

カルモデュリン蛋白質の分子進化からみた生命の進化

飯 田 陽 一*

I. はじめに

生命の進化は生物学の最も重要な課題の1つであり、現存生物の形態学的研究や古生物学的な化石の調査によって生物進化の系統関係を研究した例は本誌にも数多く報告されている。一方、最近の分子生物学の進歩によって、生命を構成している蛋白質のアミノ酸配列やDNAのヌクレオチド配列を分子レベルでかなり詳細に知ることができるようになり、それらの情報をもとに生物の系統関係を議論する研究分野が急速に発展しつつある。この報告は、カルモデュリン蛋白質の分子進化をとりあげて、そのような分野の研究の一端を紹介することを目的としている。現存生物がもっている蛋白質や遺伝子の構造に関するデータは、生命の長い進化の歴史の結果として蓄積された貴重な情報源と考えられる。そのような「生きた化石」の情報を解析することは、従来の古生物学的な化石研究にも有益な知見をもたらすのではないかと期待される。

II. カルモデュリン蛋白質のアミノ酸配列データと分子進化

いろいろな生物種で特定の蛋白質の一次構造（アミノ酸配列）データを比較して、どのアミノ酸座位が変異しているかを解析することによりそれらの生物種の系統関係を議論する方法はZuckerkandl & Pauling (1962) によってはじめられ、その後Dayhoff (1972) やその他の研究者による報告が数多くなされている。特に詳細に調べられた例はグロビンやチトクロームC等の蛋白質の分子進化であり、グロビンについてはわが国でも武居・松田 (1980) による詳しい解説がなされている。そのような研究の結果明らかになった事実の1つとして、蛋白質の種類が違くとアミノ酸の置換速度はいちじるしく異なっていることがあげられる。蛋白質における各アミノ酸座位当り1年間の置換率(k_{aa})を 1×10^{-9} 単位(1ポージング)で表わすことにする。最も置換率の大きい蛋白質としてフィブリノペプチドは $k_{aa}=9.0$ ポージングの値をもつ。一方、置換率が最小で進化に対して最も保守的と考えられる

蛋白質としてヒストンH4 ($k_{aa}=0.006$ ポージング)が知られている(武居・松田, 1980)。

この報告で述べるカルモデュリン(以下CaMと略す)は、ヒストンH4に較べると進化速度はやや大きい、それと同程度に進化に対して保守的な蛋白質である。CaMの重要な性質の1つとして、それが広く真核生物(動物、植物、カビ類)に分布しており、生物が「動く」メカニズムの基本的な制御をおこなっていることがあげられる(Wasserman et al., 1977)。CaMはカルシウム結合蛋白質の1つで、1蛋白質分子当り4個の Ca^{2+} イオンを結合する。 Ca^{2+} イオンを結合しない時のCaM分子のコンフォメーション(立体構造)と結合した時のそれが顕著に異なっているようであり、この立体構造の相違が隣接する他の酵素系を活性化する仕組みになっている。つまり細胞中では Ca^{2+} イオンが情報伝達物質になっていて、その濃度の大小によってCaMがそれと隣接している酵素系を制御していることになる。驚くべきことにCaMが制御している酵素系は1つではなく、図1に示すように数多く依存し(CaMの多機能性)、それ以外にもまだ知られていな

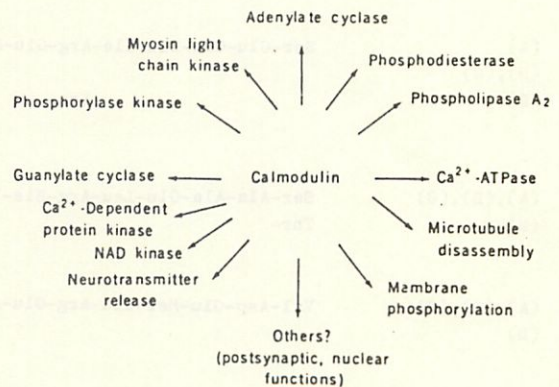


図1. 細胞内においてカルモデュリンは種々の酵素系を制御している。ここにあげた系以外にまだ多くの酵素系がカルモデュリンによって制御されているようである(Cheung, 1980)。

いがCaMが制御する酵素系は多数存在するようである(Cheung, 1980)。このような状況のもとではCaMの一次構造は非常に保守的にならざるを得ない。何故なら、もしCaMのアミノ酸座位のどれかの残基が置換するとすればそれに影響される隣接酵素系があるはずで、この変異によってその酵素系が機能を失うとなればその生物の死活にかかわることとなる。つまりCaMはそれと連動する数多くの酵素系によってとり囲まれ「がんじがらめ」にされていて、CaMのアミノ酸残基が変異しないように厳しく監視されていることになる。

このような推察をうらづけるために、図2にいくつかの系統学的に非常にはなれた動物のCaMのアミノ酸配列データを比較した結果を示す。CaMは148個のアミノ酸残基からなる蛋白質であるが、上記の理由で

生物学的系統関係が非常にはなれていてもほとんど同じアミノ酸配列をもっていることは注目に値する。例えば、哺乳類(ウシの脳)と腔腸動物(イソギンチャク)のCaMを比較しても、148個のアミノ酸残基のうちわずか3個しか置換を起していない。また、すべての脊椎動物のCaMはほぼ同一のアミノ酸配列をもっているようである。図3は(A)ウシの脳(bovine brain), (B)ホタテガイ(*Patinopecten yessoensis*), (C)イソギンチャク(*Metridium senile*)および(D)テトラヒメナ(*Tetrahymena pyriformis*)について、共通祖先法を用いてCaMの系統樹を作成した結果を示す(Iida, 1982a)。この図の作成には最大節約法を適用し、アミノ酸残基の置換が「遺伝コード」のコードンの最小ヌクレオチド置換数によって与えられると仮定して、系統樹の枝の長さをそのような数値で表わしてい

