1 and KI723.

現していると考えられる。しかし、プローブ3を用いた場合にも、弱いシグナルが認められた (Fig. 2-10, lane 9)。ここで、Fig. 2-10, lanes 1, 4, および 7 は、それぞれ未反応のプローブを泳動したものであり、各プローブの位置マーカーである。ここで得られたシグナルが RNA を調製する際に混入したプラスミド DNA による可能性もあるが、pOAD2 非保持株由来の粗 RNA 標品に精製 pOAD2 を過剰に加え、その後、pOAD2 保持株由来の粗 RNA 標品と同様に DNase I 処理を行い実験に用いたが、シグナルは検出されなかった。このことから、この実験で検出されたシグナルは、RNA 由来のものであり、持込みプラスミド DNA によるものではないと考えられる。

5 IS6100 の分布

ナイロンオリゴマー資化性菌 *Flavobacterium* sp. KI72 株が保持する pOAD2 から IS6100 を分離したが、他の細菌で IS6100 の分布を調べた。既に、IS6100 は *Mycobacterium fortuitum* と *Flavobacterium* sp. KI72 株に存在することがわかっている。そこで、KI72 株が保持する他

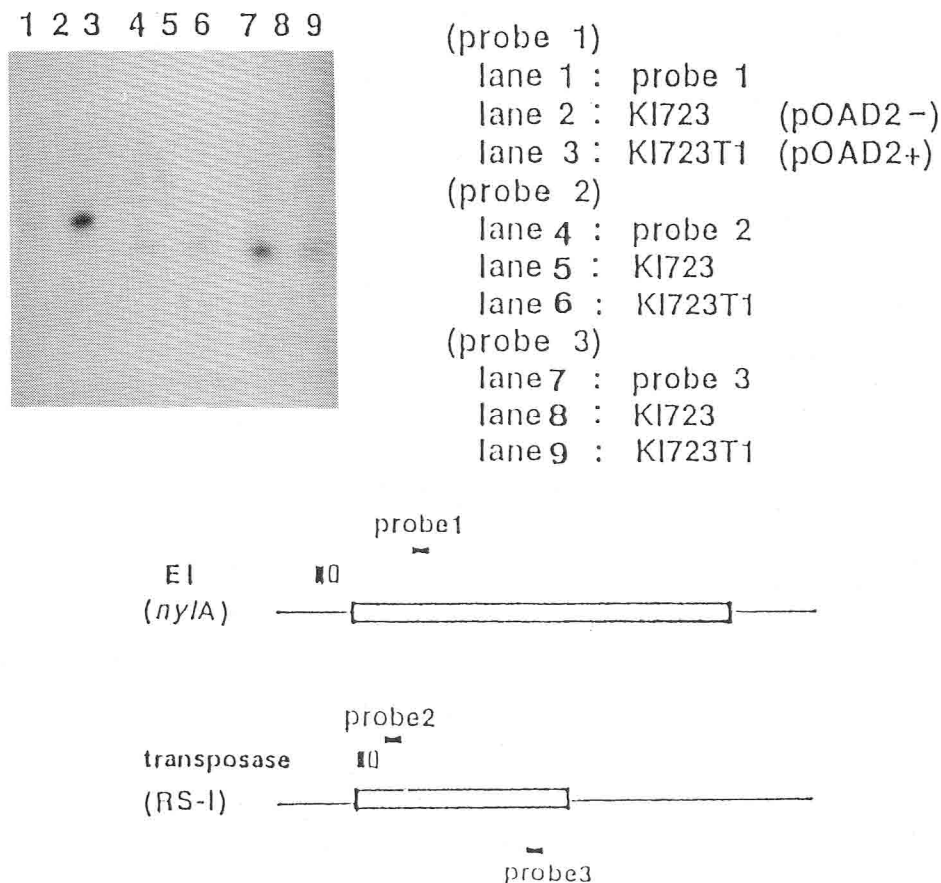


Fig. 2-10 Solution Hybridization of RNA from *Flavobacterium* sp. KI723T1 and KI723.

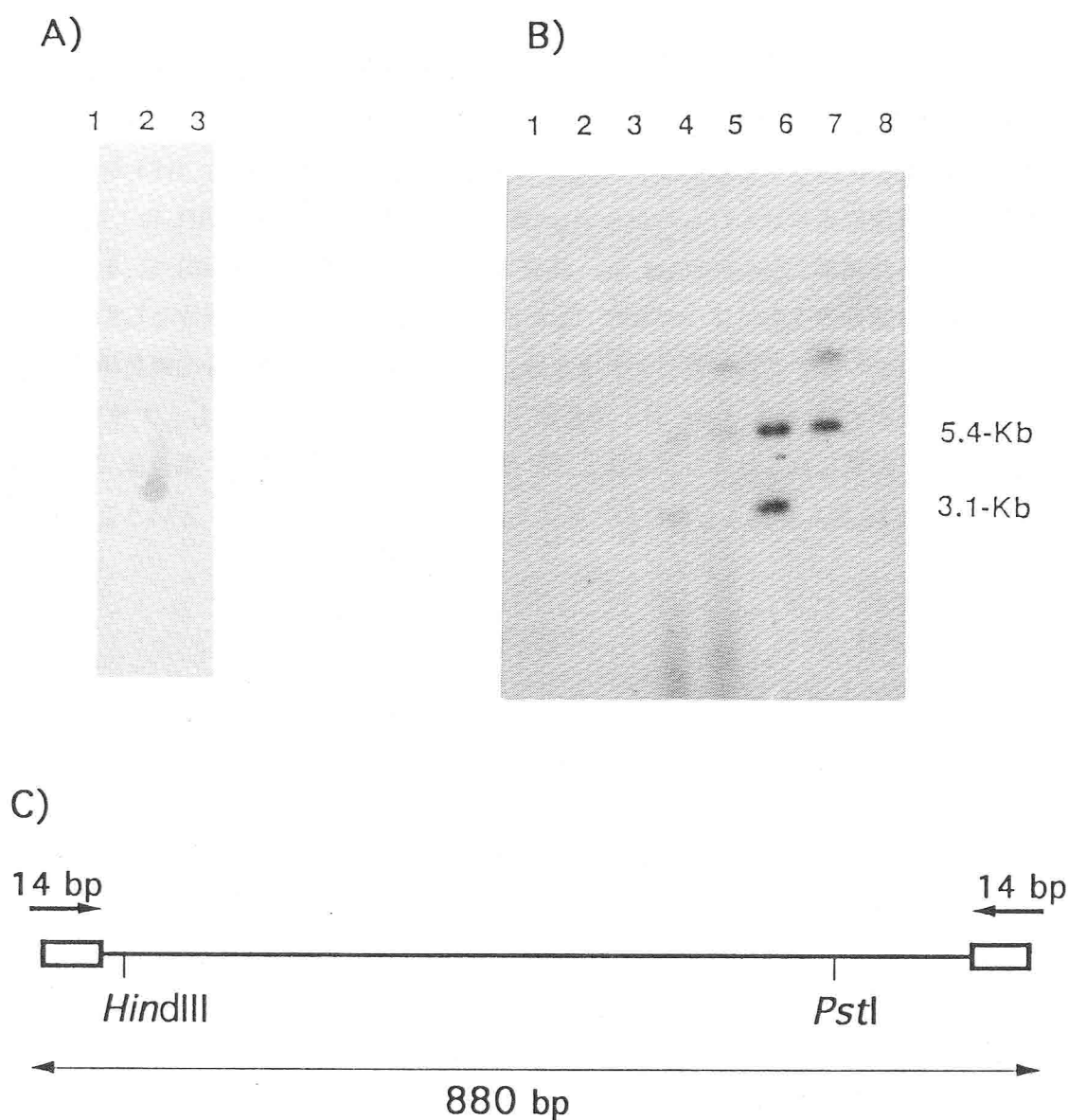


Fig. 2-11 Southern hybridization of plasmids and chromosomes using IS6100 probe. (A) Homologous DNA were detected using ^{32}P -labeled IS6100 probe (*Hind*III-*Pst*I fragment). Slot 1, plasmids (pOAD1 and pOAD3) from *Flavobacterium* sp. KI722 ; slot 2, plasmids from *Pseudomonas* sp. NK87; slot 3, λ -*Hind*III fragments. (B) Plasmid or chromosomal DNA was digested with *Hind*III (slots 1, 2, 4, 6, and 8) or *Eco*RI (slots 3, 5, and 7), and homologous fragments were detected using ^{32}P -labeled IS6100 probe. Slot 1, *E. coli* chromosome; slots 2 and 3, *Flavobacterium* sp. KI722 chromosome; slots 4 and 5, *Pseudomonas* sp. NK87 chromosome; slots 6 and 7, *Pseudomonas* sp. NK87 plasmids; slot 8 pUC19. (C) Structure of IS6100 regions. The 14 bp inverted repeats at its termini are shown by boxes.

のプラスミド (pOAD1 と pOAD3) と染色体 DNA、また、KI72 株と同じくナイロンオリゴマー資化性菌である *Pseudomonas* sp. NK87 株のプラスミド DNA と染色体 DNA、それから野生型に近い *E. coli* C600 株の染色体 DNA をそれぞれ調製し、サザンハイブリダイゼーション法で IS6100 の分布を調べた。プローブとしては、IS6100 中にある HindIII-PstI 断片 (693 bp) を pUC19 の HindIII-PstI 部位に組み込んだものから HindIII-PstI 断片を切り出し使用した。75°C でハイブリダイズを行った結果、*Pseudomonas* sp. NK87 株のプラスミド DNA (pNAD1-6) を HindIII で消化して得られる断片のうち、2つの断片 (5.4 kb と 3.1 kb) が陽性の断片として検出された (Fig. 2-11B, lane 7)。これと同じシグナルは *Pseudomonas* sp. NK87 株の染色体 DNA についても見られるが (Fig. 2-11B, lane 5)、染色体 DNA を EcoRI で消化しても、プラスミド DNA を EcoRI で消化した時に得られるシグナルと同じ挙動を示すので、これは、染色体 DNA を調製する際に、プラスミド DNA が混入したためと考えられる。5.4 kb と 3.1 kb という2つのシグナル

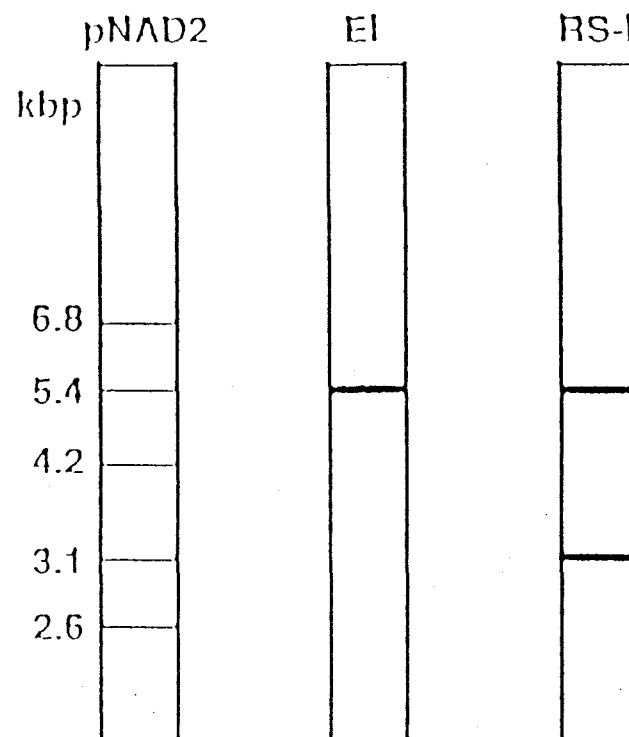


Fig. 2-12 Identification of the fragments detected at 75°C of hybridization with DNA probes of the pOAD2 in the pNAD2-HindIII fragments. HindIII digestion of the pNAD2 generated the 6.8, 5.4, 4.2, 3.1, and 2.6 kb fragments. Among these fragments, only 5.4 kb fragment was detected with the *nylA* probe, and two fragments which size were 5.4 and 3.1 kb, were detected with the IS6100 probe.

の位置は、NK87 株が持つ pNAD2 プラスミドを精製し、HindIII で消化して得られる 5 断片のうち大きい方から 2 番目と 4 番目の断片と一致する⁸⁾ (Fig. 2-12)。また、NK87 株の持つ 6 種類のプラスミドを制限酵素で切断せずに実験に用いた場合には、pNAD2 の位置にシグナルが検出できる (Fig. 2-11A, lane 2)。以上のことより、Pseudomonas sp. NK87 株が保持する pNAD2 プラスミド上の少なくとも 2 ヶ所に IS6100、もしくは IS6100 に高い相同性を持つ配列が存在することがわかった。この pNAD2 上には、Pseudomonas sp. NK87 株が持つ P-nylA が存在する。P-nylA は、Flavobacterium sp. KI72 株の pOAD2 がコードする F-nylA と 98% の相同性がある⁵¹⁾。これと対象的に、Flavobacterium sp. KI72 株が pOAD2 以外に保持する 2 つのプラスミド (pOAD1 と pOAD3) (Fig. 2-11A, lane 1), Pseudomonas sp. NK87 株が pNAD2 以外に保持する 5 つのプラスミド (pNAD1, 3, 4, 5, および 6) (Fig. 2-11A, lane 2), KI72 株, NK87 株ならびに E. coli の染色体 DNA (Fig. 2-11B, lanes 1-5) については、IS6100 と相同性を示すシグナルは検出されなかった。これらの結果は、Fig. 2-11 で示す実験よりもゆるやかな結合条件で実験した場合でも同様に検出できなかった。



Fig. 2-13 Comparison of the amino acid sequences of EII and EII' enzymes

6 EII' から EII への活性上昇に関与するアミノ酸残基置換の同定

RS-II 領域については、RS-IIA が EII 構造遺伝子, nylB, を含み、RS-IIB は EII 類似酵素 EII' をコードしている nylB' を含んでいる⁷⁾。EII (6-aminoheptanoate dimer hydrolase) は、6-aminoheptanoate linear dimer (Ald と略す) を加水分解し 6-aminoheptanoate (Ahx と略す) を生じる活性を持つ^{2,3)}。2つの遺伝子 (nylB と nylB') は非常によく似ており、両遺伝子とも 1176 bp よりなり、ORF 内で 12% が異なるのみで読み取り枠に変化は認められない。また、両者

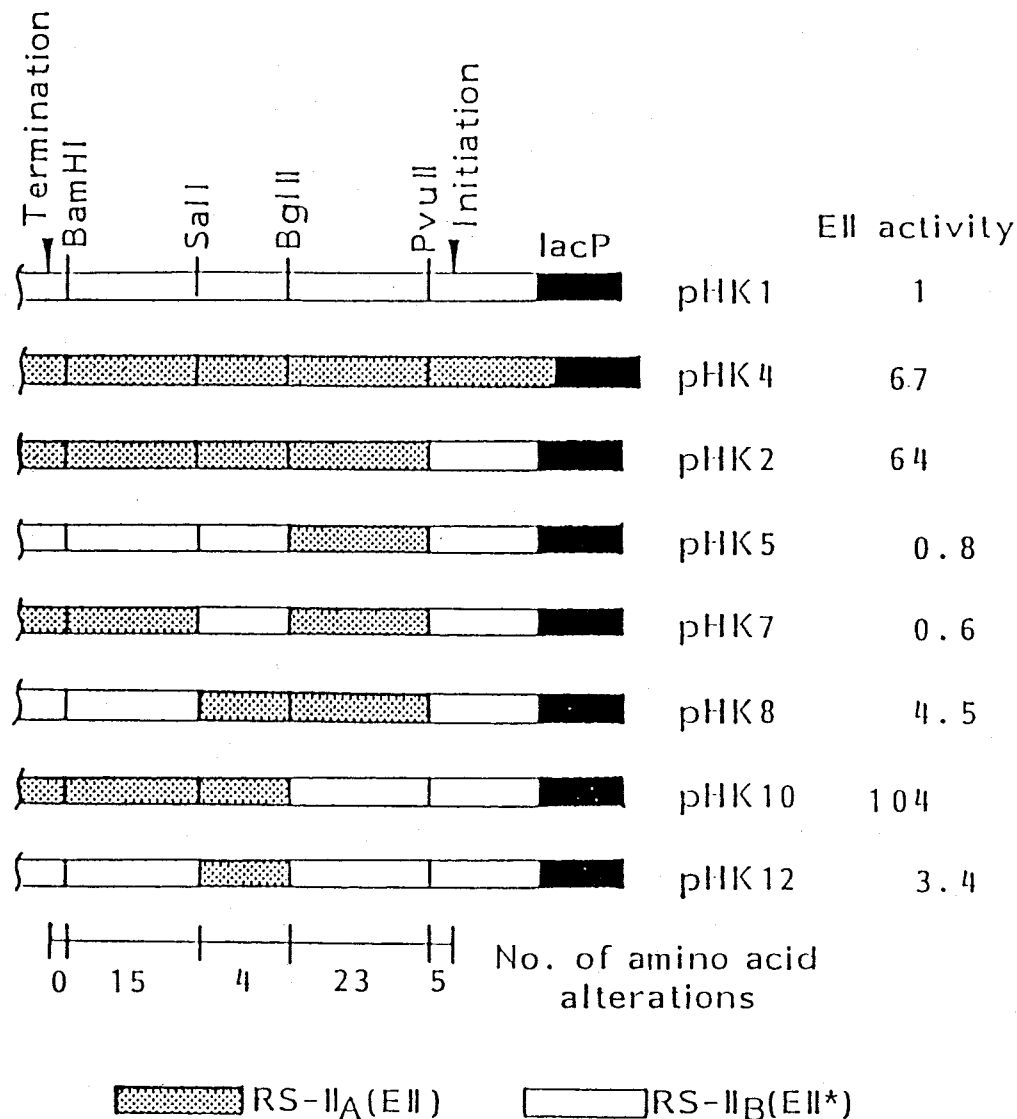


Fig. 2-14 Plasmid structures and the enzyme activities of cell extracts from *E. coli* harboring the hybrid plasmid. The antigenicity of each hybrid protein was measured by quantitative immunoelectrophoresis using anti-EII serum. The specific enzyme activity was defined as enzyme activity/quantitative antigenicity and expressed as the ratio toward the activity of *E. coli* (pHK1).

