

Title	大腸菌リボヌクレアーゼHの立体構造と機能
Author(s)	片柳, 克夫
Citation	大阪大学, 1992, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3088031
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

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

The University of Osaka

大腸菌リボヌクレアーゼHの立体構造と機能

1991

E.

片 柳 克 夫

大腸菌リボヌクレアーゼHの立体構造と機能

.

1991

.

广片 柳 克 夫

目 次

本 論

第1章 大腸菌 RNase H の立体構造 第1節 全体の分子構造 ・・・・・・・・・・・・・・・3 第2節 分子内水素結合 ・・・・・・・・・・・・・・・・14 第3節 トリプトファンの分布 ・・・・・・・・・・・・・22 第4節 システインの分布 ・・・・・・・・・・・・・・・22 第5節 ヒスチジンの分布 ・・・・・・・・・・・・・・23 第6節 2次構造 ① ジョイント予測法、円2色偏光法の結果との比較 ・・24 ② ヘリックス ・・・・・・・・・・・・・・・・26 ③ シート・・・・・・・・・・・・・・・・・・・・・・・・・・・・29 **(4)** リバース・ターン ・・・・・・・・・・・・・・・・29 第7節 水分子の分布 ・・・・・・・・・・・・・・・31 第8節 結晶のパッキング ・・・・・・・・・・・・・・・35 第9節 活性金属の結合部位 ・・・・・ ••••• 第10節 大腸菌 RNase H と

逆転写酵素 RNase H ドメインの活性中心 ・・・・41

第2章 大腸菌 RNase H の活性部位変異体の立体構造解析 第1節 活性部位変異体 D10N、E48Q、D70N ・・・・・・・46

第2節	活性部位変異体の立体構造 ・・・・・・・・・・47
第3節	2 つの非同型な結晶中での構造の比較
	— His 124 近傍のループの柔軟性 — ・・・・・53

第3章 活性中心と核酸結合部位の同定 第1節 DNase I との比較 • • • • • • • • • • 57 第2節 HIV RNase H との比較(金属結合部位) •••••61 第3節 静電ポテンシャルの分布 · · · · · 62 第4節 核酸認識部位 · · 65 結 論 ٠ 謝 辞 • • • • • • • • • • • • 72 ٠ •

実験の部

第1章の実験

第1節	RNase H の結	晶	• • • •	• •	•••	•••	•	•	•	••	• 73
第2節	X線回折強度	データ(の測定	• •		•••	•	•	•	••	• 74
第3節	重原子同型置	換法に	よる位相社	央定	•	•••	•	•	•	•••	• 76
第4節	分子モデルの	構築と	情密化	••	•••	•••	•	•	•	••	• 82
第5節	Mg ²⁺ 結合部位	の同定	• • • •	••	••	•••	•	•	•	•••	• 90
<u>第2章の</u> 実	(験・・・・	• • • •	• • • •	•••	•••	• •	•	•	•	•••	• 92
参考文献	• • • • • •	• • • •	• • • •	•••	•••	•••	•	•	•	••	• 99

(ii)

序 論

リボヌクレアーゼの立体構造は 今日まで RNase A および RNase T₁ などをはじ めとして、いくつかの酵素について研究されている。これらのリボヌクレアーゼが 単鎖RNAを基質とするのに対して、この研究の対象であるリボヌクレアーゼHは、 DNA/RNA ハイブリッド2重鎖を特異的に認識し、そのRNA鎖のみを切断し、5' リン酸基と3'水酸基を生産する。

この独特な活性は、Stein と Hansen (1969)により、仔ウシの胸腺から発見され、 リボヌクレアーゼH (RNase H)と命名された。 このHは、DNA/RNA Hybrid に由 来している。 大腸菌由来の RNase H は、 Miller ら(1973)によって単離され、 Berkower ら(1973)によってその性状が明らかにされた。この酵素は DNA/RNA ハ イブリッド2重鎖のうち、RNA鎖のみをエンドヌクレアーゼ活性により特異的に 加水分解し、その結果5'リン酸基が生ずる(Crouchら、1982)。 通常のアッセイ 系では、この酵素は基質に対し、ピリミジンの隣の開裂部に僅かな特異性を持って いる(Crouch ら、1982)。この酵素の活性にはMg²⁺が必須で、活性測定にはこのイ オンが加えられる。しかし、場合によってはMg²⁺はMn²⁺に置き換えることが可能で ある (Berkower ら、1973)。この酵素の遺伝子は、大腸菌クロモソームの中に単一 コピーとして存在することが示され (Carlら、 1980)、また dna Q 遺伝子からは 64塩基対 離れて存在することが報告されている(Maki ら、1983)。また、最近、大 腸菌から"第2の RNase H (RNase H II)"が 単離されたため (Itaya, 1990)、こ れに対して 今までの rnh A 遺伝子からつくられている RNase H は、RNase HIと 呼んで区別されることがある。DNAの塩基配列から、この酵素は155 個のアミノ 酸残基の 単一ポリペプチド鎖からなることが示された(Kanaya ら、1983)。RNase H 活性は原核生物から真核生物にいたるまで広く生物界に存在する。その生理学的

- 1 -

機能については不明な点もあるが、RNase H が、十分に分化された細胞内では酵素 活性を促進していることや、 *in vitro* での Col E1 プラスミドの複製に関与して いること、などの発見(Itoh ら、1980) は、この酵素がDNAの複製に関わってい る事を示唆するものである。さらにこの酵素は、DNA複製の開始段階で、複製の 分岐点に差し掛かるとき、その選択を調節する上で重要な役割をする事が最近示さ れている (Dasgupta ら、1987)。

また、Johnson ら(1986)は、ヒト免疫不全ウイルスを含むレトロウイルス由来 の逆転写酵素のアミノ酸配列を、コンピュータを用いて解析することによりRNA 依存性ポリメラーゼ領域の下流部分に RNase H と相同性の高いセグメントが存在 していることを示した。ウイルス由来の RNase H は、ウイルス遺伝子のRNAの 切断に重要な役割を担うものと考えられている。事実、DNA/RNA ハイブリッド2重 鎖からRNA鎖のみを RNase H によって除去することにより、 マイナス鎖RNA は、プラス鎖RNAを合成する際の鋳型と成り得ることが発見された(Varmus ら、 1988)。

この論文では、 X線結晶構造解析によって 1.48Å分解能で決定された、大腸菌 RNase H の立体構造(Katayanagi ら、1990 & 1992)について述べる。さらに、酵 素活性に必須なMg²⁺結合部位を同定し、この酵素がDNA/RNA ハイブリッド2重鎖と 特異的に結合する認識部位も決定した。NMRによる最近の研究によって、我々が 想定した RNase H と DNA/RNA ハイブリッド2重鎖の結合様式が正しいことが裏付 けられ、DNA/RNA ハイブリッド2重鎖と酵素の複合体のモデルを構築することがで きた。また、この酵素の活性アミノ酸残基を置換した3つの変異体についても、X 線結晶構造解析を行い、分子構造と機能の関係について考察した。

なお 我々とは独立に、Yang ら(1990)も、異なった位相決定法にもとづいて決 定した 2.0Å分解能における RNase H の立体構造を報告したが、 両者の立体構造 は、本質的には同一であった。

本論

第1章 大腸菌 RNase H の立体構造

第1節 全体的な分子構造

RNase H 分子全体の模式図を図1-1に、またこの酵素の1次配列を図1-2に 示す。RNase H 分子は、155 個のアミノ酸残基からなり、1つのサブユニットで構 成される。この分子のサイズは 50×45×40 Åで、不規則な楕円体である。この分 子のドメイン構造を調べるため、C α 原子間のディスタンスマップ (Nishikawaら、 1972)を計算した(図2)。その結果、残基番号で 80 番から100 番目の部分には 隣接した構造がなく、独立のドメインを構成していることがわかった。この部分は ヘリックス α III からそれに続く約10 残基の長さのループで、 当初 minor domain と呼んだが(Katayanagi ら、1990)、 逆転写酵素のドメイン構造の議論の際に紛ら わしいため、basic protrusion (塩基性の突出部)と呼ぶことにした(Katayanagi ら、1992)。 なお、Yang ら(1990)は、この領域を handle region (把手領域)と している。この領域は最近、Davies らにより解析されたH I V-1 RNase H ドメ インには欠損している部位でこの酵素にユニークな構造であるため α IIIを除いた部 分について図3に示した。

主ドメインは、4つのαヘリックス(α I、α II、α IV、α V)、および 1枚の 大きな β シートから成る。この β シートは、N末端側の3本の逆平行 β 鎖(β A、 β B、 β C)と2本の平行 β 鎖(β D, β E)より構成されている。この β シート は、他の蛋白質の β シート構造と同じようにねじれている。RNase H 分子のコアと なるこの β シートは、2本の α へリックス α I と α V に両側から挟まれ、 α V 側の 半分は表面に露出している。 α V のC 末端側半分は β シートと疎水結合しているが、

- 3 -



図<u>1-1</u> 大腸菌 RNase H の分子全体の模式図。 ヘリックスαⅢからそれに続く ループが突きだした形になっている(塩基性の突出部)。

.

- 4 -

.



90100110120WIHNWKKRGWKTADKKPVKNVDLWQRLDAALGQHQIKWEW $\vdash \alpha III - d$ $\vdash \alpha IV - d$ $\vdash \beta E$

図1-2 大腸菌 RNase H のアミノ酸配列(Kanaya & Crouch、1983)にX線構造 解析で決定した2次構造を重ねたもの。*印は逆転写酵素のアミノ酸配列で完全に 保存されているもの(表5)を示す。残基番号で80から100番の間には塩基性のア ミノ酸残基が多く分布している。また残基番号で11から23番まではグリシンが多く 分布し、柔軟な構造であることが予想される(第3章第4節)。

- 5 -



図2 大腸菌 RNase H のC & 原子の座標をもとに計算したディスタンスマップ。縦 と横の残基間の関係のうち、10人以下の距離のものは口印で、10人以上20人以下の 距離にあるものは×印で示した。残基番号が80から100 番の間には20人以下の位置 関係がほとんど無く、独立した領域であることが明瞭に示される。



図3 塩基性の突出部のループの立体構造(αIIIの部分は図10参照)。この部分は、 HIV-1RNase H ドメインにはない構造である。この鋭いターンははしご状の水 素結合で安定化している。このうち、Gly 89、Trp 90、Lys 95、Asn 100 の主鎖の 二面角は左巻きへリックスに近い角度をとっている。

- 7 -

N末端側の数残基は β シート上の残基と水素結合を形成し、あるいは静電的相互作 用をしている。 α I とこの β シートは、顕著な疎水結合を形成している。この β シ ートの疎水結合面は図4に示すように、 α II や α III と向かい合うことにより、 α I や α IV のN末端側とともに、大きな疎水結合コアの中心を構成している。その中で、 ヘリックス α I は、これを囲む β シート(α I 側)、 α I、 α II、 α III、 α III、 α IVで形 成される大きな疎水結合の中心を構成している。 α I と α IV は約25°の傾きを保ち ながら平行に並んでいるが(図5)、この2本のヘリックス間の疎水結合を見ると、 5つのロイシン、2つのアラニン、そして1つのトリプトファンが3つ組になって いる。すなわちLeu 56(α I)、Ala 110(α IV)とLeu 111(α IV)の組、Ala 52(α I)、 I1e53(α I)とLeu 107(α IV)の組、Leu 49(α I)、Leu 103(α IV)とTrp 104(α IV) の組が、顕著な疎水結合を形成している。この構造は、遺伝子の発現に関与するた くさんの転写調節蛋白質中に見いだされる、平行型のロイシンジッパー・モチーフ によく似ていて興味深い。ちなみに Leu 49 と Leu 56 は7 残基分、また Leu 103 と Leu 111 は8 残基分離れている。

RNase H の2次構造の統計値を表1に示す。この分子の主鎖のフォールディング は、厳密には他のどの蛋白質のものとも類似していないが、大局的には Levitt と Chochia (1976) らの分類に従って、 $\alpha + \beta$ 型構造に属すると考えられる。 $\alpha \parallel n$ ら $\alpha \parallel l$ にかけての部分および $\alpha \parallel l$ から $\alpha \parallel v$ にかけての部分について、*cro* リプレッ サーなどで得られている、DNA結合モチーフのヘリックス・ターン・ヘリックス 部分と構造比較したところ、r.m.s.d.が4Å以上と大きく、類似のモチーフとはい えない。したがって、ヘリックス・ターン・ヘリックス構造はこの分子には存在し ないことが、結論される。

副ドメインは、αヘリックス(αΙΙΙ)と2つのターンを含む約10残基のループか ら成っている。RNase H のディスタンスマップからは、αΙΙΙとそれに続くループは、 他の部分からはかけ離れた、小さなドメインを形成していることがわかる。αΙΙΙは



۱ 9

> 図4 RNase H の疎水結合コア。ペプチド主鎖を細線で示し、疎水性のアミノ酸残 基(Gly,Ala,Leu,Ile,Val,Phe,Trp,Met)のみを、太線で表示した。ヘリックスα I を中心とする大きな疎水結合コアとα VのN末端とβシートの下側でつくる小さな 疎水結合コアがある。ヘリックスα IIとα III はこの大きな疎水結合コアを包み込む ように折れ曲がっている。C末端の Val 153 と Pro 17 ならびに Val 155 と Pro 97は疎水的相互作用をしていると考えられ、これは隣接分子のβCと相互作用して 安定化しているN末端とは対照的である(表4)。



図5 ヘリックスαIとαⅣの疎水的相互作用。このロイシン・ジッパー様の2本 のヘリックスは約 25°の傾きを持ちながら平行に走っている。予想される核酸認識 部位(第3章第4節)はこれらのN末端側近傍である。

- 10 -

	Residues	< \$ >	<ψ>	< w>
αΙ	43 - 58	-65(5)	-40(10)	180(1)
αII	71 - 79	-64(6)	-44(8)	-179(1)
αIII	81 - 89	-73(14)	-28(17)	-179(2)
<i>α</i> IV	100 - 112	-63(8)	-41(5)	179(2)
αν	127 - 141	-62(5)	-42(6)	180(2)
helix overall	L	-65(8)	-40(10)	180(2)
mean helix fi	com data base*	-62(7)	-41(7)	179(3)
βΑ	5 - 14	-138(13)	143(9)	179(2)
βΒ	19 - 28	-132(16)	141(26)	176(4)
βС	31 - 39	-129(21)	140(18)	179(1)
βD	63 – 69	-121(18)	125(13)	180(4)
βΕ	115 - 121	-107(13)	119(12)	-179(3)
βΒ&βC		-131(17)	140(22)	178(3)
ideal anti	i parallel β †	-139	135	178
βD&βΕ		-115(17)	122(13)	180(3)
ideal para	allel β †	-119	113	180
310 I	55 - 59	-69(8)	-27(15)	180(1)
310 II	110 - 114	-74(14)	-17(13)	182(3)
310 III	138 - 143	-67(10)	-29(13)	181(1)
mean 310 from	a data base*	-71	-18	

Values in parentheses indicate standard deviations. * D.J.Barlow & J.M.Thornton J. Mol. Biol. 201, 601.(1988). † G.N.Ramachandran, C.Ramakrishnan, V.Sasisekharan, J. Mol. Biol. 7, 95.(1963).

<u>表1</u> RNase H の持つ2次構造の2面角に対する統計的評価。それぞれの数字はその2次構造内の平均値で、()内の数字はそれぞれの標準偏差を示す。

.

小さなヘリックスαIIを経由してβDと接続している。5本鎖からなるβシートと 2本のヘリックスの組(αIIとαIII)は、αIとαIVを底とする大きな裂け目を形 成しているが、この裂け目が触媒活性に何らかの形で関与しているか否かは、現在 のところわからない。ただし、後に述べるように、この裂け目は酵素の DNA/RNAハ イブリッド認識部位とは反対側に存在している。

主ドメインでは、βシートは2つのヘリックスαIとαVに挟まれていて、N末 端とその近傍のループ部分を除いてはβシートは疎水的環境にある。なお、このβ シートは他の蛋白質で見られるのと同様に、左方向にねじれている。

構造精密化の際、主鎖原子の二面角の(タ、ψ)には束縛条件はかけていない。 このようにして得られた二面角の値を用いて Ramachandran plot (Ramachandran、 1963)を計算したものを図6に示す。ほとんどのø角、ψ角は許容範囲に収まって いる様子がわかる。なお、他のアミノ酸残基より許容範囲の広い13個のグリシンは *印で示してある。この図の中で、左巻きヘリックスに対応する二面角をとってい るものは、Trp 90、Lys 95 と Asn 100 であり、それらの値は、それぞれ、(67°、 40°)、(64°、27°)そして(61°、31°)である。また、その下部に許容外領域としてプ ロットされる Arg 29(70°、120°)は、続く Gly 30(-115°、5°)と組み合わせると、 Venkatachalam(1968)の論文で指摘された、典型的なⅡ'型のβターンで 説明さ れる(後出の表3参照)。このように、構造精密化後のモデルでは155 個すべての アミノ酸残基が 合理的なコンフォメーションをとっていることが示された。RNase H には、5つのプロリン残基があるが、そのうち Pro 17 だけが、シス型をとって いることがわかった。また,他の残基においては、ω角は すべて180°近辺におさま っている (r.m.s.d.は 2.8°)。Kabsch & Sander (1983) の定義による B シート部 分の�角、�角の平均値は、 平行型、逆平行型が、それぞれ −115°と 122°、およ び -131°と140°であった。またαヘリックス部分では、それぞれ -63°と -41°であ った。



図6 RNase H の精密化後の最終構造の Ramachandran plot。左巻きヘリックスの 部分にある4残基は、Gly 89、Trp 90、Lys 95、Asn 100 でいずれも図3で示され る塩基性突出部のループ部分に属するものである。 また (70°、120°)の二面角を とるのは Arg 29 で、II ' 型のβターンである。*印はグリシン残基を示す。

図7にプロリン残基をのぞいた全アミノ酸残基の x_1 角の分布を示す。これらは 顕著に $x_1 = (-60^\circ, 60^\circ, 180^\circ)$ 付近に集中しているが、これは立体化学的に安定 な構造を取りやすい二面角のanti-periplanar(180°)あるいは \pm syn-clinal(\pm 60°) である。分子全体の x_1 の平均値はそれぞれ(-67°, 64°, -173°)また標準偏差は (13°, 6°, 16°)であった。syn-periplanar(0°)や \pm anti-clinal(\pm 120°)は完 全に避けられており、精密化後のモデルには、エネルギー的に好ましくないコンフ ォメーションが存在していないことが示された。

第2節 分子内水素結合

表2に、RNase H 分子内の水素結合の一覧を示す。 大腸菌 RNase H の場合、ア ルギニン側鎖のつくる水素結合が構造の安定化に寄与している点が目だっている。

まず Arg 46 側鎖の電子密度を図8-1に示す。この残基はヘリックスαI上に 存在するが、唯一、分子内に埋もれているアルギニンである。このグアニジノ基は 分子のほぼ中央部に存在し、αN上の Asn 100 、Asp 102 および 最後のループの Asp 148 と、3方向すべてに水素結合を形成している。したがって、αI、αNお よびC末端ループの3つのドメインを結び付ける働きをもつ。 Arg 46 側鎖原子の 温度因子も 10^{A^2} 前後と低く、多くの水素結合で支えられた安定した構造を反映し ている。この分子中央に位置する水素結合のネットワークは、分子全体のコンフォ メーションを安定化しているものと考えられる。

また β B上に存在する Arg 27 (図8-2)は、これを囲んでいる3つのグルタ ミン酸残基、 β A上の Glu 6、 β B上の Glu 32 及び α V上の Glu 130 と水素結 合をつくっている。これらの相互作用は3つの β 鎖を更に固くむすびつけていると ともに α Vを固定している。 β B上の Glu 32 は、さらに α V上の Arg 132 と水 素結合をつくっている。また、 β D上の Ser 68 は、 α V上の Asn 130 と、また β E上の Glu119 は α Vアミノ基側の His 127 と水素結合し、 上の相互作用をさ



<u>図7</u> Gly、Ala、Pro を除いた全残基の側鎖の21角の分布。

- 15 -

Residue	Atom	Residue	Atom	Distance(A)
Gln 4	Οε 2	Lys 117	Nζ	3.15
Glu 6	Οε 2	Arg 27	Nε	3.03
Glu 6	Οε 2	Arg 27	N7 2	3.07
Thr 9	0	Thr 69	0γ1	3.08
Thr 9	0γ1	Thr 69	0γ1	3.15
Asp 10	Οδ 2	Gly 11	N	2.63
Asp 10	Οδ 2	Gly 11	0	3.23
Cys 13	Sγ	Gly 15	0	3.28
Cys 13	Sγ	Gly 18	0	3.28
Cys 13	Sγ	Thr 42	0	3.38
Cys 13	N	Asn 44	08 1	3.28
Gly 22	0	Ser 36	0γ	2.92
Arg 27	N7 1	Glu 32	Οε 2	3.00
Arg 27	Nŋ 1	Glu 129	Οε 1	3.15
Arg 27	N7 2	Glu 129	Οε 2	2.57
Tyr 28	0η	Lys 60	N	3.13
Tyr 28	07	Glu 61	0ε 1	3.38
Tyr 28	0η	Glu 61	Οε 2	2.53
Glu 32	Οε 1	Arg 132	Ny 2	2.71
Thr 42	0γ1	Asp 148	08 1	2.53
Asn 44	Nδ 2	Glu 48	Οε 2	2.36 *
Arg 46	N7 1	Asn 100	0δ 1	3.02
Arg 46	N7 1	Asp 102	$0\delta 1$	2.89
Arg 46	N7 2	Asp 102	Οδ 2	2.88
Arg 46	N7 2	Asp 148	Οδ 2	2.63
Arg 46	Nε	Asp 148	0δ1	2.52
Arg 46	Nε	Asp 148	Οδ 2	3.28
Glu 48	0ε 1	Ser 71	Ογ	2.60
Glu 57	0ε 1	Arg 106	N7 2	2.68
Glu 57	Οε 2	Arg 106	Nε	2.87
Glu 61	0	Cys 63	Sγ	3.35
Glu 61	0	His 114	Νε 2	3.18
His 62	Nδ 1	Gln 113	0	2.59

<u>表2</u> 大腸菌 RNase H の 分子内水素結合の一覧(主鎖どうしの水素結合について は図12を参照)。残基番号順に並べてある。水分子は含めず、水素結合距離が 3.4 Å以内のもののみ記した。なお*印を付したもの(2つ)については、水素結合距 離としてはやや短すぎるが、これは側鎖原子の局部的なディスオーダーのためと考 えられる。事実、Asn 44 側鎖のC₇、O δ_1 、N δ_2 や、Asp 94 のO δ_1 、O δ_2 は、 周囲に存在する原子より高い温度因子を持っている。

Ser	68	0γ	Glu 1	19 O	ε 1	2.88
Ser	68	Ογ	Asn 1	30 N	δ 2	3.10
Thr	69	0γ 1	Asp 7	0 N	I	2.96
Thr	69	0γ 1	Ser 7	1 N	1	3.20
Ser	71	Ογ	Val 7	4 N	T	2.90
Tyr	73	07	Asn 1	00 N	δ 2	2.98
Tyr	73	07	Trp 1	04 N	ίε 1	3.07
Arg	75	0	Thr 7	90	y 1	2.66
Gln	76	Νε 2	Gln 8	0 O	ε 1	2.93
Thr	79	0γ 1	Gln 8	0 N	í	3.24
Gln	80	0	Asn 8	4 N	δ 2	3.15
Lys	86	NS	Asp 1	08 O	δ 2	3.40
Thr	92	0γ 1	Asp 9	4 O	δ 2	2.30 *
Thr	92	0γ 1	Lys 9	6 N	I	3.12
Thr	92	0γ 1	Lys 9	60)	3.25
Ala	110	0	His 1	14 N	18 1	3.14
Glu	119	0ε 1	His 1	27 N	ίε 2	3.10
His	127	N	Asn 1	30 O	δ 1	2.85
Asp	134	0δ 1	Arg 1	38 N	ε	3.16
Asp	148	Οδ 2	Gly 1	50 N	ſ	3.34

٠

表2 (前ページの続き)

٠

.