	1	10	20
(A),(B),(C),(D)	Ac-Ala-Asp-Gln-Leu-Thr-Glu-Glu-Gln-Ile-Ala-Glu-Phe-Lys-Glu-Ala-Phe-Ser-Leu-Phe-Asp-		
		30	40
(A),(B),(C),(D)	Lys-Asp-Gly-Asp-Gly-Thr-Ile-Thr-Thr-Lys-Glu-Leu-Gly-Thr-Val-Met-Arg-Ser-Leu-Gly-		
		50	60
(A),(B),(C),(D)	Gln-Asn-Pro-Thr-Glu-Ala-Glu-Leu-Gln-Asp-Met-Ile-Asn-Glu-Val-Asp-Ala-Asp-Gly-Asp-		
		70	80
(A),(B),(C)	Gly-Thr-Ile-Asp-Phe-Pro-Glu-Phe-Leu-Thr-Met-Met-Ala-Arg-Lys-Met-Lys-Asp-Thr-Asp-		
(D)	-Ser-Leu-		
		90	100
(A)	Ser-Glu-Glu-Glu-Ile-Arg-Glu-Ala-Phe-Arg-Val-Phe-Asp-Lys-Asp-Gly-Asp-Gly-Tyr-Ile-		
(B),(C)	-Phe-		
(D)	-Leu-	-Lys-	-Arg-
		110	120
(A),(B),(C)	Ser-Ala-Ala-Glu-Leu-Arg-His-Val-Met-Thr-Asn-Leu-Gly-Glu-TML-Leu-Thr-Asp-Glu-Glu-		
(D)	Thr-		
		130	140
(A),(B),(C)	Val-Asp-Glu-Met-Ile-Arg-Glu-Ala-Asp-Ile-Asp-Gly-Asp-Gly-Gln-Val-Asn-Tyr-Glu-Glu-		
(D)	-His-Ile-		
		148	
(A)	Phe-Val-Gln-Met-Met-Thr-Ala-Lys		
(B)	-Thr-	-Ser-	
(C)	-Lys-	-Ser-	
(D)	-Arg-	-()-Ala-	

図2. (A)ウシの脳(bovine brain), (B)ホタテガイ(*Patinopecten yessoensis*), (C)イソギンチャク(*Metridium senile*)および(D)テトラヒメナ(*Tetrahymena pyriformis*)から得られたカルモデュリン蛋白質の一次構造(アミノ酸配列)データとその比較。

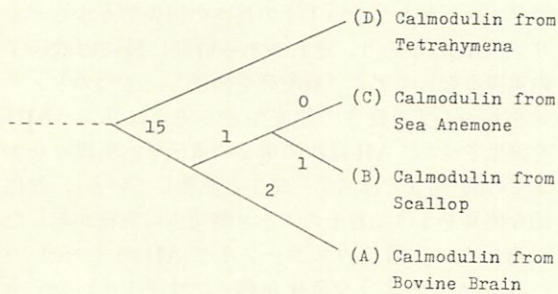


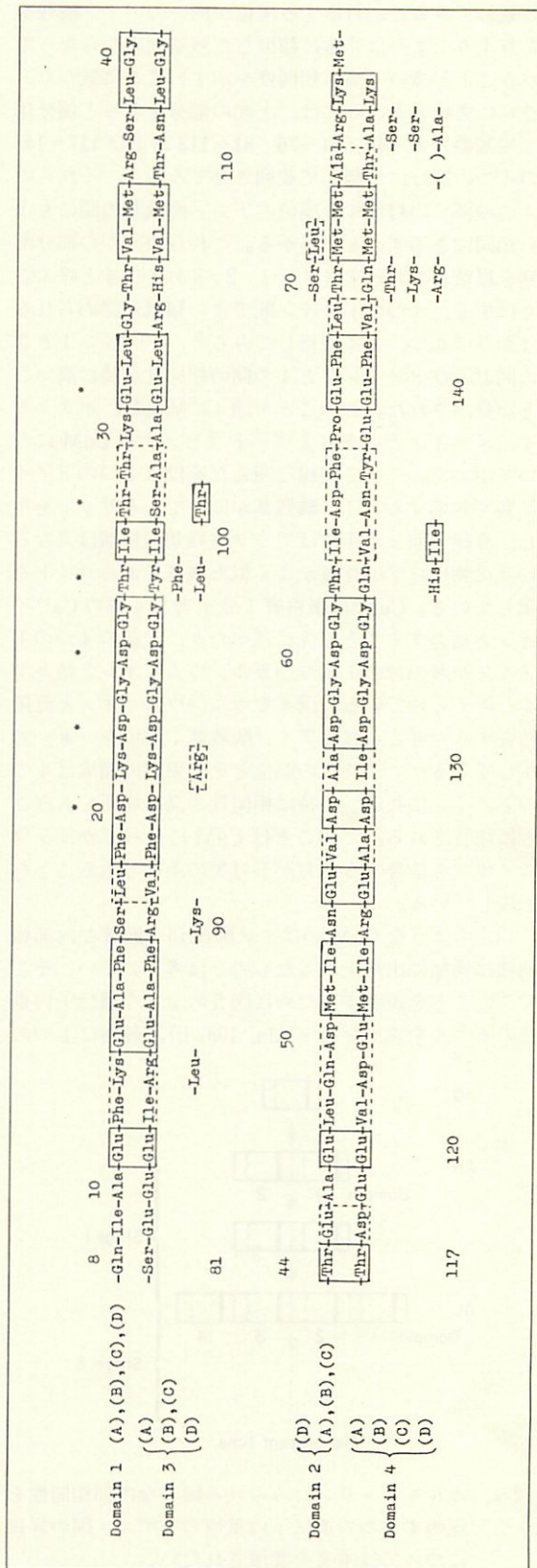
図3. いくつかの動物のカルモデュリンのアミノ酸配列データから評価された生物系統樹。(A)~(D)の生物種については図2を参照されたい。各枝の数字は最大節約法を応用した時のヌクレオチド置換数を示す。