は 392 アミノ酸残基よりなるが、47 残基が異なるのみである (Fig. 2-13)。この 47 残基が異なるため、EII' の Ald 分解活性は EII の約 1% である⁷⁾。nylB と nylB' 間で保存されている制限酵素部位 (PvuII, BglII, SalI, BamHI, SstII, NaeI, および Sau3AI) を利用し、対応するペプチド部分を交換して生じるハイブリッド酵素を作成し、Ald に対する分解活性を約 100 倍上昇させるのに関与するアミノ酸残基を評価すると、BglII-SalI 領域の Gly¹⁸¹→Asp¹⁸¹ (EII'→EII) アミノ酸残基置換と SalI-Sau3AI 領域が、活性上昇に関与している^{38,52)} (Fig. 2-14, 2-15,

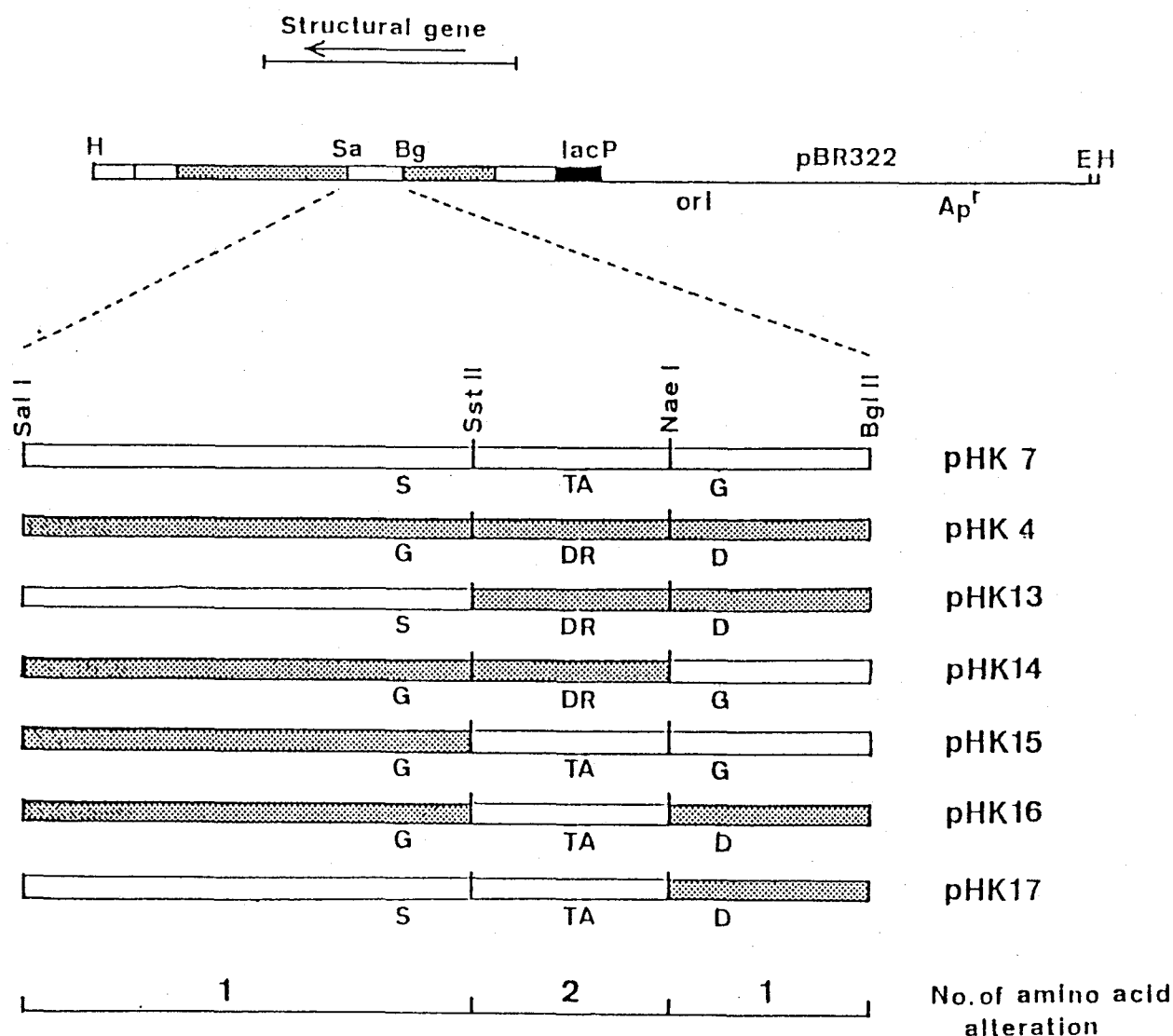


Fig. 2-15 Structure of the hybrid plasmids. Vector pBR322 is shown as solid line. Shadow boxes indicate the DNA region originated from nylB and its linked region; open boxes indicate the DNA region from nylB' and its linked region. Amino acids different between the EII and EII' are shown as one-letter codes under the map.

2-16)。BglII-SalI 領域が活性に必須であり、SalI-Sau3AI 領域がその活性上昇効果を増幅している。SalI-Sau3AI 領域には、2つのアミノ酸残基置換が存在している。そこで、合成 DNA 断片を用いて、SalI-Sau3AI 領域にアミノ酸残基置換に相当する変異を1ヶ所ずつ導入し、いずれがAld 分解活性の上昇に関与しているかを決定した。SalI-Sau3AI 領域にある SalI と BssHII の切断部位を両端に持つ 60 bp の DNA 断片を合成した。DNA 合成装置の性能も考え、30 mer ずつ計6本のオリゴヌクレオチドを合成した (Table 2-3) (Fig. 2-17)。合成した変異 EII 遺伝子断片を pHK8Δ11 に導入する手順を Fig. 2-18 に示す。pHK8Δ11 は、PvuII-SalI 領域が nylB 型である nylB' を基本とするハイブリッドプラスミドである (Fig. 2-16)。また、pHK8Δ11 が産生する酵素活性は EII' のその約4倍であり、SalI-BamHI 領域の SalI-Sau3AI 領域を nylB の該当する部分と交換すると、BglII-SalI 領域に存在する Gly¹⁸¹→Asp¹⁸¹ アミノ酸残基置換による活性上昇効果を約20倍以上増幅することがわかっている⁵²⁾ (Fig. 2-16)。

そこで、SalI-Sau3AI 領域に存在する2アミノ酸残基置換 Thr²⁵⁹→Gln²⁵⁹, His²⁶⁶→Asn²⁶⁶ (EII'→EII) に相当する変異を一方ずつ pHK8Δ11 に導入し活性上昇の有無を調べた。pHK8Δ11 を SalI と BamHI で完全に分解し、1% アガロースゲル電気泳動により、3.7 kb の断片を回収し

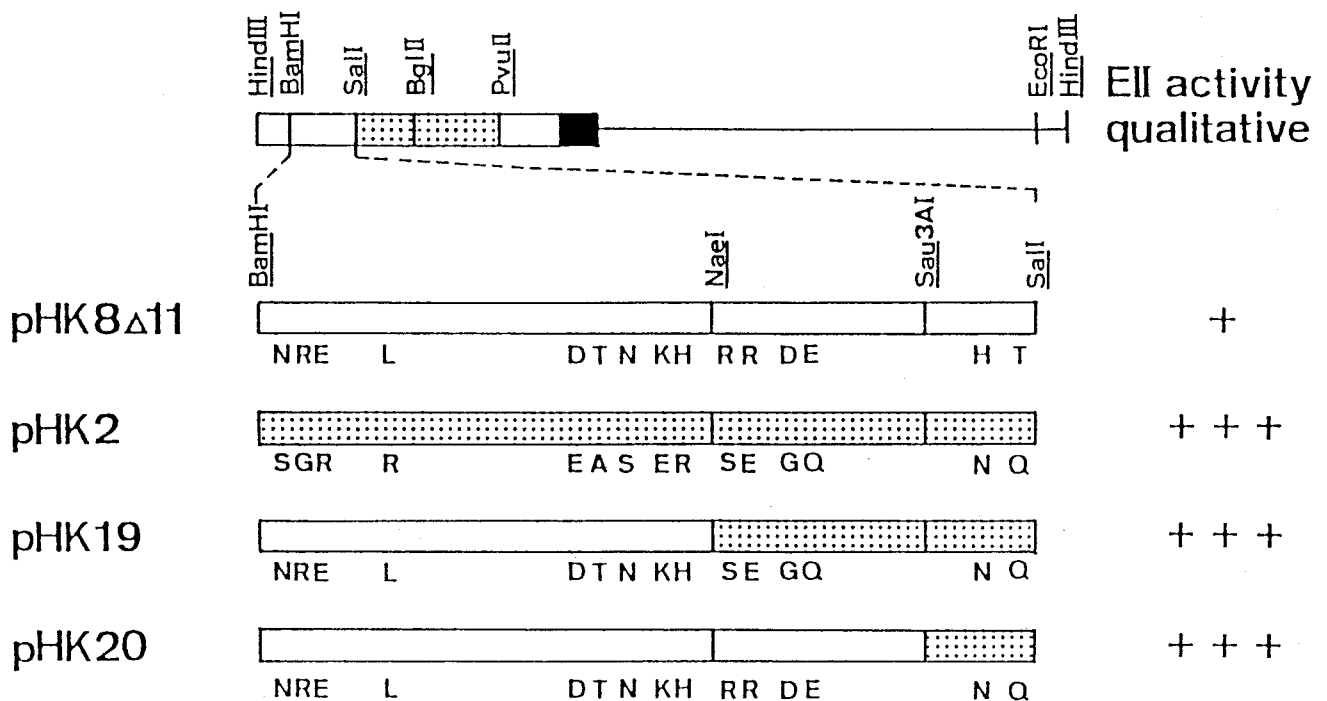


Fig. 2-16 The relationship between the plasmid structure and EII activity. Qualitative assay was carried out by using cell extracts of *E. coli* C600 harboring the hybrid plasmid. Reaction product, 6-aminohexanoate, was detected on TLC. Amino acids differing between EII and EII' enzymes are shown as one-letter codes.

Table 2-3 Synthesis of mutated oligonucleotides

	sequence
strand 1	5'-TCGACCAGAC GGGTTTCGGC TTCGCGCACG-3'
strand 2	5'-GGTCTGCCCA AAGCCGAAGC GCGTGCCGCC-3'
strand 3	5'-GCGGCGTCTC CTGCACGGCG CGGGATCTCG-3'
strand 4	5'-GCAGAGGACG TGCCGCGCCC TAGAGCGCGC-3'
strand 5	5'-TCGACACGAC GGGTTTCGGC TTCGCGAACG-3'
strand 6	5'-GTGCTGCCCA AAGCCGAAGC GCTTGCCGCC-3'

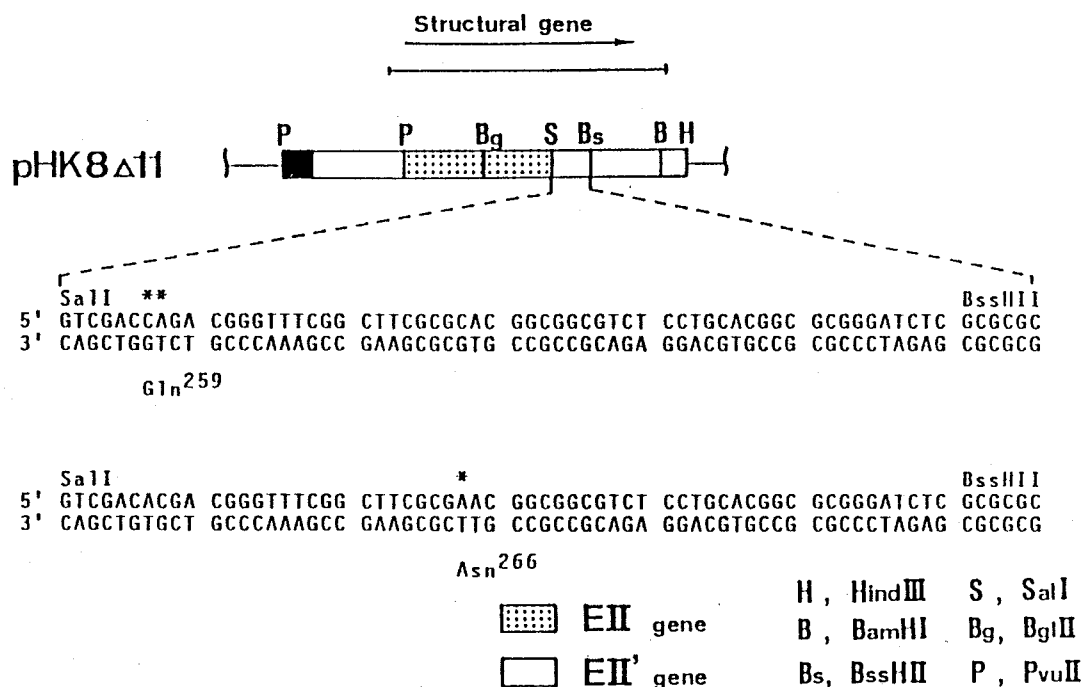


Fig. 2-17 Synthesis of mutated DNA fragments. *, the position of substituted nucleotides between nylB and nylB'.

た。同様に、pHK8 Δ 11 を *Bam*HI と *Bss*HII で完全に分解し、5% ポリアクリルアミドゲル電気泳動により、310 bp の断片を回収した。この2断片と合成した変異 EII (*Sal*I-*Bss*HII) 遺伝子断片を連結したのち、*E. coli* C600 株に導入した。得られた組換えプラスミドは His²⁶⁶→Asn²⁶⁶ (EII'→EII) に相当する変異を含むものを pHK21、Thr²⁵⁹→Gln²⁵⁹ (EII'→EII) に相当する変異を含むものを pHK22 とそれぞれ命名した。

E. coli C600 株の pHK8 Δ 11、pHK21 および pHK22 保持菌から、それぞれ粗酵素液を調製した。この粗酵素液と基質である Ald (20 mM) とを反応させ、薄層クロマトグラフィーで展開を行い、これらのハイブリッドタンパク質の Ald 分解活性を比較した (Fig. 2-19)。その結果、pHK22 保持菌は、pHK8 Δ 11 保持菌と同程度の Ald 分解活性を示したが、pHK21 保持菌の Ald 分解活性は、これら2種類の保持菌よりも顕著に高い活性を示した。His²⁶⁶→Asn²⁶⁶ (EII'→EII) の置換は活性上昇効果を示したが、Thr²⁵⁹→Gln²⁵⁹ (EII'→EII) の置換では上昇効果が見られなかったことより、*Sal*I-*Sau*3AI 領域に存在する2アミノ酸残基置換のうち、His²⁶⁶→Asn²⁶⁶ (EII'→EII) の置換が、Ald 分解活性の上昇に関与していることがわかった。

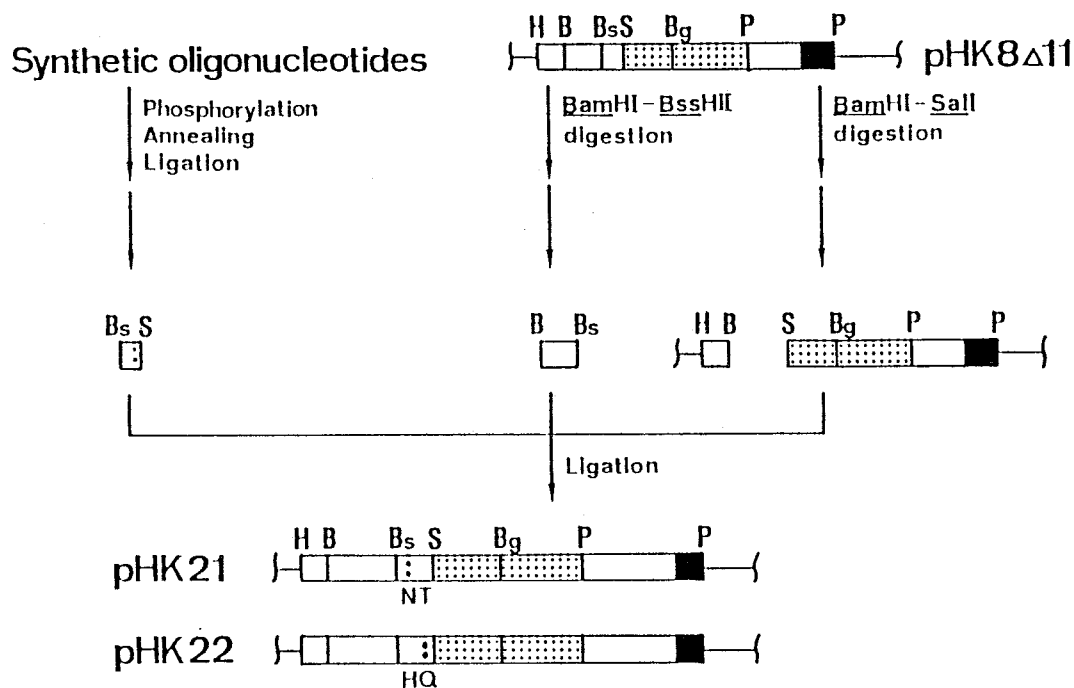
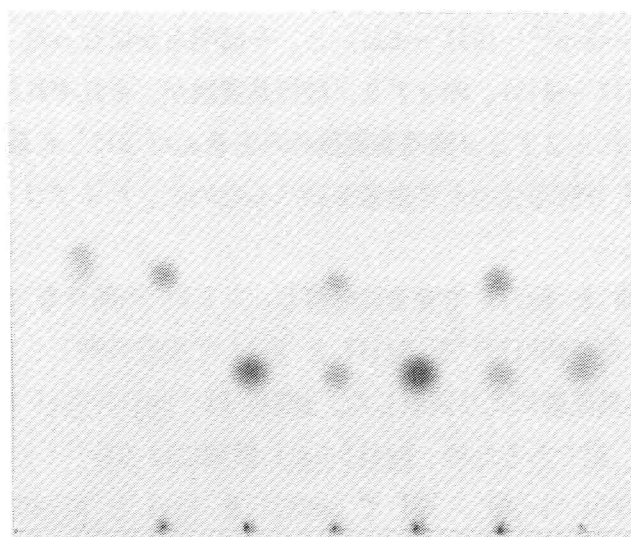


Fig. 2-18 Construction of hybrid plasmids pHK21 and pHK22.