図8-1 α I に存在する Arg 46 近傍の電子密度。等高線は1 σ の高さである。 破線は水素結合を意味し、分子中央部のヘリックス α I 上の Arg 46 が、 α IV 上の Asn 100 や Asp 102 およびC末端近傍のループにある Asp 148 と3方向へ等方的 に水素結合を形成し、分子全体のコンフォメーションを安定化している様子がわか る。Arg 46 側鎖の温度因子はどの原子も10Å²前後と小さい。

- 18 -



図8-2 β Bに存在する Arg 27 近傍の電子密度。この残基は3つのグルタミン 酸、すなわち、 β A上の Glu 6、 β C上の Glu 32、 α V上の Glu 129 に囲まれて 相互作用を受けている。図8-1の Arg 46 に比べて、等方的な水素結合ではない ため電子密度もそれよりきれいではない。この近傍では、 β シートと α VのN末端 側の水素結合として、この図以外に Ser 68-Glu 119、Ser 68-Asn130 がある。

- 19 -

らに強めるように働いている。

Arg 138 の側鎖グアジニノ基は、Asp 134 の側鎖を外側に引き出すように水素結 合している。Asp 134 は、26種類のレトロウイルス由来の逆転写酵素とその関連蛋 白質すべてに共通して保存されている、4つの酸性アミノ酸残基(第1章第9節) のうち、唯一 RNase H 活性に必須とされないものであり、またMg²⁺結合位置から 5.0 Åと最も遠く離れている。しかしながら、大腸菌 RNase H で Arg 138 に相当 する残基は、レトロウイルス由来の逆転写酵素には保存されておらず、この1本の 水素結合は、大腸菌RNase H にたまたま生じた偶然とも取れる。

第6節②項でも指摘するが、αIIIは、主鎖原子の温度因子の高いヘリックスであ り、これを固定しているのが、Lys 86 と Asp 108 の水素結合1本である。これに 比べ、αIIは Ser 71 が Gln 48 と、 また Tyr 73 が Trp 104 と(図8-3)、 というように2カ所で水素結合をしておりαIIIに比べると安定なものと考えられる。 Mg²⁺結合部位の近くにあるこのチロシンは、DNase I では核酸とスタックしている 部分に相当していると思われる部分であるが、RNase H の基質なしのものでは構造 安定化として働いている。安定化に寄与しているチロシンは、この他に Tyr 28 が Glu 61 と2本の水素結合をして、2つのループ部分を近づけている。

なお、α II の溶媒側に3 つのグルタミン残基 (Gln 72, Gln 76, Gln 80) が1 直 線に並んでいるのは興味深い。 このうち、Gln 76 と Gln 80 は側鎖間で水素結合 している。

活性に必須な Glu 48 のカルボキシル基は Ser 71 と Asn 44 に水素結合し固定 されている。更に Asn 44 のもう1方は Cys 13 のカルボニル酸素と水素結合して いる。これに比べて Asp 10、Asp 70 の必須残基は、このような固定を受けていな いことは興味深い。一般に、このMg²⁺結合部位近辺は水素結合で固められている。



図8-3 α II 上の Tyr 73 とα N上の Trp 104 の相互作用。Tyr 73 はまた、水 分子 (WAT205、B=30.05Å²)を介して Lys 99 主鎖と水素結合している。またα N はC末端側では図9に示されるように His 114 と水素結合し、 両端が支えられて いる。

第3節 トリプトファンの分布

前述したように、βシートと2本のヘリックスの組(αⅡとαⅢ)は、分子中央 部に位置するαIとαNを底とするように、大きな裂け目を形成している。RNase H 分子がもつ6つのトリプトファン残基すべてが、この裂け目の両側に集中してい る(後出の図17参照)。この裂け目は疎水結合コアの半分を占めている(図4)。 このトリプトファン残基のクラスターは、クレフトの底にある、ヘリックスαΙと αⅣの疎水性側と相互作用している。すなわち、いずれのトリプトファン残基も、 6員環が疎水結合コアであるαΙの方向を向いている。また、2組のトリプトファ ンが ずれながら重なりあい、スタックしている(Trp 81 と Trp 85 の組、および Trp 118 と Trp 120 の組)。特に Trp 81 と Trp 85 はx2角が理想値より60°ずれ ており、エネルギー的には不安定であるが、これはαⅡとαⅢの折れ曲がりにより、 Van der Waals 空間が狭まっているためと考えられる。この2つのトリプトファン の組および WAT238(後出の図13参照)を介した Gln 80 のO e1と Trp 81 のN e1と の水素結合などが、ヘリックスαⅡとαⅢの折れ曲がりの方向づけをしているもの と考えられる。αⅣのN末端にある Trp 104は、分子内部に埋もれている唯一のト リプトファン残基であるが、 αⅡの Tyr 73と 側鎖間で水素結合を形成している (図8-3)。この水素結合は、前に述べた疎水結合の他、ヘリックスαⅣの Ala 110 と水素結合している His 114 とともに、 ヘリックスαⅣを両側から構造的に 安定化している。

第4節 システインの分布

Bekower ら(1973)により、RNase H は最大の活性を得るためには、スルフヒド リル試薬を要求し、またシステイン残基の化学修飾の典型的試薬であるN-エチル マレイミド(NEM)により阻害されることが報告されており、3つのシステイン 残基の分布を考えることは興味深い。このうち Cys 13 と Cys 63 は、それぞれ異 なったループで溶媒部に露出しているが、結晶中においては、 Cys 63 は隣接分子 に近接しているため、 Cys 13 の方がこれらの試薬の影響を受けやすい (図1-1))。最後のCys 133 であるが、これはβBに直面しているαV上にあるため周囲が 完全に包囲されており、直接の化学修飾は受けにくいと考えられる。これらの環境 の特徴は、RNase H 結晶を水銀化合物に浸漬したときに、システイン残基近傍に現 れるHg²⁺サイトの占有率との結果(後出の表9参照)とよく一致する。化学修飾と 組み合わせた部位特異的変異法による最近の研究では(Kanaya ら、 1990a)、これ らのシステインの位置に対応した結果が得られているし、NEMによる酵素の不活 性化は、チオ基のブロックによるものではなく、立体障害に起因するコンフォメー ション変化によるものと考えられる。

第5節 ヒスチジンの分布

すでに立体構造解析されているヌクレアーゼでは、ヒスチジン残基がすべて活性 中心に関与している。たとえば、RNase A (Richards & Wyckoff, 1971; Carlisle ら, 1974), RNase T1 (Arni ら, 1988; Sugio ら, 1988), RNase U2 (Matsuzakiら、 私信), barnase (Maugen ら, 1982), そして DNase I (Suck & Oefner, 1986; Suck ら, 1988) である。したがって、ヒスチジンの位置を議論しておくことは、きわめ て重要である。RNase H 分子では His 62 は β Dの入り口、 His 83 は塩基性へリ ックス α III 上に存在し、いずれも溶媒表面によく露出している。His 114 は、リバ ース・ターンを構成し、His 127 は α V の N 末端側に位置する。この2 つは、Mg²⁺ 結合部位に最も近く、いずれも分子表面に露出しており、 基質の DNA/RNA ハイブ リッドとの結合に関係するとの推測も可能であるが、部位特異的変異法の結果では、 どちらのヒスチジンも RNase H の 活性には必須ではない。しかしながら His 124 の変異体では k cat の値が著しく減少している。これら4つのヒスチジンに対し、 His 114 は、 β Eの入り口にあり、5つあるヒスチジンのうち唯一、分子内に埋も

- 23 -

れている。これらの環境は、5つともすべてNMRによる pH 滴定の結果(Yoshida ら、私信)とよく一致している。また、野生型に対する非活性を見るために、5つ のヒスチジンをそれぞれ1つずつアラニンに変えたときの結果とも一致している (Kanaya ら、1991a)。

His 62 のN δ_1 原子は、Gln 113 主鎖の カルボニル酸素と水素結合している。ま たHis 127 も Glu 119 と側鎖間で水素結合している。図9に示すように、His 114 側鎖のイミダゾール環にある2個の窒素原子(N δ_1 とN ε_2)は、 それぞれ 主鎖の カルボニル酸素(Glu 61 のOと Ala 110 のO)と水素結合している。このプロト ン化しているという事実は、 pH が 9.0 という RNase H の結晶化条件と一見矛盾 する。この解釈の1つとして、His 114 は、局部的に負に荷電した環境にあるため と思われる。すなわち、このイミダゾール基は異常に高いpKa 値をもっていると考 えられる。事実、His 114 は、 Glu 57、Glu 61、Glu 64 と、3つもの酸性アミノ 酸のカルボキシル基に囲まれており、また Ala 110 と Glu 61の主鎖カルボニル酸 素に囲まれている。もう1つの解釈としては、中性のイミダゾール環が分子ごとに 別々のサイトでプロトン化し、結晶中では同じ様な占有率としてN δ_1 とN ε_2 が存在 していると考えるものである。

さらにつけ加えて、His 114 は、Cys 63 とジスルフィド結合を形成し得る距離に 存在する。この114 番目の位置にシステイン残基を導入した結果、コンフォメーシ ョンには大きな変化もなく Cys 63 とジスルフィド結合が形成され、またこの変異 体には RNase H 活性が残っていることがわかった(Kanaya ら、私信)。

第6節 二次構造

① ジョイント予測法、円2色偏光法の結果との比較。

決定された立体構造からは、αヘリックスに 42 %、βシートに 28 %のアミノ 酸残基が存在することがわかった。これらの二次構造の含有率は、円二色偏光スペ

- 24 -



図9 His 114 側鎖の 2Fo-Fc 電子密度図。 イミダゾール環の2つの窒素原子 とも、両側の主鎖カルボニル基(Glu 61 と Leu 111) と水素結合している。 クトルの実験から得た比率(α ヘリックスに 28 %、 β シートに 41 %; Kanayaら、 1990b)とは一致しなかった。Joint prediction 法 (Nishikawa & Ooi, 1986) に よる二次構造予測では、 α ヘリックスの長さや その位置は、比較的 的中したが、 β シートはあまり的中しなかった。

② ヘリックス

RNase H 分子には2つのドメインにかけて5つのαヘリックスが存在する。つま り、塩基性突出部にはαШが、主ドメインにはほかの4つのヘリックスが存在する。 これらのαヘリックスは、 α I が 43 から 59 まで、 α II が 71 から 79 まで、 αШが 81 から 89 まで、α N が 100 から 112 まで、α V が 127 から 141 まで と定義される。このほか、球状蛋白質の表面近くによくみられる 3_{10} ヘリックスは、 3つのαヘリックスの出口(α I、α N、α V)にみられる。

α II とα III は約 35°の角度で折れ曲がっているが(図10)、この接合部はプロリ ンを含まない数残基であり、データベース解析(Kabsch & Sander, 1983) によりほ かの蛋白質の立体構造と比較したが、類似の折れ曲がりは見いだせなかった。この 折れ曲がりは、前項で述べた疎水結合コアの中心をなすα I とα IVを囲むように形 作られており、内側が疎水性、外側が親水性なことから、このコアの形成に寄与し ているものと考えられる。また、ヘリックスα I は、ねじれているβシートに沿う ように疎水結合しているため、約5°とわずかながら曲がっている(ヘリックス中央 の Glu 48 のNと Ala 52 の〇の距離が 3.8Åで水素結合には長すぎる)。

各原子の温度因子や変位のパラメータから、構造上の安定性や柔軟性がわかる。 図11に、主鎖原子の等方性の温度因子を各残基ごとに平均し、残基番号でプロット したものを示した。興味深いことに、α IIIの温度因子が 25Å²を越えており、ほか のヘリックスに比べ非常に高い値を示した。そのほかの、ヘリックスやβシートの 二次構造部分では温度因子は明らかに小さい。α IIIには塩基性アミノ酸残基が多く



図10 ヘリックスαIIとαIIIの約 35°の折れ曲がり。この折れ曲がりの内側は疎水 性残基が並び、疎水性コアの形成に参加している。一方、外側には親水性残基が並 び、溶媒領域を向いており、またその1部は核酸結合部位と予想される。Thr 79、 Asn 84 の側鎖や、水分子(WAT188、B=25.32Å²)が この折れ曲がりを補助する水 素結合を形成している。



図11 構造精密化後に得られた RNase H 主鎖原子 (C α、C、N、O)の 等方性 温度因子を、各残基ごとに平均し、残基番号でプロットしたもの。二次構造部分の 温度因子は小さいが、塩基性突出部のヘリックス α Ⅲの部分のみが例外的に大きい。

- 28 -

含まれることから、核酸2重鎖との結合に何らかの役割を果たすものと考えられる。

③ βシート

この分子には5本のβ鎖より成る1枚のβシートがあり、主ドメインのコアになっている。残基番号で5から14、19から28、31から39、63から69、115から121が、それぞれβAからβEと定義され、3つの平行鎖(βA、βB、βC)と2本の逆平行鎖(βD、βE)になっている。このβシートの水素結合のダイヤ グラムを図12に示す。このβシートは通常の蛋白質に含まれるβシートと同じように左方向にねじれている。

この β シートの特徴は、これをはさんでいる2つの α ヘリックス(α I、 α V) とシートの両側で相互作用していることである。まず、 α Iや α Nサイドの例とし ては、Val 5(β A)-Leu 59(α I)、Leu 26(β B)-leu 59(α I)、Val 65(β D) -Leu 59(α I)、Ile 7(β A)-Leu 56(α I)、Ile 7(β A)-Ala 110(α IV)、 Ile 7(β A)-Leu 111(α IV)、Leu 67(β D)-Ala 55(α I)、Trp 118(β E)-Ala 55(α I) などのような疎水結合があげられる。その裏側では、 β シートは α Vの N末端と水素結合を形成したり、静電的な相互作用をしている。ただし α VのN末 端近傍部は溶媒領域に露出している。また α VのC末端側とはもうひとつの小さな 疎水結合コアを形成している。この際の疎水性残基は、 α V側が Ala 137, Ala 140, Ala 141 で、 β シート側が Gly 20, Gly 21, Gly 38 である。

④ リバース・ターン

球状蛋白質においては、ターンや折れ曲がりは2次構造の保持のために重要な要素である。βターンは一般に4つの連続した残基よりなり、その中の1番目と4番目の残基が主鎖間で水素結合しており、主鎖の方向を180°反転させる。RNase H 分子のβターンは全部で5つ、またβターンの連続した 310ヘリックスは3つ存在し

- 29 -



- 30 -

ており、それらを表3に示す。なお、これらの分類は Crawford ら(1973)の方法 によった。また、ターンの種類は2番目と3番目の残基の2面角(φ、ψ)の組み 合わせで定義される。

RNase H 分子に見られるすべてのターンはもっぱら分子の表面に存在する。興味 深いことに、塩基性突出部のループ部分には2つの β ターン(シーケンスでTAD KとVKNV)が存在し、2つのリジン(Lys 95 と Lys 99)が効率的に溶媒部に 突出している。同様な状況が β Dと β Cを結ぶターン(YRGR, Arg 29 と Arg 31)にも見られる。

α V入り口のターン付近には、 His 124 と His 127 の2つのヒスチジン残基が あり、変異体の研究からは、必須ではないが活性に大きな影響を与えていることが わかっている。特に、His 124 は主鎖のコンフォメーション変化を考えれば、Mg²⁺ 結合部位に十分近づき得るため、このターンの持つ意味は大きい。さらにC末端の ループにも3₁₀へリックスとβターン(DTGY)が存在している。後述するよう に、これら2つの領域は核酸と相互作用する面に含まれている。

また C末端付近に存在するターンは、好熱菌由来の RNase H (Itaya & Kondo、 1991)では、プロリンが4つも含まれている部分であり、このターンの多い部分が 2つの酵素間で相同な役割を持っているのかも知れない。

第7節 水分子の分布

精密化の結果、分子全体で、225 個の水分子が同定された。これらの水分子の分 布を図13に示す。ここで温度因子が 30Å²以下の低いものについては黒丸で表示し てある。同定されたすべての水分子のうち 134 個が RNase H 分子と直接に水素結 合しており、これらが第1層の水分子であることを示している。これらの水分子の うち、温度因子が 30Å²以下の低いものは構造の安定化に寄与しているものが多い。 他方で、蛋白質分子より 3.4Å以上離れている水分子は、より高い温度因子を持っ

- 31 -
| Hydrogen | bond | NO(Å) | ø 2 | ψ2 | ø 3 | ψ3 | Type § | Sequence | C α 1-C α 4 (Å) |
|------------------------|--------------|---------|------------|------------|-------------|---------|--------------|----------|------------------|
| Reverse | Turns | <u></u> | | | | | | | |
| 28 0 - | 31 N
95 N | 3.04 | 70 | -120 | -115 | 5
16 | RN-II' | YRGR | 5.54
5.14 |
| 92 0 - 1
98 0 - 1 | 01 N | 3.19 | -67 | 134 | 61
102 | 31 | RN-II | VKNV | 5.29 |
| 122 0 - 1
148 0 - 1 | 25 N
51 N | 3.19 | -53
-55 | -33
-49 | -103
-67 | -10 | RI-I
RN-I | DTGY | 5.69 |
| 310 Heli | ces | | | | | | | | |
| 31 0 – I | | <u></u> | | | | | | | |
| 55 O - | 58 N | 3.23 | -65 | -46 | -63 | -11 | RN-III | ALEA | 5.36 |
| 56 O - | 59 N | 2.98 | -63 | -11 | -81 | -23 | RN-I | LEAL | 5.60 |
| 31 o -II | | | | | | | | | |
| 110 0 - 1 | 13 N | 3.31 | -61 | -33 | -67 | -18 | RN-III | ALGQ | 5.72 |
| 111 0 - 1 | 14 N | 3.21 | -67 | -18 | -93 | -1 | RT-I | LGQH | 5.63 |
| 310-III | | | | | | | | | |
| 138 0 - 1 | 41 N | 3.25 | -59 | -44 | -62 | -40 | RT-III | RAAA | 5.30 |
| 139 0 - 1 | 42 N | 3.05 | -62 | -40 | -61 | -22 | RT-III | AAAM | 5.28 |
| 140 0 - 1 | 43 N | 2.95 | -61 | -22 | -84 | -11 | RT-I | ΑΑΜΝ | 5.35 |

§ Classified according to Crawford et al. (1973).

RT, Reverse Turn; RN, Near Reverse Turn; and O, Open Turn.

<u>表3</u> リバース・ターンと3₁₀ヘリックスの一覧。

- 32 -



図13 水分子の分布。温度因子が 30Å²より低いものについては、黒丸で表示し、 番号も添えてある。それ以外は+印で示した。この黒丸は、核酸認識部位(第3章 第4節)近辺に多く集まっている。

- 33 -

ている。

蛋白質分子内の2つの原子(とくに主鎖原子)を水素結合で結ぶ水分子のほとん どは、温度因子が小さい。これらの水分子は RNase H 分子の フォールディングを 安定化している。たとえば 780-WAT188-83N は、折れ曲がりつつ連続している2 つのヘリックス、α II とα IIIの間を固定している。同様に、360₇-WAT157-1360、 570 ε₂-WAT159-1060、100 ε₂-WAT166-700 ε₁、410-WAT173-148N、38N-WAT182 -1450₇₁、410-WAT183-1480、 730₇-WAT205-99N、 430₇₁-WAT206-100N ε₂、 90₇₁-WAT211-480 などは、2次構造やループを水素結合で結び付ける中間体とし ての役割をなしている。 WAT192 と WAT193 のペアの水分子は、α IVのC末端側付 近の3₁₀ヘリックスを安定化させている。

また、4つのチロシンに対して、その水酸基に水素結合している水分子が見いだ されている。つまり、 219WAT-Tyr 22、 199WAT-Tyr 39、 205WAT-Tyr 73、および 231WAT-Tyr 151 である。

後の章で詳細に述べる RNase H 分子の DNA/RNA ハイブリッド2重鎖との接触面 には、より多くの水分子が分布している。温度因子が 30Å²以下の水分子、たとえ ば WAT166、169、175、180、および 227 などは、 4つの重要な酸性アミノ酸残基 Asp 10、Glu 48 Asp 70 と Asp134 で形成される Mg²⁺結合のためのポケット近傍 に集中している。これらの水分子は、自身の持つ大きな双極子定数により、このポ ケットの負電荷を中和するものと思われる。C末端にあるループとα I のN末端側 で囲まれる領域にも多くの水分子が存在しており、このC末端ループのフォールデ ィングを助けているものと考えられる。

とくに、 Glu 57 付近をはじめとして、α I とα IVの溶媒側には、α V の溶媒側 より多くの水分子が見られる。これに対し、トリプトファンのクラスターで形成さ れる開列部を含め、この付近の疎水結合コアには、水分子はほとんど見られない。 第8節 結晶のパッキング

RNase H 分子の、空間群 P212121の斜方格子におけるパッキングをZ軸方向から 眺めたものを図14に示す。それぞれの酵素分子は、結晶格子中においては9つの分 子と接している。RNase H 結晶のVmは1.94で、Matthews (1968)の式から37%の溶 媒水が結晶に含まれると計算される。分子の接触部分は比較的せまく、電子密度分 布図には、大きな溶媒領域が明瞭に認識される。更に、分子間の相互作用を表4 に 示した。

RNase H 分子には、4つの酸性アミノ酸残基が集中しているMg²⁺結合部位が1ヶ 所ある。また塩基性のアミノ酸残基が集中しているところが、塩基性の突出部と、 βB→βCの折り返し部の2ヶ所ある。Protein Data Bank に登録されている 179 個の異なる蛋白質の立体構造に対して、 Nakamura と Wada の式(1985)を用いて 双極子モーメントを計算したところ、RNase H はその値が非常に高い分子であり、 上位10%内に位置づけられることが判明した。すなわち、RNase H 分子の表面は電 荷が大きく偏っており、これが結晶化の容易さと関連しているだろうと思われる。 実際、結晶中での分子の配列を見ると、ある分子の正電荷の表面が、隣の分子の負 電荷の表面に接しており、分子の電荷の偏りを打ち消すように並んでいる。