る。最近、系統関係のもっとはなれた植物やカビ類についてもCaM蛋白質の一次構造が決定されつつあり、それらのデータを利用するとより広汎な生物系統樹を作成することが可能となる。CaMのように生物界に広く分布し、かつ進化に対してきわめて保守的な蛋白質は生命の進化の「長い時間スケール」を計るよい尺度となり得る。例えば、動物、植物およびカビ類がいつそれらの共通祖先から分離したのかを議論する時には、CaMの分子進化は貴重なデータを提供するはずである。

III. カルモデュリン蛋白質の内部相同性と遺伝子内重複

CaMのもつもう1つの興味ある特徴として、そのアミノ酸配列が内部相同性 (internal homology) をもっていることがあげられる (Iida, 1982 b)。内部相同性と言うのは、1つの蛋白質のアミノ酸配列のある部分と別のある部分が相互によく類似していることを言う。CaMの場合、図2のアミノ酸残基番号8~76の部分と81~148の部分を取り出して両者を並列させ

図4. カルモデュリンのアミノ酸配列の内部相同性。アミノ酸配列を図に示すように4つの類似したドメインに分割しおのおのを比較した。ドメイン1と3およびドメイン2と4の間に特に高い相同性が認められる。実線で囲ったサイトは2つのドメイン間でアミノ酸残基が同一の場所を示し、点線で囲ったサイトは2つのドメイン間でアミノ酸残基は互に異なるがそれらの物理化学的性質はよく類似している場所を示す。*のサイトはカルシウムイオンと直接結合すると考えられるアミノ酸残基の部位を示す。生物種(A)~(D)については図2を参照されたい。



比較してみると、対応する座位が同一のアミノ酸残基を有するかまたは非常に類似した残基同志からなっていることが多い（高い相同性を示す）ことに気づく。さらに興味あることには、上記の部分をもっと細分化し残基番号8~40, 44~76, 81~113および117~148の4つの部分に分割して並列させてみると、それらの4つの部分の対応する座位のアミノ酸残基の間にもよい相同性を示すことがわかる。これらの4つの部分配列を以後おのおのドメイン1, 2, 3および4と呼ぶことにする。4つのドメイン間でよい類似が認められるばかりでなく、よく観察してみると、ドメイン1と3の間およびドメイン2と4の間の相同性が特に高いことが見出された。このことを図4に哺乳類、ホタテガイ、イソギンチャクおよびベトラヒメナのCaMについて示した。ここで実線で囲んだ座位は2つのドメイン間で対応するアミノ酸残基が同一であるサイトを示し、点線で囲んだ座位はアミノ酸残基の種類は異なるが互に物理化学的性質がよく似た残基であるサイトを示している。CaMは蛋白質1分子当たり4個のCa²⁺イオンと結合することを先に述べたが、上記の4つのドメインがおのおの1個づつカルシウムイオンと結合すると考えられている。図4でカルシウムイオンと直接結合すると考えられるアミノ酸残基の部位を(*)で示してあるが、これらの部位とその周辺の領域は4つのドメインに共通して特に相同性の高い場所であることは注目される。このことはCaMにとってカルシウムイオンを結合する能力が不可欠の条件であることを示唆している。

以上のようなCaMのアミノ酸配列の顕著な内部相同性は偶然に出来上がったものとは考えにくい。そこでこのことを説明するために図5のような遺伝子内重複のモデルを提出した(Iida, 1982 b)。最初に1つの

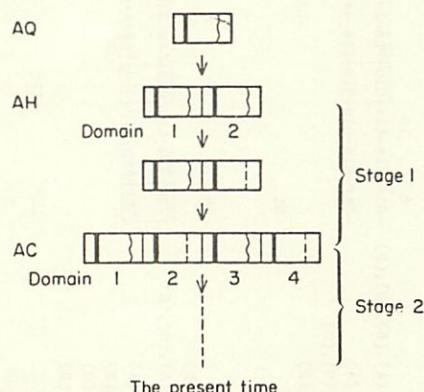


図5. カルモデュリンにみられる顕著な内部相同性を説明する為の遺伝子内重複のモデル。図の詳細については本文を参照されたい。

ドメインしかもたない1/4の長さの祖先型カルモデュリン(AQ)があり、それがある時期に最初の遺伝子内重複を起してアミノ酸配列を伸長し、2つのドメインをもつ1/2の長さの祖先型カルモデュリン(AH)を誕生させた。AHは次の第2の遺伝子内重複を起すまでの間(図5ではステージ1と記載している)、進化的な変異をうけて若干のアミノ酸座位に置換が起ったに違いない。図5のステージ1でAHの(~~~~)→(-----)がそのような置換が起ったサイトを、(——)→(——)は変異が起らずアミノ酸残基が保存されたサイトを模式的に示している。ステージ1の期間がどの程度長かったのかははっきりしないが、ある時期に進化したAHが再び遺伝子内重複を起してドメインを倍増し、4つのドメインをもった現存カルモデュリンの祖先型(AC)が誕生したと考えられる。ACの誕生から現在に到るまでの期間をステージ2とする。この期間においてもACはいくつかの進化的変異を受けたであろうが、誕生直後の祖先型ACではドメイン1と3およびドメイン2と4は互に全く同一のアミノ酸配列をもっていた(図5参照)ことに注意されたい。前節でも述べたが、CaMは進化に対してきわめて保守的な蛋白質であるので、ステージ2の長い期間の間でも変異を受けたアミノ酸座位はそれ程多くはなく、ほぼ祖先型ACのおもかげを強く残していると思われる。そのように考えると、図4の現存CaMのドメイン1と3およびドメイン2と4の間の相同性がきわめて高く、且つドメイン1, 2, 3および4の間にもよい相同性が保たれている事実が以上のモデルでよく説明される。