Lane; 1 5 mM Ald
 2 pHK 1
 3 pHK 4
 4 pHK 8 Δ 11
 5 pHK 21
 6 pHK 22
 7 5 mM Ahx

1 2 3 4 5 6 7

Fig. 2-19 Detection of EII activity on TLC. Qualitative assay was carried out by using cell extracts of *E. coli* harboring the hybrid plasmid. Reaction product, 6-aminohexanoate (Ahx), in extracts from *E. coli* harboring pHK4 and pHK21 were detected on TLC.

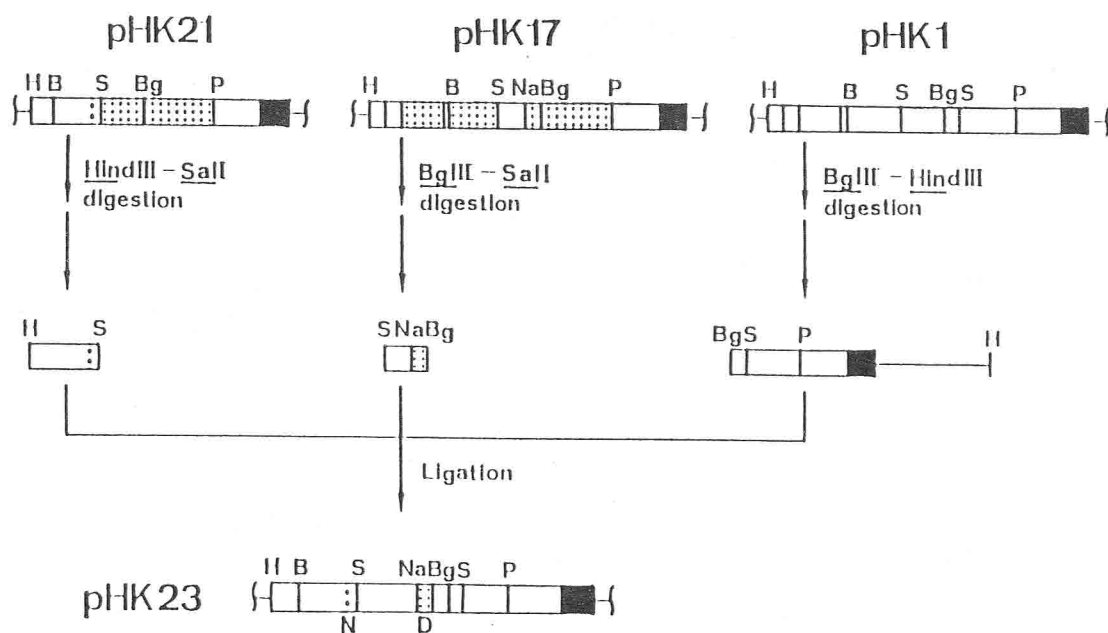


Fig. 2-20 Construction of hybrid plasmid pHK23.

7 ハイブリッドプラスミド pHK23 の構築と活性測定

EII' と EII との間に存在する 47 個のアミノ酸残基置換のうち、以前の研究で明らかになった BglII-SalI 領域に存在する Gly¹⁸¹→Asp¹⁸¹ (EII'→EII) と、今回明らかになった SalI-BamHI 領域に存在する His²⁶⁶→Asn²⁶⁶ (EII'→EII)、の 2 アミノ酸残基置換が、それぞれ活性上昇に関与することがわかった。そこで、これら 2 アミノ酸残基置換のみを含んだ EII' を基本とするハイブリッド酵素を作成し、その Ald 分解活性の上昇効果を調べるために、ハイブリッドプラスミド pHK23 を構築した (Fig. 2-20)。

nylB' よりなる pHK1³⁸⁾ を HindIII と BglII で完全に分解し、3.2 kb の断片を 1% アガロースゲル電気泳動により回収した。また、pHK17⁵³⁾ を BglII と SalI で完全分解し、290 bp の断片を 8% ポリアクリルアミドゲル電気泳動により回収した。pHK17 は、BglII-NaeI 領域が nylB 型であり、NaeI-SalI 領域が nylB' 型であるため、BglII-SalI 領域には Gly¹⁸¹→Asp¹⁸¹ のアミノ酸残基置換しか含んでいない (Fig. 2-20)。そして、pHK21 を SalI と HindIII で完全に分

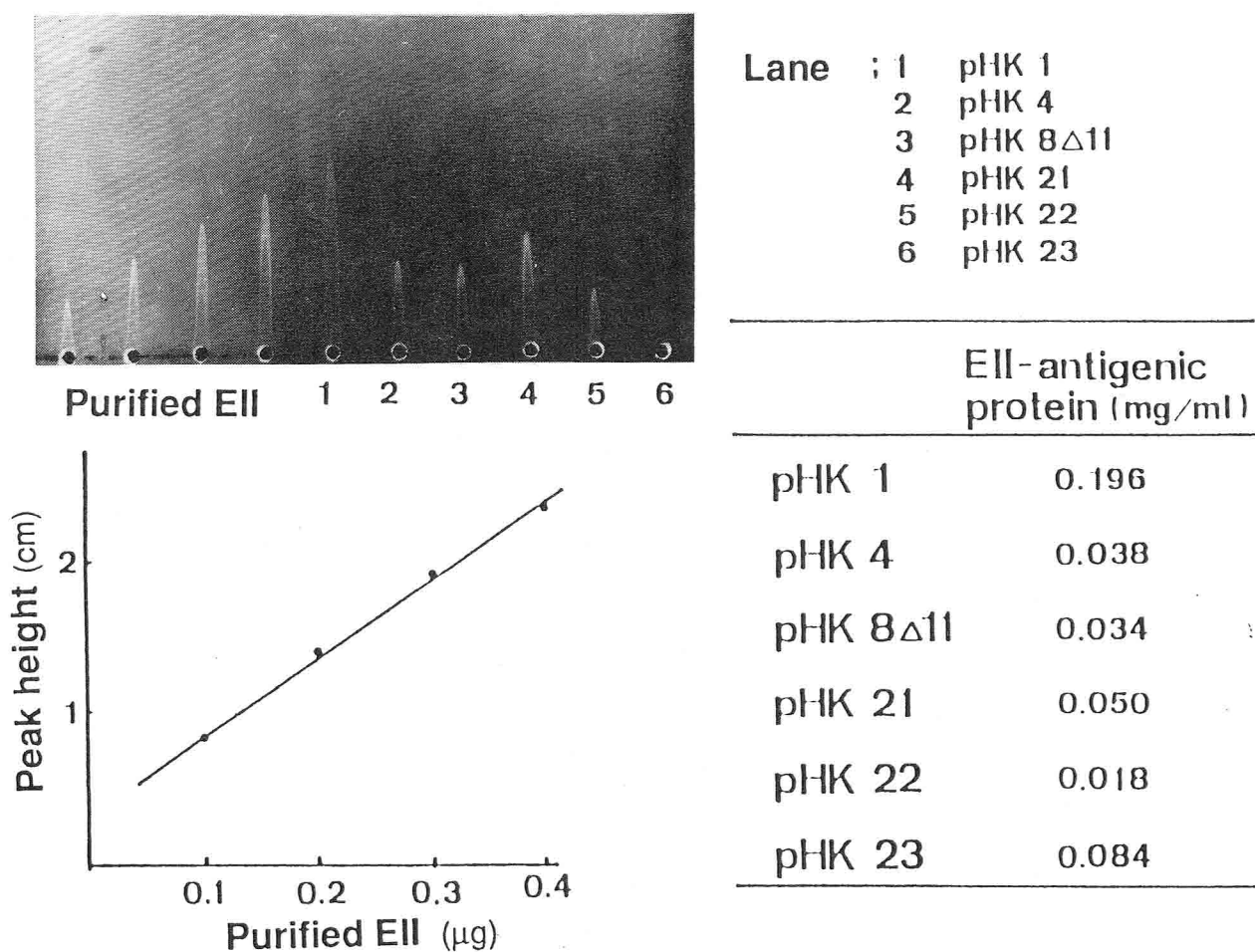


Fig. 2-21 Immunoelectrodiffusion test.

解し、His²⁶⁶→Asn²⁶⁶ (EII'→EII) に相当する変異のみを含んだ 470 bp の断片を 1.5% アガロースゲル電気泳動により回収した。これら得られた3断片 (HindIII-BglII 断片, BglII-SalI 断片, および SalI-HindIII 断片) を連結したのち、*E. coli* C600 株を用いて、pHK23 DNA を得た。

各プラスミド保持株からの粗酵素液で Ald 分解活性を測定するにあたって、粗酵素液中の変異 EII タンパク質を免疫電気拡散法を用いて求めた。抗 EII 血清と反応性を有するタンパク質を、精製 EII タンパク質を標準試料として測定した (Fig. 2-21)。Fig. 2-21 の左の4スロットは、既知濃度の精製した EII タンパク質 (0.1-0.4 μ g) を負荷したもので、タンパク質量を横軸に、ピークの高さを縦軸にとると、両者の間には直線関係が見られる。これを検量線として、各粗酵素液中の EII 抗原量を求めた。

各プラスミド保持株の粗酵素液中の Ald 分解活性は、高速液体クロマトグラフィー (HPLC) を用いて、Ald の分解速度から求めた。pHK1, pHK4, pHK8 Δ 11, pHK21, pHK22, および pHK23 の各

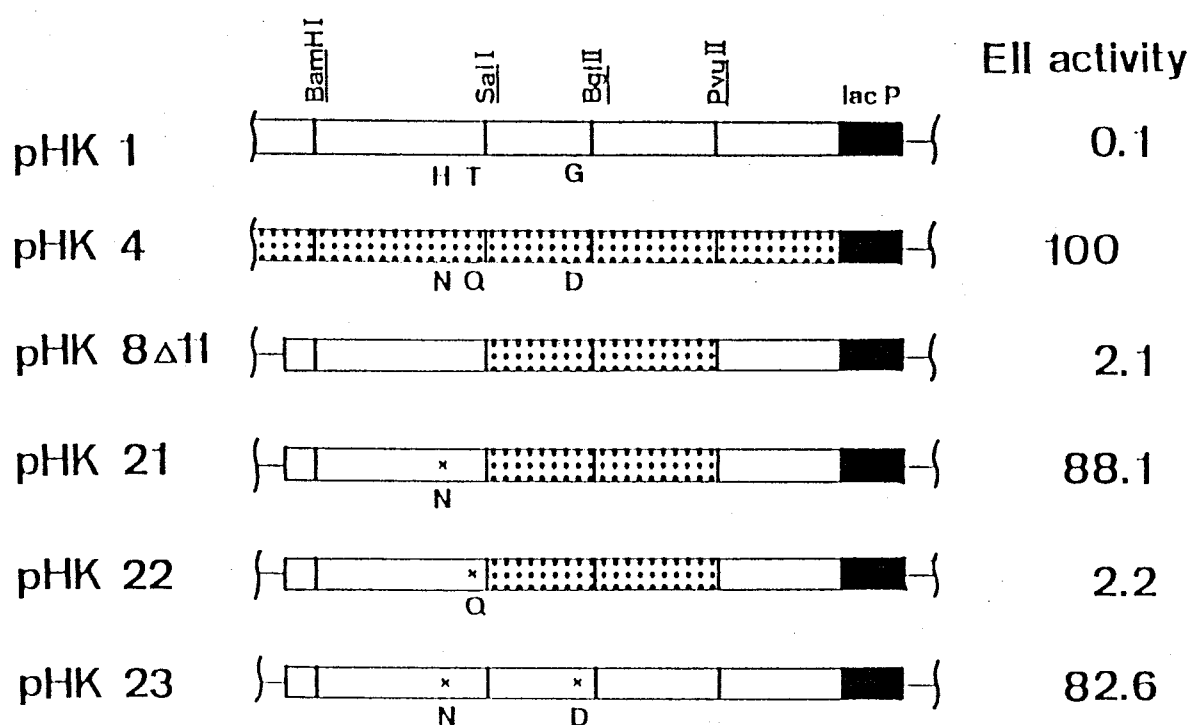


Fig. 2-22 Relationship between the plasmid structure and the enzyme activity. Enzyme activities in the cell extracts of *E. coli* harboring each pHK plasmid toward 6-amino-hexanoate dimer were examined by HPLC. Decrease of 6-amino-hexanoate dimer by the enzyme reaction was monitored by reverse-phase HPLC. The amount of antigenic EII proteins was measured by Rocket quantitative immunoelectrophoresis. The activities were expressed as percentages of the activities/amount of antigenicity towards that of *E. coli* (pHK4, EII producing). Amino acids different between EII and EII' are shown in the one-letter codes under the map.

ハイブリッドプラスミド保持株由来の粗酵素液について活性を測定し、各ハイブリッドタンパク質の比活性を EII 抗原量当りの活性として求めた。pHK1³⁸⁾は nylB' を、pHK4³⁸⁾は nylB を、それぞれ含んでおり、活性は 100 倍以上異なる。この pHK4 より生産される EII 酵素の活性を 100 として各ハイブリッドタンパク質の相対活性を Fig. 2-22 に示す。pHK21 より生産されるハイブリッドタンパク質は、BglIII-SalI 領域に存在するアミノ酸残基置換 Gly¹⁸¹→Asp¹⁸¹ による活性上昇を増幅する効果がみられ、pHK4 由来の EII と同程度の活性であった。しかし、pHK22 より生産されるハイブリッドタンパク質は、この増幅効果が見られず、pHK8△11 由来の EII 活性と同程度の活性であった。また、BglIII-SalI と SalI-BamHI 領域に存在するアミノ酸残基置換 Gly¹⁸¹→Asp¹⁸¹ と His²⁶⁶→Asn²⁶⁶ (EII'→EII) のみを含んだ pHK23 より生産されるハイブリッドタンパク質の Ald 分解活性は、EII レベルまで上昇していることがわかる。

以上のことにより、EII'-EII 間に存在する 47個のアミノ酸残基置換のうち Gly¹⁸¹→Asp¹⁸¹ と His²⁶⁶→Asn²⁶⁶ のアミノ酸残基置換のみで EII' レベルから EII レベルへの活性上昇が得られることがわかった。

第三節 考察

第一章の結果をもとに、pOAD2 上で5ヶ所に点在する RS-I 領域 (RS-IA, B, C, D, および E) を特定することができた。各 RS-I 領域の塩基配列を比較した結果、RS-IA, RS-IC, RS-ID, および RS-IE は 880 bp、RS-IB は 1300 bp からなり、RS-IA, C, D, および E の全領域と RS-IB の前半 880 bp の領域は、完全に同一の塩基配列を有していた (Fig. 2-2)。サザンハイブリダイゼーションの結果から^{24, 25)} 相同性が高いことは予想されたが、完全に一致していた。

DNA のデータベースに対し、RS-I 領域 (880 bp) について相同性検索を行った結果、*Mycobacterium fortuitum* 由来の IS6100 と全く同一の配列であった⁴⁰⁾。*Mycobacterium fortuitum* において、IS6100 はスルホンアミド耐性遺伝子を持つトランスポゾン Tn610 の一部として存在しているが (Fig. 2-5)、*Flavobacterium* sp. KI72 株が持つ pOAD2 上には Tn610 は認められなかった。

他の IS と IS6100 がコードするトランスポゼースを比較すると、IS6 ファミリーに属するものと相同性があり (Table 2-1) (Fig. 2-7)、その中でもグラム陰性細菌のグループの IS ([15]; IS15-R(P-22)⁴⁷⁾, [16]; IS176⁴⁸⁾, [17]; IS15-R(P-21)⁴⁷⁾, [18]; IS15-△IV⁴⁹⁾, [19]; IS26⁵⁴⁾) と特に高い相同性があった。IS が転移する場合、目標部位の切断などに重要な働きをするアミノ末端領域のアルギニン-チロシン残基 (IS6100 では 23-24番目)、カルボキシル末端領域のリジン-チロシン残基 (IS6100 では 165-166番目) が保存されている^{55, 56)}。Fig. 2-7 からわかるが、IS6 ファミリーに属する IS 群でもサブクラスにわけると、同種の宿主域であることがわかる。*Flavobacterium* sp. KI72 株は、当初通常のグラム染色による分類⁶⁾では、グラム陰性細菌に属するとされていたが、Takeuchi と Yokota の化学分析に基づく分類⁵⁷⁾によると、

KI72 株はむしろ、グラム陽性細菌である *Arthrobacter* に分類される。KI72 株がグラム陽性細菌であるとする、*Mycobacterium fortuitum* もグラム陽性細菌であるので、IS6100 はグラム陽性細菌に保存されていることになる。また、サザン法を用いた解析から、KI72 株と同じくナイロンオリゴマー資化性菌である、*Pseudomonas* sp. NK87 株が保持する pNAD2 プラスミド上にも、少なくとも2ヶ所の IS6100 もしくは、それと高い相同性を持つ配列が存在することがわかった (Fig. 2-11)。NK87 株はグラム陰性細菌であることから、IS6100 は比較的広い宿主域を持ち、また、KI72 株と *Mycobacterium fortuitum* で全く同じ配列であることから、その広がりには比較的近年であったと考えられる。このことは *E. coli* 由来の R-プラスミド N3 の *EcoRII* 遺伝子⁵⁸⁾の上流や、*Salmonella typhimurium* 由来の pBP11 プラスミドの IS15- Δ IV⁴⁹⁾の近接する上流に IS6100 様配列の一部が存在することからも示唆される。IS はその重要な構成要素として末端に IR を有する。この領域は他の IS でもよく保存されており、IS1⁵⁹⁾で見られるようにトランスポゼースにより認識される部位と考えられる。