図14において、分子(x, y, z)は、隣接分子(1/2+x, 1/2-y, 1-z)と2つの塩 橋、すなわち Asp 10-Lys 87 と Asp 70-Lys 87 を通じて接している。また、(1 /2+x, 1/2-y, 1-z)と(-x, 1/2+y, 1/2-z)に関しては、前者のN末端ループと後 者のβCとの間に、逆平行βシート様の水素結合が形成されている。このN末端ル ープは、実際、MIRマップでもC末端より明瞭に解釈でき、また精密化後の温度 因子もC末端側より低く抑えられている。



図14 大腸菌 RNase H 結晶の パッキング。直方体のわくは単位格子で、各方向の らせん軸は長い矢印で示した。太く短い矢印の部分は酸性アミノ酸残基の集まって いるMg²⁺結合部位に隣接分子の塩基性突出部位が入り込んでいる様子を表す。この 噛み合わせは、基質との複合体の結晶化を難しくしている原因の1つと思われる。 なお、各分子の位置関係はA(x,y,z)、B(-x,1/2+y,1/2-z)、C(1/2+x,1/2-y,1-z) D(1/2-x,1-y,1/2+z) で、EとFはそれぞれ、AとBをx方向に単位格子分だけ平 行移動したものである。

- 36 -

a) Hydrogen bonds. (≤ 3.4 Å)					
Residue Atom	Residue	Atom	Distance(Å)	Symmetry Operation	
Met 1 N	Tyr 22	07	3.01	-1/2-X,-Y,1/2+Z	
Met 1 O	Ser 36	N	2.79	-1/2-x, $-Y$, $1/2+z$	
Lys 3 N	Thr 34	0	2.90	-1/2-x, $-y$, $1/2+z$	
Glu6 Οε1	Arg 31	Ny 1	2.99	-1/2-x, $-y$, $-1/2+z$	
Arg 31 Ny 2	Glu 119	Οε 1	3.08	-1/2-x, $-y$, $-1/2+z$	
Arg 29 Ny 2	Gly 126	0	3.18	-1/2 - x - y - 1/2 + 7	
Glu 64 Oc 1	Ala 93	N	3.40	-1/2+x.1/2-y.1-7	
Asp 10 0δ 1	Lys 87	Nζ	2.81	$1/2 + x \cdot 1/2 - y \cdot 1 - 7$	
Asp 70 0δ 1	Lys 87	Nζ	3.31	$1/2 + x \cdot 1/2 - y \cdot 1 - 7$	
Asp 70 0δ 2	Lys 87	Nζ	2.57	$1/2 + x \cdot 1/2 - y \cdot 1 - z$	
Gln 76 Νε 2	His 114	o	3.12	1/2+x.1/2-y.1-z	
Gln 76 Ne 2	Gly 112	0	2.63	1/2+X, $1/2-Y$, $1-7$	
Gln 80 Οε 1	Gly 112	0	3.17	1/2+x.1/2-y.1-z	
Gln 80 Νε 2	Gly 112	0	3.04	$1/2 + X \cdot 1/2 - Y \cdot 1 - Z$	
Thr 40 Ογ 1	Lys 122	Nζ	3.34	X.Y.1-Z	
Arg 75 Ny 1	Glu 147	Οε 1	2.98	X.Y.1+Z	
Arg 75 Ny 2	Thr 149	0γ1	3.03	X.Y.1+Z	
His 124 Νε 2	Pro 144	0	3.27	X.Y.1+Z	
Lys 91 Νζ	Glu 135	Οε 2	2.95	-x.1/2+y.1/2-7	
Asp 94 0	Arg 132	N7 1	2.51	-X,1/2+Y,1/2-Z	
b) Non-bonded o	contacts.	(≦4.0	Å)		
Glu 6 Oe 2	Arg 31	N7 1	3.65	-1/2-X,-Y,1/2+Z	
Arg 31 Ne	Glu 119	0ε 1	3.58	-1/2-X, $-Y$, $1/2+Z$	
Asp 10 0δ 2	Lys 87	Nζ	3.43	1/2+X, 1/2-Y, 1-Z	
Asn 44 08 1	His 83	Νε 2	3.87	1/2+X, 1/2-Y, 1-Z	
Asn 44 Νδ 2	His 83	Νε 2	3.65	1/2+X, 1/2-Y, 1-Z	
Gln 76 Οε 1	Gly 112	0	3.81	1/2+X, 1/2-Y, 1-Z	
Trp 85 Νε 1	Gln 115	0ε 1	3.90	1/2+X, 1/2-Y, 1-Z	
Gln 4 Ne 2	Asp 94	Οδ 1	3.82	-1/2+X, $1/2-Y$, $-Z$	
His 62 N⊗ 1	Gln 80	0ε 1	3.92	-1/2+X, $1/2-Y$, $1-Z$	
Glu 64 Oe 1	Arg 88	N7 1	3.41	-1/2+X, $1/2-Y$, $1-Z$	
Glu 64 Oe 2	Ala 93	N	3.58	-1/2+X, 1/2-Y, 1-Z	
Lys 86 O	Gly 123	0	3.92	-1/2+X, $1/2-Y$, $1-Z$	
Lys 86 O	His 124	NS 1	3.84	-1/2+X, 1/2-Y, 1-Z	
Lys 87 O	His 124	N	3.41	-1/2+X, 1/2-Y, 1-Z	
Gly 89 N	His 124	Nδ 1	3.88	-1/2+X, 1/2-Y.1-Z	
Arg 75 Ny 2	Glu 147	0ε 1	3.99	X,Y,1+Z	
Arg 75 Ny 1	Cl 147	0.2	2	· · · · · · · · · · · · · · · · · · ·	
Lys 91 NS	GIU 14/	08 2	3.//	X,Y,1+Z	
	Glu 147 Glu 135	0ε 2 0ε 1	3.70	X,Y,1+Z -X,1/2+Y,1/2-Z	
Asp 94 O	Glu 147 Glu 135 Arg 132	Οε 2 Οε 1 Νε	3.70 3.59	X,Y,1+Z -X,1/2+Y,1/2-Z -X,1/2+Y,1/2-Z	

<u>表4</u> 大腸菌 RNase H の結晶学的対称分子間の水素結合。 対象となる原子どうし の距離が 3.4Å以内のものを水素結合とみなし、 3.4Å以上 4.0Å以下の範囲にあ るものは、non-bonded contact とした。 第9節 活性金属の結合部位

RNase H 活性にはMg²⁺が必須である。そのMg²⁺の結合する位置を同定するために RNase H 結晶をMg²⁺イオンを含む水溶液に浸漬させ、回折強度データを収集した。 さらに、Mg²⁺はその原子番号が小さいため、原子番号のより大きなアルカリ土類金 属、すなわち、Ca²⁺やBa²⁺についても同様の実験を行い、これらの金属が結合する 部位を同定した。

これら3つの2価の金属イオンは、 Mg^{2+} のみならず Ca^{2+} や Ba^{2+} でも、全く同じ位置に結合することが明らかになった。さらに、RNase H活性に必要な金属は、 Mg^{2+} のみではなく、 Ca^{2+} や Ba^{2+} でも、僅かながらRN Aを切断する活性を示すことが、 HPLCの実験により示された(Iwai ら、 投稿準備中)。 Mg^{2+} にイオン半径が近いという理由で、 Co^{2+} についても浸漬実験を行ったが、 Mg^{2+} 位置からは約 3.5Å離れたところにピークが見いだされた(Asp 10 と Asp 70 の側鎖に挟まれたところ)。

また、RNase H 活性はMg²⁺の他にMn²⁺でも、ある場合には促進されるという事 実から浸漬実験や共結晶化を試みたが、浸漬中に沈澱を生じやすく、やむを得ず行 った飽和濃度では、差フーリエ図上にはMg²⁺などの場合のような有意なピークは見 出せなかった。この場合、最高のピークは His 62 付近に見出されたが、その理由 は現在まで不明である。

図15にそれぞれの金属イオンの結合位置を差フーリエ図を用いて示した。また、 精密化した金属の無い RNase H の、 これに対応する領域の 2 | Fo | - | Fc | マッ プを図16に示し、金属イオンが結合していないときの環境を示した。図から明らか なように、各金属イオンの結合位置は、 イオン半径の大きさ (Shannon、1976) に 対応して、Ba²⁺、Ca²⁺、Mg²⁺の順に、キャビティの内側へシフトしている。Co²⁺は、 Asp 10 と Asp 70のカルボキシル基に挟まれる位置に存在しているが、O - Coの 距離は、どれとも 2.4Åであり、これは、Co²⁺が4つのカルボキシル酸素原子に、 配位結合している状態である。Mg²⁺のピークの位置は、重原子の位置を決定する場

- 38 -



図15 RNase H の金属の結合した結晶と金属のない結晶との差フーリエ図。この図 は (| F_{metal} | - | F_{free} |) exp ($i\alpha_{MIR}$)の式で計算した。残差電子密度の等 高面は、黄色が5.0 σ の高さのMg²⁺のものを、青色が4.0 σ の高さのCa²⁺のものを、 また赤色が14.0 σ の高さのBa²⁺のものをそれぞれ示す。また、蛋白質側の構造は、 金属を含まない野生型の構造(精密化後)を重ね合わせたものであり、紫色の骨格 はその対称分子の構造である。



図16 金属を含まない結晶について、精密化後の構造をもとにして計算した、Mg²⁺ 結合部位付近の2Fo-Fc電子密度図。水分子である WAT250 付近が金属の結合す る位置である。

!

合と同一の方法によって精密化したもので、標準偏差は0.16Åである。Mg²⁺の、近 傍の酸性アミノ酸残基のカルボキシル基からの距離は それぞれ、Asp 10 から 2.0 Å、Glu 48 から 4.2Å、Asp 70 から 4.5Å、Asp 134 から 5.0Å であった。 ま た、Asp 10 のカルボキシル酸素から 2.0Å、Gly 11 のカルボニル酸素から 2.3Å の距離で、 Mg²⁺はこれらの酸素原子に配位していることが示唆される。

第10節 大腸菌 RNase H と逆転写酵素 RNase H ドメインの活性中心

Doolittle ら(1989)は、大腸菌 RNase H と、25 種類の逆転写酵素およびその 類縁蛋白質について、アミノ酸配列の類似性を詳しく調べた。その結果、Asp 10、 Glu 48、Asp 70 と Asp 134 の4つの酸性アミノ酸残基が、計26種類の RNase H 様配列全体にわたって完全に保存されていることが、明らかとなった(表5)。し かしながら、彼らは、これらのアミノ酸残基の保存の意味については、言及できな かった。この論文が発表された時点で、我々は すでに大腸菌 RNase H の初期モデ ルの構築を終了し、このモデルから、これら4つの酸性アミノ酸残基が、1つの領 域に集中して存在していることを見いだしていた。さらにMg²⁺位置を差フーリエ図 上で同定したことから、これらの酸性アミノ酸残基が活性に関係する重要なもので あることが判明した。図17にこれら4つの酸性アミノ酸残基と、同定されたMg²⁺の、 分子全体に対する位置関係を示した。X線解析と平行して、Kanaya ら(1990b)が行 った、活性アミノ酸残基を求める一連の部位特異的変異法の実験結果からも、大腸 菌 RNase H の場合、 前3者の酸性残基については、どの1つをアミノ酸置換して も酵素活性が完全に失われることが裏付けられた(表6)。また我々の結果からは、 逆転写酵素の RNaseH 活性部位は、すべておなじ構造を取っている可能性も、同時 に示唆される。