CaMの例のように同一の遺伝子内でDNAの重複が起り、その遺伝子の蛋白質のアミノ酸配列が伸長される現象を「遺伝子内重複」と言う。このような伸長はCaM以外に他のいくつかの蛋白質の遺伝子でも起ったであろう。しかし多くの蛋白質の場合、進化速度が速くかつ長い時間スケールの中でアミノ酸置換が頻繁に起った結果、そのような遺伝子内重複のおもかげがほとんど失われてしまったのではないと思われる。遺伝子内重複とは別にDNAの重複がよくみられるのは、ある蛋白質をコードする遺伝子がもう1個コピーを作って別の遺伝子を誕生させた場合である。例えばグロビン遺伝子のように、祖先型グロビン遺伝子から α -グロビンと β -グロビンの遺伝子に分れた場合には、1個の遺伝子が伸長するのではなく重複によって遺伝子が2個生成したのである。

いずれにせよ、CaMが進化に対してきわめて保守的な蛋白質であったおかげで、カルモデュリンの祖先型AQ, AH, ACがどのようなアミノ酸配列をもっていたかをうかがい知る手がかりが得られた。もし図5

AQ: ()-()-()-Glu-()-()-Glu-()-()-()-()-Asp-
 ()-Asp-Gly-Asp-Gly-()-Ile-()-()-()-Glu-()-()-()-
 ()-Met-()-()-()-()-....

AH: AQ..AQ....

EH: ()-()-()-Glu-()-()-Glu-Ala-Phe-()-()-Phe-Asp-
 Lys-Asp-Gly-Asp-Gly-()-Ile-Thr-()-()-Glu-Leu-()-()-
 Val-Met-()-()-Leu-Gly-....-Thr-()-()-Glu-()-()-Glu-
 Met-Ile-()-Glu-()-Asp-()-Asp-Gly-Asp-Gly-()-Ile-()-
 ()-()-Glu-Phe-()-()-Met-Met-()-()-Lys-()-....

AC: EH..EH....

図6. 遺伝子内重複のモデル(図5)において予想されるカルモデュリンとその祖先型のアミノ酸配列の推定図。ここで()は当研究ではアミノ酸残基の種類を推定できない座位を示す。AQは1つのドメインしかもたない祖先型, AHとEHは2つのドメインをもつ祖先型, ACは4つのドメインをもつ原始カルモデュリンを示す。詳細については本文を参照されたい。

のモデルが正しく, アミノ酸置換がきわめて起りにくかったとすると, 現存CaMで4つのドメインの対応する座位すべてに共通して同一のアミノ酸残基がみられる場所は, AQの対応する場所でもそのようなアミノ酸残基によって占められていたに違いない。図4でこのような場所をひろいあげると(---Glu¹¹---Glu⁸⁴---Glu⁴⁷---Glu¹²⁰---), (---Asp²⁰---Asp⁹³---Asp⁵⁶---Asp¹²⁹---), (---Asp²²---Asp⁹⁵---Asp⁵⁸---Asp¹³¹---), (---Gly²³---Gly⁹⁶---Gly⁵⁹---Gly¹³²---), (---Asp²⁴---Asp⁹⁷---Asp⁶⁰---Asp¹³³---), (---Gly²⁵---Gly⁹⁸---Gly⁶¹---Gly¹³⁴---), (---Glu³¹---Glu¹⁰⁴---Glu⁶⁷---Glu¹⁴⁰---) および(---Met³⁶---Met¹⁰⁹---Met⁷²---Met¹⁴⁵---)の8ヶ所が見出された。さらに図4をよく観察すると, 3つのドメインで共通のアミノ酸残基がみられるが, 他の1つのドメインで物理化学的性質が似ている他のアミノ酸残基に置換している組合せの場所が2ヶ所見出される。第1の場所は(---Glu¹⁴---Glu⁸⁷---Asp⁵⁰---Glu¹²³---)で, この組合せは哺乳類, ホタテガイ, イソギンチャク, テトラヒメナのすべてに共通して見られる。これを図5のモデルで考えると, 原始ACでは(---Glu¹⁴---Glu⁸⁷---Glu⁵⁰---Glu¹²³---)のように4つのドメインですべて共通のアミノ酸残基であったものが, 上記4種の生物種がそれらの共通の祖先から分れる以前にすでに-Glu⁵⁰→-Asp⁵⁰と変異を起していたと考えるのが自然である。もしそうなら, これらの座位に対応するAQのサイトのアミノ酸残基も-Gluであったに違いない。第2の場所は(---Ile²⁷---Ile¹⁰⁰---Ile⁶³---Val¹³⁶---)で, これは哺乳類, ホタテ

ガイ, イソギンチャクで見られるものの, テトラヒメナでは対応する組合せは(---Ile²⁷---Ile¹⁰⁰---Ile⁶³---Ile¹³⁶---)になっている。明らかなようにテトラヒメナにみられる組合せが原始ACのおもかげをとどめたもので, 哺乳類, ホタテガイおよびイソギンチャクにみられる-Val¹³⁶はそれらの共通祖先がテトラヒメナと分離した後に-Ile¹³⁶→-Val¹³⁶と変異したと考えると都合がよい。いずれにせよ, これらの座位に対応するAQのサイトでも-Ile-が占有していたと考えられる。

