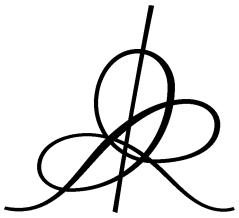
OPERATOR DYNAMICS IN MOLECULAR BIOLOGY

Tsuyoshi KATO



Institut des Hautes Études Scientifiques
35, route de Chartres
91440 – Bures-sur-Yvette (France)

Août 2001

IHES/M/01/41

OPERATOR DYNAMICS IN MOLECULAR BIOLOGY

T.KATO

Introduction

Recent develpment of molecular biology revealed some mathematical nature of the structure of genes. Basic elements of genes are DNA, which are (double helixed) sequences consisted by four bases, $X = \{A, T, G, C\}$. A primary structure of genes is just a finite sequence of X. There are higher structures, at least until fourth ones. These are related to folding of proteins, which reflects some phisical force.

One way to represent molecular biology is a mathematical formulation of molecular activity of life. Here we will use algebra A as a box which contains all informations in the formulation. Multiplication of elements of A corresponds to composition of functions of genes. Here one will use involutive algebras, in particular C^* ones, and will represent phisical states by vectors in A.

Frist we will fix a set $X = \{X_1, \ldots, X_n, \ldots\} \subset A$, whose elements will consist of *primary structure*. A primary structure is just a sequence of finite length by elements of X:

$$V = v_1 v_2 \ldots v_l, \quad v_i \in X.$$

Then we will express higher structure of V by elements $a_1, \ldots, a_{l+1} \in A$ where we will force *shift relation* as:

$$a_i a_{i+1}^* = v_i, \quad i = 1, 2, \ldots, l.$$

Combining these, we will express total structure as:

$$V = a_1[v_1]a_2[v_2]\ldots a_l[v_l]a_{l+1}.$$

We will call this as a *m-dna sequence*. The sequence $(a_1, \ldots, a_{l+1})$ will be called as a *state set* which should represent higher structures. State sets will be more flexible than the primary sets. For example folding of proteins can change by

phisical force, like catalyzing as ensymes. The energy of the state at v_i is defined as:

$$\text{trace } a_i a_i^* \in \mathbf{R} \cup \infty.$$

Thanks to functions of replication of genes, molecular lives are very active during its total life. Here we will introduce *infinitely many reflectivity* on m-dna sequences which represents possibility of replication. There are sufficiently many common operations which are shared among all lives. The aim here is, first to formulate these operations whose results are also required to be infinitely many reflective. Then after one will try to reconstruct some metabolisms, in particular Glycolysis, using particular choice of X, consisted by shift operators and projections. A basic problem will be to seek for primitive elements (metabolism) which can grow to Glycolysis. More general construction and analysis of the energy will be in the next paper.

The content is as follows:

The author has been very much influenced by a succesive seminar:

Initiation à la genomique fonctionnelle
: Aspects biologiques, mathematiques et algorithmiques

during $2000-2001$ at the institut Henri Poincar'e. He would like to thank to the organizers of the seminar. In particular he would like to express his gratitude to Professor Misha Gromov for encouragement, from whom he has received many stimulation. He also would be thankful to IHES for their hospitarity.

1. M-DNA SEQUENCES

Let A be a (non commutative) algebra with involution $*$. Let us choose subsets:

(1) $I = \{s_1, \dots, s_m\} \subset A$ and

(2) $X = \{X_1, \dots, X_n, \dots\} \subset A$. We say that I is the initial sets and X is the primary sets. We will say that (A, I, X) as a *m-dna data*.

A *m-dna sequence* is a finite length string $V = X_1 X_2 \dots X_l$, where X_i are in X with extra structures described below. Both sides of each X_i are assigned with elements a_i and a_{i+1} in A and $a_1 \in I$ which satisfy the *shift relation*:

$$X_1 = a_1 a_2^*, \quad \dots \quad X_i = a_i a_{i+1}^*, \quad \dots$$

We will denote these data as:

$$V = a_1[X_1]a_2[X_2]\dots a_l[X_l]a_{l+1}.$$

Usually one assumes $I = \{id\}$.

Let $W = X_1 \dots X_l$ be a string which are assigned with only a sequence of primary sets. Then one can assign m-dna sequences $V = a_1[X_1]a_2[X_2]\dots a_l[X_l]a_{l+1}$ if there exists a family $\{a_i\}_i$ satisfying the shift relation. We will denote the set of m-dna sequences with respect to W as:

$$V(W, A) = \{V = a_1[X_1]a_2[X_2]\dots a_l[X_l]a_{l+1} : \text{ m-dna sequence}\}.$$

We will say $W = X_1 \dots X_l$ is a *primary sequence* with respect to V.

Example 1.A: Let $V = a_1|X_1|a_2|\dots|X_l|a_{l+1}$ be a m-dna sequence. Then $V^* \equiv a_{l+1}|X_l^*|a_l|\dots a_2|X_1^*|a_1$ is also a m-dna sequence. We call this assignment $V \to V^*$ as an *involution*. We will say that V is *palindrome*, if $V = V^*$.

2. MUTATION

2.A Point mutation: Let $V = a_1|X_1|a_2\dots a_l|X_l|a_{l+1}$ be a m-dna sequence. A *point mutation* of V is another m-dna sequence:

$$V' = a_1'|X_1\dots a_{k-1}'|a_k'|X_k'|a_{k+1}'|X_{k+1}\dots a_{l+1}'$$

(V' differes from V in primary sequences only by replacing X_k by another X_k').

Example 2.A: Let P and P' be projections on a Hilbert space where the equality $PP' = P'$ holds. Then $V = id[id]id[P]P$ is a m-dna sequence. $V' = id[id]id[P']P'$ is a mutation of V.

2.B Intron: Let V be a m-dna sequence. Supose the string splits into three parts:

$$\begin{aligned} V &= V_1 \cup V_2 \cup V_3 \\ &= a_1|X_1|a_2|X_2 \dots a_{l_1}|X_{l_1}|a_{l_1+1} \cup a_{l_1+1}|X_{l_1+1}|a_{l_1+2}|X_{l_1+2} \dots a_{l_2}|X_{l_2}|a_{l_2+1} \\ &\quad \cup a_{l_2+1}|X_{l_2+1}|a_{l_2+2}|X_{l_2+2} \dots a_l|X_l|a_{l+1}. \end{aligned}$$

An *elimination* of V_2 from V consists of another m-dna sequence:

$$V_1 \cup V_3 = a'_1|X_1|a'_2|X_2 \dots a'_{l_1}|X_{l_1}X_{l_2+1}|a'_{l_2+2}|X_{l_2+2} \dots a'_l|X_l|a'_{l+1}.$$

Let us say that for a m-dna sequence V, V_2 is a *pre-intron*, if $V_1 \cup V_3$ exists. $V_1 \cup V_3$ is called as a *spliced m-dna sequence*.

2.C Amplification: Let us take m-dna sequences $V = a_1|X_1|a_2 \dots a_{l-1}|X_{l-1}|a_l$ and $V' = b_1[Y_1]b_2[Y_2] \dots [Y_k]b_{k+1}$. Suppose that there is a primary element Y such that:

$$V[Y]V' = a'_1[X_1]a'_2 \dots a'_{l-1}[X_{l-1}]a'_l[Y]b'_1[Y_1]b_2 \dots b'_{k-1}[Y_{k-1}]b'_k$$

consists of also a m-dna sequence. We will say that $V[Y]V'$ is an *amplification* and Y is a *connecting element*.

Let us take two m-dna sequences V and V'. Suppose there is connecting element Y which makes an amplification $V|Y|V'$. We will say that this mutation is *non commutative*, if $V'|Y|V$ does not admit any m-dna sequence.

Example 2.C: Let P, P', P'' be projections satisfying $PP' = P'$ and $P'P'' = P''$. Let us consider:

$$V = id|P|P, \quad V' = id|P''|P'', \quad Y = id|P'|P',$$
$$V|Y|V' = id[P]P[P']P'[P'']P'', \quad V'|Y|V = id[P'']P''[P']P'[P]P.$$

$V|Y|V'$ satisfies shift relation, but $V'|Y'|V$ does not when $P'' \neq P'$.

2.D Frame shift: Let $V = a_1|X_1|a_2| \dots a_l|X_l|a_{l+1}$ be a m-dna sequence, and $u \in A$ be a unitary, $uu^* = \mathrm{id}$. Then $V(u) = a_2u|X_1|a_3u| \dots |a_lu|X_{l-1}|a_{l+1}u$ is

also a m-dna sequence, and we will say that V' is a *frame shift* of V.

Example 2.D: Let $V = a_1|X_1|a_2|\dots a_l|X_l|a_{l+1}$ be a m-dna sequence, and take a unitary $\sigma \in A$ with $\sigma^3 = \mathrm{id}$. Then one considers another m-dna sequence:

$$\tilde{V} = \begin{pmatrix} \sigma & 0 \\ 0 & a_1 \end{pmatrix} [\begin{pmatrix} \sigma^* & 0 \\ 0 & X_1 \end{pmatrix}] \begin{pmatrix} \sigma^2 & 0 \\ 0 & a_2 \end{pmatrix} [\begin{pmatrix} \sigma^* & 0 \\ 0 & X_2 \end{pmatrix}] \begin{pmatrix} 1 & 0 \\ 0 & a_3 \end{pmatrix}$$

$$[\begin{pmatrix} \sigma^* & 0 \\ 0 & X_3 \end{pmatrix}] \begin{pmatrix} \sigma & 0 \\ 0 & a_4 \end{pmatrix} [\begin{pmatrix} \sigma^* & 0 \\ 0 & X_4 \end{pmatrix}] \begin{pmatrix} \sigma^2 & 0 \\ 0 & a_5 \end{pmatrix} [\begin{pmatrix} \sigma^* & 0 \\ 0 & X_5 \end{pmatrix}] \dots$$

Now let us choose $u = \begin{pmatrix} \sigma & 0 \\ 0 & 1 \end{pmatrix}$. Then $\tilde{V}(u)$ expresses frame shift on *codon correspondence* between DNA (RNA) and amino-acids.

2.E Inequal crossing over: Let $V = V_1V_2$ and $V' = V_1'V_2'$ be two m-dna sequences with primary sequences as $(X_1(') , \dots, X_{l(')}(') , Y_1(') , \dots, Y_{k(')}('))$ respectively. If both of another primary sequences $V_1V_2' = (X_1, \dots, X_l, Y_1', \dots, Y_{k'}')$ and $V_1'V_2 = (X_1', \dots, X_{l'}', Y_1, \dots, Y_k)$ also admit m-dna sequences U and U' respectively, then we will say that U and U' as *inequal crossing over* of V and V'.