Martin らは、Tn610 を複製能を欠失させたプラスミドに組み込み、*Mycobacterium smegmatis* 中での転移を調べ⁴⁰⁾、染色体 DNA のさまざまな場所に Tn610 が転移すること、また、それとは別の場所に IS6100 が単独で転移することを報告している。このことは Tn610 が *M. smegmatis* で転移能を持ち、トランスポゼースが発現していることを示している。そこで、IS6100 の KI72 株内での発現を転写レベルで調べると、微弱ながらその発現が認められた (Fig. 2-10)。また、その anti-sense 鎖の存在も微弱ながら認められた。これは、pOAD2 上の5ヶ所に点在している IS6100 のいずれかの下流からの read-through-mRNA を検出したものと考えられる。IS6100 の発現として検出したシグナルも上流からの read-through であったとも考えられる。KI72 株の染色体 DNA、pOAD2 以外のプラスミド DNA 上に IS6100 が存在しないことや (Fig. 2-11)、Martin らの実験では、Tn610 全体を実験に使っており IS6100 そのものを用いていないこと、通常 IS の発現は、宿主因子により制御されることが知られていることなどから、KI72 株が持つ IS6100 が実際に発現し、また、転移能を持つことについては、他の実験をする必要がある。

IS やトランスポゾンなどの移動する DNA 単位は、欠失、挿入、置換、重複を通して、染色体 DNA の再構成を行う能力を持つ⁶⁰⁾。pOAD2 上の RS-IA, RS-IB, RS-IC, RS-ID, および RS-IE によって囲まれた領域をそれぞれ領域1, 領域2, 領域3, 領域4 および領域5に分けて考えると (Fig. 2-23B)、各領域にはそれぞれ特徴がある。1) 各領域で G+C 含量に差が認められる (Fig. 2-23C)。60% から 70% まで。2) 領域4が欠失した変異型プラスミドが得られている。3) 領域1と領域2に存在する *nylB* と *nylB'* 領域 (RS-II領域) が重複している。4) RS-IB は、後半部 420 bp が重複した構造を持ち、同様のことは IS15- Δ IV⁴⁹⁾でも認められる。これらのことから、IS6100 が現在、KI72 株中で転移能を持つ IS であるかは未確認だが、pOAD2 の形成には関与したと考えられる。更に、RS-IA, B, C, D, および E のすべてにおいて、880 bp もの領域が完全に保存されていることから、それらの転移、重複が起こったのは比較的近年のことと思われる。

ナイロンオリゴマー分解酵素、6-aminohexanoate dimer hydrolase (EII) とその類似タンパク

質 (EII') は、392アミノ酸残基よりなるが、この2酵素の間で 47アミノ酸残基が異なるため、分解活性は、EII' が EII の約 1% となっている⁷⁾。どのアミノ酸残基置換が、活性の違いに関与しているかを、両者の ORF 内で保存されている制限酵素切断部位で組換え、ハイブリッドタンパク質を作成して調べた結果、EII-EII' 間に存在する 47アミノ酸残基置換のうち、2アミノ酸残基置換 Gly¹⁸¹→Asp¹⁸¹, His²⁶⁶→Asn²⁶⁶ により活性が EII' レベルから EII レベルへ上昇することがわかった。Gly¹⁸¹→Asp¹⁸¹ の置換が EII レベルの活性に必須であり、His²⁶⁶→Asn²⁶⁶ の置換がその効果を高めている。しかし、ここで行った実験方法では、すべてのアミノ酸残基置換の組合せを網羅できず、ここに示した2アミノ酸残基置換以外のアミノ酸残基置換が活性に影響していないとは言い切れない。例えば、1アミノ酸残基置換の影響を相殺する置換が、同じ領域にあったとするとこれまでの実験では検出できない。

Gly¹⁸¹→Asp¹⁸¹ および His²⁶⁶→Asn²⁶⁶ の2アミノ酸残基置換が活性に影響を及ぼすのも事実であり、この2アミノ酸残基置換の酵素分子レベルでの役割も興味深い。Hatanaka らは、EII 酵素の 181番目の Asp を Asn と Glu に換えると、Km 値がそれぞれ4倍と 1.1 倍に増加し、His

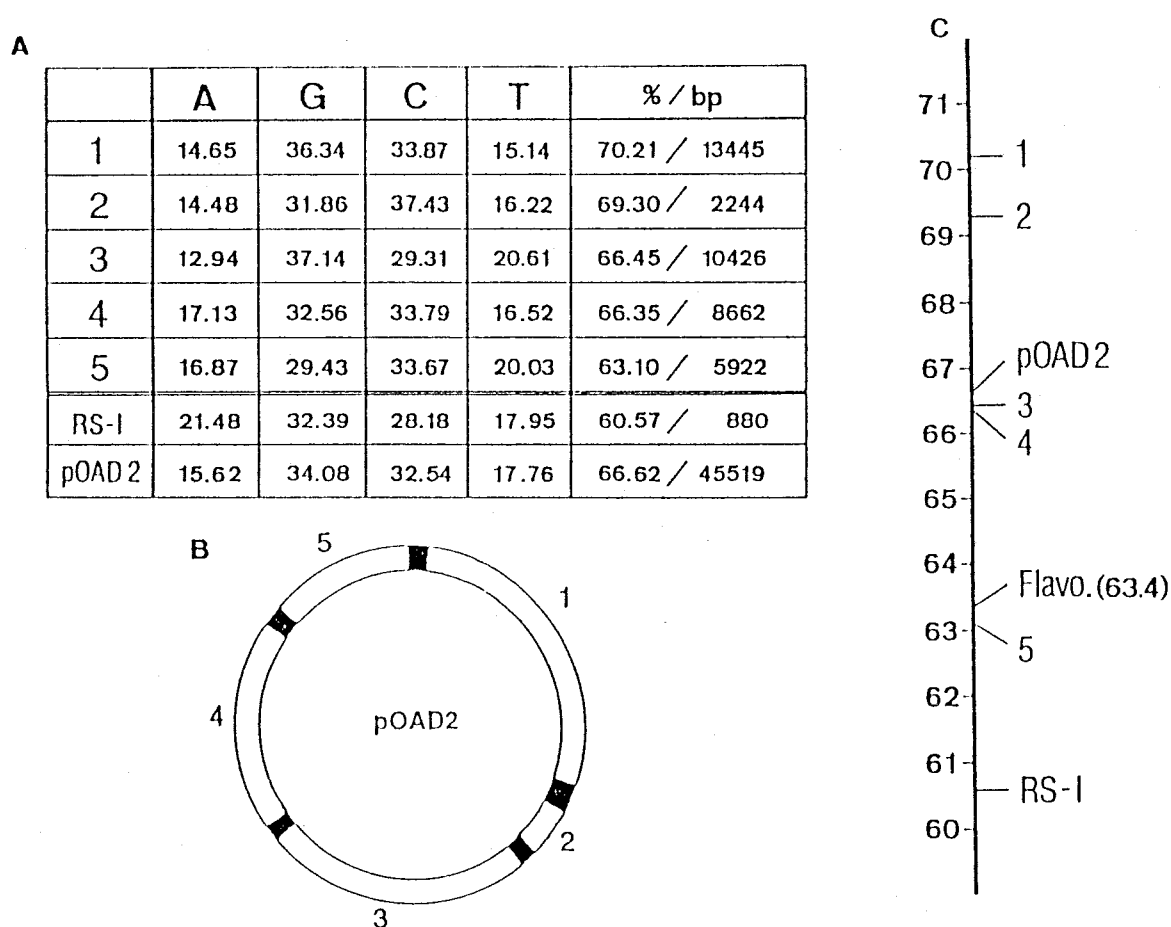


Fig. 2-23 Comparison of (G+C) content among RS-I (IS6100) regions.

や Lys に換えると活性が完全に失われると報告している⁶¹⁾。また、EII 酵素は、diisopropyl fluorophosphate (DFP) で阻害され、DFP が 112番目の Ser に結合することから、EII 酵素はこの 112番目の Ser を活性中心としたセリンプロテアーゼと類似した反応機構を持つと考えられる⁶²⁾。一般に、トリプシンやサブチリシンなどのセリンプロテアーゼは、Asp, Ser および His の3アミノ酸残基よりなる Charge relay system が活性に関係している⁶³⁾。EII 酵素は 181番目の Asp を Glu に換えても高い Kcat 値が残っている (EII 酵素の約 30%) ので、181番目の Asp はトリプシンにおける 102番目の Asp⁶⁴⁾のような活性残基ではないと考えられる。これらのことから、EII' から EII への活性上昇に関与する2アミノ酸残基置換は、主に酵素と基質の結合に影響すると考えられる。

RS-IIA や RS-IIB に加えて、RS-II 領域としては、nylB' DNA をプローブとして 55℃ で弱くハイブリダイズする領域が約 10 kb の広い範囲に観察されたが (Fig. 2-1)、この範囲に nylB' と相同性を示す領域を特定することはできなかった。しかし、第一章に示す pOAD2 の全塩基配列内には、100 bp 程度の繰り返し配列がいくつか見つかり、RS-I や RS-II 以外の繰り返し配列が存在すると思われる。

pOAD2 上に存在するナイロンオリゴマーの分解に関与する遺伝子群; EI 遺伝子 (F-nylA)、EII 遺伝子 (F-nylB) および EII' 遺伝子 (F-nylB')、と *Pseudomonas* sp. NK87 株が持つ EI 遺伝子 (P-nylA)⁵¹⁾と EII 遺伝子 (P-nylB)⁶⁵⁾、これらと IS6100 との関係を考えてみる。Nei と Gojobori の方法⁶⁶⁾により、同義的置換に基づいた進化距離を計算すると、F-nylB と P-nylB が 1.11, F-nylB と F-nylB' が 0.303, F-nylA と P-nylA が 3.6×10^{-3} であった。この値によると、F-nylB と P-nylB が最初に分岐し、それに引き続き、F-nylB と F-nylB' が、そして最後に F-nylA と P-nylA が分岐したと考えられる。1つの塩基置換もない IS6100 の分布は、これらナイロンオリゴマー分解性遺伝子の分岐ののちに起こったと考えられる。進化距離から考えると IS6100 がそれ自身に変異を蓄積しない短い期間に、分岐前のナイロンオリゴマー分解性遺伝子の複製や転座に重要な役割を果たしたとは考えられない。しかし、IS6100 が挿入されることにより、これら遺伝子の発現量が変わる可能性はある。

サザンハイブリダイゼーション法による解析から、IS6100 そのもの、もしくは、その類似配列が pOAD2 上 (KI72 株) と pNAD2 上 (NK87 株) にだけ存在し、染色体上やその他のプラスミド上に存在しないことがわかった。これらの結果から、IS6100 は KI72 株や NK87 株を宿主とし、その内部で増幅したのではなく、他の宿主において、pOAD2 の先祖型プラスミド上で増幅し、その後、そのプラスミドが直接または間接的に KI72 株や NK87 株に移動したと考えられる。先祖型プラスミド上で IS6100 が増幅する時、すでに分岐していたナイロンオリゴマー分解性遺伝子のプラスミド上への再構成には、この IS6100 が関与した可能性はある。KI72 株で、pOAD2 と他のプラスミドの挙動を比較してみると、pOAD2 は他のプラスミドに比べて脱落しやすいことや、菌体当りのコピー数が異なることなど違いがみられる。

第五節 要約

pOAD2 上の5ヶ所に点在する RS-I 領域の塩基配列を比較したところ、RS-IA, C, D, および E の全領域と RS-IB の前半 880 bp 領域が完全に同一の塩基配列を有していた。DNA のデータベースと相同性検索を行った結果、RS-I の持つ 880 bp は *Mycobacterium fortuitum* 由来の挿入配列 IS6100 と全く同一の配列であった。IS6100 がコードするトランスポゼースを他の IS のトランスポゼースと比較すると IS6 ファミリーに属する IS 群のそれと相同性があった。IS6100 は、*Mycobacterium fortuitum* 内でトランスポゾン Tn610 の一部として存在し、この Tn610 は転移能を持ち、IS6100 が発現することがわかっている。そこで、IS6100 の KI72 株内での発現を転写レベルで調べると、微弱ながらその転写物が検出できた。

IS6100 の pOAD2 以外での分布を調べると、KI72 株と同じく、ナイロンオリゴマー資化性菌である *Pseudomonas* sp. NK87 株の保持する pNAD2 プラスミド上には少なくとも2ヶ所 IS6100 そのもの、もしくは IS6100 に高い相同性を持つ相同配列が存在することがわかった。

pOAD2 上に存在する RS-II 領域 (RS-IIA と RS-IIB) がコードするナイロンオリゴマー分解酵素 (EII) とその類似タンパク質 (EII') は、392アミノ酸残基よりなるが、この2酵素間で 47アミノ酸残基が異なるため、基質である Ald の分解活性は EII' が EII の約 1/100 である。どのアミノ酸残基置換が活性の違いに関与しているかを両者の ORF 内で保存されている制限酵素切断部位で組換え、ハイブリッドタンパク質を作成して調べた結果、EII-EII' 間に存在する47アミノ酸残基置換のうち、Gly¹⁸¹→Asp¹⁸¹ と His²⁶⁶→Asn²⁶⁶ のアミノ酸残基置換のみで EII' レベルから EII レベルへの活性上昇が得られることがわかった。

pOAD2 上で IS6100 が5ヶ所で全く同一配列を有していることと、*Mycobacterium fortuitum* 由来の IS6100 とも同一の配列であることから、IS6100 の分布、転移、重複は比較的近年に起こったと考えられる。しかし、ナイロンオリゴマー分解性遺伝子群の進化距離からすると、IS6100 がこれら遺伝子の分岐前の複製や転座に重要な役割を果たしたとは考えられない。

第三章 pOAD2 上に存在する non-stop frame (NSF) の解析

第一節 緒言

第二章で、pOAD2 の構成要素である RS-I と RS-II について述べてきたが、本章では、pOAD2 上に存在する non-stop frame (NSF) について解析する。pOAD2 上には、全塩基配列の決定から 600 bp 以上の NSF が 71個存在している。新しいタンパク質あるいは酵素が出現するためには、その基となる読み取り枠が必要である。そこで、終止コドンを持たない枠 (NSF) をその候補と考え、これら NSF から予想されるアミノ酸配列をデータベースに対し相同性検索を行った。また、IS6100 (RS-I) とこれら NSF の関係を考えるため、IS6100 が持つプロモーター活性を調べた。

第二節 実験材料および方法

1 使用菌株およびプラスミド

宿主として *E. coli* JM109 株¹¹⁾をプラスミドへのサブクローニングや塩基配列の決定に用いた。プラスミドベクターとして pBluescriptII SK⁺, KS⁺^{14, 15)}および pKK232-8⁶⁷⁾を用いた。

2 相同性検索

DNA Data Bank of Japan (国立遺伝学研究所遺伝情報センター) に登録されている Gen Bank (Ver. 71), EMBL (Ver. 30) にあるデータベースを用い、DNA 配列を可能な6フレームすべてについてアミノ酸配列に翻訳し、プログラム "flat" を使い、pOAD2 上に存在する NSF をアミノ酸配列に翻訳したものの相同性検索を行った。

3 オリゴヌクレオチドの合成および精製

Polymerase chain reaction (PCR) に用いるオリゴヌクレオチドは、DNA 合成装置 (Applied Biosystems 381A DNA Synthesizer) により、ホスホトリエステル法⁶⁸⁾で合成した。オリゴヌクレオチドの精製は、OPCTM カートリッジ (Applied Biosystems) を用い、一連の操作はその使用書に従った。

4 Polymerase chain reaction (PCR) 法

PCR^{69, 70)}は PCR 装置 (ASTEC; program temp control system PC-700) を使い、変性 94℃, アニリング 55℃, 反応 72℃ で 25 cycle 行った。反応終了後、クロロホルム抽出を1回行い、その上澄液中の DNA をエタノールにより沈澱させ、遠心後、沈澱物を減圧乾固することにより PCR 標品を得た。

5 CAT (chloramphenicol acetyltransferase) 活性の検定

クロラムフェニコールの濃度が異なる3種類の LB 寒天平板培地 (0 $\mu\text{g/ml}$, 25 $\mu\text{g/ml}$, および 50 $\mu\text{g/ml}$) に最終濃度 100 $\mu\text{g/ml}$ のアンピシリンを加えたものを準備し、これら培地上でプロモーター検索用ベクター pKK232-8⁶⁷⁾ を母体とする各組換えプラスミドを保持する大腸菌の生育能を調べ、CAT 活性を検定した。

6 その他の実験材料および方法

制限酵素および修飾酵素などは、宝酒造製あるいは、東洋紡績製のものを主に使用し、実験操作については各々の使用書などに従った。プラスミド DNA の調製、アガロースゲル電気泳動ならびにゲルからの回収、形質転換、塩基配列の決定などは、第一章、第二節に示した通りに行った。

第三節 結果

1 NSF の相同性検索

全塩基配列から、pOAD2 上には 600 bp 以上にわたり終止コドンの存在しない NSF¹⁰⁾ が 71個認められた。NSF についてその長さと pOAD2 上での位置を Fig. 3-1 に示す。太線太数字で示した NSF は、時計回りの方向 (内側からフレーム 1, 2 および 3)、細線白抜き数字で示した NSF は、反時計回りの方向 (内側からフレーム 4, 6 および 5) である。フレーム 1 と 4, 2 と 6, 3 と 5 は、それぞれ読み枠の等しい相補鎖の関係である。