次に, ドメイン1と3又はドメイン2と4のみに共通して同一のアミノ酸残基が見出される座位について考える。図5で示すように, このような座位はステージ1の最後の段階で進化的変異をうけたAHが第2の遺伝子内重複を起す直前の1/2の長さのカルモデュリン(EH)でもそのようなアミノ酸残基を有していたはずである。しかしそのような2つのドメインで同一のアミノ酸残基を有する組合せでも, ある生物種では別のアミノ酸残基に置換している場合も見出された。例えば, 哺乳類, ホタテガイ, イソギンチャクでは(---Lys²¹---Lys⁹⁴---)のように同一の残基の組合せでも, テトラヒメナでは(---Lys²¹---Arg⁹⁴---)のように異なったアミノ酸残基に変異している。このような場合も, (---Lys²¹---Lys⁹⁴---)が原始ACのおもかげをとどめたもので, 哺乳類, ホタテガイおよびイソギンチャクの共通祖先とテトラヒメナが分離した後にテトラヒメナの側で-Lys⁹⁴→-Arg⁹⁴のアミノ酸残基の置換が起きたと考えるのが自然であろう。図4を参照すると, このような場所が他にも数ヶ所観察される。以上のような考察をまとめると, 原始カルモデュリン(AC)とそ

の祖先 (AQ, AH, EH) のアミノ酸配列をある程度うかがい知ることが可能となり、図6にそのような推定図を示してある。ここで () で示した場所は、以上の解析からだけではアミノ酸残基の種類を同定できない座位である。

4つのドメインをもった現存カルモデュリンの祖先型 (AC) は、動物、植物およびカビ類がその共通祖先から分れた時代 (11~13億年前) よりさらに昔にすでに誕生していたことは確実であり、2つのドメインをもった AH や1つのドメインの AQ は AC の誕生よりはるか昔にすでに存在していたはずである。そのような太古の昔に存在していたであろう蛋白質のアミノ酸配列を現存の生物種のデータから推測できるのも、CaM のもっているいくつかの特徴をうまく利用したからに他ならない。このような蛋白質の構造についての理論的な解析結果が化石の調査等によって実際に確認されれば大変有難いが、何分あまりに太古の昔の話なので実証は困難かも知れない。化石研究をおこなっておられるかたがたからの御意見をうかがえれば幸である。

IV. カルモデュリンをコードしている遺伝子の一次構造と分子進化

前節IIおよびIIIではおもにCaM蛋白質のアミノ酸配列データを議論の対象としてきた。一方、最近蛋白質をコードしている DNA の一次構造 (ヌクレオチド配列) のデータがいろいろな生物種について得られるようになった。このような DNA からの情報 (遺伝情報) は、従来のアミノ酸配列データからでは得られない貴重な情報を含んでおり、それらを解析することによって蛋白質の起源や進化についてさらに詳細な知見を得ることが可能である。このような立場から CaM をコードしている遺伝子の一次構造を解析した (Iida, 1984)。またこれと比較するために、ヒストン H 4 蛋白質をコードしている遺伝子についても若干の考察を加えた。CaM やヒストン H 4 のように進化に対してきわめて保守的な蛋白質の遺伝子の構造に興味をもたれたからである。

CaM 蛋白質をコードしている遺伝子の情報で現在利用できるものは、デンキウナギ (*Electrophorus electricus* [硬骨魚類]) とニワトリ (chicken [鳥類]) についてその mRNA から得られた cDNA (mRNA に complementary な DNA) の一次構造

CHICKEN:	ATG	GCT	GAT	CAA	CTG	ACA	GAA	GAG	CAG	ATT	GCA	GAA	TTC	AAA	GAA	GCT	TTT	TCA	CTA	TTT	GAC	20
EEL:	ATG	GCA	GAT	CAG	CTG	ACT	GAG	GAA	CAG	ATT	GCT	GAG	TTC	AAG	GAG	GCG	TTT	TCC	CTC	TTT	GAC	
																						30
	AAG	GAT	GGT	GAT	GGT	ACT	ATA	ACT	ACA	AAG	GAG	TTG	GGG	ACT	GTG	ATG	AGA	TCA	CTT	GGT		40
	AAA	GAT	GGT	GAC	GGC	ACC	ATC	ACC	ACC	AAA	GAG	CTG	GGT	ACT	GTG	ATG	CGC	TCT	CTG	GGC		
																						50
	CAG	AAC	CCC	ACA	GAA	GCA	GAA	TTA	CAG	GAC	ATG	ATC	AAT	GAA	GTA	GAC	GCT	GAT	GGC	AAT		60
	CAA	AAC	CCC	ACC	GAG	GCA	GAG	CTG	CAG	GAC	ATG	ATC	AAT	GAA	GTG	GAT	GCT	GAC	GGC	AAT		
																						70
	GGC	ACA	ATT	GAC	TTC	CCA	GAG	TTT	CTG	ACA	ATG	ATG	GCA	AGA	AAA	ATG	AAA	GAT	ACA	GAT		80
	GGA	ACA	ATA	GAC	TTC	CCG	GAG	TTC	CTG	ACC	ATG	ATG	GCC	AAG	AAA	ATG	AAG	GAC	ACA	GAC		
																						90
	AGC	GAA	GAA	GAA	ATT	AGA	GAA	GCG	TTC	CGT	GTG	TTT	GAC	AAG	GAT	GGT	AAT	GGT	TAC	ATT		100
	AGT	GAA	GAA	GAG	ATC	CGA	GAA	GCC	TTC	CGA	GTT	TTT	GAC	AAG	GAC	GGT	AAT	GGC	TAC	ATC		
																						110
	AGT	GCT	GCA	GAA	CTT	CGT	CAT	GTG	ATG	ACA	AAT	CTT	GGG	GAG	AAG	CTA	ACA	GAT	GAA	GAA		120
	AGT	GCA	GCC	GAG	TTG	CGA	CAT	GTC	ATG	ACT	AAC	TTG	GGC	GAG	AAG	CTG	ACG	GAC	GAG	GAG		
																						130
	GTT	GAT	GAA	ATG	ATT	AGG	GAA	GCA	GAC	ATT	GAT	GGT	GAT	GGT	CAA	GTA	AAC	TAT	GAA	GAG		140
	GTG	GAT	GAG	ATG	ATC	CGA	GAG	GCC	GAC	ATC	GAT	GGC	GAC	GGC	CAG	GTG	AAC	TAT	GAA	GAG		
	TTT	GTA	CAG	ATG	ATG	ACA	GCG	AAG	TGA													
	TTC	GTG	CAA	ATG	ATG	ACT	GCA	AAG	TGA													