2.F System of mutations: A system of mutation is a finite set $\mathfrak{S}$ where any element $s \in \mathfrak{S}$ represents some operation described from $2.A$ to $2.E$. Let $\bar{V}$ be a set of m-dna sequences. Then $\mathfrak{S}(\bar{V})$ gives another set of m-dna sequences. One may iterate this process as:

$$\bar{V} \subset \mathfrak{S}(\bar{V}) \subset \mathfrak{S}^2(\bar{V}) \subset \mathfrak{S}^3(\bar{V}) \subset \dots$$

This will represent a simple model of *evolution*.

Example 2.F: Let $M_n(\mathbf{C})$ be the set of $n \times n$ complex matrices, and put $M_\infty(\mathbf{C}) = lim_n M_n(\mathbf{C})$. Suppose the primary sets $X = \{X_1, \dots, X_n, \dots\} \subset M_\infty(\mathbf{C})$. Then we will call X as a *finite type*. Any set $\bar{V}$ whose primary sets consistes of finite type will be also called as finite type.

Let $\mathfrak{S}$ be the set of $2.A, 2.B, 2.C, 2.E$. Then N-th stage of evolution of $\bar{V}$ is also of finite type for any $N \geq 0$.

3. Replication

3.A Replication: Suppose the algebra A is a closed $*$ subalgebra of $B(H)$, where H is a separable Hilbert space. Let us say that an operator $a \in A$ is replicable, if there is a unitary operator $U : H \cong H \oplus H$ satisfying:

$$UaU^* = \begin{pmatrix} a' & 0 \\ 0 & a'' \end{pmatrix}.$$

We will say that a'' is a *reflective element*. Reflective elements are not unique. Let $X = \{X_1, \ldots, X_l, \ldots\}$ be the primary set. In the following, we will assume that X is *closed* under replication, in the sense that any X_i has a unitary U satisfying $UX_iU^* = \mathrm{diag}(X_j, X_k)$ for some j, k.

Let us take a m-dna data (A, I, X) where A is as above, and I and X are both consisted by replicable elements. Let us denote by Y_i as reflective elements of X_i. Let us take a m-dna sequence $V = a_1|X_1|a_2|X_2 \ldots a_l|X_l|a_{l+1}$. A *replication* of V is another m-dna sequence $b_1|Y_1|b_2| \ldots b_l|Y_l|b_{l+1}$ where each Y_i is a reflective element of X_i, using the same unitary U for all i.

Remark 3.A: (1) Suppose there are no invariant subspace for $a \in A$, and so no non trivial splitting as above. In this case, one replaces a by $\mathrm{diag}(a, a, \ldots)$ acting on $H \oplus H \oplus \ldots$ Then the latter splits as $\mathrm{diag}(a, a, \ldots) \oplus \mathrm{diag}(a, a, \ldots)$. Then even though $\mathrm{diag}(a, a, \ldots)$ contains the same information as a, this allows to make replication. (but the algebra A becomes much larger). In later examples, one assumes when a primary sequence $(X_1, \ldots, X_l)$ admits a m-dna sequence, then it always can make replication.

(2) Here one will only consider single sequences, even though in practice when genes make replication, the *double helix structure* plays an important role.

Example 3.A: (1) Let us take an orthnormal basis $H = \{v_0, v_1, v_2, \ldots\}$ and an unitary $U_0(v_{2i}) = v_i + 0, U_0(v_{2i+1}) = 0 + v_i \in H \oplus H$. Then $R : v_i \to v_{i+2}$ is replicable.

(2) Let us modify U_0 above slightly, and define $U_0(2i,0)$ by the following:

$$U_0(2i,0)(v_k) = \begin{cases} v_{2i} + 0, & k = 2i, \\ v_i + 0, & k = 4i, \\ v_j + 0, & k = 2j, j \neq i, 2i, \\ 0 + v_j, & k = 2j+1. \end{cases} \tag{3.1}$$

Then $U_0(2i,0)$ are all unitary and satisfy $U_0(2i,0)(v_{2i}) = v_{2i}$. Let $P(2i,0)$ be the projection to v_{2i}, $P(2i,0)(v_{2i}) = v_{2i}$. Then clearly $P(2i,0)$ are all replicable. One can obtain $P(i,j)$ similarly.

(3) Let us take a Hilbert space H and its basis $\{\dots, v_{-l}, \dots, v_0, v_1, v_2, \dots\}$. Let us put shift operators as follows:

$$Q : \begin{cases} v_{2i+1} \mapsto v_{2i+3} \\ v_{2i} \mapsto v_{2i+4} \end{cases}, \quad P : v_i \mapsto v_{i+1}, \quad R = P^2 : v_i \mapsto v_{i+2}.$$

Let us consider two m-dna sequences:

$$(1) \quad id[P^*]P[Q]Q', \quad (2) \quad id[R^*]R[Q]R'.$$

Using U_0 above, these have replications as follows:

$$(1) \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} 0 & 1 \\ P^* & 0 \end{bmatrix} \begin{pmatrix} 0 & P \\ 1 & 0 \end{pmatrix} \begin{bmatrix} R & 0 \\ 0 & P \end{bmatrix} \begin{pmatrix} 0 & P^* \\ P^* & 0 \end{pmatrix}.$$

Thus (1) is not reflective. For (2), we have a splitting:

$$(2) \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} P^* & 0 \\ 0 & P^* \end{bmatrix} \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \begin{bmatrix} R & 0 \\ 0 & P \end{bmatrix} \begin{pmatrix} P^* & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus the reflective element is $id[P^*]P[P]id$ (parindrome) which also consists of a m-dna sequence. The second stage of replication is as follows:

$$(2)' \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} 0 & 1 \\ P^* & 0 \end{bmatrix} \begin{pmatrix} 0 & P \\ 1 & 0 \end{pmatrix} \begin{bmatrix} 0 & P \\ 1 & 0 \end{bmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus the above reflective element is not secondary reflective. Thus the sequence (1) cannot replicate. (2) can do only once.

3.B Iterated replication: Let $X' = (X_1, \dots, X_l)$ be a primary sequence. Let

us choose an isomorphism $U : H \cong H \oplus H$ with $UX_iU^* = \mathrm{diag}(X_i(1)', X_i(1))$. If one can find another unitary U_1 as above with a decomposition $U_1X_i(1)(U_1)^* = \mathrm{diag}(X_i(2)', X_i(2))$, we say that $X_i(2)$ is a *secondary reflection element*. One may iterate this process, and find seccesively $U_2, U_3, \dots$ and $X_i(2), X_i(3), \dots$

Recall $V(X', A)$ in section 1. Let $V = a_1|X_1|a_2 \dots a_l|X_l|a_{l+1}$ be a m-dna sequence. Then V is *infinitely many reflective*, if $V((X_1(i), \dots, X_l(i)), A)$ are non empty for all i.

Example 3.B: Let us choose arbitrarilry a family $\{X_0, X_1, \dots\} \subset X$, choose any isomorphism $H \cong$ closure $\oplus_{i=0}^{\infty} H_i$ where H_i are all the same Hilbert spaces. Passing through this, one obtains an operator $X' = \mathrm{diag}(X_0, X_1, \dots)$ acting on H. Then one naturally gets a reflective element $X_1' \equiv \mathrm{diag}(X_1, X_2, \dots)$. As a secondary reflective element, one gets $X_2' \equiv \mathrm{diag}(X_2, X_3, \dots)$. Iterating one obtains n-th reflective element $X_n' \equiv \mathrm{diag}(X_n, X_{n+1}, \dots)$. Thus the realizing problem of reflective elements has a trivial solution.

Similarly choose any family $\{\dots, X_{-l}, \dots, X_0, X_1, \dots\}$ and an isomorphism $H \cong$ closure $\oplus_{i=-\infty}^{\infty} H_i$. In the both sides case also, the realizing problem has a trivial solution, from an operator $X'' = \mathrm{diag}(\dots, X_0, X_1, \dots)$.

Let us take a unitary $U : H \cong H =$ closure $\oplus H_i$, and choose any vectors $v_i \in H_i$ and $v_j \in H_j$, $i < j$. Then we will say that U *preserves the order* on (v_i, v_j), if there is a pair $i' < j'$ with $U(v_i) \in H_{i'}$ and $U(v_j) \in H_{j'}$.

Now let $\{\dots, X_{-l}, \dots, X_0, X_1, \dots\}$ be as above, and consider one side and both sides sequences:

$$X' = (X_{i_0}, X_{i_1}, \dots) \in B(H) = B(\text{ closure } \oplus_{i=0}^{\infty} H_i),$$
$$X'' = (\dots, X_{i_{-l}}, \dots, X_{i_0}, X_{i_1}, \dots) \in B(H) = B(\text{ closure } \oplus_{i=-\infty}^{\infty} H_i).$$

where the indices $I' = \{i_0, i_1 \dots\}, I'' = \{\dots, i_{-l}, \dots, i_0, i_1, \dots\} \subset \{0, 1, 2, \dots\}$ satisfy the followings:

(1) each $l \in \{0, 1, 2, \dots\}$ appears infinitely many times in both of I' and I'' in the positive direction, and

(2) for a sufficiently large j_0, $i = 0$ for all $i \leq -j_0$ in I''.

Let us consider a problem to seek for a unitary $U : H \cong H$ satisfying the following:

(1) $U^*X'U = \mathrm{diag}(X'_0, X'_1)$ where X_0 does not appear in X'_1,
(2) $X'_0 = X'$, and
(3)for any i_l, i_k in X'_m component, U preserves the order between H_{i_l} and H_{i_k} for $m = 0, 1$ respectively.

We will call such an operator as an *order preserving separating unitary*. Notice that once a family $\{X_i\}_i$ is fixed, then X' and X'' are expressed as one side or both sides sequences by $\{0, 1, \dots\}$ respectively.

An unexpected answer is that (1) in one side case, there are always no order preserving separating unitary, and (2) in both sides case, it always exists. In terms of an expression of numbers, this is the following lemma:

Lemma 3.B: Let $X' = (x_0, x_1, \dots)$ be a one side sequence in $\{0, 1, \dots\}$. Then there are no division of X' into two sequences $X' = Y \cup Z$ ($Y = (x_{m(l)})$ and $Z = (x_{n(l)})$ with $\mathbf{N} = \{m(l)\}_l \cup \{n(l)\}_l$) such that

(1) both of Y and Z consist of infinite sequences and they are order preserving,

(2) $Y = X'$ as sequences and (3) Z does not contain 0.

Let X'' be a both side sequence. Then there is a division of X'' into two sequences satisfying (1) , (2) and (3) above.

Proof: This can be seen from the next two examples. Let us consider $X'' = (\dots, 0, 0, 0, 1, 0, 1, 2, 0, 1, 2, 3, \dots)$. Then one can divide as:

$$X'' = (\dots, 0, 0, 1, 0, 1, 2, 0, 1, 2, 3, \dots) \cup (1, 2, 3, \dots).$$

In the case $X' = (0, 0, 1, 0, 1, 2, 0, 1, 2, 3, \dots)$, there are no such division.