これら NSF と相同性のあるものをデータベース上で調べた。まず、核酸のデータベースである Gen Bank と EMBL に登録されている DNA 配列を可能な6フレームすべてについてアミノ酸配列に翻訳し、新しいデータベースを作成した。この新しいデータベースに対し、71個の NSF がコードするアミノ酸配列の相同性検索を行った。すなわち、coding も non-coding も、構造遺伝子と違ったフレーム由来のアミノ酸配列も含め、すべてについて相同性検索を行った。相同性検索に ORF ではなく NSF を使ったのは、ORF が NSF の一部であり、また、データベース中の DNA と進化的関係を考える上で、より多くの情報が得られるからである。

相同性検索の結果、まず、ORF として認められた領域と相同性があったもののうち (NSF 由来のアミノ酸配列が機能を持つものと相同性があったもの) 代表的なものを挙げる (Fig. 3-2)。ここで 71個の NSF は、時計回りのものを NS、反時計回りのものを NA とし、フレーム番号と通し番号で名前が付してある。NA1-20 は、*Agrobacterium tumefaciens* 由来の pTiB6S3 プラスミドの repA⁷¹⁾、*E. coli* 由来の mini-F プラスミドの sopA⁷²⁾、P1 プラスミドの parA⁷³⁾、RK2 プラスミドの incC⁷⁴⁾ などと相同性があり、プラスミドの複製、維持あるいは分配に関与すると思われる。NA2-15 は、*Aspergillus nidulans*^{75,76)} や *Penicillium chrysogenum*⁷⁷⁾ 由来の isopenicillin-N acyltransferase penDE と相同性がある。NS3-2 は、*E. coli* 由来の ftsX⁷⁸⁾ と相同性があり、細胞分裂に関係すると思われる。NA3-9, NA1-10, NA2-11, NA3-12, および NA3-13

は、*Bacillus subtilis*^{79,80)}, *Salmonella typhimurium*⁸¹⁾, *Streptococcus pneumoniae*⁸²⁾ 由来の oligopeptide permease (OPP) オペロンと合わせて相同性がある。NA3-9 と *oppF*, NA1-10 と *oppD*, NA2-11 と *oppC*, NA3-12 と *oppB*, および NA3-13 と *oppA* である (Fig. 3-3)。この他に、NA2-22 は Bacteriophage P2 由来の DNA packaging と capsid synthesis に関与するものと、NS2-26 は *Xenopus laevis* 由来の oocyte zinc finger protein⁸³⁾ と、NA1-30 は Rat 由来の intermediate filament protein⁸⁴⁾ と、NS3-27 は *Xenopus laevis* 由来である fibrinogen γ -chain⁸⁵⁾ と、NS3-34 は Human 由来の MHC class III complement C6 protein⁸⁶⁾ と、NA3-36 は、Bovine ミトコンドリア由来の phosphotransferase⁸⁷⁾ と、NA3-37 は *Drosophila melanogaster* 由来の RNAPolymerase III⁸⁸⁾ と相同性がある (Fig. 3-4)。

次に、ORF と認められた領域以外と相同性のあったものを含め、結果を Fig. 3-5, Fig. 3-6, Fig. 3-7 に示す。ここに示すものはすべて統計的に有意と考えられる。NSF の番号、pOAD2 上での位置、相同性があった DNA 配列の機能を示している。* 印は構造遺伝子と違うフレーム由来のアミノ酸配列と相同性が認められたもの、** 印は構造遺伝子領域の外側由来のアミノ酸配列と相同性が認められたものを示す。例えば、Fig. 3-6 の NA3-9 由来のアミノ酸配列は permease と

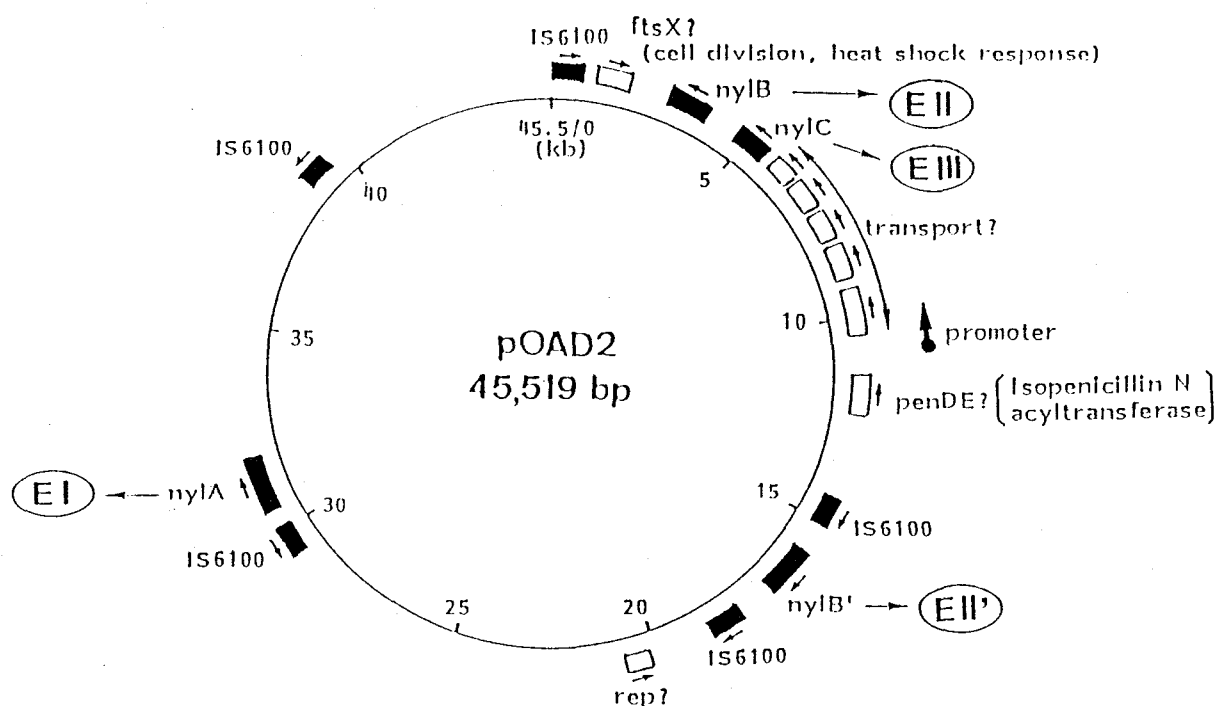


Fig. 3-2 Functional map of pOAD2 plasmid. The four nylon oligomer degradation genes (*nylA*, *nylB*, *nylB'*, and *nylC*) and IS6100 sequences appearing five times on the plasmid are shown by heavy boxes. Open boxes indicate the genes which have significant homology with the sequences of *oppA*, *oppB*, *oppC*, *oppD*, *oppF*, *ftsX*, *penDE*, and *rep* genes. Arrows indicate the direction of the genes.

相同性があるが、同じ NA3-9 は窒素固定に関与している遺伝子上流領域由来のアミノ酸配列とも相同性が認められる。Fig. 3-5, Fig. 3-6, Fig. 3-7 についてまとめると (Table 3-1)、機能領域と相同性があったもの、すなわち、遺伝子の coding 領域由来のアミノ酸配列と相同性があったものが 24個、coding 領域由来のものとは相同性がなかったものが 47個存在した。このうちの 14個は coding 領域由来のもののみと相同性があったもの、10個は coding 領域由来のものとは相同性があったが、それ以外の例えば、構造遺伝子の外側とか、違ったフレーム由来のアミノ酸配列とも相同性があったものである。また、29個は coding 以外と相同性があったもので、18個はデータベース上にあるすべての DNA 配列由来のアミノ酸配列とも相同性を示さなかったものである。相同性検索に使ったデータベースは、Gen Bank で (65100 遺伝子座, 83894652 塩基), EMBL で (63378 遺伝子座, 83574342 塩基) であり、これは、全遺伝子データの 1/3 程度の種類を含んでいると考えられる⁸⁹⁾。

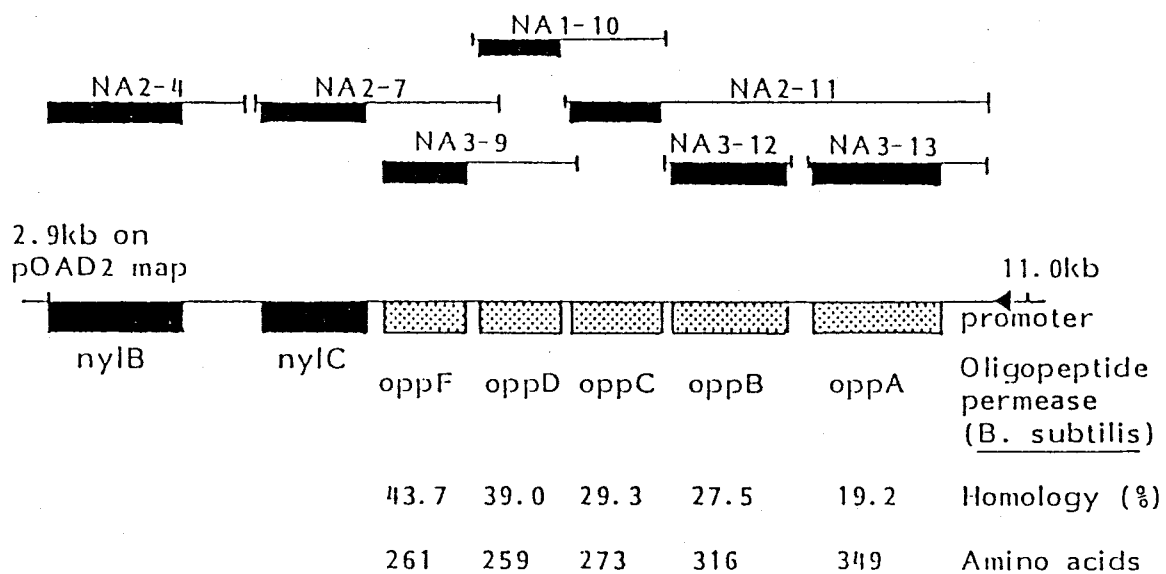
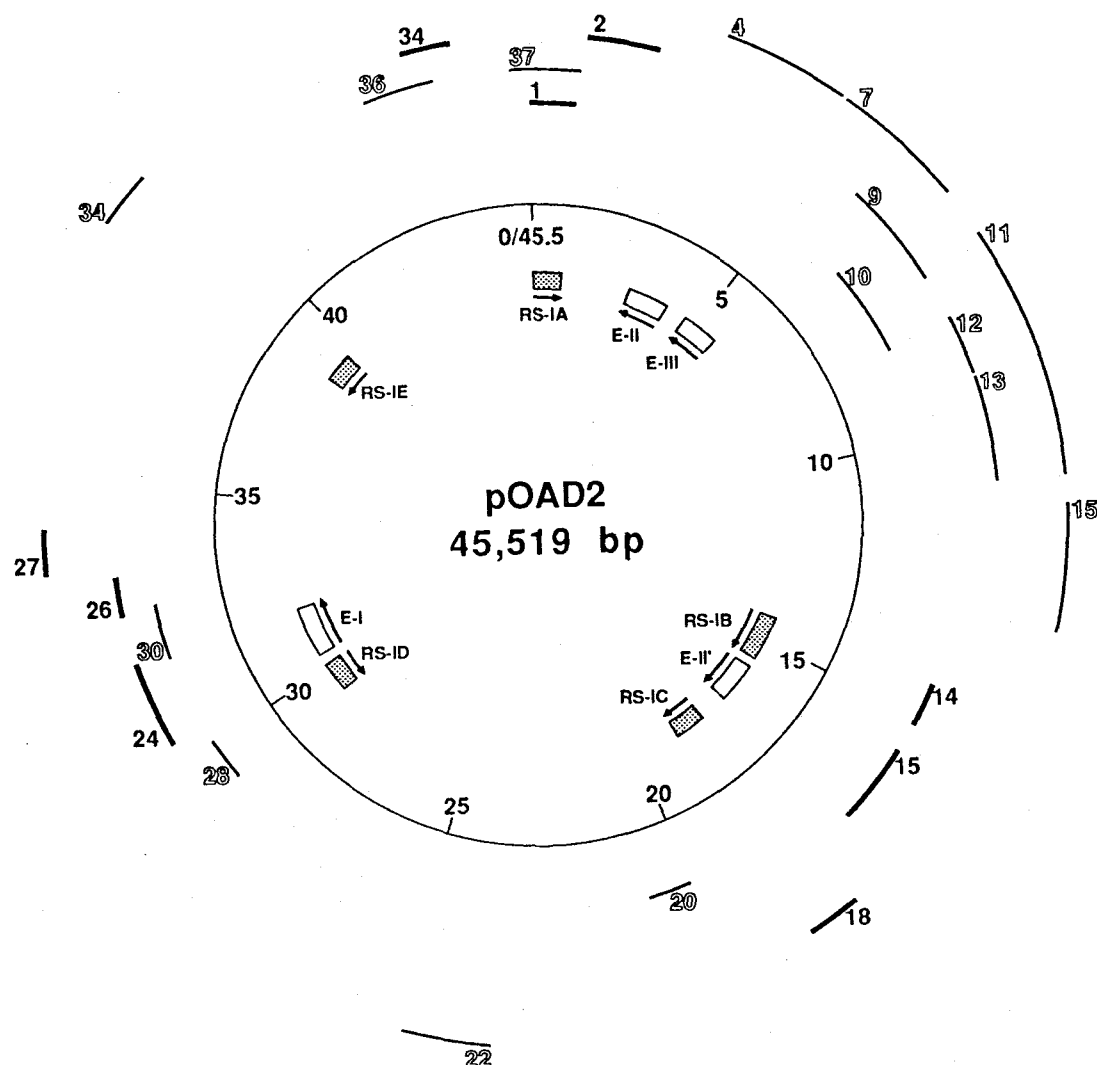


Fig. 3-3 Location of the identified and deduced genes at position 2.9-11 kb on the map of pOAD2. Among the seven NSFs, NA2-4 and NA2-7 include the *nylB* and *nylC* genes (shown by heavy boxes), respectively. The other five NSFs have significant homology with five genes responsible for the transport of oligopeptides in various microorganisms. The region which possess the homology with the *opp* genes of *Bacillus subtilis* is shown on the lower part with the extent of homology and length of the homologous amino acids. These homologous regions shown by heavy boxes are closely located each other between the *nylC* gene and a promoter region essential for expression of *nylB* in *E. coli*.



NS3-2	cell division protein	NA2-22	DNA packaging . capsid synthesis
NA3-9	permease	NA1-30	intermediate filament protein
NA1-10	permease	NS2-26	DNA-binding protein
NA2-11	permease	NS3-27	γ -fibrinogen
NA3-12	permease	NA3-36	phosphotransferase
NA3-13	permease	NS3-34	human complement component
NA2-15	acyltransferase	NA3-37	RNA polymeraseIII
NA1-20	plasmid maintenance,replication		

Fig. 3-4 Homology search for NSFs in pOAD2. These NSFs have homology sequence with functional genes.

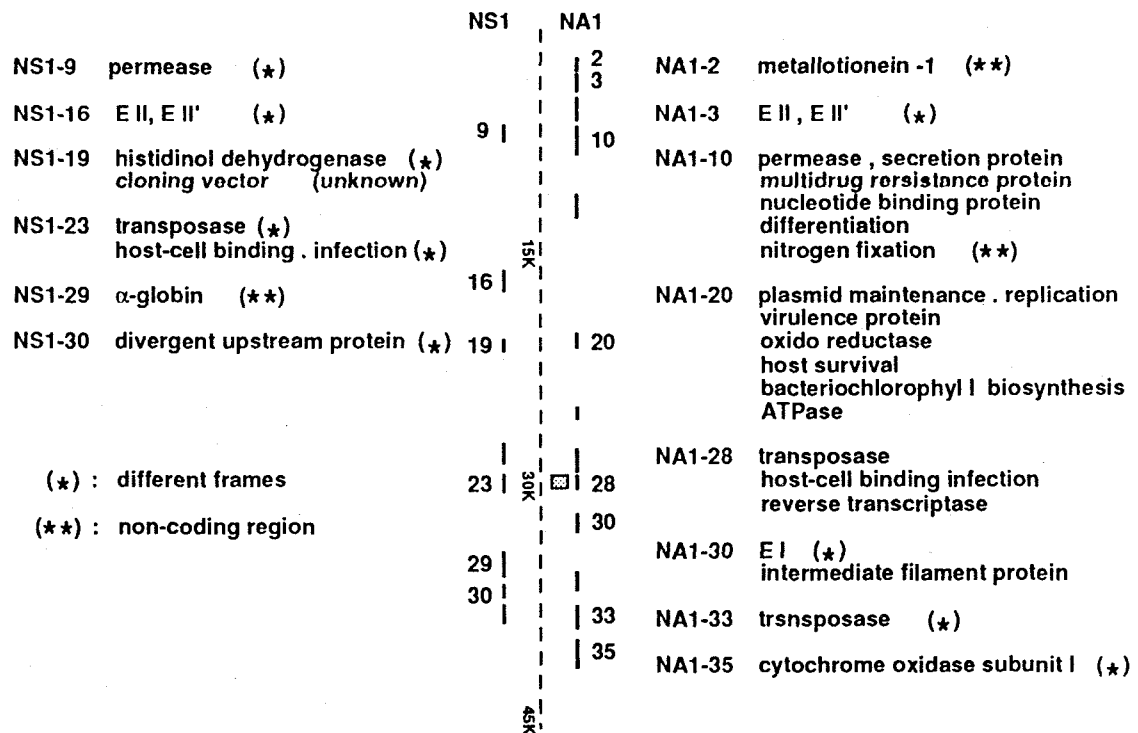


Fig. 3-5 Homology search for frames 1 and 4 in pOAD2.