なお、これらの酸性アミノ酸残基のほかにも、本章第2節で述べられた分子内相 互作用をしている残基は保存されている傾向にある。さらに、本章第1節で述べら

- 41 -

HTLV-I	INTAPCLFSDGS TSRA AYI	LWDKQIL	SQRSFPLPPPH	KS AQRAELLGLLHGLSSARS	W RCLN
HTLV-II	LDTAPCLFSDGS PQKA AYV	LWDQTIL	QQDITPLPSHET	HS AQKGELLALICGLRAAKP	V PSLN
BLY	IPAALCLFSDGA TGRG AYC	LWKDHLL	DEGAVPAP	ES AOKGELAGLLAGLAAAPP	EPLN
SRV-I	LNNALLVFTDGSSTGMA AYT	L ADTTI	KFQTNL	NS AQLVELQALIAVL SAFPI	N OPLN
IAP-18	LPDGTVVYTDGSKTGLG AYV	V KDRVI	SKOYNE	TS POVVECLIVLEVI. EAFPO	G PLN
HMTV	LEKGIVIFTDGSANGRSVTYT	OGREPTT	KENTO	NT ADDAETVAVITAFFEVS	OPEN
HERY-K	LENALTVETDGSSNGKA AYT	GPKERVI	KTPYO	S AORDEL VAVITULODED	OPTN
RSV	PVPCPTVETDASSSTHE CVV	VURECRO	WETWETAD IC	AS VOOLEADAVAMAL LLUP	TOTN
HTV.	TVCAETEVUCAANBETVLCV	ACTUTAL	COOKN NOLTH	TT NOVIEL ON THINK LLWP	
1111 1111	T TOALIF I TDUARAGINDOCKCOK	AGIVIAN	GROKY YPLIN	TT AQNIELQATILALQUS G	LEVN
0112 VICUA	IPGAEIFIIDGSCNRQSREGK	AGIVIDE	GRDKY KKLEU	TI NQUAELEAFAMALIDS G	PKVN
VISNA	LVPGPTTTTDGG KKNGK GS	LGIIAST	G EKF RIHEE	GT NQQLELRAIEEACKQG P	EKMN
EIAV	PTSGITIYTDGG KQNGEGI	AAYVTSN	GRTKQ KRLGP	VT HQVAERMAIQMALEDTRD	KQVN
MOMLY	PDADHTWYTDGSSLLQEGQRK	AGAAVTT	ETEVIWAKALDA	GTSAQRAELIALTQALKMAEG	KKLN
ECOLI	MLKQVEIFTDGSCLGNPGPGG	YGAILRY	RGREKTFSAGYT	RTTNNRMELMAAIVALEALKEI	H CEVI
HEPB	RPGLCQVFADATPTGWG LV	MGHQRMR	GTFSAPLP	INTAELLA ACFARSRS	G ANI
17.6	FTKKFTLTTDASDYALGAVLSQ	DGH PL	SYISRTLNEHEI	NYSTIEKELLAIVWATKTFRH	YLLG RHFE
297	FEKKFVLTTDASNLALGAVLSQ	NGH PI	SFISRTLNDHEL	NYSAIEKELLAIVWATKTERH	YLLG ROFL
412	FSKEFCITTDASKOACGAVLTONH	NGHOL PV	AYASRAFTKGES	NKSTTEOELAATHWATTHERP	YTYG KHET
GYPSY	FKKPFDLTTDASASGTGAVLSO	EGR PT	THISRTI KOPFO	NYATNERELLATIVALCELON	FI VESPETN
CERV	PNDKI VIETDASEEEVOOTI KATH	NCHEVIC	RYASCSEVAAER	NARSNERELLANDARTAKEST	VITO COCI
CANV	PEEKI TIETDASDOVUCCMI VAIKI	NECTNIEL IC	DVASCEEVAACK	NINGNEKELLAVINTIKKESI	ILIP SAFL
DIRS	PECKETTETERSOUTHOUTEKAINT	REGINIELIC	VTUESOUETTOE	NINSRUKETER VINTIKKEST	ILIP WHPL
TNCT	PSIDIALIIDASESGAGAILA	KUNK VI	KIWSEQWSIIQS	NMSSNRREMLALLMAIQALCH	KLNS CKLK
INGI I DIO	PREHIKLWIDGSVSLGEKLGAAAL	LHKNNTLICA	PRTGAGELSCSTR	AECVA LE IGLQRLLKWLPR	INSTPARLS
1-FAC	LEIHNFIFTDGS KINITI SFAL	TTETDVL	KIGI LP PYSSVL	TSETIAILE ALE LTK NR	R GKFI
COPD	FENKIIGYVDSDWAGSEIDRKST	TGYLFKMFDF	NLICWNTKRQNSV	AASSTEAEYMALFEAVREALW	KLE NPIK
TIEI	PDNKLVAISDASY GNQPYYKSQ	IGNIF LLNG	KVIGGKSTKASLT	CTSTTEAEIHAVSESI PLL	NKK PIIK
	· ·				
	•				-
HTLV-I	IFLD SKYLYH YL R	TLALGTE QGRSS	QAP FQALLPRLLSR	K VVYLHHVRSHTNLPDPISRLI	NALTDALLI
HTLV-I HTLV-II	IFLD SKYLYH YL R IFLD SKYLIK YL H	TLALGTF QGRSS SLAIGAF LGTSA	QAP FQALLPRLLSR HQT LQAALPPLLQG	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI	NALTDALLI NEYTDSLIL
HTLV-I Htlv-II Blv	IFLD SKYLYH YL R IFLD SKYLIK YL H IWYD SKYLYS LL R	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH	K VVYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI	NALTDALLI NEYTDSLIL NNYVDQLL
HTLV-I HTLV-II BLV SRV-I	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL	TLALCTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVCHVRSHSSASHPIASLI SIPFYICHVRAHSCLPCPIAHGI	NALTDALLI NEYTDSLIL NNYVDQLL NQKADLATK
HTLV-I HTLV-II BLV SRV-I IAP-18	IFLD SKYLYH YL R IFLD SKYLIK YL H Iwyd skylys Ll R Iytd Saylahsipl Iysd ssyvnavnl	TLALCTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI HFPVFITHVRAHSGLPGPMSLGI	NALTDALLI NEYTDSLIL NNYVDQLL NQKADLATK NDLADKATK
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVNAWNL LYTD SKYVTG LF P	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI RFPVFITHVRAHSGLPGPLSGQ GEKFYIGHIRGHTGLPGPLAQG	NALTDALLI NEYTDSLIL NNYVDQLL NQKADLATK NDLADKATK NAYADSLTR
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYYVQ AT R	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR DLNOLFNLLQOTYRKR	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIFFYIGHVRAHSGLPGPIAHGI HFFVFITHVRAHSGLPGPMSLGI QEKFYIGHIRGHTGLPGPLAQGI NFFFYITYIRAHTNLPGPLTKA	NALTDALLI NEYTDSLIL NNYVDQLL NQKADEATK NDLADKATK NAYADSLTR NEGADLL
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAFVAK ML	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMGDEGV PST	QAP FQALLPRLLSR HOT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTYRKR AAA F TLEDALSOR	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTEH P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI HFPVFITHVRAHSGLPGPMSLGI QEKFYIGHIRGHTGLPGPLAQGI NFPFYITYIRAHTNLPGPLTKAI SAMAAVI HVRSHSFVPGFFFFGI	NALTDALLI NEYTDSLIL NNYVDQLL NQKADEATK NDLADKATK NAYADSLTR NEQADLL NDVADSOAT
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLIS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VVTD SAFYAK ML L IVTD SOYALGILO A	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLAH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTYRKR AAA F ILEDALLSQR FIVDQITOL I	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPOPIAHGI RFPVFITHVRAHSGLPGPLAGGI QEKFYIGHIRGHTGLPGPLAGGI NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPCFFTEGI VFFVVI AVPAH VGIG	NALTDALLI NEYTDSLIL NNYVDQLL NQKADEATK NDLADKATK NAYADSLTR NEQADLL NDVADSQAT
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R WYTD SAFVAK ML L IVTD SQYALGIIQ A IVD SQYALGIIQ S	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST QPDK SES	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIKK QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK VINUOIFS	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI NFPVFITHVRAHSGLPGPSLGI QEKFYIGHIRGHTGLPGPLAQGI NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPCFFTEGI KEKVYLAWVPAH KGIG GI	NALTDALLI NEYTDSLIL NQKADLATK NDLADKATK NALADKATK NEQADLL NDVADSQAT NEQVDKLVS
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAFVAK ML L IVTD SQYALGIIQ A IIVD SQYALGIIQ SA	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSHDD KMGQEGV PST QPDK SES QPTE SES	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEQL IK KIVNQIIEM IK	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIFFYIGHVRAHSGLPOPHSLGI QEKFYIGHIRGHTGLPGPLAQGI NFFFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPCFFTEGI KEKIYLAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI	NALTDALLI NEYTDSLIL NNYYDQLL NGKADEATK NDLADKATK NAYADSLTR NEQADLL NDVADSQAT NEQVDKLVS NGEVDHLVS
HTLV-I HTLV-II BLV SRV-I IAP-18 MHTV HERV-K RSV HIV HIV2 VISNA FILV	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R WYTD SAYAVQ AT R IVTD SQYALGIAQ A IIVD SQYYHGISA S IVTD SRYAKEFML R	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS ELEIAGI IRSSS EIETATL SPRT DVETALIKYSHDD KMGQEGV PST QPDK SES QPTE SES NVDEEV IRN	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLAH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEQH IK VINQIIEQH HN DIQATINELY HN	K VYYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHPIASL SIPFYIGHVRAHSGLPGPHSLG NFPVFITHVRAHSGLPGPMSLG QEKFYIGHIRGHTGLPGPLAQG NFFFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPCFFTEG KEKYYLAWVPAH KGIG GI KEKIYLAWVPAH KGIG GI KEKIYLAWVPAH KGIP QI	NALTDALLI NEYTDSLIL NNYVDQLL NQKADLATK NDLADKATK NEQADLL NEQADLL NDVADSQAT NEQVDKLVS NQEVDHLVS
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MONYV	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLIS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVYQ AT R VYTD SAFYAK ML L IVTD SQYALGIIQ A IIVD SQYVHGISA S IVTD SRYAYEFML R IVTD SYYCKKNIT E	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KHQQGGV PST QPDK SES NWDEEV IRN GLGLEG PQN	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLRH TTAKLFLQCQQLIINN RVANIFQKIQAALLNR KIYTELKHLQRLINKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEEM IK PIQARIMELV HN PW WPIIQNI RE	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI HFPVFITHVRAHSGLPGPLAGGI QEKFYIGHIRGHTGLPGPLAGGI NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFFTEGI KEKIYLAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI KEKIGVHWVPGH KGIP QI KEIYYFAWVPGH KGIP QI	NALTDALLI NEYTDSLIL NNYVDQL NQKADLATK NJLADKATK NAYADSLTR NEQADLL NDVADSQAT NEQVDKLVS VEEIDRYI NQLADEAAK
HTLV-I HTLV-II BLV SRV-I IAP-I8 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV SCOLI	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAFVAK ML L IVTD SQYALGIIQ A IIVD SQYVMGISA S IVTD SRYAFFML R IVTD SRYAFFML R IVTD SRYAFFML R	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHISS LEIAGI IRSSS DVETALIKYSMDD KMGQEGV PST QPDK SES NVDEEV IRN GLGLEG PQN LLTSEG KEI	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIKK QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEEM IK PIQARIHELY HN PW WPIIQNI RE KNKDEILALLKALFLP	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI NFPYITHVRAHSGLPGPMSLGI QEKFYIGHIRGHTGLPGPLAQGI NFFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFFTEGI KEKIYLAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI KEKIGVHWVPGH KGIP QI KEKIYFAWVPGH KGIY GI K RISIHCPGHOKGHSAEARGI	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NEQADLL NEQADLL NEQADLL NEQADLL NEQVDKLVS NQEVDHLVS VGEIDRYI VQLADEAAK NFMADQAAR
HTLV-I HTLV-II BLV SRV-I IAP-18 MHTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLIS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAYAK ML L IYTD SQYALGIIQ A IIVD SQYYMGISA S IVTD SRYAFFML R IVTD SRYAFFML R IVTD SRYAFFML R IVTD SRYAFFML R	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST QPDK SES QPTE SES NWDEEV IAN GLGLEG PQN LLTSEG KEI NWKKRCWKTADKK	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEEM IK PIQARIMELV HN PW WPIIQNI RE KNKDEILALLKALFLP PVKNVDLWQRLDAALG	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIFFYIGHVRAHSGLPGPIAHGI WFPYITHVRAHSGLPGPNSLGI QEKFYIGHIRGHTGLPGPLAQGI NFFFYITIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFFTEGI KEKIYLAWVPAH KGIG GI KEALIYAWVPAH KGIG GI KELVYFAWVPGH KGIP QI KELVIFAWVPGH KGIY K RLSIIHCPGHQKGHSAEARGI QHQIKWEWVKGHAGHPE	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NEQADLL NDVADSQAT NEQVDKLVS NQEVDHLVS VEEIDRYI NQLADEAAK NRMADQAAR NERCDELAR
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IYSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVYQ AT R VYTD SAFVAK ML L IYTD SQYALGIIQ A IIVD SQYVMGISA S IVTD SRYAYEFML R IYTD SRYAFFATAHINGEIYRRRG LSTD SQYVRQGITQ WIH IGTDNSVYLSRKYTS FPW	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS ELETAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMQQEGV PST QPDK SES QPTE SES NWDEEV IRN GLGLEG PQN LLTSEG KEI NMKKRGKKTADKK LLGCAANWILRGT	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLAH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLINKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR EUNQIIEQL IK KIVNQIIEEM IK PIQARIHELV HN PW WPJIQNI RE KNKDEILALLKALFLP PVKNVQLWQRLDAALQ SFYYVPSALNPADDPS	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPOPIAHGI RFPVFITHVRAHSGLPOPLAHGI QEKYIGHIRGHTGLPOPLAGGI NFPFYITYIRAHTNLPOPLTKAN SAMAAVLHVRSHSEVPOFFTEGI KEKIYLAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI KEAIYYAWVPGH KGIP QI KEIVYFAWVPGH KGIP QI KEIVYFAWVPGH KGIP GI K RLSIIHCPGHQKGHSAEARGI QHQIKWEWKGHAGHPE H RGRLGLSRPLLRLPFRPTTGRTS	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NEQADLL NDVADSQAT NEQVDKLVS VEEIDRYI NQLADEAAK YRMADQAAR YRMADQAAR SLYADSPSV
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLIS LL R IYTD SAYLAHSIPL IYSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAFVAK ML L IVTD SQYALGIIQ A IIVD SQYVNGISA S IVTD SRYAFFML R IYTD SYYCWNIT E VYTD SRYAFFML R IYTD SYYCWNIT E IYTD SYYCWNIT F SYGVARGITQ WIH IGTDNSVVLSRKYTS FPW ISSDH QPLS W	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMQQEGV PST QPDK SES NVDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRGWKTADKK LLCCAANWILRGT LYRMK DPNS	QAP FQALLPRLLSR HOT LQAALPPLLOG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIEEM IK PIQARIMELV HN PW WPIIQNI RE KNKDEJLALLKALFLP PVKNVDLWQRLDAALG SFVYVPSALMPADDPS KLTR WRVKL S	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPHAGG MFFPYITHVRAHSGLPGPSLGQ QEKYIGHIRGHTGLPGPLAQG NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFTEGI KEKYIAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI KEXIGVHWVPGH KGIP QI KEIVIFAWVPGH KGIP QI KEISIHCPGHQKGHSAEARGI QHQIKWEWVKGHQGHSE I RGRLGLSRPLLRLPFRPTTGRTT EFDFDIKYIKGK EI	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NDLADKATK NEQADLL NDVADSQAT NEQADLL NDVADSQAT NQEVDHLVS NQEVDHLVS NGEIDRYI NQLADEAAK NEHADQAAR NERCDELAR NERCDELAR SLYADSPSY NCVADALSR
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6 297	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAFVAK ML L IVTD SQYALGIIQ A IIVD SQYVMGISA S IVTD SRYAFFML R IVTD SRYAFFML R IVTD SRYAFFML R IVTD SRYAFATAHIHGEIYRRGG LSTD SQYVRGGITQ WIH IGTDNSVVLSRKYTS FPW ISSDH QPLS W	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST QPDK SES NVDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRGWKTADKK LLGCAANWILRGT LYMK PPNS LHNLK EPGA	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIKK QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEEM IK PIQARIMELY HN PW WPIIQNI RE KNKDEILALLKALFLP PVKNVDLWQRLDAALG SFYYVPSALNPADDPS KLTR WRVKL S	K VYYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI NFPYITHVRAHSGLPGPMSLGI QEKFYIGHIRGHTGLPGPLAQGI NFFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFFTEGI KEKIYLAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI KEKIGVHWVPGH KGIP QI KEKIGVHWVPGH KGIP QI KELIYIFAWVPGH KGIP QI KELIYIFAWVPGH KGIP QI KELIYIFAWVPGH KGIP GI KRLSIIHCPGHQKGHSAEARGI QHQIKWEWVKGHAGHPE I RGRLGLSRPLLRLPFRPTTGRTS EFOPDIXYIKGK EI	NALTDALLI NALTDALLI NATADALI NAYADUL NQKADLATK NAYADSLTR NEQADLL NEQADLL NEQADLL NEQADLL NEQADLLSR YRMADQAAR NERCDELAR SLYADSPSV YSVADALSR
HTLV-I HTLV-II BLV SRV-I IAP-18 MHTV HERV-K RSV HIV HIV2 VISNA EIAV MOHLV ECOLI HEPB 17.6 297 412	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IYSD SSYVWAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAYAVQ AT R IVTD SQYALGIIQ A IIVD SQYALGIIQ A IIVD SQYYHGISA S IVTD SRYAFATAHIHGEIYRRGG LSTD SQYYRQGITQ WIH IGTDNSVVLSRKYTS FPW ISSDH QPLS W VKTDH RPLT Y	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST QPDK SES QPTE SES NWDEEV IRN GLGLEG PQN LLTSEG KEI NKKKRGWKTADKK LLGCAANWILRGT LYRMK DPNS LHNLK EPGA	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLAH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTVRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEEM IK KIVNQIIEEM IK KIVNQIIEEM IK KKKDEILALLKALFLP PVKNVDLWQRLDAALQ SFVYVPSALNPADDPS KLTR WRVKL S KLER WRVKL S KLER WRVKL S KLER WRVKL S	K VYYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHPIASL SIPFYIGHVRAHSGLPOPIAHGI MFPVFITHVRAHSGLPOPLAHGI NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPOFFTEGI KEKYYLAWVPAH KGIG GI KEKIYYAWVPAH KGIG GI KEKIGVHWVPGH KGIP QI KEIVYFAWVPGH KGIP QI KEIVYFAWVPGH KGIP GI KELSIHCPGHQKGHSAEARGI QHQIKWEWVKGHAGHPE H RGRLGLSRPLLRLPFRPTTGRTS EFDFDIXYIKGK EN EYOFKIDVIKGK EN	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NEQADLL NDVADSQAT NEQVDKLVS VEEIDRYI NQLADEAAK NRMADQAAR SLYADSPSV VCVADALSR
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6 297 412 GYPSY	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IYSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVYQ AT R VYTD SAFYAK ML L IYTD SQYALGIIQ A IIVD SQYVAGISA S IVTD SRYAFATAHHGEIYRRRG LSTD SQYVRQGITQ WIH IGTDNSVVLSRKYTS FPW ISSDH QPLS W IASDH QPLR W	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS ELEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KHQQEGV PST QPDK SES NWDEEV IRN GLGLEG PQN LLTSEG KIE NMKKRGKKTADKK LLGCAANWILRGT LYRMK DPNS LHNLK EPGA LFSMV NPSS AVADR NTNA	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLAH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLINRR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEM IK PIQARIMELV HN PW WPIIQNI RE KNKDEILALLKALFLP PVKNVOLWQRLDAALG SFYYVPSALNPADDPS KLTR WRVKL S KLTR WRVKL S KLTR WRVKL S KLTR WRVKL E KIKT WKSYI D	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI HFPVFITHVRAHSGLPGPLAHGI QEKYIGHIRGHTGLPGPLAGGI NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFFTEGI KEKYYLAWYPAH KGIG GI KEAIYVAWYPAH KGIG GI KEAIYVAWYPAH KGIG QI KEIVYFAWYPGH KGIP QI KEIVYFAWYPGH KGIP QI KEIVYFAWYPGH KGIP QI KEIVFFAWYGHAGHSAEARGI QHQIKWEWKGHAGHSE H EYGFLGLSRPLLRLPFRPTTGRTS EFDFDIKYIKGK EH EYNFTYEYLKGK DP QHNAKYFYKPGK EH	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NDLADKATK NEQADLL NDVADSQAT NQEVDHLVS VEEIDRYI NQLADEAAK RHADQAAR SLYADSPSV YCVADALSR HVADALSR
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6 297 412 GTPSY CERV	IFLD SKYLYH YL R IFLD SKYLIK YL H IWVD SKYLIS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VVTD SAFVAK ML L IVTD SQYALGIIQ A IIVD SQYVHGISA S IVTD SRYAFFML R IVTD SYYCWNIT E VYTD SRYAFFML R IVTD SYYCWNIT E VYTD SRYAFATAHIHGEIYRRRG LSTD SQYVRQGITQ WIH IGTDNSVVLSRKYTS FPW ISSDH QPLS W IASDH QPLR W VKTDH RPLT Y IFTDH QPLT F IRTDN KNFT H	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST QPDK SES NVDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRGWKTADKK LLCCAANWILRGT LYRMK DPNS LHNLK EPGA LFSMV NTSA FVNINLKGDRKQG	QAP FQALLPRLLSR HOT LQAALPPLLOG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEEM IK FJQARIMELV HN PW WPIIQNI RE KNKDEILALLKALFLP PVKNVDLWQRLDAALG SFVYVFSALNPADDPS KLTR WRVKL S KLER WRVRL S KLER WRVRL S KLER WRVRL S KLER WRVRL E KKIKR WKSYI D RLVR WQMWL S	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPIAHGI KFPVFITHVRAHSGLPGPSLGI QEKFYIGHIRGHTGLPGPLAQGI NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFFTEGI KEKVYLAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI KEXIGVHWVPGH KGIP QI KEIVIFAWVPGH KGIP QI KELSIHHCPGHQKGHSAEARGI QHQIKWEWVKGHQKGHSAEARGI QHQIKWEWVKGHQKHSE EYOFKIDYIKGK EI EYOFKIDYIKGK EI QHNAKVFYKPGK EI QHNAKVFYKPGK EI	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NDLADKATK NEQADLL NDVADSQAT NEQADLL NDVADSQAT NEQADLLSS NQEVDHLVS NQEVDHLVS NGELADEAAK NHADQAAR NERCDELAR SVADALSR VFVADALSR VVFADFLQE
HTLV-I HTLV-II BLV SRV-I IAP-18 MHTV HERV-K RSV HIV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6 297 412 GYPSI CERV CAMV	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAYAK ML L IYTD SQYALGIIQ A IIVD SQYMGISA S IVTD SRYAFEFML R IVTD SRYAFEFML R IVTD SRYAFEFML R IVTD SRYAFEFML R IVTD SRYAFEFML R IVTD SRYAFEFML R ISDH QPLS W ISSDH QPLS W VXTDH RPLT Y IFTDH QPLT F IRTDW KNFT H	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAN LQPOP LETVAQ IKHIS ELEIAGI IRSSS EIETATL SPRT DVETALIKYSHDD KMGQEGV PST QPDK SES QPTE SES NVDEEV IRN GLGLEG PQN LLTSEG KEI NVKKRGWKTADKK LLCCAANVILRGT LYRMK DPNS LHNLK EPCA LFSMV NPSS AVADR NTNA FVNINLKGDKKQG	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLAH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLIHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEQL IK KIVNQIIEQH IK PIQARIHELY HN PW WPIIQNI RE KKKDEILALKALFLP PVKNVDLWQRLDAALG SFVYVPSALNPADDPS KLTR WRVKL S KLER WRVRL S KLTR IRLEL E KIKR WKSYI D RLVR WQAWL S RNIR WQAWL S	K VYYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHFIASL SIPFYIGHVRAHSGLPGPIAHGI MFPVFITHVRAHSGLPGPHSLGI QEKFYIGHIRGHTGLPGPLAQGI NFFFFITTYIRAHTNLPGPLTKAI SAMAAVLHVRSHSEVPCFFTGGI KEKYLAWVPAH KGIG GI KEKYLAWVPAH KGIG GI KELVIFAWVPGH KGIP QI KEIVYFAWVPGH KGIP GI RGRLGLSRPLLRLPFRPTTGRTS EFDFDIKYIKGK EH EYNFTVEYLKGK DI QHNAKVFYKPGK EH YNFTVEYLKGK EH YNFTVEYLKGK EH YNFTVEYLKGK EH	NALTDALLI NETTDSLIL NNYYDQLL NOKADLATK NALADKATK NALADKATK NEQADLL NDVADSQAT NEQVDKLVS VEEIDRYI NOLADEAAK VEEIDRYI NOLADEAAK NERCDELAR SLYADSPSV VCVADALSR VFVADALSR VFVADALSR VFVADALSR VFVADALSR
HTLV-I HTLV-II BLV SRV-I IAP-18 MHTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6 297 412 GYPSY CERV CAMY DIRS	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IYSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAFVAK ML L IYTD SQYALGIIQ A IIVD SQYANGISA S IVTD SRYAFATAHIHGEIYRRGG LSTD SQYVNGITQ WIH IGTDNSVVLSRKYTS FPW ISSDH QPLS W IASDH QPLR W VKTDH RPLT Y IFTDH QPLT F IRTDN KNFT H IRTDN KNFT SY	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS ELEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMQQEGV PST QPDK SES OPTE SES NWDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRCWKTADKK LLTSEG KEI NWKKRCWKTADKK LLMLK PPOS LHNLK PPOS AVADR NTNA FVMINIKGDRXGG FVNLNYKGDSKLG	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLEHH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR EUNQIIEQL IK KIVNQIIEM IK PIQARIHELV HN PW WPJIQNI RE KNKDEILALLKALFLP PVKNVQLWQRLDAALG SFYYVPSALNPADDPS KLTR WRVKL S KLER WRVRL S KLER WRVRL S KLTR IRLEL E KIKR WKSYI D RLVR WQMWL S RNIR WQAWL S FFEQLWKQCL KK	K VYYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHPIASL SIPFYIGHVRAHSGLPOPIAHG HFPVFITHVRAHSGLPOPLAHG UGKFYIGHIRGHTGLPOPLAGG NFPFYITYIRAHTNLPOPLTKAN SAMAAVLHVRSHSEVPOFFTEG KEKVYLAWVPAH KGIG GI KEAIYVAWVPAH KGIG GI KEAIYVAWVPAH KGIP QI KEIVYFAWVPGH KGIP QI KEIVYFAWVPGH KGIP QI KEIVFINVFGH KGIP GI KEISIHCPGHQKGHSAEARGI QHQIKWEWVKGHAGHPE H RGRLGLSRPLLRLPFRPTTGRT EFOPDIKYIKGK EH EYQFKIDYIKGK EH EYQFKIDYIKGK EH QYDFOVEHIAGT KM HYSFDVEHIKGT DH KYNLIGEHIPGF FF	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NEQADLL NDVADSQAT NEQADLL NDVADSQAT NQEVDHLVS VEEIDRYI NQLADEAAK VRMADQAAR SLYADSPSV NCVADALSR VFVADALSR VFVADALSR VFFADFLQE VFFADFLQE VFFADFLQE
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV HOMLV ECOLI HEPB 17.6 297 412 GYPSY CERV CAMV DIRS INGI	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IYSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVYQ AT R WYTD SAFYAK ML L IYTD SQYLYGISA S IVTD SQYLYGISA S IVTD SQYLYGISA S IVTD SYYCKNIT E VYTD SYYCKNIT E VYTD SYYCKNIT E VYTD SYYCKNIT F ISDH QPLS W ISSDH QPLS W ISSDH QPLS W IASDH QPLS W IASDH QPLT F IRTDN KNFT H IRTDN THFK S IQTDNTTLS Y	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS EEIETATL SPRT DVETALIKYSMDD KHQQEGV PST QPDK SES NWDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRGKKTADKK LLGCAANWILRGT LYRMK DPNS LHNLK EPGA LFSMV NFSS AVADR NTNA FVNINLKGDRKQG FVNLNKCDSKLG INRCGGQGIQLSV ALQTGP LAVTD	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLRH ETAKLFLQCQQLIINR RVANIFQKIQAALLNR KIYTELKHLQRLINKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEM IK PIQARIMELV HN PW WPIIQNI RE KNKDEILALLKALFLP PVKNVOLWQRLDAALG SFYYVPSALNPADDPS KLTR WRVKL S KLTR WRVKL S KLTR IRLEL E KIKR WKSYI D RLVR WQMWL S RJFRUWWQWL S RJFRUWRLLQVQRR	K VYYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHPIASL SIPFYIGHVRAHSGLPGPIAHG HFPVFITHVRAHSGLPGPLAGG QEKYIGHIRGHTGLPGPLAGG NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFTEG KEKIYVAWVPAH KGIG GI KEKIYVAWVPAH KGIG GI KEKIYFAWPGH KGIP QH KEIYIFAWPGH KGIP QH KEIYIFAWPGH KGIP QH KEISIHCPGHQKGHSAEARG QHQIKWEWKGHAGHPE H RGRLGLSAPLLRLPFRTTGRTI EFDFDIKYIKGK EH EYNFTVEYLKGK DH QHNAKVFYKPGK EH QYDFDVEHIAGT KA HYSFLVEHIKGT DH KVNLIGEHIPGF FH	NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NDLADKATK NEQADLL NDVADSQAT NEQADLL NDVADSQAT NQEVDHLVS VEEIDRYI NQLADEAAK NERCDELAR NERCDELAR NSVADALSR VFADALSR VFADALSR VFADFLQE HVFADFLQE HVFADFLSR VFADFLSR VFXDALSR
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6 297 412 GYPSI CERV CAMV DIRS INGI I-FAC	IFLD SKYLYH YL R IFLD SKYLYK YL H IWUD SKYLYS LL R IYTD SAYLAHSIPL IVSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R WYTD SAYAK ML L IYTD SQYALGIIQ A IIVD SQYMGISA S IVTD SRYAFEFML R VYTD SRYAFEFML R IVTD SRYAFEFML R IVTD SRYAFEFML R IVTD SRYAFEFML R IVTD SRYAFEFML R IVTD SQYVRGCITQ WIH ISDDH QPLS W IASDH QPLR W VKTDH RPLT Y IFTDH QPLT F IRTDN KNFT H IRTDN THFK SS IQTDNTTILS Y IGTDNTTILS Y IFSDSLSMLT	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPDP LETVAQ IKHIS LEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST QPDK SES NVDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRGWKTADKK LLGCAANWILRGT LLGCAANWILRGT LHNLK EPGA LFSMV NPSS AVADR NTNA FVNINLKGDRKQG FVNLNYKGDSKLG INRGGGQIQDLSV ALQTGP LAVTD SIQNTN NNSFY	QAP FQALLPRLLSR HOT LQAALPPLLOG VPS YALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLHKR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEEM IK PIQARIMELV HN PW WPIIQNI RE KNKDEILALLKALFLP PVKNVDLWQRLDAALG SFVYVPSALNPADDPS KLTR WRVKL S KLER WRVKL S KLER WRVKL S KLER WRVKL S KLTR IRLEL E KIKR WKSYI D RLVR WQHWL S RNIR WQAWL S FFELLWRQL KK PILRRLWRLLQVQRR P SRIRSLITQ HAP	K VVYLHHVRSHTNLPDPISRLI K TIYLHHVRSHTNLPDPISTFI P AIFVGHVRSHSSASHPIASLI SIPFYIGHVRAHSGLPGPHAGG RFPYGITHVRAHSGLPGPSLGG QEKFYIGHIRGHTGLPGPLAGG NFPFYITYIRAHTNLPGPLTKAH SAMAAVLHVRSHSEVPGFFTEGI KEKIVYHAWPAH KGIG GI KEAIYVAWVPAH KGIG GI KEKIGVHWVPGH KGIP QI KEKIGVHWVPGH KGIP QI KEKIGVHWVPGH KGIY GI K RLSIIHCPGHOKGHSAEARGI QHQIKWEWVKGHAGHPE I RGRLGLSRPLRLPFRPTTGRTS EFDFDIKYIKGK EH EYOFKIDIKGK EH EYOFKIDIKGK EH EYOFKIDIKGK EH QYDFDVEHIAGT KM HYSFDVEHIAGT KM KYNLIGEHIPGF FF KIRIRLQFYFDHCGVK RH	NALTDALLI NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NQLADLATK NEQADLL NDVADSQAT NEQADLL NEQADLL NEQADLL NEQADLLSS NQEVDHLVS NQEVDHLVS NUERCDELAR VSVADALSR VSVADALSR VSVADALSR VFVADALSR VVFADFLQE VHFADFLQE VFFADFLQE VFFADFLQE VELADQAAK
HTLV-I HTLV-II BLV SRV-I IAP-18 MHTV HERV-K RSY HIV HIV2 VISNA EIAV HOMLV ECOLI HEPB 17.6 297 412 GTPSY CERV CCAHY DIRS INGI I-FAC COOPD	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IYSD SSYVWAVNL LYTD SKYVTG LF P IISD SAYVVQ AT R VYTD SAYAVQ AT R IVTD SQYALGIIQ A IVTD SQYALGIIQ A IVTD SQYALGIIQ A IVTD SYYLGKNIT R IVTD SYYCKNIT R IVTD SYYCKNIT R IVTD SYYCKNIT F ISDH QPLS W ISDH QPLS W VKTDH RPLT Y IFTDH QPLT F IRTDN KNFT H IRTDN THFK SS IQTDNTTILS Y IFSDSLSMLT ICSDSLSAVD IYEDNQCIS	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS EIETATL SPRT DVETALIKYSMDD KMGQEGV PST QPDK SES QPTE SES NWDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRGWKTADKK LLGCAANWILRGT LYRMK DPNS LHNLK EPGA LFSMV NPSS AVADR NTNA FVNINLKGDRKQG INRCGGQIQDLSV ALQTGP LAVTD SIQNTN NNSFY IANNPSCHK RAK	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLEH ETAKLFLQCQQLIYNR RVANIFQKIQAALLMR KIYTELKHLQRLIHKR QLNQLFNLLQQTVAKR AAA F ILEDALSQR ELVNQIEEL IK KIVNQIIEEL IK KIVNQIIEEL IK KIVNQIIEEL IK KKKDEILALLKALFU PVKNVDLWQRLDAALG SFVYVPSALNPADDPS KKLER WRVKL S KKLER WRVKL S KKLER WRVKL S KKLER WRVKL S KLTR IRLEL E KIKR WGMU S RNIR WQAWL S RNIR WQAWL S RNIR WQAWL S RNIR WQAWL S RNIR WQAWL S RNIR WQAWL S RIFRUWAQL KK PILRRUWRLLLQVQRR P SRIRSLITQ HAP HDIKXHFARE OVO	K VYYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHPIASL SIPFYIGHVRAHSGLPOFIAHG MFPVFITHVRAHSGLPOFPLAGG QEKFYIGHIRGHTGLPGPLAGG NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPOFFTEG KEKVYLAWVPAH KGIG GI KEKIGVHWVPGH KGIP QH KEIVYFAWVPGH KGIP QH KELVYFAWVPGH KGIP QH KELVYFAWVPGH KGIP GI KELSIHCPGPGKGHSAEARGH QHQIKWEWVKGHAGHPE H RGRLGLSRPLLRLPFRPTTGRTS EFDFDIKYIKGK EH EYNFTVEYLKGK EH EYNFTVEYLKGK EH EYNFTVEYLKGK EH YSFDVEHIAGT KM HYSFDVEHIKGT DD KVNLIGEHIPGF FN KIRIRLQFVFDHGGVK RM KIKIM WIPCHSGIK GH	NALTDALLI NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NEQADLL NDVADSQAT NEQVDKLVS VEEIDRYI NQLADEAAK VRHADQAAR NSVADALSR HFVADALSR HFVADALSR HFVADALSR HFVADFLSE NFFADFLSE NECCDEMAK KELADQAAK NECCDEMAK
HTLV-I HTLV-II BLV SRV-I IAP-18 MMTV HERV-K RSV HIV HIV2 VISNA EIAV MOMLV ECOLI HEPB 17.6 297 412 GYPSY CERV CAMV DIRS INGI I-FAC COPD TYEY	IFLD SKYLYH YL R IFLD SKYLYK YL H IWVD SKYLYS LL R IYTD SAYLAHSIPL IYSD SSYVNAVNL LYTD SKYVTG LF P IISD SAYVYQ AT R VYTD SAFYAK ML L IYTD SQYALGIIQ A IYVD SQYVHGISA S IVTD SQYACGIQ A IYVD SQYYHGISA S IVTD SYYCKNIT E YYTD SYYCKNIT E YYTD SYYCKNIT E YYTD SYYCKNIT F IGTDNSVYLSRKYTS FFW ISSDH QPLS W IASDH QPLS W IASDH QPLT F IRTDN KNFT H IRTDN KNFT H IRTDN KNFT H IRTDN THFK S IQTDNTTTLS Y IFSDSLSMLT ICSDSLSAVD IYEDNQCCIS GLLTDSRSTIS	TLALGTF QGRSS SLAIGAF LGTSA TLVLGAW LQPOP LETVAQ IKHIS ELEIAGI IRSSS EIETATL SPRT DVETALIKYSMDD KHQQEGV PST QPDK SES QPTE SES NWDEEV IRN GLGLEG PQN LLTSEG KEI NWKKRGKKTADKK LLGCAANWILRGT LYRMK DPNS LHNLK PCGA LFSMV NFSS AVADR NTNA FVNINLKGDRXGG FVNLNYKGDSKLG INNGGGQIQLJSV ALQTGP LAVTD SIQNTN NNSFY IANNPSCHK RAK	QAP FQALLPRLLSR HQT LQAALPPLLQG VPS XALLYKSLLRH ETAKLFLQCQQLIYNR RVANIFQKIQAALLNR KIYTELKHLQRLINR QLNQLFNLLQQTYRKR AAA F ILEDALSQR ELVNQIIEQL IK KIVNQIIEM IK PIQARIMELV HN PW WPJIQNI RE KNKDEILALLKALFLP PVKNVQLWQRLDAALG SFYYVPSALNPADDPS KLTR WRVKL S KLTR WRVKL S KLTR WRVKL S KLTR WRVKL S KLTR WQKU S KLTR WQWL S RNIR WQAWL S RNIR WQAWL S IFEQLWKQCL KK PILRRLWRLLQVQRR P SRIRSLITQ HAP HIDIKYHFARE QVQ FGGTKAMRLRD FVS	K VYYLHHVRSHTNLPDPISRL K TIYLHHVRSHTNLPDPISTF P AIFVGHVRSHSSASHPIASL SIPFYIGHVRAHSGLPGPIAHG HFPVFITHVRAHSGLPGPLAGG QEKYIGHIRGHTGLPGPLAGG NFPFYITYIRAHTNLPGPLTKAN SAMAAVLHVRSHSEVPGFFTEG KEKIYLAWYPAH KGIG GI KEAIYVAWYPAH KGIG GI KEAIYVAWYPAH KGIG QI KEIVYFAWYPGH KGIP QI KEIVYFAWYPGH KGIP QI KEIVYFAWYPGH KGIP QI KEIVYFAWYPGH KGIP QI KEIVYFAWYPGH KGIP QI KEIVYFAWYCH KGIP QI KEIVYFAWYCH KGIP QI KEIVFYFAWYCH KGIP QI KEIVFYFAWYCH KGIP QI KEIVFYFAWYCH KGIP QI KEIVFYFAWYCH KGIP QI KEIVFYFAWYCH KGIP QI KEIVFYFAWYCH KGIP QI QIQFDVEHIAGT DI KYNLIGENPGF FI KIRIRLQFYFDHCGVK RH KIKIM WIPGHSGIK GI MINVICLEYIPTE H	NALTDALLI NALTDALLI NETTDSLIL NNYVDQLL NQKADLATK NDLADKATK NDLADKATK NEQADLL NDVADSQAT NQEVDHLVS VEEIDRYI NQLADEAAK VRHADQAAR HIVADALSR VFADALSR VFADALSR VFADFLQE HHFADFLSR VFADFLQE HEVCDEMAK HELADQAAK QLADIFTK NTADWMTY