Notice that automorphism groups of one or two sides subshift of finite types have mutually very different structures ([K]).

4. Ensyme

Ensymes play roles to activate or disactivate functions of genes. They can find particular genes to which they catalyze. Then a particular domain of ensymes (activating domain) attaches to domains of genes. For this, higher dimensional structures of genes are very important.

Let us take two m-dna sequences $V = a_1|X_1|a_2|\dots a_l|X_l|a_{l+1}$ and $W = b_1|Y_1|b_2|\dots b_k|Y_k|b_{k+1}$. Let us say that W *catalyzes* (as an *ensyme*) on V (m-dna sequence) at (i,j), if the primary sequence:

$$V[W] \equiv (X_1, \dots\dots X_{i-1}, X_iY_j, X_{i+1}, \dots X_l)$$

posesses another m-dna sequence. Namely there is $(a'_1, \dots, a'_{l+1})$ with:

$$a'_1|X_1|\dots a'_{l-1}|X_{l-1}|a'_l|X_iY_j|a'_{i+1}|X_{i+1}|\dots|a'_l|X_l|a'_{l+1}$$

consists of a m-dna sequence.

4.A Inversible ensymes: Let V and V' be two m-dna sequences. If there are ensymes W and W' with the equalities:

$$V[W] = V', \quad V'[W'] = V$$

then we will say that W and W' are mutually *inverse* ensymes w.r.t. (V, V').

Example 4.A: Let $H = \{v_0, v_1, \dots\}$ be a Hilbert space. Let us take S, the standard shift $S : v_i \mapsto v_{i+1}$ and P, the projection as $P(v_i) = v_i$, $i \neq 0, 1$ and $P(v_0) = P(v_1) = 0$. Let us take two m-dna sequences:

$$V = id[PS](PS)^*, \quad V' = id[S^2](S^2)^*.$$

Notice the equality $PS^2 = S^2$. Then one can find:

$$W = id[S]S, \quad W' : \begin{cases} v_0 \to 0 \\ v_i \to v_{i-1}, \quad i \geq 1. \end{cases}$$

These give mutually inversible ensymes.

Notice that the primary structure of $V[W]$ also differs from that of V. After W has stopped catalyzing, $V[W]$ will return to the original V. So in practice catalyzing should be inversible.

4.B Promotor: Let $V_0 = a_0|X_0|a_1|X_1|a_2|\dots|a_{l-1}|X_{l-1}|a_l$ be a sequence whose primary sequence does not admit any m-dna sequences. Let $W = id|Y|Y^*$ be a m-dna sequence so that the primary sequence $(X_0Y, X_1, \dots, X_{l-1})$ has m-dna sequences. Thus V_0 cannot replicate by itself. After catalyzed by W, $V = V_0[W]$ becomes a m-dna sequence and begins replicating. One will call V_0 as a *pre m-dna sequence*. X_0 is called a *promotor*.

Example 4.B: Let $H = \{\dots, v_{-l}, \dots, v_0, \dots, v_l, \dots\}$ be a Hilbert space, $Q(i)$ and $Q(i,j)$ be the projections on $\{v_i\}$ and $\{v_i, v_j\}$ respectively. $P(i) \equiv id - Q(i)$.

Let us consider a pre m-dna sequence $V_0 = id[id]id[Q(1)]Q(1)[Q(1,2)]Q(1,2)$ and take $W = id[P(2)]P(2)$. Then the primary sequence $(P(2), Q(1), Q(1,2))$ has a m-dna sequence $id[P(2)]P(2)[Q(1)]Q(1,2)[Q(1,2)]id$.

4.C Mutually exclusive promotors: Let $V = a_1|X_1|a_2|\dots a_l|X_l|a_{l+1}$ be a pre m-dna sequence. Suppose V posseses two promotors at X_i and X_j such that two ensymes Y_i and Y_j catalyze on these regions respectively.

If the following conditions are satisfied, then these two promotors are said to be *mutually exclusive*;

(1) both of $V[Y_i]$ and $V[Y_j]$ are infinitely many reflective,
(2) $V[Y_i][Y_j]$ is only finitely many reflective.

Examples 4.C: Let $P(i_1, \dots, i_N)$ be the corank N projection by:

$$P(i_1, \dots, i_N)(v_l) = \begin{cases} 0 & l = i_j, \\ v_l & \text{others.} \end{cases} \tag{4.1}$$

One puts $Q(i_1, \dots, i_N) = 1 - P(i_1, \dots, i_N)$. Moreover let S be the shift by $v_i \mapsto v_{i+1}$. Then one considers the primary sequence:

$$X = (P(1), P(0,1), P(-1,1)).$$

$V(X, A)$ is empty. Let us choose $Y_1 = Y_2 = S$. Then one has three primary sequences:

$$X_1 = (P(1)S, P(0,1), P(-1,1)), \quad X_2 = (P(1), P(0,1)S, P(-1,1)),$$
$$X_3 = (P(1)S, P(0,1)S, P(-1,1)).$$

$V(X_1, A)$ and $V(X_2, A)$ are both non empty, on the other hand $V(X_3, A)$ is empty as shown below:

$$id[P(1)S]S^{-1}P(1)[P(0,1)]S^{-1}P(2)[P(-1,1)]S^{-1}P(0) \in V(X_1, A),$$
$$id[P(1)]P(1)[P(0,1)S]S^{-1}P(0)[P(-1,1)]S^{-1}P(2) \in V(X_2, A),$$
$$id[P(1)S]S^{-1}P(1)[P(0,1)S]S^{-2}P(2)[P(-1,1)] * .$$

One can see there are no a with $S^{-2}P(2)a = P(-1,1)$ which is equivalent to $P(2)a = S^2P(-1,1)$, since the r.h.s. contains non zero v_2 component in the image, on the other hand the l.h.s. does not for any a.

In practice, many mutually exclusive promotors are found, e.g., the switching of grobin genes.

4.D Succesive ensymes: Let $V = a_1|X_1|a_2|\dots a_l|X_l|a_{l+1}$ be a m-dna sequence, and consider the primary sequence $(X_1, \dots, X_l)$. Suppose there are two states Y_i and Y_j catalyzing on V respectively.

If the following conditions are satisfied, then these two catalyzing processes are said to be *succesive*;

(1) $V[Y_1]$ and $V[Y_j]$ are only finitely many reflective, and
(2) $V[Y_i][Y_j]$ is infinitely many reflective.

Example 4.D: Let $P(i_1, \dots, i_N)$ and $Q(i_1, \dots, i_N)$ be as before. First one considers a simple case. Let us consider a m-dna sequence:

$$V = id[id]id[P(2)]P(2)[Q(1)]Q(1).$$

Then one chooses two ensymes:

$$Y_1 = P(1), \quad Y_2 = P(1).$$

When these catalyze on V, then the corresponding primary sequences become as:

$$\begin{gathered} V[Y_1] = (P(1), P(2), Q(1)), \quad V[Y_2] = (id, P(1,2), Q(1)), \\ V[Y_1][Y_2] = (P(1), P(1,2), Q(1)). \end{gathered} \tag{4.2}$$

Let (a_1, a_2, a_3, a_4) be a sequence with $a_1 = id$. Let us try to find m-dna sequences $V = a_1|P(1)|a_2|P(2)|a_3|Q(1)|a_4$, $V' = a_1|id|a_2|P(1,2)|a_3|Q(1)|a_4$ and $V'' = a_1|P(1)|a_2|P(1,2)|a_3|Q(1)|a_4$. V'' has a solution by:

$$V'' = id|P(1)|P(1)|P(1,2)|P(2)|Q(1)|Q(1)$$

One has partial solutions for:

$$V = id|P(1)|P(1)|P(2)|*, \quad V' = id|id|id|P(1,2)|P(1,2)|Q(1)|*\,.$$

Here it is impossible to find $*$ to satisfy the shift relation.

Extension 4.D: Let us take a Hilbert space $H' = H_1 \oplus H_2 \oplus H_3$ where each H_i is also an infinite dimensional Hilbert space. Let $P(i)$, $i = 1, 2, 3$, be the projections to H_i. Let us take another projection P and a primary sequence $X = (P, 0, P(1))$, where the first two states are promotors as above. Then there is a m-dna sequence with this primary sequence, if and only if $PP(1) = P(1)$.

Notice that any projection with infinite dimensional kernel and range are all unitary equivalent. Thus there is a unitary U satisfying:

$$U^*P(1)U = \begin{pmatrix} P(1) & 0 \\ 0 & P(1) \end{pmatrix}.$$

Let us take a family of unitaries U_i, $i = 1, 2, \ldots$ so that one gets inductively equations $U_i^* P_i U_i = \mathrm{diag}(P_i, P_{i+1})$, $P_1 = P$. Let $(a_1^i, a_2^i, a_3^i, a_4^i)$ be a sequence with $a_1^i = id$, and try to judge whether V_1 is infinitely many reflective, i.e., to find a family of m-dna sequences:

$$V_i = a_1^i|P_i|a_2^i|0|a_3^i|P(1)|a_4^i.$$

Then the following can be seen; V_1 is infinitely many reflective, if and only if $P(1) \subset\subset P_i$ are satisfied for all $i = 1, 2, \ldots$

4.E Switching, special cases: Let PV be a pre m-dna sequence where two different ensymes W and W' at i and (i, j), acts on the promotor P which are mutually exclusive. Let us denote $PV(1) = PV[W]$ and $PV(2) = PV[W']$. Here one considers a special case of functions of (PV, W) as follows:

(1) W catalyzes on P and produces $PV(1)$.
(2) $W' = PV(1)$ catalyzes on P, and produces $PV[W'] = PV(2)$.

Thus when $PV(1)$ has been produced too much, then $PV(2)$ also increases.

Example 4.E: Let $P(i_1, \ldots, i_N)$ and $Q(i_1, \ldots, i_N)$ be in 4.C. Then one considers a pre m-dna sequence $PV = id[id]id[Q(1)]Q(1)[Q(1,2)]Q(1,2)$ and an ensyme $W = id[P(2)]P(2)$. Then:

$$W' = id[P(2)]P(2)[Q(1)]Q(1,2)[Q(1,2)]id$$

catalyzes on P and produces

$$\begin{aligned} PV[W'] &= id[idQ(1)]Q(1)[Q(1)Q(1,2)]Q(1,2)[Q(1,2)]id \\ &= id[Q(1)]Q(1)[Q(1)]Q(1,2)[Q(1,2)]id. \end{aligned}$$

4.F Switching: Let us take a pre m-dna sequence V, and ensymes W, W', U, U' with a diagram:

$$\begin{array}{ccccc} & & V & & \\ & & \uparrow & & \\ W & -[U]\rightarrow & W' & -[U']\rightarrow & W \end{array} \tag{4.3}$$

Thus U, U' are ensymes which act on W and W' as mutual inverses. W' can catalyze on V and W cannot.