図7. ニワトリとデンキウナギのカルモデュリン遺伝子 (コドンの領域) のヌクレオチド配列の比較。数字は N 末端からのアミノ酸残基の番号に相当し、*のサイトは2つの配列間でヌクレオチドが互に異なる場所を示す。

データである (Lagacé et al., 1983; Putkey et al., 1983)。cDNA データには蛋白質のアミノ酸配列を指定しているコドンの領域以外に、5'-untranslated および 3'-untranslated 領域のヌクレオチド配列 (以下塩基配列と略す) もある。しかし、我々は CaM 蛋白質の一次構造 (アミノ酸配列) を問題としているので、それらの翻訳されない領域については考えないで以後コドン領域のみ議論の対象とする。

図 7 にニワトリとデンキウナギの CaM 遺伝子のコドン領域の塩基配列を比較したデータを示す (与えられた数字は N 末端からのアミノ酸残基の番号に対応する)。図 7 で示された領域は各々 150 個のコドンからなり、開始コドン (ATG) と終止コドン (TGA) は両生物種で共通である。これらの 2 つのコドンを除外すると、残りの 148 個のコドンが 148 個のアミノ酸残基を指定する。コドンの配列から予想されるアミノ酸配列は両者でわずか 1 個しか異ならない (74 番目の残基がニワトリでは Arg, デンキウナギでは Lys)。しかし、コドンの塩基配列を両生物種で比較してみるといくつかの興味深い点に気付く。図 7 で星印 (*) は両者でヌクレオチドの種類が異なった場所であり、その他のヌクレオチドは両生物種で同一である。非常に保守的なアミノ酸配列から期待されるように cDNA コドン領域の塩基配列も大変よく似ており、全体として 79% の相同性がある。特にコドンの第 1, 第 2 番目のサイトが非常によく保存されている一方、コドンの第 3 番目のサイトは両生物種でかなり異なっていることが見出された (第 3 番目位置は 148 個のうち 84 個が

異なっている)。もしヌクレオチドの置換が全くランダムに起ったとすれば、両生物種の対応する位置で異なった塩基をもつ確率は $\lambda = 3/4 = 0.75$ (飽和値) となるはずである。CaM 遺伝子のコドン第 3 番目のサイトの $\lambda = 84/148 = 0.568$ は 0.75 の値に近く、ヌクレオチド置換がかなり頻繁におこなわれてきてほとんど飽和していることを示している。

両生物種でヌクレオチドの置換がどの程度起ってきたかを定量的に評価するには多重置換や復帰置換の補正をする必要がある。Kimura (1980) によって与えられた 2 パラメーターモデルがこの目的の為に有効である。サイト当りヌクレオチド全置換回数 K は、2 つの遺伝子が共通の祖先から枝分れた年数 T を使って $K = 2 Tk$ で与えられる。ここで k はサイト当り 1 年間にヌクレオチドが置換する速度を示す。2 パラメーターモデルでは transition type のヌクレオチド置換と transversion type の置換を区別して考える。前者はプリン同志 ($A \leftrightarrow G$) 又はピリミジン同志 ($T \leftrightarrow C$) の置換であり、後者はプリンとピリミジンの間の置換である。このモデルでは K は次式で与えられる。

$$K = -\frac{1}{2} \ln \left\{ (1 - 2P - Q) \sqrt{1 - 2Q} \right\}$$

ここで P と Q は 2 つの遺伝子の間で各々 transition と transversion の置換が起っている割合を示す。図 7 のニワトリとデンキウナギのデータからコドンの第 1, 第 2 および第 3 番目のサイトについて K 値 (各々 K_1, K_2 および K_3 とする) を計算した。コドンの第 1 番