<u>表5</u> レトロウイルス由来の逆転写酵素および関連蛋白質25種類と、大腸菌 RNase H とのアミノ酸配列の類似性(Doolittle ら、1989)。中央の'ECOLI'の行が、大 腸菌RNase H のアミノ酸配列(Kanaya & Crouch、1983)である。 表から明らかな ように、*印のついた4つの酸性アミノ酸残基 (大腸菌 RNase H では、Asp 10、 Glu 48、Asp 70 および Asp 134) は 26種類全体にわたって完全に保存されている。



図17 RNase H 分子全体における活性部位の位置。RNase H 分子は水色のリボンモ デルで、活性残基は黄色の骨格に赤のドットによる分子表面を付けて示した。また 同定されたMg²⁺位置は黄色の+印に黄色のドットによる球面を付けて示した。

なお、緑の骨格はトリプトファンで、RNase H の持つ6個のトリプトファンすべてがこのようにほぼ1ヶ所に集まっている(第1章第3節)。

- 43 -

Amino acid re	Relative	
Native	Mutant	activities (%)
Lys-3	Cys-3	103
Ile-7	Cys-7	90
Thr-9	Ala-9	40
Asp-10	Ala-10	< 0.1
Asp-10	Asn-10	< 0.1
Asp-10	Glu-10	5
Thr-9,Ser-12	Ala-9,Ala-12	20
Cys-13	Ser-13	190
Glu-48	Ala-48	< 0.1
Glu-48	Gln-48	< 0.1
Glu-48	Asp-48	1.7
His-62	Ala-62	100
Cys-63	Ser-63	80
Asp-70	Ala-70	< 0.1
Asp-70	Asn-70	< 0.1
Asp-70	Glu-70	< 0.8
Ser-71	Ala-71	100
Tyr-73	Leu-73	100
Ile-78	Thr-78	80
His-83	Ala-83	110
His-114	Ala-114	30
Trp-118	Phe-118	36
Lys-122	Asn-122	25
His-124	Ala-124	2.5
His-127	Ala-127	20
Asn-130	Asp-130	13
Cys-133	Ser-133	50
Asp-134	Asn-134	90
Arg-138	Cys-138	60
Pro-144	Ile-144	20
Asp-148	Asn-148	(< 1.0)

Kanaya et al. (1990) J. Biol. Chem. 265, 4615-4621. etc.

<u>表6</u> Kanaya ら (1990b) による大腸菌 RNase H のアミノ酸置換による 比活性の 変化。3つの酸性アミノ酸残基 (Asp 10、Glu 48、Asp70) の変異体は 活性が完全 に無くなっている。Asp 与 Glu の置換体は RNase H 欠損株を用いることができな いため、<0.1にはならないが、活性はないと考えられる。 れた疎水結合をしている多くの残基が、大腸菌RNase H と逆転写酵素との間で保存 されている。

我々がX線解析で決定した 大腸菌 RNase H の立体構造をもとに、これらの逆転 写酵素の中から、ヒト免疫不全ウイルス(HIV-1)と、マウス白血病ウイルス (MuLV)の2つを選び、モデルを組み立てた(Nakamura ら、1991b)。この2つ は大腸菌 RNase H と、アミノ酸配列の類似性が 最も高いものである。興味深いこ とに、HIV-1のモデルでは、DNA/RNA ハイブリッドとの結合に関与していると 思われる、塩基性の突出部を欠いている。しかし、MuLVのモデルではそれは保 存されており、大腸菌 RNase H とほぼ同じ構造であった。 最近、HIV-1由来 の逆転写酵素の RNase H ドメインの結晶構造が、Davies ら(1991)により発表さ れたが、その結晶構造は、 この予想モデルの構造と本質的には同じであり、 大腸 菌RNase H の構造とも類似性の高いものであった。また金属結合部位も Doolittle の表に示される完全に保存される4つの酸性アミノ酸を中心に構成されており、我 々の同定したMg²⁺結合部位が、逆転写酵素の場合にもそのまま当てはまることが明 らかとなった。

なお、大腸菌 RNase H と H I V - 1 RNase H ドメインとの構造の比較について は、Jacobo-Molina & Arnold (1991) と Davies ら(1991) によっても、詳細な報 告がなされている。 第2章 大腸菌 RNase H 活性部位変異体の立体構造解析

第1節 活性部位変異体について

前章では、 主として大腸菌 RNase H の立体構造の詳細と、Mg²⁺結合部位の同定 について述べた。この章では、活性に関与すると推定される酸性残基を他のアミノ 酸に置換した活性部位変異体の立体構造について述べる。 Kanaya ら (1989) は、 大腸菌 RNase H の 大量生産系を構築し、部位特異的変異の研究を可能にした。こ の結果、様々な変異体蛋白質を比較的大量に精製し、それらの結晶を調達すること ができるようになった。

さらに、我々のX線解析で決定したMg²⁺結合部位にある3つの酸性アミノ酸残基、 Asp 10、Glu 48 と Asp 70 について、Kanaya ら (1990b) は部位特異的変異法に より、どの一つを置換した場合でも活性が完全に消失することを示した。Mg²⁺結合 部位が事実、触媒作用に関与していることを証明するには、変異導入によって立体 構造が全体的には変化せず、変化がMg²⁺結合部位に限定されることを明らかにする 必要がある。なお、この際の置換の方法については、①. 側鎖カルボキシル基の電 荷を消失させたもの (Asp → Asn または、Glu → Gln の置換)、②. カルボキシ ル基を持つ他のアミノ酸に置換したもの (Asp → Glu または、Glu → Asp の置換)、③. カルボキシル基をなくしてアラニンに置換したもの、の3つについておこ なわれたが、③では官能基の部分が空洞となっていることが予想されるし、②では、 側鎖の長さがあわず基質が結合部位にうまくフィットしないと予想される。

しかし①の電荷をなくした場合の構造変化は、予測が困難であり、興味深い。通 常、核酸との接触面は塩基性のアミノ酸に富んでいるが、4つもの酸性アミノ酸が、 カルボキシル基同志の反発が予想されるわずか4Å程度の距離を保ちながら、この 塩基性の接触面の中に集中的に存在していることは、活性発現機構の観点からも興

- 46 -

味がある。このような環境下で、側鎖の電荷をなくすことがこの構造の均衡にどの 様な影響を与えるかを調べるため、結晶構造解析を行った。

以上の観点から、Mg²⁺結合部位に存在する3つの酸性残基をそれぞれアミド化した変異体、D10N、E48Q および D70N の3つを構造解析の対象として選定した。

第2節 活性部位変異体の立体構造

この研究で扱った3つの変異体は、分子の骨格構造については、いずれも大きな 変化は示さなかった。これは、前節で示したように、変異体と野生型との立体構造 を重ね合わせたときのr.m.s.d.の値が小さいことからも明らかである。したがって、 Asp 10→Asn、Glu 48→Gln、Asp 70→Asn のアミノ酸置換では、分子全体に構造変 化を及ぼすことはないと結論される。この事実は、前述(第1章、第10節)のアミ ノ酸置換を行っても酵素の基質に対するKm 値は変わらなかったという Kanaya ら (1990b)の実験結果をうまく説明するものである。

D10N 変異体では、側鎖の構造変化は、置換導入した Asn 10 近傍のみに見られ た。この様子を図18に示す。主鎖原子のみを用いて D10N 変異体と野生型との構造 の重ね合わせを行った場合、いちばん変化の大きかった原子は、 Asp 10 のO 61原 子で2.31Åであった。このカルボキシル基の構造変化により、もう一方のN 62原子 も1.53Å移動し Asp 70 の側鎖カルボキシル基のO 61原子(この原子の野生型に対 するシフトは0.69Å)に2.86Åの距離に近づいている。すなわち、置換によって新 たに導入された Asn 10 の側差のアミド基が Asp 70 のカルボキシル基に引き寄せ られて、新たな水素結合が生じたものと考えられる。この水素結合の形成によって、 側鎖がゆらいでいることが予想される溶液中の立体構造においては、活性残基であ る2つの酸性アミノ酸残基、 Asp 10 と Asp 70 の自由度に制限が加えられ、さら に基質が接するポケットの形成が妨げられてしまうことが容易に予想される。



図18 RNase H D10N 変異体の活性部位近傍の構造変化。 青色の骨格が D10N 変異 体の構造で、赤色が野生型の構造である。置換導入された Asn 10 のアミド基が、 Asp 70 のカルボキシル基に近づき、新たな水素結合が形成された。同時に Asp 70 の主鎖も Asp 10 の方に引き寄せられ、βAとβDの開裂部が狭くなった様子がわ かる。 酸性残基である Asp 10 を中性化することにより生じた電荷の均衡の崩れは、立体構造上、前述のような変化として表現された。第1章、第9節で述べたように、 Asp 10 のカルボキシル基はMg²⁺に直接的に配位しうる距離にある。 したがって、 Asp 10 の野生型における側鎖の方向は、 活性に重要な意味を持っていることが、 さらに強く示唆される。

また、RNase H 分子全体のフォールディングからみた場合、 Asp 10 と Asp 70 の2つは、それぞれ平行に走る2本のβ鎖、βAとβD、のC末端側の開裂部に存 在する。今までにβシートと活性部位との位置関係は Rossmann fold により 説明 されているが、この酵素の場合もその関係はうまく当てはまっている。図18に示さ れる活性部位の構造変化を見た場合、この2つのβ鎖はアミノ酸置換により、僅か に近寄っている。このβシートの開裂部は、DNA/RNA ハイブリッド2重鎖との結合 に際し、重要なポケットの1つを構成しているものと推測される。したがって、こ のアミノ酸置換は、①ポケットの幅を狭め、②新たに形成された水素結合でポケッ トを埋めてしまった、と解釈される。

このような構造変化は、変異体 D70N の場合でもみられた。変異体 D70N の活性 部位近傍の構造変化を図19に示す。 D10N の場合と同じように、 Asp 10 とAsn 70 の側鎖どうしは近づき、 Asp 10 の O $_{\delta_1}$ と Asn 70 の N $_{\delta_2}$ が2.97Åの距離に近 づき、水素結合を形成した。しかも、2つのβ鎖は僅かながら近づきあっており、 変異体 D10N での考察が、そのまま当てはまっていることがわかった。

さらに、変異体 E48Q の活性部位付近の構造変化を図20に示す。グルタミン酸側 鎖は、アスパラギン酸側鎖よりメチレン基1個分長いため、変異体導入による構造 変化は さらに顕著に現れた。野生型では、 Glu 48 側鎖は その両隣に位置する、 Asn 44 と Ser 71 の側鎖と水素結合をつくり、 固定されていたにもかかわらず、 新たに導入された Gln 48 の側鎖アミド基は、その近傍の Asp 10 のカルボキシル 基と新たに水素結合を形成した。それにともない、その両隣の Asn 44 や Ser 71

- 49 -



図19 RNase H D70N 変異体の活性部位近傍の構造変化。 青色の骨格が D70N 変異 体の構造で、赤色が野生型の構造である。置換導入された Asn 70 のアミド基が、 Asp 10 のカルボキシル基に近づき、新たな水素結合が形成された。 またこの方向 からは、βAとβDのいずれもやや内側にシフトし、開裂部が狭くなった様子がわ かる。



図20 RNase H E48Q 変異体の活性部位近傍の構造変化(我々の得た野生型の構造 を初期構造として、北大の田中研究室で解析された結果)。青色の骨格が E48Q 変 異体の構造で、赤色が野生型の構造である。置換導入された Gln 48 のアミド基が、 Asp 10のカルボキシル基に近づき、新たな水素結合が形成された。双方の主鎖も局 所的に内側にシフトしている。また両隣にある Ser 71 や Asn 44 との水素結合の 仕方も変化し、Gln 48 では3つの隣接するアミノ酸側鎖に支えられている。 と Glu 48 との水素結合のしかたも変化した。

以上、大腸菌 RNase H の3つの活性部位変異体について、 それぞれの立体構造 を決定した。それらの構造を見ると、いずれも全体のフォールディングは変わらず、 Mg²⁺近傍にあるアミノ酸側鎖のみ、そのコンフォメーションが変化した。置換導入 された Gln や Asn の側鎖は Asp 10 のカルボキシル基に近づいており、野生型の 大腸菌 RNase H では、反発しあってバランスを保っていたMg²⁺部位近辺のこれら の酸性アミノ酸残基が、その1つの電荷がなくなることで、場のバランスが崩れた ものと考えられる。これらの側鎖の構造変化は、Mg²⁺と酸性アミノ酸残基の静電的 結合力を減少させ、酵素活性を失わせたのだろうと思われる。さらに、これらの結 果は、RNase H の機能に対するMg²⁺の役割を暗示している。つまり、野生型では、 Mg²⁺が近づくと、負の電荷が相殺されるため側鎖間の反発力が減少し、側鎖の変位 が生ずるものと予想される。もし、前述のアミノ酸変異が導入されれば、このよう な活性にとって不都合な変位が生ずることが予想される。1つのアミノ酸残基の負 電荷が失われると、Mg²⁺結合の際にカルボキシル基のコンフォメーションを、触媒 活性に対し、安定な状態から不安定な状態にしてしまうかも知れないからである。

また、3つの場合のいずれも、アミドの導入により、 Asp 10 の側鎖が新たな水 素結合をつくり、固定されていることがわかった。前章で述べたようにMg²⁺が導入 された際に配位し得るのが Asp 10 のカルボキシル基と Gly 11 の主鎖カルボニル であるが、この Asp 10 側鎖が固定されることは興味がある。つまり、Mg²⁺との結 合に関連づけて考えると、アミノ酸置換により、酵素-Mg²⁺-基質の連結が起こり にくい構造になってしまったものとも考えられる。事実、変異体の1つ D70N では Mg²⁺結合能が無くなってしまっていることがNMRで確認されている(Oda ら、私 信)。 第3節 2つの非同型な結晶中での構造の比較

- His 124 近傍のループの柔軟性 -

前述したように、 D70N 変異体の格子定数は野生型のものとa軸が5%も異なり 明らかに非同型な回折パターンが観測された。

前節で述べたように、全体のフォールディングは D70N と野生型では変わらなか った。しかし、もう少し詳細にみると、電子密度の認識しにくいN, C未端部を別 にすれば、唯一明確に構造の異なる部分は Val 121 から Glu 129 までのループで ある。 この構造の差異は、ループ部分の中央部に属する His 124 の側鎖方向が、 2個の構造間でほぼ反対方向を向いていることから明らかである(図21-1、2、 3)。すなわち、野生型の構造では His 124 は分子の外側 つまり溶媒側を向いて いるのに対して、 D70N では、この残基は分子内部を向き、Mg²⁺結合部位を形成し ている残基の1つである Asp 70 (この変異体では Asn 70)に、約4 Åの距離に まで近づいている。さらに興味深いことは、この D70N 変異体と同じ格子定数であ る Yang ら(1990)の解析した野生型の大腸菌 RNase H の His 124 も、ほぼ同じ 構造をとっていることである。したがってこの His 124 付近のループは、 結晶内 で分子間相互作用の影響を受けやすいという点で柔軟であると言える。

Kanaya ら (1990b) による部位特異的変異法の結果では、この His 124 は、 ア ミノ酸置換すると比活性が 2.5%と小さな値まで落ちるが、しかし完全にはなくな らないため、 活性必須な残基とはいえなかった。 すなわち His124 は、 Mg²⁺ 部 位 (内側)を向くと RNase H の活性を補助的に促進させるはたらきをもつが、 溶 液状態では、内側と外側を向くという、両方の状態を取っており、そういうゆらぎ の大きい状態ではいつも促進機能を果たしているとは限らないと考えられる。次章 で詳しく述べるが、この His 124 は DNA/RNA ハイブリッド結合モデルにおいては、 分子の外側を向くと核酸のリン酸基と作用できる部分にある。したがって、現在ま での RNase H の構造解析からは、触媒活性に直接関与しないまでも、基質との結



図21-1 RNase H の野生型と D70N 変異体の全体構造の比較。 図に示すのはC a 原子のみで、それぞれの主鎖原子を用いて重ね合わせを行った後のものである。青 色が D70N 変異体、赤色が野生型である。変異体 D70N で明瞭に認識されないN末 端とC末端の残基は比較から省いた。His 124 付近のループ約10残基分に構造の違 いがみられる。 なお活性中心を示すため、 野生型の構造にのみ Asp 10、 Glu 48、 Asp 70、Asp 134 の4つの酸性アミノ酸残基の側鎖を示した。

- 54 -



図21-2 野生型 RNase H の His 124 付近のループの電子密度図。MIRより得た初期位相による電子密度図では認識しにくかったが、最終構造による電子密度図では各側鎖とも明瞭に認識できる。紫色の骨格は対称な隣接分子を示す。



図21-3 D70N 変異体の His 124 付近のループの電子密度図。黄色の骨格がD70N 変異体で、赤色の骨格が野生型の構造である。 この His 124 の側鎖の位置は、こ の D70N と同型である Yang ら (1990)の野生型の場合と一致した。

合に 何らかの重要な役割を果たすことは、 十分考えられる。 なお、 Davies ら (1991)により解析された、活性を持たないHIV-1RNase H ドメインでは、 こ の近辺のループはゆらぎが大きく構造は決められていない。

•

第3章 活性中心と核酸結合部位の同定

第1節 DNase I との比較

我々は、 大腸菌 RNase H の立体構造を、すでに立体構造の解明されているリボ ヌクレアーゼ、すなわち RNase A や、RNase T₁、barnase および staphylococcal nuclease などとの 構造比較を試みたが、類似性を見いだすことができなかった。 これらのリボヌクレアーゼは単鎖のRNAを基質としている点で、RNase H の活性 とは明確に異なっている。

RNase H が、2重ラセンの核酸を基質としていること、核酸の塩基配列に対しそ れほど基質特異性のないこと、活性に金属イオンを必要とすることを考えあわせれ ば、前述のリボヌクレアーゼよりも、すい臓由来の deoxyribonuclease I (DNase I)との類似性が高く、構造比較により 有用な知見が得られるかも知れない。しか も、DNase I については、基質であるDNA2重鎖(14-mer)との複合体との結晶構 造がわかっている。

DNase I は、257 個のアミノ酸残基からなる、単一ポリペプチド鎖の酵素である が、立体構造上からは擬対称の2回軸で対応づけられる2個のドメインより構成さ れる分子である。 我々は、この酵素と RNase H の立体構造を比較した結果、それ ぞれの酵素分子のβシート・モチーフと活性金属との位置関係に1つの共通性をみ いだすことができた。

図22に RNase H と DNase I の β シートの構造を比較した結果を示す。DNase I は、1つの活性金属であるCa²⁺が、擬対称で関係づけられる2枚の β シートに挟ま れる構造になっており、この図ではその β シートを開いて2段に示した。したがっ て、上下2段に見られる白丸は1個のCa²⁺を示す球である。RNase H の方の白丸は Mg²⁺の位置を示している。 なお、DNase I の残基番号の定義は Oefner と Suck

- 57 -



図22 RNase H と DNase I のゟシート・モチーフの比較。

- 58 -

.