4.G Relation with inequally crossing over: Let us take two m-dna sequences $V = V_0V_1$ and $U = U_0U_1$. Suppose the inequally crossing over occurs:

$$(V, U) \mapsto (V_0U_1, V_1U_0).$$

This will produce new ensymes.

4.H Inactivating: Let us take a m-dna sequence V, and ensymes W and U, U'. suppose both of U and U' catalyze on W satisfying:

(1) $W[U]$ can catalyze on V,
(2) $W[U']$ cannot do it.

The ensyme U' is said to be *inactivating*.

4.I Indirect ensymes: Let us take two m-dna sequences V and W. Suppose there is a cycle of catalyzing system:

$$W - [V] \rightarrow W[V] \equiv W', \quad V - [W'] \rightarrow V[W'] \equiv V'.$$

Then we will say that W is an *indirect ensyme* for this catalyzing system $V \rightarrow V'$.

4.J Rigidity: Let us fix a primary sequence $X = \{X_1, \ldots, X_l\}$. Recall $V(X, A)$ in section 1. We will say that X is *rigid* if there is exactly only one $V \in L(X, A)$ with $a_1 \in I$ and $V = a_1|X_1|a_2 \ldots |a_l|X_l|a_{l+1}$.

In many of the examples here, the primary sets are rigid with $I = \{id\}$.

Let us take a m-dna sequence $V = a_1|X_1|a_2|X_2\dots$ and an ensyme $W = b_1|Y|b_2$. We will say that V is *projectively rigid*, if the following is satisfied; suppose W catalyzes on V at i. If $V[W]$ consists of a m-dna sequence, then one has the equality $X_i = X_iY$.

Example 4.J: Let us put $H = H_1 \oplus H_2$ and put P_i as the projections on H_i. Let us take a m-dna sequence $V = id[P_1]P_1[P_1]id[id]id$. Let us take another projection X. Let us put $W = id[X]X$ and catalyze W on $id[P_1]P_1$ factor in V. Then one has $V[W] = id[P_1X]P_1X[P_1]id[id]id$. Thus by the definition, if $V[W]$ consists of a m-dna sequence, then the equality $P_1X = P_1$ holds.

4.K Amplification: Let PV be a m-dna sequence. Let us consider a pre m-dna sequence PVV.

Suppose that V and VV catalyze on P as ensyme. If the following two conditions are satisfied, then the amplification PVV will be called *self controlling*:

(1) $PV[V]$ is infinitely many reflective,

(2) $PVV[VV]$ is only finitely many reflective.

If the converse holds:

(1) $PV[V]$ is only finitely many reflective,

(2) $PVV[VV]$ is infinitely many reflective,

then PVV will be called *self amplification*.

Example 4.K: Let us take a Hilbert space $H = H_1 \oplus H_2$ and put the projections on H_i by P_i. Let us consider m-dna sequences:

$$P = id|id|id|id|id, \quad V = id|P_1|P_1.$$

Then the followings show that PVV is self amplification:

$$PV = id[id]id[id]id[P_1]P_1, \quad PV[V] = id[P_1]P_1[id]id[P_1]P_1,$$
$$PVV = id[id]id[id]id[P_1]P_1[P_1]P_1, \quad PVV[VV] = id[P_1]P_1[P_1]P_1[P_1]P_1[P_1]P_1.$$

5. Selective splicing

Let $V = PV_1V_2$ be a (pre) m-dna sequence. If there is an ensyme W so that $V[W]$ has a maximal sequence V_1 of infinitely many reflection, then the

procedure $V \to V[W] \to V_1$ is called *splicing*.

Example 5.A: Let us put $H = H_1 \oplus H_2 \oplus H_3$ where H_i are infinite dimensional Hilbert spaces. Let us denote by P and P' as the projections on $H_1 \oplus H_2$ and on H_1 respectively. Let $V = P_1V_1P_2V_2$ be a m-dna sequence, where:

$$P_1 = id|P|P, \quad V_1 = P|P|id, \quad P_2 = id|id|id, \quad V_2 = id|id|id.$$

Let us take an ensyme $W = id|P'|P'$ and catalyze it on P_2. Then as a maximal sequence, one gets the following:

$$V' = id[P]P[P]id[P']P'.$$

5.B Splicing systems: Let $V = PV_1 \dots V_l$ be a (pre) m-dna sequence. Suppose there are subindices $\{m_1, \dots, m_k\} \subset \{1, \dots, l\}$ so that each V_{m_j} is also infinitely many reflective. We will say that a family of ensymes $\{W_1, \dots, W_k\}$ splices $\{V_{m_j}\}_j$, if each W_j splices V_{m_j}.

Let $V = PV_1V_1'V_2V_2' \dots V_lV_l'$ be a m-dna sequence. A *succesive splicing* on V is a system of splicing as follows; there is an ensyme W catalyzing on P, and replicating V_1. Then V_1 again catalyzing on P, and replicating V_2. One can continue this process until replicating V_l.

5.C Self splicing systems: Let $V = a_1|X_1|a_2 \dots$ be a (pre) m-dna sequence. Suppose an ensyme W catalyzes on V, $V_1 = V[W]$ and it produces W_1 by splicing. Next let us catalyze W_1 on V_1 catalyzing as an ensyme, and produces W_2. Similarly let us catalyze W_2 on V_2 and produce W_3. Succesively one may obtain (V_l, W_l), $l = 1, 2, \dots$ This process will be called as a *self splicing system*.

Example 5.C: Let us define shifts $S(n)$ and projections $P(n)$ as follows:

$$S(n) : \begin{cases} v_{2^n k} \to v_{2^{n+1} k} \\ v_i \to v_i, \quad i \neq 2^n k \end{cases}, \quad P(n) : \begin{cases} v_{2^n k} \to 0 \\ v_i \to v_i, \quad i \neq 2^n k \end{cases}, \quad k = 0, 1, 2, \dots$$

Notice the relations:

$$S(n)P(n) = P(n), \quad P(n)S(n-1) = P(n-1),$$
$$P(n-1)P(n) = P(n)P(n-1) = P(n-1).$$

Let us consider a primary sequence:

$$(S(1), S(2), \ldots, S(n)).$$

This admits a m-dna sequence $V = a_1[S(1)]a_2[S(2)]\ldots[S(n)]a_{n+1}$ with $a_1 = id$. For example one determines $a_2, a_3, \ldots$ as:

$$a_2^* = S(1), \quad a_3^* : \begin{cases} v_{4k} \to v_{16k} \\ v_{4k+2} \to v_{2(4k+2)} \\ v_{2k+1} \to v_{2k+1} \\ \quad (k = 0, 1, \ldots) \end{cases}, \quad a_4^* : \begin{cases} v_{8k} \to v_{16(4k)} \\ v_{8k+4} \to v_{4(8k+4)} \\ v_{4k+2} \to v_{2(4k+2)} \\ v_{2k+1} \to v_{2k+1} \end{cases}, \ldots$$

Thus $V^* = a_{n+1}[S(n)^*]a_n[S(n-1)^*]\ldots[S(1)^*]a_1$ also consists of a m-dna sequence. Now let us consider a primary sequence:

$$S = (S(n)^*, S(n-1)^*, \ldots, S(1)^*).$$

Take an ensyme $W = id[P(n+1)]P(n+1)$ and catalyze it on V^* as follows:

$$\begin{aligned} V_1 = V^*[W] &= a_{n+1}[S(n)^*P(n+1)]a_nP(n)[S(n-1)^*]a_{n-1}[S(n-2)^*]a_{n-2}\ldots \\ &= a_{n+1}[P(n)]a_nP(n)[S(n-1)^*]a_{n-1}[S(n-2)^*]a_{n-2}\ldots \end{aligned}$$

Then by splicing, V_1 produces $W_1 = id[P(n)]P(n)$. Next let us catalyze W_1 on V_1 as follows:

$$\begin{aligned} V_2 = V_1[W_1] &= a_{n+1}[P(n)]a_nP(n)[S(n-1)^*P(n)]a_{n-1}P(n)[S(n-2)^*]a_{n-2}\ldots \\ &= a_{n+1}[P(n)]a_nP(n)[P(n-1)]a_{n-1}P(n)[S(n-2)^*]a_{n-2}\ldots \end{aligned}$$

Again by splicing, V_2 produces $W_2 = id[P(n)]P(n)[P(n-1)]P(n-1)$. One can iterate this process until obtaining W_n.

6. Convergence and divergence

Let V and V' be two infinitely many reflective m-dna sequences. Given a system of mutations $\mathfrak{S}$. Let us denote by $V_n = a_1(n)|X_1(n)|\dots a_{l(n)}(n)$ ($V'_n = a'_1(n)|X'_1(n)|\dots a_{l(n)'}(n)$) as n-th generations of V (V') w.r.t. $\mathfrak{S}$. We will say that V and V' have *convergent evolution* w.r.t. $\mathfrak{S}$, if there is n_0 so that for all $n \geq n_0$, $l(n) = l(n)'$ and $X_j(n) = X'_j(n)$ for all j.

Let us take two systems of mutations $\mathfrak{S}$ and $\mathfrak{S}'$. Let us denote by V_n and V'_n as n-th generations of V w.r.t. $\mathfrak{S}$ and $\mathfrak{S}'$. We will say that V have *divergent evolution* w.r.t. $\mathfrak{S}$ and $\mathfrak{S}$, if there is n_0 so that for all $n \geq n_0$, $l(n) \neq l(n)'$ or $l(n) = l(n)'$ and $X_j(n) \neq X'_j(n)$ for some j.

Example 6.A: Let $H = H_1 \oplus H_2 \oplus H_3$ where H_i are infinite dimensional Hilbert spaces. Let us put the projections on H_i by P_i and $Q_i = id - P_i$. Let us take a m-dna sequence $V = id[Q_3]Q_3[Q_3]id[Q_2]Q_2$. Suppose V accepts two point mutations as:

$$V \begin{cases} \nearrow V_1 = id[Q_3]Q_3[P_1]Q_2[Q_2]id, \\ \searrow V_2 = id[Q_3]Q_3[Q_3]id[P_2]P_2. \end{cases}$$

Let us take two ensymes $W_1 = id|P_1|P_1$ and $W_2 = id|P_2|P_2$ acting on V_i as follows:

$$V_1[W_1] = id[P_1]P_1[P_1]id[Q_2]Q_2, \quad V_2[W_2] = id[Q_3]Q_3[P_2]P_2[P_2]id,$$
$$V_1[W_2] = id[Q_3]Q_3[0]P_3[Q_2]*, \quad V_2[W_1] = id[P_1]P_1[Q_3]*.$$

Both of $V_i[W_i]$ are m-dna sequences, and both of $V_1[W_2]$ and $V_2[W_1]$ are not.