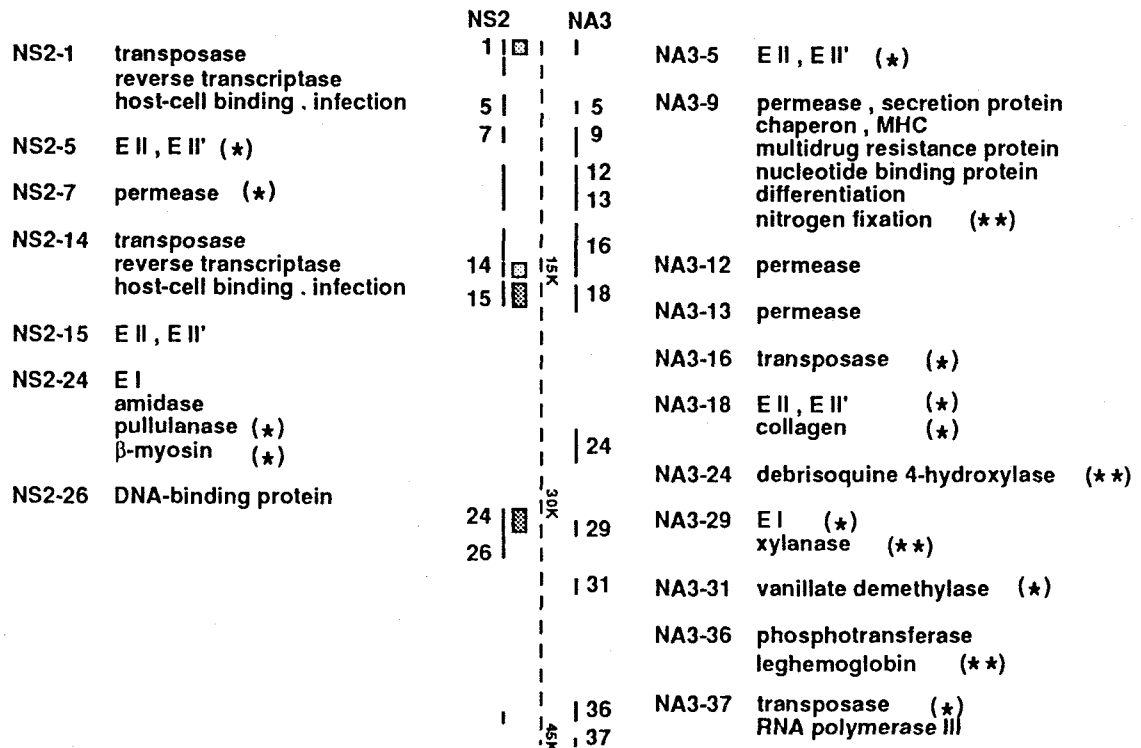


Fig. 3-6 Homology search for frames 2 and 6 in pOAD2.

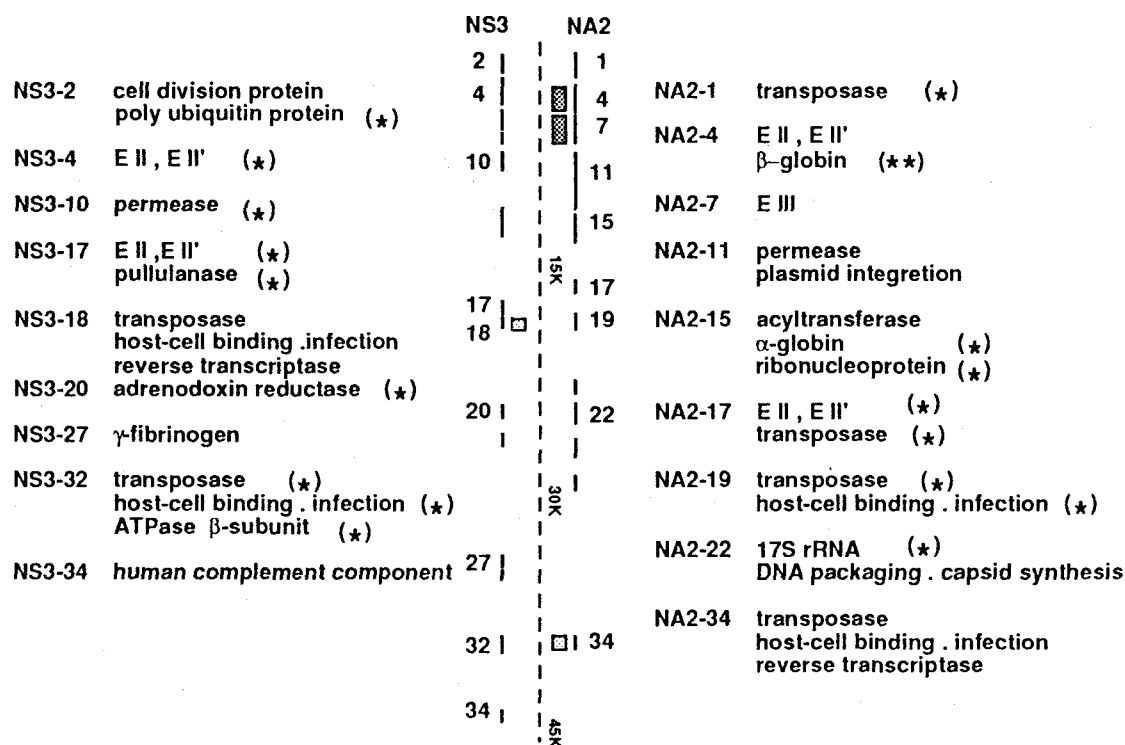


Fig. 3-7 homology search for frames 3 and 5 in pOAD2.

2 pOAD2 上での NSF 数

pOAD2 上には、200コドン以上の NSF が 71個存在したが、pOAD2 の全塩基配列が明らかになっているので、数学的に NSF 数を予想することもできる。予想される 200コドン以上の NSF 数は、塩基組成 (G+C 含量) から終止コドンの出現する確率を求め計算した場合 19.6個、pOAD2 上に存在する終止コドンの総数を数え、その数から計算すると 42.8個、実際に pOAD2 に存在する 200コドン以上の NSF 数は 71個であり、予想される NSF 数より多かった。また、NSF の長さによって、予想される数と実際とを比較すると、長い NSF ほど予想されるものとの差が大きくなった。600コドン以上の NSF では、終止コドンの数から予想される場合の 200倍であり、塩基配列から予想される場合の 3500倍もの NSF が pOAD2 上に存在していた (Fig. 3-8)。

3 IS6100 が持つプロモーター活性

NSF と IS6100 との関係を考えるため、IS6100 のプロモーターを解析した。IS などの移動する DNA 単位は、欠失、挿入、転座、重複など染色体 DNA の再構成に深く関係しているが⁶⁰⁾、挿入位置によっては、エンハンサー効果により遺伝子を活性化することも知られている⁹⁰⁾。また、IS6100 が属する IS6 ファミリーは、その両末端に持つ逆位繰返し配列 (IR) 内に大腸菌のプロモーター配列である -35領域 (コンセンサス配列; TTGACA)³⁹⁾に相当する配列 "TTGCAA" を持

	Target region	Number of NSF	
Homology to coding	Coding	14	24
	Coding & non-coding	10	
No homology to coding	Non-coding	29	47
	————	18	

Table 3-1 Characterization of NSFs in pOAD2 about homology.

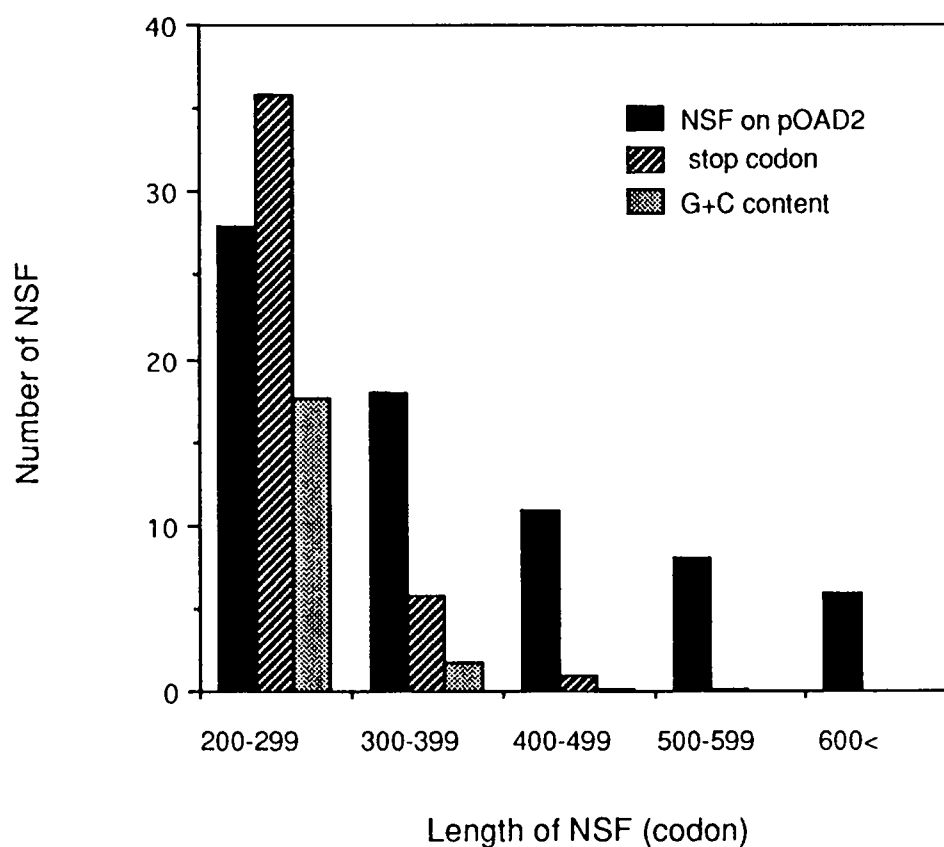


Fig. 3-8 Length distribution of NSFs in pOAD2. Closed boxes, the numbers of NSF found on pOAD2; virgule boxes, the expected numbers of NSF on pOAD2 calculated from the number of termination codons in pOAD2; shadow boxes, expected numbers of NSF in pOAD2 calculated from base composition of the plasmid.

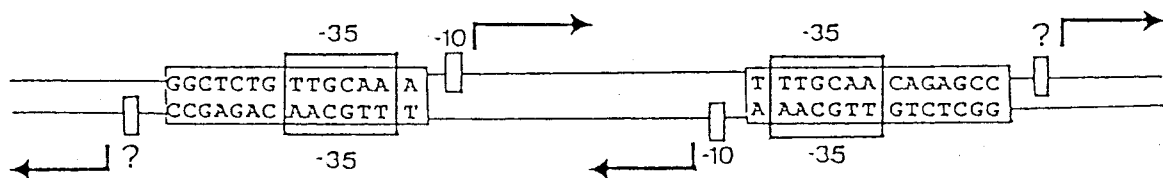


Fig. 3-9 Structure of terminal inverted repeats of IS6100. The 14 bp terminal inverted repeats are shown. -35 region and -10 region which have homology to consensus sequence with *E. coli* promoter, are indicated. Arrows indicate the direction of transcription.

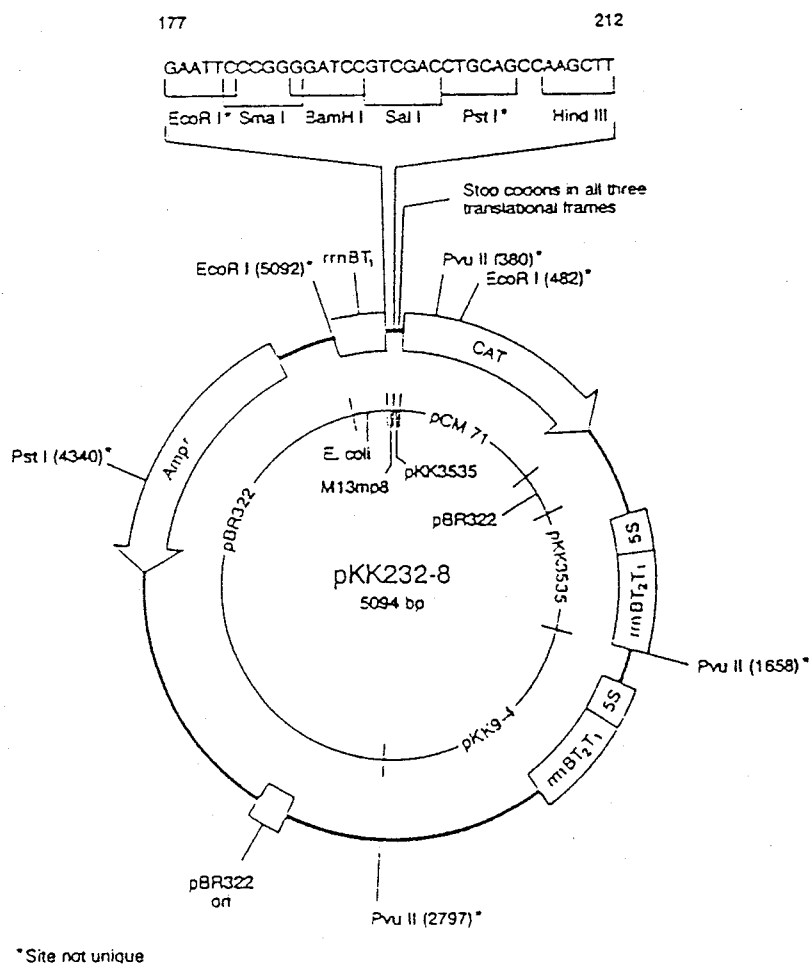


Fig. 3-10 Structural and functional map of pKK232-8. pKK232-8 is the *cat* vector for promoter analysis.

つ。“TTGCAA”はこれ自身が回転対称の配列であり、IS6100を1単位としてみると、“TTGCAA”を計4ヶ所に持つことになる。このうち1ヶ所は、IS6100自身のコードするトランスポゼースの-35領域である (Fig. 3-9)。そこで、この“TTGCAA”が-35領域として機能するかをプロモーター検索用ベクター pKK232-8 (Fig. 3-10) を使い、下流にある cat の発現で調べた。pKK232-8の BamHI 部位に IS6100 の全長が挿入できるよう IS6100 の両末端部に PCR 法を用いて BamHI 部位を導入した。IS6100 の両末端には、14 bp よりなる逆位繰返し配列があるが、この領域と BamHI 部位を合わせ、25 mer のオリゴヌクレオチド1本を両端への部位導入用プライマーとして合成した。PCR 反応後、産物を BamHI で消化し、1.5% アガロースゲル電気泳動により回収し、pBluescript IISK⁺ の BamHI 部位に導入した。BamHI 部位を導入した IS6100 については、全長の塩基配列を決定し、変異がないことを確認した。また、大腸菌のプロモーター配列 (コンセンサス配列; TTGACA, 17 bp スペース, TATAAT)³⁹⁾ によく合うよう、IS6100 の “TTGCAA” に対し、スペースの長さを考慮にいれ、pKK232-8 ベクターの SalI 部位と HindIII 部位間のギャップに組み込めるよう -10領域導入用リンカー (大腸菌プロモーターのコンセンサス配列; TATAAT を含む) を合成した。合成したリンカー内には、確認用に XhoI 部位が導入してある。合成リンカーを pKK232-8 の SalI 部位と HindIII 部位間のギャップに組み込み、その部分の塩基配列を決定することにより確認を行った。Table 3-2 に PCR 用プライマーと -10領域導入用リンカーを示す。

pKK232-8 と、これに -10領域を導入したプラスミド (Fig. 3-11; 1, 2) の BamHI 部位に、それぞれ PCR で BamHI 部位を導入した IS6100 を組み込み、4種類の組換えプラスミドを取得した (Fig. 3-11; 3, 4, 5, 6)。加えて、IS6100 の内部にある StuI 部位と pKK232-8 の SmaI 部位を用いて、前半 430 bp、もしくは、後半 450 bp に相当する領域を欠失させた組換えプラスミドを構築した (Fig. 3-11; 7, 8, 9, 10)。これら組換えプラスミドを保持する大腸菌のクロラムフェニコールを含む培地での生育能を観察すると、IS6100 を cat の上流に順方向に組み込んだものでは cat の発現が認められるが、逆方向に組み込んだものでは発現が認められなかった (Fig. 3-11; 3, 4)。IS6100 を順方向で組み込んだもの前半部分を欠失させると cat の発現がみられなくなるが、IS6100 の下流に -10領域を付加してやると cat が発現する (Fig. 3-11; 3, 7, 9)。また、IS6100 を逆方向に組み込んだものに -10領域を付加すると cat の発現がみられる。この場合、後半部分に相当する領域を欠失させてもなお cat の発現が認められる (Fig. 3-11; 4, 6, 10)。このことから IS6100 は、それ自身のプロモーター (トランスポゼース) により下流のプロモーターを持たない遺伝子を発現させることができるだけでなく、IR 内に持つ “TTGCAA” 配列に -10領域が付加されれば、プロモーター活性を示すことがわかった。

Table 3-2 Synthetic oligonucleotides primer and linker

	sequence
PCR primer	5'-CGGGATCCAGAGGGCTCTGTTGCAA-3'
-10 linker 1	5'-GATCCTATAATCTCGAGA-3'
-10 linker 2	5'-AGCTTCTCGAGATTATAG-3'

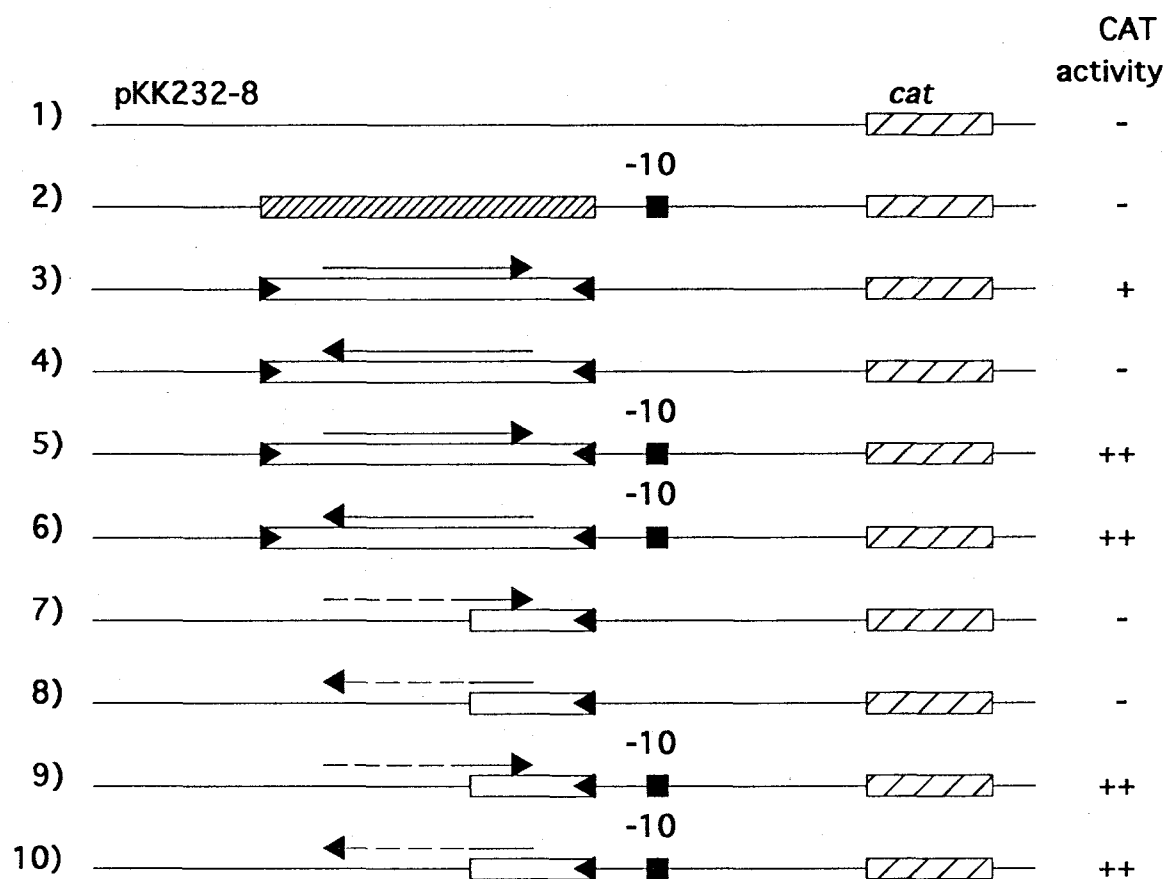


Fig. 3-11 Relationship between the plasmid structure and the enzyme activity. All plasmids contain a common vector DNA (5.1 kb pKK232-8, this vector has *cat* gene.) shown by the solid line and virgule boxes. IS6100 and its related regions are shown by open boxes. The 14 bp inverted repeats at its termini are shown by triangles. The synthetic -10 region are shown by heavy boxes. Arrows indicate the direction of IS6100.