HUMAN (I):	ATG	TCC	GGC	TGT	GGA	AAG	GGC	GGA	AAG	GGC	TTA	GGC	AAA	GGT	GGC	GCT	AAG	CGC	CAC	CGC	AAG	20
WHEAT:	ATG	TCC	GGG	CGC	GGC	AAG	GGA	GGC	AAG	GGC	CTA	GGC	AAG	GGC	GGC	GCC	AAG	CGC	CAC	CGG	AAG	
																						30
	GTC	TTG	AGA	GAC	AAC	ATT	CAG	GGC	ATC	ACC	AAG	CCT	GCC	ATT	CGG	CGT	NTA	GCT	CGG	CGT		40
	GTC	CTC	CGC	GAT	AAC	ATC	CAG	GGC	ATC	ACC	AAG	CCG	GCG	ATC	CGG	CGG	CTG	GCG	CGG	CGG		
																						50
	GGC	GGC	GTT	AAG	CGG	ATC	TCT	GGC	CTC	ATT	TAC	GAG	GAG	ACC	CGC	GGT	GTG	CTG	AAA	GTG		60
	GGC	GGC	GTG	AAG	CGC	ATC	TCG	GGG	CTC	ATC	TAC	GAG	GAG	ACC	CGC	GGC	GTG	CTC	AAG	ATC		
																						70
	TTC	TTG	GAG	AAT	GTG	ATT	CGG	GAC	GCA	GTC	ACC	TAC	ACC	GAG	CAC	GCC	AAG	CGC	AAG	ACC		80
	TTC	CTC	GAG	AAC	GTC	ATC	CGC	GAT	GCC	GTC	ACC	TAC	ACC	GAG	CAC	GCC	CGC	CGC	AAG	ACC		
																						90
	GTC	ACA	GCC	ATG	GAT	GTG	GTG	TAC	GCG	CTC	AAG	CGN	CAG	GGG	AGN	ACC	CTC	TAC	GGC	TTC		100
	GTC	ACC	GCC	ATG	GAC	GTC	GTC	TAC	GCG	CTC	AAG	CGC	CAG	GGC	CGC	ACC	CTC	TAC	GGC	TTC		
	GGA	GGC	TAG																			
	GGC	GGC	TAA																			

図 8. 人間と植物 (小麦) のヒストン H 4 遺伝子 (コドンの領域) のヌクレオチド配列の比較。記号の説明については図 7 を参照されたい。

目サイトでは $P=4/148$, $Q=3/148$ となり上式より $K_1=0.049$ (± 0.019) と計算された。ここでカッコ内の値は標準誤差を示す。同様にコドンの第2番目のサイトでは $P=1/148$, $Q=0/148$, $K_2=0.0068$ (± 0.0068), コドンの第3番目のサイトでは $P=53/148$, $Q=31/148$, $K_3=1.44$ (± 0.49) と評価された。3つのサイトすべてで $P > (1/2)Q$ となり, transition type の置換が transversion type の置換より頻繁に起っていることが示される。

以上の結果で CaM 遺伝子に特徴的なことは, コドンの第3番目サイトの K_3 が第1, 第2番目サイトの K_1 , K_2 の値よりはるかに大きく, さらに第1番目サイトの K_1 値が第2番目サイトの K_2 より大きいことである。 K_3 の値が異常に大きいのはすべてアミノ酸残基の種類を変えない同義置換によるものであり, この置換は進化に対してほとんど中立的であることを示す (Kimura, 1982)。同様にコドン第1番目のサイトのヌクレオチド置換もすべて同義置換であることが見出され, このことにより $K_1 > K_2$ の関係が説明される (遺伝コード表を参照すると, コドン第1番目の同義置換は Leu と Arg 残基のときにのみ許される)。

次に $K_2 < K_1 \ll K_3$ の関係が CaM 遺伝子だけでなく, 一般的に見られる現象かどうかを検討してみた。我々は, 非常に保守的な蛋白質であるヒストン H 4 の遺伝子を人間 (Sierra et al., 1983) と植物 (小麦) (Tabata et al., 1983) について比較したところ (図8参照), やはりコドンの第3番目サイトに多数の同義置換とコドンの第1番目サイトに数ヶ所の同義置換が観察された。CaM 遺伝子と同様に $K_2 < K_1 \ll K_3$ の関係がヒストン H 4 遺伝子でも成立つようである。

次に $K = 2 Tk$ の関係式を使って, CaM のコドンの各サイトについて k の値 (サイト当り1年間にヌクレオチドが置換する速度) を評価した。我々はデンキウナギとニワトリがその共通祖先から $T=3.5 \times 10^8$ 年前に分離したと仮定する。この値は硬骨魚類から両生類が分離した年代から推定したものである (Romer, 1966)。この値とすでに評価された K_i ($i=1, 2$ 又は 3) の値から, CaM 遺伝子のコドンの第1, 第2および第3番目のサイトに対する k の値は各々 $k_1=0.070 \times 10^{-9}$, $k_2=0.0097 \times 10^{-9}$, $k_3=2.1 \times 10^{-9}$ と計算された。

これらの結果を, 従来進化に対して最も保守的と考えられてきたヒストン H 4 と比較してみた (Kimura, 1977)。ヒストン H 4 の1年間にアミノ酸が置換する速度は $k_{aa}=0.006 \times 10^{-9}$ と評価されていることはすでに述べた。CaM の k_2 の値はほぼ k_{aa} に匹敵する程小さい値である。一方, CaM の k_1 値はヒストン H 4 の k_{aa} のほぼ10倍程度大きい, すでに述べたよう

に k_1 値が大きいのはすべてアミノ酸残基の種類を変えない同義置換によるものであることに注意しよう。このように考えると CaM 遺伝子のコドンの第1, 第2番目のサイトはヒストン H 4 遺伝子に比べるとやや進化速度は大きい, ほぼ同程度に保守的と言える。これらと対照的に, CaM のコドン第3番目のサイトの $k_3=2.1 \times 10^{-9}$ の値は, 従来最も進化速度が大きいと考えられてきたフィブリノペプチドのヌクレオチド置換速度 (2.3×10^{-9}) やヒストン H 4 遺伝子のコドン第3番目サイトのヌクレオチド置換速度 ($(3.7 \pm 1.4) \times 10^{-9}$) とほぼ同程度に大きい (Kimura, 1977)。

以上の結果から, CaM 遺伝子においてはアミノ酸残基の種類を変えない同義的なヌクレオチド置換は最も速い速度で起っている一方, アミノ酸残基を置換させるヌクレオチド置換はヒストン H 4 と同程度に遅い速度で起っていることが注目される。