7. Cancering

Let $V = a_1|X_1|\dots|a_l|X_l|a_{l+1}$ be a m-dna sequence. A *cancering* of V at i is a family of state sets:

$$\{(a_1^n, \dots, a_{l+1}^n)\}_n$$

with the following properties:

(1) each $a_1^n|X_1|a_2^n|\dots a_l^n|X_l|a_{l+1}^n$ is a m-dna sequence,

(2) $tr(a_i^n(a_i^n)^*) \to +\infty$,

(3) there is $M > 0$ with $tr(a_j^n(a_j^n)^*) \leq M$ for $i \neq j$.

A m-dna sequence V is *stable* w.r.t. the mutation system $\mathfrak{S}$, if for any mutation $X_j \to X_j'$ in $\mathfrak{S}$, the corresponding $V' = a_1|X_1|\dots X_{j-1}|a_j|X_j'|a_{j+1}\dots$ admits also a cancering.

V is *unstable*, if it is not stable.

Example 7.A: Let X be of finite type $(2.F)$. Then any primary set $(X_1, \dots, X_l)$ admits a cancering if there is a m-dna sequence $V = a_1|X_1|a_2|\dots a_l|X_l|a_{l+1}$ where the state set $\{a_2, \dots, a_{l+1}\} \subset M_\infty(\mathbf{C})$, $(a_1 = id)$.

Let $\mathfrak{S}$ be consisted by $2.A, 2.B, 2.C, 2.E$. Then N-th stage $\mathfrak{S}^N(V)$ can also admit a concering for all N.

8. Metabolism

Let us take a m-dna sequence $V = a_1|X_1|\dots a_l|X_l|a_{l+1}$, and choose an ensyme $W = b_1|Y_1|b_2|\dots b_k|Y_k|b_{k+1}$, where Y_j catalyzes on X_i. Thus there is a state set $(c_1, \dots, c_{l+1})$ so that:

$$V' = c_1|X_1|\dots c_{i-1}|X_{i-1}|c_i|X_iY_j|c_{i+1}|X_{i+1}\dots c_l|X_l|c_{l+1}$$

consists of a m-dna sequence. We will denote this catalyzing process as:

$$V \mapsto V' = V[W].$$

Let us take two m-dna sequences V and V'. A *random metabolism* M with input V and output V' consists of a family of ensymes $W = \{W_1, \dots, W_l\}$ with the following properties:

(1) one has a diagam of catalizing:

$$V \to V_1 = V[W_1] \to V_2 = V_1[W_2] \to \dots V_l = V_{l-1}[W_l] = V'.$$

Definition 8.A: Let $\{W_i\}_{i=1}^l$ be a set of ensymes. Then this family is said to be *mutually isozyme*, if there is another ensyme w and a mutation system $\mathfrak{S}$ so that any elements in $\{W_i\}_i$ are obtained by mutation in $\mathfrak{S}$ from w.

We will say w is *primitive* for the family $\{W_i\}_i$.

According to *Ureta's hypothesis*, enzymes in a metabolism form complexes which are called as *isozyme complex*.

A *constructive metabolism* is a random metabolism $M = (\{V_i\}_i, \{W_i\}_i)$ satisfying the additional properties:

(2) each W_i and W_j are mutually isozyme,

(3) $W \equiv W_1 \dots W_l$ is infinitely many reflective.

In the following examples, one will only consider random metabolisms.

8.B Tree-like metabolisms: Let $M_0 : V_0 \to V_1$ and $M_1 : V_1 \to V_2$ be two metabolisms. Then the composition of these is another:

$$M_1 \circ M_0 : V_0 \to V_1 \to V_2.$$

Let T be the trivalent graph. Let us equip an orientation on T, and let us denote the set of vertices by v_0, v_1, v_2, where v_0 is the input and the others are output ones. For all vertices v_i, let us assign m-dna sequences $\{V, V', V^1, V^2\}$.

A (random) *metabolism* $M(T)$ with input V_0 and output V_1 and V_2 consists of $\bar{W}^0$, $\bar{W}^1$ and $\bar{W}^2$ where each $\bar{W}^i$ is a family of ensymes $\bar{W}^i = \{W_1^i, \dots, W_{i(l)}^i\}$ with the following properties:

(1) one has a diagam of catalizing:

$$V \to V_1 = V[W_1] \to \dots V_{l_0} = V_{l_0-1}[W_{l_0}] = V' \begin{array}{l} \nearrow V_1^1 = V'[W_1^1] \dots V_{l_1}^1 = V^1, \\ \searrow V_1^2 = V'[W_1^2] \dots V_{l_2}^2 = V^2. \end{array}$$

When in addition, the following two properties are satisfied, then $M(T)$ will be called constructive:

(2) all W_m^i and W_n^j are mutually isozyme.

(3) Let us put $W^i = W_1^i \dots W_{i(l)}^i$. Then $W^0 \cup W^1$ and $W^0 \cup W^2$ are both infinitely many reflective.

By a similar way, one has structure of metabolism $M(T)$ for any finite tree T with a base point.

8.C Mutant metabolism: Let us have a metabolism $M : V_0 \mapsto V_1 \mapsto V_2$. Suppose a mutation occurs as $V_1 \to V_1'$ so that M changes to M' as follows:

$$\begin{array}{ccccc} M : V_0 & \mapsto & V_1 & \mapsto & V_2 \\ & \searrow & & \nearrow & \\ M' : & & V_1' & & \end{array} \tag{8.1}$$

Then $M' : V_0 \mapsto V_1' \mapsto V_2$ will be called as *mutant metabolism.*

8.D Reconstruction of metabolisms: Let $V_0 \to V_1$ and $V_2 \to V_3$ be metabolisms. Another metabolism $V_0 \to V_3$ is called a *reconstruction of metabolism*, if there are no metabolism $V_1 \to V_2$.

Let us consider a relation with selective splicing. Let $V_0 \to V_1V_2$ and $V_3 \to V_4$ be metabolisms. Suppose the following:

(1) there is a metabolism $V_1 \to V_3$, and

(2) there are no metabolism $V_1V_2 \to V_3$.

Then the composit metabolism:

$$V_0 \to V_1V_2 - \text{ selective splicing } \to V_1 \to V_3 \to V_4$$

is said to be an *extended metabolism by self splicing.*

8.E Connection by succesive self splicing: Let $M_0 : U_0 \to U_1$ and $M_1 : U_2 \to U_3$ be two metabolisms. Let us take an ensyme W catalyzing on U_1, $V_2 = U_1[W]$. Then it produces W_2 by splicing. Next let W_2 catalyze on V_2, $V_3 = V_2[W_2]$. Then again it produces W_3 by splicing. On may iterate this process, and finally obtain W_l. If W_l catalyzes on U_2 which switches on the metabolic system M_1, then this connecting process of M_0 with M_1 will be called as a *connection by succesive self splicing* (cf. 5.C). Here is a diagram:

$$\begin{array}{ccccccc} U_0 & \to & & U_1 & & U_2 & \to U_3 \\ & & W & \downarrow & W_l & \uparrow & \\ & & & V_1 & & V_l & \\ & & W_1 & \downarrow & W_{l-1} & \uparrow & \\ & & & V_2 & & V_{l-1} & \\ & & & \cdots & & \cdots & \\ & & & V_i & -W_i \to & V_{i+1} & \end{array} \tag{8.2}$$

When there are two W and W' such that $W_1 = WW'$, $W_2 = WW'W$, $W_3 = WW'WW' = (WW')^2$, $W_4 = (WW')^2W, \ldots$, then the process will be called as a *cyclic connection.*

8.F Join of metabolisms: A particular case of connection is called *join.* Let us take a metabolism system:

$$M_1 : V_0 \mapsto V_1V_2 \mapsto V_3 \mapsto \ldots, \quad M_2 : U_0 \mapsto U_1 \mapsto \ldots$$

Suppose the followings:

(1) an ensyme W catalyzes on V_1V_2 so that the self splicing $V_1V_2 \mapsto V_1$ occurs,
(2) V_1 catalyzes on U_0 with $U_0[V_1] = U_1$.

Then the following metabolism is called *join* of M_1 with M_2:

$$\begin{array}{rcl} V_0 \mapsto V_1V_2 & \mapsto & V_1 \\ & & \downarrow \\ & & U_0 \mapsto \ U_1 \mapsto U_2 \mapsto \ldots \end{array} \tag{8.3}$$

Example 8.F: Let us consider a case when M_1 and M_2 are metabolisms consisted by projection states and shift states respectively. Let us choose:

$$S\begin{cases} v_{2i} \mapsto v_{2i+2} \\ v_{2i+1} \mapsto v_{2i+1} \end{cases}, \quad P_1\begin{cases} v_0 \mapsto v_0 \\ v_{2i} \mapsto 0, \quad i \geq 1 \\ v_{2i+1} \mapsto v_{2i+1} \end{cases}, \quad P_2\begin{cases} v_{2i} \mapsto 0 \\ v_{2i+1} \mapsto v_{2i+1} \end{cases}.$$

Let us take a m-dna sequence V and consider a metabolic system:

$M_1 : V_0 \mapsto id[S]SV \equiv V_1V \mapsto V_2 \ldots$ (shift system),
$M_2 : U_0 \mapsto id[P_1]P_1U \equiv U_1 \mapsto id[P_2]P_2U \mapsto \ldots$ (projective system).

Since $P_1S = P_2$, one may catalyze as $id[P_1]P_1(id[S]S) = id[P_2]P_2$. Thus one can make join of M_1 with M_2.