第四節 考察

pOAD2 上に存在する 71個の NSF について相同性検索を行った結果、NSF 由来のアミノ酸配列が、機能を持つものと相同性が認められた NSF の内で、NA3-9, NA1-10, NA2-11, NA3-12, および NA3-13 は、*Bacillus subtilis*^{79,80)} 由来の oligopeptide permease (OPP) オペロンと相同性があったが、この領域は、EII 遺伝子 (*nylB*) と EIII 遺伝子 (*nylC*) の近接した上流に位置している。Okazaki らは、*nylB* の転写開始点を大腸菌内で決定した際、*nylB* の上流約 7 kb の位置から転写されていると報告している^{91,92)} (Fig. 3-3)。この転写開始点は、NA3-13 の 5 bp 上流にあることから、NA3-13, NA3-12, NA2-11, NA1-10, NA3-9, *nylC*, そして *nylB* は、1つの転写単位として転写されると考える。また、相同性があったのが oligopeptide permease であることから、これら NSF は、ナイロンオリゴマーの取り込みに関与すると思われる。事実、膜貫通領域や ATP 結合領域がよく保存されている。

71個の NSF すべてについて相同性検索の結果で分類すると、14個は coding 領域由来のもののみと、10個は coding 領域と non-coding 領域由来のもの、29個は coding 以外の領域由来のもの、18個はデータベース上にあるすべての DNA 配列由来のアミノ酸配列とも相同性を示さなかった。このうち、18個の NSF はもちろん、coding 領域と相同性のなかった 29個の NSF を含めて、47個の NSF は今まで機能は判っているが配列の特定されていないものに属するかもしれないが、これらは全く未知の、今まで存在しないタンパク質を作り出す可能性を持つフレームである。また、EII 酵素ファミリー (F-EII, F-EII', および P-EII) や EIII 酵素をコードする遺伝子の対応する相補鎖には、他のフレームには多くの終止コドンがあるにもかかわらず、NSF が存在する。これら遺伝子の相補鎖に NSF が存在する確率は非常に小さく、わずか 0.0001-0.0080 である。もともと *nylB* 遺伝子ファミリーの先祖型遺伝子が、その相補鎖に NSF を持っていたとしても、それが、各 *nylB* 遺伝子に受け継がれる確率は、わずか 0.007 でしかない。これら NSF の存在は、普通では考えられないことであり、何らかの力によって枠が開放されていると考えられる。興味深いことに、pOAD2 上の他の NSF もそれぞれの相補鎖に NSF を持つものが多く、*nylB*, *nylB'*, および *nylC* を含む 23組 50個もの NSF が、その相補鎖の同じ枠にも NSF を持つ (Fig. 3-1)。相同性検索の結果 coding 領域由来のもの、10個の NSF のうち 20個の NSF は、その相補鎖の同じ枠にも NSF が存在している。pOAD2 上に存在する NSF 数が、全塩基配列から予想される数よりも多いことは、相補鎖に NSF が存在していることに起因するかもしれない。もし、相補鎖の同じ枠に NSF がいないとすると、NSF 数は 44個となり、pOAD2 上に存在する終止コドン数から予想される NSF 数 42.8個とよく一致する。ある遺伝子の相補鎖に NSF が存在する例としては、ハロゲン化酢酸分解酵素遺伝子、トルエン分解酵素群の一部の遺伝子などが挙げられる。それぞれの NSF が相補鎖の同じ枠に持つ NSF は、どれもアミノ酸に翻訳した場合、データベース上のいかなる coding 領域由来のアミノ酸配列とも相同性を示さないものである。宿主にとって必要でない NSF には特別な選択圧がかからないかぎりラ

ンダムな変異が生じ、終止コドンが生まれ、その結果、NSF が消失するはずである。それぞれの遺伝子の相補鎖に NSF が存在するのは、その NSF が機能を持つタンパク質をコードしているからだとも考えられる。しかし、pOAD2 上で互いに NSF を持つ 50個の NSF すべてが機能を持つタンパク質をコードしているとは考えにくい。また、もう少し短い領域で見た場合、300 bp 以上の ORF 157個のうち、約 2/3 の ORF は、その相補鎖の同じ枠に NSF が存在していることから、やはり、何らかの力により、枠が解放されていると考えられる。枠が開放されているメカニズムはわからないが、そこに、保存されている NSF は、新規遺伝子のよい候補となるのではないかと考える。

pOAD2 上に存在する IS6100 のプロモーター活性を調べると、IS6100 自身がコードするトランスポゼースのプロモーターにより、下流に位置するプロモーターを持たない遺伝子を発現できることがわかった。また、IS6100 が IR 中に持つ "TTGCAA" に -10領域を付加してやると、ハイブリッドプロモーターを形成して IS6100 の上流に向かっても下流に向かっても、そこに -10領域があれば、プロモーター活性を示すことがわかった。すなわち、IS6100 は挿入位置によっては、それ自身のプロモーター（トランスポゼース）により、また、そこに偶然存在する -10領域様配列とハイブリッドプロモーターを形成し、下流のプロモーターを持たない遺伝子を発現できる。加えて、この IS それ自身が持つ "転移" 能力により、プロモーターを新しく付与する機会が増えると考えられる。IS6100 は *Flavobacterium* sp. K172 株や *Mycobacterium fortuitum* 内で計7ヶ所に存在するが、他の IS と違い、この7ヶ所を比較した場合、特異的な挿入部位が認められない。このことは、IS6100 は限定された部位に挿入するのではなく、不特定な広い範囲に挿入し得るわけである。遺伝子がタンパク質を産生し発現するためには、転写され翻訳されなければならない。pOAD2 上で既存の遺伝子由来のアミノ酸と相同性を有しない 47個の NSF が、もしタンパク質を産生するならば、それは既存のデータベースにない種類のタンパク質となる。しかし、これら NSF が新規遺伝子になるためには、プロモーターが付与され、転写されることが必要である。IS6100 は、このプロモーターの付与に関わっているのかもしれない。pOAD2 上で IS6100 の近傍に位置する NSF (NS2-3, NS3-2, NS2-15, NS2-16, NS2-24, NS1-33, NA3-16, NA3-18, NA1-26, および NA2-27) も IS6100 領域から転写されている可能性がある。事実、*nylB*' (NS2-15) のプロモーターが、IS6100 (RS-IB) の IR 中にある "TTGCAA" とその下流に位置する -10領域様配列よりなるハイブリッドプロモーターであることを示唆するデータがある。IS が持つ IR 中の -35領域と既存のプロモーターの -10領域とによるハイブリッドプロモーターの可能性を示す例は、*aphA7* (aminoglycoside-3',5"-phosphotransferase-I) gene と IS26⁵⁴⁾、*thyE* (thymidylate synthase) gene と IS257⁹³⁾ などで報告されている。

実際の菌体内では、必要でないタンパク質の発現があるかもしれない（我々に見えないだけであり、その種類は多いかもしれない）、pOAD2 上でも、多くの NSF がタンパク質を微弱に産生しているかもしれない。IS6100 のプロモーター活性（トランスポゼースのプロモーター、もしくは、ハイブリッドプロモーター）は、ここで調べた限り決して弱くなく、もし、IS によりプロモーター

を持たない、もしくは、非常に弱いプロモーターを持つ読み取り枠が発現したならば、それは、タンパク質として十分なものとなる。IS の関与などにより、NSF 由来のタンパク質発現量が激増し、一度発現するようになったタンパク質が菌体にとって有用であれば、その後変異を重ね、酵素として進化していくと考えられる。

第五節 要約

p0AD2 の全塩基配列から p0AD2 上には、600 bp 以上にわたり終止コドンの存在しない NSF が 71個認められた。71個の NSF すべてについて、これをアミノ酸配列に翻訳し、核酸のデータベースである Gen Bank と EMBL に登録されている DNA 配列を可能な6フレームすべてについてアミノ酸配列に翻訳したものに対し、相同性検索を行った。その結果、ORF として認められている領域由来のアミノ酸配列と相同性を持つものが 24個、相同性を持たないものが 47個であった。また、71個という NSF 数は、p0AD2 全塩基配列から予想される NSF 数よりも多く、NSF の相補鎖の同じ枠にも NSF が存在する場合が多かった。これは非常に希なことである。

NSF と IS6100 の関係を考えるため、IS6100 のプロモーターを解析した結果、IS6100 は、それ自身のプロモーター（トランスポゼース）により、下流のプロモーターを持たない遺伝子を発現できるだけでなく、IR 中に持つ "TTGCAA" 配列に -10領域が付加されれば、プロモーター活性を示すことがわかった。

総括

タンパク質あるいは酵素は非常に多種類存在する。これらタンパク質は、相同性により分類することができる。各グループに属するタンパク質は、他のグループのタンパク質と異なるそれぞれ1つの先祖型タンパク質が、長い年月の間に変異と重複を繰り返し、分岐し、現在の如く、グループ内の各タンパク質に分化したと考えられる。では、この先祖型タンパク質はどのようにして生じたのであろうか。それは、分子進化学上の興味のみならず、現在存在しないタンパク質が、どのように生まれてくるのかを考察する上で重要である。

Flavobacterium sp. KI72 株は、非天然型化合物であるナイロンオリゴマーを分解し、単一の炭素源および窒素源として資化する。ナイロンオリゴマーの分解には、3種類の酵素 EI, EII, および EIII が関与する。Flavobacterium sp. KI72 株と同じくナイロンオリゴマー資化性菌である Pseudomonas sp. NK87 株も EI と EII 酵素を持っている。このうち、EII 酵素ファミリー遺伝子, nylB, と EIII 酵素遺伝子, nylC, の相補鎖の同じ枠には、終止コドンが存在しない NSF が認められる。これら遺伝子の相補鎖に NSF が存在する確率は非常に小さく、0.0001-0.0080 であり、このような NSF が存在するには何かの機構がある。F-nylB と P-nylB の進化距離を計算すると、 1.4×10^8 年であり、この間フレームが開放され続けるためには、何か特別な力が働いているためと考えられる。新しい遺伝子が誕生するためには、読み取り枠が開放されていることが重要と考えられ、長い年月の間、終止コドンが生まれず開放されたままになっているナイロンオリゴマー分解酵素遺伝子の相補鎖に存在するような NSF は、新規遺伝子が発生する1つの場所と考えられる。nylB と nylC の相補鎖に存在する NSF 由来のアミノ酸配列は、データベース上の遺伝子由来のアミノ酸配列と相同性が認められず、もしこの NSF がタンパク質を産生したならば、今まで知られているタンパク質と進化的に関係のないタンパク質となる。nylB と nylC は Flavobacterium sp. KI72 株が保有する3種類のプラスミドの中で pOAD2 上にコードされており、本研究では、この pOAD2 の特徴を知り、新しい酵素が発現する可能性を検討した。

第一章では、ナイロンオリゴマーの分解に関与する3種の酵素 EI, EII, および EIII を支配する pOAD2 プラスミドの全塩基配列の決定を行い、以後の解析の足掛りとした。その結果、全長は、45,519 bp よりなり、その G+C 含量は 66.6% であった。300 bp 以上の ORF が 157個認められ、600 bp 以上にわたり終止コドンの存在しない NSF が 71個認められた。

第二章では、pOAD2 上に存在する2つの繰り返し配列 (RS-I と RS-II) について解析を行った。5ヶ所に存在する各 RS-I の塩基配列を比較したところ、RS-IA, RS-IC, RS-ID, および RS-IE と RS-IB の前半 880 bp は完全に同一の配列であり、RS-IB はこの 880 bp の後半 420 bp が繰り返していた。相同性検索の結果、RS-I 領域は、Mycobacterium fortuitum 由来の IS6100と全く同一の配列であった。IS6100 (RS-I) を他の IS と比較した結果、IS6 ファミリーに属すること

がわかった。IS6100 は、*Mycobacterium smegmatis* 内で トランスポゾン Tn610 として転移能を持つが、*Flavobacterium* sp. KI72 株内でも IS6100 がコードするトランスポゼースの転写が微弱ながら認められた。IS6100 の分布を調べると、KI72 株と同じくナイロンオリゴマー資化性菌である *Pseudomonas* sp. NK87 株が保有する pNAD2 プラスミド上に IS6100 と高い相同性を有する配列が存在した。しかし、KI72 株と NK87 株の染色体 DNA やこれらの菌株が保有する pOAD2 あるいは pNAD2 以外のプラスミドには、IS6100 領域は認められなかった。RS-I 領域と同じく pOAD2 上の数ヶ所に存在する繰り返し配列 (RS-II 領域) は、EII ならびに EII' 酵素をコードしている。EII と EII' はともに 392 アミノ酸残基よりなり、47 残基が異なるために EII が持つ活性は EII' のそれより 100 倍大きい。この活性の違いを与えるアミノ酸残基置換を調べると、Gly¹⁸¹→Asp¹⁸¹ と His²⁶⁶→Asn²⁶⁶ (EII'→EII) の 2 アミノ酸残基置換のみで EII' レベルから EII レベルへの活性上昇が得られることがわかった。IS6100 が、RS-II 領域を含め、ナイロンオリゴマー分解性遺伝子の複製や転座に重要な役割を果たしているとも考えられたが、各遺伝子の進化距離からすると、各遺伝子が分岐する前に IS6100 が関与した可能性は否定される。

第三章では、pOAD2 上に存在する 71 個の NSF について解析した。71 個の NSF 由来のアミノ酸配列をデータベースに登録されている DNA を可能な 6 フレームすべてについてアミノ酸に翻訳し、相同性検索を行った。その結果 coding 領域由来のものと相同性を持つものが 24 個、coding 領域由来のものとは相同性がなかったものが 47 個存在した。coding 領域由来のアミノ酸配列と相同性があったもののうち代表的なものは、プラスミドの複製あるいは分配に関与するもの、iso-penicillin-N acyltransferase、細胞分裂に関与するもの、オリゴペプチドの取り込みに関与するものなどであった。coding 領域由来のものとは相同性を持たなかった 47 個の NSF は、これがもしタンパク質を産生したならば、それは、既存のデータベースにないタンパク質である。また、pOAD2 上の NSF の総数は、予想されるより多かった。興味深いことに、NSF の相補鎖の同じ枠にも NSF が存在する場合が多く、23 組 50 個もの NSF がその相補鎖に NSF を持っている。通常、機能を持たない NSF には、特別な選択圧がかからない限りランダムな変異により終止コドンが生まれ、その結果 NSF が消失するはずである。これら NSF の存在は、何らかの力により枠が解放されているためと考えられる。新しい遺伝子が誕生するためには、読み取り枠が開放されていることが重要である。これら NSF は新規遺伝子の 1 つの候補となっていると考える。

IS6100 とこれら NSF との関係を考えるため、IS6100 のプロモーター活性を調べた。その結果、IS6100 自身のコードするトランスポゼースのプロモーターにより、下流に位置する ORF を発現できることがわかった。また、IS6100 が、その末端配列中に持つ "TTGCAA" 配列に -10 領域が付与されるとハイブリッドプロモーターを形成して IS6100 の上流に向かっても、下流に向かってもプロモーター活性を示すことがわかった。また、IS6100 には、特定の挿入部位が認められないことから、IS が持つ転移能力によりプロモーターを持たない NSF にプロモーター (トランスポゼースのプロモーター、またはハイブリッドプロモーター) を付与する機会が増えると考えられる。