V. おわりに

現存生物がもっている蛋白質のアミノ酸配列や遺伝子の塩基配列データを利用して生物進化の系統関係を研究するには, 対象とする蛋白質や遺伝子の系をどのように設定するかが問題となる。もし, 近縁の生物種の詳細な系統関係を知りたい時には, 進化速度の比較的速い蛋白質や遺伝子の系を選ぶ必要がある。このような系について, 従来の形態学的研究や化石調査から得られた系統関係と分子進化から得られた結果は矛盾しないようである (Dayhoff, 1972)。

一方, 広汎な生物進化を長い時間スケールで概観したい時には, 従来の形態学的方法や化石調査だけでは充分な知見が得られないことが多い。そのような場合でも, 分子進化の研究からは生物界に広く存在してしかも進化速度の比較的遅い蛋白質や遺伝子を対象とすることによって, かなり詳細な知見を引出すことが可能である。このような例として, 従来チトクローム C 蛋白質をもとにした系統樹の研究がさかんに行われてきた (Dayhoff, 1972)。この蛋白質より進化速度が遅く, しかも広汎に分布している蛋白質としてカルモデュリンは分子進化の研究に最適の系であると思われる。ただし, カルモデュリンは真核生物には広く存在するが, 原核生物にも存在するかどうか今のところ明らかでない, 原核生物まで含めた系統樹をカルモデュリンで作成できるかどうか問題が残る。チトクローム C 等の蛋白質にくらべてカルモデュリンがもつ非常に有利な点はアミノ酸配列が内部相同性を示すことであり, この特徴を利用すれば原始カルモデュリンが誕生する以前の蛋白質のアミノ酸配列等の貴重な情報を引出せることはこの報告で詳しくふれた。

蛋白質以外に, 遺伝子のレベルで広汎な生物進化を

論じようとすれば適切な系がいろいろある。トランスファースRNA やリボソーム RNA 遺伝子は原核、真核生物を含めて広く分布しており、これらの系を使った生物系統樹の研究もさかんに行われつつある。我々のところでも、リボソーム RNA (small and large subunits) 遺伝子のデータを解析してそのような研究をおこなっている。

以上述べたように、分子進化による生命の進化の研究は従来の現存生物の形態学的研究や古生物学的な化石の研究からでは観察しにくい側面 (例えば太古の昔の生命進化の様子や広汎な生命進化を長い時間スケールで論じる場合) を研究する時にきわめて有効な方法となり得る。従って分子進化の研究と従来の方法は矛盾するものではなく、互に相補的な役割をもっているので、両者を併用することにより生命進化のより深い理解が得られるのではないかと考えられる。

文 献

- Cheung, W. Y. (1980) Calmodulin Plays a Pivotal Role in Cellular Regulation. *Science*, 207, 19-27.
- Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Washington, D. C., Vol. 5.
- Iida, Y. (1982 a) Information Chemistry and Molecular Evolution of Protein. *Evolutionary Tree of Calmodulin*. *Bull. Chem. Soc. Japan*, 55, 2683-2684.
- (1982 b) Molecular Evolution of Protein. Internal Homology in the Amino Acid Sequence of Calmodulin. *J. Mol. Biology*, 159, 167-177.
- (1984) cDNA Sequences and Molecular Evolution of Calmodulin Genes of Chicken and Eel. *Bull. Chem. Soc. Japan*, 57, 2667-2668.
- Kimura, M. (1977) Preponderance of Synonymous Changes as Evidence for the Neutral Theory of Molecular Evolution. *Nature*, 267, 275-276.
- (1980) A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide Sequences. *J. Mol. Evol.*, 16, 111-120.
- (1982) *Molecular Evolution, Protein Polymorphism and The Neutral Theory*. Springer-Verlag, Berlin, Heidelberg, New York.
- Lagacé, L. et al. (1983) Identification of Multiple Species of Calmodulin Messenger RNA Using a Full Length Complementary DNA. *J. Biol. Chem.*, 258, 1684-1688.
- Putkey, J. A. et al. (1983) Chicken Calmodulin Genes. *J. Biol. Chem.*, 258, 11864-11870.
- Romer, A. S. (1966) *Vertebrate Paleontology*, Third Ed. The University of Chicago Press, Chicago and London.
- Sierra, F. et al. (1983) Structure and in vitro Transcription of a Human H4 Histone Gene. *Nucl. Acids Res.*, 11, 7069-7086.
- Tabata, T. et al. (1983) The Structural Organization and DNA Sequence of a Wheat Histone H4 Gene. *Nucl. Acids Res.*, 11, 5865-5875.
- 武居 洋・松田源治 (1980) ヘモグロビンの分子進化. *蛋白質核酸酵素*, 25, 1053-1068.
- Wasserman, R. H. et al. (1977) *Calcium-Binding Proteins and Calcium Function*. North-Holland Inc., New York.
- Zuckerland, E. and Pauling, L. (1962) *Horizons in Biochemistry*. Academic Press Inc., New York.

Evolution of life as studied by molecular evolution of calmodulin.

Yôichi Iida

(Abstract)

The amino acid sequences of calmodulin, one of calcium-binding proteins, taken from bovine brain, scallop, sea anemone and tetrahymena were compared to examine its molecular evolution, and its evolutionary tree was constructed. Another notable feature of the amino acid sequence of calmodulin is its internal homology. It can be subdivided into four domains having a similar amino acid sequence. An intragenic duplication model was proposed to explain such an internal homology. cDNA sequences of calmodulin genes taken from chicken and electric eel were also examined to study nucleotide changes between those genes.