8.G Maximal connecting metabolism: Let us take two m-dna sequences V_0 and V_1, and consider two metabolisms:

$$M_1 : V_0 \to U_1 \to U_2 \to \ldots U_l \to V_1, \quad M_2 : V_0 \to U_1' \to U_2' \to \ldots U_k' \to V_1.$$

Then if i is the maximal number satisfying that there is a metabolism $M : U_i \to U_i'$, then M is called a *maximal connecting metabolism*:

$$\begin{array}{ccccccccc} & & U_1 & \dots M_1 \dots & & U_i & \dots & U_l & & \\ & \nearrow & & & & & & & \searrow & \\ V_0 & & & & & M\downarrow & & & & V_1 \\ & \searrow & & & & & & & \nearrow & \\ & & U_1' & \dots M_2 \dots & & U_i' & \dots & U_k' & & \end{array} \tag{8.4}$$

Example 8.G: Let $S(n)$ and $P(n)$ be in the example 5.C:

$$S(n) : \begin{cases} v_{2^n k} \to v_{2^{n+1} k} \\ v_i \to v_i, \quad i \neq 2^n k \end{cases}, \quad P(n) : \begin{cases} v_{2^n k} \to 0 \\ v_i \to v_i, \quad i \neq 2^n k \end{cases}.$$

Let us choose two metabolisms M_1 and M_2, where M_1 is a shift system and M_2 is a projective system as:

$$\begin{aligned} M_1 :& id[id]id \mapsto id[S(3m)]S(3m)^* \mapsto id[S(3m-2)]S(3m-2)^* \mapsto \dots \\ & \mapsto id[S(3m-2j)]S(3m-2j)^* \mapsto \dots \\ & \mapsto id[S(m)]S^m \mapsto id[S(m-2)]S(m-2)^* \dots \\ & \mapsto id[S(1)]S(1)^* \mapsto id[P(1)]P(1), \\ M_2 :& id[id]id \mapsto id[P(2m)]P(2m) \mapsto id[P(2m-1)]P(2m-1) \mapsto \dots \\ & \mapsto id[P(2m-j)]P(2m-j) \mapsto \dots \\ & \mapsto id[P(m)]P(m) \mapsto id[P(m-1)]P(m-1) \cdots \mapsto id[P(1)]P(1). \end{aligned}$$

Notice that there is a metabolism from $id[S(i)]S(i)^*$ to $id[P(j)]P(j)$ if and only if $i \geq j$. Thus the maximal connecting metabolism is

$$M : id[S(m)]S(m)^* \mapsto id[P(m)]P(m).$$

8.H Convergence of metabolisms: Let us take two m-dna sequences V_0 and V_1. Suppose there are two metabolisms:

$$\begin{array}{ccccc} & & \ldots M_1 \ldots & & \\ & \nearrow & & \searrow & \\ V_0 & & & & V_1 \\ & \searrow & & \nearrow & \\ & & \ldots M_2 \ldots & & \end{array} \tag{8.5}$$

Let us take two m-dna sequences V' and V'', $V', V'' \neq V_0, V_1$. Suppose there exists another metabolism $M_3 : V' \to V''$:

$$\begin{array}{ccccccc} M_1 & : & V_0 & \to & V' & \to & V_1 \\ & & & & M_3 \downarrow & & \\ M_2 & : & V_0 & \to & V'' & \to & V_1. \end{array} \tag{8.6}$$

Then we will say that M_1 and M_2 are *connected*. If M_1 and M_2 are non connected, then we will say that these metabolisms are *convergent systems*.

Example 8.H: Let us take two operators X_1, X_2 such that there are no A and A' with $X_1 A = X_2$ and $X_2 A' = X_1$. If one can choose $V' = id[X_1]X_1^*$ and $V'' = id[X_2]X_2^*$, then M_1 and M_2 are convergent systems.

Let P_0, P_1, P_2, P_3, P_4 be projections satisfying:

$$P_0 P_3 = P_3, \quad P_1 P_3 = P_1, \quad P_1 P_2 = P_2, \quad P_2 P_4 = P_4.$$

Let us consider metabolisms:

$$M_1 : V_0 = id[P_0]P_0[P_2]P_2 \mapsto V' = id[P_1]P_1[P_2]P_2 \mapsto V_1 = id[P_1]P_1[P_4]P_4,$$
$$M_2 : V_0 = id[P_0]P_0[P_2]P_2 \mapsto V'' = id[P_0]P_0[P_4]P_4 \mapsto V_1 = id[P_1]P_1[P_4]P_4.$$

Then there are no metabolism as $V' \to V''$.

In practice there are some pairs (W, W') of ensymes such that W and W' play the same role as functions, however many of structures of genes (from primary to fourth) are very different mutually. Such pair is said to be obtained by convergent evolution. For example kimo-tripsin and subtilisin consists such pair.

8.I Divergence of metabolisms: Let us consider two metabolisms $M_1 : V_0 \to V'$ and $M_2 : V_0 \to V''$:

$$\begin{array}{ccccc} & & & & V' \\ & & & M_1 & \nearrow \\ V_0 & \to & V_1 & \to & \\ & & & M_2 & \searrow \\ & & & & V''. \end{array} \tag{8.7}$$

If there are no invertible metabolisms as below:

$$V' \to V'', \quad V'' \to V'$$

then we will say that M_1 and M_2 are *divergent systems*.

Example 8.I: Let $S(n)$ and $P(n)$ be in the example 5.C. Let us consider two metabolisms:

$$M_1 : id[S(n+2)]S(n+2)^* \mapsto id[S(n+1)]S(n+1)^* \mapsto id[S(n)]S(n)^*,$$
$$M_2 : id[P(n+2)]P(n+2) \mapsto id[P(n+1)]P(n+1) \mapsto id[P(n)]P(n).$$

The ensyme $id[P(n)]P(n)$ can catalyze on $id[S(n)]S(n)$ which produces a metabolism $id[S(n)]S(n)^* \mapsto id[P(n)]P(n)$. However there are no metabolism from $id[P(n)]P(n)$ to $id[S(n)]S(n)^*$.

In practice a pair of two ensymes is said to be obtained by divergent evolution if they arose from the same original one, and have been developped very differently.

8.J Cyclic metabolism: Let us take m-dna sequences V, V' such that both of VV' and $V'V$ are also m-dna sequences. Suppose there is a family of metabolisms $\{M_n\}_{n\geq 0}$:

$$M_0 : V \to VV', \quad M_1 : VV' \mapsto VV'V,$$
$$M_2 : VV'V \mapsto VV'VV', \quad M_3 : VV'VV' \mapsto VV'VV'V, \quad \ldots$$

Thus inductively $M_{2n} \circ \cdots \circ M_0$ produces $(VV')^n \equiv VV' \ldots VV'$ and $M_{2n+1} \circ \cdots \circ M_0$ produces $(VV')^n V$ as metabolism productions. We will say that the family $\{M_n, W_n\}_n$ is a *cyclic metabolism system*, where W_n are ensymes for M_n.

We will denote pictorically as:

$$\begin{array}{ccccc} & & \{W_{2n}\}_n & & \\ & \nearrow & & \searrow & \\ V & & & & V' \\ & \nwarrow & & \swarrow & \\ & & \{W'_{2n+1}\}_n & & \end{array} \tag{8.8}$$

We will say that the metabolic system $M = \{M_i\}_i$ is a cyclic metabolism from V to $V'' = (VV')^n$ (or to $(VV')^nV$).

Example 8.J: Here we have three types of metabolic systems:

(1) Let us take a long m-dna sequence $\tilde{V}$ which contains many copies of V and V'. When there is a family of ensymes $\{W_n\}_n$ acting on $\tilde{V}$ such that for each n, W_n catalyzes on $\tilde{V}$ which helps it to make *self splicing*.

(2) Let $\tilde{V}$ be as above. Suppose there is an ensyme W which catalyzes on $\tilde{V}$ and they produce by self splicing, VV'. Next suppose $W_2 \equiv VV'$ catalyzes on $\tilde{V}$ and they produce by self splicing, $VV'V$. Succesively $VV'V$ catalyzes on $\tilde{V}$ producing $VV'VV'$ by self splicing. One may iterate this process. This metabolism will be called as a *self production of cyclic metabolism*.

(3) Periodic inequally crossing over.

8.K Minimal metabolism for extension: Let us take a metabolism family:

$$M_0 : V_0 \to V_1, \quad M_1 : V_1 \to V_2, \ldots, M_n : V_n \to V_{n+1}, \ldots$$

Let us choose a particular m-dna sequence U. Then if i is the minimal number satisfying that there is a metabolism $M : V_0 \to U \to V_{i+1}$, then we will say that the composition $M_i \circ \cdots \circ M_0$ is called a *minimal metabolism for extension* w.r.t. $(M_i; U)$ and i is the minimal number of extension:

$$\begin{array}{ccccccccc} & & \ldots & \to & \ldots U \ldots & & & & \\ M & \nearrow & & & & & \searrow & & \\ V_0 & \to & V_1 & \ldots & & V_i & \to & V_{i+1} \to \ldots V_n \end{array} \tag{8.9}$$

Example 8.K: Let $S(m)$ and $P(m)$ be in 5.C and choose $U = id[S(m)]S(m)^*$,

$m \leq n$. Let us choose metabolisms:

$$M_0 : V_0 = id[id]id \mapsto V_1 = id[P(n)]P(n),$$
$$M_l : V_l = id[P(n-l+1)]P(n-l+1) \mapsto V_{l+1}, \quad 1 \leq l \leq n.$$

Let us find $M : id[id]id \mapsto id[S(m)]S(m)^* \mapsto id[P(m')]P(m')$. Then M exists only when $m' \leq m$. Thus the minimal number of extension is $n - m$.

8.L Locally amplification of metabolic systems: Let us consider a family of metabolisms $M_0 : V_0 \to V_1, \quad M_1 : V_1 \to V_2, \ldots$ If there is another family of metabolisms $M : V_i \to U_1 \ldots U_l \to V_{i+1}$, then the following extended system of metabolisms will be called as a *locally amplification of metabolic systems*:

$$\begin{array}{ccccccc} V_0 & \to & \ldots & V_i & \to & V_{i+1} & \to \ldots \\ & & & \downarrow & & \uparrow & \\ & & & U_1 & & U_l & \\ & & & \downarrow & M & \uparrow & \\ & & & U_2 & & U_{l-1} & \\ & & & \ldots & & \ldots & \\ & & & U_i & \to & U_{i+1} & \end{array} \tag{8.10}$$

If M is a cyclic metabolism from V_i to V_{i+1}, then we will say that the above system will be called as a *locally cyclic amplification of metabolic systems.*

Example 8.L: Let us take another m-dna sequences V_i' and V_{i+1}' so that there is a cyclic metabolism $V_i' \to V_{i+1}'$. Now suppose a mutation occurs $V_i \to V_i'$ so that a mutant metabolism M' is produced:

$$\begin{array}{lccccccc} M : V_0 & \mapsto \ldots & V_{i-1} & \mapsto & V_i & \mapsto & V_{i+1} & \mapsto \ldots \\ & & & \searrow & & \nearrow & & \\ M' : & & & & V_i' & & & \end{array} \tag{8.11}$$

If M' makes locally cyclic amplification of metabolic systems as below, then it will be called as *locally cyclic amplification of mutant metabolic systems*:

$$\begin{array}{ccccc} V_{i+1} & \leftarrow & \ldots & (V_i'V_{i+1}')^l & \ldots \\ \uparrow & \nwarrow & & & \\ V_{i-1} & \mapsto & V_i' & \mapsto V_i'V_{i+1}' \mapsto \ldots & \end{array} \tag{8.12}$$

8.M Enlargement: Let $M_0 : V_0 \to V_1 \to V_2 \to \dots$ and $M_1 : V_0^1 \to V_1^1 \to \dots$ be systems of metabolisms. Let us choose i. An *enlargement* of M_0 and M_1 at i is catalyzing of V_i^1 on V_i as an enzyme.