以上、遺伝子の相補鎖の同じ枠に NSF を持つ nylB と nylC をコードする pOAD2 プラスミドを調べると、NSF が予想されるより多く 71個存在していた。その中で 47個の NSF は、データベースに登録されている遺伝子と相同性のない、進化的に関係のないものであった。23組 50個の NSF は、その相補鎖の同じ枠にも NSF が存在しており、このうち、19組 39個の NSF は、その一方が相同性検索の結果、遺伝子として機能する可能性のあるものである。これら NSF は、KI72 株以外の微生物や pOAD2 プラスミド以外の DNA 領域で遺伝子として機能していたものが pOAD2 上に転座したものかもしれない。なぜ遺伝子の相補鎖の同じ枠が開放されているのかわからないが、pOAD2 上では、nylB と nylC の場合に限らず相補鎖の同じ枠が開放されている。新しい遺伝子が生まれるためには、読み取り枠が開放されていることが重要であるが、このように NSF が保たれているままになっているところは、新しい遺伝子が誕生するのに適している。また、この他 pOAD2 上には、IS6 ファミリーに属する IS6100 が存在していた。この IS は、プロモーターを持たない ORF にプロモーターを付与する能力がある。このプロモーター活性は決して弱くなく、もし IS が挿入された場合、その下流に位置する遺伝子の転写量は激増するはずである。その転写物が翻訳され、タンパク質として発現し、そのタンパク質が菌体にとり有用であれば、その後、変異を重ね、酵素として進化していくと考えられる。過去において機能し、これからも移動する IS6100 は、pOAD2 上に存在する NSF を活性化する可能性がある。NSF が多数存在し、その上、相補鎖の同じ枠にも NSF を持つものが多い pOAD2 プラスミドは、新しい酵素を生み出す可能性を秘めている。IS の新規酵素発生への関与を考えた場合、nylB と nylC と同じくプラスミド上の遺伝子であるハロゲン化酢酸分解酵素遺伝子、トルエン分解酵素群の遺伝子の近傍にも IS が見られ、これら遺伝子の相補鎖の同じ枠にも NSF が存在する。このことから、これらのプラスミドも pOAD2 と同様なプラスミドと考えられる。染色体外因子であるプラスミドに IS が挿入されても宿主には影響が少ない。NSF が存在し、また IS が存在するこのようなプラスミドで遺伝子が誕生するケースが多いのではないかと考える。NSF を数多く持つ細菌を使えば、新規の基質に対して活性を持つ新酵素を取得できるかもしれない。また、この時、使用する細菌で機能する IS の転移能を制御できれば、その酵素の取得効率を増加させることができるであろう。

参考文献

- 1) S. Kinoshita, S. Kageyama, K. Iba, Y. Yamada and H. Okada; *Agric. Biol. Chem.*, 39, 1219 (1975)
- 2) S. Kinoshita, S. Negoro, M. Muramatsu, V. S. Bisaria, S. Sawada and H. Okada; *Eur. J. Biochem.*, 80, 489 (1977)
- 3) S. Kinoshita, T. Terada, T. Taniguchi, Y. Takene, M. Matsunaga and H. Okada; *Eur. J. Biochem.*, 116, 547 (1981)
- 4) S. Negoro, S. Kakudo, I. Urabe and H. Okada; *J. Bacteriol.*, 174, 7948 (1992)
- 5) T. Yamada, C. J. Palm, B. Brooks and T. Kosuge; *Proc. Natl. Acad. Sci. U. S. A.*, 82, 6522 (1985)
- 6) S. Negoro, H. Sinagawa, A. Nakata, S. Kinoshita, T. Hatozaki and H. Okada; *J. Bacteriol.*, 143, 238 (1980)
- 7) H. Okada, S. Negoro, H. Kimura and S. Nakamura; *Nature (London)*, 306, 203 (1983)
- 8) K. Kanagawa, S. Negoro, N. Takada and H. Okada; *J. Bacteriol.*, 171, 3181 (1989)
- 9) M. Kimura; *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, U. K.). (1983)
- 10) T. Yomo, I. Urabe and H. Okada; *Proc. Natl. Acad. Sci. U. S. A.*, 89, 3780 (1992)
- 11) C. Yanisch-Perron, J. Vieira and J. Messing; *Gene*, 33, 103 (1985)
- 12) W. O. Bullock, J. M. Fernandez and J. M. Short; *Biotechniques* 5, 4, 376 (1987)
- 13) F. Boliva, R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Hynecker, H. W. Boyer, J. H. Crosa and S. Falkow; *Gene*, 2, 95 (1977)
- 14) J. M. Short, et al.; *Nucleic. Acids Res.*, 16, 7583 (1988)
- 15) M. A. Altling-Meese and J. M. Short; *Nucleic. Acids Res.*, 17, 9494 (1989)
- 16) J. Vieira and J. Messing; *Gene*, 19, 259 (1982)
- 17) J. Messing; *Methods in Enzymol.*, 101, 20 (1983)
- 18) T. Maniatis, E. F. Fritsh and J. Sambrook; *Molecular Cloning, A laboratory mannual*
- 19) H. C. Birnboim and J. Doly; *Nucleic. Acids Res.*, 7, 1513 (1979)
- 20) M. He, A. Wilde and M. A. Kaderbhai; *Nucleic. Acids Res.*, 18, 1660 (1989)
- 21) S. N. Cohen, A. C. Chang and L. Hsu; *Proc. Natl. Acad. Sci. U. S. A.*, 69, 2110 (1972)
- 22) B. Vogelstein and D. Gillespie; *Proc. Natl. Acad. Sci. U. S. A.*, 76, 615 (1979)
- 23) A. M. Maxam and W. Gilbert; *Methods in Enzymol.*, 65, 499 (1980)
- 24) S. Negoro, T. Taniguchi, M. Kanaoka, H. Kimura and H. Okada; *J. Bacteriol.*, 155, 22 (1983)

- 25) S. Negoro, S. Nakamura and H. Okada; *J. Bacteriol.*, 158, 419 (1984)
- 26) B. J. Bachmann; *Bacteriol. Rev.*, 36, 525 (1972)
- 27) R. K. Appleyard; *Genetics*, 39, 429 (1959)
- 28) L. J. Szabo; *J. Bacteriol.*, 157, 821 (1984)
- 29) K. Struhl, J. A. Smith, J. G. Seidman, D. D. Moore, R. E. Kingston, R. Brent and F. M. Ausubel; "CURRENT PROTOCOLS IN MOLECULAR BIOLOGY" (Greene Publishing Associates and Wiley-Interscience, 1987)
- 30) N. J. Zoller and M. Smith; *Methods in Enzymol.*, 100, 468 (1983)
- 31) A. P. Feinberg and B. Vogelstein; *Anal. Biochem.*, 137, 266 (1984)
- 32) P. S. Thomas; *Proc. Natl. Acad. Sci. U. S. A.*, 77, 5201 (1980)
- 33) M. C. O'Donovan, P. Buckland and P. McGuffin; *Nucleic. Acids Res.*, 19, 3466 (1991)
- 34) E. M. Southern; *J. Mol. Biol.*, 98, 503 (1975)
- 35) A. J. Jeffreys and R. A. Flavell; *Cell*, 12, 429 (1977)
- 36) 松崎, 成内, 臼井; 免疫学実験入門, 学会出版センター, pp102 (1981)
- 37) C. B. Laurell; *Anal. Biochem.*, 15, 45 (1966)
- 38) S. Negoro, S. Nakamura, H. Kimura, K. Fujiyama, Y. Z. Zhang, N. Kanzaki and H. Okada; *J. Biol. Chem.*, 259, 13648 (1984)
- 39) H. A. de Boer, L. A. Camstock and M. Vasser; *Proc. Natl. Acad. Sci. U. S. A.*, 80, 21 (1983)
- 40) C. Martin, J. Timm, J. Rauzier, R. Gomez-Lus, J. Davies and B. Gicquel; *Nature*, 345, 739 (1990)
- 41) J. Hein; *Methods in Enzymol.*, 183, 626 (1990)
- 42) L. Barberis-Maino, B. Berger-Baechi, H. Weber, W. D. Beck and F. H. Kayser; *Gene*, 59, 107 (1987)
- 43) D. A. Rouch, L. J. Messeroti, L. S. L. Loo, C. A. Jackson and R. A. Skurreay; *Mol. Microbiol.*, 3, 161 (1989)
- 44) L. Barberis-Maino, C. Ryffel, F. H. Kayser and B. Berger-Baechi; *Nucleic. Acids Res.*, 18, 5548 (1990)
- 45) A. J. Haandrikman, C. V. Leeuwen, J. Kok, P. Vos, W. M. de Vos and G. Venema; *Appl. Environ. Microbiol.*, 56, 1890 (1990)
- 46) K. M. Polzin and M. Shimizu-Kadota; *J. Bacteriol.*, 169, 5481 (1987)
- 47) P. Trieu-Cuot and P. Courvalin; *Gene*, 30, 113 (1984)
- 48) C. J. Wrighton and P. Dtrike; *Plasmid*, 17, 37 (1987)
- 49) E. J. Nuecken, R. B. Henschke and F. R. J. Schmidt; *DNA Sequence*, 1, 85 (1990)
- 50) A. Delecluse, C. Bourgouin, A. Klier and G. Rapoport; *Plasmid*, 21, 71 (1989)

- 51) T. Tsuchiya, S. Fukuyama, N. Kanzaki, K. Kanagawa, S. Negoro and H. Okada;
J. Bacteriol., 171, 3187 (1989)
- 52) I. D. Priyambada, S. Negoro and H. Okada; Ann. Reports ICBiotech., 10, 221 (1987)
- 53) K. Kato, K. Fujiyama, H. S. Hatanaka, I. D. Priyambada, S. Negoro, I. Urabe and
H. Okada; Eur. J. Biochem., 200, 165 (1991)
- 54) K. Y. Lee, J. D. Hopkins and M. Syvanen; J. Bacteriol., 172, 3229 (1990)
- 55) K. M. Derbyshire and N. D. F. Grindley; Cell, 47, 325 (1986)
- 56) J. C. Wang; Biochim. Biophys. Acta., 909, 1 (1987)
- 57) M. Takeuchi and A. Yokota; IFO Res. Comm., 15, 83 (1991)
- 58) A. S. Bhagwat, B. Johnson, K. Weule and R. J. Roberts; J. Biol. Chem.,
265, 767 (1990)
- 59) D. Berg, M. M. Howe (eds); Mobile DNA, American Society for Microbiology,
Washington DC. (1989)
- 60) D. Zerbib, M. Jakowec, P. Prentki, D. J. Galas and M. Chandler; EMBO J.,
6, 3163 (1987)
- 61) H. S. Hatanaka, K. Fujiyama, S. Negoro, I. Urabe and H. Okada;
J. Ferment. Bioeng., 71, 191 (1991)
- 62) S. Negoro, T. Mitamura, K. Oka, K. Kanagawa and H. Okada; Eur. J. Biochem.,
185, 521 (1989)
- 63) J. Kraut; Annu. Rev. Biochem., 46, 331 (1977)
- 64) C. S. Craik, S. Rocznik, C. Largman and W. J. Rutter; Science, 237, 909 (1987)
- 65) K. Kanagawa, M. Oishi, S. Negoro, I. Urabe and H. Okada; J. Gen. Microbiol.,
139, 787 (1993)
- 66) M. Nei and T. Gojobori; Mol. Biol. Evol., 3, 418 (1986)
- 67) J. Brosius; Gene, 27, 151 (1984)
- 68) E. Ohtuka, Y. Taniyama, R. Marumoto, H. Sato, H. Hirosaki and M. Ikehara;
Nucleic. Acids Res., 10, 2597 (1982)
- 69) K. B. Mullis and F. A. Faloon; Methods in Enzymol., 155, 335 (1987)
- 70) R. K. Saiki, S. Scharf, F. Faloon, K. B. Mullis, G. T. Horn, H. A. Erlich and
N. Arnheim; Science, 230, 1350 (1985)
- 71) S. Tabata, P. J. J. Hooykaas and A. Oka; J. Bacteriol., 171, 1665 (1989)
- 72) H. Mori, A. Kondo, A. Ohshima, T. Ogura and S. Hiraga; J. Mol. Biol.,
192, 1 (1986)
- 73) A. L. Abeles, S. A. Friedman and S. J. Austin; J. Mol. Biol., 185, 261 (1985)
- 74) J. A. Kornacki, P. J. Balderes and D. H. Figurski; J. Mol. Biol., 198, 211 (1987)

- 75) E. Montenegro, J. L. Barredo, S. Gutierrez, B. Diez, E. Alvarez and J. F. Martin;
Mol. Gen. Genet., 221, 322 (1990)
- 76) M. B. Tobin, M. D. Fleming, P. L. Skatrud and J. R. Miller; J. Bacteriol.,
172, 5908 (1990)
- 77) J. L. Barredo, P. van Solingen, B. Diez, E. Alvarez, J. M. Cantoral,
A. Kattevilder, E. B. Smaal, M. A. M. Groenen, A. E. Veenstra and J. F. Martin;
Gene, 83, 291 (1989)
- 78) D. R. Gill, G. F. Hatfull and G. P. C. Salmond; Mol. Gen. Genet., 205, 134 (1986)
- 79) M. Perego, C. F. Higgins, S. R. Pearce, M. P. Gallagher and J. A. Hoch;
Mol. Microbiol., 5, 173 (1991)
- 80) D. Z. Rudner, J. R. LeDeaux, K. Ireton and A. D. Grossman; J. Bacteriol.,
173, 1388 (1991)
- 81) I. D. Hiles, M. P. Gallagher, D. J. Jamieson and C. F. Higgins; J. Mol. Biol.,
195, 125 (1987)
- 82) G. Alloing, M. C. Trombe and J. P. Claverys; Mol. Microbiol., 4, 633 (1990)
- 83) W. Nietfeld, T. El-Baradi, H. Mentzel, T. Pieler, M. Koester, A. Poeting and
W. Knoechel; J. Mol. Biol., 208, 639 (1989)
- 84) U. Lendahl, L. B. Zimmerman and R. D. G. McKay; Cell, 60, 585 (1990)
- 85) R. L. Pastori, J. E. Moskaitis, L. H. JR. Smith and D. R. Schoenberg;
Biochemistry, 29, 2599 (1990)
- 86) R. G. Discipio and T. E. Hugli; J. Biol. Chem., 264, 16197 (1989)
- 87) M. Yamada, M. Shahjahan, T. Tanabe, F. Kishi and A. Nakazawa; J. Biol. Chem.,
264, 19192 (1989)
- 88) R. Kontermann, S. Sitzler, W. Seifarth, G. Petersen and E. K. F. Bautz;
Mol. Gen. Genet., 219, 373 (1989)
- 89) C. Chothia; Nature, 357, 543 (1992)
- 90) K. Schnetz and B. Rak; Proc. Natl. Acad. Sci. U. S. A., 89, 1244 (1991)
- 91) M. Okazaki, T. Uemura, Y. Nishimori, K. Wakamatsu, M. Shimosaka, H. Saito,
S. Negoro and H. Okada; J. Ferment. Technol., 66, 489 (1988)
- 92) Z. Y. Zhou, Y. Nishimori, M. Shimosaka and M. Okazaki; J. Ferment. Bioeng.,
70, 266 (1990)
- 93) D. A. Rouch, L. J. Messerotti, L. S. L. Loo, C. A. Jackson and R. A. Skurray;
Mol. Microbiol., 3, 161 (1989)

本論文に関係のある報告

1. Amino acid alterations essential for increasing the catalytic activity of the nylon-degradation enzyme of Flavobacterium sp.

Ko Kato, Kazuhito Fujiyama, Haruyo Sawai Hatanaka, Irfan Dwidya Priyambada, Seiji Negoro, Itaru Urabe, and Hirosuke Okada
Eur. J. Biochem., 200, 165-169 (1991)

2. Insertion sequence IS6100 on nylon oligomer degradative plasmid pOAD2

Ko Kato, Kinya Ohtsuki, Hiroyuki Mitsuda, Tetsuya Yomo, Seiji Negoro, and Itaru Urabe
J. Bacteriol., in press

3. Amino acid alterations increasing the catalytic activity of nylon oligomer degradation enzyme

Seiji Negoro, Ko Kato, Kazuhito Fujiyama, and Hirosuke Okada
submitted to Biodegradation

4. Structural analysis of nylon oligomer degradative plasmid pOAD2: whole nucleotide sequence of pOAD2

Ko Kato, Kinya Ohtsuki, Yuji Koda, Tetsu Maekawa, Tetsuya Yomo, Seiji Negoro, and Itaru Urabe
in preparation

謝辞

本研究を行うにあたり、終始御指導を賜りました卜部格教授に心から御礼申し上げます。本研究にあたり、御助言を賜りました岡田弘輔名誉教授（現、熊本工業大学教授）に厚く感謝の意を表します。本論文の作成にあたり、有益な御助言と御教示を賜りました今中忠行教授、山田靖宙教授、高野光男教授、大嶋泰治教授、菅健一教授、新名惇彦教授、塩谷捨明教授、吉田敏臣教授、二井將光教授に感謝の意を表します。

本研究に対し、直接御指導を賜りました根来誠司助教授、有益な御助言と激励の言葉を頂きました島康文氏、四方哲也博士に深く感謝致します。

本研究の共同研究者として御協力頂きました大槻欽哉氏（現、鐘淵化学工業株式会社）、幸田有司氏（現、伊藤忠商事株式会社）、前川哲氏（現、アサヒビール飲料製造株式会社）、光田浩之氏（現、宝酒造株式会社）に厚く御礼申し上げます。また、博士課程在学中有益な御助言、御討論頂きました三田村俊秀、莉沙夫妻、ならびに卜部研究室の皆様に深く感謝致します。さらに、公私にわたり有益な助言と激励を頂きました藤山和仁氏（現、大阪大学助手）、牧野泰孝氏（現、千葉大学助手）に感謝します。

最後に本研究の遂行に際して、援助して頂いた亡き父、ならびに母、さらに、激励、協力してくれた妻に心から感謝します。