(1986) の J. Mol. Biol. 605. p616 Fig. 9 によった。

この図において、それぞれ左から1から4番目のβ鎖の方向の順序に注目すると、 RNase H のCBADと、DNase I のNPACまたはLMKJの方向の順序が同一で ある。また、βシートとしてのねじれの方向も同一である。活性金属とβシートの 位置関係については、いずれも左から3番目と4番目のβ鎖のカルボキシル末端側 と左から2番目のβ鎖のアミノ末端近傍に位置する。すなわち、RNase H ではA鎖 とD鎖のカルボキシル末端側およびB鎖のアミノ末端側であり、DNase I ではA鎖 とC鎖、あるいはK鎖とJ鎖のカルボキシル末端側、およびP鎖あるいはM鎖のア ミノ末端近傍に活性金属が位置している。またいずれの酵素の場合も、β鎖を構成 しているアミノ酸残基の数は、おおよそ7から10個である。

DNase I とDNA2重鎖の複合体の結晶構造では、DNA2重鎖はβシートにほ ぼ垂直に結合しており、 RNase H の場合も、これと同様な様式で DNA/RNA ハイブ リッドと結合するものと予想される。

また Suck ら (1986) によって予想された、DNase I の加水分解反応のメカニズ ムを図23-1 に示した。 このモデルは、活性に必須な His 131 によって活性化さ れた水分子が、Ca²⁺と結合しているリン酸基を求核攻撃して切断すると、説明して いる。この機構を われわれの決定した RNase H の立体構造とMg²⁺結合位置に適用 したものが、図23-2である。このモデルに現れる3つのカルボキシル側鎖は、活 性に必須なAsp 10、Glu 48、Asp 70 のもので、 現在の立体構造からは、Mg²⁺に直 接結合しているのが Asp 10 で、 Glu 48 と Asp 70 は求核攻撃する水分子の活性 化をするか、電荷リレー系で電子の受容体になるものと考えられる。

RNase H の場合、5個のヒスチジン残基のうち 活性部位に最も近い His 124 で も、Mg²⁺位置から11Åも離れており、またこの残基のアラニン変異体では比活性が 2.5 %残ることから必須であるとはいえないため、この RNase H の His 124 を、 DNase I のモデルのヒスチジンにあてはめることはできない。



<u>図23-1</u> Suck ら(1986)により提唱された DNase I の加水分解反応のメカニズ ム。活性に必須な His 131 に活性化された水分子が、リン酸を求核攻撃する。



図23-2 我々の得た RNase H の立体構造をもとに推定される RNase H の加水分 解反応のメカニズム。 Mg²⁺に直接配位するのが Asp 10 で、Glu 48 と Asp 70 が 水分子を介してリン酸を切断すると考えられる。 His 124 は、 H124A の比活性が 2.5 %残り、活性に重要な影響を与えるが、活性に必須な残基とは言い切れないた め、このモデルには入れていない。立体構造上からは、Mg²⁺位置から約11Å離れて いるが、His 124 が活性に重要な影響が与えられるのは、この残基の属しているル ープが、結晶のパッキングが変わるとコンフォメーションが変わるくらい柔軟性の 高いもので、それによって活性部位に近づき得るためと考えられる。 第2節 HIV-1 RNase H との比較(金属結合部位)

一方では、Yang ら (1990) は、RNase H と DNA polymerase I (Derbyshire ら、 1988; Freemont ら、1988)の活性部位の環境の立体構造に基づき、この酵素には、 Mg²⁺が同時に占有する2つの別々の金属サイトが存在する可能性を強調している。

最近、 Davies ら(1991)はHIV-1逆転写酵素の RNase H ドメインの2.4Å 分解能における結晶構造を発表した。さらに、彼らはその酵素のM2+結合部位も同 定しているが、45 mM のMnCl2 水溶液に浸漬した結晶と native 結晶の回折データ から差のフーリエを計算した結果、同じ領域に2つのMn²⁺イオンがみられた。彼ら は、金属サイトが1つであるという我々の結果との違いについて、両者の結晶のパ ッキングの違いによって説明されると考えている。 我々の結果からは、 大腸菌 RNase H のMg²⁺結合部位には、2つの金属イオンが入るだけの十分な空間はある。 しかし、彼らの論文によれば、Asp 549 (大腸菌 RNase H では Asp 134 に相当)に 隣接する第2の金属サイトは、第1の金属サイトに比べてその占有率が小さい。大 腸菌 RNase H の場合、最高でも10mM 濃度のアルカリ土類金属のイオンを含む水溶 液中で浸漬したのに対し、HIV-1RNase H の場合は、遷移金属であるMn²⁺をし かも 45mM という高い濃度条件下で浸漬している。アルカリ土類金属より配位結合 を形成する傾向の強いMn²⁺を、しかも 45mM という高い濃度で浸漬したことが、こ の第2の金属サイトを形成させた、とも解釈できる。 さらには、大腸菌 RNase H の Asp 134 は、HIV-1RNase H では 第2のMn²⁺位置に隣接するAsp 549 に相 当するが、この大腸菌 RNase H の Asp 134 を、Asn に置換しても大きな活性の減 少はみられず、比活性で 90 %も残っている (Kanaya ら 1990b)。このアミノ酸置 換により負の電荷が喪失しても、活性にはほとんど影響がないことから、この第2 の金属結合部位が活性に必須のものであるとは考えにくい。以上から大腸菌 RNase H の活性には、1分子に対して 1個のMg²⁺サイトが必須である、と考えた方が良

- 61 -

さそうである。 大腸菌 RNase H に結合したMg²⁺、Ca²⁺、Ba²⁺が単一の結合部位し か持たないことは、NMRスペクトル変化の Hill plots からも、裏付けられてい る(Oda ら、1991)。

我々が同定したMg²⁺結合部位は、DNA/RNA ハイブリッドが結合していない状態の ものであり、現在のMg²⁺部位近傍の環境は、核酸が存在しているときの影響を必ず しも反映しているわけではない。加水分解の際には、ポリヌクレオチド主鎖のリン 酸部がMg²⁺の配位に加わるものと、我々は仮定している。しかしながら、核酸のな い状態での結晶形では隣接分子の Lys 86、Lys 87、Arg 88 という塩基性に富んだ ループが酸性アミノ酸残基の集まっている空のサイトにイオン性の相互作用により はまりこんだ状態になっている(図14)。このようなMg²⁺部位と隣の分子の塩基性 ループとの接触状態が、ある種の金属を高濃度で導入した際に、約2 Aの格子定数 の変化が起こる理由と考えられる。おそらくこの電荷による結晶内の分子間の相互 作用が、 Yang ら(1991)が議論しているI型からII型への結晶の相転移に関係し ているものと思われる。

RNase H と DNA/RNA ハイブリッド・合成オリゴマーとの、NMRによる結合実 験(Oda ら、私信)からも、我々の同定したMg²⁺結合部位は正しいことが裏付けら れた。さらに、Kanaya ら(1990b)による、部位特異的変異の実験から、Mg²⁺結合 部位近傍の酸性アミノ酸残基のうち、3つ(Asp 10、Glu 48、Asp 70)が触媒活性に 必須であることが示されている。

第3節 静電ポテンシャルの分布

RNase H の基質であり、 DNA-A型に類似した構造を持つ DNA/RNA ハイブリ ッド2重鎖は、酸性のリン酸骨格が表面に露出している。この核酸との結合のため には塩基性の酵素表面が必要であり、さらに金属イオンが活性に関与する場合は、 それを仲介して核酸に結合するため、酸性の分子表面も必要となる。 RNase H の場合には、Mg²⁺が活性に必要なことから、正負両方の電荷について、 それぞれ極在化している部分が必要である。このような予想をもとに、我々のX線 構造解析で得た立体構造を用いて、Nakamura & Nishida (1987)の方法により静電 ポテンシャルを計算し、それを solvent accessible surface 上に色分けして表示 したものを図24に示した。ここで、赤色は -0.1V より低い負の部分を示し、青色 は+0.1Vより高い正の部分を示す。他の色は、-0.1Vから 0.1Vまでのあいだに属 し、赤色から青色へと順次遷移していくように示されている。ここで、中性の 0V レベルは黄色で示される。なお、大腸菌 RNase H の示適pHは 8.0 であるため、静 電ポテンシャルの計算に際しては、ヒスチジン残基はプロトン化していないものと して計算してある。

RNase H 分子にはDNA-A型の立体構造にうまくフィットするような、円筒形 の大きなくぼみがある。図24ではこのくぼみが図の下側から図の前方に上ってくる ように表示してある。図からも明瞭なように、このくぼみの下半分には負に帯電し た領域が、またこのくぼみの上半分には正に帯電した領域がそれぞれ存在している。 これらの領域はそれぞれさきに述べた、Mg²⁺結合領域と塩基性の突出部に相当する。 疎水性の残基が分子表面に露出している部分は、ほとんど中性の黄色で示されてい る。このように、静電ポテンシャルの分布のみからでも、 基質である DNA/RNA 2 重鎖と結合する領域は、かなり明瞭に予測することができた。

また、RNase H には、5本のαヘリックスがあるが、αヘリックスは双極子を持ち、そのN末端は正になっている。RNase H では、ヘリックスαI、αN、αVのN末端側は、すべて、このくぼみの方を向いている(αIIとαIIIはくぼみの上にある)。多くの核酸結合蛋白質で、αヘリックスの双極子とリン酸基の負の電荷との相互作用が基質との結合の要因になっており、RNase H のヘリックスの構造もこれに適合するものと予測される。



<u>図24</u> RNase H 分子の solvent accessible surface 上の静電ポテンシャルの分布。 赤色が-0.1V、青色が+0.1V、黄色が中性である。その他は補間して配色した。手 前上部の塩基性突出部と下側の奥のMg²⁺結合部位が明瞭に見える。 第4節 核酸認識部位

前節で述べたように、RNase H の分子表面には、DNA/RNA ハイブリッド2重鎖の 外径にうまくフィットするようなくぼみがあり、このくぼみの両端にMg²⁺結合部位 と塩基性の突出部位とが存在する。全体的にはまっすぐな 円筒形の DNA/RNA ハイ ブリッドが、 この2つを結ぶ線に沿って結合するものと思われる。 この推定は、 RNase H 分子表面上の静電ポテンシャルの分布と、一連の部位特異的変異体の実験 結果から導かれる。このようにして、予想された認識部位をもとに構築した大腸菌 RNase H と DNA/RNA ハイブリッド2重鎖との結合モデルの模式図を 図25に示す。 この仮想的な結合様式は、すでに立体構造解析された すい臓の DNase I とDNA 2 重鎖複合体の結合様式と類似性がある (Suck ら、 1984; Suck & Oefner、1986; Oefner & Suck、1986; Suck ら、1988)。 さらに2つのヌクレアーゼ間のβシート モチーフの構造、とくに金属結合部位との関係には類似性がある。

図26に RNase H と 21-mer の DNA/RNA ハイブリッド2重鎖とで構築した複合体 モデルをグラフィックス上で示した。ここで DNA/RNA 2 重鎖の構造として Arnott ら (1986)により、X線繊維回折の結果から導かれたものを用いている。まず、大 腸菌 RNase H と DNA/RNA ハイブリッド、それぞれの単独な構造をグラフィックス 上で操作しながら噛み合わせる。ついで、それらの主鎖原子を固定し、プログラム AMBER (Weiner ら、 1986)による力場計算を用いて、接触の良くない側鎖の コンフォメーションをエネルギー最小化により適合させた。現在の複合体モデルで はDNA鎖は C2'-endo-puckered furanose 環に、RNA鎖は C3'-endo-puckered furanose 環に、それぞれ なっている。 なお、溶液の状態と塩基配列によっては、 DNA鎖もRNA鎖も、 C3'-endo-packered furanose 環をとっているという報告 もある(Wang ら、1982)。また、ハイブリッド2重鎖として、A型DNAの構造を 用いても複合体モデルを構築したが、以下に述べる認識部位の詳細な特徴は、本質 的には変わらない。



図25 RNase H と DNA/RNA ハイブリッドの結合を予想した模式図。 核酸の方向を 逆にした場合、2重鎖を大きく歪ませないとうまくドッキングできない。NMRの 核酸結合の実験からは、この図にみられる接触面に多くの化学シフトが観測された (Nakamuraら、1991b)。

- 66 -



図26 RNase H と DNA/RNA ハイブリッド21-mer との結合予想図。この2つの構造 のドッキングの際、RNase H の側鎖は衝突を避けるようエネルギー的に最適化され ている(Nakamuraら、1991b)。 活性残基の側鎖は赤色で、本文中で議論されている 側鎖は燈色で、その他の側鎖は紫色で示した。
この仮想的な複合体モデルに基づく核酸認識部位の正しさは、 最近、¹Nと¹⁵N を用いたNMRの実験から裏付けられた。この実験は、 RNase H 水溶液中に DNA/ RNA ハイブリッド(合成オリゴマー)を混合させ、化学シフトの変化を調べたもの で(Nakamura ら、1991b)、 化学シフトの大きなアミノ酸残基が、この結合モデル の周囲に集中して観測された。 また、Yang ら(1990)も我々とは独立に、この酵 素と核酸のドッキングモデルを発表しているが、いずれの結合様式も本質的には同 じである。

現在の1.48Å分解能で精密化された立体構造では、 酵素と DNA/RNA ハイブリッ ッドとの結合について、さらに詳細に考察することができる。まず、Mg²⁺結合部位 近傍には、グリシンに非常に富んだ領域がある。この領域のアミノ鎖配列は、G11-S12-C13-L14-G15-N16-P17-G18-P19-G20-G21-Y22-G23 であり、βAとβBとの間の 長い、柔軟なループを形成している。 このループが DNA/RNA ハイブリッドとの結 合に関与していることはNMRの測定結果から裏付けられている。現在の精密化さ れた立体構造においては、このループは、その内側にある Cys 13 のSrHが Gly 15 と Gly 18 の主鎖カルボニル酸素と 水素結合を形成することによって、構造が 安定化されている。Kanaya ら(1991b)は、我々の立体構造をもとにジスルフィド の導入可能な残基を探索し、 Cys 13 近傍に存在している Asn 44 をシステイン残 基に置換した変異体をつくり、その結果、人工的なジスルフィド結合が導入された。 新たなジスルフィド基の形成により酵素活性は大きく減少したが、これはこのルー プの僅かなコンフォメーション変化が起きたり、あるいはこのループの柔軟性が失 われたため、DNA/RNA ハイブリッドとの結合が妨害されたものと考えられる。また、 Walker ら (1982)は多くの核酸結合蛋白質にグリシンの豊富な領域があることを見 いだしており、この事実は RNase H にも 適用されるかもしれない。

複合体モデルを構築する際、まず、RNA鎖骨格の切断結合をMg²⁺結合部位に固 定した。RNA鎖上で切断部位から5'側へ隣にあるリン酸は、α IIのN末端側に

- 68 -

ある Ser 71 と Gln 72 の主鎖原子に接近している。またRNA鎖上で切断部位か ら5' 側へ2ヌクレオチド分離れたリン酸は、Lys 122 のアミノ基と Asn 44 のア ミド基に近い。さらに3' 側では、切断部位の3' 側の近傍にあるRNA骨格は、 グリシンの豊富な柔軟ループ (Gly 11-Gly 23) 上に存在する、Ser 12、Cys 13、 Leu 14 および Gly 15 に近づいている。

この複合体モデルにおいては塩基性の突出部は主にRNA鎖と相互作用している。 切断部位から5'側に 10 ヌクレオチド分だけ離れたRNA鎖のリン酸は、Asp 94 主鎖のイミノ基に面している。また、いくつかの塩基性のアミノ酸側鎖は、静電的 にRNA骨格と作用でき得る位置に存在する。

一方、DNA鎖のほうであるが、まず切断部位と塩基対をつくっている部分は、 グリシン豊富な領域にある Asn 16 や、 α I の Asn 45、および塩基性の突出部の 末端部の Lys 99 に近づいている。また、そのさらに3'側で、DNA鎖はα I と α IIの間の折れ曲がり付近にある Gln 76、Gln 80、および Trp 81 の側鎖と 相互 作用している。 Lys 122 の近傍の His 124 は、主溝側からDNA骨格に近づき、 またαV上の Arg 138 は、その His 124 と作用しているリン酸から5' 側へ2ヌ クレオチド分離れたリン酸と作用している。以上の様に、ヘリックスαVから塩基 性の突出部までの認識部位を完全にカバーするには、DNA/RNA ハイブリッド2重鎖 の長さは、約2巻き分が必要である。

Cys 13、Asn 16、Asn 44、Asn 45 および Gln72 を、それぞれアラニンにアミノ 酸置換すると、基質に対する親和性が低下している (Nakamuraら、1991b)。さらに、 Lys 122 を Asn に、His 124 を Ala に、 Arg 138 を Cys にアミノ酸置換すると、 野生型に対する比活性は、それぞれ、25 %、2.5 %、60 %に低下した(Kanaya ら、 未発表)。これらの実験事実は、この複合体モデルの正当性を裏付けている。

我々は現在まで、さまざまのオリゴヌクレオチドを用いて、浸漬法や共結晶化法 を行い、差フーリエ図上で、直接、結合した基質の電子密度を求めようとしてきた。

- 69 -

しかしながら、現在のところこれらの実験はうまく行かず、基質の確実な電子密度 は同定されていない。(d)のところで述べたように、Mg²⁺結合部位が隣接分子の 塩基性の突出部に埋められていることが、オリゴヌクレオチドのMg²⁺部位に近づく のを妨げているようである。もう1つの問題として、導入しようとしているオリゴ ヌクレオチドの RNase H の認識部位に対する親和性は、かなり小さいことが予想 される。この予想は、溶液状態でオリゴヌクレオチドの酵素に対する親和性が小さ いことを示すNMRの実験結果と一致する。複合体の形成のためには、先述したよ うに かなり長い DNA/RNA ハイブリッド2重鎖が必要なものと思われる。

- 1. 大腸菌リボヌクレアーゼHの立体構造をX線構造解析により決定し、1.48Å分 解能の精密な立体構造を得た。この構造決定は、RNase H の機能および性質を改 変するための基礎的な知見を与えた。
- 2. この酵素の触媒活性に必須なMg²⁺の結合部位を同定した。これにより、触媒活 性に関与するアミノ酸残基の位置関係が明らかにされた。また、これらの残基が 逆転写酵素 RNase H 様ドメインにおいて、1 次構造上 完全に保存されている生 物学的意義が明らかとなった。
- 3.活性アミノ酸残基(Asp 10、Glu 48、Asp 70)を置換した3種類の変異体について、それぞれの立体構造を決定し、それらの機能と構造の関係について論じた。

本研究を始める機会を賜り、終始励まして頂いた大阪大学薬学部 池原森男 名誉教 授に感謝いたします。また本研究に際し終始一貫して有益なご討論やご助言を賜り ました、蛋白工学研究所 森川耿右 第一研究部部長に感謝いたします。

研究に際し有益なご助言を賜りました、蛋白工学研究所 松島正明 第一研究部主席 研究員、Arno Pähler博士(現 大正製薬)、ならびに 三菱化成総合研 究所 松崎尹雄 部長研究員に感謝いたします。

大腸菌RNase H の試料を提供して頂いた、蛋白工学研究所 金谷茂則 第四研究部主 席研究員、 モデリングやコンピュータグラフィックスでご討論頂いた、 中村春木 第二研究部部長に感謝いたします。また、結晶化をして頂いた、第一研究部・宮川 麻由子 (現 三菱化成)、石川桃代、有吉真理子、奥村美香の各研究員、FAST の使用に際し便宜を図って頂いた第一研究部・団野真紀研究員、ディスタンスマッ プを求めて頂いた第五研究部・藤博幸博士、NMRのデータをもとにご討論頂いた 第二研究部・小田康司博士、プログラムの使用に際し便宜を図って頂いた第一研究 部・清水敏之研究員に感謝いたします。

研究に際し、ご助言、ご助力、ご支援して下さいました蛋白工学研究所第一研究部 のみなさまに感謝いたします。

最後に、本論文をまとめるにあたり、有益なご討論を賜りました、大阪大学薬学部 冨田研一 教授に感謝いたします。

実験の部

第1章の実験

第1節 RNase H の結晶

大腸菌 RNase H の至適 pH は 8.0 で あるが、 pH 9.0 以上で溶解度が急激に 下がる性質があり、これを利用して結晶 化された (Kanayaら、1989)。20 ℃にお いて、 pH 9.0 のトリス緩衝液中より生 じた微結晶を連続的に seed することに より、大きさが 0.7×0.7×0.4nm の 正 方両錐形の結晶が、比較的容易に得られ 表7.大腸菌RNase Hの結晶学的データ

Space group	P21 21 21
Cell dimensions	
a (Å)	44.06
b (Å)	86.85
c (Å)	35.47
$V (A^3)$	135,730
Z	4
Vm (Å ³ /dalton)	1.94

る。4軸型回折計CAD4 (Enraf-Nonius) により決定した結晶学的データを表7 に示す。 また Matthews (1968)の式により得られる V_m は 1.94 であり、蛋白質 結晶に含まれる溶媒水の割合は 37% で、また非対称単位に 1分子をふくむ。この 結晶は 1.5Å分解能を越える範囲までX線を回折し、高分解能の解析に適した結晶 であった。