Suppose one has a family of metabolic systems $\{M_i : V_0^i \to V_1^i \to \dots\}_i$. An enlargement of the family at i, is succesively catalyzing of V_i^l on V_i^{i+1} as an ensyme. Thus one has a diagram:

$$\begin{array}{ccccccc} M_0 : V_0 \to V_1 \to V_2 & \to & \dots & \to & V_i & \to V_{i+1} \dots & \\ & & & \nearrow & & & \\ M_1 : V_0^1 \to V_1^1 \dots & \to & V_i^1 & \to & V_{i+1}^1 & \dots & \\ & \nearrow & & & & & \\ M_2 : V_0^2 \to V_1^2 \cdots \to V_i^2 & \to & V_{i+1}^2 & \dots & & & \\ \dots & & & & & & \end{array} \tag{8.13}$$

8.N Branching of a metabolism: Let us consider a metabolism $M : V_0 \mapsto V_1 \mapsto V_2 \dots$ Suppose V_1 catalyzes on V_0 which produces another U_1. Then one gets the diagram:

$$\begin{array}{cccccc} & V_0 & \mapsto & V_1 & \mapsto V_2 \dots & \\ M : & & \searrow & & & \\ & & & U_1 & & \end{array} \tag{8.14}$$

Succesively let V_2 calatyze on U_1 which produces U_2. Inductively one considers U_i which is the production of V_i catalyzing on U_{i-1}:

$$\begin{array}{cccccc} V_0 & \mapsto & V_1 & \mapsto V_2 \dots & & \\ & \searrow & & & & \\ & & U_1 & \mapsto U_2 \mapsto U_3 \dots & & \end{array} \tag{8.15}$$

Then we will say that the above new metabolism is *the branched metabolism.*

8.O Reconstruction of metabolic systems: Let:

$$\bar{M} = \{(M_1, \dots, M_l), (\{V_{i,j}\}_{i,j \in I})\}$$

be a metabolic system where $V_{i,j}$ is a join of M_i with M_j. If there is another system $\bar{M}' = \{(M_1, \dots, M_l), (\{V'_{i,j}\}_{i,j \in I})\}$, then we will say that $\bar{M}'$ is a reconstruction of $\bar{M}$.

8.P Homeobox: Let P be a m-dna sequence which plays a role of a promotor. Let us take another family of m-dna sequences $\{V_0, V_1, \ldots, V_l\}$ where all PV_j are also m-dna sequences.

Suppose there is a metabolism:

$$M : PV_0 \to PV_1 \to \ldots PV_l.$$

If there is a mutation $P \to P'$ which consists of another metabolism:

$$M' : P'V_{i_0} \to P'V_{i_1} \to \ldots P'V_{i_l}$$

then we will say that P is a *homeobox*.

Let us take $i < j$. If M' is of the form below, then the mutation will be called *simple* :

$$\begin{array}{lllll} M : PV_0 \to \ldots PV_{i-1} \to & PV_i & \to PV_{i+1} \to & \ldots & \\ & \Updownarrow & & & \\ M' : P'V_0 \to \ldots P'V_{i-1} \to & P'V_j & \to P'V_{i+1} \to & \ldots & \end{array}$$

$$\begin{array}{llll} \ldots & PV_{j-1} \to & PV_j & \to PV_{j+1} \ldots \\ & & \Updownarrow & \\ \ldots & P'V_{j-1} & \to P'V_i & \to P'V_{j+1} \ldots \end{array}$$

8.Q Evolution of metabolisms: Let M be a metabolism. When M makes development under a mutation system $\mathfrak{S}$, then one will get a family of metabolisms. This will be called as a evolution of M under $\mathfrak{S}$.

9. Conservative system in metabolisms

Here one introduces an assumption on the *conservation of the total production by metabolisms*. Let $\bar{M} = \{M_1, \ldots, M_n\}$ be a system of metabolisms. Suppose a mutation occurs on this system, and it changes into $\bar{M}' = \{M'_1, \ldots, M'_{n'}\}$. Then the assumption saids that if $\bar{M}'$ produces new m-dna sequences, then it should be consumed as ensymes inside $\bar{M}'$.

Let us consider by an example. Suppose a situation; a metabolic system makes an amplification by mutation which produces a new ensyme W'. Then

it can be consumed to make a cyclic system. Let us take a cyclic system:

$$\begin{array}{ccccc} & & \{W_{2n}\}_n & & \\ & \nearrow & & \searrow & \\ V_i & & & & V_{i+1} \\ & \nwarrow & & \swarrow & \\ & & \{W_{2n+1}\}_n & & \end{array} \tag{9.1}$$

Then we consider a metabolic system $\bar{M} = \{M_1, M_2\}$:

$$M_0 : V_0 \mapsto \dots V_i \mapsto V_{i+1} \mapsto \dots V_n, \quad M_1 : U \mapsto \dots U'.$$

Suppose a connection between M_0 and M_1 occurs as $V_n \mapsto U$, by a help of an ensyme W. Suppose U' catalyzes on V_i which makes a loc. amp. of metabolism on $V_i \mapsto V_{i+1}$. Thus one will have the following diagram:

$$\begin{array}{lccccccc} M_0 : & V_0 \mapsto \dots & V_i & \mapsto & V_{i+1} & \mapsto \dots & \dots & V_n \\ & & \downarrow & & \uparrow & & & \\ & & V_i' V_{i+1}' & & \dots & & & \downarrow W \\ & & \downarrow & & \uparrow & & & \\ & & \dots & \mapsto & (V_i' V_{i+1}')^l & U' \leftarrow & U & : M_1 \end{array} \tag{9.2}$$

Thus in order to consume U', this metabolic system has produced a locally amplification.

10. Glycolysis

In this section, one will make a partial construction of a metabolic system, G-system, modelled on a glycolysis (Embden-Meyerhof-Parnas pathway). Let us take two m-dna sequences Gc and P_2, and two ensymes PK and PP. Then G-system is a metabolic system beginning from Gc and ending by producing P_2 where one uses PK, PP and their variations as ensyme families.

There are several steps and here we will concentrate on the following steps:

(1) Gc (glucose) $\to Gl$ (glycogen),
(2) F_1 (fructose-6-phosphate) $\to F_3$ (fructose-2,6-2-phosphate),
(3) P_1 (phospho-enol-pilbin) $\to P_2$ (pilbin).

During the above steps, several ensymes are interacting mutually. At every

step, we will use two ensymes, PK (proteinkinase) and PP (phospho-protein-phosphatase). These have some different activation states respectively.

10.A From glucose to glycogen: Let us begin from (1). An ensyme P (phospholirase) activates on Gc and produces Gl:

$$Gc - [P] \longrightarrow Gl.$$

P has three, from high to low (in reality four) activation states:

$$\text{(high) } P \longleftrightarrow P'' \longleftrightarrow P' \text{ (low).}$$

As ensymes, PP, PK and Gl act on P as the diagram below:

$$\begin{array}{ccccc} & & P' & & \\ & [PP]\nearrow & & \searrow[PK] & \\ P'' & \leftarrow & [Gl] & - & P \end{array}$$

Thus PK makes P active, and Gl do it actless.

In fact the metabolism $Gc \to Gl$ is inversible. There is another ensyme St (glycogensyntase) catalyzing as below:

$$\begin{array}{ccc} & [P] & \to \\ Gc & \longleftrightarrow & Gl \\ \leftarrow & [St] & \end{array}$$

St has also three activating states; St and St'' are both active, and S' is actless:

$$\begin{array}{ccc} & St'' & \\ \nearrow & [Gl] & \\ St' & \longleftrightarrow & St \end{array} \quad , \quad \begin{array}{ccccc} & & \leftarrow & [PK] & \\ St' & & \longleftrightarrow & & St \\ & & [PP] & \to & \end{array} \quad .$$

Example 10.A: Let H be an infinite dimensional Hilbert space, and take basis $\{\dots, v_{-l}, \dots, v_0, v_1, \dots\}$. Let S be a shift, $S : v_i \to v_{i+1}$. S is invertible. Let us begin with:

$$Gc = id[S^*]S \mapsto Gl = id[id]id, \quad PP = id[S]S^*, PK = id[S^*]S.$$

Then one deteremines P and St as:

$$P = id[S]S^*, \quad P' = S[S^2]S^*, \quad P'' = S[S]id,$$
$$St = id[S^*]S, \quad St' = St'' = id[S^{2*}]S^2.$$

Let us put this metabolism by $M_1(S)'$. In this example, one may replace S by $S^* = S^{-1}$. Thus $M_1(S^*)'$ gives another one. Let us put $H' = H \oplus H$. Then by identifying H' with H, the direct sum $T = S \oplus S^*$ also gives a metabolism. The direct sum of metabolisms will be expressed as:

$$M_1(S)' \oplus M_1(S^*)' = M_1(S \oplus S^*)'.$$

10.B Activating states on PP: There is an ensyme PI (phosphatase-inhibitar) which makes PP actless. This also has two states, high and low:

$$(\text{high})\ PP, PI \longleftrightarrow PP', PI'\ (\text{low})\ .$$

These are mutually interacting and have the following metabolic systems:

$$\begin{array}{ccc} & \leftarrow [PK] & \\ PP & \longleftrightarrow & PP' \\ & [PI] \rightarrow & \end{array} \quad , \quad \begin{array}{ccc} & \leftarrow [PK] & \\ PI & \longleftrightarrow & PI' \\ & [PP] \rightarrow & \end{array} \quad .$$

Let us consider the above metabolic system. It expresses states of PP and PI catalyzed by ensymes PK and PP. Combinig with these, one has a diagram:

$$\begin{array}{ccc} & \leftarrow PK & \\ PP & \longleftrightarrow & PP' \\ & PI \rightarrow & \end{array}$$

$$\begin{array}{ccc} PP & \downarrow\uparrow & PK \\ & PI' & \end{array}$$

Let us observe the following; for example, suppose this system begins with PP' and PI', and supply PK as an ensyme. Then the states change into PP and PI. Then PP changes into PP' catalyzed by PI. Also PI changes into PI' catalyzed by PP. Thus it comes back to the intial state in this system:

$$PK : (PP', PI') \rightarrow (PP, PI) \rightarrow (PP', PI').$$

When one changes the initial states as $(PP(')), PI(')))$, then there are similar occurence.

Definition 10.B: Let us consider a metabolic system M. If an ensyme $W(= PK)$ satisfies the above property, then it is called *stable*.

Example 10.B: Let us continue the above example. The following extension gives an example of metabolism from glucose to glycogen:

$$PI = S[S]id, \quad PI' = S[S^2]S^*, \quad PP' = S[S^2]S^*.$$

We will denote this as $M_1(S)$.

Again one may also replace S by $T = S \oplus S^*$, and get $M_1(T) = M_1(S \oplus S^*)$.

10.C Fluctose metabolic system: There are several ensymes:

$$\begin{aligned} &PFK_i \quad \text{(phospho-fluctokinase)} , \\ &FBP_i \quad \text{(fluctose-bisphosphatase)} , \end{aligned}$$

$i = 1, 2$ with the diagram:

$$\begin{array}{ccccccc} & \leftarrow & [PFK_1] & & & \leftarrow & [PFK_2] \\ F_2 & & \longleftrightarrow & F_1 \quad , \quad F_3 & & \longleftrightarrow & F_1 \quad . \\ & & [FBP_1] \quad \rightarrow & & & [FBP_2] \quad \rightarrow & \end{array}$$

Both of PFK_i and FBP_i have high and low states mutually. We will denote low state by adding dash, PFK_i' and FBP_i'. The fluctose metabolic system has a mutual interaction as:

$$\begin{array}{ccccc} & & \leftarrow & [F_1], [F_3] & \\ \text{(high) } PFK_1 & & & \longleftrightarrow & PFK_1' \text{ (low)} \quad , \\ & & & [C] \quad \rightarrow & \\ & & \leftarrow & & \\ \text{(high) } FBP_1 & & & \longleftrightarrow & FBP_1' \text{ (low)} \\ & & & [F_3] \quad \rightarrow & \end{array}$$

where C is Cuen-acid which comes from the Cuen cycle after glucose one. Here we will not take into account of the effect of C.

10.D Some properties of PFK and FBP: PFK_2 and FBP_2 have isozymes

respectively, and in order to distinguish them, we will express the isozyme pairs by:

$$(LPFK_2, HPFK_2), \quad (LFBP_2, HFBP_2).$$

PK and PP catalyze on both of L and H types which are inverse way as:

$$\begin{array}{ccc} & \leftarrow & [PP] \\ LPFK_2 & & \longleftrightarrow \\ & & [PK] \ \rightarrow \end{array} \quad LFBP_2 \quad , \quad \begin{array}{ccc} & \leftarrow & [PK] \\ HPFK_2 & & \longleftrightarrow \\ & & [PP] \ \rightarrow \end{array} \quad HFBP_2 \quad .$$

Let us take two m-dna sequences V, V' and an ensyme S catalyzing as below:

$$\begin{array}{ccc} & V & \\ S & \downarrow\uparrow & S \\ & V' & \end{array}$$

The above type of S will be called a *partial involution*.

Let us take pairs of isozymes (V, U) and (V', U'). Suppose an ensyme S catalyzes as below:

$$\begin{array}{ccccc} & V & \longleftrightarrow & U & \\ S & \downarrow & & \uparrow & S \\ & V' & \longleftrightarrow & U' & \end{array}$$

Then this type of S will be called as *quasi partial involution*.

Example 10.D:

(1) *Involution*: Let us decompose $H = H_1 \oplus H_2$, where $H_1 = \{v_{2i+1}\}_i$ and $H_2 = \{v_{2i}\}_i$. Then one puts:

$$I_0 : v_{2i} \rightarrow v_{2i+1}, \quad v_{2i+1} \rightarrow v_{2i}.$$

I_0 is an involution.

(2) *Partial involution*: H has another splitting $H = H_1' \oplus H_2' \oplus H_3'$ where H_1' is spanned by v_{4i}, H_2' by v_{4i+2} and H_3' by v_{2i+1}. Then by a similar way, there is an involution $I_0' : H_1' \cong H_2'$. Let us put:

$$I|H_1' \oplus H_2' \equiv I_0', \quad I|H_3' \equiv 0.$$

Then I is a partial involution.

(3) *Quasi partial involution*: Let I_0 be as above, and define another shift T as:

$$T : v_{2i} \to v_{2i+2}, \quad v_{2i+1} \to v_{2i-1}.$$

Then one has the following commutative diagram:

$$\begin{array}{ccccc} & & I_0^{\pm 1} & & \\ & id & \mapsto & I_0^{\pm 1} & \\ T & \downarrow & & \uparrow & T \\ & T & \mapsto & TI_0^{\pm 1} & \\ & & I_0^{\pm 1} & & \end{array}$$

I_0 is a quasi partial involution between $(id, I_0^{\pm 1})$ and $(T, TI_0^{\pm 1})$.

Now using the above example, one gets an example of fluctose metabolic system below. It will be an extension of $M_1(T)$. Let us choose:

$$\begin{aligned} &PP = id[T]T^*, \quad PK = id[T^*]T, \\ &(LFBP_2, HFBP_2) = (id[id]id, id[I_0]I_0^*), \\ &(LPFK_2, HPFK_2) = (id[T]T^*, id[TI_0]I_0^*T^*). \end{aligned}$$

This extended example of fluctose metabolism will be denoted as $M_2(T, I_0)'$.

Let S be the shift $v_i \to v_{i+1}$ as before, and have a metabolism $S^*(M_2(T, I_0)')$ by:

$$\begin{aligned} &PP = id[T]T^*, \quad PK = id[T^*]T, \\ &(LFBP_2, HFBP_2) = (id[S]S^*, id[SI_0]I_0^*S^*), \\ &(LPFK_2, HPFK_2) = (id[ST]T^*, id[STI_0]I_0^*T^*S^*). \end{aligned}$$

Let us consider interaction between F_1 and F_3 above. Using the above example, one will obtain an example of F_1 and F_3 catalyzed by $(LFBP_2, LPFK_2)$. Recall the decomposition $H = H_1 \oplus H_2$, $H_1 = \{v_{2i+1}\}_i$ and $H_2 = \{v_{2i}\}_i$. Let us denote by P_i, $i = 1, 2$ as the projections to H_i. Let us put:

$$F_1 = id[P_2]P_2, \quad F_3 = id[S^*P_1]P_1S.$$

Then from the equations:

$$P_2ST = P_2STP_1 = S^*P_1, \quad S^*P_1S = P_2$$

this gives an extension $M_2(T, I_0, S)$ of the metabolism $S^*(M_2(T, I_0)')$. Thus one obtains a system of metabolisms $S^*(M_2(T, I_0)') \subset M_2(T, I_0, S)$.

10.E Pilbin metabolic system: There is an ensyme PPK (pilbinkinase) which catalyzes from P_1 (phospho-enol-pilbin) to P_2 (pilbin). PPK has two states, activating state PPK and actless state PPK'. These have the following system:

$$P_1 \overset{[PPK]}{\longrightarrow} P_2 \quad , \qquad \text{(high)} \quad PPK \overset{\substack{\leftarrow \\ [F_2],[PP] \\ \longleftrightarrow \\ [PK] \\ \rightarrow}}{} PPK' \quad \text{(low)} \quad .$$

Example 10.E: Let us consider $M_2(T, I_0, S)$. Then one obtains an extension $M_3(T, I_0, S, S^2)$ by letting:

$$PPK = id[P_2]P_2, \quad PPK' = id[P_2T^*]TP_2.$$

11. Loop structure

Let $V = WV_1W'V_2W''$ be a m-dna sequence with a structure:

$$\begin{aligned} V_1 &= a_1|X_1|a_2 \dots a_l|X_l|a_{l+1}, \\ V_2 &= a_{l'+1}|X_{l'+1}|a_{l'+2}| \dots |a_{l+l'}|X_{l+l'}|a_{l+l'+1} \\ &\equiv b_{l+1}|Y_l|b_l| \dots b_2|Y_1|b_1. \end{aligned}$$

Suppose there exists a family of primary elements $\{Z_1, \dots, Z_l\}$ so that the shift relation is satisfied, $a_i b_i^* = Z_i$, $i = 1, \dots, l$. Then the structure:

$$\begin{array}{ccccccccccc} a_1 & |X_1| & a_2 & |X_2| & a_3 & |\dots & a_l & |X_l| & a_{l+1} & \dots \\ - & & - & & - & & - & & - & \\ Z_1 & & Z_2 & & Z_3 & & Z_l & & Z_{l+1} & \\ - & & - & & - & & - & & - & \\ b_1 & |Y_1| & b_2 & |Y_2| & b_3 & |\dots & b_l & |Y_l| & b_{l+1} & \dots \end{array} \tag{11.1}$$

is called *loop structure.*

11.A Parindrome: Let $V = a_1|X_1| \dots a_{2l-1}|X_{2l-1}|a_{2l}$ be a m-dna sequence

with parindrome. Let us rewrite V as $a_1|X_1|\dots a_l|X_l|a_{l+1}|X_l|\dots|X_1^*|a_1$. Then one obtains a loop structure by:

$$\begin{array}{cccccccccc} a_1| & X_1 & |a_2| & X_2 & |a_3| & \dots & |a_j| & X_j & |a_{j+1}| & \dots \\ & X_1^* & & X_2^* & \dots & X_{j-1}^* & & X_j^* & & \\ a_1| & X_1 & |a_2| & \dots & |a_{j-1}| & X_{j-1} & |a_j| & \dots & & \end{array} \tag{11.2}$$

References

[ABLRRW] B.Alberts, D.Bray, J.Lewis, M.Raff, K.Roberts and J.D.Watson, *Molecular biology of the cell. 2nd ed.*, Garaland Publ. Inc. New York and London, (1989).

[CG] A. Carbone and M. Gromov, *Mathematical slices of molecular biology*, IHES preprint, (2001).

[CGP] Ed. A.Carbone, M.Gromov and P.Prusinkiewicz, *Pattern formation in biology, vision and dynamics*, World Scientific (2000).

[H] K.Hori, *Isozyme no bunshi-seibutsu gaku*, Kyouritsu-shuppan (in Japanese), 1994.

[I] *Initiation la gnomique fonctionnelle : Aspects biologiques, mathmatiques et algorithmiques*, seminar guide in IHES internet, www.ihes.fr.

[K] B.P.Kitchens, *Symbolic dynamics*, Springer, (1998).

[MM] C.L.Markert and F.Mϕller, *Multiple forms of ensymes:Tissues, ontogenetic, and species patterns*, Proc. Nat. Acad. Sci. USA, 45, 753-763.

[Mu] *Museum National D'Histoire Naturelles, Galeries de Palentologie et Anatomie compar'ee*, Jardan des plants, in Paris.

[MO] Ed. C.L.Markert and Z.Ogita, *ISOZYMES: Structure, Function, and Use in Biology and Medicine*, Progr. Clin.Biol.Res., Vol 344, Wiley-Liss, New York, (1990).

[U] T.Ureta, *The role of isozymes in metabolism: A model of metabolic pathways as the basis for the biological role of isozymes*, Curr. Top. Cell. Regul., 13, 233 - 258.

Tsuyoshi Kato
Department of Mathematics
Faculty of Science
Kyoto University
Kyoto 606-01 Japan