なお、第2章でも触れるが、Yang ら(1990)が解析に用いた RNase H の結晶は、 我々が解析に用いたものよりa軸の格子定数が、約2 A短いものであった(表14)。 彼らの結晶化条件は我々と異なり、結晶化の際の沈澱剤として硫酸アンモニウムを 用いたため、非同型の結晶が得られたものと考えられる。また、我々の条件でも、 重原子置換体の探索の際や、変異体の結晶化の際には彼らの格子定数と同型の結晶 が得られている。 結局、大腸菌 RNase H の安定な結晶形は、2種類あることがわ

- 73 -

かったが、彼らの結晶が 1.7 Å 分解能までしか観測されなかったことや、彼らのと 同形な結晶は高分解能側で回折データが少なかったことから、我々の結晶の方が規 則性がよく、分解能が高い、より精密な構造決定ができたものと考える。

第2節 X線回折強度データの測定

RNase H 結晶の 回折強度データの測定には、以下の2つの装置を用いた。

まず、 native 結晶 および重原子誘導体結晶の測定のためには、 4軸型回折計 CAD4 (Enraf-Nonius)を使用し、回折強度データの処理および補正計算には低 分子解析プログラムSDP (Structure Determination Package, Enraf-Nonius)を 用いた。

この際用いた X線源は封入管式のCu管球で出力は 50 kV, 40 mA 、またグラファ イトモノクロメータにより、単色化した。回折強度は ω -scan 法で測定し、走査速 度は1-3°/分であった。結晶と検出器の間の距離は 368 mm 、測定温度は15℃ であった。結晶の X線損傷による強度の減衰補正を行うため1.5 時間おきに、各軸 方向の標準反射を測定した。強度減衰の許容値は、強度値で測定開始時の約 20 % までとした。1 個の結晶に対しX線照射時間は 35 から 45 時間であった。North ら (1968)の方法で吸収補正を行うため、回転軸方向の5 個の反射についてそれぞ れゆ角を10°ずつ変えた反射を測定し、360°にわたる補間曲線を求めた。図27に吸 収補正曲線の一例を示す。それぞれの結晶間のスケーリングは構造解析用プログラ ムPROTEIN (Steigemann, 1974)を用いて行った。このためのデータとして 各分解能の境界付近の 約 300 個の反射を重複して測定したほか、比較的強度の大 きな4.1-3.8Å分解能内の 約 300 個の反射を共通に測定した。

native 結晶に対しては 精密化を行うことを考慮して、なるべく良質の強度デー タを得るよう努力した。すなわち、1.48Å分解能までの反射を6個の結晶を用いて 強度測定をした。これらをスケーリングした際の、結晶間の回折データのRmerse



図27 吸収補正曲線の一例

- 75 -

は、5.76%であった。1.48Å分解能では、測定反射数の理論値(23,254個)にたい する割合は 84.1%(F>1 σ (F)), 78.0%(F>2 σ (F)), 70.6%(F>3 σ (F)) であった。 良好な重原子誘導体を探索するために、native 結晶との isomorphous difference $\{\Sigma (|Fnati| - |Fderi|) / \Sigma |Fnati| \}$ の大きいものに着目し、試験的 に6Å分解能での差のパターソン図を描くことにより、主要な重原子位置を決定し た。その中から良好なものとして選ばれた重原子誘導体結晶に対し、 2.5Å分解能 の回折データを測定した。

また 1.8Å分解能での精密化用のデータとして、2次元検出器FAST (Enraf-Nonius)を用いて測定した回折強度データも用いた。X線源としては Eliott 社の 回転対陰極型X線発生装置GX21を用い、45kV、65mAの線源で測定を行った。さ らに、FASTの制御とデータ処理には、プログラムMADNES(Messerschmidt & Pflugrath、1987)を用いた。この装置の機構の上からは、1.8Åが測定できる分 解能の限界であった。結晶と検出器のあいだの距離は45mmで、検出器は入射X線の 方向に対して 45°傾斜させて測定した。走査時の回転軸は1つのみであり、また当 時は auto indexのプログラムは無かったため、blind region を補う必要があり、 それぞれa*軸とb*軸とを立てた2つの結晶のデータを測定し、それらを併合したも のを解析に用いた。この時のRmerge は 6.33%であった。走査速度は、1回の走査 に対し0.2°を75secで行い、走査した角度範囲は、1個目が100°、2個目が70°であ った。1.8Å分解能までの測定した反射数は、総数で 30,561個、独立な反射にして 10,018 個であり、理論値に対し 75.1%の回折データが観測された。なお、その後 の解析で用いたCAD4によるデータでは、おなじ分解能で87.4%の回折データが 観測されている。

第3節 重原子同型置換法による位相決定

RNase H の初期構造を得るための位相は、 重原子多重同型置換法 (MIR;

- 76 -

Multiple Isomorphous Replacement) により決定した。

RNase H に重原子を導入する場合、大部分の水銀化合物や白金化合物が結晶を変 成させたり、格子定数を大きく変化させ、良好な重原子置換体を得ることは容易で はなかった。そこで、水銀化合物に対し適当量の2-メルカプトエタノールを加え ることにより、重原子が特別強く反応する-SH 基にだけ結合するように制御し、こ の問題を克服した。図28に良好な例として、5-クロロマーキュリウリジン誘導体 の native 結晶との差のパターソン図と、異常分散効果によるF+, F-間の差のパ ターソン図を示した。これらの差パターソン図のハーカー面上には明瞭なピークが 見られ、異常分散によるピークとよく一致していた。約100 種類の誘導体を探索し、 それらの中から良好な誘導体として、EMTS(ethyl mercury thiosalicylate)、 HgCl₂、および5-クロロマーキュリウリジンを選定した。 これらの重原子誘導体 結晶のデータについては表8に示す。

重原子位置は、差のパターソン図のハーカー面から拾い、さらにクロス・ベクト ルの存在も確認した。それらをもとに、プログラムPROTEIN (Steigemann、 1974)を用いて初期位相を精密化した。重原子の副サイトを求めるためにクロスの 差のフーリエ図も併用した。この手順を繰り返すことにより、解析の進度の指標で ある figure of merit が 0.57 になった。 さらに、異常分散データおよび BaCl₂ 誘導体を加えて精密化を続けた結果、figure of merit の最終値は 0.71 になった。 この時の重原子バラメータを表9に示す。また、各分解能ごとの phasing power (位相決定力;重原子の構造因子と lack of closure error との比, $\Sigma f H/\Sigma E$) と figure of merit の分布を図29に示す。

最初の電子密度図は、これらの誘導体の 2.5Å分解能の反射データを用いて計算 した。さらに BaCl2誘導体の 3.5Åデータ、および HgCl2誘導体と5-クロロマー キュリウリジン誘導体の異常分散効果をとりいれた解析により、このMIRマップ はさらに改善された。ことに、重原子サイト近傍である N末端付近の Met 1 から

- 77 -



- 78 -

data set	resolution	numbe indepe reflec	r of ndent tions	numbe cryst	r of als	R merge	av ison dif	verage Norphou Eferenc	s phasing e power	R -Cullis
CAD4 (for M.I.R.)										
native	2.5Å	5,	000	1		\$				
EMTS	2.5A	4,	586*	4		5.00 %	1	.2.3 %	2.72	0.47
# HgCl2	2.6Å	4,	325	2		6.13 %	1	.1.3 %	3.59	0.41
<pre># 5HgCl-uridine(1)</pre>	2.5Å	4,	825	3		6.21 %	2	20.5 %	3.12	0.44
<pre># 5HgCl-uridine(2)</pre>	2.5Å	4,	759*	4		5.36 %	1	.7.3 %	2.58	0.55
BaCl ₂	3.5Å	1,8	889	1		\$	1	.0.9 %	0.88	0.81
FAST (for refinement	s)									
native	1.8Å	10,0	018	2		6.33 %		—	_	—
Resolution (Å) 14.81	8.70	6.15	4.76	3.88	3.28	2.84	2.50	total	
No. of refle	ctions 33	112	241	411	630	890	1192	1491	5000	
Figure of me	rit 0.97	0.96	0.92	0.88	0.81	0.75	0.63	0.60	0.71	

* Bijvoet pairs were collected for anomalous dispersion data # soaked with 2-mercaptoethanol R merge = Σ Σ |Ih, j-<Ih>| / Σ <Ih> h

<u>表8</u> 位相決定に用いた重原子誘導体結晶。

- 79 -

Derivative	Site	x	У	Z	Relative occupancy	В
EMTS	1	0.0522	0.4330	0.6286	0.608	34.16
HgCl2	1	0.0589	0.4376	0.6045	0.324	30.65
	2	0.0435	0.4288	0.6909	0.224	29.81
	3	0.0454	0.4361	0.6447	0.148	32.96
	4	0.0407	0.4440	0.5527	0.070	26.14
	5	0.0335	0.4263	0.1868	0.020	11.05
5HgCl-uridine(1)	1	0.2673	0.1763	0.2521	0.574	44.48
	2	0.0458	0.4331	0.6308	0.448	36.25
	3	0.0569	0.4365	0.6116	0.440	30.32
	4	0.0461	0.4403	0.5819	0.250	33.41
	5	0.0258	0.4278	0.2006	0.034	19.23
	6	0.0472	0.4519	0.8963	0.028	24.38
5HgCl-uridine(2)	1	0.0507	0.4354	0.6124	0.486	24.06
	2	0.2720	0.1762	0.2480	0.054	7.96
	3	0.0499	0.4515	0.5256	0.036	38.64
BaCl ₂	1	0.3588	0.8689	0.8956	0.464	37.91

-

<u>表9</u> 最終の重原子パラメータ。

- 80 -

.



図29 各分解能に対する phasing power (実線)と figure of merit (破線、目 盛りは右端)の変化。 口は $HgCl_2$ 、 Δ は、5 - クロロマーキュリウリジン(1)、 ☆は5 - クロロマーキュリウリジン(2)、OはEMTS、+は BaCl₂誘導体をそ れぞれ示す。なお、☆と〇は異常分散データを含めて計算したものである。 Δ と☆ の違いについては表8を参照。

- 81 -

Val 4 までを明瞭に認識することができた。このマップでは、分子領域と溶媒領域 の境界が明瞭に現れており、電子密度の断面を透明なプラスチックシート上に描き、 1 Å分ごとに重ねて作成したミニマップでは、すべてのαヘリックスとβシートを 容易に同定する事ができた。また、ほとんどの側鎖、とくに芳香環を側鎖として持 つものについては、明瞭に認識することができた。揺動しやすい1つの領域(His 124 近辺のループ)を除き、主鎖の途切れるところはまったくなく、迷うことなく 主鎖のトレースができた。N末端やC末端の数残基を除き、すべてのアミノ酸残基 はミニマップ上の電子密度にうまくあてはめることができた。なお、初期モデルを 構築(次節)した結果、占有率の特に大きな2個の重原子サイトは、主サイトが、 Cys 63 やN末端近傍に、また 副サイトが Cys 13 近傍に存在していることがわか った。

第4節 分子モデルの構築と精密化

グラフィックス上で電子密度に適合する RNase H の分子モデルを構築した。 こ の際、装置としては三次元ディスプレイ PS390 (Evans & Sutherland)を使用し、 対話型のコンピュータプログラム F R O D O (Jones、1978)を用いて、モデルを電 子密度に適合させた。最初のモデルは重原子同型置換法より得た figure of merit が 0.71 の 2.5Å分解能の電子密度図を用いて、それに適合させるようにして構築 した。さらに F R O D O の最適化ルーチンを用いて逐次、モデルの改良をした。こ の操作を何度も繰り返すことで、初期の R 因子として、10 - 2.5Å分解能で 39 % の値を得た。

さらに、 束縛条件下の最小2乗法プログラムPROLSQ (Hendrickson & Konnert、1980)を用いて 構造の精密化を行った。表10に、精密化計算の全過程の うち主な部分をまとめた。また、この表のそれぞれのサイクルにおけるR因子の収 束する様子を図30にグラフで示した。精密化計算の際には、富士通のスーパーコン

Stage No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1.6	17	1.8
D-map No.				2			3	4	6	9	11	12	15	16	31	33	93	106
lower resol. range (Å) higher resol. range (Å)	5 2.5	5 2.5	10 2.2	10 2.0	10 2.5	10 2.3	10 2.0	10 2.0	6 1.8	6 1.8	6 1.8	6 1.8	6 2.0	6 1.8	10 1.8	5 1.8	6 1.48	6 1.48
No. of cycles No. of parameters No. of atoms No. of solvent atoms over all B factors (A ²) rmsd of 1-2 bond length(A No. of F-data F-data cutoff < Fo - Fc > R (initial)	6 3716 1238 - 13.01)0.008 3212 3σ 22.51	6 3716 1238 - 15.61 0.007 3212 3σ 22.01	11 3716 1238 - 12.96 0.010 5487 3σ 19.86 35.4	9 3716 1238 - 17.68 0.013 7467 3 0 17.68	9 3716 1238 - 10.88 0.008 3590 3590 3590 3590 322.25	4 3716 1238 - 10.80 0.009 4741 3 0 20.40	$ \begin{array}{r} 12\\3716\\1238\\-\\13.00\\0.009\\7467\\3\sigma\\17.01\\34\end{array} $	22 3716 1238 14.20 0.019 7467 3σ 15.78	$ \begin{array}{r} 17 \\ 3716 \\ 1238 \\ - \\ 16.20 \\ 0.013 \\ 9147 \\ 3 \sigma \\ 13.43 \\ 216 \end{array} $	8 4953 1238 - - 0.011 9147 3σ 11.39	$ \begin{array}{r} 15\\3716\\1238\\-\\13.20\\0.012\\9147\\3\sigma\\13.20\\0.22\\\end{array} $	8 4953 1238 - 0.011 9147 3 0 11.12	9 3746 1238 10 16.00 0.011 6846 3 0 14.13	$ \begin{array}{r} 12\\ 4993\\ 1238\\ 10\\ \hline 0.010\\ = 8244\\ 3\sigma\\ 11.37\\ \end{array} $	4 4004 1238 96 17.20 0.013 # 9931: 3 o 14.02	38 5337 1238 96 - 0.014 \$ 9478 3 0 9.21	22 5357 1238 101 0.015 \$16851 1σ 26.955	8 5853 1238 225 - 0.017 16851 1 σ *24.63*

Table 1 Refinement Survey

<u>表10</u> 精密化のプロセスの要所のみを示す。。PROLSQの各段階の最終パラメ ータを示す。また、構造の修正を行うために作成した電子密度図の番号を、2段目 の D-map No.に示した。D-map No.33 まではFASTデータを用いて精密化を行っ た。#印はFo の絶対値が15以下のものを弱い反射として計算の際に除いたことを 示す。これ以外はFo の絶対値が5以下のもののみを除いている。また Stage1か ら14までは反射データの処理の際に、等価な反射どうしの平均よりかけ離れている 反射((I-<I>)/<I> \geq 0.3)を除いたが、\$で示される stage 15 と 16 ではこれも含めた。stage 16 と 17 の間で本文に記されたように分解能を順次上 げていったがこの表では省略した。 また D-map No.34 以降はCAD4データを用 いた。したがって、*印を付した stage 17 と 18 の<|Fo|-|Fc|>の値は 他のものと尺度が異なる。

83

1



- 84

Т

ピュータVP400Eを用いた。精密化の初期には FASTで測定した1.8Å分解 能のデータのうち、5-2.5Å分解能のデータのみを用い、個々の原子の 温度因子は 15.0Å²に固定して精密化した。構造が収束するごとにその最終座標をもとに、 (|Fo|-|Fc|)exp(*i*α_c)と、(|Fo|-|Fc|)exp(*i*α_c)を係数とする差フーリ エ図を計算し、FRODOを用いて三次元グラフィックス上でモデルの修正を行っ た。また、電子密度の明瞭でない部分に関しては、必要に応じオミットマップも計 算した。これらの操作を繰り返し行い、さらに用いる回折データの数を、2.2、2.0、 1.8Å分解能へと、逐次増やしていった。 R因子が 28 %になったところで、水素 原子を除いた全原子の個々の温度因子をパラメータに加えて、精密化を続行した。 1.8Å分解能の FASTデータによる精密化では、最終のR因子は17.6%に収束し た。さらにこの構造を初期値として、高分解能の精密化をするため、CAD4デー タによる1.48Å分解能のデータを用いた精密化を続行した。分解能は、2.5、2.2、 2.0、1.8、1.7、1.6、1.48Åと、逐次ふやしていった。精密化の最終段階において、 モデルに水分子を加えたが、温度因子が 70Å2より大きくなるものはモデルより取 り除いた。精密化終了後のR因子は、6-1.48Å分解能に分布する 16.851 個の回折 強度データに対し 19.6 %となった。このようにして得た最終構造のデータを表11 に示した。また、精密化に用いた初期構造と、得られた最終構造を重ね合わせたも のを図31にCaモデルで示した。 さらに、各分解能領域における解析の統計を表12 に示す。なおこの際、F<1σ(F)の弱い回折データはノイズとみなし、精密化 の計算には用いなかった。

また解析で用いた精密化プログラムは、低分子で用いられる完全マトリックス型 のものではないため、精密化の際の束縛条件として用いた重みのバランス、および 得られた最終構造の、理想モデルからの平均2乗変位(r.m.s.d.)については、表13 に示す。また、原子パラメータの個々の標準偏差もこのプログラムでは得られない ため、Luzzati (1952)の方法により原子座標の平均標準偏差を求めた(図32)。各

- 85 -

Number of atoms	1,238
Number of solvent atoms	225
Number of parameters	5,853
Resolution range for refinement (A)	6.0-1.48
Final overall average B value (A^2)	23.8
Total number of crystals for data collection	ion 6
Number of observed reflections ($\infty -1.48$ Å)	17,254
Data used for refinement $(6-1.48\text{\AA})$	16,851
Reflection cutoff criterion	$F_0 > 1\sigma (F_0)$
$\langle F_{o} - F_{c} \rangle$	24.63
R - factor [†] (final)	0.196

† R - factor = $\Sigma |F_0 - F_c| / \Sigma |F_0|$.

表11 精密化計算の際のパラメータ。



図31 大腸菌 RNase H の、 精密化前の初期構造(細線)と精密化終了後の最終構造(太線)をC aモデルで示したもの。残基番号はN未端から10番ごとに付した。

Observe Number	d-reflections completeness (%)§	SigF(used)	<fo-fc></fo-fc>	Shell R*	Sphere R*
2,515	97.3	21.67	40.52	0.129	0.129
2,512	92.4	18.89	29.05	0.184	0.147
2,405	88.9	17.14	23.64	0.200	0.157
2,320	80.9	15.80	21.35	0.230	0.167
2,149	69.9	14.60	18.35	0.263	0.174
2,544	59.1	13.37	18.09	0.312	0.185
2,400	50.3	12.09	20.20	0.379	0.196
16,851	73.2		24.63		0.196
	Observe Number 2,515 2,512 2,405 2,320 2,149 2,544 2,400 16,851	Observed-reflections Number completeness 2,515 97.3 2,512 92.4 2,405 88.9 2,320 80.9 2,149 69.9 2,544 59.1 2,400 50.3 16,851 73.2	Observed-reflections Number completeness SigF(used) 2,515 97.3 21.67 2,512 92.4 18.89 2,405 88.9 17.14 2,320 80.9 15.80 2,149 69.9 14.60 2,544 59.1 13.37 2,400 50.3 12.09 16,851 73.2	Observed-reflections SigF(used) Fo-Fc> Number completeness SigF(used) <fo-fc> 2,515 97.3 21.67 40.52 2,512 92.4 18.89 29.05 2,405 88.9 17.14 23.64 2,320 80.9 15.80 21.35 2,149 69.9 14.60 18.35 2,544 59.1 13.37 18.09 2,400 50.3 12.09 20.20 16,851 73.2 24.63</fo-fc>	Observed-reflections Number completeness SigF(used) <fo-fc> Shell R* 2,515 97.3 21.67 40.52 0.129 2,512 92.4 18.89 29.05 0.184 2,405 88.9 17.14 23.64 0.200 2,320 80.9 15.80 21.35 0.230 2,149 69.9 14.60 18.35 0.263 2,544 59.1 13.37 18.09 0.312 2,400 50.3 12.09 20.20 0.379 16,851 73.2 24.63 </fo-fc>

§ Completeness of measured to possible reflections. * $R = \sum |F_{obs}(hkl) - F_{calc}(hkl)| / \sum F_{obs}(hkl)$

h k l

hk1

<u>表12</u> 各分解能における測定された反射データの割合、重みのかけかた、および最 終構造の各分解能におけるR因子の分布。

Restraints	r.m.s. deviation [‡]	σ^{\dagger}	No. of parameters
Distance (total)			3423
Bond length	0.017	0.020	1266
(1-2 neighbour)(Å)	0 036	0 020	1710
(1-3 neighbour)(1)	0.030	0.030	1/12
Planar	0.048	0.050	445
(1-4 neighbour)(Å)			
Planar groups (A_{2}^{2})	0.014	0.020	1099 *
Chiral volumes (A^3)	0.164	0.150	179
Non-bonded contacts			812
(Total)			(12908)§
Single torsion (Å)	0.189	0.300	412
Multiple torsion (Å)	0.206	0.300	282
Possible H-bond (Å)	0.270	0.300	107
Torsion angles			414
(Total)			(772)¶
Peptide plane (ω)	2.8	3.0	164
Staggered (\pm 60 or 180°)	18.6	15.0	232
Orthonormal (\pm 90°)	30.5	20.0	18
Isotropic thermal factors			
Main-chain bond	1.445	1.500	639
$(1-2 \text{ neighbour})$ (\AA^2)			
Main-chain angle	2.183	2.000	807
$(1-3 \text{ neighbour}) (\text{\AA}^2)$			
Side-chain bond $({\rm \AA}^2)$	2.409	2.000	627
Side-chain angle (A^2)	3.410	2.500	905
-			

† The weight for each restraint was $1/\sigma^2$.

r.m.s.deviation from ideality.

§ 890 actual out of 12908 possible contacts were restrained. ¶ 414 out of 772 conformational changes were restrained. * The total number of planar groups was 217.

<u>表13</u> 精密化計算の際の、ジオメトリーごとの重みのかけかた(†印)と、最終構造の各ジオメトリーごとの理想モデルに対する最小2乗偏差(‡印)。



図32 Luzzati plot (1952) による大腸菌 RNase H 最終構造の精度の評価。

分解能の殻(shell)ごとに求めたR因子は、平均誤差(Δr)が 0.14Åと0.19Åの理 論曲線の間に分布していることから、全体的には原子の標準偏差は約0.16Åと見積 もられる。

精密化後の最終構造をもとに(2|Fo|-|Fc|)を係数として計算した電子密度 図の一部を、図33に示す。図に示されるように、得られた電子密度は鮮明であり、 個々の原子が分離して現れるという、 1.5Å分解能レベルの解析の条件を良く満た した質の高いものである。

第5節 Mg²⁺結合部位の同定

RNase H 活性に必須とされるMg²⁺結合部位を同定するため、RNase H 結晶を 200 mM の Tris-HCl (pH9.0)、15 %ポリエチレングリコール、および 結合すると予想 される金属イオン (Mg²⁺、Ca²⁺、Ba²⁺、Mn²⁺、Co²⁺)とを様々の濃度で混ぜた溶液 に浸漬 (soaking) した。 ただし、Mn²⁺を含む溶液はアルカリ側でかなりの量の沈 澱が生じるため、浸漬はMn²⁺の飽和溶液を用いて行った。

重原子同型置換法で得た位相をもとに、金属無しの結晶と金属を含む結晶との差 フーリエ図({|Fmeta1|-|Ffree|}exp{*i*α_{MIR}})を計算した。その結果、Mg²⁺、 Ca²⁺、Ba²⁺のアルカリ土類金属イオンでは、共通の位置に単一の明瞭なピークを得 た。このうちMg²⁺の場合については、さらに精密化後の金属なしの結晶によるモデ ルより計算したオミットマップを計算した。このマップでも、前者の差のフーリエ 図で得た位置と同じ位置に明瞭なピークを確認できた。Mn²⁺、Co²⁺の遷移金属イオ ンでは明瞭なピークは得られなかった。



図33 最終構造に基づいて 計算された 2Fo-Fc を係数とする電子密度図。電子 密度の中心を示すため、2種類の等高線で示し、それぞれ青が1 σ、水色が3 σの 高さである。破線は水素結合を意味する。この図に示されるように、各原子が明瞭 に分離しており、 1.5Å分解能のレベルを満たした解析であることがわかる。

<u>第2章の実験</u>

大腸菌 RNase H の 3つの活性部位変異体 D10N、E48Q、D70N は、野生型の酵素 とほぼ同じ条件で結晶化された。得られた結晶はいずれも 空間群が P2₁2₁2₁ の斜 方晶系に属し、野生型のものと同じである。また格子定数については、 D10N と、 E48Q については、野生型と同一であるが、D70N については、 a軸が約2 Åだけ小 さい。興味深いことに、この D70N の格子定数は、 Yang ら (1990) が解析した大 腸菌 RNase H のそれに近いが、 これについては第2章第3節で、もう少し詳しく 述べる。これらの変異体結晶の結晶学的データ、回折強度データの測定範囲、およ び最終のR因子を表14に示す。

変異体 D10N については、4軸型回折計CAD4 (Enraf Nonius)により、10 - 2.02Å分解能の反射 7,130 個を測定した。 構造解析のために、野生型の RNase H の構造を初期座標として用いた。この際まず、rigid body refinement の計算を行ったが、この計算には野生型の座標をもとに、プログラムTRAREF (Huber & Schneider、1985)を適用した。 この、rigid body refinement は、蛋白質分子全体を1つの剛体として、微小な角度の回転や微小量の並進を行い、回折強度に最もうまく適合するモデルの位置を、結晶学的精密化を行う前の時点で決めるものである。特に、格子定数の変化が、非同型とまでは行かない1%以下の僅かな場合などには、次の精密化をうまく進める上で有効である。ついでプログラムPROLSQ (Hendrickson & Konnert、1980)を用いて束縛条件下の最小2乗法による精密化を行った結果、最終のR因子は19.9%に収束した。また、野生型の立体構造と比較するために、主鎖原子どうしが平均的に最もよく重なる場合の最小2乗変位r.m.s.d. $\{\Sigma(r_1-r_3)^2\}^{1/2}$ は0.30Å、また全原子(水素原子を除く)で重ね合わせると0.46Åと小さな値であった。

変異体 D70N については、4軸型回折計CAD4 (Enraf Nonius) により、10 -

	a(Å)	b(Å)	C(Å)	Resolution	R-factor
wild-type	44.06	86.85	35.47	6-1.48Å	19.6%
mutant D10N	44.17	87.23	35.27	10-2.02Å	19.9%
E480 §	44.0	86.5	35.8	10-1.9 Å	21.0%
D70N	42.01	86.42	35.95	10-2.0 Å	27.5% ¶
wild-type (Yang et al.)	41.79	86.34	36.31	10-2.0 A	19.8%

各変異体の比較

§北大・理(田中研究室)との共同研究 ¶温度因子はまだ精密化していない。

<u>表14</u> この研究で用いた 3つの大腸菌 RNase H 活性部位変異体の、格子定数の比較。D70N変異体は Yang ら (1990) が解析に用いた結晶と同型である。さらに、それぞれについて解析に使用した反射データの分解能と現時点での最終のR因子を示した。

- 93 -

.

2.0 Å分解能の反射 5,972 個を測定した。 この分解能における野生型と D70N 変 異体との isomorphous difference ($\Sigma | F_{nati} | - | F_{D70N} | / \Sigma | F_{nati} |$) は平均で、50.1%もあり、さらに、前述したようにa軸方向の格子が2Å(約5% 分)短いため、野生型とは明らかに非同型の結晶であると判断され、野生型の座標 はそのままでは解析の初期構造に適用することが不可能であった。そこでまず、分 子置換法を用いて初期構造の決定を行った。なお、この過程の詳細については図34 に示した。

この際、回転関数の計算にはプログラムPROTEIN (Steigemann、1974)を、 また並進関数の計算には、Crowther & Blow (1967)の理論を計算するE. Lattmann のプログラムをそれぞれ用いた。

野生型の構造をモデルとして、それを、80Å×80Å×80Åの空間群 P1 の仮想格 子中に置き、RNase H 1 分子を構成する1,238 原子の座標から分子の重心を求め、 それを回転の中心として回転関数を計算した。その結果、Huber の定義による回転 $(\psi, \theta, \phi) = (1.9^\circ, 5.4^\circ, -3.8^\circ)$ のところに、顕著な第1のピーク 角で (図35)を得た。この回転角を採用して、さらに並進関数を計算した。その結果を 図36に示すが、X,Y,Zの各方向の並進関数ともそれぞれハーカー面上に、第2 ピーク(ハーカー面以外にある一般のピークも含む)より2倍以上高い第1ピーク を得た。この時点で、結晶のパッキングを見たところ衝突はなく、また電子密度マ ップを計算したところ、ほとんどのアミノ酸残基のモデルが電子密度にうまくフィ ットしたため、このようにして得られた回転・並進関数の解を正解とした。しかし、 これらの解は、パターソン関数をもとに計算したもので、解の精度が、回転関数で 0.2°程度までに限定されるため、 さらに、 プログラム TRAREF (Huber & Schneider、1985)を用いて、 rigid body refinement を行い、解をさらに精度良 いものにした。そののちモデルの電子密度へのフィッティングとプログラムPRO LSQによる束縛条件下の最小2乗法を繰り返し、R因子を 27.5 %まで収束させ



分子置換法



図34 変異体結晶の構造解析の流れ図。



図35-1 回転関数の結果(1)。 ψ 、 θ 、 ϕ を5[®]刻みでサーチしたときの各 ϕ 層における回転関数の最大値をプロットしたもの。なお、この計算で全体の回転関 数の平均値は 4.7であった。



図35-2 回転関数の結果(2)。最終の0.2°刻みでの計算で得た回転関数の最大 値の∮角 -3.8°に対し、ψ、θを変えたときの回転関数の分布を示したもの。図35 -1とあわせ、トップのビークが顕著なものであることを示す。なお、等高線の間 隔は任意である。ただし、この層の平均の高さを0として太線で示し、それ以上の 高さのもののみを示した。



図36 並進関数の結果。上段は各並進操作における並進関数の最大値とその位置。 比較のため第2ピークも示した。A, B, C各操作における並進関数のハーカー面 を下段に示した。等高線は2σおきに引いてある。各操作において、2つの対称操 作の差のベクトルとピーク位置を比較することにより、並進ベクトルの解として、 (x,y,z)=(-0.093,0.173,-0.195)を得る。

た。なお、まだ個々の原子の温度因子はパラメータに加えていない。また、野生型 の立体構造と比較した結果、主鎖原子どうしで重ね合わせた場合r.m.s.d.は0.61Å、 側鎖原子を含めると1.33Åで、 D10N よりやや大きな値であった。これは、野生型 と非同型な結晶なので、分子表面の残基がパッキングの影響を受けたためr.m.s.d. が大きくなったと考えられる。なおこの際には、電子密度が認識しにくかったN末 端の Met 1 から Arg 3、C末端の Val 153 から Val 155、および野生型とループ 構造の異なる Val 121 から Glu 129 までの計15残基分(約1割程度)を重ね合わ せの計算から省いた。

変異体 E48Q の結晶構造解析は、北大の田中 勲 助教授との共同研究であり、デ ータ測定と解析計算は、田中助教授の研究室でおこなわれた。現在の解析状況は、 1.9 Å分解能でR因子が21%である。本論文では、構造の比較に必要なため、第2 章・第2節ではその解析結果を引用した。

参______ _ _ _ _ _ _ _ _ _ 献

- Arni, R., Heinemann, U., Tokuoka, R. & Saenger, W. (1988). Threedimensional Structure of the Ribonuclease T₁*2'-GMP Complex at 1.9-A Resolution. J. Biol. Chem. 263, 15358-15368.
- Arnott, S., Chandrasekaran, R., Millane, R.P. & Park, H.-S. (1986). DNA-RNA Hybrid Secondary Structure, J. Mol. Biol., 188, 631-640.
- Barlow, D. J. & Thornton, J. M. (1988). Helix Geometry in Proteins. J. Mol. Biol. 201, 601-619.
- Berkower, I., Leis, J. & Hurwitz, J. (1973). Isolation and Characterization of an Endonuclease from Escherichia coli Specific for Ribonucleic Acid in Ribonucleic Acid-Deoxyribonucleic Acid Hybrid Structures. J. Biol. Chem. 248, 5914-5921.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E. J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tsumi, M. (1977). The protein data bank: a computerbased archival file for macromolecular structures. J. Mol. Biol. 112, 535-542.
- Carl, P. L., Bloom, L. & Crouch, R. J. (1980). Isolation and Mapping of a Mutation in *Escherichia coli* with Altered Levels of Ribonuclease H. J. Bacteriol. 144, 28-35.
- Carlisle, H. C., Palmer, R. A., Mazmudar, K. S., Gorinsky, B. A.
 & Yeates, D. G. R. (1974). The Structure of Ribonuclease at
 2.5 Angstrom Resolution. J. Mol. Biol. 85, 1-18.
- Crawford, J. L., Lipscomb, W. N. & Schellman C. G. (1973). The reverse turn as a polypeptide conformation in globular proteins. *Proc. Nat. Acad. Sci.*, U.S.A. 70, 538-542.

-99-

- Crouch, R. J. (1990). RIBONUCLEASE H: FROM DISCOVERY TO 3D STRUCTURE. The New Biologist 2, 771-777.
- Crouch, R. J. & Dirksen, M. -L. (1982). Ribonuclease H. In 'Nuclease' (Linn, S. M. & Roberts, R. J., eds), pp. 211-241. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Crowther, R. A. & Blow, D. M. (1967). A Method of Positioning a Known Molecule in an Unknown Crystal Structure. Acta. Crystallogr. 23, 544-548.
- Dasgupta, S., Masukata, H. & Tomizawa, J. -I. (1987). Multiple Mechanisms for initiation of ColE1 DNA replication: DNA synthesis in the presence and absence of ribonuclease H. *Cell* 51, 1113-1122.
- Davies, II, J.F., Hostomska, Z., Hostomsky, Z., Jordan, S.R. & Matthews, D. (1991). Crystal Structure of the Ribonuclease H Domain of HIV-I Reverse Transcriptase. Science 252, 88-95.
- Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman J. M., Joyce, C. M. & Steitz, T. A. (1988). Genetic and Crystallographic Studies of the 3',5'-Exonucleolytic Site of DNA Polymerase I. Science 240, 199-201.
- Doolittle, R. F., Feng, D. -F., Johnson, M. S. & McClure, M. A. (1989). ORIGINS AND EVOLUTIONARY RELATIONSHIPS OF RETROVIRUSES. Quart. Rev. Biol. 64, 1-30.
- Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R. & Steitz, T. A. (1988). Cocrystal structure of an editing complex of Klenow fragment with DNA. Proc. Natl. Acad. Sci. 85, 8924-8928.
- Harrison, S. C. & Aggarwal, A. K. (1990). DNA RECOGNITION BY PROTEINS WITH THE HELIX-TURN-HELIX MOTIF. Annu. Rev. Biochem. 59. 933-969.

- Hendrickson, W. A. & Konnert, J. H. (1980). INCORPORATION OF STEREOCHEMICAL INFORMATION INTO CRYSTALLOGRAPHIC REFINEMENT. In 'Computing in Crystallography' (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp 13.01-13.23, Indian Academy of Science. Int. Union of Crystallography, Bangalore.
- Hol, W. G. J. & Wierenga. R. K. (1984). X-ray crystallography and drug action. (Horn, A. S. & DeRanter, C. J., eds), pp 151-168, Clarendon Press, Oxford.
- Itaya, M. (1990). Isolation and characterization of a second RNase H (RNase HII) of Escherichia coli K-12 encoded by the rnhB gene. Proc. Nat. Acad. Sci. U.S.A. 87, 8587-8591.
- Itaya, M. & Kondo, K. (1991). Molecular cloning of a ribonuclease H (RNase HI) gene from an extreme thermophile Thermus thermophilus HB8: a thermostable RNase H can functionally replace the Escherichia coli enzyme in vivo. Nucl. Acids Res. 19, 4443-4449.
- Itoh, T. & Tomizawa, J.-I. (1980). Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. Proc. Nat. Acad. Sci. U.S.A. 77, 2450-2454.
- Jacobo-Molina, A & Arnold, E. (1991). HIV reverse transcriptase structure-function relationships. *Biochemistry* 30, 6351-6361.
 Johnson, M. S., McClure, M. A., Feng, D. F. Gray, J. & Doolittle, R. F. (1986). Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc. Nat. Acad. Sci. U.S.A.* 83, 7648-7652.
- Jones, T. A. (1978). A Graphics Model Building and Refinement System for Macro-molecules. J. Appl. Crystallogr. 11, 268-272.

- Kabsch, W. & Sander, C. (1983). Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* 22, 2577-2637.
- Kanaya, S. & Crouch, R. J. (1983). DNA Sequence of the Gene Coding for Escherichia coli Ribonuclease H. J. Biol. Chem. 258, 1276-1281.
- Kanaya, S., Kohara, A., Miyagawa, M., Matsuzaki, T., Morikawa, K.
 & Ikehara, M. (1989). Overproduction and Preliminary Crystallographic Study of Ribonuclease H from Escherichia coli. J. Biol. Chem. 264, 11546-11549.
- Kanaya, S., Kimura, S., Katsuda, C. & Ikehara, M. (1990a). Role of cysteine residues in ribonuclease H from *Escherichia* coli: site directed mutagenesis and chemical modification. *Biochem. J.* 271, 59-66.
- Kanaya, S., Kohara, A., Miura, Y., Sekiguchi, A., Iwai, S., Inoue, H., Ohtsuka, E. & Ikehara, M. (1990b). Identification of the Amino Acid Residues Involved in an Active Site of *Escherichia coli* Ribonuclease H by Site-directed Mutagenesis. J. Biol. Chem. 265, 4615-4621.
- Kanaya, S., Katayanagi, K., Morikawa, K., Inoue, H., Ohtsuka, E. & Ikehara, M. (1991a). Effect of mutagenesis at each of five histidine residues on enzymatic activity and stability of ribonuclease H from *Escherichia coli*. *Eur. J. Biochem.* 198, 437-440.
- Kanaya, S., Katsuda, C., Kimura, S., Nakai, T., Kitakuni, E., Nakamura, H., Katayanagi, K., Morikawa, K. & Ikehara, M. (1991b). Stabilization of *Escherichia coli* Ribonuclease H by Introduction of an Artificial Disulfide Bond. J. Biol. Chem. 266, 6038-6044.

- Katayanagi, K., Miyagawa, M., Ishikawa, M., Matsushima, M., Kanaya, S., Ikehara, M., Matsuzaki, T. & Morikawa, K. (1990). Three-dimensional structure of ribonuclease H from *E. coli*. *Nature (London)* 347, 306-309.
- Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzaki, T. & Morikawa, K. (1992). Structural details of ribonuclease H from Escherichia coli refined to an atomic resolution. J. Mol. Biol. In press.
- Landschulz, W. H., Johnson P. F. & McKnight, S. L. (1988). The Leucine Zipper: A Hypothetical Structure Common to a New Class of DNA Binding Proteins. Science 240, 1759-1764.
- Levitt, M. & Chothia C. (1976). Structural patterns in globular proteins. Nature (London) 261, 552-557.
- Luzzati, V. (1952). Traitement statisque des erreurs dans la determination des structures cristallines. Acta Crystallogr. 5, 802-810.
- Maki, H., Horiuchi, T. & Sekiguchi, M. (1983). Structure and expression of the *dnaQ* mutator and the RNase H genes of *Escherichia coli*: Overlap of the promoter regions. *Proc. Nat. Acad. Sci. U.S.A.* 80, 7137-7141.
- Matthews, B. W. (1968). Solvent Content of Protein Crystals. J. Mol. Biol. 33, 491-497.
- Mauguen, Y., Hartley, R.W., Dodson, E.J., Dodson, G.G., Bricogne,
 G., Chothia, C. & Jack, A. (1982). Molecular structure of a new family of ribonucleases. Nature (London) 297, 162-164.
 Messerschmidt, A. & Pfligrath, J. W. (1987). Crystal orientation and X-ray pattern prediction routines for area-detector diffractometer systems in macromolecular crystallography. J.
Appl. Crystallogr. 20, 306-315.

Miller, H. I., Riggs, A. D. & Gill, G. N. (1973). Ribonuclease H (Hybrid) in Escherichia coli. J. Biol. Chem. 248, 2621-2624.
Nakamura, H. & Wada, A. (1985). Nature of the Charge distribution in Proteins. III. Electron Multipole Structures J. Phys. Soc. Jpn. 54, 4047-4052.

- Nakamura, H. & Nishida, S. (1987). Numerical Calculations of Electrostatic Potentials of Protein-Solvent Systems by the Self Consistent Boundary Method. J. Phys. Soc. Japan. 56, 1609-1622.
- Nakamura, H., Katayanagi, K., Morikawa, K. & Ikehara, M. (1991a). Structural models of ribonuclease H domains in reverse transcriptases from retroviruses. Nucl. Acids Res. 19, 1817-1823.
- Nakamura, H., Oda, Y., Iwai, S., Inoue, H., Ohtsuka, E., Kanaya, S., Kimura, S., Katsuda, C., Katayanagi, K., Morikawa, K., Miyashiro, H. & Ikehara, M. (1991b). How does Ribonuclease H recognize a DNA/RNA hybrid. Proc. Nat. Acad. Sci. 88, 11535-11539.
- Nishikawa, K. & Ooi, T. (1986). Amino acid sequence homology applied to the prediction of protein secondary structures, and joint prediction with existing methods. *Biochim. Biophys. Acta* 871, 45-54.
- Nishikawa, K., Ooi, T., Isogi, Y. & Saito, N. (1972). Tertiary Structure of Proteins. I. Representation and Computation of the Conformations. J. Phys. Soc. Jpn. 32, 1331-1337.
- North, A. C. T., Phillips, D. C. & Matthews, F. S. (1968). A Semi-Empirical Method of Absorption Correction. Acta Crystallogr. sect. A, 24, 351-359.

-104-

- Oas, T. G., McIntosh, L. P., O'Shea, E. K., Dahlquist, F. W. & Kim, P. S. (1990). Secondary Structure of a Leucine Zipper Determined by Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* 29, 2891-2894.
- Oda, Y., Nakamura, H., Kanaya, S. & Ikehara, M. (1991). Binding of metal ions to *E. coli* RNase HI observed by ¹H-¹⁵N heteronuclear 2D NMR. *J. Biomolec. NMR*, 1, 247-255.
- Oefner, C. & Suck, D. (1986). Crystallographic Refinement and Structure of DNase I at 2A resolution. J. Mol. Biol. 192, 605 -623.
- Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configrations. J. Mol. Biol. 7, 95-99.
- Richards, F. M. & Wyckoff, H. W. (1971). Bovine Pancreatic Ribonuclease. In 'The Enzymes' Vol.4 (Boyer, P. D., ed.), pp. 647-806, Academic Press, New York.
- Shannon, R. D. (1976). Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. Acta Crystallogr. sect. A, 32 751-767.
- Steigemann, W. (1974). Die Entwicklung und Anwendung von Rechenverfahren und Rechenprogrammen zur Strukturanalyse von Proteinen am Beispiel des Trypsin-Trypsininhibitor Komplexes, des freien Inhibitors und der L-Asparaginase. Ph. D. thesis, Technical University of Munich, F.R.G.
- Stein, H. & Hausen, P. (1969). Enzyme from calf thymus degrading the RNA moiety of DNA:RNA hybrids: Effect on DNA dependent RNA polymerase. Science 166, 393-395.
- Suck, D., Oefner, C. & Kabsch, W. (1984). Three-dimensional structure of bovine pancreatic DNase I at 2.5A resolution.

EMBO J. 3, 2423-2430.

- Suck, D. & Oefner, C. (1986). Structure of DNase I at 2.0A resolution suggests a mechanism for binding to and cutting DNA. Nature (London) 321, 620-625.
- Suck, D., Lahm, A. & Oefner, C. (1988). Structure refined of a nicked DNA octanucleotide complex with DNase I. Nature (London) 332, 464-468.
- Sugio, S., Amisaki, T., Onishi, H. & Tomita K.-I. (1988). Refined X-Ray Structure of the Law pH Form of Ribonuclease $T_1-2'-$ Guanylic Acid Complex at 1.9 A Resolution. J. Biochem. 103, 354-366.
- Varmus, H. (1988). Retroviruses. Science 240, 1427-1435.
- Venkatachalam, C. M. (1968). Stereochemical Criteria for Polypeptides and Proteins. V. Conformation of a System of Three Linked Peptide Units. *Biopolymers* 6, 1425-1436.
- Vinson, C.R., Sigler, P.B. & McKnight, S.L. (1989). Scissors-Grip model for DNA Recognition by a Family of Leucine Zipper Proteins. Science 246, 911-916.
- Wada, A. (1963). THE α -HELIX AS AN ELECTRIC MACRO-DIPOLE. Adv. Biophys. 9, 1-63.
- Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the α - and β -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945-951.
- Wang, A. H.-J., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, S. A., & Rich, A. (1982). Molecular Structure of r(GCG)d(TATACGC): a DNA-RNA Hybrid Helix Joined to Double Helical DNA. Nature (London) 299, 601-604.
- Weiner, S.J., Kollman, P.A., Nguyen, D.T. & Case, D. A. (1986).

An All Atom Force Field for Simulations of Proteins and Nucleic Acids. J. Comput. Chem. 7, 230-252.

- Yamazaki, T., Yoshida, M., Kanaya, S., Nakamura, H. & Nagayama, K.(1991). Assignments of Backbone ¹H, ¹³C, and ¹⁶N Resonances and Secondary Structure of Ribonuclease H from *Escherichia coli* by Heteronuclear Three-Dimensional NMR Spectroscopy. *Biochemistry* 30, 6036-6047.
- Yang, W., Hendrickson, W. A., Crouch, R. J. & Satow, Y. (1990). Structure of Ribonuclease H Phased at 2 A Resolution by MAD Analysis of the Selenomethionyl Protein. Science 249, 1398-1405.

